Multifractal analysis of random wavelet series with generalized Gaussian mixture statistics

W. Ben Nasr, H. Halconruy and S. Jaffard

December 23, 2025

Abstract

Random wavelet series (RWS) provide a flexible framework for modeling multiscale stochastic processes, but the classical assumption of full independence between wavelet coefficients is often unrealistic, especially across scales. We introduce semi-dependent random wavelet series (SDRWS), which preserve independence within each scale while allowing interscale dependencies at fixed spatial locations. Adopting a sample-path perspective, we investigate how this relaxed dependency structure affects pointwise and global regularity properties. We analyze the resulting multifractal behavior and show how it differs from that of fully independent models. Finally, we study the uniform Hölder regularity exponent $H^{\min}$, providing theoretical results and statistical estimation procedures that are relevant for multifractal analysis and model validation.

Contents

1 Introduction

At the end of the 1980s, when orthonormal wavelet bases were introduced, the question of identifying pertinent wavelet models for data became a major issue. In signal and image processing, this issue was rapidly addressed through denoising algorithms that explicitly account for the clustering of large wavelet coefficients, such as hidden Markov tree models and neighborhood-based shrinkage rules, see e.g. [21, 68, 70, 66]. A first step was the estimation of the distributions of wavelet coefficients for a variety of signals and images; these computations led to a first important conclusion: these distributions often match a simple parametric mixture model whose components are a Dirac mass at the origin mixed with a generalized Gaussian distribution, see [77]. The knowledge of the distributions of wavelet coefficients at available scales has been used in different contexts; see e.g. [67], where a denoising method is based on the assumption that wavelet coefficients follow a generalized Gaussian distribution. A recent application is supplied by wavelet quantile normalization (WQN) in statistics, a denoising algorithm based on the idea of *mapping*, at each scale, the distribution of wavelet coefficients of the corrupted signal onto a theoretical distribution previously computed on data available without noise, see [24, 25]. Beyond denoising, empirical and theoretical distributions of wavelet coefficients have also been used for Bayesian estimation and MAP shrinkage rules, texture analysis and synthesis, as well as hypothesis testing and model selection in the presence of scaling laws, see e.g. [73, 33, 4].

However, a major difficulty in the construction of realistic wavelet models lies in the determination of the dependencies between wavelet coefficients, see [16] for the determination of the dependency between pairs of nearby wavelet coefficients in images. This issue has become even more relevant with the recent proliferation of models supplied by IA algorithms, whose mathematical properties are not well understood; determining whether they display realistic dependency structures has thus become a major concern, see e.g. [60]. This problem cannot be addressed in full generality; nonetheless, various models based on wavelet coefficients have been proposed, which make specific assumptions on the dependencies between coefficients.

Cascade-type models lie at one end of the spectrum. They originate in turbulence, where the cascade assumption is justified by a heuristic argument proposed by Richardson to explain energy dissipation at small scales, see [30] and references therein, as well as [56] for the related model of fractional Brownian motion in multifractal time. These models were subsequently tailored to the wavelet framework, see [8, 79, 26], and they imply strong correlations between the magnitudes of neighbouring wavelet coefficients. A related model is provided by multifractal random walks (MRW), which are used in mathematical finance and present the advantage of having Gaussian marginals [10, 11]. At the other end of the spectrum are random wavelet series (RWS), which provide models where the statistics of wavelet coefficients can be freely prescribed at each scale, and the coefficients are then all drawn independently according to these distributions, both across scales and spatial positions. Intermediate models include Gaussian scale mixture models and compound Poisson or Lévy-based wavelet constructions, which allow for sparse representations and intermittency effects while retaining partial statistical structure, see e.g. [78, 13, 38].

One way to assess the relevance of such models is through the estimation of nonparametric characteristics such as those supplied by multifractal analysis. State-of-the-art techniques in this area are based on wavelet expansions, and one of their byproducts has been to highlight the crucial role played by *wavelet leaders* (defined as local suprema of wavelet coefficients) in the estimation of the multifractal spectrum of data [40, 3]. A key property of these quantities is that they make it possible to account for the clustering of large wavelet coefficients without assuming any a priori model. Indeed, permuting the locations of the wavelet coefficients at a given scale does not modify their distribution but can completely alter the distribution of wavelet leaders [41]. Therefore, wavelet leaders encapsulate information on the spatial correlations between large coefficients, and comparing their distributions with those of the coefficients themselves offers a way to validate models exhibiting various types of dependency structures. In particular, a general conclusion drawn from the determination of the Legendre multifractal spectrum of many mathematical models is the following: models with no dependency between coefficients yield spectra that increase until they reach their maximum and then fall abruptly to $-\infty$, whereas the spectra of cascade-type models exhibit a decreasing branch after the location of their maximum. Related extreme multifractal behaviors are also exhibited by lacunary series supported on Cantor-type sets, which provide classical benchmark constructions with strong sparsity and highly erratic pointwise regularity, see e.g. [37]. These observations raise the question: *How is the multifractal analysis of random wavelet series affected when the (very strong) assumption of full independence of wavelet coefficients is relaxed?*

Another motivation for investigating this question is that this independence assumption is unrealistic in applications, as demonstrated by the aforementioned studies on correlations between pairs of wavelet coefficients. Nonetheless, this assumption is often found reasonable for wavelet coefficients at a given scale, since, in contrast with pointwise values, correlations decay rapidly with the distance between the supports of the corresponding wavelets, see e.g. [29], but it is much less realistic for interscale correlations. Indeed, signals exhibiting local singularities typically display large wavelet coefficients at multiple scales and at the same spatial location. This motivates the study of models in which the independence assumption across scales is dropped, while independence at a given scale is preserved.

Beyond these modeling issues, the present work adopts a sample-path point of view, which is natural in signal and image processing applications, where observed data are interpreted as realizations of stochastic processes. From this perspective, one is interested in describing both the global and pointwise Hölder regularity properties of typical sample paths, as is customary in the multifractal analysis of stochastic processes [40, 3]. The pointwise regularity of a generic realization may exhibit strong spatial fluctuations, a phenomenon that plays a central role in multifractal analysis. In the case of random wavelet series with sufficiently heterogeneous coefficient magnitudes, this leads to a highly erratic behavior: although the Hölder exponent at any fixed point almost surely takes a constant value, the collection of pointwise exponents attained along a single sample path typically fills a whole interval, see [9]. Such erratic behaviors are not specific to wavelet-based models and also arise, for instance, in Lévy processes [38].

In this article, we extend the classical random wavelet series framework introduced in [9] by allowing dependencies between wavelet coefficients across scales, while preserving independence at each fixed scale. This more general model, referred to as semi-dependent random wavelet series (SDRWS), is motivated by applications involving stochastic processes with isolated singularities, which naturally exhibit strong correlations between wavelet coefficients at different scales but at the same spatial location. We investigate how this relaxed dependency structure impacts both pointwise and global regularity properties.

A second main objective of this work is the estimation of the parameter $H_X^{\min}$, which characterizes the uniform Hölder regularity of the process X and provides a global bound on the smoothness of its sample paths. Although often overlooked, this parameter plays a fundamental role in multifractal analysis. Its wavelet characterization, rooted in the classical theory of function spaces [57], underlies several estimation procedures based on the decay of wavelet coefficients [74, 76, 29, 2], and explains its importance in applications requiring global regularity control. In particular, $H_X^{\min}$ has been used for model selection in stochastic modeling and texture analysis [80], as well as for hypothesis testing in the presence of scaling laws or long-range dependence [4]. The present contribution complements recent developments on the estimation of multifractal parameters, such as log-cumulants [62], by providing a detailed analysis of the uniform regularity exponent in both independent and semi-dependent random wavelet series models, relying in particular on wavelet leaders [51, 81].

In line with these objectives, our main contributions are as follows:

- We refine the uniform modulus of continuity result established by [42] in the RWS framework (Proposition 2.3), and extend it to settings with dependent coefficients, both in the general SDRWS model (Proposition 2.4) and in the specific case of generalized Gaussian mixture laws (Corollary 2.5). We further investigate the pointwise modulus of continuity in both RWS and SDRWS settings (Proposition 2.8).
- We determine the multifractal spectrum of SDRWS (Corollary **??**) as a consequence of a block ubiquity theorem (Theorem 3.2). This general result is then specified for generalized Gaussian mixture models (Proposition 3.4), for which we also compute the almost-everywhere modulus of continuity (Proposition 3.5).
- We propose a statistical procedure for estimating the uniform regularity exponent $H^{\min}$, together with an explicit confidence interval (Theorem 4.3), and validate its performance empirically.

The paper is organized as follows. Section 2.1 recalls the construction of RWS, specifies the model under study, and introduces the SDRWS framework. Section 2.2 establishes the sharp uniform modulus of continuity, and Section 2.3 links it to the maximal pointwise irregularity. The multifractal analysis is revisited in Section 3.1 and extended to the SDRWS setting in Section 3.2, yielding a simple multifractality criterion. Section 3.3 and Section 3.4 address the case of generalized Gaussian mixtures and their almost-everywhere modulus of continuity. Finally, Section 4 presents the estimation procedure for $H^{\min}$ and its confidence interval.

HH: réfléchir à des exemples de modèles semi-dépendants

2 Random wavelet series and semi-dependent random wavelet series

2.1 The model

Random wavelet series (RWS) were introduced in [9] in the one-variable periodic setting, where their pointwise regularity properties were studied and their multifractal spectrum was determined. Its relevance in statistical modelling has been investigated, see [31, 32, 53, 69], and it has been used in various settings, incuding turbulence [49]. Recently, the multifractal analysis based on the Hölder exponent performed in [9] has been extended to the p-exponent in [27], and sharp results concerning the global Besov regularity of RWS were recently established in [34]. In this work, we consider the more general case of random fields, that is, non-periodic processes defined on $\mathbb{R}^d$; this setting is more relevant for applications in image modeling and processing [6, 55]. The reader can easily check that these modifications have no consequence for the results from [9] that we will use. Conversely, in some papers (see [7, 45, 50]) dealing with nonparametric estimation in a regression setting via a discrete wavelet transform, periodized wavelet bases are considered; one can easily check that the results we obtain can be readily adapted to that setting.

2.1.1 RWS and SDRWS

Random wavelet series are formally defined as follows: A smooth orthonormal wavelet basis of $L^2(\mathbb{R}^d)$ is of the form

$$\varphi(x-k) \text{ and } 2^{dj/2}\psi^{(i)}(2^jx-k), \quad x \in \mathbb{R}^d,\ j \geq 0,\ k \in \mathbb{Z}^d, i \in [2^d-1],$$

where $[n] := \{1, \dots, n\}$, φ and the $\psi^{(i)}$ are smooth functions with fast decay (the required smoothness for a given result can easily be tracked if needed). We will use the notations

$$\varphi_k(x) = \varphi(x-k) \quad \text{ and } \quad \psi^{(i)}_{j,k}(x) = \psi^{(i)}(2^jx-k).$$

Definition 2.1. *A* semi-dependent random wavelet series *(SDRWS) $X = (X_x)_{x\in\mathbb{R}^d}$ associated with a given orthonormal wavelet basis is a stochastic process such that, for each j, the wavelet coefficients $C^{(i)}_{j,k}$ of the random variable*

$$X_x = \sum_{k\in\mathbb{Z}^d} C_k\ \varphi_k(x) + \sum_{i\in[2^d-1],\ j\geq 0,\ k\in\mathbb{Z}^d} C^{(i)}_{j,k}\ \psi^{(i)}_{j,k}(x) \tag{1}$$

are independent and share a common law μ_j.
The stochastic process X is a random wavelet series *(RWS) if, additionally, the $C^{(i)}_{j,k}$ are all independent random variables.*

Remark 1. Since we are interested in regularity properties of sample paths, we make no assumption on the C_k which yield a smooth contribution to (1), and we do not mention this component in the following.
As regards the assumptions on the $C^{(i)}_{j,k}$, the RWS model is completely specified whereas the SDRWS model is not: A key advantage of this model is that no assumptions are made on the dependencies between wavelet coefficients located at different scales.

2.1.2 A key example: random wavelet mixtures (RWM) and their generalization

In both independent and semi-dependent settings, we will focus on the particular case where, for each $j \geq 0$ and $i \in [2^d - 1]$, the law of $C_{j,k}^{(i)}$ is a mixture process defined as follows. Let Y be a real random variable whose law μ has a density G with respect to the Lebesgue measure on $\mathbb{R}$. We make the following *wavelet density assumptions* on a triple $(G, (p_j)_{j\geq 0}, (C_j)_{j\geq 0})$, which involve the density G as well as the non-negative sequences $(p_j)_{j\geq 0}$ and $(C_j)_{j\geq 0}$ characterizing the sparsity and regularity of the model:

Assumption 1. *1. G is continuous in a neighbourhood of 0.*

2. $G(0) > 0$.

3. The mapping $A \mapsto \mathbb{P}(|Y| \geq A)$ has fast decay.

4. $(p_j)_{j\geq 0}$ is a nonnegative sequence such that

$$\lim_{j\to+\infty} p_j = 0 \quad \textit{and} \quad \limsup_{j\to+\infty} 2^j p_j = +\infty. \tag{2}$$

5. $(C_j)_{j\geq 0}$ is a nonnegative sequence such that

$$\exists C, \varepsilon > 0, \ \forall j \geq 0, \qquad C_j \leq C 2^{-\varepsilon j}. \tag{3}$$

Definition 2.2. *Let $(G, (p_j)_{j\geq 0}, (C_j)_{j\geq 0})$ be a triple satisfying Assumption 1. A* semi-dependent random wavelet mixture *(SDRWM) of parameters $(G, (p_j)_{j\geq 0}, (C_{j,k}^{(i)})_{j\geq 0, k\in\mathbb{Z}^d, i\in[d]})$ is a SDRWS such that the law of its wavelet coefficients $C_{j,k}^{(i)}$ satisfy*

*(i) **Sparsity:** $\forall j \geq 0$, $\forall i \in [d]$, $\forall k \in \mathbb{Z}^d$, $\mathbb{P}(C_{j,k}^{(i)} = 0) = 1 - p_j$;*

*(ii) **Amplitude distribution (conditional):** Conditionally on $C_{j,k}^{(i)} \neq 0$, $C_{j,k}^{(i)} \overset{\mathcal{L}}{=} C_j Y$ where the law of Y has density G.*

A random wavelet mixture (RWM) corresponds to the particular case where the SDRWM actually is a RWS (i.e. all wavelet coefficients of the process are independent). The law of $C_{j,k}^{(i)}$ and the law of $C_j Y$ will be denoted by ν_j and μ_j, respectively.

Remark 2. In practice, the constant C_j may depend on the wavelet index (i). Indeed, in the two-dimensional case, the three wavelets used are tensor products of the univariate functions φ and ψ:

$$\psi^{(1)}(x, y) = \psi(x)\varphi(y), \quad \psi^{(2)}(x, y) = \varphi(x)\psi(y), \text{ and } \quad \psi^{(3)}(x, y) = \psi(x)\psi(y).$$

Therefore the two first ones display cancellation in one direction only (respectively the first and the second variable) whereas the third one displays cancellation in both variables. It follows that, in the case of anisotropic textures, the statistics of the coefficients on the first and second wavelets may differ; and, in all cases, statistics that are more peaked at the origin are observed for the third wavelet because of the extra cancellation; see [55]. For notational simplicity, we do not make this possible dependence explicit; adapting our results to account for it is straightforward.

Although these processes are defined on $\mathbb{R}^d$, we restrict our analysis to the cube $K = [0,1]^d$, which entails no loss of generality with respect to local regularity properties. We work with a compactly supported wavelet basis, which ensures that for sufficiently large j, the wavelets affecting the process on K have indices satisfying $k \cdot 2^{-j} \in [0,1)^d$. Accordingly, we will focus on the corresponding

$$N_j = (2^d - 1)\, 2^{dj} \tag{4}$$

wavelet coefficients in what follows.

2.1.3 Discussion on the model and related works

Connection with lacunary wavelet series The parameters p_j quantify the sparsity of the series since

$$p_j = \mathbb{E}[\{k : \ C^{(i)}_{j,k} \neq 0\}^\sharp].$$

The first condition in (2) means that **most** wavelet coefficients of X_t vanish, and the second one implies that the series is not **too sparse**, i.e. that the number of nonvanishing wavelet coefficients tends to $+\infty$ as j tends to $+\infty$. This model is consistent with certain Bayesian wavelet techniques for nonparametric regression (see for instance [77] and the references therein) where the prior on the law ν_j takes the form

$$p_j \mu_j + (1 - p_j)\delta_0,$$

where δ_0 is a Dirac mass at 0. The hyperparameters p_j and the ones involved in the definition of μ_j, need to be appropriately specified. A choice which leads to multifractal sample paths is

$$p_j = 2^{(\eta - d)j} \quad \text{ for } \quad \eta \in (0,d). \tag{5}$$

A model where this assumption was made is supplied by *lacunary wavelet series*, see [39]; the particular case investigated corresponded to the following choice: there are $\sim 2^{\eta j}$ non-vanishing wavelet coefficients drawn at random in $[0,1]^d$ of size $2^{-\alpha j}$ (where $\alpha > 0$).

