Commun. Korean Math. Soc. **33** (2018), No. 4, pp. 1367–1376
https://doi.org/10.4134/CKMS.c170130
pISSN: 1225-1763 / eISSN: 2234-3024

DENSITY SMOOTHNESS PARAMETER ESTIMATION WITH SOME ADDITIVE NOISES

Junjian Zhao and Zhitao Zhuang

Abstract. In practice, the density function of a random variable X is always unknown. Even its smoothness parameter is unknown to us. In this paper, we will consider a density smoothness parameter estimation problem via wavelet theory. The smoothness parameter is defined in the sense of equivalent Besov norms. It is well-known that it is almost impossible to estimate this kind of parameter in general case. But it becomes possible when we add some conditions (to our proof, we can not remove them) to the density function. Besides, the density function contains impurities. It is covered by some additive noises, which is the key point we want to show in this paper.

1. Introduction

The smoothness parameter estimation of a density function is essential in studying the rate of convergence of that function's estimators. Generally speaking, density function is assumed differentiable ([15]) or in Sobolev or Besov spaces ([1,4]) with smoothness parameters s. When s is unknown, the adpative methods presented in [2,3,9] make it possible to obtain the optimal rate of convergence (or sub-optimal) of a density estimator. There are also non-adaptive methods ([6]) of estimation which require the exact value of the parameters s over certain spaces. In practice, a fundamental problem is needed for verifying the smoothness assumption and studying the smoothness tests (e.g., [10, 11]). Besides, the paper [8] gives a new light on smoothness parameter estimation and confirms the need of this estimation.

It is well-known that it is almost impossible to give effective estimation of smoothness parameter in general case ([2,13]) over such like Besov spaces. But it does not mean that it is impossible in any case. The paper [6] shows that one can effectively estimate the smoothness parameter of some classes of density

Received March 28, 2017; Revised September 20, 2017; Accepted November 7, 2017.
2010 *Mathematics Subject Classification.* 62G05, 62G07, 42C40, 35Q30, 41A15.
Key words and phrases. smoothness parameter estimation, density, wavelets, additive noise, Besov spaces.
This work was supported by the National Natural Science Foundation of China(Grant No. 11401433).

functions, such as piecewise-smooth functions. Using characterization of Besov space $B^s_{p,\infty}$ (defined below) in terms of wavelet coefficients, [6] constructed a "pseudo-consistent" estimator of the smoothness parameter, which is strongly consistent in the case of density function is a piecewise-smooth function.

Note that the model in [6] does not contain any noise, but not the case in reality. The density estimation for a statistical model with additive noise plays important roles in both statistics and econometrics. For example, Gaussian noise is most widely studied because of importance in both theory and applications ([9]). However, non-Gaussian noises appear in many areas ([14,16]). In this paper, we will study moderately ill-posed noise, whose density function φ satisfies some conditions in Section 2. With this noise, we derive a estimation of smoothness parameter. Furthermore, a non-linear wavelet estimator (defined by thresholding method) gives a better estimation [5,7] than the classical methods, due to time and frequency localization of wavelet bases. So we will also study this kind of smoothness parameter estimation under non-linear wavelet estimators' sense.

In this section, we will list a series of notations for Besov spaces. Some lemmas or preliminary results for smoothness parameter estimation will be given in Section 2. Based on this preparation, we present our main results in the last section.

Let $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{R}$ be the set of positive integers, the set of integers and the set of real numbers, respectively. Let $\mathbb{R}^d$ be the classical d-dimensional real number space. Through out this paper, we use $A \lesssim B$ to abbreviate that A is bounded by a constant multiple of B, $A \gtrsim B$ is defined as $B \lesssim A$ and $A \sim B$ means $A \lesssim B$ and $B \lesssim A$. For a Lebesgue measurable function f, the support of f means the set $\mathrm{Supp}(f) := \{x \in \mathbb{R} : f(x) \neq 0\}$, which is well-defined up to a set of measure 0.

Define $f_{j,k}(\cdot) := 2^{\frac{j}{2}} f(2^j \cdot - k)$ through out this paper except for special explanation. The classical Fourier transform is given by $\hat{f}(\xi) := \int_{\mathbb{R}^d} f(x) e^{-ix\xi} dx$ for $f \in L^1(\mathbb{R}^d)$ and a standard extension in other cases. A Muliresolution analysis (MRA) is defined below, which is a sequence of approximated spaces allowing the construction of wavelets.

