



HAL
open science

Wavelet-based density estimation in a heteroscedastic convolution model

Christophe Chesneau, Jalal M. Fadili

► **To cite this version:**

Christophe Chesneau, Jalal M. Fadili. Wavelet-based density estimation in a heteroscedastic convolution model. *Communications in Statistics - Theory and Methods*, 2011, 42 (17), pp.3085-3099. 10.1080/03610926.2011.615440 . hal-00565591v3

HAL Id: hal-00565591

<https://hal.science/hal-00565591v3>

Submitted on 30 Dec 2012

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.

Wavelet-based density estimation in a heteroscedastic convolution model

Christophe Chesneau · Jalal Fadili

Received:

Abstract We consider a heteroscedastic convolution density model under the “ordinary smooth assumption”. We introduce a new adaptive wavelet estimator based on term-by-term hard thresholding rule. Its asymptotic properties are explored via the minimax approach under the mean integrated squared error over Besov balls. We prove that our estimator attains near optimal rates of convergence (lower bounds are determined). Simulation results are reported to support our theoretical findings.

Keywords Density deconvolution · Heteroscedasticity · Rates of convergence · Wavelet bases · Hard thresholding · Lower bounds.

2000 Mathematics Subject Classification 62G07, 62G20.

1 Introduction

Problem statement and prior art The heteroscedastic deconvolution problem can be formulated as follows. Suppose we have n random variables Y_1, \dots, Y_n where, for any $v \in \{1, \dots, n\}$,

$$Y_v = X_v + \epsilon_v, \quad (1)$$

X_1, \dots, X_n are *i.i.d.* random variables and $\epsilon_1, \dots, \epsilon_n$ are independent random variables, also independent of X_1, \dots, X_n . The density of X_1 is unknown and denoted f . For any $v \in \{1, \dots, n\}$, the density of ϵ_v is known, denoted g_v and satisfies the “ordinary smooth case” (to be defined precisely in Section 2). The

Christophe Chesneau
Laboratoire de Mathématiques Nicolas Oresme, CNRS-Univ. de Caen, Campus II, Science 3, 14032 Caen, France. E-mail: chesneau@math.unicaen.fr

Jalal Fadili
GREYC CNRS-ENSICAEN-Univ. de Caen, ENSICAEN 14050 Caen Cedex, France. E-mail: Jalal.Fadili@greyc.ensicaen.fr

goal is to estimate f when only Y_1, \dots, Y_n are observed. Such a deconvolution problem arises in various areas including signal processing, biology, chemistry and economy.

In the homoscedastic case i.e. $g_1 = \dots = g_n$, (1) becomes the standard convolution density model. Various estimation techniques can be found in e.g. Carroll and Hall (1988), Devroye (1989), Fan (1991), Pensky and Vidakovic (1999), Zhang and Karunamuni (2000), Fan and Koo (2002), Butucea and Matias (2005), Hall and Qiu (2005), Comte et al. (2006), Delaigle and Gijbels (2006), Lacour (2006), Hall and Meister (2007) and Lounici and Nickl (2011). In the heteroscedastic case, (1) has been recently investigated in Delaigle and Meister (2008), Staudenmayer et al. (2008), Meister et al. (2010) and Wang et al. (2010) via kernel and Spline methods. Meister (2009) gave an elaboration on nonparametric deconvolution, including wavelet methods and Wang and Wang (2011) discussed the problem from a practical point of view.

Contributions and relation to prior work In this study, we focus our attention on nonparametric wavelet-based methods. They are attractive for nonparametric function estimation because of their ability to adapt to unknown smoothness of the function, hence allowing to estimate accurately its local features such as discontinuities and aberrations. From a theoretical point of view, they can achieve near optimal convergence rates over a wide range of function classes (typically, Besov balls) and enjoy better mean integrated squared error (MISE) properties than kernel methods. See e.g. Antoniadis (1997) and Härdle et al. (1998). The estimation of f from (1) in the homoscedastic case via wavelet-based techniques can be found in Pensky and Vidakovic (1999), Fan and Koo (2002) and Lounici and Nickl (2011).

The construction of our adaptive estimator uses a similar Fourier-wavelet methodology to the one of Pensky and Vidakovic (1999) or Fan and Koo (2002). The idea is to select the large wavelet coefficients estimators by using a term-by-term thresholding rule (hard thresholding is considered here). Our estimator enjoys the originality to be able to handle the heteroscedasticity of (1) and to operate a new “observations thresholding” rule. Its performances are evaluated via the minimax approach under the MISE over Besov balls. We determine upper and lower bounds of the minimax risk of our estimator and prove that it is nearly optimal. We also report some simulation results to illustrate the potential applicability of the estimator and to support our theoretical findings.

Paper organization The paper is organized as follows. Model assumptions on (1) are introduced in Section 2. Section 3 briefly describes compactly supported wavelet bases, as well as characterization of Besov balls through wavelet sequence norm equivalence. Our hard thresholding estimator is described in Section 4. Our main results are stated and discussed in Section 5. A simulation study is reported and commented in Section 6, before drawing conclusions in Section 7. The technical proofs are deferred to Section 8.

2 Model assumptions

Without loss of generality, we assume that the support of f is included in $[-\Omega, \Omega]$ and that there exists a constant $C_* > 0$ such that

$$\sup_{x \in [-\Omega, \Omega]} f(x) \leq C_* < \infty. \quad (2)$$

We define the Fourier transform of an integrable function h by

$$\mathcal{F}(h)(x) = \int_{-\infty}^{\infty} h(y)e^{-ixy}dy, \quad x \in \mathbb{R}.$$

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

We consider the ‘‘heteroscedastic’’ ordinary smooth assumption on g_1, \dots, g_n : there exist three constants, $C_g > 0$, $c_g > 0$ and $\delta > 1$, and n positive real numbers $\sigma_1, \dots, \sigma_n$ such that, for any $v \in \{1, \dots, n\}$ and any $x \in \mathbb{R}$,

$$|\mathcal{F}(g_v)(x)| \geq \frac{c_g}{(1 + \sigma_v^2 x^2)^{\delta/2}} \quad (3)$$

and, for any $\ell \in \{0, 1, 2\}$, the ℓ -th derivative of the Fourier transform of g_v satisfies

$$\begin{cases} \lim_{x \rightarrow 0} \sigma_v^2 |x|^{\delta+2\ell} |\mathcal{F}(g_v)(x)|^{(\ell)} \leq C_g, \\ \lim_{x \rightarrow \infty} \sigma_v^2 |x|^{\delta+\ell} |\mathcal{F}(g_v)(x)|^{(\ell)} \leq C_g. \end{cases} \quad (4)$$

In the homoscedastic case, (3) becomes the standard ordinary smooth assumption. See e.g. Pensky and Vidakovic (1999), Fan and Koo (2002) and Lounici and Nickl (2011).

