AIMS Mathematics

https://www.aimspress.com/journal/Math

AIMS Mathematics, 9(7): 17076–17088.
DOI: 10.3934/math.2024829
Received: 13 March 2024
Revised: 06 May 2024
Accepted: 08 May 2024
Published: 16 May 2024

***Research article***

# Data-driven wavelet estimations in the convolution structure density model

**Kaikai Cao***

School of Mathematics and Statistics, Weifang University, Weifang 261061, China

* **Correspondence:** Email: caokaikai@wfu.edu.cn.

**Abstract:** Based on a data-driven kernel estimator, Lepski and Willer considered the problem of adaptive $L^p$ risk estimations in the convolution structure density model in 2017 and 2019. This current paper studies the same problem with a data-driven wavelet estimator on Besov spaces, as wavelet estimations offer fast algorithm and provide more local information. Our results can reduce to the traditional adaptive wavelet estimations in the classical density model with no errors, as well as deconvolutional model.

**Keywords:** wavelets; Besov spaces; density estimations; generalized deconvolution model; adaptivity
**Mathematics Subject Classification:** 42C40, 62G07, 62G20

## 1. Introduction and preliminary

The deconvolution density estimation plays important roles in both statistics and econometrics [15]. In this paper, the generalized deconvolution density model introduced by Lepski and Willer [12, 13] is considered, which can be reduced to the classical density model with no errors and the deconvolution one.

Let $Z_1, Z_2, \cdots, Z_n$ be independent and identically distributed (i.i.d.) random variables having the same distribution as

$$Z = X + \varepsilon Y, \tag{1.1}$$

where $X$ stands for a real-valued random variable with an unknown probability density function $f$, $Y$ denotes an independent random noise (error) with a known probability density $g$ and $\varepsilon \in \{0, 1\}$ is a Bernoulli random variable with $P\{\varepsilon = 1\} = \alpha$, $\alpha \in [0, 1]$. The problem is to estimate $f$ by the observed data $Z_1, Z_2, \cdots, Z_n$ in some sense.

When $\alpha = 0$, model (1.1) reduces to the classical density model with no errors. The representative work belongs to Donoho et al. [6]. They established an adaptive and optimal $L^p$ risk estimation (up to a logarithmic factor) on Besov balls by using a nonlinear wavelet estimator. In 2019, Liu and Wu [17] provided a data-driven wavelet estimator and considered point-wise estimations on a local anisotropic Hölder space. Two years later, Cao and Zeng [2] investigated the adaptive $L^p$ risk estimations under the independence hypothesis on Besov balls. More related work can also be found in Refs. [8, 10, 11].

The model (1.1) with $\alpha = 1$ corresponds to the deconvolution model. Fan and Koo [7] considered optimal estimations with $L^2$ risk on a Besov ball. Moreover, Lounici and Nickl [19] discussed $L^\infty$ risk estimations in 2011. Three years later, Li and Liu [14] studied $L^p$ risk estimations based on linear and non-linear wavelet estimators. In 2023, Cao and Zeng [1] provided a data-driven wavelet estimator and considered $L^p$ risk estimations on Besov spaces. For more related literature, please see Refs. [4,16,22].

As in Ref [18], the density function $h$ of $Z$ in (1.1) satisfies

$$h = (1-\alpha)f + \alpha f * g,$$

where $f * g$ denotes the convolution of $f$ and $g$. Moreover, when the function $G_\alpha(t) := 1 - \alpha + \alpha g^{ft}(t)$ has nonzeros on $\mathbb{R}$, we obtain

$$f^{ft}(t) = [(1-\alpha) + \alpha g^{ft}(t)]^{-1} h^{ft}(t) = [G_\alpha(t)]^{-1} h^{ft}(t).$$

Here and throughout, $f^{ft}$ is the Fourier transform of $f \in L^1(\mathbb{R})$ defined by

$$f^{ft}(t) := \int_{\mathbb{R}} f(x) e^{-itx} dx.$$

Based on the model (1.1) with some mild assumptions on $G_\alpha$, Lepski and Willer [12] established an asymptotic lower bound estimation over $L^p$ risk. Moreover, they investigated adaptive $L^p$ risk estimations over an anisotropic Nikol'skii space based on a data-driven kernel estimator in Ref. [13]. Recently, Liu and Wu [18] provided a data-driven wavelet estimator and discussed point-wise estimations under the local Hölder spaces. Cao and Zeng [3] studied $L^p$ risk estimations by using linear and nonlinear wavelet estimators on Besov balls.

In this paper, we will introduce a data-driven wavelet estimator and study the $L^p$ $(1 \le p < \infty)$ risk estimations based on model (1.1) over Besov balls. The following conditions are necessary for our discussion:

(T1) $|G_\alpha(t)| \gtrsim (1+|t|^2)^{-\frac{\beta(\alpha)}{2}}$;

(T2) $\|(G_\alpha)^{(\ell)}\|_\infty \lesssim 1,\ \ell = 1, 2$

with $\beta(\alpha) = \beta \ge 0$ for $\alpha = 1$ and $\beta(\alpha) = 0$ for others. In fact, the condition (T1) is the same as the assumption in Refs [13, 18] and (T2) is used in Lemma 2.1. Here and after, the notation $A \lesssim B$ denotes $A \le cB$ with some fixed and independent constant $c > 0$; $A \gtrsim B$ means $B \lesssim A$; $A \sim B$ stands for both $A \lesssim B$ and $A \gtrsim B$.

It is well known that the wavelet estimation depends on an orthonormal wavelet expansion in $L^2(\mathbb{R})$. Let $\phi$ be an orthonormal scaling function and $\psi$ be the corresponding wavelet one. Subsequently, with $\vartheta_{jk}(x) := 2^{\frac{j}{2}}\vartheta(2^j x - k)$ $(\vartheta = \phi$ or $\psi)$, for $f \in L^2(\mathbb{R})$,

$$f = \sum_{k\in\mathbb{Z}} \alpha_{j_0k}\phi_{j_0k} + \sum_{j\ge j_0}\sum_{k\in\mathbb{Z}} \beta_{jk}\psi_{jk} \tag{1.2}$$

with $\alpha_{jk} := \langle f, \phi_{jk}\rangle$ and $\beta_{jk} := \langle f, \psi_{jk}\rangle$.

