ESAIM: PS 15 (2011) 69–82
DOI: 10.1051/ps/2009005

ESAIM: Probability and Statistics
www.esaim-ps.org

LOCAL ASYMPTOTIC NORMALITY PROPERTY FOR LACUNAR WAVELET SERIES MULTIFRACTAL MODEL*

Jean-Michel Loubes[1] and Davy Paindaveine[2]

Abstract. We consider a lacunar wavelet series function observed with an additive Brownian motion. Such 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 enables to prove the optimality of an estimator for the lacunarity parameter, and to build optimal (in the Le Cam sense) tests on the intensity parameter.

Mathematics Subject Classification. 60G17, 62G07.

Received October 8, 2008. Revised March 25, 2009.

1. Introduction

In the last decade random wavelet series have proved a powerful tool in signal processing since many signals, even very irregular functions, can be accurately represented in a wavelet basis using few coefficients, see for instance [2,5,6,11] or [7]. In particular random wavelet series provide good functional models for a large number of difficult issues. The distribution of these random wavelet coefficients describe the properties of the corresponding function. In many situations, a parametric distribution is chosen and inference on the parameters is thus needed to build a coherent model on observed data. In this paper, we focus on the probabilistic properties, local asymptotic normality, of a special case of lacunar random wavelet series in order to study the optimality of estimates of the distribution of the wavelet coefficients.

So, we consider the model initially described in [11]. Consider a wavelet basis ψ_{jk}, $j \geq 0$, $k = 0, \ldots, 2^j - 1$ and a couple of parameters $\boldsymbol{\theta} = (\alpha, \eta)^T \in \Theta := (0,1)^2$. Then, draw random wavelet coefficients w_{jk}, $k = 0, \ldots, 2^j - 1$ following the rescaled Bernoulli distribution $w_{jk} \sim 2^{(\eta-1)j}\delta_{2^{-\alpha j}} + (1 - 2^{(\eta-1)j})\delta_0$, where δ is the Dirac measure. One of the main interest of this yet simple model is that 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 which has multifractal properties, as studied in [1,3,11,18]. Moreover, its multifractal behaviour depends on the choice of both the lacunarity coefficient η and the intensity parameter α.

Keywords and phrases. Local asymptotic normality, lacunar wavelet series.

* *Research supported by a P.A.I. contract of the Belgian Federal Government and an Action de Recherche Concertée of the Communauté française de Belgique.*

[1] Institut de Mathématique de Toulouse, France; jean-michel.loubes@math.univ-toulouse.fr

[2] Université Libre de Bruxelles, Belgium.

Article published by EDP Sciences © EDP Sciences, SMAI 2011

https://doi.org/10.1051/ps/2009005 Published online by Cambridge University Press

In this paper, we assume that we observe noisy coefficients of this function in the following model

$$d_{jk} = w_{jk} + \epsilon_{jk},\ j = 0, \dots, j_1 = \log_2(n),\ k = 0, \dots, 2^j - 1. \tag{1.1}$$

w_{jk} are the true wavelet coefficients while ϵ_{jk} are i.i.d centered Gaussian random variables with variance $2^{-j}\sigma^2$, taken independent from the w_{jk}'s. Hence the data are given by the triangular array $\mathbf{d}_n = (d_{jk})_{1\le j\le j_1, 0\le k\le 2^j-1}$ of independent random variables, such that, for any j, k the distribution of d_{jk} is the following Gaussian mixture:

$$d_{jk} \sim 2^{(\eta-1)j}\mathcal{N}(2^{-\alpha j}, 2^{-j}\sigma^2) + (1 - 2^{(\eta-1)j})\mathcal{N}(0, 2^{-j}\sigma^2), \tag{1.2}$$

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 little is known about the estimation of this kind of multifractal functions, drawn from model (1.2). In a previous work [8], 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 parameters of the mixture (1.2) is conducted in [9]. The estimators of the parameter of interest are asymptotically normal but their rate of convergence is not the parametric rate of convergence but rather depends on the values of $\boldsymbol{\theta}$. The main difference with usual Gaussian mixtures, well studied in a large number of papers (see for instance [12,13,20] or [10]), comes from the fact that both the mixture probabilities and the mean of the distribution depend on the number of observations, which hampers the estimation and requires a particular attention.