Example: for any $v \in \{1, \dots, n\}$, let us set

$$\epsilon_v = \sum_{u=1}^p \epsilon_{u,v},$$

where $p \in \mathbb{N}^*$, $(\epsilon_{u,v})_{u \in \{1, \dots, p\}}$ are *i.i.d.* random variables having the Laplace density $\mathcal{Laplace}(0, \sigma_v)$: $h_v(x) = (1/2\sigma_v)e^{-|x|/\sigma_v}$, $x \in \mathbb{R}$. Then

$$\mathcal{F}(g_v)(x) = (\mathcal{F}(h_v)(x))^p = \frac{1}{(1 + \sigma_v^2 x^2)^p}.$$

Thus (3) is satisfied with $\delta = 2p$. Moreover, if there exists a constant $c_* > 0$ such that $\inf_{v \in \{1, \dots, n\}} \sigma_v^2 \geq c_*$, then (4) is satisfied as well.

In the sequel, we set

$$w_n = \sum_{v=1}^n \frac{1}{(1 + \sigma_v^2)^\delta}$$

and, for technical reasons, we suppose that $w_n \geq e$.

3 Wavelets and Besov balls

Let $N \in \mathbb{N}^*$, and ϕ and ψ be the Daubechies scaling and wavelet functions db N (in particular, ϕ and ψ are compactly supported). We choose N such that $\phi \in \mathcal{C}^\alpha$ and $\psi \in \mathcal{C}^\alpha$ for $\alpha > 6 + \delta$ where δ refers to (3). Define the scaled and translated versions of the scaling and wavelet functions as

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

Then there exists an integer τ and a set Λ_j of consecutive integers with a length proportional to 2^j such that, for any integer $\ell \geq \tau$, the collection $\mathcal{B} = \{\phi_{\ell,k}(\cdot), k \in \Lambda_\ell; \psi_{j,k}(\cdot); j \in \mathbb{N} - \{0, \dots, \ell - 1\}, k \in \Lambda_j\}$ is an orthonormal basis of $\mathbb{L}^2([-\Omega, \Omega]) = \{h : [-\Omega, \Omega] \rightarrow \mathbb{R}; \int_{-\Omega}^{\Omega} h^2(x)dx < \infty\}$. We refer to Cohen et al. (1993) and Mallat (2009) for a comprehensive account on wavelets.

For any integer $\ell \geq \tau$, any $h \in \mathbb{L}^2([-\Omega, \Omega])$ can be expanded on \mathcal{B} as

$$h(x) = \sum_{k \in \Lambda_\ell} \alpha_{\ell,k} \phi_{\ell,k}(x) + \sum_{j=\ell}^{\infty} \sum_{k \in \Lambda_j} \beta_{j,k} \psi_{j,k}(x),$$

where $\alpha_{j,k}$ and $\beta_{j,k}$ are the scaling and wavelet coefficients of h defined through the inner product on $\mathbb{L}^2([-\Omega, \Omega])$

$$\alpha_{j,k} = \int_{-\Omega}^{\Omega} h(x) \phi_{j,k}(x) dx, \quad \beta_{j,k} = \int_{-\Omega}^{\Omega} h(x) \psi_{j,k}(x) dx. \quad (5)$$

Let $M > 0$, $s > 0$, $p \geq 1$ and $r \geq 1$. A function h belongs to $B_{p,r}^s(M)$ if and only if there exists a constant $M^* > 0$ (depending on M) such that the associated wavelet coefficients (5) satisfy

$$2^{\tau(1/2-1/p)} \left(\sum_{k \in \Lambda_\tau} |\alpha_{\tau,k}|^p \right)^{1/p} + \left(\sum_{j=\tau}^{\infty} \left(2^{j(s+1/2-1/p)} \left(\sum_{k \in \Lambda_j} |\beta_{j,k}|^p \right)^{1/p} \right)^r \right)^{1/r} \leq M^*.$$

In this expression, s is a smoothness parameter and p and r are norm parameters. Besov balls contain the Hölder and Sobolev balls. See e.g. Meyer (1992) and Mallat (2009).

4 Hard thresholding estimator

The first step to estimate f consists in expanding f on \mathcal{B} and estimating its unknown wavelet coefficients. For any integer $j \geq \tau$ and any $k \in \Lambda_j$,

- we estimate $\alpha_{j,k} = \int_{-\Omega}^{\Omega} f(x)\phi_{j,k}(x)dx$ by

$$\hat{\alpha}_{j,k} = \frac{1}{2\pi w_n} \sum_{v=1}^n \frac{1}{(1 + \sigma_v^2)^\delta} \int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\phi_{j,k})(x)}}{\mathcal{F}(g_v)(x)} e^{-ixY_v} dx,$$

- we estimate $\beta_{j,k} = \int_{-\Omega}^{\Omega} f(x)\psi_{j,k}(x)dx$ by

$$\hat{\beta}_{j,k} = \frac{1}{w_n} \sum_{v=1}^n G_{v,j,k} \mathbb{I}_{\{|G_{v,j,k}| \leq \theta 2^{\delta j} \sqrt{\frac{w_n}{\ln w_n}}\}}, \quad (6)$$

where

$$G_{v,j,k} = \frac{1}{2\pi} \frac{1}{(1 + \sigma_v^2)^\delta} \int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\psi_{j,k})(x)}}{\mathcal{F}(g_v)(x)} e^{-ixY_v} dx,$$

for any random event \mathcal{A} , $\mathbb{I}_{\mathcal{A}}$ is the indicator function on \mathcal{A} and

$$\theta = \sqrt{(C_*/(2\pi c_g^2)) \int_{-\infty}^{\infty} (1 + x^2)^\delta |\mathcal{F}(\psi)(x)|^2 dx}.$$

We define the hard thresholding estimator \hat{f} by

$$\hat{f}(x) = \sum_{k \in \Lambda_\tau} \hat{\alpha}_{\tau,k} \phi_{\tau,k}(x) + \sum_{j=\tau}^{j_1} \sum_{k \in \Lambda_j} \hat{\beta}_{j,k} \mathbb{I}_{\{|\hat{\beta}_{j,k}| \geq \kappa \theta 2^{\delta j} \sqrt{\frac{\ln w_n}{w_n}}\}} \psi_{j,k}(x), \quad (7)$$

where $\kappa \geq 8/3 + 2 + 2\sqrt{16/9 + 4}$ and j_1 is the integer satisfying

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

The feature of the hard thresholding estimator is to only estimate the “large” unknown wavelet coefficients of f which are those encoding the main features of f . See e.g. Mallat (2009).

Our estimator (7) can be viewed as an extension of the one in Fan and Koo (2002) to the heteroscedastic case. The presence of the “observations thresholding” in (6) allows us to treat this case without any restrictive assumptions on $\sigma_1, \dots, \sigma_n$. Mention that it can be removed under some assumptions on j_1 .

