Submitted to the Annals of Statistics
arXiv: arXiv:????.????

INTERMITTENT PROCESS ANALYSIS WITH SCATTERING MOMENTS

By Joan Bruna[†], Stéphane Mallat[‡,*], Emmanuel Bacry[§], and Jean-François Muzy[§,¶]

New York University[†] Ecole Normale Supérieure[‡] CNRS, Ecole Polytechnique[§] and CNRS, Université de Corse[¶]

Scattering moments provide non-parametric models of random processes with stationary increments. They are expected values of random variables computed with a non-expansive operator, obtained by iteratively applying wavelet transforms and modulus non-linearities, which preserves the variance. First and second order scattering moments are shown to characterize intermittency and self-similarity properties of multiscale processes. Scattering moments of Poisson processes, fractional Brownian motions, Lévy processes and multifractal random walks are shown to have characteristic decay. The Generalized Method of Simulated Moments is applied to scattering moments to estimate data generating models. Numerical applications are shown on financial time-series and on energy dissipation of turbulent flows.

1. Introduction. Defining non-parametric models of non-Gaussian stationary processes remains a core issue of probability and statistics. Computing polynomial moments is a tempting strategy which suffers from the large variance of high order moment estimators. Image and audio textures are examples of complex processes with stationary increments, which can be discriminated from a single realization by the human brain. Yet, the amount of samples is often not sufficient to reliably estimate polynomial moments of degree more than 2. These non-Gaussian processes often have a long range dependency and some form of intermittency generated by randomly distributed burst of transient structures at multiple scales. Such multiscale intermittent processes appear in many domains including turbulent flows, network traffics, financial time series, geophysical and medical data.

Statistical instabilities in presence of intermittency can be reduced by calculating expected values of non-expansive operators in mean-square norm, which thus do not increase the variance. Polynomial moments do not satisfy this property, for degree different than one. Scattering moments are computed with such a non-expansive operator. They are calculated by iter-

*This work is supported by the ANR 10-BLAN-0126 and ERC InvariantClass 320959 grants.

atively applying wavelet transforms and modulus non-linearities [21]. This paper shows that they characterize self-similarity and intermittency properties of processes with stationary increments. These properties are studied by computing the scattering moments of Poisson processes, fractional Brownian motions, Levy processes and multifractal cascades, which all have very different behaviors. Scattering moments provide non parametric descriptors which reveal complex statistical properties of time series. The generalized method of simulated moments [13, 25] applied to scattering moments gives a parameter estimator for data generating models. Besides parameter estimation, a key challenge is to validate data generating models from limited data sets. Keeping sufficiently high order scattering coefficients provides a number of scattering moments which is larger than the dimensionality of the model parameter. Confidence levels for model validation can thus be computed with a χ^2 J-test [13].

Section 2 reviews the scaling properties of wavelet polynomial moments for fractal and multifractal processes. Scattering moments are defined and related to multiscale intermittency properties. Poisson processes illustrate these first results. Section 3 proves that self-similar processes with stationary increments have normalized scattering moments which are stationary across scales. Gaussian processes are discriminated from non-Gaussian processes from second order scattering moments. Results on fractional Brownian motion and stable Levy processes illustrate the analysis of multiscale intermittency properties. Section 4 extends these results to self-similar multifractal cascades [22, 23, 5].

Section 5 applies scattering moments to model parameter estimations. It introduces a scattering moment estimator whose variance is bounded. Parameters of data generating models are estimated from scattering moments with the generalized method of simulated moments [13, 25]. Scattering moments of financial time-series and turbulence energy dissipation are computed from numerical data. Models based on fractional Brownian, Levy stable and multifractal cascade processes are evaluated with a J-test. Computations can be reproduced with a software available at *www.di.ens.fr/data/software/scatnet*.

Notations: We denote $\{X(t)\}_t \overset{d}{=} \{Y(t)\}_t$ the equality of all finite-dimensional distributions. The dyadic scaling of $X(t)$ is written $L_j X(t) = X(2^{-j}t)$. If $X(t)$ is stationary then $\mathbf{E}(X(t))$ does not depend on t and is written $\mathbf{E}(X)$, and $\sigma^2(X) = \mathbf{E}(|X|^2) - \mathbf{E}(X)^2$. We denote $B(j) \simeq F(j)$, $j \to \infty$ (resp $j \to -\infty$) if there exists $C_1, C_2 > 0$ and $J \in \mathbb{Z}$ such that $C_1 \leq \frac{B(j)}{F(j)} \leq C_2$ for all $j > J$ (resp for all $j < J$).

2. Scattering Transform of Intermittent Processes.

2.1. *Polynomial Wavelet Moments.* Polynomial moments of wavelet coefficients reveal important multiscaling properties of fractals and multifractals [3, 31, 15, 14, 1, 28, 27]. We consider real valued random processes $X(t)$ having stationary increments $X(t) - X(t-\tau)$ for any $\tau \in \mathbb{R}$. A wavelet $\psi(t)$ is a function of zero average $\int \psi(t)\, dt = 0$ with $|\psi(t)| = O((1+|t|^2)^{-1})$, which is dilated:

$$\forall j \in \mathbb{Z} \quad , \quad \psi_j(t) = 2^{-j}\psi(2^{-j}t) \ .$$

The wavelet transform of $X(t)$ at a scale 2^j is defined for all $t \in \mathbb{R}$ by

$$X \star \psi_j(t) = \int X(u)\psi_j(t-u)du \ . \tag{1}$$

Since $\int \psi(t)\, dt = 0$, if X has stationary increments then one can verify that $X \star \psi_j(t)$ is a stationary process [31]. The dyadic wavelet transform of $X(t)$ is:

$$WX = \{X \star \psi_j\}_{j \in \mathbb{Z}} \ . \tag{2}$$

A wavelet ψ satisfies the Littlewood-Paley condition if its Fourier transform Ψ satisfies for all $\omega \neq 0$:

$$\sum_{j=-\infty}^{\infty} |\Psi(2^j\omega)|^2 + \sum_{j=-\infty}^{\infty} |\Psi(-2^j\omega)|^2 = 2 \ . \tag{3}$$

If $X(t)$ is a real valued stationary process with $\mathbf{E}(|X(t)|^2) < \infty$ then the wavelet transform energy equals the process variance $\sigma^2(X) = \mathbf{E}(|X|^2) - |\mathbf{E}(X)|^2$:

$$\mathbf{E}(\|WX\|^2) = \sum_{j \in \mathbb{Z}} \mathbf{E}(|X \star \psi_j|^2) = \sigma^2(X) \ . \tag{4}$$

This is proved by expressing $\mathbf{E}(|X \star \psi_j|^2)$ from the power spectrum of X and inserting (3).

The Holder regularity of a function is characterized by the decay of wavelet coefficients across scales. A wavelet $\psi(t)$ is said to have q vanishing if $\int t^k \psi(t) = 0$ for $0 \leq k < q$. If $\alpha < q$ then a function $x(t)$ is uniformly Lipschitz α if and only if $|x \star \psi_j(t)| = O(2^{j\alpha})$. This result can be localized to characterize the pointwise regularity of $x(t)$ [16].

For random processes $X(t)$, the decay of monomial wavelet moments across scales can be related to the distributions of singularities [3, 31, 15]. Moments of degree q define a scaling exponents $\zeta(q)$ such that

$$\mathbf{E}(|X \star \psi_j(t)|^q) \simeq 2^{j\zeta(q)} \ .$$

Monofractals such as fractional Brownian motions have linear scaling exponents: $\zeta(q) = q\,\zeta(1)$. These Gaussian processes are uniformly regular and are therefore not intermittent. The curvature of $\zeta(q)$ is related to the presence of singularities with different Holder exponents [14, 28]. It is used as a measurement intermittency. However, as q deviates from 1, estimations of moments become progressively more unstable which limits the application of this multifractal formalism to very large data sets.

2.2. *Scattering Moments.* Scattering moments are expected values of a non-expansive transformation of the process. They are computed with a cascade of wavelet transforms and modulus non-linearities [21]. We review their elementary properties.

Let ψ be a $\mathbf{C^1}$, complex wavelet, whose real and imaginary parts are orthogonal, and have the same $\mathbf{L}^2(\mathbb{R})$ norm. In this paper we impose that ψ has a compact support normalized to $[-1/2, 1/2]$, which simplifies the proofs. However, most results remain valid without this compact support hypothesis. We consider wavelets ψ which are nearly analytic, in the sense that their Fourier transform $\Psi(\omega)$ is nearly zero for $\omega < 0$. The compact support hypothesis prevents it from being strictly zero. All numerical computations in the paper are performed with the compactly supported complex wavelets of Selesnick [35], whose real and imaginary parts have 4 vanishing moments and are nearly Hilbert transform pairs.

Let $X(t)$ be a real valued process with stationary increments having finite first order moments: $\mathbf{E}(|X(t) - X(t-\tau)|) < \infty$ for all $\tau \in \mathbb{R}$. The wavelet transform $X \star \psi_{j_1}(t)$ is a complex stationary random process. First order scattering moments are defined by

$$\forall j_1 \in \mathbb{Z} \quad , \quad \overline{S}X(j_1) = \mathbf{E}(|X \star \psi_{j_1}|) \ .$$

First order scattering moments provide no information on the time distribution of singularities or transient structures. This information is partly provided by second order scattering moments computed from the wavelet transform of each $|X \star \psi_{j_1}(t)|$:

$$\forall (j_1, j_2) \in \mathbb{Z}^2 \quad , \quad \overline{S}X(j_1, j_2) = \mathbf{E}(||X \star \psi_{j_1}| \star \psi_{j_2}|) \ .$$

These moments measure the average multiscale time variations of $|X\star\psi_{j_1}(t)|$, with a second family of wavelets ψ_{j_2}. If $j_2 < j_1$ then $\overline{S}X(j_1,j_2)$ has a fast decay to zero as $j_1 - j_2$ increases. Its amplitudes depend on the wavelet properties as opposed to the properties of X. Indeed, if $|\psi|$ is $\mathbf{C}^p$ and has p vanishing moments then $|X \star \psi_{j_1}|$ is typically almost everywhere $\mathbf{C}^p$ so $\overline{S}X(j_1,j_2) = \mathbf{E}(||X \star \psi_{j_1}| \star \psi_{j_2}|) = O(2^{p(j_2-j_1)})$. We thus concentrate on scattering moments for $j_2 > j_1$.

The expected value of second order moments averages the time variability of $||X \star \psi_{j_1}| \star \psi_{j_2}(t)|$. This lost information can be recovered by calculating the wavelet transform of $||X \star \psi_{j_1}| \star \psi_{j_2}(t)|$ for each (j_1,j_2). Iterating this process computes scattering moments at any order $m \geq 1$:

$$\forall (j_1,...,j_m) \in \mathbb{Z}^m \quad , \quad \overline{S}X(j_1,...,j_m) = \mathbf{E}(|\ |X \star \psi_{j_1}| \star ...| \star \psi_{j_m}|)\ . \tag{5}$$

If $\mathbf{E}(|X(t) - X(t-\tau)|) < \infty$ for all $\tau \in \mathbb{R}$ then $\mathbf{E}(|X \star \psi_{j_1}|) < \infty$ and one can verify by induction on m that $\overline{S}X(j_1,...,j_m) < \infty$.

The vector of all scattering moments of X defines a non-parametric representation

$$\overline{S}X = \left\{ \overline{S}X(j_1,...,j_m)\ :\ \forall (j_1,...,j_m) \in \mathbb{Z}^m\ ,\ \forall m \in \mathbb{N}^* \right\}\ .$$

Its $\mathbf{l^2}$ norm is

$$\|\overline{S}X\|^2 = \sum_{m=1}^{\infty} \sum_{(j1,...,j_m)\in\mathbb{Z}^m} |\overline{S}X(j_1,...,j_m)|^2. \tag{6}$$

Since the wavelet transform preserves the variance in (4) and the modulus operators obviously also preserves the variance, each wavelet transform and modulus iteration preserves the variance. If $\mathbf{E}(|X|^2) < \infty$ then by applying (4), we verify [21] by induction on l that scattering coefficients satisfy

$$\sum_{m=1}^{l-1} \sum_{(j1,...,j_m)\in\mathbb{Z}^m} |\overline{S}X(j_1,...,j_m)|^2 = \mathbf{E}(|X|^2) - |\mathbf{E}(X)|^2 - \sum_{(j1,...,j_l)\in\mathbb{Z}^l} \mathbf{E}(|\ |X\star\psi_{j_1}|\star...|\star\psi_{j_l}|^2)\ .$$

It results that

$$\|\overline{S}X\|^2 \leq \sigma^2(X) = \mathbf{E}(|X|^2) - |\mathbf{E}(X)|^2\ . \tag{7}$$

Numerical experiments indicate that for large classes of ergodic stationary processes, $\sum_{(j1,...,j_l)\in\mathbb{Z}^l} \mathbf{E}(|\ |X \star \psi_{j_1}| \star ...| \star \psi_{j_l}|^2)$ converges to zero as 2^l increases. It then implies that (7) is an equality. Similarly to the Fourier

power spectrum, the $\mathbf{l}^2$ norm of scattering moments is then equal to the variance. However, this remains a conjecture [21].

The scattering norm (6) can be approximated with a summation restricted to moments of order $m = 1, 2$, because higher order scattering moments usually have a much smaller energy [2, 7]. First and second order scattering moments applied to image and audio textures provide state of the art classification errors [7, 2, 36] with a deep convolution network architecture [20], but these results are strictly numerical. In the following, we concentrate on the mathematical properties of first and second order scattering moments, which characterize self-similarity and intermittency properties.

2.3. *Normalized Scattering and Intermittency.* Scattering moments are normalized to increase their invariance. Invariance to multiplicative factors is obtained with

$$\widetilde{S}X(j_1) = \frac{\overline{S}X(j_1)}{\overline{S}X(0)} = \frac{\mathbf{E}(|X \star \psi_{j_1}|)}{\mathbf{E}(|X \star \psi|)} \ .$$

Second order scattering moments are normalized by their first order moment:

$$\widetilde{S}X(j_1, j_2) = \frac{\overline{S}X(j_1, j_2)}{\overline{S}X(j_1)} = \frac{\mathbf{E}(||X \star \psi_{j_1}| \star \psi_{j_2}|)}{\mathbf{E}(|X \star \psi_{j_1}|)} \ .$$

This can be rewritten

$$\widetilde{S}X(j_1, j_2) = \overline{S}\widetilde{X}_{j_1}(j_2) = \mathbf{E}(|\widetilde{X}_{j_1} \star \psi_{j_2}|) \ \text{ with } \ \widetilde{X}_{j_1} = \frac{|X \star \psi_{j_1}|}{\mathbf{E}(|X \star \psi_{j_1}|)} \ .$$

If X has stationary increments then $\widetilde{X}_{j_1}$ is a normalized stationary process providing the occurrence of "burst" of activity at the scale 2^{j_1}. Normalized second order moments $\widetilde{S}X(j_1, j_2)$ thus measures the time variability of these burst of activity over time scales $2^{j_2} \geq 2^{j_1}$, which gives multiscale measurements of intermittency.

Intermittency is a geometric notion which characterizes the time distribution of transient structures and singularities. It is not modified by the action of derivative operators, which are translation invariant. We verify that this invariance property holds for normalized second order moments. Let d^α be a fractional derivative defined by the multiplication by $(i\omega)^\alpha$ in the Fourier domain. Since

$$d^\alpha X \star \psi_{j_1}(t) = 2^{-\alpha j_1} \, X \star \psi^\alpha_{j_1}(t)$$

where $\psi^\alpha = d^\alpha \psi$ and $\psi^\alpha_{j_1}(t) = 2^{-j_1}\psi^\alpha(2^{-j_1}t)$, it results that

$$\overline{S}d^\alpha X(j_1) = 2^{-\alpha j_1}\mathbf{E}(|X \star \psi^\alpha_{j_1}|) \tag{8}$$

and

$$\widetilde{S}d^{\alpha}X(j_1,j_2) = \frac{\mathbf{E}(||X \star \psi^{\alpha}_{j_1}| \star \psi_{j_2}|)}{\mathbf{E}(|X \star \psi^{\alpha}_{j_1}|)} \ . \tag{9}$$

If $X(t)$ has no oscillating singularity [16] then its wavelet coefficients calculated with ψ and ψ^{α} have the same asymptotic decay, so

$$\overline{S}d^{\alpha}X(j_1) \simeq 2^{-\alpha j_1}\overline{S}X(j_1) \ \text{ and } \ \widetilde{S}d^{\alpha}X(j_1,j_2) \simeq \widetilde{S}X(j_1,j_2) \ . \tag{10}$$

Modifications of regularity produced by derivative operators affect the decay of first-order scattering moments but not the decay of normalized second order moments. Fractional Brownian motions illustrate these properties in Section 3.2.

Global intermittency parameters computed with wavelet moments can be related to normalized second order scattering moments. Section 2.1 explained that multifractal analysis quantifies intermittency from scaling properties of wavelet moments. If $\mathbf{E}(|X \star \psi_j|^q) \simeq 2^{j\zeta q}$ then intermittency is measured by the curvature of $\zeta(q)$. It can be quantified by $\zeta(2) - 2\zeta(1)$ which satisfies

$$\frac{\mathbf{E}(|X \star \psi_j|^2)}{\mathbf{E}(|X \star \psi_j|)^2} \simeq 2^{j(\zeta(2)-2\zeta(1))} \ .$$

The following proposition relates this ratio to normalized second order scattering moments.

PROPOSITION 2.1. *If X has stationary increments then for any $j_1 \in \mathbb{Z}$*

$$\frac{\mathbf{E}(|X \star \psi_{j_1}|^2)}{\mathbf{E}(|X \star \psi_{j_1}|)^2} \geq 1 + \sum_{j_2=-\infty}^{+\infty} |\widetilde{S}X(j_1,j_2)|^2 \ . \tag{11}$$

Proof: Applying the mean-square energy conservation (4) to $X \star \psi_j$ proves that

$$\mathbf{E}(|X \star \psi_j|^2) = |\mathbf{E}(|X \star \psi_j|)|^2 + \sum_{j_2=-\infty}^{+\infty} \mathbf{E}(||X \star \psi_j| \star \psi_{j_2}|^2) \ . \tag{12}$$

Applying again (4) to $||X \star \psi_j| \star \psi_{j_2}|$ proves that

$$\mathbf{E}(||X \star \psi_j| \star \psi_{j_2}|^2) = \mathbf{E}(||X \star \psi_j| \star \psi_{j_2}|)^2 + \sum_{j_3=-\infty}^{+\infty} \mathbf{E}(|||X \star \psi_j| \star \psi_{j_2}| \star \psi_{j_3}|^2).$$

Inserting this equation in (12) proves (11). □

It results from (12) that if $\sum_{j_2=-\infty}^{+\infty} \widetilde{S}X(j_1,j_2)^2 \simeq 2^{j_1\beta}$ then $\zeta(2)-2\zeta(1) \geq \beta > 0$. However, these moment computations eliminate the dependence's on the scale parameter 2^{j_2}, which provides a finer multiscale characterization of the intermittency regularity. This dependency upon 2^{j_2} is studied in the next sections and is used for model selection in Section 5.

2.4. *Scattering Poisson Processes.* The properties of scattering moments are illustrated over a Poisson process, which is a simple Lévy process with stationary increments. A homogeneous Poisson process $\{X(t)\,,\, t \geq 0\}$ has increments $X(t+\Delta)-X(t)$ which count the number of occurrence of events in $[t, t+\Delta]$, and have a Poisson distribution of intensity λ. Figure 1(a) shows an example. The following proposition gives the decay of first and second order scattering moments of Poisson processes.