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. We refer to [21] for a general definition of LAN property. In particular, we prove optimality in the Le Cam sense of an estimator of the lacunarity parameter built only with the wavelet coefficients of the last resolution level $j_1(n)$, which improves results in [9].

The paper falls into the following parts. In Section 2, we recall the properties of the random wavelet series model and describe the corresponding statistical model. In Section 3 we prove, for some values of these parameters, local asymptotic normality for the associated family of distributions, which implies, in Section 3.2, optimality of an estimator of the lacunarity parameter. Section 4 is devoted to deriving optimal tests on the intensity parameter, while the proofs of the main result are postponed to Section 5.

2. The model

Consider a lacunar random wavelet series f, as defined in [1,3,11]

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

built using a periodized wavelet basis $\psi_{jk}(x) = \sum_{l\in\mathbb{Z}} \tilde{\psi}(2^j(x-l)-k)$, $\forall j \in \mathbb{N}$, $k \in [0, 2^j - 1]$, for $\tilde{\psi}$ a mother wavelet in the Schwartz class. In this model, the wavelet coefficients are random variables whose distribution

https://doi.org/10.1051/ps/2009005 Published online by Cambridge University Press

depends on two parameters (η, α) and is given by the following relationship. Let $(X_{jk})_{j\in\mathbb{N}^*,\, k=0,\ldots,2^j-1}$ be a triangular array of independent Bernoulli random variables: for $\eta \in (0,1)$, $\mathbf{P}(X_{jk}=1) = 1-\mathbf{P}(X_{jk}=0) = 2^{(\eta-1)j}$. Hence, for $j \in \mathbb{N}^*$ and $k = 0,\ldots,2^j-1$ the random wavelet coefficients of the lacunar series f are given by

$$w_{jk} = 2^{-\alpha j} X_{jk}.$$

So we obtain

$$w_{jk} \sim 2^{(\eta-1)j}\delta_{2^{-\alpha j}} + (1-2^{(\eta-1)j})\delta_0. \tag{2.2}$$

The distribution 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. As a matter of fact, Jaffard *et al.* in their work ([1,3,11]) use such model and they show in [3] that the spectrum of singularity of the function f is almost surely

$$d_f(h) = \frac{1-\eta}{\alpha}h,\ \forall h \in \left[\alpha, \frac{\alpha}{1-\eta}\right]. \tag{2.3}$$

In a statistical framework, the parameters of interest are the parameters η and α of the lacunar wavelet series f.

We assume that, after discretization, we observe the wavelet coefficients at scale $j_1 = \log_2(n)$ in the following model

$$d_{j_1k} = w_{j_1k} + \epsilon_{j_1k},\ k = 0,\ldots,2^{j_1}-1$$

where ϵ_{j_1k} are i.i.d Gaussian variables, with variance $\frac{\sigma^2}{n}$ where n stands for the number of observations. This model mimics the discretized version of the observation of a function f, observed together with an independent Brownian motion.

Actually, the law of the observed coefficients d_{j_1k} is determined by the prior given by the model (2.2):

$$d_{j_1k} \sim 2^{(\eta-1)j_1}\mathcal{N}\left(2^{-\alpha j_1}, \frac{\sigma^2}{n}\right) + (1-2^{(\eta-1)j_1})\mathcal{N}\left(0, \frac{\sigma^2}{n}\right). \tag{2.4}$$

Hence, if we consider the rescaled observed wavelet coefficients,

$$X_{nk} = \sqrt{n} w_{\log_2(n)i},\ k = 1,\ldots,n,$$

the observations become $X_{n1},\ldots,X_{nn}$, a triangular array of observations, where X_{ni}, $i = 1,\ldots,n$ are i.i.d. random variables with common density

$$f_{\boldsymbol{\theta}} := n^{\eta-1}\phi_\alpha + (1-n^{\eta-1})\phi, \quad \boldsymbol{\theta} = (\eta,\alpha)' \in (0,1)\times(0,1), \tag{2.5}$$