5 Minimaxity results

Theorem 1 (Upper bound) *Consider (1) under (2) and (3). Let \hat{f} be as given by (7) and $r \geq 1$, $\{p \geq 2$ and $s > 0\}$ or $\{p \in [1, 2)$ and $s > (2\delta + 1)/p\}$. Then, for n a large enough, there exists a constant $C > 0$ such that*

$$\sup_{f \in B_{p,r}^s(M)} \mathbb{E} \left(\int_{-\Omega}^{\Omega} (\hat{f}(x) - f(x))^2 dx \right) \leq C \left(\frac{\ln w_n}{w_n} \right)^{2s/(2s+2\delta+1)}.$$

The proof of Theorem 1 relies on (Chesneau 2011, Theorem 2) and several probability results related to $\widehat{\alpha}_{j,k}$ and $\widehat{\beta}_{j,k}$.

To establish the minimax optimality of \widehat{f} , the minimax lower bounds must be explored. This is done in Theorem 2 below.

Theorem 2 (Lower bound) *Consider (1) under (2), (3) and (4). Assume that there exists a constant $c_* > 0$ such that $\inf_{v \in \{1, \dots, n\}} \sigma_v^2 \geq c_*$. Then there exists a constant $c > 0$ such that, for any $s > 0$, $p \geq 1$, $r \geq 1$ and n large enough,*

$$\inf_{\widetilde{f}} \sup_{f \in B_{p,r}^s(M)} \mathbb{E} \left(\int_{-\Omega}^{\Omega} (\widetilde{f}(x) - f(x))^2 dx \right) \geq c(w_n^*)^{-2s/(2s+2\delta+1)},$$

where $w_n^* = \sum_{v=1}^n 1/\sigma_v^{2\delta}$ and the infimum is taken over all possible estimators \widetilde{f} of f .

The proof of Theorem 2 builds upon (Tsybakov 2004, Theorem 2.5) and several auxiliary results.

5.1 Discussion

Naturally, in the homoscedastic case, we have $w_n = Cn$ and from Theorem 1, we obtain the same rate of convergence as the one attained by the hard thresholding estimator in (Fan and Koo 2002, Theorem 2) i.e. $(\ln n/n)^{2s/(2s+2\delta+1)}$.

As far as the lower bound is concerned, we first point out that $w_n = \sum_{v=1}^n 1/(1+\sigma_v^2)^\delta \neq \sum_{v=1}^n 1/\sigma_v^{2\delta} = w_n^*$. However, if there exists a constant $c_* > 0$ such that $\inf_{v \in \{1, \dots, n\}} \sigma_v^2 \geq c_*$, then we have $(\ln w_n/w_n)^{2s/(2s+2\delta+1)} \leq C(\ln w_n/w_n^*)^{2s/(2s+2\delta+1)}$.

Therefore, due to Theorems 1 and 2, \widehat{f} is optimal in the minimax sense up to the logarithmic term $(\ln w_n)^{2s/(2s+2\delta+1)}$.

We would like to stress the fact that some choices of $\sigma_1, \dots, \sigma_n$ can severely deteriorate the performance of \widehat{f} . As a simple example, take $\sigma_v^2 = v$ for any $v \in \{1, \dots, n\}$, and $\delta = 1$: for n large enough, we have $w_n = C \ln n$ and $w_n^* = C \ln n$, and the optimal rate of convergence becomes $(\ln n)^{-2s/(2s+3)}$.

6 Simulation results

In this simulation, $n = 10^4$ samples $(Y_v)_{v \in \{1, \dots, n\}}$ were generated according to model (1), where for $v \in \{1, \dots, n\}$ $\epsilon_v \sim \mathcal{Laplace}(0, \sigma_v)$ and σ_v were independently randomly generated in $(0, +\infty)$. X_v were i.i.d. with a common density supported in $[-\Omega, \Omega]$. More precisely, we simulated $X_v = \sum_{i=1}^m U_{v,i}$, where $U_{v,i} \sim \mathcal{U}([- \Omega/m, \Omega/m])$, i.e. $m = 1$ corresponds to the uniform distribution, $m = 2$ the triangular distribution, etc.. Thus varying m allows to show the ability of our wavelet-based estimator in adaptively recovering a large class of densities spanning a wide range of irregularities, going from discontinuous to

