NEW INSIGHTS INTO THE
ESTIMATION OF SCALING EXPONENTS *

B. Lashermes, P. Abry

CNRS UMR 5672, Laboratoire de Physique. ENS Lyon, France
tel : 33 47272 8493, fax : 33 47272 8080,
blasherm@ens-lyon.fr, pabry@ens-lyon.fr, perso.ens-lyon.fr/patrice.abry

P. Chainais

ISIMA-LIMOS UMR 6158 - Université Clermont-Ferrand II, Aubière, France
tel : 33 47340 5367, fax : 33 47340 5001,
pchainai@isima.fr, www.isima.fr/∼chainais

We study the statistical performance of multiresolution-based estimation procedures for the scaling exponents of multifractal processes. These estimators rely on the computation of multiresolution quantities such as wavelet, increment or aggregation coefficients. Estimates are obtained by linear fits performed in log of structure functions of order q versus log of scale plots. Using various and recent types of multiplicative cascades and a large variety of multifractal processes, we study and benchmark, by means of numerical simulations, the statistical performance of these estimation procedures. We show that they all undergo a systematic linearisation effect: for a range of orders q, the estimates account correctly for the scaling exponents; outside that range, the estimates significantly depart from the correct values and systematically behave as linear functions of q. The definition and characterisation of this effect are thoroughly studied. In contradiction with interpretations proposed in the literature, we provide numerical evidence leading to the conclusion that this linearisation effect is neither a finite size effect nor a infiniteness of moments effect, but that its origin should be related to the deep nature of the process itself. We comment on its importance and consequences for the practical analysis of the multifractal properties of empirical data.

Keywords: Scaling, Scaling Exponent, Scaling Exponent Estimation, Multifractal Process, Multiplicative Martingale, Multiresolution, Wavelet, Increment, Aggregation.

1. Motivation

Scaling.   During the last twenty years, scaling phenomena and scale invariance have been observed in a wide range of applications of very different natures (hydrodynamic turbulence, computer network teletrafic, body rhythms in biology,... to name but a few). In many applications, accurately measuring scaling exponents is a key issue, for classification and modelling of empirical data as well as for the analysis of the physical mechanisms producing scaling phenomena. Therefore, their detection, analysis and characterisation received considerable efforts and is still an active research area.

Most often, the practical definition of scaling in empirical time series $X(t)$ is based on multiresolution quantities (hereafter, $T_X(a,t)$), i.e., quantities that depend jointly on the time t and an analysis scale a. For instance, the $T_X(a,t)$ can stand for the increment, wavelet or box-agreggated coefficients of the process. Scaling phenomena are commonly associated to a power law dependence of statistical quantities of order q of the $|T_X(a,t)|$ with respect to the analysis scale a:

$$I\!\!E|T_X(a,t)|^q \simeq c_q|a|^{\zeta(q)}, \ \ a_m \leq a \leq a_M, \ \ q_m \leq q \leq q_M, \qquad (1.1)$$

$$\frac{1}{n_a}\sum_{k=1}^{n_a}|T_X(a,t_k)|^q \simeq c_q|a|^{\zeta(q)}, \ \ a_m \leq a \leq a_M, \ \ q_m \leq q \leq q_M. \qquad (1.2)$$

Though quite often overlooked, it is worth noting that such behaviours may be valid only within a finite range of scales $a \in [a_m, a_M]$ and for a finite range of orders $q \in [q_m, q_M]$, e.g., $I\!\!E|T_X(a,t)|^q$ may no longer exist beyond a critical q value.

Estimation procedures.   The practical study of scaling mainly consists in detecting power law behaviours as in Eqs. (1.1) and (1.2), in estimating the corresponding scaling exponents, in identifying the mathematical model (e.g., self similar processes, multifractal processes...) that better fits the data. The estimation of the scaling exponents is essentially performed in three steps. First, from the observed time series, one computes the multiresolution coefficients $T_X(a,t)$. Second, one computes the structure functions $S_n(q,a) = \frac{1}{n}\sum_{k=1}^{n}|T_X(a,k)|^q$. Third, one measures

the slope $\hat{\zeta}(q, n)$ in a $\log a$ versus $\log S_n(q, a)$ diagram.

MultiExponent MultiFractal Processes.   Unlike self-similar processes, multifractal ones cannot be defined through a single, generally valid definition. In the present work, we choose to use the following operational definition[a]: a process X is said to be multifractal when Eqs. (1.1) or (1.2) hold in the range of scales $0 \leq a \leq a_M$ and for a given range of qs. Note that this includes processes usually referred to as monofractal as special cases. However, we will not consider here all multifractal processes but will restrict ourselves to the subset defined by the fact that the scaling exponents $\zeta(q)$ depart from a strict linear behaviour in q[b]. This class will be hereafter referred to as *MultiExponent MultiFractal* (MEMF) processes for convenience, and is summarised as:

$$\text{MEMF: } \zeta(q) \neq qH. \tag{1.3}$$

The very example of such processes consists of the celebrated Mandelbrot's cascades. However, such constructions, as well as the processes that can be derived from them, suffer from important drawbacks: their increments are not wide sense stationary (a much desired property as far as the modelling of empirical data is concerned)[c] and their scaling behaviours are valid for a specific discrete set of scales only instead of holding continuously (i.e., for all scales) as suggested in Eq. (1.2). It has been suspected that some of the results reported in the literature on the behaviours of the estimators of scaling exponents observed on those specific cascades might be strongly related to their particularities. In the recent literature[11,41,7,8,15,14], new classes of multifractal processes were proposed, with known and a priori prescribed $\zeta(q)$, with stationary increments and continuous scale invariance. They are, hence, significantly renewing the possibilities and interests in studying the statistical per-

---

[a]We are aware that such a definition, led by empirical considerations, does not follow the usual definition for multifractals[22,23,36,9].
[b]Note that this class excludes a priori self-similar processes, Levy motions[22], multifractional Brownian motion[12].
[c]It implies that scaling take the form of Eq. (1.2) while the preferred form in Eq. (1.1) is not valid.

formance of the $\hat{\zeta}(q, n)$ and are motivating the present work.

Goals and methodology.   Two major classes of stochastic processes are commonly used to model scaling: self similar processes (see e.g.,[37]) versus multifractal processes (see e.g.,[36]). Estimation issues for the former class have been thoroughly addressed elsewhere (see e.g.,[1]) and are not considered here. Though considered in a restricted number of research articles[20,13,1,6], estimation procedures for the latter class received far less attention. Hence, the aims of this work are to present numerical studies that qualify, quantify and interpret the statistical performance of the multiresolution based estimation procedures $\hat{\zeta}(q, n)$, defined below, for the $\zeta(q)$, when they are applied to given classes of multifractal processes.

This is achieved by applying the $\hat{\zeta}(q, n)$ to a large number of independent realisations of identical multifractal reference processes. Statistical performance are inferred from averaging over the realisations. This benchmarking is performed for three categories of multiresolution based estimators, based on wavelets, increments and aggregation, for three types of multiplicative cascades (canonical Mandelbrot cascades, compound Poisson cascades, infinitely divisible cascades), from which four classes of processes can be constructed (density, measure, fractional Brownian motion in multifractal time, multifractal random walk).

Results.   Our first major result consists in showing the existence of a *linearisation effect* in the behaviour of the estimators as a function of q: the estimated exponents $\hat{\zeta}(q, n)$ account for scaling exponents only for values of q within an interval $q \in [q_*^-, q_*^+]$ and systematically behave according to an affine function of q outside this interval. The bounds $q_*^{\pm}$ are defined. We comment on the fact that strangely and despite its systematic nature, this effect has been almost totally overlooked in the huge literature related to the analysis of multifractal scaling in applications. However, it has originally been reported in a seminal work on cascades in turbulence by Mandelbrot[27] and then thoroughly and carefully studied in the

case of the Mandelbrot's cascades in a limited number of research papers.

In most of these works[38,39,40,42,28], the linearisation effect is related either to a finite size effect — there should exist a maximal observable singularity depending on the sampling rate — or to infiniteness of the moments of the process beyond a given statistical order q. Our second major result is to show that the numerical experiments presented here clearly and non ambiguously reject those two interpretations.

A restricted number of papers[17,21,30,31,33] studied theoretically the linearisation effect in the same specific case of the density of a Mandelbrot's cascade analysed with the only box-aggregation procedure. Our third major result lies in the fact that our characterisation of the linearisation effect not only falls in complete agreement with these theoretical studies but also suggests that they can be extended to a much wider context: new classes of cascades, new types of processes and new families of multiresolution based estimators. Following[33] and remarks borrowed from[11], we relate this linearisation effect to the very nature of the processes rather than to the estimation procedures themselves.

Outline.   Definitions of the estimators are given in Section 2. Section 3 summarises the definitions and properties of the MEMF processes (positive multiplicative martingales) actually used in the present work for the benchmarking of the estimation procedures. Empirical results and conjectures are reported in Section 4 while Section 5 proposes comments and interpretations on the origin, nature, practical importance and consequences of the linearisation effect.

2. Multiresolution based estimators for the scaling exponents

Multiresolution quantities.   Let X denote the scaling process under consideration. Let us start recalling that the multiresolution quantities $T_X(a, t; f_0)$ are obtained from comparisons, by means of inner products, between X and a collec-

tion of functions $\{f_{a,t}, t \in \mathbb{R}, a > 0\}$:

$$T_X(a, t; f_0) = \int_{\mathbb{R}} X(u) f_{a,t}(u) du, \quad \text{with } f_{a,t}(u) = \frac{1}{a} f_0 \left(\frac{u - t}{a} \right). \qquad (2.4)$$

Each specific choice of mother-function f_0 gives birth to the definition of a particular estimator. The three estimators considered here are obtained from:

$$
\begin{aligned}
&\text{EI(N), \textsc{Aggregation}}: f_0(u) = (\beta_0(u))^{*N}, \text{ where } \beta_0(u) = 1 \text{ if } 0 \le u < \tau_0 \\
&\text{EII(N), \textsc{Increment}}: \quad f_0(u) = (I_0(u))^{*N}, \text{ where } I_0(u) = \delta(u + \tau_0) - \delta(u) \\
&\text{EIII(N), \textsc{Wavelet}}: \quad f_0(u) = \psi_{0,N}(u), \quad \text{ where } \psi_{0,N}(u) \text{ is a standard mother wavelet,}
\end{aligned}
\qquad (2.5)
$$

where τ_0 is an arbitrary positive constant and $(f_0(u))^{*N}$, $N \in \mathcal{Z}^{*+}$, indicates that the function f_0 is convolved with itself $(N-1)$-times. A mother wavelet[26] is mainly characterised by its number of vanishing moments, an integer $N \ge 1$, such that:

$$\int_{\mathbb{R}} t^k \psi_{0,N}(t) dt \equiv 0, \; k = 1, \ldots, N-1, \quad \int_{\mathbb{R}} t^N \psi_{0,N}(t) dt \neq 0. \qquad (2.6)$$

It is well known that the selection of the number of vanishing moments and the possibility to vary it plays a key role in the practical analysis of scaling. This has been thoroughly discussed for the case of self similar or long range dependent processes[1]. To perform fair comparisons between estimators, it is natural to introduce N into $EI(N)$ and $EII(N)$, through $(\beta_0(u))^{*N}$ and $(I_0(u))^{*N}$, respectively[d].

