On adaptive wavelet estimation of the regression function and its derivatives in an errors-in-variables model

Christophe Chesneau

To cite this version:

Christophe Chesneau. On adaptive wavelet estimation of the regression function and its derivatives in an errors-in-variables model. 2010. <hal-00466830v2>

HAL Id: hal-00466830
https://hal.archives-ouvertes.fr/hal-00466830v2
Submitted on 29 Jun 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
On adaptive wavelet estimation of the regression function and its derivatives in an errors-in-variables model

Christophe Chesneau
University of Caen (LMNO), Caen, France. Email: chesneau@math.unicaen.fr

Abstract

We consider a regression model with errors-in-variables: \((Y, X)\), where \(Y = f(Z) + \xi\) and \(X = Z + W\). Our goal is to estimate the unknown regression function \(f\) and its derivatives under mild assumptions on \(\xi\) (only finite moments of order 2 are required). To reach this goal, we develop a new adaptive wavelet estimator based on a hard thresholding rule. Taking the minimax approach under the mean integrated squared error over Besov balls, we prove that it attains a sharp rate of convergence.

Keywords: Errors-in-variables model, Derivatives function estimation, Minimax approach, Wavelets, Hard thresholding.


1 Motivations

We observe \(n\) independent pairs of random variables \((Y_1, X_1), \ldots, (Y_n, X_n)\) in the following errors-in-variables model: for any \(v \in \{1, \ldots, n\}\),

\[
\begin{cases}
    Y_v = f(Z_v) + \xi_v, \\
    X_v = Z_v + W_v,
\end{cases}
\]

where \(f\) is an unknown function, \(X_1, \ldots, X_n\) are \(n\) i.i.d. random variables having the uniform distribution on \([0, 1]\), \(W_1, \ldots, W_n\) are \(n\) i.i.d. unobserved random variables and \(\xi_1, \ldots, \xi_n\) are \(n\) i.i.d. unobserved zero mean random variables. We assume that all these random variables are independent, the density function of \(W_1\), denoted \(g\), is known and \(\xi_1\) admits finite moments of order 2. We aim to estimate \(f\) and its derivatives from \((Y_1, X_1), \ldots, (Y_n, X_n)\).

The estimation of \(f\) from the model (1.1) has received a lot of attention. See e.g. [6, 11–13, 15, 18, 19]. In this paper, we focus on a more general problem: the adaptive estimation of the \(d\)-th derivative of \(f\): \(f^{(d)}\). This is of interest to detect possible bumps, concavity or convexity properties of \(f\). The estimation of \(f^{(d)}\) has been investigated in

This work is supported by ANR grant NatImages, ANR-08-EMER-009
several papers for various models (but never (1.1)) starting with [1]. For references using wavelet methods, see e.g. [3–5, 23].

Another feature of the study concerns $\xi_1, \ldots, \xi_n$: to estimate $f^{(d)}$ (including $f = f^{(0)}$), we only assume that $\xi_1$ has finite moments of order 2. And thus we do not need to know the distribution of $\xi_1$. Moreover, this relaxes the assumption on $\xi_1$ in [6] where finite moments of order $> 6$ are required.

Under this general framework, considering the ordinary smooth case on $g$ (see (2.2)), we estimate $f^{(d)}$ by a new wavelet estimator based on a hard thresholding rule. It has the originality to combine a singular value decomposition (SVD) approach similar to the one of [17] and some technical tools introduced in wavelet estimation theory by [7]. We evaluate its performances by taking the minimax approach under the mean integrated squared error (MISE) over a wide class of functions: the Besov balls $B^{p,r}_s(M)$ (to be defined in Section 3). We prove that our estimator attains the rate of convergence $v_n = (\ln n/n)^{2s/(2s+2d+2d+1)}$, where $\delta$ is a factor related to the ordinary smooth case. This rate of convergence is sharp in the sense that it is the one attains by the best non-realistic linear wavelet estimator up to a logarithmic term.

The paper is organized as follows. Assumptions on the model and some notations are introduced in Section 2. Section 3 briefly describes the periodized wavelet basis on $[0, 1]$ and the Besov balls. The estimators are presented in Section 4. The results are set in Section 5. The proofs are gathered in Section 6.

2 Assumptions and notations

We assume in the sequel that $f^{(d)}$ and $g$ belong to $L^2_{\text{per}}([0, 1])$, the space of periodic functions of period one that are square-integrable on $[0, 1]$:

$$L^2_{\text{per}}([0, 1]) = \left\{ h; h \text{ is 1-periodic and } \|h\|_2 = \left( \int_0^1 h^2(x) dx \right)^{1/2} < \infty \right\}.$$ 

We assume that there exists a known constant $C_* > 0$ such that

$$\|f\|_\infty = \sup_{x \in [0, 1]} |f(x)| \leq C_* < \infty. \quad (2.1)$$

Any function $h \in L^2_{\text{per}}([0, 1])$ can be represented by its Fourier series

$$h(t) = \sum_{\ell \in \mathbb{Z}} \mathcal{F}(h)(\ell) e^{2i\pi \ell t}, \quad t \in [0, 1],$$

where the equality is intended in mean-square convergence sense, and $\mathcal{F}(h)(\ell)$ denotes the Fourier coefficient given by

$$\mathcal{F}(h)(\ell) = \int_0^1 h(x) e^{-2i\pi \ell x} dx, \quad \ell \in \mathbb{Z},$$
Wavelet estimation of the regression function and its derivatives in an errors-in-variables model

whenever this integral exists. The notation $\overline{\cdot}$ will be used for the complex conjugate.

We consider the ordinary smooth case on $g$: there exist three constants, $c_g > 0$, $C_g > 0$ and $\delta > 1$, such that, for any $\ell \in \mathbb{Z}$, the Fourier coefficient of $g$, i.e. $F(g)(\ell)$, satisfies

$$\frac{c_g}{(1 + \ell^2)^{\delta/2}} \leq |F(g)(\ell)| \leq \frac{C_g}{(1 + \ell^2)^{\delta/2}}. \quad (2.2)$$

This assumption controls the decay of the Fourier coefficients of $g$, and thus the smoothness of $g$. It is a standard hypothesis usually adopted in the field of nonparametric estimation for deconvolution problems. See e.g. [14, 17, 21].

3 Wavelets and Besov balls

3.1 Periodized Meyer Wavelets

We consider an orthonormal wavelet basis generated by dilations and translations of a "father" Meyer-type wavelet $\phi$ and a "mother" Meyer-type wavelet $\psi$. The main features of such wavelets are:

1. they are bandlimited, i.e. the Fourier transforms of $\phi$ and $\psi$ have compact supports respectively included in $[-4\pi/3, 4\pi/3]$ and $[-8\pi/3, -2\pi/3] \cup [2\pi/3, 8\pi/3]$.

