A model for multiscaling and clustering of volatility in financial indexes

Alessandro Andreoli, Francesco Caravenna, Paolo Dai Pra and Gustavo Posta

Abstract— We propose a stochastic model which matches some relevant stylized facts observed in time series of financial indexes, and that are not fully captured by the models most often used in this context. These stylized facts concern with the distribution of the *log-returns* (increments of the logarithm of the index). This distribution is not Gaussian, and its moments obey peculiar scaling relations (*multiscaling*). Moreover, absolute values of log-returns in disjoint time intervals are positively correlated (*clustering of volatility*): their correlation has slow (sub-exponential) decay for moderate time distances (up to few months), and have a faster decay for larger distances. The simplicity of the model allows sharp analytic results, statistical estimation of its few parameters, and low computational effort in simulations, allowing its concrete use in applications such as option pricing.

I. INTRODUCTION

The stochastic modeling of financial indexes has challenged the scientific community for decades. The basic model, that has given rise to the celebrated Black & Scholes formula [6], [16], assumes that the logarithm X_t of the price of the underlying index, after subtracting the trend, is given by

$$\mathrm{d}X_t = \sigma \, \mathrm{d}W_t, \qquad (\mathrm{I.1})$$

where σ (the *volatility*) is a constant and $(W_t)_{t \geq 0}$ is a standard Brownian motion. This model has been widely used in practice; it shows, however, several inconsistencies with the data coming from most real financial indexes. The following is a partial list of these inconsistencies, that we support with the data of the *Dow Jones Industrial Average* (DJIA), cf. Figure 1. We have also analyzed other aggregate indexes, such as the NIKKEI, and non-aggregate ones, such as prices of single stocks; the evidence is that the following features are typical of aggregate

Alessandro Andreoli, Francesco Caravenna and Paolo Dai Pra are with the Department of Pure and Applied Mathematics, University of Padova, Padova, 35121, Italy `aandreoli@math.unipd.it`, `fcaraven@math.unipd.it`, `daipra@math.unipd.it`
Gustavo Posta is with the Department of Mathematics, Politecnico di Milano, Milano, 20133, Italy `gustavo.posta@polimi.it`

indexes, but may or may not appear in non-aggregate ones.

(1) Model (I.1) predicts a Gaussian law for the increments $X_{t+h} - X_t$ of the logarithm of the price (the *log-returns*). The empirical distributions show heavier tails (Figure 1(A)).

(2) According to model (I.1), log-returns corresponding to disjoint time intervals should be independent. The empirical evidence is that log-returns 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* (Figure 1(B)).

(3) The volatility may be estimated by computing the local empirical average of the absolute values of daily log-returns. Time dependence of the volatility may be detected by plotting the *simple moving average*. Figure 1(C) reveals a very irregular behavior of these moving averages, which have very localized peaks, in contrast with the constant volatility prescribed by (I.1).

In order to have a better fit with real data, many different models have been proposed to describe the volatility and the price process. Besides the consistency with the stylized facts above, it is desirable to have a model which preserves the basic diffusive scaling property of (I.1). Consider in fact the time series of an index $(s_i)_{1 \leq i \leq T}$ 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:

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

where $\overline{d}(i)$ is the local empirical trend of the series $(\log(s_i))_{1 \leq i \leq T}$ (see section III) and $\delta_x(\cdot)$ denotes the Dirac measure at $x \in \mathbb{R}$. The statistical analysis of various indexes shows that, for h within a suitable time scale, p_h obeys approximately a diffusive scal-

(A) *Tails of the daily log-returns empirical distribution, in log-scale. A Normal distribution would yield a concave shape $(-c \cdot x^2)$.*

(B) *Correlations: log plot (convex shape corresponds to sub-exponential decay).*
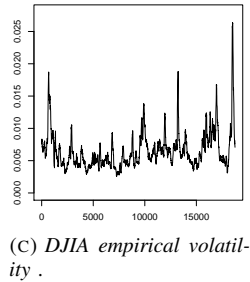