Definition 1.1. (MRA) A multiresolution analysis of $L^2(\mathbb{R})$ is a sequence of closed subspaces $\{V_j\}_{j\in\mathbb{Z}}$ verifying:

(i) $\forall j$, $V_j \subseteq V_{j+1}$; $\bigcap_j V_j = \{0\}$ and $\overline{\bigcup_j V_j} = L^2(\mathbb{R})$;

(ii) $f \in V_j \Leftrightarrow f(2\cdot) \in V_{j+1}$;

(iii) There exists a function $\phi \in V_0$ such that the family $\{\phi(x-k)\}_{k\in\mathbb{Z}}$ form a (Riesz) basis of V_0.

We can derive a wavelet function via $\psi(x) := \sum_k (-1)^k \overline{h_{1-k}} \phi_{1,k}(x)$ with $h_k = \int \phi\phi_{1,k} dx$. It is clear that both $\{\phi_{0,k}, \psi_{j,k}\}_{j\geq 0, k\in\mathbb{Z}}$ and $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ are orthonormal bases of $L^2(\mathbb{R})$. This theory will be used in characterization of Besov spaces.

Let $0 < p, q \le \infty$, $s > 0$ and $s = n + \delta$ with $n \in \mathbb{N}$, $\delta \in (0, 1]$. $W_p^n(\mathbb{R}^d)$ is the classical L^p-Sobolev spaces with $\|f\|_{W_p^n(\mathbb{R}^d)} := \|f^{(n)}\|_{L^p(\mathbb{R}^d)} + \|f\|_{L^p(\mathbb{R}^d)}$. Besov spaces are defined by

$$B_{p,q}^s(\mathbb{R}^d) := \{f \in W_p^n(\mathbb{R}^d) : \|t^{-\delta}\omega_p^2(f^{(n)}, t)\|_q^* < \infty\}.$$

Here, $\omega_p^2(f, t) = \sup\limits_{|h| \le t} \|f(\cdot + 2h) - 2f(\cdot + h) + f(\cdot)\|_{L^p(\mathbb{R}^d)}$ denotes the 2-th order smoothness modulus of f, and

$$\|f\|_q^* := \begin{cases} (\int_0^\infty |f(t)|^q \frac{dt}{t})^{\frac{1}{q}}, & 1 \le q < \infty; \\ \text{ess sup}_t |f(t)|, & q = \infty. \end{cases}$$

The Besov (quasi-)norm is given by $\|f\|_{B_{p,q}^s(\mathbb{R}^d)} := \|f\|_{W_p^n(\mathbb{R}^d)} + |f|_{B_{p,q}^s(\mathbb{R}^d)}$.

When $d = 1$, Besov space is denoted by $B_{p,q}^s(\mathbb{R})$. A function ϕ is called r-regular, if $\phi \in C^r(\mathbb{R})$ and $\phi^{(m)}(x) \le C(1 + |x|^2)^{-l}$ for each $l \in \mathbb{Z}$ and $m = 0, 1, \ldots, r$. Based on r-regular condition, characterization of Besov spaces is given below in terms of wavelet coefficients. It should be pointed out that this result is the foundation of our estimation. Meanwhile, we use $\|\lambda\|_{l_p}$ to denote $l^p(\mathbb{Z})$ norm for $\lambda := \{\lambda_k\}_{k \in \mathbb{Z}} \in l^p(\mathbb{Z})$, where $\|\lambda\|_{l_p} = (\sum_k |\lambda_k|^p)^{\frac{1}{p}}$ for $p < \infty$ and $\|\lambda\|_{l_\infty} = \sup_k |\lambda_k|$.

Lemma 1.1 ([9]). *Let ϕ be r-regular with $0 < s < r$ and ψ be the corresponding wavelets. If $f \in L^p(\mathbb{R})$, then the followings are equivalent.*

(i) $f \in B_{p,q}^s(\mathbb{R})$, $1 \le p, q \le \infty$;

(ii) $\|(\alpha_{0,k})_{k \in \mathbb{Z}}\|_{l_p} + \|(2^{j(s + \frac{1}{2} - \frac{1}{p})}\|(\beta_{j,k})_{k \in \mathbb{Z}}\|_{l_p})_{j \ge 0}\|_{l_q} < \infty$, *where* $\alpha_{0,k} = \int f(x)\phi_{0,k}(x)dx$, $\beta_{j,k} = \int f(x)\psi_{j,k}(x)dx$.