2. for any frequency in $[-2\pi, -\pi] \cup [\pi, 2\pi]$, there exists a constant $c > 0$ such that the magnitude of the Fourier transform of $\psi$ is lower bounded by $c$.

3. the functions $(\phi, \psi)$ are $C^\infty$ as their Fourier transforms have a compact support, and $\psi$ has an infinite number of vanishing moments as its Fourier transform vanishes in a neighborhood of the origin, i.e. for any $u \in \mathbb{N}$, $\int_{-\infty}^{\infty} x^u \psi(x) dx = 0$.

For the purpose of this paper, we use the periodized Meyer wavelet bases on the unit interval. For any $x \in [0, 1]$, any integer $j$ and any $k \in \{0, \ldots, 2^j - 1\}$, let

$$\phi_{j,k}(x) = 2^{j/2} \phi(2^j x - k), \quad \psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k)$$

be the elements of the wavelet basis, and

$$\phi_{j,k}^{\text{per}}(x) = \sum_{l \in \mathbb{Z}} \phi_{j,k}(x - l), \quad \psi_{j,k}^{\text{per}}(x) = \sum_{l \in \mathbb{Z}} \psi_{j,k}(x - l),$$

their periodized versions. There exists an integer $j_*$ such that the collection $B = \{\phi_{j_*,k}^{\text{per}}, k \in \{0, \ldots, 2^{j_*} - 1\}; \psi_{j_*,k}^{\text{per}}, j \in \mathbb{N} - \{0, \ldots, j_* - 1\}, k \in \{0, \ldots, 2^j - 1\}\}$ forms an orthonormal basis of $L^2_{\text{per}}([0,1])$. In what follows, the superscript "per" will be dropped to lighten the notation.
Let $j_c$ be an integer such that $j_c \geq j_*$. A function $h \in L^2_{\text{per}}([0, 1])$ can be expanded into a wavelet series as
\[
h(x) = \sum_{k=0}^{2^{j_c} - 1} \alpha_{j_c, k} \phi_{j_c, k}(x) + \sum_{j=j_c}^{\infty} \sum_{k=0}^{2^{j_c} - 1} \beta_{j, k} \psi_{j, k}(x), \quad x \in [0, 1],
\]
where
\[
\alpha_{j, k} = \int_0^1 h(x) \overline{\phi}_{j, k}(x) dx, \quad \beta_{j, k} = \int_0^1 h(x) \overline{\psi}_{j, k}(x) dx.
\] (3.1)

See [20, Vol. 1 Chapter III.11] for a detailed account on periodized orthonormal wavelet bases.

### 3.2 Besov balls

Let $M > 0$, $s > 0$, $p \geq 1$ and $r \geq 1$. Set $\beta_{j_* - 1, k} = \alpha_{j_*, k}$. A function $h$ belongs to the Besov balls $\mathcal{B}_{s, p, r}^s(M)$ if and only if there exists a constant $M^* > 0$ such that the wavelet coefficients (3.1) satisfy
\[
\left( \sum_{j=j_*-1}^{\infty} \left( 2^{j(s+1/2-1/p)} \left( \sum_{k=0}^{2^{j_c} - 1} |\beta_{j, k}|^p \right)^{1/p} \right)^r \right)^{1/r} \leq M^*.
\]
For a particular choice of parameters $s$, $p$ and $r$, these sets contain the Hölder and Sobolev balls. See [20].

### 4 Estimators

**Wavelet coefficient estimators.** The first step to estimate $f^{(d)}$ consists in expanding $f^{(d)}$ on $\mathcal{B}$ and estimating its unknown wavelet coefficients.

For any integer $j \geq j_*$ and any $k \in \{0, \ldots, 2^j - 1\}$, we estimate $\alpha_{j, k} = \int_0^1 f^{(d)}(x) \overline{\phi}_{j, k}(x) dx$ by
\[
\hat{\alpha}_{j, k} = \frac{1}{n} \sum_{v=1}^{n} \sum_{\ell \in C_j} (2i\pi \ell)^d \frac{\mathcal{F}(\phi_{j, k})(\ell)}{\mathcal{F}(g)(\ell)} Y_v e^{-2i\pi \ell X_v},
\] (4.1)

$C_j = \text{supp} (\mathcal{F}(\phi_{j, 0})) = \text{supp} (\mathcal{F}(\phi_{j, k}))$, and $\beta_{j, k} = \int_0^1 f^{(d)}(x) \overline{\psi}_{j, k}(x) dx$ by
\[
\hat{\beta}_{j, k} = \frac{1}{n} \sum_{v=1}^{n} G_v 1_{\{G_v \leq \eta_j\}},
\] (4.2)

where
\[
G_v = \sum_{\ell \in D_j} (2i\pi \ell)^d \frac{\mathcal{F}(\psi_{j, k})(\ell)}{\mathcal{F}(g)(\ell)} Y_v e^{-2i\pi \ell X_v},
\]
Wavelet estimation of the regression function and its derivatives in an errors-in-variables model

\[ D_j = \text{supp}(F(\psi_{j,0})) = \text{supp}(F(\psi_{j,k})) \]

for any random event \( A \), \( 1_A \) is the indicator function on \( A \), and the threshold \( \eta_j \) is defined by

\[ \eta_j = \theta 2^{(\delta + d)j} \sqrt{\frac{n}{\ln n}}, \quad (4.3) \]

\[ \theta = \sqrt{C_{**}(C_*^2 + \mathbb{E}(\xi_1^2))}, \]

\( C_* \) is (2.1) and \( C_{**} = 2^{\delta - 1}(2(2\pi)^2 d/c^2)(8\pi/3)^{2(\delta + d)} \). The estimators \( \hat{\alpha}_{j,k} \) and \( \hat{\beta}_{j,k} \) are constructed via a SVD approach similar to the one of [17]. The idea of the thresholding in (4.2) is to operate a selection on the observations: when, for \( v \in \{1, \ldots, n\} \), \( f \) is “too noisy” by \( \xi_v \), the observation \((Y_v, X_v)\) is neglected. Such a technique has been introduced by [7] in wavelet estimation. Statistical properties of \( \hat{\alpha}_{j,k} \) and \( \hat{\beta}_{j,k} \) are given in Propositions 6.1, 6.2 and 6.3.

We consider two wavelets estimators for \( f^{(d)} \): a linear estimator and a hard thresholding estimator.

**Linear estimator.** Assuming that \( f^{(d)} \in B_{p,r}^s(M) \) with \( p \geq 2 \), we define the linear estimator \( \hat{f}^L_d \) by

\[ \hat{f}^L_d(x) = \sum_{k=0}^{2^{j_0} - 1} \hat{\alpha}_{j_0,k} \phi_{j_0,k}(x), \quad (4.4) \]

where \( \hat{\alpha}_{j,k} \) is defined by (4.1) and \( j_0 \) is the integer satisfying

\[ 2^{-1} n^{1/(2s+2\delta+2d+1)} < 2^{j_0} \leq n^{1/(2s+2\delta+2d+1)}. \]

It is not adaptive since it depends on \( s \), the smoothness parameter of \( f^{(d)} \).