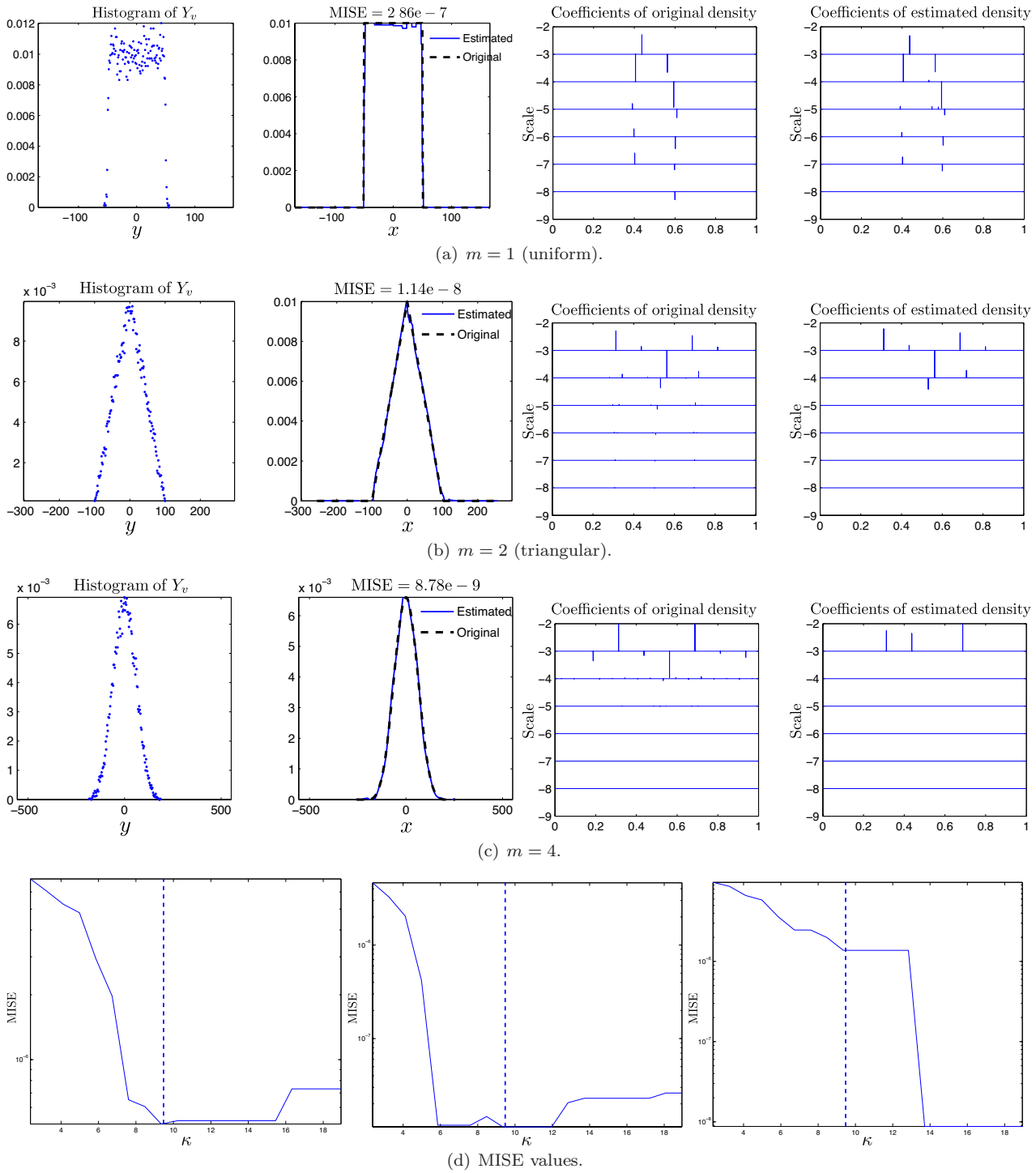


Fig. 1 Original (dashed) and estimated densities (solid) using our wavelet hard thresholding estimator from $n = 10^4$ heteroscedastic noisy samples Y_v according to (1) with $\epsilon_v \sim \mathcal{Laplace}(0, \sigma_v)$, and $X_v = \sum_{i=1}^m U_{v,i}$ are i.i.d. with $U_{v,i} \sim \mathcal{U}([- \Omega/m, \Omega/m])$, $\Omega = 50m$. (a): $m = 1$. (b): $m = 2$. (c): $m = 4$. (d): MISE values as a function of κ for $m = 1$ (left), $m = 2$ (middle) and $m = 4$ (right).

smooth densities.

Following the philosophy of reproducible research, a toolbox is made available freely for download at the address <http://www.greyc.ensicaen.fr/~jfadili/software.html>. This toolbox is a collection of Matlab functions and scripts that implement our estimator and allow to reproduce the figures reported in this paper. It requires at least WaveLab 802 (2001) to run properly.

The results are depicted in Fig. 1(a)-(c) for respectively $m \in \{1, 2, 4\}$. It can be clearly seen that our adaptive hard thresholding estimator is very effective to estimate each of the three densities. The recovered wavelet coefficients are also shown where most of the irregularities are captured in the estimated coefficients. In the figure, we also display the empirical MISE as a function of the threshold κ , where the vertical dashed line represents the lower-bound $8/3 + 2 + 2\sqrt{16/9 + 4} \approx 9.474\dots$ advocated by our estimator (7). One can see that the minimum of the MISE occurs beyond this lower-bound hence supporting the choice dictated by our theoretical procedure, although derived in an asymptotic setting.

7 Conclusion and perspectives

In this paper, considering (1), we have constructed an adaptive wavelet estimator and proved that it is nearly optimal in the minimax sense. Among possible perspectives of the present work, an interesting one would be to improve the estimation of f via another wavelet thresholding estimator such as the block thresholding one (see e.g. Pensky and Sapatinas (2010)). This needs further investigations that we leave for a future work.

8 Proofs

Proof of Theorem 1. We will apply the following general result. It is derived to (Chesneau 2011, Theorem 2).

Theorem 3 (Chesneau (2011)) *Let $\Omega > 0$. We want to estimate an unknown function f with support in $[-\Omega, \Omega]$ from n independent random variables (or vectors) U_1, \dots, U_n . We consider the wavelet basis \mathcal{B} and the notations of Section 3.*

Suppose that there exist n functions h_1, \dots, h_n and a sequence of real numbers $(\mu_v)_{v \in \mathbb{N}^}$ satisfying $\lim_{v \rightarrow \infty} \mu_v = \infty$, such that, for any $\gamma \in \{\phi, \psi\}$,*

(A1) any integer $j \geq \tau$ and any $k \in \Lambda_j$,

$$\mathbb{E} \left(\frac{1}{\mu_n} \sum_{v=1}^n h_v(\gamma_{j,k}, U_v) \right) = \int_{-\Omega}^{\Omega} f(x) \gamma_{j,k}(x) dx.$$

(A2) there exist two constants, $\theta_\gamma > 0$ and $\delta > 0$, such that, for any integer $j \geq \tau$ and any $k \in \Lambda_j$,

$$\frac{1}{\mu_n^2} \sum_{v=1}^n \mathbb{E} \left((h_v(\gamma_{j,k}, U_v))^2 \right) \leq \theta_\gamma^2 2^{2\delta j} \frac{1}{\mu_n}.$$

We define the hard thresholding estimator \hat{f} by

$$\hat{f}(x) = \sum_{k \in \Lambda_\tau} \hat{\alpha}_{\tau,k} \phi_{\tau,k}(x) + \sum_{j=\tau}^{j_1} \sum_{k \in \Lambda_j} \hat{\beta}_{j,k} \mathbb{I}_{\{|\hat{\beta}_{j,k}| \geq \kappa \theta_\psi 2^{\delta j} \sqrt{\frac{\ln \mu_n}{\mu_n}}\}} \psi_{j,k}(x),$$

where

$$\hat{\alpha}_{j,k} = \frac{1}{\mu_n} \sum_{v=1}^n h_v(\phi_{j,k}, U_v),$$

$$\hat{\beta}_{j,k} = \frac{1}{\mu_n} \sum_{v=1}^n h_v(\psi_{j,k}, U_v) \mathbb{I}_{\{|h_v(\psi_{j,k}, U_v)| \leq \theta_\psi 2^{\delta j} \sqrt{\frac{\ln \mu_n}{\mu_n}}\}},$$

$\kappa = 8/3 + 2 + 2\sqrt{16/9 + 4}$ and j_1 is the integer satisfying $(1/2)\mu_n^{1/(2\delta+1)} < 2^{j_1} \leq \mu_n^{1/(2\delta+1)}$.

