Ann Inst Stat Math (2011) 63:405–429
DOI 10.1007/s10463-009-0232-6

On adaptive wavelet estimation of a quadratic functional from a deconvolution problem

Christophe Chesneau

Received: 28 April 2008 / Revised: 16 October 2008 / Published online: 23 April 2009

Abstract We observe a stochastic process where a convolution product of an unknown function and a known function is corrupted by Gaussian noise. We wish to estimate the squared $\mathbb{L}^2$-norm of the unknown function from the observations. To reach this goal, we develop adaptive estimators based on wavelet and thresholding. We prove that they achieve (near) optimal rates of convergence under the mean squared error over a wide range of smoothness classes.

Keywords Deconvolution · Quadratic functionals · Adaptive curve estimation · Wavelets · Global thresholding

1 Motivation

We observe the stochastic process $\{Y(t);\ t \in [0,1]\}$ where

$$\mathrm{d}Y(t) = \left(\int_0^1 f(t-u)g(u)\mathrm{d}u \right) \mathrm{d}t + n^{-1/2}\mathrm{d}W(t), \quad t \in [0,1], \quad n \in \mathbb{N}^*, \tag{1}$$

$\{W(t);\ t \in [0,1]\}$ is a (non-observed) standard Brownian motion, f is an unknown one-periodic function such that $\int_0^1 f^2(t)\mathrm{d}t < \infty$, and g is a known one-periodic function such that $\int_0^1 g^2(t)\mathrm{d}t < \infty$. The goal is to estimate f, or a quantity depending on f, from $\{Y(t);\ t \in [0,1]\}$. The convolution model (1) illustrates the action of a linear time-invariant system on an input signal f when the data are corrupted with additional

C. Chesneau (✉)
Laboratoire de Mathématiques Nicolas Oresme, Université de Caen Basse-Normandie,
Campus II, Science 3, 14032 Caen, France
e-mail: chesneau@math.unicaen.fr

noise. See, for instance, Bertero and Boccacci (1998) and Neelamani et al. (2004). This is a standard inverse problem in the field of function estimation. For related results on (1), we refer to Cavalier and Tsybakov (2002), Cavalier et al. (2004), Johnstone et al. (2004) and Cavalier (2008). Extensions of (1) can be found in Willer (2005), Cavalier and Raimondo (2007) and Pensky and Sapatinas (2008).

In the literature, the main effort was spent on producing adaptive wavelet estimators of the function f. See, for instance, Cavalier and Tsybakov (2002); Johnstone et al. (2004) and Chesneau (2008). In this paper, we focus our attention on a different problem: the estimation of the squared $\mathbb{L}^2$-norm of f defined by

$$\|f\|_2^2 = \int_0^1 f^2(t)\mathrm{d}t.$$

The problem of estimating such a value is closely connected to the construction of confidence balls in nonparametric function estimation. It has already been investigated for a wide variety of models (density, regression, Gaussian model in white noise, density derivatives, ...). See, for instance, Bickel and Ritov (1988), Donoho and Nussbaum (1990), Kerkyacharian and Picard (1996), Efromovich and Low (1996), Gayraud and Tribouley (1999), Johnstone (2001a,b), Laurent (2005), Cai and Low (2005, 2006) and Rivoirard and Tribouley (2008).

To the best of our knowledge, the estimation of $\|f\|_2^2$ from a convolution problem has been firstly studied in Butucea (2007). The model considered is different to (1) and can be described as follows: i.i.d. random variables X_i, $i = 1, \dots, n$, having unknown density f are observed with additive i.i.d. noise, independent of the X_i's, with known probability density g. To estimate $\|f\|_2^2$, a kernel estimator is developed. It enjoys good statistical properties under various assumptions on g. However, it is not adaptive.

In this paper, we consider a simpler convolution model, (1), but we focus on the adaptive estimation of $\|f\|_2^2$. To reach this goal, we use wavelet estimators. As in Butucea (2007, Theorem 4), we distinguish two cases according to the nature of g: the *ordinary smooth case* where the Fourier coefficients of g decrease in a polynomial fashion, and the *supersmooth case* where they decrease in an exponential fashion.

For the ordinary smooth case, we develop an adaptive wavelet estimator $\widehat{Q}_n$ based on the global thresholding. The idea is the following: we decompose $\|f\|_2^2$ by using an appropriate wavelet basis, we estimate the associated coefficients via natural estimators, then, at each level of the wavelets, we keep all of these estimators if, and only if, the corresponding l_2 norm is greater than a fixed threshold. This technic has been initially developed by Gayraud and Tribouley (1999) for the estimation of $\|f\|_2^2$ in the standard Gaussian white noise model.

For the supersmooth case, $\widehat{Q}_n$ is not adapted. We develop another adaptive wavelet estimator Q_n^*.

We evaluate their theoretical performances via the asymptotic minimax approach under the mean squared error over Besov balls. If Q_n denotes an estimator of $\|f\|_2^2$ and $B^s_{2,\infty}(M)$ denotes the Besov balls, we aim to determine the smallest rates of convergence φ_n such that

$$\lim_{n\to\infty} \varphi_n^{-1} \sup_{f\in B^s_{2,\infty}(M)} \mathbb{E}\left(\left(Q_n - \|f\|_2^2\right)^2\right) < \infty.$$

For the ordinary smooth case, we prove that $\widehat{Q}_n$ achieves (near) optimal rates of convergence. For the supersmooth case, we prove that Q_n^* achieves (exactly) optimal rates of convergence. The upper bounds follow from a suitable decomposition of the risk and some technical inequalities (concentration, moment, …). The lower bounds are applications of a specific theorem established by Tsybakov (2004).

The paper is organized as follows. In Sect. 2, we present wavelets and Besov balls. Section 3 clarifies the assumptions made on the Fourier coefficients of g and introduces some intermediate estimators. The main estimators of the study are defined in Sect. 4. Section 5 is devoted to the results. Perspectives and open questions are set in Sect. 6. The proofs of the results are postponed in Sect. 7.

2 Wavelets and Besov balls

2.1 Wavelets

We consider an orthonormal wavelet basis generated by dilations and translations of a "father" Meyer-type wavelet ϕ and a "mother" Meyer-type wavelet ψ. The features of such wavelets are that the Fourier transforms of ϕ and ψ have bounded support. More precisely, we have supp $(F(\phi)) \subset [-4\pi 3^{-1}, 4\pi 3^{-1}]$ and supp $(F(\psi)) \subset [-8\pi 3^{-1}, -2\pi 3^{-1}] \cup [2\pi 3^{-1}, 8\pi 3^{-1}]$, where, for any function $h \in \mathbb{L}^1([0,1])$, $F(h)$ denotes the Fourier transform of h defined by $F(h)(l) = \int_0^1 h(x)e^{-2i\pi l x}dx$. Moreover, for any $l \in [-2\pi, -\pi] \cup [\pi, 2\pi]$, there exists a constant $c > 0$ such that $|F(\psi)(l)| \geq c$. For further details about Meyer-type wavelets, see Walter (1994) and Zayed and Walter (1996).

For the purposes of this paper, we use the periodized wavelet bases on the unit interval. For any $x \in [0,1]$, any integer j and any $k \in \{0, \dots, 2^j - 1\}$, let $\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k)$ and $\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$ be the elements of the wavelet basis, and

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

their periodised versions. There exists an integer τ such that the collection ζ defined by $\zeta = \{\phi^{\text{per}}_{\tau,k}, k = 0, \dots, 2^\tau - 1;\ \psi^{\text{per}}_{j,k}, j = \tau, \dots, \infty,\ k = 0, \dots, 2^j - 1\}$ constitutes an orthonormal basis of $\mathbb{L}^2_{\text{per}}([0,1])$, the set of square integrable one-periodic functions on $[0,1]$. In what follows, the superscript "per" will be suppressed from the notations for convenience.

Let m be an integer such that $m \geq \tau$. A function $f \in \mathbb{L}^2_{\text{per}}([0,1])$ can be expanded into a wavelet series as

$$f(x) = \sum_{k=0}^{2^m-1} \alpha_{m,k}\phi_{m,k}(x) + \sum_{j=m}^{\infty}\sum_{k=0}^{2^j-1} \beta_{j,k}\psi_{j,k}(x), \quad x \in [0,1],$$

where $\alpha_{m,k} = \int_0^1 f(t)\phi_{m,k}(t)\mathrm{d}t$ and $\beta_{j,k} = \int_0^1 f(t)\psi_{j,k}(t)\mathrm{d}t$. For further details about wavelet bases on the unit interval, we refer to Cohen et al. (1993).

2.2 Besov balls

Let $M \in (0,\infty)$ and $s \in (0,\infty)$. We say that a one-periodic function f belongs to the Besov balls $B^s_{2,\infty}(M)$ if, and only if, there exists a constant $M_* > 0$ such that $||f||_2^2 \leq M_*$ and the associated wavelet coefficients satisfy

$$\sup_{j\geq\tau} 2^{2js} \sum_{k=0}^{2^j-1} \beta_{j,k}^2 \leq M_*.$$

These sets contain all the Besov balls $B^s_{p,\infty}(M)$ with $p \geq 2$. See, for instance Meyer (1992).

3 Preliminary study

3.1 Assumptions on g; ordinary smooth and supersmooth cases

- *Assumption* (A_g)*: ordinary smooth case*. We suppose that there exist three constants, $c > 0$, $C > 0$ and $\delta > 2^{-1}$, such that, for any $l \in (-\infty,-1] \cup [1,\infty)$, the Fourier coefficient of g, i.e. $\mathcal{F}(g)(l)$, satisfies

$$c|l|^{-\delta} \leq |\mathcal{F}(g)(l)| \leq C|l|^{-\delta}. \tag{2}$$

For example, the square integrable one-periodic function g defined by $g(x) = \sum_{m\in\mathbb{Z}} \mathrm{e}^{-|x+m|}$, $x \in [0,1]$, satisfies (A_g). Indeed, for any $l \in \mathbb{Z}$, we have $\mathcal{F}(g)(l) = 2\left(1+4\pi^2l^2\right)^{-1}$. Hence, for any $l \in (-\infty,-1] \cup [1,\infty)$, $\mathcal{F}(g)(l)$ satisfies (2) with $c = 2(1+4\pi^2)^{-1}$, $C = (2\pi^2)^{-1}$ and $\delta = 2$.