As usual, let $P_j$ be the orthogonal projective operator from $L^2(\mathbb{R})$ onto the scaling space $V_j$ with the orthonormal basis $\{\phi_{jk}\}_{k\in\mathbb{Z}}$. Hence, for any $f \in L^2(\mathbb{R})$,

$$P_j f = \sum_{k\in\mathbb{Z}} \alpha_{jk}\phi_{jk}. \tag{1.3}$$

Moreover, the identities (1.2) and (1.3) hold in $L^p(\mathbb{R})$ for $p \geq 1$, if the scaling function $\phi$ is $\ell$-regular. Here and throughout, $\ell$-regular [5] means that $\phi \in C^\ell(\mathbb{R})$ and $|\phi^{(r)}(x)| \leq C_m(1 + |x|^2)^{-m}$ $(r = 0, 1, \cdots, \ell)$ for each $m \in \mathbb{Z}$ and some independent positive constants $C_m$.

One advantage of wavelet bases is that they can characterize Besov spaces, which contain Hölder and $L^2$-Sobolev spaces as special examples.

**Lemma 1.1.** *[21] Let $\phi$ be $\ell$-regular ($\ell > s > 0$) and $\psi$ be the corresponding wavelet. Then, for $r, q \in [1, \infty]$ and $f \in L^r(\mathbb{R})$, the following assertions are equivalent:*
$(i)$ $f \in B^s_{r,q}(\mathbb{R})$;
$(ii)$ $\{2^{js}\|P_j f - f\|_r\} \in l^q(\mathbb{Z})$;
$(iii)$ $\|\alpha_{j_0\cdot}\|_{l^r} + \|(2^{j(s-\frac{1}{r}+\frac{1}{2})}\|\beta_{j\cdot}\|_{l^r})_{j\geq j_0}\|_{l^q} < \infty$.

The Besov norm of $f$ can be defined by

$$\|f\|_{B^s_{r,q}} := \|\alpha_{j_0\cdot}\|_{l^r} + \|(2^{j(s-\frac{1}{r}+\frac{1}{2})}\|\beta_{j\cdot}\|_{l^r})_{j\geq j_0}\|_{l^q}.$$

From Lemma 1.1, we find that $\|P_j f - f\|_r \lesssim 2^{-js}$ holds for $f \in B^s_{r,q}(\mathbb{R})$ and

$$B^s_{r,q}(\mathbb{R}) \hookrightarrow B^{s'}_{p,q}(\mathbb{R})$$

for $r \leq p$ and $s' - \frac{1}{p} = s - \frac{1}{r} > 0$. Here, notation $A \hookrightarrow B$ stands for a Banach space $A$ continuously embedded in another Banach space $B$. For details, please see Refs. [9, 14].

In this paper, the notation $B^s_{r,q}(M, L)$ with $M,\ L > 0$ means that

$$B^s_{r,q}(M, L) := \{f \in B^s_{r,q}(M),\ supp\ f \subseteq [-L, L]\},$$

where $B^s_{r,q}(M) := \{f \in B^s_{r,q}(\mathbb{R}),\ f$ is a density function, and $\|f\|_{B^s_{r,q}} \leq M\}$ stands for a Besov ball.

### *1.1. Data-driven wavelet estimator and main results*

To introduce our estimator, we assume that $\phi$ satisfies $\ell$-regular with $\ell > 3\beta(\alpha) + 1$. Therefore,

$$\widehat{\alpha}_{jk} := \frac{2^{j/2}}{n}\sum_{i=1}^{n}(K_j\phi)(2^j Z_i - k) \quad \text{and} \quad (K_j\phi)(z) := \frac{1}{2\pi}\int_{\mathbb{R}} e^{itz}\frac{\phi^{ft}(t)}{G_\alpha(-2^j t)}dt \tag{1.4}$$

are well-defined under condition (T1). Clearly, $E\widehat{\alpha}_{jk} = \alpha_{jk}$, for details, please see Ref [18].

As in [18], the linear wavelet estimator for (1.1) is defined by

$$\widehat{f_j}(x) := \sum_{k\in\mathbb{Z}} \widehat{\alpha}_{jk}\phi_{jk}(x). \tag{1.5}$$

Normally, the above estimator is non-adaptive [6, 11, 14], because the parameter $j$ depends on the smoothness index $s$ of the unknown density function $f$. Motivated by the works of Lepski and

Willer [13] and Cao and Zeng [1], we provide a selection rule of $j$ deciding only by the observed data $Z_1, \cdots, Z_n$, which is so called a data-driven version.

Let $\mathcal{H} = \left\{0, 1, \cdots, \lfloor \frac{1}{2\beta(\alpha)+1} \log_2 \frac{n}{\ln n} \rfloor\right\}$ and

$$U_n(j) = \sqrt{\frac{\lambda 2^{j(2\beta(\alpha)+1)} \ln n}{n}} + \frac{\lambda 2^{j(\beta(\alpha)+1)} \ln n}{n} \tag{1.6}$$

with $\lfloor a \rfloor$ standing for the largest integer smaller or equal to $a$ and $\lambda$ being a positive constant (specified later on). Furthermore, for each $x \in \mathbb{R}$,

$$\widehat{R}_j(x) := \sup_{j' \in \mathcal{H}} \left[ |\widehat{f}_{j \wedge j'}(x) - \widehat{f}_{j'}(x)| - U_n(j \wedge j') - U_n(j') \right]_+ . \tag{1.7}$$

Here and after, the notations $a \wedge b$ denote $\min\{a, b\}$ and $a_+$ means $\max\{a, 0\}$.

