A wavelet method for unfolding sphere size distributions

Anestis Antoniadis[1], Jianqing Fan[2] and Irène Gijbels[3].

June 2000

Abstract

We consider the problem of estimating an unknown probability density function of the radii of spheres in a medium, based on their observed random cross-sections. This problem is known as Wicksell's corpuscle problem. We first convert Wicksell's integral equation which combines a size distribution of discs on a sectional plane with that of spheres to a form suitable for the application of thresholding wavelet methods for the solution of ill-posed integral equations given noisy data. The asymptotic properties of our estimators are derived and a comparison with other available methods is explored via a Monte Carlo simulation study. The newly introduced approach is also illustrated by the analysis of some real data.

Key words and phrases: Abel integral equation, adaptive estimation, Besov spaces, corpuscle problem, fractional integration, linear estimators, wavelets, wavelet thresholding.

AMS 1991 subject classifications: Primary 62G07; secondary 65Dxx.

RUNNING HEAD: Wicksell's problem.

[1] Université Joseph Fourier, Laboratoire LMC-IMAG, BP 53, 38041 Grenoble Cedex 9, France.
[2] Department of Statistics, Chinese University of Hong Kong, Shatin, Hong Kong.
[3] Institut de Statistique, Université catholique de Louvain, 20 Voie du Roman Pays, B-1348 Louvain-la-Neuve, Belgium.

1 Introduction

Consider an aggregate of non-overlapping balls of random sizes scattered in a three-dimensional specimen. In the Wicksell problem (Wicksell (1925)) one observes cross-sections, formed by a given plane Π in $\mathbb{R}^3$, of the random spheres $B_i = B_i(\mathbf{c_i}, Y_i)$, $i = 1, \ldots, n$, with centers $\mathbf{c_i}$ given the sites of a stationary Poisson point process, and with independent identically distributed radii Y_i, such that Y_i has an unknown probability density function $f_3(y)$, $y \in \mathbb{R}^+$. It is assumed that the sphere number density is small enough so that distinct spheres do not interfere with one another. We are interested in estimating f_3 at a given point $y > 0$, from a sample $X_1, \ldots, X_n$ of planar circular profiles with probability density function f_2. This problem arises in many applications in biology, stereology, stochastic geometry and so on. See, for example, Stoyan et al. (1995). Further references and related problems can be found in Hall and Smith (1988), in Krumbein (1935), in Takahashi and Sibuya (1996) and the survey paper by Hoogendoorn (1992).

The basic relation of the corpuscle problem is as follows. By a conditioning argument, and under the above assumptions, the densities f_2 and f_3 are related to each other as follows :

$$f_2(y) = W f_3(y) \tag{1}$$

where the operator W is defined as

$$Wh(y) = \begin{cases} \dfrac{y}{m} \displaystyle\int_y^T \frac{h(r)}{\sqrt{r^2 - y^2}} dr & \text{if} \quad y > 0 \\ 0 & \text{if} \quad y \leq 0, \end{cases}$$

where T is an upper bound for the maximum possible value of the three-dimensional radius (possibly infinite), and with the average three-dimensional radius denoted by

$$m = \int_0^T r f_3(r) dr. \tag{2}$$

One can easily estimate the probability density f_2 from the observed planar circular profiles $X_1, \ldots, X_n$ and then try to "unfold" the unknown density f_3 using equation (1). However, the problem of recovering estimates of f_3 from observations on f_2 is harder than might appear at first glance because it is ill-posed, meaning that small changes in f_2 may lead to substantial changes in the corresponding f_3. In particular, high frequency components in f_3 will not in general be recoverable from medium or even large samples from f_2. In order to find a solution that remains stable under small changes in the data and that is consistent both with the data and with prior assumptions, one usually relies on some kind of regularization.

To solve this problem several regularization techniques have been proposed in the literature, including Saltykov's method (Saltykov (1967)), smoothing kernel methods (Taylor (1983), Hall and Smith (1988)), smoothing spline methods (Nychka et al. (1984)), Gauss-Chebyshev type quadrature methods for solving integral equations (Mase (1995)), and singular value decomposition (SVD) methods (O'Sullivan (1986) and Silverman et al. (1990)) to cite only a few. The SVD method expands the function f_3 in a basis formed by the eigenfunctions of the self-adjoint operator W^*W where W^* is the adjoint of W. When noisy data about $(Wf_3)(y)$ are observed, the series is truncated and the coefficients of the eigenfunctions in the expansion are estimated from the data. The SVD method has certain limitations, however. The basis functions are completely derived from the operator W, not from the object of interest f_3. When the function f_3 is of inhomogeneous smoothness, the representation of f_3 by the eigenfunctions of W^*W is often inefficient and the resulting estimator does not perform well.

Wavelet bases offer efficient sparse representations for functions in a wide range of function spaces and wavelet methods have demonstrated considerable success in nonparametric function estimation in terms of spatial adaptivity and asymptotic optimality. A properly chosen wavelet basis can simultaneously quasi-diagonalize both the operator W and the functions in a range of function classes. In this paper we consider a variant of the wavelet-vaguelette decomposition (WVD) method for linear inverse problems originally proposed by Donoho (1995), which works by expanding the function f_3 in a wavelet series and producing a corresponding vaguelette series for Wf_3, and then estimating the wavelet coefficients by thresholding the empirical vaguelette coefficients.

The WVD approach of Donoho (1995) is relatively easy to describe when the integration operator in the corresponding integral equation is a homogeneous convolution operator (see e.g. Kolaczyk (1996)). Unfortunately, this is not the case for Wicksell's integral equation. In Section 2, we propose an equivalent formulation of Wicksell's problem that involves a transformation similar to the one used by Fan and Koo (1999) for the deconvolution of a density. In Section 3 a linear estimator for the linear inverse problem at hand is proposed using the appropriate type of wavelets adapted to our case.

In the spirit of wavelet denoising methods (Donoho (1995), Abramovich and Silverman (1998)), we improve the estimator by applying a hard-threshold nonlinearity to the vaguelette coefficients.

The last section is devoted to the numerical implementation of our procedure, where we also present the results of a Monte Carlo experiment designed to study the sampling properties of our estimates in presence of noise and where we also analyze a real data example.

2 The functional framework

Many different approaches have been used in the past to solve Wicksell's problem or variants of it. The oldest technique for finding a solution is by discretizing the distribution of the sphere radii and the distribution of the planar circular profiles, the last discretization being estimated by an histogram of the observed data. The integral equation becomes then a system of linear equations which is then solved by numerical methods. This method and several of its variants are related to Saltykov's method (Saltykov (1967)). The way the data are grouped (the choice of the histograms bin-widths) is critical for such methods and may give unreliable results for moderate samples.

Taylor (1983) was the first to use a kernel estimator for estimating by a smooth function the density of the planar circular profiles which is then substituted in the integral equation to obtain an estimate of the density of the sphere radius. Taylor's method, which can be viewed as a smoother operation than the one produced by the histogram binning of the Saltikov method, relies also upon a good choice of the window width of the kernel smoother. Hall and Smith (1988) have shown that Taylor's method can be efficient only if the number N of observed planar circular sections is very large, which is rarely true in practical situations. A kernel based approach for estimating the cumulative distribution of the squared three-dimensional radii has also been derived recently by Golubev and Levit (1998). For reasonable sample sizes, a data-driven choice of the window width becomes a necessity but the use of generalized cross-validation techniques for choosing the window size can be very costly in computations. Recently Feuerverger and Hall (2000) studied a thick-slice version of Wicksell's problem, and introduced an appropriate cross-validation criterion for selecting the kernel smoothing parameter.

Nychka et al. (1984) use a cross-validated spline regularization method to estimate the density of the sphere radii. The solution is obtained by approximating the unknown density by a smooth combination of natural splines in such a way that its integral transform stays close to an histogram of the data, with a number of classes much larger that the one that should be employed to estimate optimally the density of the planar sections. Such a method, which relies on global regularization, while giving satisfactory results for densities that are quite smooth, is not very well adapted when the degree of smoothness of the solution is small and variable.