Link with uniform regularity We will see that the sequence $(C_j)_{j \geq 0}$ (essentially) quantifies the uniform regularity of the sample paths of the process X. A typical choice is $C_j = 2^{-\gamma j}$ for a $\gamma > 0$. Assumption (3), together with the fast decay assumption on $\mathbb{P}(|Y| \geq A)$, imply that the process X is well defined and has some uniform regularity: for any $\varepsilon' < \varepsilon$, a.s. the sample paths of X locally belong to the Hölder space $C^{\varepsilon'}_{\text{loc}}$ see [9]; this result will be sharpened in Proposition 2.3 below.

Density function Considerable attention has been devoted to analyzing the marginals of wavelet coefficients of images. They typically exhibit heavy-tailed distributions: most are small, while a few take large values, especially when $\alpha < 2$, resulting in marginal distributions with heavier tails than the Gaussian [28, 54, 73]. Therefore, we focus on the case where the probability density G is a generalized Gaussian distribution, denoted by $\mathrm{GG}(0,1,\alpha)$, such that

$$G_\alpha(x) = \frac{\alpha}{2\Gamma(\frac{1}{\alpha})} e^{-|x|^\alpha}, \tag{6}$$

where Γ denotes the gamma function, the normalizing constant being such that $\int G(x)dx = 1$; indeed, the corresponding statistics of wavelet coefficients are commonly met in signal and image processing. This non-Gaussianity has motivated various Bayesian models, including mixtures of Gaussians [1, 19] and broader classes of Gaussian scale mixtures for image denoising [65, 71, 72, 75].
Wavelet decompositions are especially suited for natural images, which often consist of smooth regions interrupted by edges, so-called "cartoon-type" structures [23, 55]. Smooth areas yield near-zero coefficients, while edges generate sparse, high-amplitude ones. This structure explains the sharp central peaks, heavy tails, and inter-scale correlations observed in wavelet coefficient histograms [73, 65].
Although non-Gaussian image statistics have long been observed, more recent models capture these effects via linear predictors with structured uncertainty [16], sparse coding [63], or hidden Markov models [21]. Deep learning approaches have further improved modeling of inter-scale dependencies, as in [46], where a CNN-based method leverages a stationary local Markov model across scales.
These developments motivate going beyond the RWS model by relaxing the assumption of inter-scale independence, which is often unrealistic. In this work, however, we focus on analyzing the simplest consistent models - those with independence at fixed scale (RWS) or across scales (SDWS) - as a baseline for further investigation.

2.2 Uniform regularity of RWS and SDRWS

A first basic information concerning the uniform regularity of a function defined on $\mathbb{R}^d$ or of the sample paths of a stochastic process is supplied by its *uniform Hölder exponent*:

$$H_f^{\min} = \sup\{\alpha : \ f \in C^\alpha_{\text{loc}}(\mathbb{R}^d)\}. \tag{7}$$

For instance, in the case of fractional Brownian motion, the uniform Hölder exponent coincides with the Hurst exponent. However, this notion does not yield sharp estimates for the modulus of continuity. In particular, for Brownian motion, knowing that $H_B^{\min} = 1/2$ does not provide any information on the logarithmic corrections appearing in its uniform modulus of continuity. As mentioned in the introduction, the latter is given by $h \mapsto \sqrt{2|h|\log(\log(1/h))}$ (see [52]).
The uniform Hölder exponent of RWS has been determined in [9]; see also (27) below. One of our aims here is to refine this result and obtain sharp estimates for the uniform regularity of RWS and SDRWS. To this end, we will rely on the following general framework (see Sec. 1.1 of [42]).

Definition 2.3. *A* modulus of continuity *is a positive non-decreasing function $\theta : \mathbb{R}^+ \to \mathbb{R}^+$ satisfying*

$$\exists C > 0 : \quad \begin{cases} \qquad\qquad \theta(0) & = 0 \\ \forall h > 0, \quad \theta(2h) & \leq C\theta(h). \end{cases}$$

Additionally, θ is an admissible *modulus of continuity if it satisfies*

$$\exists N \in \mathbb{N},\ \exists C > 0,\ \forall J \geq 0, \quad \begin{cases} \displaystyle\sum_{j \geq J} 2^{Nj}\theta(2^{-j}) & \leq C 2^{NJ}\theta(2^{-J}) \\ \displaystyle\sum_{j \leq J} 2^{(N+1)j}\theta(2^{-j}) & \leq C 2^{(N+1)J}\theta(2^{-J}). \end{cases}$$

the corresponding N is referred to as the order of the modulus.

The relevance of admissible moduli of continuity lies in the fact that they can be characterized exactly through conditions on the absolute values of wavelet coefficients. The most commonly encountered moduli are of the form, for h small enough,

$$\theta(h) = h^{\alpha} |\log(h)|^{\beta} (\log(|\log(h)|))^{\gamma} \quad \text{for } \alpha > 0, \quad \beta, \gamma \in \mathbb{R},$$

which are admissible if and only if $\alpha \notin \mathbb{N}$ (in which case $[\alpha]$ denotes the order of the modulus).
We recall below the definition of the uniform modulus of continuity of a function, which serves to characterize its uniform regularity; see [42].

Definition 2.4. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a locally bounded function, and let θ be an admissible modulus of continuity of order $N \in \mathbb{N}$; θ is a* uniform modulus of continuity *for f if:*

1. $f \in C^N(\mathbb{R}^d)$,
2. *For every multi-index $i = (i_1, \ldots, i_d)$ with $|i| = N$, the partial derivatives $f^{(i)}$ satisfy*

$$\exists C > 0, \ \forall i \in [d], \forall x, y \in \mathbb{R}^d \text{ such that } x \neq y, \qquad |f^{(i)}(x) - f^{(i)}(y)| \leq C \frac{\theta(\|x-y\|)}{\|x-y\|^N}. \tag{8}$$

The function θ is a sharp mudulus of continuity if, in addition, no function ω which is a $o(\theta)$ is a modulus of continuity.

Remark 3. If $H_f^{\min} > 0$, then it can be rewritten

$$H_f^{\min} = \sup\{\alpha : \ h \mapsto h^{\alpha} \text{ locally is a uniform modulus of continuity of } f\}.$$

In the present study, moduli of continuity are defined up to a multiplicative constant that we will not track.

The following characterization is proved in [42].

Proposition 2.1. *(Prop. 1.2 of [42]) Let $f : \mathbb{R}^d \to \mathbb{R}$ be a locally bounded function, and let $(C_k, C_{j,k}^{(i)})_{j,k,i}$ be its coefficients in a smooth wavelet basis $(\varphi_k, \psi_{j,k}^{(i)})_{j,k,i}$. If θ is a uniform modulus of continuity for f, then the wavelet coefficients of f satisfy*

$$\exists C > 0, \quad \forall k \in \mathbb{Z}^d, \qquad |C_k| \leq C. \tag{9}$$

and

$$\exists C > 0, \quad \forall (j,k,i) \in \mathbb{N} \times \mathbb{Z}^d \times [d], \qquad |C_{j,k}^{(i)}| \leq C\theta(2^{-j}). \tag{10}$$

Conversely, assume that the wavelet coefficients of f satisfy (9) and (10); then:

1. *If θ is a admissible modulus, then (8) holds and θ is a uniform modulus of continuity for f;*

2. *Otherwise,* (8) *is replaced by*

$$\exists C > 0,\ \forall (x,y) \in \mathbb{R}^d, \quad |f^{(i)}(x) - f^{(i)}(y)| \leq C \frac{\theta(\|x-y\|)}{\|x-y\|^N}(1 + |\log(\|x-y\|)|), \tag{11}$$

and these results are optimal.

It follows that the uniform modulus of continuity of a function f is determined by the size of its largest wavelet coefficient at each scale. In the general framework of RWS, and without additional assumptions on the distribution of the wavelet coefficients, the uniform Hölder exponent $H_f^{\min}$ was established in [9]. We now refine this result by deriving a sharp uniform modulus of continuity.
To state the result, we introduce the functions f_j defined for $j \geq 0$ and $a \geq 0$ by

$$\forall j \geq 0,\ \forall a \geq 0, \qquad f_j(a) = \mathbb{P}(|C_{j,k}^{(i)}| > a), \tag{12}$$

that is, $1 - f_j$ is the cumulative distribution function of the random variable $|C_{j,k}^{(i)}|$.

Lemma 2.2. *Let X be a RWS defined on $\mathbb{R}^d$. Let $(f_j)_{j\geq 0}$ be a sequence of functions satisfying* (12) *and assume that $(a_j)_{j\geq 0}$ is a positive decreasing sequence such that*

$$f_j(a_j) = o(2^{-dj}). \tag{13}$$

- *If $\sum_j 2^{dj} f_j(a_j) < \infty$, then a.s., for j large enough, $\sup_k |C_{j,k}^{(i)}| \leq a_j$;*
- *If $\sum_j 2^{dj} f_j(a_j) = +\infty$, then a.s., for j large enough, there exists an infinite number of values of j such that $\sup_k |C_{j,k}^{(i)}| \geq a_j$.*

Proof. Since the $C \cdot 2^{dj}$ coefficients $C_{j,k}^{(i)}$ for $k \cdot 2^{-j} \in [0,1)$ are independent random variables, it follows that the probability that one of them is larger than a_j is

$$P_j = 1 - (1 - f_j(a_j))^{C\cdot 2^{dj}} = C \cdot 2^{dj} f_j(a_j) + O\left((2^{dj} f_j(a_j))^2\right).$$

Thus, $\sum_j P_j$ is finite if and only if $\sum_j 2^{dj} f_j(a_j) < \infty$. Using the independence of the wavelet coefficients, one can apply the Borel-Cantelli lemma, and Lemma 2.2 follows. □

The following result is a consequence of Proposition 2.1 and Lemma 2.2.

Proposition 2.3. *Let X be a RWS on $\mathbb{R}^d$. Let $(f_j)_{j\geq 0}$ and $(a_j)_{j\geq 0}$ be sequences satisfying* (12) *and* (13)*. Define the function $g : \mathbb{R}^+ \to \mathbb{R}^+$ by*

$$\begin{cases} g(2^{-j}) = a_j \\ g \text{ is constant on } [2^{-j}, 2\cdot 2^{-j}) \end{cases}. \tag{14}$$

- *If $\sum_j 2^{dj} f_j(a_j) < \infty$ and g is an admissible modulus of continuity, then a.s. the function $h \mapsto g(h)$ is a uniform modulus of continuity for X.*
- *If $\sum_j 2^{dj} f_j(a_j) < \infty$ and g is not an admissible modulus of continuity, then a.s. the function $h \mapsto g(h)(1 + |\log h|)$ is a uniform modulus of continuity for X.*

- *If* $\sum_j 2^{dj} f_j(a_j) = +\infty$, *then any function* $\tilde{g}$ *such that* $\tilde{g}(h) = o(g(h))$ *as* $h \to 0$ *is almost surely not a uniform modulus of continuity for* X.

Proof. In the first case, Lemma 2.2 shows that the condition $\sum_j 2^{dj} f_j(a_j) < \infty$ implies that, almost surely, for all sufficiently large j, $\sup_k |C^{(i)}_{j,k}| \leq a_j$. Using the definition (14), this can be rewritten as

$$\sup_k |C^{(i)}_{j,k}| \leq g(2^{-j}).$$

This allows us to apply the second part of Proposition 2.1. If the modulus g is admissible, then Point 1) of Proposition 2.1 gives the first statement of Proposition 2.3. Otherwise, Point 2) yields the second statement. As for the third statement of Proposition 2.3, it follows by a similar argument, this time relying on the first part of Proposition 2.1 together with the second statement of Lemma 2.2. $\square$

We now consider the setting of SDRWS, i.e. **we drop the assumption that wavelet coefficients located at different scales are independent**. Consider a sequence $(f_j)_{j\geq 0}$ satisfying (12). Fix a constant $C > 0$, and define the sequence $(b_j)_{j\geq 0}$ by

$$f_j(b_j) = C_j 2^{-dj}, \tag{15}$$

as well as the function $\bar{g} : \mathbb{R}^+ \to \mathbb{R}^+$ by

$$\bar{g}(2^{-j}) = b_j, \quad \text{and extend it as a constant on } [2^{-j}, 2 \cdot 2^{-j}). \tag{16}$$

The following result follows from Proposition 2.1 and Lemma 2.2.

Proposition 2.4. *Let* X *be a SDRWS defined on* $\mathbb{R}^d$. *Let* $(f_j)_{j\geq 0}$ *and* $(a_j)_{j\geq 0}$ *be sequences satisfying* (12) *and* (13) *and let* g *and* $\bar{g}$ *be defined respectively by* (14) *and* (16).

- *If* $\sum_j 2^{dj} f_j(a_j) < \infty$ *and* g *is an admissible modulus of continuity, then a.s. the function* $h \mapsto g(h)$ *is a uniform modulus of continuity for* X.

- *If* $\sum_j 2^{dj} f_j(a_j) < \infty$ *and* g *is not an admissible modulus of continuity, then a.s. the function* $h \mapsto g(h)(1 + |\log(h)|)$ *is a uniform modulus of continuity for* X.

- *If* $\sum_j 2^{dj} f_j(a_j) = +\infty$, *then any function* $\tilde{g} : \mathbb{R}^+ \to \mathbb{R}^+$ *such that* $\tilde{g}(h) = o(\bar{g}(h))$ *as* $h \to 0$ *is almost surely not a uniform modulus of continuity for* X.

Proof. We first observe that the first two statements of Proposition 2.3 remain valid, as they rely only on the direct part of the Borel-Cantelli lemma and therefore do not require independence.
We now turn to establishing irregularity results in this more general setting. As before, the key point is to identify, as $j \to \infty$, a sequence of "large" wavelet coefficients. For each fixed scale j, let Q_j denote the probability that all $C \cdot 2^{dj}$ wavelet coefficients at scale j are bounded by b_j. By independence, we have

$$\log(Q_j) = C' \cdot 2^{dj} \log(1 - f_j(b_j)) = -C'' j(1 + o(1)),$$

for some constant $C'' > 0$. Hence, $Q_j = e^{-C'' j(1+o(1))}$. Applying the direct part of the Borel-Cantelli lemma, we conclude that almost surely, for all sufficiently large j, at least one coefficient $c_{j,k,i}$ exceeds b_j. Since $\bar{g}$ is defined by $\bar{g}(2^{-j}) = b_j$ and extended as a constant on the interval $[2^{-j}, 2 \cdot 2^{-j})$, with $(b_j)_{j \geq 0}$ given by (15), the claim follows from the same argument as before. $\square$

Remark 4. The slight loss between the converse parts of Propositions 2.3 and 2.4 stems from the fact that we cannot apply the Borel-Cantelli lemma simultaneously to **all** wavelet coefficients; instead, a scale-by-scale argument is required. Corollary 2.5 below illustrates this phenomenon by comparing the corresponding moduli of continuity in the case where the distributions of the wavelet coefficients are generalized Gaussian mixtures.

Let us now examine how these conditions specialize in the case of the mixture models introduced in Definition 2.2.

