Hu et al. *Journal of Inequalities and Applications* (2015) 2015:208
DOI 10.1186/s13660-015-0732-2

Journal of Inequalities and Applications
a SpringerOpen Journal

RESEARCH

Open Access

L^p convergence rate of wavelet estimators for the bias and multiplicative censoring model

Lin Hu[1], Junlian Xu[2], Linshan Li[1*] and Huiying Wang[3]

*Correspondence: Sll308@163.com
[1]Department of Basic Courses, Beijing Union University, Beijing, 100101, P.R. China
Full list of author information is available at the end of the article

Abstract

Using wavelets methods, Abbaszadeh, Chesneau, Doosti studied the density estimation problem under bias and multiplicative censoring (Stat. Probab. Lett. 82:932-941, 2012), and obtains the convergence rate of wavelet estimators in L^2 norm for a density function in Besov space. This paper deals with L^p risk estimation with $1 \le p < \infty$ based on wavelet bases. Motivated by the work of Youming and Junlian (2014), we construct new estimators: a linear one and a nonlinear adaptive one; an upper bound of wavelet estimators on L^p risk for a density function in Besov space is provided, which generalizes Abbaszadeh *et al.*'s theorems. It turns out that the nonlinear adaptive estimator obtains faster rate of convergence than the linear one for $r < p$.

MSC: 62G07; 42C40; 62G20

Keywords: wavelets; multiplicative censoring; density estimation

1 Introduction and preliminary

1.1 Introductions

The density estimation plays important roles in both statistics and econometrics. This paper considers the density model under bias and multiplicative censoring, which were introduced by Abbaszadeh *et al.* [1]. Let $Z_1, Z_2, \dots, Z_n$ be independent and identically distributed (i.i.d.) random variables of

$$Z_i = U_i Y_i, \quad i = 1, \dots, n,$$

where $U_1, U_2, \dots, U_n$ are unobserved i.i.d. random variables with the common uniform distribution on $[0,1]$, $Y_1, Y_2, \dots, Y_n$ are also unobserved i.i.d. random variables and the density function f_Y is given by

$$f_Y(x) = \frac{\omega(x) f_X(x)}{\theta}, \quad x \in [0,1]. \tag{1.1}$$

Here, $\omega(x) > 0$ denotes a known weight function, $f_X(x)$ stands for an unknown density function of a random variable X and $\theta = E(\omega(X)) = \int_0^1 \omega(x) f_X(x)\,dx$ represents the unknown normalization constant (EX is the expectation of X). We suppose that U_i and Y_i are independent for each $i \in 1, 2, \ldots, n$. Our aim is to estimate f_X when only $Z_1, Z_2, \ldots, Z_n$ are observed.

In particular, when $\omega(x) = 1$, this model reduces to the classical density estimation problem under multiplicative censoring described by Vardi [2], which unifies several well-studied statistical problems, including non-parametric inference for renewal processes, certain non-parametric deconvolution problems and estimation of decreasing densities [2–4]. Many methods were proposed to deal with that problem including a series expansion method [4], the kernel method [5] and wavelet method [1, 6], *etc.* For the standard biased density model (1.1) (estimating f_X from $Y_1, Y_2, \ldots, Y_n$), we refer to [7–9]. However, the estimation of f_X from $Z_1, Z_2, \ldots, Z_n$ is a new statistical problem which has potential applications in statistics and econometrics. As far as we know, only Abbaszadeh *et al.* [1] dealt with that problem. By using wavelet method, they considered a convergence rate of estimators in L^2 norm over Besov space $B^s_{r,q}$.

It is well known that in many statistical models, the error of estimators is measured in L^p norm [10–12]. In this paper, we consider L^p $(1 < p < \infty)$ risk estimation in Besov space $B^s_{r,q}$ based on wavelet bases. We define a linear estimator and a nonlinear adaptive one motivated by Abbaszadeh *et al.* and Youming and Junlian's work. We prove that the nonlinear adaptive estimator achieves a faster rate of convergence than the linear one for $r < p$. Our results can be considered as an extension of Abbaszadeh *et al.*'s theorems from $p = 2$ to $p \in [1, +\infty)$.

Section 1.2 introduces some notations and classical results on wavelets and Besov spaces, which will be used in our discussions; the assumptions on the model and the main results are presented in Section 2. In order to prove our theorems, we show several lemmas in Section 3 and give the proofs in Section 4.

1.2 Some preparations

In recent years, the wavelet method has turned out to be effective for density estimation [1, 6, 11, 12] because of the time and frequency localization, being a fast algorithm in numerical computations. In this subsection, we introduce the wavelet basis of the real line $\mathbb{R}$ (not necessarily on the fixed interval $[0,1]$ as in [1]), which will be used in our discussions. Let $\varphi \in C_0^m(\mathbb{R})$ be an orthonormal scaling function with $m > s$. The corresponding wavelet function is denoted by ψ. It is well known that $\{\phi_{J,k}, \psi_{j,k}, j \geq J, k \in \mathbb{Z}\}$ constitutes an orthonormal basis of $L^2(\mathbb{R})$, where $f_{j,k}(x) := 2^{j/2} f(2^j x - k)$ as in wavelet analysis. Then, for each $f \in L^2(\mathbb{R})$,

$$f(x) = \sum_{k \in \mathbb{Z}} \alpha_{j,k} \phi_{J,k}(x) + \sum_{j \geq J} \sum_{k \in \mathbb{Z}} \beta_{j,k} \psi_{j,k}(x),$$

where $\alpha_{J,k} = \int f(x)\phi_{j,k}(x)\,dx$ and $\beta_{j,k} = \int f(x)\psi_{j,k}(x)\,dx$. Details on wavelet bases can be found in [13].

One of the advantages of wavelet bases is that they can characterize Besov spaces. Throughout the paper, we work within Besov space on a compact subset of the real line $\mathbb{R}$ (not necessarily on the fixed interval $[0,1]$ as [1]). To introduce those spaces, we need the well-known Sobolev spaces with integer exponents $W_p^n(\mathbb{R}) := \{f | f \in L^p(\mathbb{R}), f^{(n)} \in L^p(\mathbb{R})\}$ 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 $1 \leq p, q \leq \infty$

and $s = n + \alpha$ with $\alpha \in (0,1]$, the Besov spaces on $\mathbb{R}$ are defined by

$$B^s_{p,q}(\mathbb{R}) := \bigl\{f \in W^n_p(\mathbb{R}), \bigl\| t^{-\alpha}\omega^2_p\bigl(f^{(n)}, t\bigr)\bigr\|^*_q < \infty\bigr\},$$

where $\omega^2_p(f,t) := \sup_{|h|\le t} \|f(\cdot + 2h) - 2f(\cdot + h) + f(\cdot)\|_p$ denotes the smoothness modulus of f and

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

The associated norm $\|f\|_{B^s_{p,q}} := \|f\|_p + \|t^{-\alpha}\omega^2_p(f^{(n)}, t)\|^*_q$. It should be pointed out that Besov spaces contain Hölder spaces and Sobolev spaces with non-integer exponents for a particular choice of s, p, and r [13].

The following theorems are fundamental in our discussions.

Theorem 1.1 ([14]) *Let $f \in L^r(\mathbb{R})$ $(1 \le r \le \infty)$, $\alpha_{J,k} = \int f(x)\phi_{J,k}(x)\,dx$, $\beta_{j,k} = \int f(x)\psi_{j,k}(x)\,dx$. Then the following assertions are equivalent.*

(i) $f \in B^s_{r,q}(\mathbb{R})$, $s > 0$, $1 \le q \le \infty$;

(ii) $\{2^{js}\|P_jf - f\|_r\}_{j\ge 0} \in l^q$, *where* $P_j(x) := \sum_{k\in\mathbb{Z}} \alpha_{j,k}\phi_{J,k}(x)$ *is the projection operator to* V_j;

(iii) $\|\alpha_{J,\cdot}\|_r + \|\{2^{j(s+1/2-1/r)}\|\beta_{j,\cdot}\|_r\}_{j\ge 0}\|_q < \infty$.