Moreover,

$$\|f\|_{B_{p,q}^s(\mathbb{R})} \sim \|(\alpha_{0,k})_{k \in \mathbb{Z}}\|_{l_p} + \|(2^{j(s + \frac{1}{2} - \frac{1}{p})}\|(\beta_{j,k})_{k \in \mathbb{Z}}\|_{l_p})_{j \ge 0}\|_{l_q}.$$

Note that the constants of upper and lower bounds for equivalent norm are all only dependent on r, s, p, q.

Throughout this paper we shall denote by C and C_j $(j = 1, 2, \ldots)$ for various positive constants whose exact value may change from lines to lines but are not essential to the analysis of problem.

2. Estimation of smoothness parameter

Let us define a smoothness parameter of a function $f \in L^p(\mathbb{R})$ for some $1 \le p < \infty$ or $f \in C_b(\mathbb{R})$ in case $p = \infty$ as ($C_b(\mathbb{R})$ is the space of all continues and bounded functions)

$$s_p^* := \sup\{s : f \in B_{p,\infty}^s(\mathbb{R})\}.$$

We take $\sup \emptyset = 0$ and $\sup \mathbb{R} = \infty$. By Lemma 1.1, a function $f \in L^p(\mathbb{R})$ with the expansion $f = \sum_k \alpha_{0,k}\phi_{0,k} + \sum_{j=0}^\infty \sum_k \beta_{j,k}\psi_{j,k}$ belongs to $B_{p,\infty}^s(\mathbb{R})$ if and only if

$$\sup_{j\geq 0}(2^{j(s-\frac{1}{p}+\frac{1}{2})}\|(\beta_{j,k})_k\|_{l_p})<\infty \text{ and } \|(\alpha_{0,k})_k\|_{l_p}<\infty.$$

With the help of the above results, [6] gives an important result below.

Lemma 2.1 ([6]). *Let ϕ, ψ come from a r-regular MRA, $r \geq 1$, $1 \leq p \leq \infty$. If $0 < s_p^* < r$, then the set of indices*

$$J := \{j \geq 1 : \|(\beta_{j,k})_k\|_{l_p} \neq 0\}$$

has infinitely many elements and

$$\liminf_{j\to\infty, j\in J}\frac{-\log_2\|(\beta_{j,k})_k\|_{l_p}}{j} = s_p^* - \frac{1}{p} + \frac{1}{2}.$$

As usual, let P_j and Q_j be the orthogonal projections from $L^2(\mathbb{R})$ to V_j and W_j, respectively. Here $W_j := V_{j+1} \ominus V_j$.

$$P_j f := \sum_k \alpha_{j,k}\phi_{j,k} \text{ and } Q_j f := P_{j+1}f - P_j f = \sum_k \beta_{j,k}\psi_{j,k},$$

with $\alpha_{j,k} := \int f\phi_{j,k}dx$ and $\beta_{j,k} := \int f\psi_{j,k}dx$. We can extend these definitions of projections to $L^p(\mathbb{R})$ sense [17], i.e.,

$$P_j f := \sum_k \alpha_{j,k}\phi_{j,k} \text{ and } Q_j f := \sum_k \beta_{j,k}\psi_{j,k}$$

for $f \in L^p(\mathbb{R})$. And a lemma is followed by this definition.

Lemma 2.2 ([9]). *Let h be a scaling or wavelet function with*

$$\theta(h) := \sup_{x\in\mathbb{R}}\sum_k |h(x-k)| < \infty.$$

Then there exist $0 < C_1 \leq C_2 < \infty$ such that

$$C_1 2^{j(\frac{1}{2}-\frac{1}{p})}\|\lambda\|_{l_p} \leq \|\sum_k \lambda_k h_{j,k}\|_{L^p} \leq C_2 2^{j(\frac{1}{2}-\frac{1}{p})}\|\lambda\|_{l_p}$$

for $1 \leq p \leq \infty$, $\lambda := \{\lambda_k\}_{k\in\mathbb{Z}} \in l^p(\mathbb{Z})$.

So P_j and Q_j are well-defined for $f \in L^p(\mathbb{R})$. The Daubechies and Meyer's scaling and wavelet functions satisfy $\theta(h) < \infty$. Then, we have a corollary by replacing $\sum_k \lambda_k h_{j,k}$ with $Q_j f$.

Corollary 2.1. *Let ψ be a r-regular wavelet function, $1 \leq p \leq \infty$ and $0 < s_p^* < r$. Then*

$$\liminf_{j\to\infty, j\in J}\frac{-\log_2\|Q_j f\|_{L^p}}{j} = s_p^*.$$

We say this is "pseudo-consistent" because of "$\liminf_{j\to\infty, j\in J}$" but not "$\lim_{j\to\infty}$", and "$\lim_{j\to\infty}$" is called strongly consistent. Through out this paper, define $\log_2 0 = -\infty$. What we want to show is that whether we have similar results when $\beta_{j,\cdot}$ (or Q_jf) is changed into an estimator $\widehat{\beta}_{j,\cdot}$ (or $\widehat{Q_jf}$) with additive noise, and their definitions are in Theorem 3.1 and Remark 3.1.

