Inversion of noisy Radon transform by SVD based needlet

Gérard Kerkyacharian, George Kyriazis, Erwan Le Pennec, Pencho Petrushev, Dominique Picard

▶ **To cite this version:**

Gérard Kerkyacharian, George Kyriazis, Erwan Le Pennec, Pencho Petrushev, Dominique Picard. Inversion of noisy Radon transform by SVD based needlet. 2008. hal-00321894v1

HAL Id: hal-00321894
https://hal.science/hal-00321894v1

Preprint submitted on 19 Sep 2008 (v1), last revised 17 Aug 2009 (v2)

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

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Inversion of noisy Radon transform by SVD based needlets

Kerkyacharian Gérard, Kyriazis George, Le Pennec Erwan,
Petrushev Pencho and Picard Dominique

July 6, 2008

Abstract

A linear method for inverting noisy observations of the Radon transform is developed based on decomposition systems (needlets) with rapidly decaying elements induced by the Radon transform SVD basis. Upper bounds of the risk of the estimator are established in L^p $(1 \leq p \leq \infty)$ norms for functions with Besov space smoothness. A practical implementation of the method is given and several examples are discussed.

1 Introduction

Reconstructing images (functions) from their Radon transforms is a fundamental problem in medical imaging and more generally in tomography. The problem is to find an accurate and efficient algorithm for approximation of the function to be recovered from its Radon projections. In this paper, we consider the problem of inverting noisy observations of the Radon transform. As in many other inverse problems, there exists a basis which is fully adapted to the problem, in particular, the inversion in this basis is very stable; this is the Singular Value Decomposition (SVD) basis. The Radon transform SVD basis, however, is not quite suitable for decomposition of functions with regularities in other than L^2-related spaces. In particular, the SVD basis is not quite capable of representing local features of images, which are especially important to recover.

Our idea is to design an estimation method for inverting the Radon transform which has the advantages of maximum localization of wavelet based methods combined with the stability and computability of the SVD methods. To this end we utilize the construction from [19] (see also [13]) of localized frames based on

orthogonal polynomials on the ball, which are closely related to the Radon transform SVD basis. As shown in the simulation section the results obtained are quite promising.

A parallel method, employing a similar idea for smoothing out the projection operator but without using the needlet construction and in a no-noise framework, has been developed by Yuan Xu and his co-authors in [25, 26, 27].

The paper is structured as follows. In Section 2 we introduce the model and the Radon transform Singular Value Decomposition. In Section 3 we give the class of linear estimators built upon the SVD. We also give the needlet construction and introduce the needlet estimation algorithm. In Section 4 we establish bounds for the risk of this estimate over large classes of regularity spaces. Section 5 is devoted to the practical implementation and results of our method. Section 6 is an appendix where the proofs of some claims from Section 3 are given.

2 Radon transform and white noise model

2.1 Radon transform

Here we recall the definition and some basic facts about the Radon transform (cf. [10], [17], [14]). Denote by B^d the unit ball in $\mathbb{R}^d$, i.e. $B^d = \{x = (x_1, \dots, x_d) \in \mathbb{R}^d : |x| \leq 1\}$ with $|x| = (\sum_{i=1}^d x_i^2)^{1/2}$ and by $\mathbb{S}^{d-1}$ the unit sphere in $\mathbb{R}^d$. The Lebesgue measure on B^d will be denoted by dx and the usual surface measure on $\mathbb{S}^{d-1}$ by $d\sigma(x)$ (sometimes we will also deal with the surface measure on $\mathbb{S}^d$ which will be denoted by $d\sigma_d$). We let $|A|$ denote the measure $|A| = \int_A dx$ if $A \subset B^d$ as well as $|A| = \int_A d\sigma(x)$ if $A \subset \mathbb{S}^{d-1}$.

The Radon transform of a function f is defined by

$$Rf(\theta, s) = \int_{\substack{y \in \theta^\perp \\ s\theta + y \in B^d}} f(s\theta + y)dy, \quad \theta \in \mathbb{S}^{d-1},\ s \in [-1, 1],$$

where dy is the Lebesgue measure of dimension $d-1$ and $\theta^\perp = \{x \in \mathbb{R}^d : \langle x, \theta \rangle = 0\}$. With a slight abuse of notation, we will rewrite this integral as

$$Rf(\theta, s) = \int_{\langle y, \theta \rangle = s} f(y)dy.$$

By Fubini's theorem, we have

$$\int_{-1}^{1} Rf(\theta, s)ds = \int_{B^d} f(x)dx. \tag{2.1}$$

Hence, for $\theta \in \mathbb{S}^{d-1}$, $Rf(\theta, s)$ is well defined for almost all $s \in [-1, 1]$. Notice that $Rf(-\theta, -s) = Rf(\theta, s)$.

An obvious consequence of (2.1) is that for any function g defined on $[-1,+1]$,

$$\int_{-1}^{1} Rf(\theta,s)g(s)ds = \int_{B^d} f(x)g(\langle x,\theta\rangle)dx. \tag{2.2}$$

It is easy to see (cf. e.g. [17]) that the Radon transform is a bounded linear operator mapping $\mathbb{L}^2(B^d,dx)$ into $\mathbb{L}^2\left(\mathbb{S}^{d-1}\times[-1,1],d\mu(\theta,s)\right)$, where

$$d\mu(\theta,s) = d\sigma(\theta)\frac{ds}{(1-s^2)^{(d-1)/2}}.$$

2.2 Noisy observation of the Radon transform

We consider observations of the form

$$dY(\theta,s) = Rf(\theta,s)d\mu(\theta,s) + \varepsilon dW(\theta,s), \tag{2.3}$$

where the unknown function f belongs to $\mathbb{L}^2(B^d,dx)$. The meaning of this equation is that for any $\phi(\theta,s)$ in $\mathbb{L}^2(\mathbb{S}^{d-1}\times[-1,1],d\mu(\theta,s))$ one can observe

$$\begin{aligned} Y_\phi &= \int \phi(\theta,s)dY(\theta,s) = \int_{\mathbb{S}^{d-1}\times[-1,1]} Rf(\theta,s)\phi(\theta,s)d\mu(\theta,s) + \varepsilon\int \phi(\theta,s)dW(\theta,s) \\ &= \langle Rf,\phi\rangle_\mu + \varepsilon W_\phi. \end{aligned}$$

Here $W_\phi = \int \phi(\theta,s)dW(\theta,s)$ is a Gaussian field of zero mean and covariance

$$\mathbb{E}(W_\phi,\ W_\psi) = \int_{\mathbb{S}^{d-1}\times[-1,1]} \phi(\theta,s)\psi(\theta,s)d\sigma(\theta)\frac{ds}{(1-s^2)^{(d-1)/2}} = \langle\phi,\psi\rangle_\mu.$$

The goal is to recover the unknown function f from the observation of Y. Our idea is to devise an estimation scheme which combines the stability and computability of SVD decompositions with the superb localization and multiscale structure of wavelets. To this end we utilize a frame (essentially following the construction from [13]) with elements of nearly exponential localization which is compatible with the SVD basis of the Radon transform.

2.3 Singular Value Decomposition of the Radon transform

The SVD of the Radon transform was first established in [4, 15]. In this regard we also refer the reader to [17, 25]. In this section we record some basic facts related to the Radon SVD.

2.3.1 Adjoint operator

We next identify the adjoint operator R^* of the Radon transform operator

$$R : \mathbb{L}^2(B^d, dx) \mapsto \mathbb{L}^2(\mathbb{S}^{d-1} \times [-1,1], d\mu(\theta, s)).$$

Proposition 2.1. *For any* $g \in \mathbb{L}^2(\mathbb{S}^{d-1} \times [-1,1], d\mu(\theta, s))$

$$R^* g(x) = \int_{S^{d-1}} g(\theta, \langle x, \theta \rangle) \frac{1}{(1 - |\langle x, \theta \rangle|^2)^{(d-1)/2}} d\sigma(\theta).$$

This proposition is immediate from (2.2) (cf. [17]).

2.3.2 Jacobi and Gegenbauer polynomials

The Radon SVD bases are defined in terms of Jacobi and Gegenbauer polynomials. The Jacobi polynomials $P_n^{(\alpha,\beta)}$, $n \geq 0$, constitute an orthogonal basis for the space $\mathbb{L}^2([-1,1], w_{\alpha,\beta}(t)dt)$ with weight $w_{\alpha,\beta}(t) = (1-t)^\alpha (1+t)^\beta$, $\alpha, \beta > -1$. They are standardly normalized by $P_n^{(\alpha,\beta)}(1) = \binom{n+\alpha}{n}$ and then [1, 8, 22]

$$\int_{-1}^{1} P_n^{(\alpha,\beta)}(t) P_m^{(\alpha,\beta)}(t) w_{\alpha,\beta}(t) dt = \delta_{n,m} h_n^{(\alpha,\beta)},$$

where

$$h_n^{(\alpha,\beta)} = \frac{2^{\alpha+\beta+1}}{(2n+\alpha+\beta+1)} \frac{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}{\Gamma(n+1)\Gamma(n+\alpha+\beta+1)}. \tag{2.4}$$

The Gegenbauer polynomials C_n^λ are a particular case of Jacobi polynomials and are traditionally defined by

$$C_n^\lambda(t) = \frac{(2\lambda)_n}{(\lambda+1/2)_n} P_n^{(\lambda-1/2,\, \lambda-1/2)}(t), \quad \lambda > -1/2,$$

where by definition $(a)_n = a(a+1)\dots(a+n-1) = \frac{\Gamma(a+n)}{\Gamma(a)}$ (note that in [22] the Gegenbauer polynomial C_n^λ is denoted by P_n^λ). It is readily seen that $C_n^\lambda(1) = \binom{n+2\lambda-1}{n} = \frac{\Gamma(n+2\lambda)}{n!\Gamma(2\lambda)}$ and

$$\int_{-1}^{1} C_n^\lambda(t) C_m^\lambda(t) (1-t^2)^{\lambda-\frac{1}{2}} dt = \delta_{n,m} h_n^{(\lambda)} \quad \text{with} \quad h_n^{(\lambda)} = \frac{2^{1-2\lambda}\pi}{\Gamma(\lambda)^2} \frac{\Gamma(n+2\lambda)}{(n+\lambda)\Gamma(n+1)}. \tag{2.5}$$

2.3.3 Polynomials on B^d and $\mathbb{S}^{d-1}$

Let $\Pi_n(\mathbb{R}^d)$ be the space of all polynomials in d variables of degree $\leq n$. We denote by $\mathcal{P}_n(\mathbb{R}^d)$ the space of all homogeneous polynomials of degree n and by $\mathcal{V}_n(\mathbb{R}^d)$ the space of all polynomials of degree n which are orthogonal to lower degree polynomials with respect to the Lebesgue measure on B^d. Of course $\mathcal{V}_0$ will be the set of all constants. We have the following orthogonal decomposition:

$$\Pi_n(\mathbb{R}^d) = \bigoplus_{k=0}^{n} \mathcal{V}_k(\mathbb{R}^d).$$

Also, denote by $\mathcal{H}_n(\mathbb{R}^d)$ the subspace of all harmonic homogeneous polynomials of degree n (i.e. $Q \in \mathcal{H}_n(\mathbb{R}^d)$ if $Q \in \mathcal{P}_n(\mathbb{R}^d)$ and $\Delta Q = 0$) and by $\mathcal{H}_n(\mathbb{S}^{d-1})$ the (injective) restriction of the polynomials from $\mathcal{H}_n(\mathbb{R}^d)$ to $\mathbb{S}^{d-1}$. It is well known that

$$N_{d-1}(n) = \dim(\mathcal{H}_n(\mathbb{S}^{d-1})) = \binom{n+d-1}{d-1} - \binom{n+d-3}{d-1} \sim n^{d-2}.$$

Let $\Pi_n(\mathbb{S}^{d-1})$ be the space of restrictions to $\mathbb{S}^{d-1}$ of polynomials of degree $\leq n$ on $\mathbb{R}^d$. As is well known

$$\Pi_n(\mathbb{S}^{d-1}) = \bigoplus_{m=0}^{n} \mathcal{H}_m(\mathbb{S}^{d-1})$$

(the orthogonality is with respect of the surface measure $d\sigma$ on $\mathbb{S}^{d-1}$).