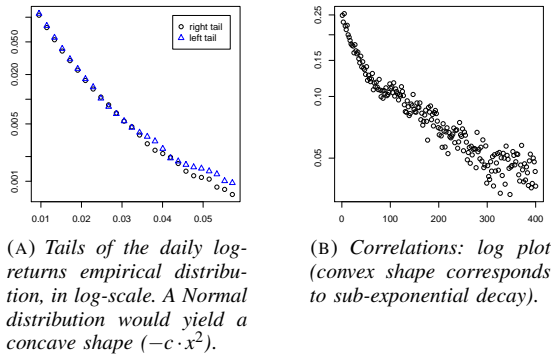(C) *DJIA empirical volatility.*

FIG. 1: *Some features of the DJIA time series (1935-2009) that are inconsistent with model* (I.1).
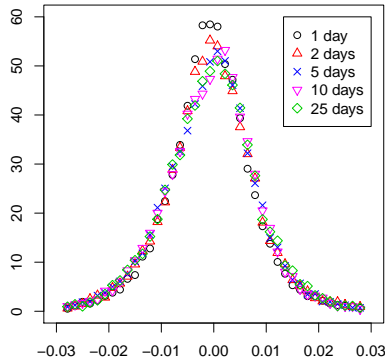

FIG. 2: *Diffusive scaling of the DJIA time series (opening prices 1935-2009)*

ing relation (cf. Figure 2):

$$p_h(\mathrm{d}r) \simeq \frac{1}{\sqrt{h}} g\left(\frac{r}{\sqrt{h}}\right) \mathrm{d}r, \qquad (\mathrm{I}.3)$$

where g is a probability density with tails heavier than Gaussian.

A subtler scaling property has been recently pointed out in [11], [4], [17]. If one considers the q-th empirical moment

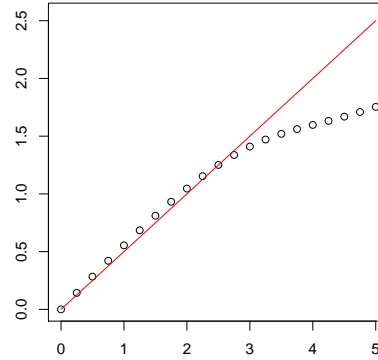$$m_q(h) := \int |r|^q \, p_h(\mathrm{d}r), \qquad (\mathrm{I}.4)$$

FIG. 3: *Multiscaling of empirical moments of the DJIA*

from relation (I.3) it is natural to guess that $m_q(h)$ should scale as $h^{q/2}$. This is indeed what one observes 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 3 (see also [11] for similar pictures concerning other indexes). This is the so-called *multiscaling of moments*.

Within the variety of models that have been proposed, both theorists and practitioners appear to have selected models that, on the one hand, are simple enough to allow statistical estimation of parameters and simulations, on the other hand, reasonably explain *some* of the facts mentioned above. In discrete-time, autoregressive models such as ARCH and GARCH [12], [7] have been widely used. In continuous time, the basic model $\mathrm{d}X_t = \sigma \mathrm{d}W_t$ has been modified by letting $\sigma = \sigma_t$ be a stochastic process, often the solution of a Stochastic Differential Equation driven by a general Lévy process. A systematic account of these *stochastic volatility models* can be found in [5]. 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 calibration with real data is itself a subject of research.

One of the most celebrated, and widely used, of these model, the GARCH ([12]), exhibits non-constant volatility, clustering of volatility and non-Gaussian distribution of log-returns. However, 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;

- correlations of log-returns decay exponentially; in contrast, in the DJIA (cf. Fig-

ure 1(C)), as well as in other indexes, like the Nikkei, the decay of correlations is slower than exponential, up to a time gap of about 100 days.

More refined version of GARCH (for instance FI-GARCH, see [3], [8]) 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.

In this paper we show that all stylized facts that we have mentioned can be all accounted for, with a striking degree of precision, by a model which is relatively simple and depends on few parameters. The basic idea is that non-constant volatility can be obtained by a (possibly random) time change in model (I.1), in other words a process of the form

$$X_t = W_{I_t},$$

for some nonnegative, increasing process I_t. In particular, localized peaks in the empirical volatility as in Figure 1(C) could correspond to singular points of $\frac{d}{dt}I_t$. The idea of using random time changes in this context is certainly not new, see for example [2], [9], [10]. In part inspired by [4], [17] we propose a specific class of time changes, where *shocks* naturally appear as points of a Poisson processes. The full definition of the model will be done in section II, where our main results are also stated. In section III we discuss the estimation of the parameters of the model, and provide comparisons with real data.

