# Continuous cascade models for asset returns

E. Bacry\* and A. Kozhemyak<sup>†</sup> CMAP, Ecole Polytechnique, 91128 Palaiseau, France

Jean-François Muzy

CNRS UMR 6134, Université de Corse, Quartier Grossetti, 20250, Corte, France<sup>‡</sup>

(Dated: April 28, 2006)

In this paper, we make a short overview of continuous cascade models recently introduced to model asset return fluctuations. We show that these models account in a very parcimonious manner for most of "stylized facts" of financial time series. We review in more details the simplest of such models namely the log-normal Multifractal Random Walk. It can simply be considered as a stochastic volatility model where the (log-) volatility memory has a peculiar "logarithmic" shape. This model possesses some appealing stability properties with respect to time aggregation. We describe how one can estimate it using a GMM method and we present some applications to volatility and VaR forecasting.

## I. INTRODUCTION

The modelling of random fluctuations of asset prices is of primary importance in finance. These flucuations turn out to reveal a very rich and non trivial statistical structure, that is to some degree universal across different assets (stocks, stock indices, currencies), regions (U.S., European, Asian) and epochs. Since Mandelbrot famous work on the fluctuations of cotton price in early sixties, it is well known that speculative price variations are poorly described by the standard "Bachelier" (geometric) Brownian motion (see e.g., [1, 2]). If return variations are, to a good approximation, uncorrelated, extreme events are more probable than in a Gaussian world and volatility fluctuations are well kwown to be of intermittent and correlated nature. This feature is known as volatility clustering [2–5]. The quest for more complex, non-Gaussian, models of asset returns aiming at capturing these "statistical" anomalies more accurately has a long history and the number of existing econometric models, continuous or discrete, from the famous GARCH family to stochastic volatility models, is wide. Our aim is to describe a class of models relying upon concepts originally introduced to describe the energy cascade in fully developed turbulence.

During the last decade, the availability of huge data sets of high frequency time series has permitted intensive statistical studies that lead to uncover new features. Many recent empirical studies have suggested that financial data share many statistical properties with turbulent velocity "intermittent" fluctuations and notably display multiscaling properties [2, 6]. Moreover the return probability density functions strongly depend on the considered time-scale. These pdf have a strong kurtosis at small scale and several studies suggest that they can be characterized by powerlaw tails. In that respect, as we shall discuss below, the phenomenology of multifractal models has provided new concepts and tools to analyze market fluctuations and inspired a particularly elegant family of models that accounts for the main observed statistical properties in a parcimonious way. Moreover these models are amenable to many analytical computations: They are easy to estimate and they lead to simple solutions to the problem of conditional risk forecasting. Such models were first introduced in the field of empirical finance by the pioneering work of Calvet, Fisher and Mandelbrot with their "Multifractal Model of Asset Returns" [7–9]. In order to improve the MMAR model, Calvet and Fisher proposed a "grid free" Poisson cascade process. In a series of econometric papers, these authors developed analytical methods to estimate their model and to forecast return volatility and risk [10-12]. The goal of this paper is to present some more recent models that belong to the same family of multifractal models and that rely on the concept of continuous cascade. We will show that continuous cascades naturally extend discrete constructions and constitute a family of parcimonious models with nice features with respect to time aggregation. We will focus on the simplest of such cascades, namely the log-normal multifractal random walk that is rich enough to reproduce faithfully most of empirical return statistical properties.

The paper is organized as follows: in section II we review the main mathematical concepts related to the notions of multifractals, multiscaling and self-similarity. In section III we introduce continuous cascade models as a natural

<sup>\*</sup>Electronic address: emmanuel.bacry@polytechnique.fr

 $<sup>^\</sup>dagger ext{Electronic address: alexey@cmapx.polytechnique.fr}$ 

<sup>‡</sup>Electronic address: muzy@univ-corse.fr

extension of discrete cascades. We then define the Multifractal Random Walk (MRW) as a stochastic volatility model for which the volatility is the simplest continuous cascade corresponding to log-normal statistics. Its main mathematical properties are briefly reviewed. A short comparison to former approaches is proposed. In section IV we consider most of empirical properties of asset returns under the light of the cascade paradigm. More precisely, the framework of multifractal analysis allows one to account for the statistical properties of return variations (moments, probability distributions) over a wide range of time-scales, the volatility correlations for various volatility definitions, the probability of extreme events and the volatility memory dynamics. In section V we address the problem of parameter estimation of a log-normal cascade and show that a simple GMM procedure can be defined. In section VI we provide some simple methods to estimate conditional volatility and Value at Risk. The so-obtained forecasts are compared to the one obtained using the classical GARCH(1,1) model. Conclusion and prospects for other problems of quantitative finance are reported in section VII. Some technical details that concern the important problem of the stability of statistics with respect to time aggregation are provided in the Appendix.

## II. MULTISCALING AND SELF-SIMILARITY

In this section we define precisely what is a multifractal process and how this notion is related to the property of self-similarity. All these concepts have been originally defined in physics for the study of fully developed turbulence (see e.g. [13] for a review). They have been first used in empirical finance by Calvet, Fisher and Mandelbrot [7–9].

## A. Multiscaling of return moments

Let p(t) be the price of some asset at time t and let us define  $\delta_l X(t)$  as the variation of the continuous compound return  $X(t) = \ln[p(t)]$  between times t and t+l:  $\delta_l X(t) = X(t+l) - X(t)$ . In the sequel, we shall consider that  $\delta_l X(t)$  is stationary and has zero mean. In particular, we do not consider seasonal effects such as the intraday modulation of the market activity.

Let us denote by M(q, l) the order q absolute moment of  $\delta_l X(t)$ ,

$$M(q,l) = \mathbb{E}\left[ |\delta_l X(t)|^q \right] . \tag{1}$$

where  $\mathbb{E}[.]$  stands for the mathematical expectation. We will say that the process is *scale-invariant*, if the scale behavior of the absolute moment M(q, l) is an exact power-law. Let us call  $\zeta(q)$  the exponent of this power law, i.e.,

$$M(q,l) = C_q l^{\zeta(q)},\tag{2}$$

where the prefactor  $C_q$  is simply the order q moment at scale l=1. When the exponent  $\zeta(q)$  is a linear function of q, i.e.,  $\exists H, \zeta(q) = qH$ , the process is referred to as a monofractal process. Let us note that the so-called self-similar processes (e.g., Brownian motion, fractional Brownian motions,  $\alpha$ -stable processes) is a particular case of monofractal processes (see next section). On the contrary, if  $\zeta(q)$  is a non-linear function of q, it is referred to as a multifractal process (or as a process displaying multiscaling or intermittency). It is important to remark that from the convexity of the moments of a random variable,  $\zeta(q)$ , as defined from the scaling behavior (2) is necessarily a concave function of q. It results that if  $\zeta(q)$  is strictly concave, then the multiscaling property cannot hold at arbitrary large scales. In general we will denote T the time scale above which the process ceases to be multifractal. This scale will be called the integral scale:

$$M(q,l) = C_q l^{\zeta(q)}, \quad \forall l \le T.$$
 (3)

One will refer to this property as a continuous scaling property. A discrete scaling property corresponds to the case this last equation only holds for some discrete sequence  $l_n$  that converges towards 0, i.e.,

$$M(q, l_n) = C_q l_n^{\zeta(q)}, \quad \forall l_n \le T, \tag{4}$$

where  $l_n \to 0$  when  $n \to +\infty$ .

## B. Self-similarity and "trading time"

1. Random self-similarity and intermittency

The multifractality as defined previously can be interpreted using a stronger assumption related to the classical notion of self-similarity. First of all, let us define what is a self-similar stochastic process. In the mathematical

literature [14], a process X(t) is called self-similar of exponent H if it has stationary increments and if  $\forall s > 0$ ,

$$X(t) = s^{-H}X(st) , (5)$$

meaning that the original process and a dilated version of it are somehow undistinguishable. According to this definition, the Brownian motion is self-similar with an exponent H = 1/2. Since X has stationary increments, if one assumes that X(0) = 0, the law of  $\delta_l X(t)$  is the same as the law of X(t) and the statement (5) is equivalent to

$$\delta_{sl}X(st) = s^H \delta_l X(t) \tag{6}$$

This definition is however too restrictive for our purpose since it excludes multifractal processes. Indeed, let us consider  $P_l(\delta X)$  the probability density function (pdf) of  $\delta_l X(t)$  If X(t) is self-similar with an exponent H, Eq. (6) directly yields:

$$P_l(\delta X) = s^H P_{sl}(s^H \delta X) . (7)$$

The moments at scale L and l = sL are related by

$$M(q,l) = \left(\frac{l}{L}\right)^{qH} M(q,L), \tag{8}$$

X(t) is therefore 'monofractal" with  $\zeta(q) = qH$ . In order to account for multifractality, one has to generalize this classical definition of self-similarity. This can be done by generalizing Eq. (6) by considering that the value of  $s^H = W_s$  is stochastic:

$$\delta X_{sl}(st) = W_s \delta X_l(t) \tag{9}$$

where  $W_s$  is a positive random factor independent of X and which law depends only on the scale ratio s. Such equation can be seen as a *random cascade* since one goes from a coarse scale to a finer one by multiplication by a random factor  $W_s$ . In the next section we will study explicit random cascade constructions that satisfy Eq. (9).

Let us consider a coarse scale L and a smaller scale l = sL (s < 1). If one denotes  $G_s(x)$  the pdf of  $\ln(W_s)$ , then Eq. (9) implies that the return pdf's at scales L and l are related by the transformation:

$$P_l(\delta X) = \int G_s(u)e^{-u}P_L(e^{-u}\delta X)du , \qquad (10)$$

This equation is a direct stochastic generalization of Eq. (7) and has been introduced in the field of fully developed turbulence by B. Castaing and co-authors [15]. It can be simply interpreted as follows: The return probability density at small scale,  $P_l(\delta X)$ , can be written as a weighted superposition of the rescaled versions of the coarse scale pdf,  $e^{-u}P_L(e^{-u}X)$ , the self-similarity kernel  $G_s(u)$  being the weight associated with each value of u. In the case of a "classical" self-similar process (Eq. (7)),  $G_s(u)$  is simply a Dirac  $\delta$ -function,  $P_l$  and  $P_L$  have the same shape and differ only by the scale factor  $e^{-u} = (l/L)^H$ . In the general case, the shapes of the pdf  $P_l$  across scales are no longer the same: When going to small scale the pdf becomes strongly leptokurtic (see Refs. [15] and fig. 4).

Let us now link  $G_s$ , the law of the weight logarithms, to the multifractal exponent spectrum  $\zeta(q)$ . Let  $F_s(k) = \ln[\int G_s(u)e^{iku}du]$  be the cumulant generating function associated with  $G_s(u)$ . A simple semi-group argument, shows that one must have  $F_s(k) = F(k)\ln(s)^{-1}$ . In that case, from Eq. (10), the q order absolute moments at scales l and L are related as:

$$M(q,l) = \hat{G}_s(-iq)M(q,L) = M(q,L)\left(\frac{l}{L}\right)^{F(-iq)}, \qquad (11)$$

Comparison with Eq. (3) leads to  $C_q = M(q, L)$ , the return moments at large scale L, while  $\zeta(q) = F(-iq)$ . The multifractal scaling is therefore directly related to the random character of the self-similarity factor when one goes from large to small scales. A nonlinear  $\zeta(q)$  spectrum implies that F is non linear and thus that  $G_s$  differs from a Dirac distribution. The simplest measure of the random character of  $\ln W_s$  is the width of its distribution, i.e., its variance  $F''(0) \ln(s)$ . Thanks to the relation between  $\zeta(q)$  and F(k), one defines the intermittency coefficient as the curvature of  $\zeta(q)$  around q = 0:

$$\lambda^2 = -\zeta''(0) \tag{12}$$

The simplest non linear case is the so-called log-normal model that corresponds to a parabolic  $\zeta(q)$  function and thus to a function  $G_s$  that is a normal distribution (see below).

<sup>&</sup>lt;sup>1</sup> This implies that  $ln(W_s)$  is infinitely divisible

## 2. "Trading time" and multifractal volatility

Let us consider  $\theta(t)$ , a non-decreasing multifractal process satisfying Eq. (9) and B(t) a self-similar process of exponent H (satisfying Eq. (6)) that is independent of  $\theta(t)$ . Following an idea of Mandelbrot [1, 7] one can interpret  $\theta(t)$  as a "multifractal time" and define the compound process:

$$X(t) = B[\theta(t)]$$

If is easy to show that X(t) is itself multifractal. Indeed, since  $\theta(st) =_{Law} W_s \theta(t)$  and  $B(st) =_{Law} s^H B(t)$ , one has, thanks to the independence of B(t) and  $\theta(t)$ ,

$$X(st) = W_s^H X(t)$$

Hence, if  $\zeta_{\theta}(q)$  and  $\zeta_{X}(q)$  are respectively the multifractal exponent spectra of  $\theta(t)$  and X(t), they are simply related as:

$$\zeta_X(q) = \zeta_\theta(qH) \tag{13}$$

Let us remark that if B(t) is a Brownian motion, then  $\delta_l X(t) = B(\theta(t+l)) - B(\theta(t)) =_{Law} [\delta_l \theta(t)]^{1/2} \delta_1 B(t)$  and thus the quantity  $\theta(t+l) - \theta(t)$  can be interpreted as the *stochastic volatility* of  $B(t+\tau) - B(t)$ . A "natural" model for return fluctuations is thus to consider a Brownian motion compound with a multifractal time  $\theta(t)$ . In this context  $\theta(t)$  is often called the *trading time*.

#### III. CONTINUOUS CASCADES AND THE MRW MODEL

#### A. From Mandelbrot cascades to continuous cascades

Since, as seen previously, one can simply built a multifractal model for return fluctuations from a multifractal trading time, let us first focus on explicit constructions of non decreasing multifractal processes. Such processes are called multifractal measures. The paradigm of multifractal measures are multiplicative cascades originally introduced by the russian school for modelling the energy cascade in fully developed turbulence. After the early works of Mandelbrot [16–18], a lot of mathematical studies have been devoted to random cascades [19–23]. Very recently, continuous versions of these processes have been defined: they share most of the original properties however they display continuous scaling (i.e., Eq. (3)) and possess stationary increments [24–28], whereas original multifractal cascades only display discrete scaling and do not possess stationary increments.