**Hard thresholding estimator.** We define the hard thresholding estimator \( \hat{f}^H_d \) by

\[ \hat{f}^H_d(x) = \sum_{k=0}^{2^{j_1} - 1} \hat{\alpha}_{j_k,k} \phi_{j_k,k}(x) + \sum_{j=j_1}^{j_0} \sum_{k=0}^{2^{j-1}} \hat{\beta}_{j,k} \mathbb{1}_{\{|\hat{\beta}_{j,k}| \geq \kappa \lambda_j\}} \psi_{j,k}(x), \quad (4.5) \]

where \( \hat{\alpha}_{j,k} \) and \( \hat{\beta}_{j,k} \) are defined by (4.1) and (4.2), \( j_1 \) is the integer satisfying

\[ 2^{-1} n^{1/(2\delta+2d+1)} < 2^{j_1} \leq n^{1/(2\delta+2d+1)}, \]

\( \kappa \geq 8/3 + 2 + 2\sqrt{16/9 + 4} \) and \( \lambda_j \) is the threshold

\[ \lambda_j = \theta 2^{(\delta + d)j} \sqrt{\frac{\ln n}{n}}, \quad (4.6) \]

The definitions of \( \eta_j \) and \( \lambda_j \) are chosen to minimize the MISE of \( \hat{f}^H_d \) and to make it adaptive. Further statistical results on the hard thresholding estimator for the standard regression model can be found in [8–10].
5 Results

**Theorem 5.1.** Consider (1.1) under the assumptions of Section 2. Suppose that \( f^{(d)} \in B_{p,r}^s(M) \) with \( s > 0, p \geq 2 \) and \( r \geq 1 \). Let \( \hat{f}^L_d \) be (4.4). Then there exists a constant \( C > 0 \) such that

\[
E \left( \int_0^1 \left( \hat{f}^L_d(x) - f^{(d)}(x) \right)^2 \, dx \right) \leq C n^{-2s/(2s+2\delta+2d+1)}.
\]

The proof of Theorem 5.1 uses a moment inequality on (4.1) and a suitable decomposition of the MISE.

Since the distribution of \( \xi_1 \) is unknown, we can not use the likelihood function related to the model and the optimal lower bound seems difficult to determine (see e.g. [16, 25]). For this reason, our benchmark will be the rate of convergence attains by the "optimal non-realistic" \( \hat{f}^L_d \), i.e. \( v_n = n^{-2s/(2s+2\delta+2d+1)} \). Note that, under the standard Gaussian assumption on \( \xi_1 \), one can prove that \( v_n \) is optimal in the minimax sense.

**Theorem 5.2.** Consider (1.1) under the assumptions of Section 2. Let \( \hat{f}^H_d \) be (4.5). Suppose that \( f^{(d)} \in B_{p,r}^s(M) \) with \( r \geq 1, \{p \geq 2 \text{ and } s > 0\} \text{ or } \{p \in [1, 2) \text{ and } s > (2\delta + 2d + 1)/p\} \). Then there exists a constant \( C > 0 \) such that

\[
E \left( \int_0^1 \left( \hat{f}^H_d(x) - f^{(d)}(x) \right)^2 \, dx \right) \leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)}.
\]

The proof of Theorem 5.2 is based on several probability results (moment inequalities, concentration inequality, . . . ) and a suitable decomposition of the MISE.

Theorem 5.2 proves that \( \hat{f}^H_d \) attains \( v_n \) up to the logarithmic term \( (\ln n)^{2s/(2s+2\delta+2d+1)} \).

**Conclusion and perspectives.** We have constructed a new adaptive estimator \( \hat{f}^H_d \) for \( f^{(d)} \) under mild assumption on \( \xi_1 \). It is based on wavelet and thresholding. It has "near-optimal" minimax properties for a wide class of functions \( f^{(d)} \). Possible perspectives of this work are

- to potentially improve the estimation of \( f^{(d)} \) by considering other kinds of thresholding rules as the block thresholding one introduced by [2],
- to investigate the random design case where the distribution of \( X_1 \) is unknown.

6 Proofs

In this section, \( C \) represents a positive constant which may differ from one term to another.
6.1 Auxiliary results

**Proposition 6.1.** For any integer \( j \geq j_\ast \) and any \( k \in \{0, \ldots, 2^j - 1\} \), let \( \alpha_{j,k} \) be the wavelet coefficient (3.1) of \( f^{(d)} \) and \( \hat{\alpha}_{j,k} \) be (4.1). Then there exists a constant \( C > 0 \) such that

\[
\mathbb{E} \left( (\hat{\alpha}_{j,k} - \alpha_{j,k})^2 \right) \leq C 2^{2(\delta + d) j} \frac{1}{n}.
\]

**Proof of Proposition 6.1.** For any \( v \in \{1, \ldots, n\} \), let us set

\[
H_v = \sum_{\ell \in C_j} (2i\pi \ell)^d \frac{\mathcal{F}(\phi_{j,k})(\ell)}{\mathcal{F}(g)(\ell)} Y_v e^{-2i\pi \ell X_v}.
\]

Since \( X_1, W_1 \) and \( \xi_1 \) are independent, using the convolution product between \( f \) and \( g \), i.e. \((f * g)(x) = \int_{-\infty}^{\infty} f(x-y)g(y)dy \), we have

\[
\mathbb{E} (Y_1 e^{-2i\pi \ell X_1}) = \mathbb{E} (f(X_1 - W_1) e^{-2i\pi \ell X_1}) + \mathbb{E} (\xi_1) \mathbb{E} (e^{-2i\pi \ell X_1})
\]

\[
= \mathbb{E} (f(X_1 - W_1) e^{-2i\pi \ell X_1}) = \mathbb{E} \left( \int_0^1 \int_0^1 (f(x-y)g(y)) e^{-2i\pi \ell x} dx dy \right) = \mathcal{F}(f * g)(\ell) = \mathcal{F}(f^{(d)})(\ell).
\]

Moreover, since \( f \) is 1-periodic, for any \( u \in \{0, \ldots, d\} \), \( f^{(u)}(0) = f^{(u)}(1) \). By \( d \) integrations by parts, for any \( \ell \in \mathbb{Z} \), we have

\[
(2i\pi \ell)^d \mathcal{F}(f)(\ell) = \mathcal{F} \left( f^{(d)} \right)(\ell).
\]