Let $Y_{l,i}$, $1 \leq i \leq N_{d-1}(l)$, be an orthonormal basis of $\mathcal{H}_l(\mathbb{S}^{d-1})$, i.e.

$$\int_{\mathbb{S}^{d-1}} Y_{l,i}(\xi)\overline{Y_{l,i'}(\xi)}d\sigma(\xi) = \delta_{i,i'}.$$

Then the natural extensions of $Y_{l,i}$ on B^d are defined by $Y_{l,i}(x) = |x|^l Y_{l,i}\big(\frac{x}{|x|}\big)$ and satisfy

$$\begin{aligned}\int_{B^d} Y_{l,i}(x)\overline{Y_{l,i'}(x)}dx &= \int_0^1 r^{d-1}\int_{\mathbb{S}^{d-1}} Y_{l,i}(r\xi)\overline{Y_{l,i'}(r\xi)}d\sigma(\xi)dr \\ &= \int_0^1 r^{d+2l-1}\int_{\mathbb{S}^{d-1}} Y_{l,i}(\xi)\overline{Y_{l,i'}(\xi)}d\sigma(\xi)dr = \delta_{i,i'}\frac{1}{2l+d}.\end{aligned}$$

For more details we refer the reader to [6].

The spherical harmonics on $\mathbb{S}^{d-1}$ and orthogonal polynomials on B^d are naturally related to Gegenbauer polynomials. Thus the kernel of the orthogonal projector onto $\mathcal{H}_n(\mathbb{S}^{d-1})$ can be written as (see e.g. [21]):

$$\sum_{i=1}^{N_{d-1}(n)} Y_{l,i}(\xi)\overline{Y_{l,i}(\theta)} = \frac{2n+d-2}{(d-2)|\mathbb{S}^{d-1}|}C_n^{\frac{d-2}{2}}(\langle\xi,\theta\rangle). \tag{2.6}$$

The "ridge" Gegenbauer polynomials $C_n^{d/2}(\langle x,\xi\rangle)$ are orthogonal to $\Pi_{n-1}(B^d)$ in $\mathbb{L}^2(B^d)$ and the kernel $L_n(x,y)$ of the orthogonal projector onto $\mathcal{V}_n(B^d)$ can be written in the form (see e.g. [18, 25])

$$L_n(x,y) = \frac{2n+d}{|\mathbb{S}^{d-1}|^2}\int_{\mathbb{S}^{d-1}} C_n^{d/2}(\langle x,\xi\rangle)C_n^{d/2}(\langle y,\xi\rangle)d\sigma(\xi) \tag{2.7}$$
$$= \frac{(n+1)_{d-1}}{2^d\pi^{d-1}}\int_{\mathbb{S}^{d-1}} \frac{C_n^{d/2}(\langle x,\xi\rangle)C_n^{d/2}(\langle y,\xi\rangle)}{\|C_n^{d/2}\|^2}d\sigma(\xi).$$

The following important identities are valid for "ridge" Gegenbauer polynomials:

$$\int_{B^d} C_n^{d/2}(\langle \xi,x\rangle)C_n^{d/2}(\langle \eta,x\rangle)dx = \frac{h_n^{(d/2)}}{C_n^{d/2}(1)}C_n^{d/2}(\langle \xi,\eta\rangle), \quad \xi,\eta\in\mathbb{S}^{d-1}, \tag{2.8}$$

and, for $x\in B^d$, $\eta\in\mathbb{S}^{d-1}$,

$$\int_{\mathbb{S}^{d-1}} C_n^{d/2}(\langle \xi,x\rangle)C_n^{d/2}(\langle \xi,\eta\rangle)d\sigma(\xi) = |\mathbb{S}^{d-1}|C_n^{d/2}(\langle \eta,x\rangle), \tag{2.9}$$

see e.g. [18]. By (2.7) and (2.9)

$$L_n(x,\xi) = \frac{(2n+d)}{|\mathbb{S}^{d-1}|}C_n^{d/2}(\langle x,\xi\rangle), \quad \xi\in\mathbb{S}^{d-1},$$

and again by (2.7)

$$\int_{\mathbb{S}^{d-1}} L_n(x,\xi)L_n(y,\xi)d\sigma(\xi) = (2n+d)L_n(x,y).$$

2.3.4 The SVD of the Radon transform

Assume that $\{Y_{l,i} : 1\le i\le N_{d-1}(l)\}$ is an orthonormal basis for $\mathcal{H}_l(\mathbb{S}^{d-1})$. Then it is standard and easy to see that the family of polynomials

$$f_{k,l,i}(x) = (2k+d)^{1/2}P_j^{(0,l+d/2-1)}(2|x|^2-1)Y_{l,i}(x),\ \ 0\le l\le k,\ k-l=2j,\ 1\le i\le N_{d-1}(l),$$

form an orthonormal basis of $\mathcal{V}_k(B^d)$, see e.g. [6]. Here as before $Y_{l,i}(x) = |x|^l Y_{l,i}(x/|x|)$. On the other hand the collection

$$g_{k,l,i}(\theta,s) = [h_k^{(d/2)}]^{-1/2}(1-s^2)^{(d-1)/2}C_k^{d/2}(s)Y_{l,i}(\theta), \quad k\ge 0,\ l\ge 0,\ 1\le i\le N_{d-1}(l),$$

is apparently an orthonormal basis of $\mathbb{L}^2(\mathbb{S}^{d-1}\times[-1,1], d\mu(\theta,s))$. Most importantly, the Radon transform $R : \mathbb{L}^2(B^d)\mapsto \mathbb{L}^2(\mathbb{S}^{d-1}\times[-1,1], d\mu(\theta,s))$ is an one-to-one mapping and

$$Rf_{k,l,i} = \lambda_k g_{k,l,i}, \quad R^* g_{k,l,i} = \lambda_k f_{k,l,i}, \quad \text{where}$$

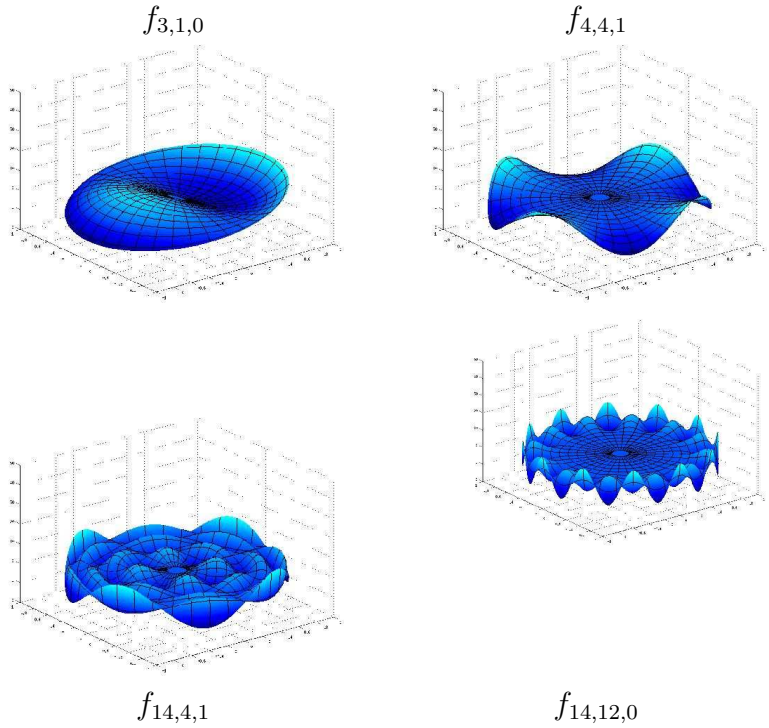

Figure 1: A few radon SVD basis elements

$$\lambda_k^2 = \frac{2^d \pi^{d-1}}{(k+1)(k+2)\dots(k+d-1)} = \frac{2^d \pi^{d-1}}{(k+1)_{d-1}} \sim k^{-d+1}. \tag{2.10}$$

Figure 1 displays a few $f_{k,l,i}$ and illustrates their lack of localization.

We are now prepared to give the SVD decomposition of the Radon transform.

Theorem 2.2. *For any* $f \in \mathbb{L}^2(B^d)$

$$Rf = \sum_{k\geq 0} \lambda_k \sum_{0\leq l\leq k,\, k-l\equiv 0 \,(\mathrm{mod}\, 2)} \sum_{1\leq i\leq N_{d-1}(l)} \langle f, f_{k,l,i}\rangle g_{k,l,i} \tag{2.11}$$

and for any $g \in \mathbb{L}^2(\mathbb{S}^{d-1}\times[-1,1], d\mu(\theta,s))$

$$R^* g = \sum_{k\geq 0} \lambda_k \sum_{0\leq l\leq k,\, k-l\equiv 0 \,(\mathrm{mod}\, 2)} \sum_{1\leq i\leq N_{d-1}(l)} \langle g, g_{k,l,i}\rangle_\mu f_{k,l,i}. \tag{2.12}$$

Furthermore, for $f \in \mathbb{L}^2(B^d)$

$$f = \sum_{k\geq 0} \lambda_k^{-1} \sum_{0\leq l\leq k,\, k-l\equiv 0 \,(\mathrm{mod}\, 2)} \sum_{1\leq i\leq N_{d-1}(l)} \langle Rf, g_{k,l,i}\rangle_\mu f_{k,l,i}. \tag{2.13}$$

In the above identities the convergence is in $\mathbb{L}^2$.

Remark : Observe that if $k \geq 0$, $0 \leq l \leq k$, $k - l \not\equiv 0 \,(\mathrm{mod}\, 2)$, and $1 \leq i \leq N_{d-1}(l)$, then $R^* f_{k,l,i} = 0$. $\star$

For the Radon SVD we refer the reader to [17] and [25]. We will only show in the appendix that λ_k^2 has the value given in (2.10).

3 Linear estimators built upon the SVD

3.1 The general idea

In a general noisy inverse model

$$dY_t = Kf dt + \varepsilon dW_t,$$

where K is a linear operator mapping $f \in \mathbb{H} \mapsto Kf \in \mathbb{K}$, and $\mathbb{H}$ and $\mathbb{K}$ are two Hilbert spaces, the SVD yields a family of linear estimators via the following classical scheme.