#### 1. "Grid bounded" cascades

The large integral scale T, below which the multiscaling (3) holds can be defined as the scale where the cascading process "starts". The simplest multifractal cascade can be constructed as follows: one starts with an interval of length T where the measure is uniformly spread (meaning that the density is constant) and split this interval in two equal parts: On each part, the density is multiplied by (positive) i.i.d. random factors W. Each of the two sub-intervals is again cut in two equal parts and the process is repeated infinitely. At construction step n, if one addresses a dyadic interval of length  $T2^{-n}$  be a kneading sequence  $k_1 \dots k_n$ , with  $k_i = 0, 1$ , the measure of this interval (denoted as  $I_{k_1 \dots k_n}$ ) is simply:

$$\theta_n(I_{k_1...k_n}) = 2^{-n} \prod_{i=1}^n W_{k_1...k_i} = 2^{-n} e^{\sum_{i=1}^n \delta \omega_{k_1...k_i}}$$
(14)

where all the  $W_k = e^{\delta \omega_k}$  are i.i.d such that  $\mathbb{E}[W] = 1$ . Since Peyrière and Kahane [19], it is well known that the previous construction converges almost surely towards a stochastic measure  $\theta_{\infty}$  provided  $\mathbb{E}[W \ln W] < 1$ . The multifractality of  $\theta_{\infty}$  (hereafter simply denoted as  $\theta$ ) directly results from its recursive construction. Indeed, from the previous definition, it is easy to show that  $\theta$  is self-similar in the generalized sense as in Eq. (9):

$$\theta[I_{k_1...k_n}] = 2^{-1} W \theta[I_{k_1...k_{n-1}}]$$
(15)

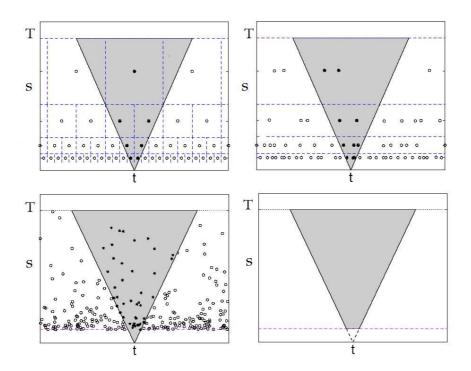


FIG. 1: From discrete to continuous cascades: (Top-Left) Mandelbrot Cascade lying on the dyadic tree. (Top-Right) "semi-continuous" Poisson construction obtained by "randomizing" the node positions at fixed scale. (Bottom-Left) Barral-Mandelbrot MPCP obtained by replacing the tree by a "random Poisson tree". (Bottom-Right) Log-infinitely divisible Bacry-Muzy construction.

If  $I_n = [t_n, t_n + l]$  is a short notation for dyadic intervals of size  $l = T2^{-n}$ , then the order q moments of  $\delta_l \theta(t_n) = \theta[I_n]$  behave has power-law:

$$\mathbb{E}\left[\theta[I_n]^q\right] = 2^{-nq} \mathbb{E}\left[W^q\right]^n \mathbb{E}\left[\theta([0,T])^q\right] \tag{16}$$

Comparison of Eqs. (16) and (4), with  $l_n = T2^{-n}$  directly yields the expression of the spectrum  $\zeta(q)$  in terms of cumulant generating function of  $\delta\omega = \ln W$ :

$$\zeta(q) = q - \ln_2(\mathbb{E}\left[W^q\right]) = q - \ln_2(\mathbb{E}\left[e^{q\delta\omega}\right]) \tag{17}$$

### 2. Continuous cascades

The previous cascade construction, though simple, does not provide a satisfying solution to model volatility fluctuations. Indeed, it is build on a fixed time interval [0, T], it is not causal and not stationary. Moreover, it involves an arbitrary fixed scale ratio s = 2. Very recently, several constructions have been proposed to generalize Mandelbrot cascades to stationary, causal continuous cascades. The idea on which such generalizations rely is illustrated in fig. 1. One starts with the discrete construction (fig. 1(a)): It can be conveniently represented is some 2D half-plane (t, s) where the parameter s can be identified as a time scale. If one associates at construction step n, with each dyadic interval  $I_n$  of size  $T2^{-n}$ , located at  $t_n = kT2^{-n}$ , a point  $(t_n, s_n = T2^{-n})$ , one constructs a dyadic tree as illustrated in fig. 1(a). The measure  $\theta(t)$  at some time t is roughly obtained as the products of weights  $W_i$  associated with each point inside a cone-like domain C(t) represented in the figure:

$$d\theta(t) = \prod_{(t_j, s_j) \in C(t)} W_j = e^{\sum_{(t_j, s_j) \in C(t)} \delta\omega_j}$$
(18)

The non stationarity of this construction appears immediately as being associated with the fixed dyadic grid corresponding to successive refinements of the interval [0,T]. A natural way to obtain a stationary model is to

replace, at each scale  $s_n = T2^{-n}$  the periodic grid by points located randomly according to a Poisson process with a rate  $r_n$  that is precisely  $r_n = s_n^{-1}$ . We will refer to this type of construction as the "semi-continuous" Poisson construction. It is illustrated in fig. 1(b) and corresponds roughly to the "Poisson Multifractal Model" proposed by Calvet and Fisher [10–12] (see sec. III D). One can however go a step further and use a "fully continuous" Poisson construction: instead of keeping the scales  $s_n$  at exact values  $T2^{-n}$ , one can draw the whole grid randomly, over the plane (t, s) using a non-homogeneous Poisson process with rate  $r(s) = s^{-1}$ . Then one associates with each Poisson point  $(t_i, s_i)$  an independent weight  $W_i$  and one can build the measure  $\theta$  according to (18); one then obtains exactly the "Multifractal Product of Cylindrical Pulses" (MPCP) introduced by Barral and Mandelbrot [25]. Finally, there is a last possible extension if one considers limits of products of MPCP: This amounts replacing the compound Poisson density by some arbitrary infinitely divisible random density in the plane (t, s). In that case, the discrete sum over Poisson points, in (18), is replaced by a stochastic integral over the cone-like domain C(t):

$$d\theta(t) = e^{\int_{(t',s)\in C(t)} d\omega(t',s)} \tag{19}$$

This construction, that involves the concept of "independently scattered random measure"  $d\omega(t,s)$ , has been proposed by Bacry and Muzy [26, 27] and allows one to build stationary random cascades with continuous scale invariance properties and with a multifractal spectrum  $\zeta(q)$  that can be associated with an arbitrary infinitely divisible law. The precise description of this construction is beyond the scope of the paper and we refer the reader to the cited references for more details. In this paper we will rather focus on the simplest of infinitely divisible law, namely the normal law, that corresponds to the case when  $d\omega(t,s)$  is a Gaussian white noise.

#### B. The Multifractal Random Walk model for asset returns

1. The log-normal multifractal trading time

Let us describe precisely the continuous cascade construction in the log-normal case. In this case the non decreasing "trading time" function  $\theta(t)$  is obtained as

$$\theta(t) = \lim_{\Delta \to 0} \int_0^t e^{2\omega_{\Delta}(u)} du \tag{20}$$

where  $\omega_{\Delta}(u)$  is called the *magnitude* processs. It is defined as the (stochastic) integral of a Gaussian white noise dW(t,s) over the cone C(t) in the (t,s) plane truncated at the large scale s=T and the small scale  $s=\Delta^{-2}$ :

$$\omega_{\Delta}(u) = \int_{\Delta}^{T} \int_{u-s}^{u+s} dW(v,s) \tag{21}$$

In order to reproduce the density of cascade points, the covariance of dW(v,s) is

$$\operatorname{Cov}\left(dW(v,s), dW(v',s')\right) = \lambda^2 \delta(v-v') \delta(s-s') \frac{dv ds}{s^2}$$

and for convergence purposes its mean is not zero but

$$\mathbb{E}\left[dW(v,s)\right] = -\lambda^2 \frac{dvds}{s^2}$$

Using these definitions, one can prove that, the limit (20) exists almost surely and, up to a low frequency component [26, 27], possesses exact multifractal properties as detailed below.

Let us mention that the trading time  $\theta(t)$  is a causal process relative to the filtration  $F_t$  generated by the values taken by the random scattered measure dW in all the "past" cones  $\cup_{t' \leq t} C(t')$ . This is even true in the general log-infinetely divisible case. In the log-normal case, one can go a step further. Indeed,  $\omega_{\Delta}$  is a 1D stationary Gaussian process, it is fully defined by its mean value and covariance function which can be computed analytically and that is

<sup>&</sup>lt;sup>2</sup> Actually, in order to get the exact continuous scaling property Eq. (3) (and not just an asymptotic one when  $l \to 0$ ), one should truncate the cone differently at large scales. One should use the cone defined by  $C_1(t) = \{(t',s) \mid (t',s) \in C(t) \text{ or } s \ge T \text{ and } |t'-t| \le T/2\}$ . See [26, 27] for more details.

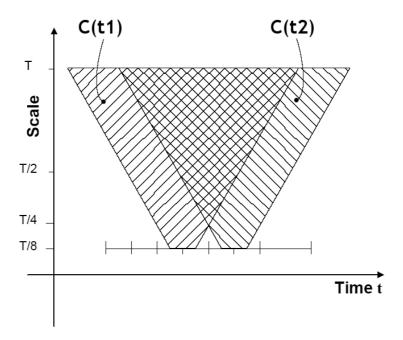


FIG. 2: The magnitude covariance in continuous cascades: In continuous cascade construction the covariance of  $\omega_{\Delta}(t_1)$  and  $\omega_{\Delta}(t_2)$  (for  $\Delta = T/8$ ) is proportional to the area of the intercept between the two cones  $C(t_1)$  and  $C(t_2)$ 

compact support (in [-T, T]). Thus it can be generated directly, i.e., without going through a 2D white noise (and this is of course much simpler). The corresponding mean value and covariance function will be detailed in the next section. Moreover, let us point out that the definition of the trading time  $\theta(t)$  at time t (Eq. (20) involves the process  $\omega_{\Delta}(u)$  at time t (Eq. (20) involves the process a causal filter to exist. Thus, the process  $\omega_{\Delta}$  is a causal gaussian process, in the sense that it is obtained by filtering a 1D white noise using a causal filter. Thus the trading time  $\theta(t)$  can be defined as a limit of a causal function of a causal 1D gaussian process.

## 2. The Multifractal Random Walk Model as a stochastic volatility model

If  $\theta(t)$  is the log-normal multifractal time built as in previous section, one can define a continuous time diffusion model for asset returns as

$$X(t) = B[\theta(t)]$$

where B(t) is a Brownian motion independent of  $\theta(t)$ .

An equivalent formulation of this model can be constructed as a stochastic volatility model. More precisely, the Multifractal Random Walk (MRW) model is constructed as the following limit

$$X(t) = \lim_{Law} \Delta_{\Delta \to 0} X_{\Delta}(t) \tag{22}$$

where the limit has to be interpreted as an equality for all finite dimensional distributions and  $X_{\Delta}(t)$  is defined as the following, step-wise constant, stochastic process:

$$X_{\Delta}(t) = \sum_{i=0}^{\lfloor t\Delta^{-1} \rfloor} \epsilon_i e^{\omega_{\Delta}(i\Delta)}$$
 (23)

where  $\lfloor x \rfloor$  stands for greatest integer smaller than x,  $\epsilon_i$  is a Gaussian white noise  $\mathcal{N}(0, \sigma^2 \Delta^{-1})$  and  $\omega_{\Delta}(t)$  is the Gaussian process previously defined (Eq. (21)).

As explained in the last part of the previous section,  $\omega_{\Delta}(t)$  can be directly defined as the Gaussian stationary process with mean

$$\mathbb{E}\left[\omega_{\Delta}(t)\right] = -\lambda^2 \ln\left(\frac{T}{\Delta}\right) \tag{24}$$

and covariance  $^3$ 

$$\mathbb{C}\text{ov}\left(\omega_{\Delta}(t), \omega_{\Delta}(t+\tau)\right) = \begin{cases}
\lambda^{2} \left(\ln\left(\frac{T}{\Delta}\right) + 1 - \frac{\tau}{\Delta}\right) & \text{if } \tau \leq \Delta \\
\lambda^{2} \ln\left(\frac{T}{\tau}\right) & \text{if } T \geq \tau \geq \Delta \\
0 & \text{if } \tau > T
\end{cases} ,$$
(25)

As illustrated in fig. 2, this last expression is simply the area (measured with the density  $s^{-2}dtds$ ) of the intercept of the two cones defining  $\omega_{\Delta}(t)$  and  $\omega_{\Delta}(t+\tau)$ . This logarithmic shape of the magnitude correlation is reminiscent of the 'tree' (i.e 'ultrametric') structure explicitly involved in the Mandelbrot dyadic construction [29].

This construction will be referred to as the *Multifractal Random Walk* (MRW) model and inherits its multiscaling properties directly from those of  $\theta(t)$ . It is important to notice that this model has only two parameters, namely the *integral scale T* that corresponds to the maximum volatility correlation time and the *intermittency coefficient*  $\lambda^2$ , related to the variance magnitude, that somehow quantifies the strength of the multifractality. A simple method to estimate these parameters will be described in section V.

## C. Scaling properties and the singularity spectrum of the MRW model

Let us briefly review the main mathematical properties of both  $\theta(t)$  and the MRW process X(t) defined in the previous sections (for proofs and details sees refs [24–27, 30]).

1. Self-similarity and multiscaling of moments

First of all,  $\theta(t)$  verifies the exact stochastic self-similarity property (9), i.e., for s < 1, t < T,

$$\theta(st) = se^{2\Gamma_s}\theta(t) \tag{26}$$

where  $\Gamma_s$  is a Gaussian random variable  $\mathcal{N}(\lambda^2 \ln(s), -\lambda^2 \ln(s))$  independent of  $\theta(t)$ . It results that its moments at scale l behave exactly as a power-law of the time scale  $(l \leq T)$ :

$$M_{\theta}(q, l) = \mathbb{E}\left[\left[\theta(t + l) - \theta(t)\right]^{q}\right] = C_{q} \left(\frac{l}{T}\right)^{\zeta_{\theta}(q)}$$
(27)

where the spectrum  $\zeta(q)$  has the log-normal expression:

$$\zeta_{\theta}(q) = q(1+2\lambda^2) - 2\lambda^2 q^2 \tag{28}$$

The constants  $C_q$  in previous equality can be interpreted as the moments at scale l = T, i.e. the moments of  $\theta(T)$  and can be computed exactly [24]. Moreover, it can be shown that if

$$q > q_c = \frac{2}{\lambda^2} \tag{29}$$

then  $C_q = \infty$ . This means that the law of  $\theta(t)$  is algebraic (fat tailed) with an exponent  $\mu = 2/\lambda^2$ , i.e., for t < T,

$$\mathbb{P}\left[\theta(t) > x\right] \underset{t \to +\infty}{\sim} x^{-\frac{2}{\lambda^2}}$$

In section IV C, we will comment on this result.

$$\omega_{\Delta}(u) = \int_{\Delta}^{+\infty} \int_{u-min(s,T)}^{u+min(s,T)} dW(v,s).$$

This process is the one that will be used all along the paper from now on

<sup>&</sup>lt;sup>3</sup> the mean value and covariance function given here corresponds actually to the definition of the process  $\omega_{\Delta}(t)$  using the cone  $C_1(t)$  defined in the footnote 2 instead of the cone C(t). Thus Eq. (21) is replaced by

All previous properties of  $\theta(t)$  can be directly translated to scaling properties of  $X(t) = B[\theta(t)]$ : X(t) is exactly self-similar,

$$\delta_{sl}X(st) \underset{Law}{=} s^{\frac{1}{2}}e^{\Gamma_s}\delta_lX(t) \tag{30}$$

where  $\Gamma_s$  is defined as previously and is independent of X. It follows that

$$M(q,l) = \mathbb{E}\left[|X(t+l) - X(t)|^q\right] = K_q \left(\frac{l}{T}\right)^{\zeta(q)}$$
(31)

with

$$\zeta(q) = \zeta_{\theta}(q/2) = q(1/2 + \lambda^2) - \frac{\lambda^2}{2}q^2$$
 (32)

It can be proved that the pdf of X(t) has also heavy tails and behaves has:

$$\mathbb{P}\left[|X(t+l) - X(t)| > x\right] \underset{\tau \to +\infty}{\sim} x^{-\frac{1}{2\lambda^2}} \tag{33}$$

## 2. Quasi-stability with respect to aggregation

One of the nice features of standard Brownian motion is stability with respect to time aggregation: At each scale, the return probability distributions remain Gaussian. Eqs (26) and (30) state that both the multifractal time  $\theta(t)$  and the MRW process X(t) are self-similar. This means that, in some sense, they possess stable properties when changing the time-scale. However these formulae, even if they are valid for all finite dimensional distributions, are of poor practical interest because they do not provide the laws at a given time scale l but simply indicate how these laws may change as l varies (Eq. (10)).