Let $r \geq 1$, $\{p \geq 2$ and $s > 0\}$ or $\{p \in [1, 2)$ and $s > (2\delta + 1)/p\}$. Then there exists a constant $C > 0$ such that

$$\sup_{f \in B_{p,r}^s(M)} \mathbb{E} \left(\int_{-\Omega}^{\Omega} (\hat{f}(x) - f(x))^2 dx \right) \leq C \left(\frac{\ln \mu_n}{\mu_n} \right)^{2s/(2s+2\delta+1)}.$$

Let us now investigate the assumptions (A1) and (A2) of Theorem 3 with, for any $v \in \{1, \dots, n\}$, $U_v = Y_v$,

$$h_v(\gamma_{j,k}, y) = \frac{1}{(1 + \sigma_v^2)^\delta 2\pi} \int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\gamma_{j,k})(x)}}{\mathcal{F}(g_v)(x)} e^{-ixy} dx$$

and $\mu_n = w_n$.

On (A1). Since X_v and ϵ_v are independent, we have

$$\mathbb{E} (e^{-ixY_v}) = \mathbb{E} (e^{-ixX_v}) \mathbb{E} (e^{-ix\epsilon_v}) = \mathcal{F}(f)(x) \mathcal{F}(g_v)(x).$$

This combined with the Fubini theorem and the Parseval-Plancherel theorem yield, for any integer $j \geq \tau$ and any $k \in A_j$,

$$\begin{aligned}
& \mathbb{E} \left(\frac{1}{\mu_n} \sum_{v=1}^n h_v(\gamma_{j,k}, Y_v) \right) \\
&= \frac{1}{w_n} \sum_{v=1}^n \frac{1}{(1 + \sigma_v^2)^\delta} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\gamma_{j,k})(x)}}{\mathcal{F}(g_v)(x)} \mathbb{E}(e^{-ixY_v}) dx \\
&= \frac{1}{w_n} \sum_{v=1}^n \frac{1}{(1 + \sigma_v^2)^\delta} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\gamma_{j,k})(x)}}{\mathcal{F}(g_v)(x)} \mathcal{F}(f)(x) \mathcal{F}(g_v)(x) dx \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \overline{\mathcal{F}(\gamma_{j,k})(x)} \mathcal{F}(f)(x) dx \left(\frac{1}{w_n} \sum_{v=1}^n \frac{1}{(1 + \sigma_v^2)^\delta} \right) \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} \overline{\mathcal{F}(\gamma_{j,k})(x)} \mathcal{F}(f)(x) dx = \int_{-\Omega}^{\Omega} f(x) \gamma_{j,k}(x) dx. \tag{8}
\end{aligned}$$

On (A2). We have

$$\begin{aligned}
& \frac{1}{\mu_n^2} \sum_{v=1}^n \mathbb{E} \left((h_v(\gamma_{j,k}, Y_v))^2 \right) \\
&= \frac{1}{w_n^2} \sum_{v=1}^n \frac{1}{(1 + \sigma_v^2)^{2\delta} (2\pi)^2} \mathbb{E} \left(\left| \int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\gamma_{j,k})(x)}}{\mathcal{F}(g_v)(x)} e^{-ixY_v} dx \right|^2 \right). \tag{9}
\end{aligned}$$

Since X_v and ε_v are independent, the density of Y_v is $q_v(x) = (f \star g_v)(x) = \int_{-\infty}^{\infty} f(t)g_v(x-t)dt$, $x \in \mathbb{R}$. Therefore

$$\begin{aligned}
\mathbb{E} \left(\left| \int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\gamma_{j,k})(x)}}{\mathcal{F}(g_v)(x)} e^{-ixY_v} dx \right|^2 \right) &= \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\gamma_{j,k})(x)}}{\mathcal{F}(g_v)(x)} e^{-ixy} dx \right|^2 q_v(y) dy \\
&= \int_{-\infty}^{\infty} \left| \mathcal{F} \left(\frac{\overline{\mathcal{F}(\gamma_{j,k})(\cdot)}}{\mathcal{F}(g_v)(\cdot)} \right) (y) \right|^2 q_v(y) dy. \tag{10}
\end{aligned}$$

Since, by (2), $\sup_{x \in \mathbb{R}} f(x) \leq C_*$ and g_v is a density, we have

$$\sup_{v \in \{1, \dots, n\}} \sup_{x \in \mathbb{R}} q_v(x) \leq C_* \sup_{v \in \{1, \dots, n\}} \int_{-\infty}^{\infty} g_v(t) dt = C_*.$$

The Parseval-Plancherel theorem and (3) imply that

$$\begin{aligned}
& \int_{-\infty}^{\infty} \left| \mathcal{F} \left(\frac{\overline{\mathcal{F}(\gamma_{j,k})(\cdot)}}{\mathcal{F}(g_v)(\cdot)} \right) (y) \right|^2 q_v(y) dy \leq C_* \int_{-\infty}^{\infty} \left| \mathcal{F} \left(\frac{\overline{\mathcal{F}(\gamma_{j,k})(\cdot)}}{\mathcal{F}(g_v)(\cdot)} \right) (y) \right|^2 dy \\
&= 2\pi C_* \int_{-\infty}^{\infty} \left| \frac{\overline{\mathcal{F}(\gamma_{j,k})(x)}}{\mathcal{F}(g_v)(x)} \right|^2 dx \leq 2\pi \frac{C_*}{c_g^2} \int_{-\infty}^{\infty} (1 + x^2 \sigma_v^2)^\delta |\mathcal{F}(\gamma_{j,k})(x)|^2 dx. \tag{11}
\end{aligned}$$

By a change of variables, we obtain $|\mathcal{F}(\gamma_{j,k})(x)| = 2^{-j/2} |\mathcal{F}(\gamma)(x/2^j)|$. Using again a change of variables and the inequality $1 + \sigma_v^2 x^2 \leq (1 + \sigma_v^2)(1 + x^2)$, we have

$$\begin{aligned} \int_{-\infty}^{\infty} (1 + x^2 \sigma_v^2)^\delta |\mathcal{F}(\gamma_{j,k})(x)|^2 dx &= 2^{-j} \int_{-\infty}^{\infty} (1 + x^2 \sigma_v^2)^\delta |\mathcal{F}(\gamma)(x/2^j)|^2 dx \\ &= \int_{-\infty}^{\infty} (1 + 2^{2j} x^2 \sigma_v^2)^\delta |\mathcal{F}(\gamma)(x)|^2 dx \leq 2^{2\delta j} \int_{-\infty}^{\infty} (1 + \sigma_v^2 x^2)^\delta |\mathcal{F}(\gamma)(x)|^2 dx \\ &\leq 2^{2\delta j} (1 + \sigma_v^2)^\delta \int_{-\infty}^{\infty} (1 + x^2)^\delta |\mathcal{F}(\gamma)(x)|^2 dx. \end{aligned} \quad (12)$$

It follows from (10), (11) and (12) that

$$\begin{aligned} &\mathbb{E} \left(\left| \int_{-\infty}^{\infty} \frac{\overline{\mathcal{F}(\gamma_{j,k})(x)}}{\mathcal{F}(g_v)(x)} e^{-ixY_v} dx \right|^2 \right) \\ &\leq \left(2\pi(C_*/c_g^2) \int_{-\infty}^{\infty} (1 + x^2)^\delta |\mathcal{F}(\gamma)(x)|^2 dx \right) 2^{2\delta j} (1 + \sigma_v^2)^\delta. \end{aligned} \quad (13)$$