Structure functions and dyadic grid.   From the multiresolution quantities $T_X(a, t)$, one defines the so-called structure functions:

$$S_n(q, a_j; f_0) = \frac{1}{n_j} \sum_{k=1}^{n_j} |T_X(a_j, t_{j,k}; f_0)|^q, \qquad (2.7)$$

where n denotes the observation duration $(0, n]$ of the process X (i.e., practically it means that X is available through its samples $\{X(1), \ldots, X(n)\}$) and n_j is the number of coefficients $T_X(a_j, t_{j,k}; f_0)$ at scale a_j, roughly $n_j \simeq n/a_j$. When the

---

[d]For the increments (i.e., for EII), N exactly is the number of vanishing moments , as in Eq. (2.6). For the aggregation procedure (i.e., for EI), the situation is different since $(\beta_0(u))^{*N}$ has strictly speaking 0 vanishing moments, whatever N. In this case, N mainly controls the regularity of the analysing function f_0, as in the wavelet case (cf.[26]). In the analysis of scaling, however, regularity plays a far less crucial role compared to that of the number of vanishing moments.

$\{T_X(a, t_k; f_0)\}_{k \in \mathcal{Z}}$ form stationary sequences at a given scale, the time averages $S_n(q, a_j; f_0)$, can be seen as estimators for the ensemble averages $\mathbb{E}|T_X(a, t)|^q$. Without loss of generality with respect to the results reported in Section 4, we have chosen to compute the multiresolution coefficients $T_X(a_j, t_{j,k}; f_0)$ on a discrete subset of points $(a_j, t_{j,k}) = (2^j, k2^j)$, known in the wavelet terminology as the dyadic grid. For $EIII(N)$, it amounts to compute the Discrete Wavelet Transform[e].

Definition of the estimators. The estimators consist in performing unweighted[f] linear regressions in log-log plots over the range of octaves $j \in [j_1, j_2]$:

$$\begin{cases} Y_{q,j} = \log_2 S_n(q, 2^j; f_0) \text{ versus } \log_2 2^j = j, \\ \hat{\zeta}(q, n) = \sum_{j=j_1}^{j_2} w_j Y_{q,j}, \\ w_j = (S_0 j - S_1)/(S_0 S_2 - S_1^2), \text{ with } S_m = \sum_{j_1}^{j_2} j^m, \ m = 0, 1, 2. \end{cases} \tag{2.8}$$

Comments. By definition, $EI(N)$ can be applied only to first order stationary processes with positive values, (it can hence be applied only to the density Q_r defined in Section 3). Therefore, the box-aggregated coefficients are strictly positive random variables with $P_{T_X(a,t)}(T = 0) \equiv 0$ and hence their moments are likely to exist a priori for all $q \in \mathbb{R}$. $EI(N)$ can hence be defined a priori with $q \in \mathbb{R}$. Conversely, by construction, the $T_X(a, t)$ for $EII(N)$ and $EIII(N)$, i.e., the increments and the wavelet coefficients of X, form stationary sequences for all non stationary processes whose increments of order smaller than N are stationary (they can hence be applied to the processes A, V_H, Y_H defined in Section 3). However, the $T_X(a, t)$ for $EII(N)$ and $EIII(N)$ are 0-mean, positive and negative random variables, such that $P_{T_X(a,t)}(T = 0) > 0$ for all the processes studied here. Therefore, their moments are finite only for $q > -1$. Hence, $EII(N)$ and $EIII(N)$ are defined only for $q > -1$.

---

[e]The reasons that led to that choice are twofold. First, it is known from the wavelet analysis of self-similar processes that the use of the dyadic grid brings close to optimal estimation performance[1,2], this has been confirmed in preliminary analysis of multifractal processes[16]. Second, the $T_X(2^j, k2^j; f_0)$ can be computed, for all three estimators, using fast Mallat type pyramidal recursive algorithms[26].
[f]This is supported by arguments developed in[13].

3. Multiplicative Processes

Except for the case of the generalised random wavelet series introduced in[4], all the MEMF processes introduced in the literature and for which a synthesis procedure is so far available are defined through a multiplicative cascade construction. Therefore, we concentrate on positive multiplicative cascades in the present work. Roughly, a multiplicative cascade consists of a recursive procedure that re-distributes the mass inside a given set according to a geometric fragmentation rule. The limiting object generated by such a procedure displays multifractal scaling behaviours as in Eqs. (1.1) or (1.2) and the scaling exponents $\zeta(q)$ are related to the generator of the cascade (i.e., the rules of re-distribution of the mass).

3.1. *Multiplicative cascades*

Canonical Mandelbrot's Cascades   Let us start by recalling the definition of the historical and celebrated canonical multiplicative Mandelbrot's cascades[27] (hereafter referred to as CMC). Two key ingredients are entering their construction: a rigid geometrical grid and independent identically distributed (i.i.d.) random multipliers. An initial interval on the real line is splitted into two[g]. This splitting procedure is then iteratively applied to each subintervals so that after J iterations one ends up with a set of dyadic intervals $\{I_{j,k} = [k2^{-j}, (k+1)2^{-j}), j = 1, \ldots, J, k = 1, \ldots, 2^j\}$. To the $I_{j,k}$ are associated multipliers $W_{j,k}$, consisting of positive i.i.d. random variable with mean equal to one and characterised by (the opposite of the logarithm of) their moment generating function,

$$\varphi(q) = -\log_2 I\!\!E W^q, (\varphi(1) = 0). \tag{3.9}$$

---

[g]Equivalent constructions based on the splitting of each interval in $b \geq 2$ subintervals instead of 2 have also been proposed.

At resolution $r = 2^{-J}$, the cascade, or density $Q_r(t)$, is obtained at time t as the product of all the $W_{j,k}$ associated to the $I_{j,k}$ containing t (see Fig. 1, top left):

$$Q_r(t) = \prod_{\{(j,k):1\leq j\leq J, t\in I_{j,k}\}} W_{j,k}. \tag{3.10}$$

Scaling behaviours for the CMC, such as those of Eq. (1.2), constitute a well-known result, cf. e.g.,[27,36]:

$$I\!E\left[\frac{1}{n}\sum_{k=1}^{n}\left(\lim_{r\to 0}\frac{1}{2^{-j}}\int_{k2^{-j}}^{(k+1)2^{-j}}Q_r(u)du\right)^q\right] = c_q|2^{-j}|^{\varphi(q)}, \ \ 2^{-j} < 1. \tag{3.11}$$

CMCs have been the first, and up to a recent past, the only construction yielding stochastic processes with a priori controlled scaling properties. However, from a data modelling point of view, they suffer from two major drawbacks. First, they possess *discrete scale invariance only*: the scaling behaviour in Eq. (3.11) above only holds for specific scales, $a_j = 2^{-j}$. Second, the density $Q_r(t)$ is not a stationary process[h]: indeed, the construction is not *time-shift invariant* since all time positions t do not occupy equivalent positions at the end of the rigid dyadic tree.

Compound Poisson Cascades.   To overcome those two major drawbacks, Barral & Mandelbrot proposed to replace the *deterministic or rigid dyadic grid* with a *random* geometry[11]. This construction starts with a Poisson random point process $(t_i, r_i)_{i\in\mathcal{I}}$, defined on a rectangle $I :, r \leq r' \leq 1, 0 \leq t' \leq T$ and with density $dm(t', r')$ (see Fig. 1, top right). Positive, with mean one, i.i.d. multipliers W_i are associated to the $(t_i, r_i)_{i\in\mathcal{I}}$. The corresponding density $Q_r(t)$, referred to a compound Poisson cascade (CPC hereafter), is then defined as the product of the W_i corresponding to points within the cone $\mathcal{C}_r(t) = \{(t', r') : r \leq r' \leq 1, t - r'/2 \leq t' \leq t + r'/2\}$, the normalisation factor ensures $I\!EQ_r = 1$:

$$Q_r(t) = \left(I\!E[\prod_{(t_i,r_i)\in\mathcal{C}_r(t)} W_i]\right)^{-1} \prod_{(t_i,r_i)\in\mathcal{C}_r(t)} W_i. \tag{3.12}$$

---

[h]Eq. (1.2) holds but Eq. (1.1) does not.

The choice $dm(t,r) = c/r^2 drdt$ together with the triangle-shaped cone $\mathcal{C}_r(t)$ ensures that the density $Q_r(t)$ presents power law behaviours as in Eq. (1.1) (cf.[11,8,14]):

$$\mathbb{E}\left[\left(\lim_{r \to 0} \frac{1}{a} \int_t^{t+a\tau_0} Q_r(u)du\right)^q\right] = c_q|a|^{\varphi(q)}, \ a \leq 1, \tag{3.13}$$

where

$$\varphi(q) = c[1 - \mathbb{E}W^q - q(1 - \mathbb{E}W)], \ (\varphi(1) = 0). \tag{3.14}$$

Infinitely Divisible Cascades.   Noting that compound Poisson distributions fall into the general class of infinitely divisible distributions, the discrete product $Q_r(t) \propto \prod_{(t_i,r_i) \in \mathcal{C}_r(t)} W_i = \exp[\sum_{(t_i,r_i) \in \mathcal{C}_r(t)} \log W_i]$ can be further generalized to the exponential of a *continuous* random measure $dM(t,r)$, with control measure $dm(t,r)$[41], cf. Fig. 1, bottom right). This leads to the definition of infinitely divisible cascades (IDC)[8,14,32]. The corresponding density Q_r reads:

$$Q_r(t) = (\mathbb{E} \exp M(\mathcal{C}_r(t)))^{-1} \exp \int_{\mathcal{C}_r(t)} dM(t',r'). \tag{3.15}$$

Again, the normalisation ensures that $\mathbb{E}Q_r(t) = 1$. The continuous measure M needs to be defined from an *independently scattered infinitely divisible distribution* G with moment generating function $\tilde{G}(q) = e^{-\rho(q)}$[18]. Again $dm(t,r) = c/r^2 drdt$ and the triangle-shaped cone $\mathcal{C}_r(t)$ imply that $Q_r(t)$ presents power-law behaviours that can be written exactly as in Eq. (3.13) above with $\varphi(q) = \rho(q) - q\rho(1)$, $(\varphi(1) = 0)$.

Positive martingales, degeneracy and divergence of moments.   From the mathematical viewpoint, the three types of cascades Q_r defined above form positive multiplicative martingales. This property rises a number of issues that will appear of important practical interest in the analysis of the results in Sections 4 and 5. The results of this paragraph were proven independently for CMC[24], for CPC[11] and for IDC[8]. By construction, the densities $Q_r(t)$ converges almost surely to 0 as the resolution r decreases to 0 so that one is led to define the corresponding measure

$A(t)$ as:

$$A(t) = \lim_{r \to 0} \int_0^t Q_r(s)ds. \tag{3.16}$$

$A(t)$ is a non degenerated process if $\varphi'(1^-) \geq -1$. Moreover, the moments of A are finite only within a range of orders q: $q \in [q_c^-, q_c^+]$, where q_c^- and q_c^+ are defined, for the three cascades[i], as:

$$q_c^+ = \sup\{q \geq 1 : q + \varphi(q) \geq 1\} \quad q_c^- = b \inf\{q \in \mathbb{R} : I\!\!E W^q < \infty\}, \tag{3.17}$$

where b corresponds to the number of splitting blocks of the CMC ($b = 2$ here), $b = 1$ for the CPC. Furthermore, $A(t)$ also presents power law behaviours as in Eqs. (1.1) and (1.2), cf. Section 3.3.