Let us first remark that statistics of the returns at scale l,  $\delta_l X(t)$  have, a priori, no reason to correspond to those of the underlying "microscopic" model  $\epsilon_i e^{\omega(t_i)}$  (Eqs. (22) and (23)) used to construct the MRW model. However, in Appendices 1 and 2, we provide results that show that, in a good approximation, when  $\lambda^2 \ll 1$ , the returns of the MRW process, can be written, in law, as (the precise meaning of the  $\simeq$  sign used in the following equation is explained in Appendices 1 and 2)

$$\delta_l X(t) \underset{Law}{\simeq} \epsilon(t) e^{\Omega_l(t)}$$
 (34)

for all  $l \leq T$ , where  $\epsilon_l$  is a Gaussian white noise of variance  $\sigma^2 l$  and the process  $\Omega_l(t)$  is a renormalized magnitude (or dressed magnitude) that is the Gaussian process which mean and covariance are the following (see Appendix):

$$\mathbb{E}\left[\Omega_l\right] = -\lambda^2 \ln\left(\frac{Te^{3/2}}{l}\right) \tag{35}$$

$$Cov(\Omega_l(0), \Omega_l(nl)) = \lambda^2 \left[ ln \left( \frac{Te^{3/2}}{l} \right) + g(n) - \frac{1}{2} (g(n+1) + g(n-1)) \right]$$
(36)

where

$$g(n) = n^2 \ln |n|$$

Notice that, for large lags, one recovers the log-correlated magnitudes of the microscopic model:

$$\operatorname{Cov}\left(\Omega_l(0), \Omega_l(nl)\right) \underset{n \gg 1}{\sim} \lambda^2 \ln\left(\frac{T}{nl}\right) \tag{37}$$

We refer the reader to the Appendices 1 and 2 for more details on what Eq. (34) means.

Therefore, whatever the aggregation scale l, the return laws of the MRW process are of the same form as the underlying microscopic stochastic volatility model. In that sense, the MRW model possesses some *quasi-stability* property with respect to time aggregation. These formula are of great practical interest. They are particularly relevant for estimation purposes (sec. V) and risk forecasting (sec. VI).

#### 3. Volatility correlation functions

Many other scaling laws can be computed and notably those of volatility correlations. In refs. [24, 30] it is shown that if

$$C(l,\tau) = \mathbb{E}\left[\delta_l X(\tau)^2 \delta_l X(0)^2\right] \tag{38}$$

then for  $0 \le \tau < T$ ,  $0 \le \tau + l < T$ , we have,

$$C(l,\tau) = K(|l+\tau|^{2-4\lambda^2} + |l-\tau|^{2-4\lambda^2} - 2|\tau|^{2-4\lambda^2})$$
(39)

where

$$K = \frac{\sigma^4 T^{4\lambda^2}}{(1 - 4\lambda^2)(2 - 4\lambda^2)}.$$

Such formulae could be useful for the computation of the optimal Wiener filter in the linear volatility estimation or forecasting problem [31].

Let us note that in the usual case  $0 \le l \ll \tau$ , one gets

$$C(l,\tau) \simeq \sigma^4 l^2 \left(\frac{\tau}{T}\right)^{-4\lambda^2}$$
 (40)

i.e., for fixed l, the volatility correlation function scales as the power-law

$$C(\tau) \sim \tau^{-2\nu} \tag{41}$$

with  $\nu = 2\lambda^2$ 

It is interesting to extend the previous computation of the correlation function to the power of returns  $|X(t+l) - X(t)|^p$ . Indeed some authors, in the econometrics literature, refer to other proxys for volatility and several empirical works have concerned the study of such "generalized volatilities". People often noticed variations of amplitude of the correlation and of the power-law exponent  $\nu_p$  when varying the order p (see refs. [3, 4]). In Ref. [30] (see also Appendix), it is shown that the quantity,

$$C_p(l,\tau) = \mathbb{E}\left[ |\delta_l X(\tau)|^p |\delta_l X(0)|^p \right] , \tag{42}$$

behaves, when l is small enough, as a power-law

$$C_p(l,\tau) \sim K_p^2 \left(\frac{l}{T}\right)^{2\zeta_p} \left(\frac{\tau}{T}\right)^{-\lambda^2 p^2}$$
 (43)

when the constant  $K_p$  can be computed analytically. One sees that for multifractal processes, the "long range memory" of the volatility depends on the definition one chooses for volatility.

## 4. Mutifractal formalism

Let us end this review by briefly describing the so-called multifractal formalism. This formalism has been introduced in the early eighties by Parisi and Frisch (see e.g. [13]) in order to interpret the above multiscaling properties of the moments in terms of pointwise regularity properties of the paths of the process  $\theta(t)$ . In this section, we only describe the multifractal formalism for the trading time  $\theta(t)$ . All the considerations below naturally extend to analog results for the MRW paths X(t). For sake of simplicity, we do not mention the reference to  $\theta$  in the function  $\zeta_{\theta}(q)$  and simply denote it, in this section,  $\zeta(q)$ .

Let us introduce the local Hölder exponent  $\alpha(t_0)$  at point (or time)  $t_0$  as

$$\delta_l \theta(t_0) = \theta(t_0 + l) - \theta(t) \underset{l \to 0}{\sim} l^{\alpha(t_0)}$$
(44)

The limit  $l \to 0$  means  $l \ll T$  where T is the integral scale. The singularity spectrum  $f^*(\alpha)$  can be introduced as the fractal (Haussdorf or Packing) dimension of the iso-Hölder exponents sets:

$$f^{\star}(\alpha) = Dim\{t, \alpha(t) = \alpha\} \tag{45}$$

Roughly speaking, this equation means that at scale  $l \ll T$ , the number of points where  $\delta_l \theta(t) \sim l^{\alpha}$  is

$$N(l,\alpha) \sim l^{-f^{\star}(\alpha)}$$
 (46)

The multifractal formalism states that  $f^*(\alpha)$  and  $\zeta(q)$  as defined in Eq. (3) are basically Legendre transforms of each other. More precisely, if we define  $f(\alpha)$  as the Legendre transform of  $\zeta(q)$ , i.e.,

$$f(\alpha) = 1 + \min_{q} (q\alpha - \zeta(q))$$
  
$$\zeta(q) = 1 + \min_{\alpha} (q\alpha - f(\alpha)),$$

then

$$f^{\star}(\alpha) = f(\alpha), \quad \forall \alpha \in [\alpha_{\star}, \alpha^{\star}]$$
  
=  $-\infty$  otherwise

where  $\alpha_{\star}$  and  $\alpha^{\star}$  are defined by

$$\alpha_{\star} = argmin\{\alpha, f^{\star}(\alpha) = 0\}$$
  
 $\alpha^{\star} = argmax\{\alpha, f^{\star}(\alpha) = 0\}.$ 

The validity of these equalities has been rigorously established for both discrete [21] and continuous cascades [25] In section IV C, we will use the fact that q can be interpreted as a value of the derivative of  $f(\alpha)$  and conversely  $\alpha$  is a value of the derivative of  $\zeta(q)$ : for a given value of  $q = q_0$  one has

$$f(\alpha_0) = 1 + q_0 \alpha_0 - \zeta(q_0)$$

$$\alpha_0 = \frac{d\zeta}{dq}(q_0)$$

$$q_0 = \frac{df}{d\alpha}(\alpha_0).$$
(47)

Let us note that,  $f^*(\alpha)$  can be seen as the Legendre transform of the function  $\zeta^*(q)$  simply defined as

$$\zeta^{\star}(q) = \begin{cases}
\zeta(q) & \text{for } q^{\star} \leq q \leq q_{\star} \\
q/\alpha_{\star} & \text{for } q > q_{\star}, \\
q/\alpha^{\star} & \text{for } q < q^{\star},
\end{cases} ,$$
(48)

where

$$q_{\star} = \frac{df}{d\alpha}(\alpha_{\star}) \tag{49}$$

$$q^* = \frac{df}{d\alpha}(\alpha^*) \tag{50}$$

It is important to point out that, experimentally, under usual conditions, only  $\zeta^*(q)$  (and not  $\zeta(q)$ ) can be estimated (see e.g. refs. [21, 32, 33]). Indeed, if one defines the following partition function at scale l, as an estimator of  $M_{\theta}(q, l)$  (Eq. (27)):

$$S(q,l) = N^{-1} \sum_{k=1}^{N} \delta_l \theta(kl)^q \tag{51}$$

then, almost surely, when  $l \ll T$ ,

$$S(q,l) \sim l^{\zeta_{\star}(q)} \tag{52}$$

In the log-normal case, all these quantities are very easy to compute. If  $\zeta(q)$  is defined as in Eq. (28), then, from Eqs (47), its Legendre transform  $f(\alpha)$  is simply:

$$\alpha(q) = 1 + 2\lambda^2 - 4\lambda^2 q \tag{53}$$

$$f(q) = 1 - 2\lambda^2 q^2 \tag{54}$$

Therefore  $q_{\star}$ ,  $q^{\star}$ ,  $\alpha_{\star}$  and  $\alpha^{\star}$  are simply

$$q_{\star} = \sqrt{\frac{2}{\lambda^2}} \Rightarrow \alpha_{\star} = 1 + 2\lambda^2 - 4\sqrt{2}\lambda$$
 (55)

$$q^{\star} = -\sqrt{\frac{2}{\lambda^2}} \Rightarrow \alpha_{\star} = 1 + 2\lambda^2 + 4\sqrt{2}\lambda \tag{56}$$

## D. Qualitative comparison to other approaches and models

### 1. Similar multifractal models

Let us first compare, on a qualitative ground, the MRW model to some similar asset return models recently defined or studied by several authors. We do not refer here to the *shape* of the multifractal spectrum  $\zeta(q)$  for which a wide number of solutions have been proposed (see sec. IV A) but rather focus on bonafide *models* of asset returns.

As recalled previously, the idea to compound a Brownian motion with a multifractal measure  $\theta(t)$  is due to Mandelbrot, Calvet and Fisher [7]. These authors proposed a simple discrete "grid-bound" cascade (as described in sec. III A 1) for  $\theta(t)$ . Because of the lack of stationarity and causality of this construction, Calvet and Fisher [10] have introduced a generalization called "Poisson multifractal". Without entering into details, they construct a multifractal time as the product:

$$\theta_N(t) = \prod_{i=1}^N \theta_i(t) \tag{57}$$

where each volatility "component"  $\theta_i(t)$  is equal to a random multiplier  $W_k$  drawn from some fixed distribution and renewed at a Poisson rate  $r_0 2^{-i}$ . Calvet and Fisher proved the stochastic convergence  $\theta_N \to \theta$  in the limit  $N \to \infty$  and the asymptotic multiscaling of the limit measure [10]. They also proposed a discretized version of their model which has been specifically studied by Lux [34]. Notice that the Calvet & Fisher "Multifractal Poisson" construction is very close to the "semi-continuous" Poisson construction illustrated in fig. 1 top-right. Indeed, Eq. (57) can be rewritten as

$$\theta_N = e^{\omega_N(t)} = e^{\sum_{i=1}^N \delta\omega_i(t)}$$

In the "semi-continuous" cascade version described in section III A 2, the magnitude component  $\delta\omega_i(t)$  at scale  $2^{-i}$  is the sum of random variables  $\ln(W_k)$  associated with Poisson points within the interval  $[t-2^{-i},t+2^{-i}]$  while in the Calvet & Fisher construction, it is just the value  $\ln(W_k)$  corresponding to the last Poisson point occurring before time  $t+2^{-i}$ . However, since in both cases the Poisson rate is the same and is precisely proportional to  $2^{-i}$ , one expects both constructions to be very close. In that respect, the Barral-Mandelbrot MPCP and the Bacry-Muzy log-infinitely divisible extension, can be considered as generalizations of the "semi-continous" Poisson construction or of the Calvet-Fisher approach that involve only dyadic scales.

Let us stress that like Calvet-Fisher construction [11, 34], the MRW model is the weak limit (when the discretization step goes to zero) of a discrete time model. Moreover, our model benefits from an additional appealing property: it possess some stability property when sampled at various sampling periods (see section III C 2). This allows one to control its properties whatever the considered sampling frequency (intradaily, daily, monthly,...).

## 2. Other econometric models

Multifractal models aim at describing the return statistical properties parcimoniously at all time-scales; this contrasts with classical GARCH or stochastic volatility models which aggregation properties are not easy to control and which parameters may change when changing the sampling period. The advantages of the multifractal models have been already discussed elsewhere, so we will not further discuss this question.

It is interesting to compare the MRW model to the most popular stochastic volatility model defined as:

$$X(n) - X(n-1) = e^{\tilde{\omega}_n} \epsilon_n$$

where  $\tilde{\omega}_n \equiv AR(1)$ , which mean that its  $MA(\infty)$  formulation involves coefficients decreasing exponentially fast.