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 $X_{n1},\ldots,X_{nn}$.

https://doi.org/10.1051/ps/2009005 Published online by Cambridge University Press

Remark 2.1. Note that, in this paper, we deal with the estimation of $\boldsymbol{\theta}$, considering only the last level in the wavelet decomposition. Indeed, using all the levels, as studied in [9] does not help in the estimation issue and is more a drawback, leading to complicate expressions for the estimators.

3. Local asymptotic normality (LAN)

3.1. **Main result**

Our aim is to establish Local Asymptotic Normality for the statistical model with i.i.d observations $X_{ni}, i = 1, \ldots, n$ with joint distribution $\mathrm{P}^n_{\boldsymbol{\theta}}$, where $\boldsymbol{\theta} = (\eta, \alpha)' \in (0,1) \times (0,1)$, with

$$\mathrm{P}_{\boldsymbol{\theta}} \sim \mathcal{N}(n^{\frac{1}{2}-\alpha}, \sigma^2) + (1 - n^{\eta-1})\mathcal{N}(0, \sigma^2).$$

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 $\boldsymbol{\Theta}^+ = (0,1) \times (1/2,1)$. Actually, we prove that the family of distributions $\mathcal{P}_n^- := \left\{ \mathrm{P}^n_{\boldsymbol{\theta}} \,\middle|\, \boldsymbol{\theta} \in \boldsymbol{\Theta}^- \right\}$ 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

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

the associated local log-likelihood ratio.

The following result states that the submodel $\mathcal{P}_n^-$ is LAN.

Theorem 3.1. *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*

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

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

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

$$L^n_{\boldsymbol{\theta}+\boldsymbol{\nu}_n(\boldsymbol{\theta})\boldsymbol{\tau}_n/\boldsymbol{\theta}} = \boldsymbol{\tau}'_n \boldsymbol{\Delta}^n_{\boldsymbol{\theta}} - \frac{1}{2}\boldsymbol{\tau}'_n \boldsymbol{\Gamma}\, \boldsymbol{\tau}_n + o_{\mathrm{P}}(1) \quad \text{and} \quad \boldsymbol{\Delta}^n_{\boldsymbol{\theta}} \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma}) \tag{3.1}$$

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

The proof of this result is technical and is postponed to Section 5.

We point out that inference in $\mathcal{P}_n^+$ is totally different. Indeed, for $\boldsymbol{\theta} \in \boldsymbol{\Theta}^+$, 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 (2.5). So, 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.

https://doi.org/10.1051/ps/2009005 Published online by Cambridge University Press

3.2. **Local and asymptotic efficiency of a thresholded estimation procedure**

If we consider the model described in (2.5), 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 [9], 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

$$\hat{\eta}_n = 1 + \frac{1}{\log n} \log(S_n). \tag{3.1}$$

The asymptotics and the optimality of this estimator are given in the following theorem which can be found in [9].

Theorem 3.2. *Under the assumption $\boldsymbol{\theta} \in \boldsymbol{\Theta}^-$, the estimator (3.1) is asymptotically normal. More precisely,*

$$n^{\frac{\eta}{2}} \log(n)\,(\hat{\eta}_n - \eta) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1). \tag{3.2}$$

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 (3.2) and the value of $\boldsymbol{\Gamma}$ (see Thm. 3.1).