Theorem 1.2 ([14]) *Let ϕ be a scaling function or a wavelet with $\theta(\phi) := \sup_{x\in\mathbb{R}} |\phi(x-k)| < \infty$. Then*

$$\Bigl\|\sum_{k\in\mathbb{Z}} \lambda_k\phi_{j,k}\Bigr\|_p \sim 2^{j(\frac{1}{2}-\frac{1}{p})}\|\lambda\|_p$$

for $\lambda = \{\lambda_k\} \in l^p(\mathbb{Z})$ and $1 \le p \le \infty$, where

$$\|\lambda\|_p := \begin{cases} (\sum_{k\in\mathbb{Z}} |\lambda_k|^p)^{\frac{1}{p}}, & \textit{if } p < \infty, \\ \sup_{k\in\mathbb{Z}} |\lambda_k|, & \textit{if } q = \infty. \end{cases}$$

Here and after, $A \lesssim B$ denotes $A \le CB$ for some constant $C > 0$; $A \sim B$ stands for both $A \lesssim B$ and $B \lesssim A$. Clearly, Daubechies and Meyer's scaling and wavelet functions satisfy the conditions $\theta(\phi) < \infty$.

2 Main results

This section is devoted to the statement of our main results. To do that, we make the following assumptions as described in [1]:

(A1) The two density functions f_X and f_Y have the support $[0,1]$ and f_X belongs to the Besov ball $B^s_{r,q}(H)$ $(H > 0)$ defined as

$$B^s_{r,q}(H) := \bigl\{f \in B^s_{r,q}(\mathbb{R}), f \text{ is a probability density and } \|f\|_{B^s_{r,q}} \le H\bigr\}.$$

(A2) The density of Z_i is

$$f_Z(x) = \int_x^1 \frac{f_Y(y)}{y}\,dy.$$

(A3) There exists a constant $C > 0$ such that

$$\sup_{x\in[0,1]} f_X(x) \le C, \qquad \sup_{x\in[0,1]} f_Z(x) \le C.$$

(A4) There exist a constant $C > 0$ and $c > 0$ such that

$$\sup_{x\in[0,1]} \omega(x) \le C, \qquad \sup_{x\in[0,1]} \omega'(x) \le C, \qquad \inf_{x\in[0,1]} \omega(x) \ge c.$$

To introduce the wavelet estimator, we define the operator T by

$$T(h)(x) := \frac{h(x)\omega(x) + xh'(x)\omega(x) - xh(x)\omega'(x)}{\omega^2(x)}$$

for $h \in C^1(\mathbb{R})$, the function set of all differential functions on $\mathbb{R}$. Then the linear estimator is given as follows:

$$\hat{f}^{\mathrm{lin}}(x) := \sum_{k\in\wedge} \hat{\alpha}_{j_0,k}\phi_{j_0,k}(x), \tag{2.1}$$

where j_0 is chosen such that $2^{j_0} \sim n^{\frac{1}{2s+3}}$ and $\wedge := \{k \in \mathbb{Z}, \mathrm{supp} f_X \cap \mathrm{supp}\,\phi_{j_0,k} \neq \varnothing\}$.

To obtain a nonlinear estimator, we take j_0 and j_1 such that $2^{j_1} \sim \frac{n}{\ln n}$ and $2^{j_0} \sim n^{\frac{1}{2m+3}}$ with $m > s$. By definition

$$\hat{\alpha}_{j,k} = \frac{\hat{\theta}}{n}\sum_{i=1}^{n} T(\phi_{j,k})(Z_i), \qquad \hat{\beta}_{j,k} = \frac{\hat{\theta}}{n}\sum_{i=1}^{n} T(\psi_{j,k})(Z_i) \tag{2.2}$$

are the estimators of $\alpha_{j,k} = \int f(x)\phi_{j,k}(x)\,dx$ and $\beta_{j,k} = \int f(x)\psi_{j,k}(x)\,dx$, respectively, with

$$\hat{\theta} := \left[\frac{1}{n}\sum_{i=1}^{n} \frac{\omega(Z_i) - Z_i\omega'(Z_i)}{\omega^2(Z_i)}\right]^{-1}.$$

Then the nonlinear estimator is given by

$$\hat{f}^{\mathrm{non}}(x) := \sum_{k\in\wedge} \hat{\alpha}_{j_0,k}\phi_{j_0,k}(x) + \sum_{j=j_0}^{j_1}\sum_{k\in\wedge_j} \hat{\beta}_{j,k} 1_{\{|\hat{\beta}_{j,k}|>\lambda\}}\psi_{j,k}(x), \tag{2.3}$$

where $\wedge := \{k \in \mathbb{Z}, \mathrm{supp} f_X \cap \mathrm{supp}\,\phi_{j_0,k} \neq \varnothing\}$, $\wedge_j := \{k \in \mathbb{Z}, \mathrm{supp} f_X \cap \mathrm{supp}\,\psi_{j,k} \neq \varnothing\}$ and 1_D denotes the indicator function on the set D with $\lambda := T2^j\sqrt{\frac{\ln n}{n}}$.

Remark 2.1 From the definition of $\hat{f}_n^{\mathrm{non}}$, we find that the nonlinear estimator has the advantage of being adaptive, since it does not depend on the indices s, r, q, and H in its construction.

Remark 2.2 The definitions of $\hat{f}^{\mathrm{lin}}$ and $\hat{f}^{\mathrm{non}}$ are essentially same as in [1]. However, the selection of j_0 and j_1 is different from that in [1] and the wavelet functions are defined on the real line $\mathbb{R}$ not necessarily on $[0,1]$.

Then we have the following approximation result, which extends Abbaszadeh *et al.*'s theorems [1] from $p = 2$ to $p \in [1, +\infty)$.

Theorem 2.1 *Let* $f_X(x) \in B^s_{r,q}(H)$ $(s > \frac{1}{r}, r, q \geq 1)$ *and* $\hat{f}^{\mathrm{lin}}$ *be the estimator defined by* (2.1). *If* (A1)-(A4) *hold, then for each* $1 \leq p < \infty$, $s' = s - (1/r - 1/p)_+$, *and* $x_+ := \max(x, 0)$,

$$\sup_{f_X \in B^s_{r,q}(H)} E\left\|\hat{f}^{\mathrm{lin}}(x) - f_X(x)\right\|_p^p \lesssim n^{-\frac{s'p}{2s'+3}}.$$

Remark 2.3 The condition $s > \frac{1}{p}$ seems natural, since $B^s_{r,q}(\mathbb{R}) \subseteq C(\mathbb{R})$ for $sp > 1$, where $C(\mathbb{R})$ denotes the function set of all continuous functions on $\mathbb{R}$.

Remark 2.4 If $r \geq 2$ and $p = 2$, then Abbaszadeh *et al.*'s Theorem 4.1 [1] follows directly from our theorem, in this case $s' = s$. That is, Theorem 2.1 extends the corresponding theorem of [1] from $p = 2$ to $p \in [1, +\infty)$.

Remark 2.5 When $\omega(x) = 1$ and $\theta = E(\omega(X)) = 1$, the model reduces to the standard multiplicative censoring one considered by Abbaszadeh *et al.* [6]. In [6], they estimate the convergence rate of wavelet estimators in L_p norm for a density and its derivatives in Besov space. Our result is consistent with Theorem 4.1 [6] taken with $m = 0$.

