Extreme values and fat tails of multifractal fluctuations

J. F. Muzy*
SPE UMR 6134, CNRS, Université de Corse, 20250 Corte, France

E. Bacry† and A. Kozhemyak‡
CMAP, Ecole Polytechnique, 91128 Palaiseau, France

(Received 13 September 2005; published 12 June 2006)

In this paper we discuss the problem of the estimation of extreme event occurrence probability for data drawn from some multifractal process. We also study the heavy (power-law) tail behavior of probability density function associated with such data. We show that because of strong correlations, the standard extreme value approach is not valid and classical tail exponent estimators should be interpreted cautiously. Extreme statistics associated with multifractal random processes turn out to be characterized by non-self-averaging properties. Our considerations rely upon some analogy between random multiplicative cascades and the physics of disordered systems and also on recent mathematical results about the so-called multifractal formalism. Applied to financial time series, our findings allow us to propose an unified framework that accounts for the observed multiscaling properties of return fluctuations, the volatility clustering phenomenon and the observed "inverse cubic law" of the return pdf tails.

DOI: 10.1103/PhysRevE.73.066114 PACS number(s): 02.50.−r, 05.40.−a, 05.45.−a, 89.65.Gh

I. INTRODUCTION

Statistics of extremes is an issue of prime importance in many situations where extreme events may appear to have disastrous effects or to govern the main observations. Such situations can be found in a wide range of fields from physics (e.g., disordered systems at low temperature), geology (earthquakes), meteorology (rainfalls, storms), insurance, or finance [1–3]. Extreme events are particularly relevant for random phenomena involving a probability density function (pdf) which tails decrease very slowly and roughly follow a power-law. Such heavy tailed distributions have been observed in many natural phenomena. An important question concerns therefore the estimation and interpretation of pdf tail exponent as well as the identification of mechanisms leading to them. These problems are at the heart of an increasing number of works [3–6]. Probabilistic and statistical questions related to very high or very low values of random variables are addressed within the framework of extreme value theory. This theory has been originally developed for independent identically distributed (iid) random variables and more recently extended to stationary processes where independence condition has been relaxed [1]. However, when correlations are not weak enough very few results are known.

In this paper we aim at studying the statistics of extreme events and the (fat) tail exponent of fluctuations associated with multifractal random processes. Multifractal extreme fluctuations are interesting because they represent an example of strongly correlated random variables that do not satisfy standard mixing conditions of extreme value theory. But multifractals are also interesting because they are widely used to model of self-similar phenomena displaying multiscaling properties. Our purpose is to study probability of extreme event of multifractal processes and the associated power-law tail exponent. We will show that for such processes, the pdf tail exponent value observed (estimated) from experimental data may be different from the value associated with the unconditional theoretical pdf. We examine different experimental conditions depending upon the size of the observed sample L, the correlation length T, and the observation scale τ. We emphasize that, under most usual conditions, the estimated tail exponent is smaller than the exponent one would expect without correlations. This result is intimately related to some non-self averaging property of usual tail exponent estimators and is the analog of the glassy behavior observed at low temperatures in disordered systems. We apply our phenomenological framework to multifractal models of asset return fluctuations and show that the well known "inverse cubic law" of pdf tails can be naturally explained in terms of volatility correlations.

The paper is organized as follows. In Sec. II, we briefly review the main results about extreme value theory and the commonly used tail exponent estimators. Multifractal processes, multiplicative cascades and their main mathematical properties are reviewed in Sec. III. In Sec. IV we build an extreme value theory for multifractal cascades. The problem of the estimation of the power-law tail exponent associated with the cumulative probability distribution of multifractal fluctuations is addressed in Sec. V. In Sec. VI we illustrate our phenomenology by numerical examples of continuous as well as discrete cascades. Application to finance is considered in Sec. VII while Sec. VIII contains concluding remarks and questions for future research. Auxiliary computations or technical material are reported in Appendixes.

II. FAT TAILS AND EXTREME VALUE STATISTICS

Let us briefly review the main estimators used to characterize the power-law tail behavior of some probability distri-

*Electronic address: muzy@univ-corse.fr
†Electronic address: emmanuel.bacry@polytechnique.fr
‡Electronic address: alexey@cmapx.polytechnique.fr

bution. Let $F(x)=\mathbb{P}[Z\leqslant x]$ be the cumulative probability distribution (cdf) of some random variable Z. The variable Z is said to be of power-law type tail or Pareto type tail if, when $x\to+\infty$,

$$1-F(x)\sim Cx^{-\mu}, \tag{1}$$

where C is a positive normalization constant or a slowly varying function. The exponent $\mu>0$ is called the *tail exponent* of the distribution. The problem addressed in this section concerns the estimation of this exponent from empirical data.

A simple, widely used method relies on the so-called "Zipf" or "rank-frequency" plots (see, e.g., Refs. [3,5,7]): Let $Z_1\ldots Z_N$ be N iid samples characterized by the same distribution function $F(x)$. Let us denote $X_1\geqslant X_2\geqslant\cdots\geqslant X_N$ the rank ordered values of Z_i (sorted in descending order). Then, if the asymptotic behavior of $F(x)$ is Pareto-like [as in Eq. (1)], one has for $1\leqslant j\leqslant k\ll N$

$$1-F(X_j)\sim\frac{j}{N}\Rightarrow X_j\sim\left(\frac{j}{N}\right)^{-1/\mu} \tag{2}$$

and therefore μ can be simply estimated as the slope of the Zipf plot $[\ln X_j,\ln(j)]$, $j=1,\ldots,k$ (see Ref. [7] for exact results). In the following, we refer to this estimator as the power-law fit estimator [8]. Since this estimator is biased [8,5], one should use alternative estimators.

Actually, there are many alternative tail exponent estimators [2]. The most commonly used are Hill or Pickands estimators that are defined as follows. Let $k(N)=o(N)$ be the maximum X rank used to estimate μ, the Hill estimator is simply

$$\mu_H(k,N)=\frac{k-1}{\sum_{i=1}^{k-1}\ln(X_i/X_k)} \tag{3}$$

while the Pickands estimator is

$$\mu_P(k,N)=\frac{\ln(2)}{\ln\left(\frac{X_k-X_{2k}}{X_{2k}-X_{4k}}\right)}. \tag{4}$$

The mathematical study of these estimators (consistency, bias, asymptotic normality) relies upon extreme value theory [1,2], i.e., the theory that deals with maxima and minima properties of random variables. According to this theory, the maximum value Y of N iid random variables (normalized properly), has a probability density function that asymptotically belongs to the Fisher-Tippett's extreme value distribution class. According to the shape of the pdf of Y, the pdf of the maximum can either be of the Frechet type $[F_Y(x)=e^{-x^{-\mu}}]$, of the Gumbel type $[F_Y(x)=e^{-e^{-x}}]$ or of the Weibull class [1].

Extreme value theory has also been extended to dependent (or correlated) stationary random processes [1]. Under mixing conditions ensuring asymptotic independence of maxima, it can be shown that the limit theorems established in the iid case still hold. The key difference is that the number N of independent variables is replaced by an "effective" number $N\theta$. The value $0<\theta<1$ that quantifies the effect of dependence is called the "extremal index" [1]. In the case of Gaussian processes, these theorems hold provided the covariance function $\rho(x)$ decreases sufficiently fast for large lags x, i.e.,

$$\rho(x)\ln(x)\underset{x\to+\infty}{\to}0. \tag{5}$$

A simple intuitive justification of $1/\ln(x)$ as the limiting case for the validity of standard extreme value theorems is provided in Ref. [9]. The main purpose of this paper is to try to understand how these theoretical results apply to data sampled from a multifractal process.

III. MULTIFRACTAL PROCESSES

Multifractal processes are random functions that possess non trivial scaling properties. They are now widely used models in many areas of applied and fundamental fields. Well known examples are turbulence, Internet traffic, rainfall distributions or finance. For the sake of simplicity we will consider only non-decreasing multifractal processes (often referred to as multifractal *measures* though their variations are not bounded) denoted hereafter $M(t)$. More general multifractal processes can be conveniently built as a simple Brownian motion $B(t)$ compound with the measure $M(t)$ considered as a stochastic time: $X(t)=B[M(t)]$. The statistical properties of $X(t)$ can be directly deduced from those of $M(t)$ (see, e.g., Ref. [10,11]).

A. Multiscaling

Multifractal processes are characterized by multiscaling properties of their variations. More precisely, if one defines the increments of $M(t)$ at scale τ, $M(t,\tau)=M(t+\tau)-M(t)$, multifractality can be loosely defined from the scaling behavior of the moments of $M(t,\tau)$:

$$\mathbb{E}[\,M(t,\tau)^q)]\underset{\tau\to0}{\sim}\tau^{\zeta(q)}, \tag{6}$$

where $q\in\mathbb{R}$ (more precisely q must belong to the range where previous moments are finite) is the order of the moment and the exponent $\zeta(q)$ is some nonlinear convex function often called the multifractal exponent spectrum. The simplest example of such function is the so-called log-normal spectrum for which $\zeta(q)$ is a simple parabola

$$\zeta(q)=\left(1+\frac{\lambda^2}{2}\right)q-\frac{\lambda^2}{2}q^2. \tag{7}$$

The coefficient λ^2 quantifies the curvature of $\zeta(q)$ (and hence the multifractality of the process) that is constant in the log-normal case. In the general case, one often calls $-\zeta''(0)$ the intermittency coefficient. Let us notice that, because of Hölder inequality for moments the scaling (6) with a nonlinear convex $\zeta(q)$ cannot hold at arbitrary large scales τ but is valid only in a domain bounded by some large scale T [actually the limit $\tau\to0$ in Eq. (6) must be understood as $\tau/T\ll1$]. This scale T will be called the "integral scale."

