Guo and Liu *Journal of Inequalities and Applications* 2014, 2014:236
http://www.journalofinequalitiesandapplications.com/content/2014/1/236

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH

Open Access

Wavelet estimations for densities and their derivatives with Fourier oscillating noises

Huijun Guo and Youming Liu*

*Correspondence:
liuym@bjut.edu.cn
Department of Applied Mathematics, Beijing University of Technology, Pingle Yuan 100, Beijing, 100124, P.R. China

Abstract

By developing the classical kernel method, Delaigle and Meister provide a nice estimation for a density function with some Fourier-oscillating noises over a Sobolev ball $W_2^s(L)$ and over L^2 risk (Delaigle and Meister in Stat. Sin. 21:1065-1092, 2011). The current paper extends their theorem to Besov ball $B_{r,q}^s(L)$ and L^p risk with $p, q, r \in [1, \infty]$ by using wavelet methods. We firstly show a linear wavelet estimation for densities in $B_{r,q}^s(L)$ over L^p risk, motivated by the work of Delaigle and Meister. Our result reduces to their theorem, when $p = q = r = 2$. Because the linear wavelet estimator is not adaptive, a nonlinear wavelet estimator is then provided. It turns out that the convergence rate is better than the linear one for $r \leq p$. In addition, our conclusions contain estimations for density derivatives as well.

Keywords: Besov space; density derivative; density function; Fourier-oscillating noises; wavelet estimation

1 Introduction and preliminary

One of the fundamental deconvolution problems is to estimate a density function f_X of a random variable X, when the available data $W_1, W_2, \ldots, W_n$ are independent and identically distributed (i.i.d.) with

$$W_j = X_j + \delta_j \quad (j = 1, 2, \ldots, n).$$

We assume that all X_j and δ_j are independent and the density function f_δ of the noise δ is known.

Let the Fourier transform f^{ft} of $f \in L(\mathbb{R})$ be defined by $f^{ft}(t) = \int_{\mathbb{R}} f(x)e^{itx}\,dx$ in this paper. When f_δ^{ft} satisfies

$$\left|f_\delta^{ft}(t)\right| \geq c\left(1 + |t|\right)^{-\alpha} \tag{1}$$

with $c > 0$ and $\alpha > 0$, there exist lots of optimal estimations for f_X [1–6]. However, many noise densities f_δ have zeros in the Fourier transform domain, *i.e.*, the inequality (1) does not hold. For example, Sun *et al.* described an experiment where data on the velocity of halo stars in the Milky Way are collected, and where the measurement errors are assumed to be uniformly distributed [7]. The classical kernel method provides a slower convergence rate in that case [8–10]. Delaigle and Meister [11] developed a new method for a density f_δ

with

$$\left|f_\delta^{ft}(t)\right| \geq c\left|\sin\left(\frac{\pi t}{\lambda}\right)\right|^\nu \left(1+|t|\right)^{-\alpha}. \tag{2}$$

Here $c, \lambda, \alpha > 0$ and $\nu \in \mathbb{N}$ (non-negative integer set). Such noises are called Fourier-oscillating. Clearly, (2) allows f_δ^{ft} having zeros for $\nu \geq 1$. When $\nu = 0$, (2) reduces to (1) (non-zero case).

Delaigle and Meister defined a kernel estimator $\hat{f}_n$ for a density f_X in a Sobolev space and prove that with EX denoting the expectation of a random variable X,

$$\sup_{f_X \in W_2^s(L)} E\int_a^b \left|\hat{f}_n(x) - f_X(x)\right|^2 dx = O\left(n^{-\frac{2s}{2s+2\alpha+1}}\right) \tag{3}$$

under the assumption (2) (Theorem 4.1 in [11]). Here, $W_2^s(L)$ stands for the Sobolev ball with the radius L. This above convergence rate attains the same one as in the non-zero case [1–4, 6]. In particular, it does not depend on the parameter ν.

It seems that many papers deal with L^2 estimations. However, L^p estimations ($1 \leq p \leq +\infty$) are important [5, 12]. On the other hand, Besov spaces contain many classical spaces (*e.g.*, L^2 Sobolev spaces and Hölder spaces) as special examples. The current paper extends (3) from $W_2^s(L)$ to the Besov ball $B_{r,q}^s(L)$, and from L^2 to L^p risk estimations. In addition, our results contain estimations for dth derivatives $f_X^{(d)}$ of f_X. The next section provides a linear wavelet estimation for $f_X^{(d)}$ over a Besov ball $B_{r,q}^s(L)$ and over L^p risk ($p, q, r \geq 1$). It turns out that our estimation reduces to (3), when $d = 0$, $p = r = q = 2$. Moreover, we show a nonlinear wavelet estimation which improves the linear one for $r \leq p$ in the last part.

1.1 Wavelet basis

The fundamental method to construct a wavelet basis comes from the concept of multiresolution analysis (MRA [13]). It is defined as a sequence of closed subspaces $\{V_j\}$ of the square integrable function space $L^2(\mathbb{R})$ satisfying the following properties:

(i) $V_j \subset V_{j+1}$, $j \in \mathbb{Z}$ (the integer set);
(ii) $f(x) \in V_j$ if and only if $f(2x) \in V_{j+1}$ for each $j \in \mathbb{Z}$;
(iii) $\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L^2(\mathbb{R})$ (the space $\bigcup_{j\in\mathbb{Z}} V_j$ is dense in $L^2(\mathbb{R})$);
(iv) There exists $\varphi(x) \in L^2(\mathbb{R})$ (scaling function) such that $\{\varphi(x-k)\}_{k\in\mathbb{Z}}$ forms an orthonormal basis of $V_0 = \overline{\operatorname{span}}\{\varphi(x-k)\}_{k\in\mathbb{Z}}$.

With the standard notation $h_{j,k}(x) := 2^{\frac{j}{2}} h(2^j x - k)$ in wavelet analysis, we can find a corresponding wavelet function

$$\psi(x) = \sum_{k\in\mathbb{Z}} (-1)^{k+1} \overline{h}_{1-k} \varphi_{1,k}(x) \quad \text{with } h_k = \langle \varphi, \varphi_{1,k} \rangle$$

such that, for a fixed $j \in \mathbb{Z}$, $\{\psi_{j,k}\}_{k\in\mathbb{Z}}$ constitutes an orthonormal basis of the orthogonal complement W_j of V_j in V_{j+1} [13]. Then each $f \in L^2(\mathbb{R})$ has an expansion in $L^2(\mathbb{R})$ sense,

$$f = \sum_{k\in\mathbb{Z}} \alpha_{j,k}\varphi_{j,k} + \sum_{l=j}^{\infty}\sum_{k\in\mathbb{Z}} \beta_{l,k}\psi_{l,k},$$

where $\alpha_{j,k} = \langle f, \varphi_{j,k} \rangle$, $\beta_{l,k} = \langle f, \psi_{l,k} \rangle$.

A family of important examples are Daubechies wavelets $D_{2N}(x)$, which are compactly supported in time domain [14]. They can be smooth enough with increasing supports as N gets large, although D_{2N} do not have analytic formulas except for $N = 1$.

As usual, let P_j and Q_j be the orthogonal projections from $L^2(\mathbb{R})$ to V_j and W_j, respectively,

$$P_j f = \sum_{k\in\mathbb{Z}} \alpha_{j,k}\varphi_{j,k}, \qquad Q_j f = \sum_{k\in\mathbb{Z}} \beta_{j,k}\psi_{j,k} = (P_{j+1} - P_j)f.$$

The following simple lemma is fundamental in our discussions. We use $\|f\|_p$ to denote $L^p(\mathbb{R})$ norm for $f \in L^p(\mathbb{R})$, and $\|\lambda\|_p$ do $l_p(\mathbb{Z})$ norm for $\lambda \in l_p(\mathbb{Z})$, where

$$l_p(\mathbb{Z}) := \begin{cases} \{\lambda = \{\lambda_k\}, \sum_{k\in\mathbb{Z}} |\lambda_k|^p < +\infty\}, & 1 \le p < \infty; \\ \{\lambda = \{\lambda_k\}, \sup_{k\in\mathbb{Z}} |\lambda_k| < +\infty\}, & p = \infty. \end{cases}$$

By using Proposition 8.3 in [15], we have the following conclusion.