THEOREM 2.2. *If X is a Poisson process of intensity λ and $\bar{\psi}(t) = \int_0^t \psi(u)\,du$ then for all $j_1 \leq j_2$*

$$\overline{S}X(j_1) = 2^{j_1}\,\lambda \|\bar{\psi}\|_1 \Big(1 + O(2^{j_1}\lambda)\Big)\ , \tag{13}$$

$$\lim_{j_1 \to \infty} 2^{-j_1/2} \overline{S}X(j_1) = C\lambda^{1/2} > 0\ , \tag{14}$$

where C depends only upon the wavelet ψ, and

$$\tilde{S}X(j_1,j_2) = \frac{\| |\bar{\psi}| \star \psi_{j_2-j_1}\|_1}{\|\bar{\psi}\|_1} \Big(1 + O(\lambda 2^{j_1}) + O(\lambda 2^{j_2})\Big) \tag{15}$$

$$\lim_{j_2 \to \infty} 2^{j_2/2} \tilde{S}X(j_1,j_2) = C' > 0\ . \tag{16}$$

The proof is in Appendix A, but these results can be explained as follows. A Poisson distribution has on average 1 jump on intervals of size λ^{-1}. When $2^{j_1} \ll \lambda^{-1}$ then the probability that there is one jump on the support of ψ_{j_1} is $2^{j_1}\lambda$, and there is at most one jump with overwhelming probability. We have $\overline{S}X(j_1) \approx \lambda\, 2^{j_1} \|\bar{\psi}\|_1$ because the wavelet transform of a jump satisfies $\int |1_{[0,\infty)} \star \psi_{j_1}(t)|\,dt = \|\bar{\psi}\|_1$. For $2^{j_1} \gg \lambda^{-1}$, the average number of jumps on the support of ψ_{j_1} is $2^{j_1}\lambda^{-1}$. Since the jump locations are independent, Appendix A proves that $X \star \psi_j(t)$ converges to the wavelet transform of a Gaussian white noise of variance $\lambda\, 2^{j_1}$. The first order scattering moments are thus equal to the first order moments of these Gaussian random variables, which are proportional $\lambda^{1/2}\, 2^{j_1/2}$. Figure 1(b) verifies that $\log_2 \widetilde{S}X(j_1)$ has a slope of 1 for $j_1 \ll -\log_2(\lambda)$ and a slope of 1/2 for $j_1 \gg -\log_2(\lambda)$.

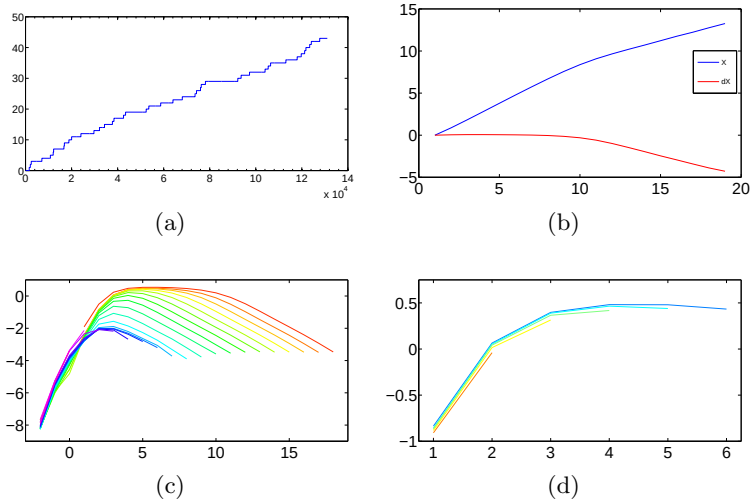

FIG 1. *(a): Realization of a Poisson process $X(t)$ of intensity $\lambda = 10^{-4}$. (b):* $\log_2 \widetilde{S}X(j)$ *and* $\log_2 \widetilde{S}dX(j)$ *as a function of j. (c):* $\log_2 \widetilde{S}X(j_1, j_2)$ *as a function of $j_2 - j_1$ for several values of j_1. (d): The same curves as in (c), but restricted to $j_2 < -\log_2(\lambda) - 1$.*

Second order moments $\widetilde{S}X(j_1, j_2)$ depend upon the signal intermittency. Their behavior is totally different when $2^{j_2} \ll \lambda^{-1}$ and $2^{j_2} \gg \lambda^{-1}$. If $2^{j_2} \ll \lambda^{-1}$ then $\widetilde{S}X(j_1, j_2)$ is approximately equal to $\||\bar{\psi}| \star \psi_{j_2 - j_1}\|_1 / \|\bar{\psi}\|_1$, which is a function of $j_2 - j_1$. This is a self-similarity property. Indeed, it depends on the wavelet transform of a jump $1_{[0,\infty)}(t)$, which is a self-similar function. When $j_2 - j_1$ increases

$$\lim_{j_1 \to -\infty} \widetilde{S}X(j_1, j_2) = \|\psi\|_1 \left(1 + O(2^{j_2}\lambda)\right) .$$

This convergence to a constant indicates a high degree of intermittency corresponding to isolated singularities. Figure 1(d) shows $\log_2 \widetilde{S}X(j_2 - j_1, j_2)$ as a function of $j_2 - j_1$, for $j_2 \leq -\log(\lambda)$. As expected, these curves overlap for different j_1, and converge to $\|\psi\|_1$.

If $j_2 \gg -\log_2(\lambda^{-1})$ then $\widetilde{S}X(j_1, j_2) \simeq 2^{-j_2/2}$. This decay is characteristic of Gaussian stationary processes, which are uniformly regular and thus have no intermittency. This is further studied in Section 3.2 for fractional Brownian motions. Figure 1(c) verifies that $\widetilde{S}X(j_2 - j_1, j_2)$ decays with a slope of $-1/2$ as a function of $j_2 - j_1$. This behavior is typical of random processes having an "integral scale," here equal to λ^{-1}, beyond which the process converges to a Gaussian process.

When going from X to dX then the sum of jumps is replaced by a measure

which is a sum of Diracs. We verify from Appendix A that $\overline{S}dX(j_1) \simeq 2^{-j_1}\overline{S}X(j_1)$. This reflects the change of regularity. Figure 1(b) shows that the difference between the slopes of $\log_2 \widetilde{S}X(j_1)$ and $\log_2 \widetilde{S}dX(j_1)$ is indeed equal to 1. The situation is different for normalized second order moments where $\widetilde{S}dX(j_1, j_2)$ is nearly equal to $\widetilde{S}X(j_1, j_2)$. When $2^{j_1} \ll \lambda^{-1}$ then dX and X have isolated singularities occurring with same probability distribution so have the same intermittency. When $2^{j_1} \gg \lambda^{-1}$ then $dX \star \psi_j$ and $X \star \psi_j$ converge to stationary Gaussian processes having different power spectrum and hence different regularity, but they have the same intermittency because their regularity is uniform.

3. Self-Similar Processes. Second order scattering moments of self-similar processes are proved to be stationary across scales. Fractional Brownian motions and Lévy stable processes are studied in Sections 3.2 and 3.3.

3.1. *Scattering Self-Similarity.* Self-similar processes of Hurst exponent H are stochastic processes $X(t)$ which are invariant in distribution under a scaling of space or time:

$$\forall\, s > 0\ ,\ \{X(st)\}_t \overset{d}{=} \{s^H X(t)\}_t\ . \tag{17}$$

We consider self-similar processes having stationary increments. Fractional Brownian motions and α-stable Lévy processes are examples of Gaussian and non-Gaussian self-similar processes with stationary increments.

If X is self-similar, then applying (17) with a change of variable $u' = 2^{-j}u$ in (1) proves that

$$\forall j \in \mathbb{Z}\quad ,\quad \{X \star \psi_j(t)\}_t \overset{d}{=} 2^{jH}\,\{X \star \psi(2^{-j}t)\}_t\ .$$

The following proposition proves that normalized second order scattering moments can be written as a univariate function.

PROPOSITION 3.1. *If X is a self-similar process with stationary increments then for all $j_1 \in \mathbb{Z}$*

$$\widetilde{S}X(j_1) = 2^{j_1 H}\,, \tag{18}$$

and for all $(j_1, j_2) \in \mathbb{Z}^2$

$$\widetilde{S}X(j_1, j_2) = \overline{S}\widetilde{X}(j_2 - j_1) \quad \textit{with} \quad \widetilde{X}(t) = \frac{|X \star \psi(t)|}{\mathbf{E}(|X \star \psi|)}\ . \tag{19}$$

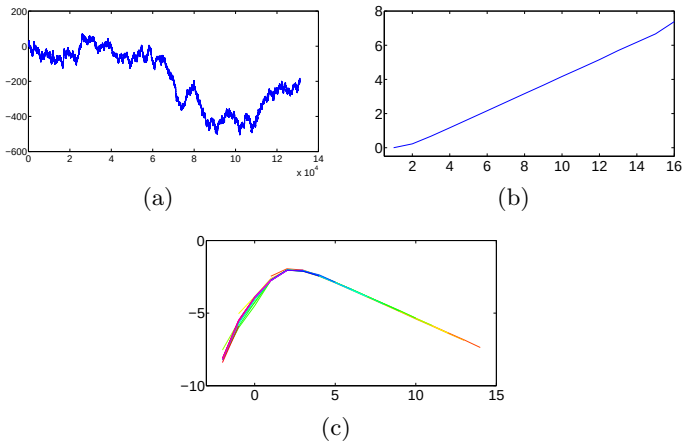

FIG 2. *(a): Realization of a Brownian motion $X(t)$. (b):* $\log_2 \widetilde{S}X(j_1)$ *as a function of* j_1. *(c) The curves* $\log_2 \widetilde{S}X(j_1, j_1 + l)$ *as a function of* l *are identical for different* j_1.

Proof: We write $L_j x(t) = x(2^{-j}t)$. Since $\psi_{j_1} = 2^{-j_1} L_{j_1}\psi$, a change of variables yields $L_{j_1}|X \star \psi| = |L_{j_1} X \star \psi_{j_1}|$, and hence

$$|X \star \psi_{j_1}| = L_{j_1}|L_{-j_1} X \star \psi| \overset{d}{=} 2^{j_1 H}\, L_{j_1}|X \star \psi| \ . \tag{20}$$

If $Y(t)$ is stationary then $\mathbf{E}(L_j Y(t)) = \mathbf{E}(Y(t))$, which proves (18).

By cascading (20) we get

$$\forall\, (j_1, j_2) \ , \ ||X \star \psi_{j_1}| \star \psi_{j_2}| \overset{d}{=} 2^{j_1 H}\, L_{j_1}||X \star \psi| \star \psi_{j_2 - j_1}| \ , \tag{21}$$

so $\overline{S}X(j_1, j_2) = 2^{j_1 H}\, \mathbf{E}(||X \star \psi| \star \psi_{j_2 - j_1}|)$. Together with (18) it proves (19). □

Property (19) proves that if X is self-similar then $\widetilde{S}X(j_1, j_1 + l)$ is a function of l, which can be interpreted as a stationary property across scales. This function of l is a scattering intermittency measure of the random process. A Brownian motion is a Gaussian self-similar process with a Hurst exponent $H = 1/2$. It results from (18) that $\log_2 \widetilde{S}(j_1) = j_1/2$, which is illustrated by Figure 2 (b). Figure 2 (c) displays $\widetilde{S}X(j_1, j_2)$ expressed as a function of $j_2 - j_1$, for different j_1. The curves for different j_1 are equal, as proved by (18). When $j_2 - j_1 < 0$, $\widetilde{S}X(j_1, j_2)$ increases with a slope which does not depend on X but on the number of vanishing moments and on the regularity of the wavelet ψ. For $j_2 - j_1 \geq 0$, the decay depends upon the property of X and satisfies $\widetilde{S}X(j_1, j_2) \simeq 2^{-(j_2 - j_1)/2}$. Next section proves this result in the more general context of fractional Brownian motions, and shows that it reflects the fact that a Brownian motion is a Gaussian process.

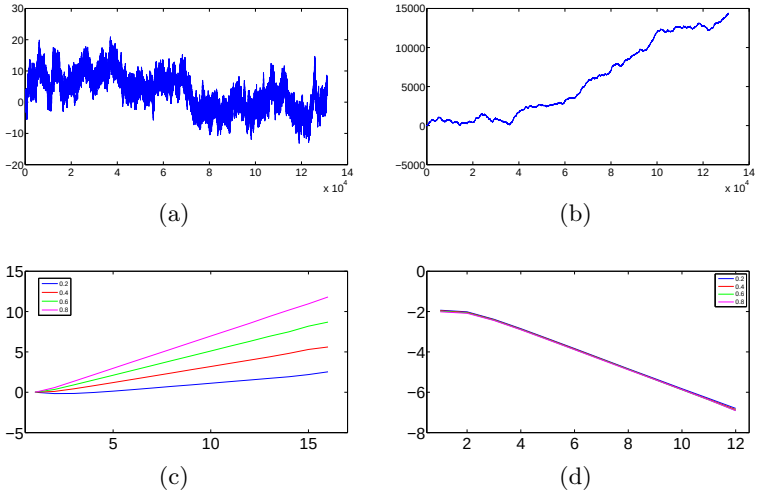

FIG 3. *(a,b) Realizations of fractional Brownian motions $X(t)$ with $H = 0.2$ in (a) and $H = 0.8$ in (b). (c) $\log_2 \widetilde{S}X(j_1)$ as a function of j_1, for $H = 0.2$, 0.4, 0.6, 0.8. Slopes are equal to H. (d) $\log_2 \widetilde{S}X(j_1, j_1 + l)$ as a function of l do not depend on j_1 for all H.*

3.2. *Fractional Brownian Motions.* We compute the normalized scattering representation of Fractional Brownian Motions, which are the only self-similar Gaussian processes with stationary increments. A fractional Brownian motion of Hurst exponent $0 < H < 1$ is defined as a zero mean Gaussian process $\{X(t)\}$, satisfying

$$\forall\, t, s > 0\ , \mathbf{E}(X(t)X(s)) = \frac{1}{2}\left(t^{2H} + s^{2H} - |t-s|^{2H}\right)\mathbf{E}(X(1)^2)\ .$$

It is self-similar and satisfies

$$\forall\, s > 0\ ,\ \{X(st)\}_t \overset{d}{=} s^H \{X(t)\}_t\ .$$

Proposition 3.1 proves in (18) that

$$\widetilde{S}X(j_1) = 2^{Hj_1}\ .$$

This is verified by Figure 3(a) which shows that $\log_2 \widetilde{S}X(j_1) = H\, j_1$ for several fractional Brownian motions with $H = 0.2, 0.4, 0.6, 0.8$.

Figure 3(c) displays $\log_2 \widetilde{S}(j_1, j_2)$, which is a function of $j_2 - j_1$, as proved by (19). Modulo a proper initialization at $t = 0$, if X is a fractional Brownian motion of exponent H then $d^\alpha X$ is a fractional Brownian motion of exponent $H - \alpha$. We thus expect from (10) that $\log_2 \widetilde{S}X(j_2 - j_1)$ nearly does not

depend upon H. This is shown by Figure 3(c) where all curve superimpose for $j_2 - j_1 > 0$, with a slope of $-1/2$. This result is proved by the following theorem.

THEOREM 3.2. *Let $X(t)$ be a Fractional Brownian Motion with Hurst exponent $0 < H < 1$. There exists a constant $C > 0$ such that for all $j_1 \in \mathbb{Z}$*

$$\lim_{l \to \infty} 2^{-l/2} \widetilde{S}X(j_1, j_1 + l) = C \ . \tag{22}$$

Proof: Proposition 3.1 proves in (19) that $\widetilde{S}X(j_1, j_1 + j) = \mathbf{E}(|\widetilde{X} \star \psi_j|)$ and $\widetilde{X}(t) = |X \star \psi(t)|/\mathbf{E}(|X \star \psi|)$. We denote $\widetilde{S}_2 X(j) = \mathbf{E}(|\widetilde{X} \star \psi_j|)$. Let $B(t)$ be a Brownian motion and $dB(t)$ be the Wiener measure. The two processes $X \star \psi(t)$ and $d^{H-1}dB \star \psi(t)$ are Gaussian stationary processes having same power spectrum so

$$\{|X \star \psi(t)|\}_t \stackrel{d}{=} \{|d^{H-1}dB \star \psi(t)|\}_t \stackrel{d}{=} \{|dB \star d^{H-1}\psi(t)|\}_t \ .$$

It results that

$$\widetilde{S}_2 X(j) = \frac{\mathbf{E}(||dB \star d^{H-1}\psi| \star \psi_j|)}{\mathbf{E}(|X \star \psi|)} \tag{23}$$

Since ψ is $\mathbf{C}^1$, with a compact support and two vanishing moments, one can verify that $|d^{H-1}\psi(u)| = O((1 + |u|^2)^{-1})$. It results that $|dB \star d^{H-1}\psi|$ is stationary process whose autocorrelation has some decay. As the scale 2^j increases, the second convolution with ψ_j performs a progressively wider averaging. By applying a central-limit theorem for dependent random variables, the following lemma applied to $\varphi = d^{H-1}\psi$ proves that $2^{j/1}|dB \star d^{H-1}\psi| \star \psi_j$ converges to a Gaussian processes and that its first moment converges to a constant when j goes to ∞. The theorem result (22) stating that $2^{j/2}\widetilde{S}_2 X(j)$ converges to a constant results from (23).

LEMMA 3.3. *If $\varphi(u) = O((1 + |u|^2)^{-1})$ then*

$$2^{j/2}|dB \star \varphi| \star \psi_j(t) \xrightarrow[j \to \infty]{l} \mathcal{N}(0, \sigma^2 \mathrm{Id}) \ , \tag{24}$$

with $\sigma^2 = \|\psi\|_2^2 \int R_Y(\tau) d\tau$ and

$$\lim_{j \to \infty} \mathbf{E}(|2^{j/2}|dB \star \varphi| \star \psi_j|) = \sigma \sqrt{\frac{\pi}{2}} \ . \quad \square \tag{25}$$

For a fractional Brownian motions, $\log_2 \widetilde{S}X(j_1, j_1 + l)$ do not depend on j_1 or H, and their slopes is thus equal to $-1/2$ when l increases. This value is characteristic of wide-band Gaussian stationary processes. It indicates that there is no intermittency phenomenon at all scales.

3.3. *α-stable Lévy Processes.* In this section, we compute the scattering moments of α-stable Lévy processes and analyze their intermittency behavior for $1 < \alpha \leq 2$. These processes have finite polynomial moments only for degree strictly smaller than $\alpha \leq 2$. The Lévy-Khintchine formula [19] characterizes infinitely divisible distributions from their characteristic exponents. Self-similar Lévy processes have stationary increments with heavy tailed distributions. Their realizations contain rare, large jumps, which are responsible for the blow up of moments larger than α. They induce a strongly intermittency behavior.

For $\alpha > 1$, an α-stable Lévy process $X(t)$ has stationary increments and $\mathbf{E}(|X(t) - X(t-\tau|) < \infty$ for any $\tau \in \mathbb{R}$. Its scattering moments are thus well defined at all orders. This process satisfies the self-similarity relation

$$\{X(st)\}_t \overset{d}{=} s^{\alpha^{-1}} \{X(t)\}_t \ , \tag{26}$$

so Proposition 3.1 proves that

$$\widetilde{S}X(j_1) = 2^{j_1 \alpha^{-1}} \ . \tag{27}$$

This is verified in Figure 4 which shows that $\log_2 \widetilde{S}X(j_1) = \alpha^{-1} j_1$. First order moments thus do not differentiate a Lévy stable processes from fractional Brownian motions of Hurst exponent $H = \alpha^{-1}$.