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

$$(\boldsymbol{\nu}_n(\boldsymbol{\theta}))^{-1}(\tilde{\boldsymbol{\theta}}_n - \boldsymbol{\theta}) = O_{\mathrm{P}}(1), \tag{3.3}$$

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 notable that our estimator $\hat{\boldsymbol{\theta}}_n$ does not need this one-step improvement, since it does directly reach the efficiency bound.*

4. 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?

https://doi.org/10.1051/ps/2009005 Published online by Cambridge University Press

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 3.1 – that we need a preliminary estimator $\tilde{\alpha}_n$ such that

$$\log(n)n^{\frac{\eta}{2}+\frac{1}{2}-\alpha}(\tilde{\alpha}_n-\alpha)=O_{\mathrm{P}}(1), \tag{4.1}$$

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

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

$$\tilde{\alpha}_n=\frac{1}{\log n}\frac{\sum_{k=1}^n X_{nk}}{\sum_{k=1}^n X_{nk}^2-\sigma^2} \tag{4.2}$$

that was proposed in [9], 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 (4.1), 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 (1.2), 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 literature 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 enables to define locally and asymptotically optimal (maximin, actually) tests about the intensity parameter. More specifically, consider the testing problem (at asymptotic level $\beta\in(0,1)$)

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

for some fixed $\alpha_0<1/2$. Then the LAN result in Theorem 3.1 and the consistency (at the appropriate rate) of our estimator $\hat{\eta}$ in (3.1) straightforwardly yield the following.

Theorem 4.1. *Let $\phi^{(n)}$ be the test that rejects $\mathcal{H}_0^{(n)}$ in favor of $\mathcal{H}_1^{(n)}$ iff*

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

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).

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 difficult to handle in this setup.

https://doi.org/10.1051/ps/2009005 Published online by Cambridge University Press

5. Appendix

Proving Theorem 3.1 is a two step procedure.

- First, we prove that the central sequence defined in Theorem 3.1, $\boldsymbol{\Delta}^n_{\boldsymbol{\theta}} := (\Delta^n_{\boldsymbol{\theta},I}, \Delta^n_{\boldsymbol{\theta},II})'$ is asymptotically normal with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Gamma}$. Asymptotic normality will follow from the Lindeberg-Feller Central Limit Theorem for an array of random variables while Lemma 5.1 provides the asymptotic variance of the central sequence.
- Then we prove that the second-order stochastic expansion in (3.1) holds. Swensen (1985)'s Lemma in [19] enables to conclude, provided its assumptions are satisfied.

Lemma 5.1. *Let*

$$\begin{aligned} Z_{nk} = Z_{nk}(\boldsymbol{\theta}) &:= \frac{1}{2}(\boldsymbol{\nu}_n(\boldsymbol{\theta})\boldsymbol{\tau}_n)'\nabla_{\boldsymbol{\theta}} \log f_{\boldsymbol{\theta}}(X_{nk}) \\ &= \frac{n^{\frac{\eta}{2}-1}}{2f_{\boldsymbol{\theta}}(X_{nk})}\left\{ s_n(\phi_\alpha - \phi)(X_{nk}) + \frac{t_n}{\sigma^2}(n^{\frac{1}{2}-\alpha} - X_{nk})\phi_\alpha(X_{nk})\right\} \\ &= \frac{1}{2}\big\{ s_n D^{nk}_{\boldsymbol{\theta},I} + t_n D^{ni}_{\boldsymbol{\theta},II}\big\}. \end{aligned}$$

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

$$\mathrm{E}\Big[Z^2_{n1}\mathbb{I}(|Z_{n1}| > \delta)\Big] = o(n^{-1}). \tag{5.1}$$

Proof of Lemma 5.1. (i) The first claims are trivial since $\mathrm{E}\big[D^{n1}_{\boldsymbol{\theta},I}\big] = n^{\frac{\eta}{2}-1}\int(\phi_\alpha - \phi)(x)\,\mathrm{d}x = 0$ and $\mathrm{E}\big[D^{n1}_{\boldsymbol{\theta},II}\big] = \sigma^{-2}n^{\frac{\eta}{2}-1}\int(n^{\frac{1}{2}-\alpha} - x)\phi_\alpha(x)\,\mathrm{d}x = 0$.

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

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

where

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

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

and

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

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

https://doi.org/10.1051/ps/2009005 Published online by Cambridge University Press

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

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

where

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

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

and

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

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

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

where

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

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

and

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

Now, $|R_{n1}| \le C \int_{-\infty}^{-r_n} |y|\phi(y)\,\mathrm{d}y$ 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 $R_{n3} = -\sigma^{-2}(\int y\phi(y)\,\mathrm{d}y + o(1)) = o(1)$ as $n \to \infty$.