Lemma 1.1 *Let h be a Daubechies scaling function or the corresponding wavelet. Then there exists $c_2 \ge c_1 > 0$ such that, for $\lambda = \{\lambda_k\} \in l_p(\mathbb{Z})$ and $1 \le p \le \infty$,*

$$c_1 2^{j(\frac{1}{2}-\frac{1}{p})}\|\lambda\|_p \le \Big\|\sum_{k\in\mathbb{Z}} \lambda_k h_{j,k}\Big\|_p \le c_2 2^{j(\frac{1}{2}-\frac{1}{p})}\|\lambda\|_p.$$

1.2 Besov spaces

One of the advantages of wavelet bases is that they can characterize Besov spaces. To introduce those spaces (see [15]), we need the Sobolev spaces with integer order $W_p^n(\mathbb{R}) := \{f \in L^p(\mathbb{R}), f^{(n)} \in L^p(\mathbb{R})\}$ and $\|f\|_{W_p^n} := \|f\|_p + \|f^{(n)}\|_p$. Then $L^p(\mathbb{R})$ can be considered as $W_p^0(\mathbb{R})$.

For $1 \le p, q \le \infty$, $s = n + \alpha$ with $n \in \mathbb{N}$ and $\alpha \in (0, 1]$, the Besov spaces are defined by

$$B_{p,q}^s(\mathbb{R}) := \big\{f \in W_p^n(\mathbb{R}), \big\{2^{j\alpha}\omega_p^2(f^{(n)}, 2^{-j})\big\}_{j\in\mathbb{Z}} \in l_q(\mathbb{Z})\big\},$$

with the associated norm $\|f\|_{B_{p,q}^s} := \|f\|_{W_p^n} + \|\{2^{j\alpha}\omega_p^2(f^{(n)}, 2^{-j})\}_{j\in\mathbb{Z}}\|_{l_q}$, where

$$\omega_p^2(f, t) := \sup_{|h|\le t} \big\|f(\cdot + 2h) - 2f(\cdot + h) + f(\cdot)\big\|_p$$

stands for the smoothness modulus of f. It should be pointed out that $B_{2,2}^s(\mathbb{R}) = W_2^s(\mathbb{R})$.

According to Theorem 9.6 in [15], the following result holds.

Lemma 1.2 *Let $\varphi = D_{2N}$ be a Daubechies scaling function with large N and ψ be the corresponding wavelet. If $f \in L^p(\mathbb{R})$, $1 \le p, q \le \infty$, $s > 0$, $\alpha_{0,k} = \langle f, \varphi_{0,k}\rangle$, and $\beta_{j,k} = \langle f, \psi_{j,k}\rangle$, then the following assertions are equivalent*:

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

(ii) $\|\alpha_{0,\cdot}\|_{l_p} + \|\{2^{j(s+\frac{1}{2}-\frac{1}{p})}\|\beta_{j,\cdot}\|_{l_p}\}_{j\ge0}\|_{l_q} < +\infty$;

(iii) $\{2^{js}\|P_j f - f\|_p\}_{j\ge0} \in l_q$, *where P_j is the projection operator to V_j.*

In each case,

$$\|f\|_{B^s_{p,q}} \sim \|\alpha_{0,\cdot}\|_{l_p} + \left\| \left\{ 2^{j(s+\frac{1}{2}-\frac{1}{p})} \|\beta_{j,\cdot}\|_{l_p} \right\}_{j\geq 0} \right\|_{l_q} \sim \|P_0 f\|_p + \left\| \left\{ 2^{js} \|P_j f - f\|_p \right\}_{j\geq 0} \right\|_{l_q}.$$

Here and throughout, $A \lesssim B$ denotes $A \leq CB$ for some constant $C > 0$; $A \gtrsim B$ means $B \lesssim A$; we use $A \sim B$ standing for both $A \lesssim B$ and $B \lesssim A$.

Note that l_{p_1} is continuously embedded into l_{p_2} for $p_1 \leq p_2$. Then the above lemma implies that

$$B^{s_1}_{p_1,q}(\mathbb{R}) \subset B^{s_2}_{p_2,q}(\mathbb{R}) \quad \text{for } p_1 \leq p_2 \text{ and } s_1 - \frac{1}{p_1} = s_2 - \frac{1}{p_2}.$$

2 Linear wavelet estimation

We shall provide a linear wavelet estimation for a compactly supported density function f_X and its derivatives $f_X^{(d)}$ under Fourier-oscillating noises in this section, motivated by the work of Delaigle and Meister. It turns out that our result generalizes their theorem.

As in [11], we define

$$p(x) = \sum_{m=0}^{\nu} \binom{\nu}{m} (-1)^{\nu-m} f_X\left(x - \frac{2\pi m}{\lambda}\right). \tag{4}$$

Then $p \in L(\mathbb{R})$ and $p^{ft}(t) = (e^{i\frac{2\pi t}{\lambda}} - 1)^{\nu} f_X^{ft}(t)$. Delaigle and Meister found that

$$f_X(x) = \sum_{m=0}^{J} \eta_m p\left(x - \frac{2\pi m}{\lambda}\right), \tag{5}$$

where J and η_m depend only on ν and the support length of f_X.

Let $\varphi = D_{2N}$ be the Daubechies scaling function with N large enough. Since both f_X and φ have compact supports, the set $K_j := \{k \in \mathbb{Z} : \langle f_X, \varphi_{j,k}\rangle \neq 0\}$ is finite and the cardinality $|K_j| \sim 2^j$. Then with $\alpha_{j,k} = \langle f_X^{(d)}, \varphi_{j,k}\rangle$,

$$P_j f_X^{(d)} = \sum_{k\in K_j} \alpha_{j,k}\varphi_{j,k}.$$

It is easy to see $\alpha_{j,k} = (-1)^d \langle f_X, (\varphi_{j,k})^{(d)}\rangle$. This with (5) and the Plancherel formula leads to

$$\alpha_{j,k} = \frac{(-1)^d}{2\pi} \left\langle \sum_{m=0}^{J} \eta_m e^{i\frac{2\pi mt}{\lambda}} p^{ft}(t), \left[(\varphi_{j,k})^{(d)}\right]^{ft}(t) \right\rangle. \tag{6}$$

Note that $p^{ft}(t) = (e^{\frac{2\pi it}{\lambda}} - 1)^{\nu} \frac{f_W^{ft}(t)}{f_\delta^{ft}(t)}$ and $[(\varphi_{j,k})^{(d)}]^{ft}(t) = 2^{-\frac{j}{2}+dj} e^{ik2^{-j}t} [\varphi^{(d)}]^{ft}(2^{-j}t)$. Then the identity (6) reduces to

$$\alpha_{j,k} = \frac{(-1)^d}{2\pi} \int 2^{-\frac{j}{2}+dj} \xi(t) \frac{\overline{[\varphi^{(d)}]^{ft}(2^{-j}t)}}{f_\delta^{ft}(t)} e^{-ik2^{-j}t} f_W^{ft}(t)\, dt \tag{7}$$

where $\xi(t) = \sum_{m=0}^{J} \eta_m e^{i\frac{2\pi mt}{\lambda}}(e^{i\frac{2\pi t}{\lambda}} - 1)^\nu$. Since the empirical estimator for f_W^{ft} is $\frac{1}{n}\sum_{l=1}^{n} e^{iW_l t}$, it is natural to define a linear wavelet estimator

$$\hat{f}_{n,d}^{lin}(x) = \sum_{k\in K_j} \hat{\alpha}_{j,k}\varphi_{j,k}(x), \tag{8}$$

with

$$\hat{\alpha}_{j,k} = \frac{1}{n}\sum_{l=1}^{n} \frac{(-1)^d}{2\pi} \int 2^{-\frac{j}{2}+dj}\xi(t)\frac{\overline{[\varphi^{(d)}]^{ft}(2^{-j}t)}}{f_\delta^{ft}(t)} e^{-ik2^{-j}t} e^{iW_l t}\,dt. \tag{9}$$

When $\nu = d = 0$, $p(x) = f_X(x)$ and $\eta_m = \delta_{m,0}$. Then our estimator $\hat{f}_{n,d}^{lin}$ reduces to the classical linear estimator for the case f_δ^{ft} having no zeros (see *e.g.*, [2–5]).