The self-similarity implies that $\widetilde{S}X(j_1, j_1 + l)$ does not depend on j_1. However, they have a very different behavior than second order scattering moments of fractional Brownian motion. Figure 4 shows that $\log_2 \widetilde{S}X(j)$ has a slope which tends to $\alpha^{-1} - 1$ and hence that when l increases

$$\widetilde{S}X(j_1, j_1 + l) \simeq 2^{l(\alpha^{-1}-1)} \ . \tag{28}$$

For $\alpha < 2$ then $\alpha^{-1} - 1 > -1/2$ so $\widetilde{S}(j_1, j_1 + l)$ has a slower decay for α-stable Lévy processes than for fractional Brownian motion, which corresponds to the fact that these processes are highly intermittent and the intermittency increases when α decreases. For $\alpha = 2$, the Lévy process X is a Brownian motion and we recover that $\widetilde{S}X(j_1, j_1 + l) \simeq 2^{-l/2}$ as proved in Theorem 3.2.

The scaling property (28) is explained qualitatively, without formal proof. We proved in (19) that

$$\widetilde{S}X(j_1, j_2) = \frac{\mathbf{E}(||X \star \psi(t)| \star \psi_l|)}{\mathbf{E}(|X \star \psi|)} \quad \text{for} \quad l = j_2 - j_1. \tag{29}$$

The stationary process $|X \star \psi(t)|$ measures the amplitude of local variations of the process X. It is dominated by a sparse sum of large amplitude bumps

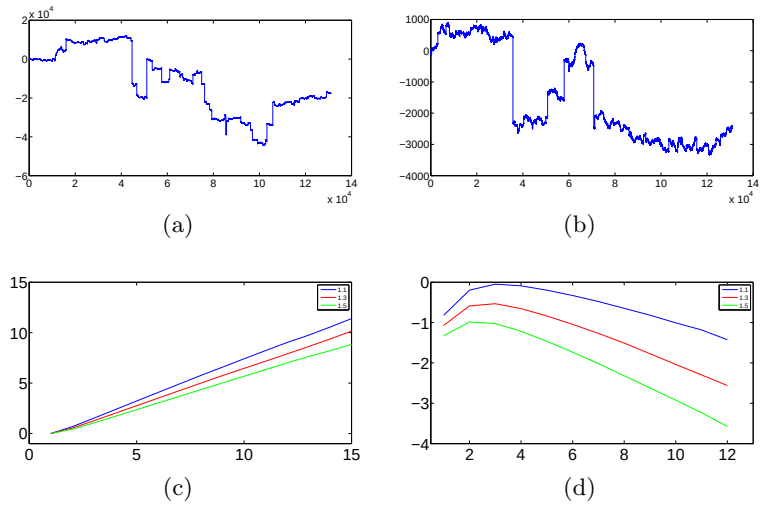

FIG 4. *(a,b): Realizations of α-stable Lévy processes $X(t)$ with $\alpha = 1.1$ and $\alpha = 1.5$. (c) $\log_2 \widetilde{S}X_\alpha(j_1)$ as a function of j_1 with $\alpha = 1.1, 1.2, 1.3$. Slopes are equal to α^{-1}. (d) $\log_2 \overline{S}X_\alpha(j_1, j_1 + l)$ as a function of l do not depend on j_1. Slopes tend to $\alpha^{-1} - 1$ when l increases.*

of the form $a\,|\psi(t-u)|$, where a and u are the random amplitudes and positions of rare large amplitude jumps in $X(t)$, distributed according to the Lévy measure. It results that

$$\mathbf{E}(||X \star \psi| \star \psi_l|) \simeq \mathbf{E}(|dX \star |\bar{\psi}| \star \psi_l|) \text{ with } \overline{\psi}(t) = \int_0^t \psi(u)\,du. \tag{30}$$

If $2^l \gg 1$ then $|\bar{\psi}| \star \psi_l \approx \|\bar{\psi}\|_1\, \psi_l$, and $\mathbf{E}(|dX \star \psi_l|) \simeq 2^{l(\alpha^{-1}-1)}$ because the Lévy measure $dX(t)$ satisfies the self-similarity property

$$\{dX(st)\}_t \overset{d}{=} s^{\alpha^{-1}-1}\{dX(t)\}_t \ .$$

Inserting (30) in (29) gives the scaling property (28).

4. Scattering Moments of Multiplicative Cascades. We study the scattering representation of multifractal processes which satisfy a stochastic scale invariance property. Section 4.2 studies the particularly important case of log-infinitely divisible multiplicative processes.

4.1. *Stochastic Self-Similar Processes.* We consider processes with stationary increments which satisfy the following stochastic self-similarity:

$$\forall\, 1 \geq s > 0\ ,\ \{X(st)\}_{t \leq 2^L} \overset{d}{=} A_s \cdot \{X(t)\}_{t \leq 2^L}\ , \tag{31}$$

where A_s is a log-infinitely divisible random variable independent of $X(t)$ and the so-called *integral scale* 2^L is chosen (for simplicity) as a power of 2. The Multifractal Random Measures (MRM) introduced by [30, 6] are important examples of such processes. Let us point out that MRM's are stationary increments versions of the multiplicative cascades initially introduced by Yaglom [38] and Mandelbrot [22, 23], and further studied by Kahane and Peyriere [17]. Strictly speaking, these multiplicative cascades do not satisfy (31). However they do satisfy an extremely similar equation when sampled on a dyadic-grid, and consequently, all the results that we obtained on MRM's are easily generalized to multiplicative cascades. For the sake of conciseness, we did not include them here.

Since X has stationary increments and satisfies (31), with a change of variables we verify that

$$\forall j \leq L \quad , \quad \{X \star \psi_j(t)\}_t \stackrel{d}{=} A_{2^j} \{X \star \psi(2^{-j}t)\}_t \ ,$$

and hence for all $q \in \mathbb{Z}$ and $j \leq L$

$$\mathbf{E}(|X \star \psi_j|^q) = \mathbf{E}(|A_{2^j}|^q)\, \mathbf{E}\{|X \star \psi|^q\} \simeq C_q 2^{j\zeta(q)} \ , \tag{32}$$

where $\zeta(q)$ is a priori a non-linear function of q. Since the self-similarity is upper bounded by an integral scale, the convexity of moments [14] implies that $\zeta(q)$ is a concave function of q. Similarly to Proposition 3.1, the following proposition shows that normalized scattering moments capture stochastic self-similarity with a univariate function.

PROPOSITION 4.1. *If X is randomly self-similar in the sense of (31) with stationary increments then for all $j_1 \leq L$*

$$\widetilde{S}X(j_1) = \mathbf{E}(|A_{2^{j_1}}|)\,. \tag{33}$$

If $2^{j_1} + 2^{j_2} \leq L$ then

$$\widetilde{S}X(j_1, j_2) = \overline{S}\widetilde{X}(j_2 - j_1) \quad \textit{with} \quad \widetilde{X}(t) = \frac{|X \star \psi(t)|}{\mathbf{E}(|X \star \psi|)} \ . \tag{34}$$

Proof: Property (18) is particular case of (32) for $q = 1$. If $2^{j_1} + 2^{j_2} \leq L$, with the same derivations as for (21), we derive from (31) that

$$||X \star \psi_{j_1}| \star \psi_{j_2}| \stackrel{d}{=} A_{2^{j_1}}\, L_{j_1} ||X \star \psi| \star \psi_{j_2 - j_1}| \ , \tag{35}$$

so $\overline{S}X(j_1, j_2) = \mathbf{E}(A_{2^{j_1}})\, \mathbf{E}(||X \star \psi| \star \psi_{j_2 - j_1}|)$. Together with (33) it proves (34). □.

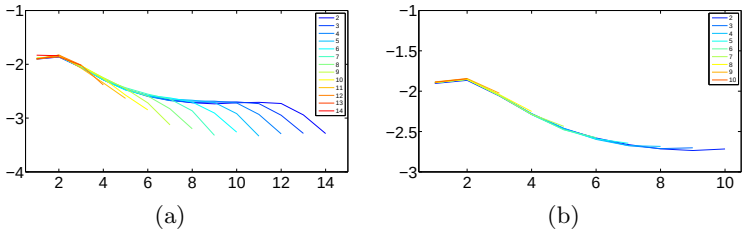

FIG 5. *(a):* $\log_2 \overline{S} X_\alpha(j_1, j_1 + l)$ *as a function of* l *for a Multifractal Random Measure (MRM) with* $\lambda^2 = 0.04$ *and an integral scale* $2^L = 2^{13}$*. Different colors stand for different values of* j_1*. (b): Same curves restricted to* $j_2 = j_1 + l < L - 1$*.*

Figure 5 shows the normalized scattering of a multiplicative cascade process described in Section 4.2, with an integral scale $2^L = 2^{17}$. When $2^{j_2} \geq 2^L$ is beyond the integral scale, as for a Poisson process, wavelet coefficients converge to Gaussian processes. It results that $\log_2 \widetilde{S}(j_1, j_2)$ decays with a slope $-1/2$ as a function of $j_2 - j_1$ for $j_2 > L$, as shown in Figure 5(a). If $j_1 < j_2 < L$ then (34) proves that $\widetilde{S}X(j_1, j_2)$ only depends on $j_2 - j_1$, and all curves in Figure 5(b) superimpose in this range.

Propositions 3.1 and 4.1 show that the stationary property $\widetilde{S}X(j_1, j_2) = \widetilde{S}X(j_2 - j_1)$ can be used to detect the presence of self-similarity, both deterministic and stochastic. This necessary condition is an alternative to the scaling of the q-order moments, $\mathbf{E}(|X \star \psi_j|^q) \simeq C_q 2^{j\zeta(q)}$, which is difficult to verify empirically for $q \geq 2$ or $q < 0$.

Multifractals typically become decorrelated beyond their integral scale. At scales $2^{j_2} > 2^L$, wavelet coefficients converge to Gaussian processes. The resulting normalized second order scattering thus converge to that of a Gaussian process, which decays like $2^{-j_2/2}$ as seen in Section 3.2. If $2^{j_1} > 2^L$ then $X \star \psi_{j_1}(t)$ already becomes Gaussian and second order moments thus decay like $2^{(j_1 - j_2)/2}$. Consequently, the resulting normalized second order scattering is

$$\widetilde{S}X_L(j_1, j_2) \approx \begin{cases} \widetilde{S}X(j_2 - j_1) & \text{if } j_1 < j_2 < L \\ C\, 2^{(j_1 - j_2)/2} & \text{if } L < j_1 < j_2 \end{cases} . \tag{36}$$

4.2. *Log-infinitely divisible Multifractal Random Processes.* Multiplicative cascades as introduced by Mandelbrot in [22, 23] are built as an iterative process starting at scale 2^L with the Lebesgue measure on the interval $[0, 2^L]$. The iteration consists in cutting this interval at the middle point and multiplying the mass on each interval by a log-infinitely divisible variable (iid versions are used on each interval). One then gets a random measure "living"

at scale 2^{L-1} (i.e., it is uniform on intervals of length 2^{L-1} : $[0, 2^{L-1}]$ and $[2^{L-1}, 2^L]$). The iteration is then applied recursively on each sub-interval. At the nth iteration, one gets a random measure living at scale 2^{L-n} (i.e., which is uniform on each interval of the form $[k2^{L-n}, (k+1)2^{L-n}]$). The object of interest is the weak limit ($n \to +\infty$) of this random measure. At a given point t it can be written as an infinite product $\prod_{n=1}^{+\infty} W_n$ where W_n are iid log-infinitely divisible variables.

Multifractal Random Measures, introduced in [30, 6], can be seen as stationary increments versions of these multiplicative cascades. They are obtained by infinitely decomposing each iteration step, i.e., no longer going directly from scale $s = 2^{L-n}$ to scale $s = 2^{L-n-1}$ but going from an arbitrary scale s to a scale rs where r is infinitely close to 1. Thus, they are built using an infinitely divisible random noise dP distributed in the half-plane (t, s) ($s > 0$). Using the previous notations, the noise around (t, s) can be seen as the equivalent of the infinitely divisible variable $\log_2 W_{\log s}(t)$. More precisely, if $\omega_l^{2^L}(t) = \int_{\mathcal{A}_l^{2^L}(t)} dP$ where $\mathcal{A}_l^{2^L}(t)$ is the cone in the (t, s) half-plane pointing to point $(0, s)$ and truncated for $s < l$, the MRM is defined as the weak limit

$$dM(t) = \lim_{l \to 0} e^{\omega_l^{2^L}(t)} dt. \tag{37}$$

For a rigorous definition of $\omega_l^{2^L}$ and of a Multifractal Random Measure, we refer the reader to [6]. In this section, we will study the scaling properties of scattering moments associated with $X = dM$. Thanks to the discussion in Section 2.3, one can easily show that all our results can be extended to $X(t) = M(t) = \int_0^t dM$.

One can prove that dM is a stochastic self-similar process in the sense of (31). It is multifractal in the sense that

$$\mathbf{E}(|X \star \psi_j|^q) = \mathbf{E}(|A_{2^j}|^q)\, \mathbf{E}\{|X \star \psi|^q\} \simeq C_q 2^{j\zeta(q)} \ ,$$

where $\zeta(q)$ is a non-linear function which is uniquely defined by infinitely divisible law chosen for dP. Notice that if dP is Gaussian, then dM is a log-Normal random cascade. In this case [6]:

$$\zeta(q) = (1 + \frac{\lambda^2}{2})q - \frac{\lambda^2}{2} q^2 \ . \tag{38}$$

The curvature of the concave function $\zeta(q)$ at $q = 0$ is considered as an intermittency factor in the multifractal formalism [14], which is here equal to λ^2. The bigger λ^2, the more intermittency.

The self-similarity properties of dM are mainly direct consequences of a "global" self-similarity properties of $\omega_l^{2^L}$, which is true for all L and all $s > 0$:

$$\{\omega_{sl}^{s2^L}(st)\}_t \overset{law}{=} \{\omega_l^{2^L}(t)\}_t. \tag{39}$$

But also of a stochastic self-similarity property, which is true for all L and $s < 1$

$$\{\omega_{sl}^{2^L}(su)\}_{u<T} \overset{law}{=} \{\Omega_s + \omega_l^{2^L}(u)\}_{u<T} \ , \tag{40}$$

where Ω_s is an infinitely divisible random variable independent of $\omega_l^{2^L}(u)$ such that $E(e^{q\Omega_s}) = e^{-(q-\zeta(q))\ln(s)}$. More precise results used in the proofs are stated in Appendix C. The following proposition computes the first order normalized scattering moments of dM.

PROPOSITION 4.2. *If dM is a multifractal random measure then for all $j < L$*

$$\widetilde{S}dM(j) = 1. \tag{41}$$

Proof: We first prove that for all $t \in [0, 2^L - 1]$ and all $j < L$

$$\frac{dM \star \psi_j}{\mathbf{E}(|dM \star \psi|)} \overset{law}{=} e^{\Omega_{2^{j-L}}} \xi(t2^{-j}) \ , \tag{42}$$

where Ω_s is the random variable defined in (40) and $\xi_1(t)$ is a normalized 1-dependent, stationary random process independent of $\Omega_{2^{j-L}}$ defined as:

$$\xi(t) = K^{-1}\epsilon_1(t) \tag{43}$$

where $\epsilon_1(t) = \lim_{l\to 0} \int \psi(u-t)e^{\omega_l^1(u)}du$ and $K = \mathbf{E}(|\epsilon_1(t)|)$.

Let $j < L$ and $dM_l = e^{\omega_l^{2^L}(t)}dt$. From (39), one has:

$$dM_l \star \psi_j \overset{law}{=} 2^{-j} \int e^{\omega_{2^{j-L}l}^{2^j}(u2^{j-L})} \psi(2^{-j}(t-u))du \ ,$$

and thus, by setting $s = 2^{j-L}$, from (40) and (39):

$$dM_l \star \psi_j \overset{law}{=} 2^{-j} e^{\Omega_s} \int \psi(2^{-j}(t-u)) e^{\omega_l^{2^j}(u)} du \overset{law}{=} e^{\Omega_s} \int \psi(2^{-j}t - u) e^{\omega_{l2^{-j}}^1(u)} du \ .$$

Taking the limit $l \to 0$ for a fixed j in the last equation shows that

$$dM \star \psi_j \overset{law}{=} e^{\Omega_{2^{j-L}}} \epsilon_1(t2^{-j}) \ , \tag{44}$$

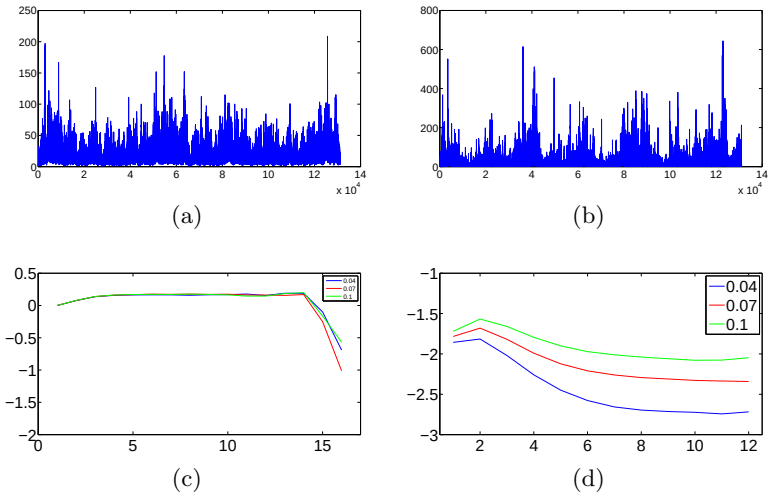

FIG 6. *(a,b) Realizations dM of log-normal Multifractal Random Measures with* $\lambda^2 = 0.04$ *and* $\lambda^2 = 0.1$. *(c)* $\log_2 \widetilde{S}dM(j_1)$ *with* $\lambda^2 = 0.04$, $\lambda^2 = 0.07$ *and* $\lambda^2 = 0.1$. *(d)* $\log_2 \widetilde{S}dM(j_1, j_1 + l)$, *for* $\lambda^2 = 0.04$, $\lambda^2 = 0.07$ *and* $\lambda^2 = 0.1$, *as a function of* l, *for* $j_1 + l < L$ *where* $2^L = 2^{13}$ *is the integral scale.*

with

$$\epsilon_T(t) = \lim_{l \to 0} \int \psi(u - t) e^{\omega_l^T(u)} du. \tag{45}$$

Normalizing $\epsilon_1(t)$ by $E(|dM \star \psi|)$ proves (42).

Since ψ has a compact support of size 1, the process $\epsilon_1(t)$ (and therefore the process $\xi_1(t)$) is a 1-dependent process, i.e., $\forall \tau > 1$, $\epsilon_1(t+\tau)$ is independent from $\{\epsilon_1(t')\}_{t' \leq t}$. Equation (41) is a direct consequence of (42) and of the fact that $E(e^{\Omega_s}) = 1$. □

The following theorem proves that normalized second order moments $\widetilde{S}(j_1, j_2)$ converge to a constant when j_1 decreases.