Suppose K has an SVD

$$Kf = \sum_m \sigma_m \langle f, e_m \rangle e_m^*, \quad f \in \mathbb{H},$$

where $\{e_m\}$ and $\{e_m^*\}$ are orthonormal bases for $\mathbb{H}$ and $\mathbb{K}$, respectively, and $Ke_m = \sigma_m e_m^*$ and $K^* e_m^* = \sigma_m e_m$ with K^* being the adjoint operator of K. We also assume that $\sigma_m \to 0$. Then, if $\sigma_m \neq 0$,

$$\begin{aligned} \int e_m^* dY_t &= \int Kf \cdot e_m^* dt + \varepsilon \int e_m^* dW_t = \int f \cdot K^* e_m^* dt + \varepsilon \int e_m^* dW_t \\ &= \sigma_m \int f e_m dt + \varepsilon \int e_m^* dW_t \end{aligned}$$

and hence

$$\frac{1}{\sigma_m} \int e_m^* dY_t = \langle f, e_m \rangle + \frac{\varepsilon}{\sigma_m} \int e_m^* dW_t. \tag{3.1}$$

In going further, suppose that $\{\phi_l\}$ is a tight frame for $\mathbb{H}$. Therefore, for any $f \in \mathbb{H}$

$$f = \sum_l \alpha_l \phi_l, \quad \alpha_l = \langle f, \phi_l \rangle.$$

We can represent ϕ_l in the basis $\{e_m\}$:

$$\phi_l = \sum_m \langle \phi_l, e_m \rangle e_m = \sum_m \gamma_l^m e_m$$

and hence

$$\alpha_l = \sum_m \gamma_l^m \langle f, e_m \rangle.$$

On account of (3.1) this leads to the estimator

$$\hat{f}_N = \sum_{l \le N} \hat{\alpha}_l \phi_l \quad \text{with} \quad \hat{\alpha}_l = \sum_m \gamma_l^m \frac{1}{\sigma_m} \int e_m^* dY_t, \tag{3.2}$$

where N is a parameter. By (3.1) we have

$$\hat{\alpha}_l = \sum_m \gamma_l^m \langle f, e_m \rangle + \sum_m \gamma_l^m \frac{\varepsilon}{\sigma_m} \int e_m^* dW_t = \alpha_l + Z_l,$$

where Z_l has a normal distribution $N(0, \sum_m (\gamma_l^m)^2 \frac{\varepsilon^2}{\sigma_m^2})$. In this scheme the factors $\frac{1}{\sigma_m^2}$, which are inherent to the inverse model, bring instability by inflating the variance.

The selection of the frame $\{\phi_l\}$ is critical for the method described above. The standard SVD method corresponds to the choice $\phi_l = e_l$. This SVD method is very attractive theoretically and can be shown to be asymptotically optimal in many situations (see Dicken and Maass [5], Mathé and Pereverzev [16] together with their nonlinear counterparts Cavalier and Tsybakov [3], Cavalier et al [2], Tsybakov [23], Goldenschluger and Pereverzev [9], Efromovich and Kolchinskii [7]). It also has the big advantage of performing a quick and stable inversion of the operator K. However, while the SVD bases are fully adapted to describe the operator K, they are usually not quite appropriate for accurate description of the solution of the problem with a small number of parameters. Although the SVD method is suitable for estimating the unknown function f with an L^2-loss, it is inappropriate for other losses. It is also restricted to functions which are well represented in terms of the Sobolev space associated to the SVD basis. Switching to an arbitrary frame $\{\phi_m\}$, however, may yield additional instability through the factors $(\gamma_l^m)^2$'s.

Our idea is to utilize a frame $\{\phi_l\}$ which is compatible with the SVD basis $\{e_m\}$, allowing to keep the variance within reasonable bounds, and has elements with superb space localization and smoothness, guaranteeing excellent approximation of the unknown function f. In the following we implement the above described method to the inversion of the Radon transform. We shall build upon the frames constructed in [19] and called "needlets".

3.2 Construction of needlets on the ball

In this part we construct the building blocks of our estimator. We will essentially follow the construction from [19].

3.2.1 The orthogonal projector L_k on $\mathcal{V}_k(B^d)$.

Let $\{f_{k,l,i}\}$ be the orthonormal basis of $\mathcal{V}_k(B^d)$ defined in §2.3.4. Denote by T_k the index set of this basis, i.e. $T_k = \{(l,i) : 0 \le l \le k,\ l \equiv k(\mathrm{mod}\,2),\ 0 \le i \le N_{d-1}(l)\}$. Also, set $\nu = d/2 - 1$. Then the orthogonal projector of $\mathbb{L}^2(B^d)$ onto $\mathcal{V}_k(B^d)$ can be written in the form

$$L_k f = \int_{B^d} f(y) L_k(x,y)dy \quad \text{with} \quad L_k(x,y) = \sum_{l,i \in T_k} f_{k,l,i}(x) f_{k,l,i}(y).$$

Using (2.6) $L_k(x,y)$ can be written in the form

$$\begin{aligned} &L_k(x,y) \qquad\qquad (3.3)\\ &= (2k+d) \sum_{l\le k,\, k-l\equiv 0(2)} P_j^{(0,l+\nu)}(2|x|^2-1)|x|^l P_j^{(0,l+\nu)}(2|y|^2-1)|y|^l \sum_i Y_{l,i}\Big(\frac{x}{|x|}\Big) Y_{l,i}\Big(\frac{y}{|y|}\Big)\\ &= \frac{(2k+d)}{|\mathbb{S}^{d-1}|} \sum_{l\le k,\, k-l\equiv 0(2)} P_j^{(0,l+\nu)}(2|x|^2-1)|x|^l P_j^{(0,l+\nu)}(2|y|^2-1)|y|^l \Big(1+\frac{l}{\nu}\Big) C_l^\nu\Big(\Big\langle \frac{x}{|x|}, \frac{y}{|y|}\Big\rangle\Big). \end{aligned}$$

Another representation of $L_k(x,y)$ has already be given in (2.7). Clearly

$$\int_{B^d} L_m(x,z) L_k(z,y)dz = \delta_{m,k} L_m(x,y) \qquad (3.4)$$

and for $f \in \mathbb{L}^2(B^d)$

$$f = \sum_{k\ge 0} L_k f \quad \text{and} \quad \|f\|_2^2 = \sum_k \|L_k f\|_2^2 = \sum_k \langle L_k f, f\rangle. \qquad (3.5)$$

3.2.2 Smoothing

Let $a \in C^\infty[0,\infty)$ be a cut-off function such that $0 \le a \le 1$, $a(t) = 1$ for $t \in [0,1/2]$ and $\mathrm{supp}\, a \subset [0,1]$. We next use this function to introduce a sequence of operators on $\mathbb{L}^2(B^d)$. For $j \ge 0$ write

$$A_j f(x) = \sum_{k\ge 0} a\Big(\frac{k}{2^j}\Big) L_k f(x) = \int_{B^d} A_j(x,y) f(y)dy \quad \text{with} \ \ A_j(x,y) = \sum_k a\Big(\frac{k}{2^j}\Big) L_k(x,y).$$

Also, we define $B_j f = A_{j+1} f - A_j f$. Then setting $b(t) = a(t/2) - a(t)$ we have

$$B_j f(x) = \sum_k b\Big(\frac{k}{2^j}\Big) L_k f(x) = \int_{B^d} B_j(x,y) f(y)dy \quad \text{with} \ \ B_j(x,y) = \sum_k b\Big(\frac{k}{2^j}\Big) L_k(x,y).$$

Evidently, for $f \in \mathbb{L}^2(B^d)$

$$\langle A_j f, f\rangle = \sum_k a\Big(\frac{k}{2^j}\Big)\langle L_k f, f\rangle \le \|f\|_2^2 \tag{3.6}$$

and

$$\lim_{j\to\infty} \|A_j f - f\|_2 = \lim_{j\to\infty} \|(A_0 + \sum_{m=0}^{j-1} B_m)f - f\|_2 = 0. \tag{3.7}$$

An important result from [19] (see also [13]) asserts that the kernels $A_j(x,y)$, $B_j(x,y)$ have nearly exponential localization, namely, for any $M > 0$ there exists a constant $c_M > 0$ such that

$$|A_j(x,y)|, |B_j(x,y)| \le C_M \frac{2^{jd}}{(1+2^j d(x,y))^M \sqrt{W_j(x)}\sqrt{W_j(y)}}, \quad x, y \in B^d, \tag{3.8}$$

where $W_j(x) = 2^{-j} + \sqrt{1-|x|^2}$, $|x|^2 = |x|_d^2 = \sum_{i=1}^d x_i^2$, and

$$d(x,y) = \text{Arccos}\,(\langle x, y\rangle + \sqrt{1-|x|^2}\sqrt{1-|y|^2}), \quad \langle x, y\rangle = \sum_{i=1}^d x_i y_i. \tag{3.9}$$

The left part of Figure 2 illustrates this concentration: it displays the influence of a point x to the value of $B_j f$ at a second point y_0, namely the values of $B_j(x, y_0)$ for a fixed y_0. This influence peaks at y_0 and vanishes exponentially fast to 0 as soon as one goes away from y_0. The right part of Figure 2 shows the lack of concentration of B_j when the cut-off function a used is far from being C^∞. The resulting kernel still peaks at y_0 but the value of $B_j f$ at y_0 is strongly influenced by values far away from y_0.

Remark : At this point it is important to notice the following correspondence which will be used in the sequel. For $\mathbb{S}_+^d = \{(x,z) \in \mathbb{R}^d \times \mathbb{R}^+,\ |x|_d^2 + z^2 = 1\}$, we have the natural bijection

$$x \in B^d \mapsto \tilde{x} = (x, \sqrt{1-|x|^2}) \quad \text{and} \quad d(x,y) = d_{\mathbb{S}_+^d}(\tilde{x}, \tilde{y}),$$

where $d_{\mathbb{S}_+^d}$ is the geodesic distance on $\mathbb{S}_+^d$. ⋆

3.2.3 Approximation

Here we discuss the approximation properties of the operators $\{A_j\}$. We will show that in a sense they are operators of "near best" polynomial L^p-approximation. Denote by $E_n(f,p)$ the best L^p-approximation of $f \in \mathbb{L}^p(B^d)$ from Π_n, i.e.