Let ψ be the Daubechies wavelet function corresponding to the scaling function D_{2N} and $\beta_{j,k} = \langle f^{(d)}, \psi_{j,k}\rangle$. Similar to (9), we define

$$\hat{\beta}_{j,k} = \frac{1}{n}\sum_{l=1}^{n} \frac{(-1)^d}{2\pi} \int 2^{-\frac{j}{2}+dj}\xi(t)\frac{\overline{[\psi^{(d)}]^{ft}(2^{-j}t)}}{f_\delta^{ft}(t)} e^{-ik2^{-j}t} e^{iW_l t}\,dt. \tag{10}$$

Then it can easily be seen that $E\hat{\alpha}_{j,k} = \alpha_{j,k}$, $E\hat{\beta}_{j,k} = \beta_{j,k}$ and $E\hat{f}_{n,d}^{lin} = P_j f_X^{(d)}$.

For $a < b$, $L > 0$, we consider the subset of $B_{p,q}^s(\mathbb{R})$,

$$B_{p,q}^s(a,b,L) := \left\{ f \in B_{p,q}^s(\mathbb{R}) : \|f\|_{B_{p,q}^s} \le L, f(x) \ge 0, \int_{\mathbb{R}} f(x)\,dx = 1, \operatorname{supp} f \subset [a,b] \right\}$$

in this paper.

Lemma 2.1 *Let $\varphi = D_{2N}$ (N large enough), ψ be the corresponding wavelet and f_δ^{ft} satisfy* (2). *If $f_X \in B_{r,q}^{s+d}(a,b,L)$, $r, q \in [1,+\infty]$, $s > \frac{1}{r}$ and $2^j \le n$, then, for $p \in [1,+\infty)$,*

$$E|\hat{\alpha}_{j,k} - \alpha_{j,k}|^p \lesssim n^{-\frac{p}{2}} 2^{jp(\alpha+d)}, \qquad E|\hat{\beta}_{j,k} - \beta_{j,k}|^p \lesssim n^{-\frac{p}{2}} 2^{jp(\alpha+d)}.$$

Proof One shows only the first inequality; the second one is similar. Define

$$Z_{l,k} := \frac{(-1)^d}{2\pi} \int 2^{-\frac{j}{2}+dj}\xi(t)\frac{\overline{[\varphi^{(d)}]^{ft}(2^{-j}t)}}{f_\delta^{ft}(t)} e^{-ik2^{-j}t} e^{iW_l t}\,dt.$$

Then $\hat{\alpha}_{j,k} = \frac{1}{n}\sum_{l=1}^{n} Z_{l,k}$ and

$$\hat{\alpha}_{j,k} - \alpha_{j,k} = \frac{1}{n}\sum_{l=1}^{n}(Z_{l,k} - EZ_{l,k}) := \frac{1}{n}\sum_{l=1}^{n} Y_{l,k}. \tag{11}$$

Clearly, $EY_{l,k} = 0$. One estimates $|Y_{l,k}|$ and $E|Y_{l,k}|^2$ in order to use the Rosenthal inequality: By the assumption (2), $|Z_{l,k}| \lesssim \int 2^{-\frac{j}{2}+dj}|e^{i\frac{2\pi t}{\lambda}} - 1|^\nu \frac{|[\varphi^{(d)}]^{ft}(2^{-j}t)|}{|f_\delta^{ft}(t)|}\,dt \lesssim \int 2^{-\frac{j}{2}+dj}(1 + |t|)^\alpha |[\varphi^{(d)}]^{ft}(2^{-j}t)|\,dt = \int 2^{\frac{j}{2}+dj}(1 + |2^j t|)^\alpha |[\varphi^{(d)}]^{ft}(t)|\,dt \lesssim 2^{j(\alpha+d+\frac{1}{2})}\int(1 + |t|)^\alpha |[\varphi^{(d)}]^{ft}(t)|\,dt$. Because $\varphi = D_{2N}$, the last integration is finite for large N. Hence,

$$|Y_{l,k}| \le 2|Z_{l,k}| \lesssim 2^{j(\alpha+d+\frac{1}{2})}. \tag{12}$$

Since $f_X \in B^{s+d}_{r,q}(\mathbb{R}) \subset B^{s+d-\frac{1}{r}}_{\infty,q}(\mathbb{R})$ $(s > \frac{1}{r})$, $\|f_X\|_\infty < +\infty$ and $\|f_W\|_\infty = \|f_X * f_\delta\|_\infty \leq \|f_X\|_\infty \times \|f_\delta\|_1 < +\infty$. This with the Parseval identity shows

$$\begin{aligned} EZ^2_{l,k} &= \int \left| \frac{(-1)^d}{2\pi} \int 2^{-\frac{j}{2}+dj}\xi(t) \frac{\overline{[\varphi^{(d)}]^{ft}(2^{-j}t)}}{f^{ft}_\delta(t)} e^{-ik2^{-j}t} e^{ity}\, dt \right|^2 f_W(y)\, dy \\ &\lesssim \|f_W\|_\infty \int \left| 2^{-\frac{j}{2}+dj}\xi(t) \frac{\overline{[\varphi^{(d)}]^{ft}(2^{-j}t)}}{f^{ft}_\delta(t)} e^{-ik2^{-j}t} \right|^2 dt. \end{aligned}$$

Furthermore, one obtains $EZ^2_{l,k} \lesssim \int |2^{-\frac{j}{2}+dj}(1+|t|)^\alpha [\varphi^{(d)}]^{ft}(2^{-j}t)|^2\, dt \lesssim 2^{2j(\alpha+d)}$ thanks to (2). Hence,

$$EY^2_{l,k} = E|Z_{l,k} - EZ_{l,k}|^2 \leq EZ^2_{l,k} \lesssim 2^{2j(\alpha+d)}. \tag{13}$$

According to (11) and the Rosenthal inequality,

$$E|\hat{\alpha}_{j,k} - \alpha_{j,k}|^p \lesssim \begin{cases} n^{-p}(\sum_{l=1}^n E|Y_{l,k}|^2)^{\frac{p}{2}}, & p \in [1,2), \\ n^{-p}\sum_{l=1}^n E|Y_{l,k}|^p + n^{-p}(\sum_{l=1}^n E|Y_{l,k}|^2)^{\frac{p}{2}}, & p \in [2,+\infty). \end{cases}$$

Combining this with (13), one obtains $E|\hat{\alpha}_{j,k} - \alpha_{j,k}|^p \lesssim n^{-\frac{p}{2}} 2^{jp(\alpha+d)}$ for $1 \leq p < 2$, which is the desired conclusion. When $p \geq 2$,

$$E|Y_{l,k}|^p \lesssim \|Y_{l,k}\|^{p-2}_\infty E|Y_{l,k}|^2 \lesssim 2^{j(\alpha+d+\frac{1}{2})(p-2)} 2^{2j(\alpha+d)} = 2^{j(\frac{p}{2}-1+\alpha p+dp)}$$

due to (12) and (13). Moreover, $E|\hat{\alpha}_{j,k} - \alpha_{j,k}|^p \lesssim n^{1-p}2^{j(\frac{p}{2}-1+\alpha p+dp)} + n^{-\frac{p}{2}}2^{jp(\alpha+d)}$. Since $2^j \leq n$,

$$n^{1-p}2^{j(\frac{p}{2}-1+\alpha p+dp)} = n^{-\frac{p}{2}} \left(\frac{2^j}{n}\right)^{\frac{p}{2}-1} 2^{jp(\alpha+d)} \leq n^{-\frac{p}{2}} 2^{jp(\alpha+d)}.$$

Finally, $E|\hat{\alpha}_{j,k} - \alpha_{j,k}|^p \lesssim n^{-\frac{p}{2}}2^{jp(\alpha+d)}$. This completes the proof of the first part of Lemma 2.1. □

Now, we are in a position to state our first theorem.

Theorem 2.1 *Let* $\hat{f}^{lin}_{n,d}$ *be defined by* (8)-(9), $r,q \in [1,+\infty]$, $p \in [1,+\infty)$ *and* $s > \frac{1}{r}$. *Then with* $s' := s - (\frac{1}{r} - \frac{1}{p})_+$ *and* $x_+ = \max\{0,x\}$,

$$\sup_{f_X \in B^{s+d}_{r,q}(a,b,L)} E\big\|\hat{f}^{lin}_{n,d} - f^{(d)}_X\big\|^p_p \lesssim n^{-\frac{s'p}{2s'+2\alpha+2d+1}}.$$

Proof It is easy to see that $\|\hat{f}^{lin}_{n,d} - f^{(d)}_X\|_p \lesssim \|\hat{f}^{lin}_{n,d} - f^{(d)}_X\|_r$ for $r \geq p$, because $L_r([a,b])$ is continuously embedded into $L_p([a,b])$. Moreover, $E\|\hat{f}^{lin}_{n,d} - f^{(d)}_X\|^p_p \lesssim E\|\hat{f}^{lin}_{n,d} - f^{(d)}_X\|^p_r \leq (E\|\hat{f}^{lin}_{n,d} - f^{(d)}_X\|^r_r)^{\frac{p}{r}}$ thanks to Jensen's inequality.