We will study a moderately ill-posed noise ε, whose density function φ satisfies that for some $\nu \geq 0$ (see [12] for more details, here we only need the third condition):

(C) $\widehat{\varphi}(t) \geq C(1+|t|^2)^{-\frac{\nu}{2}}$.

When φ degenerates to the Dirac functional, its Fourier transform $\widehat{\varphi} \equiv 1$ and (C) hold automatically ($\nu = 0$). Then the study goes back to the classical one in [6].

Let $Y_1, Y_2, \ldots, Y_n$ be independent and identically distributed (i.i.d) random variables of

$$Y = X + \varepsilon,$$

where X stands for real-valued random variable with unknown probability density $f : \mathbb{R} \to \mathbb{R}^+$ ($\mathbb{R}^+$ is the nonnegative real number set) and ε denotes an independent random noise with the probability density φ. The problem is to estimate the smoothness parameter of f via $Y_1, Y_2, \ldots, Y_n$ in some cases. If the density of Y is g, as a deconvolution problem, density g equals to the convolution of f and φ. That is $g = f * \varphi := \int f(\cdot - x)\varphi(x)dx$.

As in [12], we introduce

$$\mathbf{K}_j h(y) := \frac{1}{2\pi}\int_{\mathbb{R}} e^{ity}\frac{\widehat{h}(t)}{\widehat{\varphi}(-2^j t)}dt,\ \widehat{\gamma}_{j,k} := \frac{1}{n}\sum_{i=1}^{n}(\mathbf{K}_j h)_{j,k}(Y_i).$$

Replacing $f \in B^{s'}_{p,q}(\mathbb{R})$ by $\|f\|_{L^\infty(\mathbb{R})} < \infty$ in Lemma 2.3 of [12], and then put it into Remark 2.2 in [12], we will have:

Lemma 2.3. *Let φ satisfy* (C)*, h be $r+2$-regular scaling or wavelet function and h, f have compact supports with $\|f\|_{L^\infty} < \infty$, $1 \leq p < \infty$. Define $\widehat{\gamma}_{j,\cdot} := (\widehat{\gamma}_{j,k})_k, \gamma_{j,\cdot} := (\gamma_{j,k})_k$ and $E\widehat{\gamma}_{j,k} = \gamma_{j,k}$ with $j2^j \leq n$. Then*

$$E\|\widehat{\gamma}_{j,\cdot} - \gamma_{j,\cdot}\|^p_{l_p} \lesssim n^{-\frac{p}{2}} 2^{j(rp+1)}.$$

Lemma 2.4. *By Lemma 2.4 in* [12]*, when h is the Meyer scaling or wavelet function, without assuming compact support of f, we also have the same result for $p \geq 2$ and $\|f\|_{L^p(\mathbb{R})} < \infty$ (because of $f * \varphi \in L^{\frac{p}{2}}$).*

3. Main results and proofs

Now, it's time to present our Main results. Theorem 3.1 shows that we can use $\widehat{s^*_p} := \frac{-\log_2 \|\widehat{\beta}_{j,\cdot}\|_{l_p}}{j} + \frac{1}{p} - \frac{1}{2}$ as an estimator of s^*_p. This result can be extended to the case of by using Meyer scaling (wavelet) functions, and even by the thresholding methods.

Theorem 3.1. *Let ϕ, ψ come from a $(r+2)$-regular MRA with compact supports, $r > 0$, $1 \leq p \leq \infty$. Let $Y_1, Y_2, \ldots, Y_n$ be a sequence of i.i.d random variables with density $f * \varphi$ (n is the size of experiment) and φ satisfy condition*

(C). *Assume that* $\|f\|_{L^\infty(\mathbb{R})} < \infty$ *and* f *has compact support. If* $0 < s_p^* < r$, $n \sim 2^{2j(2r+\frac{1}{2})}$, *then*

$$\liminf_{j\to\infty, j\in J} \frac{-\log_2 \|\widehat{\beta}_{j,\cdot}\|_{l_p}}{j} = s_p^* - \frac{1}{p} + \frac{1}{2}.\ a.e.,$$

where $\widehat{\beta}_{j,k} := \frac{1}{n}\sum_{i=1}^{n}(\mathbf{K}_j\psi)_{j,k}(Y_i)$, $\widehat{\beta}_{j,\cdot} := (\widehat{\beta}_{j,k})_k$ *and* $\beta_{j,\cdot} := (\beta_{j,k})_k$.