The Parseval-Plancherel theorem gives

\[
\mathbb{E}(H_1) = \sum_{\ell \in C_j} (2i\pi \ell)^d \frac{\mathcal{F}(\phi_{j,k})(\ell)}{\mathcal{F}(g)(\ell)} \mathbb{E} (Y_1 e^{-2i\pi \ell X_1})
\]

\[
= \sum_{\ell \in C_j} (2i\pi \ell)^d \frac{\mathcal{F}(\phi_{j,k})(\ell)}{\mathcal{F}(g)(\ell)} \mathcal{F}(f)(\ell) \mathcal{F}(g)(\ell)
\]

\[
= \sum_{\ell \in C_j} \mathcal{F}(\phi_{j,k})(\ell)(2i\pi \ell)^d \mathcal{F}(f)(\ell) = \sum_{\ell \in C_j} \mathcal{F}(\phi_{j,k})(\ell) \mathcal{F} \left( f^{(d)} \right)(\ell)
\]

\[
= \int_0^1 \bar{\phi}_{j,k}(x) f^{(d)}(x) dx = \alpha_{j,k}.
\]

Hence \( \mathbb{E} (\hat{\alpha}_{j,k} - \alpha_{j,k})^2 \leq C 2^{2(\delta + d) j} \frac{1}{n} \mathbb{E}(H_1) \leq \frac{1}{n} \mathbb{E} \left( H_1^2 \right) \).

\[
\mathbb{E} \left( (\hat{\alpha}_{j,k} - \alpha_{j,k})^2 \right) \leq C 2^{2(\delta + d) j} \frac{1}{n} \mathbb{E}(H_1) \leq \frac{1}{n} \mathbb{E} \left( H_1^2 \right).
\]

(6.2)
Since $X_1$, $W_1$ and $\xi_1$ are independent with $\mathbb{E}(\xi_1) = 0$ and $|f(X_1 - W_1)| \leq \|f\|_\infty \leq C_* < \infty$, we have

$$
\mathbb{E}(H_1^2) = \mathbb{E}\left(f^2(X_1 - W_1) \left( \sum_{\ell \in \mathcal{C}_j} (2i\pi \ell)^d \frac{F_j(\phi_{j,k})(\ell)}{F(g)(\ell)} e^{-2i\pi \ell X_1} \right)^2 \right)
$$

$$
+ 2\mathbb{E}(\xi_1) \mathbb{E}\left(f(X_1 - W_1) \left( \sum_{\ell \in \mathcal{C}_j} (2i\pi \ell)^d \frac{F_j(\phi_{j,k})(\ell)}{F(g)(\ell)} e^{-2i\pi \ell X_1} \right)^2 \right)
$$

$$
+ \mathbb{E}(\xi_1^2) \mathbb{E}\left( \left( \sum_{\ell \in \mathcal{C}_j} (2i\pi \ell)^d \frac{F_j(\phi_{j,k})(\ell)}{F(g)(\ell)} e^{-2i\pi \ell X_1} \right)^2 \right)
$$

$$
= \mathbb{E}\left(f^2(X_1 - W_1) \left( \sum_{\ell \in \mathcal{C}_j} (2i\pi \ell)^d \frac{F_j(\phi_{j,k})(\ell)}{F(g)(\ell)} e^{-2i\pi \ell X_1} \right)^2 \right)
$$

$$
+ \mathbb{E}(\xi_1^2) \mathbb{E}\left( \left( \sum_{\ell \in \mathcal{C}_j} (2i\pi \ell)^d \frac{F_j(\phi_{j,k})(\ell)}{F(g)(\ell)} e^{-2i\pi \ell X_1} \right)^2 \right)
$$

$$
\leq \left( C_*^2 + \mathbb{E}(\xi_1^2) \right) \mathbb{E}\left( \left( \sum_{\ell \in \mathcal{C}_j} (2i\pi \ell)^d \frac{F_j(\phi_{j,k})(\ell)}{F(g)(\ell)} e^{-2i\pi \ell X_1} \right)^2 \right). \quad (6.3)
$$

The assumption $(2.2)$ implies

$$
\sup_{\ell \in \mathcal{C}_j} \frac{(2\pi \ell)^{2d}}{F(g)(\ell)^2} \leq \sup_{\ell \in \mathcal{C}_j} \ell^{2d} \left( 1 + \ell^2 \right)^{\delta} \leq 2^{\delta - 1} \frac{(2\pi)^{2d}}{c_g^2} \sup_{\ell \in \mathcal{C}_j} \ell^{2d} (1 + \ell^{2\delta}) \leq \frac{8\pi}{3} 2^{(\delta + d)} 2^{2(\delta + d)j} = C_* 2^{2(\delta + d)j}. \quad (6.4)
$$

Using the fact that $(e^{-2i\pi \ell x})_{\ell \in \mathbb{Z}}$ is an orthonormal basis of $L^2_{per}([0,1])$, (6.4) and the Parseval-Plancherel theorem, we obtain

$$
\mathbb{E}\left( \left( \sum_{\ell \in \mathcal{C}_j} (2i\pi \ell)^d \frac{F_j(\phi_{j,k})(\ell)}{F(g)(\ell)} e^{-2i\pi \ell X_1} \right)^2 \right)
$$

$$
= \int_0^1 \left( \sum_{\ell \in \mathcal{C}_j} (2i\pi \ell)^d \frac{F_j(\phi_{j,k})(\ell)}{F(g)(\ell)} e^{-2i\pi \ell x} \right)^2 dx = \sum_{\ell \in \mathcal{C}_j} (2\pi \ell)^{2d} \frac{|F_j(\phi_{j,k})(\ell)|^2}{|F(g)(\ell)|^2}
$$

$$
\leq C_* 2^{2(\delta + d)j} \sum_{\ell \in \mathcal{C}_j} |F_j(\phi_{j,k})(\ell)|^2 = C_* 2^{2(\delta + d)j} \int_0^1 |\phi_{j,k}(x)|^2 dx
$$

$$
= C_* 2^{2(\delta + d)j}. \quad (6.5)
$$
Putting (6.3) and (6.5) together, we obtain

\[ \mathbb{E}(H_1^2) \leq \theta^2 2^{2(d+j)}j. \] (6.6)

It follows from (6.2) and (6.6) that

\[ \mathbb{E}\left( (\hat{\alpha}_{j,k} - \alpha_{j,k})^2 \right) \leq C 2^{2(d+j)} \frac{1}{n}. \]

\[ \square \]

**Proposition 6.2.** For any integer \( j \geq j_\ast \) and any \( k \in \{0, \ldots, 2^j - 1\} \), let \( \beta_{j,k} \) be the wavelet coefficient (3.1) of \( f(d) \) and \( \hat{\beta}_{j,k} \) be (4.2). Then there exists a constant \( C > 0 \) such that

\[ \mathbb{E}\left( \left( \hat{\beta}_{j,k} - \beta_{j,k} \right)^4 \right) \leq C 2^{4(d+j)} \frac{(\ln n)^2}{n^2}. \]