Theorem 2.2 *Let* $f_X(x) \in B^s_{r,q}(H)$ $(\frac{1}{r} < s < m,\ r, q \geq 1)$ *and* $\hat{f}_n^{\mathrm{non}}$ *be the estimator given by* (2.3). *If* (A1)-(A4) *hold, then there exists* $C > 0$ *such that for each* $1 \leq p < \infty$ *and* $\alpha := \min\{\frac{s}{2s+3}, \frac{s-\frac{1}{r}+\frac{1}{p}}{2(s-\frac{1}{r})+3}\}$,

$$\sup_{f_X \in B^s_{r,q}(H)} E\left\|\hat{f}_n^{\mathrm{non}}(x) - f_X(x)\right\|_p^p \lesssim (\ln n)^p \left(\frac{\ln n}{n}\right)^{\alpha p}.$$

Remark 2.6 When $p = 2$ and $r \geq 2$ or $\{1 \leq r < 2, s > 3/r\}$, our result is the exactly same as Theorem 4.2 [1], ignoring the log factor. In this case, $\alpha = \frac{s}{2s+3}$. In other words, our theorem can be considered as an extension of Theorem 4.2 in [1].

Remark 2.7 When $\omega(x) = 1$ and $\theta = E(\omega(X)) = 1$, our result coincides with Theorem 4.2 [6] taken with $m = 0$, ignoring the log factor. In this case, the model reduces to the standard density estimation problem under multiplicative censoring.

Remark 2.8 When $r < p$, the nonlinear estimator attains a better rate of convergence than that of the linear one, due to $\frac{s'p}{2s'+3} < \alpha p$ because of $\frac{s'p}{2s'+3} < \frac{sp}{2s+3}$ and $\frac{s'p}{2s'+3} < \frac{(s-\frac{1}{r}+\frac{1}{p})p}{2(s-\frac{1}{r})+3}$. When $r \geq p$, the nonlinear estimator does the same rate of convergence to that of the linear one, *i.e.*, $n^{-\frac{sp}{2s+3}}$, ignoring the log factor. However, taking into account that the nonlinear estimator is adaptive, it is preferable to the linear one in the estimation of f_X.

3 Lemmas

We present several important lemmas in this section, which will be needed for the proofs of our main theorems. To show Lemma 3.1, we need Rosenthal's inequality [14].

Rosenthal's inequality Let $X_1, X_2, \ldots, X_n$ be independent random variables such that $E(X_i) = 0$ and $E|x_i|^p < \infty$ $(i = 1, 2, \ldots, n)$. Then

$$E\left|\sum_{i=1}^{n} X_i\right|^p \le \begin{cases} C_p[\sum_{i=1}^n E|X_i|^p + (\sum_{i=1}^n E|X_i|^2)^{\frac{1}{2}p}], & p \ge 2, \\ C_p(\sum_{i=1}^n E|X_i|^2)^{\frac{1}{2}p}, & 0 < p \le 2. \end{cases}$$

Lemma 3.1 *Let* $\hat{\alpha}_{j,k}$, $\hat{\beta}_{j,k}$ *be defined by* (2.2). *If* (A1)-(A4) *hold, then there exists a constant* $C > 0$ *such that*

$$E|\hat{\alpha}_{j,k} - \alpha_{j,k}|^p \le 2^{pj} n^{-\frac{1}{2}p} \quad \textit{and} \quad E|\hat{\beta}_{j,k} - \beta_{j,k}|^p \le 2^{pj} n^{-\frac{1}{2}p}$$

for $1 \le p < \infty$ *and* $2^j \le n$.

Remark 3.1 When $p = 4$, Lemma 3.1 reduces to Proposition 4.1 in [1].

Proof of Lemma 3.1 One only proves the first inequality, the second one is similar. Clearly,

$$\hat{\alpha}_{j,k} - \alpha_{j,k} = \frac{\hat{\theta}}{\theta} \frac{1}{n} \sum_{i=1}^{n} \left[\theta T(\phi_{j,k})(Z_i) - \alpha_{j,k}\right] + \alpha_{j,k} \hat{\theta} \left(\frac{1}{\theta} - \frac{1}{\hat{\theta}}\right).$$

By (A1) and (A3), $|\alpha_{j,k}| \le \int_0^1 f_X(x)|\phi_{j,k}(x)|\,dx \lesssim \int_0^1 |\phi_{j,k}(x)|\,dx \lesssim 1$. On the other hand, $c \le |\theta| = |E(\omega(X))| = |\int_0^1 \omega(x) f_X(x)\,dx| \le C$ and $|\hat{\theta}| \lesssim 1$ thanks to (A4). Therefore,

$$E|\hat{\alpha}_{j,k} - \alpha_{j,k}|^p \le E\left|\frac{1}{n} \sum_{i=1}^{n} \left[\theta T(\phi_{j,k})(Z_i) - \alpha_{j,k}\right]\right|^p + E\left|\frac{1}{\hat{\theta}} - \frac{1}{\theta}\right|^p := T_1 + T_2.$$

To estimate T_1, one defines $\xi_i := \theta T(\phi_{j,k})(Z_i) - \alpha_{j,k}$. Then $T_1 = E|\frac{1}{n}\sum_{i=1}^n \xi_i|^p$ and $E(\xi_i) = 0$ by Lemma 4.2 in [1]. By the definition of the operator T,

$$E\left|\theta T(\phi_{j,k})(Z_i)\right|^p = |\theta|^p \int_0^1 \frac{|\phi_{j,k}(x)\omega(x) + x(\phi_{j,k})'(x)\omega(x) - x\phi_{j,k}(x)\omega'(x)|^p}{\omega^2(x)} f_Z(x)\,dx.$$

Moveover, $E|\theta T(\phi_{j,k})(Z_i)|^p \lesssim \int_0^1 (|\phi_{j,k}(x)|^p + |\phi'_{j,k}(x)|^p)\,dx \lesssim 2^{j(\frac{3}{2}p-1)}$ due to (A3) and (A4). Note that $|\alpha_{j,k}|^p \lesssim 1$. Then for $p \ge 1$,

$$E|\xi_i|^p \lesssim E\left|\theta T(\phi_{j,k})(Z_i)\right|^p + |\alpha_{j,k}|^p \lesssim 2^{j(\frac{3}{2}p-1)}. \tag{3.1}$$

This with Rosental's inequality leads to

$$E\left|\frac{1}{n}\sum_{i=1}^{n} \xi_i\right|^p \lesssim n^{-p} \max\left\{nE|\xi_i|^p, \left(nE|\xi_i|^2\right)^{\frac{1}{2}p}\right\} \lesssim n^{-p} \max\left\{n2^{j(\frac{3}{2}p-1)}, n^{\frac{1}{2}p} 2^{pj}\right\}.$$

Using the assumption $2^j \le n$, one obtains $T_1 = E|\frac{1}{n}\sum_{i=1}^n \xi_i|^p \lesssim 2^{pj} n^{-p/2}$. To end the proof, one needs only to show

$$T_2 = E\left|\frac{1}{\hat{\theta}} - \frac{1}{\theta}\right|^p \lesssim 2^{pj} n^{-\frac{1}{2}p}. \tag{3.2}$$

Denote

$$\eta_i := \frac{\omega(Z_i) - Z_i\omega'(Z_i)}{\omega^2(Z_i)} - \frac{1}{\theta} \quad (i = 1, 2, \ldots, n).$$

Then $E|\eta_i|^p \le C$ due to (A4). By Theorem 4.1 [1], $\eta_1, \eta_2, \ldots, \eta_n$ are i.i.d. and $E(\eta_i) = 0$. Then it follows from Rosental's inequality that

$$T_2 = n^{-p} E\left|\sum_{i=1}^n \eta_i\right|^p \lesssim n^{-p}\left(\sum_{i=1}^n E|\eta_i|^2\right)^{\frac{1}{2}p} \lesssim n^{-\frac{1}{2}p} \le 2^{pj} n^{-\frac{1}{2}p}$$

for $1 \le p \le 2$ and

$$T_2 = n^{-p} E\left|\sum_{i=1}^n \eta_i\right|^p \lesssim n^{-p}\left[\sum_{i=1}^n E|\eta_i|^p + \left(\sum_{i=1}^n E|\eta_i|^2\right)^{\frac{1}{2}p}\right] \lesssim n^{-(p-1)} + n^{-\frac{1}{2}p} \lesssim n^{-\frac{1}{2}p}$$

for $p \ge 2$, which proves the desired conclusion (3.2). This finishes the proof of Lemma 3.1. □

The well-known Bernstein inequality [14] is needed in order to prove Lemma 3.2.