Proof. Let $1 \le p < \infty$, using mathematical expectation's definition, we have

$$P\{\|\widehat{\beta}_{j,\cdot} - \beta_{j,\cdot}\|_{l_p}^p > 2^{-jp(s-\frac{1}{p}+\frac{1}{2})}\} \le \frac{E\|\widehat{\beta}_{j,\cdot} - \beta_{j,\cdot}\|_{l_p}^p}{2^{-jp(s-\frac{1}{p}+\frac{1}{2})}}.$$

This with Lemma 2.3,

$$P\{\|\widehat{\beta}_{j,\cdot} - \beta_{j,\cdot}\|_{l_p} > 2^{-j(s-\frac{1}{p}+\frac{1}{2})}\} \le Cn^{-\frac{p}{2}}2^{j(rp+1)}2^{jp(s-\frac{1}{p}+\frac{1}{2})}.$$

Considering the choice of n,

$$\sum_{j=0}^{\infty} P\{\|\widehat{\beta}_{j,\cdot} - \beta_{j,\cdot}\|_{l_p} > 2^{-j(s-\frac{1}{p}+\frac{1}{2})}\} < \infty$$

holds for $0 < s < r$. Since $P\{\|\widehat{\beta}_{j,\cdot} - \beta_{j,\cdot}\|_{l_p} > 2^{-j(s-\frac{1}{p}+\frac{1}{2})}\} \ge P\{|\|\widehat{\beta}_{j,\cdot}\|_{l_p} - \|\beta_{j,\cdot}\|_{l_p}| > 2^{-j(s-\frac{1}{p}+\frac{1}{2})}\} \ge P\{\|\widehat{\beta}_{j,\cdot}\|_{l_p} > 2^{-j(s-\frac{1}{p}+\frac{1}{2})} + \|\beta_{j,\cdot}\|_{l_p}\}$, it follows that

$$\sum_{j=0}^{\infty} P\{\|\widehat{\beta}_{j,\cdot}\|_{l_p} - \|\beta_{j,\cdot}\|_{l_p} > C2^{-j(s-\frac{1}{p}+\frac{1}{2})}\} < \infty$$

and

$$\sum_{j=0}^{\infty} P\{\|\widehat{\beta}_{j,\cdot}\|_{l_p} > 2^{-j(s-\frac{1}{p}+\frac{1}{2})} + \|\beta_{j,\cdot}\|_{l_p}\} < \infty \tag{1}$$

for $0 < s < r$. By the definition of s_p^*, for each $0 < s < s_p^*$, there exists $N_s \in \mathbb{N}$, $\|\beta_{j,\cdot}\|_{l_p} \le 2^{-j(s-\frac{1}{p}+\frac{1}{2})}$ for $j \ge N_s$. Then $\sum_{j=N_s}^{\infty} P\{\|\widehat{\beta}_{j,\cdot}\|_{l_p} \ge 2\cdot 2^{-j(s-\frac{1}{p}+\frac{1}{2})}\} < \infty$, which means

$$\sum_{j=N_s}^{\infty} P\{\frac{-\log_2 \|\widehat{\beta}_{j,\cdot}\|_{l_p}}{j} \le s - \frac{1}{p} + \frac{1}{2} - \frac{1}{j}\} < \infty.$$

Apply Borel-Cantelli Lemma to this inequality,

$$\frac{-\log_2 \|\widehat{\beta}_{j,\cdot}(\omega)\|_{l_p}}{j} > s - \frac{1}{p} + \frac{1}{2} - \frac{1}{j}$$

for almost all $\omega \in \Omega$ with $j \ge N_s$. So

$$\liminf_{j\to\infty} \frac{-\log_2 \|\widehat{\beta}_{j,\cdot}\|_{l_p}}{j} \ge s - \frac{1}{p} + \frac{1}{2} \text{ a.e.} \tag{2}$$

for $0 < s < s_p^*$.

It is left to show the supreme of s is $\liminf_{j\to\infty}\frac{-\log_2\|\widehat{\beta}_{j,\cdot}\|_{l_p}}{j}+\frac{1}{p}-\frac{1}{2}$.