**Proof of Proposition 6.2.** Proceeding as in (6.1) (with \( \psi \) instead of \( \phi \)), we have

\[ \beta_{j,k} = \int_0^1 f(d)(x) \psi_{j,k}(x)dx = \mathbb{E}(G_e) = \mathbb{E}(G_e 1_{\{|G_e| \leq \eta_j\}}) + \mathbb{E}(G_1 1_{\{|G_1| > \eta_j\}}). \] (6.7)

We have

\[ \mathbb{E}\left( \left( \hat{\beta}_{j,k} - \beta_{j,k} \right)^4 \right) \]

\[ = \mathbb{E}\left( \left( \frac{1}{n} \sum_{v=1}^n (G_e 1_{\{|G_e| \leq \eta_j\}} - \mathbb{E}(G_e 1_{\{|G_e| \leq \eta_j\}}) - \mathbb{E}(G_1 1_{\{|G_1| > \eta_j\}}) \right)^4 \right) \]

\[ \leq 8(A + B), \] (6.8)

where

\[ A = \mathbb{E}\left( \left( \frac{1}{n} \sum_{v=1}^n (G_e 1_{\{|G_e| \leq \eta_j\}} - \mathbb{E}(G_e 1_{\{|G_e| \leq \eta_j\}}) \right)^4 \right) \]

and

\[ B = \left( \mathbb{E}(|G_1| 1_{\{|G_1| > \eta_j\}}) \right)^4. \]

Let us bound \( A \) and \( B \), in turn. To bound \( A \), we need the Rosenthal inequality presented in lemma below (see [24]).

**Lemma 6.1 (Rosenthal’s inequality).** Let \( p \geq 2 \), \( n \in \mathbb{N}^* \) and \( (U_v)_{v \in \{1, \ldots, n\}} \) be \( n \) zero mean i.i.d. random variables such that \( \mathbb{E}(|U_1|^p) < \infty \). Then there exists a constant \( C > 0 \) such that

\[ \mathbb{E}\left( \left| \sum_{v=1}^n U_v \right|^p \right) \leq C \max \left( n \mathbb{E}(|U_1|^p), (n \mathbb{E}(U_1^2))^{p/2} \right). \]
Applying the Rosenthal inequality with \( p = 4 \) and, for any \( v \in \{1, \ldots, n\} \),
\[
U_v = G_v 1_{\{|G_v| \leq n_j\}} - \mathbb{E} \left( G_v 1_{\{|G_v| \leq n_j\}} \right),
\]
we obtain
\[
A = \frac{1}{n^4} \mathbb{E} \left( \left( \sum_{v=1}^{n} U_v \right)^4 \right) \leq C \frac{1}{n^4} \max \left( n \mathbb{E} \left( U_1^4 \right), (n \mathbb{E} \left( U_1^2 \right))^2 \right).
\]

Using (6.7), we have
\[
\text{Proof of Proposition 6.3.}
\]

For any integer \( \beta \in \{2, 4\} \),
\[
C \frac{1}{n} \mathbb{E} \left( |G_1| 1_{\{|G_1| \leq n_j\}} \right) \leq 2^{a_n} \eta_j^{a - 2} \mathbb{E} \left( G_1^2 \right) \leq 2^{a_n} \eta_j^{a - 2} 2^{2(\delta + d)j}.
\]
Hence
\[
A \leq C \frac{1}{n} \ln n \frac{2(\delta + d)j}{\eta_j} \mathbb{E} \left( |G_1| 1_{\{|G_1| \geq n_j\}} \right) \leq \frac{1}{\theta 2^{(\delta + d)j} \ln n} \frac{2(\delta + d)j}{\eta_j} \mathbb{E} \left( |G_1| 1_{\{|G_1| \geq n_j\}} \right).
\]

Let us now bound \( B \). Using again (6.6) (with \( \psi \) instead of \( \phi \)), we obtain
\[
\mathbb{E} \left( |G_1| 1_{\{|G_1| \geq n_j\}} \right) \leq \mathbb{E} \left( G_1^2 \right) \leq \frac{1}{\theta 2^{(\delta + d)j} \ln n} \frac{2(\delta + d)j}{\eta_j} \mathbb{E} \left( |G_1| 1_{\{|G_1| \geq n_j\}} \right).
\]

Hence
\[
B \leq C 2^{4(\delta + d)j} \frac{(\ln n)^2}{n^2}.
\]

Combining (6.8), (6.9) and (6.11), we have
\[
\mathbb{E} \left( \left| \beta_{j,k} - \beta_{j,k} \right|^4 \right) \leq C \left( 2^{4(\delta + d)j} \frac{1}{n^2} + 2^{4(\delta + d)j} \frac{(\ln n)^2}{n^2} \right) \leq C 2^{4(\delta + d)j} \frac{(\ln n)^2}{n^2}.
\]

\( \square \)

**Proposition 6.3.** For any integer \( j \geq j_0 \) and any \( k \in \{0, \ldots, 2^j - 1\} \), let \( \beta_{j,k} \) be the wavelet coefficient (3.1) of \( \theta_{j,k} \) and \( \beta_{j,k} \) be (4.2). Then, for any \( \kappa \geq 8/3+2+2\sqrt{16/9}+4 \),
\[
\mathbb{P} \left( \left| \beta_{j,k} - \beta_{j,k} \right| \geq \kappa \lambda_j / 2 \right) \leq 2n^{-2}.
\]

**Proof of Proposition 6.3.** Using (6.7), we have
\[
\left| \beta_{j,k} - \beta_{j,k} \right| = \left| \frac{1}{n} \sum_{v=1}^{n} \left( G_v 1_{\{|G_v| \leq n_j\}} - \mathbb{E} \left( G_v 1_{\{|G_v| \leq n_j\}} \right) \right) - \mathbb{E} \left( G_1 1_{\{|G_1| > n_j\}} \right) \right| + \mathbb{E} \left( |G_1| 1_{\{|G_1| > n_j\}} \right).
\]
Using (6.10), we obtain
\[ \mathbb{E} \left( |G_1| \mathbb{1}_{\{|G_1| > \eta_j\}} \right) \leq \theta^2(\delta + d)j \sqrt{\frac{\ln n}{n}} \lambda_j. \]

Hence
\[ S = \mathbb{P} \left( |\hat{\beta}_{j,k} - \beta_{j,k}| \geq \kappa \lambda_j/2 \right) \leq \mathbb{P} \left( \left| \frac{1}{n} \sum_{v=1}^{n} (G_v 1_{\{|G_v| \leq \eta_j\}} - \mathbb{E}(G_v 1_{\{|G_v| \leq \eta_j\}})) \right| \geq (\kappa/2 - 1)\lambda_j \right). \]

Now we need the Bernstein inequality presented in the lemma below (see [22]).