$$E_n(f,p) = \inf_{P\in\Pi_n} \|f - P\|_p. \tag{3.10}$$

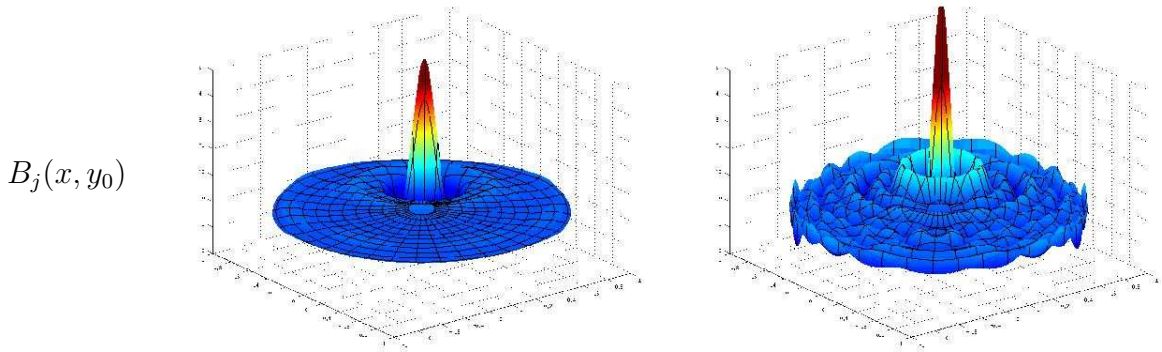

Figure 2: Smoothing kernel $B_j(x, y_0)$ for a fixed y_0 for (1) a C^∞ cut-off function a and (2) a non smooth cut-off function a

Estimate (3.8) yields (cf. ([19, Proposition 4.5])

$$\int_{B^d} |A_j(x, y)|dy \le c^*, \quad x \in B^d,\ j \ge 0,$$

where c^* is a constant depending only on d. Therefore, the operators A_j are (uniformly) bounded on $\mathbb{L}^1(B^d)$ and $\mathbb{L}^\infty(B^d)$, and hence, by interpolation, on $\mathbb{L}^p(B^d)$, $1 \le p \le \infty$, i.e.

$$\|A_j f\|_p \le c^* \|f\|_p, \quad f \in \mathbb{L}^p(B^d). \tag{3.11}$$

On the other hand, since $a(t) = 1$ on $[0, 1/2]$ we have $A_j P = P$ for $P \in \Pi_{2^{j-1}}$. We use this and (3.11) to obtain, for $f \in \mathbb{L}^p(B^d)$ and an arbitrary polynomial $P \in \Pi_{2^{j-1}}$,

$$\|f - A_j f\|_p = \|f - P + P - A_j f\|_p \le \|f - P\|_p + \|A_j(P - f)\|_p \le (1 + c^*)\|f - P\|_p = K\|f - P\|_p.$$

Consequently, $\|f - A_j f\|_p \le K E_{2^{j-1}}(f, p)$. In the opposite direction, evidently, $A_j f \in \Pi_{2^j}$ and hence $E_{2^j}(f, p) \le \|f - A_j f\|_p$. Therefore, for $f \in \mathbb{L}^p(B^d)$, $1 \le p \le \infty$,

$$E_{2^j}(f, p) \le \|f - A_j f\|_p \le K E_{2^{j-1}}(f, p). \tag{3.12}$$

These estimates do not tell the whole truth about the approximation power of A_j. It is rather obvious that because of the superb localization of the kernel $A_j(x, y)$ the operator A_j provides far better rates of approximation than $E_{2^{j-1}}(f, \infty)$ away from the singularities of f.

In contrast, the kernel $S_j(x, y) = \sum_{0 \le k \le 2^j} L_k(x, y)$ of the orthogonal projector S_j onto Π_{2^j} is poorly localized and hence S_j is useless for approximation in L^p, $p \ne 2$. This partially explains the fact that the traditional SVD estimators perform poorly in L^p-norms when $p \ne 2$.

3.2.4 Splitting procedure

Let us define

$$C_j(x,y) = \sum_m \sqrt{a\Big(\frac{m}{2^j}\Big)} L_m(x,z)) \quad \text{and} \quad D_j(x,y) = \sum_m \sqrt{b\Big(\frac{m}{2^j}\Big)} L_m(x,z).$$

Note that C_j and D_j have the same localization as the localization of A_j, B_j in (3.8) (cf. [19]). Using (3.4), we get the desired splitting

$$A_j(x,y) = \int_{B^d} C_j(x,z) C_j(z,y) dz \tag{3.13}$$

and

$$B_j(x,y) = \int_{B^d} D_j(x,z) D_j(z,y) dz. \tag{3.14}$$

Obviously $z \mapsto C_j(x,z)C_j(z,y)$ is a polynomial of degree $< 2^{j+1}$ and $z \mapsto D_j(x,z)D_j(z,y)$ is a polynomial of degree $< 2^{j+2}$. The next step is to discretize the kernels $A_j(x,y)$ and $B_j(x,y)$.

3.2.5 Cubature formula and discretization

To construct the needlets on B^d we need one more ingredient - a cubature formula on B^d exact for polynomials of a given degree.

Recall first the bijection between the ball B^d (equipped with the usual Lebesgue measure) and the unit upper hemisphere in $\mathbb{R}^{d+1}$:

$$\mathbb{S}^d_+ = \{(x,y),\ x \in \mathbb{R}^d,\ 0 \le y \le 1,\ \ |x|_d^2 + y^2 = 1\}$$

equipped with $d\sigma$ the usual surface measure.

$$T : (x,y) \in \mathbb{S}^d_+ \mapsto x \in \mathbb{R}^d$$

and

$$T^{-1} : x \in \mathbb{R}^d \mapsto \tilde{x} = (x, \sqrt{1 - |x|_d^2}) \in \mathbb{S}^d_+$$

Applying the substitution T one has (see e.g. [24])

$$\int_{\mathbb{S}^d_+} F(x,y) d\sigma(x,y) = \int_{B^d} F(x, \sqrt{1-|x|^2}) \frac{dx}{\sqrt{1-|x|^2}} \tag{3.15}$$

and hence for $f : \mathbb{R}^d \mapsto \mathbb{R}$

$$\int_{(x,y)\in S^d_+} f(x) y d\sigma_d(x,y) = \int_{x \in B^d} f(x) \sqrt{1-|x|_d^2} \frac{dx}{\sqrt{1-|x|_d^2}} = \int_{B^d} f(x) dx. \tag{3.16}$$

Therefore, given a cubature formula on $\mathbb{S}^d$ one can easily derive a cubature formula on B^d. Indeed, suppose we have a cubature formula on $\mathbb{S}^d_+$ exact for all polynomials of degree $n+1$, i.e, there exist $\tilde{\chi}_n \subset \mathbb{S}^d_+$ and coefficients $\omega_{\tilde{\xi}} > 0$, $\tilde{\xi} \in \tilde{\chi}_n$, such that

$$\int_{\mathbb{S}^d_+} P(u)d\sigma(u) = \sum_{\tilde{\xi}\in\tilde{\chi}_n} \omega_{\tilde{\xi}} P(\tilde{\xi}) \quad \forall P \in \Pi_{n+1}(\mathbb{R}^{d+1}).$$

If $P \in \Pi_n(\mathbb{R}^d)$ then $P(x)y \in \Pi_{n+1}(\mathbb{R}^{d+1})$ and hence

$$\sum_{\tilde{\xi}\in\tilde{\chi}_n} \omega_{\tilde{\xi}} P(\xi)\sqrt{1-\xi^2} = \int_{\mathbb{S}^d_+} P(x)yd\sigma = \int_{B^d} P(x)dx.$$

Thus the projection χ_n of $\tilde{\chi}_n$ on B^d is the set of nodes and the associated coefficients given by $\omega_\xi = \sqrt{1-\xi^2}\omega_{\tilde{\xi}}$ induce a cubature formula on B^d exact for $\Pi_n(\mathbb{R}^d)$.

The following proposition follows from results in [19] and [24].

Proposition 3.1. *Let $\{B(\tilde{\xi}_i, \rho) : i \in I\}$ be a maximal family of disjoint spherical caps of radius $\rho = \tau 2^{-j}$ with centers on the hemisphere $\mathbb{S}^d_+$. Then for sufficiently small $0 < \tau \leq 1$ the set of points $\chi_j = \{\xi_i : i \in I\}$ obtained by projecting the set $\{\tilde{\xi}_i : i \in I\}$ on B^d is a set of nodes of a cubature formula which is exact for $\Pi_{2j+2}(B^d)$. Moreover, the coefficients ω_{ξ_i} of this cubature formula are positive and $\omega_{\xi_i} \sim W_j(\xi_i)2^{-jd}$. Also, the cardinality $\#\chi_j \sim 2^{jd}$.*

3.2.6 Needlets

Going back to identities (3.13) and (3.14) and applying the cubature formula described in Proposition 3.1, we get

$$A_j(x,y) = \int_{B^d} C_j(x,z)C_j(z,y)dz = \sum_{\xi\in\chi_j} \omega_\xi C_j(x,\xi)C_j(y,\xi) \quad \text{and}$$
$$B_j(x,y) = \int_{B^d} D_j(x,z)D_j(z,y)dz = \sum_{\xi\in\chi_j} \omega_\xi D_j(x,\xi)D_j(y,\xi).$$

We define the **father needlets** $\phi_{j,\xi}$ and the **mother needlets** $\psi_{j,\xi}$ by

$$\phi_{j,\xi}(x) = \sqrt{\omega_\xi}C_j(x,\xi) \quad \text{and} \quad \psi_{j,\xi}(x) = \sqrt{\omega_\xi}D_j(x,\xi), \quad \xi \in \chi_j,\ j \geq 0.$$

We also set $\psi_{-1,0} = \frac{\mathbb{1}_{B^d}}{|B^d|}$ and $\chi_{-1} = \{0\}$. From above it follows that

$$A_j(x,y) = \sum_{\xi\in\chi_j} \phi_{j,\xi}(x)\phi_{j,\xi}(y), \quad B_j(x,y) = \sum_{\xi\in\chi_j} \psi_{j,\xi}(x)\psi_{j,\xi}(y).$$

Therefore,

$$A_j f(x) = \int_{B^d} A_j(x,y) f(y) dy = \sum_{\xi \in \chi_j} \langle f, \phi_{j,\xi} \rangle \phi_{j,\xi} = \sum_{\xi \in \chi_j} \alpha_{j,\xi} \phi_{j,\xi}, \quad \alpha_{j,\xi} = \langle f, \phi_{j,\xi} \rangle. \tag{3.17}$$

and

$$B_j f(x) = \int_{B^d} B_j(x,y) f(y) dy = \sum_{\xi \in \chi_j} \langle f, \psi_{j,\xi} \rangle \psi_{j,\xi} = \sum_{\xi \in \chi_j} \beta_{j,\xi} \psi_{j,\xi}, \quad \beta_{j,\xi} = \langle f, \psi_{j,\xi} \rangle. \tag{3.18}$$

By (3.17) and (3.6) we have

$$\|\phi_{j,\xi}\|_2^2 \geq \langle A_j \phi_{j,\xi}, \phi_{j,\xi} \rangle = \langle \sum_{\xi' \in \chi_j} \langle \phi_{j,\xi}, \phi_{j,\xi'} \rangle \phi_{j,\xi'}, \phi_{j,\xi} \rangle = \sum_{\xi' \in \chi_j} |\langle \phi_{j,\xi}, \phi_{j,\xi'} \rangle|^2 \geq \|\phi_{j,\xi}\|_2^4$$

and hence

$$\|\phi_{j,\xi}\|_2 \leq 1. \tag{3.19}$$

From (3.5) and the fact that $\sum_{j \geq 0} b(t 2^{-j}) = 1$ for $t \in [1, \infty)$, it readily follows that

$$f = \sum_{j \geq -1} \sum_{\xi \in \chi_j} \langle f, \psi_{j,\xi} \rangle \psi_{j,\xi}, \quad f \in \mathbb{L}^2(B^d),$$

and taking inner product with f this leads to

$$\|f\|_2^2 = \sum_j \sum_{\xi \in \chi_j} |\langle f, \psi_{j,\xi} \rangle|^2,$$

which in turn shows that the family $\{\psi_{j,\xi}\}$ is a tight frame for $\mathbb{L}^2(B^d)$ and consequently

$$\|\psi_{j,\xi}\|_2^2 \geq \|\psi_{j,\xi}\|_2^4, \quad \text{i.e.} \quad \|\psi_{j,\xi}\|_2 \leq 1. \tag{3.20}$$

Observe that using the properties of the cubature formula from Proposition 3.1, estimate (3.8) leads to the localization estimate (cf. [19]):

$$|\phi_{j,\xi}(x)|, |\psi_{j,\xi}(x)| \leq C_M \frac{2^{jd/2}}{\sqrt{W_j(\xi)}(1 + 2^j d(x,\xi))^M} \quad \forall M > 0. \tag{3.21}$$

Nontrivial lower bounds for the norms of the needlets are obtained in [13]. More precisely, in [13] it is shown that for $0 < p \leq \infty$

$$\|\psi_{j,\xi}\|_p \sim \|\phi_{j,\xi}\|_p \sim \Big(\frac{2^{jd}}{W_j(\xi)}\Big)^{1/2-1/p}, \quad \xi \in \chi_j. \tag{3.22}$$

We next record some properties of needlets which will be needed later on. For convenience we will denote in the following by $h_{j,\xi}$ either $\phi_{j,\xi}$ or $\psi_{j,\xi}$.