In the MRW model, at each sampling scale l, the memory of  $\Omega_l$ , as defined in sec. III C 2 is wider since it can be written as a MA(p) process with  $p = \lfloor l^{-1}T \rfloor$  involving a slowly decreasing filter:

$$\Omega_l[nl] = \alpha_0 + \sum_{i=k}^p \alpha_k \xi_{n-k-1} \tag{58}$$

where  $\xi$  is a Gaussian white noise and the filter  $\alpha_k$  behaves, for large k, as [35]:

$$\alpha_k \sim k^{-1/2} \tag{59}$$

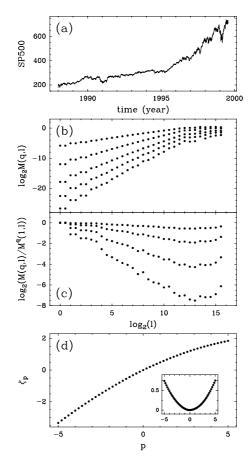


FIG. 3: Multifractal Analysis of the intraday S&P500 future index over the period 1988-1999. (a) Plot of the original index time serie. The analyzed time series is the detrended and de-seasonalized logarithm of this series. (b) Log-log plots of M(q, l) versus l for q = 1, 2, 3, 4, 5. The time scales l range from 10 minutes to 6 years. (c)  $\log_2(M(q, l)/M(1, l)^q)$  for q = 2, 3, 4, 5. Such plots should be horizontal for a process that is not multifractal. (d)  $\zeta(q)$  spectrum for the S&P 500 fluctuations. The plot in the inset is the parabolic nonlinear part of  $\zeta(q)$ .

Finally, let us mention that a GARCH-like process, the so-called HARCH process [36], has been proposed in order to account for multiple time scales in the volatility dynamics. Zumbach and Lynch [37] and, more recently, L. Borland [38, 39] have introduced a long memory generalization of such regressive model that seems to present many of the features of multifractal models. The main difference is that in these models the volatility is a deterministic function of past returns. For a precise comparison of these approaches to the models reviewed in this paper, see [40].

### IV. COMPARISON TO EMPIRICAL DATA

## A. Multiscaling of moments and volatility correlations

As suggested by many recent studies, multifractality as defined in Eq. (3), is likely to be a pertinent concept to account for the price fluctuations in financial time series. First empirical evidences of return moment multiscaling were given in [6] and [8] on FX rates. These pioneering works have been followed by a a very large number of studies that concerned many different markets. We will not make an exhaustive review of these studies in this section but simply provide some key references. Evidences for multiscaling have been reported on FX rates [6, 8, 41], commodity markets [42], stock markets [43], future markets [24, 30], emerging markets [44],...

Let us illustrate the multifractal nature of return variations on the S&P500 index. In fig. 3, the  $\zeta(q)$  function has been estimated for the S&P500 future index over the period 1988-1999. The intraday time series sampled at

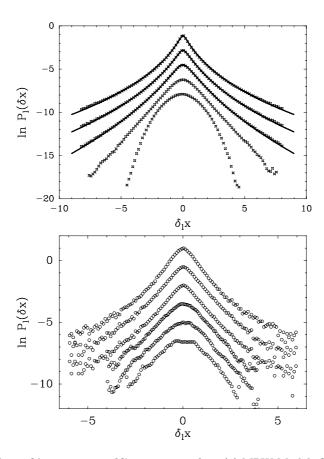


FIG. 4: Continuous deformation of increment pdf's across scales. (a) MRW Model. Standardized pdf's (in logarithmic scale) of  $\delta_l X(t)$  for 5 different time scales, l=16,128,2048,8192,32768. These pdf's have been estimated on 500 MRW realizations of  $2^{17}$  sampled points with  $\lambda^2=0.03$ ,  $\Delta t=1/16$  and T=8192. One can see the continuous deformation and the appearance of fat tails when going from large to fine scales. In solid line, we have superimposed the deformation of the large scale pdf using Castaing's equation (10) with the normal self-similarity kernel. This provides an excellent fit of the data. (b) S&P 500 future. Standardized pdf's at scales (from top to bottom) l=10,40,160 min, 1 day, 1 week and one month. As in fig. (a) the scale is logarithmic and plots have been arbitrarily shifted along the vertical axis for illustration purposes. Notwithstanding the small size of the statistical sample (as compared to (a)), one clearly sees the same phenomenon as for the MRW. The fact that Castaing's equation (10) allows one to describe the pdf's deformation across time scales of financial assets has originally been reported in ref. [6] where similar plots for FX rates can be found.

10mm rate (fig. 3(a)) has been detrended and de-seasonalized <sup>4</sup>. The  $\zeta(q)$  spectrum in fig. 3(c) has been obtained using linear regression fit of "log-log" representation of the behavior of the qth order moment versus the time scale as illustrated in fig. 3(b). In this figure, the scales span an interval from a few minutes to 6 years. From the linear behavior of such curves, one clearly sees that the scale invariance hypothesis is satisfied over nearly 3 decades. In fig. 3(b) we have plotted  $\log_2 \frac{M(q,l)}{M(1,l)^q}$  versus  $\log_2(l)$ . The fact that such plots are not constant reflects the nonlinearity the  $\zeta(q)$  spectrum. The S&P500 future index can thus be considered, at least at this description level, as a multifractal signal. Let us notice that we have computed, in fig. 3(d) the  $\zeta(q)$  values for q-order moments that include negative values of q. This can be achieved using a wavelet based technique that has been introduced in refs. [45, 46]. This spectrum turns out to be well fitted by a parabolic shape  $\zeta(q) = 0.53q - 0.015q^2$ . The non linear parabolic component of  $\zeta(q)$  has been plotted in the inset of fig. 3(d).

In all the above mentionned studies, the authors provide some estimate of the multifractal exponent spectrum  $\zeta(q)$  from the scaling properties of the moments <sup>5</sup>. of returns computed at different scales (as in fig. 3). We will not discuss

<sup>&</sup>lt;sup>4</sup> The amplitude of the return variations in each intraday period is normalized according to the estimated U-shaped intraday r.m.s.

<sup>&</sup>lt;sup>5</sup> Let us notice that, the estimation of the multifractal parameters are less reliable using the moment scaling than using the GMM estimation method detailed in section V for the log-normal model

the analytic shape that fits better the estimated  $\zeta(q)$  function. It has been found to be well fitted by a log-Normal, log-Poisson (compound), log-Levy,... From a statistical point of view the precise determination of this function from a limited set of data is a difficult task (see section IV C). However, one can summarize these studies by the fact the multiscaling appears as an universal feature of financial time series. In that respect, the log-normal model advocated in this review is the simplest<sup>6</sup> one that accounts for the "intermittent" nature of return variations.

Let us mention that Ding, Engle and Granger [3, 4] were the first to observe a non trivial dependence of the volatility statistics as a function of the power q of the return chosen to define the "volatility". However these authors did not compute the scaling properties of return moments but rather were interested in the behavior of the volatility correlation function  $\mathbb{E}[|\delta_l X(t)|^q |\delta_l X(t+\tau)|^q]$  as defined in Eq. (42). Although they did not explicitly refer to the framework of multifractal analysis, they observed a power-law behavior with an exponent that is a non linear function of q. Such behavior has been confirmed by further studies (see e.g., [47]) and notably Pasquini and Serva [48] that get estimates on stock indices and FX markets compatible with the MRW prediction (43). Notice that the previous correlation function, converges, when  $q \to 0$ , to the correlation function of absolute return logarithms. This quantity has been specifically estimated in ref. [29] and shown to be very close to the cascade logarithmic shape prediction (37).

Multifractality, when interpreted as a "stochastic self-similarity", also precisely describes how the return law changes when one goes from large to small scales (Eq. 10). These pdf's will evolve from a "quasi-Gaussian" shape at large scale ( $l \simeq T$ ) to pdf's with higher kurtosis at small scales. It is noteworthty that this observation is at the origin of the analogy between the energy cascade in hydrodynamical turbulence and some 'information cascade' in financial markets as proposed in ref. [6]. This transformation of the pdf's is illustrated in fig. 4(a) where are plotted, in logarithmic scale, the standardized pdf's (their variance has been set to one) for different time scales in the range [1, 4T]. The pdf's have been estimated for 500 realizations of size  $2^{17}$  of MRW with parameters  $\lambda^2 = 0.03$  and  $T = 2^{13}$ . In solid line, we have superimposed the Castaing's transformation (10) obtained from the coarse scale pdf (at scale T) using the appropriate normal self-similarity kernel. If fig. 4(b) we have reproduced a similar analysis for the S&P 500 futures variations. Besides statistical convergence limitations, one can observe the same features as in fig. 4(a).

#### B. Market response to volatility shocks

In refs. [35], the MRW model has been used to study the specific conditional volatility relaxation either after a strong "exogeneous" event or as a mean response to some past volatility shock, wathever its value. Indeed, as explained in sec. III D 2, the magnitude  $\Omega_l(t)$  at each time resolution l can be written as a moving average process with a filter  $\alpha_k$  that decreases algebraically, as  $k^{-1/2}$  (Eqs. (58) and (59)). Let us assume that a new major "piece of information"  $\eta(t) = \omega_0 \ \delta(t)$  disturbs the market at some time (taken without loss of generality to be t = 0).  $\omega_0$  is the amplitude of the external shock. Then, if  $\omega_0$  is large enough, the temporal variation of the log-volatility  $\Omega_l(t)$  will correspond to the response to a Delta-function a therefore will reflect the power-law dependence of the filter  $\alpha_p \sim p^{-1/2}$ . The volatility response conditional on this incoming major information is thus (see [35] for details)

$$\delta_l X(t)^2 |\omega_0 \sim e^{2K_0 \sqrt{\frac{\lambda^2 T}{t}}} \tag{60}$$

For time t large enough, the volatility relaxes to its unconditional average value. Such a response of the volatility to an external shock has been verified empirically in ref. [35] on various situations where the market has been significantly "perturbated" by an external shock.

On the other hand, an "endogeneous" shock is the result of the cumulative effect of many small "news", each one looking relatively benign taken alone, but when taken all together collectively along the full path of information can add up coherently due to the long-range memory of the volatility dynamics to create a large "endogeneous" shock. It is this set of small news prior to some (large or small) "shock" that not only led to it but also continues to influence the dynamics of the volatility time series and creates an anomalously slow relaxation. In ref. [35] is given the derivation of the law of this relaxation according to the MRW model. More precisely, the conditional "response"  $\mathbb{E}\left[\sigma^2(\tau)|s\right] \equiv \mathbb{E}\left[\delta_l X(t+\tau)^2 \mid \delta_l X(t) = \sigma^2 e^{2s}\right]$ , reads:

$$\mathbb{E}\left[\sigma^2(\tau) \mid s\right] \sim \tau^{-\alpha(s)} \tag{61}$$

<sup>&</sup>lt;sup>6</sup> 'simplest' must be understood as the simplest non linear  $\zeta(q)$  function since this model has a quadratic spectrum

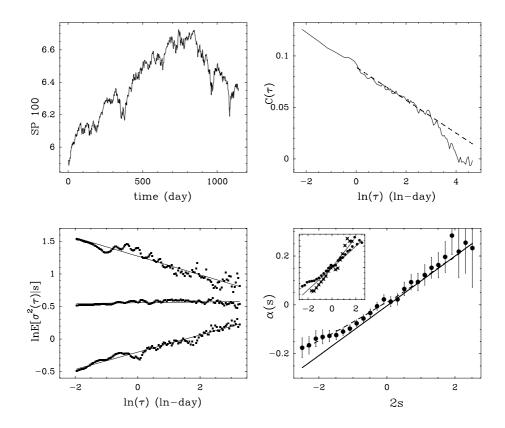


FIG. 5: Measuring the conditional volatility response (Top Left) The 5 minute S&P 100 intradaily time series from 04/08/1997 to 12/24/2001. 40 minutes and daily volatilities are estimated by aggregating 5 minute squared returns. (Top right) 40 minute log-volatility covariance as a function of the logarithm of the lag  $\tau$ . The MRW theoretical curve with  $\lambda^2 \simeq 0.02$  and T=1 year (dashed line) provides an excellent fit of the data up to lags of one month. (Bottom left) conditional volatility response  $\ln(\mathbb{E}\left[\sigma^2(\tau)\mid s\right])$  as a function of  $\ln(\tau)$  for three shocks s=-1,0,1. (Bottom right) Estimated exponent  $\alpha(s)$  for l=40 minutes ( $\bullet$ ) as a function of s. The solid line is the MRW prediction as given by Eq. (62). The dashed line corresponds to empirical MRW estimate using Monte-Carlo trials. In the inset,  $\alpha(s)$  is compared for l=40 minutes ( $\bullet$ ) and l=1 day ( $\times$ ).

where

$$\alpha(s) = \frac{2s}{\ln(\frac{Te^{3/2}}{l})}\tag{62}$$

Figure 5 reports empirical estimates of the conditional volatility relaxation for the S&P100 intradaily series made of 5 minute close prices during the period from 04/08/1997 to 12/24/2001 (figure 1(a)). The original intraday squared returns have been preprocessed in order to remove intradaily seasonalities. Figure 5(b) shows that the MRW model provides a very good fit of the empirical volatility covariance in a range of time scales from 5 minutes to one month. fig. 5(c) plots in a double logarithmic representation, for the time scale l=40 minutes, the estimated conditional volatility responses for s=1,0,-1, where the endogeneous shocks are parameterized by  $\sigma^2 e^{2s}$ . A value s>0 (resp. s<0) corresponds to a positive bump (resp. negative) of the volatility above (resp. below)  $\sigma^2$ . The straight lines are the predictions (Eqs. (61,62)) of the MRW model and qualify power law responses whose exponents  $\alpha(s)$  are continuous function of the shock amplitude s. Figure 1(d) plots the conditional response exponent  $\alpha(s)$  as a function of s for the two time scales l=40 minutes and l=1 day (inset). For l=40 minutes, we observe that  $\alpha$  varies between -0.2 for the largest positive shocks to +0.2 for the largest negative shocks, in excellent agreement with MRW estimates (dashed line) and, for  $\alpha \geq 0$ , with Eq. (62) obtained without any adjustable parameter.

The error bars represent the 95 % confidence intervals estimated using 500 trials of synthetic MRW with the same parameters as observed for the S&P 100 series. By comparing  $\alpha(s)$  for different l (inset), we can see that the MRW model is thus able to recover not only the s-dependence of the exponent  $\alpha(s)$  of the conditional response function to endogeneous shocks but also its dependence as respect to the chosen time scale l: this exponent increases as one goes from fine to coarse scales. Similar results are obtained for other intradaily time series (Nasdaq, FX-rates, etc.). We also obtain the same results for 17 years of daily return times series of various indices (French, German, Canadian, Japanese, etc.).

## C. Explaining the heavy tails of return distribution from multifractal formalism

It is well known that return distributions are characterized by heavy tails. Even if the precise shape of these tails is still a matter of debate [2, 49], the most commonly admitted form is a power-law with a tail exponent  $\mu$  within the interval [3, 5] (see e.g., refs [2, 50]). Empirically, it has been determined using data from several markets, from various countries, that the intermittency coefficient characterizing the multifractal statistics of the volatility is close to  $\lambda^2 = 0.03$  while the integral scale T is typically around 1 year (see section V). This observation leads to one of the main objections raised against the MRW model for asset returns [2]. Indeed, the *unconditional* pdf of the volatility associated with a log-normal MRW model of typical intermittency coefficient  $\lambda^2 \simeq 0.03$  has, according to Eq. (33), a tail exponent  $\mu \simeq 66$ , i.e., more than ten times the observed value! Note that this would mean that the tail exponent of the volatity pdf is around  $\mu_{\theta} \approx 16$ .