(v) Define

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

https://doi.org/10.1051/ps/2009005 Published online by Cambridge University Press

Then since

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

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

we obtain that, for all x, $\mathbb{I}_n(x+2r_n) = o(1)$ as $n \to \infty$. Write then $n\mathrm{E}\big[Z_{n1}^2\mathbb{I}(|Z_{n1}| > \delta)\big] \le Cn\big(\mathrm{E}\big[\big(D^{n1}_{\boldsymbol{\theta},I}\big)^2\mathbb{I}(|Z_{n1}| > \delta)\big] + \mathrm{E}\big[\big(D^{n1}_{\boldsymbol{\theta},II}\big)^2\mathbb{I}(|Z_{n1}| > \delta)\big]\big) =: C(U_{n1}+U_{n2})$. 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 5.1(ii), except that the corresponding integrands are multiplied by $\mathbb{I}_n(x)$ in each case. Clearly, the same argument as in Lemma 5.1(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 $\widetilde{T}_{n3} = \int_{r_n}^{\infty}\phi_\alpha(x)\,\mathbb{I}_n(x)\,\mathrm{d}x = \int_{-r_n}^{\infty}\phi(y)\,\mathbb{I}_n(y+2r_n)\,\mathrm{d}y = o(1)$, as $n \to \infty$.

Similarly, defining $\widetilde{S}_{ni}$, $i = 1,2,3$ as in the proof of Lemma 5.1(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 5.1(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 $\widetilde{S}_{n3} = \sigma^{-4}\int_{-r_n}^{\infty} y^2\phi(y)\,\mathbb{I}_n(y+2r_n)\,\mathrm{d}y = o(1)$, as $n \to \infty$. □

Hence, it follows directly from Lemma 5.1 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 (3.1) holds, we use Swensen (1985)'s Lemma ([19]). Defining the quantities

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

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

Lemma 5.2 (Swensen)**.** *Assume that (i)* $\sum_{k=1}^n \mathrm{E}[(Z_{nk}-\xi_{nk})^2] = o(1)$, *(ii)* $\sup_n \sum_{k=1}^n \mathrm{E}[Z_{nk}^2] < \infty$, *(iii)* $\max_{1\le k\le n}|Z_{nk}| = o_{\mathrm{P}}(1)$, *(iv)* $\sum_{k=1}^n Z_{nk}^2 - \frac{1}{4}\boldsymbol{\tau}_n'\boldsymbol{\Gamma}\boldsymbol{\tau}_n = o_{\mathrm{P}}(1)$, *(v)* $\sum_{k=1}^n \mathrm{E}[Z_{nk}^2\mathbb{I}(|Z_{nk}| > 1/2)] = o(1)$, *(vi)* $\mathrm{E}[Z_{nk}] = 0$, *and (vii)* $\sum_{k=1}^n \mathrm{E}[\xi_{nk}^2+2\xi_{nk}] = o(1)$ *(where all expectations and convergences in probability are taken under* $\mathrm{P}^n_{\boldsymbol{\theta}}$*).*

Then the second-order stochastic expansion in (3.1) holds.

This Lemma enables to conclude the proof of Theorem 3.1. Hence it remains to check that the assumptions hold for our particular model.

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 $\sum_{k=1}^n \mathrm{E}[(Z_{nk}-\xi_{nk})^2] = n\mathrm{E}[(Z_{n1}-\xi_{n1})^2]$ is given by

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

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

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

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

https://doi.org/10.1051/ps/2009005 Published online by Cambridge University Press

and

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

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

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

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

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

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

$$V_{n\eta} \leq C n^{-\eta} (\log n)^2 \int \int_\eta^{\eta_n} (\eta_n - \lambda) \, g_{\lambda,\alpha_n}^2(x) \, \frac{\left[g_{\lambda,\alpha_n}(x) + 2 \right]^2}{\left[g_{\lambda,\alpha_n}(x) + 1 \right]^3} \, \phi(x) \, \mathrm{d}\lambda \, \mathrm{d}x, \tag{5.2}$$

