Journal of Multivariate Analysis 101 (2010) 1458–1470

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

# An adaptive wavelet shrinkage approach to the Spektor–Lord–Willis problem

Bogdan Ćmiel

*AGH University of Science and Technology, Faculty of Applied Mathematics, Al. Mickiewicza 30, 30-059 Cracow, Poland*

ARTICLE INFO

*Article history:*
Received 13 May 2009
Available online 29 January 2010

*AMS subject classifications:*
62G05
45Q05

*Keywords:*
Spektor–Lord–Willis problem
Inverse problem
Rate of convergence
Minimax risk
Adaptive estimator

ABSTRACT

The stereological problem of unfolding the sphere size distribution from linear sections is considered. A minimax estimator of the intensity function of a Poisson process that describes the problem is introduced and an adaptive estimator is constructed that achieves the optimal rate of convergence over Besov balls to within logarithmic factors. The construction of these estimators uses Wavelet–Vaguelette Decomposition (WVD) of the operator that defines our inverse problem.

© 2010 Elsevier Inc. All rights reserved.

## 1. Introduction

Let us suppose that spheres of random radii are randomly distributed in some opaque medium in $\mathbb{R}^3$. The centers of the spheres form a homogeneous Poisson process on $\mathbb{R}^3$, and the radii have a distribution $Q$ on $[0; 1]$, independent of the centers and absolutely continuous with respect to the Lebesgue measure with probability density $q$. Let $c$ denote the expected number of sphere centers per unit volume. As the sphere radii cannot be observed directly, we take a linear section through the medium and observe the line segments that are intersections of the line and the spheres. From that observation we want to estimate $f := cq$. Let $n$ be the "size of the experiment". We thus observe a Poisson process $G_g^n$ on $[0; 1]$ with intensity $ng$ and it can be shown (see [1]) that

$$g(u) = 2u \int_u^1 f(x)\mathrm{d}x =: (Gf)(u), \tag{1}$$

where $G : L^2([0; 1], \mathrm{d}x) \to L^2([0; 1], \mathrm{d}u)$. The problem of unfolding $f$ from linear sections is known in the literature as the Spektor–Lord–Willis (SLW) problem, and its degree of ill-posedness is higher than for the related and better-known Wicksell's problem of unfolding the sphere size distribution from planar sections. The exact inverse of the operator $G$ leads to the equation

$$f(x) = \frac{1}{2}\left[\frac{g(x)}{x^2} - \frac{g'(x)}{x}\right].$$

As noted in [1], inverse estimation of $f$ in $L^2([0; 1], \mathrm{d}x)$ roughly corresponds to direct estimation of the intensity $g$ in $L^2([0; 1], u^{-4}\mathrm{d}u)$ and of its derivative in $L^2([0; 1], u^{-2}\mathrm{d}u)$, which explains the statistical difficulty of the problem. Eq. (1)

*E-mail address:* cmielbog@mat.agh.edu.pl.

0047-259X/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2010.01.009

can be used as a model of some measurements in material sciences (see [2,3]) and in isotropic cases as a model of linear intercept measurements on polished metallographic sections (cf. sintering processes in [4] and methodological remarks in [5]). It can also be applied in metallurgy (see, e.g., [6]). The SLW problem was analysed in [1] where B-spline sieved quasi-maximum likelihood estimators were used. Construction of a spectral estimator that is, up to a logarithmic factor, asymptotically minimax over a Sobolev-type class of functions can be found in [7]. In this paper we fill a gap in the rates by exploiting the properties of the wavelets.

In many aspects, this paper is based technically on [8] where a two-dimensional problem of positron emission tomography (PET) was considered. However, the main results of that paper could not be directly applied to the SLW problem. Firstly, the Wavelet–Vaguelette Decomposition (WVD) of the operator $G$ had to be constructed. Secondly, a lower bound of the estimator risk had to be found. For the Radon operator in [8] it was possible to construct a class of the functions with the images separated from zero, which was essential in the construction of the lower bounds. For the operator $G$, however, $g(1) = 0$ for all $f$. This paper resolves that problem by using the Assouad's cube technique to obtain a lower bound for the risk. A direct application of the methods from [8] for upper bounding the risk was not possible neither, because of the shape of the function $\gamma$, which is discussed later. This problem is resolved by restricting the domain to radii larger than some positive minimal detection level $\varepsilon$ in which case exact minimaxity can be achieved. The influence of $\varepsilon$ on the upper bound for the risk is studied in detail, which enables a construction of an almost minimax estimator for a model with $\varepsilon = 0$. In that case, however, there is still, as in [7], a logarithmic gap between the lower and the upper bound for the risk.

In the next section we will use a Wavelet–Vaguelette Decomposition to construct an estimator of the function $f$. To circumvent some problems with the construction of the WVD, we change the dominating measure in the image space, which also changes the operator itself. Dividing both sides of (1) by $u^2$, one obtains

$$h(u) := \frac{g(u)}{u^2} = \frac{2}{u}\int_u^1 f(x)\mathrm{d}x =: (Kf)(u),$$

where $K : L^2([0;1], \mathrm{d}x) \to L^2([0;1], \mathrm{d}\mu)$, $\mathrm{d}\mu = u^2\mathrm{d}u$. Note that the operators $K$ and $G$ are compact Hilbert–Schmidt operators. Consequently, their inverses are not bounded and the unfolding problem is ill-posed in the Hadamard sense. In the following sections we will denote by $\langle\cdot,\cdot\rangle$ and $[\cdot,\cdot]$ the inner products in $L^2(\mathrm{d}x)$ and $L^2(\mathrm{d}\mu)$, respectively.

## 2. Wavelet–Vaguelette Decomposition

The Wavelet–Vaguelette Decomposition will be constructed along the lines described in [9], Sec. 5.2. For the construction of the WVD, $K$ will be considered as an operator from $L^2(\mathbb{R}, \mathrm{d}x)$ to $L^2(\mathbb{R}, \mathrm{d}\mu)$ and the domain $\mathbb{R}$ will be omitted in the notation. Let $\psi$ be a smooth mother wavelet that satisfies the conditions: supp $\psi = [0; N]$, $\int_{-\infty}^{\infty}\psi(x)\mathrm{d}x = 0$, $\int_{-\infty}^{\infty}\psi^2(x)\mathrm{d}x = 1$, $\psi \in C^2$, $\|\psi'\|_{L^2(\mathrm{d}x)} < \infty$ and $\|\psi^{(-1)}\|_{L^2(\mathrm{d}x)} < \infty$, where $\psi^{(-1)}(u) := \int_{-\infty}^{u}\psi(x)\mathrm{d}x$. The set of functions $\psi_{jk}(x) = 2^{j/2}\psi(2^jx - k)$, with $j$ and $k$ integers, is a complete, orthonormal system in $L^2(\mathrm{d}x)$. Define

$$\varepsilon_{jk}(u) := (K\psi_{jk})(u) = \frac{2}{u}\int_u^{\infty}\psi_{jk}(x)\mathrm{d}x = -\frac{2}{u}\psi_{jk}^{(-1)}(u),$$

and

$$\gamma_{jk}(u) := \frac{\psi'_{jk}(u)}{2u} = 2^{\frac{3}{2}j}\,\frac{\psi'(2^ju - k)}{2u}. \tag{2}$$

It is easy to see that for any $j$ and $k$, $\varepsilon_{jk} \in L^2(\mathrm{d}\mu)$ and $\gamma_{jk} \in L^2(\mathrm{d}\mu)$. Write

$$[\gamma_{jk}, \varepsilon_{j'k'}] = -\int_{-\infty}^{\infty}\left[\frac{\psi'_{jk}(u)}{2u}\,\frac{2}{u}\int_{-\infty}^{u}\psi_{j'k'}(x)\mathrm{d}x\right]u^2\mathrm{d}u.$$

Integrating by parts, we have

$$[\gamma_{jk}, \varepsilon_{j'k'}] = \int_{-\infty}^{\infty}\psi_{jk}(u)\psi_{j'k'}(u)\mathrm{d}u = \langle\psi_{jk}(u), \psi_{j'k'}(u)\rangle = \delta_{(jk),(j'k')},$$

where $\delta$ denotes the Kronecker delta. Let us observe that

$$\left\|2^{-j}\gamma_{jk}\right\|_{L^2(\mathrm{d}\mu)} = \frac{1}{4}\left\|\psi'\right\|_{L^2(\mathrm{d}u)} = Const$$

and

$$\left\|2^{j}\varepsilon_{jk}\right\|_{L^2(\mathrm{d}\mu)} = 4\left\|\psi^{(-1)}\right\|_{L^2(\mathrm{d}u)} = Const.$$