II. THE MODEL AND ITS MAIN PROPERTIES

Given a real number $\lambda > 0$ and a probability ν on $(0,\infty)$, 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 $\mathscr{T} = (\tau_n)_{n\in\mathbb{Z}}$ on $\mathbb{R}$ with intensity λ;

- a sequence $\Sigma = (\sigma_n)_{n\geq 0}$ of i.i.d. (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 $W, \mathscr{T}, \Sigma$ are defined on some probability space $(\Omega, \mathscr{F}, P)$ and that they are independent. By convention, we label the points of $\mathscr{T}$ so that $\tau_0 < 0 < \tau_1$. We will actually need only the points of $\mathscr{T} \cap [\tau_0, \infty)$, 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 i.i.d. $Exp(\lambda)$ (exponentially distributed with mean λ^{-1}). Although some of our results would hold for more general distributions of $\mathscr{T}$, we stick for simplicity to the (rather natural) choice of a Poisson process. For $t \geq 0$, define

$$i(t) := \sup\{n \geq 0 : \tau_n \leq t\} = \#\{\mathscr{T} \cap (0,t]\}, \quad \text{(II.1)}$$

so that $\tau_{i(t)}$ is the location of the last point in $\mathscr{T}$ before t. Consider also a strictly increasing, *concave* function $\theta : [0,+\infty) \to [0,+\infty)$, such that $\theta(0) = 0$. Now we introduce the process $I = (I_t)_{t\geq 0}$ by

$$I_t := \sigma_{i(t)}^2 \theta\left(\lambda(t - \tau_{i(t)})\right)$$
$$+ \sum_{k=1}^{i(t)} \sigma_{k-1}^2 \theta\left(\lambda(\tau_k - \tau_{k-1})\right) - \sigma_0^2 \theta\left(-\lambda\tau_0\right), \quad \text{(II.2)}$$

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

$$X_t := W_{I_t}. \quad \text{(II.3)}$$

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

- The points of $\mathscr{T}$ are the times at which a "shock" in the market occurs.

- Between two consecutive points of $\mathscr{T}$, say τ_k and τ_{k+1}, the log-price X evolves as a Brownian motion with possibly random volatility σ_k, and *subject to the deterministic time change* $t - \tau_k \mapsto \theta(\lambda(t - \tau_k))$.

As we have mentioned above, the idea of modeling the reaction of the market to a shock by a nonlinear time change is not new. The prototype of such a time change, as proposed in [4], [17], is give by the function $\theta(s); = s^{2D}$ for some $D \leq 1/2$; see also [15] for applications of this specific time change.

The key point of this work is that, by specifying the distribution of the shocks, we obtain a simple model, for which the agreement with stylized facts can be checked both numerically and by rigorous analysis.

In what follows we state the most relevant property of the model, under the following assumption on the function θ.

Assumption A. The function $\theta : [0,+\infty) \to [0,+\infty)$ satisfies the following properties:

(1) θ is concave, strictly increasing, $\theta(0) = 0$, and $\theta(s) \to +\infty$ as $s \to +\infty$.

(2) θ is $\mathscr{C}^1$ on $(0, +\infty)$, and there exists a constant $0 < D < \frac{1}{2}$ such that

$$\lim_{s \downarrow 0} \frac{\theta'(s)}{2Ds^{2D-1}} = 1.$$

In particular, $\theta(s)$ behaves as s^{2D} near $s = 0$.

All proofs can be found in [1].

A. BASIC FACTS

We begin by stating some properties on the distribution of X_t. From now on we write $E(\sigma^q)$ for $\int \sigma^q \nu(d\sigma)$.

