Parametric estimation of a multifractal function

Jean-Michel Loubes*
and
Davy Paindaveine†‡

May 2007

Abstract

Multifractal functions are widely used to model irregular signals such as turbulence, data stream or road traffic. Here, we consider multifractal functions defined as lacunar wavelet series observed in a white noise model. These random functions are statistically characterized by two parameters. The first parameter governs the lacunarity of the wavelet coefficients while the second one governs its intensity. In this paper, we establish the local and asymptotic normality (LAN) of the model, with respect to this couple of parameters. This allows for providing an estimator for the lacunarity parameter, that is asymptotically optimal in the Le Cam sense, as well as optimal (still in the Le Cam sense) tests on the intensity parameter.

AMS 1980 subject classification : 60G17, 62G07.
Key words and phrases : Multifractal analysis, Wavelet Bases, LAN

1 Introduction

This work deals with statistical analysis of a probabilistic model for multifractal functions. We observe in a white noise regression framework a multifractal function defined as a realization of a random lacunar wavelet series. This series is defined by two parameters, η a lacunarity parameter and α an intensity parameter. These two parameters govern the multifractal properties of such functions. We prove that this model is LAN under some conditions on the range of α and discuss optimality (in the Le Cam sense) of the estimators of the parameters of interest. We also derive locally and asymptotically optimal tests.

In the last decade much emphasis has been placed on modelling of very irregular signals. Indeed it is well known that standard Brownian motion fails in explaining certain time series arising from finance, biology or turbulence theory. In [12] for example is emphazised that LAN/WAN traffic traces for which the aggregation level is insufficient presents multifractal properties. Such phenomenon is also pointed out for the stock market in [7]. So, roughly speaking, a multifractal function is a function whose local Hölder regularity index is not constant. That means that the function may be very regular in some areas while it is very irregular in

*CNRS and Université Montpellier 2.
†E.C.A.R.E.S., Institut de Recherche en Statistique, and Département de Mathématique, Université Libre de Bruxelles, Belgium. Paindaveine is also member of ECORE, the recently created association between CORE and ECARES.
‡Research supported by a Crédit d'Impulsion of the Fonds National de la Recherche Scientifique, Communauté française de Belgique.

others. Such functions, with rapid changes of regularity, have been first introduced to model physical phenomena by [4], or road traffic or data traffic on networks by [17].

To construct multifractal functions, [2], [3] or [18] and [19] have recently shown that some lacunary random series built on wavelets have multifractal properties. As a matter of fact, consider a wavelet basis ψ_{jk}, $j \geq 0$, $k = 0, \dots, 2^j - 1$ and a couple of parameters $\boldsymbol{\theta} = (\alpha, \eta)^T \in \Theta := (0,1)^2$. Then, draw the random wavelet coefficient w_{jk}, $k = 0, \dots, 2^j - 1$ following the rescaled Bernoulli distribution:

$$w_{jk} \sim 2^{(\eta-1)j}\delta_{2^{-\alpha j}} + (1 - 2^{(\eta-1)j})\delta_0, \tag{1.1}$$

where δ is the Dirac measure. Then, the corresponding random function $f_n = \sum_{j=1}^{\log_2 n} \sum_{k=0}^{2^j-1} w_{jk}\psi_{jk}$ converges to a random function having multifractal properties. The multifractal properties of general sparse random series have been studied in [3], [18], [11], [1] and depend on the choice of both the lacunarity coefficient η and the intensity parameter α.

In this paper we consider n observations of a multifractal function together with a white noise. Hence after discretization, the statistical model is the following

$$d_{jk} = w_{jk} + \epsilon_{jk},\ j = 0, \dots, j_1 = \log_2(n),\ k = 0, \dots, 2^j - 1. \tag{1.2}$$

w_{jk} are the true wavelet coefficients while ϵ_{jk} are i.i.d centered Gaussian random variables with variance σ^2, taken independent from the w_{jk}. So, we observe the triangular array $\mathbf{d}_n = (d_{jk})_{1 \leq j \leq j_1, 0 \leq k \leq 2^j - 1}$ of independent random variables. For any j, k the distribution of d_{jk} is the Gaussian mixture:

$$d_{jk} \sim 2^{(\eta-1)j}\mathcal{N}(2^{-\alpha j}, \frac{\sigma^2}{n}) + (1 - 2^{(\eta-1)j})\mathcal{N}(0, \frac{\sigma^2}{n}), \tag{1.3}$$

where $\boldsymbol{\theta} = (\alpha, \eta)^T \in (0,1)^2$ is the unknown parameter vector, $\sigma > 0$ is known and, as usual, $\mathcal{N}(m, \xi^2)$ denotes the Gaussian distribution with mean m and standard deviation ξ.

In the statistical literature, very few is known about the estimation of multifractal functions. In a previous work [9], nonparametric estimation of a realization of such a random process is tackled in a Bayesian setting. Roughly speaking, the Bayesian nonparametric posterior estimate is built on a ranked thresholding procedure, and its rate of convergence is different from the usual rates in nonparametric estimation of smooth functions. In the parametric framework, a first study to construct estimators of the unknown parameters of the mixture (1.1) is conducted in [10]. The estimators of the parameter of interests are asymptotically normal but their rate of convergence is not the parametric rate of convergence but rather depends on the values of $\boldsymbol{\theta}$. Hence, in this work, we focus here on statistical inference for empirical estimation of the hyperparameters and on the efficiency of this estimation procedure. More precisely, we aim at proving Local Asymptotic Normality (LAN) for the observation model and comparing the different rates of convergence for some estimators. In particular, we prove optimality in the Le Cam sense of an estimator of the lacunarity paramter built only with the wavelet coefficients of the last resolution level $j_1(n)$, which improves results in [10].

The paper falls into the following parts. After recalling the properties of the model we consider in Section 2 and the corresponding statistical model in Section 3, we prove, for some values of these parameters, local asymptotic normality for the associated family of distributions in Section 4. This enable us to prove, in Section 5, optimality of an estimator of the lacunarity parameter and to derive optimal tests on the intensity parameter in Section 6. Finally, Section 7 presents some simulation results.

2 Multifractal wavelet models

In this paper, we will always work with functions on $[0,1]$. To begin with, let us first introduce some useful definitions around multifractal functions.
The Haussdorf dimension $d_H(A)$ of a set A is defined as follows. Let $C(A,\delta)$ be the set of all $\delta-$covering (C_i) of A with open sets C_i of diameter $|C_i| \leq \delta$. Let also

$$\begin{aligned} H_{s,\delta}(A) &= \inf_{(C_i)_i \in C(A,\delta)} \sum_i |C_i|^s \\ H_s(A) &= \lim_{\delta \to 0} H_{s,\delta}(A) \\ d_H(A) &= \inf\{s \ : H_s(A) = 0\} \end{aligned}$$

Definition 1 *Let f be a function on $[0,1]$.*

1) Let $x_0 \in [0,1]$ and $h \geq 0$, the set $C^h(x_0)$ is the set of all functions f on $[0,1]$ such that there exist a polynomial P of degree less than h and a neighborhood V of x_0 satisfying

$$|f(x) - P(x)| = O\left(x - x_0\right)^h \quad (x \in V).$$

2) Let $h_f(x_0) = \sup\{h \geq 0, f \in C^h(x_0)\}$ and

$$E_h = \{x \in [0,1],\ h_f(x) = h\} \quad (h \geq 0).$$

The spectrum of singularity d_f of f is the function on $\mathbb{R}^+$ which associates to each $h \geq 0$ the Haussdorf dimension of the set E_h.

Multifractal analysis of a function was first introduced in a physical framework in [8]. Given a function f, one of the main goal of this analysis is the computation of the spectrum of singularities d_f. When d_f does not vanish in at least two points we say that f is multifractal. The spectrum of singularities of a function f is a relevant quantity to describe the smoothness variation of f. Multifractal functions can be constructed using their decomposition into an appropriate wavelet basis as described in [2] and [11]. Since we restrict attention to functions on $[0,1]$, we will only consider periodized wavelets in the Schwartz class. This implies that all moments of the wavelets vanish. In an equivalent way, we could have used compactly supported wavelets as it is stated in [11] but the results are heavier to state.
Following the construction provided in [16] we define a wavelet $\tilde{\psi}$ in the Schwartz class, and construct the periodized wavelet $\psi(x) = \sum_{l \in \mathbb{Z}} \tilde{\psi}(x-l)$. The functions $\psi_{jk} = \psi(2^j x - l), \quad \forall j \in \mathbb{N},\ k \in [0, 2^j - 1]$ are obtained from the first wavelet by dilatation and translation. Then $(2^{j/2}\psi_{jk})_{(j,k)}$ provides an orthonormal basis of the Hilbert space $L^2([0,1])$ (observe the presence of a normalizing factor $2^{j/2}$). Let $f \in L^2([0,1])$ on one hand, its wavelet coefficients may be computed as

$$w_{jk} = 2^j \int_0^1 f(t)\psi_{jk}(t)dt \ \ (j \in \mathbb{N},\ k \in [0, 2^j - 1]).$$