Now, with $u_{jk} := 2^{-j}\gamma_{jk}$ and $v_{jk} := 2^j\varepsilon_{jk}$ one has the WVD of the operator $K$ (see [9], sec. 1.5)

$$K\psi_{jk} = 2^{-j}v_{jk},$$
$$K^*u_{jk} = 2^{-j}\psi_{jk}.$$

The functions $u_{jk}$ and $v_{j'k'}$ are biorthogonal:

$$[u_{jk}, v_{j'k'}] = \delta_{(jk),(j'k')}.$$

For completeness, one can also prove (see [9], sec. 4) the near orthogonality relations

$$\left\| \sum_{jk} a_{jk} v_{jk} \right\|_{L^2(\mathrm{d}\mu)} = 4 \left\| \sum_{jk} a_{jk} 2^{\frac{j}{2}} \psi^{(-1)}(2^j u - k) \right\|_{L^2(\mathrm{d}u)} \asymp \|(a_{jk})\|_{l_2},$$

$$\left\| \sum_{jk} a_{jk} u_{jk} \right\|_{L^2(\mathrm{d}\mu)} = \frac{1}{4} \left\| \sum_{jk} a_{jk} 2^{\frac{j}{2}} \psi'(2^j u - k) \right\|_{L^2(\mathrm{d}u)} \asymp \|(a_{jk})\|_{l_2},$$

which are, however, not necessary for our purpose. Considering the WVD for the original operator $G$ in $L^2(\mathrm{d}x)$, one can see that the function $\gamma_{jk}$ for the operator $G$ is the same as for the operator $K$ but, unfortunately, for some $k \in \mathbb{Z}$, $\gamma_{jk}$ is not an $L^2(\mathrm{d}u)$ object in this case. The second point is that $\|\gamma_{jk}\|_{L^2(\mathrm{d}u)}$ depends on $k$, which results in problems with finding quasi-singular values of the operator $G$ depending only on resolution level $j$.

Let $\phi$ be a father wavelet that satisfies the conditions: supp $\phi = [0; N]$, $\int_{-\infty}^{\infty} \phi(x)\mathrm{d}x = 1$ and $\phi \in C^2$. The function $f$ has the following inhomogeneous wavelet expansion (cf. [10] Ch. 3.2):

$$f = \sum_{k\in\mathbb{Z}} \langle f, \phi_{j_1 k}\rangle \phi_{j_1 k} + \sum_{j=j_1}^{\infty} \sum_{k\in\mathbb{Z}} \langle f, \psi_{jk}\rangle \psi_{jk},$$

where $j_1$ is any fixed integer and $\phi_{jk} = 2^{j/2}\phi(2^j x - k)$. Let us define functionals $c_{jk}(\cdot)$ as

$$c_{jk}(g) = [\gamma_{jk}, g],$$

so that $\gamma_{jk}$ is the Riesz representer of the functional $c_{jk}$. Observe that

$$c_{jk}(Kf) = [\gamma_{jk}, Kf] = \langle K^*\gamma_{jk}, f\rangle = \langle \psi_{jk}, f\rangle. \tag{3}$$

Since $\phi_{j_1 k_1} \in L^2(\mathrm{d}x)$ for $k_1 \in \mathbb{Z}$, it has a homogeneous wavelet expansion

$$\phi_{j_1 k_1} = \sum_{j,k} a_{(j_1 k_1)jk} \psi_{jk}.$$

Define a linear functional by

$$b_{j_1 k_1}(\cdot) = \sum_{j,k} a_{(j_1 k_1)jk} c_{jk}(\cdot)$$

and observe that

$$b_{j_1 k_1}(Kf) = \sum_{j,k} a_{(j_1 k_1)jk} c_{jk}(Kf) = \sum_{j,k} a_{(j_1 k_1)jk} \langle \psi_{jk}, f\rangle = \langle \phi_{j_1 k_1}, f\rangle. \tag{4}$$

Now, using (3) and (4), we have the reproducing formula

$$f = \sum_{k\in\mathbb{Z}} b_{j_1 k}(Kf)\phi_{j_1 k} + \sum_{j=j_1}^{\infty} \sum_{k\in\mathbb{Z}} c_{jk}(Kf)\psi_{jk}.$$

As discussed in [9], p. 111, although $K\phi \notin L^2(\mathrm{d}\mu)$ the reproducing formula remains valid at least for functions $f$ with only finite number of nonzero terms in the inhomogeneous wavelet expansion. Let $\tilde{\gamma}_{j_1 k}$ be the Riesz representer of the functional $b_{j_1 k}$. Then

$$f = \sum_{k\in\mathbb{Z}} [Kf, \tilde{\gamma}_{j_1 k}]\phi_{j_1 k} + \sum_{j=j_1}^{\infty} \sum_{k\in\mathbb{Z}} [Kf, \gamma_{jk}]\psi_{jk}. \tag{5}$$

Because $K^*\tilde{\gamma}_{j_1 k} = \phi_{j_1 k}$ and $(K^*\tilde{\gamma}_{j_1 k})(x) = 2\int_{-\infty}^{x} u\tilde{\gamma}_{j_1 k}(u)\mathrm{d}u$ we have

$$\tilde{\gamma}_{j_1 k}(u) = \frac{\phi'_{j_1 k}(u)}{2u} = 2^{\frac{3}{2}j_1} \frac{\phi'(2^{j_1}u - k)}{2u}. \tag{6}$$

## 3. The minimax risk

Besov spaces admit a characterization in terms of wavelet coefficients. A function $f = \sum_{k\in\mathbb{Z}} \alpha_{j_1 k}\phi_{j_1 k} + \sum_{j=j_1}^{\infty} \sum_{k\in\mathbb{Z}} \beta_{jk}\psi_{jk}$ belongs to the Besov ball $B_{\sigma pq}(M)$ if and only if

$$\|\alpha_{j_1\cdot}\|_{l_p} + \left( \sum_{j=j_1}^{\infty} (2^{j(\sigma+1/2-1/p)} \|\beta_{j\cdot}\|_{l_p})^q \right)^{1/q} \leqslant M.$$

Usually one cannot observe spheres with radii $r < \varepsilon$. We call $\varepsilon$ the minimal detection radius, assume $\varepsilon \in (0; 1/2)$ and restrict the domain of the radii to $[\varepsilon; 1]$. Define

$$\mathcal{K}_j^\varepsilon(\phi) := \left\{ k \in \mathbb{Z} : \text{supp } \phi_{jk} \cap [\varepsilon; 1] \neq \emptyset \right\},$$
$$\mathcal{K}_j^\varepsilon(\psi) := \left\{ k \in \mathbb{Z} : \text{supp } \psi_{jk} \cap [\varepsilon; 1] \neq \emptyset \right\},$$
$$\mathcal{F}_{\sigma pq}(M, [\varepsilon; 1]) = \left\{ f \cdot \mathbb{1}_{[\varepsilon;1]} : f \in B_{\sigma pq}(M), f \cdot \mathbb{1}_{[\varepsilon;1]} \geqslant 0 \right\}.$$

Then, for $f \in \mathcal{F}_{\sigma pq}(M, [\varepsilon; 1])$, $f = \sum_{k \in \mathcal{K}_{j_1}^\varepsilon(\phi)} \alpha_{j_1 k}\phi_{j_1 k} + \sum_{j=j_1}^{\infty} \sum_{k \in \mathcal{K}_j^\varepsilon(\psi)} \beta_{jk}\psi_{jk}$ on the interval $[\varepsilon, 1]$. It is easy to see that $|\mathcal{K}_j^\varepsilon(\phi)| \leqslant A_N 2^j$ and $|\mathcal{K}_j^\varepsilon(\psi)| \leqslant A_N 2^j$, where $A_N$ is a constant independent of $j$. The following theorem is an immediate consequence of propositions proved in Sections 3.1 and 3.2.

**Theorem 1.** *Let $p \geqslant 1$. Assume that $\sigma > 3(1/p - 1/2)$ if $p < 4/3$ and $\sigma > 1/p$ if $p \geqslant 4/3$. Then*

$$\forall\, \varepsilon \in \left(0; \frac{1}{2}\right) \quad \inf_{\hat{f}_n} \sup_{f \in \mathcal{F}_{\sigma pq}(M,[\varepsilon;1])} \mathbf{E}_f \|\hat{f}_n - f\|^2_{L^2([\varepsilon;1])} \asymp n^{-\frac{2\sigma}{2\sigma+3}},$$

*where $\hat{f}_n$ denotes any estimator of the intensity function $f$.*

The theorem remains valid if $f \in g_0 + \mathcal{F}_{\sigma pq}(M, [\varepsilon; 1])$, where $g_0 \geqslant 0$ is a known and fixed function with support in $[\varepsilon; 1]$. Note that we estimate the function $f$ on a support separated from zero. This is forced by the shape of the functions $\gamma_{jk}$. If supp $\psi_{jk}$ contained zero as an interior point, then $\|\gamma_{jk}\|_\infty$ would be infinite (cf. (2)), which would cause problems with the upper bound for the risk. We avoid that by taking $j_1$ large enough. The functions $f$ in Theorem 1 may by quite irregular. The assumptions are satisfied by, e.g., the sample paths of the Brownian motion supported on $[\varepsilon; 1]$. It can be shown that they belong to $B_{1/2p\infty}$ for any $1 \leqslant p < \infty$ (see [10], Ch. 9). It should be remarked that, as in [8], for $p \in [1; 2)$ the minimax linear risk on $\mathcal{F}_{\sigma pq}(M, [\varepsilon; 1])$ is higher than the minimax risk for all estimators. It can be shown that

$$\sup_{f \in \mathcal{F}_{\sigma pq}(M,[\varepsilon;1])} \mathbf{E}_f \|\hat{f}_L - f\|^2_{L^2([\varepsilon;1])} \geqslant C n^{-\frac{2\sigma'}{2\sigma'+3}},$$

where $\sigma' = \sigma + 1/2 - 1/\min(2, p)$, and $\hat{f}_L$ denotes any linear estimator of the intensity function $f$.