Proposition 1: For every $q > 0$ we have

$$E(|X_t|^q) < \infty \text{ for some (hence any) } t > 0$$
$$\Longleftrightarrow \quad E(\sigma^q) < \infty. \quad (\text{II.4})$$

Moreover, regardless of the distribution of σ, for every $q > (1-D)^{-1}$ we have

$$E\left[\exp\left(\gamma |X_t|^q\right)\right] = \infty, \qquad \forall t > 0, \ \forall \gamma > 0. \quad (\text{II.5})$$

Note that, for $D < 1/2$, $(1-D)^{-1} < 2$; thus equation (II.5) implies that the tails of X_t are heavier that Gaussian. More detailed relations between the tails of ν and those of X_t can be found in [1].

The statistical estimation of the parameters of the model, that will be discussed in section III, is based on the ergodic behavior of the empirical averages of log-returns, that is guaranteed by the following result.

Proposition 2: The process $(X_t)_{t \geq 0}$ is a zero-mean martingale (provided $E(\sigma) < +\infty$), with stationary and ergodic increments. In particular, for every $\delta > 0$, $k \in \mathbb{N}$ and for every choice of the intervals $(a_1, b_1), \ldots, (a_k, b_k) \subseteq (0, \infty)$, the limit

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} F(X_{n\delta+b_1} - X_{n\delta+a_1}, \ldots, X_{n\delta+b_k} - X_{n\delta+a_k})$$
$$= E\left[F(X_{b_1} - X_{a_1}, \ldots, X_{b_k} - X_{a_k})\right], \quad (\text{II.6})$$

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

We remark that the martingale property of (X_t) is essential for the use of the model in pricing options in absence of arbitrage.

B. MAIN RESULTS

The next result shows that the distribution of log-returns obeys an approximate diffusive scaling, with a limiting law with power-law tails.

Theorem 3: As $h \downarrow 0$ we have the convergence in distribution

$$\frac{(X_{t+h} - X_t)}{\sqrt{h}} \xrightarrow[h \downarrow 0]{d} f(x) \, dx, \quad (\text{II.7})$$

where f is a mixture of centered Gaussian densities, namely

$$f(x) = \int_0^\infty \nu(d\sigma) \int_0^\infty dt \, \lambda e^{-\lambda t} \frac{t^{1/2-D}}{\sigma \sqrt{4D\pi}} \exp\left(-\frac{t^{1-2D}x^2}{4D\sigma^2}\right). \quad (\text{II.8})$$

We stress that the function f appearing in (II.7)–(II.8), 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 (II.4) that the same holds for $(X_{t+h} - X_t)$. However, from (II.8) and a simple change of variables we get

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

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

This singular behavior of f produces peculiar scaling properties on the moments of X_t, in agreement with what shown in Figure 3.

Theorem 4 (Multiscaling of moments): Let $q > 0$, and assume $E(\sigma^q) < +\infty$. 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 $h \downarrow 0$:

$$m_q(h) \sim \begin{cases} C_q \, (\lambda h)^{\frac{q}{2}} & \text{if } q < q^* \\ C_q \, (\lambda h)^{\frac{q}{2}} \log(\frac{1}{\lambda h}) & \text{if } q = q^* \\ C_q \, (\lambda h)^{Dq+1} & \text{if } q > q^* \end{cases},$$

where $q^* := \frac{1}{(\frac{1}{2}-D)}$.
The constant $C_q \in (0, \infty)$ is given by one of the following expressions

$$\begin{cases} E(|W_1|^q) E(\sigma^q) \int_0^{+\infty} (\theta'(s))^{q/2} e^{-s} ds & \text{if } q < q^* \\ E(|W_1|^q) E(\sigma^q) (2D)^{q/2} & \text{if } q = q^* \\ E(|W_1|^q) E(\sigma^q) \left[\int_0^\infty ((1+x)^{2D} - x^{2D})^{\frac{q}{2}} \, dx + \frac{1}{Dq+1} \right] & \text{if } q > q^* \end{cases} \quad (\text{II.9})$$

It follows, in particular, that the following relation holds true:

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

(II.10)

The explicit form (II.9) of the multiplicative constant C_q will be used in section III for the estimation of the parameters of our model on the DJIA time series.