B. Singularity spectrum and multifractal formalism

Let us recall some classical results about the multifractal formalism. This formalism was introduced in the early 1980's by Parisi and Frisch (see, e.g., Ref. [12,13]) in order to interpret the above multiscaling properties of the moments in terms of pointwise regularity properties of the paths of the process $M(t)$. Let us define the local Hölder exponent $\alpha(t_0)$ at point (or time) t_0 as [52,53]

$$M(t_0,\tau) \underset{\tau\to 0}{\sim} \tau^{\alpha(t_0)}. \tag{8}$$

The limit $\tau\to 0$ means $\tau \ll T$ where T is the integral scale. The singularity spectrum $f^\star(\alpha)$ can be introduced as the fractal (Hausdorff or packing) dimension of the iso-Hölder exponents sets

$$f^\star(\alpha) = \dim\{t, \alpha(t) = \alpha\}. \tag{9}$$

Roughly speaking, this equation means that at scales $\tau \ll T$, the number of points where $M(t,\tau) \sim \tau^\alpha$ is

$$N(\tau,\alpha) \sim \tau^{-f^\star(\alpha)}. \tag{10}$$

When then multifractal formalism holds, $f^\star(\alpha)$ and $\zeta(q)$ as defined in Eq. (6) are basically Legendre transform one to 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

$$f^\star(\alpha) = f(\alpha), \quad \forall\ \alpha^\star \geqslant \alpha \geqslant \alpha_\star,$$

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

$$\alpha_\star = \inf\{\alpha, f^\star(\alpha) = 0\},$$

$$\alpha^\star = \sup\{\alpha, f^\star(\alpha) = 0\}.$$

In the following sections, we will use the fact that q can be interpreted as a value of the derivative of $f(\alpha)$ and conversely α is a value of the derivative of $\zeta(q)$: for a given value of $q=q_0$ one has, from previous Legendre transform relationship and thanks to the convexity of $\zeta(q)$:

$$f(\alpha_0) = 1 + q_0\alpha_0 - \zeta(q_0), \tag{11}$$

$$\alpha_0 = \frac{d\zeta}{dq}(q_0), \tag{12}$$

$$q_0 = \frac{df}{d\alpha}(\alpha_0). \tag{13}$$

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

$$\zeta^\star(q) = \begin{cases} \zeta(q) & \text{for } q^\star \leqslant q \leqslant q_\star, \\ \alpha_\star q & \text{for } q > q_\star, \\ \alpha^\star q & \text{for } q < q^\star, \end{cases} \tag{14}$$

where

$$q_\star = \frac{df}{d\alpha}(\alpha_\star), \tag{15}$$

$$q^\star = \frac{df}{d\alpha}(\alpha^\star). \tag{16}$$

It is important to point out that, experimentally, under usual conditions, only $\zeta^\star(q)$ [and not $\zeta(q)$] can be estimated (see, e.g., Refs. [14–16]).

C. Cascades

The paradigm of multifractal measures are multiplicative cascades originally introduced by the Russian school [17–19] for modeling the energy cascade in fully developed turbulence. After the early works of Mandelbrot [20–22], a lot of studies have been devoted to discrete random cascades on both physical [13,23–25] and mathematical grounds [14,26–30]. Very recently, continuous versions of these cascades have been defined: they share exact multifractal scaling with discrete cascades but they display continuous scaling and possess stationary increments [10,11,31–34]. Let us summarize the main properties of these constructions and set some notations.

The simplest discrete multifractal cascade can be constructed as follows: one starts with an interval of length T where the measure is uniformly spread, and split the interval in two equal parts: On each part, the density is multiplied by (positive) iid random factors W. Each of the two subintervals 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 $T2^{-n}$ by a binary sequence $k_1\cdots k_n$, with $k_i=0,1$, the "mass" of this interval (denoted as $I_{k_1\cdots k_n}$) is simply

$$M_n(I_{k_1\cdots k_n}) = 2^{-n}\prod_{i=1}^{n} W_{k_1\cdots k_i} = 2^{-n} e^{\sum_{i=1}^{n}\omega_{k_1\cdots k_i}} \tag{17}$$

where all the $W_{k_1\cdots k_i}=e^{\omega_{k_1\cdots k_i}}$ are iid such that $\mathbb{E}[W]=1$. Peyrière and Kahane [26] proved that this construction converges almost surely towards a stochastic non decreasing process M_∞ provided $\mathbb{E}[W\ln W]<1$. The multifractality of M_∞ (hereafter simply denoted as M) and the validity of the previously described multifractal formalism have been studied by many authors (see, e.g., Refs. [14,30]). An interesting additional property of cascades is that they are self-similar in the following stochastic sense

$$M[I_{k_1\cdots k_n}] \underset{\text{law}}{=} 2^{-1} W M[I_{k_1\cdots k_{n-1}}] \tag{18}$$

and therefore the order q moments of $M(t,\tau)$ behave as a power law

$$\mathbb{E}[M[0,T2^{-n}]^q] = 2^{-nq}\mathbb{E}[W^q]^n\mathbb{E}[M[0,T]^q]. \quad (19)$$

Comparison of Eqs. (19) and (6) with $\tau = T2^{-n}$ directly yields the expression of the spectrum $\zeta(q)$ in terms of cumulant-generating function of $\omega = \ln W$:

$$\zeta(q) = q - \ln_2(\mathbb{E}[W^q]). \quad (20)$$

Let us mention that the validity of the multifractal formalism has been rigorously proved for cascades. In Appendix A we provide explicit expressions of $\zeta(q)$ for various laws of W.

Let us note that in Refs. [35,36], the correlation function (covariance) $\rho(x)$ of the logarithm of the measure $[\ln M(I_n)]$ associated with such cascades has been shown to decrease slowly as

$$\rho(x) \simeq \ln(T/x) \text{ for lags } x \leqslant T. \quad (21)$$

If we consider that the data come from a sampling of successive independent cascades, the correlation function is 0 for lags above T, i.e.,

$$\rho(x) = 0, \quad x \geqslant T. \quad (22)$$

The integral scale T where cascading process "starts" can therefore be interpreted as a correlation length for the variations of the $\ln M$.

Because the previous construction involves dyadic intervals, and a "top-bottom" construction, it is far from being stationary. In order to get rid of this drawback, as already mentioned, some continuous cascade constructions have been recently proposed and studied on a mathematical ground [10,11,31–34]. Without entering into details, we just want to mention that such continuous cascades involve a family of infinitely divisible random processes $\omega_l(t)$ which correlation function basically follows Eqs. (21) and (22). The process $e^{\omega_l(t)}$ is the analog of the density satisfying the self-similarity:

$$e^{\omega_{sl}(st)} \underset{\text{law}}{=} e^{\Omega_s} e^{\omega_l(t)},$$

where Ω_s is a random variable of same law as $\omega_l(t)$ and independent of it. Martingale theory allows one to prove the convergence of the continuous process $M[0,t] = \lim_{l\to 0}\int_0^t e^{\omega_l(u)}du$.

Let us mention that the validity of the multifractal formalism has been established by Molchan by for discrete cascades [14,28] and generalized by Barral and Mandelbrot for continuous cascades [33]. In remaining of the paper, we will be using indifferently the classical top-bottom model (i.e., we will consider that the data come from a sampling of a succession of independent realizations of the same cascade process) and the continuous cascade model (i.e., we will consider that the data simply come from a sampling of a continuous cascade). Most of the arguments will be done using the first model while most of numerical examples will be performed on data of the second model.

D. Cascades have fat tails

Let us first emphasize that, as first originally noticed by Mandelbrot [20], the unconditional law of M can have a power law tail (see also Ref. [37]). Indeed, a simple argument involving the self-similarity of the limit measure allows one to obtain a simple bound: since the measure is additive, $M[0,1] = M[0,1/2] + M[1/2,1]$, for $q > 1$ one gets

$$\mathbb{E}[M[0,1]^q] \geqslant 2\mathbb{E}[M[0,1/2]^q] = 2^{1-\zeta(q)}\mathbb{E}[M[0,1]^q] \quad (23)$$

and finally $\mathbb{E}[M[0,1]^q] < +\infty \Rightarrow \zeta(q) \geqslant 1$. The reverse implication is basically true. It is, however, trickier to obtain and we refer the reader to Refs. [26,27] for a precise proof.

Since the power law tail exponent of a distribution is directly related to the maximum order finite moments, if one defines

$$\mu = 1 + \sup\{q, q > 1, \zeta(q) > 1\}, \quad (24)$$

then, cascades (discrete as well as continuous) have thick tails of exponent μ:

$$\mathbb{P}[M(t,\tau) \geqslant x] \underset{x\to+\infty}{\sim} x^{-\mu}. \quad (25)$$

(Let us note that, in full rigor, without any further assumption on $\mathbb{P}[M \geqslant x]$, previous implication should be written as $\mathbb{P}[M \geqslant x] = O(x^{-r})$, $\forall r < \mu$.) In the sequel, we will show that the exponent μ (of the unconditional law of M) is hardly observable in practice. As announced in the introduction, we shall see that the observed exponent depends on experimental conditions and turns out to be smaller than μ.

E. Defining the asymptotic limit $N \to +\infty$

As we have seen [Eq. (21)], the covariance function $\rho(x)$ of $\ln(M)$, the logarithm of the multifractal measure, has a very slow decay [slower than Eq. (5)] up to the integral scale T above which data are independent. Thus we expect that mixing conditions are not valid in the range from the sampling scale τ to the integral scale T while they hold when "looking" above scale T. Let us try to be more precise: the mixing conditions are conditions in the limit when the total number of samples N goes to $+\infty$. However, there are several ways to reach this asymptotic limit. Let L be the length of the whole sequence and τ the sampling scale. The total number of samples is therefore

$$N = \frac{L}{\tau},$$

while the number of integral scales N_T and the number of samples per integral scales N_τ are, respectively,

$$N_T = \frac{L}{T}, \quad N_\tau = \frac{T}{\tau}, \quad N = N_\tau N_T. \quad (26)$$

In order to control the relative values of N_T and N_τ, let us define the exponent χ as follows:

$$N_T \sim N_\tau^\chi \quad (27)$$

Let us note that this exponent χ has been already introduced by B.B. Mandelbrot as an "embedding dimension" [38] in order to discuss the concept of negative dimension (see below).