THEOREM 4.3. *Let* $j_1, j_2 < L$. *If* $\zeta(2) > 1$ *then* $\widetilde{S}dM(j_1, j_2)$ *depends only on* $j_1 - j_2$ *and there exists* $\tilde{K} > 0$ *such that for for each* $j_2 \leq L$

$$\lim_{j_1 \to -\infty} \widetilde{S}dM(j_1, j_2) = \tilde{K}. \tag{46}$$

The proof is in Appendix D. Let us illustrate the results of Proposition 4.2 and Theorem 4.3 in the log-normal case. Figures 6(a,b) displays two realizations of log-Normal MRM cascades for $\lambda^2 = 0.04$ and and $\lambda^2 = 0.07$,

with an integral scale $2^L = 2^{13}$. Figure 6(c) shows estimations of normalized first order scattering moments for $\lambda^2 = 0.04$, $\lambda^2 = 0.07$ and $\lambda^2 = 0.1$. As predicted by Proposition 4.2, $\log_2 \widetilde{S}dM(j_1) = 0$ for $j_1 < L = 13$. The second order scattering moments for the same values of λ^2 are displayed in Figure 6(d). As expected from Theorem 4.3, $\log_2 \widetilde{S}dM(j_1, j_2)$ only depends on $j_2 - j_1$ for $j_2 < L$. It converges to a constant $\widetilde{K}$ when $j_2 - j_1$ increases.

With a Taylor expansion, one can show that, for large $j_2 - j_1$, $\widetilde{K}$ is a linear function of λ up to some $O(\lambda^2)$ additive term. This is numerically verified by Monte Carlo simulations which shows that $\widetilde{K} \approx 0.82\,\lambda$. We see here again the correspondence between scattering coefficients and intermittency measurements. The constant 0.82 depends upon the choice of wavelet ψ.

Another important class of a stochastic self-similar processes is obtained by performing a change of variable in a Brownian motion $B(t)$ with a Multifractal Random Measure, which defines $X(t) = B(M(t))$. Such kind of processes, called Multifractal Random Walks (MRW), were introduced in [29, 6]. It can be shown [6] that a MRW process can be obtained as the limit when l goes to 0 of

$$X_l(t) = l^{\frac{2-\zeta(2)}{2}} \int_0^t e^{\omega_l^{2^L}(u)} dB(u) \ , \tag{47}$$

where $dB(u)$ is the standard Wiener noise. Accordingly, $B(M(t))$ can be considered as a stochastic volatility model, where the associated MRM,

$$\frac{dM_l(u)}{du} = l^{2-\zeta(2)} e^{2\omega_l^{2^L}(u)} \tag{48}$$

corresponds is the local stochastic variance of a Brownian motion. In that respect, such a model can account for asset price fluctuations in financial markets by mimicking the stochastic behavior of asset volatility [29, 4, 5].

Since $B(t)$ is self-similar, one can verify that $X(t)$ inherits the stochastic self-similarity of $M(t)$ and satisfies (32). In particular, one can show (see e.g. [6]) that the multifractal spectrum $\zeta(q)$ of the MRW $X(t)$ defined in (47) is related to the spectrum $\zeta_M(q)$ of the MRM $M(t)$ defined in (48) through:

$$\zeta(q) = \zeta_M(q/2) \ . \tag{49}$$

Another consequence is that scattering moments of a multifractal random walk behave similarly as the scattering moments of dM. The following theorem proves that the results of Proposition 4.2 and Theorem 4.3 also apply to Multifractal Random Walks.

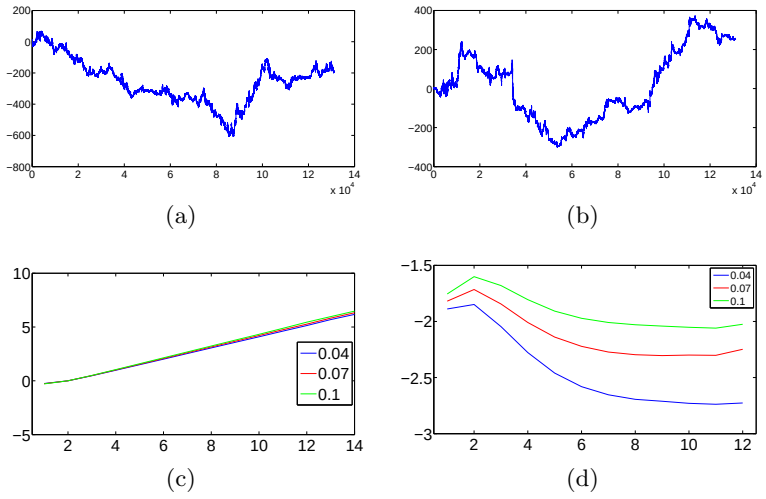

FIG 7. *(a,b) Realizations $X(t)$ of log-normal Multifractal Random Walks with $\lambda^2 = 0.04$ and $\lambda^2 = 0.1$. (c) $\log_2 \widetilde{S}X(j_1)$ with $\lambda^2 = 0.04$, $\lambda^2 = 0.07$ and $\lambda^2 = 0.1$. (d) $\log_2 \widetilde{S}X(j_1, j_1 + l)$, for $\lambda^2 = 0.04$, $\lambda^2 = 0.07$ and $\lambda^2 = 0.1$, as a function of l, for $j_1 + l < L$ where $2^L = 2^{13}$ is the integral scale.*

THEOREM 4.4. *Let $X(t)$ be a Multifractal Random Walk as defined in (47) of integral scale 2^L and scaling exponents $\zeta(q)$. For all $j < L$*

$$\tilde{S}X(j) = 2^{j\frac{3-\zeta(2)}{2}}. \tag{50}$$

If $\zeta(2) > 1$ and $j_1, j_2 < L$ then $\widetilde{S}X(j_1, j_2)$ depends only on $j_1 - j_2$, and for each $j_2 < L$:

$$\lim_{j_1 \to -\infty} \widetilde{S}X(j_1, j_2) = \tilde{K} \tag{51}$$

where the constant $\tilde{K}$ is the constant of theorem 4.3 of the Multifractal Random Measure associated with $e^{\omega_l^{2^L}}$ in (47).

The proof is very similar to the proof of Proposition 4.2 and Theorem 4.3 for Multifractal Random Measures dM so we only provide in Appendix E the main steps without entering into details.

Figure 7 shows the scattering moments of multifractal random walks $X(t) = B(dM(t))$ for a log-normal random measure dM, with $\lambda^2 = 0.04$, $\lambda^2 = 0.07$ and $\lambda^2 = 0.1$. In the log-normal case, it results from (49) that

$$\zeta(q) = (1 + 2\lambda^2)\frac{q}{2} - \frac{\lambda^2}{2}q^2 \ , \tag{52}$$

Thereby $\zeta(2) = 1$. As expected from (50), Figure 7(c) shows that $\log_2 \widetilde{S}X(j_1) = j_1(\frac{3-\zeta(2)}{2})$. As expected from (51), Figure 7(d), compared to Figure 6(d), shows that second order scattering moments satisfy $\widetilde{S}X(j_1, j_2) \approx \widetilde{S}dM(j_1, j_2)$, for the three values of λ^2. If one uses the same wavelet ψ, we check that they converge to the same constant $\widetilde{K}$ as in the MRM case which is proportional to the intermittency parameter λ: $\widetilde{K} \approx 0.82\,\lambda$ in the displayed examples.

5. Parametric Model Estimation with Scattering Moments. Section 5.1 introduces estimators of scattering moments. Section 5.2 applies the generalized method of simulated moments to scattering moments to estimate the parameters of data generating models. Section 5.3 and 5.4 analyze the scattering moments of turbulence data and financial time series to evaluate fractional Brownian, Lévy stable and multifracal cascade models. Computations are performed with a Selesnik compactly supported wavelet [35].

5.1. *Estimation of Scattering Moments.* We study scattering moment estimators introduced in [21], and compute upper bounds of their mean-square error. A scattering moment $\overline{S}X(j_1, ..., j_m) = \mathbf{E}(|\,|X\star\psi_{j_1}|\star...|\star\psi_{j_m}|)$ is estimated by replacing the expected value by a time averaging at a scale 2^M. It is calculated with a time window $\phi_M(t) = 2^{-M}\phi(2^{-M}t)$ with $\int \phi(t)\,dt = 1$. For any $(j_1, ..., j_m) \in \mathbb{Z}^m$ with $j_k \leq M$, the estimator is

$$\widehat{S}X(j_1, ..., j_m) = |\,|X \star \psi_{j_1}| \star ...| \star \psi_{j_m}| \star \phi_M(t_0)\ . \tag{53}$$

where t_0 is typically in the middle of the domain where $X(t)$ is known. Since $\int \phi_M(t)\,dt = 1$, this estimator is unbiased $\mathbf{E}(\widehat{S}X(j_1, ..., j_m)) = \overline{S}X(j_1, ..., j_m)$.

THEOREM 5.1. *Suppose that the Fourier transform $\Phi(\omega)$ of ϕ satisfies*

$$|\Phi(\omega)|^2 \leq \frac{1}{2}\sum_{j=1}^{\infty}\left(|\Psi(2^j\omega)|^2 + |\Psi(-2^j\omega)|^2\right) \text{ with } \Phi(0) = 1\ . \tag{54}$$

If X has stationary increment and $\mathbf{E}(|X \star \psi_{j_1}|^2) < \infty$ then

(55)

$$\forall\, j_1 \leq M\ ,\ \mathbf{E}(|\widehat{S}X(j_1)|^2) + \sum_{m=2}^{\infty}\sum_{-\infty < j_2, ..., j_m \leq M} \mathbf{E}(|\widehat{S}X(j_1, ..., j_m)|^2) \leq \mathbf{E}(|X\star\psi_{j_1}|^2)\ .$$

Proof: The Fourier transform $\Psi(\omega)$ of ψ satisfies the Littlewood-Paley condition (3)

$$\sum_{j=-\infty}^{\infty}\left(|\Psi(2^j\omega)|^2 + |\Psi(-2^j\omega)|^2\right) = 2\ .$$

It results that

$$|\Phi(2^M\omega)|^2 + \frac{1}{2}\sum_{j=-\infty}^{M}\left(|\Psi(2^j\omega)|^2 + |\Psi(-2^j\omega)|^2\right) \leq 1 \ . \tag{56}$$

Let Y be a stationary process with $E(|Y(t)|^2) < \infty$, and $P_Y(\omega)$ be its power spectrum. Multiplying (56) by $P_Y(\omega)$ and integrating in ω gives

$$\mathbf{E}(|Y \star \phi_M|^2) + \sum_{j=-\infty}^{M} \mathbf{E}(|Y \star \psi_j|^2) \leq \mathbf{E}(|Y|^2) \ . \tag{57}$$

Let us prove by induction that for any $q \geq 2$

$$\begin{aligned}\mathbf{E}(|\widehat{S}X(j_1)|^2) + \sum_{m=2}^{q-1}\sum_{-\infty<j_2,...,j_m\leq M} \mathbf{E}(|\widehat{S}X(j_1,...,j_m)|^2) + \\ \sum_{-\infty<j_2,...,j_q\leq M} \mathbf{E}(|\ |X\star\psi_{j_1}|\star...|\star\psi_{j_q}|^2) \quad \leq \quad \mathbf{E}(|X\star\psi_{j_1}|^2) \ .\end{aligned} \tag{58}$$

Applying (57) to $Y = |X \star \psi_{j_1}|$ proves that (58) for $q = 2$. If (58) is valid for q we prove it for $q+1$ by applying (57) to each $Y = |\ |X \star \psi_{j_1}| \star ...| \star \psi_{j_q}|$. Taking the limit of (57) as q goes to ∞ proves (55). □

If ψ has a compact support then one can verify that there exists ϕ having the same support as ψ, whose Fourier transform satisfies (54), with an equality on the Fourier transform modulus. Since $\mathbf{E}(\widehat{S}X(j_1,...,j_m)) = \overline{S}X(j_1,...,j_m)$, the mean-square estimation error at the scale 2^{j_1} is

$$\epsilon(j_1) \stackrel{def}{=} \mathbf{E}(|\widehat{S}X(j_1) - \overline{S}X(j_1)|^2) + \sum_{m=2}^{\infty}\sum_{-\infty<j_2,...,j_m\leq M} \mathbf{E}(|\widehat{S}X(j_1,...,j_m) - \overline{S}X(j_1,...,j_m)|^2) \ .$$

It results from (55) that

$$\epsilon(j_1) \leq \sigma^2(|X\star\psi_{j_1}|) - \sum_{m=2}^{\infty}\sum_{-\infty<j_2,...,j_m\leq M} |\overline{S}X(j_1,...,j_m)|^2 \ . \tag{59}$$

When j_1 is close to M then $|X \star \psi_{j_1}(t)|$ decorrelates slowly relatively to the averaging window scale 2^M so $\epsilon(j_1)$ is large, but it is bounded by $\sigma^2(|X\star\psi_{j_1}|)$.

Large variance estimators $\widehat{S}X(j_1,...,j_m)$ are eliminated by keeping only small scales $j_k \leq J$ for $1 \leq k \leq m$, with $M - J$ sufficiently large. For most classes of random processes, including fractional Brownian motions and multifractal random walks, we observe numerically that $\epsilon(j_1)$ converges

to zero as the averaging scale 2^M goes to ∞. Equation (59) proves that it is the case if for all j_1

$$\sigma^2(|X \star \psi_{j_1}|) = \sum_{m=2}^{\infty} \sum_{-\infty < j_2,...,j_m \leq \infty} |\overline{S}X(j_1, ..., j_m)|^2 \ .$$

This energy conservation has been conjectured for large classes of processes in [21], but it is not proved.

For n independent realizations $\{X_k(t)\}_{1 \leq k \leq n}$, we compute an averaged scattering estimator

$$\widehat{S}X = n^{-1} \sum_{k=1}^{n} \widehat{S}X_k \ . \tag{60}$$

Its variance is thus reduced by n^{-1}. When n goes to ∞, the central limit theorem proves that $\widehat{S}X - \overline{S}X$ converges to a zero-mean normal distribution whose variance goes to 0.

5.2. *Generalized Method of Simulated Moments.* The generalized method of simulated moments computes parameter estimators for data generative models, from arbitrary families of moments. It is applied to scattering moments and results are numerically evaluated on synthetic data generated from multifractal random measures.

Suppose that $\{X_k\}_{1 \leq k \leq n}$ are n independent realizations of $X(t)$. If the $X_k(t)$ are realizations of a parametric model Y_θ then $\widehat{S}X_k$ is an unbiased estimator of $\overline{S}Y_\theta$ so

$$m(\theta) = \mathbf{E}(\widehat{S}X_k) - \overline{S}Y_\theta = 0 \ .$$

The generalized method of moments estimates this moment condition by an empirical average defined by

$$\widehat{m}(\theta) = n^{-1} \sum_{k=1}^{n} \widehat{S}X_k - \overline{S}Y_\theta = \widehat{S}X - \overline{S}Y_\theta. \tag{61}$$

When n goes to ∞ the central limit theorem proves that $\widehat{m}(\theta)$ converges to a normal distribution. The generalized method of moments finds the parameter $\widehat{\theta}$ such that:

$$\widehat{\theta} = \operatorname*{argmin}_{\theta} \widehat{m}(\theta)\, W \widehat{m}(\theta)^T \tag{62}$$

for appropriate matrices W. Setting $W = Id$ gives

$$\widehat{\theta}_1 = \operatorname*{argmin}_{\theta} \|\widehat{S}X - \overline{S}Y_\theta\|^2 \ . \tag{63}$$

The two-step generalized method of moment updates the first estimator $\widehat{\theta}_1$ by setting $W = \widehat{W}_{\hat{\theta}_1}$, where $\widehat{W}_\theta$ is the inverse of the empirical covariance

$$\widehat{W}_\theta = \Big(n^{-1}\sum_{k=1}^{n}(\widehat{S}X_k - \overline{S}Y_\theta)(\widehat{S}X_k - \overline{S}Y_\theta)^T\Big)^{-1} \ . \tag{64}$$

It computes

$$\widehat{\theta} = \operatorname*{argmin}_{\theta} \widehat{m}(\theta)\, \widehat{W}_{\hat{\theta}_1}\, \widehat{m}(\theta)^T \ . \tag{65}$$

Since we can not compute $\overline{S}Y_\theta$ analytically, according to the simulated method of moments [25], $\overline{S}Y_\theta$ is replaced in (61) and (64) by an estimator $\widehat{S}Y_\theta$ calculated with a Monte Carlo integration. This estimator is computed with $n' \gg n$ realizations which are adjusted in order to yield a negligible mean-square error $\mathbf{E}(\|\widehat{S}Y_\theta - \overline{S}Y_\theta\|^2)$.

In numerical computations, the dimension p of the vector of scattering moments is limited by keeping only the finest scale moments. This reduces the variance of the estimator (65). The estimation is often over-identified in the sense that the dimension p is larger than the dimension d of the parameter vector θ. A p-value is computed for the null hypothesis which supposes that the parametrized model is valid. The J-test [13] is a chi-squared goodness of fit test, which we normalize by the $p - d$ degrees of freedom:

$$\chi^2_{\text{red}} = (p-d)^{-1}\, n\, \widehat{m}(\widehat{\theta})\, \widehat{W}_{\hat{\theta}}\, \widehat{m}(\widehat{\theta})^T \ . \tag{66}$$

Under the null hypothesis, $(p-d)\chi^2_{\text{red}}$ asymptotically follows a chi-squared distribution with $p - d$ degrees of freedom.

In all numerical computations, we limit ourself to first and second order scattering coefficients, which carry most of the energy of the scattering vector. Computations are performed from discretized data sequences whose sampling interval is normalized to 1. Wavelet coefficients can thus be computed at scales $2^j > 1$. However, we often need to eliminate the finest scale coefficients $j_1 \leq J_0$ to remove high frequency errors due to aliasing, discretization or to some data smoothing. As explained in Section 5.1 we only keep coefficient below a maximum scale $j_1 \leq J$ and $j_2 \leq J$, to eliminate the largest variance scattering estimators. Second order coefficients $\overline{S}X(j_1, j_2)$

for $j_2 \leq j_1$ are also eliminated, because they have a negligible amplitude which carries little information on X, as explained Section 2.2. The resulting scattering vector is thus

$$\overline{S}X = \left(\overline{S}X(j_1)\,\overline{S}X(j_1,j_2)\right)_{J_0<j_1\leq J, j_1<j_2\leq J} .$$

It has $J - J_0$ first order scattering moments and $(J - J_0 - 1)(J - J_0)/2$ second order moments. They are estimated by the scattering estimator $\widehat{S}X$ computed in (53) and in (60) for multiple realizations.

The dimension of the scattering vector $\widehat{S}X$ is $p = (J - J_0 + 1)(J - J_0)/2$. As J increases, the dimension p of the scattering vector increases but the variance of this vector also increases. The method of generalized moments requires that $\widehat{W_\theta}$ in (64) gives an accurate estimation of the inverse of the covariance matrix of $\overline{S}Y_\theta$ whose dimension is p. For large p it is sufficient to have a number n of independent realizations such that $n \simeq p \log_2 p$ [37]. This condition limits the maximum value of J.