When $r \leq p$, $s' - \frac{1}{p} = s - \frac{1}{r}$ and $B^{s+d}_{r,q}(a,b,L) \subset B^{s'+d}_{p,q}(a,b,L)$. Then

$$\sup_{f_X \in B^{s+d}_{r,q}(a,b,L)} E\big\|\hat{f}^{lin}_{n,d} - f^{(d)}_X\big\|^p_p \leq \sup_{f_X \in B^{s'+d}_{p,q}(a,b,L)} E\big\|\hat{f}^{lin}_{n,d} - f^{(d)}_X\big\|^p_p.$$

Therefore, it suffices to prove the theorem, for $r = p$,

$$\sup_{f_X\in B^{s+d}_{p,q}(a,b,L)} E\big\|\hat{f}^{lin}_{n,d} - f^{(d)}_X\big\|^p_p \lesssim n^{-\frac{sp}{2s+2\alpha+2d+1}}. \tag{14}$$

If $f_X \in B^{s+d}_{p,q}(\mathbb{R})$, then $f^{(d)}_X \in B^s_{p,q}(\mathbb{R})$ and

$$\big\|P_jf^{(d)}_X - f^{(d)}_X\big\|^p_p \lesssim 2^{-jsp} \tag{15}$$

due to Lemma 1.2. On the other hand, $\hat{f}^{lin}_{n,d} - P_jf^{(d)}_X = \sum_{k\in K_j}(\hat{\alpha}_{j,k} - \alpha_{j,k})\varphi_{j,k}$ and

$$E\big\|\hat{f}^{lin}_{n,d} - P_jf^{(d)}_X\big\|^p_p \lesssim 2^{j(\frac{p}{2}-1)}\sum_{k\in K_j} E|\hat{\alpha}_{j,k} - \alpha_{j,k}|^p \lesssim 2^{j\frac{p}{2}} \sup_{k\in K_j} E|\hat{\alpha}_{j,k} - \alpha_{j,k}|^p$$

because of Lemma 1.1 and $|K_j| \sim 2^j$. This with Lemma 2.1 leads to $E\|\hat{f}^{lin}_{n,d} - P_jf^{(d)}_X\|^p_p \lesssim (\frac{2^j}{n})^{\frac{p}{2}} 2^{jp(\alpha+d)}$. Combining this with (15), one obtains

$$E\big\|\hat{f}^{lin}_{n,d} - f^{(d)}_X\big\|^p_p \lesssim E\big\|\hat{f}^{lin}_{n,d} - P_jf^{(d)}_X\big\|^p_p + \big\|P_jf^{(d)}_X - f^{(d)}_X\big\|^p_p \lesssim \left(\frac{2^j}{n}\right)^{\frac{p}{2}} 2^{jp(\alpha+d)} + 2^{-jsp}.$$

Take $2^j \sim n^{\frac{1}{2s+2\alpha+2d+1}}$. Then the inequality (14) follows, and the proof of Theorem 2.1 is finished. □

Remark 2.1 If $p = q = r = 2$ and $d = 0$, then $s' = s$, $B^{s+d}_{r,q}(a,b,L) = W^s_2(a,b,L)$, Theorem 2.1 reduces to Theorem 4.1 in [11].

Remark 2.2 From the choice $2^j \sim n^{\frac{1}{2s+2\alpha+2d+1}}$ in the proof of Theorem 2.1, we find that our estimator is not adaptive, because it depends on the parameter s of $B^s_{r,q}(\mathbb{R})$. In order to avoid that shortcoming, we study a nonlinear estimation in the next part.

3 Nonlinear estimation

This section is devoted to an adaptive nonlinear estimation, which also improves the convergence rate of the linear one in some cases. The idea of proof comes from [12]. Choose $r_0 > s$,

$$2^{j_0} \sim n^{\frac{1}{2r_0+2\alpha+2d+1}} \quad \text{and} \quad 2^{j_1} \sim n^{\frac{1}{2\alpha+2d+1}}. \tag{16}$$

Let $\hat{\alpha}_{j,k}$ and $\hat{\beta}_{j,k}$ be defined by (9) and (10), respectively, and

$$\tilde{\beta}_{j,k} = \begin{cases} \hat{\beta}_{j,k}, & |\hat{\beta}_{j,k}| > T2^{j(\alpha+d)}\sqrt{j/n}, \\ 0, & \text{otherwise}, \end{cases}$$

where the constant T will be determined in the proof of Theorem 3.1. Then we define a nonlinear wavelet estimator

$$\hat{f}^{non}_{n,d}(x) := \sum_{k\in K_{j_0}} \hat{\alpha}_{j_0,k}\varphi_{j_0,k}(x) + \sum_{j=j_0}^{j_1}\sum_{k\in I_j} \tilde{\beta}_{j,k}\psi_{j,k}(x),$$

where $K_{j_0} := \{k \in \mathbb{Z} : \langle f_X, \varphi_{j_0,k}\rangle \neq 0\}$, and $I_j := \{k \in \mathbb{Z} : \langle f_X, \psi_{j,k}\rangle \neq 0\}$. Clearly, the cardinality $|I_j| \sim 2^j$ since both f_X and ψ have compact supports.

Lemma 3.1 *If $j2^j \le n$, then there exists $c_0 > 0$ such that, for each $T \ge T_0 > 0$,*

$$P\left\{|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{T}{2} 2^{j(\alpha+d)} \sqrt{j/n}\right\} \lesssim 2^{-c_0 Tj}.$$

Proof By the definitions of $\beta_{j,k}$ and $\hat{\beta}_{j,k}$, $\hat{\beta}_{j,k} - \beta_{j,k} = \frac{1}{n}\sum_{l=1}^{n}(Z_{l,k} - EZ_{l,k}) := \frac{1}{n}\sum_{l=1}^{n} Y_{l,k}$, where

$$Z_{l,k} := \frac{(-1)^d}{2\pi}\int 2^{-\frac{j}{2}+dj}\sum_{m=0}^{J}\eta_m e^{it\frac{2\pi m}{\lambda}}\left(e^{\frac{2\pi it}{\lambda}} - 1\right)^{\nu}\frac{\overline{[\psi^{(d)}]^{ft}(2^{-j}t)}}{f_\delta^{ft}(t)} e^{-i2^{-j}kt} e^{iW_l t}\,dt.$$

Then $EY_{l,k} = 0$ and with $\lambda = \frac{T}{2}2^{j(\alpha+d)}\sqrt{\frac{j}{n}}$,

$$P\left\{|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{T}{2} 2^{j(\alpha+d)} \sqrt{\frac{j}{n}}\right\} \le 2\exp\left(-\frac{n\lambda^2}{2(EY_{l,k}^2 + \|Y_{l,k}\|_\infty \lambda/3)}\right) \tag{17}$$

thanks to the classical Bernstein inequality in [15]. On the other hand, $EY_{l,k}^2 + \|Y_{l,k}\|_\infty\lambda/3 \lesssim 2^{2j(\alpha+d)} + 2^{(\alpha+d+\frac{1}{2})j}\frac{T}{6}2^{j(\alpha+d)}\sqrt{\frac{j}{n}} \le CT2^{2j(\alpha+d)}$ because of (12), (13), and $j2^j \le n$. Hence, $\frac{n\lambda^2}{2(EY_{l,k}^2 + \|Y_{l,k}\|_\infty\lambda/3)} \ge \frac{n\frac{T^2}{4}\frac{j}{n}2^{2j(\alpha+d)}}{2CT2^{2j(\alpha+d)}} = \frac{T}{8C}j$, and (17) reduces to

$$P\left\{|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{T}{2} 2^{j(\alpha+d)} \sqrt{j/n}\right\} \lesssim 2e^{-\frac{T}{8C}j} = 2^{-c_0 Tj}$$

with $c_0 = \frac{1}{8C}\log_2 e$. This completes the proof of Lemma 2.1. □

Theorem 3.1 *Under the assumptions of Theorem* 2.1, *there exist $\theta > 0$ and $T_0 > 0$ such that, for $T \ge T_0$,*

$$\sup_{f_X \in B_{r,q}^{s+d}(a,b,L)} E\left\|\hat{f}_{n,d}^{non} - f_X^{(d)}\right\|_p^p \lesssim \begin{cases} (\ln n)^\theta n^{-\frac{sp}{2s+2\alpha+2d+1}}, & r \in (\frac{(2\alpha+2d+1)p}{2s+2\alpha+2d+1}, p]; \\ (\ln n)^\theta n^{-\frac{s'p}{2(s-1/r)+2\alpha+2d+1}}, & r \in [1, \frac{(2\alpha+2d+1)p}{2s+2\alpha+2d+1}]. \end{cases} \tag{18}$$