Theorem 3.2. *Let $1 \le p \le \infty$ and $j \ge 1$. The following inequalities hold*

$$\sum_{\xi \in \chi_j} \|h_{j,\xi}\|_p^p \le c 2^{j(dp/2+(p/2-2)_+)} \quad \text{if} \quad p \ne 4, \tag{3.23}$$

$$\sum_{\xi \in \chi_j} \|h_{j,\xi}\|_p^p \le c j 2^{jdp/2} \quad \text{if} \quad p = 4, \tag{3.24}$$

and for any collection of complex numbers $\{d_\xi\}_{\xi \in \chi_j}$

$$\| \sum_{\xi \in \chi_j} d_\xi h_{j,\xi} \|_p \le c \Big(\sum_{\xi \in \chi_j} |d_\xi|^p \|h_{j,\xi}\|_p^p \Big)^{1/p}. \tag{3.25}$$

Here $c > 0$ is a constant depending only on d, p, and τ.

To make our presentation more fluid we relegate the proof of this theorem to the appendix.

3.3 Linear needlet estimator

Our motivation for introducing the estimator described below is the excellent approximation power of the operators A_j defined in §3.2.2 and its compatibility with the Radon SVD. We begin with the following representation of the unknown function f

$$f = \sum_{k,l,i} \langle f, f_{k,l,i} \rangle \, f_{k,l,i},$$

where the sum is over the index set $\{(k,l,i) : k \ge 0, 0 \le l \le k, l \equiv k (\mathrm{mod}\, 2), 1 \le i \le N_{d-1}(l)\}$. Combining this with the definition of A_j we get

$$A_j f = \sum_{\xi \in \chi_j} \langle f, \phi_{j,\xi}(y) \rangle \, \phi_{j,\xi} = \sum_{\xi \in \chi_j} \alpha_{j,\xi} \phi_{j,\xi},$$

where

$$\alpha_{j,\xi} = \langle f, \phi_{j,\xi} \rangle = \sum_{k,l,i} \gamma_{k,l,i}^{j,\xi} \langle f, f_{k,l,i} \rangle = \sum_{k,l,i} \gamma_{k,l,i}^{j,\xi} \frac{1}{\lambda_k} \langle R(f), g_{k,l,i} \rangle_\mu = \sum_{k,l,i} \gamma_{k,l,i}^{j,\xi} \frac{1}{\lambda_k} \int g_{k,l,i} R(f) d\mu.$$

Here $\gamma_{k,l,i}^{j,\xi} = \langle f_{k,l,i}, \phi_{j,\xi}(y) \rangle$ can be precomputed.

It seems natural to us to define an estimator $\widehat{f_j}$ of the unknown function f by

$$\widehat{f_j} = \sum_{\xi \in \chi_j} \widehat{\alpha}_{j,\xi} \phi_{j,\xi}, \tag{3.26}$$

where

$$\widehat{\alpha}_{j,\xi} = \sum_{k,l,i} \gamma_{k,l,i}^{j,\xi} \frac{1}{\lambda_k} \int g_{k,l,i} dY. \tag{3.27}$$

Here the summation is over $\{(k,l,i) : 0 \leq k < 2^j, 0 \leq l \leq k, l \equiv k (\mathrm{mod}\, 2), 1 \leq i \leq N_{d-1}(l)\}$ and j is a parameter.

Some clarification is needed here. The father and mother needlets, introduced in §3.2.6, are closely related but play different roles. Both $\phi_{j,\xi}$ and $\psi_{j,\xi}$ have superb localization, however, the mother needlets $\{\psi_{j,\xi}\}$ have multilevel structure and, therefore, are an excellent tool for nonlinear n-term approximation of functions on the ball, whereas the father needlets are perfectly well suited for linear approximation. So, there should be no surprise that we use the father needlets for our linear estimator.

Furthermore, even if the needlets are central in the analysis of the estimator, the estimator $\widehat{f}_j$ can be defined without them. Indeed,

$$\widehat{f}_j = \sum_{\xi \in \chi_j} \sum_{k,l,i} \gamma_{k,l,i}^{j,\xi} \frac{1}{\lambda_k} \int g_{k,l,i} dY \, \phi_{j,\xi}$$

as all the sum are finite, their order can be interchanged, yielding

$$\widehat{f}_j = \sum_{k,l,i} \frac{1}{\lambda_k} \int g_{k,l,i} dY \sum_{\xi \in \chi_j} \gamma_{k,l,i}^{j,\xi} \phi_{j,\xi} = \sum_{k,l,i} \frac{1}{\lambda_k} \int g_{k,l,i} dY \, A_j f_{k,l,i}$$

and thus the estimator is obtained by a simple componentwise multiplication on the SVD coefficients

$$\widehat{f}_j = \sum_{k,l,i} \frac{a\left(\frac{k}{2^j}\right)}{\lambda_k} \int g_{k,l,i} dY \, f_{k,l,i}.$$

4 The risk of the needlet estimator

In this section we estimate the risk of the needlet estimator introduced above in terms of the Besov smoothness of the unknown function.

4.1 Besov spaces

We introduce the Besov spaces of positive smoothness on the ball as spaces of L^p-approximation from algebraic polynomials. As in §3.2.3 we will denote by $E_n(f,p)$ the best L^p-approximation of $f \in \mathbb{L}^p(B^d)$ from Π_n. We will mainly use the notations from [13].

Definition 4.1. *[13] Let $0 < s < \infty$, $1 \le p \le \infty$, and $0 < q \le \infty$. The space $B^{s,0}_{p,q}$ on the ball is defined as the space of all functions $f \in \mathbb{L}^p(B^d)$ such that*

$$|f|_{B^{s,0}_{p,q}} = \Big(\sum_{n\ge 1} (n^s E_n(f,p))^q \frac{1}{n} \Big)^{1/q} < \infty \quad \text{if } q < \infty,$$

and $|f|_{B^{s,0}_{p,q}} = \sup_{n\ge 1} n^s E_n(f,p) < \infty$ if $q = \infty$. The norm on $B^{s,0}_{p,q}$ is defined by

$$\|f\|_{B^{s,0}_{p,q}} = \|f\|_p + |f|_{B^{s,0}_{p,q}}.$$

Remark : From the monotonicity of $\{E_n(f,p)\}$ it readily follows that

$$\|f\|_{B^{s,0}_{p,q}} \sim \|f\|_p + \Big(\sum_{j\ge 0} (2^{js} E_{2^j}(f,p))^q \Big)^{1/q}$$

with the obvious modification when $q = \infty$. ⋆

There are several different equivalent norms on the Besov space $B^{s,0}_{p,q}$.

Theorem 4.2. *With indexes s, p, q as in the above definition the following norms are equivalent to the Besov norm $\|f\|_{B^{s,0}_{p,q}}$:*

$$\begin{aligned}
&(i) \quad \mathcal{N}_1(f) = \|f\|_p + \|(2^{js}\|f - A_j f\|_p)_{j\ge 0}\|_{l^q},\\
&(ii) \quad \mathcal{N}_2(f) = \|f\|_p + \|(2^{js}\|B_j f\|_p)_{j\ge 1}\|_{l^q},\\
&(iii) \quad \mathcal{N}_3(f) = \|f\|_p + \|(2^{js} \sum_{\xi\in\chi_j} |\langle f, \psi_{j,\xi}\rangle|^p \|\psi_{j,\xi}\|_p^p)_{j\ge -1}\|_{l^q}.
\end{aligned}$$

Proof. The equivalence $\mathcal{N}_1(f) \sim \|f\|_{B^{s,0}_{p,q}}$ is immediate from (3.12).

To prove that $\mathcal{N}_2(f) \sim \mathcal{N}_1(f)$, we recall that $B_j = (A_{j+1} - A_j)$ (see §3.2.2) and hence $\|B_j f\|_p \le \|f - A_{j+1} f\|_p + \|f - A_j f\|_p$ which readily implies $\mathcal{N}_2(f) \le c\mathcal{N}_1(f)$. In the other direction, we have

$$\|f - A_j f\|_p = \|\sum_{l=j}^{\infty} B_l f\|_p \le \sum_{l=j}^{\infty} \|B_l f\|_p.$$

Assuming that $\mathcal{N}_2(f) < \infty$ we have $\|B_l(f)\|_p = \alpha_l 2^{-ls}$ with $\{\alpha_l\} \in l^q$. Hence

$$\sum_{l=j}^{\infty} \|B_l(f)\|_p = \sum_{l=j}^{\infty} \alpha_l 2^{-ls} = 2^{-js} \sum_{l=j}^{\infty} \alpha_l 2^{-(l-j)s} =: 2^{-js}\beta_j$$

and by the convolution inequality $\{\beta_j\} \in l^q$. Therefore, $\mathcal{N}_1(f) \le c\mathcal{N}_2(f)$.