In ref. [33], we have shown that multifractal fluctuations, because of their long-range memory, do not obey classical results of extreme value statistics. In particular it can be proved that that the estimators of the pdf tail exponent not only depend on the number of "extreme values" involved in their computation, but are also not asymptotically equal the the value expected from the unconditional pdf (i.e., as if the data where independent). Let us be a little more precise: we suppose that we have a volatility series of total length N and time resolution  $\tau$ , i.e., we have measured the series,  $Z_1 = \theta(l) - \theta(0), Z_2 = \theta(2l) - \theta(l), \ldots, Z_n = \theta(Nl) - \theta((N-1)l)$ . Let us denote by  $X_1 \ldots X_N$  the same series sorted in descending order, i.e.  $X_1 \geq X_2 \geq \ldots \geq X_N$ . The most commonly used tail exponent estimators are Hill and Pickands estimators. They both involve a maximum rank k(N) = o(N) used to estimate  $\mu_{\theta}$ . For example, the Pickands estimator is:

$$\hat{\mu}_{\theta}(k, N) = \frac{\ln(2)}{\ln\left(\frac{X_k - X_{2k}}{X_{2k} - X_{4k}}\right)}$$
(63)

The asymptotic properties of these estimators rely upon classical results of Extreme Value Theory [51] (EVT). However, this theory exclusively applies to independent or weakly correlated stationary random processes [5, 51]. Therefore, multifractal fluctuations, for which the covariance decreases very slowly, are beyond the domain of validity of EVT. Estimated tail exponents should be different from those expected in the i.i.d. case. In ref. [33], we have addressed this problem and shown that it can simply be solved using the multifractal formalism. One can, for example, compute the expected value of the Pickands tail exponent estimator as follows: Let us denote  $k(N) = N^{\nu}$  the rank used to compute the tail estimator and define  $\alpha_{\nu}$  such that

$$k = N^{\nu} \sim \left(\frac{l}{T}\right)^{-f^{\star}(\alpha_{\nu})}$$

where  $\alpha_{\nu}$ ,  $f^{*}(\alpha_{\nu})$  are respectively the local Hölder exponent and the singularity spectrum defined in Eqs (44, 46) associated with the volatility  $\theta(t)$ . According to these equations,  $\alpha_{\nu}$  satisfies

$$f^{\star}(\alpha_{\nu}) = \nu \tag{64}$$

One can also consider  $\alpha'_{\nu} = \alpha_{\nu} + \epsilon_1$  such that

$$2k \sim \left(\frac{l}{T}\right)^{-f^{\star}(\alpha_{\nu}')}$$

i.e.,  $2k \sim \left(\frac{l}{T}\right)^{-f^{\star}(\alpha_{\nu})+\epsilon_1 q_{\nu}}$ , thus  $2^{-1/q_{\nu}} \sim \left(\frac{l}{T}\right)^{\epsilon_1}$  where

$$q_{\nu} = \frac{\partial f^{*}(\alpha)}{\partial \alpha}|_{\alpha = \alpha_{\nu}} . \tag{65}$$

Along the same line if  $\alpha''_{\nu} = \alpha_{\nu} + \epsilon_2$  such that

$$4k \sim \left(\frac{l}{T}\right)^{-f^{\star}(\alpha_{\nu}^{"})}$$

we have  $4^{-1/q_{\nu}} \sim \left(\frac{l}{T}\right)^{\epsilon_2}$ . Thanks to the fact that,

$$X_k - X_{2k} \sim X_k (1 - (l/T)^{\epsilon_1})$$
  
 $X_{2k} - X_{4k} \sim X_k ((l/T)^{\epsilon_1} - (l/T)^{\epsilon_2})$ 

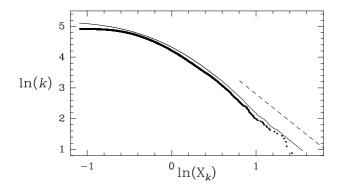


FIG. 6: Rank-Frequency plot of CAC40 daily volatility estimates (•) as compared to similar plot for a log-Gamma continuous cascade with T=253 days,  $\lambda^2=0.03$ . The log-Gamma MRW fits extreme tail better than the log-Normal MRW model in the sense that it provides a wider scaling range. The fit of the tail provides an estimation  $\mu_{\theta} \simeq 2$  (dashed line).

one finally gets for expression (63):

$$\hat{\mu_{\theta}}(\nu) \simeq \ln(2) \left( \ln \frac{1 - 2^{-1/q_{\nu}}}{2^{-1/q_{\nu}} - 4^{-1/q_{\nu}}} \right)^{-1} = q_{\nu}$$
 (66)

Eqs. (65, 64) can be solved using the expression of the  $\zeta_{\theta}(q)$  spectrum using the multifractal formalism, as explained in sec. III C 4 (Eqs 47). We then come to the conclusion that the Pickands tail estimator for a multifractal process does not, in general, provides an estimate of  $\mu_{\theta}$  the exponent of the unconditional volatility pdf and explicitly depends on the choice of the rank  $k = N^{\nu}$ . This is a strong difference with standard theory for i.i.d. random variables. In the log-normal case, from Eqs. (53, 54), one simply obtains:

$$\hat{\mu_{\theta}}(\nu) = \sqrt{\frac{2(1-\nu)}{\lambda^2}} \tag{67}$$

Remark that if  $\nu = 0$ , i.e., one uses a "logarithmically small" number of extremes in the estimator, one has  $\hat{\mu}_{\theta}(0) = q_{\star}$  where  $q_{\star}$  is defined in Eq. (55).

Typically, in finance  $l \simeq 10^{-2} - 1$  day and  $T \simeq 1 - 2$  years. Therefore, a rough estimate of  $\nu$  can be  $\nu \simeq 0.5$  (if one uses around 5 % of the data to estimate  $\mu_{\theta}$ ). Using this value, together with  $\lambda^2 = 0.03$ , in Eq. (67), one finds a typical value for the estimated tail exponent within the log-normal model that is  $\hat{\mu_{\theta}} \simeq 2.8$  that corresponds to a tail exponent for the returns  $\hat{\mu} \simeq 5.6$ . We thus recover the correct order of magnitude as compared to empirical estimates. Even if, in ref. [33], we show that a MRW model relying on a log-Gamma statistics with the same intermittency coefficient  $\lambda^2 = 0.03$ , provides a better fit of the data than the simplest log-normal version (as illustrated in fig. 6 for volatilities computed from stocks composing the CAC40 french index), the main message of this section is that the extreme tail estimated from the data mainly reflects the long-ranged volatility correlations and does not correspond to the tail of the unconditional law. In other words, one observes strongly non-ergodic behavior on quantities involving extreme values (like e.g. moments of high order q). This situation is analog to the physics of low temperature disordered systems [33]. Let us finally mention that the failure of tail estimators has been also discussed by Calvet and Fisher in [12].

## V. GMM ESTIMATION

Notwithstanding the wide number of mathematical studies devoted to multifractal and cascade models, very few of them address statistical issues like parameter estimation (see however [52]). In their pioneering works, Calvet and Fisher [10, 12] introduced multifractal modelling in econometrics and notably formulated a Markov-Switching version of their cascade model that is amenable to Maximum Likelihood estimation. These authors were also the first to propose a simulated Generalized Method of Moments (GMM) to estimate the parameters of a MMAR model. T. Lux [34, 53] also proposed a simple GMM method to estimate discrete cascade or a variant of Calvet & Fisher Poisson model parameters. In this section, we describe a GMM method for estimating the MRW parameters that proceeds along the same line. Since many of statistical moments associated with the MRW model can be analytically computed, a GMM procedure can be easily devised. As we shall discuss below, we introduce a set of moments the possess ergodic properties in both large sampled period size and small sampling rate limits. It results that our estimation becomes more and more reliable as the sampling scale is refined.

## A. Parameters of the log-normal MRW

As we have seen in section IIIB2, the log-normal MRW is fully defined by 3 parameters:

- $\sigma^2$  which represents the mean of the volatility increment at scale 1, i.e.,  $\mathbb{E}\left[|\delta_1 X(t)|^2\right] = \sigma^2$ . (Let us note that it is also the variance of the noise  $\epsilon_l$  in Eq. (23))
- T, the integral scale which corresponds to the decorrelation scale (actually, two increments with a lag greater than T are not only decorrelated, they are independent; see Eq. (25)).
- $\lambda^2$ , the intermittency factor which is involved in the non-linear part of  $\zeta_q$  (Eq. (32)) and which corresponds, for large lags, to the slope of the "renormalized magnitude"  $\Omega_l$  (see Eqs (36) and (37)) as a function of the logarithm of the lag. (Let us note that it is also the slope of the microscopic magnitude  $\omega = 2\tilde{\omega}$  (see Eq. (25))).

We see that all three parameters are linked to some moments of the process increments or of their logarithm. It is thus very natural to estimate them using the Generalized Methods of Moment. However, as discussed in sec. IV C, the simple moments of the MRW returns M(q,l) (as defined in Eq. (31)) are, for large q values, not amenable to meaningful estimation because of non self-averaging effects. In order to obtain a reliable estimation method, we therefore use moments for low q values, and more precisely the correlation function of the logarithms of absolute returns.

## B. GMM principles

The Generalized Methods of Moment (GMM) theory was introduced by Hansen in 1982 [54]. It consists in estimating the parameters of a process using moment estimations. The Method of Moment uses as many moments as there are parameters (it ends up solving an equation) whereas the GMM uses more moments than parameters. The parameter estimations are obtained by finding the values that correspond to the "best" fit (minimizing a quadratic error) of the chosen moments.

More precisely, if the process  $Y_{\Theta=\Theta_0}[n]$  depends on a p dimensional parameter  $\Theta$ , the moment conditions are given through an r (> p) dimensional function  $f(Y_{\Theta_0}[n], \Theta)$  such that

$$\mathbb{E}\left[f(Y_{\Theta_0}[n],\Theta)\right] = 0 \Longleftrightarrow \Theta = \Theta_0.$$

The moment conditions are estimated using the empirical means

$$g_N = \frac{1}{N} \sum_{t=1}^{N} f(Y_{\Theta_0}[n], \Theta).$$

The GMM-estimator is then define as

$$\hat{\Theta} = argmin_{\Theta} \left( g_N^* W_N g_N \right), \tag{68}$$

where  $W_N$  is a weighting matrix that must converge when  $N \to +\infty$  towards a positive definite matrix  $W_\infty$ . Then [54] if

- (i)  $\{Y[n]\}_n$  is ergodic,
- (ii) the series  $f(Y_{\Theta_0}[n], \Theta)$  satisfies a central limit theorem, i.e.,

$$\frac{1}{\sqrt{N}} \sum_{n=1}^{N} f(Y_{\Theta_0}[n], \Theta) \longrightarrow \mathcal{N}(0, V_{\Theta}),$$

where  $V_{\Theta}$  is naturally defined by

$$V_{\Theta} = \lim_{K \to +\infty} \sum_{k=-K}^{K} \mathbb{E}\left[f(Y_{\Theta_0}[n], \Theta) f(Y_{\Theta_0}[n-k], \Theta)^*\right]. \tag{69}$$

(iii) the  $r \times p$  dimensional matrix

$$Dg_N = \frac{\partial g_N}{\partial \Theta},$$

has full-rank (q) and converges towards

$$Df = \mathbb{E}\left[\frac{\partial f(Y_{\Theta_0}[n], \Theta)}{\partial \Theta}\right].$$

then the estimator  $\hat{\Theta}$  is consistent and

$$\sqrt{N}(\hat{\Theta} - \Theta_0) \longrightarrow \mathcal{N}(0, \Sigma),$$

where

$$\Sigma = (Df^*W_{\infty}Df)^{-1}Df^*W_{\infty}V_{\Theta_0}W_{\infty}Df(Df^*W_{\infty}Df)^{-1}.$$

**Optimal estimator.** One can easily prove that the optimal estimator is obtained when  $W_N \to W_\infty = V_{\Theta_0}^{-1}$  in Eq. (68). In this case, the asymptotic covariance of the GMM estimator (68) is

$$\Sigma_{opt} = (Df^*V_{\Theta_0}^{-1}Df)^{-1}.$$

Of course, from a practical point of view, one cannot use directly the weighting matrix  $V_{\Theta_0}^{-1}$  since  $\Theta_0$  is unknown. Thus, usually, the GMM optimal estimator is obtained using an iterative procedure :

- a) Set the weighting matrix  $W_N$  to be an arbitrary weighting matrix (one can choose Id or  $V_{\hat{\Theta}}^{-1}$  where  $\hat{\Theta}$  is an apriori estimate of  $\Theta$ )
- b) Compute the GMM estimator  $\hat{\Theta}$  (Eq. (68)) using  $W_N$
- c) Set the weighting matrix  $W_N = V_{\hat{\Theta}}^{-1}$  where  $\hat{\Theta}$  is the GMM estimator obtained in b).
- d) Go to b) unless the successive GMM estimates are close one to each other (a threshold is used).

### C. GMM estimation of log-normal MRW

For log-normal MRW, using notations of the previous section, we set the process Y to be a discretized version (at a fixed scale l) of the increments of the MRW X(t), i.e.,

$$Y[n] = \delta_l X(ln).$$

In the Appendices 1 and 2 (see also Eq. (34)), it is shown that, when  $\lambda^2 \ll 1$  (which is the case for financial time-series), in some sense, one can write

$$Y[n] \underset{Law}{\simeq} \epsilon[n] e^{\Omega_l[n]},$$

where  $\epsilon[n]$  is a gaussian white noise of variance  $\sigma^2$  and  $\Omega_l[n]$  is the "renormalized (Gaussian) magnitude" defined by Eqs (35) and (36). We choose the parameter vector  $\Theta$  (p=3) to be

$$\Theta = \{ \ln \sigma, \ \lambda^2, \ \ln T \}.$$

For the moment condition, following the discussion in section VA, it seems natural to choose (i) the variance of Y that should lead to an estimator of  $\sigma^2$  and (ii) the autocorrelation function of  $\ln |Y|$  for different lags that should lead to both an estimator of  $\lambda^2$ , i.e., the slope of the autocorrelation function of the magnitude  $\Omega_l$  (cf Appendix) and the decorrelation (integral) scale T. Thus we choose:

$$f(Y[n], \Theta) = \begin{pmatrix} Y[n]^2 - \sigma^2 \\ (\ln|Y[n]| - m)(\ln|Y[n - n_1]| - m) - C_{n_1} \\ (\ln|Y[n]| - m)(\ln|Y[n - n_2]| - m) - C_{n_2} \\ & \dots \\ (\ln|Y[n]| - m)(\ln|Y[n - n_K]| - m) - C_{n_K} \end{pmatrix},$$
(70)

where

$$m = \mathbb{E} \left[ \ln |Y[n]| \right] = \ln \sigma + \mathbb{E} \left[ \ln |\epsilon| \right] + \mathbb{E} \left[ \Omega \right],$$

and

$$C_k = \mathbb{C}\text{ov}\left(\ln|Y[n]|, \ln|Y[n-k]|\right),$$

and  $\{n_1, n_2, \ldots, n_K\}$  are K different positive lags. Let us note that the analytical expressions of m and  $C_k$  as functions of  $\sigma^2$ ,  $\lambda^2$  and T are obtained using the log-normal approximation described in Appendix. Using the same approximation, one can get [55] an analytical expression of the matrix  $V_{\Theta}$  defined in Eq. (69). Explicit expressions and more details about GMM approximation for log-normal MRW can be found in [55]

Let us point out that N, the total length of the series Y[n], depends on both L the length of sampled period and l, the sampling time scale. In general asymptotic convergence is achieved in the limit  $L \to +\infty$ . However, as explained in sec. IV C and illustrated below, the volatility correlation length (our integral scale T parameter) is of the order of the year, so, in order to "realize" the limit  $L \to +\infty$ , one should have time series sampled over many years. Since the MRW model is scale invariant, we have been able to show that an ergodic theorem for the lagged correlation functions of return logarithms can also be obtained in the limit  $l \to 0$ . That simply means that the finer is the time resolution, the better will be the model estimation. This is a very nice property of our model as compared to models that focus on a single time scale. So the hypothesis (i) and (ii) of last sections are verifyied in both large L and small l limits. One can also check [55] that condition (iii) is satisfied.