Thus, when $N \to +\infty$, if, for instance, $\chi = 0$, it means that we are in the case $\tau \to 0$ (while L and T are fixed), i.e., most

of the data are lying between the lags τ and T. Consequently, we do not expect the mixing conditions to hold. On the contrary, if $\chi=+\infty$, it means that we are in the case $L\to+\infty$ (while τ and T are fixed), i.e., most of the data are lying between the lags T and L and consequently the mixing conditions are satisfied.

Thus, as it will be discussed in the next two sections, in the first case ($\chi=0$), nothing guarantees that classical results of extreme value theory can be applied and that exponent tail estimators provide the expected values, whereas in the second case ($\chi=+\infty$), we expect the iid extreme value theorems to hold and the exponent tail estimators to be consistent. As we will see, one can go continuously from the first case to the second one. Actually we will show that both the extreme value distribution associated with cascades and the corresponding tail exponent estimator strongly depend on the value of χ.

IV. MULTIFRACTAL EXTREME VALUE STATISTICS

A. Cumulative probability distribution of the maximum

Let $\tau=T2^{-n}$. We call $X_1(N)$ the maximum value of $\ln(M(I_n)/\tau)$, where I_n is a short notation of the dyadic intervals $I_{k_1\cdots k_n}$ of size τ. Let $P(x,N)=\mathbb{P}[X_1(N)<x]$ be the cumulative distribution function (cdf) of $X_1(N)$, i.e., the probability that $X_1(N)$ is smaller than x.

Let us recall that we consider that the data come from a sampling of successive independent realizations of the same cascade measure. Thus T (the integral scale) is fixed whereas L (the total length of the data) and τ (the sampling scale) are varying. We want to study the statistics $P(x,N)$ of the data in the limit $N\to+\infty$. As we explained in the previous section, these statistics will strongly depend on χ [Eq. (27)], i.e., on the way N_τ and N_T go to $+\infty$. We fix $\chi=r/p$ and we choose the following parametrization:

$$N_\tau\sim 2^{pm},\quad N_T\sim 2^{rm},\quad N=N_\tau N_T\sim 2^{(p+r)m},$$

when the integer parameter $m\to+\infty$. Thus $P(x,m)\equiv P[x,N(m)]$ is the cdf of $\ln M(I_{pm})/\tau$ where $\tau=T2^{-pm}$.

It is easy to see that $P(x,m)$ can be simply expressed as

$$P(x,m)=[P'(x,pm)]^{2^{rm}}, \tag{28}$$

where $P'(x,n)$ is the cdf associated with the maximum of $\ln[M(I_n)/\tau]$ on a single integral scale instead of the whole data. In Appendix B, we show that this cdf satisfies the renormalization equation

$$P'(x,n+1)=[P'(x,n)*g(x)]^2, \tag{29}$$

where $*$ stands for the convolution product and $g(x)$ is the probability density of $\omega=\ln(W)$.

Let us notice that the initial condition $P(x,0)$ is precisely given by the law of $\ln(M)$ which exponential tail is described by Eq. (25):

$$P(x,0)\underset{x\to+\infty}{\sim}1-Ce^{-\mu x}. \tag{30}$$

At this point, let us make two important remarks. The first one concerns the extension of previous analysis to continuous cascades. Since this problem will be specifically addressed in a forthcoming study, let us just mention that such an extension can be performed along the same line as in Ref. [9] where the authors considered the equivalent of a log-normal continuous cascade (see also Ref. [39]). In that case, Eqs (28) and (29) become partial differential equations that turns to be the famous Kolmogorov-Petrovsky-Piscounov (KPP) equation which solutions are traveling fronts [40]. It is precisely these traveling front solutions and the techniques developed to study KPP-like equations, that will be used to solve Eq. (28). We thus expect the results established in the following to be valid for both continuous and discrete cascades.

Our second remark concern numerical simulations of discrete (or continuous) cascades. It is important to notice that, from a practical point of view, one cannot handle M_∞ but one builds a finite approximation of the cascade M, i.e., one builds the cascade up to step $n_{\max}$, introducing an ultraviolet cutoff $l=2^{-n_{\max}}$. It is easy to prove that the previous renormalization iteration and therefore, all the arguments developed in the sequel, remain valid provided the cutoff l is rescaled with the observation scale $\tau=T2^{-n}$, i.e., $n_{\max}=n+n_0$ [54]. The main difference with the limit $n_{\max}\to\infty$, is that the initial condition $P(x,0)$ is no longer given the asymptotic law (25) but depends on the chosen value of n_0.

B. Traveling front solutions

1. Case $\chi=+\infty$

When $r=1$ and $p=0$, Eq. (28) is simply the recurrence for the maximum cdf of iid random variables which solutions are Fisher-Tippett's fixed points reviewed in Sec. II. More precisely, since the initial condition is exponentially decreasing [Eq. (30)], when $N\to+\infty$, $P(x,N)$ will have a Gumbel shape. Consequently, the law of the maximum value of M will belong to the Frechet class with a tail exponent μ as defined in Eq. (24).

2. Equation (29) and the KPP equation

In order to solve the nonlinear problem (28) in the general case, let us first study Eq. (29) and its solutions. We will show that these solutions are traveling fronts so that, in some moving frame, Eq. (28) reduces to a standard (iid) extreme value problem.

Equation (29) is exactly satisfied by the cdf associated with the maximum value of random variables hierarchically correlated (generated additively along a Cayley tree). Such an equation has been studied in Refs. [39,41]. As recalled previously, in Ref. [9], the authors have considered a continuous log-normal analog that is nothing but a KPP equation. Thus Eq. (29), in the case where $g(x)$ is Gaussian, can be seen as a "discretized" version of the KPP equation.

It is well known that the KPP equation has traveling wavefront solutions connecting the homogeneous stable state to the unstable one. These solutions can be studied using linear stability analysis. As reviewed in Ref. [40], most of KPP features are somehow universal in the sense that the same analysis can be performed for a wide variety of

"reaction-diffusion" problems involving nonlinear integrodifferential or integrodifference equations. In particular the famous marginal stability criterion (see, e.g., Refs. [40,42,43] and below) for the selected front velocity can be generically applied for a large class of equations. Therefore, as explained in Refs. [41–43], one can apply the same techniques to get solutions of Eq. (29). Though these studies do not rely on a fully rigorous mathematical ground (such as KPP), we reproduce a similar analysis in the following.

Let us first notice that both Eqs (28) and Eqs (29) admit two homogeneous solutions $P=0$ and $P=1$, the first one being (linearly) stable while the second one is unstable. A given cdf $P(x,N)$ [or $P'(x,n)$] will therefore connect the stable state to the unstable one. As in the above cited references, one can consider a traveling front solution of Eq. (29) $P'(x,n)=P'_t(x-v'n)$ where the suffix t stands for "traveling" and v' is the front velocity. In order to compute this velocity, one performs a linear analysis in the vicinity of the unstable solution, i.e., in the limit $x\to+\infty$ where $P'_t(x-v'n)\to 1$.

If we denote

$$Q'_t(x) \equiv 1 - P'_t(x) \ll 1$$

then, to the first order in $Q'_t(x)$, Eq. (29) becomes

$$Q'_t(x-v') = 2Q'_t(x) * g(x) + O(Q'^2_t).$$

If one seeks for exponential solutions

$$Q'_t(x) = Ce^{-qx},$$

then

$$P'_t(x,n) = 1 - Ce^{-q[x-v'(q)n]}, \tag{31}$$

where v' and q satisfy the "dispersion" relation

$$v'(q) = q^{-1}\ln\left[2\int g(x)e^{qx}dx\right].$$

Notice that $g(x)$ is the law of the logarithms of the weights of the cascade construction, and then, from Eq. (20)

$$\ln\left(2\int g(x)e^{qx}dx\right) = \ln(2)[1+q-\zeta(q)]$$

which yields

$$v'(q) = \frac{\ln(2)[1+q-\zeta(q)]}{q} \tag{32}$$

In the next section, one shall see which velocity (or which q value) is selected

3. Case $\chi \neq +\infty$

From Eqs. (28) and (31), one gets the following traveling front solution for P:

$$P(x,m) = (1 - Ce^{-q[x-v'(q)pm]})^{2^{rm}},$$

where $v'(q)$ is given by Eq. (32).

If $y=x-v'(q)pm-q^{-1}r\ln(2)m$, one has

$$\ln P(y,m) = 2^{rm}\ln(1-Ce^{-qy}2^{-rm}) \underset{m\to+\infty}{\to} -Ce^{-qy}.$$

The cdf of y therefore converges to the Gumbel shape

$$P(y,m) \underset{m\to+\infty}{\to} e^{-Ce^{-qy}}$$

and thus $P(x,m)$ is itself a traveling Gumbel front

$$P(x,m) = e^{-Ce^{-q[x-mv(q)]}} \tag{33}$$

with the dispersion relation

$$v(q) = \frac{\ln(2)\{r+p[1+q-\zeta(q)]\}}{q}. \tag{34}$$

One can reproduce the same kind of analysis as in Refs. [41–45]. In Appendix C, we provide a sketch of proof that one can use a standard Aronson-Weinberger stability criterion to compute the velocity and q value which are selected: Let $q_{\min}$ be the unique positive q value such that

$$v(q_{\min}) = \min_{q>0} v(q) \tag{35}$$

and let

$$q_{\star,\chi} = \min(\mu, q_{\min}). \tag{36}$$

Then the selected velocity is simply $v(q_{\star,\chi})$ and the shape of the traveling front is (up to a subdominant correction described in Appendix C)

$$P(x,m) = e^{-Ce^{-q_\star[x-v(q_{\star,\chi})m]}}.$$

Let us recall (see Sec. IV A) that $P(x,m)$ corresponds to the cdf of $\ln M(I_n)/\tau$, where $\tau=T2^{-n}$ and $n=pm$. Let us notice that from a practical point of view, the cascade step n is not observable and it is therefore useful to translate these results as functions of the "real" scale τ.