On the other hand, f may be reconstructed using its wavelet coefficients

$$f = \sum_{j=0}^{\infty} \sum_{k=0}^{2^j - 1} w_{jk}\psi_{jk} \tag{2.1}$$

Using this wavelet representation, we now turn on the construction of random functions exhibiting multifractal properties. This will be done considering sparse random wavelet series. Let ρ_j, $j \in \mathbb{N}^*$ be a repartition function on $\mathbb{R}$. Further, let $Z_j = (Z_{jk})_{k=0,\ldots,2^j-1}$ be 2^j independent random vectors having common distribution ρ_j. Now, build a random function F using the reconstruction formula (2.1) where for any $j \in \mathbb{N}^*, k = 0, \ldots, 2^j - 1$ $|w_{jk}| = 2^{-jZ_{jk}}$. To study the multifractal properties of the random function F, [3], [11] introduced the following functions:

$$\begin{aligned}
\tilde{\rho}(\alpha, \epsilon) &= \limsup_{j\to\infty} \frac{\log_2(2^j \rho_j[\alpha - \epsilon, \alpha + \epsilon])}{j} \\
&= 1 + \limsup_{j\to\infty} \frac{\log \mathbf{P}(Z_{jk} \in [\alpha - \epsilon, \alpha + \epsilon])}{j} \\
\tilde{\rho}(\alpha) &= \inf_{\epsilon > 0} \tilde{\rho}(\alpha, \epsilon)
\end{aligned}$$

Under some assumptions on $(\rho_j)_{j\in\mathbb{N}^*}$, which can be found in [11], or [3], Jaffard et al prove that the spectrum of singularity of F can be calculated. Indeed, they show that, for all $h > 0$

$$d_F(h) = h \sup_{\alpha\in(0,h]} \frac{\tilde{\rho}(\alpha)}{\alpha} \text{ (a.s.).} \tag{2.2}$$

In this paper, we focus on the simplest statistical model derived from the ones described in the last paragraph. Let $(X_{jk})_{j\in\mathbb{N}^*,\, k=0,\ldots,2^j-1}$ be a triangular array of independent Bernoulli random variables: for $\eta \in (0, 1)$

$$\mathbf{P}(X_{jk} = 1) = 1 - \mathbf{P}(X_{jk} = 0) = 2^{(\eta-1)j}.$$

Further, take for $j \in \mathbb{N}^*$ and $k = 0, \ldots, 2^j - 1$ random wavelet coefficients $w_{jk} = 2^{-\alpha j} X_{jk}$ for $\alpha \in (0, 1)$. So we get

$$w_{jk} \sim 2^{(\eta-1)j} \delta_{2^{-\alpha j}} + (1 - 2^{(\eta-1)j})\delta_0. \tag{2.3}$$

The corresponding function f is then defined by its wavelet decomposition into the basis ψ_{jk} by $f = \sum_{j=0}^{\infty} \sum_{k=0}^{2^j-1} w_{jk}\psi_{jk}$. So, this simple multifractal model is characterized by two parameters η and α in $(0, 1)$. On one hand η describes the lacunarity of the wavelet series (that is its sparsity). On the other hand the coefficient α is inversely proportional to the intensity of the value of the wavelet coefficients. These parameters completely characterizes the spectrum of singularity of the random functions involved. Generating a function with this method may seem too restrictive. However, such processes appear naturally when studying multifractal processes. As a matter of fact, Jaffard and al. in their work ([3], [1], [11]) use such modelization and they show in [3] that the spectrum of singularity of the function f is almost surely

$$d_f(h) = \frac{1-\eta}{\alpha} h, \ \forall h \in [\alpha, \frac{\alpha}{1-\eta}]. \tag{2.4}$$

In figure 1 we plot a realization of such a multifractal function. The lacunarity coefficient is $\eta = 0.4$ while $\alpha = 0.3$.

In this paper, we will build and study estimators of these two parameters when the observation is a wavelet series observed in a white noise model.

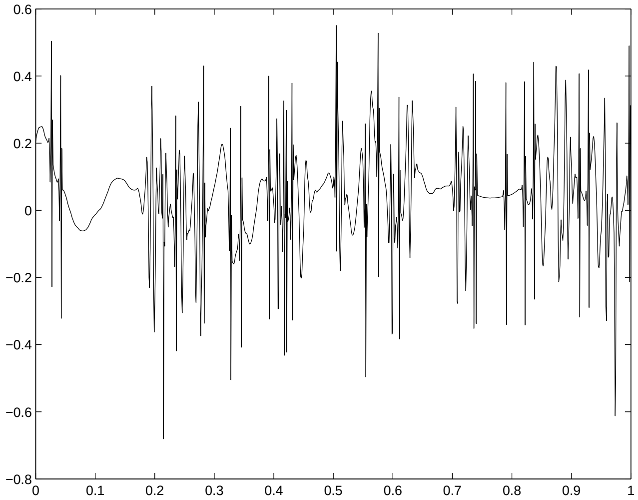

Figure 1: Multifractal function

3 The model

We aim at estimating the parameters η and α of a multifractal function f observed with some measurement errors. Throughout the paper, we make the assumption that the wavelet basis, in which the function has the decomposition (2.3), is known.

In the white noise model, we observe the wavelet coefficients w_{jk} of the function f together with a Gaussian white noise ϵ_{jk} having variance $\frac{\sigma^2}{n}$ where n is the number of observations. We assume that the observations are dyadic and $n = 2^{j_1}$, $(j_1 > 0)$. Recall that the wavelet coefficients are obtained from discrete regression model

$$Y_i = f(i/2^{j_1}) + W_i,\ i = 1, \dots, n \tag{3.1}$$

with $W_i \overset{\text{i.i.d}}{\sim} \mathcal{N}(0, \sigma^2)$, by performing the Discrete Wavelet Transform (DWT). Such transform is performed by Mallat's fast algorithm [14] that requires only $O(n)$ operations. Hence, the observations are drawn from the following model:

$$d_{jk} = w_{jk} + \epsilon_{jk},\ j = 1, \dots, j_1,\ k = 0, \dots, 2^j - 1$$

As a result, the law of the observed coefficients d_{jk} is determined by the prior given by the model (2.3):

$$d_{jk} \sim 2^{(\eta-1)j}\mathcal{N}(2^{-\alpha j}, \frac{\sigma^2}{n}) + (1 - 2^{(\eta-1)j})\mathcal{N}(0, \frac{\sigma^2}{n}) \tag{3.2}$$

In this paper, we will only consider the last level in the wavelet decomposition. Indeed, using all the levels, as done in [10] does not help in the estimation issue and is more a drawback, leading to complicate expressions for the estimators.

Hence consider the rescaled observed wavelet coefficients,

$$X_{nj} = \sqrt{n} w_{\log_2(n)j},\ j = 1, \dots, n.$$

We then observe $X_{n1}, \dots, X_{nn}$, a triangular array of observations, where X_{ni}, $i = 1, \dots, n$ are i.i.d. with common density

$$f_{\boldsymbol{\theta}} := n^{\eta-1}\phi_\alpha + (1 - n^{\eta-1})\phi, \quad \boldsymbol{\theta} = (\eta, \alpha)' \in (0,1) \times (0,1), \tag{3.3}$$

where ϕ_α (resp., ϕ) stands for the density of a Gaussian random variable with mean $n^{\frac{1}{2}-\alpha}$ (resp., 0) and variance σ^2. For fixed n, the observations are therefore generated by a Gaussian mixture of the form $n^{\eta-1}\mathcal{N}(n^{\frac{1}{2}-\alpha},\sigma^2)+(1-n^{\eta-1})\mathcal{N}(0,\sigma^2)$. We aim at studying the statistical problem of estimating the parameters $\boldsymbol{\theta}$ using the observations $X_{n1},\ldots,X_{nn}$.

4 Local asymptotic normality (LAN)

Partition the parametric space into $\boldsymbol{\Theta}=\boldsymbol{\Theta}^-\cup\boldsymbol{\Theta}^\pm\cup\boldsymbol{\Theta}^+$, with $\boldsymbol{\Theta}^-=(0,1)\times(0,1/2)$, $\boldsymbol{\Theta}^\pm=(0,1)\times\{1/2\}$, and $\boldsymbol{\Theta}^+=(0,1)\times(1/2,1)$. In this section, we prove that the family of distributions $\mathcal{P}_n^-:=\{\mathrm{P}^n_{\boldsymbol{\theta}}\,|\,\boldsymbol{\theta}\in\boldsymbol{\Theta}^-\}$ is LAN.

For any bounded sequence $\boldsymbol{\tau}_n=(s_n,t_n)'$, consider the corresponding local perturbation $\boldsymbol{\theta}_n=\boldsymbol{\theta}+\boldsymbol{\nu}_n(\boldsymbol{\theta})\boldsymbol{\tau}_n=(\eta_n,\alpha_n)'$ of the parameter value $\boldsymbol{\theta}=(\eta,\alpha)'$, where $\boldsymbol{\nu}_n(\boldsymbol{\theta})=\mathrm{diag}(c_n(\boldsymbol{\theta}),d_n(\boldsymbol{\theta}))=\mathrm{diag}(n^{-\frac{\eta}{2}}(\log n)^{-1},n^{-\frac{\eta}{2}+\alpha-\frac{1}{2}}(\log n)^{-1})$ and denote by

$$L^n_{\boldsymbol{\theta}_n/\boldsymbol{\theta}}=\frac{d\mathrm{P}^n_{\boldsymbol{\theta}_n}}{d\mathrm{P}^n_{\boldsymbol{\theta}}}=\sum_{i=1}^n[\log f_{\boldsymbol{\theta}_n}(X_{ni})-\log f_{\boldsymbol{\theta}}(X_{ni})]$$

the associated local log-likelihood ratio.

The following result states that the submodel $\mathcal{P}_n^-$ is LAN. We point out that inference in $\mathcal{P}_n^+$ is totally different; see Section 5 for a discussion.