For the equivalence $\mathcal{N}_3(f) \sim \|f\|_{B^{s,0}_{p,q}}$, see [13, Theorem 5.4]. □

4.1.1 Comparison with the "classical" Besov spaces

The classical Besov space $B^s_{p,q}(B^d)$ is defined through the L^p-norm of the finite differences:

$$\Delta_h f(x) = (f(x+h) - f(x))\mathbb{1}_{x\in B^d}\mathbb{1}_{x+h\in B^d}$$

and in general

$$\Delta_h^N f(x) = \mathbb{1}_{x\in B^d}\mathbb{1}_{x+Nh\in B^d}\sum_{k=0}^{N}(-1)^{N-k}\binom{N}{k}f(x+kh).$$

Then the Nth modulus of smoothness in L^p is defined by

$$\omega_p^N(f,t) = \sup_{|h|\le t}\|\Delta_h^N f\|_p, \quad t>0.$$

For $0<s<N$, $1\le p\le\infty$, and $0<q\le\infty$, the classical Besov space $B^s_{p,q}$ is defined by the norm

$$\|f\|_{B^s_{p,q}} = \|f\|_p + \Big(\int_0^1 [t^s\omega_p^N(f,t)]^q\frac{dt}{t}\Big)^{1/q} \sim \|f\|_p + \Big(\sum_{j=0}^{\infty}[2^{js}\omega_p^N(f,2^{-j})]^q\Big)^{1/q}$$

with the usual modification for $q=\infty$. It is well known that the definition of $B^s_{p,q}$ does not depend on N as long as $s<N$ [11]. Moreover, the imbedding

$$B^s_{p,q}\subseteq B^{s,0}_{p,q}, \tag{4.1}$$

is immediate from the estimate $E_n(f,p)\le c\omega_p^N(f,1/n)$ [11].

4.2 Upper bound for the risk of the needlet estimator

Theorem 4.3. *Let $1\le p\le\infty, 0<s<\infty$, and assume that $f\in B^{s,0}_{p,\infty}$ with $\|f\|_{B^{s,0}_{p,\infty}}\le M$. Let*

$$\widehat{f}_J = \sum_{\xi\in\chi_J}\widehat{\alpha}_{J,\xi}\phi_{j,\xi}$$

be the needlet estimator introduced in §3.3, where J is selected depending on the parameters as described below.

1. *If $M2^{-J(s+d)}\sim\varepsilon$ when $p=\infty$, then*

$$\mathbb{E}\|f-\widehat{f}_J\|_\infty \le c_\infty M^{\frac{d}{s+d}}\varepsilon^{\frac{s}{s+d}}\sqrt{\log M/\varepsilon}.$$

2. *If* $M2^{-Js} \sim \varepsilon 2^{J(d-2/p)}$ *when* $4 \leq p < \infty$, *then*

$$\mathbb{E}\|f - \widehat{f}_J\|_p^p \leq c_p M^{\frac{(d-2/p)p}{s+d-2/p}} \varepsilon^{\frac{sp}{s+d-2/p}},$$

where when $p = 4$ *there is an additional factor* $\ln(M/\varepsilon)$ *on the right.*

3. *If* $M2^{-Js} \sim \varepsilon 2^{J(d-1/2)}$ *when* $1 \leq p < 4$, *then*

$$\mathbb{E}\|f - \widehat{f}_j\|_p^p \leq c_p M^{\frac{(d-1/2)p}{s+d-1/2}} \varepsilon^{\frac{sp}{s+d-1/2}}.$$

Remarks :

- It will be shown in a forthcoming paper that the following rates of convergence are, in fact, minimax, i.e. there exist positive constants c_1 and c_2 such that

$$\sup_{\|f\|_{B^{s,0}_{p,\infty}} \leq M} \inf_{\tilde{f} \text{ estimator}} \mathbb{E}\|f - \tilde{f}\|_p^p \geq c_1 \max\{\varepsilon^{\frac{sp}{s+d-2/p}},\ \varepsilon^{\frac{sp}{s+d-1/2}}\},$$

$$\sup_{\|f\|_{B^{s,0}_{\infty,\infty}} \leq M} \inf_{\tilde{f} \text{ estimator}} \mathbb{E}\|f - \tilde{f}\|_\infty \geq c_2 \varepsilon^{\frac{s}{s+d}} \sqrt{\log 1/\varepsilon}.$$

- The case $p = 2$ above corresponds to the standard SVD method which involves Sobolev spaces. In this setting, minimax rates have already been established (cf. [5], [16] [3], [2], [23], [9], [7]); these rates are $\varepsilon^{\frac{2s}{s+d-1/2}}$. Also, it has been shown that the SVD algorithms yield minimax rates. These results extend (using straightforward comparisons of norms) to L^p losses for $p < 4$, but still considering the Sobolev ball $\{\|f\|_{B^{s,0}_{2,\infty}} \leq M\}$ rather than the Besov ball $\{\|f\|_{B^{s,0}_{p,\infty}} \leq M\}$. Therefore, our results can be viewed as an extension of the above results, allowing a much wider variety of regularity spaces.

- The Besov spaces involved in our bounds are in a sense well adapted to our method. However, the embedding results from Section 4.1.1 shows that the bounds from Theorem 4.3 hold in terms of the standard Besov spaces as well. This means that in using the Besov spaces described above, our results are but stronger.

- In the case $p \geq 4$ we exhibit here new minimax rates of convergence, related to the ill posedness coefficient of the inverse problem $\frac{d-1}{2}$ along with edge effects induced by the geometry of the ball. These rates have to be compared with similar phenomena occurring in other inverse problems involving Jacobi polynomials (e.g. Wicksell problem), see [12].

$\star$

4.3 Proof of Theorem 4.3

Assume $f \in B^{s,0}_{p,\infty}$ and $\|f\|_{B^{s,0}_{p,\infty}} \le M$. Then by Theorem 4.2,

$$\|A_j f - f\|_p \le c\|f\|_{B^{s,0}_{p,\infty}} 2^{-js} \le cM2^{-js}. \tag{4.2}$$

Now from

$$dY = Rf d\mu + \varepsilon dW$$

we have

$$\begin{aligned}\int g_{k,l,i}\, dY &= \int_Z Rf\, g_{k,l,i}\, d\mu + \varepsilon \int g_{k,l,i}\, dW = \int_{B^d} f\, R^* g_{k,l,i}\, dx + \varepsilon\, Z_{k,l,i} \\ &= \lambda_k \int_{B^d} f\, f_{k,l,i}\, dx + \varepsilon\, Z_{k,l,i}\end{aligned}$$

and hence

$$\frac{1}{\lambda_k}\int g_{k,l,i}\, dY = \int_{B^d} f\, f_{k,l,i} dx + \frac{\varepsilon}{\lambda_k}\, Z_{k,l,i}.$$

On account of (3.27) this leads to

$$\begin{aligned}\widehat{\alpha}_{j,\xi} &= \sum_{k,l,i} \gamma^{j,\xi}_{k,l,i} \int_{B^d} f f_{k,l,i} dx + \sum_{k,l,i} \gamma^{j,\xi}_{k,l,i} \frac{\varepsilon}{\lambda_k}\, Z_{k,l,i} \\ &= \alpha_{j,\xi} + Z_{j,\xi}.\end{aligned}$$

Here the summation is over $\{(k,l,i) : 0 \le k < 2^j, 0 \le l \le k, l \equiv k (\mathrm{mod}\, 2), 1 \le i \le N_{d-1}(l)\}$. Since $Z_{k,l,i}$ are independent $N(0,1)$ random variables, $Z_{j,\xi} \sim N(0, \sigma^2_{j,\xi})$ with

$$\sigma^2_{j,\xi} = \varepsilon^2 \sum_{k,l,i} |\gamma^{j,\xi}_{k,l,i}|^2 \frac{(k)_d}{\pi^{d-1} 2^d k} \le \frac{(2^j)_{d-1}}{\pi^{d-1}2^d} \le c 2^{j(d-1)} \varepsilon^2 \tag{4.3}$$

with $c = (d/2\pi)^{d-1}$. Here we used that $\{f_{k,l,i}\}$ is an orthonormal basis for $\mathbb{L}^2$ and hence $\sum_{k,l,i} |\gamma^{j,\xi}_{k,l,i}|^2 = \|\phi_{j,\xi}\|_2^2 \le 1$.

From (3.26) $\widehat{f}_j = \sum_{\xi \in \chi_j} \widehat{\alpha}_{j,\xi} \phi_{j,\xi}$and using (4.2) we have, whenever $1 \le p < \infty$,

$$\begin{aligned}\mathbb{E}\|f - \widehat{f}_j\|_p^p &\le 2^{p-1}\{\|f - A_j f\|_p^p + \mathbb{E}\|A_j f - \widehat{f}_j\|_p^p\} \\ &\le 2^{p-1}\{cM^p 2^{-jsp} + \mathbb{E}\|A_j f - \widehat{f}_j\|_p^p\}\end{aligned} \tag{4.4}$$

and, for $p = \infty$,

$$\begin{aligned}\mathbb{E}\|f - \widehat{f}_j\|_\infty &\le \|f - A_j f\|_\infty + \mathbb{E}\|A_j f - \widehat{f}_j\|_\infty \\ &\le cM2^{-js} + \mathbb{E}\|A_j f - \widehat{f}_j\|_\infty.\end{aligned} \tag{4.5}$$

On the other hand, using inequality (3.25) of Theorem 3.2 we obtain, if $1 \le p < \infty$,

$$\|A_j f - \widehat{f}_j\|_p^p = \|\sum_{\xi\in\chi_j}(\alpha_{j,\xi} - \widehat{\alpha}_{j,\xi})\phi_{j,\xi}\|_p^p \le c\sum_{\xi\in\chi_j}|\alpha_{j,\xi} - \widehat{\alpha}_{j,\xi}|^p\|\phi_{j,\xi}\|_p^p$$

and hence

$$\mathbb{E}\|A_j f - \widehat{f}_j\|_p^p \le c\sum_{\xi\in\chi_j}\mathbb{E}|Z_{j,\xi}|^p\|\phi_{j,\xi}\|_p^p \le c(\varepsilon 2^{j(d-1)/2})^p\sum_{\xi\in\chi_j}\|\phi_{j,\xi}\|_p^p, \tag{4.6}$$

where we used that $\mathbb{E}|Z_{j,\xi}|^p \le c(\varepsilon 2^{j(d-1)/2})^p$. Similarly, for $p = \infty$,

$$\|A_j f - \widehat{f}_j\|_\infty = \|\sum_{\xi\in\chi_j}(\alpha_{j,\xi} - \widehat{\alpha}_{j,\xi})\phi_{j,\xi}\|_\infty \le c\max_{\xi\in\chi_j}|\alpha_{j,\xi} - \widehat{\alpha}_{j,\xi}|\|\phi_{j,\xi}\|_\infty$$

and hence

$$\begin{aligned}\mathbb{E}\|A_j f - \widehat{f}_j\|_\infty &\le c\mathbb{E}\{\max_{\xi\in\chi_j}|Z_{j,\xi}|\|\phi_{j,\xi}\|_\infty\}\\ &\le c\varepsilon 2^{j(d-1)/2}\max_{\xi\in\chi_j}\|\phi_{j,\xi}\|_\infty\sqrt{2\log_2 2^{jd}} \le c\varepsilon 2^{jd}\sqrt{j}.\end{aligned} \tag{4.7}$$

For the second inequality above we used Pisier's lemma: If $Z_j \sim N(0,\sigma_j^2)$, $\sigma_j \le \sigma$, then

$$\mathbb{E}(\sup_{1\le j\le N}|Z_j|) \le \sigma\sqrt{2\log_2 N}.$$

We also used that $\max_{\xi\in\chi_j}\|\phi_{j,\xi}\|_\infty \le c2^{j(d+1)/2}$, which follows by inequality (3.23) of Theorem 3.2.