### *3.1. The lower bound*

In this section we use the Assouad's cube technique to prove the following proposition.

**Proposition 1.** *Let $p \geqslant 1$, $\sigma > 1/p$, and $0 \leqslant \varepsilon < 1/2$. Then for any estimator $\hat{f}_n$ of function $f$, there exists some constant $C$ such that*

$$\sup_{f \in \mathcal{F}_{\sigma pq}(M,[\varepsilon;1])} \mathbf{E}_f \|\hat{f}_n - f\|^2_{L^2([\varepsilon;1])} \geqslant C n^{-\frac{2\sigma}{2\sigma+3}}.$$

**Proof.** Let

$$\mathcal{G}_j^\varepsilon(\psi) := \left\{ k \in \mathbb{Z} : \text{supp } \psi_{jk} \subset \left[\varepsilon, \frac{1}{2}\right] \right\},$$

and

$$\mathcal{G}_{\sigma pq}^\varepsilon(j) = \left\{ f_\omega \geqslant 0 : f_\omega = f_0 + \delta_j \sum_{k \in \mathcal{G}_j^\varepsilon(\psi)} \omega_k \psi_{jk} \right\},$$

where $\omega_k \in \{0, 1\}$, $f_0 \in \mathcal{F}_{\sigma pq}(M/2, [\varepsilon; 1])$, $\int_{1/2}^1 f_0(x)\mathrm{d}x = C_1 > 0$ and $\delta_j \leqslant \min\left\{2^{-j(\sigma+1/2)}M/2, 2^{-j/2}C_1/(2A_N N\|\psi\|_\infty)\right\}$. It is easy to see that, for any constant $C_1$, $\mathcal{G}_{\sigma pq}^\varepsilon(j) \subset \mathcal{F}_{\sigma pq}(M, [\varepsilon; 1])$. Denote as $\mathcal{L}(G_{Kf})$ the distribution of $G_{Kf}$—the Poisson process with the intensity function $Kf$ with respect to $\mu$. The Hellinger affinity between the distributions takes the form

$$\rho(\mathcal{L}(G_{Kf_\omega}), \mathcal{L}(G_{Kf_{\omega'}})) = \exp[-H^2(Kf_\omega, Kf_{\omega'})],$$

where $H^2(Kf_\omega, Kf_{\omega'}) = \frac{1}{2}\int_0^1 \left(\sqrt{Kf_\omega} - \sqrt{Kf_{\omega'}}\right)^2 \mathrm{d}\mu$ (cf. [11], Ch. 3.2).

Let $f_\omega, f_{\omega'} \in \mathcal{G}^{\varepsilon}_{\sigma pq}(j)$ and $\Delta(\omega, \omega') = 1$, where $\Delta(\cdot, \cdot)$ denotes the Hamming distance. Notice that, by the construction of $\mathcal{G}^{\varepsilon}_{\sigma pq}(j)$, the support of the function $(Kf_\omega - Kf_{\omega'})$ is contained in $[\varepsilon; 1/2]$ which will ease the evaluation of $H^2(Kf_\omega, Kf_{\omega'})$:

$$\begin{aligned}H^2(Kf_\omega, Kf_{\omega'}) &= \frac{1}{2}\int_0^1 \frac{u^2\,(Kf_\omega - Kf_{\omega'})^2}{\left(\sqrt{Kf_\omega} + \sqrt{Kf_{\omega'}}\right)^2}\mathrm{d}u \\ &= \delta_j^2 \int_\varepsilon^{1/2} \frac{u\left(\int_u^1 \psi_{jk}(x)\mathrm{d}x\right)^2}{\int_u^1 f_0(x)\mathrm{d}x}\left(\sqrt{\frac{\int_u^1 f_\omega(x)\mathrm{d}x}{\int_u^1 f_0(x)\mathrm{d}x}} + \sqrt{\frac{\int_u^1 f_{\omega'}(x)\mathrm{d}x}{\int_u^1 f_0(x)\mathrm{d}x}}\right)^{-2}\mathrm{d}u.\end{aligned}$$

Since $u \in [\varepsilon; 1/2]$, it is clear that $\int_u^1 f_0(x)\mathrm{d}x \geqslant C_1$ and

$$\begin{aligned}\frac{\int_u^1 f_\omega(x)\mathrm{d}x}{\int_u^1 f_0(x)\mathrm{d}x} &= 1 + \frac{\delta_j \int_u^1 \sum_{k\in\mathcal{G}_j^\varepsilon(\psi)} \omega_k \psi_{jk}(x)\mathrm{d}x}{\int_u^1 f_0(x)\mathrm{d}x} \geqslant 1 - \frac{\delta_j \sum_{k\in\mathcal{G}_j^\varepsilon(\psi)} \int_u^1 |\psi_{jk}(x)|\mathrm{d}x}{\int_u^1 f_0(x)\mathrm{d}x} \\ &\geqslant 1 - \delta_j A_N 2^j N 2^{-j/2}\|\psi\|_\infty / C_1 \geqslant \frac{1}{2}.\end{aligned}$$

We have

$$\begin{aligned}H^2(Kf_\omega, Kf_{\omega'}) &\leqslant \frac{1}{2}\,\delta_j^2 \int_{\mathrm{supp}\psi_{jk}} \frac{u\left(\int_u^1 \psi_{jk}(x)\mathrm{d}x\right)^2}{\int_u^1 f_0(x)\mathrm{d}x}\mathrm{d}u \\ &\leqslant C_2\delta_j^2 N 2^{-j}\left(\|\psi\|_\infty N 2^{-j/2}\right)^2 \leqslant C_3 2^{-j(2\sigma+3)}.\end{aligned}$$

Let $2^j \asymp n^{1/(2\sigma+3)}$. Then, using the Assouad Lemma (see [7]), we get the result. □

### 3.2. The upper bound

In this section we construct an estimator that achieves the optimal rate of convergence. The construction of that estimator and the evaluation of its risk are much the same as in [8] with two important differences: the dominating measure in the image space is not the Lebesgue measure and there is an additional parameter $\varepsilon$ which influences the support of the function $f$ and, consequently, the number of wavelets used for the estimation. Let $G_h^n$ denote the observed Poisson process with intensity function $nh$ with respect to $\mathrm{d}\mu$, and let $\nu_h^n$ denote the intensity measure of that process. With $n = 1$, we write $\nu_h$ rather than $\nu_h^1$. We consider the following estimator of $f$ on the interval $[\varepsilon; 1]$:

$$\hat{f}_n^\varepsilon((\lambda_j), j_1(\varepsilon), j_2(n)) = \sum_{k\in\mathcal{K}^\varepsilon_{j_1(\varepsilon)}(\phi)} \hat{\alpha}_{j_1(\varepsilon)k}\phi_{j_1(\varepsilon)k} + \sum_{j=j_1(\varepsilon)}^{j_2(n)} \sum_{k\in\mathcal{K}_j^\varepsilon(\psi)} \delta_S(\hat{\beta}_{jk}, \lambda_j)\psi_{jk} \tag{7}$$

where

$$\hat{\alpha}_{j_1(\varepsilon)k} = \frac{1}{n}\int_0^1 \tilde{\gamma}_{j_1(\varepsilon)k}\mathrm{d}G_h^n, \qquad \hat{\beta}_{jk} = \frac{1}{n}\int_0^1 \gamma_{jk}\mathrm{d}G_h^n, \tag{8}$$

and the nonnegative sequence $(\lambda_j)$ defines a soft-threshold rule:

$$\delta_S(\hat{\beta}_{jk}, \lambda_j) = \mathrm{sgn}(\hat{\beta}_{jk})(|\hat{\beta}_{jk}| - \lambda_j)_+.$$