The Generalized Method of Simulated Moments can still be applied if we observe a single realization $X(t)$, when sufficiently far away wavelet coefficients become independent. This is indeed sufficient to guarantee that $\widehat{S}X$ becomes asymptotically normal. Let us consider scattering vector estimators computed at intervals Δ:

(67)
$$\widehat{S}X_k(j_1) = |X \star \psi_{j_1}| \star \phi_M(k\Delta) \;\text{ and }\; \widehat{S}_k X(j_1,j_2) = ||X \star \psi_{j_1}| \star \psi_{j_2}| \star \phi_M(k\Delta) .$$

The goodness of fit J-test supposes that the variable (66) follows a chi-squared distribution with $p - d$ degrees of freedom, which requires that the estimators $\widehat{S}X_k$ are independent for different k. This is the case if wavelet coefficients for different k are computed from independent values of X. If X has an integral scale T, as in multifractal cascades, then increments are independent at distances larger than T. One can thus set $\Delta = 2T$. Other processes, such as fractional Brownian motions have no integral scales but their wavelet coefficients become nearly independent at distances much larger than the scale. Nearly independent estimators are thus obtained if $\Delta \gg 2^M$.

The situation is different if we only compute the parameter estimator $\widehat{\theta}$ with (65). Its consistency requires that $\widehat{S}X$ converges to a normal distribution, but we can estimate its covariance up to an unknown multiplicative factor. It is then unnecessary to estimate the decorrelation properties of wavelet coefficients, and one can set $\Delta = 1$ to be the process sampling interval. Indeed, although the $\widehat{S}X_k$ are correlated, their average $\widehat{S}X$ still becomes asymptotically normal if there exists a Δ for which the $\widehat{S}X_k$ become independent. Reducing Δ is equivalent to averaging several correlated estimators,

where each one is computed from independent realizations and converges to a normal distribution as N increases. It introduces a multiplicative factor in the covariance estimation, which does not affect estimator $\widehat{\theta}$ in (65).

The properties of the Scattering Method of Moments are illustrated on the estimation of the intermittency parameter $\theta = \lambda^2$ for multifractal random measures. Section 4.2 proves that normalized second order scattering moments converge to a constant $\widetilde{K}$ which is proportional to λ, which shows that the intermittence λ^2 is characterized by first and second order scattering moments. However, the information is not just carried by this asymptotic value, which is why all scattering moments are used for the estimation. The scattering estimation is compared with two estimators dedicated to this particular estimation problem [5].

Scattering moment estimators are computed from n independent realizations of size 2^{11} of a multifractal random measure having an integral scale $T = 2^{10}$. The total number of data points is $N = n \cdot 2^{11}$, and we set $J_0 = 0$. For different values of $N = n \cdot 2^{11}$, we report in Table 1 the value of J which minimizes the mean squared error $\mathbf{E}(|\widehat{\theta} - \theta|^2)$, estimated with Monte Carlo simulations. We also give the average value of the reduced χ^2_{red} test in (66) and the model p-value. For small values of n, the covariance of $\widehat{S}X$ is computed up to a multiplicative constant, from correlated scattering coefficients calculated within each realization with $\Delta = 1$ in (67). It leads to a good estimation of $\widehat{\theta}$ but the model p-value can not be estimated.

The intermittency parameter of multifractal random measures can also be estimated directly from wavelet coefficients. Section 4.2 explains that the scaling exponent of wavelet moments of order q is $\zeta(q) = \left(\frac{1}{2} + \lambda^2\right) q - \frac{\lambda^2}{2} q^2$. It results that $\lambda^2 = 2\zeta(1) - \zeta(2)$. The intermittency parameter can thus be estimated with a linear regression on the estimated first and second order moments of wavelet coefficients at scales $2^j < 2^L$:

$$2 \log_2 \mathbf{E}(|X \star \psi_j|^2) - \log_2 \mathbf{E}(|X \star \psi_j|)^2 \approx j(\zeta(2) - 2\zeta(1)) + C \ . \tag{68}$$

The wavelet moments $\mathbf{E}(|X \star \psi_j|^2)$ and $\mathbf{E}(|X \star \psi_j|)$ are estimated with empirical averages of $|X \star \psi_j|$ and $|X \star \psi_j|^2$, calculated from the N data samples. An improved estimator has been introduced in [4, 5] with a regression on the covariance of the log of the multifractal random measure. One can indeed prove that

$$\mathrm{Cov}\left(\log |X \star \psi_j(t)|, \log |X \star \psi_j(t+l)|\right) \simeq -\lambda^2 \ln\left(\frac{l}{2^L}\right) + o\left(\frac{j}{l}\right) \ , \tag{69}$$

which leads to lower variance estimations.

TABLE 1

Estimation of λ^2 for a multifractal random measure. The table gives the mean and the standard deviation of estimators computed with a wavelet moment regressions (68), a log covariance regression (69), and the method of simulated scattering moments, for several values of λ^2 and several sample sizes $N = n \cdot 2^{11}$.

λ^2	N	$\widehat{\theta}$ Wavelet	$\widehat{\theta}$ log-cov	$\widehat{\theta}$ Scattering	J	χ^2_{red}	p-value
0.02	10^6	$0.025 \pm 2 \cdot 10^{-3}$	$0.02 \pm 2 \cdot 10^{-4}$	$0.02 \pm 2 \cdot 10^{-4}$	7	1.1 ± 0.3	0.7 ± 0.3
0.05	10^6	$0.055 \pm 2 \cdot 10^{-3}$	$0.05 \pm 6 \cdot 10^{-4}$	$0.05 \pm 3 \cdot 10^{-4}$	6	0.8 ± 0.3	0.5 ± 0.3
0.1	10^6	$0.105 \pm 4 \cdot 10^{-3}$	0.1 ± 10^{-3}	0.1 ± 10^{-3}	5	0.8 ± 0.5	0.5 ± 0.3
0.1	10^5	0.109 ± 10^{-2}	$0.1 \pm 3 \cdot 10^{-3}$	$0.1 \pm 2 \cdot 10^{-3}$	5	0.7 ± 0.3	0.3 ± 0.3
0.1	10^4	$0.12 \pm 3 \cdot 10^{-2}$	$0.1 \pm 1.3 \cdot 10^{-2}$	$0.1 \pm 9 \cdot 10^{-3}$	5	N/A	N/A

Table 1 shows that the scattering moment estimation of λ^2 has a smaller variance than the regression of first and second order wavelet moments. This is due to the low variance of the scattering estimators which are computed with non-expansive operators. It gives comparable results with the log-covariance estimator, which was optimized for this problem [5]. The J-test validates the multifractal model, since we obtain a normalized J-test with mean and standard deviation close to $1 \pm \sqrt{\frac{2}{p-1}}$, corresponding to mean and standard deviation of a chi-squared distribution with $p-1$ degrees of freedom. The resulting p-values for rejecting the true model are of the order of 0.5. As expected, reducing the maximum scattering scale J improves the estimation of λ^2 for high intermittences. It removes large variance coefficients. However, numerical experiments confirm that the generalized method of moments is robust to the choice of J, because the inverse covariance $\widehat{W}$ in (65) reduces the impact of high variance coefficients.

5.3. *Turbulence Energy Dissipation.* Turbulent regimes that appear in a wide variety of experimental situations, are characterized by random fluctuations over a wide range of time and space scales. The main physical picture behind this complexity was introduced by Richardson and Kolmogorov [32, 18]: the fluid receives kinetic energy at large scales and dissipates this energy at small scales where fluctuations are well known to be of intermittent nature. The overall range of scales between injection and dissipation is called the inertial range and only depends on the Reynolds number. Making a theory of this "energy cascade" across the inertial range remains of the famous challenges in classical physics [12].

Normalized scattering moments are computed over dissipative measurements of a turbulent gaz, to analyze their self-similarity and intermittency

properties. The precision of fractional Brownian motion, Lévy stable and multifractal random measure models are evaluated with a J-test resulting from scattering moments. This study does not pretend evaluating general turbulence physical models. However, it shows that one can have confident model evaluations from data sets, despite intermittency phenomena.

The data we used have been recorded by the group of B. Castaing in Grenoble in a low temperature gazeous Helium jet which Taylor scale based Reynolds number is $R_\lambda = 703$ [9]. A single probe provides measures of velocity temporal variations at a fixed space location that involve both Lagrangian and Eulerian fluctuations. If one supposes the flow homogeneous and isotropic, the local energy dissipation rate at a given time and location is given by:

$$X(\vec{x},t) = 15\nu \left(\frac{\partial v_{||}}{\partial x}\right)^2 ,$$

where $v_{||}(\vec{x},t)$ is the stream-wise component of velocity and ν stands for the kinematic viscosity constant. The Taylor frozen-flow hypothesis [12] assumes that the stream-wise mean velocity V is very large so one can neglect temporal lagrangian fluctuations. It results that the temporal variations of a field $X(\vec{x}_0,t)$, at some fixed location $\vec{x_0}$, correspond to the spatial fluctuations advected by the mean flow, i.e., $X(\vec{x}_0,t) = X(\vec{x_0} - \vec{V}t, 0)$. This hypothesis implies that a surrogate of the dissipation field can be obtained from the temporal evolution of longitudinal velocity field as:

$$X(t) \simeq \left(\frac{\partial v_{||}}{\partial t}\right)^2_{;} . \tag{70}$$

Figure 8-(a) shows a sample of the dissipation field $X(t)$ as a function of time, estimated from the experimental velocity records. The Kolmogorov (dissipative) scale η is observed at approximately 2^2 sample points, whereas the integral scale is approximately $2^L = 2^{11}$ sample points.

Estimators of normalized scattering moments are computed as ratio of the first and second order scattering estimators $\widehat{S}X(j_1)$ and $\widehat{S}X(j_1,j_2)$ in (53). First order scattering coefficients are renormalized at the finest scale defined by $j_1 = 2$. The normalized moments $\widetilde{S}X(j_1) = \overline{S}X(j_1)/\overline{S}X(2)$ and $\widetilde{S}X(j_1,j_2) = \overline{S}X(j_1,j_2)/\overline{S}X(j_1)$ are thus estimated by

$$\widehat{\widetilde{S}}X(j_1) = \frac{\widehat{S}X(j_1)}{\widehat{S}X(2)} \quad \text{and} \quad \widehat{\widetilde{S}}X(j_1,j_2) = \frac{\widehat{S}X(j_1,j_2)}{\widehat{S}X(j_1)} . \tag{71}$$

These estimators are biased because $\widehat{S}X(j_1,j_2)$ and $\widehat{S}X(j_1)$ are not independent but the bias tends to zero as the variance of scattering moment estimators goes to zero.

First order scattering coefficients are displayed in Figure 8(b). In the inertial range $2^1 = 2^{J_0} < 2^{j_1} \leq 2^L = 2^{11}$ the scaling law of the exponents is $\widetilde{S}X(j_1) \simeq 2^{-0.25 j_1}$. If $2^{j_1} \geq 2^L$ then $\widetilde{S}X(j_1) \simeq 2^{-j_1/2}$ because the low frequencies of a turbulence flow becomes Gaussian and independent beyond the integral scale.

Figure 8(c) gives normalized second order coefficients $\log_2 \widetilde{S}X(j_1, j_1 + l)$ as a function of l, for different j_1. For $j_2 = j_1 + l > L$, the slopes increase up to $-1/2$ because beyond the integral scale, wavelet coefficients converge to Gaussian random processes. Below the integral scale, $j_2 = j_1 + l < L - 1$ Figure 8(d) shows that the curves $\log_2 \widetilde{S}X(j_1, j_1 + l)$ with error bars giving the standard deviations of each estimated values. In this inertial range, the average slope of all curves is -0.2. This slope is very different from the $-1/2$ decay of Gaussian processes, which indicate the presence of intermittent phenomena. Although these curves are similar, one can observe that they differ significantly compared to the error bars, which indicates the the self-similarity of turbulence data is violated.

This non-self-similarity is likely to originate from the fact that Taylor hypothesis does not rigorously hold. A single probe provides measures of velocity temporal variations at a fixed space location, which involve both lagrangian and eulerian fluctuations. This point has already been observed in ref. [11] and discussed in details by Castaing in ref. [8] in order to explain the behavior of the correlation functions of velocity amplitude variations. This example illustrates the ability of scattering moments to finely evaluate self-similarity properties.

Let us study how 3 simple models can be compared to fit the experimental data. The square of a Fractional Gaussian noise can be considered as a first model of turbulent dissipative energy. It is calculated as the square of the "derivative" of a fractional Brownian motion parametrized by the index $\theta = H$. A second model is computed as the square of the derivative of an α-stable Lévy process, which is parametrized by $\theta = \alpha$. The third class of models we consider is standard in the literature for the dissipation field [12, 26]: they are random cascade models as represented by log-normal the multifractal random measures, parametrized by the intermittency parameter $\theta = \lambda^2$. These models are evaluated from scattering moments computed in the inertial range. Setting $J_0 = 1$ eliminates coefficients below the diffusion scale. We have $N = 4 \cdot 10^6$ data samples, divided into 4 realizations. Within each realization, since the integral scale is $T = 2 \cdot 10^3$, samples are independent at a distance larger than T. The maximum scale is set to $2^J = 2^8$ but its modification has a marginal impact on the estimation. The size of the resulting scattering vector is $p = (J - J_0 + 1)(J - J_0)/2 = 28$.

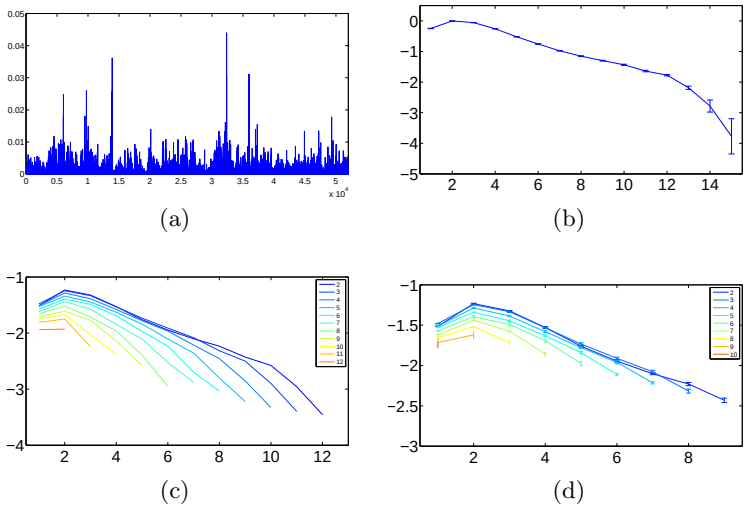

FIG 8. *(a) Realization of dissipation* $X(t) = \left(\frac{\partial v}{\partial t}\right)^2$ *in a turbulent flow. (b) Estimation* $\log_2 \widehat{\widetilde{S}} X(j_1)$ *as a function of* j_1*, calculated from* 4 *realizations of* 2^{19} *samples each. (c)* $\log_2 \widehat{\widetilde{S}} X(j_1, j_1 + l)$ *as a function of* l*, for* $2 \leq j_1 \leq 12$*. (d)* $\log_2 \widehat{\widetilde{S}} X(j_1, j_1 + l)$ *in the inertial range* $j_1 + l < L - 1 = 10$*. We plot the confidence intervals corresponding to the standard deviation of the estimated* $\log_2 \widehat{\widetilde{S}} X(j_1, j_1 + l)$*.*

TABLE 2
Parameter estimation for the turbulence data in Figure 8(a), calculated from Fractional Brownian Noise measures (FBN), Lévy stable measures (LS), and Multifractal Random Measures (MRM).

	FBN	LS	MRM
$\hat{\theta}$	$H = 0.9$	$\alpha = 1.98$	$\lambda^2 = 0.09$
χ^2_{red}	28	29	29
p-value	$< 10^{-6}$	$< 10^{-6}$	$< 10^{-6}$

For each model family, Table 2 estimates an optimal parameter $\widehat{\theta}$ with the method of simulated scattering moments presented in Section 5.2. Table 2 also gives the value of the χ^2_{red} goodness of fit test in (66), together with its p-value. All models are rejected with very high confidence. For nearly the same number of data values, with integral scales of same size, Table 1 gave much higher p-values for valid multifractal random measure models of same intermittency.

The main source of errors of each model clearly appears by analyzing the normalized scattering moments in Figure 8. Fractional Brownian motions have first order coefficients which can mimic the decay of the first order scattering coefficients but their second order coefficients decay with a slope of $-1/2$ as opposed to -0.2. Levy stable processes have first and second order scattering coefficients which decay with a slope of $\alpha^{-1} - 1$. To match the slopes in Figure 8(c,d), respectively equal to -0.25 and -0.2, we would need that $\alpha \approx 1.2$. However, the resulting second order coefficients in Figure 4 have a much larger amplitude than in Figure 8(d). Table 1 estimation $\alpha = 1.98$ is close to 2 and thus has a much lower intermittency, but it does not match the decay of first and second order coefficients. Multifractal random measure model misfit comes from their first order coefficients which remain constant whereas turbulence data coefficients decay with a slope close to -0.2.

5.4. *Financial Time-Series Analysis.* Since Mandelbrot's pioneering work on the fluctuations of cotton price in the early sixties, it is well known that market price variations have some form self-similar properties, which are poorly described by a Brownian motion [24]. Extreme events are more probable than in a Gaussian world and variance fluctuations are well known to be of intermittent and correlated nature. In the following, we analyze the normalized scattering moments of two financial time series: high-frequency Euro-Bund data and intraday S&P 100 index data. We also evaluate fractional Brownian motion, Levy stable and multifractal random walk models on these data.

Euro-Bund is one of the most actively traded financial asset in the world. It corresponds to a future contract on an interest rate of the Euro-zone and it is traded on the Eurex electronic market (in Germany). The typical number of trades is around 40.000 per day and in this study we have used 800 trading days going from May 2009 May to September 2012. Each trade occurs at a given price, whose logarithm is notated $X(t)$.

Every single day, the logarithmic returns of the price (i.e., the increments of $X(t)$) are computed on rolling 10 second intervals, after preprocessing the

microstructure noise using the technique advocated in [33]. Each day corresponds to 9 hours of trading and hence 3240 increments. Intraday financial data are subject to strong seasonal intraday effects. Thus, for instance, the variance is systematically larger at opening and settlement time than at lunch time. These effects are removed with a standard "deseasonalizing" algorithm which normalizes the returns by the square root of the intraday seasonal variance. It is obtained by estimating the variance every 5mn during the day, with an averaging across days.

Figure 9(a) shows the resulting "deseasonalized" Euro-Bund log-price $X(t)$ for a particular day. Figure 9(b) shows $\log_2 \widetilde{S}X(j_1)$ as a function of j_1. The slope is 0.48. Scattering coefficients are computed at scales smaller than a day, and are averaged in time within a day and across days. Figure 9(c) shows $\log_2 \widetilde{\widetilde{S}}X(j_1, j_1 + l)$ as a function of l for different values of j_1 up to largest available scales. The decay of second order coefficients is $\log_2 \widetilde{\widetilde{S}}X(j_1, j_1 + l) \sim -0.2l$ for all $j_2 = j_1 + l$. Contrarily to turbulence data, we do not see an integrable scale, beyond which second order coefficients would have a fast decay of $-0.5l$. This is not surprising since the integral scale is known to be larger than few months [5]. Figure 9(d) gives intra-day second order coefficients $j_2 = j_1 + l < 9$. The variance of $\widetilde{\widetilde{S}}X(j_1, j_1 + l)$ is indicated with vertical error bars. Observe that the variations of $\log_2 \widetilde{\widetilde{S}}X(j_1, j_1 + l)$ has small variations as a function of j_1, which is a strong indication of self-similarity.