Then, the selection of $j = j_0$ in (1.5) is given by

$$j_0 = j_0(x) = \operatorname*{arginf}_{j \in \mathcal{H}} \left[ \widehat{R}_j(x) + 2U_n(j) \right]. \tag{1.8}$$

Moreover, the data-driven wavelet estimator is obtained by

$$\widehat{f}_n(x) := \widehat{f}_{j_0}(x) = \sum_{k \in \mathbb{Z}} \widehat{\alpha}_{j_0 k} \phi_{j_0 k}(x). \tag{1.9}$$

To introduce Theorem 1.1, let

$$B_j(x, f) := |E\widehat{f}_j(x) - f(x)| \text{ and } \xi_n(x, j) := \widehat{f}_j(x) - E\widehat{f}_j(x) \tag{1.10}$$

be the bias and the stochastic error of $\widehat{f}_j$, respectively. Furthermore, we define

$$B_j^*(x, f) := \sup_{j' \in \mathcal{H},\ j' \geq j} B_{j'}(x, f) \quad \text{and} \quad v_n(x) := \sup_{j \in \mathcal{H}} \left[ |\xi_n(x, j)| - U_n(j) \right]_+, \tag{1.11}$$

where $U_n(j)$ is given by (1.6). Then the following point-wise oracle inequality holds.

**Theorem 1.1.** *For any $x \in \mathbb{R}$, the estimator $\widehat{f}_n(x)$ in (1.9) satisfies that*

$$|\widehat{f}_n(x) - f(x)| \leq \inf_{j \in \mathcal{H}} \left\{ 5B_j^*(x, f) + 5U_n(j) \right\} + 5v_n(x),$$

*where $B_j^*(x, f)$, $v_n(x)$ are defined in (1.11) and $U_n(j)$ is given by (1.6).*

It is important to point out that the proof of Theorem 1.1 only depends on the selection rule of $j_0$ in (1.8) and does not need any assumptions on the unknown density $f$ (except for the restrictions ensuring the existence of the model and of the risk). For more detail, please see Section 3. Therefore, the point-wise oracle inequality in Theorem 1.1 is especially useful and plays an important role in the following $L^p$ risk estimation.

**Theorem 1.2.** *Let $\phi$ be $\ell$-regular ($\ell > 3\beta(\alpha)+1$) and $\widehat{f_n}$ be given by (1.9). If the conditions (T1) and (T2) hold, $0 < s < \ell$ and $r, q \in [1,\infty]$, then for each $p \in [1,\infty)$, we have*

$$\sup_{f\in B^s_{r,q}(M,L)\cap L^\infty(M)} E\|\widehat{f_n}I_{[-L,L]} - f\|_p^p \lesssim \Big(\frac{\ln n}{n}\Big)^{\theta p},$$

*where*

$$\theta := \begin{cases} \frac{s}{2s+2\beta(\alpha)+1}, & 1 \le p < \frac{2sr}{2\beta(\alpha)+1} + r; \\ \frac{sr}{(2\beta(\alpha)+1)p}, & p \ge \frac{2sr}{2\beta(\alpha)+1} + r,\ s \le \frac{1}{r}; \\ \frac{s-\frac{1}{r}+\frac{1}{p}}{2(s-\frac{1}{r})+2\beta(\alpha)+1}, & p \ge \frac{2sr}{2\beta(\alpha)+1} + r,\ s > \frac{1}{r}. \end{cases}$$

**Remark 1.1.** *When $\alpha = 1$, $\beta(\alpha) = \beta$ and $\theta = \min\{\frac{s}{2s+2\beta+1}, \frac{s-\frac{1}{r}+\frac{1}{p}}{2(s-\frac{1}{r})+2\beta+1}\}$ for $s > \frac{1}{r}$, which coincides with Theorem 4 of Li and Liu [14]. On the other hand, $\beta(\alpha) = 0$ with $\alpha = 0$, then $\theta = \min\{\frac{s}{2s+1}, \frac{s-\frac{1}{r}+\frac{1}{p}}{2(s-\frac{1}{r})+1}\}$ with $s > \frac{1}{r}$, while the conclusion of Theorem 3 of Donoho et al. [6] can follow directly from Theorem 1.2. Moreover, the estimation for $s \le \frac{1}{r}$ is considered, whereas there is none in Refs. [6, 14].*

**Remark 1.2.** *The conclusion of Theorem 1.2 with $\alpha = 1$ can be reduced to Theorem 4.1 of Cao and Zeng [1]. In addition, the condition $\beta \ge 0$ in Theorem 1.2 is weaker than the condition $\beta > 1$ in Ref. [1].*

## 2. Two propositions

This section provides two necessary propositions that play important roles in the proof of Theorem 1.2.

With $K_j\phi$ being given by (1.4), we denote

$$K_j^*(t,x) := 2^j \sum_{k\in\mathbb{Z}} (K_j\phi)(2^j t - k)\phi(2^j x - k). \tag{2.1}$$

Then, the following lemma holds.

**Lemma 2.1.** *Let $f \in L^\infty(M)$, $\phi$ be $\ell$-regular ($\ell > 3\beta(\alpha)+1$), and conditions (T1) and (T2) hold. Then $K_j^*(t,x)$ given by (2.1) satisfies that*

$$|K_j^*(t,x)| \le M_1 2^{j(\beta(\alpha)+1)} \ \ \textit{and}\ \ E|K_j^*(Z_1,x)|^2 \le M_1 2^{j(2\beta(\alpha)+1)},$$

*where $M_1 > 0$ is some constant.*

*Proof.* According to (1.4) and (2.1), there exists some constant $M_0 > 0$ such that

$$|K_j^*(t,x)| = \Big|2^j \sum_{k\in\mathbb{Z}} (K_j\phi)(2^j t - k)\phi(2^j x - k)\Big| \le M_0 2^{j(\beta(\alpha)+1)} \tag{2.2}$$

thanks to conditions (T1) and the regularity of $\phi$.