Our final results establishes the clustering of volatility. Note that, since $(X_t)_{t \geq 0}$ is a zero-mean martingale, its increments have zero correlation. We show here sharp asymptotics for the correlations of the *absolute* log-returns. We let

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

denote the correlation between two random variables X and Y.

Theorem 5 (Volatility autocorrelation): Assume that $E(\sigma^2) < \infty$. Define, for $x \geq 0$

$$\phi(x) := Cov\big(\sigma \sqrt{\theta'(S)}, \sigma \sqrt{\theta'(S+x)}\big), \quad \text{(II.11)}$$

where $\sigma \sim \nu$ and $S \sim Exp(1)$ are independent. The correlation of the increments of the process X has the following asymptotic behavior as $h \downarrow 0$:

$$\lim_{h \downarrow 0} \rho(|X_{s+h} - X_s|, |X_{t+h} - X_t|)$$

$$= \frac{2}{\pi Var\big[\sigma |W_1| \sqrt{\theta'(S)}\big]} e^{-\lambda |t-s|} \phi(\lambda |t - s|).$$

(II.12)

This shows that the volatility autocorrelation of our process X decays exponentially fast for time scales greater than the mean distance $1/\lambda$ between the epochs τ_k, while for shorter time scales the decay is governed by the function $\phi(\cdot)$. Under our assumptions on $\theta(\cdot)$, it can be shown that for $x = O(1)$, the decay of $\phi(x)$ is faster than polynomial but slower than exponential (see Figure 5).

III. ESTIMATION OF PARAMETERS

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 real time series of the Dow Jones Industrial Average (DJIA) over a period of 75 years: more precisely, we have taken the DJIA opening prices from 1 Jan 1935 to 31 Dec 2009, for a total of 18849 data. 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. Theorem 4.

(b) The *volatility autocorrelation decay*, cf. Theorem 5.

(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.

In its full generality, the model has one scalar parameter, λ, and two functional parameters, the measure ν and the function $\theta(\cdot)$. In this paper we restrict the analysis to the case, proposed in [4], [17], of

$$\theta(s) := s^{2D},$$

with $0 < D < 1/2$. With this choice, the quantities $A(\cdot)$ (see (II.10)), C_1, C_2 (see (II.9)), and the asymptotic volatility autocorrelation

$$\rho(t) := \frac{2}{\pi Var\big[\sigma |W_1| \sqrt{\theta'(S)}\big]} e^{-\lambda t} \phi(\lambda t) \quad \text{(III.1)}$$

(see (II.12)), are all functions of $D, \lambda, E(\sigma), E(\sigma^2)$. We therefore set up the following procedure for estimating those four scalar parameters.

Step 1 The DJIA time series is denoted by $(s_i)_{0 \leq i \leq N}$ (where $N = 18848$) and the corresponding detrended log-DJIA time series will be denoted by $(x_i)_{0 \leq i \leq N}$, where

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

and $\overline{d}(i)$ is the mean log-DJIA price on the previous 250 days from time i (in other terms $\overline{d}(i) = \frac{1}{250} \sum_{k=i-250}^{i-1} \log(s_i)$).

Step 2 The empirical evaluation of the scaling exponent $A(q)$, cf. (II.10), 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 $\widehat{m}_q(h)$ over h days, defined by

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

By Theorem 4, for h small the following relation holds:

$$(\log \widehat{m}_q(h)) \sim A(q)(\log \lambda h) + \log(C_q)$$
$$= A(q)(\log h) + \log(K_q),$$

with

$$K(q) = \lambda^{A(q)} C_q. \quad \text{(III.2)}$$

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 K_q on the DJIA time series, that we call $\widehat{A}(q)$ and $\widehat{K}_q$.

Step 3 The computation in step 2 provides, in particular, the estimates $\widehat{K}_1$ and $\widehat{K}_2$; they can be compared with their theoretical values K_1, K_2 given by (III.2) and (II.9), which only depends on $D, \lambda, E(\sigma), E(\sigma^2)$.

Step 4 We then compute from data the empirical volatility autocorrelation $\widehat{\rho}(t)$, that is plotted in Figure 1(B), that has to be compared with its theoretical counterpart $\rho(t)$ given in (III.1).