Proof Similar to [12], one defines

$$\mu := \min\left\{\frac{s}{2s+2\alpha+2d+1}, \frac{s'}{2(s-1/r)+2\alpha+2d+1}\right\},$$

and

$$\omega := -sr + \left(\alpha + d + \frac{1}{2}\right)(p-r). \tag{19}$$

It is easy to check that $\omega < 0$ holds if and only if $r > \frac{(2\alpha+2d+1)p}{2s+2\alpha+2d+1}$, and $\mu = \frac{s}{2s+2\alpha+2d+1}$ as well as $\omega \geq 0$ if and only if $r \leq \frac{(2\alpha+2d+1)p}{2s+2\alpha+2d+1}$, and $\mu = \frac{s'}{2(s-1/r)+2\alpha+2d+1}$. Then the conclusion of Theorem 3.1 can be rewritten as

$$\sup_{f_X \in B^{s+d}_{r,q}(a,b,L)} E\big\|\hat{f}^{non}_{n,d} - f^{(d)}_X\big\|^p_p \lesssim (\ln n)^\theta n^{-\mu p}. \tag{20}$$

Choose $j_0(s,r,q)$ and $j_1(s,r,q)$ such that

$$2^{j_0(s,r,q)} \backsimeq n^{\frac{1-2\mu}{2\alpha+2d+1}} \quad \text{and} \quad 2^{j_1(s,r,q)} \backsimeq n^{\frac{\mu}{s'}}.$$

Then it can easily be shown by (16) that $2^{j_0} \lesssim 2^{j_0(s,r,q)} \lesssim 2^{j_1(s,r,q)} \lesssim 2^{j_1}$. Clearly,

$$\big\|\hat{f}^{non}_{n,d} - f^{(d)}_X\big\|^p_p \lesssim \big\|P_{j_0}\big(\hat{f}^{non}_{n,d} - f^{(d)}_X\big)\big\|^p_p + \big\|D_{j_0,j_1}\big(\hat{f}^{non}_{n,d} - f^{(d)}_X\big)\big\|^p_p + \big\|P_{j_1}f^{(d)}_X - f^{(d)}_X\big\|^p_p, \tag{21}$$

where

$$D_{j_0,j_1}f = \sum_{j=j_0}^{j_1}\sum_{k\in\mathbb{Z}} \beta_{j,k}\psi_{j,k}.$$

By the assumption $r \leq p$, $s' := s - (\frac{1}{r} - \frac{1}{p}) = s - \frac{1}{r} + \frac{1}{p}$ and $B^{s+d}_{r,q}(a,b,L)$ is continuously embedded into $B^{s'+d}_{p,q}(a,b,L)$. Since $f_X \in B^{s'+d}_{p,q}(\mathbb{R})$, $f^{(d)}_X \in B^{s'}_{p,q}(\mathbb{R})$ and $\|P_{j_1}f^{(d)}_X - f^{(d)}_X\|^p_p \lesssim 2^{-j_1 s' p}$ thanks to Lemma 1.2. This with $2^{j_1(s,r,q)} \lesssim 2^{j_1}$ and the definition of μ leads to

$$\big\|P_{j_1}f^{(d)}_X - f^{(d)}_X\big\|^p_p \lesssim 2^{-j_1(s,r,p)s'p} \lesssim n^{-\mu p}. \tag{22}$$

Note that $P_{j_0}(\hat{f}^{non}_{n,d} - f^{(d)}_X) = \sum_{k\in K_{j_0}}(\hat{\alpha}_{j_0,k} - \alpha_{j_0,k})\varphi_{j_0,k}$. Then $E\|P_{j_0}(\hat{f}^{non}_{n,d} - f^{(d)}_X)\|^p_p \lesssim 2^{j_0(\frac{p}{2}-1)} \times \sum_{k\in K_{j_0}} E|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p \lesssim 2^{j_0\frac{p}{2}} n^{-\frac{p}{2}} 2^{j_0 p(\alpha+d)}$ due to Lemma 1.1, $|K_{j_0}| \sim 2^{j_0}$ and Lemma 2.1. By $2^{j_0} \lesssim 2^{j_0(s,r,q)}$ and the choice $2^{j_0(s,r,q)} \backsimeq n^{\frac{1-2\mu}{2\alpha+2d+1}}$,

$$E\big\|P_{j_0}\big(\hat{f}^{non}_{n,d} - f^{(d)}_X\big)\big\|^p_p \lesssim \left(\frac{2^{j_0(s,r,q)}}{n}\right)^{p/2} 2^{j_0(s,r,q)p(\alpha+d)} \lesssim n^{-\mu p}. \tag{23}$$

According to (21)-(23), it is sufficient to prove $E\|D_{j_0,j_1}(\hat{f}^{non}_{n,d} - f^{(d)}_X)\|^p_p \lesssim (\ln n)^\theta n^{-\mu p}$: Define

$$\hat{B}_j := \big\{k : |\hat{\beta}_{j,k}| > T2^{j(\alpha+d)}\sqrt{j/n}\big\}, \qquad \hat{S}_j = \hat{B}^c_j;$$

$$B_j := \left\{k : |\beta_{j,k}| > \frac{T}{2}2^{j(\alpha+d)}\sqrt{j/n}\right\}, \qquad S_j = B^c_j;$$

$$B'_j := \big\{k : |\beta_{j,k}| > 2T2^{j(\alpha+d)}\sqrt{j/n}\big\}, \qquad S'_j = B'^c_j.$$

Then $D_{j_0,j_1}(\hat{f}^{non}_{n,d} - f^{(d)}_X) = \sum_{j=j_0}^{j_1}\sum_{k\in I_j}(\hat{\beta}_{j,k} - \beta_{j,k})[I\{k \in \hat{B}_j \cap S_j\} + I\{k \in \hat{B}_j \cap B_j\}]\psi_{j,k} - \sum_{j=j_0}^{j_1}\sum_{k\in I_j}\beta_{j,k}[I\{k \in \hat{S}_j \cap S'_j\} + I\{k \in \hat{S}_j \cap B'_j\}]\psi_{j,k} = e_{bs} + e_{sb} + e_{bb} + e_{ss}$, where

$$e_{bs} := \sum_{j=j_0}^{j_1}\sum_{k\in I_j}(\hat{\beta}_{j,k} - \beta_{j,k})I\{k \in \hat{B}_j \cap S_j\}\psi_{j,k}, \qquad e_{sb} := \sum_{j=j_0}^{j_1}\sum_{k\in I_j}\beta_{j,k}I\big\{k \in \hat{S}_j \cap B'_j\big\}\psi_{j,k},$$

$$e_{bb} := \sum_{j=j_0}^{j_1} \sum_{k\in I_j} (\hat{\beta}_{j,k} - \beta_{j,k}) I\{k \in \hat{B}_j \cap B_j\} \psi_{j,k}, \qquad e_{ss} := \sum_{j=j_0}^{j_1} \sum_{k\in I_j} \beta_{j,k} I\{k \in \hat{S}_j \cap S'_j\} \psi_{j,k}.$$

In order to conclude Theorem 3.1, one needs only to show that

$$E\|e_{bs}\|_p^p + E\|e_{sb}\|_p^p + E\|e_{bb}\|_p^p + E\|e_{ss}\|_p^p \lesssim (\ln n)^\theta n^{-\mu p}. \tag{24}$$