Clearly, $\|h\|_\infty \le (1-\alpha)\|f\|_\infty + \alpha\|f * g\|_\infty \le \|f\|_\infty$. This with (1.4) and (2.1) has

$$E|K_j^*(Z_1,x)|^2 \;\le\; \|f\|_\infty 2^{2j}\int_{\mathbb{R}}\Big[\sum_k \phi(2^jx-k)\frac{1}{2\pi}\int e^{it(2^jz-k)}\frac{\phi^{ft}(t)}{G_\alpha(-2^jt)}dt\Big]^2 dz.$$

Combining (T1) and (T2) and Parseval identity, one obtains

$$E|K_j^*(Z_1,x)|^2 \;\le\; \|f\|_\infty 2^j\Big(\sum_k |\phi(2^jx-k)|\Big)^2\int_{\mathbb{R}}\Big[\frac{\phi^{ft}(t)}{G_\alpha(-2^jt)}\Big]^2 dt \le M_0' 2^{j(2\beta(\alpha)+1)}, \tag{2.3}$$

where $M_0' > 0$ is some constant. The proof is completed by choosing $M_1 := \max\{M_0, M_0'\}$ and (2.2), (2.3). □

In order to prove Proposition 2.1, the following inequality is necessary:
**Bernstein's inequality.** [20] *Let $X_1, \cdots, X_n$ be i.i.d. random variables with $EX_i^2 \le \sigma^2$ and $|X_i| \le M$ $(i = 1, 2, \cdots, n)$. Then, for any $t > 0$,*

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^n (X_i - EX_i)\right| \ge \sqrt{\frac{2\sigma^2 t}{n}} + \frac{4Mt}{3n}\right\} \le 2e^{-t}.$$

**Proposition 2.1.** *Let $f \in L^\infty(M)$, $\phi$ be $\ell$-regular with $\ell > 3\beta(\alpha) + 1$ and conditions (T1) and (T2) hold. Then, for each $x \in \mathbb{R}$ and $\gamma > 0$, there exists $\lambda > 4M_1\gamma$ such that*

$$E[v_n(x)]^\gamma \lesssim n^{-\frac{\gamma}{2}},$$

*where $v_n(x)$ is given by (1.11) and $M_1$ is the positive constant in Lemma 2.1.*

*Proof.* With $\lambda_j = \max\{(\beta(\alpha) + 2)\gamma j \ln 2, \frac{1}{4}\}$ and $j \in \mathcal{H}$, one defines

$$\overline{U_n}(j) := \sqrt{\frac{2M_1 2^{j(2\beta(\alpha)+1)}}{n}\lambda_j} + \frac{4M_1 2^{j(\beta(\alpha)+1)}}{3n}\lambda_j. \tag{2.4}$$

It is easy to see that $\lambda \ln n \ge 2M_1\lambda_j$ for large $n$ follows from $\lambda > 4M_1\gamma$ and $j \in \mathcal{H}$. Thus, one concludes $\overline{U_n}(j) \le U_n(j)$, thanks to (1.6) and (2.4). Therefore,

$$\Big[|\xi_n(x,j)| - U_n(j)\Big]_+ \le \Big[|\xi_n(x,j)| - \overline{U_n}(j)\Big]_+. \tag{2.5}$$

Note that $\Big\{[|\xi_n(x,j)| - \overline{U_n}(j)]_+ > t\Big\} = \Big\{|\xi_n(x,j)| - \overline{U_n}(j) > t\Big\}$ for each $t \ge 0$. Thus,

$$E\Big[|\xi_n(x,j)| - \overline{U_n}(j)\Big]_+^\gamma = \gamma\int_0^\infty t^{\gamma-1}P\Big\{|\xi_n(x,j)| - \overline{U_n}(j) > t\Big\}dt.$$

This with a change of variables $t = v\omega$ and $\omega := \sqrt{\frac{2M_1 2^{j(2\beta(\alpha)+1)}}{n}} + \frac{4M_1 2^{j(\beta(\alpha)+1)}}{3n}$ obtains

$$\begin{aligned} &E\Big[|\xi_n(x,j)| - \overline{U_n}(j)\Big]_+^\gamma \\ &\le \gamma\int_0^\infty (v\omega)^{\gamma-1}P\left\{|\xi_n(x,j)| > \sqrt{\frac{2M_1 2^{j(2\beta(\alpha)+1)}}{n}}(\sqrt{v+\lambda_j}) + \frac{4M_1 2^{j(\beta(\alpha)+1)}}{3n}(v+\lambda_j)\right\}\omega dv \end{aligned} \tag{2.6}$$

by $\lambda_j = \max\{(\beta(\alpha)+2)\gamma j \ln 2, \frac{1}{4}\} \geq \frac{1}{4}$.

According to (1.5), (1.10), and (2.1), one gets

$$\xi_n(x,j) = \frac{1}{n}\sum_{i=1}^{n}[K_j^*(Z_i,x) - EK_j^*(Z_i,x)].$$

Moreover, Lemma 2.1 says that

$$|K_j^*(Z_i,x)| \leq M_1 2^{j(\beta(\alpha)+1)} \text{ and } E|K_j^*(Z_i,x)|^2 \leq M_1 2^{j(2\beta(\alpha)+1)}.$$

These with Bernstein's inequality,

$$P\left\{|\xi_n(x,j)| > \sqrt{\frac{2M_1 2^{j(2\beta(\alpha)+1)}}{n}}(\sqrt{v+\lambda_j}) + \frac{4M_1 2^{j(\beta(\alpha)+1)}}{3n}(v+\lambda_j)\right\} \leq 2e^{-(v+\lambda_j)}.$$

Substituting this above estimate into (2.6), one knows

$$E\Big[|\xi_n(x,j)| - \overline{U_n}(j)\Big]_+^\gamma \leq 2\gamma\omega^\gamma \int_0^\infty v^{\gamma-1}e^{-(v+\lambda_j)}dv \lesssim \omega^\gamma e^{-\lambda_j}.$$

Therefore,

$$E\Big[|\xi_n(x,j)| - \overline{U_n}(j)\Big]_+^\gamma \lesssim \Big(\frac{2^{j(\beta(\alpha)+1)}}{\sqrt{n}}\Big)^\gamma 2^{-(\beta(\alpha)+2)\gamma j} \lesssim n^{-\frac{\gamma}{2}}2^{-\gamma j} \tag{2.7}$$

due to $\omega = \sqrt{\frac{2M_1 2^{j(2\beta(\alpha)+1)}}{n}} + \frac{4M_1 2^{j(\beta(\alpha)+1)}}{3n}$ and $e^{-\lambda_j} \leq 2^{-(\beta(\alpha)+2)\gamma j}$.