Bernstein's inequality Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables with $E(X_i) = 0$, $\|X_i\|_\infty \le M$. Then, for each $\gamma > 0$,

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^n X_i\right| > \gamma\right\} \le 2\exp\left(-\frac{n\gamma^2}{2[E(X_i^2) + \|X\|_\infty \gamma/3]}\right).$$

Lemma 3.2 *Let* $\hat{\beta}_{j,k}$ *be defined by* (2.2). *If* $2^j \le n/\ln n$ *and* (A1)-(A4) *hold, then for each* $\varepsilon > 0$, *there exists* $T > 0$ *such that*

$$P\left\{|\hat{\beta}_{j,k} - \beta_{j,k}| > \frac{T}{2} 2^j \sqrt{\frac{\ln n}{n}}\right\} \lesssim 2^{-\varepsilon j}. \tag{3.3}$$

Proof By the definition of $\hat{\beta}_{j,k}$,

$$\hat{\beta}_{j,k} - \beta_{j,k} = \frac{\hat{\theta}}{n}\sum_{i=1}^n T(\psi_{j,k})(Z_i) - \beta_{j,k} = \frac{\hat{\theta}}{\theta}\frac{1}{n}\sum_{i=1}^n \left[\theta T(\psi_{j,k})(Z_i) - \beta_{j,k}\right] + \hat{\theta}\beta_{j,k}\left(\frac{1}{\theta} - \frac{1}{\hat{\theta}}\right).$$

Then

$$|\hat{\beta}_{j,k} - \beta_{j,k}| \le \left|\frac{\hat{\theta}}{\theta}\right|\left|\frac{1}{n}\sum_{i=1}^n \left[\theta T(\psi_{j,k})(Z_i) - \beta_{j,k}\right]\right| + |\theta||\beta_{j,k}|\left|\frac{1}{\hat{\theta}} - \frac{1}{\theta}\right|. \tag{3.4}$$

The proof of Lemma 3.1 shows $|\hat{\theta}| \le C$ and $c \le |\theta| \le C$. By (A1) and (A3), $|\beta_{j,k}| = |\int_0^1 f_X(x)\psi_{j,k}(x)\,dx| \lesssim \int_0^1 |\psi_{j,k}(x)|\,dx \lesssim 1$. Then (3.4) reduces to

$$|\hat{\beta}_{j,k} - \beta_{j,k}| \le C\left|\frac{1}{n}\sum_{i=1}^n \left[\theta T(\psi_{j,k})(Z_i) - \beta_{j,k}\right]\right| + C\left|\frac{1}{\hat{\theta}} - \frac{1}{\theta}\right|.$$

Furthermore,

$$P\left\{|\hat{\beta}_{j,k}-\beta_{j,k}|>\frac{T}{2}2^{j}\sqrt{\frac{\ln n}{n}}\right\}\le I_1+I_2, \tag{3.5}$$

where

$$I_1:=P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\left[\theta T(\psi_{j,k})(Z_i)-\beta_{j,k}\right]\right|>\frac{T}{4C}2^{j}\sqrt{\frac{\ln n}{n}}\right\}$$

and

$$I_2:=P\left\{\left|\frac{1}{\hat{\theta}}-\frac{1}{\theta}\right|>\frac{T}{4C}2^{j}\sqrt{\frac{\ln n}{n}}\right\}.$$

By (3.5), one needs only to prove

$$|I_i|\lesssim 2^{-\varepsilon j}\quad (i=1,2) \tag{3.6}$$

for the desired conclusion (3.3).

To estimate I_1, one defines $U_i=\theta T(\psi_{j,k})(Z_i)-\beta_{j,k}$. Then

$$\left|T(\psi_{j,k})(x)\right|=\omega^{-2}(x)\left|\psi_{j,k}(x)\omega(x)+x(\psi_{j,k})'(x)\omega(x)-x\psi_{j,k}(x)\omega'(x)\right|$$

by the definition of the operator T. Using (A4), $|T(\psi_{j,k})(x)|\lesssim|\psi_{j,k}(x)|+|(\psi_{j,k})'(x)|\lesssim 2^{\frac{3}{2}j}$ for $x\in[0,1]$ and

$$|U_i|=\left|\theta T(\psi_{j,k})(Z_i)-\beta_{j,k}\right|\lesssim|\theta|\left|T(\psi_{j,k})(Z_i)\right|+|\beta_{j,k}|\lesssim 2^{\frac{3}{2}j}$$

due to $|\theta|\lesssim 1$ and $|\beta_{j,k}|\lesssim 1$. Moreover, using (3.1) with ψ instead of ϕ and $p=2$, one obtains $E|U_i|^2\lesssim 2^{2j}$. Because $U_1,U_2,\dots,U_n$ are i.i.d. and $E(U_i)=0$ $(i=1,2,\dots,n)$ thanks to Lemma 4.2 [1], Bernstein's inequality tells us that

$$I_1:=P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}U_i\right|>\frac{T}{4C}2^{j}\sqrt{\frac{\ln n}{n}}\right\}\le 2\exp\left(-\frac{n\gamma^2}{2[E(U_i^2)+\frac{\gamma}{3}\|U\|_\infty]}\right)$$

with $\gamma=\frac{T}{4C}2^{j}\sqrt{\frac{\ln n}{n}}$. It is easy to see that $\frac{n\gamma^2}{2[E(U_i^2)+\frac{\gamma}{3}\|U\|_\infty]}\ge\frac{n\frac{T^2}{16C^2}2^{2j}\frac{\ln n}{n}}{2(2^{2j}+\frac{T}{12C}2^{\frac{5}{2}j}\sqrt{\frac{\ln n}{n}})}\ge\frac{T^2\ln n}{32C^2(1+\frac{T}{12C})}$ because of $2^{\frac{j}{2}}\sqrt{\frac{\ln n}{n}}\le 1$ by the assumption $2^{j}\le\frac{n}{\ln n}$. Note that $\ln n>j\ln 2$ due to $n\ge 2^{j}\ln n>2^{j}$. Hence, $\frac{n\gamma^2}{2[E(U_i^2)+\frac{\gamma}{3}\|U\|_\infty]}\ge\frac{T^2\ln 2}{32C^2(1+\frac{T}{12C})}j$. One chooses $T>0$ such that $\frac{T^2\ln 2}{32C^3(1+\frac{T}{12C})}>\varepsilon$. Then $\frac{T^2\ln 2}{32C^2(1+\frac{T}{12C})}>\varepsilon$ due to $C\ge 1$ and $I_1\lesssim\exp(-\frac{T^2\ln 2}{32C^2(1+\frac{T}{12C})}j)\lesssim 2^{-\varepsilon j}$, which shows (3.6) for $i=1$.

Next, one estimates I_2: Define $W_i:=\frac{\omega(Z_i)-Z_i\omega'(Z_i)}{\omega^2(Z_i)}-\frac{1}{\theta}$. Then $W_1,W_2,\dots,W_n$ are i.i.d. and $E(W_i)=0$. On the other hand, (A3) implies $|W_i|\le C$ and $E|W_i|^2\le C$. Applying Bernstein's inequality, one obtains

$$I_2=P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}W_i\right|>\frac{T}{4C}2^{j}\sqrt{\frac{\ln n}{n}}\right\}\le 2\exp\left(-\frac{n\gamma^2}{2[E(W_i^2)+\frac{\gamma}{3}\|W\|_\infty]}\right)$$

with $\gamma = \frac{T}{4C}2^j\sqrt{\frac{\ln n}{n}}$. Note that $\frac{n\gamma^2}{2[E(W_i^2)+\frac{\gamma}{3}\|W\|_\infty]} \geq \frac{n\frac{T^2}{16C^2}2^{2j}\frac{\ln n}{n}}{2(C+\frac{TC}{12C}2^j\sqrt{\frac{\ln n}{n}})} \geq \frac{\frac{T^2}{16C^2}2^j\ln n}{2C(1+\frac{T}{12C})} \geq \frac{T^2\ln 2}{32C^3(1+\frac{T}{12C})}j$ because of $\sqrt{\frac{\ln n}{n}} \leq 1$ and $\ln n > j\ln 2$. The desired conclusion (3.6) ($i = 2$) follows by taking $T > 0$ such that $\frac{T^2\ln 2}{32C^3(1+\frac{T}{12C})} > \varepsilon$. This completes the proof of Lemma 3.2. □

4 Proofs

This section is devoted to the proof of Theorems 2.1 and 2.2, based on the knowledge of Section 3. We begin with the proof of Theorem 2.1.