By (16), $j_1 - j_0 \sim \ln n$ and $\| \sum_{j=j_0}^{j_1} g_j \|_p^p \lesssim (j_1 - j_0 + 1)^{p-1} \sum_{j=j_0}^{j_1} \|g_j\|_p^p \lesssim (\ln n)^{p-1} \sum_{j=j_0}^{j_1} \|g_j\|_p^p$. This with Lemma 1.1 shows that, for $\hat{f}(x) = \sum_{j=j_0}^{j_1} \sum_{k\in I_j} \hat{f}_{j,k} \psi_{j,k}(x)$, there exists $\theta > 0$ such that

$$E\|\hat{f}\|_p^p \lesssim (\ln n)^\theta \sup_{j_0 \le j \le j_1} 2^{j(\frac{p}{2}-1)} \sum_{k\in I_j} E|\hat{f}_{j,k}|^p. \tag{25}$$

To estimate $E\|e_{bs}\|_p^p$, one takes $\hat{f}_{j,k} := (\hat{\beta}_{j,k} - \beta_{j,k}) I\{k \in \hat{B}_j \cap S_j\}$. Then, for each $k \in \hat{B}_j \cap S_j$, $|\hat{\beta}_{j,k} - \beta_{j,k}| \ge |\hat{\beta}_{j,k}| - |\beta_{j,k}| > \frac{T}{2} 2^{j(\alpha+d)} \sqrt{\frac{j}{n}}$ and $\hat{B}_j \cap S_j \subset \hat{D}_j := \{k : |\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{T}{2} 2^{j(\alpha+d)} \sqrt{\frac{j}{n}}\}$. This with the Hölder inequality shows

$$E|\hat{f}_{j,k}|^p \le E|\hat{\beta}_{j,k} - \beta_{j,k}|^p I\{k \in \hat{D}_j\} \le \left(E|\hat{\beta}_{j,k} - \beta_{j,k}|^{2p}\right)^{\frac{1}{2}} \left(EI\{k \in \hat{D}_j\}\right)^{\frac{1}{2}}.$$

Clearly, $EI\{k \in \hat{D}_j\} = P\{|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{T}{2} 2^{j(\alpha+d)} \sqrt{\frac{j}{n}}\}$. Furthermore, using $|I_j| \lesssim 2^j$, Lemma 2.1, and Lemma 3.1, one obtains $\sum_{k\in I_j} E|\hat{f}_{j,k}|^p \lesssim 2^j n^{-\frac{p}{2}} 2^{jp(\alpha+d)} 2^{-\frac{c_0 T}{2} j}$. Moreover, by (25),

$$E\|e_{bs}\|_p^p \lesssim (\ln n)^\theta \sup_{j_0 \le j \le j_1} n^{-\frac{p}{2}} 2^{j(\frac{p}{2}+\alpha p + dp - \frac{c_0 T}{2})}.$$

Choose $T \ge \frac{p+2\alpha p+2dp}{c_0}$. Then $\frac{p}{2} + \alpha p + dp - \frac{c_0 T}{2} \le 0$, and $\sup_{j_0\le j\le j_1} n^{-\frac{p}{2}} 2^{j(\frac{p}{2}+\alpha p+dp-\frac{c_0 T}{2})} \lesssim n^{-\frac{p}{2}} 2^{j_0(\frac{p}{2}+\alpha p+dp-\frac{c_0 T}{2})} \lesssim n^{-\frac{p}{2}} 2^{j_0 \frac{p}{2}} 2^{j_0 p(\alpha+d)}$. Similar to (23), one has

$$E\|e_{bs}\|_p^p \lesssim (\ln n)^\theta n^{-\mu p}. \tag{26}$$

In the proof of (26), one needs to choose $T_0 \ge c_0^{-1}(p + 2\alpha p + 2dp)$.

Now, one considers $E\|e_{sb}\|_p^p$: For $k \in \hat{S}_j \cap B'_j$, $|\hat{\beta}_{j,k} - \beta_{j,k}| \ge |\beta_{j,k}| - |\hat{\beta}_{j,k}| > T 2^{j(\alpha+d)} \sqrt{\frac{j}{n}}$ and $\hat{S}_j \cap B'_j \subset \hat{D}_j := \{k : |\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{T}{2} 2^{j(\alpha+d)} \sqrt{\frac{j}{n}}\}$. By Lemma 3.1, $EI\{k \in \hat{S}_j \cap B'_j\} \le EI\{k \in \hat{D}_j\} \lesssim 2^{-c_0 Tj}$. Since $f_X^{(d)} \in B^s_{r,q}(\mathbb{R}) \subset B^{s'}_{p,q}(\mathbb{R})$, $\|\beta_{j,\cdot}\|_p := \|\langle f_X^{(d)}, \psi_{j,\cdot}\rangle\|_p \lesssim \|f_X^{(d)}\|_{B^{s'}_{p,q}} 2^{-j(s'+1/2-1/p)}$ and

$$\sum_{k\in I_j} |\beta_{j,k}|^p EI\{k \in \hat{S}_j \cap B'_j\} \lesssim \|\beta_{j,\cdot}\|_p^p 2^{-c_0 Tj} \lesssim \|f_X^{(d)}\|_{B^{s'}_{p,q}}^p 2^{-j(s'p+\frac{p}{2}-1+c_0 T)}.$$

Moreover, it follows from the definition of e_{sb} and (25) that

$$E\|e_{sb}\|_p^p \lesssim (\ln n)^\theta \sup_{j_0\le j\le j_1} 2^{-j(s'p+c_0 T)} \le (\ln n)^\theta 2^{-j_0(s'p+c_0 T)}.$$

By (16), one can choose $T \ge \frac{j_1 - j_0}{c_0 j_0} s'p$ (independent of n) so that $j_1 s'p \le j_0(s'p + c_0 T)$. Hence, this above inequality reduces to $E\|e_{sb}\|_p^p \lesssim (\ln n)^\theta 2^{-j_1 s'p}$. Similar arguments to (22) lead to

$$E\|e_{sb}\|_p^p \lesssim (\ln n)^\theta n^{-\mu p}. \tag{27}$$

In this above proof, one needs to choose $T_0 \geq (c_0 j_0)^{-1}(j_1 - j_0)s'p$.

For $E\|e_{bb}\|_p^p \lesssim (\ln n)^\theta n^{-\mu p}$, one uses $|I_j| \lesssim 2^j$, (25), and Lemma 2.1 to find

$$\mathrm{I} := E\left\|\sum_{j=j_0}^{j_0(s,r,q)}\sum_{k\in I_j}(\hat{\beta}_{j,k} - \beta_{j,k})I\{k \in \hat{B}_j \cap B_j\}\psi_{j,k}\right\|_p^p \lesssim (\ln n)^\theta \sup_{j_0\leq j\leq j_0(s,r,q)} n^{-\frac{p}{2}}2^{jp(\alpha+d+\frac{1}{2})}.$$

Recall that $j_0 \leq j_0(s,r,q)$. Then $I \lesssim (\ln n)^\theta (n^{-1}2^{j_0(s,r,q)})^{\frac{p}{2}}2^{j_0(s,r,q)p(\alpha+d)}$, which reduces to $\mathrm{I} \lesssim (\ln n)^\theta n^{-\mu p}$ by similar arguments of (23). It remains to show

$$\mathrm{II} := E\left\|\sum_{j=j_0(s,r,q)}^{j_1}\sum_{k\in I_j}(\hat{\beta}_{j,k} - \beta_{j,k})I\{k \in \hat{B}_j \cap B_j\}\psi_{j,k}\right\|_p^p \lesssim (\ln n)^\theta n^{-\mu p}. \tag{28}$$

By Lemma 2.1 and the definition of B_j, $\sum_{k\in I_j}E|(\hat{\beta}_{j,k} - \beta_{j,k})I\{k \in \hat{B}_j \cap B_j\}|^p \lesssim \sum_{k\in I_j}E|\hat{\beta}_{j,k} - \beta_{j,k}|^p I\{k \in B_j\} \lesssim n^{-\frac{p}{2}}2^{jp(\alpha+d)}\sum_{k\in B_j}|2\beta_{j,k}T^{-1}2^{-j(\alpha+d)}\sqrt{nj^{-1}}|^r$. According to Lemma 1.2, $\|\beta_{j,\cdot}\|_r^r \lesssim \|f_X^{(d)}\|_{B^s_{r,q}}^r 2^{-j(sr+r/2-1)}$. Hence,

$$\sum_{k\in I_j}E\left|(\hat{\beta}_{j,k} - \beta_{j,k})I\{k \in \hat{B}_j \cap B_j\}\right|^p \lesssim n^{-\frac{p-r}{2}}(\ln n)^\theta 2^{-j(sr+\frac{r}{2}-1-\alpha p-dp+\alpha r+dr)}.$$

Combining this above inequality with (25), one obtains

$$\mathrm{II} \lesssim (\ln n)^\theta n^{-\frac{p-r}{2}} \sup_{j_0(s,r,q)\leq j\leq j_1} 2^{j\omega} := A_n,$$

where $\omega := -sr + (\alpha + d + \frac{1}{2})(p - r) = -sr - \frac{r}{2} + \frac{p}{2} + \alpha p + dp - \alpha r - dr$ as defined in (19).