Combining (1.11), (2.5), and (2.7), one has

$$E[v_n(x)]^\gamma \lesssim E\sup_{j\in\mathcal{H}}\Big[|\xi_n(x,j)| - \overline{U_n}(j)\Big]_+^\gamma \lesssim \sum_{j\in\mathcal{H}} E\Big[|\xi_n(x,j)| - \overline{U_n}(j)\Big]_+^\gamma \lesssim n^{-\frac{\gamma}{2}},$$

because $\mathcal{H}$ is a discrete set. The proof is done. □

In order to introduce Proposition 2.2, we need the following notations:

$$U_f(x) := \inf_{j\in\mathcal{H}}\{B_j^*(x,f) + U_n(j)\}, \tag{2.8}$$

$$\Omega_m := \{x \in [-L,L],\ 2^m\delta_n < U_f(x) \leq 2^{m+1}\delta_n\}, \tag{2.9}$$

where $\delta_n = (\frac{C\ln n}{n})^{\frac{s}{2s+2\beta(\alpha)+1}}$ and $C > 1$ is some constant.

Clearly, $U_f(x) \leq \sup_x U_f(x) := c_0$ for $f \in L^\infty(M)$. Therefore, there exists

$$m_2 := \min\{m \in \mathbb{Z},\ 2^m\delta_n \geq c_0\} \tag{2.10}$$

satisfying $\Omega_m = \emptyset$ for any $m > m_2$. Moreover, it is easy to verify that $m_2 > 0$ for a large $n$.

**Proposition 2.2.** *Let $f \in B^s_{r,q}(M) \cap L^\infty(M)$, $\phi$ be $\ell$-regular ($\ell > 3\beta(\alpha)+1$) and conditions (T1) and (T2) hold. Then, for $m \in \mathbb{Z}$ satisfying $0 \le m \le m_2$,*

$$\int_{\Omega_m} [U_f(x)]^p dx \lesssim 2^{m(p-r-\frac{2sr}{2\beta(\alpha)+1})} \delta_n^p;$$

*Furthermore, if $s > \frac{1}{r}$ and $r \le p$, then with $s' := s - \frac{1}{r} + \frac{1}{p}$,*

$$\int_{\Omega_m} [U_f(x)]^p dx \lesssim 2^{-\frac{2ms'p}{2\beta(\alpha)+1}} \delta_n^{\frac{s'}{s}p},$$

*where $U_f(x)$ and $\Omega_m$ are defined in (2.8) and (2.9), respectively.*

*Proof.* Let $j_1$ satisfy

$$c_1 2^{\frac{2m}{2\beta(\alpha)+1}} \delta_n^{-\frac{1}{s}} \le 2^{j_1} \le c_2 2^{\frac{2m}{2\beta(\alpha)+1}} \delta_n^{-\frac{1}{s}} \tag{2.11}$$

with two positive constants $c_1, c_2$ satisfying $(2M)^{\frac{1}{s}} I_{\{r=\infty\}} < c_1 < c_2 < \min\left\{\frac{C}{4c_0^2}, \frac{C}{4(\sqrt{\lambda}+\lambda)^2}\right\}^{\frac{1}{2\beta(\alpha)+1}}$. Then $j_1 \in \mathcal{H}$ and $U_n(j_1) \le 2^{m-1}\delta_n$ for $0 < m \le m_2$ and large $n$. In fact, (2.10) leads to $2^{m_2} \le 2c_0\delta_n^{-1}$. According to $0 < m \le m_2$, (2.11) and $\delta_n = (\frac{C\ln n}{n})^{\frac{s}{2s+2\beta(\alpha)+1}}$, one concludes that

$$1 < c_1\delta_n^{-\frac{1}{s}} \le 2^{j_1} \le c_2 2^{\frac{2m_2}{2\beta(\alpha)+1}} \delta_n^{-\frac{1}{s}} \le c_2(2c_0)^{\frac{2}{2\beta(\alpha)+1}} \delta_n^{-(\frac{1}{s}+\frac{2}{2\beta(\alpha)+1})} < \Big(\frac{n}{\ln n}\Big)^{\frac{1}{2\beta(\alpha)+1}}.$$

Hence, $j_1 \in \mathcal{H}$. On the other hand,

$$\begin{aligned} U_n(j_1) &= \sqrt{\frac{\lambda 2^{j_1(2\beta(\alpha)+1)}\ln n}{n}} + \frac{\lambda 2^{j_1(\beta(\alpha)+1)}\ln n}{n} \\ &\le (\sqrt{\lambda}+\lambda)\sqrt{\frac{2^{j_1(2\beta(\alpha)+1)}\ln n}{n}} \\ &\le (\sqrt{\lambda}+\lambda)\sqrt{c_2^{2\beta(\alpha)+1} 2^{2m} \delta_n^{-\frac{2\beta(\alpha)+1}{s}} \frac{\ln n}{n}} \\ &\le (\sqrt{\lambda}+\lambda)\sqrt{c_2^{2\beta(\alpha)+1}/C}\, 2^m \delta_n. \end{aligned}$$

Thus, $U_n(j_1) \le 2^{m-1}\delta_n$ follows from $c_2 < [\frac{C}{4(\sqrt{\lambda}+\lambda)^2}]^{\frac{1}{2\beta(\alpha)+1}}$.