Putting (9) and (13) together, we obtain

$$\begin{aligned} \frac{1}{\mu_n^2} \sum_{v=1}^n \mathbb{E} \left((h_v(\gamma_{j,k}, Y_v))^2 \right) &\leq \theta_\gamma^2 2^{2\delta j} \frac{1}{w_n^2} \sum_{v=1}^n \frac{1}{(1 + \sigma_v^2)^{2\delta}} (1 + \sigma_v^2)^\delta \\ &= \theta_\gamma^2 2^{2\delta j} \frac{1}{w_n^2} w_n = \theta_\gamma^2 2^{2\delta j} \frac{1}{w_n}, \end{aligned} \quad (14)$$

where $\theta_\gamma = \sqrt{(C_*/(2\pi c_g^2)) \int_{-\infty}^{\infty} (1 + x^2)^\delta |\mathcal{F}(\gamma)(x)|^2 dx}$.

It follows from Theorem 3, (8) and (14) that the hard thresholding estimator (7) satisfies, for any $r \geq 1$, $\{p \geq 2$ and $s > 0\}$ or $\{p \in [1, 2)$ and $s > (2\delta + 1)/p\}$,

$$\sup_{f \in B_{p,r}^s(M)} \mathbb{E} \left(\int_{-\Omega}^{\Omega} (\widehat{f}(x) - f(x))^2 dx \right) \leq C \left(\frac{\ln w_n}{w_n} \right)^{2s/(2s+2\delta+1)}.$$

The proof of Theorem 1 is complete. □

Proof of Theorem 2. We will apply the following general result. It is (Tsybakov 2004, Theorem 2.5).

Theorem 4 (Tsybakov (2004)) *Let (\mathcal{F}, d) be a metric space, $(\mathcal{X}, \mathcal{A}, (\mathbb{P}_\theta)_{\theta \in \mathcal{F}})$ be a probability space, $m \in \mathbb{N} - \{0, 1\}$, $\Theta \subseteq \mathcal{F}$ be a set containing $m+1$ elements $\theta_0, \dots, \theta_m$ and, for any $j \in \{0, \dots, m\}$, $\mathbb{P}_j = \mathbb{P}_{\theta_j}$. We make the following assumptions:*

(H1) For any $(j, k) \in \{0, \dots, m\}^2$ with $j \neq k$, there exists $\delta > 0$ such that

$$d(\theta_j, \theta_k) \geq 2\delta.$$

(H2) Let K be the Kullback divergence defined by

$$K(\mathbb{P}, \mathbb{Q}) = \begin{cases} \int \ln \left(\frac{d\mathbb{P}}{d\mathbb{Q}} \right) d\mathbb{P} & \text{if } \mathbb{P} \ll \mathbb{Q}, \\ \infty & \text{otherwise.} \end{cases}$$

There exists $\alpha \in (0, 1/8)$ such that

$$\mathcal{K}_m = \inf_{v \in \{0, \dots, m\}} \frac{1}{m} \sum_{\substack{k \in \{0, \dots, m\} \\ k \neq v}} K(\mathbb{P}_k, \mathbb{P}_v) \leq \alpha \log m.$$

Then there exists a constant $c > 0$ such that

$$\sup_{\hat{\theta}} \sup_{\theta \in \Theta} \mathbb{P}_{\theta}(d(\hat{\theta}, \theta) \geq \delta) \geq c.$$

Consider the Besov balls $B_{p,r}^s(M)$. Let j_0 be an integer suitably chosen below. For any $\varepsilon = (\varepsilon_k)_{k \in \Lambda_{j_0}} \in \{0, 1\}^{\text{Card}(\Lambda_{j_0})}$, set

$$h_{\varepsilon}(x) = \rho(x) + M_* 2^{-j_0(s+1/2)} \sum_{k \in \Lambda_{j_0}} \varepsilon_k \psi_{j_0, k}(x),$$

where $M_* > 0$ is a constant,

$$\rho(x) = \frac{C_0}{(1+x^2)^{r_0}},$$

with $r_0 \in (1/2, 1)$ and $C_0 > 0$ is such that ρ is a density. Then h_{ε} is a density and, with a suitable M_* , $h_{\varepsilon} \in B_{p,r}^s(M)$ (see (Fan and Koo 2002, Lemma 4)).

The Varshamov-Gilbert theorem (see (Tsybakov 2004, Lemma 2.7)) asserts that there exist a set $E_{j_0} = \{\varepsilon^{(0)}, \dots, \varepsilon^{(T_{j_0})}\}$ and two constants, $c \in]0, 1[$ and $\alpha \in]0, 1[$, such that, for any $u \in \{0, \dots, T_{j_0}\}$, $\varepsilon^{(u)} = (\varepsilon_k^{(u)})_{k \in \Lambda_{j_0}} \in \{0, 1\}^{\text{Card}(\Lambda_{j_0})}$ and any $(u, v) \in \{0, \dots, T_{j_0}\}^2$ with $u < v$, the following hold:

$$\sum_{k \in \Lambda_{j_0}} |\varepsilon_k^{(u)} - \varepsilon_k^{(v)}| \geq c 2^{j_0}, \quad T_{j_0} \geq e^{\alpha 2^{j_0}}. \quad (15)$$

Let us now consider the set $\Theta = \{h_{\varepsilon^{(u)}}(x); u \in \{0, \dots, T_{j_0}\}\}$ and the \mathbb{L}^2 -distance $d(h, k) = \left(\int_{-\Omega}^{\Omega} (h(x) - k(x))^2 dx \right)^{1/2}$ for $(h, k) \in (\mathbb{L}^2([-\Omega, \Omega]))^2$. Note that, due to the Markov inequality, for any real number $\delta > 0$, we have

$$\inf_{\tilde{f}} \sup_{f \in B_{p,r}^s(M)} \mathbb{E} \left(\int_{-\Omega}^{\Omega} (\tilde{f}(x) - f(x))^2 dx \right) \geq p\delta^2, \quad (16)$$

where

$$p = \inf_{\tilde{f}} \sup_{u \in \{0, \dots, T_{j_0}\}} \mathbb{P}_{h_{\varepsilon^{(u)}}} \left(d(\tilde{f}, h_{\varepsilon^{(u)}}) \geq \delta \right)$$

and $\mathbb{P}_f = \times_{v=1}^n \mathbb{P}_f^v$ where \mathbb{P}_f^v is the probability measure related to (1).