Theorem 4.1 *The family of distributions $\mathcal{P}_n^-$ is LAN, with central sequence $\boldsymbol{\Delta}^n_{\boldsymbol{\theta}}:=(\Delta^n_{\boldsymbol{\theta},I},\Delta^n_{\boldsymbol{\theta},II})'$, where*

$$\Delta^n_{\boldsymbol{\theta},I}:=\sum_{i=1}^n D^{ni}_{\boldsymbol{\theta},I}:=n^{\frac{\eta}{2}-1}\sum_{i=1}^n\frac{\phi_\alpha-\phi}{n^{\eta-1}[\phi_\alpha-\phi]+\phi}(X_{ni}),$$

$$\Delta^n_{\boldsymbol{\theta},II}:=\sum_{i=1}^n D^{ni}_{\boldsymbol{\theta},II}:=\sigma^{-2}n^{\frac{\eta}{2}-1}\sum_{i=1}^n(n^{\frac{1}{2}-\alpha}-X_{ni})\frac{\phi_\alpha}{n^{\eta-1}[\phi_\alpha-\phi]+\phi}(X_{ni}),$$

and information matrix $\boldsymbol{\Gamma}:=\mathrm{diag}\left(1,\sigma^{-2}\right)$. More precisely, for any $\boldsymbol{\theta}\in\boldsymbol{\Theta}^-$ and any bounded sequence $\boldsymbol{\tau}_n=(s_n,t_n)'$, we have

$$L^n_{\boldsymbol{\theta}+\boldsymbol{\nu}_n(\boldsymbol{\theta})\boldsymbol{\tau}_n/\boldsymbol{\theta}}=\boldsymbol{\tau}_n'\boldsymbol{\Delta}^n_{\boldsymbol{\theta}}-\frac{1}{2}\boldsymbol{\tau}_n'\boldsymbol{\Gamma}\boldsymbol{\tau}_n+o_{\mathrm{P}}(1)\quad\text{and}\quad\boldsymbol{\Delta}^n_{\boldsymbol{\theta}}\xrightarrow{\mathcal{L}}\mathcal{N}\left(\mathbf{0},\boldsymbol{\Gamma}\right)\tag{4.1}$$

under $\mathrm{P}^n_{\boldsymbol{\theta}}$, as $n\to\infty$.

We start the proof of Theorem 4.1 with the following lemma.

Lemma 4.1 *Under $\mathrm{P}^n_{\boldsymbol{\theta}}$, as $n\to\infty$, (i) $\mathrm{E}[D^{n1}_{\boldsymbol{\theta},I}]=\mathrm{E}[D^{n1}_{\boldsymbol{\theta},II}]=0$, (ii) $\mathrm{Var}[D^{n1}_{\boldsymbol{\theta},I}]=n^{-1}(1+o(1))$, (iii) $\mathrm{Var}[D^{n1}_{\boldsymbol{\theta},II}]=\sigma^{-2}n^{-1}(1+o(1))$, (iv) $\mathrm{Cov}[D^{n1}_{\boldsymbol{\theta},I},D^{n1}_{\boldsymbol{\theta},II}]=o(n^{-1})$, and (v) for all $\delta>0$,*

$$\mathrm{E}\Big[Z^2_{n1}\mathbb{I}(|Z_{n1}|>\delta)\Big]=o(n^{-1}),\tag{4.2}$$

where $Z_{ni}=Z_{ni}(\boldsymbol{\theta}):=\frac{1}{2}(\boldsymbol{\nu}_n(\boldsymbol{\theta})\boldsymbol{\tau}_n)'\nabla_{\boldsymbol{\theta}}\log f_{\boldsymbol{\theta}}(X_{ni})=\frac{1}{2}\{s_nD^{ni}_{\boldsymbol{\theta},I}+t_nD^{ni}_{\boldsymbol{\theta},II}\}$.

Proof of Lemma 4.1. (i) The first claims are trivial since $\mathrm{E}[D^{n1}_{\boldsymbol{\theta},I}] = n^{\frac{\eta}{2}-1}\int(\phi_\alpha-\phi)(x)\,dx = 0$ and $\mathrm{E}[D^{n1}_{\boldsymbol{\theta},II}] = \sigma^{-2}n^{\frac{\eta}{2}-1}\int(n^{\frac{1}{2}-\alpha}-x)\phi_\alpha(x)\,dx = 0$.

(ii) Define $r_n := n^{\frac{1}{2}-\alpha}/2$. First note that $\frac{\phi_\alpha}{\phi}(x) = \exp\left[2r_n(x-r_n)/\sigma^2\right]$ for all x. Therefore

$$
\begin{aligned}
n\mathrm{Var}[D^{n1}_{\boldsymbol{\theta},I}] &= n^{\eta-1}\int \frac{(\phi_\alpha-\phi)^2}{n^{\eta-1}[\phi_\alpha-\phi]+\phi}(x)\,dx \\
&= \int \frac{\left(\exp\left[2r_n(x-r_n)/\sigma^2\right]-1\right)^2}{\exp\left[2r_n(x-r_n)/\sigma^2\right]-1+n^{1-\eta}}\phi(x)\,dx \\
&= T_{n1}-T_{n2}+T_{n3},
\end{aligned}
$$

where

$$T_{n1} := \int_{-\infty}^{r_n} \frac{\left(\exp\left[2r_n(x-r_n)/\sigma^2\right]-1\right)^2}{\exp\left[2r_n(x-r_n)/\sigma^2\right]-1+n^{1-\eta}}\phi(x)\,dx,$$

$$T_{n2} := \int_{r_n}^{\infty}\left[\exp\left[2r_n(x-r_n)/\sigma^2\right] - \frac{\left(\exp\left[2r_n(x-r_n)/\sigma^2\right]-1\right)^2}{\exp\left[2r_n(x-r_n)/\sigma^2\right]-1+n^{1-\eta}}\right]\phi(x)\,dx,$$

and

$$T_{n3} := \int_{r_n}^{\infty}\exp\left[2r_n(x-r_n)/\sigma^2\right]\phi(x)\,dx.$$

Now, Lebesgue's DCT yields $T_{n1} = n^{\eta-1}\left(\int\phi(x)\,dx + o(1)\right) = o(1)$ as $n\to\infty$. As for T_{n2}, one easily shows that, for some constant C, $|T_{n2}| \leq C\int_{r_n}^{\infty}\phi(x)\,dx = o(1)$ as $n\to\infty$. Eventually, letting $y = x-2r_n$, we obtain $T_{n3} = \int_{r_n}^{\infty}\phi_\alpha(x)\,dx = \int_{-r_n}^{\infty}\phi(y)\,dy = 1+o(1)$, as $n\to\infty$.

(iii) Letting $y = x-2r_n$, we have

$$
\begin{aligned}
n\mathrm{Var}[D^{n1}_{\boldsymbol{\theta},II}] &= \sigma^{-4}n^{\eta-1}\int\frac{(2r_n-x)^2\phi_\alpha^2(x)}{n^{\eta-1}[\phi_\alpha-\phi]+\phi}(x)\,dx \\
&= \sigma^{-4}\int\frac{(2r_n-x)^2\phi_\alpha(x)}{1+(n^{1-\eta}-1)\phi(x)/\phi_\alpha(x)}\,dx \\
&= \sigma^{-4}\int\frac{y^2\phi(y)}{1+(n^{1-\eta}-1)\exp\left[-2r_n(y+r_n)/\sigma^2\right]}\,dy \\
&= S_{n1}+S_{n2}+S_{n3},
\end{aligned}
$$

where

$$S_{n1} := \sigma^{-4}\int_{-\infty}^{-r_n}\frac{y^2\phi(y)}{1+(n^{1-\eta}-1)\exp\left[-2r_n(y+r_n)/\sigma^2\right]}\,dy,$$

$$S_{n2} := \sigma^{-4}\int_{-r_n}^{\infty}\left[\frac{1}{1+(n^{1-\eta}-1)\exp\left[-2r_n(y+r_n)/\sigma^2\right]}-1\right]y^2\phi(y)\,dy,$$

and

$$S_{n3} := \sigma^{-4}\int_{-r_n}^{\infty}y^2\phi(y)\,dy.$$

Again, Lebesgue's DCT shows that both S_{n1} and S_{n2} are $o(1)$ as $n\to\infty$. This yields the result since, clearly, $S_{n3} = \sigma^{-2}+o(1)$ as $n\to\infty$.

(iv) Letting $y = x - 2r_n$ again, we have

$$
\begin{aligned}
n\mathrm{Cov}[D^{n1}_{\boldsymbol{\theta},I}, D^{n1}_{\boldsymbol{\theta},II}] &= \sigma^{-2} n^{\eta-1} \int \frac{(2r_n - x)\phi_\alpha(x)(\phi_\alpha(x) - \phi(x))}{n^{\eta-1}[\phi_\alpha - \phi] + \phi}(x)\, dx \\
&= \sigma^{-2} \int \frac{(2r_n - x)\phi_\alpha(x)(1 - \phi(x)/\phi_\alpha(x))}{1 + (n^{1-\eta} - 1)\phi(x)/\phi_\alpha(x)}\, dx \\
&= -\sigma^{-2} \int \frac{y\phi(y)(1 - \exp\big[-2r_n(y + r_n)/\sigma^2\big])}{1 + (n^{1-\eta} - 1)\exp\big[-2r_n(y + r_n)/\sigma^2\big]}\, dy \\
&= R_{n1} + R_{n2} + R_{n3},
\end{aligned}
$$

where

$$
R_{n1} := -\sigma^{-2} \int_{-\infty}^{-r_n} \frac{y\phi(y)(1 - \exp\big[-2r_n(y + r_n)/\sigma^2\big])}{1 + (n^{1-\eta} - 1)\exp\big[-2r_n(y + r_n)/\sigma^2\big]}\, dy,
$$

$$
R_{n2} := \sigma^{-2} \int_{-r_n}^{\infty} \frac{n^{1-\eta}\exp\big[-2r_n(y + r_n)/\sigma^2\big]}{1 + (n^{1-\eta} - 1)\exp\big[-2r_n(y + r_n)/\sigma^2\big]}\, y\phi(y)\, dy,
$$

and

$$
R_{n3} := -\sigma^{-2} \int_{-r_n}^{\infty} y\phi(y)\, dy.
$$

Now, $|R_{n1}| \leq C \int_{-\infty}^{-r_n} |y|\phi(y)\, dy$ for some constant C, so that R_{n1} is $o(1)$ as $n \to \infty$. And, by Lebesgue's DCT, so is R_{n2}. This yields the result since $R_{n3} = -\sigma^{-2}(\int y\phi(y)\, dy + o(1)) = o(1)$ as $n \to \infty$.