If $s>s_p^*$, there exists a subsequence $\{j_k\}_{k=1}^{\infty}$ such that

$$\|\beta_{j_k,\cdot}\|_{l_p}\geq 2^{-j_k(s-\frac{1}{p}+\frac{1}{2})}$$

and $P\{\|\widehat{\beta}_{j_k,\cdot}\|_{l_p}<\frac{1}{2}\cdot 2^{-j_k(s-\frac{1}{p}+\frac{1}{2})}\}=P\{\|\beta_{j,\cdot}\|_{l_p}-\|\widehat{\beta}_{j_k,\cdot}\|_{l_p}>\|\beta_{j,\cdot}\|_{l_p}-\frac{1}{2}\cdot 2^{-j_k(s-\frac{1}{p}+\frac{1}{2})}\}\leq P\{|\|\beta_{j_k,\cdot}\|_{l_p}-\|\widehat{\beta}_{j_k,\cdot}\|_{l_p}|>\frac{1}{2}\cdot 2^{-j_k(s-\frac{1}{p}+\frac{1}{2})}\}$. Applying the same computation procedure of (1), one has $\sum_{k=1}^{\infty}P\{\|\widehat{\beta}_{j_k,\cdot}\|_{l_p}<\frac{1}{2}\cdot 2^{-j_k(s-\frac{1}{p}+\frac{1}{2})}\}<\infty$, i.e.,

$$\sum_{k=1}^{\infty}P\{\frac{-\log_2\|\widehat{\beta}_{j_k,\cdot}\|_{l_p}}{j_k}>s-\frac{1}{p}+\frac{1}{2}+\frac{1}{j_k}\}<\infty$$

for $s_p^*<s<r$. Also by Borel-Cantelli Lemma,

$$\liminf_{j\to\infty}\frac{-\log_2\|\widehat{\beta}_{j,\cdot}\|_{l_p}}{j}\leq s-\frac{1}{p}+\frac{1}{2}\text{ a.e.}\tag{3}$$

for $s_p^*<s<r$. Next, two class of spaces are defined below,

$$A_s:=\{\omega\in\Omega:\liminf_{j\to\infty}\frac{-\log_2\|\widehat{\beta}_{j,\cdot}\|_{l_p}}{j}\geq s-\frac{1}{p}+\frac{1}{2}\}$$

and

$$B_s:=\{\omega\in\Omega:\liminf_{j\to\infty}\frac{-\log_2\|\widehat{\beta}_{j,\cdot}\|_{l_p}}{j}\leq s-\frac{1}{p}+\frac{1}{2}\}.$$

Thus,

$$P(A_s)=1\text{ for any }0<s<s_p^*\text{ and }P(B_s)=1\text{ for any }s_p^*<s<r$$

with the help of (2) and (3). Because

$$\{\omega\in\Omega:\liminf_{j\to\infty}\frac{-\log_2\|\widehat{\beta}_{j,\cdot}\|_{l_p}}{j}=s_p^*-\frac{1}{p}+\frac{1}{2}\}=(\bigcap_{k=1}^{\infty}A_{s_p^*-\frac{1}{k}})\bigcap(\bigcap_{k=1}^{\infty}B_{s_p^*+\frac{1}{k}}).$$

By DeMorgan's Law,

$$P\{\omega\in\Omega:\liminf_{j\to\infty}\frac{-\log_2\|\widehat{\beta}_{j,\cdot}\|_{l_p}}{j}=s_p^*-\frac{1}{p}+\frac{1}{2}\}$$
$$\geq 1-\sum_{k=1}^{\infty}P(A^C_{s_p^*-\frac{1}{k}})-\sum_{k=1}^{\infty}P(B^C_{s_p^*+\frac{1}{k}})=1.$$

Then the result holds for $1\leq p<\infty$.

To end of the proof, we have to show the case of $p=\infty$. Note that $P\{\|\widehat{\beta}_{j,\cdot}-\beta_{j,\cdot}\|_{l_\infty}>2^{-j(s+\frac{1}{2})}\}\leq P\{\|\widehat{\beta}_{j,\cdot}-\beta_{j,\cdot}\|_{l_2}>2^{-j(s+\frac{1}{2})}\}=P\{\|\widehat{\beta}_{j,\cdot}-\beta_{j,\cdot}\|^2_{l_2}>2^{-2j(s+\frac{1}{2})}\}\leq 2^{2j(s+\frac{1}{2})}E\|\widehat{\beta}_{j,\cdot}-\beta_{j,\cdot}\|^2_{l_2}$.

Then we get the result analogously and the proof is completed. □

Meanwhile, one can get a similar result by using Meyer wavelets.