- *Assumption* (A_g^*)*: supersmooth case*. We suppose that there exist four constants, $c > 0$, $C > 0$, $a > 0$ and $b > 0$, such that, for any $l \in (-\infty,-1] \cup [1,\infty)$, the Fourier coefficient of g, i.e. $\mathcal{F}(g)(l)$, satisfies

$$c\mathrm{e}^{-a|l|^b} \leq |\mathcal{F}(g)(l)| \leq C\mathrm{e}^{-a|l|^b}. \tag{3}$$

For example, the square integrable one-periodic function g defined by $g(x) = \sum_{m\in\mathbb{Z}} \mathrm{e}^{-(x+m)^2/2}$, $x \in [0,1]$, satisfies (A_g^*). Indeed, for any $l \in \mathbb{Z}$, we have $\mathcal{F}(g)(l) = \sqrt{2\pi}\mathrm{e}^{-2\pi^2l^2}$. Hence, for any $l \in (-\infty,-1] \cup [1,\infty)$, $\mathcal{F}(g)(l)$ satisfies (3) with $c = C = \sqrt{2\pi}$, $a = 2\pi^2$ and $b = 2$.

Further details and examples about these cases can be found in Pensky and Vidakovic (1999) and Fan and Koo (2002).

3.2 Preliminary to the estimation of $\|f\|_2^2$

Thanks to the orthonormality of the wavelet basis, for any integer $m \geq \tau$, the unknown $\|f\|_2^2$ can be decomposed as

$$\|f\|_2^2 = \sum_{k=0}^{2^m-1} \alpha_{m,k}^2 + \sum_{j=m}^{\infty} \sum_{k=0}^{2^j-1} \beta_{j,k}^2.$$

Thus, the first step to estimate $\|f\|_2^2$ consists in estimating the unknown coefficients $(\alpha_{m,k}^2)_k$ and $(\beta_{j,k}^2)_{j,k}$. In what follows, we investigate the estimation of $(\beta_{j,k}^2)_{j,k}$. First of all, we use the singular value decomposition (SVD) of the convolution operator $T(f)(t) = (f \star g)(t) = \int_0^1 f(t-u)g(u)\mathrm{d}u$. We aim to write the model (1) in the Fourier domain. For any $l \in \mathbb{Z}$, we have $F(T(f))(l) = F(f)(l)F(g)(l)$. Therefore, it we set $y_l = \int_0^1 \mathrm{e}^{-2\pi ilt}\mathrm{d}Y(t)$, $f_l = F(f)(l)$, $g_l = F(g)(l)$, and $e_l = \int_0^1 \mathrm{e}^{-2\pi ilt}\mathrm{d}W(t)$, it follows from (1) that

$$y_l = f_l g_l + n^{-1/2} e_l. \tag{4}$$

The Parseval theorem and the translation theorem of the Fourier transform give

$$\begin{aligned}\sum_{l\in\mathbb{Z}} y_l g_l^{-1} F(\psi_{j,k})(l) &= \sum_{l\in\mathbb{Z}} f_l F(\psi_{j,k})(l) + n^{-1/2} \sum_{l\in\mathbb{Z}} g_l^{-1} F(\psi_{j,k})(l) e_l \\ &= \beta_{j,k} + n^{-1/2} \sum_{l\in\mathbb{Z}} g_l^{-1} \mathrm{e}^{2i\pi lk/2^j} F(\psi_{j,0})(l) e_l.\end{aligned}$$

Since $n^{-1/2} \sum_{l\in\mathbb{Z}} g_l^{-1} \mathrm{e}^{2i\pi lk/2^j} F(\psi_{j,0})(l) e_l \sim \mathcal{N}\left(0, n^{-1} \sum_{l\in\mathbb{Z}} |g_l|^{-2} |F(\psi_{j,0})(l)|^2\right)$, we have

$$\widehat{\beta}_{j,k} = \sum_{l\in\mathbb{Z}} y_l g_l^{-1} F(\psi_{j,k})(l) \sim \mathcal{N}\left(\beta_{j,k}, n^{-1} \sum_{l\in\mathbb{Z}} |g_l|^{-2} |F(\psi_{j,0})(l)|^2\right).$$

Therefore, the random variable $\widehat{\theta}_{j,k}$ defined by

$$\widehat{\theta}_{j,k} = \widehat{\beta}_{j,k}^2 - n^{-1} \sum_{l\in\mathbb{Z}} |g_l|^{-2} |F(\psi_{j,0})(l)|^2$$

is an unbiased (and a natural) estimator of $\beta_{j,k}^2$. We are now in the position to describe the main estimators of the study.

4 Estimators

According to the ordinary smooth case and the supersmooth case, we describe two different estimators of $Q(f) = \|f\|_2^2$. We use the notations introduced in Sect. 3.2.

4.1 Estimator: the ordinary smooth case

Suppose that (A_g) is satisfied. Let j_0 and j_1 be two integers such that

$$2^{-1}\left(n(\log n)^{-1}\right)^{1/(4\delta+1)} < 2^{j_0} \leq \left(n(\log n)^{-1}\right)^{1/(4\delta+1)}$$

and

$$2^{-1}n^{1/(2\delta+1/2)} < 2^{j_1} \leq n^{1/(2\delta+1/2)}.$$

For any real number a, set $(a)_+ = \max(a, 0)$. Let κ be a positive real number. We define the thresholding estimator $\widehat{Q}_n$ by

$$\widehat{Q}_n = \sum_{k=0}^{2^{j_0}-1}\left(\widehat{\alpha}_{j_0,k}^2 - n^{-1}\eta_{j_0}^2\right) + \sum_{j=j_0}^{j_1}\left(\sum_{k=0}^{2^j-1}(\widehat{\beta}_{j,k}^2 - n^{-1}\sigma_j^2) - \kappa n^{-1}\sigma_j^2\left(j2^j\right)^{1/2}\right)_+, \tag{5}$$

where

$$\widehat{\alpha}_{j_0,k} = \sum_{l\in\mathcal{D}_{j_0}} y_l g_l^{-1} F(\phi_{j_0,k})(l), \quad \widehat{\beta}_{j,k} = \sum_{l\in\mathcal{C}_j} y_l g_l^{-1} F(\psi_{j,k})(l),$$
$$\eta_{j_0}^2 = \sum_{l\in\mathcal{D}_{j_0}} |g_l|^{-2} F^2(\phi_{j_0,0})(l), \quad \sigma_j^2 = \sum_{l\in\mathcal{C}_j} |g_l|^{-2}|F(\psi_{j,0})(l)|^2. \tag{6}$$

Here, for any $k \in \{0, \ldots, 2^j - 1\}$, $\mathcal{D}_{j_0}$ denotes the support of $F(\phi_{j_0,k})(l)$ and $\mathcal{C}_j$ the support of $F(\psi_{j,k})(l)$.

Let us briefly explain the construction of $\widehat{Q}_n$. Firstly, we estimate the approximation term, $\sum_{k=0}^{2^{j_0}-1}\alpha_{j_0,k}^2$, by $\sum_{k=0}^{2^{j_0}-1}(\widehat{\alpha}_{j_0,k}^2 - n^{-1}\eta_{j_0}^2)$. Secondly, we estimate the "detail term", $\sum_{j=j_0}^{j_1}\sum_{k=0}^{2^j-1}\beta_{j,k}^2$, via the global thresholding rule described as follows. At each level $j \in \{j_0, \ldots, j_1\}$, we keep $\sum_{k=0}^{2^j-1}(\widehat{\beta}_{j,k}^2 - n^{-1}\sigma_j^2) - \kappa n^{-1}\sigma_j^2\left(j2^j\right)^{1/2}$ if, and only if, we have $\sum_{k=0}^{2^j-1}(\widehat{\beta}_{j,k}^2 - n^{-1}\sigma_j^2) \geq \kappa n^{-1}\sigma_j^2\left(j2^j\right)^{1/2}$. Then we sum all the selected terms. The idea is to only estimate the unknown coefficients $\beta_{j,k}^2$ which characterized completely $\| f \|_2^2$. As mentioned in Sect. 1, the global thresholding technic has been initially developed by Gayraud and Tribouley (1999) for the estimation of the squared $\mathbb{L}^2$-norm of f in the standard Gaussian white noise model.

4.2 Estimator: the supersmooth case

Suppose that (A_g^*) is satisfied. In this case, the estimator $\widehat{Q}_n$ defined by (5) is not appropriate. One can prove that, in our minimax framework, it achieves a "suboptimal" bound. For this reason, we propose another estimator. It is described as follows. Let j_* be an integer such that

$$2^{-1}(16\pi 3^{-1}a)^{-1} (\log n)^{1/b} < 2^{j_*} \leq (16\pi 3^{-1}a)^{-1/b} (\log n)^{1/b}.$$

We define the linear (or projection) estimator Q_n^* by

$$Q_n^* = \sum_{k=0}^{2^{j_*}-1} \left(\widehat{\alpha}_{j_*,k}^2 - n^{-1}\eta_{j_*}^2\right), \tag{7}$$

where $\widehat{\alpha}_{j_*,k}$ and η_{j_*} are defined by (6) with j_* instead of j_0. The idea of Q_n^* is to estimate the approximation term $\sum_{k=0}^{2^{j_*}-1} \alpha_{j_*,k}^2$ of $\| f \|_2^2$ (until the level j_*), and to neglect the detail term.

Remark The consideration of two different estimators for each case (ordinary smooth and supersmooth) is standard for the adaptive wavelet estimation of f from convolution model. See, for instance, Pensky and Vidakovic (1999) and Fan and Koo (2002). However, to the best of our knowledge, this is new for the adaptive wavelet estimation of $\| f \|_2^2$.

Important remark The estimators $\widehat{Q}_n$ and Q_n^* do not require any a priori knowledge on f in their constructions; they are adaptive.