(v) Define

$$
\mathbb{I}_n(x) := \mathbb{I}\left[\frac{n^{\frac{\eta}{2}-1}}{2} \left| \frac{s_n(\phi_\alpha(x) - \phi(x)) + t_n\sigma^{-2}(2r_n - x)\phi_\alpha(x)}{n^{\eta-1}[\phi_\alpha(x) - \phi(x)] + \phi(x)} \right| > \delta \right].
$$

Then since

$$
\mathbb{I}_n(x + 2r_n) \leq \mathbb{I}\left[\frac{|1 - \frac{\phi}{\phi_\alpha}(x + 2r_n)| + |x|}{n^{\eta-1} + (1 - n^{\eta-1})\frac{\phi}{\phi_\alpha}(x + 2r_n)} > Cn^{1-\frac{\eta}{2}} \right]
$$

$$
\leq \mathbb{I}\left[\left|1 - \frac{\phi}{\phi_\alpha}(x + 2r_n)\right| + |x| > Cn^{\frac{\eta}{2}} \right] \leq \mathbb{I}\left[\left|1 - \exp\big[-2r_n(x + r_n)/\sigma^2\big]\right| + |x| > Cn^{\frac{\eta}{2}} \right],
$$

we obtain that, for all x, $\mathbb{I}_n(x + 2r_n) = o(1)$ as $n \to \infty$. Write then $n\mathrm{E}\big[Z_{n1}^2\mathbb{I}(|Z_{n1}| > \delta)\big] \leq Cn\big(\mathrm{E}\big[(D^{n1}_{\boldsymbol{\theta},I})^2\mathbb{I}(|Z_{n1}| > \delta)\big] + \mathrm{E}\big[(D^{n1}_{\boldsymbol{\theta},II})^2\mathbb{I}(|Z_{n1}| > \delta)\big]\big) =: C(U_{n1} + U_{n2})$. Decompose U_{n1} into $U_{n1} = \widetilde{T}_{n1} + \widetilde{T}_{n2} + \widetilde{T}_{n3}$, where $\widetilde{T}_{ni}$, $i = 1, 2, 3$ are defined as in the proof of Lemma 4.1(ii), except that the corresponding integrands are multiplied by $\mathbb{I}_n(x)$ in each case. Clearly, the same argument as in Lemma 4.1(ii) show that both $\widetilde{T}_{n1}$ and $\widetilde{T}_{n2}$ are $o(1)$ as $n \to \infty$. As for $\widetilde{T}_{n3}$, the absolute continuity of the Lebesgue integral implies that $\widetilde{T}_{n3} = \int_{r_n}^{\infty} \phi_\alpha(x)\, \mathbb{I}_n(x)\, dx = \int_{-r_n}^{\infty} \phi(y)\, \mathbb{I}_n(y + 2r_n)\, dy = o(1)$, as $n \to \infty$.

Similarly, defining $\widetilde{S}_{ni}$, $i = 1, 2, 3$ as in the proof of Lemma 4.1(iii), except that the corresponding integrands are multiplied by $\mathbb{I}_n(y + 2r_n)$ in each case, we decompose U_{n2} into $U_{n2} = \widetilde{S}_{n1} + \widetilde{S}_{n2} + \widetilde{S}_{n3}$. Again, working as in Lemma 4.1(iii), $\widetilde{S}_{n1}$ and $\widetilde{S}_{n2}$ are seen to be $o(1)$ as $n \to \infty$. Eventually, by absolute continuity again, we have $\widetilde{S}_{n3} = \sigma^{-4} \int_{-r_n}^{\infty} y^2\phi(y)\, \mathbb{I}_n(y + 2r_n)\, dy = o(1)$, as $n \to \infty$. □

It follows directly from Lemma 4.1 and the Lindeberg-Feller CLT that, under $\mathrm{P}^n_{\boldsymbol{\theta}}$, the central sequence $\boldsymbol{\Delta}^n_{\boldsymbol{\theta}}$ is asymptotically normal with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Gamma}$. In order to prove that the second-order stochastic expansion in (4.1) holds, we use Swensen (1985)'s Lemma. Defining the quantities

$$\xi_{ni} = \xi_{ni}(\boldsymbol{\theta}) := \left(\frac{f_{\boldsymbol{\theta}_n}(X_{ni})}{f_{\boldsymbol{\theta}}(X_{ni})}\right)^{1/2} - 1,$$

Swensen's lemma, in this i.i.d. context, takes the following form.

Lemma 4.2 (Swensen) *Assume that (i)* $\sum_{i=1}^n \mathrm{E}[(Z_{ni} - \xi_{ni})^2] = o(1)$, *(ii)* $\sup_n \sum_{i=1}^n \mathrm{E}[Z_{ni}^2] < \infty$, *(iii)* $\max_{1\leq i\leq n} |Z_{ni}| = o_{\mathrm{P}}(1)$, *(iv)* $\sum_{i=1}^n Z_{ni}^2 - \frac{1}{4}\boldsymbol{\tau}'_n \boldsymbol{\Gamma}\, \boldsymbol{\tau}_n = o_{\mathrm{P}}(1)$, *(v)* $\sum_{i=1}^n \mathrm{E}[Z_{ni}^2 \mathbb{I}(|Z_{ni}| > 1/2)] = o(1)$, *(vi)* $\mathrm{E}[Z_{ni}] = 0$, *and (vii)* $\sum_{i=1}^n \mathrm{E}[\xi_{ni}^2 + 2\xi_{ni}] = o(1)$ *(where all expectations and convergences in probability are taken under* $\mathrm{P}^n_{\boldsymbol{\theta}}$*). Then the second-order stochastic expansion in (4.1) holds.*

Proof of Lemma 4.2. Unless otherwise stated, all convergences in probability, expectations and variances, in this proof, are with respect to $\mathrm{P}^n_{\boldsymbol{\theta}}$.

(i) Since $\frac{1}{2} f_{\boldsymbol{\theta}}^{1/2} \nabla_{\boldsymbol{\theta}} \log f_{\boldsymbol{\theta}} = \nabla_{\boldsymbol{\theta}} f_{\boldsymbol{\theta}}^{1/2}$, we see that $\sum_{i=1}^n \mathrm{E}[(Z_{ni} - \xi_{ni})^2] = n\mathrm{E}[(Z_{n1} - \xi_{n1})^2]$ is given by

$$n \int \left(f_{\boldsymbol{\theta}_n}^{1/2}(x) - f_{\boldsymbol{\theta}}^{1/2}(x) - \left\{ c_n s_n (\partial_\eta f_{\boldsymbol{\theta}}^{1/2})(x) + d_n t_n (\partial_\alpha f_{\boldsymbol{\theta}}^{1/2})(x) \right\} \right)^2 dx.$$

Therefore $\sum_{i=1}^n \mathrm{E}[(Z_{ni} - \xi_{ni})^2] \leq 3n(V_{n\eta} + V_{n\alpha} + V_{n\eta\alpha})$, where

$$V_{n\eta} := \int \left(f_{\boldsymbol{\theta}_n}^{1/2}(x) - f_{\eta,\alpha_n}^{1/2}(x) - c_n s_n (\partial_\eta f_{\eta,\alpha_n}^{1/2})(x) \right)^2 dx,$$

$$V_{n\alpha} := \int \left(f_{\eta,\alpha_n}^{1/2}(x) - f_{\boldsymbol{\theta}}^{1/2}(x) - d_n t_n (\partial_\alpha f_{\boldsymbol{\theta}}^{1/2})(x) \right)^2 dx,$$

and

$$V_{n\eta\alpha} := c_n^2 s_n^2 \int \left((\partial_\eta f_{\eta,\alpha_n}^{1/2})(x) - (\partial_\eta f_{\boldsymbol{\theta}}^{1/2})(x) \right)^2 dx.$$

We proceed by proving that (a) $V_{n\eta}$, (b) $V_{n\alpha}$, and (c) $V_{n\eta\alpha}$ all are $o(n^{-1})$, as $n \to \infty$.