Corollary 3.1. *Under the same assumptions of Theorem 3.1 except for* $\|f\|_{L^p(\mathbb{R})} < \infty$ *with* $p \geq 2$ (f *without compact support*), *and* ϕ (ψ) *is Meyer scaling* (*wavelet*) *function. Then by Lemma 2.4, we have*

$$\liminf_{j\to\infty, j\in J} \frac{-\log_2 \|\widehat{\beta}_{j,\cdot}\|_{l_p}}{j} = s_p^* - \frac{1}{p} + \frac{1}{2}, \text{ a.e.}$$

As an example of Theorem 3.1, we have an estimation for piecewise-smooth function below.

Corollary 3.2. *Let* ψ *be Daubechies wavelet, let* f *be a density satisfying* $f \in C^{m-1}(\mathbb{R})$ ($C^m(\mathbb{R})$ *is the set of* m*-order continuous differentiable functions with* $m \in \mathbb{N}$), $f \in C^{m+1}((-\infty, a])$ *and* $f \in C^{m+1}([a, +\infty,))$ *but* $f^{(m)}(a^-) \neq f^{(m)}(a^+)$. *The other conditions are the same as Theorem 3.1, then*

$$\lim_{j\to\infty} \frac{-\log_2 \|\widehat{\beta}_{j,\cdot}\|_{l_p}}{j} = m + \frac{1}{p} = s_p^* - \frac{1}{p} + \frac{1}{2}, \text{ a.e.}$$

The proof of this corollary is similar to Corollary 4.2 in [6], so we omit it here. It should be pointed out that we replace Theorem 4.1 of [6] with Theorem 3.1 in the proof's procedure.

Moreover, the thresholding methods can also get a good estimation.

Theorem 3.2. *Let* ϕ, ψ *come from a* $(r+2)$*-regular MRA,* $r > 0$, $1 \leq p \leq \infty$. *Let* $Y_1, Y_2, \ldots, Y_n$ *be a sequence of i.i.d random variables with density* $f * \varphi$ (n *is the size of experiment*). *Assume that* $\|f\|_{L^\infty(\mathbb{R})} < \infty$ *and* f *has compact support. If* $0 < s_p^* < r$, *then*

$$\liminf_{j\to\infty, j\in J} \frac{-\log_2 \|\widetilde{\beta}_{j,\cdot}\|_{l_p}}{j} = s_p^* - \frac{1}{p} + \frac{1}{2}, \text{ a.e.},$$

where $\widetilde{\beta}_{j,k} := \begin{cases} \widehat{\beta}_{j,k}, & |\widehat{\beta}_{j,\cdot}| > \tau_{j,n}; \\ 0, & |\widehat{\beta}_{j,\cdot}| \leq \tau_{j,n}, \end{cases}$ *and* $\tau_{j,n} \sim 2^{jr}(\frac{1}{n})^{\frac{1}{2}}$, $n \sim 2^{2j(2r+\frac{1}{2})}$.

Proof. With the help of Theorem 3.1, what we only to show is the estimation

$$E\|\widetilde{\beta}_{j,\cdot} - \beta_{j,\cdot}\|_{l_p}^p \lesssim n^{-\frac{p}{2}} 2^{j(rp+1)}.$$

By triangle inequality, $E\|\widetilde{\beta}_{j,\cdot} - \beta_{j,\cdot}\|_{l_p}^p \lesssim E\|\widehat{\beta}_{j,\cdot} - \beta_{j,\cdot}\|_{l_p}^p + E\|\widetilde{\beta}_{j,\cdot} - \widehat{\beta}_{j,\cdot}\|_{l_p}^p$, the left is to show $E\|\widetilde{\beta}_{j,\cdot} - \widehat{\beta}_{j,\cdot}\|_{l_p}^p \lesssim n^{-\frac{p}{2}} 2^{j(rp+1)}$. In fact,

$$E\|\widetilde{\beta}_{j,\cdot} - \widehat{\beta}_{j,\cdot}\|_{l_p}^p = E\|\widehat{\beta}_{j,\cdot} I\{|\widehat{\beta}_{j,\cdot}| \leq \tau_{j,n}\}\|_{l_p}^p \lesssim n^{-\frac{p}{2}} 2^{j(rp+1)}$$

because of the choice of $\tau_{j,n}$, and $I\{S\}$ is the classical indicator function on a set S. It is similar when $p = \infty$ and the result holds. □

Remark 3.1. It is easy to show

$$\liminf_{j\to\infty, j\in J} \frac{-\log_2 \|\widehat{Q_j f}\|_{L^p}}{j} = s_p^*, \text{ a.e.}$$

with $\widehat{Q_j f} := \sum_k \widehat{\beta}_{j,k}\psi_{j,k}$; Note that when using Meyer scaling and wavelet functions, we need only assume $f \in L^\infty(\mathbb{R})$ without compact support, and then the range of p is $p \geq 2$ (by Lemma 2.4 of [12]), which is the same as Theorem 4.1 of [6].