4. From $\ln M(I_n)/T2^{-n}$ to $\ln M([0,\tau])$

The velocity $v(q)$ is a velocity related to the parameter m, i.e., $P(x,m)=P[x-v(q)m]$. For practical applications, it is useful to compute the velocity $v_\tau(q)$ as respect to the "observable" scale $\ln(T/\tau)$, i.e., $P(x,m)=P[x-v_\tau(q)\ln(T/\tau)]$. Clearly, $v_\tau(q)=v(q)m/\ln(T/\tau)$. Since $\tau=T2^{-n}$ and $n=pm$, one gets $v_\tau(q)=v(q)/p\ln(2)$. Now, if we switch from the cdf of $\ln M(I_n)/\tau$ to the cdf of $\ln M(I_n)$, one finally obtains the velocity

$$v_{\ln M}(q) \equiv v(q)/p\ln(2) - 1 = -\frac{\zeta(q)-1-\chi}{q}.$$

Let us note that the minimum of $v_{\ln M}(q)$ [i.e., the minimum of $v(q)$] is reached for $q_{\min}$ satisfying

$$-\chi = 1 + q_{\min}\frac{d\zeta(q_{\min})}{dq} - \zeta(q_{\min}).$$

One can also notice that $1+q\frac{d\zeta(q)}{dq}-\zeta(q)$ is the Legendre transform of $\zeta(q)$. We know from Sec. III B that this Legendre transform is nothing but the spectrum $f(\alpha)$. According

to the results of this section [and particularly Eq. (11)] it can be easily seen that this value of $q_{\min}$ corresponds to a singularity exponent $\alpha_{\min}$ with

$$f(\alpha_{\min}) = -\chi \tag{37}$$

$$q_{\min} = \frac{df}{d\alpha}(\alpha_{\min}) \tag{38}$$

$$\alpha_{\min} = v_{\ln M}(q_{\min}). \tag{39}$$

Let us remark that for $\chi=0$, the value of $q_{\min}$ is exactly the value $q_\star$ that is involved in the multifractal formalism as described at the end of Sec. III B [Eq. (15)]. The multifractal formalism reviewed in this section allows one to link a statistical quantity [$\zeta(q)$] to a geometrical one [the dimension $f^\star(\alpha)$]. The previous result can somehow be considered as an extension of the multifractal formalism provided one gives to $f(\alpha)$ a probabilistic interpretation. Such a probabilistic recasting of $f(\alpha)$ and the notion of "negative dimension" have been discussed more than a decade ago by Mandelbrot [38], Schertzer and Lovejoy [46]. In these works the exponent χ has been interpreted using a codimension as respect to the dimension of a "supersample." Let us notice that in Ref. [46], the authors already obtained the previous relations with an approach involving phenomenological arguments.

In summary, we established that the law of the maximum of $\ln(M)$ is (up to subdominant corrections) Gumbel and therefore, by a direct change of variable, that the law of the maximum of M at scale τ of a sample of length $\tau^{-\chi}$ is (up to subdominant corrections) Frechet when $\tau \to 0$ with a tail exponent $q_{\star,\chi}$ that depends on χ:

$$q_{\star,\chi} = \min(q_{\min}, \mu),$$

$$q_{\min} = \frac{df}{d\alpha}(\alpha_{\min}),$$

$$f(\alpha_{\min}) = -\chi, \tag{40}$$

where we have accounted for the influence of the initial condition through the exponent μ (see Appendix C).

Numerical evidence of the so-obtained results are reported in Sec. VI A.

V. ORDER STATISTICS, TAIL EXPONENT ESTIMATION AND MULTIFRACTAL FORMALISM

A. Notations

The idea underlying tail exponent estimation is to study statistics of the $k(N)$ observed extreme values in the asymptotic regime $N \to +\infty$ and $k(N) \to +\infty$. As in the previous section, the limit $N \to +\infty$ is taken as explained in Sec. III E, using the χ exponent [see Eqs. (26) and (27)].

For the $k(N) \to +\infty$ limit, let us parametrize $k(N)$ as

$$k(N) \sim N^\nu, \tag{41}$$

where $0 \leqslant \nu \leqslant 1$, the value $\nu=0$ being interpreted as $k(N) \sim \ln(N)$. We will denote $\hat{\mu}(\nu,\chi)$ the estimated tail exponent using one of the estimators reviewed in Sec. II for some given χ and ν (χ is defined in Sec. III E). At scale $\tau = T2^{-n}$, as in Sec. II, we denote $X_1 \cdots X_N$ the rank ordered values of $M(I_n)$ over dyadic intervals.

B. Tail exponent estimators

Let us compute the expected value of the tail exponent estimator (Power-law fit, Pickands, or Hill estimator) for fixed values of ν and χ. For the sake of simplicity we focus on Pickands estimator but the same argument equally applies to other estimators (e.g., Hill). Our heuristics will rely upon the multifractal formalism.

As recalled in Sec. III B, the positive part of the Legendre transform of $\zeta(q)$ corresponds to the singularity spectrum, i.e., the Hausdorff dimension of sets of isoregularity. Mandelbrot has proposed a long time ago, a probabilistic interpretation of negative dimensions [i.e., negative values of $f(\alpha)$] in terms of large deviation spectrum [38]. In this section, we will use this formalism in order to study the order statistics of multifractal fluctuations.

The Pickands estimator is simply defined as

$$\hat{\mu}(\nu,\chi) = \ln(2)\left(\ln \frac{X_k - X_{2k}}{X_{2k} - X_{4k}}\right)^{-1}. \tag{42}$$

Let us define $\alpha_{\nu,\chi}$ such that

$$k = N^\nu \sim N_T \left(\frac{\tau}{T}\right)^{-f(\alpha_{\nu,\chi})}.$$

Using the definition of χ according to which $N_T \sim (\tau/T)^{-\chi}$, it follows that $\alpha_{\nu,\chi}$ satisfies

$$f(\alpha_{\nu,\chi}) = \nu - \chi(1-\nu). \tag{43}$$

Let us now consider $\alpha'_{\nu,\chi} = \alpha_{\nu,\chi} + \epsilon_1$ such that

$$2k \sim N_T \left(\frac{\tau}{T}\right)^{-f(\alpha'_{\nu,\chi})},$$

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

$$q_{\nu,\chi} = \frac{df}{d\alpha}(\alpha_{\nu,\chi}). \tag{44}$$

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

$$4k \sim N_T \left(\frac{\tau}{T}\right)^{-f(\alpha''_{\nu,\chi})}$$

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

$$X_k - X_{2k} \sim X_k[1 - (\tau/T)^{\epsilon_1}]$$

$$X_{2k} - X_{4k} \sim X_k[(\tau/T)^{\epsilon_1} - (\tau/T)^{\epsilon_2}]$$

one finally gets for expression (42):

$$\hat{\mu}(\nu,\chi) \simeq \ln(2)\left(\ln \frac{1 - 2^{-1/q_{\nu,\chi}}}{2^{-1/q_{\nu,\chi}} - 4^{-1/q_{\nu,\chi}}}\right)^{-1} = q_{\nu,\chi}. \tag{45}$$

We then come to the conclusion that the Pickands tail estimator for a multifractal process strongly depends on the

choice of the rank $k=N^{\nu}$ and the exponent χ. This is another difference with standard theory for iid random variables. The same kind of phenomenology can also be applied to the power-law fit or Hill estimator. Let us notice that when $\nu=0$ we consider only the "extreme" of the "extremes" for the tail exponent estimation and therefore we recover the tail of the law of the maximum value as discussed previously. Indeed, from Eqs. (43)–(45), we get

$$\hat{\mu}(0,\chi) = f'(\alpha_{0,\chi}),$$

$$f(\alpha_{0,\chi}) = -\chi,$$

that is, exactly Eqs. (40). Numerical evidence illustrating these results are provided in the next section.

VI. NUMERICAL EXAMPLES

Most of our considerations rely on phenomenological scaling and asymptotic limit arguments. We therefore neglected prefactors and slowly varying corrections, i.e., finite size effects. In order to test our approach as well as to quantify the importance of finite-size corrections, it is therefore interesting to perform numerical simulations. In this section, we illustrate our purpose on specific examples such as those mentioned in Appendix A. Let us add that most of these simulations are performed on continuous cascades which construction algorithm is described in Ref. [10]. Indeed, as already explained at the end of Sec. IV A, one expects all our phenomenological results to remain valid when one goes from discrete to continuous cascades (mainly because the log-correlation structure of the logs remains the same).