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 (5.2) into $C(V_{n\eta,1} + V_{n\eta,2} + V_{n\eta,3})$, where

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

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

and

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

https://doi.org/10.1051/ps/2009005 Published online by Cambridge University Press

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

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

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

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

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

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

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

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

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

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

Now, $\partial_\alpha f_{\eta,\lambda}(x)=\sigma^{-2}n^{\eta-\lambda-\frac{1}{2}}(\log n)(n^{\frac{1}{2}-\lambda}-x)\phi_\lambda(x)$ and $\partial^2_\alpha f_{\eta,\lambda}(x)=\sigma^{-4}n^{\eta-2\lambda}(\log n)^2\phi_\lambda(x)\big\{(n^{\frac{1}{2}-\lambda}-x)^2+\sigma^2 n^{\lambda-\frac{1}{2}}x-2\sigma^2\big\}$, so that

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

https://doi.org/10.1051/ps/2009005 Published online by Cambridge University Press

Letting $y = x - 2\tilde{r}_n$ (with $\tilde{r}_n := n^{\frac{1}{2}-\lambda}/2$) and using $\frac{\phi}{\phi_\lambda}(y + 2\tilde{r}_n) = \exp\big[-2\tilde{r}_n(y+\tilde{r}_n)/\sigma^2\big]$, we obtain that

$$\begin{aligned}
V_{n\alpha} &\le Cn^{-\frac{\eta}{2}-\frac{1}{2}+3\alpha}(\log n)\int\int_{\alpha}^{\alpha_n} n^{-4\lambda}\phi(y) \\
&\quad\times \frac{\big[2\big[1+(n^{1-\eta}-1)\exp\big[-2\tilde{r}_n(y+\tilde{r}_n)/\sigma^2\big]\big]^2\big\{y^2+\sigma^2 y/(2\tilde{r}_n)-\sigma^2\big\}-y^2\big]^2}{\big[1+(n^{1-\eta}-1)\exp\big[-2\tilde{r}_n(y+\tilde{r}_n)/\sigma^2\big]\big]^3}\,\mathrm{d}\lambda\,\mathrm{d}y \\
&\le Cn^{-\frac{\eta}{2}-\frac{1}{2}+3\alpha}(\log n)\int\int_{\alpha}^{\alpha_n} n^{-4\lambda}\phi(y)\max(1,y^4)\,\mathrm{d}\lambda\,\mathrm{d}y \\
&\le Cn^{-\frac{\eta}{2}-\frac{1}{2}-\alpha}(1-\exp(-4n^{-\frac{\eta}{2}-\frac{1}{2}+\alpha}t_n)) \le Cn^{-\eta-1},
\end{aligned}$$

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

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

respectively, we have

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

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

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

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

and

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

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 Lem. 5.1(ii)). As for $V_{n\eta\alpha,2}$, proceeding as in the proof of Lem. 5.1(ii) yields

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

https://doi.org/10.1051/ps/2009005 Published online by Cambridge University Press

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 5.1(ii), we obtain that $V_{n\eta\alpha,3}$ is – up to $o(1)$ terms (as $n \to \infty$) – successively equal to

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

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 $Z_{nk}^2 \le C\big\{(D_{\boldsymbol{\theta},I}^{nk})^2 + (D_{\boldsymbol{\theta},II}^{nk})^2\big\}$. Consequently, Lemma 5.1(i)–(iii) yield $\sum_{k=1}^n \mathrm{E}[Z_{nk}^2] \le C\big\{\mathrm{Var}\big[D_{\boldsymbol{\theta},I}^{nk}\big] + \mathrm{Var}\big[D_{\boldsymbol{\theta},II}^{nk}\big]\big\} = O(1)$, as $n \to \infty$.