Step 5 We consider a *loss function* of the following form

$$L(D,\lambda,E(\sigma),E(\sigma^2)) := \alpha_1 \left(\frac{\widehat{K}_1}{K_1} - 1 \right)^2$$

$$+ \alpha_2 \left(\frac{\widehat{K}_2}{K_2} - 1 \right)^2$$

$$+ \frac{\alpha_3}{20} \sum_{k=1}^{20} \left(\frac{\widehat{A}(k/4)}{A(k/4)} - 1 \right)^2$$

$$+ \frac{\alpha_4}{\sum_{t=1}^{400} e^{-t/T}} \sum_{t=1}^{400} e^{-t/T} \left(\frac{\widehat{\rho}(t)}{\rho(t)} - 1 \right)^2,$$

where $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ are suitable *weights* and T controls a discount factor in long-range correlations. Once $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ and T are chosen, we define the estimates of the parameters of the model by

$$(\widehat{D}, \widehat{\lambda}, \widehat{E(\sigma)}, \widehat{E(\sigma^2)}) := \text{argmin } L(D,\lambda,E(\sigma),E(\sigma^2)).$$

It should be remarked that the empirical distribution of log-returns is not fully used in the procedure above: in fact, we only obtain estimates for the first two moments of ν; in particular, the tails of ν are not estimated. One could alternatively proceed by choosing ν within a certain parametric class, and estimate the parameters as illustrated above. For the DJIA, this turns out to be not relevant since, for "reasonable" values of $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ and T, $\widehat{E(\sigma^2)} \sim \left(\widehat{E(\sigma)} \right)^2$, so that σ is nearly constant.

We have then proceeded to the numerical study of the loss function $L(D,\lambda,E(\sigma),E(\sigma^2))$, with the choice $a_1 = a_2 = \frac{1}{2}$, $a_3 = a_4 = 1$, $T = 40$. The
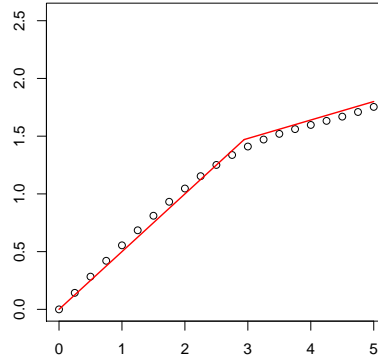

FIG. 4: *Multiscaling of empirical moments of the DJIA: the red curve is the piecewise linear A(q) in (II.10).*
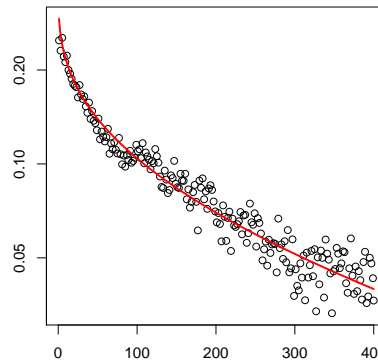

FIG. 5: *Volatility autocorrelation of the DJIA (log plot); the red curve is the graph of the function ρ in (III.1).*

function L appears to be quite regular and convex, and the value of the argmin varies slowly if we change the choice of $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ and T. We approximately obtain the following estimates:

$$\widehat{D} \simeq 0.16, \quad \widehat{\lambda} \simeq 9.7 \times 10^{-4},$$

$$\widehat{E(\sigma)} \simeq \sqrt{\widehat{E(\sigma^2)}} \simeq 0.33.$$

In particular, the mean time between two consecutive shock is $\simeq \widehat{\lambda}^{-1} \simeq 1000$ (working) days, i.e. about four years.