**Lemma 6.2 (Bernstein’s inequality).** Let \( n \in \mathbb{N}^+ \) and \((U_v)_{v \in \{1, \ldots, n\}}\) be \( n \) zero mean i.i.d. random variables such that there exists a constant \( M > 0 \) satisfying, for any \( v \in \{1, \ldots, n\}, |U_v| \leq M < \infty \). Then, for any \( \lambda > 0 \), holds
\[ \mathbb{P} \left( \left| \sum_{v=1}^{n} U_v \right| \geq \lambda \right) \leq 2 \exp \left( -\frac{\lambda^2}{2 (n\mathbb{E}(U_1^2) + \frac{\lambda M}{3})} \right). \]

Let us set, for any \( v \in \{1, \ldots, n\}, \)
\[ U_v = G_v 1_{\{|G_v| \leq \eta_j\}} - \mathbb{E}(G_v 1_{\{|G_v| \leq \eta_j\}}). \]

Then \( \mathbb{E}(U_1) = 0, \)
\[ |U_v| \leq |G_v 1_{\{|G_v| \leq \eta_j\}}| + \mathbb{E}(\{G_v 1_{\{|G_v| \leq \eta_j\}}\}) \leq 2\eta_j \]
and, using again (6.6) (with \( \psi \) instead of \( \phi \)),
\[ \mathbb{E}(U_1^2) = \mathbb{V}(G_v 1_{\{|G_v| \leq \eta_j\}}) \leq \mathbb{E}(G_1^2) \leq \theta^2 2^{2(\delta + d)j}. \]

It follows from the Bernstein inequality that
\[ S \leq 2 \exp \left( -\frac{\lambda^2}{2 (\theta^2 n 2^{2(\delta + d)j} + \frac{2n(\kappa/2 - 1)\lambda_j \eta_j}{3})} \right). \]

Since
\[ \lambda_j \eta_j = \theta^2(\delta + d)j \sqrt{\frac{\ln n}{n}} \theta^2(\delta + d)j \sqrt{\frac{n}{\ln n}} = \theta^2 2^{2(\delta + d)j} \frac{\ln n}{n}, \]
we have, for any \( \kappa \geq 8/3 + 2 + 2\sqrt{16/9 + 4}, \)
\[ S \leq 2 \exp \left( -\frac{(\kappa/2 - 1)^2 \ln n}{2 \left( 1 + \frac{2(\kappa/2 - 1)}{3} \right)} \right) \leq 2n^{-2}. \]

\( \square \)
6.2 Proofs of the main results

Proof of Theorem 5.1. We expand the function \( f^{(d)} \) as

\[
f^{(d)}(x) = \sum_{k=0}^{2^{j_0} - 1} \alpha_{j_0, k} \phi_{j_0, k}(x) + \sum_{j=j_0}^{2^j-1} \sum_{k=0}^{2^j-1} \beta_{j, k} \psi_{j, k}(x),
\]

where

\[
\alpha_{j_0, k} = \int_0^1 f^{(d)}(x) \phi_{j_0, k}(x) dx, \quad \beta_{j, k} = \int_0^1 f^{(d)}(x) \psi_{j, k}(x) dx.
\]

We have

\[
\hat{f}^{(d)}_L(x) - f^{(d)}(x) = \sum_{k=0}^{2^{j_0} - 1} (\hat{\alpha}_{j_0, k} - \alpha_{j_0, k}) \phi_{j_0, k}(x) - \sum_{j=j_0}^{2^j-1} \sum_{k=0}^{2^j-1} \beta_{j, k} \psi_{j, k}(x).
\]

Hence

\[
\mathbb{E}\left( \int_0^1 \left( \hat{f}^{(d)}_L(x) - f^{(d)}(x) \right)^2 dx \right) = A + B,
\]

where

\[
A = \sum_{k=0}^{2^{j_0} - 1} \mathbb{E}\left( (\hat{\alpha}_{j_0, k} - \alpha_{j_0, k})^2 \right), \quad B = \sum_{j=j_0}^{2^j-1} \sum_{k=0}^{2^j-1} \beta_{j, k}^2.
\]

Using Proposition 6.1, we obtain

\[
A \leq C 2^{j_0(1+2\delta+2d)} \frac{1}{n} \leq C n^{-2s/(2s+2\delta+2d+1)}.
\]

Since \( p \geq 2 \), we have \( B^p(M) \subseteq B^{q}_{2,\infty}(M) \). Hence

\[
B \leq C 2^{-2j_0s} \leq C n^{-2s/(2s+2\delta+2d+1)}.
\]

So

\[
\mathbb{E}\left( \int_0^1 \left( \hat{f}^{(d)}_L(x) - f^{(d)}(x) \right)^2 dx \right) \leq C n^{-2s/(2s+2\delta+2d+1)}.
\]

The proof of Theorem 5.1 is complete.

\[ \square \]

Proof of Theorem 5.2. We expand the function \( f^{(d)} \) as

\[
f^{(d)}(x) = \sum_{k=0}^{2^{j_0} - 1} \alpha_{j_0, k} \phi_{j_0, k}(x) + \sum_{j=j_0}^{2^j-1} \sum_{k=0}^{2^j-1} \beta_{j, k} \psi_{j, k}(x),
\]
Wavelet estimation of the regression function and its derivatives in an errors-in-variables model

where

\[ \alpha_{j,k} = \int_0^1 f^{(d)}(x) \phi_{j,k}(x) \, dx, \quad \beta_{j,k} = \int_0^1 f^{(d)}(x) \psi_{j,k}(x) \, dx. \]

We have

\[
\hat{f}_d^H(x) - f^{(d)}(x) \\
= \sum_{k=0}^{2^j-1} (\hat{\alpha}_{j,k} - \alpha_{j,k}) \phi_{j,k}(x) + \sum_{j=j_1}^{j_1-1} \sum_{k=0}^{2^j-1} (\hat{\beta}_{j,k} 1_{\{|\beta_{j,k}| \geq \kappa \lambda_j\}} - \beta_{j,k}) \psi_{j,k}(x) \\
- \sum_{j=j_1+1}^{\infty} \sum_{k=0}^{2^{j-1}-1} \beta_{j,k} \psi_{j,k}(x).
\]

Hence

\[
E \left( \int_0^1 \left( \hat{f}_d^H(x) - f^{(d)}(x) \right)^2 \, dx \right) = R + S + T, \tag{6.12}
\]

where

\[
R = \sum_{k=0}^{2^j-1} E \left( (\hat{\alpha}_{j,k} - \alpha_{j,k})^2 \right), \quad S = \sum_{j=j_1}^{j_1-1} \sum_{k=0}^{2^j-1} E \left( (\hat{\beta}_{j,k} 1_{\{|\beta_{j,k}| \geq \kappa \lambda_j\}} - \beta_{j,k})^2 \right)
\]

and

\[
T = \sum_{j=j_1+1}^{\infty} \sum_{k=0}^{2^{j-1}-1} \beta_{j,k}^2.
\]

Let us bound \( R, T \) and \( S \), in turn.