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

$$\mathrm{P}\Big[\max_{1\le k\le n} |Z_{nk}| > \delta\Big] = \mathrm{P}\Big[\sum_{k=1}^n Z_{nk}^2 \mathbb{I}(|Z_{nk}| > \delta) > \delta^2\Big] \le \delta^{-2} \sum_{k=1}^n \mathrm{E}\Big[Z_{nk}^2 \mathbb{I}(|Z_{nk}| > \delta)\Big],$$

so that the result follows from Lemma 5.1(v).

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

$$\sum_{k=1}^n Z_{nk}^2 - \frac{1}{4}\boldsymbol{\tau}_n' \boldsymbol{\Gamma}\, \boldsymbol{\tau}_n = \frac{1}{4}\boldsymbol{\tau}_n' \left\{ \left(\sum_{k=1}^n \mathbf{D}_{\boldsymbol{\theta}}^{nk} \mathbf{D}_{\boldsymbol{\theta}}^{nk\prime} \right) - \boldsymbol{\Gamma} \right\} \boldsymbol{\tau}_n,$$

which, by using the boundedness of $(\boldsymbol{\tau}_n)$ and Lemma 5.1(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 5.1(v).

(vi) This is a direct consequence of Lemma 5.1.

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

References

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

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

[3] J.M. Aubry and S. Jaffard, Random wavelet series. *Comm. Math. Phys.* **227** (2002) 483–514.

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

[5] Z. Chi, Construction of stationary self-similar generalized fields by random wavelet expansion. *Probab. Theory Related Fields* **121** (2001) 269–300.

[6] A. Durand, Random wavelet series based on a tree-indexed Markov chain. *Comm. Math. Phys.* **283** (2008) 451–477.

[7] P. Flandrin, Wavelet analysis and synthesis of fractional Brownian Motion. *IEEE Trans. Inform. Theory* **38** (1992) 910–917.

https://doi.org/10.1051/ps/2009005 Published online by Cambridge University Press

[8] F. Gamboa and J.-M. Loubes, Bayesian estimation of multifractal wavelet function. *Bernoulli* (2005) 34–57.
[9] F. Gamboa and J.-M. Loubes, Estimation of the parameters of a multifractal wavelet function. *Test* **16** (2007) 383–407.
[10] C. Genovese and L. Wasserman, Rates of convergence for the Gaussian mixture sieve. *Ann. Statist.* **28** (2000) 1105–1127.
[11] S. Jaffard, On lacunary wavelet series. *The Annals of Applied Probability* **10** (2000) 313–329.
[12] B. Lindsay, The geometry of mixture likelihoods: a general theory. *Ann. Statist.* **11** (1983) 86–94.
[13] G. McLachlan and K. Basford, Mixture models. Inference and applications to clustering. *Statistics: Textbooks and Monographs* **84.** Marcel Dekker, Inc., New York (1988).
[14] S.G. Mallat, Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbf{r})$. *Transactions of the American Mathematical Society* **315** (1989) 69–87.
[15] D.L. McLeish and C.G. Small, Likelihood methods for the discrimination problem. *Biometrika* **73** (1986) 397–403.
[16] Y. Meyer, *Ondelettes et Opérateurs*. Hermann (1990).
[17] R.H. Riedi, M.S. Crouse, V.J. Ribeiro and R.G. Baraniuk, A multifractal wavelet model with application to network traffic. *Institute of Electrical and Electronics Engineers. Transactions on Information Theory* **45** (1999) 992–1018.
[18] F. Roueff, Almost sure haussdorff dimensions of graphs of random wavelet series. *J. Fourier Analysis and App.* **9** (2003).
[19] A.R. Swensen, The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend. *J. Multivariate Anal.* **16** (1985) 54–70.
[20] S. van de Geer, Rates of convergence for the maximum likelihood estimator in mixture models. *J. Nonparametr. Statist.* **6** (1996) 293–310.
[21] A.W. van der Vaart, *Asymptotic statistics*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge (1998). ISBN 0-521-49603-9; 0-521-78450-6.

https://doi.org/10.1051/ps/2009005 Published online by Cambridge University Press