3.2. *Multiscaling Stochastic processes*

Despite their possessing a priori prescribed multiscaling properties, the processes $Q_r(t)$ and $A(t)$ may not be general enough for the modelling of empirical data since the former takes positive values only and the latter displays non negative variations only. This section discusses two alternatives recently introduced in the literature[j].

Fractional Brownian motion in Multifractal time   Following an idea that goes back to Mandelbrot[29] and was further developed in[36], one can define a process with prescribed scaling exponents as well as positive and negative fluctuations.

Let A be a measure obtained from one of the three densities Q_r (CMC, CPC, IDC) defined above, and let B_H denote fractional Brownian motion[37] with Hurst parameter $0 < H < 1$. The process obtained by warping the time variable according to

---

[i]No theoretical result is available for q_c^- in the IDC case.
[j]Random wavelet cascades[3] also provide a solution. Because they suffer from the same drawbacks as CMCs, we did not consider them here.

$t \to A(t)$ *is called* fractional Brownian motion in Multifractal time (MF-FBM)[k]*:*

$$V_H(t) = B_H(A(t)), \quad t \in \mathbb{R}^+. \tag{3.18}$$

Multifractal Random Walk   Another possible choice was proposed by Bacry *et al.*[7,32,8]. It consists whenever it is mathematically sound to perform a stochastic integration of a density Q_r as defined above, against fractional Brownian motion. This integration contains numerous mathematical involved issues not discussed here (the process is well defined only when $2H + \varphi(2) - 1 > 0$, cf.[32]). Hence, Y_H is practically defined through the limit of Riemann sums (following[7,32], this process will be referred to as *Multifractal Random Walk* (MRW)[l]):

$$Y_H(t) = \lim_{\Delta t \to 0} Y_{H,\Delta t}(t) = \lim_{\Delta t \to 0} \sum_{k=0}^{t/\Delta t} Q_r(k\Delta t)(B_H(k\Delta t) - B_H((k-1)\Delta t)). \tag{3.19}$$

3.3. *Scaling properties*

The properties of the processes A, V_H and Y_H are directly inherited from those of the Q_r they are defined from. This implies that when constructed from CMCs, their increments suffer from non stationarity and discrete scale invariance, while based on CPCs or IDCs, they possess stationary increments and continuous scale. The scaling properties of the different processes can be expressed, when $a\tau_0 < 1$, as follows (these results are gathered from[36,11,8,14]). For CMCs, one has[m]:

$$\left.\begin{array}{lll} Q_r, & \frac{1}{n}\sum_{k=1}^{n}(\lim_{r\to 0}\frac{1}{2^{-j}}\int_{k2^{-j}}^{(k+1)2^{-j}}Q_r(u)du)^q = S_n(q,2^{-j};\beta_0) \simeq c_q|2^{-j}|^{\varphi(q)}, \\ A, & \frac{1}{n}\sum_{k=1}^{n}|A((k+1)2^{-j}) - A(k2^{-j})|^q = S_n(q,2^{-j};I_0) \simeq c_q|2^{-j}|^{q+\varphi(q)}, \\ V_H, & \frac{1}{n}\sum_{k=1}^{n}|V_H((k+1)2^{-j}) - V_H(k2^{-j})|^q = S_n(q,2^{-j};I_0) \simeq c_q'|2^{-j}|^{qH+\varphi(qH)}, \\ Y_H, & \frac{1}{n}\sum_{k=1}^{n}|Y_H((k+1)2^{-j}) - Y_H(k2^{-j})|^q = S_n(q,2^{-j};I_0) \simeq c_q''|2^{-j}|^{qH+\varphi(q)}. \end{array}\right\} \tag{3.20}$$

---

[k]In the original definition the label *fractional Brownian motion in Multifractal time* was used only for the case where $A(t)$ was obtained from a Mandelbrot's cascade (CMC); the definition here is therefore an extension to the CPCs and IDCs. It was labelled *Infinitely Divisible Cascading random walk* and *log-infinitely divisible multifractal random walk* in[14,32], respectively.

[l]Though the soundness of this extension has not yet been proven in general[7,32], we extend this definition to the three types of densities Q_r described here. Since numerical simulations are discrete by nature, the simulation of $Y_{H,\Delta t}$ is easy and can be used as a surrogate definition.

[m]The symbol $\simeq$ stands for the fact that one would have an exact $=$ in the limits $n \to +\infty$ and $2^{-j} \to 0$, cf.[36] for precise details. For practical purposes, this $\simeq$ cannot be distinguished from a strict $=$.

For CPCs and IDCs, one has[n],

$$
\left.
\begin{aligned}
Q_r, \quad & \mathbb{E}(\lim_{r\to 0} \tfrac{1}{a}\int_t^{t+a\tau_0} Q_r(u)du)^q = \mathbb{E}|T_{Q_0}(a,t;\beta_0)|^q \simeq c_q|a|^{\varphi(q)}, \\
A, \quad & \mathbb{E}|A(t+a\tau_0) - A(t)|^q \quad = \mathbb{E}|T_A(a,t;I_0)|^q \simeq c_q|a|^{q+\varphi(q)}, \\
V_H, \quad & \mathbb{E}|V_H(t+a\tau_0) - V_H(t)|^q = \mathbb{E}|T_{V_H}(a,t;I_0)|^q \simeq c'_q|a|^{qH+\varphi(qH)}, \\
Y_H, \quad & \mathbb{E}|Y_H(t+a\tau_0) - Y_H(t)|^q = \mathbb{E}|T_{Y_H}(a,t;I_0)|^q \simeq c''_q|a|^{qH+\varphi(q)}.
\end{aligned}
\right\} \quad (3.21)
$$

Examples of sample path for Q_r, A, V_H, Y_H are shown in Fig. 2.

3.4. *Synthesis procedures - Number of integral scales versus resolution (or depth) of the cascade*

We developed MATLAB procedures to synthesise the processes defined above. They are documented in[14,15] and available upon request. This section does not intend to detail them but rather to put the emphasis on two key issues: the resolution r of the cascades and the number n_L of integral scales.

The constructions of multiplicative cascades described above imply that scaling hold from a maximal (or integral) scale L, down to a minimal scale corresponding to the resolution r of the cascade. This means that the scaling reported in Eqs. (3.20) or (3.21) are practically valid in the range $r < a\tau_0 < L$. Actually, only the ratio L/r — the depth of the cascade — matters, hence L was arbitrarily labelled $L = 1$ above (cf. Section 3.1). Practically, we are working with discrete time time series, with sampling period T_s. It is natural[14] to tie the sampling period to the resolution, $T_s = r$. The number of samples n corresponding to the observation duration reads $n = n_L L/r$ where n_L stands for the number of integrals scale. Varying n amounts either to decrease the resolution $r \to 0$ for a given n_L or to increase the number of observed integral scales n_L for a given resolution. For ease of exposition, those two different asymptotic behaviours will be referred to as — given n_L, $r \to 0$ — and — given r, $n_L \to +\infty$ —, respectively. The results and interpretations reported in Sections 4 and 5 investigate both limits and are valid for these two cases.

---

[n]The symbol $\simeq$ stands for the fact that one would have an exact $=$ in the limits $a \to 0$, cf.[8,14,15] for precise details. For practical purposes, this $\simeq$ cannot be distinguished from a strict $=$.

4. Linearisation effect: analysis and formulation

4.1. *Methodology.*

The estimation performance of the $\hat{\zeta}(q,n)$ are obtained from numerical simulations: the estimators $\hat{\zeta}(q,n)$ are applied to *nbreal* copies of a chosen process X. The statistical characteristics (expectations, variances, ...) of the $\hat{\zeta}(q,n)$ are deduced from averaging over realisations. In the present work, we used standard orthonormal Daubechies wavelets[26], $N = 1, \ldots, 10$. We set by convention $r = T_s = \tau_0 = 1$, $n = 2^J, L = 2^{J_L}, n_L = 2^{J-J_L}, J_L = 10, \ldots, 16, J = 8, \ldots, 17, nbreal = 1000$, $j_1 = 3$, $j_2 = \min(J - (2N+1), J_L - 1)$.

4.2. *Empirical findings*

Linearisation effect.   The application of the estimation procedures to a very large number of realisations of a studied process leads to the observation of the following fundamental empirical fact. While q belongs to a specific interval $q \in [\hat{q}_o^-, \hat{q}_o^+]$, the estimates $\hat{\zeta}(q,n)$ account for the $\zeta(q)$, given by the theoretical considerations on the studied process developed in Section 3 ; when q is outside that interval, $q \notin [\hat{q}_o^-, \hat{q}_o^+]$, the $\hat{\zeta}(q,n)$ significantly depart from the theoretical $\zeta(q)$ and, besides that, systematically behave as a linear function of q: $\hat{\zeta}(q,n) = \hat{\alpha}_o^\pm + \hat{\beta}_o^\pm q$. We refer to this behaviour as to a *linearisation effect* of the $\hat{\zeta}(q,n)$ with respect to q. We put the emphasis on the fact that this occurs for each and every single realisation of the process and not only on average. The quantities $q_o^\pm, \hat{\alpha}_o^\pm$ and $\hat{\beta}_o^\pm$ are random variables, taking values that depend on each realisation. It is illustrated in Fig. 3, top row, on (5 realisations of) two specific examples: left column, CMC, Q_r, $EI(1)$; right column, CPC, V_H, $EIII(3)$.

Legendre transforms.   Because it turns a straight line into a single point, the Legendre transform can be thought of as an interesting tool to study the linearisation effect. Let d denote the Euclidean dimension of the space over which the

process X is defined[o], the Legendre transform $D(h)$ of the function $\zeta(q)$ is defined[p] as (see e.g.,[36]):

$$D(h) = d + \text{Inf}_q \ (qh - \zeta(q)). \quad (4.22)$$

Fig. 3, middle row, compares the Legendre Transforms, $\hat{D}(h, n)$, of the estimates $\hat{\zeta}(q, n)$ obtained from each single observation to that, $D(h)$, of the theoretical function $\zeta(q)$. It shows that, within an interval $h \in [\hat{h}_o^-, \hat{h}_o^+]$, the $\hat{D}(h, n)$ tend (to superimpose) to $D(h)$. It also shows that the $\hat{D}(h, n)$ are abruptly ended by *accumulation points*, whose coordinates in the plane (h, D) are labelled $(\hat{h}_o^\pm, \hat{D}_o^\pm)$. The existence of these accumulation points constitutes another evidence for the linearisation effect and the Legendre transform immediately yields $\hat{\alpha}_o^\pm = d - \hat{D}_o^\pm$ and $\hat{\beta}_o^\pm = \hat{h}_o^\pm$. Fig. 3, bottom row, shows accumulation points, $(\hat{h}_o^+, \hat{D}_o^+)$, obtained from 1000 realisations of the same process. Again, they consist in random variables depending on realisations. However, one can notice that they are spread in the neighbourhood of the theoretical curve $D(h)$ and mainly concentrate around the right and left zeroes of $D(h)$: $D(h) = 0$.