**Proposition 2.** *Let $p \geqslant 1$ and $0 < \varepsilon < 1/2$. Assume that $\sigma > 3(1/p - 1/2)$ if $p < 4/3$ and $\sigma > 1/p$ if $p \geqslant 4/3$. Then for any $f \in \mathcal{F}_{\sigma pq}(M, [\varepsilon; 1])$, there exist some constant $C$ and sequences $(j_2(n))$ and $(\lambda_j)$ such that*

$$\mathbf{E}\|\hat{f}_n^\varepsilon((\lambda_j), j_1(\varepsilon), j_2(n)) - f\|^2_{L^2([\varepsilon;1])} \leqslant Cn^{-\frac{2\sigma}{2\sigma+3}}.$$

**Proof.** Let us assume that

$$\frac{\varepsilon}{4} \leqslant N 2^{-j_1(\varepsilon)} \leqslant \frac{\varepsilon}{2}. \tag{9}$$

It is easy to see that

$$\forall k \in \mathcal{K}^\varepsilon_{j_1(\varepsilon)}(\phi) \quad \mathrm{supp}\ \tilde{\gamma}_{j_1(\varepsilon)k} \subset \left[\frac{\varepsilon}{2}; 1 + \frac{\varepsilon}{2}\right]$$

and

$$\forall j \geqslant j_1(\varepsilon)\ \forall k \in \mathcal{K}_j^{\varepsilon}(\psi) \qquad \operatorname{supp} \gamma_{jk} \subset \left[\frac{\varepsilon}{2}; 1 + \frac{\varepsilon}{2}\right].$$

It is known (see [12], Ch. 3.2) that

$$\mathbf{E}\, \hat{\alpha}_{j_1(\varepsilon)k} = \frac{1}{n}\int_0^1 \tilde{\gamma}_{j_1(\varepsilon)k} \mathrm{d}\nu_h^n = b_{j_1(\varepsilon)k}(Kf) = \langle f, \phi_{j_1(\varepsilon)k}\rangle := \alpha_{j_1(\varepsilon)k},$$

$$\mathbf{E}\, \hat{\beta}_{jk} = \frac{1}{n}\int_0^1 \gamma_{jk} \mathrm{d}\nu_h^n = c_{jk}(Kf) = \langle f, \psi_{jk}\rangle := \beta_{jk},$$

$$\operatorname{Var} \hat{\beta}_{jk} = \frac{1}{n}\int_0^1 \gamma_{jk}^2 \mathrm{d}\nu_h := \frac{\sigma_{jk}^2}{n}.$$

With

$$l_{jk} := \gamma_{jk}/\sigma_{jk},$$

one has

$$\int_0^1 l_{jk}^2 \mathrm{d}\nu_h = 1. \tag{10}$$

Since $k \in \mathcal{K}_j^{\varepsilon}(\psi)$, using (2) we have

$$\|l_{jk}\|_\infty \leqslant \frac{2^{\frac{3}{2}j} C_1}{\varepsilon \sigma_{jk}}. \tag{11}$$

In order to evaluate the risk of the estimator (7), a Gaussian approximation will be constructed in the sequence space. Let

$$\hat{\eta}_{jk} = \beta_{jk} + n^{-\frac{1}{2}} \sigma_{jk} z_{jk}, \tag{12}$$

where $z_{jk}$ is a Gaussian variable specified below. We evaluate $\mathbf{E}[\hat{\beta}_{jk} - \hat{\eta}_{jk}]^2$ by considering two cases. First, let us assume that for some constant $C_2$

$$\sigma_{jk}^2 \geqslant \frac{1}{n} C_2 \varepsilon^{-2} 2^{3j} \log^3 n. \tag{13}$$

Then

$$\mathbf{E}[\hat{\beta}_{jk} - \hat{\eta}_{jk}]^2 = \frac{\sigma_{jk}^2}{n}\, \mathbf{E}\left[n^{-\frac{1}{2}} \int_0^1 l_{jk} d(G_h^n - \nu_h^n) - z_{jk}\right]^2.$$

Define $V_n = \int_0^1 l_{jk} d(G_h^n - \nu_h^n) := \int_0^1 l_{jk} \mathrm{d}\bar{G}_h^n$. We have

$$\mathbf{E}[\hat{\beta}_{jk} - \hat{\eta}_{jk}]^2 = \frac{\sigma_{jk}^2}{n}\, \mathbf{E}\left[n^{-\frac{1}{2}} V_n - z_{jk}\right]^2.$$

We now need the following lemma (see [8], Lemma V.1):

**Lemma 1.** *Suppose that $\nu_h([0; 1]) = 1$ and $\int_0^1 l^2 \mathrm{d}\nu_h = 1$. Let $\|l\|_\infty \leqslant L$ and $V_n = \int_0^1 l(\mathrm{d}G_h^n - \mathrm{d}\nu_h^n)$, where $G_h^n$ is a Poisson process with intensity measure $\nu_h^n = n\nu_h$. Then, there exist absolute constants $D_1$ and $D_2$ such that, whenever $L^2 n^{-1} \log^3 n \leqslant D_1$, there exists a random variable $Z \sim \mathcal{N}(0, 1)$ such that*

$$\mathbf{E}\left(n^{-\frac{1}{2}} V_n - Z\right)^2 \leqslant D_2 L^2 n^{-1}.$$

Lemma 1 and formulas (10), (11) and (13) prove the existence of $z_{jk} \sim \mathcal{N}(0, 1)$ such that

$$\mathbf{E}[\hat{\beta}_{jk} - \hat{\eta}_{jk}]^2 \leqslant C_3 \varepsilon^{-2} \frac{2^{3j}}{n^2}. \tag{14}$$

We still have to check the case when condition (13) is not valid, i.e. when

$$\sigma_{jk}^2 < \frac{1}{n} C_2 \varepsilon^{-2} 2^{3j} \log^3 n.$$

In that case, we take any $z_{jk} \sim \mathcal{N}(0, 1)$ and use the inequality $(a - b)^2 \leqslant 2a^2 + 2b^2$ to obtain

$$\mathbf{E}[\hat{\beta}_{jk} - \hat{\eta}_{jk}]^2 \leqslant 4\sigma_{jk}^2 n^{-1} < C_2 n^{-2} \varepsilon^{-2} 2^{3j} \log^3 n. \tag{15}$$

Formulas (14) and (15) show that for fixed $j$, $k$ there exists a Gaussian variable $z_{jk}$ such that

$$\mathbf{E}[\hat{\beta}_{jk} - \hat{\eta}_{jk}]^2 \leqslant C_4 n^{-2} \varepsilon^{-2} 2^{3j} \log^3 n. \tag{16}$$

We now evaluate the risk of the estimator (7) for $f \in \mathcal{F}_{\sigma pq}(M, [\varepsilon; 1])$:

$$\mathbf{E}\|\hat{f}_n^{\varepsilon}((\lambda_j), j_1(\varepsilon), j_2(n)) - f\|^2_{L^2([\varepsilon;1])} \leqslant \sum_{k \in \mathcal{K}^{\varepsilon}_{j_1(\varepsilon)}(\phi)} \mathbf{E}[\hat{\alpha}_{j_1(\varepsilon)k} - \alpha_{j_1(\varepsilon)k}]^2 + \sum_{j=j_1(\varepsilon)}^{j_2(n)} \sum_{k \in \mathcal{K}^{\varepsilon}_j(\psi)} \mathbf{E}[\delta_S(\hat{\beta}_{jk}, \lambda_j) - \beta_{jk}]^2 + \sum_{j=j_2(n)+1}^{\infty} \sum_{k \in \mathcal{K}^{\varepsilon}_j(\psi)} \beta^2_{jk} := L_n(f) + S_n(f) + T_n(f).$$

Using (6) and (9) we obtain

$$L_n(f) \leqslant C_5 \frac{1}{n} \sum_{k \in \mathcal{K}^{\varepsilon}_{j_1(\varepsilon)}(\phi)} \frac{2^{2j_1(\varepsilon)}}{\varepsilon^2} \leqslant C_6 n^{-1} \varepsilon^{-5}. \tag{17}$$

For the last term we have

$$T_n(f) \leqslant \sup\{T_n(f), f \in \mathcal{F}_{\sigma pq}(M, [\varepsilon; 1])\} = C_7 2^{-2j_2(n)\left(\sigma + \frac{1}{2} - \frac{1}{p}\right)}. \tag{18}$$

To evaluate $S_n(f)$ we will use the random variable $\hat{\eta}_{jk}$ defined in (12):

$$S_n(f) \leqslant \sum_{j=j_1(\varepsilon)}^{j_2(n)} \sum_{k \in \mathcal{K}^{\varepsilon}_j(\psi)} 2\mathbf{E}[\delta_S(\hat{\beta}_{jk}, \lambda_j) - \delta_S(\hat{\eta}_{jk}, \lambda_j)]^2 + \sum_{j=j_1(\varepsilon)}^{j_2(n)} \sum_{k \in \mathcal{K}^{\varepsilon}_j(\psi)} 2\mathbf{E}[\delta_S(\hat{\eta}_{jk}, \lambda_j) - \beta_{jk}]^2.$$

Since for every $\lambda$ the mapping $y \to \delta_S(y, \lambda)$ is a contraction, i.e. $|\delta_S(y_1, \lambda) - \delta_S(y_2, \lambda)| < |y_1 - y_2|$, we have

$$\begin{aligned} S_n(f) &\leqslant \sum_{j=j_1(\varepsilon)}^{j_2(n)} \sum_{k \in \mathcal{K}^{\varepsilon}_j(\psi)} 2\mathbf{E}[\hat{\beta}_{jk} - \hat{\eta}_{jk}]^2 + \sum_{j=j_1(\varepsilon)}^{j_2(n)} \sum_{k \in \mathcal{K}^{\varepsilon}_j(\psi)} 2\mathbf{E}[\delta_S(\hat{\eta}_{jk}, \lambda_j) - \beta_{jk}]^2 \\ &:= a_n(f) + b_n(f). \end{aligned}$$

Using (16), we obtain

$$a_n(f) \leqslant 4C_4 A_N \varepsilon^{-2} n^{-2} \log^3 n\, 2^{4j_2(n)}. \tag{19}$$

To evaluate $b_n(f)$ we will use the following lemma (see [13], Lemma 3).