As already noted in the introduction, if one focuses on building optimal estimators for a choice of densities that better model the spatial inhomogeneity of complex signals such as those observed in practice then it is natural to proceed to the identification with a thresholding in an orthogonal wavelet basis which provides a sparse signal representation and nearly diagonalizes the integral operator. This is indeed the theoretical framework that we have chosen in this paper to justify and implement an algorithm in an appropriate wavelet basis. In order to proceed we will first transform the original problem into a relationship between the densities of the squared circle and sphere radius, which has a structure easier to work with and which is perhaps more natural in practice as well.

Define $g(y) = (2y^{1/2})^{-1} f_3(y^{1/2})$ and $f(x) = (2x^{1/2})^{-1} f_2(x^{1/2})$; these are the densities of squared radii for spheres and discs, respectively. The integral equation (1) becomes (see also Golubev and Levit (1998)):

$$f(x) = \frac{1}{2m} \int_x^{+\infty} \frac{1}{\sqrt{y-x}} g(y)dy, \tag{3}$$

and its solution, when it exists is given by

$$g(y) = -2m\pi^{-1} \frac{d}{dy} \left\{ \int_y^{+\infty} \frac{1}{\sqrt{x-y}} f(x)dx \right\}. \tag{4}$$

Setting

$$\bar{g}(y) = \frac{g(y)\Gamma(1/2)}{2m},$$

equation (3) may be written as

$$f(x) = \frac{1}{\Gamma(1/2)} \int_x^{T} \frac{1}{\sqrt{y-x}} \bar{g}(y)dy, \quad 0 \le x \le T, \tag{5}$$

where T is an upper bound of the support of g. Such an equation is a form of Abel's integral equation, closely connected to fractional integro-differentiation of order 1/2 (see e.g. Samko et al. (1993) or Gorenflo and Vessela (1980)). Note also that by a change of variable the following relation holds

$$f(T-x) = \frac{1}{\Gamma(1/2)} \int_0^x \frac{1}{\sqrt{x-y}} \bar{g}(T-y)dy, \quad 0 \leq x \leq T,$$

i.e.

$$\tilde{f}(x) = \frac{1}{\Gamma(1/2)} \int_0^x \frac{1}{\sqrt{x-y}} \tilde{\bar{g}}(y)dy,$$

where $\tilde{h}(x) = h(T-x)$. Since both f and g are compactly supported with support within $[0,T]$, the above equation is a particular case of the problem

$$I_{-\infty}^{1/2} u = v, \quad \mathrm{Supp}(u) \subset [0,+\infty[, \ \mathrm{Supp}(v) \subset [0,+\infty[, \tag{6}$$

where $I_{-\infty}^{1/2}$ denotes the operator of fractional integration of order 1/2.
Note also that formally (see Samko et al. (1993), p. 29), equation (6) admits the following inversion formula

$$u(x) = -\frac{1}{\Gamma(1/2)} \frac{d}{dx} \int_x^\infty \frac{v(t)}{\sqrt{t-x}}\, dt = (D_+^{1/2} v)(x),$$

where $D_+^{1/2}$ denotes the operator of fractional differentiation of order 1/2.

Recall that the Sobolev space $H^s(\mathbb{R})$, $s \in \mathbb{R}$, is the space of tempered distributions v such that

$$\|v\|_s^2 = \int_{\mathbb{R}} (1+|\xi|^2)^s |\hat{v}(\xi)|^2 \, d\xi < \infty,$$

where

$$\hat{v}(\xi) = \int_{\mathbb{R}} e^{i\xi t} v(t)\, dt,$$

denotes the Fourier transform of v.

Now, according to Theorem 6.3.5 of Gorenflo and Vessela (1980), if $v \in H^s(\mathbb{R})$, $s \in \mathbb{R}$, $\mathrm{Supp}(v) \subset [0,+\infty[$, then the unique solution u of (6) in the space of tempered distributions is in the Sobolev space $H^{s-1/2}(\mathbb{R})$ and

$$\|u\|_{s-1/2} \leq \|v\|_s.$$

Looking at the proof of the above mentioned theorem, equation (6) is equivalent, in a distributional sense, to

$$(i\xi)^{-1/2} \hat{u}(\xi) = \hat{v}(\xi), \tag{7}$$

or equivalently,

$$\hat{u}(\xi) = (i\xi)^{1/2}\hat{v}(\xi) \tag{8}$$

where $(i\xi)^{-1/2} = |\xi|^{-1/2}e^{-i\mathrm{sgn}(\xi)\pi/4}$ (see e.g. Gorenflo and Vessela (1980)).

In order to take full advantage of Theorem 6.3.5 and of sparse representations of functions in Besov spaces by wavelet approximations we will assume hereafter that the density f (and therefore v) belongs to the set

$$\mathrm{D}_{\sigma p2}(M,T) = \{f; \int f = 1, f \geq 0, \|f\|_{\sigma p2 \leq M}, \mathrm{Supp}(f) \subset [0,T]\},$$

where $\|\cdot\|_{\sigma p2}$ denotes the norm of the Besov space $B_{\sigma p2}(\mathbb{R})$ with $1 \leq p \leq 2$ and $\sigma = s+\frac{1}{p}-1$, with $s > 1/2$. Indeed, under such conditions it is known that the following continuous inclusion holds

$$B_{\sigma p2}(\mathbb{R}) \subset B_{ss2}(\mathbb{R}) = H^s(\mathbb{R}),$$

and therefore equation (6) makes perfectly sense. Moreover the condition $s > 1/2$ insures that the solution u will lie in $H^{s-1/2}(\mathbb{R})$ which is then a Sobolev space with positive index. Note that when $s < 1$, functions from $H^s(\mathbb{R})$ can be very rough and include hat functions or step functions. Another reason of seeking a solution in a Sobolev space with positive index is that the level of smoothness of a density is naturally described in terms of its characteristic function Φ, since it is the maximum index r for which $\int_{\mathbb{R}}(1+|\xi|^2)^r|\Phi(\xi)|^2\,d\xi < \infty$.

Before closing this section, note that at this stage we have assumed that the constant m is known. In practice it will be unknown, but the formula (see Gorenflo and Vessela (1980))

$$\mathbb{E}\{X^{-1}\} = \pi/(2m),$$

suggests the estimator

$$\hat{m} = n\pi/2\left(\sum_{\ell} X_i^{-1}\right)^{-1}.$$

It may be shown (see Hall and Smith (1988) and Ripley (1981)) that

$$\hat{m} - m = \mathcal{O}_P(n^{-1/2+\delta}),$$

for any $\delta > 0$. In other words, the rate of convergence of $\hat{m}$ to m is arbitrary close to the order $n^{-1/2}$, which will be a faster rate of convergence than that of the nonparametric estimator of $\bar{g}$ (or u) which we will obtain in the following sections. Consequently, $\hat{m}$ may be substituted for m in the definition of $\widehat{\bar{g}}$, with no change in the asymptotic behaviour of the estimator.

3 Wavelet estimation

We start this section by briefly reviewing some useful facts from basic wavelet theory, that will be used to derive our estimators. A general introduction to the theory of wavelets can be found in Chui (1992), Daubechies (1992) and Walter (1994).

The construction of orthonormal wavelet bases for $L^2(\mathbb{R})$ is now well understood. There are many families of wavelets. Throughout this paper we will consider compactly supported wavelets such as Daubechies' orthogonal wavelets. For the construction of orthonormal bases of compactly supported wavelets for $L^2(\mathbb{R})$, one starts with a couple of special, compactly supported functions known as the scaling function ϕ and the wavelet ψ. The collection of functions $\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$, $j, k \in \mathbb{Z}$, then constitutes an orthonormal basis for $L^2(\mathbb{R})$. For fixed $j \in \mathbb{Z}$, the $\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k)$, $k \in \mathbb{Z}$ are an orthonormal basis for a subspace $V_j \subset L^2(\mathbb{R})$; the spaces V_j constitute a multiresolution analysis, meaning in particular that

$$\cdots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \cdots$$

with

$$\cap_j V_j = \{0\}, \qquad \overline{\cup_j V_j} = L^2(\mathbb{R}).$$

Denoting by $P_{V_j} f = \sum_{k \in \mathbb{Z}} < f, \phi_{j,k} > \phi_{j,k}$ the orthogonal projection of f on the approximation space V_j, we have

$$P_{V_{j+1}} f = P_{V_j} f + \sum_{k \in \mathbb{Z}} < f, \psi_{j,k} > \psi_{j,k}.$$

The multiresolution analysis is said to be r-regular if ϕ is C^r, and if both ϕ and its derivatives have a fast decay,

$$|d^\alpha \phi(x)| \leq C_m (1 + |x|)^{-m}, \quad \forall m \in \mathbb{N}, \quad 0 \leq \alpha \leq r.$$

One can prove that if a multiresolution analysis is r-regular, the wavelet ψ is also C^r and has vanishing moments up to the order r (see e.g. Daubechies (1992), Corollary 5.2):

$$\int_{-\infty}^{+\infty} x^k \psi(x) dx = 0 \quad \text{for} \quad 0 \leq k \leq r.$$

The converse is generally false, and the number of vanishing moments is usually larger than the regularity of the multiresolution analysis.