## D. Applications to real data

We applied the GMM method described in the last section on real financial time-series. Figure 7 illustrates the GMM estimation on both CAC40 french index and Italian index. Daily (close) data from 1973 to 1997 (i.e.,  $\simeq 6200$  points) were used. The lags in Eq. (70) were chosen as K=35 and  $n_k=2k-1$ . For the CAC40, one gets the estimation  $\sigma^2 \simeq 0.011$ ,  $\lambda^2 \simeq 0.022$  and  $T \simeq 268$  days, whereas the Italian index parameters are  $\sigma^2 \simeq 0.013$ ,  $\lambda^2 \simeq 0.029$  and  $T \simeq 500$ . Figure 7 displays a GMM fit of the covariance of the magnitude  $\Omega$  as a function of the logarithm of the lag ( $\lambda^2$  corresponds roughly to the slope of the fit and T to the decorrelation scale).

Let us note that the largest time-scale (i.e., the size of the data, about 24 years) is not that large compared to the correlation scale (2 years for the italian index). It means that the GMM asymptotic has not really been reached yet. However some more detailed studies [55] on GMM estimations of numerical simulations of MRW prove that one can "trust" the obtained estimator though one should not rely on the error bars given by the GMM estimation. The gaussian asymptotics has not been reached yet. The error bars should be computed using Monte-Carlo methods. The 5% error bar on  $\sigma^2$  is of the order of  $\pm 10\%$ , the one on  $\lambda^2$  of the order of  $\pm 40\%$  and the one on T is very large and is of the order of -25%, +300%. We refer the reader to [55] for more detailed discussions on error bars.

The quality of the fit (i.e., of the estimation) depends of course on the number of points in the time-series. Figure 8 illustrates the GMM estimation on an intraday S&P 500 time-series (5mn ticks) from 1996 to 1998 (i.e., more than 43000 points). The fit reproduces really well the correlation behavior starting from intraday lags. One gets  $\sigma^2 \simeq 0.01$ ,  $\lambda^2 \simeq 0.02$  and  $T \simeq 62$  days. Let us notice that the value  $\lambda^2 = 0.02$  is slightly smaller than the reported value  $\lambda^2 = 0.03$  in section IV A. However, as mentionned previously and as shown by Monte-Carlo experiments on MRW trials, estimation relying on moment multiscaling has relatively large errors.

The table I shows the result of GMM estimation on various financial time-series: intraday S&P500 index data, daily exchange rate data for the Canadian Dollar (CAD), the Japanese Yen (JPY), the Swiss Franc (CHF) and the British Pound (GBP), all against the US Dollar, and daily data for the French index (CAC40), Italian index (MIB), Canadian index (TSX), German index (DAX), UK index (FTSE100), Hong-Kong index. The exchange rate series will be used to estimate performances of MRW model in volatility and Value at Risk forecasting in section VI. These series begin on July 1st 1977 and run until March 20 2006. We have used data until December 28, 1989 for estimation procedure (in-sample interval). The remaining of data will be used in next sections for backtesting purposes (out-of-sample).

Despite some wide dispersion in the integral scale values (probably due to very large estimation errors), we see that the typical values are  $T \simeq 1$  year for the integral scale and  $\lambda^2 \simeq 0.03$  for the intermittency coefficient.

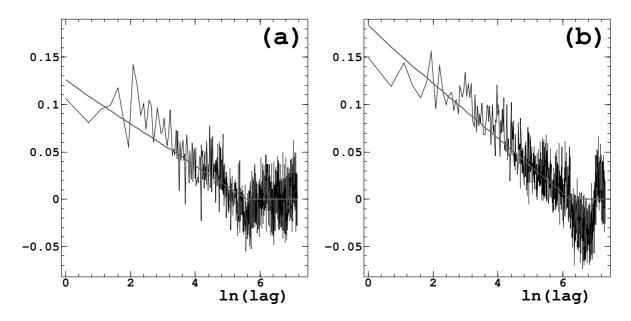


FIG. 7: GMM estimation of log-normal MRW parameters on daily (close) financial time-series. The data range from 1973 to 1997 which corresponds approximately to 6200 points. The lags in Eq. (70) are such that K=35 and  $n_k=2k-1$ . The covariance function of the magnitude  $Cov(\Omega)[lag]$  is displayed as a function of  $\ln(lag)$ . The fit corresponds to the theoretical curve (cf Appendix) computed using the parameters  $\lambda^2$  and T estimated using the GMM. For discussion on error bars see text. (a) CAC40 french index. One gets  $\sigma^2 \simeq 0.011$ ,  $\lambda^2 \simeq 0.022$  and  $T \simeq 268$ . (b) Italian index. One gets  $\sigma^2 \simeq 0.013$ ,  $\lambda^2 \simeq 0.029$  and  $T \simeq 500$ .

Series	Type	$\lambda^2$	T
S&P500 index	intraday data	0.02	3 months
CAD/USD exchange rate	daily data	0.024	1 year
JPY/USD exchange rate	daily data	0.026	9 months
CHF/USD exchange rate	daily data	0.021	7 months
GBP/USD exchange rate	daily data	0.018	4.5 years
French index	daily data	0.022	1 year
Italian index	daily data	0.029	2 years
Canadian index	daily data	0.032	1.5 years
German index	daily data	0.027	6 years
UK index	daily data	0.023	6 years
Hong-Kong index	daily data	0.037	6 years

TABLE I: GMM estimation on various intraday (5mn ticks) or daily financial time-series.

# VI. CONDITIONAL VOLATILITY AND VAR FORECASTING

## A. Volatility forecasting

In this section we evaluate the volatility forecasting performances of MRW. In ref. [12], Calvet and Fisher have already shown that their cascade model provides better volatility forecasts as compared to GARCH(1,1) or Markov-Switching GARCH. Lux [34] provided additional evidences of the performances of this model. Here we proceed along the same line and show that the simplest linear volatility forecast provided by the MRW model outperforms GARCH(1,1) models.

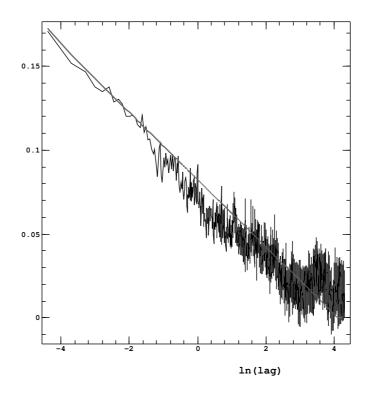


FIG. 8: GMM estimation of log-normal MRW parameters on intraday (5mn ticks) SP500 time-series. The data range from 1996 to 1998 which corresponds approximately to 43000 points. The lags in Eq. (70) are such that K=35 and  $n_k=2k-1$ . The covariance function of the magnitude  $Cov(\Omega)[lag]$  is displayed as a function of  $\ln(lag)$ . The fit corresponds to the theoretical curve (cf Appendix) computed using the parameters  $\lambda^2$  and T estimated using the GMM. For discussion on error bars see text. One gets  $\sigma^2 \simeq 0.01$ ,  $\lambda^2 \simeq 0.02$  and  $T \simeq 62$ .

### 1. What do we mean by volatility forecasting

Volatility is a model dependent notion. For instance, for GARCH models, at a given time the conditional volatility (to all the observed past) is a deterministic number whereas for stochastic volatility models or for the MRW model it is a random variable. Thus, "volatility forecasting" is a model dependent notion. In order to compare different models on what is generally referred to as "volatility forecasting", one needs to define a common problematic. The natural problematic is forecasting of estimation of realized volatility given by squared returns over forecasted period. Thus we try to solve the problem of forecasting  $\sum_{k=1}^{l_1/l} |\delta_l X(t_0 + kl)|^2$  (with  $l_1 \geq l$ ) knowing all the past data  $\{\delta_l X(t)\}_{t \leq t_0}$ . The parameter  $l_1$  will be referred to as the prediction scale.

## 2. Forecasting volatility of Forex exchange rates

The simplest way to forecast volatility using MRW is to use a linear regression of the sum of square log-return over forecasted interval  $\sum_{k=1}^{l_1/l} |\delta_l X(t_0+kl)|^2$  on past square log-returns  $\{|\delta_l X(t)|^2\}_{t \leq t_0}$ . The optimal linear prediction filter can be computed just using the expression of the volatility auto-covariance function (39). Let us notice that more sophisticated forecasts can be devised: for instance if volatility is observable, by e.g. aggregating squared intraday returns, one can compute the full conditional expectation of future squared returns. However, mean squared linear forecasting is sufficient to the purpose of this review paper and we refer to ref. [31] for alternative MRW forecasting methods and more details.

The forecasting performances of MRW filter will be compared to those of GARCH(1,1) model. Though GARCH(1,1) might not be the best known models for volatility prediction, they are widely used in the literature and they have been proven to be among the best models [56]. Let us note that GARCH(1,1) is also used as reference model by Calvet, Fisher to investigate the performance of their Markov-Schitwing model [12]. This is the reason why we choose

	n	GARC	H	tGARCH				
	$\omega$	$\alpha_1$	$\beta_1$	$\nu$	$\omega$	$\alpha_1$	$\beta_1$	
CAD/USD	0.0020	0.1345	0.8451	6.6857	0.0015	0.1164	0.8675	
JPY/USD	0.0215	0.0913	0.8629	3.8769	0.0114	0.1040	0.8878	
CHF/USD	0.0160	0.0904	0.8879	5.8048	0.0116	0.0772	0.9091	
$_{\mathrm{GBP}/\mathrm{USD}}$	0.0134	0.0661	0.9053	4.9151	0.0021	0.0722	0.9278	

TABLE II: Maximum likelihood estimation results for GARCH(1,1) model with normal 'nGARCH' and t-Student 'tGARCH' errors for the four exchange rate series.

1-day		MAE	
forecast	tGARCH	nGARCH	MRW
CAD/USD	0.140	0.141	0.135
JPY/USD	0.558	0.521	0.517
CHF/USD	0.556	0.554	0.548
GBP/USD	0.360	0.368	0.356
5-days		MAE	
forecast	tGARCH	nGARCH	MRW
CAD/USD	0.394	0.399	0.374
JPY/USD	1.770	1.582	1.547
CHF/USD	1.557	1.554	1.515
$_{\mathrm{GBP}/\mathrm{USD}}$	1.052	1.083	1.035
00.1		3.5.4.75	
20-days		MAE	
20-days forecast	tGARCH	nGARCH	MRW
=	tGARCH 1.097		MRW <b>1.067</b>
forecast		nGARCH	
$\frac{\rm forecast}{\rm CAD/USD}$	1.097	nGARCH 1.114	1.067
forecast CAD/USD JPY/USD	1.097 5.679	nGARCH 1.114 4.461	1.067 4.301
forecast CAD/USD JPY/USD CHF/USD	1.097 5.679 4.327	nGARCH 1.114 4.461 4.288	1.067 4.301 4.049
forecast CAD/USD JPY/USD CHF/USD GBP/USD	1.097 5.679 4.327	nGARCH 1.114 4.461 4.288 3.334 MAE	1.067 4.301 4.049 2.976
forecast CAD/USD JPY/USD CHF/USD GBP/USD 50-days	1.097 5.679 4.327 3.231	nGARCH 1.114 4.461 4.288 3.334 MAE	1.067 4.301 4.049 2.976
forecast CAD/USD JPY/USD CHF/USD GBP/USD 50-days forecast	1.097 5.679 4.327 3.231 tGARCH	nGARCH 1.114 4.461 4.288 3.334 MAE nGARCH	1.067 4.301 4.049 2.976
forecast CAD/USD JPY/USD CHF/USD GBP/USD 50-days forecast CAD/USD	1.097 5.679 4.327 3.231 tGARCH 2.632	nGARCH 1.114 4.461 4.288 3.334 MAE nGARCH 2.674	1.067 4.301 4.049 2.976 MRW 2.557
forecast CAD/USD JPY/USD CHF/USD GBP/USD 50-days forecast CAD/USD JPY/USD	1.097 5.679 4.327 3.231 tGARCH 2.632 15.361	nGARCH 1.114 4.461 4.288 3.334 MAE nGARCH 2.674 9.686	1.067 4.301 4.049 2.976 MRW 2.557 9.210

TABLE III: Mean absolute error (MAE) for out-of-sample volatility forecasting for exchange rate series using GARCH(1,1) models with normal 'nGARCH' and t-Students 'tGARCH' errors and MRW model 'MRW'. MRW values are in bold if smaller than both GARCH values.

to compare MRW volatility prediction versus GARCH(1,1) volatility prediction, with normal and t-Student errors. As in [11, 12], we choose to compare on the exchange data. These series have been already used in section V. They consist of daily prices of FX rates associated with Canadian Dollar (CAD), Japanes Yen (JPY), Swiss Franc (CHF) and British Pound (GBP) against the USD. The in-sample data beginning on 1st July 1977 and ending on 28 December 1989 has been used for the estimation of each model. We report in table II the maximum likelihood in-sample estimates of normal and t-Student GARCH(1,1) models for each FX series.

The out-of-sample period extends from 29 December 1989 to 20 March 2006. The forecasting performances of each model are compared at four different prediction scales:  $l_1 = 1,5,20$  and 50 days. Table III presents the Mean Absolute prediction Error (MAE) for each series and each prediction scale. We have chosen to print MRW result in bold face if its error is smaller than those both GARCH models. As one can see, MRW predictions outperform clearly GARCH(1,1) predictions for all series and at all prediction scales.