Corollary 2.5. *Let X be a RWM with parameters $C_j = 2^{-\gamma j}$ and $p_j \geq 2^{(\varepsilon-1)j}$ for some $\varepsilon > 0$. If $\gamma \notin \mathbb{N}$, then a sharp uniform modulus of continuity of X almost surely is*

$$h \mapsto g(h) = h^{\gamma} |\log h|^{1/\alpha}. \tag{17}$$

If $\gamma \in \mathbb{N}$, then

$$h \mapsto g(h) = h^{\gamma} |\log h|^{1+1/\alpha} \tag{18}$$

*is a uniform modulus of continuity for X, and any function that is $o(h^{\gamma} |\log h|^{1/\alpha})$ is **not** a uniform modulus of continuity for X.*
*Now let X be a SDRWM with the same assumptions on C_j and p_j. Then the function in (18) is a uniform modulus of continuity for X, and the function $h \mapsto h^{\gamma} |\log h|^{(1/\alpha)-1}$ is **not** a uniform modulus of continuity for X.*

Proof. Since, for a large,

$$F_{\alpha}(a) := \int_a^{\infty} e^{-x^{\alpha}} dx = \frac{e^{-a^{\alpha}}}{\alpha a^{\alpha-1}} (1 + o(1)), \tag{19}$$

it follows that

$$\mathbb{P}(|C_{j,k}^{(i)}| \geq a_j) := f_j(a_j) = C p_j \frac{e^{-(a_j/C_j)^{\alpha}}}{(a_j/C_j)^{\alpha-1}} (1 + o(1)),$$

and the uniform regularity of X will be given according to the convergence or divergence of the series

$$\sum_j 2^{dj} \, p_j \frac{e^{-(a_j/C_j)^{\alpha}}}{(a_j/C_j)^{\alpha-1}}. \tag{20}$$

Since $C_j = 2^{-\gamma j}$ with $\gamma > 0$, Theorem 1 of [9] yields that the uniform Hölder exponent $H_X^{\min}$ of the corresponding RWS is γ so that we expect a slight correction of a uniform modulus of the form $h \mapsto \theta(h) = h^{\gamma}$. For this reason, we take a_j under the form $a_j = a 2^{-\gamma j} j^b$. The general term of the series (20) boils down to

$$2^{dj} \, p_j \frac{e^{-a^{\alpha} j^{b\alpha}}}{a^{\alpha-1} j^{b(\alpha-1)}}.$$

We assume that $p_j \geq 2^{(\varepsilon-1)j}$ for an $\varepsilon > 0$ (which covers the specific case we considered in (5)). Then we pick $b = 1/\alpha$: For a large enough, the series is convergent, whereas for a small it is divergent. These choices yield moduli of continuity of the form $h \mapsto g(h) = Ch^\gamma |\log h|^{1/\alpha}$. Hence Corollary 2.5 follows from Corollaries 2.3 and 2.4 applied to these specific settings. $\square$

We can contrast these results with those obtained in the case of Gaussian processes that can be expanded into wavelet series with *pure* (in contrast to *mixtures*) Gaussian statistics such as fractional Brownian motion (fBm).
Indeed, the fBm with Hurst index $H \in (0,1)$ can be expanded (see [58]) as

$$B_t^H = \sum_{j\in\mathbb{N}} \sum_{k\in\mathbb{Z}} 2^{-Hj} \xi_{j,k} \psi_{H+1/2}(2^j t - k) + R_t =: Z_t^H + R_t,$$

where R is a smooth process and $(\xi_{j,k},\ j \in \mathbb{N}, k \in \mathbb{Z})$ is a sequence of i.i.d. standard normal variables, and the $\psi_{H+1/2}$ is a biorthogonal wavelet basis. The uniform modulus of continuity of the RWS Z_t is $h \mapsto |h|^H |\log h|^{1/2}$. We retrieve the uniform modulus of continuity of the fBm as the specific case $p_j = 1$ (no mixture), with $C_{j,k} = \xi_{j,k}$, $C_j = 2^{-\gamma j} = 2^{-Hj}$ and $\alpha = 2$.

2.3 Most irregular points of RWS and SDRWS

The technique which led to the determination of the uniform regularity of RWS and SDRWS also yields their pointwise irregularity in a sharp way at the most irregular points. In order to state such results, we first recall the notion of pointwise modulus of continuity.

Definition 2.5. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a locally bounded function, and let θ be a modulus of continuity of order N. We say that θ is a modulus of continuity of f at x_0 if there exists a polynomial P_{x_0} of degree at most N and constants $C > 0$ and $\delta > 0$ such that, for all x with $|x - x_0| < \delta$,*

$$|f(x) - P(x - x_0)| \leq C\theta(|x - x_0|). \tag{21}$$

The function θ is a sharp mudulus of continuity at x_0 if, in addition, no function ω which is a $o(\theta)$ is a modulus of continuity at x_0.

We will use the following notation. To index wavelets and their coefficients, we will interchangeably use dyadic cubes

$$\lambda_{j,k} = \left[\frac{k_1}{2^j}, \frac{k_1+1}{2^j}\right] \times \cdots \times \left[\frac{k_d}{2^j}, \frac{k_d+1}{2^j}\right] \text{ where } k = k_1, \cdots, k_d \tag{22}$$

and we will write $C_\lambda^{(i)}$ or $C_{j,k}^{(i)}$ without distinction. For a dyadic cube λ, let 3λ denote the cube with the same center and three times the side length.
If $f : \mathbb{R}^d \to \mathbb{R}$ is locally bounded, the *wavelet leaders* of f are defined by

$$\ell_\lambda = \sup_{\lambda' \subset 3\lambda} |C_{\lambda'}^{(i)}|.$$

For $x \in \mathbb{R}^d$, denote by $\lambda_j(x)$ the unique dyadic cube of side length 2^{-j} containing x. A key fact is that the pointwise modulus of continuity can be recovered from the wavelet leaders, as stated in the following result; see [40, 42].

Theorem 2.6. *Let $f \in C^{\varepsilon}_{\text{loc}}(\mathbb{R}^d)$ for an $\varepsilon > 0$ and let $\psi^{(i)}_{j,k}$ be a smooth wavelet basis. If θ is a modulus of continuity of f at x_0, then the wavelet leaders of f satisfy*

$$\exists C > 0, \quad \forall j, k, \qquad \ell_{\lambda_j(x_0)} \leq C\theta(2^{-j}). \tag{23}$$

Conversely, assume that the wavelet coefficients of f satisfy (23); then

$$h \mapsto \theta(h)(1 + |\log(h)|)$$

is a modulus of continuity of f at x_0.

As a consequence of Lemma 2.2, let us prove the following result.

Proposition 2.7. *Let X be a RWS. If its uniform modulus of continuity is admissible, then there exists a dense set of points where the pointwise modulus of continuity of X is equal to its uniform modulus of continuity.*

Remark 5. Before proving this result, let us motivate its statement; indeed, a common belief is that the uniform modulus of continuity coincides with the largest pointwise modulus of continuity met in the data, in which case this result would not be relevant. However, it is not the case, as shown by *chirps* of the form $x \mapsto |x|^{\alpha} \sin(1/|x|^{\beta})$ for $\alpha, \beta > 0$, see [42] for a detailed analysis this question.

Proof of Proposition 2.7. Let λ be an arbitrary dyadic cube contained in $[0,1)^d$. The proof of the second statement of Lemma 2.2 can be carried out inside λ instead of $[0,1)^d$. In particular, it yields a sub-cube $\lambda' \subset \lambda$ such that $|C_{\lambda'}| \geq a_j$. Iterating this argument within λ', we construct a sequence $(\lambda_n)_{n\geq 0}$ of dyadic cubes with generations $(j_n)_{n\geq 0}$ such that

$$\lambda_{n+1} \subset \lambda_n \quad \text{ and } \quad |C_{\lambda_n}| \geq a_{j_n}.$$

This nested sequence of cubes converges to a point x_0, at which Theorem 2.6 implies that the modulus of continuity of f cannot be $o(\theta(h))$. Since the pointwise modulus of continuity is always bounded above by the uniform modulus of continuity, it follows that θ is a sharp modulus of continuity at $x_0 \in \lambda$. As λ was arbitrary, the set of such points is dense in $[0,1)^d$. $\square$

We now consider the case of SDRWS. The argument for the determination of the most irregular points is the same as for RWS, except that the sequence $(a_j)_{j\geq 0}$ is replaced by the sequence $(b_j)_{j\geq 0}$ defined by (15). The following result follows.

Proposition 2.8. *Let X be a SDRWS. If its uniform modulus of continuity is admissible, then there exists a dense set of points where the function $\tilde{g}$ defined in Corollary 2.4 is not a pointwise modulus of continuity of X.*

3 Multifractal analysis

In this section we first recall the results of [9] concerning the multifractal analysis of RWS and then we show how they extend to the more general setting of SDRWS.

We begin by recalling several notions from multifractal analysis. Let $\gamma \geq 0$ and let $f : \mathbb{R}^d \to \mathbb{R}$ be a locally bounded function. The function f belongs to $C^\gamma(x_0)$ if $h \mapsto h^\gamma$ is a pointwise modulus of continuity for f at x_0.
The *Hölder exponent* of f at x_0 is

$$h_f(x_0) = \sup\left\{\gamma : \quad f \quad \text{is} \quad C^\gamma(x_0)\right\}. \tag{24}$$

The *multifractal spectrum* of f describes the size of the sets of points sharing the same Hölder exponent, the so-called *isohölder sets*

$$\mathcal{I}_f(H) = \{x : \quad h_f(x) = H\}, \tag{25}$$

see [64]. It is defined by

$$\mathcal{D}_f(H) = \dim(\mathcal{I}_f(H)),$$

where dim denotes the Hausdorff dimension (with the convention $\dim(\emptyset) = -\infty$).
The *multifractal support* of f is the set

$$\mathcal{MS}_f = \{H : \quad \mathcal{D}_f(H) \geq 0\} = \{H : \quad \mathcal{I}_f(H) \neq \emptyset\}.$$

3.1 Multifractal analysis of RWS

The a.s. multifractal spectrum of the sample paths of RWS has been determined in [9], and we now recall these results; we keep the same notations as in this article, making the necessary adjustments required by the d-variable setting; denote by $\boldsymbol{\rho_j}$ the common probability measure of the 2^{dj} random variables $X_{j,k} := -\log_2(|C^{(i)}_{j,k}|)/j$. Thus $\boldsymbol{\rho_j}$ satisfies

$$\mathbb{P}\Big(|C^{(i)}_{j,k}| \geq 2^{-\alpha j}\Big) = \boldsymbol{\rho_j}((-\infty, \alpha]).$$

Note that it follows from (12) that $\boldsymbol{\rho_j}((-\infty, \alpha]) = f_j(2^{-\alpha j})$.
We note for $\alpha \geq 0$

$$\boldsymbol{\rho}(\alpha, \varepsilon) := \limsup_{j \to +\infty} \frac{\log_2\left(2^{dj}\boldsymbol{\rho_j}([\alpha - \varepsilon, \alpha + \varepsilon])\right)}{j}$$

and the *wavelet large deviation spectrum* of the corresponding RWS is

$$\boldsymbol{\rho}(\alpha) := \inf_{\varepsilon > 0} \boldsymbol{\rho}(\alpha, \varepsilon). \tag{26}$$

As in [9], we suppose that $\boldsymbol{\rho}(\alpha)$ takes a positive value for at least one value of α. Let

$$W = \Big\{\alpha : \quad \forall \varepsilon > 0, \qquad \sum_{j \in \mathbb{N}} 2^j \boldsymbol{\rho}_j([\alpha - \varepsilon, \alpha + \varepsilon]) = +\infty\Big\}.$$

Let

$$H_X^{\max} := \left(\sup_{\alpha>0} \frac{\boldsymbol{\rho}(\alpha)}{\alpha}\right)^{-1}.$$

The assumptions made on RWS imply that $H_X^{\min} > 0$, and the uniform Hölder exponent of a RWS is positive and given by

$$H_X^{\min} := \inf_{\alpha \geq 0} W. \tag{27}$$

The following result yields the multifractal spectrum of almost every sample paths of a RWS.

Theorem 3.1. *(Theorem 2 of [9]). Let X be a random wavelet series. With probability one, the sample paths of X satisfy the following properties:*

1. *Their multifractal support is $\mathcal{S}_X = [H_X^{\min}, H_X^{\max}]$;*
2. *Their multifractal spectrum is given by*

$$\forall H \in \mathcal{S}_X, \qquad \mathcal{D}_X(H) = H \sup_{\alpha \in (0,H]} \frac{\boldsymbol{\rho}(\alpha)}{\alpha}; \tag{28}$$

3. *For almost every x,*

$$h_X(x) = H_X^{\max}. \tag{29}$$

Remark 6.
- The first statement implies that the Hölder exponent at the most irregular points is $H_X^{\min}$. Prop. 2.7, in conjunction with Theorem 2.6, sharpens this result by yielding the sharp modulus of continuity at these points. In the case of generalized Gaussian mixture models, the last statement will also be sharpened in Sec. 3.4 which will yield the almost everywhere sharp modulus of continuity.
- The function

$$H \mapsto \sup_{\alpha \in (0,H]} H \frac{\boldsymbol{\rho}(\alpha)}{\alpha}$$

is increasing on $(0, H_X^{\max}]$ and takes the value 1 for $h = H_X^{\max}$, which is in accordance with (39).

3.2 Multifractal analysis of SDRWS

Multifractal properties of RWS were first obtained as a by-product of a general framework introduced in [39], now referred to as *ubiquity methods*. Let $K = [0,1]^d$, let $S = (x_n)_{n \geq 0}$ be a sequence of points in K, and let $L = (l_n)_{n \geq 0}$ be a sequence of positive numbers with $l_n \to 0$ as $n \to \infty$. The pair (S, L) satisfies the *ubiquity condition* if almost every point of K belongs to the set

$$E = \limsup_n B(x_n, l_n). \tag{30}$$

It satisfies the *strong ubiquity condition* if **every** point of K belongs to E. Ubiquity methods provide lower bounds on the Hausdorff dimension of the sets

$$E_\eta = \limsup_n B(x_n, (l_n)^\eta). \tag{31}$$

The first results of this type were obtained in [39, 9] when $(x_n)_{n\geq 0}$ is an i.i.d. sequence equidistributed with respect to the Lebesgue measure on K. These results were used to derive the multifractal spectra of lacunary wavelet series and subsequently of RWS. In these cases, each x_n identifies the unique dyadic cube λ at generation j corresponding to the wavelet coefficient C_λ.
In such random settings, the ubiquity condition follows directly from the Borel-Cantelli lemma, while the strong ubiquity condition is ensured by almost-sure covering results for random sets; see [47] and the references therein. Since then, the ubiquity framework has been considerably extended; see, for example, [15, 14, 5] and ref. therein.
In the more general setting of SDRWS, however, the sequence $(x_n)_{n\geq 0}$ is no longer i.i.d., and the independence assumption is replaced by the weaker notion of *random block independence*, where the scale structure is emphasized.

Definition 3.1. *A random block independent dyadic sequence in K is a sequence $(\mathbf{K_j})$, each $\mathbf{K_j}$ being constituted by dyadic cubes $\lambda_{j,k} \subset K$, of generation j (i.e. of width 2^{-j}), such that for each j, the cubes $\lambda_{j,k}$ are independently drawn with the same probability p_j among the 2^{dj} dyadic subcubes of K of generation j.*

Note that no assumption is made on possible correlations between the locations of the cubes $\lambda_{j,k} \subset \mathbf{K_j}$ across different generations j.
The expectation of the number of elements of $\mathbf{K_j}$ is

$$\mathbb{E}(\mathrm{Card}(K_j)) = 2^{dj} p_j := N_j. \tag{32}$$

We impose the following exponential growth condition on the N_j: There exists a subsequence (j_n) such that

$$\exists \alpha, \beta \text{ with } 0 < \alpha < \beta < d \quad \text{such that, for } n \text{ large enough,} \quad 2^{\alpha j_n} \leq N_{j_n} \leq 2^{\beta j_n}. \tag{33}$$

We can now state the following block ubiquity result.

Proposition 3.2. *Let $(\lambda_{j,k})$ be a random block independent dyadic sequence satisfying (33). Let $L = (l_j)$ be the sequence*

$$l_j = \left(\frac{C \log(N_j)}{N_j} \right)^{1/d}, \tag{34}$$

where $C \geq 4$ is picked so that $1/l_j$ is a power of 2. Then, a.s. $\exists N$ such that, for all $\forall n \geq N$, the dyadic cubes of width l_{j_n} that include $\lambda_{j_n,k}$ for $k \in \mathbf{K_j}$ form a covering of K.

In particular, the couple constituted by the centers of the cubes $\lambda_{j,k}$ and the radii $2^d l_j$ almost surely satisfies the strong ubiquity condition.

Proof. Let j be one of the j_n. We partition the cube $K = [0,1]^d$ into dyadic subcubes of sidelength l_j as defined in (34). Let $C_{j,\ell}$ be one of these subcubes; it contains

$$m_j = 2^{dj} (l_j)^d \tag{35}$$

dyadic cubes of generation j; therefore, the probability that none of the $(\lambda_{j,k})_{k\in\mathbf{K_j}}$ is inside $C_{j,\ell}$ is

$$(1-p_j)^{m_j}.$$

Since there are $(l_j)^{-d}$ cubes $C_{j,\ell}$, it follows that the probability that **at least one** of the subcubes $C_{j,\ell}$ contains no element of $\mathbf{K_j}$ is bounded by

$$P_j = (l_j)^{-d}(1-p_j)^{m_j}.$$

We will show that $\sum P_{j_n} < \infty$ which, by the Borel-Cantelli Lemma, will imply that, for n large enough, all subcubes $C_{j_n,\ell}$ will contain at least one element of $\mathbf{K_{j_n}}$.