With these values of the parameters, we can now check to what extent our model fits the stylized that have been mentioned in the introduction. To begin with, the following plots show that multiscaling of moments (Figure 4) and clustering of volatility (Figure 5) agree with the corresponding theoretical behaviors within a striking degree of precision. Although the agreement with those curves is exactly what we have aimed for in minimizing the loss L, the fact of obtaining this precision with the calibration of four scalar parameters is, we believe, remarkable.
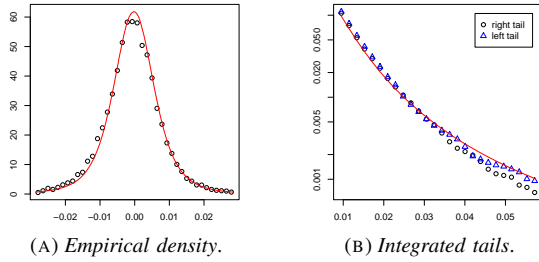
(A) *Empirical density.*   (B) *Integrated tails.*

FIG. 6: *Distribution of 1-day log-returns in the DJIA time series (1935-2009).*
(A) The density of the empirical distribution $\hat{p}_1(\cdot)$ and the theoretical prediction (line).
(B) The integrated right and left tails of the empirical distribution $\hat{p}_1(\cdot)$ and the theoretical prediction (line).

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

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

In Figure 6(A) we have plotted the bulk of the distributions $p_1(\cdot)$ and $\hat{p}_1(\cdot)$ 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 $\hat{p}_1(\cdot)$ (i.e., the empirical standard deviation of the daily log returns evaluated on the DJIA time series). In Figure 6(B) we have plotted the integrated tail of $p_1(\cdot)$, that is the function $z \mapsto P(X_1 > z) = P(X_1 < -z)$ (note that $X_t \sim -X_t$ for our model) and the right and left integrated empirical tails $\hat{R}(z)$ and $\hat{L}(z)$ of $\hat{p}_1(\cdot)$, defined for $z \ge 0$ by

$$\hat{L}(x) := \frac{\#\{i \in \{1, N\} : x_i - x_{i-1} < -z\}}{N}$$

$$\hat{R}(x) := \frac{\#\{i \in \{1, N\} : x_i - x_{i-1} > z\}}{N},$$

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.*

We now would like to point out a further aspect, which goes beyond what we have discussed so far. The parameters estimation has been based on various estimators. In particular, we have emphasized the role of $\hat{A}(q)$ and $\hat{\rho}(t)$, which show the main facts our model fits with. If we compute those
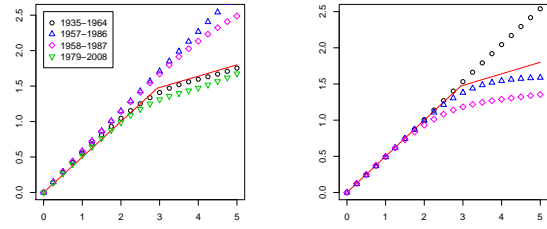

(A) *Subperiods of 30 years of the DIJA time series.*   (B) *Subperiods of 30 years of a simulated time series.*

FIG. 7: *Multiscaling of moments in the DJIA time series (1935-2009) and in a sample from the law of our model.*
(A) The empirical scaling exponent $\hat{A}(q)$ evaluated in subperiods of 30 years (see legend) and the theoretical scaling exponent $A(q)$ for $D = 0.16$ (line).
(B) The empirical scaling exponent $\hat{A}(q)$ evaluated in subperiods of 30 years of a simulated sample, and the theoretical scaling exponent $A(q)$ for $D = 0.16$ (line).

estimators on *different sub-periods* of the 75 years of data, we would observe a remarkable variability, in particular in the curve $\hat{A}(q)$. In Figure 7(A) we have plotted $\hat{A}(q)$ evaluated in different sub-periods of 30 years. This shows that whatever model is to be used for the DJIA, 30 years are not enough for the empirical measure of increments to get close to its ergodic limit. This is indeed consistent with the model we propose. The Poisson process $\mathscr{T}$ of the shocks has an estimated mean inter-arrival time of about 4 years. Thus, in a 30 years period there are too few events for empirical averages to stabilize. In Figure 7(B) we plot $\hat{A}(q)$ evaluated in different sub-periods of 30 years of a single 75-years long trajectory, sample from our model. We can see that fluctuations in the simulated path are consistent with what seen on real data. A similar analysis on $\hat{\rho}(t)$ leads to a similar conclusion, see [1] for details.