Proof of Theorem 2.1 Clearly, $\hat{f}_n^{\mathrm{lin}} - f^X = (\hat{f}_n^{\mathrm{lin}} - P_{j_0}f_X) + (P_{j_0}f_X - f_X)$ and

$$E\|\hat{f}_n^{\mathrm{lin}} - f_X\|_p^p \leq \|P_{j_0}f_X - f_X\|_p^p + E\|\hat{f}_n^{\mathrm{lin}} - P_{j_0}f_X\|_p^p. \tag{4.1}$$

It follows from the proof of Theorem 4.1 [15] that

$$\|P_{j_0}f_X - f_X\|_p^p \lesssim 2^{-j_0ps'} \lesssim n^{-\frac{ps'}{2s'+3}} \tag{4.2}$$

due to $2^{j_0} \sim n^{\frac{1}{2s'+3}}$. By (4.1) and (4.2), it is sufficient to show

$$E\|\hat{f}_n^{\mathrm{lin}} - P_{j_0}f_X\|_p^p \lesssim n^{-\frac{ps'}{2s'+3}} \tag{4.3}$$

for the conclusion of Theorem 2.1. By the definition of $\hat{f}_n^{\mathrm{lin}}$, $\hat{f}_n^{\mathrm{lin}} - P_{j_0}f_X = \sum_{k\in\wedge}(\hat{\alpha}_{j_0,k} - \alpha_{j_0,k})\phi_{j_0,k}$. Then $\|\hat{f}_n^{\mathrm{lin}} - P_{j_0}f_X\|_p^p \lesssim 2^{j_0(\frac{1}{2}p-1)}\sum_{k\in\wedge}|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p$ thanks to Theorem 1.2. This with Lemma 3.1 and the choice of j_0 leads to

$$E\|\hat{f}_n^{\mathrm{lin}} - P_{j_0}f^X\|_p^p \lesssim 2^{\frac{p}{2}j_0}E|\hat{\alpha}_{j_0,k} - \alpha_{j_0,k}|^p \lesssim 2^{\frac{3}{2}pj_0}n^{-\frac{1}{2}p} \lesssim n^{-\frac{ps'}{2s'+3}},$$

which is the desired conclusion (4.3). This finishes the proof of Theorem 2.1. □

Next, we prove Theorem 2.2.

Proof of Theorem 2.2 It is sufficient to prove the case $r \leq p$. In fact, when $r > p$, $\hat{f}_n^{\mathrm{non}}$ has compact support because of ϕ, ψ, and f having the same property. Then

$$E\|\hat{f}_n^{\mathrm{non}}(x) - f_X(x)\|_p^p \lesssim \left(E\|\hat{f}_n^{\mathrm{non}}(x) - f_X(x)\|_r^r\right)^{\frac{p}{r}}$$

using the Hölder inequality. For $f_X \in B_{r,q}^s(H)$, using Theorem 2.2 for the case $r = p$, one has

$$\sup_{f_X\in B_{r,q}^s(H)} E\|\hat{f}_n^{\mathrm{non}}(x) - f_X(x)\|_r^r \lesssim (\ln n)^r\left(\frac{\ln n}{n}\right)^{\alpha r}$$

and

$$\sup_{f_X\in B_{r,q}^s(H)} E\|\hat{f}_n^{\mathrm{non}}(x) - f_X(x)\|_p^p \lesssim (\ln n)^p\left(\frac{\ln n}{n}\right)^{\alpha p}.$$

Now, one estimates the case $r \le p$. By the definition of $\hat{f}_n^{\mathrm{non}}$,

$$\hat{f}_n^{\mathrm{non}} - f = \left(\hat{f}_n^{\mathrm{lin}} - P_{j_0} f\right) + (P_{j_1+1} f - f) + \sum_{j=j_0}^{j_1} \sum_{k\in\wedge_j} (\hat{\beta}_{j,k} 1_{\{|\hat{\beta}_{j,k}|>\lambda\}} - \beta_{j,k})\psi_{j,k}.$$

Then

$$E\left\|\hat{f}_n^{\mathrm{non}} - f\right\|_p^p \lesssim E\left\|\hat{f}_n^{\mathrm{lin}} - P_{j_0} f\right\|_p^p + \|P_{j_1+1} f - f\|_p^p + E\left\|\sum_{j=j_0}^{j_1} \sum_{k\in\wedge_j} (\hat{\beta}_{j,k} 1_{\{|\hat{\beta}_{j,k}|>\lambda\}} - \beta_{j,k})\psi_{j,k}\right\|_p^p. \tag{4.4}$$

From the proof of Theorem 2.1, one knows

$$\|P_{j_1+1} f - f\|_p^p \lesssim 2^{-j_1 s' p}, \qquad E\left\|\hat{f}_n^{\mathrm{lin}} - P_{j_0} f\right\|_p^p \lesssim 2^{\frac{3}{2} p j_0} n^{-\frac{p}{2}}.$$

Note that $2^{j_0} \sim n^{\frac{1}{2m+3}}$, $2^{j_1} \sim \frac{n}{\ln n}$, and $\alpha = \min\{\frac{s}{2s+3}, \frac{s-\frac{1}{r}+\frac{1}{p}}{2(s-\frac{1}{r})+3}\} \le s - \frac{1}{r} + \frac{1}{p} = s'$ thanks to $s > \frac{1}{r}$. Then

$$\|P_{j_1+1} f - f\|_p^p \lesssim \left(\frac{\ln n}{n}\right)^{\alpha p}, \qquad E\left\|\hat{f}_n^{\mathrm{lin}} - P_{j_0} f\right\|_p^p \lesssim \left(\frac{\ln n}{n}\right)^{\alpha p}. \tag{4.5}$$

To estimate $E\| \sum_{j=j_0}^{j_1} \sum_{k\in\wedge_j} (\hat{\beta}_{j,k} 1_{\{|\hat{\beta}_{j,k}|>\lambda\}} - \beta_{j,k})\psi_{j,k}\|_p^p$, one defines

$$\sum_{j=j_0}^{j_1} \sum_{k\in\wedge_j} (\hat{\beta}_{j,k} 1_{\{|\hat{\beta}_{j,k}|>\lambda\}} - \beta_{j,k})\psi_{j,k}(x) := T_1 + T_2 + T_3 + T_4,$$

where

$$T_1 := \sum_{j=j_0}^{j_1} \sum_{k\in\wedge_j} (\hat{\beta}_{j,k} - \beta_{j,k})\psi_{j,k}(x) 1_{\{|\hat{\beta}_{j,k}|>\lambda, |\beta_{j,k}|<\lambda/2\}},$$

$$T_2 := \sum_{j=j_0}^{j_1} \sum_{k\in\wedge_j} (\hat{\beta}_{j,k} - \beta_{j,k})\psi_{j,k}(x) 1_{\{|\hat{\beta}_{j,k}|>\lambda, |\beta_{j,k}|\ge\lambda/2\}},$$

$$T_3 := \sum_{j=j_0}^{j_1} \sum_{k\in\wedge_j} \beta_{j,k}\psi_{j,k}(x) 1_{\{|\hat{\beta}_{j,k}|\le\lambda, |\beta_{j,k}|>2\lambda\}},$$

$$T_4 := \sum_{j=j_0}^{j_1} \sum_{k\in\wedge_j} \beta_{j,k}\psi_{j,k}(x) 1_{\{|\hat{\beta}_{j,k}|\le\lambda, |\beta_{j,k}|\le 2\lambda\}}.$$

Then $E\| \sum_{j=j_0}^{j_1} \sum_{k\in\wedge_j} (\hat{\beta}_{j,k} 1_{\{|\hat{\beta}_{j,k}|>\lambda\}} - \beta_{j,k})\psi_{j,k}\|_p^p \lesssim \sum_{i=1}^4 E\|T_i\|_p^p$. By (4.4) and (4.5), it is sufficient to show

$$E\|T_i\|_p^p \lesssim (\ln n)^p \left(\frac{\ln n}{n}\right)^{\alpha p} \quad (i = 1, 2, 3, 4) \tag{4.6}$$

for the conclusion of Theorem 2.2.