We now estimate P_j. We have

$$\log(P_j) = -d\log(l_j) - m_j p_j + O(m_j(p_j)^2).$$

Since the sequence l_j decays exponentially, and using (32), (34), and (35), we obtain

$$\begin{aligned}\log(P_j) &= \log(N_j) - \log(C\log(N_j)) - (l_j)^d N_j(1+o(1))\\ &= (1-C)\log(N_j) + o(\log(N_j))).\end{aligned}$$

Since $C \geq 4$, it follows from the exponential growth assumption (33) that $\sum_n P_{j_n} < \infty$. By the direct part of the Borel-Cantelli lemma (which does not require independence), we conclude that, with probability 1, for all sufficiently large n, each subcube $C_{j_n,\ell}$ contains at least one cube $\lambda_{j_n,k}$. $\square$

We will now apply this framework to the setting supplied by SDRWS. We first consider the value of the uniform Hölder exponent $H_X^{\min}$. In the case of RWS, it is determined as follows.

$$\text{Let} \quad W = \left\{ \alpha : \quad \forall \varepsilon > 0 : \quad \sum_j 2^{dj}\rho_{\mathbf{j}}[\alpha-\varepsilon,\alpha+\varepsilon] = +\infty \right\}, \text{ and} \qquad H_1^{min} = \inf W. \tag{36}$$

If X is a RWS, then a.s.

$$H_X^{\min} = H_1^{min}.$$

As regards the uniform Hölder exponent of SDRWS, the value of $H_X^{\min}$ depends on the following quantity which, may differ from H_1^{min}.

$$\text{Let} \quad \tilde{W} = \left\{ \alpha : \ \forall\varepsilon>0, \exists\eta>0, \ \exists j_n \to +\infty : \quad 2^{dj_n}\rho_{\mathbf{j_n}}[\alpha-\varepsilon,\alpha+\varepsilon] \geq 2^{\eta j_n} \right\}, \text{ and} \quad H_2^{min} = \inf \tilde{W}.$$

Clearly, $H_1^{min} \leq H_2^{min}$. We will first check the following result, which replaces Proposition 4.8 and Appendix A of [27] for RWS: If X is a SDRWS, then

$$\text{a.s.} \qquad H_1^{min} \leq H_X^{\min} \leq H_2^{min}. \tag{37}$$

Proof: The first inequality follows from the fact that the multifractal spectrum of a SDRWS is bounded by (28) (this is actually true for any wavelet series, see [9]). The second inequality follows for the fact that, by definition of H_2^{min}, there exists collections of wavelet coefficients of exponential cardinality and of size $\sim 2^{-H_2^{min} j}$, and the result follows from the block ubiquity result.

The following result, which extends Theorem 3.1 to the setting of SDRWS, is a consequence of the previous random block ubiquity property.

Theorem 3.3. *Let X be a SDRWS. With probability one, the sample paths of X satisfy the following properties:*

1. *Their multifractal support is $\mathcal{S}_X = [H_X^{\min}, H_X^{\max}]$;*
2. *Their multifractal spectrum is given by*

$$\forall H \in \mathcal{S}_X, \qquad \mathcal{D}_X(H) = H \sup_{\alpha \in (0,H]} \frac{\boldsymbol{\rho}(\alpha)}{\alpha}; \tag{38}$$

3. *For almost every x,*

$$h_X(x) = H_X^{\max}. \tag{39}$$

Remark: Note that this result leaves open the precise value of $H_X^{\min}$; the only information we have is that it takes value in the interval $[H_1^{min}, H_2^{min}]$. In other words, the multifractal spectrum of a SDRWS may differ from the corresponding RWS (with the same distributions of wavelet coefficients at each scale) only on the interval $[H_1^{min}, H_X^{\min}]$ where it takes the value 0 for RWS and it takes the value $-\infty$ for SDRWS. For $H \geq H_X^{\min}$, the spectra of both processes coincide.

We now sketch how the proof of Theorem 3.1 has to be modified in order to apply to SDRWS. We first note that, in [9], the upper bounds for the multifractal spectra in (38) are obtained in terms of large deviation spectra derived form the scale by scale distributions, and therefore are derived without specific dependence assumptions, and thus remain valid for SDRWS.

In order to obtain the lower bounds, we will show how block ubiquity techniques apply to the setting of SDRWS. We won't detail the full proof, but rather point to the locations where the scale by scale independence is used in the RWS case, and show how to replace the corresponding arguments using block ubiquity. We will actually refer to the more recent proof of [27] which fully uses the more general versions of ubiquity techniques, and also uses the notion of wavelet p-leaders, which were introduced in the meantime, and therefore it will allow for a simpler and more pedagogical derivation; note that in [27], RWS are processes indexed by $\mathbb{R}$, but the reader will easily check that the results adapt to the d-variable random fields.

The key argument for the derivation of the lower bound of the multfractal spectrum in [27] is supplied by Theorem 4.13 which yields the existence of a a gauge function with the right scaling properties. It is derived in the specific case of RWS by. a classical a.s. covering lemma of $[0, 1]$, and it is replaced here, as explained before, by Proposition 3.2. In the case of SDRWS, in order to apply this Proposition, we have to verify that, at the relevant values for H for which the supremum is attained in (38), the number of corresponding wavelet coefficients of size $\sim 2^{-Hj}$ is exponentially large, which allows to apply Proposition 3.2. But this simply follows from the fact that the function defined by the right hand side of (38) is increasing, and therefore, as soon as it is positive, it corresponds to the exponentially large case mentioned above.

3.3 Generalized Gaussian mixture models

We now check what these general results yield in the case of generalized Gaussian mixture models supplied by Definition 2.2. We make the following additional assumptions on the triple $(G, (p_j)_{j\geq 0}, (C^{(i)}_{j,k})_{j\geq 0, k\in\mathbb{Z}^d, i\in[d]})$:

Assumption 2.
- *$C_j = 2^{-\gamma j}$ for a $\gamma > 0$;*
- *$\exists \beta \in (0,1)$ and $\delta > 0$: $\qquad 2^{(\beta-1)j} \leq p_j \leq \dfrac{1}{j^\delta}$;*
- *The distribution G is a generalized Gaussian.*

Theorem 3.1 shows that, in order to obtain the multifractal spectrum of the sample paths of this model, one has to determine its large deviation spectrum $\boldsymbol{\rho}(\alpha)$. Denote

$$p(\alpha,\varepsilon,j) = \mathbb{P}(2^{-(\alpha+\varepsilon)j} \leq |C^{(i)}_{j,k}| \leq 2^{-(\alpha-\varepsilon)j}).$$

Then

$$p(\alpha,\varepsilon,j) = p_j\ \mathbb{P}(C_j \cdot |Y| \in [2^{-(\alpha+\varepsilon)j}, 2^{-(\alpha-\varepsilon)j}]).$$

Since $C_j = 2^{-\gamma j}$, it follows that

$$p(\alpha,\varepsilon,j) = p_j \left(F_\alpha(2^{-(\alpha-\gamma+\varepsilon)j}) - F_\alpha(2^{-(\alpha-\gamma-\varepsilon)j}) \right),$$

where the function F_α is defined by (19). The decay rate of $F_\alpha(a)$ when $a \to +\infty$ implies that

$$\text{If} \quad \alpha - \gamma < 0, \text{ then} \qquad F_\alpha(2^{-(\alpha-\gamma+\varepsilon)j}) << 2^{-Nj}, \quad \forall N > 0$$

so that

$$\text{if} \quad \alpha < \gamma, \text{ then} \qquad \boldsymbol{\rho}(\alpha) = -\infty.$$

If $\alpha - \gamma = 0$, then $p(\alpha,\varepsilon,j) \sim p_j F_\alpha(0)$ when $j \to +\infty$ so that

$$\text{if} \quad \alpha = \gamma, \text{ then} \qquad \boldsymbol{\rho}(\alpha) = 1 + \limsup_{j\to+\infty} \frac{\log(p_j)}{\log(2^j)}.$$

If $\alpha - \gamma > 0$, then using the fact that generalized Gaussians have a continuous nonvanishing density in the neighbourhood of 0, we obtain that $p(\alpha,\varepsilon,j) \sim p_j 2^{-(\alpha-\gamma-\varepsilon)j}$ when $j \to +\infty$ so that

$$\text{if} \quad \alpha > \gamma, \text{ then} \qquad \boldsymbol{\rho}(\alpha) = 1 - \alpha + \gamma + \limsup_{j\to+\infty} \frac{\log(p_j)}{\log(2^j)}.$$

The multifractal spectrum of the sample paths of X stated in the following proposition follows then from (28): Let

$$\omega = \limsup_{j\to+\infty} \frac{\log(p_j)}{\log(2^j)}.$$

The assumptions on the sequence p_j imply that $\omega \in [-1, 0]$.

Proposition 3.4. *Let X be a RWS given by the generalized Gaussian mixture model with $C_j = 2^{-\gamma j}$ and $p_j \geq 2^{(\varepsilon-1)j}$ for an $\varepsilon > 0$. The multifractal spectrum of almost every the sample paths of X is given by*

$$\begin{cases} \mathcal{D}_X(H) &= \dfrac{H(1+\omega)}{\gamma} \qquad \text{if } H \in \left[\gamma, \dfrac{\gamma}{1+\omega}\right] \\ &= -\infty \qquad\qquad \text{else.} \end{cases}$$

Remark 7. This result highlights a qualitative distinction between the cases $\omega = 0$ and $\omega \neq 0$. When $\omega = 0$, the sample paths are monohölder: the Hölder exponent is constant and equal to $H = \gamma$ everywhere. By contrast, when $\omega \neq 0$, the sample paths become multifractal, and the Hölder exponent ranges over the entire interval $[\gamma, \gamma/(1+\omega)]$.

We plan to investigate in future work the borderline case $\omega = 0$. A specific example of interest is when $p_j = 1/j^\delta$ for some $\delta > 0$. In this regime, although the sample paths remain monohölder, the modulus of continuity exhibits logarithmic fluctuations.

3.4 Almost everywhere modulus of continuity of GMM

In this section, we sharpen the last statement of Theorem 3.1 by determining the sharp a.e. modulus of continuity of generalized Gaussian mixtures. A consequence of this result will be that the smallest and largest possible pointwise moduli of continuity will have been determined exactly. We will prove the following result.

Proposition 3.5. *Let X be a SDRWM where $Y_{j,k}$ is a normalized generalized Gaussian and additionally C_j and $j \cdot 2^j p_j$ are decreasing sequences. Let θ be the modulus of continuity defined by the conditions*

$$\theta(l_j) = C_j \qquad \text{and } \theta \text{ is constant on } [l_j, l_{j-1}],$$

then the a. e. modulus of continuity of X cannot be a $o(\theta(h))$ and $\theta(h)|\log(h)|$ is an a. e. modulus of continuity of X.

Proof. Let $b > 0$ be defined by the condition

$$\mathbb{P}(|Y_{j,k}| \geq b) = 1/2.$$

We now consider the random collection $\mathcal{C}$ of couples (j,k) such that $|Y_{j,k}| \geq b$. They are also drawn at random and independently with probability $q_j = p_j/2$. Let us now consider the intervals $I_{j,k}$ of width $l_j = 1/j(2^j p_j)$ which are centered at such a point $k \cdot 2^{-j}$ where $(j,k) \in \mathcal{C}$. A given point $x_0 \in (0,1)$ has probability $r_{j,k} = 1/(j \cdot 2^j q_j)$ to belong to one of the $I_{j,k}$; since a given couple (j,k) has probability q_j to be chosen, and since there are 2^j coefficients at each scale,

$$\sum_{j,k} r_{j,k} q_j = \sum_j 2^j r_{j,k} q_j = \sum_j 1/j = +\infty.$$

It follows from the Borel-Cantelli lemma that almost every x_0 belongs to an infinite number of intervals $I_{j,k}$. Let now x_0 be such a point. First, we remark that Theorem 2.6 can be rewritten as follows (see

[42]) : If θ is a modulus of continuity of f at x_0, then

$$\forall j,k, \qquad |i| \leq C\theta\left(2^{-j} + \left|x_0 - \frac{k}{2^j}\right|\right). \tag{40}$$

Applying this criterium to $(j,k) \in \mathcal{C}$, we see that, if θ_1 is a modulus of continuity at x_0 then $\theta_1(l_j)$ cannot be a $o(C_j)$. Therefore, the a. e. modulus of continuity of X cannot be a $o(\theta(h))$.

In order to obtain a modulus of continuity which holds almost everywhere, we now pick intervals $I_{j,k}$ of width $l_j = 1/j^2(2^j p_j)$ which are centered at the point $k \cdot 2^{-j}$ where $y - j,k \neq 0$. A given point $x_0 \in (0,1)$ has probability $r_{j,k} = 1/(j^2 \cdot 2^j q_j)$ to belong to one of the $I_{j,k}$; by the same argument as above, we now have

$$\sum_{j,k} r_{j,k} q_j = \sum_j 2^j r_{j,k} q_j = \sum_j 1/j^2 < +\infty.$$

It follows from the Borel-Cantelli lemma that almost every x_0 belongs to a finite number of intervals $I_{j,k}$. Let now x_0 be such a point. Using now the converse part of Theorem 2.6, we obtain that $h \mapsto \theta_2(h)|\log(h)|$ is an a. e. modulus of continuity of X. $\square$

4 Estimation of the uniform Hölder exponent

To simplify the presentation of the method, we assume in this section that $d = 1$. Since our motivation comes from stochastic processes, we now consider $X = (X_t)_{t \in \mathbb{R}_+}$.

In [62] we provided estimations of the multifractality parameters c_1 and c_2 which encapsulate key information, respectively on the location of the maximum of the multifractal spectrum (which can be interpreted as the regularity exponent most often met in the data) and on the width of the spectrum (and therefore on the range of regularity exponents that are met in the data). In the present article, we complement this study by providing and estimation of the third most important multifractality parameter: the uniform Hölder exponent $H_X^{\min}$ which describes the uniform Hölder regularity of a function, a measure or a Schwartz distribution X. Formally, it is defined (see for instance [44]) by

$$H_X^{\min} = \sup\left\{\alpha : X \in C^\alpha(\mathbb{R}_+)\right\}.$$

This parameter proves highly practical due to its physical interpretation: in turbulence, $H_X^{\min}$ highlights the most singular structures within a turbulent flow, often associated with sharp gradients or extreme dissipation events; in finance, it captures abrupt changes or volatility spikes in time series data; and in image analysis, $H_X^{\min}$ identifies the sharpest edges or transitions in fractal-like textures. It has been used in many applications as a classification parameter, and therefore the question of its statistical estimation is important.

Related work. The estimation of multifractality parameters has a long history. Early contributions focused on increment- and variation-based estimators of local Hölder exponents [35], while global multiscale methods such as the Wavelet Transform Modulus Maxima (WTMM) [61], Multifractal Detrended Fluctuation Analysis (MFDFA) [48], and wavelet-leader regressions [2] have been used to estimate extremal pointwise exponents ($h_{\min}$, $h_{\max}$), the spectral mode h_{pic}, and the log-cumulants

c_q, including in particular c_1 and c_2 [36, 83, 20]. These approaches rely on log-log regressions across scales, which require sufficiently long signals and may lead to large estimation variances, especially in multivariate settings [4, 76]. To address these limitations, Bayesian methods have been proposed for the estimation of multifractality parameters in images and multivariate fields, relying on Whittle approximations and gamma Markov random field priors [20, 82]. In parallel, significant work has focused on estimating structural parameters in random wavelet series and multifractal stochastic models, including log-normal and log-infinitely-divisible cascades [12, 18] and Markov-switching multifractal processes [17]. These parametric models complement nonparametric approaches and place the estimation of $H_X^{\min}$ within a broader inference framework.

4.1 Estimation procedure

In contrast with the parameters c_1 and c_2 whose estimates were based on the laws of the log-leaders of the wavelet coefficients, the quantity $H_X^{\min}$ is derived directly from the wavelet coefficients: it can be computed through a log-log plot regression as

$$H_X^{\min} = \liminf_{j\to+\infty} \frac{\log\left(\sup_k |C_{j,k}|\right)}{\log(2^{-j})}. \tag{41}$$

We now study the uniform Hölder exponent of a semi-dependent random wavelet mixture (SDRWM) $X = (X_t)_{t\in\mathbb{R}+}$ given by

$$X_t = \sum_{k\in\mathbb{Z}^d} C_k \,\varphi_k(t) + \sum_{j\geq 0,\ k\in\mathbb{Z}^d} C_{j,k}\,\psi_{j,k}(t) \tag{42}$$

where the $(C_{j,k})_{j,k}$ satisfy Definition (2.2). More precisely, we focus on the wavelet-coefficient model $(G, (p_j)_{j\geq 0}, (C_{j,k})_{j\geq 0, k\in\mathbb{Z}})$, which is assumed to satisfy the following conditions:

Assumption 3. *1. For each resolution level $j \geq 0$ and spatial index $k \in \mathbb{Z}$,*

(i) $\mathbb{P}(C_{j,k} = 0) = 1 - p_j$, where we assume $p_j = 2^{(\eta-1)j}$ for some $\eta \in (0, 1)$;

(ii) Conditionally on $C_{j,k} \neq 0$, one has $C_{j,k} \overset{\mathcal{L}}{=} C_j, Y$, where Y has density G.

2. The coefficients are of the form

$$C_{j,k} = 2^{-\alpha j} D X_{j,k}, \tag{43}$$

where $\alpha > 0$ is a fixed decay exponent and $D > 0$ is a scaling constant.

3. For any $j \geq 0$ the random variables $(X_{j,k})_k$ are i.i.d. and generalized Gaussian *(see (6)), with density*

$$f_\beta(x) = \frac{\beta}{2\Gamma(\frac{1}{\beta})} e^{-|x|^\beta} = \kappa_\beta e^{-|x|^\beta}, \tag{44}$$

for some $\beta > 0$.