Using Proposition 6.1, we have

\[
R \leq C 2^{j_1(1+2\delta+2d)} \frac{1}{n} \leq C \frac{1}{n} \leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)}. \tag{6.13}
\]

For \( r \geq 1 \) and \( p \geq 2 \), we have \( B^s_{p,r}(M) \subseteq B^s_{2,\infty}(M) \). So

\[
T \leq C \sum_{j=j_1+1}^{\infty} 2^{-2js} \leq C 2^{-2js} \leq C n^{-2s/(2\delta+2d+1)} \leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)}.
\]

For \( r \geq 1 \) and \( p \in [1, 2) \), we have \( B^s_{p,r}(M) \subseteq B^{s+1/2-1/p}_{2,\infty}(M) \). Since \( s > (2\delta+2d+1)/p \), we have \((s+1/2-1/p)/(2\delta+2d+1) > s/(2s+2\delta+2d+1)\). So

\[
T \leq C \sum_{j=j_1+1}^{\infty} 2^{-2js} \leq C 2^{-2js} \leq C n^{-2(s+1/2-1/p)/(2\delta+2d+1)} \leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)}.
\]
Hence, for $r \geq 1$, \{ $p \geq 2$ and $s > 0$ \} or \{ $p \in [1, 2)$ and $s > (2\delta + 2d + 1)/p$ \}, we have

\[
T \leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)}.
\]

(6.14)

The term $S$ can be decomposed as

\[
S = e_1 + e_2 + e_3 + e_4,
\]

(6.15)

where

\[
e_1 = \sum_{j=j_*}^{2^{j_1}-1} \sum_{k=0}^{2^{j_1}-1} \mathbb{E} \left( \left( \beta_{j,k} - \hat{\beta}_{j,k} \right)^2 1_{\{ |\beta_{j,k}| \geq \kappa \lambda_j \}} 1_{\{ |\beta_{j,k}| < \kappa \lambda_j/2 \}} \right),
\]

\[
e_2 = \sum_{j=j_*}^{2^{j_1}-1} \sum_{k=0}^{2^{j_1}-1} \mathbb{E} \left( \left( \beta_{j,k} - \hat{\beta}_{j,k} \right)^2 1_{\{ |\beta_{j,k}| \geq \kappa \lambda_j \}} 1_{\{ |\beta_{j,k}| \geq \kappa \lambda_j/2 \}} \right),
\]

\[
e_3 = \sum_{j=j_*}^{2^{j_1}-1} \sum_{k=0}^{2^{j_1}-1} \mathbb{E} \left( \beta_{j,k}^2 1_{\{ |\beta_{j,k}| < \kappa \lambda_j \}} 1_{\{ |\beta_{j,k}| \geq \kappa \lambda_j \}} \right)
\]

and

\[
e_4 = \sum_{j=j_*}^{2^{j_1}-1} \sum_{k=0}^{2^{j_1}-1} \mathbb{E} \left( \beta_{j,k}^2 1_{\{ |\beta_{j,k}| < \kappa \lambda_j \}} 1_{\{ |\beta_{j,k}| < 2\kappa \lambda_j \}} \right).
\]

Let us analyze each term $e_1$, $e_2$, $e_3$ and $e_4$ in turn.

**Upper bounds for $e_1$ and $e_3$.** We have

\[
\left\{ |\hat{\beta}_{j,k}| < \kappa \lambda_j, |\beta_{j,k}| \geq 2\kappa \lambda_j \right\} \subseteq \left\{ |\hat{\beta}_{j,k} - \beta_{j,k}| > \kappa \lambda_j/2 \right\},
\]

\[
\left\{ |\hat{\beta}_{j,k}| \geq \kappa \lambda_j, |\beta_{j,k}| < \kappa \lambda_j/2 \right\} \subseteq \left\{ |\hat{\beta}_{j,k} - \beta_{j,k}| > \kappa \lambda_j/2 \right\}
\]

and

\[
\left\{ |\hat{\beta}_{j,k}| < \kappa \lambda_j, |\beta_{j,k}| \geq 2\kappa \lambda_j \right\} \subseteq \left\{ |\beta_{j,k}| \leq 2|\hat{\beta}_{j,k} - \beta_{j,k}| \right\}.
\]

So

\[
\max(e_1, e_3) \leq C \frac{\ln n}{n^2}.
\]

It follows from the Cauchy-Schwarz inequality and Propositions 6.2 and 6.3 that

\[
\mathbb{E} \left( \left( \beta_{j,k} - \hat{\beta}_{j,k} \right)^2 1_{\{ |\beta_{j,k} - \hat{\beta}_{j,k}| > \kappa \lambda_j/2 \}} \right)
\]

\[
\leq \left( \mathbb{E} \left( \beta_{j,k} - \hat{\beta}_{j,k} \right)^4 \right)^{1/2} \left( \mathbb{P} \left( |\hat{\beta}_{j,k} - \beta_{j,k}| > \kappa \lambda_j/2 \right) \right)^{1/2}
\]

\[
\leq C 2^{2(\delta + d)j} \frac{\ln n}{n^2}.
\]
Wavelet estimation of the regression function and its derivatives in an errors-in-variables model \(5\)

Hence
\[
\max(e_1, e_3) \leq C \frac{\ln n}{n^2} \sum_{j=j_*}^{j_1} 2^{j(1+2\delta+2d)} \leq C \frac{\ln n}{n^2} 2^{j_1(1+2\delta+2d)} \\
\leq C \frac{\ln n}{n} \leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)} .
\]

(6.16)

Upper bound for the term \(e_2\). Using Proposition 6.2 and the Cauchy-Schwarz inequality, we obtain
\[
\mathbb{E} \left( \left( \widehat{\beta}_{j,k} - \beta_{j,k} \right)^2 \right) \leq \left( \mathbb{E} \left( \left( \widehat{\beta}_{j,k} - \beta_{j,k} \right)^4 \right) \right)^{1/2} \leq C 2^{2(\delta+d)j} \frac{\ln n}{n} .
\]

Hence
\[
e_2 \leq C \frac{\ln n}{n} \sum_{j=j_*}^{j_1} 2^{2(\delta+d)j} \sum_{k=0}^{2^j-1} 1\{ |\beta_{j,k}| > \kappa \lambda_j / 2 \} .
\]

Let \(j_2 \) be the integer defined by
\[
2^{-1} \left( \frac{n}{\ln n} \right)^{1/(2s+2\delta+2d+1)} < 2^j_2 \leq \left( \frac{n}{\ln n} \right)^{1/(2s+2\delta+2d+1)} .
\]