**Lemma 2.** *If $\sigma_{jk} \leqslant 2^j \varepsilon^{-1} C_8$ for all $j$, $k$, then*

$$\mathbf{E}\left[\delta_S\left(\beta_{jk} + n^{-\frac{1}{2}} \sigma_{jk} z_{jk}, \lambda_j\right) - \beta_{jk}\right]^2 \leqslant 2\mathbf{E}\left[\delta_S\left(\beta_{jk} + n^{-\frac{1}{2}} 2^j \varepsilon^{-1} C_8 z_{jk}, \lambda_j\right) - \beta_{jk}\right]^2.$$

We have

$$b_n(f) \leqslant \sum_{j=j_1(\varepsilon)}^{j_2(n)} \sum_{k \in \mathcal{K}^{\varepsilon}_j(\psi)} 4\mathbf{E}\left[\delta_S\left(\beta_{jk} + n^{-\frac{1}{2}} 2^j C_8 \varepsilon^{-1} z_{jk}, \lambda_j\right) - \beta_{jk}\right]^2.$$

For an appropriate choice of $(\lambda_j)$ it can be shown (see [9] sec. 8) that

$$b_n(f) \leqslant C_9 n^{-\frac{2\sigma}{2\sigma+3}} \varepsilon^{-\frac{4\sigma}{2\sigma+3}} \leqslant C_9 n^{-\frac{2\sigma}{2\sigma+3}} \varepsilon^{-2}. \tag{20}$$

Using (17)–(20) we obtain

$$\begin{aligned} &\mathbf{E}\|\hat{f}_n^{\varepsilon}((\lambda_j), j_1(\varepsilon), j_2(n)) - f\|^2_{L^2([\varepsilon;1])} \\ &\quad \leqslant C_6 \varepsilon^{-5} n^{-1} + C_7 2^{-2j_2(n)\left(\sigma + \frac{1}{2} - \frac{1}{p}\right)} + C_{10} \varepsilon^{-2} n^{-2} \log^3 n\, 2^{4j_2(n)} + C_9 \varepsilon^{-2} n^{-\frac{2\sigma}{2\sigma+3}}. \end{aligned} \tag{21}$$

We can choose $j_2(n)$ such that

$$\frac{\sigma \log_2 n}{(2\sigma + 3)(\sigma + 1/2 - 1/p)} \ll j_2(n) \ll \frac{\sigma + 3}{2(2\sigma + 3)} \log_2 n - \frac{3}{4} \log_2 \log n,$$

where $a_n \ll b_n$ means that $\lim_{n\to\infty} b_n - a_n = \infty$. This is possible when

$$\frac{\sigma}{(2\sigma+3)(\sigma+1/2-1/p)} < \frac{\sigma+3}{2(2\sigma+3)}.$$

Since $p \geqslant 1$ it is easy to check that this condition is true if we assume that $\sigma > 1/p$. With that choice of $j_2(n)$ we have

$$C_7 2^{-2j_2(n)\left(\sigma+\frac{1}{2}-\frac{1}{p}\right)} < C_7 2^{\frac{-2\sigma}{2\sigma+3}\log_2 n} = C_7 n^{-\frac{2\sigma}{2\sigma+3}}$$

and

$$C_{10}\varepsilon^{-2} n^{-2} \log^3 n\, 2^{4j_2(n)} < C_{10}\varepsilon^{-2} n^{-\frac{2\sigma}{2\sigma+3}}.$$

Using this in (21) we finally obtain

$$\mathbf{E}\|\hat{f}_n^{\varepsilon}((\lambda_j), j_1(\varepsilon), j_2(n)) - f\|^2_{L^2([\varepsilon;1])} \leqslant C_{11}\left(\varepsilon^{-2} n^{-\frac{2\sigma}{2\sigma+3}} + \varepsilon^{-5} n^{-1}\right). \tag{22}$$

Since $\varepsilon$ is fixed, this completes the proof of Proposition 2. □

It should be remarked that, in some special cases, the estimator $\hat{f}_n^{\varepsilon}$ can achieve good rates of convergence also without the restriction of the domain to $[\varepsilon; 1]$. Let us evaluate

$$\mathbf{E}\|\hat{f}_n^{\varepsilon} - f\|^2_{L^2([0;1])} = \mathbf{E}\|\hat{f}_n^{\varepsilon} - f\|^2_{L^2([0;\varepsilon])} + \mathbf{E}\|\hat{f}_n^{\varepsilon} - f\|^2_{L^2([\varepsilon;1])}.$$

If $f \in \mathcal{F}_{\sigma pq}(M, [0; 1])$ then $f \cdot \mathbb{1}_{[\varepsilon;1]} \in \mathcal{F}_{\sigma pq}(M, [\varepsilon; 1])$ and using (22) we have

$$\mathbf{E}\|\hat{f}_n^{\varepsilon} - f\|^2_{L^2([0;1])} \leqslant C_{11}\left(\varepsilon^{-2} n^{-\frac{2\sigma}{2\sigma+3}} + \varepsilon^{-5} n^{-1}\right) + C_{12}\varepsilon \max_{x\in[0;\varepsilon]} f^2(x).$$

If we assume that, for some $\alpha > 0$,

$$\lim_{x\to 0^+} f(x) e^{\frac{1}{x^\alpha}} < \infty, \tag{23}$$

then

$$\mathbf{E}\|\hat{f}_n^{\varepsilon} - f\|^2_{L^2([0;1])} \leqslant C_{11}\left(\varepsilon^{-2} n^{-\frac{2\sigma}{2\sigma+3}} + \varepsilon^{-5} n^{-1}\right) + C_{13}\varepsilon e^{-\frac{2}{\varepsilon^\alpha}}.$$

Now, if we take $\varepsilon = (\log n)^{-\frac{1}{\alpha}}$ and denote the corresponding $\hat{f}_n^{\varepsilon}$ as $\hat{f}_n^{0}$, then

$$\mathbf{E}\|\hat{f}_n^{0} - f\|^2_{L^2([0;1])} \leqslant C_{14} n^{-\frac{2\sigma}{2\sigma+3}} (\log n)^{\frac{2}{\alpha}}.$$

Combining this with Proposition 1 we conclude that $\hat{f}_n^{0}$ achieves almost the optimal rate of convergence to within logarithmic terms.

## 4. The adaptive estimator

The form of the minimax estimator from the previous section depends on the parameters of the space $\mathcal{F}_{\sigma pq}$ that our intensity function belongs to (consider the choice of $j_2(n)$, for example). In practical estimation problems we do not know the parameters of the space $\mathcal{F}_{\sigma pq}$. Because of that, we need an estimator, with the best possible rate of convergence, that does not use the values of those parameters. We will call it an adaptive estimator. As in Section 3.2 we closely follow the derivation from [8], with modifications forced by the changed dominating measure and the restricted domain.