Combining (4.5) and (4.7) we obtain, for $p = \infty$,

$$\mathbb{E}\|f - \widehat{f}_j\|_\infty \le c\{M2^{-js} + \varepsilon 2^{jd}\sqrt{j}\}$$

and if $M2^{-j(s+d)} \sim \varepsilon$, then

$$\mathbb{E}\|f - \widehat{f}_j\|_\infty \le cM^{\frac{d}{s+d}}\varepsilon^{\frac{s}{s+d}}\sqrt{\log M/\varepsilon}.$$

Similarly, combining estimate (3.23) of Theorem 3.2 with (4.6) and inserting the resulting estimate in (4.4) we obtain in the case $4 \le p < \infty$

$$\begin{aligned}\mathbb{E}\|f - \widehat{f}_j\|_p^p &\le c\{M2^{-jsp} + (\varepsilon 2^{j(d-1)/2})^p 2^{jdp/2+p/2-2}\}\\ &= c\{M^p 2^{-jsp} + \varepsilon^p 2^{j(dp-2)}\}.\end{aligned}$$

If $M2^{-js} \sim \varepsilon 2^{j(d-2/p)}$ this yields

$$\mathbb{E}\|f - \widehat{f}_j\|_p^p \le cM^{\frac{(d-2/p)p}{s+d-2/p}}\varepsilon^{\frac{sp}{s+d-2/p}}.$$

Accordingly, for $p = 4$ we combine inequality (3.24) with (4.6) and insert the result in (4.4) to obtain

$$\begin{aligned}\mathbb{E}\|f - \widehat{f}_j\|_p^p &\leq c\{M^p 2^{-jsp} + (\varepsilon 2^{j(d-1)/2})^p j 2^{jdp/2}\} \\ &= c\{M^p 2^{-jsp} + j(\varepsilon 2^{j(d-1/2)})^p\}\end{aligned}$$

and if $M2^{-js} \sim \varepsilon 2^{j(d-1/2)}$ this yields

$$\mathbb{E}\|f - \widehat{f}_j\|_p^p \leq cM^{\frac{(d-2/p)p}{s+d-2/p}} \varepsilon^{\frac{sp}{s+d-1/2}} \log M/\varepsilon.$$

Finally, if $1 \leq p < 4$ as above we obtain using (3.23), (4.6), and (4.4)

$$\begin{aligned}\mathbb{E}\|f - \widehat{f}_j\|_p^p &\leq c\{M^p 2^{-jsp} + (\varepsilon 2^{j(d-1)/2})^p 2^{jdp/2}\} \\ &= c\{M^p 2^{-jsp} + (\varepsilon 2^{j(d-1/2)})^p\}.\end{aligned}$$

So, if $M2^{-js} \sim \varepsilon 2^{j(d-1/p)}$, then

$$\mathbb{E}\|f - \widehat{f}_j\|_p^p \leq cM^{\frac{(d-1/2)p}{s+d-1/2}} \varepsilon^{\frac{sp}{s+d-1/2}}.$$

This completes the proof of the theorem. □

5 Application to the Fan Beam Tomography

5.1 Radon and $2d$ Fan Beam Tomography

We have implemented this scheme for $d = 2$. This case corresponds to the fan beam Radon transform used in Computed Axial Tomography (CAT). As shown if Figure 3, an object is positioned in the middle of the device. X rays are sent from a pointwise source $S(\theta_1)$ located on the boundary and making an angle θ_1 with the horizontal. They go through the object and are received on the other side on uniformly sampled array of receptors $R(\theta_1, \theta_2)$. The log decay of the energy from the source to a receptor is proportional to the integral of the density f of the object along the ray and thus one finally measures

$$\tilde{R}f(\theta_1, \theta_2) = \int_{e_{\theta_1} + \lambda e_{\theta_1 - \theta_2} \in B^2} f(x) d\lambda$$

with $e_\theta = (\cos\theta, \sin\theta)$ or equivalently the classical Radon transform

$$Rf(\theta, s) = \int_{\substack{y \in \theta^\perp \\ s\theta + y \in B^1}} f(s\theta + y) dy, \quad \theta \in \mathbb{S}^1,\ s \in [-1, 1],$$

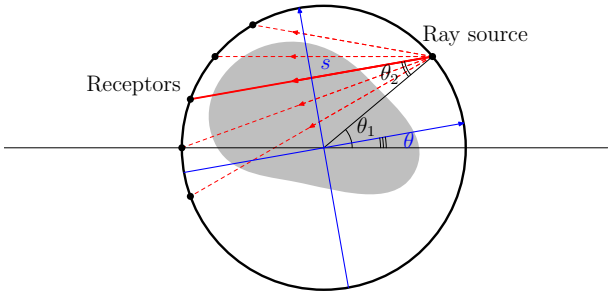

Figure 3: Simplified CAT device

for $\theta = \theta_1 - \theta_2$ and $s = \sin\theta_2$. The device is then rotated to a different angle θ_1 and the process is repeated. Note that $d\theta\frac{ds}{(1-s^2)}$ is nothing but the measure corresponding to the uniform $d\theta_1 d\theta_2$ by the change of variable that maps (θ_1, θ_2) into (θ, s). The white noise model studied here can thus be seen as a proxy for a regression problem with the uniform design on θ_1 and θ_2 that is implemented in real devices.

The Fan Beam Radon SVD basis of the disk is tensorial in polar coordinates:

$$f_{k,l,i}(r,\theta) = (2k+2)^{1/2} P_j^{(0,l)}(2|r|^2-1)|r|^l Y_{l,i}(\theta), \;\; 0 \le l \le k, \; k-l = 2j, \; 1 \le i \le 2,$$

where $P_j^{0,l}$ is the corresponding Jacobi polynomial, and $Y_{l,1}(\theta) = c_l \cos(l\theta)$ and $Y_{l,2}(\theta) = c_l \sin(\theta)$ with $c_0 = \frac{1}{\sqrt{2\pi}}$ and $c_l = \frac{1}{\sqrt{\pi}}$ otherwise. The basis of $S^2 \times [-1,1]$ has a similar tensorial structure as it is given by

$$g_{k,l,i}(\theta,t) = [h_k]^{-1/2}(1-t^2)^{1/2} C_k^1(t) Y_{l,i}(\theta), \quad k \ge 0, \; l \ge 0, \; 1 \le i \le 2,$$

where C_k^1 is the Gegenbauer of parameter 1 and degree k. The corresponding eigenvalues are

$$\lambda_k = \frac{2\sqrt{\pi}}{\sqrt{k+1}}.$$

The estimator $\widehat{f_j}$ is thus simply defined by

$$\widehat{f_j} = \sum_{k,l,i} \frac{a\left(\frac{k}{2^j}\right)}{\lambda_k} \int g_{k,l,i} dY \; f_{k,l,i} = \sum_{k,l,i} \frac{a\left(\frac{k}{2^j}\right)\sqrt{k+1}}{2\sqrt{\pi}} \int g_{k,l,i} dY \; f_{k,l,i}$$

which only requires the knowledge of $\int g_{k,l,i} dY$, that is exactly the data that are supposed to be known in the white noise model.

5.2 Numerical results

To illustrate the advantages of the linear needlet estimator over the linear SVD estimator, we have compared their performances on three synthetic examples, the classical Logan Shepp phantom[20] and two smoothed version, for different L^p norm and different noise level. The Logan Shepp phantom is a synthetic image used as a benchmark in the tomography community. It is a simple toy model for human body structures simplified as a piecewise constant function with discontinuities along ellipsoids (see Figure 4). This example is not regular in a classical sense. Indeed, it belongs to $B^{1,0}_{1,1}$ but not to any $B^{s,0}_{p,q}$ with $s > 1$. We propose to study two smoothed versions that are more regular and thus closer to the theoretical setting of this paper. They have been obtained by convolving the original phantom with a Gaussian kernel with different widths. Note that they are very regular as they belong to every $B^{s,0}_{p,q}$, but that the corresponding norms are large for large s.

To conduct the experiments, we have adopted the following scheme. Denote by f any of the three functions presented above ('Logan', 'smoothed Logan', 'very smoothed Logan'), we have approximated their decomposition in the SVD basis $f_{k,l,i}$ up to degree $k_0 = 128$ with a numerical quadrature χ valid for polynomial of degree $8 \times k_0 = 1024$,

$$\langle f, f_{k,l,i} \rangle \simeq \sum_{(r_i,\theta_i)\in\chi} \omega_{(r_i,\theta_i)} f(r_i,\theta_i) f_{k,l,i}(r_i,\theta_i) = c_{k,l,i}$$

and used this value to approximate the SVD coefficients of $R(f)$, the noiseless Radon transform of f,

$$\langle R(f), f_{k,l,i} \rangle \simeq \lambda_k c_{k,l,i}.$$

A noisy observation $\int g_{k,l,i} dY$ is thus generated by

$$\int g_{k,l,i} dY \simeq \lambda_k c_{k,l,i} + \epsilon W_{k,l,i}$$

where ϵ is the noise level and $W_{k,l,i}$ a iid sequence of standard Gaussian random variable. Our linear needlet estimator $\widehat{f}_J$ of level $J = \log_2(k^N)$ is then computed as

$$\widehat{f}_J = \sum_{k\le k_0,l,i} a\left(\frac{k}{2^J}\right)(c_{k,l,i} + \frac{\epsilon}{\lambda_k} W_{k,l,i}) f_{k,l,i}$$

while the linear SVD estimator $\widehat{f}^S_{k^S}$ of degree k^S is defined as

$$\widehat{f}^S_{k^S} = \sum_{k\le k^S,l,i} (c_{k,l,i} + \frac{\epsilon}{\lambda_k} W_{k,l,i}) f_{k,l,i}.$$

We also consider the naive inversion up to degree k_0 $\widehat{f}^I$ which is equal to $\widehat{f}^S_{k_0}$. The L^p estimation error is measured by reusing the initial quadrature formula,

$$\|f - \widehat{f}\|^p \simeq \sum_{(r_i,\theta_i)\in\chi} \omega_{(r_i,\theta_i)} |f(r_i,\theta_i) - \widehat{f}(r_i,\theta_i)|^p.$$

For each noise level and each norm, the best level J and the best degree K has been selected as the one minimizing the average error over 50 realizations of the noise. Table 1, 2 and 3 display the relative errors

$$\frac{\|f - \hat{f}\|_p}{\|f\|_p}$$

as well as the standard deviation in parenthesis. A dark bold value correspond to an estimator of minimal risk for a given noise level and norm while gray values indicate ties. If the winner wins from more than 20% it is shown with a larger font. The power of two in front of each number gives the optimal maximum degree k^N and k^S used in the chosen estimator. Figures 4, and 5 show typical realizations of the estimators.

Tables 1, 2 and 3 show that, except for the very low noise case and the original Logan Shepp phantom, both the linear SVD estimator and the linear Needlet estimators reduce the error over a naive inversion linear SVD estimate up to the maximal available degree k_0. They also show that the Needlet estimator outperforms the SVD estimator in a large majority of cases from the norm point of view and almost always from the visual point of view. The localization of the needlet also 'localizes' the errors and thus the "simple" smooth regions are much better restored with the needlet estimate than with the SVD because the errors are essentially concentrated along the edges for the needlet.