(a) By using successively the integral form for the remainder of the first order Taylor expansion and Jensen's inequality, we obtain

$$\begin{aligned} V_{n\eta} &= \int \left(\int_\eta^{\eta_n} (\eta_n - \lambda)\, (\partial_\eta^2 f_{\lambda,\alpha_n}^{1/2})(x)\, d\lambda \right)^2 dx \\ &\leq \frac{1}{2}(c_n s_n)^2 \int \int_\eta^{\eta_n} (\eta_n - \lambda)\, (\partial_\eta^2 f_{\lambda,\alpha_n}^{1/2})^2(x)\, d\lambda\, dx, \end{aligned}$$

since $\int_\eta^{\eta_n} (\eta_n - \lambda)\, d\lambda = (c_n s_n)^2/2$. Therefore,

$$\begin{aligned} V_{n\eta} &\leq C c_n^2 \int \int_\eta^{\eta_n} (\eta_n - \lambda) \left\{ \frac{2 f_{\lambda,\alpha_n} (\partial_\eta^2 f_{\lambda,\alpha_n}) - (\partial_\eta f_{\lambda,\alpha_n})^2}{f_{\lambda,\alpha_n}^{3/2}}(x) \right\}^2 d\lambda\, dx \\ &\leq C c_n^2 (\log n)^4 \int \int_\eta^{\eta_n} (\eta_n - \lambda) \\ &\qquad \left\{ \left[n^{\lambda-1}(\phi_{\alpha_n} - \phi) \right]^2 \frac{\left[n^{\lambda-1}(\phi_{\alpha_n} - \phi) + 2\phi \right]^2}{\left[n^{\lambda-1}(\phi_{\alpha_n} - \phi) + \phi \right]^3} \right\}(x)\, d\lambda\, dx \end{aligned}$$

since $\partial_\eta f_{\lambda,\alpha_n} = n^{\lambda-1}(\log n)(\phi_{\alpha_n} - \phi)$ and $\partial^2_\eta f_{\lambda,\alpha_n} = n^{\lambda-1}(\log n)^2(\phi_{\alpha_n} - \phi)$. Using again $\frac{\phi_\alpha}{\phi}(x) = \exp\left[2r_n(x-r_n)/\sigma^2\right]$, we obtain

$$V_{n\eta} \leq Cn^{-\eta}(\log n)^2 \int\int_\eta^{\eta_n} (\eta_n - \lambda)\, g^2_{\lambda,\alpha_n}(x)\, \frac{\left[g_{\lambda,\alpha_n}(x)+2\right]^2}{\left[g_{\lambda,\alpha_n}(x)+1\right]^3}\, \phi(x)\, d\lambda\, dx, \tag{4.3}$$

with $g_{\eta,\alpha}(x) := n^{\lambda-1}\big(\exp\left[2\bar{r}_n(x-\bar{r}_n)/\sigma^2\right] - 1\big)$, $\bar{r}_n = n^{\frac{1}{2}-\alpha_n}/2$. Decompose the upper bound in (4.3) into $C(V_{n\eta,1} + V_{n\eta,2} + V_{n\eta,3})$, where

$$V_{n\eta,1} := n^{-\eta}(\log n)^2 \int_{-\infty}^{\bar{r}_n}\int_\eta^{\eta_n} (\eta_n - \lambda)\, g^2_{\lambda,\alpha_n}(x)\, \frac{\left[g_{\lambda,\alpha_n}(x)+2\right]^2}{\left[g_{\lambda,\alpha_n}(x)+1\right]^3}\, \phi(x)\, d\lambda\, dx,$$

$$\begin{aligned} V_{n\eta,2} := n^{-\eta}(\log n)^2 \int_{\bar{r}_n}^{\infty}\int_\eta^{\eta_n} (\eta_n - \lambda) \Bigg\{ & g^2_{\lambda,\alpha_n}(x)\, \frac{\left[g_{\lambda,\alpha_n}(x)+2\right]^2}{\left[g_{\lambda,\alpha_n}(x)+1\right]^3} \\ & -n^{\lambda-1}\exp\left[2\bar{r}_n(x-\bar{r}_n)/\sigma^2\right]\Bigg\}\, \phi(x)\, d\lambda\, dx, \end{aligned}$$

and

$$V_{n\eta,3} := n^{-\eta}(\log n)^2 \int_{\bar{r}_n}^{\infty}\int_\eta^{\eta_n} (\eta_n - \lambda)\, n^{\lambda-1}\exp\left[2\bar{r}_n(x-\bar{r}_n)/\sigma^2\right]\phi(x)\, d\lambda\, dx.$$

Now, uniformly in λ and x, we have $|g_{\lambda,\alpha_n}(x)|\mathbb{I}(x < \bar{r}_n) < 1/2$, so that

$$\begin{aligned} V_{n\eta,1} &\leq Cn^{-\eta}(\log n)^2 \int\int_\eta^{\eta_n} (\eta_n - \lambda)\, n^{2(\lambda-1)}\, \phi(x)\, d\lambda\, dx \\ &\leq Cn^{-1}\Big\{\big(\exp(2s_n n^{-\eta/2}) - 1\big)n^{\eta-1} - 2s_n n^{\frac{\eta}{2}-1}\Big\}, \end{aligned}$$

which shows that $V_{n\eta,1}$ is indeed $o(n^{-1})$ as $n \to \infty$. As for $V_{n\eta,2}$, one can check that

$$\left| g^2_{\lambda,\alpha_n}(x)\, \frac{\left[g_{\lambda,\alpha_n}(x)+2\right]^2}{\left[g_{\lambda,\alpha_n}(x)+1\right]^3} - n^{\lambda-1}\exp\left[2\bar{r}_n(x-\bar{r}_n)/\sigma^2\right] \right|$$

is bounded in n (uniformly in λ, x), so that

$$\begin{aligned} V_{n\eta,2} &\leq n^{-\eta}(\log n)^2 \left(\int_{\bar{r}_n}^{\infty} \phi(x)\, dx\right)\left(\int_\eta^{\eta_n} (\eta_n - \lambda)\ d\lambda\right) \\ &\leq Cn^{-2\eta}\left(\int_{\bar{r}_n/\sigma}^{\infty} \exp(-y^2/2)\, dy\right) \leq Cn^{-2\eta}\exp[-\bar{r}_n/(2\sigma)], \end{aligned}$$

which shows that $V_{n\eta,2}$ also is $o(n^{-1})$ as $n \to \infty$. Eventually, $V_{n\eta,3}$ is also $o(n^{-1})$ as $n \to \infty$, since

$$\begin{aligned} V_{n\eta,3} &= n^{-\eta}(\log n)^2 \left(\int_{\bar{r}_n}^{\infty} \phi_{\alpha_n}(x)\, dx\right)\left(\int_\eta^{\eta_n} (\eta_n - \lambda)\, n^{\lambda-1}\, d\lambda\right) \\ &\leq n^{-1}\big(\exp(s_n n^{-\frac{\eta}{2}}) - 1 - s_n n^{-\frac{\eta}{2}}\big). \end{aligned}$$

(b) Working as for $V_{n\eta}$, we obtain (since $\int_\alpha^{\alpha_n}(\alpha_n-\lambda)^2\,d\lambda=(d_nt_n)^3/3$)

$$\begin{aligned}
V_{n\alpha} &= \int\left(\int_\alpha^{\alpha_n}(\alpha_n-\lambda)\,(\partial^2_\alpha f^{1/2}_{\eta,\lambda})(x)\,d\lambda\right)^2\,dx\\
&\le \frac{1}{3}(d_nt_n)^3\int\int_\alpha^{\alpha_n}(\partial^2_\alpha f^{1/2}_{\eta,\lambda})^2(x)\,d\lambda\,dx\\
&\le Cd_n^3\int\int_\alpha^{\alpha_n}\left\{\frac{2f_{\eta,\lambda}(\partial^2_\alpha f_{\eta,\lambda})-(\partial_\alpha f_{\eta,\lambda})^2}{f^{3/2}_{\eta,\lambda}}(x)\right\}^2\,d\lambda\,dx
\end{aligned}$$

Now, $\partial_\alpha f_{\eta,\lambda}(x)=\sigma^{-2}n^{\eta-\lambda-\frac12}(\log n)(n^{\frac12-\lambda}-x)\phi_\lambda(x)$ and $\partial^2_\alpha f_{\eta,\lambda}(x)=\sigma^{-4}n^{\eta-2\lambda}(\log n)^2\phi_\lambda(x)\{(n^{\frac12-\lambda}-x)^2+\sigma^2n^{\lambda-\frac12}x-2\sigma^2\}$, so that

$$\begin{aligned}
V_{n\alpha} &\le Cd_n^3n^{\eta+1}(\log n)^4\int\int_\alpha^{\alpha_n}n^{-4\lambda}\phi_\lambda(x)\\
&\frac{\left[2[1+(n^{1-\eta}-1)\frac{\phi}{\phi_\lambda}(x)]^2\{(n^{\frac12-\lambda}-x)^2+\sigma^2n^{\lambda-\frac12}x-2\sigma^2\}-(n^{\frac12-\lambda}-x)^2\right]^2}{[1+(n^{1-\eta}-1)\frac{\phi}{\phi_\lambda}(x)]^3}\,d\lambda\,dx.
\end{aligned}$$

Letting $y=x-2\tilde r_n$ (with $\tilde r_n:=n^{\frac12-\lambda}/2$) and using $\frac{\phi}{\phi_\lambda}(y+2\tilde r_n)=\exp\left[-2\tilde r_n(y+\tilde r_n)/\sigma^2\right]$, we obtain that

$$\begin{aligned}
V_{n\alpha} &\le Cn^{-\frac{\eta}{2}-\frac12+3\alpha}(\log n)\int\int_\alpha^{\alpha_n}n^{-4\lambda}\phi(y)\\
&\frac{\left[2[1+(n^{1-\eta}-1)\exp\left[-2\tilde r_n(y+\tilde r_n)/\sigma^2\right]]^2\{y^2+\sigma^2y/(2\tilde r_n)-\sigma^2\}-y^2\right]^2}{[1+(n^{1-\eta}-1)\exp\left[-2\tilde r_n(y+\tilde r_n)/\sigma^2\right]]^3}\,d\lambda\,dy\\
&\le Cn^{-\frac{\eta}{2}-\frac12+3\alpha}(\log n)\int\int_\alpha^{\alpha_n}n^{-4\lambda}\phi(y)\max(1,y^4)\,d\lambda\,dy\\
&\le Cn^{-\frac{\eta}{2}-\frac12-\alpha}(1-\exp(-4n^{-\frac{\eta}{2}-\frac12+\alpha}t_n))\le Cn^{-\eta-1},
\end{aligned}$$

so that $V_{n\alpha}=o(n^{-1})$ as $n\to\infty$.