Table IV presents the summary of  $L^2$  measures such as Mean Square Error (MSE),  $R^2$  measure and the coefficients

1-day		MSE			$R^2$			$\gamma_0$			$\gamma_1$	
forecast	tGARCH	nGARCH	MRW	tGARCH	nGARCH	MRW	tGARCH	nGARCH	MRW	tGARCH	nGARCH	MRW
CAD/USD	0.064	0.064	0.064	0.103	0.095	0.109	0.030	0.034	0.008	0.814	0.784	1.083
JPY/USD	1.320	1.299	1.287	0.044	0.059	0.068	0.104	0.049	-0.019	0.687	0.903	1.055
CHF/USD	0.822	0.824	0.814	0.025	0.023	0.034	0.126	0.142	0.014	0.683	0.664	0.906
GBP/USD	0.419	0.416	0.414	0.071	0.076	0.081	0.056	-0.054	-0.027	0.761	1.028	1.012
5-days		MSE			$R^2$			$\gamma_0$			$\gamma_1$	
forecast	tGARCH	nGARCH	MRW	tGARCH	nGARCH	MRW	tGARCH	nGARCH	MRW	tGARCH	nGARCH	MRW
CAD/USD	0.370	0.380	0.366	0.343	0.325	0.350	0.131	0.145	-0.018	0.854	0.833	1.213
JPY/USD	10.394	9.632	9.341	0.049	0.119	0.146	0.761	0.473	0.059	0.583	0.807	0.992
CHF/USD	5.081	5.114	4.866	0.073	0.067	0.112	0.668	0.727	-0.122	0.660	0.646	0.965
GBP/USD	2.854	2.771	2.718	0.189	0.212	0.227	0.336	-0.293	-0.169	0.720	1.026	1.022
20-days		MSE			$R^2$			$\gamma_0$			$\gamma_1$	
forecast	tGARCH	nGARCH	MRW	tGARCH	nGARCH	MRW	tGARCH	nGARCH	MRW	tGARCH	nGARCH	MRW
CAD/USD	2.691	2.791	2.938	0.520	0.502	0.476	0.416	0.403	-0.279	0.926	0.935	1.372
JPY/USD	77.832	59.638	56.360	-0.020	0.218	0.261	2.963	0.455	-0.861	0.543	0.963	1.114
CHF/USD	34.544	33.897	30.102	0.081	0.099	0.200	2.394	2.168	-2.505	0.653	0.679	1.123
GBP/USD	24.350	21.748	20.047	0.178	0.266	0.323	1.827	-2.040	-0.700	0.624	1.090	1.005
50-days		MSE			$R^2$			$\gamma_0$			$\gamma_1$	
forecast	tGARCH	nGARCH	MRW	tGARCH	nGARCH	MRW	tGARCH	nGARCH	MRW	tGARCH	nGARCH	MRW
CAD/USD	16.319	16.869	17.928	0.444	0.426	0.390	0.961	0.612	-0.831	0.994	1.060	1.491
$\rm JPY/USD$	452.099	262.717	238.996	-0.402	0.185	0.259	7.436	-5.403	-4.256	0.477	1.247	1.214
CHF/USD	179.165	165.453	135.537	-0.142	-0.055	0.136	5.861	3.360	-8.249	0.614	0.706	1.167
$_{\mathrm{GBP}/\mathrm{USD}}$	146.488	114.414	92.190	-0.112	0.132	0.300	6.002	-9.656	-1.089	0.506	1.249	0.949

TABLE IV: . Mean square error (MSE),  $R^2$  measure equals to the MSE divided by the sum of squared demeaned squared returns in the out of sample period and two coefficients  $\gamma_0$  and  $\gamma_1$  from the Mincer-Zarnowitz OLS regression, for out-of-sample volatility forecasting for exchange rate series using GARCH(1,1) models with normal 'nGARCH' and t-Students 'tGARCH' errors and MRW model 'MRW'. MRW MSE values are in bold if smaller than both GARCH MSE values. MRW  $R^2$  values are in bold if closer to 1 than both GARCH  $r^2$  values. MRW  $\gamma_0$  (resp.  $\gamma_1$ ) values are in bold if closer to 0 (resp. 1) than both GARCH ones.

 $\gamma_0$  and  $\gamma_1$  from the Mincer-Zarnowitz OLS regression. The "bold face" convention for MSE and  $R^2$  is the same as in previous table. For the Mincer-Zarnowitz regression, the MRW  $\gamma_0$  (resp.  $\gamma_1$ ) value is in bold if it is closer to 0 (resp. 1) than both GARCH ones. As in the case of the MAE, if one accounts for the MSE or  $R^2$  criteria, one clearly see that the MRW prediction outperforms both GARCH(1,1) models: except for the CAD/USD rate at scales 20 and 50 days, the MRW prediction is better than GARCH predictions. Let us notice that Mincer-Zarnowitz values are globally closer to the expected values  $\gamma_0 = 0$  and  $\gamma_1 = 1$  for the MRW model than for GARCH models. A computation of asymptotic standard errors (not reported) shows that the hypothesis  $\gamma_0 = 0$  and  $\gamma_1 = 1$  is accepted at a 5 % confidence level in the MRW case for all series at all prediction scales.

Let us notice that those forecasting performances of MRW have been confirmed on other series like daily return of stocks composing the french CAC40 index. Morevoer, those performances are naturally improved when on uses return at intraday time scales (see ref. [31]).

## B. Conditional Value at Risk forecasting

The MRW model allows us to predict not only the volatility but other measures of risk. Among all risk measures, Value at Risk (VaR) is widely used (see the definition below). In ref. [57], Calvet and Fisher developed a bivariate extension of their Markov-Switching model and showed that it compares favorably to a GARCH model (CC-GARCH) for Value at Risk forecasting of FX rate data. Here, we show, in agreement with results on volatility, MRW outperforms univariate GARCH(1,1) models on conditional VaR prediction.

#### 1. VaR forecasting

By definition, the unconditional value-at-risk  $VaR_{l,p}$  at scale l at level p, is the value which verifies the following probability equality

$$\mathbb{P}\left[\delta_{l}X(t) < -VaR_{l,n}\right] = p, \quad \forall t. \tag{71}$$

Roughly speaking,  $VaR_{l,p}$  can be interpreted as the most probable amplitude of the worst loss at a time-scale l over an horizon of  $p^{-1}$  periods [2]. The conditional value at risk  $VaR_{l,p}(t_0)$  at time  $t_0$  is the number which satisfies

$$\mathbb{P}\left[\delta_{l}X(t_{0}+l)|_{X(t),t< t_{0}} < -VaR_{l,p}(t_{0})\right] = p, \quad \forall t.$$
(72)

In this section, we try to solve the problem of forecasting  $VaR_{l_1,p}(t_0)$  (with  $l_1 \geq l$ ) knowing all the past data  $\{\delta_l X(t)\}_{t \leq t_0}$ . The parameter  $l_1$  will be referred to as the prediction scale.

### 2. Value at Risk forecasting with MRW model

In Appendix, we explain, using an edgeworth expansion of the law of  $\Omega_l^{\theta}(t)/\lambda$ , why the (unconditional) cumulative distribution function of  $\delta_l X(t)$  can be approximated (first order in  $\lambda^2$ ) by the cumulative distribution of a log normal process, i.e.,

$$\mathbb{P}\left[\delta_{l}X(t) < z\right] \simeq \mathbb{P}\left[\varepsilon e^{\gamma\lambda g - \gamma^{2}\lambda^{2}} < z\right],$$

where  $g = \mathcal{N}(0,1)$  and  $\gamma$  is defined in the Appendix. By numerical inversion of the following equation

$$\mathbb{P}\left[\varepsilon e^{\gamma\lambda g - \gamma^2\lambda^2} < -VaR_{l,p}\right] = \mathbb{E}\left[N\left(-\frac{VaR_{l,p}}{\sigma\sqrt{l}}e^{\gamma\lambda g - \gamma^2\lambda^2}\right)\right] = p \tag{73}$$

where N(x) is the cumulative distribution function of a Gaussian variable, one can get an estimation of the (unconditional) VaR.

Using the same kind of arguments [58] (on conditional laws), one can compute estimations of the conditional VaR based on estimation of the conditional mean and variance of  $\Omega_l^{\theta}$  as performed in the former section.

## 3. Forecasting VaR of Forex exchange rates

As for the volatility forecasting, we choose to compare the MRW prediction versus the GARCH(1,1) (normal or t-Students) predictions using daily FX rate series. In our tests we used estimated parameters reported in the tables I and II. The backtesting consisted in computing how many times the forecasted VaR at level p was reached. It should be reached exactly p% of the time. To evaluate VaR forecasts we use the Markov binomial test of Bernoulli series described in [59]. Let us note that for large scales  $l_1$ , the length of corresponding series becomes small that leads to acceptance of the majority of statistical tests. Therefore, we present the results only for 1-day VaR forecasting.

The table V presents the comparison of 1-day conditional value-at-risk forecasting for different models for various probability levels p. Accepted Markov binomial test at 5 % confidence level are reported in bold face. We see that the GARCH(1,1) model with t-Students errors does not provide satisfactory results whatever the level p. This could be explained by wrong fitting of the distribution tails by this model. The GARCH(1,1) model with normal innovations provides good results around the level p=5 % but fails at larger and smaller probability levels. The MRW model prediction is always accepted at all probability levels excepted for the JPY/USD series at small p. But we see that even in that case, it provides a backtested p value closer to the expected one than GARCH models.

In figure 9 is displayed the difference between the backtested frequency and the probability level p as a function of p. The level p is varying from 0.5% to 30% with 0.5% step. One clearly sees on the right figure, that if GARCH(1,1) fits relatively well probability levels between 5 and 12 %, it fails to account correctly for low or high probability values p. This is not the case for the MRW prediction that remains within the 5 % confidence interval in the whole range of values of p. This means that the MRW model provides satisfactory results at all asked value-at-risk levels.

1-day		0.5%	
forecast	tGARCH	nGARCH	MRW
CAD/USD	0.002 (0.005)	0.012 (0.000)	0.004 (0.313)
JPY/USD	$0.001 \ (0.000)$	$0.013\ (0.000)$	$0.008 \; (0.010)$
CHF/USD	0.001 (0.000)	0.011 (0.000)	$0.004\ (0.313)$
USD/GBP	$0.001 \ (0.000)$	$0.006\ (0.321)$	$0.003 \ (0.079)$
1-day		1%	
forecast	tGARCH	nGARCH	MRW
CAD/USD	0.005 (0.001)	0.017 (0.000)	0.008 (0.275)
JPY/USD	0.004 (0.000)	$0.020\ (0.000)$	$0.015 \ (0.003)$
CHF/USD	$0.003 \ (0.000)$	$0.015 \ (0.003)$	0.009 (0.663)
USD/GBP	0.002 (0.000)	0.010 (0.844)	0.007 (0.076)
1-day		5%	
forecast	tGARCH	nGARCH	MRW
CAD/USD	0.034 (0.000)	$0.056 \ (0.086)$	$0.052 \ (0.507)$
JPY/USD	$0.023 \ (0.000)$	$0.053\ (0.381)$	$0.054\ (0.277)$
CHF/USD	$0.030 \ (0.000)$	$0.055\ (0.132)$	$0.048 \ (0.482)$
USD/GBP	$0.024\ (0.000)$	$0.039\ (0.000)$	$0.043\ (0.050)$
1-day		10%	
forecast	tGARCH	nGARCH	MRW
CAD/USD	0.077 (0.000)	0.098 (0.660)	0.109 (0.053)
JPY/USD	$0.053 \ (0.000)$	$0.087 \ (0.004)$	0.098 (0.660)
CHF/USD	$0.072\ (0.000)$	$0.093\ (0.121)$	0.098 (0.660)
USD/GBP	0.061 (0.000)	$0.077 \ (0.000)$	$0.093\ (0.164)$

TABLE V: . The frequency of returns that exceed the value-at-risk forecasted by the GARCH(1,1) models with normal 'nGARCH' and t-Student 'tGARCH' errors and the MRW model 'MRW'. Between the parenthesis p-values of Markov statistical tests are reported. The accepted statistical tests at 5% significance level appear in bold.

### VII. CONCLUSION AND PROSPECTS

In this paper we have proposed a survey of a class of mathematical models for asset returns relying upon the concept of "continuous cascade". More specifically, we have seen that the simplest of such models, the log-normal MRW model, allows one to account, with few parameters, for most of universal features empirically observed on real markets. Its two key parameters describe respectively the length of (log-) volatility correlations (the integral scale T) and the strength of (log-) volatility fluctuations (the intermittency parameter  $\lambda^2$ ). We have seen that a lot of statistical quantities can be obtained in closed forms or simply approximated. We introduced a simple method to estimate its parameters and have shown that it provides accurate risk forecasts. Let us point out that our purpose was not to claim that return fluctuations are likely to result from a cascade model and that among these models, the log-normal we proposed is the best one. We simply aimed at illustrating that, the simplest of cascade models, is rich enough to account for a wide variety of complex properties observed on empirical data and can be successfully used to address classical problems of financial engineering.

From a practical point of view, MRW-like models while capturing the "heteroskedastic" nature of return fluctuations, still preserve, in some sense, the nice stability property across time scales of the Brownian motion. This sharply contrasts with classical econometric models that can hardly be controlled as the time-scale is changed. In spite of its limitations and drawbacks (as descrided below) there is another major breakthrough associated with the model we presented: they allow one to rationalize many different "stylized facts" and show that multiscaling of moments, multiscaling of return correlation functions, slow convergence towards a Gaussian shape as returns are aggregated, fat tails are likely to be different facets of the same reality, namely the existence of magnitude (i.e., volatility logarithms) long range correlations. In that respect, this model can also be considered as an analyzing and computational tool to explore other, more complex models.