To prove (4.6) for $i = 1$, applying Theorem 1.2, one has

$$E\|T_1\|_p^p \lesssim (j_1 - j_0 + 1)^{p-1} \sum_{j=j_0}^{j_1} 2^{j(\frac{p}{2}-1)} \sum_{k\in\wedge_j} E\big[|\hat{\beta}_{j,k} - \beta_{j,k}|^p 1_{\{|\hat{\beta}_{j,k}-\beta_{j,k}|\geq\lambda/2\}}\big]$$

due to the fact that $\{|\hat{\beta}_{j,k}| > \lambda, |\beta_{j,k}| < \lambda/2\} \subseteq \{|\hat{\beta}_{j,k} - \beta_{j,k}| \geq \lambda/2\}$. By the Hölder inequality,

$$\begin{aligned} E\|T_1\|_p^p &\lesssim (j_1 - j_0 + 1)^{p-1} \sum_{j=j_0}^{j_1} 2^{j(\frac{p}{2}-1)} \sum_{k\in\wedge_j} \big(E|\hat{\beta}_{j,k} - \beta_{j,k}|^{2p}\big)^{\frac{1}{2}} \big[E(1_{\{|\hat{\beta}_{j,k}-\beta_{j,k}|\geq\lambda/2\}})\big]^{\frac{1}{2}} \\ &\lesssim (j_1 - j_0 + 1)^{p-1} \sum_{j=j_0}^{j_1} 2^{j(\frac{p}{2}-1)} \sum_{k\in\wedge_j} \big(E|\hat{\beta}_{j,k} - \beta_{j,k}|^{2p}\big)^{\frac{1}{2}} \big[P\big(|\hat{\beta}_{j,k} - \beta_{j,k}| \geq \lambda/2\big)\big]^{\frac{1}{2}}. \end{aligned}$$

This with Lemma 3.1 and Lemma 3.2 leads to

$$E\|T_1\|_p^p \lesssim (j_1 - j_0 + 1)^{p-1} n^{-\frac{p}{2}} \sum_{j=j_0}^{j_1} 2^{\frac{3}{2}pj} 2^{-\frac{1}{2}\varepsilon j} \lesssim (\ln n)^{p-1} n^{-\frac{1}{2}p} \sum_{j=j_0}^{j_1} 2^{(\frac{3}{2}p-\frac{1}{2}\varepsilon)j}$$

thanks to $j_1 - j_0 \sim \ln n$ by the choice of j_0 and j_1. Take ε such that $\varepsilon > 3p$. Then $E\|T_1\|_p^p \lesssim (\ln n)^{p-1} n^{-\frac{p}{2}} 2^{\frac{3}{2}pj_0} \lesssim (\ln n)^{p-1} n^{-\frac{ps}{2s+3}} \lesssim (\ln n)^p n^{-\alpha p}$ due to the choice of j_0 and $\alpha \leq \frac{s}{2s+3}$. That is, (4.6) holds for $i = 1$.

To show (4.6) for $i = 3$, one uses the fact that $\{|\hat{\beta}_{j,k}| \leq \lambda, |\beta_{j,k}| > 2\lambda\} \subseteq \{|\hat{\beta}_{j,k} - \beta_{j,k}| \geq \lambda/2\}$. Hence,

$$E\|T_3\|_p^p \lesssim (j_1 - j_0 + 1)^{p-1} \sum_{j=j_0}^{j_1} 2^{j(\frac{p}{2}-1)} \sum_{k\in\wedge_j} E\big[|\beta_{j,k}|^p 1_{\{|\hat{\beta}_{j,k}-\beta_{j,k}|>\lambda/2\}}\big]$$

thanks to Theorem 1.2. When $|\hat{\beta}_{j,k}| \leq \lambda < |\beta_{j,k}|/2$, $|\hat{\beta}_{j,k} - \beta_{j,k}| \geq |\beta_{j,k}| - |\hat{\beta}_{j,k}| > |\beta_{j,k}|/2 > \lambda$. Then

$$E\|T_3\|_p^p \lesssim (j_1 - j_0 + 1)^{p-1} \sum_{j=j_0}^{j_1} 2^{j(\frac{p}{2}-1)} \sum_{k\in\wedge_j} E\big[|\hat{\beta}_{j,k} - \beta_{j,k}|^p 1_{\{|\hat{\beta}_{j,k}-\beta_{j,k}|>\lambda/2\}}\big]$$

due to Theorem 1.2. The same arguments as above shows $E\|T_3\|_p^p \lesssim (\ln n)^p n^{-\alpha p}$, which is the desired conclusion (4.6) with $i = 3$.

In order to estimate $E\|T_2\|_p^p$ and $E\|T_4\|_p^p$, one defines

$$2^{j_0^*} \sim \left(\frac{n}{\ln n}\right)^{\frac{1-2\alpha}{3}}, \qquad 2^{j_1^*} \sim \left(\frac{n}{\ln n}\right)^{\frac{\alpha}{s-\frac{1}{r}+\frac{1}{p}}}.$$

Recall that $2^{j_0} \sim n^{\frac{1}{2m+3}}$, $2^{j_1} \sim \frac{n}{\ln n}$, and $\alpha := \min\{\frac{s}{2s+3}, \frac{s-\frac{1}{r}+\frac{1}{p}}{2(s-\frac{1}{r})+3}\}$. Then $\frac{1-2\alpha}{3} \geq \frac{1}{2s+3} > \frac{1}{2m+3}$ and $\frac{\alpha}{s-\frac{1}{r}+\frac{1}{p}} \leq \frac{1}{2(s-\frac{1}{r})+3} \leq 1$. Hence, $2^{j_0} \leq 2^{j_0^*}$ and $2^{j_1^*} \leq 2^{j_1}$. Moreover, a simple computation shows $\frac{1-2\alpha}{3} \leq \frac{\alpha}{s-\frac{1}{r}+\frac{1}{p}}$, which implies $2^{j_0^*} \leq 2^{j_1^*}$.