(c) Eventually, we complete the proof of (i) by proving that $V_{n\eta\alpha}$ is also $o(n^{-1})$ as $n\to\infty$. Defining $h_{n\eta\alpha}(x)$ and $h_{n\eta\alpha_n}(x)$ as

$$\frac{\exp\left[2r_n(x-r_n)/\sigma^2\right]-1}{\left[\exp\left[2r_n(x-r_n)/\sigma^2\right]-1+n^{1-\eta}\right]^{1/2}}\quad\text{and}\quad\frac{\exp\left[2\bar r_n(x-\bar r_n)/\sigma^2\right]-1}{\left[\exp\left[2\bar r_n(x-\bar r_n)/\sigma^2\right]-1+n^{1-\eta}\right]^{1/2}},$$

repspectively, we have

$$\begin{aligned}
V_{n\eta\alpha} &\le c_n^2n^{2\eta-2}(\log n)^2\int\left\{\frac{(\phi_{\alpha_n}-\phi)}{n^{\eta-1}(\phi_{\alpha_n}-\phi)+\phi}-\frac{(\phi_\alpha-\phi)}{n^{\eta-1}(\phi_\alpha-\phi)+\phi}\right\}^2(x)\,dx\\
&\le Cn^{\eta-2}\int\left\{\frac{\frac{\phi_{\alpha_n}}{\phi}-1}{\left[n^{\eta-1}(\frac{\phi_{\alpha_n}}{\phi}-1)+1\right]^{1/2}}-\frac{\frac{\phi_\alpha}{\phi}-1}{\left[n^{\eta-1}(\frac{\phi_\alpha}{\phi}-1)+1\right]^{1/2}}\right\}^2(x)\,\phi(x)\,dx\\
&\le Cn^{-1}\int\left\{h_{n\eta\alpha_n}(x)-h_{n\eta\alpha}(x)\right\}^2\phi(x)\,dx,
\end{aligned}$$

which we decompose into $Cn^{-1}(V_{n\eta\alpha,1}+V_{n\eta\alpha,2}+V_{n\eta\alpha,3})$, where

$$V_{n\eta\alpha,1} := \int_{-\infty}^{\min(r_n,\bar{r}_n)} \Big\{ h_{n\eta\alpha_n}(x) - h_{n\eta\alpha}(x) \Big\}^2 \phi(x)\, dx,$$

$$V_{n\eta\alpha,2} := \int_{(r_n;\bar{r}_n)} \Big\{ h_{n\eta\alpha_n}(x) - h_{n\eta\alpha}(x) \Big\}^2 \phi(x)\, dx,$$

and

$$V_{n\eta\alpha,3} := \int_{\max(r_n,\bar{r}_n)}^{\infty} \Big\{ h_{n\eta\alpha_n}(x) - h_{n\eta\alpha}(x) \Big\}^2 \phi(x)\, dx;$$

here, $(a;b)$ denotes the interval $(\min(a,b),\max(a,b))$. Now, Lebesgue's DCT shows that $V_{n\eta\alpha,1} = o(1)$ as $n\to\infty$ (as for T_{n1} in the proof of Lemma 4.1(ii)). As for $V_{n\eta\alpha,2}$, proceeding as in the proof of Lemma 4.1(ii) yields

$$\begin{aligned} V_{n\eta\alpha,2} &\leq C\Big\{ \int_{(r_n;\bar{r}_n)} h^2_{n\eta\alpha_n}(x)\,\phi(x)\,dx + \int_{(r_n;\bar{r}_n)} h^2_{n\eta\alpha}(x)\,\phi(x)\,dx \Big\} \\ &\leq C\Big\{ \int_{(r_n;\bar{r}_n)} \phi^{1/2}_{\alpha_n}(x)\,dx + \int_{(r_n;\bar{r}_n)} \phi^{1/2}_{\alpha}(x)\,dx \Big\} \\ &= C\Big\{ \int_{(-r_n;\bar{r}_n-2r_n)} \phi^{1/2}(x)\,dx + \int_{(r_n-2\bar{r}_n;-\bar{r}_n)} \phi^{1/2}(x)\,dx \Big\}, \end{aligned}$$

which, since one can easily check that $\bar{r}_n = r_n + o(1)$ as $n\to\infty$, is $o(1)$ as $n\to\infty$. Eventually, working exactly as for T_{n2} and T_{n3} in the proof of Lemma 4.1(ii), we obtain that $V_{n\eta\alpha,3}$ is—up to $o(1)$ terms (as $n\to\infty$)—successively equal to

$$\begin{aligned} &\int_{\max(r_n,\bar{r}_n)}^{\infty} \Big\{ \big(\exp\big[2\bar{r}_n(x-\bar{r}_n)/\sigma^2\big]\big)^{1/2} - \big(\exp\big[2r_n(x-r_n)/\sigma^2\big]\big)^{1/2} \Big\}^2 \phi(x)\,dx \\ &\qquad = \int_{\max(r_n,\bar{r}_n)}^{\infty} \Big\{ \phi^{1/2}_{\alpha_n}(x) - \phi^{1/2}_{\alpha}(x) \Big\}^2 dx \\ &\quad = \int_{\max(-r_n,\bar{r}_n-2r_n)}^{\infty} \Big\{ \phi^{1/2}(y+2(r_n-\bar{r}_n)) - \phi^{1/2}(y) \Big\}^2 dy, \end{aligned}$$

which is seen to be $o(1)$ as $n\to\infty$, by using again Lebesgue's DCT and the fact that $\bar{r}_n = r_n + o(1)$ as $n\to\infty$. Therefore, we have shown that $V_{n\eta\alpha,i}$, $i=1,2,3$ all are $o(1)$ as $n\to\infty$, which yields that $V_{n\eta\alpha}$ is $o(n^{-1})$ as $n\to\infty$.

(ii) The boundedness of $(\boldsymbol{\tau}_n)$ implies that $Z^2_{ni} \leq C\{(D^{ni}_{\boldsymbol{\theta},I})^2 + (D^{ni}_{\boldsymbol{\theta},II})^2\}$. Consequently, Lemma 4.1(i)-(iii) yield $\sum_{i=1}^n \mathrm{E}[Z^2_{ni}] \leq C\{\mathrm{Var}[D^{ni}_{\boldsymbol{\theta},I}] + \mathrm{Var}[D^{ni}_{\boldsymbol{\theta},II}]\} = O(1)$, as $n\to\infty$.

(iii) Note that by using Markov's inequality,

$$\mathrm{P}\Big[\max_{1\leq i\leq n}|Z_{ni}|>\delta\Big] = \mathrm{P}\Big[\sum_{i=1}^n Z^2_{ni}\mathbb{I}(|Z_{ni}|>\delta)>\delta^2\Big] \leq \delta^{-2}\sum_{i=1}^n \mathrm{E}\Big[Z^2_{ni}\mathbb{I}(|Z_{ni}|>\delta)\Big],$$

so that the result follows from Lemma 4.1(v).

(iv) Letting $\mathbf{D}^{ni}_{\boldsymbol{\theta}} := (D^{ni}_{\boldsymbol{\theta},I}, D^{ni}_{\boldsymbol{\theta},II})'$, note that

$$\sum_{i=1}^n Z^2_{ni} - \frac{1}{4}\boldsymbol{\tau}'_n\boldsymbol{\Gamma}\,\boldsymbol{\tau}_n = \frac{1}{4}\boldsymbol{\tau}'_n\Big\{\Big(\sum_{i=1}^n \mathbf{D}^{ni}_{\boldsymbol{\theta}}\mathbf{D}^{ni\prime}_{\boldsymbol{\theta}}\Big) - \boldsymbol{\Gamma}\Big\}\boldsymbol{\tau}_n,$$

which, by using the boundedness of $(\boldsymbol{\tau}_n)$ and Lemma 4.1(i)-(iv), is seen to be $o_{\mathrm{P}}(1)$ as $n \to \infty$.

(v) This is a particular case of the convergence result in Lemma 4.1(v).

(vi) This is a direct consequence of Lemma 4.1.

(vii) This is trivial since $\mathrm{E}[\xi_{ni}^2 + 2\xi_{ni}] = \mathrm{E}[\xi_{ni}(\xi_{ni}+2)] = \mathrm{E}[\frac{f_{\boldsymbol{\theta}_n}(X_{ni})}{f_{\boldsymbol{\theta}}(X_{ni})} - 1] = 0.$ □

5 Local and asymptotic efficiency of a thresholded estimation procedure

If we consider the model described in (3.3), we observe that the parameter η balances the mixture of the two Gaussian random variables. Under the range $\boldsymbol{\theta} \in \boldsymbol{\Theta}^-$, the two components of the mixture are asymptotically separated since, in this case, the mean of the first Gaussian variable goes to infinity ($n^{\frac{1}{2}-\alpha} \to +\infty$). Hence, we propose the following thresholding procedure to build an estimator of the lacunarity of a multifractal function. We aim at counting the number of coefficients above a given level, growing to infinity, but at a smaller rate of convergence than the mean of the second group. We point out that, unlike the estimate defined in [10], which relies on the whole wavelet coefficients, here we only use the coefficients on well chosen resolution level $j_1(n) = \log_2(n)$. Hence the estimate is more easily computable and still has the same asymptotic behaviour.

Set $S_n = \frac{1}{n}\sum_{k=1}^n \mathbb{I}(X_{nk} \geq \log n)$ and define the following estimator

$$\hat{\eta}_n = 1 + \frac{1}{\log n}\log(S_n). \tag{5.1}$$

The asymptotics of this estimator is given in the following theorem.

Theorem 5.1 *Under the assumption $\boldsymbol{\theta} \in \boldsymbol{\Theta}^-$, the estimator* (5.1) *is asymptotically normal. More precisely,*

$$n^{\frac{\eta}{2}}\log(n)\,(\hat{\eta}_n - \eta) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1). \tag{5.2}$$

Hence, $\hat{\eta}_n$ is asymptotically optimal in the Le Cam sense, that is, locally and asymptotically efficient.