According to $\Omega_m = \{x \in [-L, L],\ 2^m\delta_n < U_f(x) \le 2^{m+1}\delta_n\}$, one has

$$\int_{\Omega_m} [U_f(x)]^p dx \le (2^{m+1}\delta_n)^p |\Omega_m|, \tag{2.12}$$

where $|\Omega_m|$ stands for the Lebesgue measure of the set $\Omega_m$. Recall that $U_f(x) = \inf_{j\in\mathcal{H}}\{B_j^*(x, f) + U_n(j)\}$ in (2.8). Therefore,

$$\begin{aligned} |\Omega_m| &\le |\{x \in [-L, L],\ U_f(x) > 2^m\delta_n\}| \\ &\le |\{x \in [-L, L],\ B_{j_1}^*(x, f) + U_n(j_1) > 2^m\delta_n\}| \end{aligned}$$

$$\leq \quad |\{x \in [-L, L],\ B^*_{j_1}(x, f) > 2^{m-1}\delta_n\}| \tag{2.13}$$

thanks to $j_1 \in \mathcal{H}$ and $U_n(j_1) \leq 2^{m-1}\delta_n$.

When $1 \leq r < \infty$, by using Chebyshev's inequality, (1.11), (2.13), and $f \in B^s_{r,q}(M)$,

$$\begin{aligned} |\Omega_m| &\leq |\{x \in [-L, L],\ B^*_{j_1}(x, f) > 2^{m-1}\delta_n\}| \\ &\leq \sum_{j \in \mathcal{H}, j \geq j_1} |\{x \in [-L, L],\ B_j(x, f) > 2^{m-1}\delta_n\}| \\ &\leq \sum_{j \in \mathcal{H}, j \geq j_1} \frac{\|B_j(\cdot, f)\|_r^r}{(2^{m-1}\delta_n)^r} \lesssim 2^{-mr}\delta_n^{-r}2^{-j_1 sr}. \end{aligned} \tag{2.14}$$

This with (2.12) and $2^{j_1} \sim 2^{\frac{2m}{2\beta(\alpha)+1}}\delta_n^{-\frac{1}{s}}$ implies that

$$\int_{\Omega_m} [U_f(x)]^p dx \lesssim (2^{m+1}\delta_n)^p 2^{-mr}\delta_n^{-r}2^{-j_1 sr} \lesssim 2^{m(p-r)}\delta_n^{p-r}2^{-j_1 sr} \lesssim 2^{m(p-r-\frac{2sr}{2\beta(\alpha)+1})}\delta_n^p. \tag{2.15}$$

When $r = \infty$, according to $f \in B^s_{r,q}(M)$ and $m > 0$, one gets

$$B^*_{j_1}(x, f) = \sup_{j' \geq j_1} B_{j'}(x, f) \leq M2^{-j_1 s} \leq Mc_1^{-s}2^{-\frac{2ms}{2\beta(\alpha)+1}}\delta_n \leq 2^{m-1}\delta_n$$

by the choice of $2^{j_1} \geq c_1 2^{\frac{2m}{2\beta(\alpha)+1}}\delta_n^{-\frac{1}{s}}$ with $c_1 > (2M)^{\frac{1}{s}}$. Hence, $|\Omega_m| = 0$ by (2.13). This with (2.12) obtains

$$\int_{\Omega_m} [U_f(x)]^p dx \leq (2^{m+1}\delta_n)^p|\Omega_m| = 0. \tag{2.16}$$

Finally, the situation of $s > \frac{1}{r}$ and $r \leq p$ needs to be considered. Note that $B^s_{r,q} \hookrightarrow B^{s'}_{p,q}$ with $s' = s - \frac{1}{r} + \frac{1}{p}$. With the same arguments as (2.14), one has

$$|\Omega_m| \leq \sum_{j \in \mathcal{H}, j \geq j_1} |\{x \in [-L, L],\ B_j(x, f) > 2^{m-1}\delta_n\}| \leq \sum_{j \in \mathcal{H}, j \geq j_1} \frac{\|B_j(\cdot, f)\|_p^p}{(2^{m-1}\delta_n)^p} \lesssim 2^{-mp}\delta_n^{-p}2^{-j_1 s' p}.$$

Then it follows from (2.12) and $2^{j_1} \sim 2^{\frac{2m}{2\beta(\alpha)+1}}\delta_n^{-\frac{1}{s}}$ that

$$\int_{\Omega_m} [U_f(x)]^p dx \lesssim (2^{m+1}\delta_n)^p 2^{-mp}\delta_n^{-p}2^{-j_1 s' p} \lesssim 2^{-j_1 s' p} \lesssim 2^{-\frac{2ms'p}{2\beta(\alpha)+1}}\delta_n^{\frac{s'}{s}p}.$$

This with (2.15) and (2.16) leads to the desired conclusions. □

## 3. Proofs

In this section, we shall provide the proofs of Theorems 1.1 and 1.2.

*Proof of Theorem 1.1.* Note that $U_n(j)$ in (1.6) is monotonically increasing with $j$. Together with (1.7), one shows

$$|\widehat{f}_{j\wedge j_0}(x) - \widehat{f}_{j_0}(x)| \leq \widehat{R}_j(x) + U_n(j \wedge j_0) + U_n(j_0) \leq \widehat{R}_j(x) + 2U_n(j_0).$$

Similarly,

$$|\widehat{f}_{j_0\wedge j}(x)-\widehat{f}_j(x)|\le \widehat{R}_{j_0}(x)+2U_n(j).$$

These with the selection of $j_0$ in (1.8) and $\widehat{f}_{j_0\wedge j}=\widehat{f}_{j\wedge j_0}$ imply that

$$|\widehat{f}_{j_0\wedge j}(x)-\widehat{f}_{j_0}(x)|+|\widehat{f}_{j_0\wedge j}(x)-\widehat{f}_j(x)|\le 2\widehat{R}_j(x)+4U_n(j). \tag{3.1}$$

Moreover, by (1.7) and (1.10),

$$\begin{aligned}\widehat{R}_j(x) &= \sup_{j'\in\mathcal{H}}\Big[|\widehat{f}_{j\wedge j'}(x)-\widehat{f}_{j'}(x)|-U_n(j\wedge j')-U_n(j')\Big]_+\\ &\le \sup_{j'\in\mathcal{H}}\Big[|E\widehat{f}_{j\wedge j'}(x)-E\widehat{f}_{j'}(x)|+|\xi_n(x,j\wedge j')|-U_n(j\wedge j')+|\xi_n(x,j')|-U_n(j')\Big]_+.\end{aligned}$$