When $\omega \leq 0$, $A_n \lesssim (\ln n)^\theta n^{-\frac{p-r}{2}}2^{j_0(s,r,q)\omega}$. By the choice $2^{j_0(s,r,q)} \sim n^{\frac{1-2\mu}{2\alpha+2d+1}}$,

$$A_n \lesssim (\ln n)^\theta n^{-\frac{p-r}{2}} n^{\frac{1-2\mu}{2\alpha+2d+1}[-sr+(\alpha+d+1/2)(p-r)]} = (\ln n)^\theta n^{-\mu p} n^{r(\mu-\frac{1-2\mu}{2\alpha+2d+1}s)}.$$

Since $\mu = \frac{s}{2s+2\alpha+2d+1}$ for $\omega \leq 0$, $\mu - \frac{1-2\mu}{2\alpha+2d+1}s = 0$ and $A_n \lesssim (\ln n)^\theta n^{-\mu p}$. Then one obtains the desired inequality (28).

When $\omega > 0$, $r < \frac{(2\alpha+2d+1)p}{2s+2\alpha+2d+1}$ and $\mu = \frac{s'}{2(s-1/r)+2\alpha+2d+1}$. Take

$$p_1 = \frac{2\alpha p + 2dp + p - 2}{2(s - \frac{1}{r}) + 2\alpha + 2d + 1}.$$

Then $\frac{p-p_1}{2} = \mu p$ and $r < p_1$ in that case. With $\hat{f}_{j,k} = (\hat{\beta}_{j,k} - \beta_{j,k})I\{k \in \hat{B}_j \cap B_j\}$, one knows from Lemma 2.1 and the definition of B_j that

$$\sum_{k\in I_j}E|\hat{f}_{j,k}|^p \leq n^{-\frac{p}{2}}2^{jp(\alpha+d)}\sum_{k\in B_j}\left|\frac{2\beta_{j,k}}{T}2^{-j(\alpha+d)}\sqrt{\frac{n}{j}}\right|^{p_1} \lesssim (\ln n)^\theta n^{-\frac{p-p_1}{2}}2^{j(p-p_1)(\alpha+d)}\|\beta_{j,\cdot}\|_{p_1}^{p_1}.$$

Since $r < p_1$, $\|\beta_{j,\cdot}\|_{p_1} \le \|\beta_{j,\cdot}\|_r$. This with (25) leads to

$$E\|e_{bb}\|_p^p \lesssim (\ln n)^\theta n^{-\frac{p-p_1}{2}} \sup_{j_0\le j\le j_1} \bigl[2^{j(\frac{p-2+2\alpha p+2dp}{2p_1})}2^{-j(\alpha+d)}\|\beta_{j,\cdot}\|_r\bigr]^{p_1}.$$

Note that $2^{j\frac{p-2+2\alpha p+2dp}{2p_1}}2^{-j(\alpha+d)} = 2^{j(s-\frac{1}{r}+\frac{1}{2})}$ due to the definition of p_1, as well as $f_X^{(d)} \in B^s_{r,q}(a,b,L)$ implies $2^{j(s-\frac{1}{r}+\frac{1}{2})}\|\beta_{j,\cdot}\|_r \lesssim \|f_X^{(d)}\|_{B^s_{r,q}}$. Then

$$E\|e_{bb}\|_p^p \lesssim (\ln n)^\theta n^{-\frac{p-p_1}{2}} = (\ln n)^\theta n^{-\mu p}. \tag{29}$$

Finally, one estimates $E\|e_{ss}\|_p^p$: Define $\hat{f}_{j,k} := \beta_{j,k}I\{k \in \hat{S}_j \cap S'_j\}$. Then

$$\sum_{k\in I_j}|\hat{f}_{j,k}|^p \le \sum_{k\in S'_j}|\beta_{j,k}|^{p-r}|\beta_{j,k}|^r \le \left(2T2^{j(\alpha+d)}\sqrt{\frac{j}{n}}\right)^{p-r}2^{-j(s+\frac{1}{2}-\frac{1}{r})r}\|f_X^{(d)}\|^r_{B^s_{r,q}}$$

due to $r \le p$ and the definition of S'_j. Using (25) and $\omega := -sr + (\alpha + d + \frac{1}{2})(p - r)$ in (19), one obtains

$$E\|e_{ss}\|_p^p = E\left\|\sum_{j=j_0}^{j_1}\sum_{k\in I_j}\hat{f}_{j,k}\psi_{j,k}\right\|_p^p = \begin{cases}(\ln n)^\theta 2^{j_0\omega}n^{-\frac{p-r}{2}}, & \omega \le 0,\\ (\ln n)^\theta 2^{j_1\omega}n^{-\frac{p-r}{2}}, & \omega > 0.\end{cases} \tag{30}$$

When $\omega \le 0$, $E\|\sum_{j=j_0(s,r,q)}^{j_1}\sum_{k\in I_j}\beta_{j,k}I\{k \in \hat{S}_j \cap S'_j\}\psi_{j,k}\|_p^p \lesssim (\ln n)^\theta n^{-\frac{p-r}{2}}2^{j_0(s,r,q)\omega}$. Recall that $\omega \le 0$ holds if and only if $\mu = \frac{s}{2s+2\alpha+2d+1}$ and $2^{j_0(s,r,q)} \sim n^{\frac{1-2\mu}{2\alpha+2d+1}}$. Then it can be checked that $2^{j_0(s,r,q)\omega}n^{-\frac{p-r}{2}} = n^{-\mu p}$. Hence,

$$E\left\|\sum_{j=j_0(s,r,q)}^{j_1}\sum_{k\in I_j}\beta_{j,k}I\{k \in \hat{S}_j \cap S'_j\}\psi_{j,k}\right\|_p^p \lesssim (\ln n)^\theta n^{-\mu p}. \tag{31}$$

On the other hand, Lemma 1.1 tells that

$$E\left\|\sum_{j=j_0}^{j_0(s,r,q)}\sum_{k\in\hat{S}_j\cap S'_j}\beta_{j,k}\psi_{j,k}\right\|_p^p \lesssim \sum_{j=j_0}^{j_0(s,r,q)}2^{j(\frac{p}{2}-1)}\sum_{k\in\hat{S}_j\cap S'_j}E|\beta_{j,k}|^p.$$

Since $k \in S'_j$, $|\beta_{j,k}| \le 2T2^{j(\alpha+d)}\sqrt{\frac{j}{n}}$ and $\sum_{k\in\hat{S}_j\cap S'_j}|\beta_{j,k}|^p \lesssim \sum_{k\in I_j}2^{jp(\alpha+d)}(\ln n)^\theta n^{-\frac{p}{2}} \lesssim (\ln n)^\theta \times n^{-\frac{p}{2}}2^j 2^{jp(\alpha+d)}$. Moreover,

$$\begin{aligned}
&E\left\|\sum_{j=j_0}^{j_0(s,r,q)}\sum_{k\in I_j}\beta_{j,k}I\{k \in \hat{S}_j \cap S'_j\}\psi_{j,k}\right\|_p^p\\
&\quad\lesssim (\ln n)^\theta n^{-\frac{p}{2}}\sup_{j_0\le j\le j_0(s,r,q)}2^{j(\frac{p}{2}-1)}2^j 2^{jp(\alpha+d)}\\
&\quad\lesssim (\ln n)^\theta n^{-\frac{p}{2}}2^{j_0(s,r,q)(\alpha+d+\frac{1}{2})p} \lesssim (\ln n)^\theta n^{-\mu p},
\end{aligned} \tag{32}$$

where the last inequality comes from the choice $2^{j_0(s,r,q)} \sim n^{\frac{1-2\mu}{2\alpha+2d+1}}$. Combining this with (31), one has, for $\omega \le 0$,

$$E\|e_{ss}\|_p^p \lesssim (\ln n)^\theta n^{-\mu p}. \tag{33}$$

It remains to show $E\|e_{ss}\|_p^p \lesssim n^{-\mu p}$ for $\omega > 0$. By Lemma 1.1,

$$\begin{aligned}\Big\|\sum_{k\in I_j}\beta_{j,k}I\{k\in\hat{S}_j\cap S_j\}\psi_{j,k}\Big\|_p &\sim 2^{j(\frac{1}{2}-\frac{1}{p})}\Big(\sum_{k\in I_j}\big|\beta_{j,k}I\{k\in\hat{S}_j\cap S_j\}\big|^p\Big)^{1/p}\\ &\le 2^{j(\frac{1}{2}-\frac{1}{p})}\|\beta_{j,\cdot}\|_{l_p}.\end{aligned}$$