Hence, we have constructed a consistent estimator of the lacunarity parameter. Its rate of convergence is nonparametric since it depends on the true values of the unknown parameter $n^{\frac{\eta}{2}}\log(n)$. As for local asymptotic optimality, recall that an estimator of $\hat{\boldsymbol{\theta}}_n$ is said to be *locally and asymptotically efficient* (over $\mathcal{P}_n^-$) iff it satisfies $(\boldsymbol{\nu}_n(\boldsymbol{\theta}))^{-1}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) = \boldsymbol{\Gamma}^{-1}\Delta_{\boldsymbol{\theta}}^n + o_{\mathrm{P}}(1)$ at any $\mathrm{P}_{\boldsymbol{\theta}}^n \in \mathcal{P}_n^-$. The asymptotic distribution, under $\mathrm{P}_{\boldsymbol{\theta}}^n$, of such an estimator is therefore given by

$$(\boldsymbol{\nu}_n(\boldsymbol{\theta}))^{-1}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}^{-1}).$$

Local asymptotic optimality of our estimator $\hat{\eta}_n$ thus follows directly from (5.2) and the value of $\boldsymbol{\Gamma}$ (see Proposition 4.1).

Assume that we have at our disposal an estimator $\tilde{\boldsymbol{\theta}}_n$ that converges at the optimal rate, that is, which satisfies

$$(\boldsymbol{\nu}_n(\boldsymbol{\theta}))^{-1}(\tilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) = O_{\mathrm{P}}(1), \tag{5.3}$$

at any $\mathrm{P}_{\boldsymbol{\theta}}^n \in \mathcal{P}_n^-$. Le Cam's one-step methodology then consists in relying on the estimator

$$\hat{\boldsymbol{\theta}}_n = \tilde{\boldsymbol{\theta}}_n + \boldsymbol{\nu}_n(\tilde{\boldsymbol{\theta}}_n)\boldsymbol{\Gamma}^{-1}\Delta_{\tilde{\boldsymbol{\theta}}_n}^n,$$

which can easily be shown to be locally asymptotically optimal (in the above sense). This method thus allows for transforming an arbitrary preliminary consistent estimator $\tilde{\boldsymbol{\theta}}_n$ (only optimal in terms of consistency rate) into a locally asymptotically optimal one $\hat{\boldsymbol{\theta}}_n$ (optimal both in terms of consistency rate and limiting variance). *It is quite remarkable that our estimator* $\hat{\boldsymbol{\theta}}_n$ *does not need this one-step improvement, since it does directly reach the efficiency bound.*

The proof of Theorem 5.1 follows the guidelines of Theorem 3.2 in [10]. For sake of completness, we recall here the frame of the proof.

We start the proof with the following lemma.

Lemma 5.1 *Set* $F(n) = \mathrm{P}(\mathcal{N}(0,1) \geq \log n)$. *Under the assumption that* $\boldsymbol{\theta} \in \boldsymbol{\Theta}^-$, *we get that*

$$n^{\frac{\eta}{2}} n^{1-\eta} \left(S_n - F(n)\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1). \tag{5.4}$$

Proof of Lemma 5.1. First note that $nS_n \sim \mathcal{B}(n, P(X_{n1} \geq \log n))$, where $\mathcal{B}(N,p)$ denotes the Binomial distribution with parameters N and p. We claim that the following expansions holds for large n.

$$\begin{aligned} \mathrm{E}\left[n^{1-\eta}(S_n - F(n))\right] &= 1 + o(1) \\ \mathrm{Var}\left[n^{\frac{\eta}{2}} n^{1-\eta}(S_n - F(n))\right] &= 1 + o(1). \end{aligned}$$

Now write

$$\mathrm{E}\left(\exp\left[itn^{\frac{\eta}{2}}(n^{1-\eta}(S_n - F(n)) - 1)\right]\right) = \exp\left[-itn^{\frac{\eta}{2}}(1 + n^{1-\eta}F(n))\right] T_n,$$

where the following asymptotic expansion holds

$$T_n = \exp\left[itn^{\frac{\eta}{2}}(1 + n^{1-\eta}F(n)) - \frac{t^2}{2}\right] + o(1).$$

Hence we get as n grows to infinity that

$$\mathrm{E}\left(\exp\left[itn^{\frac{\eta}{2}}(n^{1-\eta}(S_n - F(n)) - 1)\right]\right) \to \exp\left(-\frac{t^2}{2}\right),$$

proving the result. □

Now we are ready to prove Theorem 5.1.

Proof of Theorem 5.1. First note that $\log(n)(\hat{\eta}_n - \eta) = \log(n^{1-\eta} S_n)$. Since Lemma 5.1 provides the asymptotic distribution for $n^{\frac{\eta}{2}} n^{1-\eta} \left(S_n - F(n)\right)$, the Δ method, cf [22], yields the result. As for local asymptotic optimality, it directly follows from (5.2) as we showed above. □

For $\boldsymbol{\theta} \in \boldsymbol{\Theta}^+$, the situation is different. Indeed, the rescaled coefficients are drawn from a mixture distribution composed of a centered Gaussian variable and a Gaussian variable with mean decreasing to 0, as defined in (3.3). Hence, in this case it is not possible to estimate the number of non zero coefficients, neither to distinguish asymptotically non zero and zero coefficients. Hence the model is not asymptotically identifiable anymore.

6 Locally and asymptotically optimal tests for intensity

In the previous section, we have provided a locally and asymptotically optimal estimator for the lacunarity parameter η. Although η is the parameter of interest in the analysis of multifractal models, as cited in [1] for instance, a natural question still is: is it possible to use LAN to build a locally and asymptotically optimal estimator of α?

The answer is unfortunately negative, due to the lack of a suitable preliminary estimator for α. As we have explained in the previous section, the Le Cam one-step methodology indeed requires some $(\boldsymbol{\nu}_n(\boldsymbol{\theta}))^{-1}$-consistent estimator for $\boldsymbol{\theta}$. For the intensity parameter, this means—in view of the LAN property given in Theorem 4.1—that we need a preliminary estimator $\tilde{\alpha}_n$ such that

$$\log(n) n^{\frac{\eta}{2}+\frac{1}{2}-\alpha}(\tilde{\alpha}_n - \alpha) = O_{\mathrm{P}}(1), \tag{6.1}$$

at any $\mathrm{P}^n_{\boldsymbol{\theta}} \in \mathcal{P}^-_n$.

Now, to the authors knowledge, none of the existing estimators for α is consistent at this very rate. For instance, the estimator

$$\tilde{\alpha}_n = \frac{1}{\log n} \frac{\sum_{k=1}^n X_{nk}}{\sum_{k=1}^n X^2_{nk} - \sigma^2} \tag{6.2}$$

that was proposed in [10], satisfies (under the additional assumption that $\eta > 2\alpha$)

$$\log(n) n^{\frac{\eta}{2}} (\tilde{\alpha}_n - \alpha) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2),$$

which shows that $\tilde{\alpha}_n$ admits the same rate of convergence as $\hat{\eta}_n$. But since $\alpha < 1/2$, this rate is unfortunately larger than the rate in (6.1), which means that $\tilde{\alpha}_n$ cannot be used as a preliminary estimator in the Le Cam one-step methodology.

Actually, an optimal estimation procedure should focus only on the mean without relying first on the composition of the mixture, ruled by the parameter η. But, due to the construction of model (1.3), it seems difficult to estimate independently the mean of a mixture without knowing the proportion of the mixture, so that it seems difficult to build a suitable preliminary estimator for α. optimal through our LAN property). To the authors knowledge, few is done in the statistical litterature in this direction. We point out that standard loglikelihood estimators are too difficult to handle in this case.

However, while the lack of preliminary estimators for α prevents the LAN theory from providing locally and optimal estimators, LAN still allows for defining locally and asymptotically optimal (maximin, actually) tests about the intensity parameter. More specifically, consider the testing problem (at asymptotic level $\beta \in (0,1)$)

$$\begin{cases} \mathcal{H}_0^{(n)} : \alpha \leq \alpha_0 \\ \mathcal{H}_1^{(n)} : \alpha > \alpha_0 \end{cases},$$

for some fixed $\alpha_0 < 1/2$. Then the LAN result in Theorem 4.1 and the consistency (at the appropriate rate) of our estimator $\hat{\eta}$ in (5.1) straightforwardly yield the following.

Theorem 6.1 *Let $\phi^{(n)}$ be the test that rejects $\mathcal{H}_0^{(n)}$ in favour of $\mathcal{H}_1^{(n)}$ iff*

$$\sigma\Delta^n_{\hat{\eta}_n,\alpha_0,II} = \sigma^{-1} n^{\frac{\hat{\eta}_n}{2}-1} \sum_{i=1}^n (n^{\frac{1}{2}-\alpha_0} - X_{ni}) \frac{\phi_{\alpha_0}}{n^{\hat{\eta}_n - 1}[\phi_{\alpha_0} - \phi] + \phi}(X_{ni}) \Phi^{-1}(1-\beta),$$

where Φ denotes the cdf of the standard normal distribution. Then the sequence of tests $\phi^{(n)}$ is locally and asymptotically maximin (at asymptotic level β).