A. Extreme value statistics

In order to check the results of Sec. IV, we have generated $N=5000$ independent realizations of log-normal continuous cascades of intermittency parameter $\lambda^2=0.2$ for various values of the parameters N_τ and N_T. From these independent samples, we have estimated the cdf of the maximum value of $\ln(M)$ at scale $\tau=T2^{-n}$. In Fig. 1(a) the cdf $P(x,n)$ [also referred to as $P(x,\tau)$] is plotted for $N_T=1$ and $T/\tau=8,16,32,64$. According to Eqs. (33) and (40), we expect to observe, as the scale $\ln(\tau/T)$ decreases, a front traveling towards negative x at velocity $\alpha_{\star,\chi=0}=\alpha_\star$ [one defines $\alpha_{\star,\chi}$ as the "conjugate" of $q_{\star,\chi}$, i.e., $\alpha_{\star,\chi}=\zeta'(q_{\star,\chi})$]. This behavior is well verified and can be quantitatively checked in Fig. 1(b) where all the fronts merge when plotted versus $x-\alpha_\star \ln(\tau/T)$ with $\alpha_\star \simeq 0.47$ as given according to Eqs. (40) and (7):

$$\alpha_\star = 1 + \frac{\lambda^2}{2} - \lambda\sqrt{2(1+\chi)}.$$

In Fig. 1(c) the same analysis is performed for $N_T=T/\tau$ and therefore $\chi=1$. One observes that the velocity decreases as expected (in that case $\alpha_{\star,\chi=1}\simeq 0.2$). For χ large enough, $\alpha_{\star,\chi}$ can become negative as illustrated in Fig. 1(c) where we have reproduce the plot of Fig. 1(a) for $\tau=8,16,32$ and $\chi=3$. According to previous formula, $\alpha_{\star,\chi=3}\simeq -0.16$, a value

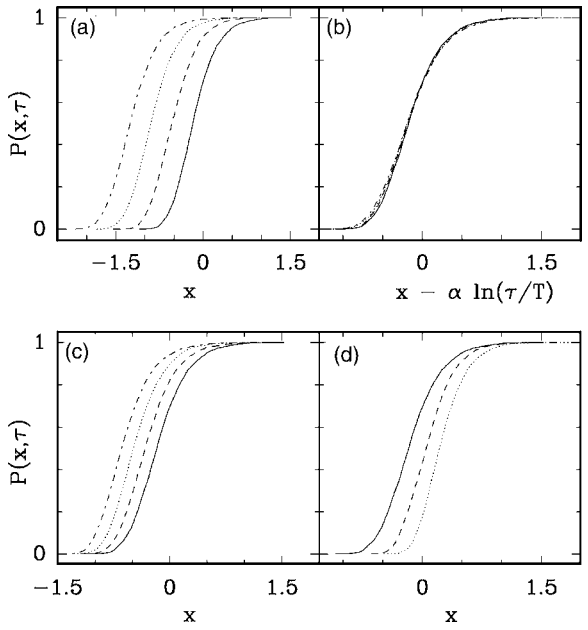

FIG. 1. Cumulative probability distribution of the maximum of $\ln(M)$ at scale $\tau=T2^{-n}$, $P(x,n)$ [also referred to as $P(x,\tau)$], as a function of x for various values of τ for a continuous log-normal cascade with $\lambda^2=0.2$. (a) $N_T=1$ and $\mathrm{In}_2(T/\tau)=3$ (continuous line), 4, 5, 6 (dotted lines). According to the formalism of Sec. IV, the cdf should be a front moving toward $x<0$ at a "velocity" $\alpha_{\star,\chi=0}=\alpha_\star \simeq 0.47$. (b) All the cdf of Fig. (a) merge to a single curve when plotted in the "moving referential." (c) Same plot as (a) but with $\chi=1$. One expects a smaller velocity $\alpha_{\star,\chi=1}\simeq 0.2$. (d) Same plot as (a) and (c) for scales $\mathrm{In}_2(T/\tau)=3,4,5$ with $\chi=3$. One observes, as expected, a negative velocity $\alpha_{\star,\chi=3}\simeq -0.16$.

compatible with observations where one sees the front moving towards positive x values as the scale decreases. The fact that $\alpha_{\star,\chi}$ is negative for χ large enough can be easily understood as follows: because the measure is continuous, as τ decreases, the maximum is expected to go to zero, but if in the same time, the number of independent integral scales is increased, the maximum is expected to increase. The exponent χ controls the balance between these two opposite effects. For χ large enough, one explores smaller negative $f(\alpha)$ that can be associated with negative α's. Notice that in all plots one observes a slight change in the shape of the front as the scale goes to zero: this is not surprising because the asymptotic shape of the front depends on χ and is a priori different from the initial front at scale $T/\tau=8$ considered in Fig. 1.

According to Eq. (33), the asymptotic shape of the fronts should be Gumbel, i.e.,

$$P(x) = e^{-Ce^{-q_{\star,\chi}x}}$$

with parameter $q_{\star,\chi}$ for the log-normal model:

$$q_{\star,\chi} = \sqrt{\frac{2(1+\chi)}{\lambda^2}}.$$

Notice that $q_{\star,\chi}$ is an increasing function of χ. In Fig. 2, we have plotted $\ln\{-\ln[P(x)]\}$ versus x for various values of

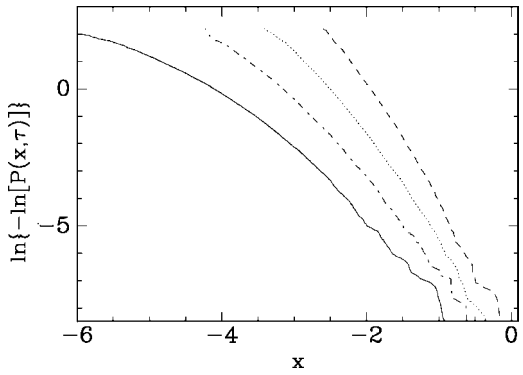

FIG. 2. "Gumbel plots" of the cumulative probability distribution of the maximum of $\ln(M)$ at scale $\tau=T2^{-n}$, $P(x,n)$ [also referred to as $P(x,\tau)$] for different values of χ. One sees that as χ increases from 0 to 1.5, $q_{\star,\chi}$, the slope in the tail increases significantly.

$T/\tau=N_\tau$ and N_T. We have chosen the values ($N_\tau=512$, $N_T=1$), ($N_\tau=512$, $N_T=8$), ($N_\tau=512$, $N_T=64$), and ($N_\tau=64$, $N_T=512$) that correspond, respectively (if one neglects prefactors in our analysis), to $\chi=0$, $\chi=0.33$, $\chi=0.67$, $\chi=1.5$. For an exact Gumbel law, one would expect a straight line. One clearly observes strong deviations to the Gumbel shape because the asymptotic regime is not reached but the straight part of the tails are relatively well estimated by the theoretical $q_{\star,\chi}$ values associated with the values of χ: one sees a systematic increase of $q_{\star,\chi}$ when one goes from $\chi=0$ to $\chi=1.5$.

B. Tail exponent estimators

Let us now check our results on the tail behavior of the estimated pdf from measured samples. Let us first compute $\hat{\mu}_{\nu,\chi}$ as given by Eq. (45), i.e., solve Eqs. (43) and (44) for, respectively, log-normal, log-Poisson, and log-gamma examples. After some simple algebra, we find in the log-normal case

$$\hat{\mu}_{\ln}(\nu,\chi)=\sqrt{\frac{2(1+\chi)(1-\nu)}{\lambda^2}}. \tag{46}$$

Let us notice that for $\nu=0$, one recovers previous tail exponent of the extreme values $\hat{\mu}(0,\chi)=q_{\star,\chi}$. In the case of log-Poisson statistics [Eq. (A3)], the computation leads to an expression for $\hat{\mu}(\nu,\chi)$ that involves Lambert $W(x)$ function. A perturbation series in the limit $\delta\to 0$ of this expression gives

$$\hat{\mu}_{\mathrm{lp}}(\nu,\chi)=\sqrt{\frac{2(1+\chi)(1-\nu)}{\lambda^2}}+2\delta\frac{(1+\chi)(\nu-1)-1}{3\lambda^2}+\cdots. \tag{47}$$

Along the same line, the value of $\hat{\mu}(\nu,\chi)$ can be computed in the case of the log-gamma cascade [Eq. (A4)] which again involves the second branch of the Lambert function W_{-1}:

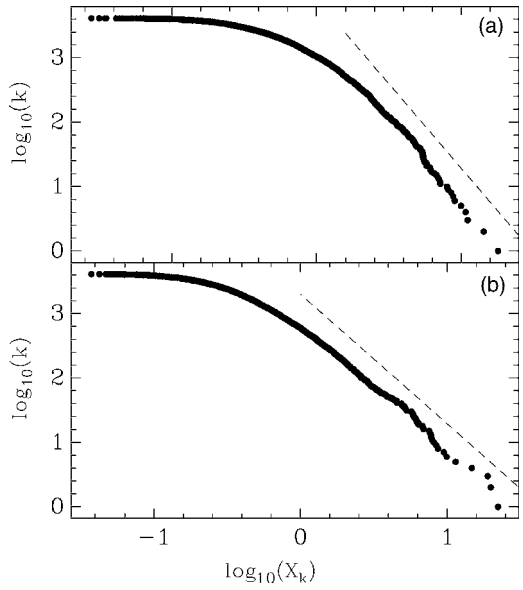

FIG. 3. Rank-frequency (Zipf) plots in log-log representation for a log-normal cascade (top) and a log-gamma cascade (bottom). The slope of the right linear part provides an estimate of the tail exponent μ. Dashed lines indicate analytical expectations (see text).

$$\hat{\mu}_{\mathrm{lg}}(\nu,\chi)=\beta[1-e^{1+W_{-1}(-e^{-1-(1+\chi)(\nu-1)/\lambda^2\beta^2})+(1+\chi)(1-\nu)/\lambda^2\beta^2}]. \tag{48}$$

A series expansion in the limit $\beta\to+\infty$ gives

$$\hat{\mu}_{\mathrm{lg}}(\nu,\chi)=\sqrt{\frac{2(1+\chi)(1-\nu)}{\lambda^2}}-4\frac{(\nu-1)(1+\chi)}{3\beta\lambda^2}+\cdots. \tag{49}$$

In Fig. 3, the rank ordering of the a log-normal and a log-gamma (with $\beta=4$) cascade processes with an intermittency coefficient $\lambda^2=0.2$ are plotted in doubly logarithmic scale (such plots are often referred to as "rank-frequency" plots or "Zipf" plots [3,5]). One clearly sees that the rightmost part of each distribution behaves as a power law. From Eqs. (46) and (48) (with $\chi=0$ and $\nu=0$), the slope of the plots should be, respectively, $\mu\simeq 3.1$ and $\mu\simeq 2.0$. One sees that these behaviors, reported on the figures as dashed lines, fit the data relatively well. One can observe that the scaling range associated with the log-gamma cascade is wider than for the log-Normal measure. Indeed, according to Appendix D [see Eq. (D1)], in the log-normal case, this range should be around $0.2pq_\star\ln(T/\tau)\simeq 0.7$ for $p=0.1$ while its values for log-gamma is expected to be $1.13pq_\star\ln(T/\tau)\simeq 2$ for $p=0.1$, i.e., more than two times wider than for the log-normal case. This difference can be visually checked in Fig. 3.