In order to bound p , let us now investigate the assumptions (H1) and (H2) of Theorem 4 with the set $\Theta = \{h_{\varepsilon^{(u)}}(x); u \in \{0, \dots, T_{j_0}\}\}$ previously defined (so $m = T_{j_0}$) and the \mathbb{L}^2 -distance.

On (H1). For any $(u, v) \in \{0, \dots, T_{j_0}\}^2$ with $u \neq v$, using the orthonormality of \mathcal{B} , the fact that, for any $k \in A_{j_0}$, $|\varepsilon_k^{(u)} - \varepsilon_k^{(v)}| \in \{0, 1\}$ and (15), we have

$$\begin{aligned} d(h_{\varepsilon^{(u)}}, h_{\varepsilon^{(v)}}) &= \left(\int_{-\Omega}^{\Omega} (h_{\varepsilon^{(u)}}(x) - h_{\varepsilon^{(v)}}(x))^2 dx \right)^{1/2} \\ &= M_* 2^{-j_0(s+1/2)} \left(\int_{-\Omega}^{\Omega} \left(\sum_{k \in A_{j_0}} (\varepsilon_k^{(u)} - \varepsilon_k^{(v)}) \psi_{j_0, k}(x) \right)^2 dx \right)^{1/2} \\ &= M_* 2^{-j_0(s+1/2)} \left(\sum_{k \in A_{j_0}} |\varepsilon_k^{(u)} - \varepsilon_k^{(v)}| \right)^{1/2} \geq c 2^{-j_0(s+1/2)} 2^{j_0/2} = c 2^{-j_0 s}. \end{aligned}$$

Therefore, if we set $\delta = (c/2) 2^{-j_0 s}$, we have

$$d(h_{\varepsilon^{(u)}}, h_{\varepsilon^{(v)}}) \geq 2\delta. \quad (17)$$

On (H2). Let us now bound $\mathcal{K}_{T_{j_0}}$. Let \star be the convolution product. Let χ^2 be the chi-square divergence defined by

$$\chi^2(\mathbb{P}, \mathbb{Q}) = \begin{cases} \int \left(\frac{d\mathbb{P}}{d\mathbb{Q}} - 1 \right)^2 d\mathbb{Q} & \text{if } \mathbb{P} \ll \mathbb{Q}, \\ \infty & \text{otherwise,} \end{cases}$$

and set

$$\mathcal{F}_{j_0}(\xi) = \left\{ h \in \mathbb{L}^2([-\Omega, \Omega]); h(x) = \rho(x) + \sum_{k \in A_{j_0}} \lambda_{j_0, k} \psi_{j_0, k}(x) : |\lambda_{j_0, k}| \leq \xi \right\}.$$

For any $f_2 \in \mathcal{F}_{j_0}(\xi)$ with $\xi \leq C_1 2^{-j_0(s+1/2)}$ where $C_1 > 0$ denotes a suitable constant, we have $\sup_{x \in [-\Omega, \Omega]} |f_2(x) - \rho(x)| \leq (1/2) \inf_{x \in [-\Omega, \Omega]} \rho(x)$ and, a fortiori,

$$(f_2 \star g_v)(x) \geq \frac{1}{2} (\rho \star g_v)(x). \quad (18)$$

By the Fatou lemma, observe that

$$\begin{aligned} \liminf_{|x| \rightarrow \infty} (1+x^2)^{r_0} (\rho \star g_v)(x) &\geq C_0 \int_{-\infty}^{\infty} \liminf_{|x| \rightarrow \infty} \frac{(1+x^2)^{r_0}}{(1+(x-y)^2)^{r_0}} g_v(y) dy \\ &= C_0 \int_{-\infty}^{\infty} g_v(y) dy = C_0. \end{aligned}$$

Therefore, for any $v \in \{1, \dots, n\}$ and any $x \in \mathbb{R}$,

$$(\rho \star g_v)(x) \geq C_0(1+x^2)^{-r_0}. \quad (19)$$

Using (3), (4) and $\inf_{v \in \{1, \dots, n\}} \sigma_v^2 \geq c_*$, we can apply (Fan and Koo 2002, Lemma 1) with $2^{j_0} \sigma_v$ instead of 2^{j_0} . This yields the existence of a constant $C > 0$ such that, for any $x \in \mathbb{R}$,

$$\left| \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}(\psi)(y) \mathcal{F}(g_v)(2^{j_0} y) e^{-ixy} dy \right| \leq C \frac{2^{-\delta j_0}}{\sigma_v^\delta (1+|x|)^2}$$

and, for any sequence of real numbers $(u_{j_0, k})_{k \in \Lambda_{j_0}}$ such that $\sup_{k \in \Lambda_{j_0}} |u_{j_0, k}| \leq L$,

$$\left| \left(\sum_{k \in \Lambda_{j_0}} u_{j_0, k} \psi_{j_0, k} \star g_v \right) (x) \right| \leq C 2^{j_0/2} L \frac{2^{-\delta j_0}}{\sigma_v^\delta (1+|x|)^2}. \quad (20)$$

Putting (18), (19) and (20) in (Fan and Koo 2002, Proof of Lemma 4), for any functions f_1 and f_2 in $\mathcal{F}_{j_0}(\xi)$ with $\xi \leq C_1 2^{-j_0(s+1/2)}$, we have, for any $v \in \{1, \dots, n\}$,

$$\begin{aligned} \chi^2(\mathbb{P}_{f_1}^v, \mathbb{P}_{f_2}^v) &= \int_{-\infty}^{\infty} \frac{((f_1 \star g_v)(x) - (f_2 \star g_v)(x))^2}{(f_2 \star g_v)(x)} dx \\ &\leq 2 \int_{-\infty}^{\infty} \frac{(((f_1 - f_2) \star g_v)(x))^2}{(\rho \star g_v)(x)} dx \\ &\leq C 2^{j_0} \xi^2 2^{-2\delta j_0} \frac{1}{\sigma_v^{2\delta}} \int_{-\infty}^{\infty} \frac{(1+x^2)^{r_0}}{(1+|x|)^4} dx \\ &\leq C 2^{j_0} \xi^2 2^{-2\delta j_0} \frac{1}{\sigma_v^{2\delta}} \leq C 2^{-2j_0(s+1/2+\delta)} 2^{j_0} \frac{1}{\sigma_v^{2\delta}}. \end{aligned} \quad (21)$$

Using the elementary inequality: $K(P, Q) \leq \chi^2(P, Q)$, and (21), we have, for any $(u, v) \in \{0, \dots, T_{j_0}\}^2$ with $u \neq v$,

$$\begin{aligned} K(\mathbb{P}_{h_\varepsilon(u)}^v, \mathbb{P}_{h_\varepsilon(v)}^v) &= \sum_{v=1}^n K(\mathbb{P}_{h_\varepsilon(u)}^v, \mathbb{P}_{h_\varepsilon(v)}^v) \leq \sum_{v=1}^n \chi^2(\mathbb{P}_{h_\varepsilon(u)}^v, \mathbb{P}_{h_\varepsilon(v)}^v) \\ &\leq C 2^{-2j_0(s+1/2+\delta)} 2^{j_0} \sum_{v=1}^n \frac{1}{\sigma_v^{2\delta}} \\ &= C w_n^* 2^{-2j_0(s+1/2+\delta)} 2^{j_0}. \end{aligned}$$