One estimates $E\|T_2\|_p^p$ by dividing T_2 into

$$T_2 = \sum_{j=j_0}^{j_1}\sum_{k\in\wedge_j}(\hat{\beta}_{j,k}-\beta_{j,k})\psi_{j,k}(x)1_{\{|\hat{\beta}_{j,k}|>\lambda,|\beta_{j,k}|\geq\lambda/2\}} = \sum_{j=j_0}^{j_0^*} + \sum_{j=j_0^*+1}^{j_1} =: t_1 + t_2. \tag{4.7}$$

Then $E\|t_1\|_p^p \lesssim (j_1-j_0+1)^{p-1}\sum_{j=j_0}^{j_0^*}2^{j(\frac{1}{2}p-1)}\sum_{k\in\wedge_j}E|\hat{\beta}_{j,k}-\beta_{j,k}|^p$ due to Theorem 1.2. This with Lemma 3.1 and the definition of j_0^* leads to

$$E\|t_1\|_p^p \lesssim (j_1-j_0+1)^{p-1}\sum_{j=j_0}^{j_0^*}2^{\frac{3}{2}pj}n^{-\frac{1}{2}p} \lesssim (\ln n)^{p-1}n^{-\frac{1}{2}p}2^{\frac{3}{2}pj_0^*} \lesssim (\ln n)^{\frac{1}{2}p-1}\left(\frac{\ln n}{n}\right)^{\alpha p}. \tag{4.8}$$

To estimate $E\|t_2\|_p^p$, one observes that $1_{\{|\hat{\beta}_{j,k}|>\lambda,|\beta_{j,k}|\geq\lambda/2\}} \leq 1_{\{|\beta_{j,k}|\geq\lambda/2\}} \leq (\frac{|\beta_{j,k}|}{\lambda/2})^r$. Then it follows from Theorem 1.2 that

$$E\|t_2\|_p^p \lesssim (j_1-j_0+1)^{p-1}\sum_{j=j_0^*+1}^{j_1}2^{j(\frac{1}{2}p-1)}\sum_{k\in\wedge_j}E|\hat{\beta}_{j,k}-\beta_{j,k}|^p\left(\frac{|\beta_{j,k}|}{\lambda/2}\right)^r. \tag{4.9}$$

By Lemma 3.1, $E|\hat{\beta}_{j,k}-\beta_{j,k}|^p \leq n^{-\frac{p}{2}}2^{pj}$. On the other hand, $\|\beta_{j,\cdot}\|_r \leq 2^{-j(s+\frac{1}{2}-\frac{1}{r})}$ for $f\in B^s_{r,q}(H)$ due to Theorem 1.1. Then (4.9) reduces to

$$E\|t_2\|_p^p \lesssim (j_1-j_0+1)^{p-1}n^{-\frac{p}{2}}\sum_{j=j_0^*+1}^{j_1}2^{-j(sr+\frac{1}{2}r-\frac{3}{2}p)}\lambda^{-r}.$$

Substituting $\lambda = T2^j\sqrt{\frac{\ln n}{n}}$ into the above inequality, one gets

$$E\|t_2\|_p^p \lesssim (\ln n)^{p-\frac{1}{2}r-1}n^{\frac{1}{2}(r-p)}\sum_{j=j_0^*+1}^{j_1}2^{-j(sr+\frac{3}{2}r-\frac{3}{2}p)}$$

due to $j_1-j_0\sim\ln n$. Denote $\theta := sr+\frac{3}{2}r-\frac{3}{2}p$. When $\theta>0$, $r>\frac{3p}{2s+3}$, and

$$E\|t_2\|_p^p \lesssim (\ln n)^{p-\frac{1}{2}r-1}n^{\frac{1}{2}(r-p)}2^{-j_0^*(sr+\frac{3}{2}r-\frac{3}{2}p)} \lesssim (\ln n)^{\frac{1}{2}p-1}\left(\frac{\ln n}{n}\right)^{\alpha p} \tag{4.10}$$

thanks to the definition of $2^{j_0^*}$.

Moreover, (4.10) also holds for $\theta\leq 0$. In fact, the same analysis as (4.9) produces

$$E\|t_2\|_p^p \lesssim (j_1-j_0+1)^{p-1}\sum_{j=j_0^*+1}^{j_1}2^{j(\frac{1}{2}p-1)}\sum_{k\in\wedge_j}E|\hat{\beta}_{j,k}-\beta_{j,k}|^p\left(\frac{|\beta_{j,k}|}{\lambda/2}\right)^{r_1},$$

where $r_1 := (1-2\alpha)p$. When $\theta\leq 0$, $\alpha = \frac{s-\frac{1}{r}+\frac{1}{p}}{2(s-\frac{1}{r})+3} \leq \frac{s}{2s+3}$, and $r\leq\frac{3p}{2s+3}\leq(1-2\alpha)p = r_1$. Then $\|\beta_{j,\cdot}\|_{r_1} \leq \|\beta_{j,\cdot}\|_r \leq 2^{-j(s-\frac{1}{r}+\frac{1}{2})}$ for $f\in B^s_{r,q}(H)$ thanks to Theorem 1.1. Using Lemma 3.1 and the definition of λ, one has

$$E\|t_2\|_p^p \lesssim (\ln n)^{p-1-\frac{1}{2}r_1} n^{\frac{1}{2}(r_1-p)} \sum_{j=j_0^*+1}^{j_1} 2^{j[\frac{3}{2}p-1-(s-\frac{1}{r}+\frac{3}{2})r_1]}.$$

Note that $\frac{3}{2}p-1-(s-\frac{1}{r}+\frac{3}{2})r_1 = 0$ because of $r_1 = (1-2\alpha)p$ and $\alpha = \frac{s-\frac{1}{r}+\frac{1}{p}}{2(s-\frac{1}{r})+3}$. Hence, $E\|t_2\|_p^p \lesssim (\ln n)^{p-\frac{1}{2}r_1} n^{\frac{1}{2}(r_1-p)} \lesssim (\ln n)^{\frac{1}{2}p}(\frac{\ln n}{n})^{\alpha p}$, which shows (4.10). Combing (4.7), (4.8), and (4.10), one obtains the desired conclusion (4.6) with $i = 2$.

To end the proof, it is sufficient to estimate $E\|T_4\|_p^p$. When $\theta > 0$, one splits T_4 into

$$T_4 = \sum_{j=j_0}^{j_1} \sum_{k\in\wedge_j} \beta_{j,k}\psi_{j,k}(x) 1_{\{|\hat{\beta}_{j,k}|\le\lambda, |\beta_{j,k}|\le 2\lambda\}} = \sum_{j=j_0}^{j_0^*} + \sum_{j=j_0^*+1}^{j_1} =: e_1 + e_2. \tag{4.11}$$

Since $|\beta_{j,k}| 1_{\{|\hat{\beta}_{j,k}|\le 2\lambda, |\beta_{j,k}|\le 2\lambda\}} \le 2|\lambda|$, $E\|e_1\|_p^p \lesssim (j_1 - j_0 + 1)^{p-1} \sum_{j=j_0}^{j_0^*} 2^{j(\frac{1}{2}p-1)} 2^j |\lambda|^p$ due to Theorem 1.2. Note that $\lambda = T2^j\sqrt{\frac{\ln n}{n}}$, $2^{j_0^*} \sim (\frac{n}{\ln n})^{\frac{1-2\alpha}{3}}$, and $\alpha = \frac{s}{2s+3}$ when $\theta > 0$. Then

$$E\|e_1\|_p^p \lesssim (\ln n)^{\frac{3}{2}p-1} n^{-\frac{p}{2}} \sum_{j=j_0}^{j_0^*} 2^{\frac{3}{2}pj} \lesssim (\ln n)^{\frac{3}{2}p-1} n^{-\frac{p}{2}} 2^{\frac{3}{2}pj_0^*} \lesssim (\ln n)^{p-1} \left(\frac{\ln n}{n}\right)^{\alpha p}. \tag{4.12}$$