Remark 8. Condition (i) ensures that there are approximately $2^{\eta j}$ non-vanishing wavelet coefficients, randomly located in $[0, 1]$, each of magnitude $2^{-\alpha j}$ (with $\alpha > 0$). When $\eta = 1$, the coefficients are non-lacunary, and the model reduces to the standard SDRWS case. Condition 3 corresponds to the semi-dependent case: the wavelet coefficients are independent at a given scale j, but no independence assumption is imposed between coefficients located at different scales.

The following result, which may be viewed as a corollary of Proposition 2.4, provides an explicit expression for the uniform modulus of continuity of the RWM satisfying Assumption 3.

Proposition 4.1. *Let X be the SDRWM with coefficients satisfying Assumption 3. Then, $H_X^{\min} = \alpha$.*

Proof. Let us assume, without loss of generality, that $\beta > 1$. The case $\beta \leq 1$ can be handled in the same way by using the version of Mills' ratio given by (59) ($\beta < 1$) or exact computation (60) (*beta* $=$). Let for any $(j,k) \in \mathbb{N} \times \mathbb{Z}$, $C_{j,k} := 2^{-\alpha j} D X_{j,k}$ (with $\beta > 1$). For any $x \in \mathbb{R}_+^*$, by independence of the wavelet coefficients $(C_{j,k})_k$,

$$\begin{aligned}\mathbb{P}\big(\sup_k |C_{j,k}| \leq x\big) &= \prod_k \mathbb{P}(|C_{j,k}| \leq x)\\ &= \Big[1 - p_j \mathbb{P}(|X_{j,k}| \geq 2^{\alpha j} x/D)\Big]^{N_j},\end{aligned}$$

where $N_j = C_1 p_j 2^j = C_1 2^{\eta j}$. In the one hand, choosing x such that $p_j \mathbb{P}(|X_{j,k}| \geq 2^{\alpha j} x/D) = 1/N_j$, we get

$$\lim_{j\to\infty} \mathbb{P}\big(\sup_k |C_{j,k}| \leq x\big) = e^{-1}.$$

On the other hand, it follows from Mills ratio (56) applied to generalized Gaussian random variables with $\beta > 1$, that, for any $x > 0$,

$$\mathbb{P}(|X_{j,k}| \geq 2^{\alpha j} x/D) \underset{j\to\infty}{\sim} \frac{2 f_\beta(2^{\alpha j} x/D)}{\beta (2^{\alpha j} x/D)^{\beta-1}},$$

Ignoring the factor $\Gamma(1/\beta)(2^{\alpha j} x/D)^{\beta-1}$ at first order and setting

$$x = \left[\frac{D^\beta\big((2\eta-1) j \log(2) + \log(C_1)\big)}{2^{\alpha j \beta}}\right]^{1/\beta},$$

we get

$$\log\big(\sup_k |C_{j,k}|\big) \underset{j\to\infty}{\sim} \log\big((\log(2)(2\eta-1)D^\beta)^{1/\beta} j^{1/\beta} 2^{-\alpha j}\big),$$

so that $H_X^{\min} = \liminf_{j\to+\infty} \log\big(\sup_k |C_{j,k}|\big)/\log(2^{-j}) = \alpha$. Hence the result. □

To construct an estimator of $H_X^{\min}$, we need a more precise understanding of the behavior of the random variables $\log\big(\sup_k |C_{j,k}|\big)/\log(2^{-j})$ defined in (41) for large j. This amounts to providing a quantitative version of Proposition 4.1. The estimation procedure relies on the following key lemma.

Lemma 4.2. *Let X be the RDWM with coefficients defined by (43) and (44). Let $\varepsilon > 0$. For $\beta > 0$, set $\theta_1 = 3(1-\beta)/(2\beta)$, $\theta_2 = (1-\beta)/(2\beta)$, $\delta = -(1-\beta)/(2\beta)$, $c' = \kappa_\beta/(2\beta-1)$ and $c'' = \kappa_\beta/\beta$. For $c > 0$, set $C_{\eta,3}(c) = cC_1 \log(2)^{(1-\beta)/\beta}$, $C_{\eta,4}(c) = c\log(2)^{(1-\beta)/\beta}$.*

- ***Case*** $\beta > 1$***.*** *Set* $j_1(\varepsilon)$ *and* $j_2(\varepsilon)$ *such that*

$$\mathrm{e}^{-C_3(c')j_1(\varepsilon)^\delta} = \varepsilon/2 \quad \textit{and} \quad \mathrm{e}^{-C_{\eta,3}(c'')j_2(\varepsilon)^{-\delta}/(1-C_{\eta,4}(c'')2^{-j_2(\varepsilon)}j_2(\varepsilon)^{-\delta})} = 1-\varepsilon/2. \tag{45}$$

For any $j \geq \max(j_1(\varepsilon), j_2(\varepsilon))$*, the following holds with probability at least* $1-\varepsilon$*:*

$$\sup_k |C_{j,k}| \in \Big[2^{-\alpha j}D(\eta j\log(2) + \theta_1 \log(j))\,,\, 2^{-\alpha j}D(\eta j \log(2) + \theta_2 \log_2(j))\Big].$$

- ***Case*** $\beta < 1$***.*** *Set* $j_1(\varepsilon)$ *and* $j_2(\varepsilon)$ *such that*

$$\mathrm{e}^{-C_{\eta,3}(c'')j_1(\varepsilon)^{-\delta}} = \varepsilon/2 \quad \textit{and} \quad \mathrm{e}^{-C_{\eta,3}(c'')\left[1+C_{\eta,4}(c)j^\delta/(2(1-C_{\eta,4}(c)j^\delta)\right]} = 1-\varepsilon/2.$$

For any $j \geq \max(j_1(\varepsilon), j_2(\varepsilon))$*, the following holds with probability at least* $1-\varepsilon$*:*

$$\sup_k |C_{j,k}| \in \Big[2^{-\alpha j}D(\eta j\log(2) + \theta_2 \log(j))\,,\, 2^{-\alpha j}D(\eta j \log(2) + \theta_1 \log_2(j))\Big].$$

- *Case* $\beta = 1$*. Set* $j_1(\varepsilon)$ *and* $j_2(\varepsilon)$ *such that*

$$e^{-C_1 j_1(\varepsilon)} = \varepsilon/2 \quad \textit{and} \quad e^{-C_1 j_2(\varepsilon)^{-1}\left[1-C_1 j_2(\varepsilon)^{-1}/(2(1-C_1 j_2(\varepsilon)^{-1}))\right]} = 1-\varepsilon/2.$$

For any $j \geq \max(j_1(\varepsilon), j_2(\varepsilon))$*, the following holds with probability at least* $1-\varepsilon$*:*

$$\sup_k |C_{j,k}| \in \Big[2^{-\alpha j}D(\eta j\log(2) - \log(j))\,,\, 2^{-\alpha j}D(\eta j \log(2) + \log_2(j))\Big].$$

Proof. The proof is postponed to Appendix C.2. □

Estimation of $H_X^{\min}$ from quantiles We directly deduce from the Lemma 4.2 that for any $j \geq \max(j_1(\varepsilon/2), j_2(\varepsilon/2))$, we have with probability at least $1-\varepsilon/2$:

$$\sup_k |C_{j,k}| \in \Big[2^{-\alpha j}D(\eta j\log(2) + \theta_\ell \log(j))\,,\, 2^{-\alpha j}D(\eta j \log(2) + \theta_u \log(j))\Big].$$

as well as

$$\sup_k |C_{2j,k}| \in \Big[2^{-2\alpha j}D(2\eta j\log(2) + \theta_\ell \log(2j))\,,\, 2^{-2\alpha j}D(2\eta j \log(2) + \theta_u \log(2j))\Big],$$

where $\theta_\ell = \theta_1 \mathbf{1}_{\{\beta\geq 1\}} + \theta_2\mathbf{1}_{\{\beta<1\}}$ and $\theta_u = \theta_2 \mathbf{1}_{\{\beta\geq 1\}} + \theta_1\mathbf{1}_{\{\beta<1\}}$. For any $j \geq 0$, define

$$Z_j = \frac{\log\left(\sup_k |C_{2j,k}|/\sup_k |C_{j,k}|\right)}{\log(2^{-j})}. \tag{46}$$

Then, we have

$$\mathbb{P}(Z_j \geq \ell_j) \geq 1-\varepsilon \quad \text{and} \quad \mathbb{P}(Z_j \leq u_j) \geq 1-\varepsilon, \tag{47}$$

where we set

$$\ell_j = \frac{\log\left(\frac{2^{-2\alpha j} D(2\eta j \log(2)+\theta_\ell \log(j))}{2^{-\alpha j} D(\eta j \log(2)+\theta_u \log(j))}\right)}{\log(2^{-j})} = \frac{\log\left(2^{-\alpha j} \frac{2\eta j \log(2)+\theta_\ell \log(j)}{\eta j \log(2)+\theta_u \log(j)}\right)}{\log(2^{-j})} = \alpha + r_j^\ell, \tag{48}$$

and

$$u_j = \frac{\log\left(\frac{2^{-2\alpha j} D(2\eta j \log(2)+\theta_u \log(j))}{2^{-\alpha j} D(\eta j \log(2)+\theta_\ell \log(j))}\right)}{\log(2^{-j})} = \frac{\log\left(2^{-\alpha j} \frac{2\eta j \log(2)+\theta_u \log(j)}{\eta j \log(2)+\theta_\ell \log(j)}\right)}{\log(2^{-j})} = \alpha + r_j^u, \tag{49}$$

with

$$r_j^\ell = \frac{\log\left(\frac{2\eta j \log(2)+\theta_\ell \log(j)}{\eta j \log(2)+\theta_u \log(j)}\right)}{\log(2^{-j})} \quad \text{and} \quad r_j^u = \frac{\log\left(\frac{2\eta j \log(2)+\theta_u \log(j)}{\eta j \log(2)+\theta_\ell \log(j)}\right)}{\log(2^{-j})}. \tag{50}$$

To estimate $H_X^{\min} = \alpha$, it is then enough to have estimates of ℓ_j and u_j and we have

$$H_X^{\min} = \frac{(\ell_j - r_j^\ell) + (u_j - r_j^u)}{2}.$$

Note moreover that follows from (47) that

$$q_p^j := q_{1-\varepsilon}^j := \frac{\ell_j + u_j}{2}$$

is the $(1-\varepsilon)$-quantile of the random variable Z_j.

Estimation of ℓ_j and u_j Consider i.i.d. copies $X^1, \dots, X^n$ of the process X defined in (1). For any $l \in \{1, \dots, n\}$, set

$$Z_j^l = \frac{\log\left(\sup_k |C_{2j,k}^l| / \sup_k |C_{j,k}^l|\right)}{\log(2^{-j})}.$$

As a reminder, the empirical cumulative distribution function associated to the sample $\{Z_j^1, \dots, Z_j^n\}$ is defined as:

$$\widehat{F}_n^j(x) = \frac{1}{n} \sum_{l=1}^n \mathbf{1}_{\{Z_j^l \le x\}}.$$

The Z_j^l are i.i.d. random variables distributed as the variable Z_j defined by (46). Let $p = p_\varepsilon = 1 - \varepsilon$. As explained in the Appendix, the idea of the so-called Peaks-Over-Threshold (POT) method is to first extract the excesses $Y_j^i = Z_j^{(n-k_n^j+i)} - Z_j^{(n-k_n^j)}$ for $i \in \{1, \dots, k_n^j\}$. Under the Pickands-Balkema-de Haan theorem the conditional excesses sequence $(Y_j^i)_{i \ge 1}$ converges in distribution to a random variable Y_j distributed according to a generalized Pareto distribution

$$F_{Y_j}(y) = \mathbb{P}(Y_j > y | Y_j > 0) = 1 - (1 + \xi_j y / \sigma_j)^{-1/\xi_j}, \qquad y > 0,\ 1 + \xi_j y / \sigma > 0,$$

where $\xi_j > 0$ is the shape (tail index) and $\sigma_j = \sigma_{j,n}$ is a scale that may depend on v_n. Define the POT estimators (see Appendix B) by

$$\widehat{\ell}_j := v_n + \frac{\widehat{\sigma}_j}{\widehat{\xi}_j}\left[\left(\frac{k_n^j}{np}\right)^{-\widehat{\xi}_j} - 1\right] \text{ and } \widehat{u}_j := \widehat{\ell}_j - \left(Z_j^{(n-k_n^j+1)} - Z_j^{(n-k_n^j)}\right). \tag{51}$$

where the estimators $(\widehat{\xi}_j, \widehat{\sigma}_j)$ can be obtained by the maximum of likelihood estimator

$$(\widehat{\xi}_j, \widehat{\sigma}_j) := (\widehat{\xi}_j^{k_n^j}, \widehat{\sigma}_j^{k_n^j}) = \underset{(\xi,\sigma)}{\operatorname{argmax}} \sum_{i=1}^{n} \log\left(f_{Y_j}(y_i; \xi, \sigma)\right)$$

<u>Final estimator</u> Then, we consider

$$\widehat{H_X^{\min}} = \frac{1}{2|J(\varepsilon)|} \sum_{j \in J(\varepsilon)} \left[(\widehat{\ell}_j - r_j^{\ell}) + (\widehat{u}_j - r_j^{u})\right], \tag{52}$$

where $J(\varepsilon) = \{j,\, j \geq \max(j_1(\varepsilon), j_2(\varepsilon))\}$.

Remark 9. The choice of the random variables Z_j (46) may appear unconventional. In the context of log-log regressions, one typically considers the ratio

$$\frac{\log\left(\sup_k |C_{j,k}|\right)}{\log(2^{-j})}.$$

In contrast, we focus on the numerator,

$$\log\left(\frac{\sup_k |C_{2j,k}|}{\sup_k |C_{j,k}|}\right),$$

which allows us to mitigate the potentially detrimental influence of the normalizing constant D and thereby enhances the robustness of our procedure with respect to it.

We can deduce a confidence interval for the estimation of $H_X^{\min}$:

Theorem 4.3. *Let $\delta \in (0,1)$. With probability $1-\delta$,*

$$H_X^{\min} \in \left[\frac{1}{2|J(\varepsilon)|} \sum_{j \in J(\varepsilon)} \frac{(\widehat{\ell}_j - r_j^{\ell}) + (\widehat{u}_j - r_j^{u})}{1 + z_{1-\delta/(2/|J(\varepsilon)|)} \sqrt{\widehat{V}_{j,k}/k_n^j}}\, ,\, \frac{1}{2|J(\varepsilon)|} \sum_{j \in J(\varepsilon)} \frac{(\widehat{\ell}_j - r_j^{\ell}) + (\widehat{u}_j - r_j^{u})}{1 - z_{1-\delta/(2/|J(\varepsilon)|)} \sqrt{\widehat{V}_{j,k}/k_n^j}}\right]. \tag{53}$$

where $z_{1-\delta/2}$ is the $(1-\delta/2)$-quantile of the standard normal distribution and $\widehat{V}_{j,k} = \widehat{\xi}_{j,k}^2 + (1+\widehat{\xi}_{j,k})^2 \varepsilon/(2(1-\varepsilon/2))$.