Fix an integer $r_0 > 3/2$ and suppose that the parameters $(\sigma, p, q)$ belong to the class

$$J = \left\{(\sigma, p, q) : \max\left\{\frac{1}{p}, 3\left(\frac{1}{p} - \frac{1}{2}\right)\right\} < \sigma < r_0,\ 1 \leqslant p, q \leqslant \infty\right\}.$$

Consider the following estimator:

$$\tilde{f}_n^{\varepsilon} = \sum_{k\in\mathcal{K}^{\varepsilon}_{j_3(n)}(\phi)} \hat{\alpha}_{j_3(n)k}\phi_{j_3(n)k} + \sum_{j=j_3(n)}^{j_4(n)} \sum_{k\in\mathcal{K}^{\varepsilon}_j(\psi)} \delta_H(\hat{\beta}_{jk}, T(\varepsilon)c_j)\psi_{jk}, \tag{24}$$

where the coefficients $\hat{\alpha}_{j_3(n)k}$ and $\hat{\beta}_{jk}$ are defined in (8),

$$c_j = 2^j\sqrt{\frac{j}{n}}, \qquad 2^{j_3(n)} \asymp n^{1/(2r_0+3)}, \qquad 2^{j_4(n)} \asymp n/\log_2 n$$

and

$$\delta_H(\hat{\beta}_{jk}, T(\varepsilon)c_j) = \begin{cases} \hat{\beta}_{jk}, & \text{if } |\hat{\beta}_{jk}| > T(\varepsilon)c_j \\ 0, & \text{if } |\hat{\beta}_{jk}| \leqslant T(\varepsilon)c_j \end{cases}$$

is a hard-threshold rule with a constant $T(\varepsilon)$. We will prove the following theorem:

**Theorem 2.** *Let $(\sigma, p, q) \in J$. Then*

$$\forall\, \varepsilon \in \left(0; \frac{1}{2}\right) \quad \sup_{f \in \mathcal{F}_{\sigma pq}(M,[\varepsilon;1])} \mathbf{E}_f \|\tilde{f}_n^{\varepsilon} - f\|^2_{L^2([\varepsilon;1])} \leqslant C \left(\frac{\log n}{n}\right)^{\frac{2\sigma}{2\sigma+3}}.$$

The theorem says that the estimator $\tilde{f}_n^{\varepsilon}$ achieves almost the optimal rate of convergence to within logarithmic terms. Since $\tilde{f}_n^{\varepsilon}$ does not depend on the parameters $\sigma$, $p$, $q$, it is an adaptive estimator.

**Proof of Theorem 2.** Recall that supp $\psi = [0; N]$ and assume that

$$N2^{-j_3(n)} \leqslant \frac{\varepsilon}{2}.$$

It is easy to see that

$$\forall k \in \mathcal{K}^{\varepsilon}_{j_3(n)}(\phi) \qquad \operatorname{supp} \tilde{\gamma}_{j_3(n)k} \subset \left[\frac{\varepsilon}{2}; 1 + \frac{\varepsilon}{2}\right]$$

and

$$\forall j \geqslant j_3(n)\ \forall k \in \mathcal{K}^{\varepsilon}_j(\psi) \qquad \operatorname{supp} \gamma_{jk} \subset \left[\frac{\varepsilon}{2}; 1 + \frac{\varepsilon}{2}\right].$$

We write some useful inequalities. Using (2) we have

$$\int_0^1 |\gamma_{jk}(u)|^m \mathrm{d}\mu = \int_0^1 2^{\frac{3}{2}jm} \left|\frac{\psi'(2^j u - k)}{2u}\right|^m \mathrm{d}\mu \leqslant C_1 \varepsilon^{-m} 2^{j(\frac{3}{2}m-1)}, \tag{25}$$

and from (25) we obtain

$$\mathbf{E}[\hat{\beta}_{jk} - \beta_{jk}]^2 = \frac{1}{n} \int_0^1 \gamma_{jk}^2 \mathrm{d}\nu_h \leqslant C_2 \varepsilon^{-2} \frac{2^{2j}}{n}. \tag{26}$$

From the equation (see [8] Lemma A.1)

$$\mathbf{E}\left(\int_0^1 \gamma_{jk} d(G_h^n - \nu_h^n)\right)^4 = \int_0^1 \gamma_{jk}^4 \mathrm{d}\nu_h^n + 3\left(\int_0^1 \gamma_{jk}^2 \mathrm{d}\nu_h^n\right)^2$$

and from (25) we have

$$\mathbf{E}[\hat{\beta}_{jk} - \beta_{jk}]^4 = \mathbf{E}\left[\frac{1}{n}\int_0^1 \gamma_{jk} d(G_h^n - \nu_h^n)\right]^4 \leqslant C_3 \varepsilon^{-4} \left(\frac{2^{5j}}{n^3} + \frac{2^{4j}}{n^2}\right). \tag{27}$$

We now use the following lemma (see [8] Lemma A.2):

**Lemma 3.** *If $\int_0^1 \gamma_{jk}^2 \mathrm{d}\nu_h^n \leqslant V$ and $\|\gamma_{jk}\|_\infty \leqslant H$ then*

$$P\left(\left|\int_0^1 \gamma_{jk} d(G_h^n - \nu_h^n)\right| \geqslant \lambda\right) \leqslant 2\exp\left[-\frac{1}{2}\frac{\lambda^2}{V + H\lambda/3}\right].$$

Since $\int_0^1 \gamma_{jk}^2 \mathrm{d}\nu_h^n \leqslant C_4 \varepsilon^{-2} 2^{2j} n$ and $\|\gamma_{jk}\|_\infty \leqslant C_4 \varepsilon^{-1} 2^{\frac{3}{2}j}$ we have

$$P\left(\left|\hat{\beta}_{jk} - \beta_{jk}\right| > \frac{T(\varepsilon)}{2} c_j\right) \leqslant 2\exp\left[-\frac{1}{8}\frac{\varepsilon T^2(\varepsilon) j 2^{2j}}{C_4 \varepsilon^{-1} 2^{2j} + C_4 2^{2j} T(\varepsilon)/6}\right].$$

Let $T(\varepsilon) = C_5 \eta(\varepsilon)$, where $C_5^2 \geqslant 8C_4(1 + C_5/6)\log 2$. Then,

$$P\left(\left|\hat{\beta}_{jk} - \beta_{jk}\right| > \frac{T(\varepsilon)}{2} c_j\right) \leqslant 2\exp\left[-\varepsilon\eta(\varepsilon) j \frac{1 + C_5/6}{1/(\varepsilon\eta(\varepsilon)) + C_5/6} \log 2\right].$$

If we now choose $\eta(\varepsilon) \geqslant \varepsilon^{-1}$, we obtain

$$\forall \eta(\varepsilon) \geqslant \varepsilon^{-1} \quad P\left(\left|\hat{\beta}_{jk} - \beta_{jk}\right| > \frac{T(\varepsilon)}{2} c_j\right) \leqslant 2^{-\varepsilon\eta(\varepsilon)j+1}. \tag{28}$$

Let us define

$$E_j f = \sum_{k \in \mathcal{K}_j^\varepsilon(\phi)} \alpha_{jk}\phi_{jk},$$

$$D_{j_3(n)j_4(n)} f = \sum_{j=j_3(n)}^{j_4(n)} D_j f = \sum_{j=j_3(n)}^{j_4(n)} \sum_{k \in \mathcal{K}_j^\varepsilon(\psi)} \beta_{jk}\psi_{jk},$$

$$\hat{E}_{j_3(n)} = \sum_{k \in \mathcal{K}_{j_3(n)}^\varepsilon(\phi)} \hat{\alpha}_{j_3(n)k}\phi_{j_3(n)k},$$

$$\hat{D}_{j_3(n)j_4(n)} = \sum_{j=j_3(n)}^{j_4(n)} \sum_{k \in \mathcal{K}_j^\varepsilon(\psi)} \delta_H(\hat{\beta}_{jk}, T(\varepsilon)c_j)\psi_{jk}.$$

Since $f = E_{j_3(n)}f + D_{j_3(n)j_4(n)}f + f - E_{j_4(n)}f$ and $\tilde{f}_n^\varepsilon = \hat{E}_{j_3(n)} + \hat{D}_{j_3(n)j_4(n)}$, we have

$$\mathbf{E}\|\tilde{f}_n^\varepsilon - f\|^2_{L^2([\varepsilon;1])} \leqslant 3\mathbf{E}\|\hat{E}_{j_3(n)} - E_{j_3(n)}f\|^2_{L^2([\varepsilon;1])} + 3\mathbf{E}\|\hat{D}_{j_3(n)j_4(n)} - D_{j_3(n)j_4(n)}f\|^2_{L^2([\varepsilon;1])} + 3\mathbf{E}\|f - E_{j_4(n)}f\|^2_{L^2([\varepsilon;1])}. \tag{29}$$

First let us evaluate $\mathbf{E}\|f - E_{j_4(n)}f\|^2_{L^2([\varepsilon;1])}$:

$$\|f - E_{j_4(n)}f\|^2_{L^2([\varepsilon;1])} = \left\|\sum_{j=j_4(n)}^{\infty} D_j f\right\|^2_{L^2([\varepsilon;1])} \leqslant \left(\sum_{j=j_4(n)}^{\infty} \|D_j f\|_{L^2([\varepsilon;1])}\right)^2.$$

Since $B_{\sigma pq} \subset B_{\sigma' 2\infty} (\sigma' = \sigma + 1/2 - 1/p)$, we have $\|f\|_{\sigma' 2\infty} \leqslant \|f\|_{\sigma pq}$, so

$$2^{j\sigma'}\|D_j f\|_{L^2([\varepsilon;1])} \leqslant \sup_{j \geqslant j_4(n)} 2^{j\sigma'}\|D_j f\|_{L^2([\varepsilon;1])} + \|E_{j_4(n)}f\|_{L^2([\varepsilon;1])} \leqslant \|f\|_{\sigma pq}.$$