5 Results

This section is divided into two parts. The first part is devoted to the adaptive estimation of $|| f ||_2^2$ for the ordinary smooth case (A_g), and the second part, for the supersmooth case (A_g^*).

5.1 Results: ordinary smooth case

Under (A_g), Theorem 1 below determines the rates of convergence achieved by $\widehat{Q}_n$ under the mean squared error over the Besov balls $B_{2,\infty}^s(M)$.

Theorem 1 *Consider the convolution model defined by (1). Suppose that (A_g) is satisfied. Let $\widehat{Q}_n$ be the estimator defined by (5) with a large enough κ. Then there exists a constant $C > 0$ such that, for n large enough,*

$$\sup_{f \in B^s_{2,\infty}(M)} \mathbb{E}\left(\left(\widehat{Q}_n - \|f\|_2^2\right)^2\right) \leq C\varphi_n,$$

where

$$\varphi_n = \begin{cases} n^{-1}, & \textit{when} \quad s > \delta + 4^{-1}, \\ \left(n(\log n)^{-1/2}\right)^{-4s/(2s+2\delta+1/2)}, & \textit{when} \quad s \in (0, \delta + 4^{-1}], \end{cases}$$

i.e., $\varphi_n = \max\left(n^{-1}, \left(n(\log n)^{-1/2}\right)^{-4s/(2s+2\delta+1/2)}\right)$.

The proof of Theorem 1 uses several auxiliary results. In particular, we establish a functional upper bound for the risk similar to Proposition 3 of Rivoirard and Tribouley (2008). It is obtained by proving two technical concentration inequalities satisfied by the estimators $(\widehat{\beta}_{j,k})_{k=0,\ldots,2^j-1}$.

The result of Theorem 1 raises the following question: are the rates of convergence φ_n the optimal? Theorem 2 below provides the answer by determining the minimax lower bounds.

Theorem 2 *Consider the convolution model defined by (1). Suppose that (A_g) is satisfied. Then there exists a constant $c > 0$ such that, for n large enough,*

$$\inf_{\widetilde{Q}_n} \sup_{f \in B^s_{2,\infty}(M)} \mathbb{E}\left(\left(\widetilde{Q}_n - \|f\|_2^2\right)^2\right) \geq c\varphi_n^*,$$

where $\inf_{\widetilde{Q}_n}$ *denotes the infimum over all the possible estimators of* $\|f\|_2^2$, *and*

$$\varphi_n^* = \begin{cases} n^{-1}, & \textit{when} \quad s > \delta + 4^{-1}, \\ n^{-4s/(2s+2\delta+1/2)}, & \textit{when} \quad s \in (0, \delta + 4^{-1}]. \end{cases}$$

The proof of Theorem 2 is based on a general theorem established by Tsybakov (2004).

The results of Theorems 1 and 2 show that, under (A_g), $\widehat{Q}_n$ is near optimal under the mean squared error over the Besov balls $B^s_{2,\infty}(M)$. "Near" is due to the case $s \in (0, \delta + 4^{-1}]$ where there is an extra logarithmic term.

5.2 Results: supersmooth case

Under (A_g^*), Theorem 1 below determines the rates of convergence achieved by Q_n^* under the mean squared error over the Besov balls $B^s_{2,\infty}(M)$.

Theorem 3 *Consider the convolution model defined by* (1)*. Suppose that* (A_g^*) *is satisfied. Let* Q_n^* *be the estimator defined by* (7)*. Then there exists a constant* $C > 0$ *such that, for n large enough,*

$$\sup_{f \in B^s_{2,\infty}(M)} \mathbb{E}\left(\left(Q_n^* - \|f\|_2^2\right)^2\right) \leq C(\log n)^{-4s/b}.$$

The proof of Theorem 3 is based on a good decomposition of the risk and technical inequalities (l_2 norm, moment, ...).

As in Theorem 2, but under (A_g^*), Theorem 4 below determines the minimax lower bounds of the model.

Theorem 4 *Consider the convolution model defined by* (1)*. Suppose that* (A_g^*) *is satisfied. Then there exists a constant* $c > 0$ *such that, for n large enough,*

$$\inf_{\widetilde{Q}_n} \sup_{f \in B^s_{2,\infty}(M)} \mathbb{E}\left(\left(\widetilde{Q}_n - \|f\|_2^2\right)^2\right) \geq c(\log n)^{-4s/b},$$

where $\inf_{\widetilde{Q}_n}$ *denotes the infimum over all the possible estimators of* $\|f\|_2^2$.

The proof of Theorem 4 is based on a general theorem proved by Tsybakov (2004).

Theorems 3 and 4 show that, under (A_g^*), $\upsilon_n = (\log n)^{-4s/b}$ is the minimax rate of convergence under the mean squared error over the Besov balls $B^s_{2,\infty}(M)$. Moreover the adaptive estimator Q_n^* is optimal.

6 Conclusion and open questions

Our rates of convergence (see Table 1) are similar to those determined in Butucea (2007, Theorem 4) which considers the density convolution, a non-adaptive estimator based on kernel, and the risk: $R(h, f) = \mathbb{E}(|h - \|f\|_2^2|)$, $h \in \mathbb{R}$. In comparison to this result, we consider another convolution model, simpler to manipulate, but we provide a contribution to the adaptive estimation of $\|f\|_2^2$.

Some perspectives and open questions are presented below:

– A possible extension of this work is to adapt our estimators to other convolution models. For instance, the density convolution model: i.i.d. random variables X_i,

Table 1 Minimax results of $\widehat{Q}_n$ and Q_n^* in the ordinary smooth and supersmooth cases

Case	Ordinary smooth	Supersmooth
Estimator	$\widehat{Q}_n$	Q_n^*
Upper bound	$\max\left(n^{-1}, \left(n(\log n)^{-1/2}\right)^{-4s/(2s+2\delta+1/2)}\right)$	$(\log n)^{-4s/b}$
Optimality	$\begin{cases} \text{optimal}, & \text{when } s > \delta + 4^{-1}, \\ \text{near optimal}, & \text{when } s \in (0, \delta + 4^{-1}]. \end{cases}$	Optimal

$i = 1, \ldots, n$, having unknown density f are observed with additive i.i.d. noise, independent of the X_i's, with known probability density g. Another interesting convolution model is the one studied in Pensky and Sapatinas (2008): one observes n random variables $\{Y_i;\ i = 1, \ldots, n\}$ where $Y_i = \int_0^1 f(i/n - u)g(u)\mathrm{d}u + \epsilon_i$, $i = 1, \ldots, n$, $\{\epsilon_i;\ i = 1, \ldots, n\}$ are (non-observed) i.i.d. Gaussian standard random variables, f is an unknown one-periodic function such that $\int_0^1 f^2(t)\mathrm{d}t < \infty$, and g is a known one-periodic function such that $\int_0^1 g^2(t)\mathrm{d}t < \infty$. Despite the "equivalences" which exist between these models and (1), they are more difficult to manipulate. In particular, they can not be directly written as a sequence model of the form (4) with "$f_l \times g_l$" and/or "e_l" independent $\mathcal{N}(0, 1)$. For instance, this last point is an obstacle to the application of the Cirelson inequality used in the proof of Theorem 1 (see Lemma 3). Thus, the estimation of $\|f\|_2^2$ from these models via our estimators requires technical tools certainly not used in our study.
- Our minimax results can be extended to the Besov balls $B^s_{p,\infty}(M)$ with $p \geq 2$. However, the case $p < 2$ is beyond the scope of the paper. As underlined in Cai and Low (2006) for the standard Gaussian sequence model (p. 2300, lines 18–19): "*The sparse case where $p < 2$ presents some major new difficulties which requires a novel approach for the construction of adaptive procedures*". In response to these difficulties, Cai and Low (2006) have developed a sophisticated adaptive estimator which uses both block thresholding and term-by-term thresholding. Taking the minimax approach under the mean squared error over $B^s_{p,\infty}(M)$, it is (near) optimal for $p \geq 1$. An open question is: for the model (1), more complex than the standard Gaussian sequence model due to the presence of g, how can we adapt this estimator in order to keep these optimal properties over $B^s_{p,\infty}(M)$ for any $p \geq 1$?
- And, finally: can we construct only one adaptive estimator Q_n which are (near) optimal for both ordinary smooth case and supersmooth case?

All these aspects need further investigation that we leave for a future work.

7 Proofs

In this section, c and C denote positive constants which can take different values for each mathematical term. They are independent of f and n.

Proof of Theorem 1 Proposition 1 below presents a "functional" upper bound for the mean squared risk of $\widehat{Q}_n$ without any assumption on the smoothness of f.

Proposition 1 *Let $\widehat{Q}_n$ be the estimator defined by (5) with a large enough κ. There exist two constants $C_1 > 0$ and $C_2 > 0$ such that*

$$\mathbb{E}\left(\left(\widehat{Q}_n - \|f\|_2^2\right)^2\right) \leq C_1(\mathbf{A} + \mathbf{B} + \mathbf{C} + \mathbf{D}),$$

where

$$\mathbf{A} = n^{-2}2^{(1+4\delta)j_0}, \quad \mathbf{B} = n^{-1}\|f\|_2^2, \quad \mathbf{C} = \left(\sum_{j=j_1+1}^{\infty}\sum_{k=0}^{2^j-1}\beta_{j,k}^2\right)^2$$

and

$$\mathbf{D} = \left(\sum_{j=j_0}^{j_1}\min\left(\sum_{k=0}^{2^j-1}\beta_{j,k}^2, C_2\kappa 2^{2\delta j}n^{-1}(j2^j)^{1/2}\right)\right)^2.$$

Therefore, to prove Theorem 1, it is enough to investigate the upper bounds for the terms **A**, **B**, **C** and **D** defined in Proposition 1.