Dependence with the number of samples n. Since the parameters, $\hat{q}_o, \hat{\alpha}_o, \hat{\beta}_o, \hat{D}_o, \hat{h}_o$, defining the linearisation effect are random variables taking values that depend on each realisation, we now investigate the dependence of their statistics with respect to the number of samples n of the process. The measures of $\hat{h}_o, \hat{\alpha}_o, \hat{\beta}_o, \hat{D}_o$ are straightforward ; $\hat{q}_o$ is computed by equating $|\hat{\zeta}(q, n) - \zeta(q)|$ and $|\hat{\zeta}(q, n) - (\hat{\alpha}_o + \hat{\beta}_o q)|$. Fig. 4 shows the means and standard deviations of $\hat{q}_o, \hat{D}_o, \hat{h}_o$ as a function of $\log_2(n)$ for the asymptotic behaviours — given $n_L, r \to 0$ — (right column) and — given $r, n_L \to +\infty$ — (left column). Major conclusions can be drawn from these plots[q]. First, the mean values of the parameters characterising the linearisation effect **do not** depend on n. This implies that the critical q, above

---

[o]For a $1D$ time series: $d = 1$.

[p]When $\zeta(q)$ is a continuously differentiable function, $D(h) = q(h)\frac{d\zeta}{dq}(q(h)) - \zeta(q(h))$, where $q(h)$ is derived from $h = \frac{d\zeta}{dq}(q)$.

[q]Plots for $\hat{\alpha}_o, \hat{\beta}_o$ are not shown for space reasons but would yield identical conclusions.

which the linearisation occurs, has no functional dependence with the observation duration n, the average linear function $\alpha + \beta q$ on which the $\hat{\zeta}(q, n)$ collapse does not vary when n is increased. **Therefore, the linearisation effect is in no way a finite size effect, that would weaken or disappear when $n \to +\infty$.** Second, for a given number of integral scales (right column), the standard deviations of the fluctuations of the parameters characterising the linearisation effect decrease as $r \to 0$. In the ideal limit $r \to 0$, $\hat{q}_o, \hat{\alpha}_o, \hat{\beta}_o, \hat{D}_o, \hat{h}_o$ may be exactly identical for all realisations of a same process. Third, for a given resolution (left column), the standard deviations of the fluctuations of the parameters characterising the linearisation effect remain constant as soon as the observation duration is larger than a single integral scales $n > n_L$ (or $J > J_L$) and do not decrease with the increase of the number of available scales n_L (cf. Fig. 4, where $J_L = 11$). This implies that **the linearisation effect is not a low performance statistical estimation issue**, this is a key information with respect to parameter estimation issues under current investigations (see[25]). Note however that Fig. 4 also indicates that in situations where the observation duration is too short, i.e., when it does not even cover a single integral scale of the analysed process (in our notations, $n \leq L/r$ or $J \leq J_L$), the linearisation effect may be hidden by dominant estimation issues: $\hat{\zeta}(q, n)$ and q_*^+ are then poorly estimated and their low statistical performance depend on n.

Generality. We now wish to put the emphasis on the fact that the experimental findings reported above hold systematically for the three types of cascades (CMC, CPC, IDC), for the four declinations of MEMF processes (Q_r, A, V_H, Y_H), described in Section 3 as well as for the three families of multiresolution estimators $EI(N), EII(N)$ and $EIII(N)$, and for both types of asymptotic behaviours — given n_L, $r \to 0$ — and — given r, $n_L \to +\infty$ —: it does not disappear in the limit of an infinite observation duration and its parameters do not depend on the observation duration nor on the depth of the cascade.

We also mention that other declinations on multiresolution estimators, such as

e.g., the Wavelet Transform Modulus Maxima[10], are subject to an identical linearisation effect.

As a further extension, let us mention that the linearisation effect occurs identically for random fields defined in dimension higher than 1. Fig. 5 shows an example of the linearisation effect observed on a two dimensional cascade (a 2D-CMC), in that case, $d = 2$ in the definition of the Legendre transform in Eq. (4.22).

4.3. *Conjecture*

Based on the fact that the empirical observations reported above are obtained from large numbers of realisations, for a wide variety of cascades, with various choices for $\varphi(q)$, for different MEMF-processes and with various types of estimators, we are led to formulate the following conjecture regarding the behaviour of the $\hat{\zeta}(q, n)$.

Critical points.   From the Legendre transform of the function $\zeta(q)$, we define the critical values, $D_*^\pm, h_*^\pm, q_*^\pm$, that will enter the theoretical characterisation of the linearisation effect. Note that **these critical quantities depend on the definition of the process itself only** and not on the estimation procedures:

$$\left.\begin{array}{rll} (h_*^\pm, D_*^\pm) & \text{such that} \quad D_*^\pm = 0 \quad \text{ and } D(h_*^\pm) = 0, \\ q_*^\pm & \text{such that} \quad h_*^\pm = \quad (d\zeta(q)/dq)_{q=q_*^\pm}, \\ \alpha_*^\pm, \beta*^\pm & \text{such that} \quad \alpha_*^\pm = d - D_*^\pm, \; \beta*^\pm = h_*^\pm. \end{array}\right\} \qquad (4.23)$$

For sake of simplicity and without loss of generality, we restrict here the analysis to the cases where $q_*^- \leq -1$.

Conjecture.   For any MEMF-process (cf. definition in Section 1 and Eq. (1.3)), the $\hat{\zeta}(q, n)$ behave as:

$$EI \begin{cases} \hat{\zeta}(q, n) = \hat{\alpha}_o^- + \hat{\beta}_o^- q \to d - D_*^- + h_*^- q, \; q \leq q_*^-, \\ \qquad \hat{\zeta}(q, n) \qquad \to \zeta(q), \qquad q_*^- \leq q, \leq q_*^+, \\ \hat{\zeta}(q, n) = \hat{\alpha}_o^+ + \hat{\beta}_o^+ q \to d - D_*^+ + h_*^+ q, \; q \geq q_*^+, \end{cases} \qquad (4.24)$$

$$EII\&III \begin{cases} \qquad \hat{\zeta}(q, n) \qquad \to \zeta(q), \qquad -1 < q \leq q_*^+, \\ \hat{\zeta}(q, n) = \hat{\alpha}_o^+ + \hat{\beta}_o^+ q \to \; d - D_*^+ + h_*^+ q, q_*^+ \leq q. \end{cases} \qquad (4.25)$$

Eqs. (4.24) and (4.25) above mean that $(\hat\zeta(q,n) - \zeta(q) - B(q,n))/\sqrt{V(q,n)}$ when $q \in [q_*^-, q_*^+]$ and $(\hat\zeta(q,n) - (d - D_*^\pm + h_*^\pm q) - B(q,n))/\sqrt{V(q,n)}$ when $q \notin [q_*^-, q_*^+]$ converge to random variables with unit variance zero mean limiting laws $\mathcal{P}$ and $\mathcal{P}'$. The terms $B(q,n)$ and $V(q,n)$ account for biases and variances of the $\hat\zeta(q,n)$. Their precise forms depend on the exact nature of the studied processes, on q, on the chosen estimator and require specific case by case formulations. However, generic features, valid for all processes and all estimators, have been observed. The limiting laws may significantly depart from a normal law. In the limit of large n, the $V(q,n)$ decrease as power laws of n, $V(q,n) \sim \Lambda_q n^{\gamma(q)}$, where $\gamma(q) \simeq -1$ for $q \simeq 0$ and $\gamma(q) \simeq -0.1$ for $|q| \geq q_*^\pm$, corresponding to an evolution from the usual n^{-1} fast decrease to a very slow $n^{-0.1}$ decrease. This will be detailed in forthcoming works. Note that MEMF processes such that the critical point $D(h_*^\pm) = 0$ is not defined simply correspond to $q_*^\pm = \infty$. Fig. 6 shows the excellent agreement between the prediction of this conjecture and the corresponding empirical observations.

5. Linearisation effect: interpretations and comments

5.1. *Overview of the literature*

Strangely enough, despite its very systematic and robust nature, this linearisation effect has been almost completely overlooked in the fairly large number of papers dedicated to the study of the multifractal properties of empirical data obtained from various applications. However, following the inspiring contribution of Mandelbrot[27], a limited number of works reported the existence of the linearisation effect in empirical data[28,38,39,40,42] and relate it theoretically and empirically either to finiteness of moments issues or to finite size effects. Another set of papers studies theoretically this effect in the specific case of CMC[r], for positive multiplicative martingale Q_r, and the $EI(1)$ estimator[17,21,30,31,33,34].

---

[r]The results of[33] were recently reformulated[35] for conservative Mandelbrot's cascades, Q_r process analysed with the Haar wavelet (i.e., a specific subcase of $EIII(1)$).

Finiteness of moments.   In his seminal work on multiplicative cascades (CMC) in turbulence[27], Mandelbrot indicates that, when the resolution r tends to 0, the moments of the box-aggregated coefficients $T_{Q_r}(a, t; \beta_0)$, or equivalently the first order increment coefficients of A, $T_A(a, t; I_0)$, are infinite when $q > q_c^+$ (cf. Eq. (3.17)). Following this, the issue of the relations between infiniteness of the moments and the linearisation of the $\hat{\zeta}(q)$ has been addressed in e.g.,[38,39,40,42]. The results and analysis reported therein relate, in a number of cases referred to as *first order multifractal phase transitions,* the critical order q beyond which the linearisation effect occurs to the order q_c^+ beyond which moments are infinite.

For all the cascades and processes studied here, the quantities $q_c^{\pm}$ and $q_*^{\pm}$ can be derived theoretically (cf. Eqs. (3.17) and Eqs. (4.23), respectively). Furthermore, the convexity of $\varphi(q)$ enables to show[s] that $q_c^- \leq q_*^- < 0 < 1 \leq q_*^+ \leq q_c^+$ for all types of cascades. The numerical simulations reported in the present work are clearly and unambiguously showing that the critical order q beyond which the linearisation effect is observed corresponds to $q_*^{\pm}$ and not to $q_c^{\pm}$. Hence, **the linearisation effect is not related to a divergence of moments issue**.

Finite size effects.   A second category of analysis of the linearisation effect is embedded in the multifractal formalism, for a thorough introduction, see e.g.,[36,22,23]. A major property of the processes Q_r, A, V_H, Y_H lies in their sample paths being multifractal, i.e., consisting, in the limit of small scales, in a hierarchical collection of singularities characterised by their (Hölder) exponent h, this translates to[t] $|X(t + a\tau_0) - X(t)| \sim c(t) a^{h(t)}$, $a \to 0$. The multifractal spectrum $D(h)$ consists of the Hausdorff dimension of the set of points t on the real line where the singularities have the same exponent h. It can be related to the scaling exponents in Eqs. (1.1) and (1.2) through a Legendre transform.

References[28,38,39,40,42] claim that, for any finite duration time series of a given

---

[s]The case $q_*^+ \equiv q_c^+$ corresponds only to trivial processes where $\varphi(q) \equiv 0$.
[t]This translation implies a slight oversimplification aiming at simplicity of exposition, for precise statements, see e.g.,[36].

realisation of a multifractal process, there exists a maximal observable singularity (characterised by a minimal h) and hence a stopping point in the estimated multifractal spectrum. This yields linearisation through an inverse Legendre transformation. Pursuing this analysis, these papers indicate that the maximal observable singularity is dependent both on the sampling rate of the data (in our words, this corresponds to the depth or resolution of the cascade and asymptotic behaviour — given n_L, $r \to 0$ —) and on the number realisations of the process available for the analysis, usually referred to as *supersampling*, (in our work, it corresponds to the number of integral scales available and asymptotic behaviour — given r, $n_L \to +\infty$ —). The linearisation effect is hence explained as a finite size effect.