The smoother wavelets provide not only orthonormal bases for $L^2(\mathbb{R})$, but also unconditional bases for several function spaces including Besov spaces Triebel (1983). In particular one can determine whether a function f belongs to such spaces simply by examining its wavelet coefficients. Moreover, it has been proved (see Meyer (1990), Lemma 4 of Section 6.2) that fractional integrals and fractional derivatives of regular enough wavelets, provided they are viewed in the context of generalized function theory, form Riesz bases of $L^2(\mathbb{R})$.

Let ϕ and ψ be, respectively, a scaling function and a wavelet generated by an orthonormal multiresolution decomposition of $L^2(\mathbb{R})$. For any integer m, the function u of (6) admits the following decomposition:

$$u = \sum_{k\in\mathbb{Z}} \alpha_{m,k}\phi_{m,k} + \sum_{j=m}^{\infty}\sum_{k\in\mathbb{Z}} \beta_{j,k}\psi_{j,k}, \tag{9}$$

and the coefficients $\alpha_{m,k}$ and $\beta_{j,k}$ have the forms:

$$\alpha_{m,k} = \int_{\mathbb{R}} \phi_{m,k}(x)u(x)dx \quad \text{and} \quad \beta_{j,k} = \int_{\mathbb{R}} \psi_{j,k}(x)u(x)dx.$$

In linear wavelet estimation procedures one usually estimates the scaling coefficients of the function to be estimated. In order to propose unbiased estimators of the unknown scaling coefficients of u, we are going to give an alternative representation of the coefficients $\alpha_{m,k}$. Observe that, for regular enough wavelets, using expression (8) and the existing results of Meyer (1990), we have:

$$\begin{aligned}
\alpha_{m,k} &= \int_{\mathbb{R}} \phi_{m,k}(x)u(x)dx = \int_{\mathbb{R}} \phi_{m,k}(x)\left(D_+^{1/2}v\right)(x)dx \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\phi_{m,k}}(-\xi)\widehat{D_+^{1/2}v}(\xi)d\xi = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\phi_{m,k}}(-\xi)(i\xi)^{1/2}\hat{v}(\xi)d\xi,
\end{aligned}$$

where $\widehat{\phi_{m,k}}$ denotes the Fourier transform of $\phi_{m,k}$. It follows that the scaling coefficients admit the following representation:

$$\alpha_{m,k} = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\phi_{m,k}}(-\xi)(i\xi)^{1/2}\Phi_X(\xi)d\xi, \tag{10}$$

where Φ_X is the characteristic function of the squared random disc radius.

Now, let $\hat{\Phi}_n(\xi)$ be the empirical characteristic function of the observed squared radii sections

$$\hat{\Phi}_n(\xi) = \frac{1}{n}\sum_{\ell=1}^{n} \exp(i\xi X_\ell),$$

and define the scaling coefficient estimator

$$\hat{\alpha}_{m,k} = \frac{1}{2\pi}\int_{\mathbb{R}} \widehat{\phi_{m,k}}(-\xi)(i\xi)^{1/2}\hat{\Phi}_n(\xi)d\xi.$$

Relation (10) implies that $\hat{\alpha}_{m,k}$ is an unbiased estimator of $\alpha_{m,k}$. Note also that we can rewrite

$$\hat{\alpha}_{m,k} = \frac{1}{n}\sum_{\ell=1}^{n} \left(\mathcal{K}_m^{-}\phi\right)_{m,k}(X_\ell), \tag{11}$$

where the operator $\mathcal{K}_m^{-}$ is formally defined by

$$\left(\mathcal{K}_m^{-}h\right)(x) = \frac{1}{2\pi}\int_{\mathbb{R}} e^{i\xi x}\hat{h}(-\xi)(i2^m\xi)^{1/2}d\xi. \tag{12}$$

Assuming that the scaling function ϕ is r-regular with $r > 3/2$, then

$$|\hat{\phi}(\xi)| \le C(1+|\xi|)^{-r},$$

where C is a generic constant (see Daubechies (1992), page 225). The last inequality implies that $\mathcal{K}_m^{-}\phi$ is well defined. Since ψ inherits the same smoothness properties than ϕ the wavelet coefficients can be estimated similarly by

$$\hat{\beta}_{m,k} = \frac{1}{n}\sum_{\ell=1}^{n} \left(\mathcal{K}_m^{-}\psi\right)_{m,k}(X_\ell). \tag{13}$$

Therefore, estimating $\alpha_{m,k}$ by (13) and $\beta_{j,k}$ by (15) we obtain a linear wavelet estimator of u by

$$\hat{u}_{m,n} = \sum_{k\in\mathbb{Z}} \hat{\alpha}_{m,k}\phi_{m,k}, \tag{14}$$

and a nonlinear wavelet estimator by

$$\hat{u}_{j_0,j_1,n} = \sum_{k\in\mathbb{Z}} \hat{\alpha}_{j_0,k}\phi_{j_0,k} + \sum_{j=j_0}^{j_1}\sum_{k\in\mathbb{Z}} \hat{\beta}_{j,k}\psi_{j,k}, \tag{15}$$

where $j_0 < j_1$ are integers.

The asymptotic properties of the above estimators will be good if one chooses properly the resolution indices m, j_0 and j_1. We will first investigate the properties of the linear wavelet estimator of u, given by (14) with an appropriately chosen truncation index m. We will show that with such a choice of m the linear wavelet estimator achieves asymptotic rates similar to the optimal minimax rates of convergence of linear wavelet deconvolution density estimators established by Fan and Koo (1999) (see also Pensky and Vidakovic (1999)).

Since both u and v are compactly supported, note that by expression (11), when the scaling function has a bounded support, we only need k's in (14) for which $< u, \phi_{m,k} >\neq 0$. The number of such k's in this case is less than $C2^m$ for some constant C and therefore $\hat{u}_{m,n}$ involves only a finite number of terms in its definition, when compactly supported wavelets are used.

We shall measure the performances of an estimator $\hat{u}$ of u by its MISE defined by

$$\mathrm{MISE}(\hat{u}) = \mathbb{E}\left[\|\hat{u} - u\|_2^2\right].$$

The following theorem whose proof is given in the Appendix states an asymptotic MISE rate for our linear wavelet estimators when Daubechies' wavelets of regularity $r > 3/2$ are used.

Theorem 3.1 *Suppose that $1 \leq p \leq 2$ and that v belongs to the set $D_{\sigma p 2}(M,T)$ with $\sigma = s + \frac{1}{p} - 1$ and $s > 1$. Let the estimator $\hat{u}_{m,n}$ be defined by means of a Daubechies' wavelet of regularity $r > 3/2$. Then*

$$\sup_{u \in D_+^{1/2}(D_{\sigma p 2}(M,T))} \left[\mathbb{E}\|\hat{u}_{m,n} - u\|_2^2\right]^{1/2} \leq C n^{-\sigma'/(2\sigma'+2)},$$

where $\sigma' = s - 1/2$, provided that $2^m = \mathcal{O}(n^{1/(2s+1)})$.