In Fig. 4, we have plotted the mean value of the tail exponent estimator (power-law fit estimator) as a function of the intermittency parameter λ^2 in the case of a log-normal (continuous) cascade for various values of k, N_T, and N_τ. The theoretical prediction (46) for $\nu=0$ is shown in continuous line for comparison. The mean has been evaluated on 10^3 independent cascade samples. In Fig. 4(a), one can see that for all parameter values, the curves have the same decreasing

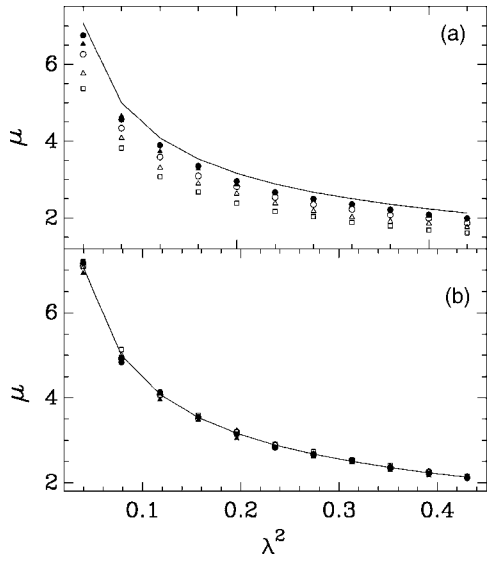

FIG. 4. Mean estimated tail exponent as a function of the intermittency coefficient λ^2 for continuous log-normal cascades. Each mean value is computed using 1000 cascade samples of length L. (a) Power-law fit estimator as a function of λ^2 for $T=1024$, $L=4096$, $k=16$ (●), $k=32$ (○), $k=64$ (△), $k=128$ (□), and $T=512$, $L=8192$, $k=32$ (▲). The continuous line corresponds to the theoretical prediction $\mu_{\ln}=\sqrt{2/\lambda^2}$ [Eq. (46)]. (b) The same as in (a) but each curve has been rescaled by a factor $\sqrt{2(1-\nu)(1+\chi)}$ according to Eq. (46) (see text).

behavior. All these curves collapse on the theoretical prediction $\sqrt{2/\lambda^2}$ as shown in Fig. 4(b), if one rescales each one by a factor $\sqrt{(1-\nu)(1+\chi)}$ where ν and χ are computed from expressions (41) and (27) by assuming that the prefactors are trivially 1. Empirically we find that the asymptotic phenomenology works quite well and prefactors or slowly varying behavior are negligible. As previously emphasized, the exact computation of prefactor values and finite size effects is beyond the scope of this paper and should involve more sophisticated mathematical tools (see Appendix C).

In Fig. 5, we compare the expression (48) to the behavior of the tail exponent estimator as a function of β in the case of a discrete log-gamma cascade: in that case, λ^2 is fixed ($\lambda^2=0.2$) while β varies. Once again, we can see that if we take

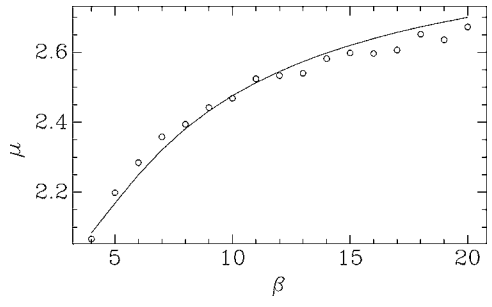

FIG. 5. Mean estimated (power-law fit) tail exponent (○) as a function of β for discrete log-gamma cascades with $\lambda^2=0.2$ and $T=2048$. The Hill parameter is $k=32$ and each mean value is computed from 1000 cascade samples. The continuous line corresponds to the analytical expression (48).

into account finite k value through the value of ν and χ close to their expected values, the analytical expression provides a very good fit of the data.

We have therefore illustrated, on two specific examples, that the phenomenology developed in previous section allows us to predict with a relative precision the tail behavior of multifractal measure samples. Let us now show how financial time series fluctuations can be described within this framework.

VII. APPLICATION TO FINANCIAL DATA

Multifractal models are, with many regards, well suited to account for return fluctuation of financial assets [7,47]. Among the "stylized facts" characterizing the asset return fluctuations, the phenomenon of "volatility clustering" (called heteroscedastisity in econometrics) is the most important one. One of the key points raised in Refs. [31,48] is that these volatility correlations are found empirically very close to the "log correlations" of continuous multifractal cascades [35]. Therefore the observed multiscaling properties of returns can be simply explained in terms of volatility persistence. In this section we want to stress that, for the same reason, namely the logarithmic shape of log-volatility correlations, the pdf of return fluctuations appear empirically as fat tailed with a rather small tail exponent.

In Refs. [31,48], we have shown that a parsimonious model of $X(t)$, some asset return value at time t, can be constructed as follows:

$$X(t) = B(M(t)) \tag{50}$$

where $B(t)$ is the standard Brownian motion and $M(t)$ is a multifractal continuous cascade as defined in Refs. [10,11,49]:

$$M(t) = \lim_{l\to 0} \int_0^t e^{\omega_l(t)} dt.$$

The process $\omega_l(t)$ plays exactly the same role as $\omega=\ln(W)$ in discrete cascades. It is easy to see that the variations of $M(t)$ in (50) can be interpreted as a stochastic variance. This quantity is called the "volatility" in finance [47].

If $M(t)$ has multiscaling (e.g. log-normal) properties then so do $X(t)$. 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 $\lambda^2=0.2$ while the integral scale T is typically around 1 year [31]. On the other hand, many studies relying on high frequency data or on thousands of daily stock returns, have revealed that the financial return pdf have heavy tails with a tail exponent in the interval [3,5]. This is the famous "inverse cubic" law for return fluctuations [47,50]. This observation lead to one of the main objections raised against the previous multifractal model for asset returns [47]. Indeed, the unconditional pdf of the volatility associated with a log-normal multifractal cascades of coefficient $\lambda^2=0.15$ has a tail exponent $\mu\approx 13$ [Eq. (A2)]. Within the "subordinated" model (50), this would mean that the tail of the return pdf is around $2\mu\approx 26$, i.e., close to ten times

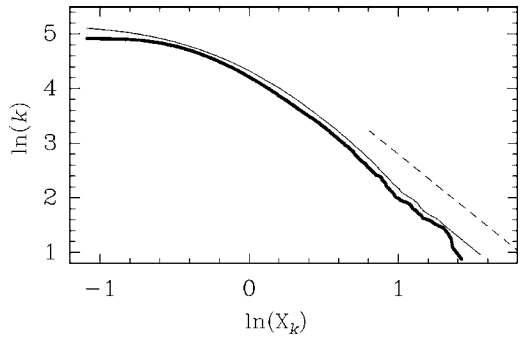

FIG. 6. Rank-frequency plot of CAC40 daily volatility estimates (●) as compared to similar plot for a log-gamma continuous cascade with $T=253$ days, $\lambda^2=0.05$, and $\beta=4$ (thin line). The fit of the extreme tail provides an estimation $\mu\simeq 2$ (dashed line).

the observed value! However, the main message of this paper is that for multifractal fluctuations, the observed extreme events are far from being distributed as they were independent. In particular we have shown in Sec. V that the estimators of the pdf tail exponent strongly depend on ν and χ and are "generically" smaller than μ and one must use, in the log-normal case, Eq. (46) instead of Eq. (A2). Typically, in finance $\tau\simeq 10^{-2}-1$ day, $T\simeq 1-2$ yr, and $L\simeq 10$ yr. Therefore, a rough estimate of χ and ν can be $\chi\simeq\nu\simeq 0.5$. If one uses these values in Eq. (46), one finds a typical value for the estimated tail within the log-normal model that is $2\hat{\mu}\simeq 6$. This is a value closer to the observations. In order to have a better fit of the tail behavior one could use a log-gamma model with the same intermittency coefficient and $\beta=4$ (see Appendix A). In that case, Eq. (48) gives an estimator value $2\hat{\mu}\simeq 3.6$ that agrees with observations.

In Fig. 6 is reported a rank-frequency plot of estimated daily volatilities associated with the 40 stock values composing to the French CAC 40 index. The data are daily "open," "high," "low," "close" quotes over a mean period of 10 yr. The daily volatilities are estimated using the widely used Garman-Klass method [51] and each volatility sample mean has been normalized to 1. For comparison the Zipf plot associated with a multifractal log-gamma measure with $T=1$ yr, $\lambda^2=0.2$, and $\beta=4$ has been also reported. One can see that both curves behave very similarly with a power-law tail exponent $\mu\simeq 2$. Let us note that for small volatilities values, the behavior of CAC40 volatility pdf is slightly different from the log-gamma cascade probably because of the height frequency noise in the Garman-Klass volatility estimates. This figure illustrates very well our result: there is no discrepancy between the value of the intermittency coefficient and the estimated pdf power-law behavior.