We conclude this section by mentioning some facts, concerning the DJIA, that are *not* properly reproduced by the model we have just calibrated. Given the simplicity of the model and the few parameters in the game, a perfect matching would be unrealistic.

- One of the stylized facts motivating our work is the behavior of the empirical volatility, plotted in Figure 1(C). A similar plot on simulated paths shows peaks distributed according to a Poisson law, which is consistent with data, but having nearly the same height, unlike in Figure 1(C). The main reason is that the law v is our model, as calibrated to the DJIA, is nearly deterministic. We have

441

numerical evidence that this mismatch can be avoided by using a different scaling time function $\theta(\cdot)$ instead of t^{2D}, like the following one:

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

with c a constant parameter to be estimated.

- Figure 6(B) seems to suggest that our model slightly overestimate tails of the distribution of the daily log-return. This overestimation becomes relevant for n-days log-returns with n of the order of few months.

- Figure 6(B) shows also a small but nonzero skewness in the distribution of the daily log-return. Skewness has been analyzed for many indexes, and roughly follows from the evidence that "large log-returns are more likely to be negative that positive". In our model, X_t and $-X_t$ are equally distributed, so no skewness is admitted. At the discrete-time level, and therefore at the level of simulations, skewness could be easily introduced in the framework of the model as follows. *Conditionally* on $\mathcal{T}$, the increments ΔX of the log-returns are Gaussian; one could replace the distribution of the increments which follow closely an event of $\mathcal{T}$ (those that are most likely to be large) by a skew "deformation" of a Gaussian.

IV. CONCLUSIONS AND FUTURE WORKS

We have introduced a stochastic model that, on the one hand, reproduces relevant stylized facts of several financial indexes, and, on the other hand, is simple enough to allow estimation and simulations. Further current studies are concentrated of the following points:

- we apply the model to option pricing, as alternative to Black & Scholes and related models; the comparison with market data and with the results obtained form other models, could reveal how our model can detect the essential features of the markets;

- we are considering various modifications of

the model, in particular by considering different *time-change functions* $\theta(\cdot)$;

- we study the asymptotic properties (consistency, normality) of the estimators we use (see section III).

REFERENCES

[1] A. Andreoli, F. Caravenna, P. Dai Pra, G. Posta, *Scaling and multiscaling in financial indexes: a simple model*, preprint, 2010.

[2] T. Ané, H. Geman, *Order Flow, Transaction Clock, and Normality of Asset Returns*, The Journal of Finance **LV** (2000), 2259–2284.

[3] R. T. Baillie, *Long memory processes and fractional integration in econometrics*, J. Econometrics **73** (1996), 5–59.

[4] F. Baldovin, A. Stella, *Scaling and efficiency determine the irreversible evolution of a market*, PNAS **104**, n. 50 (2007), 19741–19744.

[5] 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.

[6] F. Black and M. Scholes, *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy **81**, n.3 (1973), 637-654

[7] T. Bollerslev, *Generalized Autoregressive Conditional Heteroskedasticity*, J. Econometrics **31** (1986), 307–327.

[8] T. Bollerslev, H.O. Mikkelsen, *Modeling and pricing long memory in stock market volatility*, J. Econometrics **31** (1996), 151–184.

[9] P. Carr, H. Geman, D.B. Madan, M. Yor, *Stochastic volatility for Lévy processes*, Mathematical Finance **13**, n. 3 (2003), 345–382.

[10] P. Carr, L. Wu, *T*ime-changed Lévy processes and option pricing, Journal of Financial Economics **71** (2004), 113141

[11] 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.

[12] R. Engle, *Estimates of the Variance of U.S. Inflation Based on the ARCH Model*, J. Money, Credit and Banking **15** (1983), 286–301.

[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] T. Kokholm, E. Nicolato, *Sato Processes in Default Modeling*, Applied Mathematical Finance, to appear.

[16] R. Merton, *Theory of Rational Option Pricing*, Bell Journal of Economics and Management Science 4 n.1 (1973), 141-183.

[17] A. L. Stella, F. Baldovin, *Role of scaling in the statistical modeling of finance*, Pramana **71** (2008), 341–352.