The results reported in the present work unambiguously disagree with that analysis. They clearly show that the linearisation effect does not disappear either when — given n_L, $r \to 0$ — or — given r, $n_L \to +\infty$ —. They also show that the (average values of the) parameters entering its description do not depend on the resolution of the cascade nor on the number of integral scales, but only on quantities entering the theoretical definition of the process. **Therefore, the linearisation effect is not a finite size effect**.

Multiplicative positive martingales and MEMF processes.   Elaborating on[17], two independent and recent works by Molchan[30,31] and Waymire et al.[21,33,34] studied theoretically the linearisation effect in the specific context of CMC, Q_r and $EI(1)$. The experimental observations described here as well as the definitions of the critical parameters $D_*^{\pm}, h_*^{\pm}, q_*^{\pm}$ are in perfect agreement with the theoretical analysis proposed in these papers. The results reported in the present work can therefore be read and understood as evidence in favour of the extension, mutatis mutandis, of the characterisation of the *linearisation effect* to new types of cascades, CMC, CPC, IDC, to new types of MEMF processes, Q_r, A, V_H, Y_H, derived from these cascades, to new families of multiresolution estimators, $EI(N), EII(N), EIII(N)$, and to different types of asymptotic behaviours.

Furthermore, it is worth noting that the MEMF processes tested here, despite many important differences, are all based, deeply in nature, in multiplicative positive martingales. Therefore, it could be conjectured that the linearisation effect is deeply rooted in the multiplicative martingale nature of the process and not in the estimation procedures themselves. It could also be suspected however that the results reported here even more crucially depend on the MEMF classification of the processes and occur identically for MEMF processes not defined from multiplicative positive martingales. Such issues are under current investigations using the recently introduced class of multifractal processes referred to as random wavelet series[23,4,5].

5.2. *Picturing the linearisation effect*

Let Q_r denote a CPC density[u]. The definition of Q_r as a multiplicative positive martingale implies the existence of, not a single one but, two functions of q and two different power law behaviours. On the one hand, from the multipliers W, comes in the function $\varphi(q)$, cf. Eq. (3.14). Let $D_\varphi(h)$ denote the Legendre transform of $\varphi(q)$. When $r \to 0$, the following power law behaviour can be proven[8,14,11]:

$$\mathbb{E}Q_r(t)^q = r^{\varphi(q)}, \;\; q \in \mathbb{R}. \tag{5.26}$$

Note that this is not a relation describing scaling since it involves the synthesis resolution of the cascade r and not an analysis scale a. On the other hand, the power law behaviours of the moments of order q of the multiresolution coefficients[v] $T_{Q_0}(t, a; \beta_0) = T_A(t, a; I_0)/a$ yield another function of q, $\zeta(q)$, defined as (with $a\tau_0 < L$):

$$\left.\begin{array}{ll} \mathbb{E}T_{Q_0}(t, a; \beta_0)^q = c_q|a|^{\zeta(q)}, \; q \in [q_c^-, q_c^+], \\ \mathbb{E}T_{Q_0}(t, a; \beta_0)^q = \infty, & \text{else,} \end{array}\right\} \tag{5.27}$$

---

[u]CPC is chosen because stationarity makes the statement of the arguments easier and clearer. For sake of simplicity, we assume that $\varphi(q)$ is defined for $q \in \mathbb{R}$, in other words that the multipliers W have have finite moments of any order.

[v]For ease of notations, $T_{Q_0}(t, a; \beta_0) = \lim_{r \to 0}(1/a)\int_t^{t+a\tau_0} Q_r(s)ds$.

with:

$$\left.\begin{array}{l}
\zeta(q) = 1 + qh_*^-, \; q \in [q_c^-, q_*^-], \\
\zeta(q) = \varphi(q), \qquad q \in [q_*^-, q_*^+], \\
\zeta(q) = 1 + qh_*^+, \; q \in [q_*^+, q_c^+].
\end{array}\right\} \tag{5.28}$$

Let $D_\zeta(h)$ stand for the Legendre transform of $\zeta(q)$. For CMCs, Eqs. (5.27) and (5.28), where the expectation $I\!E$ is replaced by a time average $\lim_{n \to +\infty} 1/n \sum_{k=1}^n$, are proven in[33]. For CPCs, they have not been stated elsewhere as such, however, they can be deduced from a collection of theorems in[11] which contains them implicitly. Up to our knowledge, no theoretical result is available for IDC.

In the specific case of CMCs, the fact that there exist two different types of power laws had already been pointed out by Mandelbrot in[27] and is referred to as the behaviours of the *bare cascade*, Q_r, and of the *dressed cascade*, $T_{Q_0}(t, a; \beta_0)$, respectively[38,40,42,38,39]. Those last papers had noticed and established that the moments of the dressed cascade remain finite only within a range of values of q, $q \in [q_c^-, q_c^+]$ but fail to distinguish between $\zeta(q)$ and $\varphi(q)$ and to define theoretically $q_*^\pm$. They explained the linearisation effect by finite size effects arguments combined together with divergence of moments (implying q_c^+) issues.

Up to our knowledge, [30,33] are the first contributions that established theoretically for CMCs[w] that $\varphi(q)$ and $\zeta(q)$ coincide only for $q \in [q_*^-, q_*^+]$ and that gave explicit definitions for $q_*^\pm$. Equivalently, it implies that $D_\varphi(h)$ and $D_\zeta(h)$ coincide only when they are both positive.

The numerical simulations reported here suggest that this picture can be extended straightforwardly to IDC. Furthermore, they also indicates that this picture also extends, mutatis mutandis, to the processes $X = A, V_H, Y_H$ and to their increment or wavelet coefficients $T_X(t, a; (I_0)^{*N})$ or $T_X(t, a; \psi_{0,N})$. It is understood as the fact that their definitions are deeply tied to multiplicative positive martingale constructions. Finally, it is worth noting that the multiresolution estimators $\hat{\zeta}(q, n)$ are totally *blind* or *unsensitive* to the critical values $q_c^\pm$ corresponding to the finite-

---

[w]This was also implicitly stated in the theoretical analysis of the regularity of the sample paths of the so-called random wavelet cascades[3].

ness of the moments and extend the linear behaviour of $\hat{\zeta}(q,n)$ beyond $q_c^{\pm}$. This picture is summarised and sketched in Fig. 7.

Because in applications, most of, if not all, the MEMF processes used are based on multiplicative positive martingales, practical multifractal analysis and modelling have been strongly misled by a potential confusion between the two functions $\varphi(q)$ and $\zeta(q)$. The former is related to the construction (or synthesis) of the model while the latter traduces scaling behaviours and is hence connected to the empirical analysis point of view. The scaling analysis with the multiresolution estimators $\hat{\zeta}(q,n)$ can only capture $\zeta(q)$: it should therefore be decided whether one seeks to estimate $\varphi(q)$ or $\zeta(q)$.

Major remaining questions are: how is this picture modified for the case of MEMF processes that are not defined from cascade constructions, in which case the function $\varphi(q)$ no longer exists? Are the definitions of the critical $D_*^{\pm}, h_*^{\pm}, q_*^{\pm}$ strongly tied to the multiplicative positive martingale nature of the processes or do they hold for a much wider range of processes, as we suspect ? The random wavelet series introduced in[4,5] constitute an excellent and versatile model to address these issues. As the CMCs, they suffer from non stationarity and non continuous scale invariance, but, up to our knowledge, they provide the only known, both theoretically and practically, MEMF processes avoiding multiplicative constructions. This is under current investigations.

5.3. *Major consequences of the linearisation effect*

In the practical analysis of multifractal processes, the existence of this linearisation effect has a major implication: it compels to recast the usual goal — estimate the scaling exponents $\zeta(q)$ — into a new and more accurate one: estimate the critical points $D_*^{\pm}, h_*^{\pm}, q_*^{\pm}$ and $\zeta(q)$ only within the interval $q \in [q_*^{-}, q_*^{+}]$. This is addressed in[25].

In the analysis and modelling of empirical data, large orders q are often used as

potentially rich information to discriminate between various candidate models for the function $\zeta(q)$ (this is the case e.g., in hydrodynamic turbulence[19]). The existence of the linearisation effect indicates that estimated scaling exponents for large qs must be used with care.

Furthermore, an important question raised by the analysis of scaling in empirical data is: should one resort to mono- or multi-fractal models for an accurate description. In most cases, this issue is addressed through the practical rephrasing: do the estimated scaling exponents $\hat{\zeta}(q, n)$ follow a linear behaviour? And the corresponding heuristic answer is: if yes, use a monofractal, if no, use a multifractal. From the analysis reported in the present work, we see that such an empirical answer makes no sense unless the linearisation effect of the $\hat{\zeta}(q, n)$ is precisely accounted for.

6. Conclusion and perspectives

Measuring the scaling exponents on empirical data has been the subject of a considerable amount of research works, spread in a large variety of domains. Surprisingly, the linearisation effect studied here remained so far widely overlooked. It might be because of the involved nature of the theoretical studies of the estimation procedures for the still partially understood multifractal processes. We have indicated here that the theoretical results obtained for CMC, Q_r, $EI(1)$ in[17,30,33] and characterising the linearisation effect can be extended to a much wider context: to new types of multiplicative positive cascades, to new sets of multi-exponents multifractal processes, to new families of multiresolution estimators, to asymptotic behaviours of different natures. It can also straightforwardly be extended to higher dimensions. We also clear out the facts that the linearisation effect is not related to infiniteness of moments, that this is not a finite size effect nor an estimation difficulty effect. We indicate that the linearisation effect is a limitation that is intrinsic to the nature of the multi-exponents multifractal processes and shed a new light on how the estimation of scaling exponents must be thought of and estimates used to

analyse data and draw physical conclusions.

A considerable amount of work is still to be done. The design of relevant practical estimators for the critical parameters $q_*^\pm, h_*^\pm, D_*^\pm$ is under study[25]. Detailed analysis of the bias and variance of the $\hat\zeta(q,n)$ are being performed. The incorporation of that effect into a practical procedure aiming at discriminating between mono- and multi-fractal may also prove of valuable help. Finally, the existence of the linearisation effect and the measurements of the corresponding critical points for empirical data in the fields of hydrodynamic turbulence and computer network telefraffic are being considered.

References

1. P. Abry, P. Flandrin, M. Taqqu, and D. Veitch. Wavelets for the analysis, estimation and synthesis of scaling data. In *Self-similar Network Traffic and Performance Evaluation, K. Park and W. Willinger, Eds.* Wiley, 2000.

2. P. Abry and D. Veitch. Wavelet analysis of long-range dependent traffic. *IEEE Trans on Info. Theory*, 44(1):2–15, 1998.

3. A. Arneodo, E. Bacry, and J.-F. Muzy. Random cascades on wavelet dyadic trees. *J. Math Phys.*, 39(8):4142–4164, 1998.

4. J.M. Aubry and S. Jaffard. Random wavelet series. *Com. in Mathematicals Physics*, 227:483–514, 2002.

5. J.M. Aubry and S. Jaffard. Random wavelet series; theory and applications. *Internal Report 14, CNRS UMR 8050*, 2003.

6. B. Audit, E. Bacry, J.F. Muzy, and A. Arneodo. Wavelets based estimators of scaling behavior. *IEEE Trans. on Info. Theory*, 48(11):2938–2954, 2002.