The same scattering computations are performed on S&P 100 index, sampled every 5 minutes from April 8th 1997 to December 17th 2001. It opens 6.5 hours a day, from 9:30am to 4:00pm (i.e., 78 samples every day). The S&P 100 Index is a stock market index of United States stocks maintained by Standard & Poor's. It is a subset of the S&P 500 index. The same preprocessing is applied as for the Euro-Bund data, but microstructure noise is non-existent because the data is sampled at a lower frequency. We used high and low values on each 5mn interval to compute an estimation on the 5mn variance. All days are concatenated and the overnight period has been preprocessed with the same deseasonalizing algorithm as intraday periods.

Figure 10(a) shows the deseasonalized log price $X(t)$. Panel (b) displays the estimated first order scattering moments. Each trading day has $78 \simeq 2^6$ samples. The deseasonalizing algorithm eliminates opening and closing artifacts and $\log_2 \widetilde{S}X(j_1)$ remains regular for j_1 close to 6. Figure 10(c) shows the estimated second order moments. For $j_2 = j_1 + l = 6$ the coefficient $\log_2 \widetilde{S}X(j_1, j_1 + l)$ is higher than expected, relatively to other coefficients, which means a higher level of intermittency. As explained in Section 5.1,

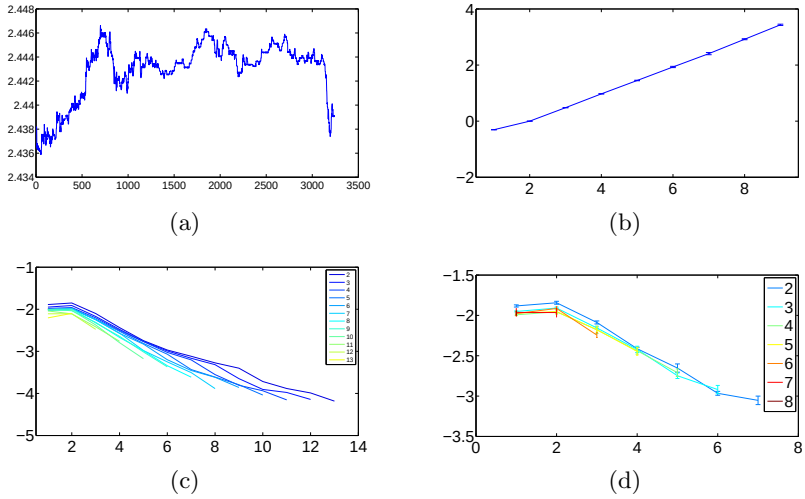

FIG 9. *(a) One day of the deseasonalized Euro-Bund log-price $X(t)$. (b) Estimated $\log_2 \widehat{\widetilde{S}}X(j_1)$. (c) Estimated $\log_2 \widehat{\widetilde{S}}X(j_1, j_1 + l)$. (d) Estimated $\log_2 \widehat{\widetilde{S}}X(j_1, j_1 + l)$ for $j_1 + l < 9$.*

TABLE 3

The left and right parts of the table correspond to Euro-Bund and S&P 100 time series. The first row gives the estimated parameter value $\hat{\theta}$ for Fractional Brownian Motion(FBM), Levy stable processes (LS) and Multifractal Random Walks (MRW).

S	Euro-Bund			S&P		
	FBM	LS	MRW	FBM	LS	MRW
$\hat{\theta}$	$H = 0.5$	$\alpha = 1.95$	$\lambda^2 = 0.03$	$H = 0.5$	$\alpha = 1.8$	$\lambda^2 = 0.08$
χ^2_{red}	29	26	23	17	16	10

large scales 2^{j_1} and 2^{j_2} have coefficients of higher variance.

The method of simulated scattering moments is applied to estimate model parameters for these financial data. We consider fractional Brownian motions models with $\theta = H$, Levy stable processes parametrized with $\theta = \alpha$ and multifractal random walks with $\theta = \lambda^2$. For each model family, Table 2 estimates an optimal parameter $\widehat{\theta}$ from first and second order scattering coefficients. They are computed from a total of $N = 3 \cdot 10^6$ samples for the Euro-Bund and $N = 10^5$ samples for S&P 100. The maximum scale 2^J is adjusted to $J = 8$ for the Euro-Bund and $J = 6$ for the S&P 100 dataset. We set $J_0 = 1$ to eliminate discretization effects in both cases. For fractional Brownian motions, the estimated parameter $\widehat{\theta} = H = 0.5$ corresponds to a Brownian motion. Brownian motion models explain the power-spectrum

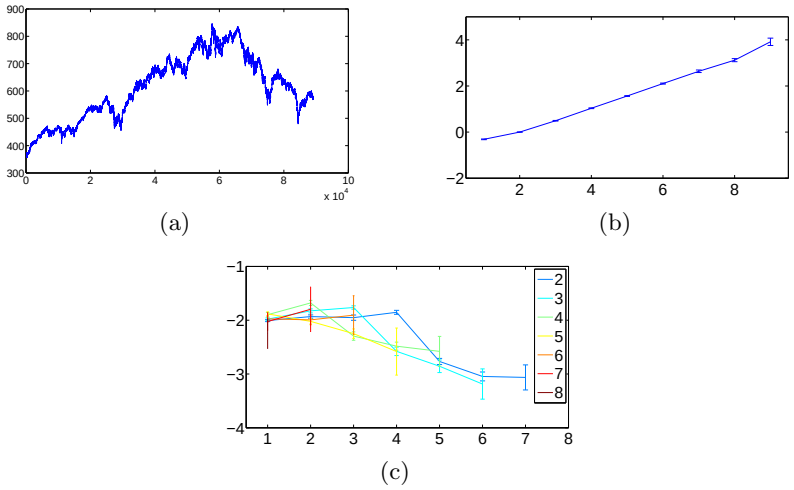

FIG 10. *(a) Three years of the deseasonalized S&P 100 index log-price $X(t)$. (b) Estimated $\log_2 \widehat{\widehat{S}}X(j_1)$. (c) Estimated $\log_2 \widehat{\widehat{S}}X(j_1, j_1 + l)$ for $j_1 + l < 9$.*

decay of these processes but are known not to be appropriate because they do not take into account the intermittency behavior of financial markets. This appears in the second order scattering coefficients of Figure 9(d) and 10(c), which have a much slower decay than Brownian motions. The Levy-stable parameters α in Table 2 are close to 2 (order 2 moment of financial time-series are finite). Estimated models of multifractal random walks also show the existence of intermittency which is larger for the S&P 100 data set than for the Euro-Bund data.

For each model, Table 2 gives the value of the J-test variable χ^2_{red} computed with (66). Multifractal random walks have the lowest value χ^2_{red} for the Euro-Bund and S&P 100 data, which means that these models better fit the data. However, one can not compute a p-value because the empirical covariance matrix is computed from correlated scattering estimators $\widehat{S}X_k$ in (67). Because the integral scale is too large, one cannot fix an interval Δ providing independent scattering values. One can still verify numerically that the empirical covariance of $\widehat{S}X_k$ converges up to a multiplicative factor, by computing the variations of empirical covariance $\widehat{W}^{-1}_{\theta,\Delta}$ as a function of Δ, and by verifying that

$$\|\widehat{W}^{-1}_{\theta,\Delta} - \eta \widehat{W}^{-1}_{\theta,\Delta'}\| \ll \|\widehat{W}^{-1}_{\theta,\Delta}\|$$

for some $\eta(\Delta, \Delta') \in \mathbb{R}^+$. The ratio between both terms is of the order of 0.05 for the Euro-Bund data.

Appendix A: Proof of Theorem 2.2

Let us first prove (13). If X is a Poisson process then $X \star \psi_j = 2^j dX \star \bar{\psi}_j$ where $\bar{\psi}(t) = \int_0^t \psi(u)\,du$ has a support in $[-1/2, 1/2]$, and $\bar{\psi}_j(t) = 2^{-j}\bar{\psi}(2^{-j}t)$. Since $dX(t) = \sum_i \delta(t - \tau_i)$ we get $X \star \psi_j(t) = 2^j \sum_i \bar{\psi}_j(t - \tau_i)$. We write

$$|X \star \psi_j(t)| = 2^j |dX \star \bar{\psi}_j|(t) = 2^j (dX \star |\bar{\psi}_j|(t) + e_j(t)) \ . \tag{72}$$

The first term satisfies

$$\mathbf{E}(dX \star |\bar{\psi}_j|) = \lambda \|\bar{\psi}_j\|_1 = \lambda \|\bar{\psi}\|_1 \ . \tag{73}$$

Let us show that $\mathbf{E}(|e_j(t)|) = O(\lambda^2 2^j)$. Let $N_j(t)$ be the number of events counted by $X(t)$ in the interval $[t - 2^j, t + 2^j)$. We decompose

$$\mathbf{E}(|e_j(t)|) = \mathbf{E}\left(|e_j(t)| \Big/ N_j(t) \leq 1\right) \mathrm{Prob}(N_j(t) \leq 1) + \mathbf{E}\left(|e_j(t)| \Big/ N_j(t) > 1\right) \mathrm{Prob}(N_j(t) > 1) \ .$$

If $N_j(t) \leq 1$, since $\bar{\psi}_j$ has support $[-2^{j-1}, 2^{j-1}]$, it results that $e_j(t) = 0$, and hence $\mathbf{E}\left(|e_j(t)| \Big/ N_j(t) \leq 1\right) = 0$. Since

$$|e_j(t)| \leq 2(dX \star |\bar{\psi}_j|(t)) \leq 2^{j+1} \|\bar{\psi}\|_\infty N_j(t) \ ,$$

it follows that

$$\mathbf{E}(|e_j(t)|) \leq 2\|\bar{\psi}\|_\infty 2^j \mathbf{E}\left(N_j(t) \Big/ N_j(t) > 1\right) \mathrm{Prob}(N_j(t) > 1) \ . \tag{74}$$

Since $N_j(t)$ is a Poisson random variable of parameter $\lambda 2^j$, we verify that

$$\mathbf{E}\left(N_j(t) \Big/ N_j(t) > 1\right) \mathrm{Prob}(N_j(t) > 1) = \lambda 2^j (1 - e^{-\lambda 2^j}) \ ,$$

which implies from (74) that

$$\mathbf{E}(|e_j(t)|) \leq 2\lambda \|\bar{\psi}\|_\infty (1 - e^{-\lambda 2^j}) = O(\lambda^2 2^j) \ , \tag{75}$$

and, together with (73) and (72) proves (13).

Property (14) is proved by showing that $2^{-j/2} X \star \psi_j(t)$ converges to a Gaussian random process at large scales 2^j. The convergence of $2^{-j/2} X \star \psi_j(t)$ relies on the use of a central-limit theorem for real dependent random variables. The extension to the two-dimensional complex random variables is done by considering arbitrary linear combinations of its real and imaginary

parts. The Cramer-Wold theorem proves that if $X_j = 2^{-j/2} X \star \psi_j(t) = 2^{-j/2} Re(X_j) + i\, 2^{-j/2} Im(X_j)$ satisfies

$$\forall (\alpha, \beta) \in \mathbb{R}^2 \ , \ \alpha Re(X_j) + \beta Im(X_j) \xrightarrow[j\to\infty]{l} \alpha A_1 + \beta A_2 \tag{76}$$

then $X_j \xrightarrow[j\to\infty]{l} A_1 + iA_2$. The random variables A_1 and A_2 are zero-mean Gaussian random variables if and only if $\alpha A_1 + \beta A_2$ is a centered Gaussian random variable for all $(\alpha, \beta) \in \mathbb{R}^2$. But

$$\alpha Re(X_j) + \beta Im(X_j) = X \star (\alpha Re(\psi_j) + \beta Im(\psi_j)) \ ,$$

so the convergence of X_j to a complex Gaussian variable will follow by showing that $2^{-j/2} X \star \tilde{\psi}_j \to \mathcal{N}(0, \sigma^2)$ for any wavelet of the form $\tilde{\psi}_j = \alpha Re(\psi_j) + \beta Im(\psi_j)$.

We thus concentrate in the real case, and we denote the real wavelet ψ_j to simplify notations. Assuming $j > 0$,

$$X \star \psi_j(t) = 2^j dX \star \bar{\psi}_j(t) = 2^j \int_{-2^{j-1}}^{2^{j-1}} \bar{\psi}_j(u-t)\, dX(u) = \sum_{i=-2^{j-1}}^{2^{j-1}-1} S_{i,j} \ ,$$

where $S_{i,j} = \int_i^{i+1} \bar{\psi}(2^{-j}(u-t)) dX(u)$ are a collection of zero-mean independent random variables. We apply the Berk central limit theorem [**?**], to this sum of independent random variables.

THEOREM A.1 (Berk Central Limit). *For any $j \in \mathbb{N}$, let $\{S_{i,j}\}_{i=1,\dots,n_j}$ be a sequence of zero mean random variables such that for any $i \leq n_j$ $S_{i,j}$ is independent of $S_{i+r,j}$ for $r \geq m_j$. If the following properties are satistied*

(i) $\exists \delta > 0$, $\lim_{j\to\infty} n_j^{-1} m_j^{2+2/\delta} = 0$
(ii) $\exists M > 0$, $\forall i, j > 0$, $\mathbf{E}(|S_{i,j}|^{2+\delta}) \leq M$
(iii) $\exists K > 0$, $\forall i, j, l > k > 0$, $Var(\sum_{i=k+1}^{k+l} S_{i,j}) \leq l\, K$
(iv) $\lim_{j\to\infty} n_j^{-1} Var(\sum_{i=1}^{n_j} S_{i,j}) = \sigma^2 > 0$

then

$$n_j^{-1/2} \sum_{i=1}^{n_j} S_{i,j} \xrightarrow[j\to\infty]{l} \mathcal{N}(0, \sigma^2) \ . \tag{77}$$

Let us now verify the hypothesis of this central limit theorem, with $m_j = 1$, $n_j = 2^j$ and $\delta = 1$. Since the variables $(S_{i,j})_i$ are independent, hypothesis (i) is verified with $m_j = 1$. Moreover, we verify that

$$\mathbf{E}(|S_{i,j}|^q) \leq \|\psi\|_\infty^q \mathbf{E}(|N_0|^q)$$

where N_0 is the number of jumps of dX in an interval of length 1. Since it follows a Poisson distribution of parameter λ, it has finite moments. It results that hypothesis (ii) is verified for $\delta = 1$. Since the $S_{i,j}$ are independent, $\mathrm{Var}(\sum_{i=k+1}^{k+l} u_{i,j}) \leq \|\psi\|_\infty^2 l \mathbf{E}(|N_0|^2)$, which verifies hypothesis (iii). Finally, since dX is a white noise of power spectrum λ

$$2^{-j} \sum_{|i|\leq 2^j} \mathbf{E}(|S_{i,j}|^2) = 2^j \mathbf{E}(|dX \star \bar{\psi}_j|^2) = 2^j \sigma^2(dX) \|\bar{\psi}_j\|^2 = \lambda \, \|\bar{\psi}\|^2 \ .$$

It verifies the last hypothesis (iv). Applying (A.1) and the Cramer-Wald theorem proves that $2^{-j/2} X \star \psi_j(t)$ converges to a complex Gaussian distribution of total variance $\lambda \, \|\bar{\psi}\|^2$. In order to control the limit of first order moments, we use the following lemma on uniform integrability of sequences of random variables:

LEMMA A.2. *([10], thm 6.1-6.2) Let $\{X_j\} j \in \mathbb{N}$ be a sequence of random variables. If $X_j \stackrel{d}{\to} X_\infty$ and*

$$\sup_j \mathbf{E}(|X_j|^{1+\delta}) < \infty \quad \textit{for} \quad \delta > 0 \ ,$$

then

$$\lim_{j\to\infty} \mathbf{E}(X_j) = \mathbf{E}(X_\infty) \ .$$

As a result, for any $\alpha \leq 2$, $\mathbf{E}(|2^{-j/2} X \star \psi_j|^\alpha) \to \mathbf{E}(|Z_1 + iZ_2|^\alpha)$, where Z_1 and Z_2 are Gaussian random variables with total variance $\lambda \, \|\bar{\psi}\|^2$. For $\alpha = 1$, it results that there exists a constant C, depending only on the wavelet ψ, such that

$$\lim_{j\to\infty} 2^{-j/2} \mathbf{E}(|X \star \psi_j|) = \lambda^{1/2} \, \|\bar{\psi}\| \, C \ .$$

which proves (14).

The proof of (15) is very similar to the proof of (13). The key property is that $||X \star \psi_{j_1}| \star \psi_{j_2}|$ only depends on values of X over an interval of size $2^{j_1} + 2^{j_2}$. From (75), it results that

(78)
$$|X \star \psi_{j_1}| \star \psi_{j_2}(t) = 2^{j_1} |dX \star \bar{\psi}_{j_1}| \star \psi_{j_2}(t) \stackrel{d}{=} 2^{j_1} (dX \star (|\bar{\psi}_{j_1}| \star \psi_{j_2})(t) + e_{j_1} \star \psi_{j_2}) \ ,$$

with $\mathbf{E}(|e_{j_1} \star \psi_{j_2}|) = O(\lambda^2 2^{j_1})$. As a consequence, (78) and (13) imply that

$$\begin{aligned} \mathbf{E}(|\,|X \star \psi_{j_1}| \star \psi_{j_2}(t)|) &= 2^{j_1} (\mathbf{E}(|dX \star (|\bar{\psi}_{j_1}| \star \psi_{j_2})|(t)) + O(\lambda^2 2^{j_1})) \ , \\ &= 2^{j_1} \lambda \|(|\bar{\psi}_{j_1}| \star \psi_{j_2})\|_1 (1 + O(\lambda 2^{j_1}) + O(\lambda 2^{j_2})) \ . \end{aligned} \tag{79}$$

Using again (13), we conclude that

$$\begin{aligned}\tilde{S}X(j_1,j_2) &= \frac{\mathbf{E}(|\,|dX\star\bar{\psi}_{j_1}|\star\psi_{j_2}(t)|)}{\mathbf{E}(|dX\star\bar{\psi}_{j_1}|)}\\ &= \frac{\|(|\bar{\psi}|\star\psi_{j_2-j_1})\|_1}{\|\bar{\psi}\|_1}\left(1+O(\lambda 2^{j_1})+O(\lambda 2^{j_2})\right)\ ,\end{aligned}$$

which proves (15). Finally, in order to prove (16), observe that $|X\star\psi_j|$ is a stationary process with lag 2^j. As a result, by using the same Central Limit argument to prove (14), one can verify that $2^{j_2/2}(|X\star\psi_{j_1}|\star\psi_{j_2})$ converges in distribution towards a Gaussian distribution as $j_2\to\infty$, which yields a decay on the normalized second order scattering of the form $\tilde{S}dX(j_1,j_2)\simeq 2^{-j_2/2}$ as $j_2\to\infty$. □

Appendix B: Proof of lemma 3.3

Let $Y_j(t)=2^{j/2}|dB\star\varphi|\star\psi_j(t)$. To prove that $\mathbf{E}(|Y_j|)$ converges to a constant, we shall prove that the distribution of Y_j is asymptotically Gaussian:

$$Y_j(t)\xrightarrow[j\to\infty]{l} A=A_1+iA_2 \tag{80}$$

where A_1 and A_2 are two zero-mean Gaussian distributions of total variance $\sigma_1^2+\sigma_2^2=\|\psi\|_2^2\int R_{|dB\star\varphi|}(\tau)d\tau$, which is the first result of Lemma 3.3. We shall also prove that

$$\lim_{j\to\infty}\mathbf{E}(|Y_j|^2)=\mathbf{E}(|A|^2). \tag{81}$$

Using again lemma (A.2), we conclude that

$$\lim_{j\to\infty}\mathbf{E}(|Y_j|)=\mathbf{E}(|A|)>0\ . \tag{82}$$

and hence finish the proof of Lemma 3.3.