The upper bound for **A**. Thanks to the definition of j_0, we have

$$\mathbf{A} = n^{-2}2^{(1+4\delta)j_0} \le n^{-1} \le \max\left(n^{-1}, \left(n(\log n)^{-1/2}\right)^{-4s/(2s+2\delta+1/2)}\right) = \varphi_n.$$

The upper bound for **B**. Since $f \in B^s_{2,\infty}(M)$, we have $\|f\|_2^2 \le M_*$. Therefore

$$\mathbf{B} = n^{-1}\|f\|_2^2 \le C\max\left(n^{-1}, \left(n(\log n)^{-1/2}\right)^{-4s/(2s+2\delta+1/2)}\right) = C\varphi_n.$$

The upper bound for **C**. Since $f \in B^s_{2,\infty}(M)$, we have

$$\mathbf{C} = \left(\sum_{j=j_1+1}^{\infty}\sum_{k=0}^{2^j-1}\beta_{j,k}^2\right)^2 \le C\left(\sum_{j=j_1+1}^{\infty}2^{-2js}\right)^2 \le C2^{-4j_1 s} \le Cn^{-4s/(2\delta+1/2)}$$
$$\le C\max\left(n^{-1}, \left(n(\log n)^{-1/2}\right)^{-4s/(2s+2\delta+1/2)}\right) = C\varphi_n.$$

The upper bound for **D**. We distinguish the case $s > \delta + 4^{-1}$ and the case $s \in (0, \delta + 4^{-1}]$.

– *The case $s > \delta + 4^{-1}$*. Since $f \in B^s_{2,\infty}(M)$, the definition of j_0 yields

$$\mathbf{D} = \left(\sum_{j=j_0}^{j_1}\min\left(\sum_{k=0}^{2^j-1}\beta_{j,k}^2, C_2\kappa 2^{2\delta j}n^{-1}(j2^j)^{1/2}\right)\right)^2 \le C\left(\sum_{j=j_0}^{j_1}\sum_{k=0}^{2^j-1}\beta_{j,k}^2\right)^2$$
$$\le C\left(\sum_{j=j_0}^{j_1}2^{-2js}\right)^2 \le C2^{-4j_0 s} \le C\left(n(\log n)^{-1}\right)^{-4s/(4\delta+1)} \le Cn^{-1}.$$

– *The case* $s \in (0, \delta + 4^{-1}]$. Let j_2 be an integer such that

$$2^{-1}\left(n(\log n)^{-1/2}\right)^{1/(2s+2\delta+1/2)} < 2^{j_2} \leq \left(n(\log n)^{-1/2}\right)^{1/(2s+2\delta+1/2)}.$$

Notice that $j_0 \leq j_2 \leq j_1$. We have the following decomposition

$$\mathbf{D} = (\mathbf{L} + \mathbf{M})^2,$$

where

$$\mathbf{L} = \sum_{j=j_0}^{j_2} \min\left(\sum_{k=0}^{2^j-1} \beta_{j,k}^2, C_2\kappa 2^{2\delta j} n^{-1}(j2^j)^{1/2}\right)$$

and

$$\mathbf{M} = \sum_{j=j_2+1}^{j_1} \min\left(\sum_{k=0}^{2^j-1} \beta_{j,k}^2, C_2\kappa 2^{2\delta j} n^{-1}(j2^j)^{1/2}\right).$$

Let us study the upper bounds for $\mathbf{L}$ and $\mathbf{M}$ in turn.

It follows from the definition of j_2 that

$$\mathbf{L} \leq Cn^{-1}\sum_{j=j_0}^{j_2} j^{1/2}2^{(2\delta+1/2)j} \leq Cn^{-1}(\log n)^{1/2}2^{(2\delta+1/2)j_2}$$
$$\leq C\left(n(\log n)^{-1/2}\right)^{-2s/(2s+2\delta+1/2)}.$$

Since $f \in B^s_{2,\infty}(M)$, the definition of j_2 yields

$$\mathbf{M} \leq \sum_{j=j_2+1}^{j_1}\sum_{k=0}^{2^j-1} \beta_{j,k}^2 \leq C\sum_{j=j_2+1}^{j_1} 2^{-2js} \leq C2^{-2j_2 s}$$
$$\leq C\left(n(\log n)^{-1/2}\right)^{-2s/(2s+2\delta+1/2)}.$$

Putting these two inequalities together, we obtain

$$\mathbf{D} = (\mathbf{L} + \mathbf{M})^2 \leq C\left(n(\log n)^{-1/2}\right)^{-4s/(2s+2\delta+1/2)}.$$

It follows from Proposition 1 and the obtained upper bounds for $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ that

$$\mathbb{E}\left(\left(\widehat{Q}_n - \|f\|_2^2\right)^2\right) \leq C_1(\mathbf{A} + \mathbf{B} + \mathbf{C} + \mathbf{D}) \leq C\varphi_n.$$

Theorem 1 is proved.

Proof of Proposition 1 The proof of Proposition 1 is similar to the proof of Proposition 3 of Rivoirard and Tribouley (2008). If we analyze this proof, Proposition 1 follows from Lemmas 1 and 2 below.

Lemma 1 *If* (A_g) *is satisfied, then here exist two constants* $c > 0$ *and* $C > 0$ *such that, for any integer* $j \geq \tau$*, we have*

$$c2^{2\delta j} \leq \sigma_j^2 \leq C2^{2\delta j}.$$

Moreover, we have $\eta_{j_0}^2 \leq C2^{2\delta j_0}$.

Proof of Lemma 1 Using the inequalities $\sup_{l\in\mathcal{C}_j} |F(\psi_{j,0})(l)|^2 \leq C2^{-j}$, $\mathrm{Card}(\mathcal{C}_j) \leq C2^j$ and (A_g), we have

$$\sigma_j^2 = \sum_{l\in\mathcal{C}_j} |g_l|^{-2}|F(\psi_{j,0})(l)|^2 \leq C2^{-j}\sum_{l\in\mathcal{C}_j}|g_l|^{-2} \leq C\sup_{l\in\mathcal{C}_j}|g_l|^{-2} \leq C2^{2\delta j}.$$

In the same way, we prove that $\eta_{j_0}^2 \leq C2^{2\delta j_0}$.

Recall that, for any $l \in [-2\pi, -\pi] \cup [\pi, 2\pi]$, there exists a constant $c > 0$ such that $|F(\psi)(l)| \geq c$. Therefore, there exists a set $\mathcal{H}_j$ such that $\inf_{l\in\mathcal{H}_j} |F(\psi_{j,0})(l)|^2 \geq c2^{-j}$, $\mathrm{Card}(\mathcal{H}_j) \geq c2^j$ and, thanks to (A_g), $\inf_{l\in\mathcal{H}_j} g_l^{-2} \geq c2^{2\delta j}$. Hence,

$$c2^{2\delta j} \leq c\inf_{l\in\mathcal{H}_j} g_l^{-2} \leq c2^{-j}\sum_{l\in\mathcal{H}_j}|g_l|^{-2} \leq \sum_{l\in\mathcal{C}_j}|g_l|^{-2}|F(\psi_{j,0})(l)|^2 = \sigma_j^2.$$

This ends the proof of Lemma 1.

Lemma 2 below presents two concentration inequalities satisfied by the estimator $\widehat{\beta}_{j,k}$.

Lemma 2 *Let* $\widehat{\beta}_{j,k}$ *be the estimator defined by* (6)*. For* λ *large enough, there exists a constant* $C > 0$ *such that, for any integer* $j \geq \tau$*, we have*

$$\mathbb{P}\left(\left|\sum_{k=0}^{2^j-1}\beta_{j,k}(\widehat{\beta}_{j,k}-\beta_{j,k})\right| \geq \lambda n^{-1/2}j^{1/2}2^{\delta j}\left(\sum_{k=0}^{2^j-1}\beta_{j,k}^2\right)^{1/2}\right) \leq 2\exp\left(-C\lambda^2 j\right) \tag{8}$$

and

$$\mathbb{P}\left(\left(\sum_{k=0}^{2^j-1}(\widehat{\beta}_{j,k}-\beta_{j,k})^2\right)^{1/2} \geq \lambda n^{-1/2}2^{j/2}2^{\delta j}\right) \leq 2\exp\left(-C\lambda^2 2^j\right). \tag{9}$$

Proof of Lemma 2 The first part of the proof is devoted to the bound (8) and the second part, to the bound (9).

– *Proof of the bound (8).* Since the random variable $(e_l)_l$ are i.i.d. with $e_l \sim \mathcal{N}(0, 1)$, we have

$$\sum_{k=0}^{2^j-1} \beta_{j,k}(\widehat{\beta}_{j,k} - \beta_{j,k}) = n^{-1/2} \sum_{l \in C_j} e_l g_l^{-1} \sum_{k=0}^{2^j-1} \beta_{j,k} F(\psi_{j,k})(l) \sim \mathcal{N}\left(0, \theta_j^2\right),$$

where

$$\theta_j^2 = n^{-1} \sum_{l \in C_j} |g_l|^{-2} \left(\sum_{k=0}^{2^j-1} \beta_{j,k} F(\psi_{j,k})(l) \right)^2 .$$

(A_g) yields $\sup_{l \in \mathcal{C}_j} |g_l|^{-2} \leq C 2^{2\delta j}$. It follows from this, the Plancherel theorem and the orthonormality of the wavelet basis that

$$\begin{aligned}
\theta_j^2 &= n^{-1} \sum_{l \in C_j} |g_l|^{-2} \left(F \left(\sum_{k=0}^{2^j-1} \beta_{j,k} \psi_{j,k} \right)(l) \right)^2 \\
&\leq C n^{-1} 2^{2\delta j} \sum_{l \in C_j} \left(F \left(\sum_{k=0}^{2^j-1} \beta_{j,k} \psi_{j,k} \right)(l) \right)^2 \\
&= C n^{-1} 2^{2\delta j} \int_0^1 \left(\sum_{k=0}^{2^j-1} \beta_{j,k} \psi_{j,k}(x) \right)^2 \mathrm{d}x = C n^{-1} 2^{2\delta j} \sum_{k=0}^{2^j-1} \beta_{j,k}^2.
\end{aligned}$$

Applying a standard Gaussian concentration inequality, we obtain

$$\begin{aligned}
&\mathbb{P}\left(\left| \sum_{k=0}^{2^j-1} \beta_{j,k}(\widehat{\beta}_{j,k} - \beta_{j,k}) \right| \geq \lambda n^{-1/2} j^{1/2} 2^{\delta j} \left(\sum_{k=0}^{2^j-1} \beta_{j,k}^2 \right)^{1/2} \right) \\
&\quad \leq 2 \exp\left(- \left(\lambda n^{-1/2} j^{1/2} 2^{\delta j} \left(\sum_{k=0}^{2^j-1} \beta_{j,k}^2 \right)^{1/2} \right)^2 / (2\theta_j^2) \right) \leq 2 \exp\left(-C\lambda^2 j \right).
\end{aligned}$$

This ends the proof of the bound (8).