7. E. Bacry, J. Delour, and J.F. Muzy. Multifractal random walk. *Phys. Rev. E*, 64:026103, 2001.

8. E. Bacry and J.F. Muzy. Log-infinitely divisible multifractal processes. *Comm. in Math. Phys.*, 236:449–475, 2003.

9. E. Bacry, J.F. Muzy, and A. Arneodo. Singularity spectrum of fractal signals from wavelet analysis: exact results. *J. Stat. Phys.*, 70:635–674, 1993.

10. E. Bacry, J.F. Muzy, and A. Arneodo. Singularity spectrum of fractal signals from wavelet analysis: Exact results. *J. Stat. Phys.*, 70:635–674, 1994.

11. J. Barral and B. Mandelbrot. Multiplicative products of cylindrical pulses. *Probab. Theory Relat. Fields*, 124:409–430, 2002.

12. A. Benassi, A. Cohen, and J. Istas. identifying the multifractional function of a Gaussian process. In *Fractals and engineering*, pages 115–123. Springer Verlag, 1997.

13. P. Chainais. *Cascades log-infiniment divisibles et analyse multirésolution. Application à l'étude des intermittences en turbulence.* PhD thesis, E.N.S. Lyon, 2001.

14. P. Chainais, R. Riedi, and P. Abry. On non scale invariant infinitely divisible cascades. 2003. preprint.

15. P. Chainais, R. Riedi, and P. Abry. Scale invariant infinitely divisible cascades. In *Int. Symp. on Physics in Signal and Image Processing, Grenoble, France*, January 2003.

16. P. Chainais, S. Roux, P. Abry, and D. Veitch. Analyse et modélisation de séries temporelles à l'aide de cascades. Application à l'étude du trafic internet. In *Actes du GRETSI*, Toulouse, 2001.

17. P. Collet and F. Koukiou. Large deviations for multiplicative chaos. *Comm. Math. Phys.*, 147:329–342, 1992.

18. W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 2. John Wiley and Sons, Inc., New-York, London, Sidney, 1966.

19. U. Frisch. *Turbulence: the legacy of A. N. Kolmogorov.* Cambridge University Press, Cambridge, U. K., Cambridge, U. K., 1995.

20. P. Goncalves, R. Riedi, and R. Baraniuk. A simple statistical analysis of wavelet-based multifractal spectrum estimation. In *Proc. of the 32nd Conf. on Signals, Systems and Computers*, Asilomar, USA, 1998.

21. V.K. Gupta and E. Waymire. A statistical analysis of mesoscale rainfall as a random cascade. *J. Appl. Meteor.*, 32:251–267, 1993.

22. S. Jaffard. The multifractal nature of levy processes. *Probability Theory and Related Fields*, 114:207–227, 1999.

23. S. Jaffard. On lacunary wavelet series. *The Annals of Applied Probability*, 10:313–329, 2000.

24. J.-P. Kahane and J. Peyrière. Sur certaines martingales de Benoit Mandelbrot. *Adv. Math.*, 22:131–145, 1976.

25. B. Lashermes, P.Abry, and P.Chainais. Scaling exponents estimation for multiscaling processes. In *Int. Conference on Acoustics, Speech and Signal Processing, Montreal, Canada*, 2004.

26. S. Mallat. *A Wavelet Tour of Signal Processing.* Academic Press, San Diego, CA, 1998.

27. B. B. Mandelbrot. Intermittent turbulence in self similar cascades: Divergence of high moments and dimension of the carrier. *J. Fluid. Mech.*, 62:331, 1974.

28. B. B. Mandelbrot. negative fractal dimensions an multifractals. *Physica A*, 163:306–315, 1990.

29. B. B. Mandelbrot. A multifractal walk down wall street. *Scientific American*, 280(2):70–73, Feb. 1999.

30. G. M. Molchan. Scaling exponents and multifractal dimensions for independent random cascades. *Comm. Math. Phys.*, 179:681, 1996.

31. G. M. Molchan. Turbulent cascades: Limitations and a statistical test of the lognormal hypothesis. *Phys. Fluids*, 9(8):2387–2396, 1997.

32. J.F. Muzy and E. Bacry. Multifractal stationary random measures and multifractal random walks with log-infinitely divisible scaling laws. *Phys. Rev. E*, 66, 2002.

33. M. Ossiander and E.C. Waymire. Statistical estimation for multiplicative cascades. *The Annals of Statistics*, 28(6):1533–1560, 2000.

34. M. Ossiander and E.C. Waymire. On estimation theory for multiplicative cascades. *The Indian Journal of Statistics*, 64A:323–343, 2002.

35. S. Resnick, G. Samorodnitsky, A. Gilbert, and W. Willinger. Wavelet analysis of conservative cascades. *Bernoulli*, 9(1):97–135, 2003.

36. R. H. Riedi. Multifractal processes. *in: "Theory and applications of long range dependence ", eds. Doukhan, Oppenheim and Taqqu*, pages 625–716, 2003.

37. G. Samorodnitsky and M. Taqqu. *Stable non-Gaussian random processes*. Chapman and Hall, New York ISBN 0-412-05171-0, 1994.

38. D. Schertzer and S. Lovejoy. Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. *J. Geophys. Res.*, 92:9693, 1987.

39. D. Schertzer and S. Lovejoy. Hard and soft multifractal process. *physica A*, 185:187–194, 1992.

40. F. Schmitt. Intermittence et turbulences : analyse de données, validation de modèles et applications. *Habilitation à diriger des Recherches, Université Paris VI, France*, 2001.

41. F. Schmitt and D. Marsan. Stochastic equations generating continuous multiplicative cascades. *Eur. Phys. J. B*, 20:3–6, 2001.

42. F. Schmitt, D. Schertzer, S. Lovejoy, and Y. Brunet. Empirical study of multifractal phase transitions in atmospheric turbulence. *Nonlinear Processes in Geophysics*, 1:95–104, 1994.