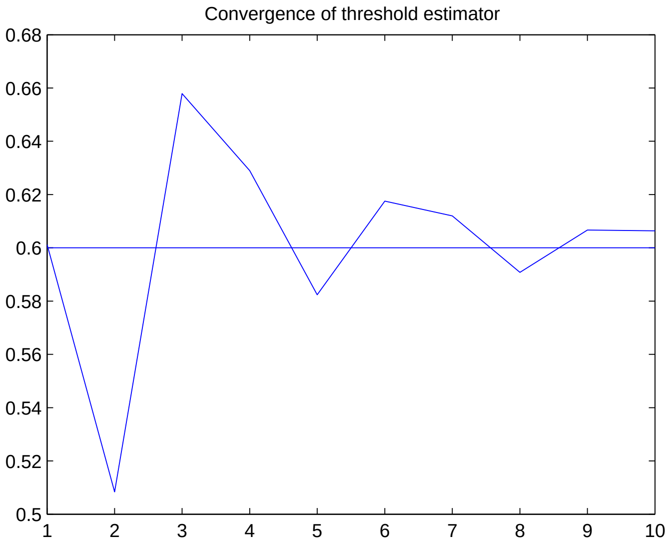

Figure 2: Lacunarity Estimation with threshold estimator.

One defines optimal two-sided tests in the same way. Of course, one could think of defining (possibly optimal) estimators of α by inverting such two-sided tests. This, however, would give again standard likelihood estimators of α, which, as mentioned above, are extremely diffcult to handle in this setup.

7 Numerical study of lacunarity estimation of a multifractal function

In this section, we compute the estimator of the lacunarity parameter of a multifractal function observed with a Gaussian noise with variance $\sigma^2 = 1$. In all the simulations, the intensity parameter is fixed $\alpha = .3$. We will consider three estimators. The first one is the threshold estimator defined in (5.1). In Figure 2, we plot the estimator of $\eta = .6$ when $n = 2^k$ increases for $k = 2, \ldots, 10$. The estimator is convergent with a quite slow rate.

To compare the estimation procedure, then we compute a moment estimator defined in [10] as

$$\tilde{\eta}_n = \tilde{\alpha}_n + \frac{1}{\log n} \log \left(\sum_{j=1}^{\log_2 n} d_{jk} \right).$$

Finally, we compute the maximum likelihood estimator

$$\hat{\eta}_{MLE} = \arg\max_{\eta} \sum_{i=1}^{n} \left[\log f_{\boldsymbol{\theta}_n}(X_{ni}) \right].$$

Figure 3 allows for comparing estimations (based on 30 replications) of the normalized mean square errors associated with the 3 estimators, for different values of $\eta = .01, \ldots, .99$, when the number of observations is large enough (here $n = 1024$). The threshold estimator is represented

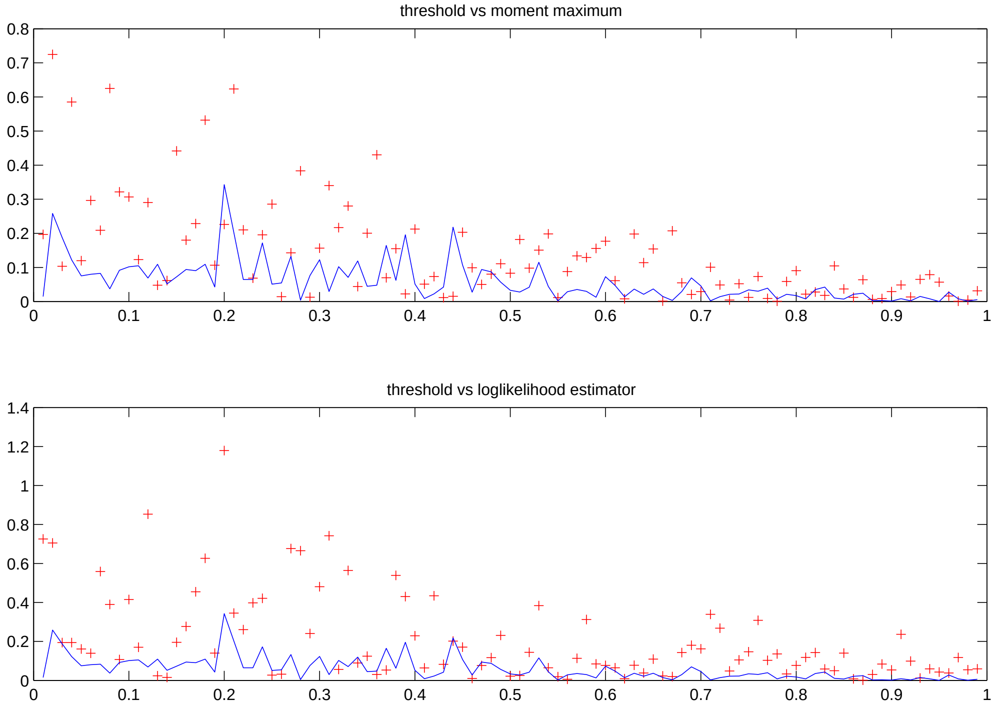

Figure 3: Normalized MSE with different estimators with $n = 1024$.

in straight lines while the two others are drawn with cross lines in the two different plots.

First, we point out that, when η increases, the estimation is more accurate as expected. Indeed, the difficulty in the estimation is given by the mixture property of the distribution, which vanishes when η is close to 1. For small η, the signal is very close to white noise, which prevents an efficient estimation.

When comparing the different estimates, the optimality of the threshold estimator appears clearly with a smaller variance than the two others. It seems also that numerically, the moment estimator is outperformed by the maximum likelihood estimator. Nevertheless, we can conclude that thresholding is here a very efficient procedure to estimate the lacunarity parameter in noisy data. This is the starting point of any modelling by multifractal processes. Testing multifractality is done in [10].

Here, the LAN property shows the optimality of the estimator of the lacunarity parameter, which is highlighted by the simulation results. But it did not help providing optimal estimators of the intensity parameter α (since we do not know any suitable preliminary estimator for α, that is an estimator being consistent at the rate $\log(n)n^{\frac{\eta}{2}+\frac{1}{2}-\alpha}$). Nevertheless, LAN allowed for defining optimal (in the Le Cam sense) tests on α.

References

[1] Arneodo, A., Bacry, E., Jaffard, S., and Muzy, J. F. (1998). Singularity spectrum of multifractal functions involving oscillating singularities. *The Journal of Fourier Analysis and Applications*, 4(2):159–174.

[2] Arneodo, A., Bacry, E., Jaffard, S., and Muzy, J. F. (1999). Oscillating singularities and fractal functions. In *Spline functions and the theory of wavelets (Montreal, PQ, 1999)*, pp. 315–329. Amer. Math. Soc., Providence, RI.

[3] Aubry, J. M. and Jaffard, S. (2001). Random wavelet series. Preprint.

[4] Bacry, E., Arneodo, A., Frisch, U., Gagne, Y., and Hopfinger, E. (1989). Wavelet analysis of fully developed turbulence data and measurement of scaling exponents. In *Turbulence and coherent structures (Grenoble, 1989)*, pp. 203–215. Kluwer Acad. Publ., Dordrecht.

[5] Benassi, A., Cohen, S., and Istas, J. (1998). Identifying the multifractional function of a Gaussian process. *Statistics & Probability Letters*, 39(4):337–345.

[6] Billingsley, P. (1995). *Probability and measure*. John Wiley & Sons Inc., New York, 3rd ed. ISBN 0-471-00710-2. A Wiley-Interscience Publication.

[7] Fillol, J. (2003). Multifractality: Theory and Evidence. An Application to the French Stock Market. *Economics Bulletin*, **3**, 31, 1–12.

[8] Frisch, U. and Parisi, G. (1985). On the singularity structure of fully developped turbulence; appendix to fully developped turbulence and intermittency. *Proc. Int. Summer School Phys. Enrico Fermi* (1):84–88.

[9] Gamboa, F. and Loubes, J.-M. (2005). Bayesian estimation of multifractal wavelet function. Bernoulli, (2), 34-57.

[10] Gamboa, F. and Loubes, J.-M. (2005). Estimation of the parameters of a multifractal wavelet function. Test, to appear.

[11] Jaffard, S. (2000). On lacunary wavelet series. *The Annals of Applied Probability*, 10(1):313–329.

[12] Leland, W. and Taqqu, M. and Willinger, W. and Wilson, D. (1994). On the self-similar nature of Etherned traffic. *IIE/ACM Transactions on Networking*, 2:1–15.

[13] Linton, O. (1993). Adaptive estimation in ARCH models. *Econometric Theory* **9** 539-569.

[14] Mallat, S. G. (1989). Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbf{r})$. *Transactions of the American Mathematical Society*, 315(1):69–87.

[15] McLeish, D. L. and Small, C. G. (1986). Likelihood methods for the discrimination problem. *Biometrika*, 73(2):397–403.

[16] Meyer, Y. (1990). *Ondelettes et Opérateurs*. Hermann.

[17] Riedi, R. H., Crouse, M. S., Ribeiro, V. J., and Baraniuk, R. G. (1999). A multifractal wavelet model with application to network traffic. *Institute of Electrical and Electronics Engineers. Transactions on Information Theory*, 45(3):992–1018.

[18] Roueff, F. (2001). Thèse de l'ecole nationale des télécommunications.

[19] Roueff, F. (2003). Almost sure haussdorff dimensions of graphs of rqndom wavelet series. *J. of Fourier Analysis and App.* 9(3).

[20] Shorack, G. R. (2003). *Probability for Statisticians.* Springer-Verlag, New York.

[21] Swensen, A. R. (1985). The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend. *J. Multivariate Anal.* **16** 54-70.

[22] van der Vaart, A. W. (1998). *Asymptotic statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge. ISBN 0-521-49603-9; 0-521-78450-6.

Jean-Michel Loubes
Equipe de probabilites et statistique,
Département de Mathématique,
UMR5149,
Université Montpellier 2,
Place Eugène Bataillon
34000 Montpellier France
Jean-Michel.Loubes@math.univ-montp2.fr

Davy Paindaveine
E.C.A.R.E.S,
Institut de Recherche en Statitique,
and Département de Mathématique
Université Libre de Bruxelles
50, avenue Roosevelt, CP 114
B-1050 Bruxelles Belgium
dpaindav@ulb.ac.be