– *Proof of the bound (9).* To obtain the bound (9), we need the Cirelson inequality presented in Lemma 3 below.

Lemma 3 (Cirelson et al. (1976)) *Let D be a subset of $\mathbb{R}$. Let $(\eta_t)_{t\in D}$ be a centered Gaussian process. If $\mathbb{E}\left(\sup_{t\in D}\eta_t\right) \leq N$ and $\sup_{t\in D}\mathbb{V}(\eta_t) \leq V$ then, for any $x > 0$, we have*

$$\mathbb{P}\left(\sup_{t\in D}\eta_t \geq x + N\right) \leq \exp\left(-x^2/(2V)\right).$$

We have $\widehat{\beta}_{j,k} - \beta_{j,k} = n^{-1/2}\sum_{l\in\mathcal{C}_j} e_l g_l^{-1} F(\psi_{j,k})(l) \sim \mathcal{N}(0, n^{-1}\sigma_j^2)$, where $\sigma_j^2 = \sum_{l\in\mathcal{C}_j}|g_l|^{-2}|F(\psi_{j,0})(l)|^2$. Set $\Omega = \{a = (a_k) \in \mathbb{R};\ \sum_{k=0}^{2^j-1} a_k^2 \leq 1\}$. For any $a \in \Omega$, let $Z(a)$ be the centered Gaussian process defined by

$$Z(a) = \sum_{k=0}^{2^j-1} a_k(\widehat{\beta}_{j,k} - \beta_{j,k}) = n^{-1/2}\sum_{l\in C_j} e_l g_l^{-1}\sum_{k=0}^{2^j-1} a_k F(\psi_{j,k})(l).$$

By an argument of duality, we have $\sup_{a\in\Omega} Z(a) = (\sum_{k=0}^{2^j-1}(\widehat{\beta}_{j,k} - \beta_{j,k})^2)^{1/2}$. Now, let us determine the values of N and V which appeared in the Cirelson inequality (see Lemma 3).

Value of N. The Cauchy–Schwarz inequality and Lemma 1 imply that

$$\mathbb{E}\left(\sup_{a\in\Omega} Z(a)\right) = \mathbb{E}\left(\left(\sum_{k=0}^{2^j-1}(\widehat{\beta}_{j,k} - \beta_{j,k})^2\right)^{1/2}\right) \leq \left(\sum_{k=0}^{2^j-1}\mathbb{E}\left((\widehat{\beta}_{j,k} - \beta_{j,k})^2\right)\right)^{1/2}$$
$$\leq C\left(n^{-1}\sum_{k=0}^{2^j-1}\sigma_j^2\right)^{1/2} \leq Cn^{-1/2}2^{j/2}2^{\delta j}.$$

Hence $N = Cn^{-1/2}2^{j/2}2^{\delta j}$.
Value of V. Using the fact that, for any $l, l' \in \mathbb{Z}$, we have $\mathbb{E}\left(e_l\overline{e_{l'}}\right) = \int_0^1 \mathrm{e}^{-2i\pi(l-l')t}\,\mathrm{d}t = 1_{\{l=l'\}}$, (A_g), the Plancherel inequality and the orthonormality of the wavelet basis, we obtain

$$\sup_{a\in\Omega}\mathbb{V}(Z(a))$$
$$= \sup_{a\in\Omega}\left(\mathbb{E}\left(\sum_{k=0}^{2^j-1}\sum_{k'=0}^{2^j-1} a_k(\widehat{\beta}_{j,k} - \beta_{j,k})\overline{a_{k'}(\widehat{\beta}_{j,k'} - \beta_{j,k'})}\right)\right)$$
$$= n^{-1}\sup_{a\in\Omega}\left(\sum_{k=0}^{2^j-1}\sum_{k'=0}^{2^j-1} a_k a_{k'}\sum_{l\in C_j}\sum_{l'\in C_j} g_l^{-1}F(\psi_{j,k})(l)(\overline{g_{l'}})^{-1}\overline{F(\psi_{j,k'})(l')}\mathbb{E}\left(e_l\overline{e_{l'}}\right)\right)$$
$$= n^{-1}\sup_{a\in\Omega}\left(\sum_{k=0}^{2^j-1}\sum_{k'=0}^{2^j-1} a_k a_{k'}\sum_{l\in C_j}|g_l|^{-2}F(\psi_{j,k})(l)\overline{F(\psi_{j,k'})(l)}\right)$$

$$\leq Cn^{-1}2^{2\delta j}\sup_{a\in\Omega}\left(\sum_{k=0}^{2^j-1}\sum_{k'=0}^{2^j-1}a_k a_{k'}\sum_{l\in C_j}F(\psi_{j,k})(l)\overline{F(\psi_{j,k'})(l)}\right)$$

$$= Cn^{-1}2^{2\delta j}\sup_{a\in\Omega}\left(\sum_{k=0}^{2^j-1}\sum_{k'=0}^{2^j-1}a_k a_{k'}\int_0^1\psi_{j,k}(x)\overline{\psi_{j,k'}(x)}\mathrm{d}x\right)$$

$$= Cn^{-1}2^{2\delta j}\sup_{a\in\Omega}\left(\sum_{k=0}^{2^j-1}a_k^2\right)\leq C2^{2\delta j}n^{-1}.$$

Hence $V = C2^{2\delta j}n^{-1}$. By taking λ large enough and $x = 2^{-1}\lambda n^{-1/2}2^{j/2}2^{\delta j}$, the Cirelson inequality described in Lemma 3 yields

$$\mathbb{P}\left(\left(\sum_{k=0}^{2^j-1}(\widehat{\beta}_{j,k}-\beta_{j,k})^2\right)^{1/2}\geq\lambda n^{-1/2}2^{j/2}2^{\delta j}\right)$$
$$\leq\mathbb{P}\left(\sup_{a\in\Omega}Z(a)\geq x+N\right)\leq\exp\left(-x^2/(2V)\right)\leq 2\exp\left(-C\lambda^2 2^j\right).$$

The bound (9) is proved. The proof of Lemma 2 is complete.

This ends the proof of Proposition 1.

Proof of Theorem 2 We distinguish the case $s > \delta+4^{-1}$ and the case $s \in (0, \delta+4^{-1}]$.

The case $s > \delta + 4^{-1}$. We consider the two functions

$$f_0(x) = 1, \quad f_1(x) = 1 + n^{-1/2}, \quad x \in [0, 1].$$

Clearly, f_0 and f_1 belong to $B^s_{2,\infty}(M)$. Moreover, we have

$$\left|\|f_0\|_2^2 - \|f_1\|_2^2\right| = 2n^{-1/2} + n^{-1} \geq n^{-1/2} = \kappa_n.$$

Now, in order to apply Theorem 2.12 (iii) of Tsybakov (2004), we aim to bound (by a constant C) the chi-square divergence $\chi^2(\mathbb{P}_{f_1}, \mathbb{P}_{f_0})$ defined by

$$\chi^2(\mathbb{P}_{f_1}, \mathbb{P}_{f_0}) = \int\left(\frac{\mathrm{d}\mathbb{P}_{f_1}}{\mathrm{d}\mathbb{P}_{f_0}}\right)^2\mathrm{d}\mathbb{P}_{f_0} - 1, \qquad (10)$$

where $\mathbb{P}_h$ denotes the probability distribution of $\{Y(t);\ t \in [0, 1]\}$ indexed by the function h.

Let $\star$ be the standard convolution product on the unit interval. The Girsanov theorem yields

$$\frac{\mathrm{d}\mathbb{P}_{f_1}}{\mathrm{d}\mathbb{P}_{f_0}} = \exp\left(n\int_0^1 ((f_1\star g)(t)-(f_0\star g)(t))\,\mathrm{d}Y(t)\right.$$
$$\left.-2^{-1}n\int_0^1\left((f_1\star g)^2(t)-(f_0\star g)^2(t)\right)\mathrm{d}t\right)$$
$$=\exp\left(n^{1/2}\int_0^1 g(u)\mathrm{d}u\int_0^1 \mathrm{d}Y(t)-2^{-1}n\left(\int_0^1 g(u)\mathrm{d}u\right)^2(2n^{-1/2}+n^{-1})\right).$$

So, under $\mathbb{P}_{f_0}$, we have

$$\frac{\mathrm{d}\mathbb{P}_{f_1}}{\mathrm{d}\mathbb{P}_{f_0}} = \exp\left(\int_0^1 g(u)\mathrm{d}u\int_0^1 \mathrm{d}W(t)-2^{-1}\left(\int_0^1 g(u)\mathrm{d}u\right)^2\right).$$

Using the equality $\int\exp(2\int_0^1 g(u)\mathrm{d}u\int_0^1 \mathrm{d}W(t))\mathrm{d}\mathbb{P}_{f_0} = \exp(2(\int_0^1 g(u)\mathrm{d}u)^2)$, the Cauchy–Schwarz inequality and the fact that $\int_0^1 g^2(t)\mathrm{d}t < \infty$, we have

$$\int\left(\frac{\mathrm{d}\mathbb{P}_{f_1}}{\mathrm{d}\mathbb{P}_{f_0}}\right)^2\mathrm{d}\mathbb{P}_{f_0} = \exp\left(\left(\int_0^1 g(u)\mathrm{d}u\right)^2\right)\le\exp\left(\int_0^1 g^2(u)\mathrm{d}u\right)\le C.$$

Therefore, there exists a constant $C > 0$ such that

$$\chi^2(\mathbb{P}_{f_1},\mathbb{P}_{f_0}) = \int\left(\frac{\mathrm{d}\mathbb{P}_{f_1}}{\mathrm{d}\mathbb{P}_{f_0}}\right)^2\mathrm{d}\mathbb{P}_{f_0}-1\le C<\infty.$$

It follows from Theorem 2.12 (iii) of Tsybakov (2004) that

$$\inf_{\widetilde{Q}_n}\sup_{f\in B^s_{2,\infty}(M)}\mathbb{E}\left(\left(\widetilde{Q}_n-\|f\|_2^2\right)^2\right)\ge c\kappa_n^2 = cn^{-1} = c\varphi_n^*. \tag{11}$$

The case $s\in(0,\delta+4^{-1}]$. We consider the two functions

$$f_0(x)=0,\quad f_1(x)=\gamma_{j_*}\sum_{k_*=0}^{2^{j_*}} w_{k_*}\psi_{j_*,k_*}(x),\quad x\in[0,1], \tag{12}$$

where j_* is an integer to be chosen below, γ_{j_*} is a quantity to be chosen below, and $w_0,\ldots,w_{2^{j_*}-1}$ are i.i.d. Rademacher random variables (i.e., for any $u\in\{0,\ldots,2^{j_*}-1\}$, $\mathbb{P}(w_u=1)=\mathbb{P}(w_u=-1)=2^{-1}$).