28   *Lashermes, Abry, Chainais*
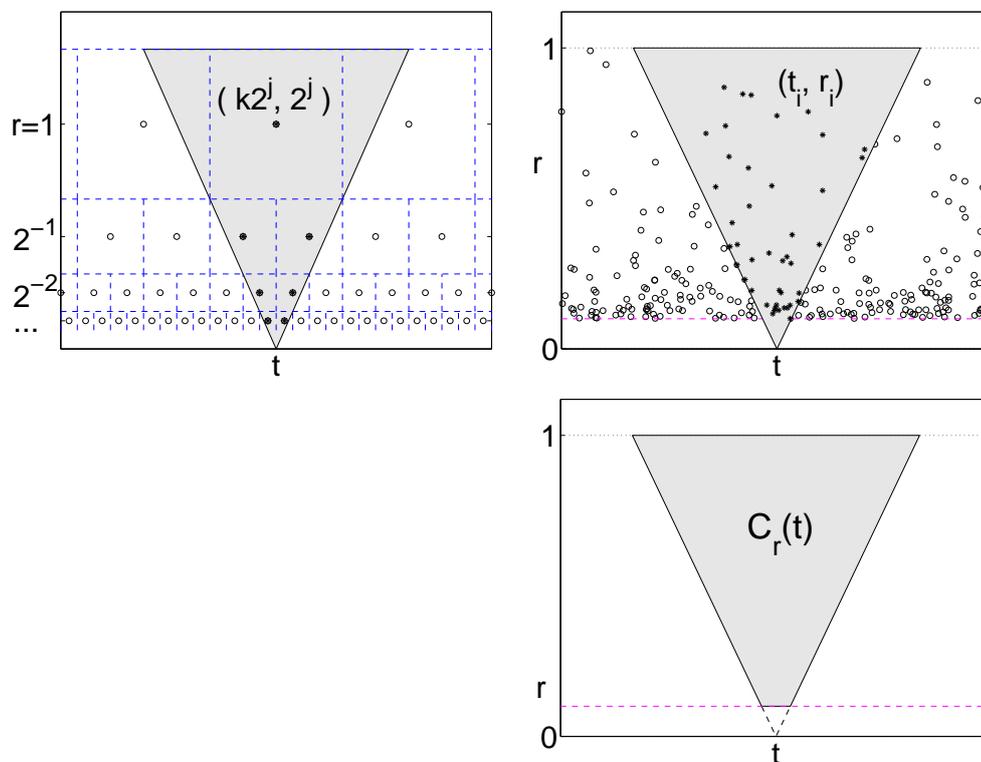
Fig. 1. **Compared illustrations for the "time-scale" construction of multiplicative cascades** between, **top left**, Canonical Mandelbrot's Multiplicative Cascade (CMC), **top right**, Compound Poisson Cascade (CPC), **bottom right**, Infinitely Divisible Cascade (IDC). The grey region indicates the cone containing multipliers that determine the value of the density at time t. The geometrically rigid grid underlying the construction of the CMC is replaced with random ones, a point process for the CPC, a continuously scattered measure for the IDC.
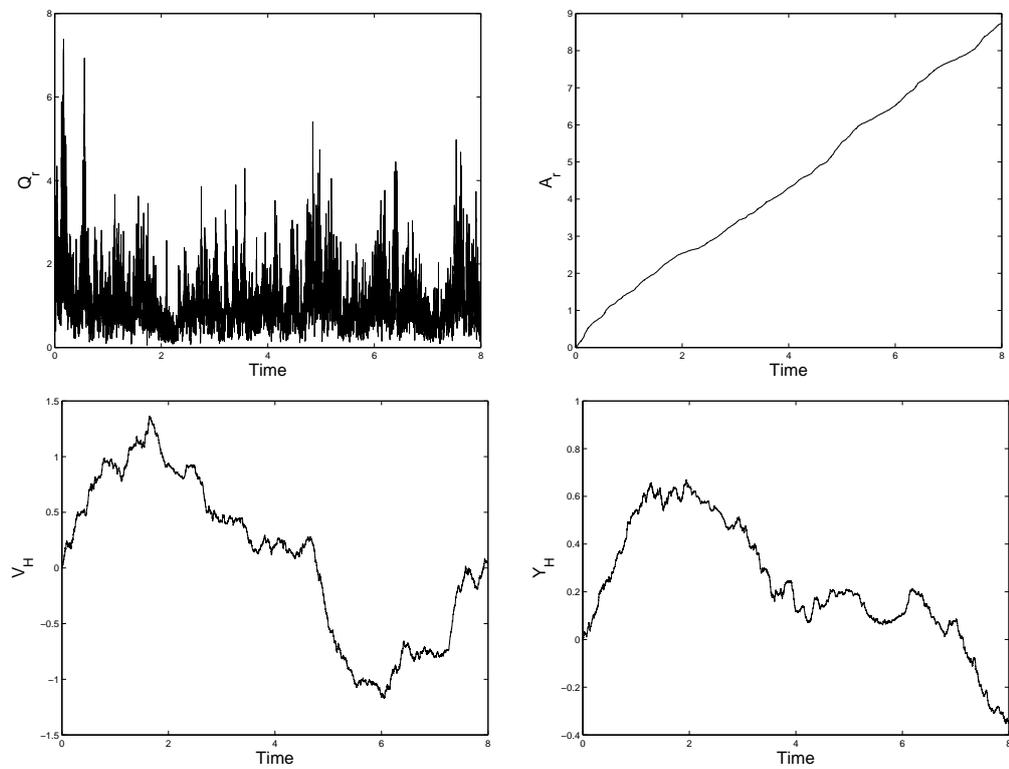
Fig. 2. **Examples of Sample Paths of multifractal processes.** They are obtained from a Compound Poisson Cascade (CPC), with $I\!\!E W^q = \exp(\mu q + \sigma^2 q^2)$ ($\mu = 0.1, \sigma = 0.03$) and with $H = 3/4$ for V_H and Y_H: **top left**, $Q_r(t)$, **top right**, $A(t)$, **bottom left**, $V_H(t)$ and **bottom right**, $Y_H(t)$.
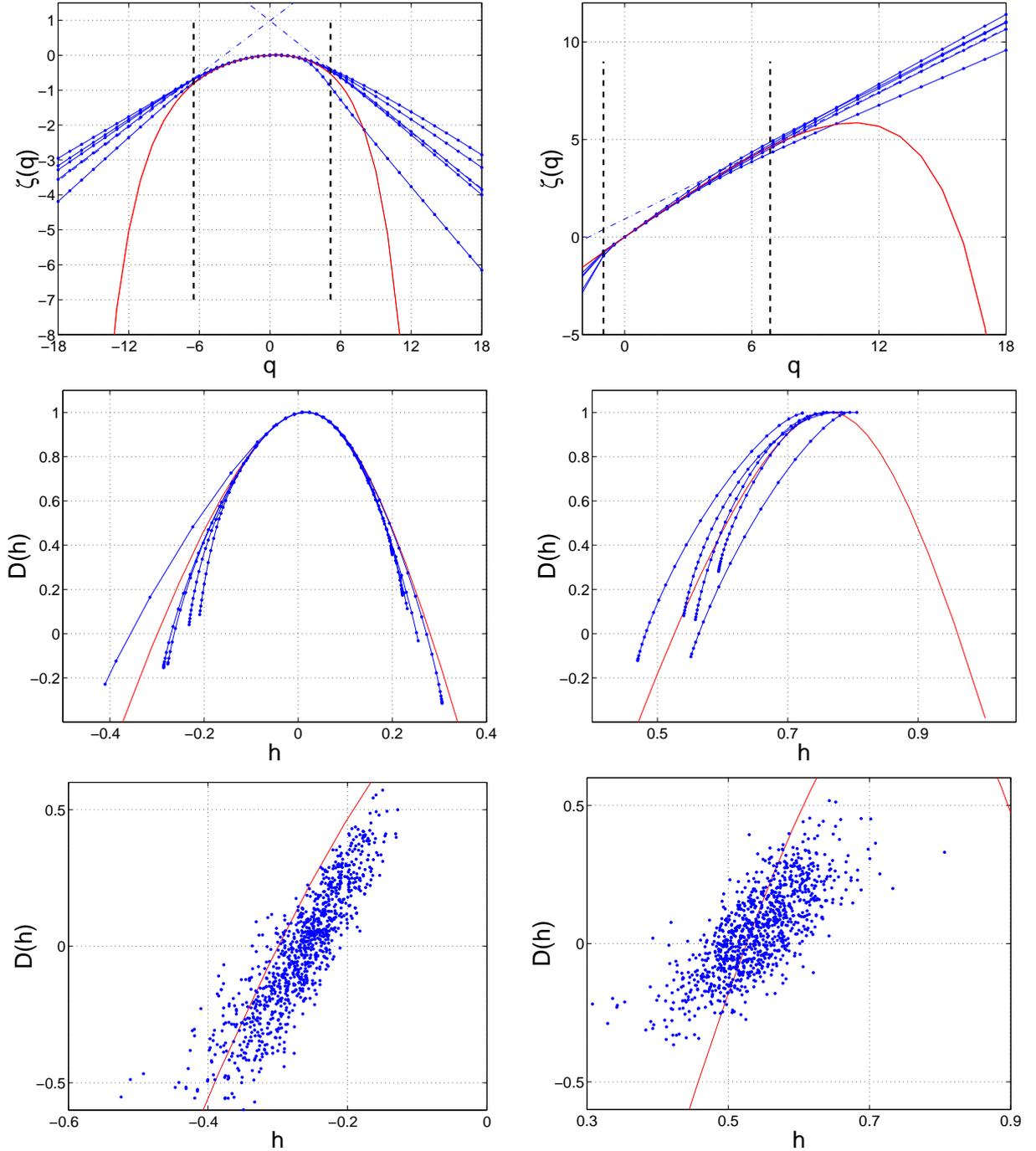
Fig. 3. **linearisation Effect: empirical facts.** Left column, $EI(1)$ applied to a CPC, Q_r (with $\mathbb{E}W^q = \exp(\mu q + \sigma^2/2q^2)$, $\mu = 0.04, \sigma^2 = 0.03$). Right column, $EIII(3)$ applied to the corresponding V_H, (with $H = 3/4$). Top row, linearisation effect observed on 5 independent realisations. Solid curves stand for the theoretical function $\zeta(q)$, derived from $\varphi(q), q \in \mathbb{R}$, according to Eqs. (3.20) or (3.21), while dotted ones denote the estimated $\hat{\zeta}(q, n)$. Beyond critical values, the estimates systematically follow linear behaviours in q. The vertical dashed lines indicate the positions of the theoretical $q_*^\pm$. Middle row, Legendre transforms $D(h)$ of $\zeta(q)$ (solid line) and $\hat{D}(h, n)$ of $\hat{\zeta}(q)$ (dotted line) for the same 5 realisations. $\hat{D}(h, n)$ and $D(h)$ roughly superimpose for a range of hs but $\hat{D}(h, n)$ is abruptly ended by accumulation points. Bottom row, accumulation points obtained from 1000 realisations of the same processes. The accumulation points are widely spread still along $D(h)$ (solid line) and centered around the point defined as $D(h) = 0$.
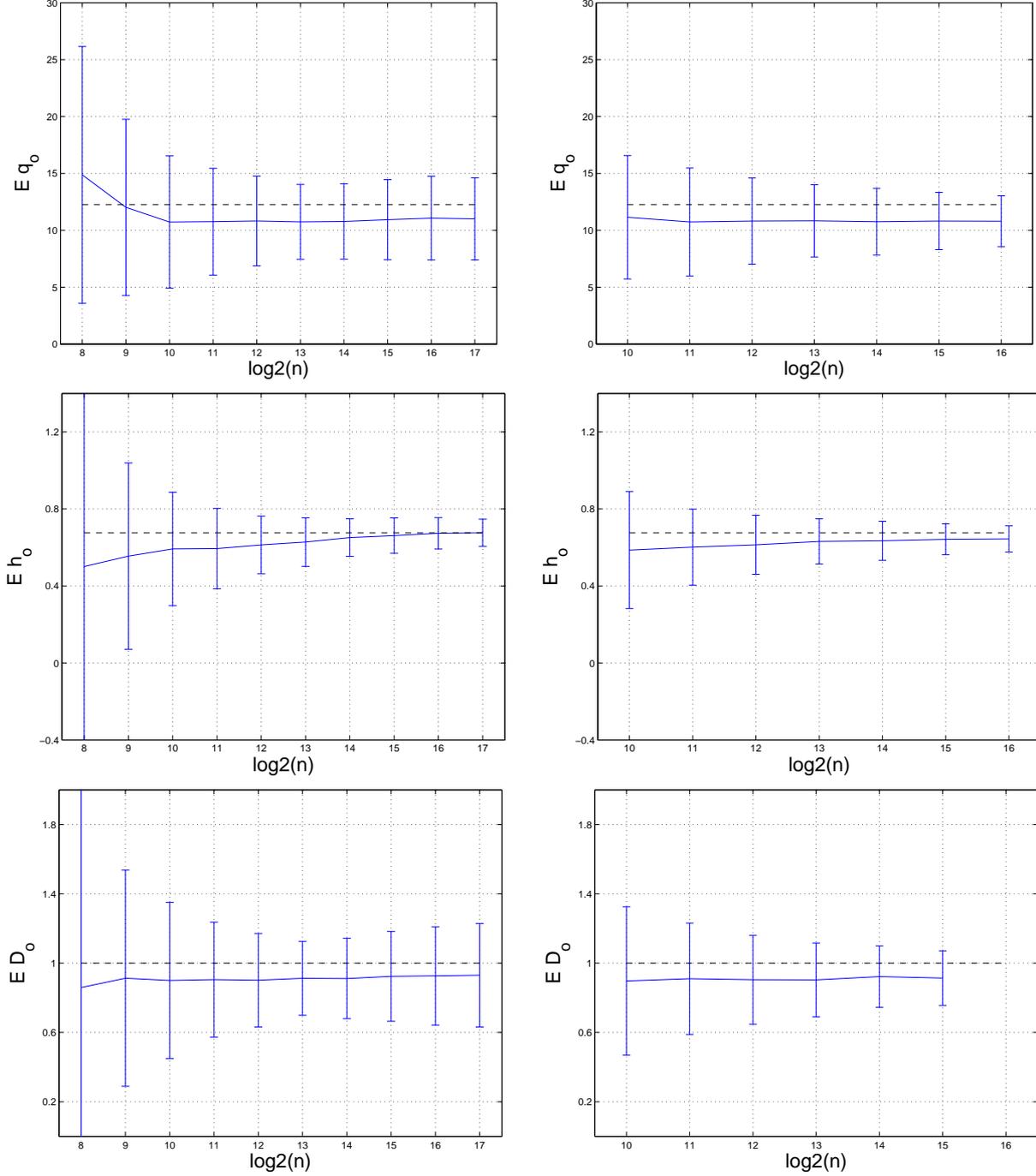
Fig. 4. **Dependence on n:** For a $CPC(I\!\!EW^q = \exp(\mu q + \sigma^2/2q^2), \ \mu = 0.05, \sigma^2 = 0.005), \ V_H(H = 3/4), EIII(3)$, left column, dependence on the number of integral scales, asymptotic behaviour — given $r, \ n_L \to +\infty$ —, the integral scale corresponds to 2^{J_L} with $J_L = 11$. Right column, dependence on the cascade resolution, asymptotic behaviour — given $n_L, \ r \to 0$ —, one integral scale, $n_L = 1$. Means plus and minus two standard deviations of the parameters characterising the linearisation effect as a function of $\log_2(n)$: top, $\hat{q}_o^+$, middle, $\hat{h}_o^+$, bottom, $\hat{D}_o^+$. Dashed lines denote the theoretical critical values D_*^+, h_*^+, q_*^+. Key observations are: the mean values of the parameters characterising the linearisation effect do not depend on n; for asymptotic behaviour — given $r, \ n_L \to +\infty$ — the variances decrease as long as $n \leq n_L = 2^{J_L}$ (i.e., the observation duration is shorter then the integral scale) but remain constant as soon as $n \gtrsim n_L$ (i.e., when the observation duration is larger than or equal to the integral scale).