Proof. Let $\delta > 0$. Follows from Theorem B.1, that for all $j \in [J(\varepsilon)]$ with probability at least $1 - \delta/|J(\varepsilon)|$,

$$q_p^j = \frac{\ell_j + u_j}{2} \in \left[\frac{(\widehat{\ell}_j + \widehat{u}_j)/2}{1 + z_{1-\delta/(2|J(\varepsilon)|)} \sqrt{\widehat{V}_{j,k}/k_n^j}}\, ,\, \frac{(\widehat{\ell}_j + \widehat{u}_j)/2}{1 - z_{1-\delta/(2|J(\varepsilon)|)} \sqrt{\widehat{V}_{j,k}/k_n^j}}\right], \tag{54}$$

where $z_{1-\delta/2}$ is the $(1-\delta/2)$-quantile of the standard normal distribution and $\widehat{V}_{j,k} = \widehat{\xi}_{j,k}^2 + (1+\widehat{\xi}_{j,k})^2 p/p$. For all $j \in [J(\varepsilon)]$, let

$$E_j = \left\{\frac{(\widehat{\ell}_j + \widehat{u}_j)/2}{1 + z_{1-\delta/(2/|J(\varepsilon)|)} \sqrt{\widehat{V}_{j,k}/k_n^j}} \leq \frac{\ell_j + u_j}{2} \leq \frac{(\widehat{\ell}_j + \widehat{u}_j)/2}{1 - z_{1-\delta/(2/|J(\varepsilon)|)} \sqrt{\widehat{V}_{j,k}/k_n^j}}\right\}$$

Note that $\mathbb{P}(E_j) \geq 1 - \delta/|J(\varepsilon)|$ for all $j \in [J(\varepsilon)]$. Consider the event

$$A = \Bigg\{ \frac{1}{2|J(\varepsilon)|} \sum_{j \in J(\varepsilon)} \frac{(\widehat{\ell}_j - r_j^\ell) + (\widehat{u}_j - r_j^u)}{1 + z_{1-\delta/(2/|J(\varepsilon)|)} \sqrt{\widehat{V}_{j,k}/k_n}} \leq H_X^{\min} \leq \frac{1}{2|J(\varepsilon)|} \sum_{j \in J(\varepsilon)} \frac{(\widehat{\ell}_j - r_j^\ell) + (\widehat{u}_j - r_j^u)}{1 - z_{1-\delta/(2/|J(\varepsilon)|)} \sqrt{\widehat{V}_{j,k}/k_n}} \Bigg\}$$

$$= \Bigg\{ \frac{1}{2|J(\varepsilon)|} \sum_{j \in J(\varepsilon)} \frac{(\widehat{\ell}_j - r_j^\ell) + (\widehat{u}_j - r_j^u)}{1 + z_{1-\delta/(2/|J(\varepsilon)|)} \sqrt{\widehat{V}_{j,k}/k_n}} \leq \frac{1}{2|J(\varepsilon)|} \sum_{j \in J(\varepsilon)} [(\ell_j - r_j^\ell) + (u_j - r_j^u)] \leq \frac{1}{2|J(\varepsilon)|} \sum_{j \in J(\varepsilon)} \frac{(\widehat{\ell}_j - r_j^\ell) + (\widehat{u}_j - r_j^u)}{1 - z_{1-\delta/(2/|J(\varepsilon)|)} \sqrt{\widehat{V}_{j,k}/k_n}} \Bigg\}$$

Using a union bound argument, we get

$$\begin{aligned}
\mathbb{P}(A) &\geq \mathbb{P}\Big(\bigcap_{j=1}^{J(\epsilon)} E_j \Big) \\
&\geq 1 - \sum_{j=1}^{J(\epsilon)} \mathbb{P}(E_j^c) \\
&\geq 1 - |J(\epsilon)|/(\delta|J(\epsilon)|) = 1 - \delta.
\end{aligned}$$

Hence the result. □

4.2 Numerical experiments

In this section, we empirically assess the estimation procedure for $H^{\min}$ for random wavelet series described in the previous section. More specifically, we investigate the finite-sample behavior of the proposed peaks-over-threshold (POT)-based estimator (52) of the uniform Hölder exponent. Synthetic multiscale coefficient fields with known theoretical regularity are generated, and we evaluate the accuracy of both the point estimator (52) and the associated confidence interval (53). Particular attention is devoted to the choice of the intermediate sequence $(k_n)_n$, which is selected via bootstrap minimization of the mean squared error (MSE) of the Hill estimator.

Procedure For given parameters $\alpha > 0$, $D > 0$, and $\beta > 0$, we simulate independent generalized Gaussian wavelet coefficients

$$C_{j,k}^\ell = D 2^{-\alpha j} X_{j,k}^{(\ell)}, \qquad 1 \leq \ell \leq n, \quad 1 \leq j \leq J_{\max}, \quad 1 \leq k \leq 2^j,$$

where $X_{j,k}^\ell \sim \mathrm{GG}(0, 1, \beta)$ has density $f : x \mapsto \beta/(2\Gamma(1/\beta)) \exp(-|x|^\beta)$.
Throughout, we set $n = 10^5$ and $J_{\max} = 14$, which is the largest scale level for which the simulation remains computationally feasible. Beyond this point, the exponential growth of the number of coefficients makes the procedure considerably slower and unstable in practice.

For each replicate, we compute the quantities Z_j^ℓ defined in (46), up to scale $j = 7$ according to their construction. We then retain the set of scales $J(\varepsilon) = \{4, 5, 6, 7\}$, which correspond to the highest available resolutions. In the theoretical analysis, the scale index j is considered asymptotically large, while in practice only a finite number of scales can be used. Accordingly, $J(\varepsilon)$ is chosen to consist of the largest available scales to closely approximate this asymptotic behavior. The exceedance probability is fixed at $\varepsilon = 10^{-5}$.

Tuning Parameter As described in Appendix B, the Peaks-Over-Threshold (POT) method starts by extracting the exceedances $Y_i = Z_{(n-k_n+i)} - v_n,\ i = 1, \ldots, k_n$, where the threshold is $v_n = Z_{(n-k_n)}$. The choice of the number of upper order statistics k_n plays a central role in tail estimation: selecting too few observations results in high variance, whereas including too many introduces bias. Following the data-driven strategy proposed by [22], we determine k_n through a bootstrap-based selection procedure. For each candidate value of $k_n = k$, the tail index ξ is first estimated from the top k exceedances. Then, for this same k, we generate B bootstrap resamples drawn with replacement from the original sample, recompute the tail index on each bootstrap sample, and evaluate the corresponding mean squared error (MSE) between the bootstrap estimates and the original estimate. The optimal k_n is selected as the value of k that minimizes this bootstrap MSE, thus achieving a principled trade-off between bias and variance. In our study, we consider a regularly spaced grid $K_{\text{grid}} = 10 : 10 : \left\lfloor \frac{n}{5} \right\rfloor$, and perform $B = 300$ bootstrap replications. This choice provides a sufficiently fine resolution for selecting k_n, while ensuring the standard requirement $k_n/n \to 0$ for the consistency of semi-parametric tail estimators. It also remains computationally feasible, offering a good compromise between statistical accuracy and numerical efficiency.

Results and interpretation

Table 1 reports the estimated 95% confidence intervals for the minimal Hölder exponent $H_X^{\min}$ under various choices of the parameters α, β, and D. These numerical experiments are designed to assess the finite-sample performance of the proposed Peaks-Over-Threshold (POT) estimator and to evaluate its robustness with respect to the underlying model parameters.
We first examine the effect of the tail parameter β by fixing $D = 0.5$ and considering both $\beta < 1$ and $\beta > 1$. The resulting confidence intervals exhibit only minor variations across different values of β, indicating that the estimator is largely insensitive to the tail behavior of the generalized Gaussian coefficients. This robustness with respect to β suggests that the POT-based approach performs reliably across a broad range of heavy-tailed regimes.

Next, we analyze the influence of the scale decay parameter α. For each fixed β, the width of the confidence intervals increases monotonically with α, in agreement with the theoretical scaling of the wavelet coefficients (43). Smaller values of α correspond to slower decay across scales, leading to larger extreme coefficients and consequently narrower confidence intervals. In contrast, larger α values induce faster decay, yielding smaller extremes and increased relative variability in their estimation, which results in wider intervals. Therefore, the decay rate α directly controls the magnitude of the

Table 1: Estimated 95% confidence intervals for the uniform Hölder exponent $H_X^{\min}$ with fixed amplitude $D = 0.5$ and varying scale decay α and tail parameter β.

β	$\alpha = H_X^{\min}$	$\widehat{H_X^{\min}}$	CI	Size
0.8	0.5	0.501	$[0.499, 0.504]$	0.005
	1	0.998	$[0.993, 1.003]$	0.010
	5	4.983	$[4.958, 5.008]$	0.049
	10	9.991	$[9.939, 10.043]$	0.104
	50	49.99	$[49.728, 50.255]$	0.527
3	0.5	0.498	$[0.495, 0.501]$	0.005
	1	0.997	$[0.992, 1.001]$	0.009
	5	4.998	$[4.975, 5.021]$	0.046
	10	9.997	$[9.950, 10.044]$	0.094
	50	49.996	$[49.759, 50.236]$	0.476

Table 2: Estimated 95% confidence intervals for the uniform Hölder exponent $H_X^{\min}$ with fixed tail parameter $\beta = 4$ and varying scale decay α and amplitude D.

D	$\alpha = H_X^{\min}$	$\widehat{H_X^{\min}}$	CI	Size
0.5	0.5	0.498	$[0.495, 0.500]$	0.005
	1	0.998	$[0.993, 1.003]$	0.010
	5	4.998	$[4.972, 5.024]$	0.052
	10	9.998	$[9.946, 10.050]$	0.104
	50	49.998	$[49.740, 50.258]$	0.518
5	0.5	0.498	$[0.495, 0.500]$	0.005
	1	0.998	$[0.993, 1.003]$	0.010
	5	4.998	$[4.972, 5.024]$	0.052
	10	9.998	$[9.946, 10.050]$	0.104
	50	49.998	$[49.740, 50.258]$	0.518

extreme wavelet coefficients and, in turn, governs the precision of the POT-based estimator.

In Table 2, we investigate the effect of the amplitude parameter D for fixed $\beta = 4$, considering two representative values, $D = 0.5$ and $D = 5$. The corresponding confidence intervals are nearly identical across these choices, demonstrating that the estimator is essentially invariant with respect to the amplitude of the coefficients. As in the previous experiments, the interval width remains primarily driven by α, confirming that the decay parameter plays a dominant role in controlling estimation variability.

Overall, these experiments demonstrate that the proposed estimator is robust with respect to both the tail parameter β and the amplitude D. Moreover, the estimation error scales predictably with α, in agreement with the theoretical behavior of the wavelet coefficients and the asymptotic properties of the POT estimator. In summary, the proposed estimator provides accurate and reliable confidence intervals for $H_X^{\min}$ across a wide range of parameter settings. The estimator remains accurate and stable over a broad range of heavy-tail parameters β and amplitudes D, with confidence interval lengths consistently remaining small, thereby confirming the effectiveness of the POT-based approach.

5 Concluding remarks

In this paper, we refined classical uniform and pointwise regularity results concerning random wavelet series (RWS) and we extended them to settings with dependent coefficients. We focused on the particular case commonly met in applications of generalized Gaussian mixture models. We broadened the classical independent-scale setting to a semi-dependent random wavelet series (SDRWS) framework, permitting arbitrary cross-scale dependence while preserving within-scale independence. In this unified setting we specified known uniform moduli of continuity, we extended them to SDRWS and mixture models, and we characterized the associated pointwise worst-case regularity. A new block ubiquity theorem allowed to derive multifractal spectrum for SDRWS, which we specialized to generalized Gaussian mixtures together with the almost-everywhere modulus of continuity. These results elucidate in a coherent manner how cross-scale dependence shapes Hölder and multifractal features. Finally, exploiting wavelet leaders, we constructed an estimator of the uniform Hölder (minimal) exponent $H^{\min}$ and derived a theoretically grounded confidence interval. Empirical experiments confirm the reliability of the inference and its confidence interval. Together with the estimation of the other classical multifractality parameters c_1 and c_2 performed in the previous article [62], the addition of the third one $H_X^{\min}$ performed in the present paper now allows to consider questions that were addressed previously in very particular parametric settings such as: *Are the data monohölder or multifractal?*

We considered laws of wavelet coefficients that are commonly met in the signal and image processing litterature. However, a natural question is to determine if these hypotheses are compatible with the natural consistency requirement that this hypothesis remains invariant under a change of wavelet basis. Such verifications are common for the definition of function spaces defined by conditions on the wavelet coefficients [57], but this problem does not seem to have been considered in other settings. Another question of interest is to investigate applications of these results to models which satisfy the the SDRWS assumptions; the conclusions of the present paper state that multifractal properties are not really affected if the cross-scale independence assumption is dropped. A natural question is to determine of some more refined multifractal analysis would be able to put in evidence such results. Natural candidates are supplied by bivariate multifractal analyses of several different pointwise regularity quantities derived from the data, see [43].

A Mills ratio

In probability theory, the Mills ratio [59] states that for a continuous real random variable X with density f and for any $x \in \mathbb{R}$,

$$\frac{I(x)}{f(x)} = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \mathbb{P}(x < X \leq x + \varepsilon | X > x), \tag{55}$$

where for all $x \in \mathbb{R}$, we define $I(x) = \mathbb{P}(X > x)$. Bounding (55) provides insights on the distribution of the tails of a random variable. For instance, if X has a generalized Gaussian distribution with density (44), for all $x > 0$,

$$I(x) \underset{x \to \infty}{\sim} \frac{x^{1-\beta} \mathrm{e}^{-x^{\beta}}}{2\Gamma(1/\beta)}. \tag{56}$$

This result can be refined by providing non-asymptotic bounds. We recall the following lemma, proved in the companion paper.

Lemma A.1 (General bounds for Mills ratio)**.** *Let $g : \mathbb{R} \to \mathbb{R}$ be a twice differentiable positive function such that g' is positive on $\mathbb{R}_+^*$ and g'' has a constant sign on $\mathbb{R}_+$. Assume, moreover, that there exists* *a function $M : \mathbb{R}_+ \to \mathbb{R}_+$ such that for $x > 0$,*

$$\sup_{t\in[x,\infty)} \left| \frac{g''(t)}{(g'(t))^2} \right| \le M(x).$$

Let X be a real random variable with density $f = \kappa e^{-g}$, where $\kappa > 0$ is a normalisation constant. For all $x \in \mathbb{R}$, we define $I(x) = \mathbb{P}(X > x) = \int_x^\infty \kappa e^{-g(t)} dt$. Then,

$$\forall x \in \mathbb{R}_+^*, \qquad \frac{f(x)}{g'(x)(1 + M(x)\mathbf{1}_{\{g''>0\}})} \le I(x) \le \frac{f(x)}{g'(x)(1 - M(x)\mathbf{1}_{\{g''<0\}})}. \tag{57}$$

Proof. For $x \in \mathbb{R}_+^*$,

$$\begin{aligned}
\kappa^{-1} I(x) &= \int_x^\infty \frac{g'(t)}{g'(t)} e^{-g(t)} dt = \lim_{A\to\infty} \left[-\frac{1}{g'(t)} e^{-g(t)} \right]_x^A - \int_x^\infty \frac{g''(t)}{(g'(t))^2} e^{-g(t)} dt \\
&= \frac{1}{g'(x)} e^{-g(x)} - \int_x^\infty \frac{g''(t)}{(g'(t))^2} e^{-g(t)} dt.
\end{aligned}$$

If $g'' > 0$, we get that for any $x \in \mathbb{R}_+^*$,

$$\frac{e^{-g(x)}}{g'(x)} - \kappa^{-1} M(x) I(x) \le \kappa^{-1} I(x) \le \frac{e^{-g(x)}}{g'(x)},$$

whereas if $g'' < 0$,

$$\frac{e^{-g(x)}}{g'(x)} \le \kappa^{-1} I(x) \le \frac{e^{-g(x)}}{g'(x)} + \kappa^{-1} m(x) I(x).$$

The two inequalities lead to (57). □

Example 1. 1. **Standard Gaussian.** Taking $g(x) = x^2/2$ and $\kappa = (2\pi)^{-1/2}$, we find $M(x) = 1/x^2$. This boils down to the well-known result:

$$\forall x > 0, \qquad \frac{1}{\sqrt{2\pi}} \frac{\mathrm{e}^{-x^2/2}}{x(1 + 1/x^2)} \le I(x) \le \frac{1}{\sqrt{2\pi}} \frac{\mathrm{e}^{-x^2/2}}{x}.$$

2. **Generalized Gaussian with light tails ($\beta > 1$).** Taking $g(x) = |x|^\beta$ and $\kappa_\beta = \beta/(2\Gamma(1/\beta))$ with $\beta > 1$, clearly $g' > 0$ and $g'' > 0$ on $\mathbb{R}_+^*$. We have that for all $x > 0$, $M(x) = (\beta - 1)/(\beta x^\beta)$ and

$$\frac{1}{2\Gamma(1/\beta)} \frac{\mathrm{e}^{-x^\beta}}{x^{\beta-1}(1 + (\beta-1)/(\beta x^\beta))} \le I(x) \le \frac{1}{2\Gamma(1/\beta)} \frac{\mathrm{e}^{-x^\beta}}{x^{\beta-1}}. \tag{58}$$

3. **Generalized Gaussian with heavy tails ($0 < \beta < 1$).** Taking $g(x) = |x|^\beta$ and $\kappa = \beta/(2\Gamma(1/\beta))$ with $\beta < 1$, we check that $g' > 0$ and $g'' < 0$ on $\mathbb{R}_+^*$. We have that for all $x > 0$, $M(x) = (1-\beta)/(\beta x^\beta)$ and that

$$\frac{x^{1-\beta}\mathrm{e}^{-x^\beta}}{2\Gamma(1/\beta)} \le I(x) \le \frac{1}{2\Gamma(1/\beta)} \frac{x^{1-\beta}\mathrm{e}^{-x^\beta}}{(1 - (1-\beta)/(\beta x^\beta))}. \tag{59}$$

In the case of a **Laplace random variable** ($\beta = 1$), there is no need to invoke Mills' ratio; the integral can be computed explicitly. Indeed, for any $x \ge 0$, one has

$$I(x) = \frac{e^{-x}}{2}. \tag{60}$$

B Extreme quantile estimation

B.1 Definition and confidence interval

Let $p \in (0,1)$ and F be the cumulative distribution of a random variable Z. The p-quantile can be defined as

$$q_p = \frac{1}{2}\Big[\inf\{x : F(x) \geq p\} + \sup\{x : F(x) \leq p\}\Big] =: \frac{1}{2}[q_p^+ + q_p^-].$$

Consider $Z_1, \dots, Z_n$ i.i.d. random variables. The *empirical cumulative distribution function* is defined as:

$$\widehat{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{Z_i \leq x\}}.$$