This with Lemma 1.2 shows that

$$\begin{aligned}&\Bigg[\sum_{j=j_1(s,r,q)}^{j_1}\Big(2^{js'}\Big\|\sum_{k\in I_j}\beta_{j,k}I\{k\in\hat{S}_j\cap S_j\}\psi_{j,k}\Big\|_p\Big)^q\Bigg]^{1/q}\\ &\quad\le\Bigg[\sum_{j=j_1(s,r,q)}^{j_1}\big(2^{j(s'+\frac{1}{2}-\frac{1}{p})}\|\beta_{j,\cdot}\|_{l_p}\big)^q\Bigg]^{1/q}\le\|f_X^{(d)}\|_{B^{s'}_{p,q}}.\end{aligned} \tag{34}$$

When $q = 1$,

$$\begin{aligned}&\sum_{j=j_1(s,r,q)}^{j_1}\Big\|\sum_{k\in I_j}\beta_{j,k}I\{k\in\hat{S}_j\cap S_j\}\psi_{j,k}\Big\|_p\\ &\quad\lesssim\sum_{j=j_1(s,r,q)}^{j_1}2^{js'}\Big\|\sum_{k\in I_j}\beta_{j,k}I\{k\in\hat{S}_j\cap S_j\}\psi_{j,k}\Big\|_p 2^{-j_1(s,r,q)s'}\lesssim 2^{-j_1(s,p,q)s'}.\end{aligned}$$

When $q = +\infty$,

$$\begin{aligned}&\sum_{j=j_1(s,r,q)}^{j_1}\Big\|\sum_{k\in I_j}\beta_{j,k}I\{k\in\hat{S}_j\cap S_j\}\psi_{j,k}\Big\|_p\\ &\quad\lesssim\sum_{j=j_1(s,r,q)}^{j_1}2^{-js'}\|f_X^{(d)}\|_{B^{s'}_{p,q}}\lesssim(\ln n)^\theta 2^{-j_1(s,p,q)s'}.\end{aligned}$$

For $1 < q < +\infty$, by the Hölder inequality,

$$\begin{aligned}&\sum_{j=j_1(s,r,q)}^{j_1}\Big\|\sum_{k\in I_j}\beta_{j,k}I\{k\in\hat{S}_j\cap S_j\}\psi_{j,k}\Big\|_p\\ &\quad\lesssim\Bigg[\sum_{j=j_1(s,r,q)}^{j_1}2^{-js'q'}\Bigg]^{\frac{1}{q'}}\Bigg[\sum_{j=j_1(s,r,q)}^{j_1}\Big(2^{js'}\Big\|\sum_{k\in I_j}\beta_{j,k}I\{k\in\hat{S}_j\cap S_j\}\psi_{j,k}\Big\|_p\Big)^q\Bigg]^{\frac{1}{q}}\\ &\quad\lesssim 2^{-j_1(s,r,q)s'}.\end{aligned}$$

Using (34) and the choice $2^{j_1(s,r,q)} \sim n^{\frac{\mu}{s'}}$, one obtains

$$E\left\| \sum_{j=j_1(s,r,q)}^{j_1} \sum_{k\in I_j} \beta_{j,k} I\{k \in \hat{S}_j \cap S_j\} \psi_{j,k} \right\|_p^p \lesssim (\ln n)^\theta n^{-\mu p}.$$

On the other hand,

$$E\left\| \sum_{j=j_0}^{j_1(s,r,q)} \sum_{k\in I_j} \beta_{j,k} I\{k \in \hat{S}_j \cap S_j\} \psi_{j,k} \right\|_p^p \lesssim (\ln n)^\theta 2^{j_1(s,r,q)\omega} n^{-\frac{p-r}{2}}$$

thanks to (30). According to the choice of $2^{j_1(s,r,q)} \sim n^{\frac{\mu}{s'}}$ and

$$\mu = \frac{s - 1/r + 1/p}{2(s - 1/r) + 2\alpha + 2d + 1} = \frac{s'}{2(s - 1/r) + 2\alpha + 2d + 1}$$

for $\omega > 0$, one finds $2^{j_1(s,r,q)\omega} n^{-\frac{p-r}{2}} \lesssim n^{-\mu p}$ by direct computations. Hence, $E\|e_{ss}\|_p^p \lesssim n^{-\mu p}$ in each case, which is (33). Now, (24) follows from (26), (27), (29), and (33). The proof is done. □

Remark 3.1 We find easily from Theorem 2.1 and Theorem 3.1 that the nonlinear wavelet estimator converges faster than the linear one for $r \le p$. Moreover, the nonlinear estimator is adaptive, while the linear one is not.

Remark 3.2 This paper studies wavelet estimations of a density and its derivatives with Fourier-oscillating noises. The remaining problems include the optimality of the above estimations, numerical experiments as well as the corresponding regression problems. We shall investigate those problems in the future.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
All authors contributed equally and significantly in writing this article. All authors read and approved the final manuscript.

Acknowledgements
This paper is supported by the National Natural Science Foundation of China (No. 11271038).

Received: 24 March 2014 Accepted: 28 May 2014 Published: 08 Jun 2014

References

1. Walter, GG: Density estimation in the presence of noise. Stat. Probab. Lett. **41**, 237-246 (1999)
2. Pensky, M, Vidakovic, B: Adaptive wavelet estimator for nonparametric density deconvolution. Ann. Stat. **27**, 2033-2053 (1999)
3. Fan, J: On the optimal rates of convergence for nonparametric deconvolution problems. Ann. Stat. **19**, 1257-1272 (1991)
4. Fan, J, Koo, J: Wavelet deconvolution. IEEE Trans. Inf. Theory **48**, 734-747 (2002)
5. Lounici, K, Nickl, R: Global uniform risk bounds for wavelet deconvolution estimators. Ann. Stat. **39**, 201-231 (2011)
6. Li, R, Liu, Y: Wavelet optimal estimations for a density with some additive noises. Appl. Comput. Harmon. Anal. **36**, 416-433 (2014)
7. Sun, J, Morrison, H, Harding, P, Woodroofe, M: Density and mixture estimation from data with measurement errors. Technical report (2002)
8. Devroye, L: Consistent deconvolution in density estimation. Can. J. Stat. **17**, 235-239 (1989)
9. Hall, P, Meister, A: A ridge-parameter approach to deconvolution. Ann. Stat. **35**, 1535-1558 (2007)
10. Meister, A: Deconvolution from Fourier-oscillating error densities under decay and smoothness restrictions. Inverse Probl. **24**, 1-14 (2008)

11. Delaigle, A, Meister, A: Nonparametric function estimation under Fourier-oscillating noise. Stat. Sin. **21**, 1065-1092 (2011)
12. Donoho, DL, Johnstone, IM, Kerkyacharian, G, Picard, D: Density estimation by wavelet thresholding. Ann. Stat. **24**, 508-539 (1996)
13. Hernández, E, Weiss, G: A First Course on Wavelets. CRC Press, Boca Raton (1996)
14. Daubechies, I: Ten Lectures on Wavelets. SIAM, Philadelphia (1992)
15. Härdle, W, Kerkyacharian, G, Picard, D, Tsybakov, A: Wavelets, Approximation, and Statistical Applications. Lecture Notes in Statist. Springer, New York (1998)

10.1186/1029-242X-2014-236
Cite this article as: Guo and Liu: Wavelet estimations for densities and their derivatives with Fourier oscillating noises. *Journal of Inequalities and Applications* 2014, 2014:236