Note that the condition $0 < s_p^* < r$ is needed in the above two theorems. The reason comes from the characterization result of Lemma 1.1, which plays an important role in this paper. In some sense, we get a necessary condition for the smoothness parameter estimation of density function.

Besides, the proof of Theorem 3.1 is very similar to Theorem 4.1 in [6]. But the range of p in our result is larger because of the help of Lemma 2.3. And our density function is covered by additive noises φ, i.e., random variable sequence $(X_1, X_2, \ldots, X_n)$ is replaced by $(Y_1, Y_2, \ldots, Y_n)$. So we also have a different choice of size experiment n. Of course, in order to prove Theorem 3.2 more succinctly, we give a detailed proof for Theorem 3.1.

References

[1] T. T. Cai, *Adaptive wavelet estimation: a block thresholding and oracle inequality approach*, Ann. Statist. **27** (1999), no. 3, 898–924.

[2] T. T. Cai and M. G. Low, *Adaptive confidence balls*, Ann. Statist. **34** (2006), no. 1, 202–228.

[3] E. Chicken and T. T. Cai, *Block thresholding for density estimation: local and global adaptivity*, J. Multivariate Anal. **95** (2005), no. 1, 76–106.

[4] D. L. Donoho and I. M. Johnstone, *Minimax estimation via wavelet shrinkage*, Ann. Statist. **26** (1998), no. 3, 879–921.

[5] D. L. Donoho, I. M. Johnstone, G. Kerkyacharian, and D. Picard, *Density estimation by wavelet thresholding*, Ann. Statist. **24** (1996), no. 2, 508–539.

[6] K. Dziedziul and B. Ćmiel, *Density smoothness estimation problem using a wavelet approach*, ESAIM Probab. Stat. **18** (2014), 130–144.

[7] J. Fan and J.-Y. Koo, *Wavelet deconvolution*, IEEE Trans. Inform. Theory **48** (2002), no. 3, 734–747.

[8] A. Gloter and M. Hoffmann, *Nonparametric reconstruction of a multifractal function from noisy data*, Probab. Theory Related Fields **146** (2010), no. 1-2, 155–187.

[9] W. Härdle, G. Kerkyacharian, D. Picard, and A. Tsybakov, *Wavelets, Approximation, and Statistical Applications*, Lecture Notes in Statistics, **129**, Springer-Verlag, New York, 1998.

[10] L. Horvath and P. Kokoszka, *Change-point detection with non-parametric regression*, Statistics **36** (2002), no. 1, 9–31.

[11] Y. Ingster and N. Stepanova, *Estimation and detection of functions from anisotropic Sobolev classes*, Electron. J. Stat. **5** (2011), 484–506.

[12] R. Li and Y. Liu, *Wavelet optimal estimations for a density with some additive noises*, Appl. Comput. Harmon. Anal. **36** (2014), no. 3, 416–433.

[13] M. G. Low, *On nonparametric confidence intervals*, Ann. Statist. **25** (1997), no. 6, 2547–2554.

[14] A. Martinez, *Communication by energy modulation: the additive exponential noise channel*, IEEE Trans. Inform. Theory **57** (2011), no. 6, 3333–3351.

[15] S. J. Sheather and M. C. Jones, *A reliable data-based bandwidth selection method for kernel density estimation*, J. Roy. Statist. Soc. Ser. B **53** (1991), no. 3, 683–690.

[16] J. Sun, H. Morrison, P. Harding, and M. Woodroofe, *Density and mixture estimation from data with measurement errors*, Technical report, 2002.

[17] J. Zhao, *The convergence of wavelet expansion with divergence-free properties in vector-valued Besov spaces*, Appl. Math. Comput. **251** (2015), 143–153.

Junjian Zhao
Department of Mathematics
College of Science
Tianjin Polytechnic University
No. 399 Binshuixi Street, Xiqing District, Tianjin 300387, P. R. China
Email address: zhaojunjian@emails.bjut.edu.cn; zhaojunjian@tjpu.edu.cn

Zhitao Zhuang
College of Mathematics and Information Science
North China University of Water Resources and Electric Power
Zhengzhou, 450011, P. R. China
Email address: zhuangzhitao@ncwu.edu.cn