For that purpose, we follow the same strategy as in the proof of (14), applied to the process $Y_j=2^{j/2}|dB\star\varphi|\star\bar{\psi}_j(t)$, where $\bar{\psi}_j$ is any linear combination of real and imaginary parts of ψ_j.

Let us write $\varphi_\Delta=\varphi\,\mathbf{1}_{[-\Delta/2,\Delta/2]}$. We shall limit ϕ to a compact support by defining $\{\Delta_j\}_{j\geq 0}$ with $\lim_{j\to\infty}\Delta_j=\infty$ and decompose

$$|dB\star\varphi(t)|=|dB\star\varphi_{\Delta_j}+dB\star(\varphi-\varphi_{\Delta_j})|\ .$$

As a result

$$|dB\star\varphi(t)|=|dB\star\varphi_{\Delta_j}|+Z_j(t)$$

with $\mathbf{E}(|Z_j|) \leq \mathbf{E}(|dB \star (\varphi - \varphi_{\Delta_j})|)$. Since dB is the Wiener measure, if $\theta \in \mathbf{L}^2(\mathbb{R})$ then

$$\mathbf{E}(|dB \star \theta|) \leq \mathbf{E}(|dB \star \theta|^2)^{1/2} = \|\theta\|_2 \ , \tag{83}$$

so $\mathbf{E}(|Z_j|) \leq \|\varphi - \varphi_{\Delta_j}\|_2$. It results that

$$|dB \star \varphi| \star \bar{\psi}_j(t) = |dB \star \varphi_{\Delta_j}| \star \bar{\psi}_j(t) + Z_j \star \bar{\psi}_j(t) \ , \tag{84}$$

and

$$\mathbf{E}(|Z_j \star \bar{\psi}_j|) \leq \mathbf{E}(|Z_j|)\|\bar{\psi}_j\|_1 \leq \|\varphi - \varphi_{\Delta_j}\|_2 \, \|\bar{\psi}\|_1 \ .$$

Since $\lim_{j\to\infty} \Delta_j = \infty$, $\lim_{j\to\infty} \|\varphi - \varphi_{\Delta_j}\|_2 = 0$ so $Z_j \star \bar{\psi}_j$ converges to 0 in probability when j increases. So the limits of $|dB \star \varphi| \star \bar{\psi}_j(t)$ and $|dB \star \varphi_{\Delta_j}| \star \bar{\psi}_j(t)$ are equal.

We now prove (24) by applying Berk central limit Theorem A.1, to show that $\bar{Y}_j = |dB \star \varphi_{\Delta_j}| \star \bar{\psi}_j(t)$ converges to a normal distribution. Since $|dB \star \varphi_{\Delta_j}| \star \bar{\psi}_j(t)$ is stationary, its distribution can be evaluated at $t = 0$

$$|dB \star \varphi_{\Delta_j}| \star \bar{\psi}_j(0) = \int |dB \star \varphi_{\Delta_j}|(u)\bar{\psi}_j(-u)du.$$

The central-limit theorem is applied by dividing this integral into disjoint integrals

$$S_{i,j} = 2^j \int_{2^j b_{i,j}}^{2^j b_{i+1,j}} |dB \star \varphi_{\Delta_j}(u)| \, \bar{\psi}_j(-u)du \ , \tag{85}$$

where for each $j \in \mathbb{Z}$, $\{b_{i,j}\}_{1\leq i\leq n_j}$ is an increasing sequence of points in $\mathbb{R} \cup \{\pm\infty\}$ such that

$$\forall i \ , \ \int_{b_{i,j}}^{b_{i+1,j}} |\bar{\psi}(-u)|du = 2^{-j}\|\bar{\psi}\|_1 \ . \tag{86}$$

Since $\bar{\psi}$ is $\mathbf{C^1}$ and bounded, we verify that $n_j \simeq 2^j$. Summing these random variables gives

$$2^{-j/2} \sum_{i=1}^{n_j} S_{i,j} = 2^{j/2} \, |dB \star \varphi_{\Delta_j}| \star \bar{\psi}_j(0) \ . \tag{87}$$

We now show that the $S_{i,j}$ satisfy the hypothesis of the Beck central-limit theorem so that we can apply the convergence result (77) which implies (24).

Let us first prove that $S_{i,j}$ is independent of $S_{i+r,j}$ for $r \geq m_j$ which satisfies (i). Since $\bar{\psi}$ is bounded, it results that $\inf_{i,j} 2^j |b_{i,j} - b_{i+1,j}| = \eta > 0$.

Since φ_{Δ_j} has a support of size Δ_j and dB is a Wiener Noise, it follows that $|dB \star \varphi_{\Delta_j}|(u)$ is independent of $|dB \star \varphi_{\Delta_j}|(u')$ for $|u-u'| > \Delta_j$ and hence that $S_{i,j}$ is independent of $S_{i+r,j}$ for $r \geq m_j = \Delta_j/\eta$.

To verify (i) let us set $\delta = 1$. Since $n_j \simeq 2^j$, if we choose $\Delta_j = 2^{j/5}$ then

$$\lim_{j\to\infty} \frac{m_j^4}{n_j} \leq \eta^{-4} \lim_{j\to\infty} 2^{j(4/5-1)} = 0 \; . \tag{88}$$

We now verify condition (ii) with $\delta = 1$. Since $\bar\psi_j(u)$ has a zero integral, one can replace $|dX \star \varphi_{\Delta_j}(u)|$ by $Q_j(u) = |dX \star \varphi_{\Delta_j}|(u) - \mathbf{E}(|dX \star \varphi_{\Delta_j}|)$ in the definition (85) of $S_{i,j}$. It results that

$$\begin{aligned}
\mathbf{E}(|S_{i,j}|^3) &\leq \iiint \mathbf{E}(Q_j(u)Q_j(u')Q_j(u''))\, 2^{3j} |\bar\psi_j(-u)|\, |\bar\psi_j(-u')|\, |\bar\psi_j(-u'')|\, du\, du'\, du'' \\
&\leq \mathbf{E}(|dB \star \varphi_{\Delta_j}|^3) \|\bar\psi\|_1^3 = 2^{5/2}\pi^{-1/2} \|\varphi_{\Delta_j}\|_2^3 \|\bar\psi\|_1^3 \leq 2^{5/2}\pi^{-1/2} \|\varphi\|_2^3 \|\bar\psi\|_1^3 \; .
\end{aligned}$$

Let us now verify condition (iii). The sum $A_{k,l,j} = \sum_{i=k}^{k+l} S_{i,j}$ is by definition

$$A_{k,l,j} = 2^j \int_{2^j b_{k,j}}^{2^j b_{k+l,j}} |dB \star \varphi_{\Delta_j}(u)|\, \bar\psi_j(-u) du = \int_{\mathbb{R}} |dB \star \varphi_{\Delta_j}(u)|\, h_{k,l,j}(u)\, du$$

with $h_{k,l,j}(u) = 2^j \bar\psi_j(-u) 1_{[2^j b_{k,j}, 2^j b_{k+l,j}]}(u)$. It results that

$$\mathrm{Var}(A_{k,l,j}) \leq \|R_{|dB\star\varphi_{\Delta_j}|}\|_1 \, \|h_{k,l,j}\|_2^2 \; . \tag{89}$$

But, with a change of variable and applying (86) we get

$$\|h_{k,l,j}\|_2^2 = \int_{2^j b_{k,j}}^{2^j b_{k+l,j}} |\bar\psi(2^{-j}u)|^2\, du \leq \|\bar\psi\|_\infty \int_{b_{k,j}}^{b_{k+l,j}} 2^j |\bar\psi(u)|\, du \leq \|\bar\psi\|_\infty \, \|\bar\psi\|_1 \, l \; .$$

We are now going to bound $\|R_{|dB\star\varphi_{\Delta_j}|}\|_1$ by using the decay $\varphi(u) = O((1+|u|^{-2}))$.

$$R_{|dB\star\varphi_{\Delta_j}|}(\Delta) = \mathbf{E}(|dB \star \varphi_{\Delta_j}(\Delta)|\; |dB \star \varphi_{\Delta_j}(0)|) - \mathbf{E}(|dB \star \varphi_{\Delta_j}|)^2 \; .$$

If $|\Delta| > |\Delta_j|$ then since the support of $\varphi_{\Delta_j}(u)$ and $\varphi_{\Delta_j}(u-\Delta)$ does not overlap, $|dB \star \varphi_{\Delta_j}(\Delta)|$ and $|dB \star \varphi_{\Delta_j}(0)|$ are independent random variables so $R_{|dB\star\varphi_{\Delta_j}|}(\Delta) = 0$. Otherwise, we decompose

$$|dB \star \varphi_{\Delta_j}(u)| = |dB \star \varphi_{\Delta}(u) + dB \star (\varphi_{\Delta_j} - \varphi_{\Delta}(u)| \; .$$

Since $|dB \star \varphi_\Delta(0)|$ and $|dB \star \varphi_\Delta(\Delta)|$ are independent random variables,

$$\begin{aligned} |R_{|dB\star\varphi_{\Delta_j}|}(\Delta)| &\leq |\mathbf{E}(|dB\star\varphi_\Delta|)^2 - \mathbf{E}(|dB\star\varphi_{\Delta_j}|)^2| + \\ &\quad +2\,\mathbf{E}(|dB\star\varphi_\Delta|)\mathbf{E}(|dB\star(\varphi_{\Delta_j}-\varphi_\Delta)|) + \mathbf{E}(|dB\star(\varphi_{\Delta_j}-\varphi_\Delta)|)^2 \ . \end{aligned}$$

Since $\mathbf{E}(|dB\star\theta|) \leq \mathbf{E}(|dB\star\theta|^2)^{1/2} \leq \|\theta\|_2$, by applying this to $\theta = \varphi_\Delta$ and $\theta = \varphi_{\Delta_j} - \varphi_\Delta$ one can verify that

$$|R_{|dB\star\varphi_{\Delta_j}|}(\Delta)| \leq 6\|\varphi\|_2\,\|\varphi-\varphi_\Delta\|_2 \ . \tag{90}$$

Since $\varphi(u) = O((1+|u|)^{-2})$ it results that $\|\varphi-\varphi_\Delta\|_2 = O((1+|\Delta|)^{-3/2})$ so $\|R_{|dB\star\varphi_{\Delta_j}|}\|_1$ is bounded independently of j. Inserting this in (89) proves the theorem hypothesis (iii).

Let us now verify the hypothesis (iv). It results from (87) that

$$2^{-j}\mathrm{Var}(\sum_i S_{i,j}) = 2^j\mathrm{Var}(|dX\star\varphi_{\Delta_j}|\star\bar\psi_j) = 2^j\int \widehat R_{|dX\star\varphi_{\Delta_j}|}(\omega)|\widehat{\bar\psi}(2^j\omega)|^2 d\omega \ .$$

We proved (90) that $R_{|dB\star\varphi_{\Delta_j}|} \in \mathbf{L}^1$ but the same inequality is valid for $R_{|dB\star\varphi_\Delta|}$ which proves that it is also in $\mathbf{L}^1$. It results that $\widehat R_{|dX\star\varphi|}$ is continuous. Since φ_{Δ_j} converges to φ in $\mathbf{L}^2\cap\mathbf{L}^1$ as $j\to\infty$, $\widehat R_{|dX\star\varphi^j|}(0)$ converges to $\widehat R_{|dX\star\varphi|}(0)$. Since $2^j|\widehat{\bar\psi}(2^j\omega)|^2$ converges to $\|\bar\psi\|_2^2\delta(\omega)$ when j goes to ∞

$$\lim_{j\to\infty} 2^{-j}\mathrm{Var}(\sum_i S_{i,j}) = \widehat R_{|dX\star\varphi|}(0)\|\bar\psi\|_2^2 = \bar\sigma^2,$$

which proves condition (iv).

We can thus apply Theorem A.1 which proves that $2^{j/2}|dB\star\varphi|\star\bar\psi_j(t)$ converges in distribution to $\mathcal{N}(0\,\bar\sigma^2)$ and hence (24). Finally, by following the same reasoning used in Theorem 2.2, we apply Lemma A.2 to conclude that $\lim_{j\to\infty} 2^{j/2}\mathbf{E}(|\,|dB\star\varphi|\star\psi_j|) = C > 0$, which proves (25) □.

Appendix C: Various results on the MRM measure

LEMMA C.1. *The process $\omega_l^T(t)$ used for the construction of the MRM dM is an infinitely-divisible process whose two-points characteristic function reads:*

$$\mathbf{E}\left(e^{p_1\omega_l^T(t_1)+p_2\omega_l^T(t_2)}\right) = e^{[F(p_1)+F(p_2)]\rho_l^T(0)+[F(p_1+p_2)-F(p_1)-F(p_2)]\rho_l^T(t_2-t_1)} \tag{91}$$

where $F(-ip)$ is the cumulant generating function characterizing the infinitely divisible law as provided by the Levy-Khintchine formula where the drift term is chosen such that

$$F(1) = 0 \ , \tag{92}$$

and where the function $\rho_l^T(\tau)$ is defined by:

$$\rho_l^T(\tau) = \begin{cases} \ln(T/l) + 1 - |\tau|/l \ , & \text{if } |\tau| \leq l \ , \\ \ln(T/|\tau|) \ , & \text{if } l \leq |\tau| < T \ , \\ 0 \ , & \text{otherwise} \ . \end{cases} \tag{93}$$

Moreover, the function $\zeta(p)$ which satisfies (32) (with $X = dM$ where dM is the associated MRM) is given by

$$\zeta(p) = p - F(p).$$

The proof of this Lemma is given in [6].
The next Lemma uses an alternative MRM measure considered in Ref. [6] defined by

$$d\tilde{M}(t) = \lim_{l \to 0} e^{\tilde{\omega}_l^T(t)} dt$$

where $\tilde{\omega}_l^T$ is defined exactly as the process ω_l^T but only differs by its ρ function which is replaced by : $\tilde{\rho}_l^T(\tau) = \rho_l^T(\tau) + \frac{\tau}{T} - 1$, for $\tau \leq T$. One can then easily show that $\tilde{\omega}_l^T$ is linked with ω_l^T by the following cascade property :

$$\forall l \leq a \leq T, \ \omega_l^T(u) \stackrel{a.s.}{=} \tilde{\omega}_l^a(u) + \omega_a^T(u) \tag{94}$$

where $\tilde{\omega}_l^a$ and ω_a^T are independent copies of the processes defined previously. Moreover, $\tilde{\omega}_l^T$ satisfies both (39) and (39).

We are now ready to state the last Lemma we will need.

LEMMA C.2. *Let ω_l^T the infinitely divisible process associated with the MRM dM and ψ be a wavelet of support in $[0,1]$ such that $\|\psi\|_\infty < \infty$. For all α such that $0 < \alpha < 1$, one has:*

$$\forall l < 2^j \ , \ (\psi_j \star e^{\omega_l^T})(t) = e^{\omega_{2^{j\alpha}}^T(t)} \left(\psi_j \star e^{\tilde{\omega}_l^{2^{j\alpha}}} \right) + \eta_{l,j}(t) \ , \tag{95}$$

where the process $\eta_{l,j}(t)$ has a limit process $\lim_{l \to 0} \eta_{l,j}(t) = \eta_j(t)$ which satisfies, in the limit $j \to -\infty$,

$$\mathbf{E}(|\eta_j(t)|) = O(2^{j \frac{1-\alpha(1+F(2))}{2}}) \tag{96}$$

and

$$\mathbf{E}(|\eta_j(t)|^2) = O(2^{j(3-F(2)-\alpha)}) \ . \tag{97}$$

Without loss of generality we fix $t = 0$. Let us consider $0 < \alpha < 1$ and l and j small enough and such that:

$$l < 2^j < 2^{j\alpha} < T$$

Let us first remark that, for $u < 2^{j\alpha}$, one has from (91):

$$\mathbf{E}\left(e^{p(\omega^T_{2^{j\alpha}}(u)+\omega^T_{2^{j\alpha}}(0))}\right) = 2^{-j\alpha F(2p)} T^{F(2p)} e^{F(2p)(1-u2^{-j\alpha})} \tag{98}$$

where $F(p) = \varphi(-ip) = p - \zeta(p)$. Hence, we have:

$$\begin{aligned} \mathbf{E}\left(e^{2\omega^T_{2^{j\alpha}}(u)}\right) &= 2^{-j\alpha F(2)} T^{F(2)} e^{F(2)} \\ \mathbf{E}\left(e^{\omega^T_{2^{j\alpha}}(u)+\omega^T_{2^{j\alpha}}(0)}\right) &= 2^{-j\alpha F(2)} T^{F(2)} e^{F(2)(1-u2^{-j\alpha})} \ . \end{aligned}$$

One defines $\eta_{l,j}$ as:

$$\eta_{l,j}(0) = 2^{-j} \int \psi(u2^{-j}) \left(e^{\omega^T_l(u)} - e^{\tilde{\omega}^{2^{j\alpha}}_l(u)+\omega^T_{2^{j\alpha}}(0)}\right) du \tag{99}$$

Using dominated convergence, (94), $\mathbf{E}(e^{\tilde{\omega}^{2^{j\alpha}}_l}) = 1$ and the fact that ψ is a bounded function of support $[0, 1]$ one has:

$$\begin{aligned} \mathbf{E}(\lim_{l\to 0} |\eta_{l,j}|) &= \lim_{l\to 0} \mathbf{E}(|\eta_{l,j}|) \\ &\leq ||\psi||_\infty 2^{-j} \int_0^{2^j} \sqrt{\mathbf{E}\left[\left(e^{\omega^T_{2^{j\alpha}}(u)} - e^{\omega^T_{2^{j\alpha}}(0)}\right)^2\right]} du \\ &= ||\psi||_\infty 2^{-j} \int_0^{2^j} \sqrt{\mathbf{E}\left(e^{2\omega^T_{2^{j\alpha}}(0)} + e^{2\omega^T_{2^{j\alpha}}(u)} - 2e^{\omega^T_{2^{j\alpha}}(u)+\omega^T_{2^{j\alpha}}(0)}\right)} du \\ &= \sqrt{2}||\psi||_\infty 2^{-\frac{j\alpha F(2)}{2}} T^{\frac{F(2)}{2}} e^{\frac{F(2)}{2}} \int_0^1 \left(1 - e^{-F(2)u2^{j(1-\alpha)}}\right)^{\frac{1}{2}} du \\ &\underset{j\to-\infty}{=} O(2^{j\frac{1-\alpha(1+F(2))}{2}}) \end{aligned}$$

which proves (96). In order to bound the second moment, we consider

$$\begin{aligned}
& \mathbf{E}(\lim_{l\to 0}|\eta_{l,j}|^2) = \lim_{l\to 0}\mathbf{E}(|\eta_{l,j}|^2) \\
= \;& 2^{-2j}\iint_0^{2^j}\psi(2^{-j}u)\psi(2^{-j}u')\mathbf{E}(e^{\tilde\omega_l^{2^{j\alpha}}(u)+\tilde\omega_l^{2^{j\alpha}}(u')})\mathbf{E}\left((e^{\omega_{2^{j\alpha}}^T(u)}-e^{\omega_{2^{j\alpha}}^T(0)})(e^{\omega_{2^{j\alpha}}^T(u')}-e^{\omega_{2^{j\alpha}}^T(0)})\right)dudu' \\
\leq \;& 2^{-2j}2^{-j\alpha F(2)}(Te)^{F(2)}\iint_0^{2^j}|\psi(2^{-j}u)|\,|\psi(2^{-j}u')|\mathbf{E}(e^{\tilde\omega_l^{2^{j\alpha}}(u)+\tilde\omega_l^{2^{j\alpha}}(u')})\cdot \\
\cdot \;& \left|e^{-F(2)|u-u'|^{2^{-j\alpha}}}+1-e^{-F(2)|u|^{2^{-j\alpha}}}-e^{-F(2)|u'|^{2^{-j\alpha}}}\right|dudu' \\
= \;& 2^{-j\alpha F(2)}(Te)^{F(2}\iint_0^1|\psi(u)||\psi(u')|\mathbf{E}(e^{\tilde\omega_l^{2^{j\alpha}}(2^ju)+\tilde\omega_l^{2^{j\alpha}}(2^ju')})\cdot \\
\cdot \;& \left|e^{-F(2)|u-u'|^{2^{j(1-\alpha)}}}+1-e^{-F(2)|u|^{2^{j(1-\alpha)}}}-e^{-F(2)|u'|^{2^{j(1-\alpha)}}}\right|dudu' \\
\underset{j\to-\infty}{=} \;& O(2^{j(1-\alpha(1+F(2)))})\mathbf{E}(|e^{\tilde\omega_l^{2^{j\alpha}}(2^ju)}\star|\psi||^2) \\
= \;& O(2^{j(1-\alpha(1+F(2))+\zeta(2)+\alpha F(2)}) = O(2^{j(3-F(2)-\alpha)})\;.\;\square
\end{aligned}$$

Let us remark that one could obtain a smaller error with a smoother variant of the ω_l. Indeed, as shown in [34] it is possible to choose the way ω_l is regularized at scale l. One can thus define a MRM process using ω_l with a covariance function that is C^2 at $\tau = 0$. In that case, in (98), the function $\rho_l(u)$ would be proportional to $2^{-2j\alpha}u^2$ and the error mean absolute value could be bounded by $2^{j(1-\alpha-F(2)/2)}$.