Then the empirical p-quantiles are given by

$$\widehat{q_p^+} := \inf\{x : \widehat{F}_n(x) \geq p\} \text{ and } \widehat{q_p^-} := \sup\{x : \widehat{F}_n(x) \leq p\} \tag{61}$$

We set

$$\widehat{q_p^+} = Z^{(k)} \text{ for } p \in ((k-1)/n, k/n] \text{ and } \widehat{q_p^-} = Z^{(k)} \text{ for } p \in [k/n, (k+1)/n),$$

where $Z^{(1)} \leq Z^{(2)} \leq \cdots \leq Z^{(n)}$ are the order statistics. To remain consistent with Lemma (4.2), we set $p = p_\varepsilon = 1 - \varepsilon/2$, where ε is very small, so that we target an *extreme* quantile. The (usual) empirical estimator (61) based on the order statistic quickly becomes impractical, because on average it needs $1/(1-p)$ observations lying *beyond* that quantile, that is tens of thousands when $p = 0.9999$, and millions or even billions as p creeps closer to 1. To bypass this data-hungry bottleneck, we cannot simply observe the tail; we must instead *extrapolate* it. The idea of the so-called Peaks-Over-Threshold (POT) method is to first extract the excesses $Y_i = Z^{(n-k_n+i)} - v_n$ for $i \in \{1, \dots, k_n\}$, by choosing for instance $v_n = Z^{(n-k_n)}$. With this choice, exactly k_n observations exceed v_n, so that $Y_i > 0$. Under the Pickands-Balkema-de Haan theorem the conditional excesses sequence $(Y_i)_{i \geq 1}$ converges in distribution to a random variable distributed according to a generalized Pareto distribution

$$F_Y(y) = \mathbb{P}(Y > y | Y > 0) = 1 - (1 + \xi y/\sigma)^{-1/\xi}, \qquad y > 0,\ 1 + \xi y/\sigma > 0,$$

where $\xi > 0$ is the shape (tail index) and $\sigma = \sigma_n$ is a scale that may depend on v_n. Estimators $(\widehat{\xi}_k, \widehat{\sigma}_k)$ are typically obtained by the maximum of likelihood estimator

$$(\widehat{\xi}_k, \widehat{\sigma}_k) := (\widehat{\xi}_{k_n}, \widehat{\sigma}_{k_n}) = \underset{(\xi,\sigma)}{\operatorname{argmax}} \sum_{i=1}^{k_n} \log\big(f_Y(y_i; \xi, \sigma)\big) =: \underset{(\xi,\sigma)}{\operatorname{argmax}}\ \ell(\xi, \sigma)$$

where f_Y is the density function associated to F_Y. Then, the *POT estimator* of q_p^+ is defined by

$$\widehat{q}_p^{\text{POT},+} = v_n + \frac{\widehat{\sigma}_k}{\widehat{\xi}_k}\left[\left(\frac{k_n}{np}\right)^{-\widehat{\xi}_k} - 1\right]. \tag{62}$$

Assumption (T1) (First-order tail) Let for $t > 1$, $U(t) = F^{\leftarrow}(1 - 1/t)$, be the generalized inverse function of F. Assume

$$\frac{U(tx)}{U(t)} \xrightarrow[t\to\infty]{} x^{\xi}, \qquad x > 0, \quad \text{with } \xi > -1/2.$$

Assumption (T2) (Second-order tail) There exist a function A such that $A(t) \to 0$ and a constant $\rho \le 0$ such that

$$\frac{U(tx)/U(t) - x^{\xi}}{A(t)} \xrightarrow[t\to\infty]{} x^{\xi}\,\frac{x^{\rho}-1}{\rho}, \qquad x > 0.$$

Assumption (K) Let a sequence $(k_n)_{n\ge 0}$ satisfying

$$k_n \xrightarrow[n\to\infty]{} \infty, \qquad \frac{k_n}{n} \xrightarrow[n\to\infty]{} 0, \qquad \sqrt{k_n}\,A\Big(\frac{n}{k_n}\Big) \xrightarrow[n\to\infty]{} 0.$$

Assumption (E) Set the threshold $v_n = Z^{(n-k)}$ and $\sigma_k := \sigma_{k_n}$, let $(\widehat{\xi}_k, \widehat{\sigma}_k)$ be the MLE of (ξ, σ). Assume

$$\sqrt{k_n}\begin{pmatrix} \widehat{\xi}_k - \xi \\ \widehat{\sigma}_k/\sigma_k - 1 \end{pmatrix} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}\big(0,\ \Sigma_{\xi,\sigma}\big), \qquad \text{and } \ \xi > -1/2.$$

We have the following result:

Theorem B.1. *Let $Z_1, \dots, Z_n$ be i.i.d. real-valued random variables with cumulative distribution function F. Under Assumptions (T1), (T2), (K) and (E), we have*

$$\sqrt{k_n}\,\frac{\widehat{q}_p^{\,\mathrm{POT},+} - q_p^+}{q_p^+} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0, V(p,\xi))$$

with $V(p,\xi) = \xi^2 + (1+\xi)^2(1-p)/p$.

We can define the left-hand version of the POT estimator:

$$\widehat{q}_p^{\,\mathrm{POT},-} = \widehat{q}_p^{\,\mathrm{POT},+} - \big(Z^{(n-k+1)} - Z^{(n-k)}\big),$$

as well as the average POT-estimator:

$$\widehat{q}_p^{\,\mathrm{POT}} = \frac{\widehat{q}_p^{\,\mathrm{POT},+} + \widehat{q}_p^{\,\mathrm{POT},-}}{2}$$

The gap between the adjacent order statistics that straddle the threshold satisfies (classical spacing theory) $Z^{(n-k+1)} - Z^{(n-k)} = O_{\mathbb{P}}(n^{-1})$. As $k_n = o(n)$ we get $\sqrt{k_n}(\widehat{q}_p^{\,\mathrm{POT},+} - \widehat{q}_p^{\,\mathrm{POT},-}) \xrightarrow{\mathbb{P}} 0$ so that the two POT estimators are indistinguishable on the $\sqrt{k_n}$-scale. Using Slutsky's lemma, we get then

$$\sqrt{k_n}\,\frac{\widehat{q}_p^{\,\mathrm{POT}} - q_p}{q_p} \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0, V(p,\xi)),$$

with $V(p,\xi) = \xi^2 + (1+\xi)^2(1-p)/p$. We can deduce a confidence interval for q_p:

Corollary B.2. *Let $\delta > 0$. Under the Assumptions of Theorem B.1, with probability at least $1-\delta$,*

$$q_p \in \left[\frac{\widehat{q}_p^{\,\mathrm{POT}}}{1 + z_{1-\delta/2}\sqrt{\widehat{V}_k/k_n}},\ \frac{\widehat{q}_p^{\,\mathrm{POT}}}{1 - z_{1-\delta/2}\sqrt{\widehat{V}_k/k_n}}\right], \tag{63}$$

where $z_{1-\delta/2}$ is the $(1-\delta/2)$-quantile of the standard normal distribution and $\widehat{V}_k = \widehat{\xi}_k^2 + (1+\widehat{\xi}_k)^2 p/p$.

C Proofs

C.1 Proof of Lemma 4.2 (case $\beta > 1$)

Let for any $(j,k) \in \mathbb{N} \times \mathbb{Z}$, $C_{j,k} := 2^{-\alpha j} D X_{j,k}$ (with $\beta > 1$). The proof in the case $\beta < 1$ easily adapts by using the Mills ratio (59).

Proceeding as in the proof of Proposition 4.1, we have, for any for any $x \in \mathbb{R}_+^*$,

$$\mathbb{P}(\sup_k |C_{j,k}| \geq x) = 1 - \Big[1 - \mathbb{P}(|X_{j,k}| \geq 2^{\alpha j} x/D)\Big]^{N_j},$$

where $N_j = C_1 2^{\eta j}$. Note that, from Mills ratio (56) applied to generalized Gaussians with $\beta > 1$, we have for any $t > 0$,

$$\frac{f_\beta(t)}{(2\beta - 1)t^{\beta-1}} \leq \mathbb{P}(|X_{j,k}| \geq t) \leq \frac{f_\beta(t)}{\beta t^{\beta-1}}.$$

Then, we get

$$1 - \Big[1 - \frac{1}{2\beta - 1} \frac{f_\beta(2^{\alpha j}x/D)}{(2^{\alpha j}x/D)^{\beta-1}}\Big]^{N_j} \leq \mathbb{P}(\sup_k |C_{j,k}| \geq x) \leq 1 - \Big[1 - \frac{1}{\beta} \frac{f_\beta(2^{\alpha j}x/D)}{(2^{\alpha j}x/D)^{\beta-1}}\Big]^{N_j}.$$

Define $A_j(c,x) = \Big[1 - c\mathrm{e}^{-(2^{\alpha j}x/D)^\beta}(2^{\alpha j}x/D)^{1-\beta}\Big]^{N_j}$ for any $c > 0$ and $x > 0$. Then we have

$$\mathbb{P}\big(\sup_k |C_{j,k}| \geq x\big) \geq 1 - A_j(c', x) \quad \text{and} \quad \mathbb{P}\big(\sup_k |C_{j,k}| \leq x\big) \geq A_j(c'', x), \tag{64}$$

with $c' = \kappa_\beta/(2\beta - 1)$ and $c'' = \kappa_\beta/\beta$. Note that $f_\beta(x) = \kappa_\beta 2^{-x^\beta/\log(2)}$. For any $x > 0$, define

$$H(x) = c2^{-h(x)/\log(2)}(h(x))^{(1-\beta)/\beta} \quad \text{and} \quad h(x) = (2^{\alpha j}x/D)^\beta.$$

Using that for $u > 0$ small we have

$$-u - \frac{u^2}{2(1-u)} \leq \log(1-u) \leq -u, \tag{65}$$

we get, for x large enough,

$$-C_1 2^{\eta j} H(x)\Big[1 + \frac{H(x)}{2(1 - H(x))}\Big] \leq \log(A_j(c,x)) \leq -C_1 2^{\eta j} H(x).$$

Choose x such that $h(x)/\log(2) = (2^{\alpha j}x/D)^\beta/\log(2) = \eta j + \theta \log_2(j)$ where θ will be chosen later on. Then, let us define

$$C_{\eta,3}(c) = c\eta^{(1-\beta)/\beta} C_1 \log(2)^{(1-\beta)/\beta} \quad \text{and} \quad C_{\eta,4}(c) = c\eta^{(1-\beta)/\beta} \log(2)^{(1-\beta)/\beta}. \tag{66}$$

We have

$$\begin{aligned} -C_1 2^{\eta j} H(x) &= -C_1 2^{\eta j} c 2^{-(\eta j + \theta \log_2(j))}\big[\log(2)(\eta j + \theta \log_2(j))\big]^{(1-\beta)/\beta} \\ &= -C_{\eta,3}(c)\eta^{-(1-\beta)/\beta} j^{-\theta}[\eta j + \theta \log_2(j)]^{(1-\beta)/\beta} \\ &\underset{j \to \infty}{\sim} -C_{\eta,3}(c) j^{-\theta + (1-\beta)/\beta}, \end{aligned}$$

and

$$c2^{-C_2(2^{\alpha j}x/D)^\beta}(2^{\alpha j}x/D)^{1-\beta} = c2^{-(\eta j+\theta\log_2(j))}\big[\log(2)(\eta j+\theta\log_2(j))\big]^{(1-\beta)/\beta}$$
$$\underset{j\to\infty}{\sim} -C_{\eta,4}(c)j^{-\theta+(1-\beta)/\beta},$$

Let $\delta = (\beta-1)/(2\beta) > 0$.

Case 1 Let $\theta = \theta_1 := 3(1-\beta)/(2\beta)$ We get, that for j large enough,

$$-C_{\eta,3}(c)j^\delta\left[1+\frac{C_{\eta,4}(c)j^\delta}{2(1-C_{\eta,4}(c)2^{-j})j^\delta}\right] \le \log(A_j(c,x)) \le -C_{\eta,3}(c)j^\delta.$$

Then, for j large enough, by (64),

$$\mathbb{P}(\sup_k |C_{j,k}| \ge x) \ge 1 - A_j(c',x) \ge 1 - \mathrm{e}^{-C_3(c')j^\delta},$$

where, as a reminder, x is such that $(2^{\alpha j}x/D)^\beta = \eta j\log(2) + \theta_1\log(j)$, $C_{\eta,3}(c)$ and $C_{\eta,4}(c)$ are given by (66). For sufficiently large j, the following holds with high probability:

$$\sup_k |C_{j,k}| \ge 2^{-\alpha j}D\big[\log(2)(\eta j+\theta_1\log_2(j))\big]^{1/\beta}.$$

Case 2 Let $\theta_2 = (1-\beta)/(2\beta)$ We get, that for j large enough,

$$-C_{\eta,3}(c)j^\delta\left[1+\frac{C_{\eta,4}(c)j^\delta}{2(1-C_{\eta,4}(c)j^\delta}\right] \le \log(A_j(c'',x)) \le -C_{\eta,3}(c'')j^{-\delta}.$$

Then, for j large enough, by (64),

$$\mathbb{P}\big(\sup_k |C_{j,k}| \le x\big) \ge A_j(c'',x) \ge \mathrm{e}^{-C_{\eta,3}(c'')\left[1+C_{\eta,4}(c)j^\delta/(2(1-C_{\eta,4}(c)j^\delta)\right]}$$

where, as a reminder, x is such that $(2^{\alpha j}x/D)^\beta = \eta j\log(2) + \theta_2\log(j)$, $C_{\eta,3}(c'')$ and $C_{\eta,4}(c'')$ are given in (66). For sufficiently large j, the following holds with high probability:

$$\sup_k |C_{j,k}| \le 2^{-\alpha j}D\big[\log(2)(\eta j+\theta\log_2(j))\big]^{1/\beta}.$$

We can then deduce a confidence interval for $\sup_k |C_{j,k}|$. Let $\varepsilon \in (0,1)$ be small.
First, let $j(\varepsilon)$ such that $\mathrm{e}^{-C_3(c')j(\varepsilon)^\delta} = \varepsilon/2$. From Case 1, we know that for any $j \ge j(\varepsilon)$,

$$\mathbb{P}\big(\sup_k |C_{j,k}| \ge x\big) \ge 1 - \mathrm{e}^{-C_3(c')j^\delta} = 1 - \frac{\varepsilon}{2},$$

where x is such that $(2^{\alpha j}x/D)^\beta = \eta j\log(2) + \theta_1\log(j)$.
Second, let $j_2(\varepsilon)$ such that $\exp\big(-C_{\eta,3}(c'')j_2(\varepsilon)^{-\delta}/(1-C_{\eta,4}(c'')2^{-j_2(\varepsilon)}j_2(\varepsilon)^{-\delta})\big) = 1-\varepsilon/2$. From Case 2, we know that for any $j \ge j_2(\varepsilon)$,

$$\mathbb{P}\big(\sup_k |C_{j,k}| \le x\big) \ge \mathrm{e}^{-C_{\eta,3}(c'')\left[1+C_{\eta,4}(c)j^\delta/(2(1-C_{\eta,4}(c)j^\delta)\right]} = 1 - \frac{\varepsilon}{2},$$

where x is such that $(2^{\alpha j}x)^\beta = \eta j\log(2) + \theta\log_2(j)$. Then for any $j \ge \max(j_1(\varepsilon), j_2(\varepsilon))$, we have with probability at least $1-\varepsilon$,

$$\sup_k |C_{j,k}| \in \Big[2^{-\alpha j}D(\eta j\log(2)+\theta_1\log(j))\,,\, 2^{-\alpha j}D(\eta j\log(2)+\theta_2\log_2(j))\Big].$$

Hence the result.

C.2 Proof of Lemma 4.2 (case $\beta = 1$)

Proceeding as in the proof of Proposition 4.1, we have, for any for any $x \in \mathbb{R}_+^*$,

$$\mathbb{P}(\sup_k |C_{j,k}| \geq x) = 1 - \Big[1 - \mathbb{P}(|X_{j,k}| \geq 2^{\alpha j} x/D)\Big]^{N_j} \quad \text{and} \quad \mathbb{P}(\sup_k |C_{j,k}| \leq x) = \Big[1 - \mathbb{P}(|X_{j,k}| \geq 2^{\alpha j} x/D)\Big]^{N_j}$$

where $N_j = C_1 2^{\eta j}$. Note that for any $t > 0$,

$$\mathbb{P}(|X_{j,k}| \geq t) = \frac{e^{-t}}{2}$$

Then, we get

$$\mathbb{P}(\sup_k |C_{j,k}| \geq x) = 1 - \Big[1 - \exp(-2^{\alpha j} x/D)\Big]^{N_j} =: 1 - A_j(x) \quad \text{and} \quad \mathbb{P}(\sup_k |C_{j,k}| \leq x) = A_j(x).$$

Using (65), we get

$$-H(x)\Big[1 + \frac{H(x)}{2(1 - H(x))}\Big] \leq \log(A_j(x)) \leq -H(x)$$

with

$$H(x) := C_1 2^{\eta j} e^{-2^{\alpha j} x/D} = C_1 2^{\eta j} 2^{-\eta j + \theta \log_2(j)} = C_1 j^{-\theta}.$$

Then, we have

$$\mathbb{P}(\sup_k |C_{j,k}| \geq x) \geq 1 - e^{-C_1 j^{-\theta_1^1}} \quad \text{and} \quad \mathbb{P}(\sup_k |C_{j,k}| \leq x) \geq e^{-C_1 j^{-\theta_2^1}\left[1 - C_1 j^{-\theta_2^1}/(2(1 - C_1 j^{-\theta_2^1}))\right]},$$

where we choose $\theta_1^1 = -1$ and $\theta_2^1 = 1$. Consider $j_1(\varepsilon), j_2(\varepsilon)$ such that

$$1 - e^{-C_1 j_1(\varepsilon)} = 1 - \varepsilon/2 \quad \text{and} \quad e^{-C_1 j_2(\varepsilon)^{-1}\left[1 - C_1 j_2(\varepsilon)^{-1}/(2(1 - C_1 j_2(\varepsilon)^{-1}))\right]} = 1 - \varepsilon/2.$$

Hence the result.