It is not our purpose here to show that the asymptotic rate given in Theorem 3.1 is an optimal minimax rate among linear estimators over the class $\mathrm{D}_{\sigma p 2}(M,T)$. Note however, that, while derived in a different setting, the above asymptotic rate under integrated square error loss coincides with the optimal minimax rate over the class of linear deconvolution estimators with an ordinary smooth convolution kernel, established by Fan (1993) (see also Carroll and Hall (1988)).

In order to derive a MISE rate, similar to the minimax one obtained in deconvolution estimation problems over the class $\mathrm{D}_{\sigma p2}(M,T)$ with ordinary smooth kernels, we will now use a thresholding nonlinear wavelet estimation procedure.

When regular enough Daubechies' wavelets are used in the decomposition of u we have already seen that the estimators of the wavelet coefficients $\beta_{j,k}$ are given by

$$\hat{\beta}_{m,k} = \frac{1}{n}\sum_{\ell=1}^{n}\left(\mathcal{K}_m^{-}\psi\right)_{m,k}(X_\ell),$$

are unbiased.

Let us now, in a manner similar to the one in Fan and Koo (1999), employ a hard-thresholding:

$$\tilde{\beta}_{j,k} = \begin{cases} \hat{\beta}_{j,k}, & \text{if } |\hat{\beta}_{j,k}| > \lambda c_j n^{-1/2}, \\ 0, & \text{if } |\hat{\beta}_{j,k}| \le \lambda c_j n^{-1/2} \end{cases}, \tag{16}$$

where $c_j = 2^{j/2}\sqrt{j}$, and λ is a thresholding constant to be chosen appropriately. Then the nonlinear wavelet thresholding estimator defined formally by (15), associated to the parameters $j_0(n) < j_1(n)$ and λ is given by:

$$\hat{u}_{\lambda,n} = \hat{u}_{j_0,n} + \sum_{j=j_0}^{j_1}\sum_{k\in\mathbb{Z}}\tilde{\beta}_{j,k}\psi_{j,k}. \tag{17}$$

Now, in analogy with ordinary nonlinear wavelet density estimation, as studied by Donoho et al. (1996) (see also Donoho and Johnstone (1995)), take, j_0 and j_1 such that

$$2^{j_0(n)} \simeq \left(n(\log n)^{(2-p)/p)}\right)^{1/(2\sigma+2)} \quad \text{and} \quad 2^{j_1(n)} \simeq n^{\gamma/\sigma'},$$

where

$$\gamma = \frac{\sigma}{2\sigma+2} = \frac{s+1/p-1}{2(s+1/p)}.$$

We then have

Theorem 3.2 *Suppose that $1 \le p \le 2$ and that v belongs to the set $D_{\sigma p2}(M,T)$ with $\sigma = s+\frac{1}{p}-1$ and $s > 1$. Let the estimator $\hat{u}_{\lambda,n}$ be defined by means of Daubechies' wavelets of regularity $r \ge 2$. Then there exist constants $C = C(s,p,T)$ and $\lambda_0 = \lambda_0(s,p,T)$ such that, for $\lambda > \lambda_0$,*

$$\sup_{u\in D_+^{1/2}(D_{\sigma p2}(M,T))}\left[\mathbb{E}\|\hat{u}_{\lambda,n} - u\|_2^2\right]^{1/2} \le Cn^{-\gamma}(\log n)^{((2-p)/((s-1)p+1))\gamma}.$$

The proof of this theorem is given in the Appendix and relies upon some moment bounds and some large deviation results on the sequence of the estimated wavelet coefficients. Observe that the rate of convergence in Theorem 3.2 coincides with the optimal rate of convergence for nonlinear wavelet deconvolution estimation (see Fan and Koo (1999)). It is however, non adaptive, i.e. the choice of the parameters m, j_0, and j_1 depends on the unknown smoothness s of the density v. This difficulty can be overcome by using Daubechies' wavelets of regularity $r > 2$. Indeed, the modified adaptive estimator $\hat{u}^A_{\lambda,n}$ is obtained as before but using now resolution levels j_0 and j_1 defined by:

$$2^{j_0(n)} \simeq n^{1/(2r+2)} \quad \text{and} \quad 2^{j_1(n)} \simeq n/\log_2 n.$$

We then have

Theorem 3.3 *Suppose that* $1 \leq p \leq 2$ *and that* v *belongs to the set* $D_{\sigma p2}(M,T)$ *with* $\sigma = s + \frac{1}{p} - 1$ *and* $s > 1$. *Let the estimator* $\hat{u}^A_{\lambda,n}$ *be defined by means of Daubechies' wavelets of regularity* $r > 2$. *Then, for all* $s > 1$, $1 \leq p$ *and* $0 < T < \infty$, *there exist constants* $c = c(s,p,T)$ *and* λ_0 *such that, for* $\lambda > \lambda_0$,

$$\sup_{u \in D_+^{1/2}(D_{\sigma p2}(M,T))} \left[\mathbb{E}\|\hat{u}^A_{\lambda,n} - u\|_2^2\right]^{1/2} \leq c(\log n/n)^{\gamma}.$$

To conclude this section, let us say that we constructed linear and nonlinear wavelet estimators of the solution to Wicksell's problem, based on compactly supported wavelets, that enjoy good asymptotic rates for $v \in H^s(\mathbb{R})$. Moreover, we have proved that the nonlinear estimators are globally adaptive, i.e. the choice of smoothing parameters is independent of the unknown parameter s.

4 Simulations and a real example

In order to investigate the sampling properties of our estimator, we have conducted a small Monte-Carlo simulation study.

4.1 Preliminary algorithmic considerations

The main algorithmic advantage of using a multiresolution analysis lies in the recursive relations between the wavelet $\psi_{j,k}$ and the scaling function $\phi_{j,k}$ which provide a fast algorithm to compute the wavelet coefficients. We shall prove that such a relation is still true under our setting.

It is well known (see Daubechies (1992)) that ϕ and ψ verify the following scaling relations:

$$\begin{aligned}\phi(x) &= \sum_\ell h_\ell \phi(2x-\ell)\\ \psi(x) &= \sum_\ell g_\ell \phi(2x-\ell)\end{aligned}$$

where h and g are the impulse response values of quadratic mirror filters and ℓ runs over a finite set. These relations are equivalent to

$$\begin{aligned}\hat{\phi}(\xi) &= \frac{1}{2}H(\frac{\xi}{2})\hat{\phi}(\frac{\xi}{2})\\ \hat{\psi}(\xi) &= \frac{1}{2}G(\frac{\xi}{2})\hat{\psi}(\frac{\xi}{2})\end{aligned}$$

where H and G are the discrete-time Fourier transforms of the sequences h and g:

$$\begin{aligned}H(\xi) &= \sum_\ell e^{i\ell\xi} h_\ell\\ G(\xi) &= \sum_\ell e^{i\ell\xi} g_\ell.\end{aligned}$$

Using the above relations, it is relatively easy to show that

$$\hat{\alpha}_{m,k} = \sqrt{2}\sum_\ell h_{\ell-2k}\hat{\alpha}_{m+1,k} \tag{18}$$

$$\hat{\beta}_{m,k} = \sqrt{2}\sum_\ell g_{\ell-2k}\hat{\alpha}_{m+1,k}. \tag{19}$$

Thanks to the above scaling equation, we only need to estimate the scaling coefficient $\hat{\alpha}_{m,k}$ at the finest possible scale m. The estimators of the other scaling coefficients and of the wavelet coefficients can then be computed via Mallat's fast pyramidal algorithm.
Our fast algorithm will therefore depend on an efficient evaluation of $\hat{\alpha}_{m,k}$. Recall here that, by using the properties of the Fourier transform, an alternative expression of $\hat{\alpha}_{m,k}$ is given by:

$$\hat{\alpha}_{m,k} = 2^{-m/2}\frac{1}{2\pi}\int e^{-i2^{-m}k\xi}\hat{\phi}(-2^{-m}\xi)(i\xi)^{1/2}\hat{\Phi}_n(\xi)d\xi. \tag{20}$$

A very natural, simple and convenient idea is to use the Fast Fourier Transform (FFT) for the numerical evaluation of (20). However, note that the support of the integrand in

(20) is infinite, which certainly is not adapted to numerical evaluation (see e.g. Gasquet and Witomski (1995)). To do so we will therefore apply the mollification method of Hào et al. Hào et al. (1994). More precisely, let V_ν be the de la Vallée-Poussin kernel defined by

$$V_\nu(x) = \frac{1}{\nu}\frac{\cos(\nu x) - \cos(2\nu x)}{x^2}, \quad \nu > 0. \tag{21}$$

The function V_ν is an entire function of exponential type of degree 2ν (see Hào et al. (1994)), absolutely integrable and bounded and satisfies the following property:

$$\hat{V}_\nu(\xi) = \begin{cases} \pi & \text{if } |\xi| \le \nu, \\ \frac{2\nu - |\xi|}{\nu} & \text{if } \nu < |\xi| \le 2\nu, \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

Figure 4.1 diplays the graph of V_ν and its Fourier transform $\hat{V}_\nu$ for $\nu = 2$. Note that each function V_ν has a Fourier transform that is compactly supported within the interval $[-2\nu, 2\nu]$.