Clearly, f_0 belongs to $B^s_{2,\infty}(M)$. Thanks to the orthogonality of the wavelet basis, for any integer $j\ge\tau$ and any $k\in\{0,\ldots,2^j-1\}$, we have $\beta_{j,k}=\int_0^1 f_1(x)\psi_{j,k}(x)$

$\mathrm{d}x = \gamma_{j_*} w_k$ for $j = j_*$, and 0 otherwise. Thus, if $\gamma_{j_*}^2 = M_* 2^{-j_*(2s+1)}$, then

$$2^{2j_*s} \sum_{k=0}^{2^{j_*}-1} \beta_{j_*,k}^2 = 2^{2j_*s} \gamma_{j_*}^2 \sum_{k=0}^{2^{j_*}-1} w_k^2 = \gamma_{j_*}^2 2^{j_*(2s+1)} = M_*.$$

Therefore, for such a choice of γ_{j_*}, we have $f_1 \in B_{2,\infty}^s(M)$.

Let j_* be an integer such that

$$2^{-1} n^{1/(2s+2\delta+1/2)} \leq 2^{j_*} \leq n^{1/(2s+2\delta+1/2)}. \tag{13}$$

With the previous value of γ_{j_*}, the orthonormality of the wavelet basis gives

$$\begin{aligned}\left| \| f_0 \|_2^2 - \| f_1 \|_2^2 \right| &= \| f_1 \|_2^2 = \gamma_{j_*}^2 \sum_{k=0}^{2^{j_*}-1} w_k^2 = \gamma_{j_*}^2 2^{j_*} \geq C 2^{-2j_*s} \\ &\geq C n^{-2s/(2s+2\delta+1/2)} = \kappa_n.\end{aligned}$$

In what follows, we consider the function f_1 described by (19) with the integer j_* defined by (13) and the quantity $\gamma_{j_*} = M_*^{1/2} 2^{-j_*(s+1/2)}$. In order to apply Theorem 2.12 (iii) of Tsybakov (2004), we aim to bound the chi-square divergence $\chi^2(\mathbb{P}_{f_1}, \mathbb{P}_{f_0})$ (see (10)) by a constant.

Let $\star$ be the standard convolution product. Due to the definitions of f_0, f_1, the random variables $w_0, \ldots, w_{2^{j_*}-1}$, and the Girsanov theorem, we have

$$\begin{aligned}\frac{\mathrm{d}\mathbb{P}_{f_1}}{\mathrm{d}\mathbb{P}_{f_0}} = \prod_{k=0}^{2^{j_*}-1} \Bigg[& 2^{-1} \exp\left(n\gamma_{j_*} \int_0^1 (\psi_{j_*,k} \star g)(t) \mathrm{d}Y(t) - 2^{-1} n \gamma_{j_*}^2 \int_0^1 (\psi_{j_*,k} \star g)^2(t) \mathrm{d}t \right) \\ & + 2^{-1} \exp\left(-n\gamma_{j_*} \int_0^1 (\psi_{j_*,k} \star g)(t) \mathrm{d}Y(t) - 2^{-1} n \gamma_{j_*}^2 \int_0^1 (\psi_{j_*,k} \star g)^2(t) \mathrm{d}t \right) \Bigg].\end{aligned}$$

So, under $\mathbb{P}_{f_0}$, we have

$$\begin{aligned}&\frac{\mathrm{d}\mathbb{P}_{f_1}}{\mathrm{d}\mathbb{P}_{f_0}} \\ &= \prod_{k=0}^{2^{j_*}-1} \Bigg[2^{-1} \exp\left(n^{1/2}\gamma_{j_*} \int_0^1 (\psi_{j_*,k} \star g)(t) \mathrm{d}W(t) - 2^{-1} n \gamma_{j_*}^2 \int_0^1 (\psi_{j_*,k} \star g)^2(t) \mathrm{d}t \right) \\ &\quad + 2^{-1} \exp\left(-n^{1/2}\gamma_{j_*} \int_0^1 (\psi_{j_*,k} \star g)(t) \mathrm{d}W(t) - 2^{-1} n \gamma_{j_*}^2 \int_0^1 (\psi_{j_*,k} \star g)^2(t) \mathrm{d}t \right) \Bigg]\end{aligned}$$

$$= \prod_{k=0}^{2^{j_*}-1} 2^{-1} \exp\left(-2^{-1} n\gamma_{j_*}^2 \int_0^1 (\psi_{j_*,k} \star g)^2(t)\mathrm{d}t\right)$$
$$\times \left[\exp\left(n^{1/2}\gamma_{j_*} \int_0^1 (\psi_{j_*,k} \star g)(t)\mathrm{d}W(t)\right)\right.$$
$$\left. + \exp\left(-n^{1/2}\gamma_{j_*} \int_0^1 (\psi_{j_*,k} \star g)(t)\mathrm{d}W(t)\right)\right].$$

Hence,

$$\int \left(\frac{\mathrm{d}\mathbb{P}_{f_1}}{\mathrm{d}\mathbb{P}_{f_0}}\right)^2 \mathrm{d}\mathbb{P}_{f_0} = \prod_{k=0}^{2^{j_*}-1} 2^{-2} \exp\left(-n\gamma_{j_*}^2 \int_0^1 (\psi_{j_*,k} \star g)^2(t)\mathrm{d}t\right)$$
$$\times \int \left[\exp\left(n^{1/2}\gamma_{j_*} \int_0^1 (\psi_{j_*,k} \star g)(t)\mathrm{d}W(t)\right)\right.$$
$$\left. + \exp\left(-n^{1/2}\gamma_{j_*} \int_0^1 (\psi_{j_*,k} \star g)(t)\mathrm{d}W(t)\right)\right]^2 \mathrm{d}\mathbb{P}_{f_0}.$$

Since

$$\int \exp\left(2n^{1/2}\gamma_{j_*} \int_0^1 (\psi_{j_*,k} \star g)(t)\mathrm{d}W(t)\right) \mathrm{d}\mathbb{P}_{f_0}$$
$$= \int \exp\left(-2n^{1/2}\gamma_{j_*} \int_0^1 (\psi_{j_*,k} \star g)(t)\mathrm{d}W(t)\right) \mathrm{d}\mathbb{P}_{f_0}$$
$$= \exp\left(2n\gamma_{j_*}^2 \int_0^1 (\psi_{j_*,k} \star g)^2(t)\mathrm{d}t\right),$$

we have

$$\int \left(\frac{\mathrm{d}\mathbb{P}_{f_1}}{\mathrm{d}\mathbb{P}_{f_0}}\right)^2 \mathrm{d}\mathbb{P}_{f_0} = \prod_{k=0}^{2^{j_*}-1} 2^{-1} \left(\exp\left(n\gamma_{j_*}^2 \int_0^1 (\psi_{j_*,k} \star g)^2(t)\mathrm{d}t\right)\right.$$
$$\left. + \exp\left(-n\gamma_{j_*}^2 \int_0^1 (\psi_{j_*,k} \star g)^2(t)\mathrm{d}t\right)\right).$$

The inequality $2^{-1}(\mathrm{e}^x + \mathrm{e}^{-x}) \leq \mathrm{e}^{x^2/2}$, $x \in \mathbb{R}$, implies that

$$\int \left(\frac{\mathrm{d}\mathbb{P}_{f_1}}{\mathrm{d}\mathbb{P}_{f_0}}\right)^2 \mathrm{d}\mathbb{P}_{f_0} \leq \prod_{k=0}^{2^{j_*}-1} \exp\left(2^{-1} n^2 \gamma_{j_*}^4 \left(\int_0^1 (\psi_{j_*,k} \star g)^2(t)\mathrm{d}t\right)^2\right). \quad (14)$$

Using (A_g) and the Plancherel theorem, we obtain

$$\begin{aligned}\int_0^1 (\psi_{j_*,k} \star g)^2(t)\mathrm{d}t &= \sum_{l\in\mathbb{Z}} \left(F\left(\psi_{j_*,k}\star g\right)\right)^2(l) = \sum_{l\in\mathcal{C}_{j_*}} |F\left(\psi_{j_*,k}\right)(l)|^2|g_l|^2 \\ &\le C2^{-2\delta j_*}\sum_{l\in\mathcal{C}_{j_*}} |F\left(\psi_{j_*,k}\right)(l)|^2 = C2^{-2\delta j_*}\int_0^1 \psi^2_{j_*,k}(x)\mathrm{d}x \\ &= C2^{-2\delta j_*}.\end{aligned}$$

We deduce that

$$\int \left(\frac{\mathrm{d}\mathbb{P}_{f_1}}{\mathrm{d}\mathbb{P}_{f_0}}\right)^2 \mathrm{d}\mathbb{P}_{f_0} \le \prod_{k=0}^{2^{j_*}-1} \exp\left(Cn^2\gamma^4_{j_*}2^{-4\delta j_*}\right) = \exp\left(Cn^2\gamma^4_{j_*}2^{(1-4\delta)j_*}\right).$$