(6.17)

We have
\[
e_2 \leq e_{2,1} + e_{2,2} ,
\]

where
\[
e_{2,1} = C \frac{\ln n}{n} \sum_{j=j_*}^{j_2} 2^{(1+2\delta+2d)} \leq C \frac{\ln n}{n} 2^{j_2(1+2\delta+2d)} \leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)} .
\]

We have
\[
e_{2,1} \leq C \frac{\ln n}{n} \sum_{j=j_*}^{j_2} 2^{j(1+2\delta+2d)} \leq C \frac{\ln n}{n} 2^{j_2(1+2\delta+2d)} \leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)} .
\]

For \(r \geq 1\) and \(p \geq 2\), since \(B^p_{p,r}(M) \subseteq B^p_{2,\infty}(M)\),
\[
e_{2,2} \leq C \frac{\ln n}{n} \sum_{j=j_*}^{j_2} 2^{(\delta+d)j} \frac{1}{\lambda_j} \sum_{k=0}^{2^j-1} \beta^2_{j,k} = C \sum_{j=j_*}^{\infty} \sum_{k=0}^{2^j-1} \beta^2_{j,k} \leq C 2^{-2j_2} .
\]

\[
e_{2,2} \leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)} .
\]
For \( r \geq 1, p \in [1, 2) \) and \( s > (2\delta + 2d + 1)/p \), since \( B_{p,r}^s(M) \subseteq B_{2,\infty}^{s+1/2-1/p}(M) \) and 
\( (2s + 2\delta + 2d + 1)(2 - p)/2 + (s + 1/2 - 1/p + \delta + d - 2(\delta + d)/p)p = 2s \), we have

\[
e_{2,2} \leq C \frac{\ln n}{n} \sum_{j=\lfloor s/p \rfloor + 1}^{\lceil \delta \rceil / \lambda_j} 2^{2j(\delta+d)} \sum_{k=0}^{2j-1} |\beta_{j,k}|^p
\]

\[
\leq C \left( \frac{\ln n}{n} \right)^{(2-p)/2} \sum_{j=\lfloor s/p \rfloor + 1}^{\lceil \delta \rceil / \lambda_j} 2^{j(\delta+d)(2-p)} 2^{-j(s+1/2-1/p)p}
\]

\[
\leq C \left( \frac{\ln n}{n} \right)^{(2-p)/2} 2^{-j(s+1/2-1/p+\delta+d-2(\delta + d)/p)p}
\]

\[
\leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)}
\]

So, for \( r \geq 1, \{p \geq 2 \text{ and } s > 0\} \) or \( \{p \in [1, 2) \text{ and } s > (2\delta + 2d + 1)/p\} \),

\[
e_2 \leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)}.
\]

\textit{Upper bound for the term } e_4. \text{ We have}

\[
e_4 \leq \sum_{j=\lfloor s/p \rfloor + 1}^{\lceil \delta \rceil / \lambda_j} 2^{j-1} \sum_{k=0}^{2j-1} \beta_{j,k}^2 1_{\{|\beta_{j,k}| < 2\lambda_j\}}.
\]

Let \( j_2 \) be the integer \((6.17)\). We have

\[
e_4 \leq e_{4,1} + e_{4,2},
\]

where

\[
e_{4,1} = \sum_{j=\lfloor s/p \rfloor + 1}^{j_2} 2^{j-1} \sum_{k=0}^{2j-1} \beta_{j,k}^2 1_{\{|\beta_{j,k}| < 2\lambda_j\}}, \quad e_{4,2} = \sum_{j=\lfloor s/p \rfloor + 1}^{j_2} 2^{j-1} \sum_{k=0}^{2j-1} \beta_{j,k}^2 1_{\{|\beta_{j,k}| < 2\lambda_j\}}.
\]

We have

\[
e_{4,1} \leq C \sum_{j=\lfloor s/p \rfloor + 1}^{j_2} 2^{j \lambda_j^2} = C \frac{\ln n}{n} \sum_{j=\lfloor s/p \rfloor + 1}^{j_2} 2^{j(1+2\delta+2d)} \leq C \frac{\ln n}{n} 2^{j_2(1+2\delta+2d)}
\]

\[
\leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)}.
\]

For \( r \geq 1 \) and \( p \geq 2 \), since \( B_{p,r}^s(M) \subseteq B_{2,\infty}^s(M) \), we have

\[
e_{4,2} \leq \sum_{j=\lfloor s/p \rfloor + 1}^{\infty} \sum_{k=0}^{2j-1} \beta_{j,k}^2 \leq C 2^{-2j_2 s} \leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)}.
\]
Wavelet estimation of the regression function and its derivatives in an errors-in-variables model

For $r \geq 1$, $p \in [1, 2)$ and $s > (2\delta + 2d + 1)/p$, since $B^{r}_{p, r}(M) \subseteq B^{s+1/2-1/p}_{2, \infty}(M)$ and $(2 - p)(2s + 2\delta + 2d + 1)/2 + (s + 1/2 - 1/p + \delta + d - 2(\delta + d)/p)p = 2s$, we have

$$e_{4,2} \leq C \sum_{j=j_2+1}^{j_1} \lambda_j^{2-p} \sum_{k=0}^{2^j-1} |\beta_{j,k}|^p = C \left( \frac{\ln n}{n} \right)^{(2-p)/2} \sum_{j=j_2+1}^{j_1} 2^{j(2-p)} \sum_{k=0}^{2^j-1} |\beta_{j,k}|^p \leq C \left( \frac{\ln n}{n} \right)^{(2-p)/2} \sum_{j=j_2+1}^{\infty} 2^{j(\delta+d)(2-p)} 2^{-j(s+1/2-1/p)p} \leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)}.$$

So, for $r \geq 1$, $\{p \geq 2$ and $s > 0\}$ or $\{p \in [1, 2)$ and $s > (2\delta + 2d + 1)/p\}$,

$$e_4 \leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)}.$$

It follows from (6.15), (6.16), (6.18) and (6.19) that

$$S \leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)}.$$

Combining (6.12), (6.13), (6.14) and (6.20), we have, for $r \geq 1$, $\{p \geq 2$ and $s > 0\}$ or $\{p \in [1, 2)$ and $s > (2\delta + 2d + 1)/p\}$,

$$\mathbb{E} \left( \int_0^1 (\hat{f}^H(x) - f^{(d)}(x))^2 \, dx \right) \leq C \left( \frac{\ln n}{n} \right)^{2s/(2s+2\delta+2d+1)}.$$

The proof of Theorem 5.2 is complete.

□

This work is supported by ANR grant NatImages, ANR-08-EMER-009.

References

Wavelet estimation of the regression function and its derivatives in an errors-in-variables model.