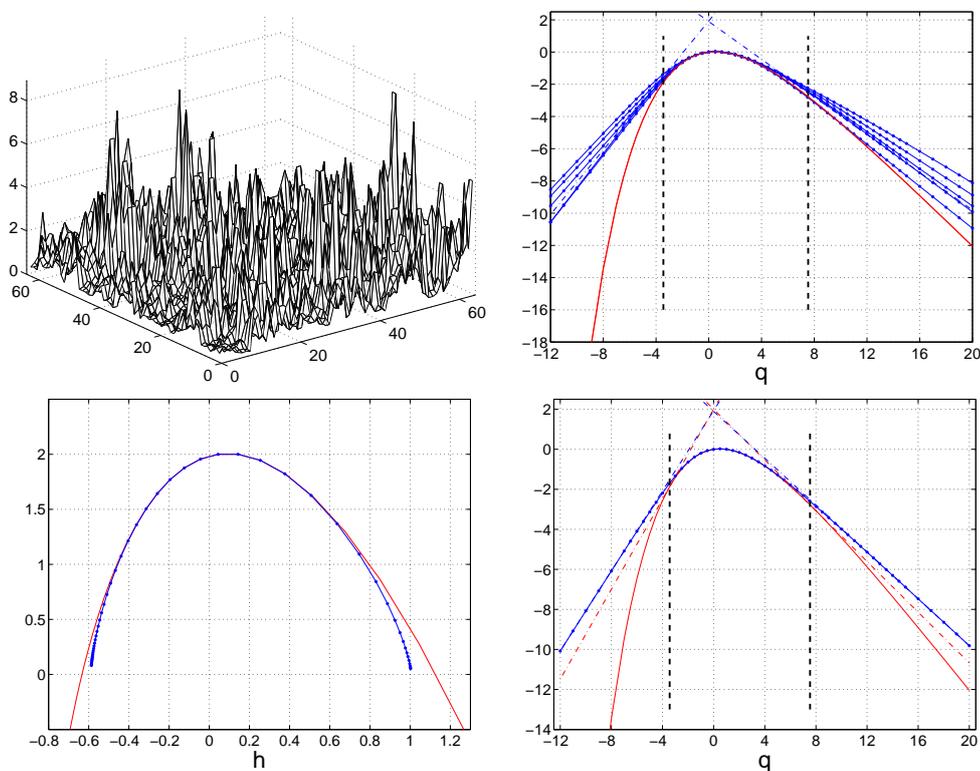
Fig. 5. **Linearisation effect: a 2D example.** Top left, a 2D CMC density Q_r. Right column, solid line: theoretical $\zeta(q)$, dotted line: $\hat{\zeta}(q, n)$, top: for 5 realisations, bottom, dotted line is obtained by averaging over thousands of realisations, dashed line corresponds theoretical ensemble average $I\!E\hat{\zeta}(q, n)$. Bottom left, Legendre transform of $\zeta(q)$ (derived from $\varphi(q), q \in \mathbb{R}$, according to Eqs. (3.20) or (3.21), solid line) and of $I\!E\hat{\zeta}(q, n)$ (dotted line). One clearly sees that a linearisation effect occurs as in the 1D case. The formulation of its characterisation can be straightforwardly extended from the 1D to the 2D case, mutatis mutandis.
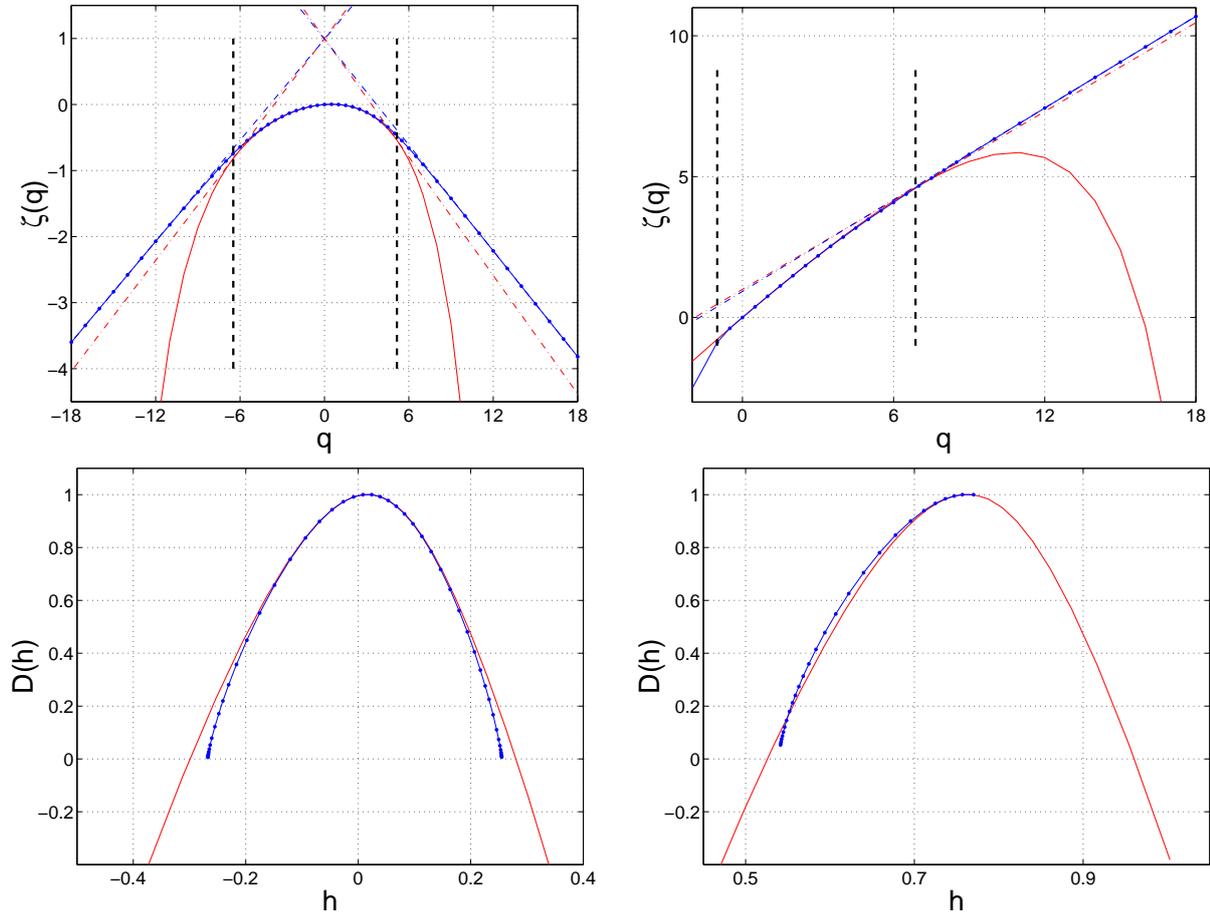
Fig. 6. **Linearisation Effect: illustration of the major result.** Left, on a CPC Q_r and $EI(1)$ and, right, on the corresponding V_H and $EIII(3)$, as in Fig. 3. Top row, Scaling exponents: solid line, theoretical $\zeta(q)$, derived from $\varphi(q), q \in \mathbb{R}$, according to Eqs. (3.20) or (3.21); mixed line, theoretical ensemble average $\mathbb{E}\hat{\zeta}(q, n)$ of the estimates $\hat{\zeta}(q, n)$, the corresponding linear behaviour is extended to the origin for clarity, dotted line, estimated value of $\mathbb{E}\hat{\zeta}(q, n)$ obtained by performing averages over a large number of realisations. The vertical dashed lines indicate the positions of the theoretical $q_*^{\pm}$. Bottom row, Legendre transform: solid line, theoretical $D(h)$, derived from $\varphi(q), q \in \mathbb{R}$, dotted line, estimated value of $\mathbb{E}\hat{D}(h, n)$ obtained from average over realisations.
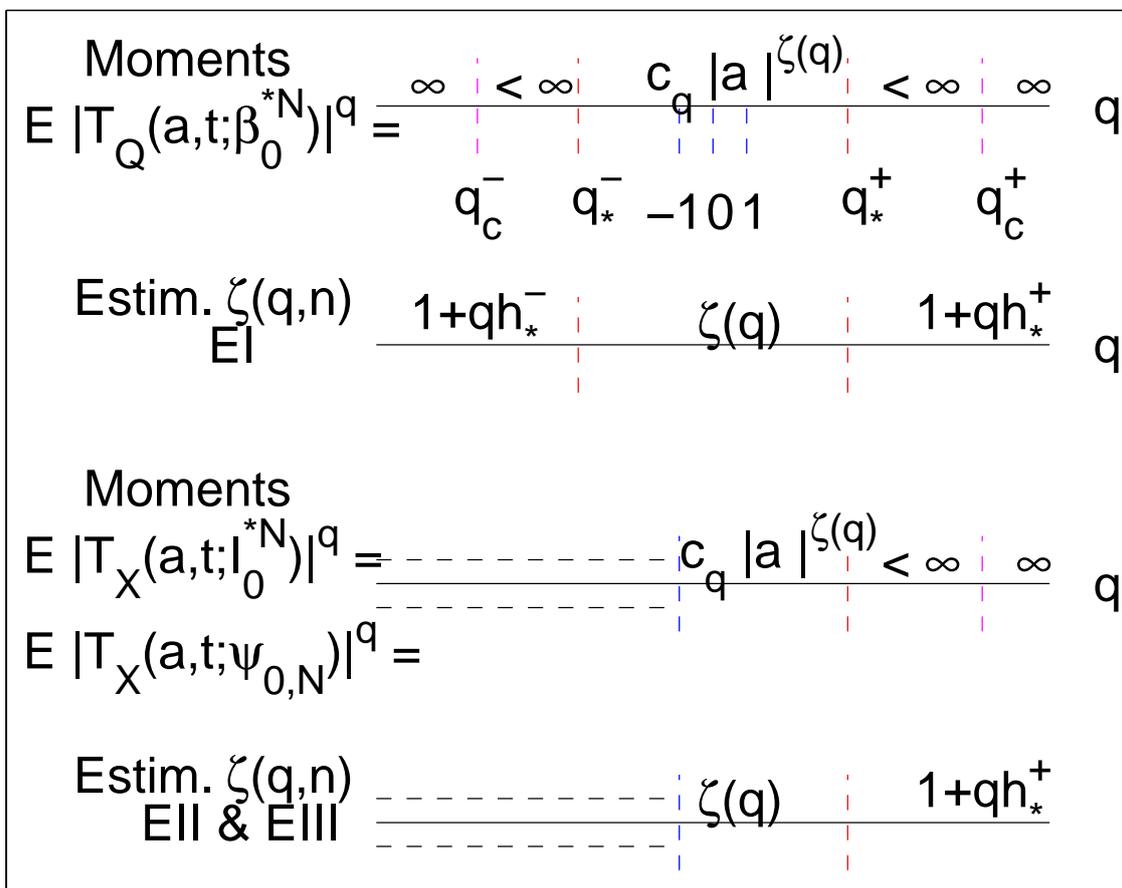
Fig. 7. **Linearisation Effect: Sketched Interpretation.** Let Q_r denote a density obtained from a multiplicative cascade and X a corresponding multifractal process (A, V_H, Y_H) . The moments of order q of the multiresolution coefficients $T_{Q_0}(a, t)$ (respectively, $T_X(a, t)$) are finite only when $q \in [q_c^-, q_c^+]$. Moreover, they behave as power laws of the scale a, with theoretical scaling exponents as given in Eqs. (3.20) or (3.21) only within a subinterval $q \in [q_*^-, q_*^+] \subset [q_c^-, q_c^+]$, and with scaling exponents $1 + qh_*^+$ (resp., $1 + qh_*^-$) when $q \in [q_*^+, q_c^+]$ (resp., $q \in [q_c^-, q_*^-]$). The estimators $\hat{\zeta}(q, n)$ account for the scaling exponents only within $q \in [q_*^-, q_*^+]$ and behave on average as $1 + qh_*^\pm$ elsewhere. The $\hat{\zeta}(q, n)$ are totally blind to $q_c^\pm$.