VIII. SUMMARY AND PROSPECTS

In this paper we have addressed the problem of extreme value statistics for multifractal processes. This problem is non trivial and possesses a rich phenomenology involving nonergodic behavior. In the case of multifractal processes, two important parameters govern the asymptotics: the overall sample length and the scale at which data are sampled. The exponent χ we have introduced, precisely quantifies how one defines the asymptotic limit as one changes these two parameters. The observed extreme value statistics result from a "competition" between an increase of the number of independent samples (which tends to increase typical extreme values) and a decrease of the sampling scale (which tends to decrease typical observed values). Consequently, the law of extremes continuously depends on χ, an exponent that turns out to corresponds to a value of the $f(\alpha)$ spectrum in the negative side, often interpreted as a "negative dimension."

This exponent naturally plays an important role when one wants to estimate the tail of the probability law associated with multifractal fluctuations. Using the phenomenology of the multifractal formalism we have shown that tail exponent estimators continuously depend on χ. They also depend on another exponent ν that quantifies "how many" extremes one uses for estimation. Under usual experimental conditions we have notably shown that the obtained exponent is usually smaller than the exponent expected from the unconditional law. Such nonergodic behavior are similar to those observed in the thermodynamics of disordered systems, at low temperature, under the freezing transition [9,39].

The arguments and methods we used in this paper are mostly phenomenological and rely upon large deviation type arguments, multifractal formalism and traveling front solution of nonlinear iteration equations. Beyond the need for sitting it on rigorous bases, they are other appealing mathematical prospects raised by our approach, such as the possibility to address finite size effects in multifractal scaling laws or to define a precise statistical framework for studying them.

As far as applications are concerned, we have provided a direct use of our results in the field of econophysics where multifractal models for asset returns are popular. We have shown that the observed fat tails of return pdf are well reproduced by a multifractal model designed to account for the volatility clustering phenomenon. Other fields where multifractal processes are involved can be potentially investigated along the same line. Conversely and perhaps more importantly, within this framework, multifractality appears as an alternative that can invoked to explain the origin of fat tails as observed in many fields of applied science [3–5].

ACKNOWLEDGMENTS

We thank J.P. Bouchaud for helpful discussions about the analogy of some multifractal problems addressed in this paper and similar problems arising in physics of disordered systems. This work has been supported by contract ACI *nouvelles interfaces des mathématiques*, Nb 54-03, from french ministère de l'éducation nationale.

APPENDIX A: LOG-NORMAL, LOG-POISSON, AND LOG-GAMMA MULTIFRACTAL MEASURES

In this appendix we provide three examples of multifractal statistics to which we will refer all along the paper. We consider the three infinitely divisible laws of $\ln W$ (or equivalently ω): normal, gamma and Poisson.

In the simplest case, $\omega=\ln W$ is normal of variance $\lambda^2/\ln(2)$ [and mean $-\lambda^2/2\ln(2)$ because $E(W)=1$]. Then $\mathbb{E}[W^q]=e^{-q\lambda^2/2+\lambda^2q^2/2}$ and we recover the expression of Eq. (7):

$$\zeta_{\ln}(q)=q(1+\lambda^2/2)-\lambda^2q^2/2. \tag{A1}$$

The intermittency coefficient in this case is simply λ^2. If one solves $\zeta(q)=1$, $q>1$ one gets

$$\mu=\frac{2}{\lambda^2}. \tag{A2}$$

In the second example, n is a Poisson random variable of intensity $\gamma\ln(2)$ and $\omega=m_0\ln(2)+n\delta$. Then $q-\ln_2\mathbb{E}[e^{q\omega}]=q(1-m_0)+\gamma(1-e^{q\delta})$. If one sets $\zeta(1)=1$ and $-\zeta''(0)=\lambda^2$, one obtains the spectrum

$$\zeta_{\mathrm{lp}}(q)=q\left(1+\frac{\lambda^2}{\delta^2}(e^{\delta}-1)\right)+\frac{\lambda^2}{\delta^2}(1-e^{q\delta}). \tag{A3}$$

Notice that when $\delta\to 0$, $\zeta_{\mathrm{lp}}\to\zeta_{\ln}$. It is easy to show that, when δ is negative and small enough, for all $q>1$, $\zeta_{\mathrm{lp}}(q)>1$ and therefore $\mu=+\infty$ in that case.

In the third example, ω is Gamma distributed: if x is a random variable of pdf $\beta^{\alpha\ln(2)}x^{\alpha\ln(2)-1}e^{-\beta x}/\Gamma[\alpha\ln(2)]$ and $\omega=x+m_0\ln(2)$, then $q-\ln_2\mathbb{E}[e^{q\omega}]=q(1-m_0)+\alpha\ln(1-q/\beta)$. By setting $\zeta(1)=1$ and $-\zeta''(0)=\lambda^2$, we have

$$\zeta_{\mathrm{lg}}(q)=q\left[1-\lambda^2\beta^2\ln\left(\frac{\beta-1}{\beta}\right)\right]+\lambda^2\beta^2\ln\left(\frac{\beta-q}{\beta}\right). \tag{A4}$$

Notice that $\zeta_{\mathrm{lg}}\to\zeta_{\ln}$ when $\beta\to+\infty$. The solutions of $\zeta_{\mathrm{lg}}(q)=1$ can be obtained in terms of Lambert W function (satisfying $W(x)e^{W(x)}=x$) and therefore the value of μ can be exactly computed as a function of λ^2 and β.

APPENDIX B: PROOF OF EQ. (28)

In this appendix we prove Eq. (28). Let us first study how the law of the maximum of $\ln[M(I_n)/\tau]$ inside one integral scale varies when one changes the scale $\tau=T2^{-n+1}$ to $\tau=T2^{-n}$. Let $I_n(k)$, $k=0\ldots 2^n-1$ denote the dyadic intervals of size $T2^{-n}$ of the interval $[0,T]$. From the cascade construction $\forall n$, the following stochastic equality can be easily proved:

$$M[I_n(k)]\underset{\mathrm{fdd}}{=}\frac{W_1}{2}M_1(I_{n-1}(k)),\quad k\in[0,2^{n-1}),$$

$$M[I_n(k)]\underset{\mathrm{fdd}}{=}\frac{W_2}{2}M_2[I_{n-1}(k-2^{n-1})],\quad k\in[2^{n-1},2^n],$$

where W_1, W_2, M_1, and M_2 are independent copies of W and M, respectively. The symbol fdd means an equality in law for all finite dimensional distributions. Therefore, if

$$X(n)=\max_k\{\ln[M(I_n(k))/\tau]\},$$

one has

$$X(n)\underset{\mathrm{law}}{=}\max[\ln(W_1)+X_1(n-1),\ln(W_2)+X_2(n-1)].$$

If $g(x)$ denote the law of $\omega=\ln(W)$ and $P'(x,n)=\mathbb{P}[X(n)<x]$, the previous equality can be rewritten as

$$P'(x,n)=\left[\int P'(x-z,n-1)g(z)dz\right]^2=[g*P'(x,n-1)]^2.$$

Now if one has $N_T=2^{rm}>1$ independent integral scales, the cdf of the maximum over 2^{rm} independent copies of law $P'(x,n)$ yields, when $n=pm$, Eq. (28).

APPENDIX C: ARONSON-WEINBERGER CRITERION AND FINITE SIZE EFFECTS ON FRONT SOLUTIONS OF EQ. (28)

In this Appendix we provide some additional technical details on the solutions of Eq. (28). We do not establish rigorous proofs but mostly recast some results from Refs. [40–45] to our problem. If one linearizes Eq. (28) in the tail region $x\to+\infty$, one obtains the following recursion for $Q(x)=1-P(x,m)$:

$$Q(x,m+1)=2^{p+r}Q(x,m)*g^{(p)}(x), \tag{C1}$$

where $g^{(p)}(x)$ is simply the product of p convolutions $g(x)*\cdots*g(x)$. If one decomposes $Q(x,m)$ on Fourier modes:

$$\hat{Q}(k,m)=\int e^{-ikx}Q(x,m)dx$$

then Eq. (C1) becomes

$$\hat{Q}(k,m+1)=2^{p+r}\hat{Q}(k,m)e^{pF(ik)},$$

where $F(ik)=\ln\mathbb{E}[e^{-ik\omega}]$ is the cumulant generating function of $\omega=\ln(W)$ the logarithm of cascade weights. The solution is therefore

$$\hat{Q}(k,m)=A(k)e^{m\ln(2)\{r+p[1+F(ik)/\ln(2)]\}},$$

where $A(k)$ is simply the Fourier transform of the initial condition. $Q(x,m)$ is obtained as the inverse Fourier transform

$$Q(x,m)=(2\pi)^{-1}\int e^{ikx}A(k)e^{m\ln(2)\{r+p[1+F(ik)/\ln(2)]\}}dk.$$

In a referential moving at velocity v (i.e., $x_m=x_0+vm$), the previous integral can be computed using a steepest descent method: One deforms the integral over the real axis to a contour in the complex plane of constant phase. In the limit $m\to+\infty$, the main contribution comes from saddle point of the function $F(ik)$ along this path

$$Q(x_m,m)\sim A(k_\star)e^{ik_\star x_m+m\ln(2)\{r+p[1+F(ik_\star)/\ln(2)]\}},$$

where $k_\star$ satisfies

$$-ik_\star v=\left.\frac{dF(ik)}{dk}\right|_{k=k_\star}. \tag{C2}$$

Moreover, if the front is stationary in the moving frame, the selected velocity should be such that the real part of the exponent is zero:

$$\mathrm{Re}\left\{ ik_\star mv + m\ln(2)\left[r + p\left(1 + \frac{F(ik_\star)}{\ln(2)} \right) \right] \right\} = 0 \quad \text{(C3)}$$

and thus, if v is real, by setting $k=iq$, and thanks to the equality

$$F(-q) = \ln(2)[q - \zeta(q)]$$

one can rewrite respectively Eqs. (C3) and (C2) as follows:

$$v(q) = \frac{\ln(2)\{r + p[1 + q - \zeta(q)]\}}{q},$$

$$\left.\frac{\partial v(q)}{\partial q}\right|_{q=q_{\star,\chi}} = 0.$$

The first equation is the dispersion relationship (34) while the second is the standard Aronson-Weinberger criterion stating that the selected velocity is the minimum velocity, a velocity that corresponds to the marginally stable solution (35).