Thanks to the definitions of j_* and γ_{j_*}, we have

$$n^2\gamma^4_{j_*}2^{(1-4\delta)j_*} \le Cn^2 2^{-j_*(4s+2)}2^{(1-4\delta)j_*} \le Cn^2 2^{-j_*(4s+1+4\delta)} \le C.$$

It follows the existence of a constant $C > 0$ such that

$$\chi^2(\mathbb{P}_{f_1}, \mathbb{P}_{f_0}) = \int \left(\frac{\mathrm{d}\mathbb{P}_{f_1}}{\mathrm{d}\mathbb{P}_{f_0}}\right)^2 \mathrm{d}\mathbb{P}_{f_0} - 1 \le C < \infty.$$

Theorem 2.12 (iii) of Tsybakov (2004) yields

$$\inf_{\widetilde{Q}_n} \sup_{f\in B^s_{2,\infty}(M)} \mathbb{E}\left(\left(\widetilde{Q}_n - \|f\|_2^2\right)^2\right) \ge c\kappa_n^2 = cn^{-4s/(2s+2\delta+1/2)} = c\varphi^*_n. \tag{15}$$

Putting (11) and (15) together, we prove the existence of a constant $c > 0$ such that

$$\inf_{\widetilde{Q}_n} \sup_{f\in B^s_{2,\infty}(M)} \mathbb{E}\left(\left(\widetilde{Q}_n - \|f\|_2^2\right)^2\right) \ge c\varphi^*_n,$$

where $\varphi^*_n = \max\left(n^{-1}, n^{-4s/(2s+2\delta+1/2)}\right)$. Theorem 2 is proved.

Proof of Theorem 3 We need the following lemma.

Lemma 4 *If (A^*_g) is satisfied, then there exists a constant $C > 0$ such that, for any integer $j \ge \tau$,*

$$\eta_j^2 \le C\mathrm{e}^{(8\pi 3^{-1}a)2^{bj}}.$$

Proof of Lemma 4 Using the inequalities $\sup_{l\in\mathcal{D}_j}|F(\phi_{j,0})(l)|^2 \leq C2^{-j}$, $Card(\mathcal{D}_j)\leq C2^j$, (A_g^*) and the fact that $\sup_{l\in\mathcal{D}_j}|g_l|^{-2}\leq C\sup_{l\in\mathcal{D}_j}\mathrm{e}^{2a|l|^b}\leq C\mathrm{e}^{(8\pi 3^{-1}a)2^{bj}}$, we have

$$\eta_j^2=\sum_{l\in\mathcal{D}_j}|g_l|^{-2}|F(\phi_{j,0})(l)|^2\leq C2^{-j}\sum_{l\in\mathcal{D}_j}|g_l|^{-2}\leq C\sup_{l\in\mathcal{D}_j}|g_l|^{-2}\leq C\mathrm{e}^{(8\pi 3^{-1}a)2^{bj}}.$$

The decomposition $\|f\|_2^2=\sum_{k=0}^{2^{j_*}-1}\alpha_{j_*,k}^2+\sum_{j=j_*}^{\infty}\sum_{k=0}^{2^j-1}\beta_{j,k}^2$ and the elementary inequality: $(x+y)^2\leq 2(x^2+y^2)$, $(x,y)\in\mathbb{R}^2$, we have

$$\mathbb{E}\left(\left(Q_n^*-\|f\|_2^2\right)^2\right)\leq 2(\mathbf{E}+\mathbf{F}), \tag{16}$$

where

$$\mathbf{E}=\mathbb{E}\left(\left(\sum_{k=0}^{2^{j_*}-1}\left(\widehat{\alpha}_{j_*,k}^2-n^{-1}\eta_{j_*}^2-\alpha_{j_*,k}^2\right)\right)^2\right),\quad \mathbf{F}=\left(\sum_{j=j_*}^{\infty}\sum_{k=0}^{2^j-1}\beta_{j,k}^2\right)^2.$$

The upper bound for **F**. Since $f\in B_{2,\infty}^s(M)$, we have

$$\mathbf{F}=\left(\sum_{j=j_*}^{\infty}\sum_{k=0}^{2^j-1}\beta_{j,k}^2\right)^2\leq C\left(\sum_{j=j_*}^{\infty}2^{-2js}\right)^2\leq C2^{-4j_*s}\leq C(\log n)^{-4s/b}. \tag{17}$$

The upper bound for **E**. It follows from the decomposition $\widehat{\alpha}_{j_*,k}=\alpha_{j_*,k}+n^{-1/2}\sum_{l\in\mathbb{Z}}g_l^{-1}\mathrm{e}^{2i\pi lk/2^{j_*}}F(\phi_{j_*,0})(l)e_l$ and again: $(x+y)^2\leq 2(x^2+y^2)$, $(x,y)\in\mathbb{R}^2$ that

$$\mathbf{E}\leq 2\left(\mathbf{G}+\mathbf{H}\right),$$

where

$$\mathbf{G}=n^{-2}\mathbb{E}\left(\left(\sum_{k=0}^{2^{j_*}-1}\left(\left(\sum_{l\in\mathbb{Z}}g_l^{-1}\mathrm{e}^{2i\pi lk/2^{j_*}}F(\phi_{j_*,0})(l)e_l\right)^2-\eta_{j_*}^2\right)\right)^2\right)$$

and

$$\mathbf{H}=4n^{-1}\mathbb{E}\left(\left(\sum_{k=0}^{2^{j_*}-1}|\alpha_{j_*,k}|\left(\sum_{l\in\mathbb{Z}}g_l^{-1}\mathrm{e}^{2i\pi lk/2^{j_*}}F(\phi_{j_*,0})(l)e_l\right)\right)^2\right).$$

The upper bound for **G**. Using the Minkowski inequality, the fact that the random variables $(e_l)_{l\in\mathbb{Z}}$ are independent $\mathcal{N}(0, 1)$, (A_g^*) and Lemma 4, we obtain

$$\begin{aligned}
\mathbf{G} &\le n^{-2}\left(\sum_{k=0}^{2^{j_*}-1}\left(\mathbb{E}\left(\left(\left(\sum_{l\in\mathbb{Z}} g_l^{-1}\mathrm{e}^{2i\pi lk/2^{j_*}}F(\phi_{j_*,0})(l)e_l\right)^2-\eta_{j_*}^2\right)^2\right)\right)^{1/2}\right)^2\\
&= n^{-2}\left(\sum_{k=0}^{2^{j_*}-1}\left(\mathbb{V}\left(\left(\sum_{l\in\mathbb{Z}} g_l^{-1}\mathrm{e}^{2i\pi lk/2^{j_*}}F(\phi_{j_*,0})(l)e_l\right)^2\right)\right)^{1/2}\right)^2\\
&\le n^{-2}\left(\sum_{k=0}^{2^{j_*}-1}\left(\mathbb{E}\left(\left(\sum_{l\in\mathbb{Z}} g_l^{-1}\mathrm{e}^{2i\pi lk/2^{j_*}}F(\phi_{j_*,0})(l)e_l\right)^4\right)\right)^{1/2}\right)^2\\
&\le Cn^{-2}\left(\sum_{k=0}^{2^{j_*}-1}\eta_{j_*}^2\right)^2 = Cn^{-2}2^{2j_*}\eta_{j*}^4 \le Cn^{-2}2^{2j_*}\mathrm{e}^{(16\pi 3^{-1}a)2^{bj_*}}\\
&\le Cn^{-2}(\log n)^{2/b}n = Cn^{-1}(\log n)^{2/b} \le C(\log n)^{-4s/b}.
\end{aligned}$$

The upper bound for **H**. Using the Hölder inequality, the inequality $\sum_{k=0}^{2^{j_*}-1}\alpha_{j_*,k}^2 \le ||f||_2^2 \le M_*$, the fact that the random variables $(e_l)_{l\in\mathbb{Z}}$ are independent $\mathcal{N}(0, 1)$, (A_g^*) and Lemma 4, we obtain

$$\begin{aligned}
\mathbf{H} &= 4n^{-1}\mathbb{E}\left(\left(\sum_{k=0}^{2^{j_*}-1}|\alpha_{j_*,k}|\left(\sum_{l\in\mathbb{Z}} g_l^{-1}\mathrm{e}^{2i\pi lk/2^{j_*}}F(\phi_{j_*,0})(l)e_l\right)\right)^2\right)\\
&\le 4n^{-1}\sum_{k=0}^{2^{j_*}-1}\alpha_{j_*,k}^2\sum_{k=0}^{2^{j_*}-1}\mathbb{E}\left(\left(\sum_{l\in\mathbb{Z}} g_l^{-1}\mathrm{e}^{2i\pi lk/2^{j_*}}F(\phi_{j_*,0})(l)e_l\right)^2\right)\\
&\le Cn^{-1}\sum_{k=0}^{2^{j_*}-1}\eta_{j_*}^2 = Cn^{-1}2^{j_*}\eta_{j*}^2 \le Cn^{-1}2^{j_*}\mathrm{e}^{(8\pi 3^{-1}a)2^{bj_*}}\\
&\le Cn^{-1}(\log n)^{1/b}n^{1/2} = Cn^{-1/2}(\log n)^{1/b} \le C(\log n)^{-4s/b}.
\end{aligned}$$

It follows from the obtained upper bounds for **G** and **H** that

$$\mathbf{E} \le 2\,(\mathbf{G}+\mathbf{H}) \le C(\log n)^{-4s/b}. \tag{18}$$

Putting (16), (17) and (18) together, we obtain

$$\mathbb{E}\left(\left(Q_n^* - \|f\|_2^2\right)^2\right) \le 2(\mathbf{E}+\mathbf{F}) \le C(\log n)^{-4s/b}.$$

This ends the proof of Theorem 3.