From this we conclude that

$$\|f - E_{j_4(n)}f\|^2_{L^2([\varepsilon;1])} \leqslant C_6\|f\|^2_{\sigma pq} 2^{-2j_4(n)\sigma'} \leqslant C_7 n^{-\frac{2\sigma}{2\sigma+3}}. \tag{30}$$

Now we evaluate

$$\mathbf{E}\|\hat{E}_{j_3(n)} - E_{j_3(n)}f\|^2_{L^2([\varepsilon;1])} \leqslant \frac{C_8}{n} \sum_{k \in \mathcal{K}_{j_3(n)}^\varepsilon(\phi)} \int_0^1 \tilde{\gamma}^2_{j_3(n)k} \mathrm{d}\nu_h.$$

Using (6) we can show, exactly as in (25), that

$$\int_0^1 |\tilde{\gamma}_{jk}(u)|^m \mathrm{d}\mu \leqslant C_9 \varepsilon^{-m} 2^{j(\frac{3}{2}m-1)},$$

and from this we conclude that

$$\mathbf{E}\|\hat{E}_{j_3(n)} - E_{j_3(n)}f\|^2_{L^2([\varepsilon;1])} \leqslant \frac{C_{10}}{n\varepsilon^2} \sum_{k \in \mathcal{K}_{j_3(n)}^\varepsilon(\phi)} 2^{2j_3(n)} \leqslant C_{11}\varepsilon^{-2} n^{-\frac{2\sigma}{2\sigma+3}}. \tag{31}$$

To evaluate $\mathbf{E}\|\hat{D}_{j_3(n)j_4(n)} - D_{j_3(n)j_4(n)}f\|^2_{L^2([\varepsilon;1])}$ let us define

$$\hat{B}_j = \{k \in \mathcal{K}_j^\varepsilon(\psi) : |\hat{\beta}_{jk}| > T(\varepsilon)c_j\}, \quad \hat{S}_j = \hat{B}_j^c$$
$$B_j = \{k \in \mathcal{K}_j^\varepsilon(\psi) : |\beta_{jk}| > T(\varepsilon)c_j/2\}, \quad S_j = B_j^c$$
$$B'_j = \{k \in \mathcal{K}_j^\varepsilon(\psi) : |\beta_{jk}| > 2T(\varepsilon)c_j\}, \quad S'_j = B'^{c_j}$$
$$e_{bs} = \sum_{j=j_3(n)}^{j_4(n)} \sum_{k \in \hat{B}_j S_j} (\hat{\beta}_{jk} - \beta_{jk})\psi_{jk}$$

$$e_{bb} = \sum_{j=j_3(n)}^{j_4(n)} \sum_{k\in\hat{B}_jB_j} (\hat{\beta}_{jk} - \beta_{jk})\psi_{jk}$$

$$e_{sb} = \sum_{j=j_3(n)}^{j_4(n)} \sum_{k\in\hat{S}_jB'_j} \beta_{jk}\psi_{jk}$$

$$e_{ss} = \sum_{j=j_3(n)}^{j_4(n)} \sum_{k\in\hat{S}_jS'_j} \beta_{jk}\psi_{jk}.$$

We have

$$\hat{D}_{j_3(n)j_4(n)} - D_{j_3(n)j_4(n)}f = e_{bs} + e_{bb} - e_{sb} - e_{ss}. \tag{32}$$

First we evaluate $\mathbf{E}\|e_{bs}\|^2_{L^2([\varepsilon;1])}$. Let

$$G_j = \{k \in \mathcal{K}^\varepsilon_j(\psi) : |\hat{\beta}_{jk} - \beta_{jk}| > T(\varepsilon)c_j/2\}.$$

Since $\hat{B}_jS_j \subset G_j$ we have

$$\mathbf{E}\|e_{bs}\|^2_{L^2([\varepsilon;1])} \leqslant \sum_{j=j_3(n)}^{j_4(n)} \sum_{k\in\mathcal{K}^\varepsilon_j(\psi)} [\mathbf{E}(\hat{\beta}_{jk} - \beta_{jk})^4]^{\frac{1}{2}}[\mathrm{P}(|\hat{\beta}_{jk} - \beta_{jk}| > T(\varepsilon)c_j/2)]^{\frac{1}{2}}.$$

We now use (27) and (28) and we get

$$\mathbf{E}\|e_{bs}\|^2_{L^2([\varepsilon;1])} \leqslant 2\frac{C_{12}A_N}{\varepsilon^2 n} \sum_{j=j_3(n)}^{j_4(n)} 2^{(3-\varepsilon\eta(\varepsilon)/2)j}.$$

If we now choose $\eta(\varepsilon) \geqslant 8\varepsilon^{-1}$ we conclude that

$$\mathbf{E}\|e_{bs}\|^2_{L^2([\varepsilon;1])} \leqslant \frac{C_{13}}{\varepsilon^2 n}. \tag{33}$$

Let us evaluate $\mathbf{E}\|e_{sb}\|^2_{L^2([\varepsilon;1])}$. Since $\hat{S}_jB'_j \subset G_j$ we get

$$\mathbf{E}\|e_{sb}\|^2_{L^2([\varepsilon;1])} \leqslant \sum_{j=j_3(n)}^{j_4(n)} \sum_{k\in\mathcal{K}^\varepsilon_j(\psi)} \beta^2_{jk}\mathrm{P}(|\hat{\beta}_{jk} - \beta_{jk}| > T(\varepsilon)c_j/2).$$

Using (28) we have

$$\mathbf{E}\|e_{sb}\|^2_{L^2([\varepsilon;1])} \leqslant 2\sum_{j=j_3(n)}^{j_4(n)} \sum_{k\in\mathcal{K}^\varepsilon_j(\psi)} \beta^2_{jk}2^{-\varepsilon\eta(\varepsilon)j} = 2\sum_{j=j_3(n)}^{j_4(n)} 2^{-\varepsilon\eta(\varepsilon)j}\|\beta_{j\cdot}\|^2_{l_2}.$$

Since $2^{2j(\sigma'+1/2-1/p)}\|\beta_{j\cdot}\|^2_{l_2} \leqslant C_{14}\|f\|^2_{\sigma'2\infty} \leqslant C_{14}\|f\|^2_{\sigma pq}$, then

$$\mathbf{E}\|e_{sb}\|^2_{L^2([\varepsilon;1])} \leqslant 2C_{14}\|f\|^2_{\sigma pq} \sum_{j=j_3(n)}^{j_4(n)} 2^{-j(2\sigma'+1-\frac{2}{p}+\varepsilon\eta(\varepsilon))}.$$

We notice that $2\sigma' + 1 - \frac{2}{p} + \varepsilon\eta(\varepsilon) > 0$, and from this it follows that

$$\mathbf{E}\|e_{sb}\|^2_{L^2([\varepsilon;1])} \leqslant C_{15}2^{-j_3(n)(2\sigma'+1-\frac{2}{p}+\varepsilon\eta(\varepsilon))} \leqslant C_{16}n^{-\frac{2\sigma'+1-2/p+\varepsilon\eta(\varepsilon)}{2r_0+3}}.$$

If we take $\eta(\varepsilon) \geqslant \varepsilon^{-1}(2r_0 + 2)$, we obtain

$$\mathbf{E}\|e_{sb}\|^2_{L^2([\varepsilon;1])} \leqslant C_{16}n^{-\frac{2\sigma}{2\sigma+3}}. \tag{34}$$

Let us now evaluate $\mathbf{E}\|e_{bb}\|^2_{L^2([\varepsilon;1])}$. We choose $j_0(n)$ such that $2^{j_0(n)} \asymp n^{1/(2\sigma+3)}$. We notice that

$$2^{j_3(n)} \asymp n^{\frac{1}{2r_0+3}} \leqslant n^{\frac{1}{2\sigma+3}} < \frac{n}{\log_2 n} \asymp 2^{j_4(n)},$$

so $j_3(n) \leqslant j_0(n) \leqslant j_4(n)$. We have

$$\begin{aligned}\mathbf{E}\|e_{bb}\|^2_{L^2([\varepsilon;1])} &\leqslant \mathbf{E}\left\|\sum_{j=j_3(n)}^{j_0(n)}\sum_{k\in\hat{B}_jB_j}(\hat{\beta}_{jk}-\beta_{jk})\psi_{jk}\right\|^2_{L^2([\varepsilon;1])} + \mathbf{E}\left\|\sum_{j=j_0(n)}^{j_4(n)}\sum_{k\in\hat{B}_jB_j}(\hat{\beta}_{jk}-\beta_{jk})\psi_{jk}\right\|^2_{L^2([\varepsilon;1])} \\ &= \mathbf{E}\|e_{bba}\|^2_{L^2([\varepsilon;1])} + \mathbf{E}\|e_{bbb}\|^2_{L^2([\varepsilon;1])}.\end{aligned}$$