References

[1] F. Abramovich, T. Sapatinas, and B. Silverman. Wavelet thresholding via a bayesian approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60(4):725–749, 1998.

[2] P. Abry, P. Gonçalvès, and P. Flandrin. *Wavelets, spectrum estimation and $1/f$ processes*, chapter 103. Springer-Verlag, New York, 1995. Wavelets and Statistics, Lecture Notes in Statistics.

[3] P. Abry, S. Jaffard, and H. Wendt. Irregularities and scaling in signal and image processing: Multifractal analysis. *Benoit Mandelbrot: A Life in Many Dimensions, M. Frame and N. Cohen, Eds., World scientific publishing*, pages 31–116, 2015.

[4] Patrice Abry, Patrick Flandrin, Murad S Taqqu, Darrl Veitch, et al. Self-similarity and long-range dependence through the wavelet lens. *Theory and applications of long-range dependence*, 1:527–556, 2003.

[5] Demi Allen and Edouard Daviaud. A survey of recent extensions and generalisations of the mass transference principle. In *Conference of Fractals and Related Fields*, pages 17–47. Springer, 2022.

[6] J.-P. Antoine, R. Murenzi, P. Vandergheynst, and S. T. Ali. *Two-dimensional wavelets and their relatives*. Cambridge University Press, 2008.

[7] A. Antoniadis, J. Bigot, and T. Sapatinas. Wavelet estimators in nonparametric regression: a comparative simulation study. *Journal of statistical software*, 6:1–83, 2001.

[8] Alain Arneodo, Emmanuel Bacry, and Jean-François Muzy. Random cascades on wavelet dyadic trees. *Journal of Mathematical Physics*, 39(8):4142–4164, 1998.

[9] J.M. Aubry and S. Jaffard. Random wavelet series. *Communications In Mathematical Physics*, 227(3):483–514, 2002.

[10] E. Bacry, J. Delour, and J.F. Muzy. Multifractal random walk. *Phys. Rev. E*, 64(2):026103, 2001.

[11] E. Bacry and J.F. Muzy. Log-infinitely divisible multifractal processes. *Commun. Math. Phys.*, 236:449–475, 2003.

[12] Emmanuel Bacry and Jean François Muzy. Log-infinitely divisible multifractal processes. *Communications in Mathematical Physics*, 236(3):449–475, 2003.

[13] J. Barral and B. Mandelbrot. Multifractal products of cylindrical pulses. *Probability Theory and Related Fields*, 124(3):409–430, 2002.

[14] Julien Barral and Stéphane Seuret. Heterogeneous ubiquitous systems in d and hausdorff dimension. *Bulletin of the Brazilian Mathematical Society, New Series*, 38:467–515, 2007.

[15] Victor Beresnevich and Sanju Velani. A mass transference principle and the duffin-schaeffer conjecture for hausdorff measures. *Annals of mathematics*, pages 971–992, 2006.

[16] R. Buccigrossi and E. Simoncelli. Image compression via joint statistical characterization in the wavelet domain. *IEEE transactions on Image processing*, 8(12):1688–1701, 1999.

[17] Laurent E Calvet and Adlai J Fisher. How to forecast long-run volatility: Regime switching and the estimation of multifractal processes. *Journal of Financial Econometrics*, 2(1):49–83, 2004.

[18] Pierre Chainais. Infinitely divisible cascades to model the statistics of natural images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(12):2105–2119, 2007.

[19] H. Chipman, E. Kolaczyk, and R. McCulloch. Adaptive bayesian wavelet shrinkage. *Journal of the American Statistical Association*, pages 1413–1421, 1997.

[20] Sébastien Combrexelle, Herwig Wendt, Nicolas Dobigeon, Jean-Yves Tourneret, Stephen McLaughlin, and Patrice Abry. Bayesian estimation of the multifractality parameter for image

texture using a whittle approximation. *IEEE Transactions on Image Processing*, 24(8):2540–2551, 2015.

[21] Matthew S Crouse, Robert D Nowak, and Richard G Baraniuk. Wavelet-based statistical signal processing using hidden markov models. *IEEE Transactions on signal processing*, 46(4):886–902, 2002.

[22] J. Daníelsson, L. de Haan, L. Peng, and C. G. de Vries. Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of Multivariate Analysis*, 76(2):226–248, 2001.

[23] D. Donoho and I. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *biometrika*, 81(3):425–455, 1994.

[24] Matteo Dora and David Holcman. Adaptive single-channel eeg artifact removal for real-time clinical monitoring. *IEEE Transactions on Neural Systems and Rehabilitation EngineeringF*, 2022.

[25] Matteo Dora, Stéphane Jaffard, and David Holcman. The wqn algorithm for eeg artifact removal in the absence of scale invariance. *IEEE Transactions on Signal Processing*, 72:2663–2676, 2024.

[26] Arnaud Durand. Random wavelet series based on a tree-indexed markov chain. *Communications in mathematical physics*, 283(2):451–477, 2008.

[27] Céline Esser, Thelma Lambert, and Vedel Béatrice. The p-spectrum of random wavelet series. *arXiv preprint arXiv:2510.00622*, 2025.

[28] David J Field. What is the goal of sensory coding? *Neural computation*, 6(4):559–601, 1994.

[29] P. Flandrin. Wavelet analysis and synthesis of fractional Brownian motions. *IEEE Trans. Info. Theory*, 38:910–917, 1992.

[30] Uriel Frisch and Andreĭ Kolmogorov. *Turbulence: the legacy of AN Kolmogorov.*

[31] Fabrice Gamboa and Jean-Michel Loubes. Wavelet estimation of a multifractal function. *Bernoulli*, 11(2):221–246, 2005.

[32] Fabrice Gamboa and Jean-Michel Loubes. Estimation of parameters of a multifractal process. *Test*, 16(2):383–407, 2007.

[33] David J Heeger and James R Bergen. Pyramid-based texture analysis/synthesis. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 229–238, 1995.

[34] Andreas Horst, Thomas Jahn, and Felix Voigtlaender. Besov regularity of random wavelet series. *arXiv preprint arXiv:2411.18155*, 2024.

[35] Jacques Istas and Gabriel Lang. Quadratic variations and estimation of the local hölder index of a gaussian process. In *Annales de l'Institut Henri Poincare (B) probability and statistics*, volume 33, pages 407–436. Elsevier, 1997.

[36] Garry Jacyna, Damon Frezza, David M Slater, and James R Thompson. An improved model of wavelet leader covariance for estimating multifractal properties. *ACM Transactions on Modeling and Computer Simulation*, 34(1):1–22, 2023.

[37] S. Jaffard. Multifractal formalism for functions, Part 1: Results valid for all functions, Part 2: Selfsimilar functions. *SIAM Journal of Mathematical Analysis*, 28(4):944–998, 1997.

[38] S. Jaffard. The multifractal nature of Levy processes. *Probability Theory and Related Fields*, 114(2):207–227, 1999.

[39] S. Jaffard. On lacunary wavelet series. *Annals of Applied Probability*, 10:313–329, 2000.

[40] S. Jaffard. Wavelet techniques in multifractal analysis. In M. Lapidus and M. van Frankenhuijsen, editors, *Fractal Geometry and Applications: A Jubilee of Benoît Mandelbrot, Proc. Symp. Pure Math.*, volume 72(2), pages 91–152. AMS, 2004.

[41] S. Jaffard. Beyond Besov spaces, part 2: Oscillation spaces. *Constructive Approximation*, 21(1):29–61, 2005.

[42] S. Jaffard and Y. Meyer. Wavelet methods for pointwise regularity and local oscillations of functions. *Memoirs of the A.M.S.*, 123(587), 1972.

[43] S. Jaffard, S. Seuret, H. Wendt, R. Leonarduzzi, and P. Abry. Multifractal formalisms for multivariate analysis. *Proc. Royal Society A*, 475(2229), 2019.

[44] Stephane Jaffard. Wavelet techniques in multifractal analysis. In *Proceedings of symposia in pure mathematics*, volume 72, pages 91–152, 2004.

[45] I. M. Johnstone. Minimax bayes, asymptotic minimax and sparse wavelet priors. In *Statistical Decision Theory and Related Topics V*, pages 303–326. Springer, 1994.

[46] Z. Kadkhodaie, F. Guth, S. Mallat, and E. Simoncelli. Learning multi-scale local conditional probability models of images. *arXiv preprint arXiv:2303.02984*, 2023.

[47] J.-P. Kahane. *Some random series of functions*. Cambridge University Press, 1985.

[48] J. W. Kantelhardt, S. A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, and H. E. Stanley. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A*, 316(1):87–114, 2002.

[49] Christopher J Keylock. Multifractal surrogate-data generation algorithm that preserves pointwise hölder regularity structure, with initial applications to turbulence. *Physical Review E*, 95(3):032123, 2017.

[50] A. Kovac and B. Silverman. Extending the scope of wavelet regression methods by coefficient-dependent thresholding. *Journal of the American Statistical Association*, 95(449):172–183, 2000.

[51] B. Lashermes, S. Jaffard, and P. Abry. Wavelet leader based multifractal analysis. *2005 Ieee International Conference On Acoustics, Speech, and Signal Processing, Vols 1-5*, pages 161–164, 2005.

[52] Paul Lévy. *Processus Stochastiques et Mouvement Brownien.* Gauthier-Villars, Paris, 1948.

[53] Jean-Michel Loubes and Davy Paindaveine. Local asymptotic normality property for lacunar wavelet series multifractal model. *ESAIM: Probability and Statistics*, 15:69–82, 2011.

[54] S. Lyu and E. Simoncelli. Modeling multiscale subbands of photographic images with fields of gaussian scale mixtures. *IEEE Transactions on pattern analysis and machine intelligence*, 31(4):693–706, 2008.

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

[56] B. Mandelbrot. *Fractals and scaling in finance.* Selected Works of Benoit B. Mandelbrot. Springer-Verlag, New York, 1997. Discontinuity, concentration, risk, Selecta Volume E, With a foreword by R.E. Gomory.

[57] Y. Meyer. *Ondelettes et Opérateurs.* Hermann, Paris, 1990. English translation, *Wavelets and operators*, Cambridge University Press, 1992.

[58] Y. Meyer, F. Sellan, and M. Taqqu. Wavelets, generalized white noise and fractional integration: the synthesis of fractional brownian motion. *Journal of Fourier Analysis and Applications*, 5:465–494, 1999.

[59] J. Mills. Table of the ratio: area to bounding ordinate, for any portion of normal curve. *Biometrika*, 18(3/4):395–400, 1926.

[60] Rudy Morel, Stéphane Mallat, and Jean-Philippe Bouchaud. Path shadowing monte carlo. *Quantitative Finance*, 24(9):1199–1225, 2024.

[61] Jean-François Muzy, Emmanuel Bacry, and Alain Arneodo. Multifractal formalism for fractal signals: The structure-function approach versus the wavelet-transform modulus-maxima method. *Physical review E*, 47(2):875, 1993.

[62] Wejdene Ben Nasr, Hélène Halconruy, and Stéphane Jaffard. Questioning normality: A study of wavelet leaders distribution: W. ben nasr et al. *Statistics and Computing*, 36(1):28, 2026.

[63] Bruno A Olshausen and David J Field. Sparse coding of sensory inputs. *Current opinion in neurobiology*, 14(4):481–487, 2004.

[64] G. Parisi and U. Frisch. Fully developed turbulence and intermittency. In M. Ghil, R. Benzi, and G. Parisi, editors, *Turbulence and Predictability in geophysical Fluid Dynamics and Climate Dynamics*, Proc. of Int. School, page 84, Amsterdam, 1985. North-Holland.

[65] J. Portilla, V. Strela, M. Wainwright, and E. Simoncelli. Adaptive Wiener denoising using a Gaussian scale mixture model in the wavelet domain. In *Proceedings 2001 International Conference on Image Processing (Cat. No. 01CH37205)*, volume 2, pages 37–40. IEEE, 2001.

[66] Javier Portilla, Vasily Strela, Martin J Wainwright, and Eero P Simoncelli. Image denoising using scale mixtures of gaussians in the wavelet domain. *IEEE Transactions on Image processing*, 12(11):1338–1351, 2003.

[67] Radu Ranta, Valerie Louis-Dorr, Christian Heinrich, and Didier Wolf. Iterative wavelet-based denoising methods and robust outlier detection. *IEEE Signal Processing Letters*, 12(8):557–560, 2005.

[68] Justin K Romberg, Hyeokho Choi, and Richard G Baraniuk. Bayesian tree-structured image modeling using wavelet-domain hidden markov models. *IEEE Transactions on image processing*, 10(7):1056–1068, 2001.

[69] Abd-Krim Seghouane. An adaptive bayesian wavelet thresholding approach to multifractal signal denoising. In *AIP Conference Proceedings*, volume 735, pages 297–304. American Institute of Physics, 2004.

[70] Levent Sendur and Ivan W Selesnick. Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency. *IEEE Transactions on signal processing*, 50(11):2744–2756, 2002.

[71] Ali Siahkoohi, Rudy Morel, Randall Balestriero, Erwan Allys, Grégory Sainton, Taichi Kawamura, and Maarten V de Hoop. Martian time-series unraveled: A multi-scale nested approach with factorial variational autoencoders. *arXiv preprint arXiv:2305.16189*, 2023.

[72] E. Simoncelli. Statistical models for images: Compression, restoration and synthesis. In *Conference Record of the Thirty-First Asilomar Conference on Signals, Systems and Computers (Cat. No. 97CB36136)*, volume 1, pages 673–678. IEEE, 1997.

[73] E. Simoncelli, E.and Adelson. Noise removal via bayesian wavelet coring. In *Proceedings of 3rd IEEE International Conference on Image Processing*, volume 1, pages 379–382. IEEE, 1996.

[74] Skander Soltani, Patrice Simard, and Daniel Boichu. Estimation of the self-similarity parameter using the wavelet transform. *Signal Processing*, 84(1):117–123, 2004.

[75] V. Strela, J. Portilla, and E. Simoncelli. Image denoising using a local Gaussian scale mixture model in the wavelet domain. In *Wavelet Applications in Signal and Image Processing VIII*, volume 4119, pages 363–371. SPIE, 2000.

[76] Darryl Veitch and Patrice Abry. A wavelet-based joint estimator of the parameters of long-range dependence. *IEEE Transactions on Information Theory*, 45(3):878–897, 2002.

[77] B. Vidakovic. *Statistical modeling by wavelets*. John Wiley & Sons, 2009.

[78] M. Wainwright and E. Simoncelli. Scale mixtures of gaussians and the statistics of natural images. *Advances in neural information processing systems*, 12, 1999.

[79] Martin J Wainwright, Eero P Simoncelli, and Alan S Willsky. Random cascades on wavelet trees and their use in analyzing and modeling natural images. *Applied and Computational Harmonic Analysis*, 11(1):89–123, 2001.

[80] H. Wendt, P. Abry, S. Jaffard, H. Ji, and Z. Shen. Wavelet leader multifractal analysis for texture classification. In *Proc. IEEE Int. Conf. Image Proc. (ICIP)*, Cairo, Egypt, 2009.

[81] H. Wendt, S.G. Roux, P. Abry, and S. Jaffard. Wavelet leaders and bootstrap for multifractal analysis of images. *Signal Processing*, 89(6):1100–1114, 2009.

[82] Herwig Wendt, Sébastien Combrexelle, Yoann Altmann, Jean-Yves Tourneret, Stephen McLaughlin, and Patrice Abry. Multifractal analysis of multivariate images using gamma markov random field priors. *SIAM Journal on Imaging Sciences*, 11(2):1294–1316, 2018.

[83] Herwig Wendt, Nicolas Dobigeon, Jean-Yves Tourneret, and Patrice Abry. Bayesian estimation for the multifractality parameter. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 6556–6560. IEEE, 2013.