This with $\sup_{j'\in\mathcal{H}}|E\widehat{f}_{j\wedge j'}(x)-E\widehat{f}_{j'}(x)|\le\sup_{\{j'\in\mathcal{H},\ j'\ge j\}}\{B_{j\wedge j'}(x,f)+B_{j'}(x,f)\}$ and (1.11) implies that

$$\widehat{R}_j(x)\le 2B^*_j(x,f)+2v_n(x).$$

Furthermore,

$$|\widehat{f}_{j_0\wedge j}(x)-\widehat{f}_{j_0}(x)|+|\widehat{f}_{j_0\wedge j}(x)-\widehat{f}_j(x)|\le 4B^*_j(x,f)+4v_n(x)+4U_n(j) \tag{3.2}$$

by (3.1).

On the other hand, $|\xi_n(x,j)|\le\Big[|\xi_n(x,j)|-U_n(j)\Big]_+ +U_n(j)\le v_n(x)+U_n(j)$ by (1.11). This with (1.10) and (1.11) leads to

$$|\widehat{f}_j(x)-f(x)|\le B_j(x,f)+|\xi_n(x,j)|\le B^*_j(x,f)+v_n(x)+U_n(j). \tag{3.3}$$

Clearly, $|\widehat{f}_{j_0}(x)-f(x)|\le|\widehat{f}_{j_0\wedge j}(x)-\widehat{f}_{j_0}(x)|+|\widehat{f}_{j_0\wedge j}(x)-\widehat{f}_j(x)|+|\widehat{f}_j(x)-f(x)|$. It follows from (3.2) and (3.3) that

$$|\widehat{f}_{j_0}(x)-f(x)|\le 5B^*_j(x,f)+5v_n(x)+5U_n(j)$$

holds for any $j\in\mathcal{H}$. Moreover,

$$|\widehat{f}_{j_0}(x)-f(x)|\le\inf_{j\in\mathcal{H}}\Big\{5B^*_j(x,f)+5U_n(j)\Big\}+5v_n(x).$$

Finally, the desired conclusion is completed by $\widehat{f}_n(x)=\widehat{f}_{j_0}(x)$ in (1.9).

*Proof of Theorem 1.2.* According to Theorem 1.1, for any $x\in[-L,L]$,

$$|\widehat{f}_n(x)-f(x)|\lesssim U_f(x)+v_n(x).$$

Let $\Omega_0^-:=\{x\in[-L,L],\ U_f(x)\le\delta_n\}$ with $\delta_n=(\frac{C\ln n}{n})^{\frac{s}{2s+2\beta(\alpha)+1}}$. Then, for each $p\in[1,\infty)$,

$$E\|\widehat{f}_nI_{[-L,L]}-f\|_p^p \lesssim E\int_{-L}^{L}[U_f(x)+v_n(x)]^pdx\lesssim\int_{-L}^{L}[U_f(x)]^pdx+\int_{-L}^{L}E[v_n(x)]^pdx$$

$$
\begin{aligned}
&\lesssim \int_{\Omega_0^-}[U_f(x)]^p dx + \sum_{m=0}^{m_2}\int_{\Omega_m}[U_f(x)]^p dx + n^{-\frac{p}{2}} \\
&\lesssim \sum_{m=0}^{m_2}\int_{\Omega_m}[U_f(x)]^p dx + \delta_n^p
\end{aligned}
\tag{3.4}
$$

thanks to $f \in B^s_{r,q}(M, L)$ and Proposition 2.1.

Recall that $2^{m_2} \sim \delta_n^{-1}$ and $\delta_n \sim (\frac{\ln n}{n})^{\frac{s}{2s+2\beta(\alpha)+1}}$ by (2.9) and (2.10). By using Proposition 2.2, one obtains the following estimations:

(i) For $1 \le p < \frac{2sr}{2\beta(\alpha)+1} + r$,

$$
\sum_{m=0}^{m_2}\int_{\Omega_m}[U_f(x)]^p dx + \delta_n^p \lesssim \delta_n^p \lesssim \Big(\frac{\ln n}{n}\Big)^{\frac{sp}{2s+2\beta(\alpha)+1}}. \tag{3.5}
$$

(ii) For $p \ge \frac{2sr}{2\beta(\alpha)+1} + r$,

$$
\sum_{m=0}^{m_2}\int_{\Omega_m}[U_f(x)]^p dx + \delta_n^p \lesssim 2^{m_2(p-r-\frac{2sr}{2\beta(\alpha)+1})}\delta_n^p + \delta_n^p \lesssim \Big(\frac{\ln n}{n}\Big)^{\frac{sr}{2\beta(\alpha)+1}}. \tag{3.6}
$$

(iii) For the cases $p \ge \frac{2sr}{2\beta(\alpha)+1} + r$ and $s > \frac{1}{r}$. Let $m_1 \in \mathbb{Z}$ satisfy

$$
2^{m_1} \sim \delta_n^{\frac{s' p(\frac{1}{s}-\frac{1}{s'})}{(\frac{2s'}{2\beta(\alpha)+1}+1)p-\frac{2sr}{2\beta(\alpha)+1}-r}}. \tag{3.7}
$$

Hence, $0 < m_1 < m_2$ follows from $r < p,\ p \ge \frac{2sr}{2\beta(\alpha)+1} + r$ and $s > \frac{1}{r}$. Thus,

$$
\begin{aligned}
\sum_{m=0}^{m_2}\int_{\Omega_m}[U_f(x)]^p dx + \delta_n^p &\le \sum_{m=0}^{m_1}\int_{\Omega_m}[U_f(x)]^p dx + \sum_{m=m_1}^{m_2}\int_{\Omega_m}[U_f(x)]^p dx + \delta_n^p \\
&\lesssim 2^{m_1(p-r-\frac{2sr}{2\beta(\alpha)+1})}\delta_n^p + 2^{-\frac{2m_1 s' p}{2\beta(\alpha)+1}}\delta_n^{\frac{s'}{s}p} + \delta_n^p.
\end{aligned}
$$

This with (3.7), $\delta_n \sim (\frac{\ln n}{n})^{\frac{s}{2s+2\beta(\alpha)+1}}$ and $s' = s - \frac{1}{r} + \frac{1}{p}$ leads to

$$
\sum_{m=0}^{m_2}\int_{\Omega_m}[U_f(x)]^p dx + \delta_n^p \lesssim \Big(\frac{\ln n}{n}\Big)^{\frac{s' p}{2(s-\frac{1}{r})+2\beta(\alpha)+1}}. \tag{3.8}
$$

The proof of Theorem 1.2 is completed due to (3.4)–(3.6) and (3.8).