Appendix D: Proof of Theorem 4.3

As for the first order, using first (39) and then (40) with $s = 2^{j_2-L}$ we obtain :

$$\begin{aligned}
|\psi_{j_2}\star|\psi_{j_1}\star dM_l||(t) &\overset{law}{=} |\psi_{j_2}\star|\psi_{j_1}\star dM_l||(0) \\
&\overset{law}{=} 2^{-j_2}e^{\Omega_{2^{j_2-L}}}\left|\int\psi(-u2^{-j_2})2^{-j_1}\left|\int\psi(\frac{u-v}{2^{j_1}})e^{\omega_l^{2^{j_2}}(v)}dv\right|du\right|\;.
\end{aligned}$$

Making the changes of variables $u' = u2^{-j_2}$ and $v' = v2^{-j_1}$ and using (39), leads to

$$|\psi_{j_2}\star|\psi_{j_1}\star dM_l||(t) \overset{law}{=} e^{\Omega_{2^{j_2-L}}}\left|\int\psi(-u)\left|\int\psi(2^{j_2-j_1}u-v)e^{\omega_{2^{-j_1}l}^{2^{j_2-j_1}}(v)}dv\right|du\right|\;.$$

Since j_2 is fixed, with no loss of generality, in the following we can set $j_2 = 0$. Using (39), one gets

$$|\psi\star|\psi_{j_1}\star dM_l(0)|| \overset{law}{=} e^{\Omega_{2^{-L}}}\left|\int\psi(-u)\left|\psi_{j_1}\star e^{\omega_l^1}(u)\right|du\right|\;. \tag{100}$$

We now use the Lemma C.2 proved in Appendix C with $\alpha = \frac{1-2\nu}{1+F(2)}$ ($\nu < 1/2$). We get :

$$\mathbf{E}(|\eta_{j_1}|) = O(2^{j_1\nu}) \ , \tag{101}$$

and

$$\begin{aligned}
\psi_{j_1} \star e^{\omega_l^1}(u) &= 2^{-j_1} e^{\omega^1_{2^{j_1}\alpha}(u)} \int \psi(\frac{u-v}{2^{j_1}}) e^{\tilde{\omega}_l^{2^{j_1}\alpha}(v)} dv + \eta_{j_1,l}(u) \\
&\overset{law}{=} e^{\omega^1_{2^{j_1}\alpha}(u)} \int \psi(u2^{-j_1} - v) e^{\tilde{\omega}^1_{l2^{-j_1}\alpha}(v2^{j_1(1-\alpha)})} dv + \eta_{j_1,l}(u) \\
&\overset{law}{=} e^{\omega^1_{2^{j_1}\alpha}(u)} \int \psi(u2^{-j_1} - v) e^{\tilde{\omega}^{2^{j_1(\alpha-1)}}_{l2^{-j_1}}(v)} dv + \eta_{j_1,l}(u) \\
&\underset{l\to 0}{\to} e^{\omega^1_{2^{j_1}\alpha}(u)} \tilde{\epsilon}_{2^{j_1(\alpha-1)}}(2^{-j_1}u) + \eta_{j_1}(u) \ ,
\end{aligned}$$

where we used property (39) for $\tilde{\omega}_l^T$ and we defined the T-dependent noise:

$$\tilde{\epsilon}_T(t) = \lim_{l\to 0} \int \psi(t-v) e^{\tilde{\omega}^T_{l2^{-j_1}}(v)} dv \ . \tag{102}$$

If follows that :

$$\lim_{l\to 0} \psi \star |\psi_{j_1} \star dM_l(0)| \overset{law}{=} e^{\Omega_{2^{-L}}} \int \psi(-u) e^{\omega^1_{2^{j_1}\alpha}(u)} |\tilde{\epsilon}_{2^{j_1(\alpha-1)}}(2^{-j_1}u)| du + \int \psi(-u)\tilde{\eta}_{j_1}(u) du \ ,$$

where

$$\tilde{\eta}_{j_1}(u) = |e^{\omega^1_{2^{j_1}\alpha}(u)} \tilde{\epsilon}_{2^{j_1(\alpha-1)}}(2^{-j_1}u) + \eta_{j_1}(u)| - |\tilde{\epsilon}_{2^{j_1(\alpha-1)}}(2^{-j_1}u)| \ . \tag{103}$$

Along the same line as in ref [30, 6], it is easy to prove that in the limit $T \to \infty$:

$$\mathbf{E}(|\tilde{\epsilon}_T(t)|^q) \simeq \widetilde{K}_q T^{q-\zeta(q)} \ , \tag{104}$$

where $\tilde{K}_q$ does not depend on T (thanks to the stationarity of $\tilde{\epsilon}_T(t)$, it does not depend on t either). Since $\zeta(1) = 1$, let $\tilde{K}_1 = \tilde{K} = \mathbf{E}(|\tilde{\epsilon}_T(t)|)$ and let us define the centered process: $\bar{\epsilon}_T(t) = |\tilde{\epsilon}_T(t)| - \tilde{K}$. Let us remark that, when $T \to \infty$,

$$\mathbf{E}\left(\bar{\epsilon}_T^2\right) \simeq \mathbf{E}\left(\tilde{\epsilon}_T^2\right) \simeq T^{2-\zeta(2)} \ . \tag{105}$$

Thus we can write

$$\lim_{l\to 0} \psi \star |\psi_{j_1} \star dM_l|(0) \overset{law}{=} e^{\Omega_{2^{-L}}}(I + II + III) \ , \tag{106}$$

where

$$I = \tilde{K} \int \psi(-u) e^{\omega^1_{2^{j_1}\alpha}(u)} du \ , \tag{107}$$

$$II = \int \psi(-u) e^{\omega^1_{2^{j_1}\alpha}} \bar{\epsilon}_{2^{j_1(\alpha-1)}}(2^{-j_1}u) du \ , \tag{108}$$

$$III = \int \psi(-u) \tilde{\eta}_{j_1}(u) du \ . \tag{109}$$

Since $\int \psi(u) e^{\omega^1_{2^{j_1}\alpha}} du$ converges in law, when $j_1 \to -\infty$, towards $\epsilon_1(t)$, (where $\epsilon_1(t)$ is an independent copy of the process defined in (45)), we have, in the limit $j_1 \to -\infty$:

$$\mathbf{E}(|I|) \to K\tilde{K} \ , \tag{110}$$

where $K = \mathbf{E}(|\epsilon_1(t)|)$. Thus, since $\overline{S}dM(j_1, 0) = \mathbf{E}(|I + II + III|)$,

$$|\overline{S}dM(j_1, 0) - \widetilde{K}K| \leq |\mathbf{E}(|I + II|) - \mathbf{E}(|I|)| + \mathbf{E}(|III|) \ .$$

From the Lemma, we know that $\lim_{j_1 \to -\infty} \mathbf{E}(|\eta_{j_1}|) = 0$ and consequently $\lim_{j_1 \to -\infty} \mathbf{E}(|\tilde{\eta}_{j_1}|) = 0$ which leads to $\lim_{j_1 \to -\infty} \mathbf{E}(|III|) = 0$. Moreover

$$|\mathbf{E}(|I + II|) - \mathbf{E}(|I|)| \leq \mathbf{E}(|II|) \leq \sqrt{\mathbf{E}(|II^2|)} \ .$$

From the expression of II and the fact that $\bar{\epsilon}_{2^{j_1(\alpha-1)}}(2^{-j_1}u)$ is a $2^{j_1\alpha}$-dependent process, we have, when $j_1 \to -\infty$:

$$\begin{aligned} \mathbf{E}(|II|^2) &\leq ||\psi||^2_\infty \int_0^1 \int_0^1 \mathbf{E}\left(e^{\omega^1_{2^{j_1}\alpha}(u) + \omega^1_{2^{j_1}\alpha}(v)}\right) \mathbf{E}\left(\bar{\epsilon}_{2^{j_1(\alpha-1)}}(2^{-j_1}u) \bar{\epsilon}_{2^{j_1(\alpha-1)}}(2^{-j_1}v)\right) du dv \\ &\leq ||\psi||_\infty \mathbf{E}(\bar{\epsilon}^2_{2^{j_1(\alpha-1)}}) 2^{j_1\alpha} \mathbf{E}(e^{2\omega^1_{2^{j_1}\alpha}}) \simeq 2^{j_1(\alpha - F(2)} \ , \end{aligned}$$

which goes to 0 provided whe choose $1 > \alpha > F(2)$. Thus $\overline{S}dM(j_1, 0)$ converges to $\widetilde{K}K$ which proves (46).

Appendix E: Sketch of proof of Theorem 4.4

The proof of Theorem 4.4 can be established exactly along the same line of the proofs of Proposition 4.2 and Theorem 4.3. First let us remark that if $\overline{\psi}(t) = \int_{-\infty}^t \psi(u) du$, then a simple integration by parts allows one to show that:

$$|X_l \star \psi_j| = 2^j |dX_l \star \overline{\psi}_j| \ , \tag{111}$$

where $X_l(t)$ is defined in (47)) and $dX_l(t) = l^{\frac{2-\zeta(2)}{2}} e^{\omega_l^{2^L}(t)} dB(t)$. Then, in order to study the behavior of first and second order scattering moments of $X_l(t)$, one can adapt the proofs of the MRM case to the MRW case by replacing $e^{\omega_l(u)}du$ by $l^{\frac{2-\zeta(2)}{2}} e^{\omega_l^{2^L}(u)} dB(u)$.

For the first order moment, the Wiener noise scaling $dB(su) \overset{law}{=} s^{-1/2} dB(u)$ and (40), leads to $dX_{sl}(su) \overset{law}{=} e^{\Omega_s} s^{1/2-\zeta(2)/2}$. Thanks to (111), one gets (50).

As far as the second order scattering moment is concerned, a simple adaptation of Lemma C.2 allows one to follow the same steps as in Appendix D. One is lead to the same decomposition as in (106):

$$\lim_{l\to 0} \psi \star |\psi_{j_1} \star X_l|(0) \overset{law}{=} e^{\Omega_{2^{-L}}} (I + II + III) \ , \tag{112}$$

where

$$I = \tilde{K}' 2^{j_1 \frac{3-\zeta(2)}{2}} \int \psi(-u) e^{\omega^1_{2^{j_1}\alpha}(u)} du \ , \tag{113}$$

$$II = 2^{j_1} \int \psi(-u) e^{\omega^1_{2^{j_1}\alpha}} \bar{\epsilon}_{2^{j_1(\alpha-1)}}(2^{-j_1}u) du \ , \tag{114}$$

$$III = 2^{j_1} \int \psi(-u) \tilde{\eta}'_{j_1}(u) du \ , \tag{115}$$

where η'_j is the noise term corresponding to η_j in Lemma C.2 and

$$\tilde{K}' = \mathbf{E}\left(\left| \lim_{l\to 0} \int \overline{\psi}(u) l^{\frac{2-\zeta(2)}{2}} e^{\omega_l^1(u)} dB(u) \right| \right) \ .$$

Since from (104),

$$\mathbf{E}\left(\left| \lim_{j_1 \to -\infty} \int \psi(-u) e^{\omega^1_{2^{j_1}\alpha}(u)} du \right| \right) = \tilde{K} \ ,$$

the term $\mathbf{E}(|I|)$ behaves, when $j_1 \to -\infty$ as $\tilde{K}'\tilde{K} 2^{\frac{3j_1-\zeta(2)}{2}}$. The contribution of the terms II and III can be shown to be negligible following the same arguments as in Appendix D. Since the first order scattering moment behaves like $\tilde{K}' 2^{j_1 \frac{3-\zeta(2)}{2}}$ we obtain (51) with the same constant as in Theorem 4.3.

References

[1] P. Abry, P. Flandrin, M. Taqqu, and D. Veitch. Wavelets for the Analysis, estimation and synthesis of scaling data. *K. Park and W. Willinger eds*, 2000.

[2] J. Anden and S. Mallat. Deep scattering spectrum. *Submitted to IEEE transactions of Signal Processing*, 2013.

[3] A. Arneodo and S. Jaffard. L'Analyse Multi-fractale des Signaux. *Journal du CNRS*, 2004.
[4] E. Bacry, A. Kozhemyak, and J. F. Muzy. Continuous cascade models for asset returns. *Journal of Economic Dynamics and Control*, 32:156–199, 2008.
[5] E. Bacry, A. Kozhemyak, and J. F. Muzy. Log-normal continuous cascades: aggregation properties and estimation. application to financial time series. *Quantitative Finance*, 13:795–818, 2013.
[6] E. Bacry and J.F. Muzy. Log-infinitely divisible multifractal process. *Comm. in Math. Phys.*, 236:449–475, 2003.
[7] J. Bruna and S. Mallat. Invariant scattering convolution networks. *IEEE transactions of PAMI*, 2012.
[8] B. Castaing. Lagrangian and eulerian velocity intermittency. *Eur. Phys. J. B*, 29:357–358, 2002.
[9] O. Chanal, B. Chabaud, B. Castaing, and B. Hebral. *Eur. Phys. J. B*, 17:309–317, 2000.
[10] A. DasGupta. *Asymptotic Theory of Statistics and Probability*. Springer, 2008.
[11] J. Delour, J. F. Muzy, and A. Arneodo. Intermittency of 1d velocity spatial profiles in turbulence : a magnitude cumulant analysis. *Eur. Phys. J. B*, 23:243, 2001.
[12] U. Frisch. *Turbulence*. Cambridge University Press, 1995.
[13] L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica*, 50:1029–1054, 1982.
[14] S. Jaffard. Multifractal Formalism for functions, parts i and ii. *SIAM J. Math Anal.*, 1997.
[15] S. Jaffard. Wavelet Expansions, Function Spaces and Multifractal Analysis. 2000.
[16] S. Jaffard and Y. Meyer. Wavelet Methods for Pointwise Regularity and Local Oscillations of Functions. *Memoirs of the AMS*, 1996.
[17] J.P. Kahane and J. Peyriere. *Sur certaines martingales de B. Mandelbrot*. Departement de mathematique, 1975.
[18] A.N. Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large reynolds numbers. *Dokl. Akad. Nauk. SSSR*, 30:301, 1941.
[19] A. Kyprianou. *An Introduction to the Theory of Lévy Processes*. 2007.
[20] Y. LeCun, K. Kavukvuoglu, and C. Farabet. Convolutional networks and applications in vision. *Proc. IEEE Int. Symposium Circuits and Systems*, 2010.
[21] S. Mallat. Group Invariant Scattering. *Communications in Pure and Applied Mathematics*, 2012.
[22] B.B. Mandelbrot. On Intermittent Free Turbulence. *Turbulence of Fluids and Plasmas*, 1969.
[23] B.B. Mandelbrot. Intermittent Turbulence in self-similar Cascades: divergence of high moments and dimension of the carrier. *Journal of Fluid Mechanics*, 1974.
[24] B.B. Mandelbrot. *Fractals and scaling in finance. Discontinuity, concentration, risk*. Springer, 1997.
[25] D. McFadden. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica*, 57:995–1026, 1989.
[26] C. Meneveau and K.R. Sreenivasan. The Multifractal nature of Turbulent energy dissipation. *J. Fluid Mechanichs*, 224:429–484, 1991.
[27] J-F. Muzy, E. Bacry, and A. Arneodo. Multifractal Formalism for fractal signals: The structure-function approach versus the wavelet-transform modulus-maxima method. *Phys. Rev E*, 47:875, 1993.
[28] J-F. Muzy, E. Bacry, and A. Arneodo. Singularity Spectrum of Fractal signals from wavelet analysis: exact results. *J. Stat. Phys*, 70:635, 1993.

[29] J. F. Muzy, J. Delour, and E. Bacry. *Eur. J. Phys. B*, 17:537–548, 2000.
[30] J.F. Muzy and E. Bacry. Multifractal stationary random measures and multifractal random walks using log-infinitely divisible laws. *Phys. Rev. E*, 66, 2002.
[31] G. Oppenheim P.Doukhan and M. Taqqu. *Theory and Applications of Long-Range Dependence*. Birkhauser, Boston, 2002.
[32] L.F. Richardson. *Weather Prediction bt Numerical Process*. Cambridge Universtity Press, 1922.
[33] C. Robert and M. Rosenbaum. A new approach for the dynamics of ultra high frequency data: the model with uncertainty zones. *Journal of Financial Econometrics*, 9:344–366, 2011.
[34] R. Robert and V. Vargas. Gaussian multiplicative chaos revisited. *Annals of Probability*, 38(2):605–631, 2010.
[35] G. Selesnick, I. Baraniuk and Kingsbury N. A dual tree complex wavelet transform. *IEEE Signal Processing Magazine*, 123, 2005.
[36] L. Sifre and S. Mallat. Combined Scattering for Rotation Invariant Texture Analysis. *ESANN*, 2012.
[37] R. Vershynin. How close is the sample covariance matrix to the actual covariance matrix ? *Journal of Theoretical Probability*, 25(3):655–686, 2012.
[38] A.M. Yaglom. The influence of fluctuation in energy dissipation on the shape of turbulent characteristics in the inertial interval. *Soviet Physics Dockl.*, 1966.

Courant Institute of Mathematical Sciences
Department of Computer Science
715 Broadway
New York, NY, 10012 E-mail: bruna@cims.nyu.edu

Departement d'Informatique,
École Normale Supérieure
Rue d'Ulm
Paris 75005 France E-mail: mallat@cmap.polytechnique.fr

Centre des Mathématiques Appliquées ,
École Polytechnique
Route de Saclay
Palaiseau, France E-mail: bacry@cmap.polytechnique.fr E-mail: muzy@univ-corse.fr