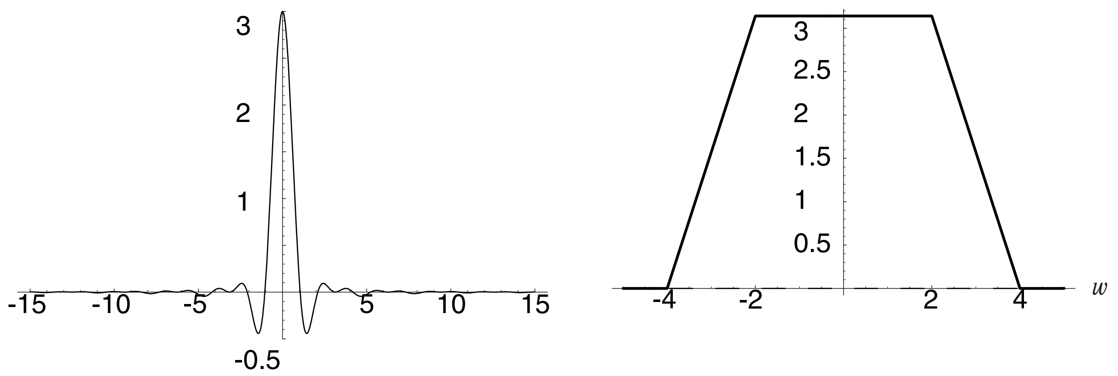

Figure 4.1: The graphs of V_ν and its Fourier transform $\hat{V}_\nu$ for $\nu = 2$.

Now, for a function $f \in L^p(\mathbb{R})$, $p \ge 1$, the mollification of f by V_ν is defined by

$$(M_\nu f)(x) = (V_\nu * f)(x) = \int V_\nu(x-t) f(t) dt.$$

Instead of calculating the fractional derivative of the scaling function ϕ, hidden in the expression (20) of $\hat{\alpha}_{m,k}$, we will use the fractional derivative of $M_\nu \phi$. It is known (see Hào et al. (1994)) that, if $\nu = 2^{\alpha m/r}$, where r is the regularity of ϕ and α is an arbitrary number in $]0,1[$, $\|D^{1/2}\phi - D^{1/2}M_\nu\phi\|_2 \to 0$, as m goes to infinity. The above suggests therefore the following method to compute $\hat{\alpha}_{m,k}$. Use the FFT to evaluate the factor

$$2^{-m/2}\frac{1}{2\pi}e^{-i2^{-m}k\xi}\hat{V}_\nu(-\xi)\hat{\phi}(-2^{-m}\xi), \tag{23}$$

in (20), since the support of $\hat{V}_\nu$ is compact and derive, by standard numerical integration, an approximation to $\hat{\alpha}_{m,k}$ given by

$$\tilde{\hat{\alpha}}_{m,k} = 2^{-m/2}\frac{1}{2\pi}\int e^{-i2^{-m}k\xi}\hat{V}_\nu(\xi)\hat{\phi}(-2^{-m}\xi)(i\xi)^{1/2}\hat{\Phi}_n(\xi)d\xi. \tag{24}$$

Now, note that expression (23) is the Fourier transform of $V_\nu * \phi_{m,k}$. If therefore one uses Coiflets (see e.g. Daubechies (1992)), it is easy to show that the following approximation holds:

$$\langle V_\nu(x-\cdot), \phi_{m,k}(\cdot)\rangle_2 \simeq 2^{-m/2}V_\nu\left(x - \frac{k}{2^m}\right).$$

In order to check numerically this approximation, we have approximated the scaling coefficients of $V_\nu(x-\cdot)$, $\nu = 2$ onto the approximation space at level 7, generated by coiflets of order 4, and, by inverting the discrete wavelet transform, obtained an approximation to $V_\nu(x-\cdot)$. Figure 4.2 displays the function $V_\nu(4-\cdot)$ and its Coiflet4 approximation (slightly shifted, to allow the visual comparison).

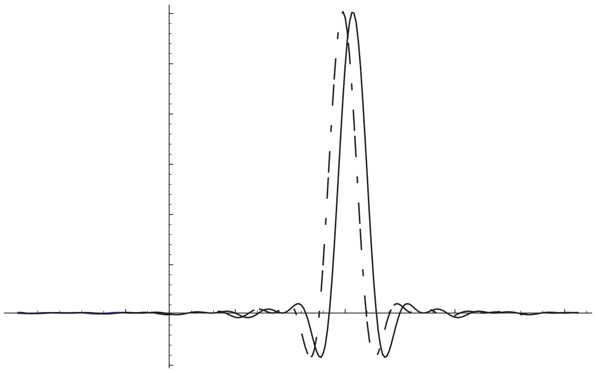

Figure 4.2: The graphs of $V_\nu(4-\cdot)$ and its Coiflet4 transform approximation, slightly shifted to the right.

To obtain the factors (23), all that remains to be done is to compute the discrete FFT of the functions $V_\nu(x-k/2^m)$, for x in a sufficiently large grid and for all $-A \le k \le 2^mT$, where $-A$ is the lower bound of the Coiflet support and T is the upper bound of the squared radii.

This is the algorithm we have used for the simulations and the example that follow. Note that the above calculations are to be done only once for a given support T and a given resolution m.

4.2 Monte Carlo experiments

The geometry of the Monte Carlo experiment is exactly that described in the introduction of this paper. We consider a block medium of dimensions $[0, \ell] \times [0, \ell] \times [0, h]$, $\ell = 750$, $h = 238$. A pseudo-random number of spheres was generated according to a Poisson distribution with mean $\lambda_V = 800$. If the number of spheres is n_3, then n_3 centers are uniformly distributed throughout the entire block. For each center a random radius was generated according to a theoretical density f_3. A parallel slice, at height $hs = 120$, was made through the medium, and the radii of all (two-dimensional) intersections were recorded. The parameters above were chosen in such a way that approximately one expects to observe about 270 such intersections.

We present the results of two Monte Carlo experiments. In each experiment, four replicates were performed. A replicate consists of drawing a sample of spheres, slicing the block and recording the observed cross-sectional radii. In order to compare our method with other available methods, we have first estimated the density of the squared radii, and then converted it by a change of variables to an estimate of the radii themselves. For each replication we have graphically compared our estimation with the estimates produced by a variant of Saltykov's method (see Taylor (1983)) and the spline method of Nychka et al. (1984). In experiment 1, the theoretical density f_3 was taken a Gamma distribution with shape parameter 2 and scale parameter 1/14. In experiment 2, the theoretical density f_3 was taken as a mixture of a Gamma distribution with shape parameter 2.2 and scale parameter 1/14, and a Normal distribution of mean the mean of the density of the first experiment shifted to the right by 40 and of standard deviation 8. The relative proportions in the mixture were 0.7 and 0.3 respectively.

Figure 4.3 below reproduces the results of the first experiment. On each panel, the true density f_3 is displayed (solid curve) along with the spline cross-validated estimation (dashed-dot), Saltykov's estimation (step function) and our linear wavelet estimator using coiflets of order 4 and a projection onto the approximation space of order 2, following the theoretical developments of the previous sections. Since the simulated densities are smooth, and given the obtained results we have not used a nonlinear wavelet estimate here. At each panel of the figure we also have reproduced the cross-validation criterion used for choosing the right amount of smoothing for splines.