## 4. Conclusions

Based on a data-driven wavelet estimator, we study the adaptive $L^p$ risk estimations in the convolution structure density model (see (1.1)). When $\alpha = 0$, the model (1.1) reduces to the classical density model with no errors, and our results coincide with the conclusions of Theorem 3 of Donoho et al. [6]. On the other hand, the model (1.1) with $\alpha = 1$ corresponds to the deconvolution model, and Theorem 4 of Li and Liu [14] and Theorem 4.1 of Cao and Zeng [1] can follow directly from our Theorem 1.2.

## Use of AI tools declaration

The author declares that he has not used Artificial Intelligence (AI) tools in the creation of this article.

## Acknowledgments

The work is supported by the National Natural Science Foundation of China (No. 12101459).

## Conflict of interest

The author declares no conflict of interest.

## References

1. K. K. Cao, X. C. Zeng, A data-driven wavelet estimator for deconvolution density estimations, *Results Math.*, **78** (2023), 156. https://doi.org/10.1007/s00025-023-01928-0
2. K. K. Cao, X. C. Zeng, Adaptive wavelet density estimation under independence hypothesis, *Results Math.*, **76** (2021), 196. https://doi.org/10.1007/s00025-021-01506-2
3. K. K. Cao, X. C. Zeng, Adaptive wavelet estimations in the convoultion structure density model, *Mathematics*, **8** (2020), 139. https://doi.org/10.3390/math8091391
4. F. Comte, C. Lacour, Anisotropic adaptive kernel deconvolution, *Ann. I. H. Poincaré Probab. Stat.*, **49** (2013), 569–609. https://doi.org/10.1214/11-AIHP470
5. I. Daubechies, *Ten lectures on wavelets*, SIAM, Philadelphia, 1992.
6. D. L. Donoho, I. M. Johnstone, G. Kerkyacharian, D. Picard, Density estimation by wavelet thresholding, *Ann. Stat.*, **24** (1996), 508–539. https://doi.org/10.1214/aos/1032894451
7. J. Fan, J. Y. Koo, Wavelet deconvolution, *IEEE T. Inform. Theory*, **48** (2002), 734–747. https://doi.org/10.1109/18.986021
8. A. Goldenshluger, O. Lepski, On adaptive minimax density estimation on $\mathbb{R}^d$, *Probab. Theory Rel.*, **159** (2014), 479–543. https://doi.org/10.1007/s00440-013-0512-1
9. W. Härdle, G. Kerkyacharian, D. Picard, A. Tsybakov, *Wavelets, approximation and statistical applications*, Springer: New York, NY, USA, 1998.
10. A. Juditsky, S. Lambert-Lacroix, On minimax density estimation on $\mathbb{R}$, *Bernoulli*, **10** (2004), 187–220. https://doi.org/10.3150/bj/1082380217
11. G. Kerkyacharian, D. Picard, Density estimation in Besov spaces, *Stat. Probab. Lett.*, **13** (1992), 15–24. https://doi.org/10.1016/0167-7152(92)90231-S
12. O. Lepski, T. Willer, Lower bounds in the convolution structure density model, *Bernoulli*, **23** (2017), 884–926. https://doi.org/10.3150/15-BEJ763
13. O. Lepski, T. Willer, Oracle inequalities and adaptive estimation in the convolution structure density model, *Ann. Stat.*, **47** (2019), 233–287. https://doi.org/10.1214/18-AOS1687

14. R. Li, Y. M. Liu, Wavelet optimal estimations for a density with some additive noises, *Appl. Comput. Harmon. Anal.*, **36** (2014), 416–433. https://doi.org/10.1016/j.acha.2013.07.002

15. Q. Li, J. S. Racine, *Nonparametric econometrics: Theory and practice*, Princeton University Press, Princeton, 2007.

16. Y. M. Liu, X. C. Zeng, Asymptotic normality for wavelet deconvolution density estimators, *Appl. Comput. Harmon. Anal.*, **48** (2020), 321–342. https://doi.org/10.1016/j.acha.2018.05.006

17. Y. M. Liu, C. Wu, Point-wise estimation for anistropic densities, *J. Multivariate Anal.*, **171** (2019), 112–125. https://doi.org/10.1016/j.jmva.2018.11.014

18. Y. M. Liu, C. Wu, Point-wise wavelet estimation in the convolution structure density model, *J. Fourier Anal. Appl.*, **26** (2020), 81. https://doi.org/10.1007/s00041-020-09794-y

19. K. Lounici, R. Nickl, Global uniform risk bounds for wavelet deconvolution estimators, *Ann. Stat.*, **39** (2011), 201–231. https://doi.org/10.1214/10-AOS836

20. P. Massart, *Concentration inequalities and model selection*, In: Lectures from the 33rd Summer School on Probability Theory held in Saint-Flour, Springer, Berlin, 2007.

21. Y. Meyer, *Wavelets and operators*, Cambridge University Press, Cambridge, 1992.

22. M. Pensky, B. Vidakovic, Adaptive wavelet estimator for nonparametric density deconvolution, *Ann. Stat.*, **27** (1999), 2033–2053. https://doi.org/10.1214/aos/1017939249

AIMS Press