To estimate $E\|e_2\|_p^p$ with $e_2 = \sum_{j=j_0^*+1}^{j_1} \sum_{k\in\wedge_j} \beta_{j,k}\psi_{j,k}(x) 1_{\{|\hat{\beta}_{j,k}|\le\lambda, |\beta_{j,k}|\le 2\lambda\}}$, one uses the fact that $1_{\{|\hat{\beta}_{j,k}|\le\lambda, |\beta_{j,k}|\le 2\lambda\}} \le (\frac{2\lambda}{|\beta_{j,k}|})^{p-r}$ because of $r \le p$. Then

$$\begin{aligned} E\|e_2\|_p^p &\lesssim (j_1 - j_0 + 1)^{p-1} \sum_{j=j_0^*+1}^{j_1} 2^{j(\frac{1}{2}p-1)} \sum_{k\in\wedge_j} |\beta_{j,k}|^p \left(\frac{2\lambda}{|\beta_{j,k}|}\right)^{p-r} \\ &\lesssim (j_1 - j_0 + 1)^{p-1} \sum_{j=j_0^*+1}^{j_1} 2^{j(\frac{1}{2}p-1)} |\lambda|^{p-r} \sum_{k\in\wedge_j} |\beta_{j,k}|^r \end{aligned} \tag{4.13}$$

due to Theorem 1.2. By $f \in B_{r,q}^s(H)$ and Theorem 1.1, $\|\beta_{j,\cdot}\|_r \le 2^{-j(s-\frac{1}{r}+\frac{1}{2})}$. Furthermore,

$$\begin{aligned} E\|e_2\|_p^p &\lesssim (j_1 - j_0 + 1)^{p-1} \left(\frac{\ln n}{n}\right)^{\frac{1}{2}(p-r)} n^{\frac{r-p}{2}} \sum_{j=j_0^*+1}^{j_1} 2^{-j(sr+\frac{3}{2}r-\frac{3}{2}p)} \\ &\lesssim (\ln n)^{p-1} \left(\frac{\ln n}{n}\right)^{\frac{1}{2}(p-r)} 2^{-j_0^*(sr+\frac{3}{2}r-\frac{3}{2}p)} \lesssim (\ln n)^{p-1} \left(\frac{\ln n}{n}\right)^{\alpha p}. \end{aligned} \tag{4.14}$$

In the last inequality, one used the assumption $2^{j_0^*} \sim (\frac{n}{\ln n})^{\frac{1-2\alpha}{3}}$ and $\alpha = \frac{s}{2s+3}$ for $\theta > 0$. This with (4.11) and (4.12) leads to

$$E\|T_4\|_p^p \lesssim (\ln n)^{p-1} \left(\frac{\ln n}{n}\right)^{\alpha p}. \tag{4.15}$$

Then one needs only to show that (4.15) holds for $\theta \le 0$. Similarly, one divides T_4 into

$$T_4 = \sum_{j=j_0}^{j_1} \sum_{k \in K_j} \beta_{j,k} \psi_{j,k}(x) 1_{\{|\hat{\beta}_{j,k}| \le \lambda, |\beta_{j,k}| \le 2\lambda\}} = \sum_{j=j_0}^{j_1^*} + \sum_{j=j_1^*+1}^{j_1} =: e_1^* + e_2^*.$$

For the first sum, proceeding as (4.13) and (4.14), one has $E\|e_1^*\|_p^p \lesssim (j_1 - j_0 + 1)^{p-1} \times (\frac{\ln n}{n})^{\frac{p-r}{2}} \sum_{j=j_0}^{j_1^*} 2^{-j(sr+\frac{3}{2}r-\frac{3}{2}p)} \lesssim (j_1 - j_0 + 1)^{p-1} (\frac{\ln n}{n})^{\frac{p-r}{2}} 2^{-j_1^*(sr+\frac{3}{2}r-\frac{3}{2}p)}$. Note that $j_1 - j_0 \sim \ln n$ and $2^{j_1^*} \sim (\frac{n}{\ln n})^{\frac{\alpha}{s-\frac{1}{r}+\frac{1}{p}}}$. Then $E\|e_1^*\|_p^p \lesssim (\ln n)^{p-1} (\frac{\ln n}{n})^{\alpha p}$ due to $\alpha = \frac{s-\frac{1}{r}+\frac{1}{p}}{2(s-\frac{1}{r})+3}$ for $\theta \le 0$.

To estimate the second sum, using Theorem 1.2,

$$E\|e_2^*\|_p^p \lesssim (j_1 - j_0 + 1)^{p-1} \sum_{j=j_1^*+1}^{j_1} 2^{j(\frac{1}{2}p-1)} \sum_{k \in \wedge_j} |\beta_{j,k}|^p.$$

By $f \in B_{r,q}^s(H)$ and Theorem 1.1, $\|\beta_{j,\cdot}\|_p \le \|\beta_{j,\cdot}\|_r \lesssim 2^{-j(s-\frac{1}{r}+\frac{1}{2})}$. Hence, $E\|e_2^*\|_p^p \lesssim (j_1 - j_0 + 1)^{p-1} \sum_{j=j_1^*+1}^{j_1} 2^{-j(s-\frac{1}{r}+\frac{1}{p})p} \lesssim (j_1 - j_0 + 1)^{p-1} 2^{-j_1^*(s-\frac{1}{r}+\frac{1}{p})p}$. By the choice of j_1^*, $E\|e_2^*\|_p^p \lesssim (\ln n)^{p-1} (\frac{\ln n}{n})^{\alpha p}$ because of $\alpha = \frac{s-\frac{1}{r}+\frac{1}{p}}{2(s-\frac{1}{r})+3}$, when $\theta \le 0$. Then the desired conclusion (4.15) follows. This completes the proof of Theorem 2.2. □

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

LH, LL, JX, and HW completed this work together. All four authors read and approved the final manuscript.

Author details

[1]Department of Basic Courses, Beijing Union University, Beijing, 100101, P.R. China. [2]Department of Mathematics, Baoji University of Arts and Sciences, Baoji, 721013, P.R. China. [3]School of Applied Mathematics, Central University of Finance and Economics, Beijing, 100081, P.R. China.

Acknowledgements

This work is supported by the National Natural Science Foundation of China, Tian Yuan Foundation (11426040), the National Natural Science Foundation of China (11301558), Beijing Municipal Education Commission Project (KM201511417002) and Beijing Talents Found (2014000020124G065).

Received: 24 January 2015 Accepted: 11 June 2015 Published online: 23 June 2015

References

1. Abbaszadeh, M, Chesneau, C, Doosti, H: Nonparametric estimation of a density under bias and multiplicative censoring via wavelet methods. Stat. Probab. Lett. **82**, 932-941 (2012)
2. Vardi, Y: Multiplicative censoring, renewal processes, deconvolution and decreasing density: nonparametric estimation. Biometrika **76**, 751-761 (1989)
3. Vardi, Y, Zhang, CH: Large sample study of empirical distributions in a random multiplicative censoring model. Ann. Stat. **20**, 1022-1039 (1992)
4. Andersen, K, Hansen, M: Multiplicative censoring: density estimation by a series expansion approach. J. Stat. Plan. Inference **98**, 137-155 (2001)
5. Asgharian, M, Carone, M, Fakoor, V: Large-sample study of the kernel density estimators under multiplicative censoring. Ann. Stat. **40**, 159-187 (2012)
6. Abbaszadeh, M, Chesneau, C, Doosti, H: Multiplicative censoring: estimation of a density and its derivatives under the L_p-risk. REVSTAT Stat. J. **11**(3), 255-276 (2013)
7. Efromovich, S: Density estimation for biased data. J. Stat. Plan. Inference **32**(3), 1137-1161 (2004)
8. Chesneau, C: Wavelet block thresholding for density estimation in the presence of bias. J. Korean Stat. Soc. **39**, 43-53 (2010)
9. Ramirez, P, Vidakovic, B: Wavelet density estimation for stratified size-biased sample. J. Stat. Plan. Inference **140**, 419-432 (2010)
10. Chesneau, C: Regression with random design: a minimax study. Stat. Probab. Lett. **77**, 40-53 (2007)
11. Donoho, DL, Johnstone, IM, Kerkyacharian, G, Picard, D: Density estimation by wavelet thresholding. Ann. Stat. **24**, 508-539 (1996)

12. Li, LY: On the minimax optimality of wavelet estimators with censored data. J. Stat. Plan. Inference **137**, 1138-1150 (2007)
13. Daubechies, I: Ten Lectures on Wavelets. SIAM, Philadelphia (1992)
14. Härdle, W, Kerkyacharian, G, Picard, D, Tsybakov, A: Wavelets, Approximations, and Statistical Applications. Lecture Notes in Statistics. Springer, Berlin (1998)
15. Liu, YM, Xu, JL: Wavelet density estimation for negatively associated stratified size-biased sample. J. Nonparametr. Stat. **26**(3), 537-554 (2014)