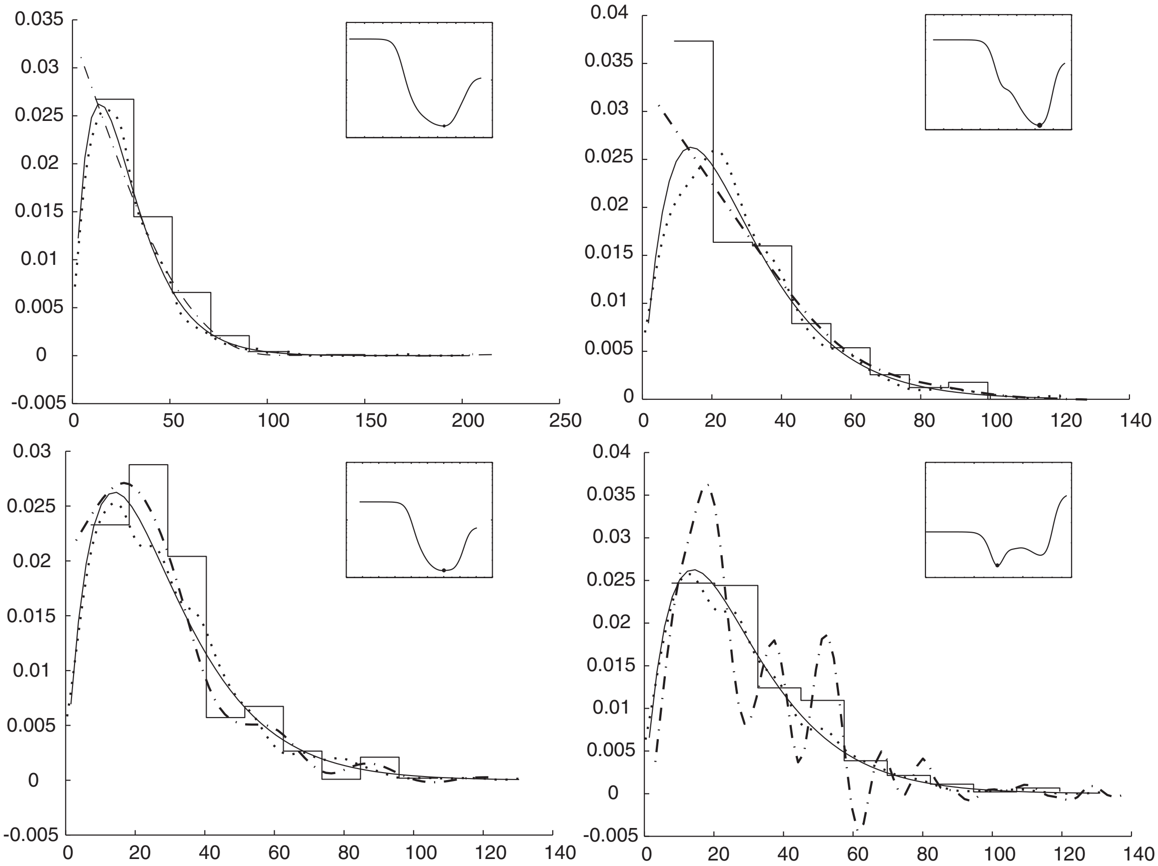

Figure 4.3: Experiment 1: The true density (solid curve); The spline cross-validated estimation (dashed-dot); Saltykov's estimation (step function); Linear wavelet estimator. Each top-right panel displays the cross-validation criterion used for choosing the right amount of smoothing for splines.

Overall the agreement between the true density and the wavelet estimate is good. The spline estimate is smoother, although its behaviour is not as good as expected near the origin, due to the way smoothness is defined by the spline method (extrapolation at the boundary is mainly linear).

In experiment 2, we wished to see the behaviour of the wavelet estimate when the data is multimodal with a more pronounced peak. The results of this experiment are displayed in Figure 4.4.

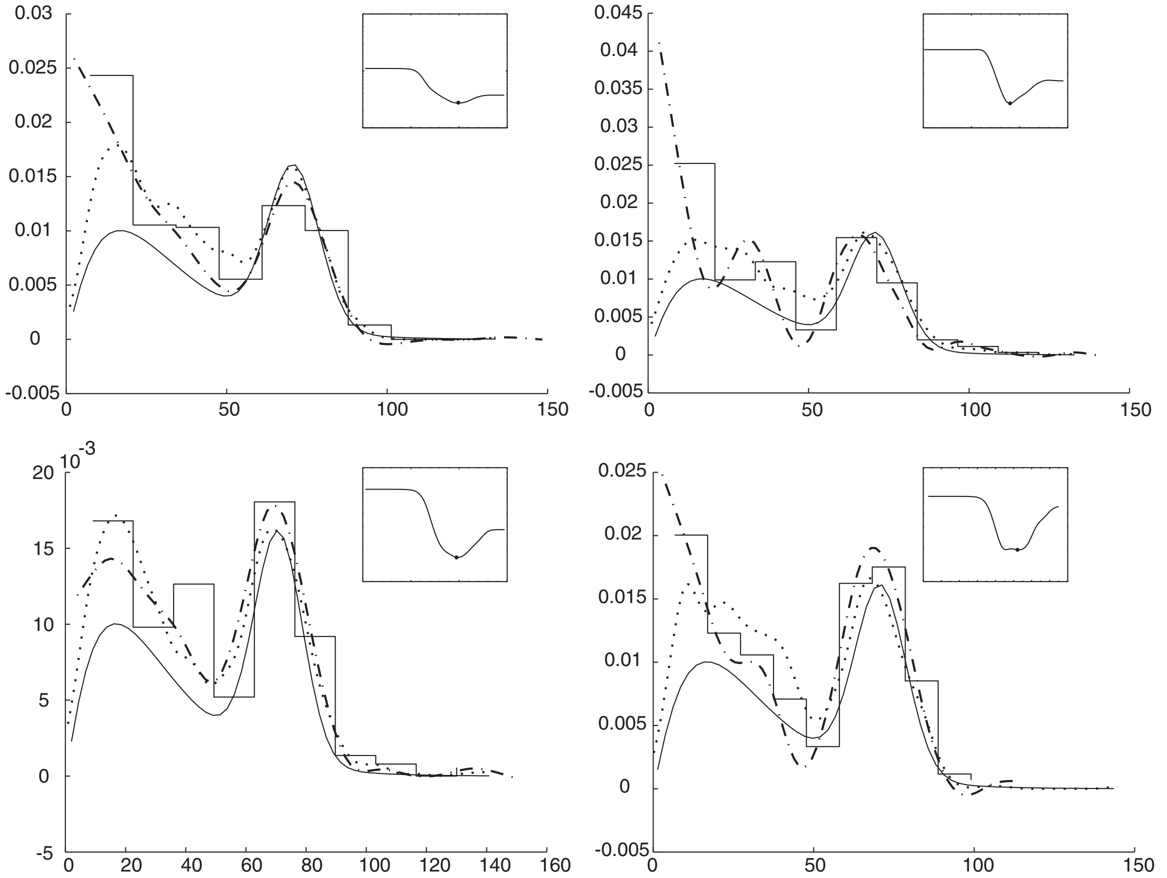

Figure 4.4: Experiment 2: The true mixture density (solid curve); The spline cross-validated estimation (dashed-dot); Saltykov's estimation (step function); Linear wavelet estimator. Each top-right panel displays the cross-validation criterion used for choosing the right amount of smoothing for splines.

In this experiment one can see that most replicates overestimate f_3 near its first mode, whatever method is used. On the other mode (far right), the overall behaviour of the wavelet estimator is quite good, particularly in comparison with the other methods. Note the ability of the estimate to resolve distinct peaks which is one of the reasons one may wish to use wavelets.

We conclude this simulation study by saying that the wavelet approach to the Wicksell problem appears to be successful, although its behaviour is not as good as might be hoped for mixtures with peaks of different magnitudes for small sample sizes.