Proof of Theorem 4 The proof is similar to the proof of Theorem 2 in the case $s \in (0, \delta + 4^{-1}]$. We consider the two functions

$$f_0(x) = 0, \quad f_1(x) = \gamma_{j_*} \sum_{k_*=0}^{2^{j_*}} w_{k_*} \psi_{j_*,k_*}(x), \quad x \in [0, 1], \tag{19}$$

where j_* is an integer to be chosen below, γ_{j_*} is a quantity to be chosen below, and $w_0, \ldots, w_{2^{j_*}-1}$ are i.i.d. Rademacher random variables. Clearly, f_0 belongs to $B^s_{2,\infty}(M)$. If $\gamma^2_{j_*} = M_* 2^{-j_*(2s+1)}$, then $f_1 \in B^s_{2,\infty}(M)$. Let j_* be an integer such that

$$\begin{aligned}&(8\pi 3^{-1} a)^{-1/b} \left(2 \log n - ((4s+1)/b) \log(\log n)\right)^{1/b} \\ &\quad \le 2^{j_*} < 2(8\pi 3^{-1} a)^{-1/b} \left(2 \log n - ((4s+1)/b) \log(\log n)\right)^{1/b}. \end{aligned} \tag{20}$$

With the previous value of γ_{j_*}, the orthonormality of the wavelet basis gives

$$\begin{aligned}\left| \|f_0\|_2^2 - \|f_1\|_2^2 \right| &= \|f_1\|_2^2 = \gamma^2_{j_*} \sum_{k=0}^{2^{j_*}-1} w_k^2 = \gamma^2_{j_*} 2^{j_*} \ge C 2^{-2j_* s} \\ &\ge C (\log n)^{-2s/b} = \kappa_n.\end{aligned}$$

In what follows, we consider the function f_1 described by (19) with the integer j_* defined by (20) and the quantity $\gamma_{j_*} = M_*^{1/2} 2^{-j_*(s+1/2)}$. In order to apply Theorem 2.12 (iii) of Tsybakov (2004), we aim to bound the chi-square divergence $\chi^2(\mathbb{P}_{f_1}, \mathbb{P}_{f_0})$ by a constant. It follows from (14) that

$$\begin{aligned}\chi^2(\mathbb{P}_{f_1}, \mathbb{P}_{f_0}) &= \int \left(\frac{\mathrm{d}\mathbb{P}_{f_1}}{\mathrm{d}\mathbb{P}_{f_0}} \right)^2 \mathrm{d}\mathbb{P}_{f_0} - 1 \le \int \left(\frac{\mathrm{d}\mathbb{P}_{f_1}}{\mathrm{d}\mathbb{P}_{f_0}} \right)^2 \mathrm{d}\mathbb{P}_{f_0} \\ &\le \prod_{k=0}^{2^{j_*}-1} \exp\left(2^{-1} n^2 \gamma^4_{j_*} \left(\int_0^1 (\psi_{j_*,k} \star g)^2(t) \mathrm{d}t \right)^2 \right).\end{aligned}$$

Using (A_g^*) and the Plancherel theorem, we obtain

$$\begin{aligned}\int_0^1 (\psi_{j_*,k} \star g)^2(t) \mathrm{d}t &= \sum_{l \in \mathbb{Z}} \left(F\left(\psi_{j_*,k} \star g\right) \right)^2 (l) = \sum_{l \in \mathcal{C}_{j_*}} \left| F\left(\psi_{j_*,k}\right)(l) \right|^2 |g_l|^2 \\ &\le C \mathrm{e}^{-(4\pi 3^{-1} a) 2^{b j_*}} \sum_{l \in \mathcal{C}_{j_*}} \left| F\left(\psi_{j_*,k}\right)(l) \right|^2 \\ &= C \mathrm{e}^{-(4\pi 3^{-1} a) 2^{b j_*}} \int_0^1 \psi^2_{j_*,k}(x) \mathrm{d}x = C \mathrm{e}^{-(4\pi 3^{-1} a) 2^{b j_*}}.\end{aligned}$$

We deduce that

$$\chi^2(\mathbb{P}_{f_1}, \mathbb{P}_{f_0}) \le \prod_{k=0}^{2^{j_*}-1} \exp\left(Cn^2\gamma_{j_*}^4 \mathrm{e}^{-(8\pi 3^{-1}a)2^{bj_*}}\right)$$
$$= \exp\left(Cn^2\gamma_{j_*}^4 2^{j_*} \mathrm{e}^{-(8\pi 3^{-1}a)2^{bj_*}}\right).$$

Thanks to the definitions of j_* and γ_{j_*}, we have

$$n^2\gamma_{j_*}^4 2^{j_*}\mathrm{e}^{-(8\pi 3^{-1}a)2^{bj_*}} \le Cn^2 2^{-j_*(4s+1)}\mathrm{e}^{-(8\pi 3^{-1}a)2^{bj_*}}$$
$$\le C\frac{(\log n)^{(4s+1)/b}}{(2\log n - ((4s+1)/b)\log(\log n))^{(4s+1)/b}} \le C.$$

It follows the existence of a constant $C > 0$ such that

$$\chi^2(\mathbb{P}_{f_1}, \mathbb{P}_{f_0}) = \int \left(\frac{\mathrm{d}\mathbb{P}_{f_1}}{\mathrm{d}\mathbb{P}_{f_0}}\right)^2 \mathrm{d}\mathbb{P}_{f_0} - 1 \le C < \infty.$$

Theorem 2.12 (iii) of Tsybakov (2004) yields

$$\inf_{\widetilde{Q}_n} \sup_{f\in B^s_{2,\infty}(M)} \mathbb{E}\left(\left(\widetilde{Q}_n - \|f\|_2^2\right)^2\right) \ge c\kappa_n^2 = c(\log n)^{-4s/b}.$$

Theorem 4 is proved.

Acknowledgments We thank the Associate Editor and the two referees for detailed and constructive comments which have helped to improve the presentation of the paper.

References

Bertero, M., Boccacci, P. (1998). *Introduction to inverse problems in imaging*. Bristol and Philadelphia: Institute of Physics.

Bickel, P. J., Ritov, Y. (1988). Estimating integrated squared density derivatives: Sharp best order of convergence estimates. *Sankhya Series A, 50*, 381–393.

Butucea, C. (2007). Goodness-of-fit testing and quadratic functional estimation from indirect observations. *The Annals of Statistics, 35*(5), 1907–1930.

Cai, T., Low, M. (2005). Non-quadratic estimators of a quadratic functional. *The Annals of Statistics, 33*, 2930–2956.

Cai, T., Low, M. (2006). Optimal adaptive estimation of a quadratic functional. *The Annals of Statistics, 34*, 2298–2325.

Cavalier, L. (2008). Nonparametric statistical inverse problems. *Inverse Problems, 24*, 034004.

Cavalier, L., Raimondo, M. (2007). Wavelet deconvolution with noisy eigenvalues. *IEEE Transactions on Signal Processing, 55*, 2414–2424.

Cavalier, L., Tsybakov, A. (2002). Sharp adaptation for inverse problems with random noise. *Probability Theory and Related Fields, 123*, 323–354.

Cavalier, L., Golubev, Y., Lepski, O., Tsybakov, A. (2004). Block thresholding and sharp adaptive estimation in severely ill-posed inverse problems. *Probability Theory and Related Fields, 48*(3), 426–446.

Chesneau, C. (2008). Wavelet estimation via block thresholding: A minimax study under the $\mathbb{L}^p$ risk. *Statistica Sinica, 18*(3), 1007–1024.
Cirelson, B., Ibragimov, I., Sudakov, V. (1976). *Norm of Gaussian sample functions*. Berlin: Springer.
Cohen, A., Daubechies, I., Jawerth, B., Vial, P. (1993). Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis, 24*(1), 54–81.
Donoho, D. L., Nussbaum, M. (1990). Minimax quadratic estimation of a quadratic functional. *Journal of Complexity, 6*, 290–323.
Efromovich, S. Y., Low, M. (1996). On optimal adaptive estimation of a quadratic functional. *The Annals of Statistics, 24*, 1106–1125.
Fan, J., Koo, J. Y. (2002). Wavelet deconvolution. *IEEE Tractions on Information Theory, 48*, 734–747.
Gayraud, G., Tribouley, K. (1999). Wavelet methods to estimate an integrated quadratic functional: Adaptivity and asymptotic law. *Statistics and Probability Letters, 44*, 109–122.
Johnstone, I. M. (2001a). Chi-square oracle inequalities. In *State of the art in probability and statistics* (Leiden, 1999), *Lecture notes—monograph series* (Vol. 36, pp. 399–418). Beachwood: Institute of Mathematical Statistics.
Johnstone, I. M. (2001b). Thresholding for weighted chi-square. *Statistica Sinica, 11*, 691–704.
Johnstone, I., Kerkyacharian, G., Picard, D., Raimondo, M. (2004). Wavelet deconvolution in a periodic setting. *Journal of the Royal Statistical Society, Series B, 66*(3), 547–573.
Kerkyacharian, G., Picard, D. (1996). Estimating nonquadratic functionals of a density using haar wavelets. *The Annals of Statistics, 24*(2), 485–507.
Laurent, B. (2005). Adaptive estimation of a quadratic functional of a density by model selection. *ESAIM PS, 9*, 1–18.
Meyer, Y. (1992). *Wavelets and operators*. Cambridge: Cambridge University Press.
Neelamani, R., Choi, H., Baraniuk, R. (2004). Forward: Fourier-wavelet regularized deconvolution for ill-conditioned systems. *IEEE Transactions on Signal Processing, 52*, 418–433.
Pensky, M., Sapatinas, T. (2009). Functional deconvolution in a periodic setting. *The Annals of Statistics* (to appear).
Pensky, M., Vidakovic, B. (1999) Adaptive wavelet estimator for nonparametric density deconvolution. *The Annals of Statistics, 27*, 2033–2053.
Rivoirard, V., Tribouley, K. (2008). The maxiset point of view for estimating integrated quadratic functionals. *Statistica Sinica, 18*, 255–279.
Tsybakov, A. (2004). *Introduction à l'estimation nonparametrique*. Berlin: Springer.
Walter, G. (1994). *Wavelets and other orthogonal systems in applications*. Boca Raton: CRC Press.
Willer, T. (2005). Deconvolution in white noise with a random blurring effect. Preprint LPMA.
Zayed, A., Walter, G. (1996). Characterization of analytic functions in term of their wavelet coefficients. *Complex Variables, 29*, 265–276.