Hence

$$\begin{aligned} \mathcal{K}_{T_{j_0}} &= \inf_{v \in \{0, \dots, T_{j_0}\}} \frac{1}{T_{j_0}} \sum_{\substack{u \in \{0, \dots, T_{j_0}\} \\ u \neq v}} K(\mathbb{P}_{h_\varepsilon(u)}^v, \mathbb{P}_{h_\varepsilon(v)}^v) \\ &\leq C w_n^* 2^{-2j_0(s+1/2+\delta)} 2^{j_0}. \end{aligned}$$

Choosing j_0 such that

$$2^{-j_0(s+1/2+\delta)} = c_0 \frac{1}{(w_n^*)^{1/2}} \quad (\text{i.e. } 2^{j_0} = (w_n^*)^{1/(2s+2\delta+1)}), \quad (22)$$

where c_0 denotes a suitable constant, (15) implies the existence of $\alpha_* \in (0, 1/8)$ satisfying

$$\mathcal{K}_{T_{j_0}} \leq Cc_0^2 2^{j_0} \leq \alpha_* \log T_{j_0}. \quad (23)$$

It follows from Theorem 4, (17), (22) and (23) that

$$p = \inf_{\tilde{f}} \sup_{u \in \{0, \dots, T_{j_0}\}} \mathbb{P}_{h_{\varepsilon(u)}} \left(d(\tilde{f}, h_{\varepsilon(u)}) \geq \delta \right) \geq c > 0$$

and, by (16),

$$\begin{aligned} \inf_{\tilde{f}} \sup_{f \in B_{p,r}^s(M)} \mathbb{E} \left(\int_{-\Omega}^{\Omega} (\tilde{f}(x) - f(x))^2 dx \right) &\geq p\delta^2 \geq c2^{-2j_0s} \\ &= c(w_n^*)^{-2s/(2s+2\delta+1)}. \end{aligned}$$

The proof of Theorem 2 is complete. □

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

References

- Antoniadis, A. (1997). Wavelets in statistics: a review (with discussion), *Journal of the Italian Statistical Society Series B* 6, 97-144.
- Butucea, C. and Matias, C. (2005). Minimax estimation of the noise level and of the signal density in a semiparametric convolution model. *Bernoulli*, 11, 2, 309-340.
- Caroll, R.J. and Hall, P. (1988). Optimal rates of convergence for deconvolving a density. *J. Amer. Statist. Assoc.*, 83, 1184-1186.
- Chesneau, C. (2011). Adaptive wavelet estimator for a function and its derivatives in an indirect convolution model. *Journal of Statistical Theory and Practice*, 5, 2, 303-326.
- Cohen, A., Daubechies, I., Jawerth, B. and Vial, P. (1993). Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis*, 24, 1, 54-81.
- Comte, F., Rozenholc, Y. and Taupin, M.-L. (2006). Penalized contrast estimator for density deconvolution. *The Canadian Journal of Statistics*, 34, 431-452.
- Delaigle, A. and Gijbels, I. (2006). Estimation of boundary and discontinuity points in deconvolution problems. *Statistica Sinica*, 16, 773-788.
- Delaigle, A. and Meister, A. (2008). Density estimation with heteroscedastic error. *Bernoulli*, 14, 562-579.
- Devroye, L. (1989). Consistent deconvolution in density estimation. *Canad. Journ. Statist.*, 17, 235-239.
- Fan, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problem. *Ann. Statist.*, 19, 1257-1272.
- Fan, J. and Koo, J.Y. (2002). Wavelet deconvolution. *IEEE transactions on information theory*, 48, 734-747.

- Hall, P. and Qiu, P. (2005). Discrete-transform approach to deconvolution problems. *Biometrika* **92** 135-148.
- Hall, P. and Meister, A. (2007). A ridge-parameter approach to deconvolution. *Ann. Statist.*, **35** 1535-1558.
- Härdle, W., Kerkycharian, G., Picard, D. and Tsybakov, A. (1998). *Wavelet, Approximation and Statistical Applications*, Lectures Notes in Statistics New York 129, Springer Verlag.
- Lacour, C. (2006). Rates of convergence for nonparametric deconvolution. *C. R. Acad. Sci. Paris Ser. I Math.*, **342** (11), 877-882.
- Lounici, K. and Nickl, R. (2011). Global Uniform Risk Bounds for Wavelet Deconvolution Estimators. *The Annals of Statistics*, **39**, 201-231.
- Mallat, S. (2009). A wavelet tour of signal processing. Elsevier/ Academic Press, Amsterdam, third edition. The sparse way, With contributions from Gabriel Peyré.
- Meister, A. (2009). *Deconvolution Problems in Nonparametric Statistics*. Lecture Notes in Statistics. Springer, New York.
- Meister, A., Stadtmüller, U. and Wagner, C. (2010). Density deconvolution in a two-level heteroscedastic model with unknown error density. *Electron. J. Statist.*, **4**, 36-57.
- Meyer, Y. (1992). Wavelets and Operators. *Cambridge University Press*, Cambridge.
- Pensky, M. and Vidakovic, B. (1999). Adaptive wavelet estimator for nonparametric density deconvolution. *The Annals of Statistics*, **27**, 2033-2053.
- Pensky, M. and Sapatinas, T. (2010). On convergence rates equivalency and sampling strategies in functional deconvolution models. *Ann. Statist.*, **38**, 3, 1793-1844.
- Staudenmayer, J., Ruppert, D. and Buonaccorsi, J. (2008). Density estimation in the presence of heteroskedastic measurement error, *JASA*, **103**, 726-736.
- Tsybakov, A. (2004). *Introduction à l'estimation nonparamétrique*. Springer Verlag, Berlin.
- Wang, X.F., Fan, Z. and Wang, B. (2010). Estimating smooth distribution function in the presence of heterogeneous measurement errors. *Computational Statistics and Data Analysis*, **54**, 25-36.
- Wang, X.F. and Wang, B. (2011). Deconvolution estimation in measurement error models: The R package decon. *Journal of Statistical Software*, **39**(10), 1-24.
- Wavelab toolbox. <http://www-stat.stanford.edu/~wavelab>, 2001.
- Zhang, S. and Karunamuni, R. (2000). Boundary bias correction for nonparametric deconvolution. *Ann. Inst. Statist. Math.*, **52**, 612-629.