4.3 A real example

We would like to apply our method to some experimental data obtained from a series of polymer blends which were kindly supplied to us for this purpose by J. Honerkamp and Tobias Roth from the Albert Ludwig University of Freiburg, Germany. The blends consist of 92.5 % polymethylmethacrylate and 7.5 % polystyrene with the former component being the matrix and the latter one building the inclusions (see, e.g. Weese et al. (1997)). In order to characterize the morphology of the blends, thin slices of 60 nm thickness were prepared and transmission electron microscopy images were taken. The radii of the profiles in the images were determined using an image processing software. The data set B075 at our disposal contains 4104 observed sections. We have applied our linear wavelet estimation procedure with coiflets of order 4 and a resolution level 3.

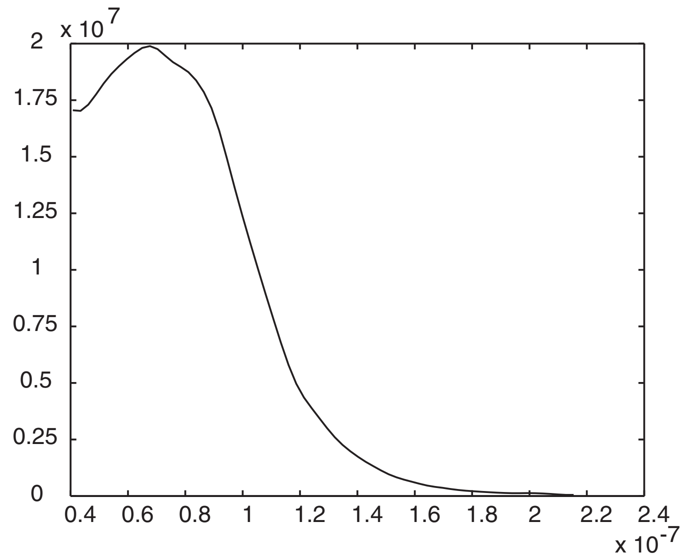

Figure 4.5: Linear wavelet estimator of the profile density for the blend B075.

Figure 4.5 displays the estimated density which is not far from the one obtained by Weese et al. Weese et al. (1997) on the same data set using some kernel methods. Note the lower bound of the support for the estimated density. This is due to the fact that profiles with radii below this lower bound could not be measured. To avoid a boundary bias for the wavelet method we have assumed that this lower bound was also the lower bound of the true density which is certainly not true here. The wavelet estimate is computationally fast and the results show that our approach may be a valuable tool for analyzing such problems.

Acknowledgments

This research was supported by 'Projet d'Actions de Recherche Concertées', No. 98/03–217, and 'Projet IDOPT, INRIA'. The first author would like to thank the Institute of Pure and Applied Mathematics and the Institute of Statistics of the Université Catholique de Louvain, for the financial support and their hospitality. We would also like to thank J. Honerkamp and T. Roth for making the transmission electron microscopy data available.

5 Appendix

The proofs of all major statements are based on several auxiliary results. Hereafter C will denote a generic constant which may change from one inequality to another.

Recall that the estimates of the scaling and wavelet coefficients are respectively defined via the transforms

$$\left(\mathcal{K}_m^- \phi\right)_{m,k}(X_\ell), \quad \text{and} \quad \left(\mathcal{K}_j^- \psi\right)_{j,k}(X_\ell),$$

where the operator $\mathcal{K}_m^-$ has been defined by expression (12). Now since both ϕ and ψ are supposed to be r-regular with $r > 3/2$, it is easy to see by integration by parts that for all integer $0 \leq \nu \leq r$, we have:

$$|\hat{\phi}^{(\nu)}(\xi)| \leq C(1+|\xi|)^{-r} \quad \text{and} \quad |\hat{\psi}^{(\nu)}(\xi)| \leq C(1+|\xi|)^{-r}. \tag{25}$$

It therefore follows that for all integer j, we have:

$$|\left(\mathcal{K}_j^- \phi\right)(x)| \leq C 2^{j/2} \quad \text{and} \quad |\left(\mathcal{K}_j^- \psi\right)(x)| \leq C 2^{j/2}. \tag{26}$$

Indeed,

$$\begin{aligned} |\left(\mathcal{K}_j^- \phi\right)(x)| &\leq |\frac{1}{2\pi} \int e^{ix\xi} \hat{\phi}(-\xi)(i2^j\xi)^{1/2} d\xi| \\ &\leq C' 2^{j/2} \int (1+|\xi|)^{-r+1/2} d\xi \leq C 2^{j/2}, \end{aligned}$$

since $r > 3/2$ and the same argument can be applied for $\mathcal{K}_j^- \psi$. Note also that the following inequality holds, whenever ψ is a wavelet of regularity $r \geq 2$:

$$|\left(\mathcal{K}_j^- \psi\right)(x)| \leq C 2^{j/2} (1+|x|)^{-2}. \tag{27}$$

When $|x| \leq 1$, by inequality (26), inequality (27), is obviously true with another generic constant C. Suppose therefore that $|x| > 1$. Let $\eta(\xi) = \hat{\psi}(-\xi)(i2^j\xi)^{1/2}$. Since ψ has regularity $r \geq 2$, the function η is at least twice differentiable. By integration by parts, it follows that

$$|\left(\mathcal{K}_j^- \psi\right)(x)| \leq C|x|^{-2} \int |\eta''(\xi)| d\xi.$$

Now, by a standard computation, we have

$$\int |\eta''(\xi)| d\xi \leq 2^{j/2} \int |\xi|^{1/2} (|\hat{\psi}''(-\xi)| + |\xi|^{-1} |\hat{\psi}'(\xi)| + |\xi|^{-2} |\hat{\psi}(\xi)|) d\xi,$$

which, by our assumptions on ψ is finite. Inequality (27) therefore follows.

We are now ready to prove Theorem 3.1. Note first that since $v \in D_{\sigma p2}(M,T)$, we have $\|v\|_\infty \leq C$ from which it follows, using Theorem 6.3.5 of Gorenflo and Vessela (1980), that $\|u\|_\infty \leq C$.

Proof of Theorem 3.1: We have already noticed that since both u and v are compactly supported, when the scaling function has a bounded support, the number of indices k for which $\alpha_{m,k} \neq 0$ is less than $C2^m$ for some constant C. Recall also that $\hat{\alpha}_{m,k}$ are unbiased estimators of the corresponding $\alpha_{m,k}$. We therefore have, using inequality (26):

$$\begin{aligned} \mathbb{E}\left((\hat{\alpha}_{m,k} - \alpha_{m,k})^2\right) &= \frac{1}{n} \operatorname{Var}\left((\mathcal{K}_m^- \phi)_{m,k}(X)\right) \\ &\leq \frac{1}{n} \int \left((\mathcal{K}_m^- \phi)_{m,k}(x)\right)^2 v(x) dx \leq C \frac{2^m}{n} \int v(x) dx = C \frac{2^m}{n}. \end{aligned} \tag{28}$$

Let P_{V_m} denote the projection onto the approximation space V_m of the multiresolution decomposition of $L^2(\mathbb{R})$ defined by ϕ. From inequality (28) and the fact that the number of indices k for which $\alpha_{m,k} \neq 0$ is of the order 2^m, it follows that

$$\mathbb{E}\left(\|\hat{u}_{m,n} - P_{V_m} u\|_2^2\right) \leq C \frac{2^{2m}}{n}. \tag{29}$$

It is also know (see e.g. Mallat (1989)), that since $u \in H^{s-1/2}(\mathbb{R})$, we have:

$$\|u - P_{V_m} u\|_2^2 \leq C \|u\|_{s-1/2}^2 2^{-2m(s-1/2)}. \tag{30}$$

Combining (29) and (30), leads to

$$\mathbb{E}\left(\|\hat{u}_{m,n} - u\|_2^2\right) \leq C \frac{2^{2m}}{n} + C' 2^{-2m(s-1/2)}. \tag{31}$$

Choosing now m such that $2^m = \mathcal{O}(n^{1/(2s+1)})$ completes the proof. ■

Proof of Theorem 3.2: The proof, once properly adapted, is very similar to the proof of Theorem 7 Fan and Koo (1999). For convenience of the reader we outline below the key ideas.