Notice that previous argument assumes that both $F(ik)$ and $A(k)$ are analytical functions. If the initial condition decreases exponentially, i.e., $A(k)$ has a complex pole which imaginary part is μ [as in Eq. (30)], then, as discussed in Ref. [40], $q_{\star,\chi}$ must be replaced by $\min(q_{\star,\chi},\mu)$. This yields Eq. (36). As far as $F(ik)$ is concerned, it is easy to show that it is analytical in a strip around the real axis $\{z, \mathrm{Im}(z) \leqslant \mu\}$.

Let us finally remark that a general solution traveling at velocity v of Eq. (C1) can be written as

$$Q(z_m) = A_1 e^{-q_1 z_m} + A_2 e^{-q_2 z_m},$$

where $z_m = z_0 + vm$ and q_1 and q_2 are the two complex conjugated roots of the dispersion relation (34). When $v=v(q_{\star,\chi})$, the selected minimum velocity, these two roots merge and generically the solution behaves as [9,40,42]:

$$Q(z_m) = (Az_m + B)e^{-q_{\star,\chi} z_m}.$$

We see that the asymptotic shape of the front is not precisely Gumbel but has a subdominant correction factor.

Notice that the matching of the previous functional shape and the shape obtained in the previous saddle point analysis, when one accounts for the Gaussian corrections in the integration around the saddle point, leads to famous Bramson logarithmic correction to the front velocity [9,40,42]:

$$x_m = mv(q_{\star,\chi}) - \frac{3}{2q_{\star,\chi}}\ln(m).$$

This "universal" logarithmic correction to the velocity is well known for solutions of KPP equation. A specific analysis of these finite-size corrections to scaling is beyond the scope of this paper and will be reported in a forthcoming study.

APPENDIX D: SCALING RANGE FOR TAIL ESTIMATION

One question that naturally arises from the analysis made in section V is the question of stability of $\hat{\mu}(\nu,\chi)=q_{\nu,\chi}$ as a function of ν (and thus of k) and χ. This question is linked to the question of the scaling range associated with the log-log representation of Eq. (2). This problem is important for practical purpose. Let $p=\Delta q/q$ be the precision above which one can detect a tail exponent variation (for the sake of simplicity we will consider $p \simeq 0.1$). Let $\ln(S)$ be the scaling range over which one observes the power-law. According to our description, this corresponds to a variation of the Hölder exponent, i.e.,

$$\ln(S) \simeq \Delta\alpha \ln(T/\tau).$$

This variation corresponds therefore to a variation of q that is

$$\Delta q \simeq \Delta\alpha \frac{\partial q}{\partial h(q)} = \frac{\Delta\alpha}{\zeta''(q)|_{q=q_{\nu,\chi}}}$$

and finally, if the scaling range is such that $\Delta q = pq$, one obtains

$$\ln(S) = pq_{\nu,\chi}\ln(T/\tau)\zeta''(q_{\nu,\chi}). \quad \text{(D1)}$$

We see that this scaling range is related to the global scaling range rescaled by the factor that involves the intermittency coefficient evaluated at the value $q=q_{\nu,\chi}$.

[1] M. R. Leadbetter, G. Lingren, and H. Rotzén, *Extremes and Related Properties of Random Sequences and Processes* (Springer Verlag, Berlin, 1983).

[2] P. Embrechts, C. Kluppelberg, and T. Mikosch, *Modelling Extremal Events for Insurance and Finance* (Springer, Berlin, 1997).

[3] D. Sornette, *Critical Phenomena in Natural Sciences* (Springer, Heidelberg, 2003).

[4] J. P. Bouchaud, Quant. Finance **1**, 105 (2001).

[5] M. E. J. Newman, Contemp. Phys. **46**, 323 (2005).

[6] U. Frisch and D. Sornette, J. Phys. I **7**, 1155 (1997).

[7] B. B. Mandelbrot, *Fractals and Scaling in Finance. Discontinuity, Concentration, Risk* (Springer, New-York, 1997).

[8] M. I. Goldstein, S. Morris, and G. G. Yen, Eur. Phys. J. B **41**, 255 (2004).

[9] D. Carpentier and P. LeDoussal, Phys. Rev. E **63**, 026110 (2001).

[10] J. F. Muzy and E. Bacry, Phys. Rev. E **66**, 056121 (2002).

[11] E. Bacry and J. F. Muzy, Commun. Math. Phys. **236**, 449 (2003).

[12] U. Frisch, *Turbulence: The Legacy of A. N. Kolmogorov* (Cambridge University Press, Cambridge, 1995).

[13] T. C. Halsey, M. H. Jensen, L. P. Kadanoff, I. Procaccia, and B. I. Shraiman, Phys. Rev. A **33**, 1141 (1986).

[14] G. M. Molchan, Commun. Math. Phys. **179**, 681 (1996).

[15] B. Lahermes, P. Abry, and P. Chainais, Int. J. Wavelets, Multiresolution Inf. Proc. **2**, 497 (2004).

[16] J. F. Muzy and E. Bacry, (unpublished).

[17] A. M. Yaglom, Sov. Phys. Dokl. **2**, 26 (1966).
[18] A. Kolmogorov, J. Fluid Mech. **13**, 82 (1962).
[19] A. M. Obukhov, J. Fluid Mech. **13**, 77 (1962).
[20] B. B. Mandelbrot, J. Fluid Mech. **62**, 1057 (1974).
[21] B. B. Mandelbrot, C. R. Acad. Sci. Paris **278**, 289 (1974).
[22] B. B. Mandelbrot, J. Stat. Phys. **110**, 739 (2003).
[23] R. Benzi, G. Paladin, G. Parisi, and A. Vulpiani, J. Phys. A **7**, 3521 (1984).
[24] D. Schertzer and S. Lovejoy, in *Turbulence and Chaotic Phenomena in Fluids*, edited by T. Tatsumi (Elsevier, Amsterdam, 1984), pp. 505–512.
[25] C. Meneveau and K. R. Sreenivasan, Phys. Rev. Lett. **59**, 1424 (1987).
[26] J. P. Kahane and J. Peyrière, Adv. Math. **22**, 131 (1976).
[27] Y. Guivarc'h, C. R. Acad. Sci. Paris **305**, 139 (1987).
[28] G. M. Molchan, Phys. Fluids **9**, 2387 (1997).
[29] Q. S. Liu, Asian J. Math. **6**, 145 (2002).
[30] M. Ossiander and E. C. Waymire, Ann. Stat. **28**, 1533 (2000).
[31] J. F. Muzy, J. Delour, and E. Bacry, Eur. Phys. J. B **17**, 537 (2000).
[32] F. Schmitt and D. Marsan, Eur. Phys. J. B **20**, 3 (2001).
[33] J. Barral and B. B. Mandelbrot, Probab. Theory Relat. Fields **124**, 409 (2002).
[34] P. Chainais, R. Riedi, and P. Abry, IEEE Trans. Inf. Theory **51**, 1063 (2005).
[35] A. Arneodo, J. F. Muzy, and D. Sornette, Eur. Phys. J. B **2**, 277 (1998).
[36] A. Arneodo, E. Bacry, S. Manneville, and J. F. Muzy, Phys. Rev. Lett. **80**, 708 (1998).
[37] D. Schertzer and S. Lovejoy, J. Geophys. Res. **92**, 9692 (1987).
[38] B. B. Mandelbrot, Physica A **163**, 306 (1990).
[39] B. Derrida and H. Spohn, J. Stat. Phys. **51**, 817 (1988).
[40] W. Van-Saarloos, Phys. Rep. **386**, 29 (2003).
[41] M. Mobilia and P. A. Bares, Phys. Rev. E **64**, 045101(R) (2001).
[42] E. Brunet, Ph.D. thesis, Université de Paris VII, Paris, 2000.
[43] S. N. Majumdar and P. L. Krapivsky, Phys. Rev. E **65**, 036127 (2002).
[44] E. Brunet and B. Derrida, Phys. Rev. E **62**, 2597 (2000).
[45] S. N. Majumdar and P. L. Krapivsky, Phys. Rev. E **62**, 7735 (2000).
[46] D. Schertzer and S. Lovejoy, Physica A **185**, 187 (1992).
[47] J. P. Bouchaud and M. Potters, *Theory of Financial Risk and Derivative Pricing* (Cambridge University Press, Cambridge, 2003).
[48] E. Bacry, J. Delour, and J. F. Muzy, Phys. Rev. E **64**, 026103 (2001).
[49] B. B. Mandelbrot, Sci. Am. **280**, 50 (1999).
[50] P. Gopikrishnam, M. Meyer, L. A. N. Amaral, and H. E. Stanley, Eur. Phys. J. B **3**, 139 (1998).
[51] B. M. Garman and M. J. Klass, J. Business **3**, 67 (1980).
[52] S. Jaffard, SIAM J. Math. Anal. **28**, 944 (1997).
[53] In full rigor, the definition of pointwise Hölder regularity should account for possible local polynomial behavior of the function but we do not consider such eventuality here (see, e.g., Ref. [52]).
[54] In all numerical simulations considered in Sec. VI, typical values of n are in the range [3,12] and $n_{\max}$ is chosen within the interval [12,17].