For $k \in B_j$, $|2\beta_{jk}/(T(\varepsilon)c_j)|^p > 1$, so using (26) we get

$$\mathbf{E}\|e_{bbb}\|^2_{L^2([\varepsilon;1])} \leqslant \frac{C_2}{\varepsilon^2 n^{1-p/2}}\sum_{j=j_0(n)}^{j_4(n)} 2^{j(2-p)} j^{-p/2}\|\beta_{j\cdot}\|^p_{l_p}.$$

Since $j^{-p/2} \leqslant 1$ and $2^{j\sigma' p}\|\beta_{j\cdot}\|^p_{l_p} \leqslant \|f\|^p_{\sigma pq}$, we have

$$\mathbf{E}\|e_{bbb}\|^2_{L^2([\varepsilon;1])} \leqslant \frac{C_{17}}{\varepsilon^2 n^{1-p/2}} 2^{-j_0(n)(\sigma' p+p-2)} \leqslant C_{18}\varepsilon^{-2} n^{-\frac{2\sigma}{2\sigma+3}}. \tag{35}$$

To evaluate $\mathbf{E}\|e_{bba}\|^2_{L^2([\varepsilon;1])}$ we use (26):

$$\mathbf{E}\|e_{bba}\|^2_{L^2([\varepsilon;1])} \leqslant C_2 \sum_{j=j_3(n)}^{j_0(n)} \sum_{k\in\mathcal{K}_j^{\varepsilon}(\psi)} \frac{2^{2j}}{\varepsilon^2 n} \leqslant C_{19}\frac{2^{3j_3(n)}}{\varepsilon^2 n} \leqslant C_{20}\varepsilon^{-2} n^{-\frac{2\sigma}{2\sigma+3}}. \tag{36}$$

From (35) and (36) we obtain

$$\mathbf{E}\|e_{bb}\|^2_{L^2([\varepsilon;1])} \leqslant C_{21}\varepsilon^{-2} n^{-\frac{2\sigma}{2\sigma+3}}. \tag{37}$$

We now evaluate $\mathbf{E}\|e_{ss}\|^2_{L^2([\varepsilon;1])}$. Let $\tau_{jk} = 2^{-j}\beta_{jk}$ and

$$\tau = (\tau_{jk}), \qquad \beta = (\beta_{jk}), \quad j \geqslant j_3(n), k \in \mathcal{K}_j^{\varepsilon}(\psi).$$

If $\beta \in B_{\sigma pq}(M)$, then $\tau \in B_{\bar{\sigma} pq}(M)$, where $\bar{\sigma} = \sigma + 1$. We have

$$\|e_{ss}\|_{L^2([\varepsilon;1])} \leqslant \|\{\tau_{jk} : j_3(n) \leqslant j \leqslant j_4(n),\ k \in S'_j\}\|_{122}.$$

For $k \in S'_j$, $2^j|\tau_{jk}| \leqslant 2T(\varepsilon)2^j\sqrt{j/n}$, so

$$|\tau_{jk}| \leqslant 2T(\varepsilon)\sqrt{\frac{j}{n}} \leqslant 2T(\varepsilon)\sqrt{\frac{j_4(n)}{n}} := \Lambda_n. \tag{38}$$

Let

$$\Omega(\Lambda; \|\cdot\|; A) = \sup\{\|\tau\| : \tau \in A, |\tau_{jk}| < \Lambda\}$$

and

$$\Omega_n = \Omega(\Lambda_n; \|\cdot\|_{122}; B_{\bar{\sigma} pq}(M)). \tag{39}$$

Using (38), (39) and ([14], Theorem 3) we get

$$\Omega_n \leqslant M^{1-\bar{\alpha}}\left(2T(\varepsilon)\sqrt{\frac{j_4(n)}{n}}\right)^{\bar{\alpha}},$$

where

$$\bar{\alpha} = \frac{\bar{\sigma}-1}{\bar{\sigma}+1/2} = \frac{\sigma}{\sigma+3/2} = \frac{2\sigma}{2\sigma+3}.$$

Since $j_4(n) \leqslant C_{22}\log n$ and $T(\varepsilon) = C_5\max\{8\varepsilon^{-1}, (2r_0+2)\varepsilon^{-1}\}$ we have

$$\mathbf{E}\|e_{ss}\|^2_{L^2([\varepsilon;1])} \leqslant \Omega_n^2 \leqslant C_{23}\left(\frac{\log n}{\varepsilon^2 n}\right)^{\frac{2\sigma}{2\sigma+3}}. \tag{40}$$

Using (33), (34), (37) and (40) we finally obtain

$$\sup_{f\in\mathcal{F}_{\sigma pq}(M,[\varepsilon;1])} \mathbf{E}_f \|\tilde{f}_n^{\varepsilon} - f\|^2_{L^2([\varepsilon;1])} \leqslant C_{24}\left[\left(\frac{\log n}{\varepsilon^2 n}\right)^{\frac{2\sigma}{2\sigma+3}} + \varepsilon^{-2} n^{-\frac{2\sigma}{2\sigma+3}}\right]. \tag{41}$$

Since $\varepsilon$ is fixed we have

$$\sup_{f\in\mathcal{F}_{\sigma pq}(M,[\varepsilon;1])} \mathbf{E}_f \|\tilde{f}_n^{\varepsilon} - f\|^2_{L^2([\varepsilon;1])} \leqslant C\left(\frac{\log n}{n}\right)^{\frac{2\sigma}{2\sigma+3}},$$

which proves Theorem 2. □

As in Section 3, the applicability of the estimator can be extended to the case with minimal detection radius equal to zero. For $f \in \mathcal{F}_{\sigma pq}(M, [0; 1])$, we assume (23), then we define $\tilde{f}_n^0$ by taking $\varepsilon = (\log n)^{-\frac{1}{\alpha}}$ and using (41) we have

$$\mathbf{E}_f \|\tilde{f}_n^{0} - f\|^2_{L^2([0;1])} \leqslant C_{25} n^{-\frac{2\sigma}{2\sigma+3}} (\log n)^s,$$

where $s = \max\{(1 + 2/\alpha)(2\sigma/(2\sigma + 3)), 2/\alpha\}$. Combining this with Proposition 1 we conclude that $\tilde{f}_n^0$ achieves almost the optimal rate of convergence to within logarithmic terms.

## Acknowledgments

The author wishes to express his gratitude to Professor Zbigniew Szkutnik for suggesting the problem and for many stimulating conversations.

## References

[1] Z. Szkutnik, Unfolding spheres size distribution from linear sections with B-splines and EMDS algorithm, Opuscula Mathematica 27 (2007) 151–165.
[2] A.G. Spektor, Analysis of distribution of spherical particles in nontransparent structures, Zavodskaja Laboratorija 16 (1950) 173–177.
[3] G.W. Lord, T.F. Willis, Calculation of air bubble distribution from results of a Rosiwal traverse of aerated concrete, ASTM Bulletin 56 (1951) 177–187.
[4] J.-H. Han, D.-Y. Kim, Determination of three-dimensional grain size distribution by linear intercept measurement, Acta Materialia 46 (1998) 2021–2028.
[5] J.E. Hilliard, L.R. Lawson, Stereology and Stochastic Geometry, Kluwer, Dordrecht, 2003.
[6] M. Barthel, P. Klimanek, D. Stoyan, Stereological substructure analysis in hot-deformed metals from TEM-images, Czechoslovak Journal of Physics B, 35, 265–268.
[7] A. Dudek, Z. Szkutnik, Minimax unfolding spheres size distribution from linear sections, Statistica Sinica 18 (2008) 1063–1080.
[8] L. Cavalier, J.-Y. Koo, Poisson intensity estimation for tomographic data using a wavelet shrinkage approach, IEEE Transactions on Information Theory 48 (2002) 2794–2802.
[9] D.L. Donoho, Nonlinear solution of linear inverse problems by wavelet–vaguelette decomposition, Applied and Computational Harmonic Analysis 2 (1995) 101–126.
[10] W. Härdle, G. Kerkyacharian, D. Picard, A.B. Tsybakov, Wavelets, Approximation and Statistical Applications, Springer-Verlag, New York, 1998.
[11] R.D. Reiss, A Course on Point Processes, Springer, New York, 1993.
[12] J.F.C. Kingman, Poisson Processes, Oxford, University Press, 1993.
[13] D.L. Donoho, I.M. Johnstone, G. Kerkyacharian, D. Picard, Density estimation by wavelet thresholding, The Annals of Statistics 24 (1996) 508–539.
[14] D.L. Donoho, I.M. Johnstone, G. Kerkyacharian, D. Picard, Universal near minimaxity of wavelet shrinkage, in: Festschrift for Lucien Le Cam, Springer-Verlag, New York, 1996.