One may apply Rosenthal's inequality (see Rosenthal (1972)) to the sequence of i.i.d. random variables

$$Z_\ell = (\mathcal{K}_j^- \psi)_{j,k}(Y_\ell) - \beta_{j,k}, \quad \ell = 1, \dots, n,$$

since $\hat{\beta}_{j,k}$ is an unbiased estimator of $\beta_{j,k}$ and since, by inequality (26), one has, for $a \geq 2$:

$$\mathbb{E}\{|Z_\ell|^a\} \leq C 2^{(a-1)j}.$$

It follows then that, for $a \geq 2$:

$$\mathbb{E}\left(|\hat{\beta}_{j,k} - \beta_{j,k}|^a\right) \leq C \left\{ \frac{2^j}{n^{a/2}} + \frac{2^{(a-1)j}}{n^{a-1}} \right\}. \tag{32}$$

By inequality (26), one can also easily show that $\operatorname{Var} Z \leq M 2^j$ and that $\| \left(\mathcal{K}_j^- \psi\right)_{j,k} \|_\infty \leq M 2^j$. Using now Bernstein's inequality (see e.g. Pollard (1984), p. 192), for any $\epsilon \geq 1$, there exists $\tau = c(M)\epsilon$, such that, if $j 2^j \leq n$, we have

$$\mathbb{P}\left(|\hat{\beta}_{j,k} - \beta_{j,k}| > 2^{-1}\tau\sqrt{j 2^j / n}\right) \leq 2^{-\epsilon j}. \tag{33}$$

Using now (32) and (33), the proof is completed exactly in the same way as in Fan and Koo, by writing

$$\mathbb{E}(\|\hat{u}_{\lambda,n} - u\|_2^2) \leq 3\left(\mathbb{E}(\|\hat{u}_{j_0,n} - P_{V_{J_0}} u\|_2^2) + \mathbb{E}(\|\hat{Q}_{j_0,j_1} - Q_{j_0,j_1} u\|_2^2) + \|u - P_{V_{J_1}} u\|_2^2\right),$$

where

$$Q_{j_0,j_1} u = \sum_{j_0}^{j_1} \sum_k \beta_{j,k} \psi_{j,k},$$

and

$$\hat{Q}_{j_0,j_1} u = \sum_{j_0}^{j_1} \sum_k \hat{\beta}_{j,k} \psi_{j,k}.$$

■

The proof of Theorem 3.3 follows exactly the same lines as the proof of Theorem 8 of Fan and Koo and is hence is omitted.

References

Abramovich, F. and B. Silverman, 1998: Wavelet decomposition approaches to statistical inverse problems. *Biometrika*, **85**, 115–129.

Carroll, R. J. and P. Hall, 1988: Optimal rates of convergence for deconvolving a density. *J. Amer. Statist. Assoc.*, **83**, 1184–1186.

Chui, C. K., 1992: *An introduction to Wavelets.* Academic Press, New York.

Daubechies, I., 1992: *Ten Lectures on Wavelets*, volume 91 of *CBMS-NSF Regional Conference series in Applied Mathematics.* SIAM, Philadelphia.

Donoho, D. L., 1995: Nonlinear solution of linear inverse problems by wavelet-vaguelette decomposition. *Appl. Comput. Harmonic Anal.*, **2**, 101–126.

Donoho, D. L. and I. M. Johnstone, 1995: Adapting to unknown smoothness via wavelet shrinkage. *J. Amer. Statist. Assoc.*, **90**, 1200–1224.

Donoho, D. L., I. M. Johnstone, G. Kerkyacharian, and D. Picard, 1996: Density estimation by wavelet thresholding. *Ann. Statist.*, **24**, 508–539.

Fan, J., 1993: Adaptively local one-dimensional subproblems with application to a deconvolution problem. *Ann. Statist.*, **21**, 600–610.

Fan, J. and J.-K. Koo, 1999: Wavelet deconvolution. Technical report, University of California at Los Angeles.

Feuerverger, A. and P. Hall, 2000: Methods for density estimation in thick-slice versions of wicksell's problem. *J. Amer. Statist. Assoc.*, **95**, 535–546.

Gasquet, G. and P. Witomski, 1995: *Analyse de Fourier et Applications: Filtrage, Calcul numérique, Ondelettes.* Masson, Paris, Milan, Barcelone.

Golubev, G. K. and B. Y. Levit, 1998: Asymptotically efficient estimation in the wicksell problem. *Ann. Stat.*, **26**, 2407–2419.

Gorenflo, R. and S. Vessela, 1980: *Abel Integral Equations: Analysis and applications*, volume 1461 of *Lecture Notes in Mathematics.* Springer-Verlag, New York, Berlin.

Hall, P. and R. Smith, 1988: The kernel method for unfolding sphere size distributions. *J. of Computat. Phys.*, **74**, 409–421.

Hào, D.-N., H.-J. Reinhardt, and F. Seiffarth, 1994: Stable numerical fractional differentiation by mollification. *Numer. Funct. Anal. and Optimiz.*, **15**, 635–659.

Hoogendoorn, A. W., 1992: Estimating the weight undersize distribution for the Wicksell problem. *Statist. Neerlandica*, **4**, 259–282.

Kolaczyk, E. D., 1996: A wavelet shrinkage approach to tomographic image reconstruction. *J. Amer. Statist. Assoc.*, **91**, 1079–1090.

Krumbein, W., 1935: Thin section mechanical analysis of inducted sediments. *Journal of Geology*, **43**, 482–496.

Mallat, S., 1989: A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**, 674–693.

Mase, S., 1995: Stereological estimation of particle size distributions. *Adv. Appl. Prob.*, **27**, 350–366.

Meyer, Y., 1990: *Ondelettes et opérateurs Vol. I and II.* Hermann et Cie, Paris.

Nychka, D., G. Wahba, S. Goldfarb, and T. Pugh, 1984: Cross-validated spline methods for the estimation of three-dimensional tumor size distributions from observations on two-dimensional cross sections. *J. Amer. Statist. Assoc.*, **79**, 832–846.

O'Sullivan, F., 1986: A statistical perspective on ill-posed problems. *Statist. Sci.*, **1**, 502–518.

Pensky, M. and B. Vidakovic, 1999: Adaptive wavelet estimator for nonparametric density deconvolution. *Ann. Statist.*, **27**, 2033–2053.

Pollard, D., 1984: *Convergence of Stochastic Processes.* Springer-Verlag, New York.

Ripley, B., 1981: *Spatial Statistics.* Wiley, New York.

Rosenthal, H. P., 1972: On the span in l_p of sequences of independent random variables. *Israel Math. J.*, **8**, 273–303.

Saltykov, S., 1967: *The determination of the size distribution of particles in an opaque material from measurement of the size distribution of their sections*. Lecture Notes in Stereology. Springer-Verlag, New York.

Samko, S. G., A. A. Kilbas, and O. I. Marichev, 1993: *Fractional Integrals and Derivatives: Theory and Applications*. Gordon and Breach, Switzerland.

Silverman, B. W., M. C. Jones, J. D. Wilson, and D. W. Nychka, 1990: A smoothed EM approach to indirect estimation problems with particular reference to stereology and emission tomography. *J. R. Statist. Soc., B*, **52**, 271–324.

Stoyan, D., W. S. Kendall, and J. Mecke, 1995: *Stochastic Geometry and its Applications*. Wiley, New York, 2nd edition.

Takahashi, R. and M. Sibuya, 1996: The maximum size of the planar sections of random spheres and its application to metallurgy. *Ann. Inst. Statist. Math.*, **48**, 127–144.

Taylor, C., 1983: A new method for unfolding sphere size distributions. *J. Microscopy*, **132**, 57–66.

Triebel, H., 1983: *Theory of function spaces II*, volume 78 of *Monographs in Mathematics*. Birkhauser Verlag, Basel.

Walter, G. G., 1994: *Wavelets and Other Orthogonal Systems with Applications*. CRC Press, Boca Raton.

Weese, J., E. Korat, D. Maier, and J. Honerkamp, 1997: Unfolding sphere size distributions with a density estimator based on tikhonov regularization. *J. of Comput. Physics*, **138**, 331–353.

Wicksell, S. D., 1925: The corpuscule problem. a mathematical study of a biometric problem. *Biometrika*, **17**, 84–99.