Many interesting questions for forthcoming studies are related to drawbacks of the MRW model and mainly to its incapacity to account for the so-called "leverage effect" according to which past returns have impacts on future

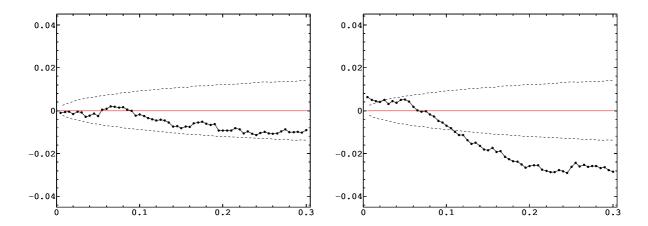


FIG. 9: . Errors for conditional value-at-risk forecasting (from 0.5% to 30% with 0.5% step) on the exchange rate series of Swiss Franc against the US Dollar. The left figure corresponds to MRW prediction, the right one corresponds to nGARCH(1,1) prediction. Dashed lines present 5% confidence interval of Markov statistical test.

volatilities. This phenomenon is well documented and by definition cannot be accounted by a cascade model that is time reversal symmetric. An interesting extension of the MRW that explicitly breaks time reversal invariance has been proposed in ref. [60]. However this "skewed" MRW model "loses" its skewness in the continuous time limit and can only be used as a discrete time model. The construction of skewed cascade processes remains therefore an open mathematical problem. Another interesting related issue, concerns the possible extension of the MRW approach of volatility memory to multiscale ARCH-like models along the same line as the models studied in [37, 40]. As far as other prospects for future researh are concerned, most of them naturally concern classical problems of mathematical finance like portfolio management and option pricing. Notice that the latter one as already been considered in ref. [60]. To address the portfolio optimization problem, one needs a multivariate version of the models described in this paper. Even if some attempts have been proposed in that direction [24], a lot of mathematical as well as practical problems remain unsolved. Finally, we also need to further develop statistical studies of MRW model in relation to asset return modelling. In doing so, we hope that this model will find a place in the toolbox of financial engineering beside other popular econometric models.

### Acknowledgments

This work has been supported by contract ACI nouvelles interfaces des mathématiques, Nb 54-03, from french ministère de l'éducation nationale. We are grateful to D. Farmer and T. Lux for inviting us to contribute to this special issue. We also thank both anonymous referees for their helpful remarks.

## **APPENDIX:** MRW LIMIT BEHAVIOR FOR $\lambda^2 \ll 1$

#### 1. The characteristic function

Let  $\theta(t)$  be the trading time for log-normal MRW (as defined in Eq. (20)), and let  $\Omega_l^{\theta}(t)$  be half the magnitude of  $\theta$  at size l:

$$\Omega_l^{\theta}(t) \equiv \frac{\ln(\theta(t+l) - \theta(t))}{2}.$$
(74)

If  $0 \le t_1 < t_1 + l_1 \le t_2 < t_2 + l_2 \le ... \le t_n < t_n + l_n \le T$ , then, one can show [58] that the characteristic function of the n variables  $\{\Omega_{l_j}^{\theta}(t_j)\}_{j=1}^n$  is given by

$$\mathbb{E}\left[\exp\left(i\sum_{j=1}^{n}q_{j}\Omega_{l_{j}}^{\theta}(t_{j})\right)\right] = \exp\left(i\sum_{j=1}^{n}q_{j}m_{l_{j}} - \sum_{j,k=1}^{n}\frac{q_{j}q_{k}}{2}r_{l_{j},l_{k}}(t_{j} - t_{k})\right) + O\left(\lambda^{4}\right)$$

$$(75)$$

where the mean value  $m_l$  and the variance  $r_{l,l}(0)$  are

$$m_l = -\lambda^2 \ln \frac{Te^{3/2}}{l}$$
  $r_{l,l}(0) = \lambda^2 \ln \frac{Te^{3/2}}{l}$  (76)

and the autocorrelation function  $r_{l_1,l_2}(t)$  is

$$r_{l_1,l_2}(t) = \frac{\lambda^2}{2l_1l_2} \left( 2l_1l_2 \ln(Te^{3/2}) + g(t) - g(t+l_2) - g(t-l_1) + g(t+l_2-l_1) \right), \tag{77}$$

where

$$g(n) = n^2 \ln |n|.$$

We refer the reader to [58] for the exact meaning of  $O(\lambda^4)$  in Eq. (75).

Thus, up to a second order term in  $\lambda^2$ , the characteristic function Eq. (75) of  $\{\Omega_l^{\theta}(t)\}_{l,t}$  is the one of the "renormalized magnitude"  $\Omega_l(t)$  which is a gaussian variable defined by

$$\{\Omega_l(t)\}_{l,t} = \mathcal{N}(m_l, r_{l_1 l_2}(t))$$
 (78)

where the mean  $m_l$  is defined by Eq. (76) and the covariance  $r_{l_1,l_2}(t)$  by Eq. (77).

In that sense, the trading time  $\theta(t)$  can be written, in good approximation, as a log-normal trading time, i.e.,

$$\theta(t+l) - \theta(t) = e^{2\Omega_l^{\theta}(t)} \underset{Law}{\simeq} e^{2\Omega_l(t)},$$

or, for the MRW process X(t) itself,

$$\delta_l X(t) \underset{Law}{=} \epsilon(t) e^{\Omega_l^{\theta}(t)} \underset{Law}{\simeq} \epsilon(t) e^{\Omega_l(t)},$$

where  $\epsilon(t)$  is a gaussian white noise of variance  $\sigma^2 l$ .

This is one of the meaning we give to the  $\simeq_{Law}$  sign in Eq. (34). One can actually prove that these last approximations have other meanings than just a first order equality of the characteristic function. These meanings are stated in the next appendix. The corresponding proves can be found in [58].

## 2. The law and the moments

Actually, it can be shown [58] that,

- when  $\lambda^2 \to 0$ ,  $\{\Omega_l^{\theta}(t)/\lambda\}_{l,t}$  converges in law to the centered gaussian variable with covariance  $r(t, l_1, l_2)/\lambda^2$  as defined in Eq. (77),
- when  $\lambda^2 \ll 1$ , the n-order p-point (eventually conditional) moments of  $\Omega_l^{\theta}(t)$  are equal (first order term in  $\lambda^2$ ) to those of the renormalized magnitude  $\Omega_l(t)$  as defined in Eq. (78).
- when  $\lambda^2 \ll 1$ , as long as they are finite, the n-order p-point (eventually conditional) moments of  $\delta_l X(t) = \epsilon(t) e^{\Omega_l^{\theta}(t)}$  are equal (first order term in  $\lambda^2$ ) to those of the process  $\epsilon(t) e^{\Omega_l(t)}$  where  $\Omega_l(t)$  is the renormalized magnitude as defined in Eq. (78).

## 3. Cumulative distribution function of X(t)

Hereafter, we will note  $\gamma = \sqrt{\ln \frac{Te^{3/2}}{l}}$ , g a standardized gaussian variable (centered, with variance 1), n(x) and N(x) respectively the standard Gaussian density and cumulative distribution .  $He_i(x)$  will stand for the i-th Hermite polynomials.

We suppose the validity of the Edgeworth expansion for  $\frac{\Omega_t^{\theta}(t)}{\lambda \gamma}$ , i.e., there exists a constant C non depending on x which verifies

$$\left| \mathbb{P}\left[ \frac{\Omega_l^{\theta}(t)}{\gamma \lambda} < x \right] - N(x) - n(x)\gamma \lambda - n(x)R(x,\lambda) \right| < C\lambda^3$$
 (79)

where

$$R(x,\lambda) = \alpha_1 H e_1(x) \lambda^2 + \alpha_2 H e_2(x) \lambda + \alpha_3 H e_3(x) \lambda^2 + \alpha_5 H e_5(x) \lambda^2$$
(80)

So we can bound the difference of the cumulative distribution functions by

$$\left| \mathbb{P} \left[ \delta_l X(t) < z \right] - \mathbb{P} \left[ \varepsilon e^{\lambda \gamma g} < z \right] \right| \leqslant C \lambda^2 \tag{81}$$

Let us modify slightly the expression (79) by introducing the variable  $x - \gamma \lambda$ 

$$\left| \mathbb{P} \left[ \frac{\Omega_l^{\theta}(t)}{\gamma \lambda} + \gamma \lambda < x \right] - N(x) - n(x) \tilde{R}(x, \lambda) \right| < C\lambda^3$$
 (82)

where

$$\tilde{R}(x,\lambda) = \alpha_1 H e_1(x) \lambda^2 + \alpha_2 H e_2(x) \lambda + (\alpha_3 + \gamma \alpha_2) H e_3(x) \lambda^2 + \alpha_5 H e_5(x) \lambda^2$$
(83)

That leads us to the next estimation

$$\left| \mathbb{P}\left[ \delta_l X(t) < z \right] - \mathbb{P}\left[ \varepsilon e^{\gamma \lambda g - \gamma^2 \lambda^2} < z \right] \right| \leqslant C \lambda^3. \tag{84}$$

Let us note that  $\varepsilon e^{\gamma \lambda g - \gamma^2 \lambda^2} =_{law} \epsilon e^{\Omega_l(t)}$ , where  $\Omega_l(t)$  is the renormalized magnitude defined by Eq. (78).

- [1] B. B. Mandelbrot, Fractals and scaling in finance. Discontinuity, concentration, risk (Springer, New-York, 1997).
- [2] J. P. Bouchaud and M. Potters, *Theory of Financial Risk and Derivative Pricing* (Cambridge University Press, Cambridge, 2003).
- [3] Z. Ding, C. W. J. Granger, and R. F. Engle, J. of Empirical Finance 62, 83 (1993).
- [4] Z. Ding and C. W. J. Granger, Journal of Econometrics 73, 185 (1996).
- [5] R. Cont, Quantitative Finance 1, 223 (2001).
- [6] S. Ghashghaie, W. Breymann, J. Peinke, P. Talkner, and Y. Dodge, Nature 381, 767 (1996).
- [7] B. B. Mandelbrot, L. Calvet, and A. Fisher (1997), preprint, Cowles Foundation Discussion paper 1164.
- [8] A. Fisher, L. Calvet, and B. B. Mandelbrot (1997), preprint, Cowles Foundation Discussion paper 1166.
- [9] L. Calvet, A. Fisher, and B. B. Mandelbrot (1997), preprint, Cowles Foundation Discussion paper 1165.
- [10] L. Calvet and A. Fisher, Journal of Econometrics 105, 27 (2001)
- [11] L. Calvet and A. Fisher, Review of Economics & Statistics 84, 381 (2002).
- [12] L. Calvet and A. Fisher, Journal of Financial Econometrics 2, 49 (2004).
- [13] U. Frisch, Turbulence: The legacy of A.N. Kolmogorov (Cambridge University Press, Cambridge, 1995).
- [14] M. S. Taqqu and G. Samorodnisky, Stable non-Gaussian random processes (Chapman & Hall, New-York, 1994).
- [15] B. Castaing, Y. Gagne, and E. Hopfinger, Physica D 46, 177 (1990).
- [16] B. B. Mandelbrot, J. Fluid Mech. **62**, 1057 (1974).
- [17] B. B. Mandelbrot, C.R. Acad. Sci. Paris 278, 289 (1974).
- [18] B. B. Mandelbrot, J. Stat. Phys. 110, 739 (2003).
- [19] J. P. Kahane and J. Peyrière, Adv. in Mathematics 22, 131 (1976).
- [20] Y. Guivarc'h, C.R. Acad. Sci. Paris **305**, 139 (1987).
- [21] G. M. Molchan, Comm. in Math. Phys. 179, 681 (1996).
- [22] G. M. Molchan, Phys. Fluids 9, 2387 (1997).
- [23] Q. S. Liu, Asian J. of Math. 6, 145 (2002).
- [24] J. F. Muzy, J. Delour, and E. Bacry, Eur. J. Phys. B 17, 537 (2000).
- [25] J. Barral and B. B. Mandelbrot, Prob. Theory and Relat. Fields 124, 409 (2002).
- [26] J. F. Muzy and E. Bacry, Phys. Rev. E 66, 056121 (2002).
- [27] E. Bacry and J. F. Muzy, Comm. in Math. Phys. 236, 449 (2003).
- [28] P. Chainais, R. Riedi, and P. Abry, IEEE Trans. on Inf. Th. 51, 1063 (2005).
- [29] A. Arneodo, J. F. Muzy, and D. Sornette, Eur. Phys. J. B 2, 277 (1998).
- [30] E. Bacry, J. Delour, and J. F. Muzy, Phys. Rev. E 64, 026103 (2001).
- [31] E. Bacry, A. Kozhemyak, and J. F. Muzy (2006), preprint.
- [32] B. Lahermes, P. Abry, and P. Chainais, Int. J. of Wavelets, Multiresolution and Inf. Proc. 2, 497 (2004).
- [33] J. F. Muzy, E. Bacry, and A. Kozhemyak (2005), preprint.
- [34] T. Lux (2004), preprint, University of Kiel.
- [35] D. Sornette, Y. Malevergne, and J. F. Muzy, Risk Magazine (2002).
- [36] R. B. O. M. Dacorogna, U. Muller and O. Pictet, Nonlinear Modelling of High Frequency Financial Time Series (1998).
- [37] G. Zumbach and P. Lynch, Quantitative Finance 3 (2003).
- [38] L. Borland (2004), preprint, arXiv.org/cond-mat/0412526.

- [39] L. Borland, J. P. Bouchaud, J. F. Muzy, and G. Zumback, Wilmott Magazine (2005).
- [40] L. Borland and J. P. Bouchaud (2005), preprint, arXiv.org/cond-mat/0507073.
- [41] F. Schmitt, D. Schertzer, and S. Lovejoy, App. Stoch. Models and Data Anal. 15, 29 (1999).
- [42] Y. A. K. Matia and H. E. Stanley, Europhys. Lett. **61**, 422 (2003).
- [43] N. Vandewalle and M. Ausloos, Eur. Phys. J. B 4, 257 (1998).
- [44] T. D. Matteo, T. Aste, and M. Dacorogna, J. of Banking and Finance 29, 827 (2005).
- [45] J. F. Muzy, E. Bacry, and A. Arneodo, Phys. Rev. Lett. **67**, 3515 (1991).
- [46] E. Bacry, J. F. Muzy, and A. Arneodo, J. Stat. Phys. **70**, 635 (1993).
- [47] T. Lux, App. Econ. Lett. 3, 701 (1996).
- [48] M. Pasquini and M. Serva, Eur. Phys. J. B 16, 195 (2000).
- [49] Y. Malevergne, V. Pisarenko, and D. Sornette (2003), preprint, arXiv.org/physics/0305089.
- [50] P. Gopikrishnam, M. Meyer, L. A. N. Amaral, and H. E. Stanley, Eur. Phys. J. B 3, 139 (1998).
- [51] M. R. Leadbetter, G. Lingren, and H. Rotzén, Extremes and related properties of random sequences and processes (Springer Verlag, Berlin, 1983).
- [52] M. Ossiander and E. C. Waymire, The Ann. of Stat. 28, 1533 (2000).
- [53] T. Lux, Quantitative finance 1, 632 (2001).
- [54] L. Hansen, Econometrica **50**, 1029 (1982).
- [55] E. Bacry, A. Kozhemyak, and J. F. Muzy (2006), preprint.
- [56] P. R. Hansen and A. Lunde (2004), preprint, Brown University, Department of Economics Working Paper 01-04.
- [57] L. Calvet and A. Fisher, Journal of Econometrics 131, 179 (2006).
- [58] A. Kozhemyak, E. Bacry, and J. F. Muzy (2005), preprint.
- [59] P. Christoffersen, International Economic Review 39, 841 (1998).
- [60] B. Pochart and J. P. Bouchaud, Quantitative Finance p. 303 (2002).