The choice of the maximum degree is very important to obtain a good estimator. In our proposed scheme, and in the Theorem, this parameter is set by the user according to some expected properties of the unknown function or using some oracle. Nevertheless, an adaptive estimator, which does not require this input, can already be obtained from this family, for example, by using some aggregation technique. A different way to obtain an adaptive estimator based on thresholding is under investigation by some of the authors.

	Inversion $\hat{f}^I = \hat{f}^S_{2^7}$	k^N_{opt}	Needlet $\hat{f}^N_{k^N_{\text{opt}}}$	k^S_{opt}	SVD $\hat{f}^S_{k^S_{\text{opt}}}$
$\boldsymbol{\epsilon = 2}$					
L^1	**1.23e-01 (6.35e-04)**	2^8	**1.25e-01 (6.08e-04)**	2^7	**1.25e-01 (6.08e-04)**
L^2	**9.78e-02 (4.24e-04)**	2^8	**9.78e-02 (4.24e-04)**	2^7	**9.78e-02 (4.24e-04)**
L^3	**8.75e-02 (4.35e-04)**	2^8	**8.75e-02 (4.35e-04)**	2^7	**8 .75e-02 (4.35e-04)**
L^4	**8.59e-02 (6.03e-04)**	2^8	**8.59e-02 (6.03e-04)**	2^7	**8.59e-02 (6.03e-04)**
L^6	**9.66e-02 (1.63e-03)**	2^8	**9.66e-02 (1.63e-03)**	2^7	**9.66e-02 (1.63e-03)**
L^{10}	**1.29e-01 (5.77e-03)**	2^8	**1.29e-01 (5.77e-03)**	2^7	**1.29e-01 (5.77e-03)**
L^∞	2.60e-01 (2.97e-02)	2^7	**1.68e-01 (5.69e-03)**	2^6	2.59e-01 (3.11e-02)
$\boldsymbol{\epsilon = 4}$					
L^1	2.47e-01 (1.27e-03)	2^7	**2.05e-01 (7.16e-04)**	2^6	2.45e-01 (1.25e-03)
L^2	1.96e-01 (8.49e-04)	2^7	**1.68e-01 (4.27e-04)**	2^6	1.95e-01 (8.40e-04)
L^3	1.75e-01 (8.70e-04)	2^7	**1.74e-01 (9.92e-06)**	2^6	1.95e-01 (8.60e-04)
L^4	1.72e-01 (1.21e-03)	2^7	**1.44e-01 (5.91e-04)**	2^6	1.71e-01 (1.22e-03)
L^6	1.93e-01 (3.25e-03)	2^7	**1.43e-01 (9.85e-04)**	2^6	1.92e-01 (3.57e-03)
L^{10}	2.58e-01 (1.15e-02)	2^7	**1.58e-01 (4.90e-03)**	2^6	2.57e-01 (1.26e-02)
L^∞	5.20e-01 (5.95e-02)	2^7	**2.90e-01 (3.63e-02)**	2^5	**3.63e-01 (6.20e-02)**
$\boldsymbol{\epsilon = 8}$					
L^1	4.93e-01 (2.54e-03)	2^6	**3.21e-01 (9.71e-04)**	2^5	3.46e-01 (1.03e-03)
L^2	3.91e-01 (1.70e-03)	2^7	**2.61e-01 (1.02e-03)**	2^5	3.08e-01 (4.43e-04)
L^3	3.50e-01 (1.74e-03)	2^6	**2.93e-01 (5.16e-05)**	2^5	3.08e-01 (6.14e-04)
L^4	3.44e-01 (2.41e-03)	2^7	**2.19e-01 (1.45e-03)**	2^5	2.89e-01 (9.47e-04)
L^6	3.87e-01 (6.51e-03)	2^7	**2.32e-01 (4.38e-03)**	2^5	2.92e-01 (1.57e-03)
L^{10}	5.17e-01 (2.31e-02)	2^7	**2.97e-01 (1.61e-02)**	2^5	**3.07e-01 (2.63e-03)**
L^∞	1.04e+00 (1.19e-01)	2^6	4.77e-01 (9.42e-03)	2^5	**3.99e-01 (1.83e-02)**
$\boldsymbol{\epsilon = 16}$					
L^1	9.87e-01 (5.08e-03)	2^6	**3.92e-01 (2.13e-03)**	2^5	4.75e-01 (2.28e-03)
L^2	7.83e-01 (3.39e-03)	2^6	**3.75e-01 (9.11e-04)**	2^5	3.89e-01 (1.39e-03)
L^3	7.00e-01 (3.48e-03)	2^6	**3.50e-01 (1.34e-04)**	2^5	3.89e-01 (1.55e-03)
L^4	6.88e-01 (4.82e-03)	2^6	3.74e-01 (1.51e-03)	2^5	**3.34e-01 (2.07e-03)**
L^6	7.73e-01 (1.30e-02)	2^6	3.84e-01 (2.49e-03)	2^5	**3.35e-01 (3.84e-03)**
L^{10}	1.03e+00 (4.62e-02)	2^6	**4.08e-01 (4.10e-03)**	2^5	**3.74e-01 (1.58e-02)**
L^∞	2.08e+00 (2.38e-01)	2^6	**5.20e-01 (2.00e-02)**	2^4	**6.24e-01 (8.55e-02)**
$\boldsymbol{\epsilon = 24}$					
L^1	1.48e+00 (7.62e-03)	2^6	**4.73e-01 (3.34e-03)**	2^4	5.05e-01 (2.39e-03)
L^2	1.17e+00 (5.09e-03)	2^6	**4.18e-01 (1.69e-03)**	2^4	4.86e-01 (6.12e-04)
L^3	1.05e+00 (5.22e-03)	2^6	**4.38e-01 (2.52e-04)**	2^4	4.86e-01 (1.01e-03)
L^4	1.03e+00 (7.23e-03)	2^6	**3.92e-01 (2.35e-03)**	2^5	4.16e-01 (3.71e-03)
L^6	1.16e+00 (1.95e-02)	2^6	**4.00e-01 (3.98e-03)**	2^5	4.29e-01 (9.31e-03)
L^{10}	1.55e+00 (6.92e-02)	2^6	**4.30e-01 (8.41e-03)**	2^5	5.29e-01 (3.26e-02)
L^∞	3.12e+00 (3.57e-01)	2^6	**6.51e-01 (9.35e-02)**	2^4	**6.43e-01 (1.27e-01)**
$\boldsymbol{\epsilon = 32}$					
L^1	1.97e+00 (1.02e-02)	2^5	**5.05e-01 (2.78e-03)**	2^4	5.41e-01 (3.13e-03)
L^2	1.57e+00 (6.79e-03)	2^6	**4.72e-01 (2.57e-03)**	2^4	5.04e-01 (1.05e-03)
L^3	1.40e+00 (6.96e-03)	2^5	**4.96e-01 (4.49e-04)**	2^4	5.04e-01 (1.38e-03)
L^4	1.38e+00 (9.64e-03)	2^6	**4.22e-01 (3.33e-03)**	2^4	4.98e-01 (2.22e-03)
L^6	1.55e+00 (2.60e-02)	2^6	**4.31e-01 (6.38e-03)**	2^4	5.13e-01 (3.60e-03)
L^{10}	2.07e+00 (9.23e-02)	2^6	**4.87e-01 (2.37e-02)**	2^4	5.45e-01 (5.44e-03)
L^∞	4.16e+00 (4.76e-01)	2^4	**6.92e-01 (1.47e-02)**	5	**6.73e-01 (4.76e-02)**

Table 1: Logan : Linear Needlet and linear SVD relative estimation error in norm L^1, L^2, L^3, L^4, L^6, L^{10} and L^∞ for several noise level ϵ.

	Inversion $\hat{f}^I = \hat{f}^S_{2^7}$	k^N_{opt}	Needlet $\hat{f}^N_{k^N_{\text{opt}}}$	k^S_{opt}	SVD $\hat{f}^S_{k^S_{\text{opt}}}$
$\boldsymbol{\epsilon = 2}$					
L^1	1.28e-01 (5.68e-04)	2^6	5.73e-02 (2.25e-04)	2^5	**5.11e-02 (3.56e-04)**
L^2	1.18e-01 (3.74e-04)	2^6	5.69e-02 (1.38e-04)	2^5	**4.60e-02 (2.45e-04)**
L^3	1.19e-01 (4.36e-04)	2^6	**4.53e-02 (2.87e-07)**	2^5	4.60e-02 (2.74e-04)
L^4	1.24e-01 (7.16e-04)	2^6	6.07e-02 (2.17e-04)	2^5	**4.57e-02 (3.86e-04)**
L^6	1.42e-01 (2.09e-03)	2^6	6.17e-02 (3.50e-04)	2^5	**5.00e-02 (1.02e-03)**
L^{10}	1.86e-01 (6.85e-03)	2^6	**6.28e-02 (6.48e-04)**	2^5	**6.30e-02 (3.14e-03)**
L^∞	3.48e-01 (3.16e-02)	2^6	**7.67e-02 (7.00e-03)**	2^5	**1.11e-01 (1.26e-02)**
$\boldsymbol{\epsilon = 4}$					
L^1	2.57e-01 (1.14e-03)	2^6	**7.70e-02 (4.99e-04)**	2^5	9.35e-02 (7.63e-04)
L^2	2.35e-01 (7.47e-04)	2^6	**7.14e-02 (3.25e-04)**	2^5	8.51e-02 (5.31e-04)
L^3	2.37e-01 (8.72e-04)	2^6	**8.50e-02 (9.28e-07)**	2^5	**8.51e-02 (5.97e-04)**
L^4	2.47e-01 (1.43e-03)	2^6	**7.11e-02 (4.86e-04)**	2^5	8.74e-02 (8.48e-04)
L^6	2.85e-01 (4.18e-03)	2^6	**7.29e-02 (1.15e-03)**	2^5	9.82e-02 (2.15e-03)
L^{10}	3.73e-01 (1.37e-02)	2^6	**8.42e-02 (4.67e-03)**	2^5	1.25e-01 (6.41e-03)
L^∞	6.95e-01 (6.32e-02)	2^6	**1.46e-01 (1.88e-02)**	2^4	**1.75e-01 (2.63e-02)**
$\boldsymbol{\epsilon = 8}$					
L^1	5.14e-01 (2.27e-03)	2^6	**1.23e-01 (1.09e-03)**	2^5	1.82e-01 (1.57e-03)
L^2	4.71e-01 (1.49e-03)	2^6	**1.12e-01 (7.29e-04)**	2^5	1.67e-01 (1.09e-03)
L^3	4.75e-01 (1.74e-03)	2^6	**1.67e-01 (5.15e-06)**	2^5	**1.67e-01 (1.22e-03)**
L^4	4.94e-01 (2.86e-03)	2^6	**1.12e-01 (1.29e-03)**	2^4	1.67e-01 (6.71e-04)
L^6	5.70e-01 (8.37e-03)	2^6	**1.26e-01 (3.80e-03)**	2^4	1.65e-01 (1.12e-03)
L^{10}	7.46e-01 (2.74e-02)	2^6	**1.65e-01 (1.06e-02)**	2^4	**1.66e-01 (2.09e-03)**
L^∞	1.39e+00 (1.26e-01)	2^5	**2.51e-01 (4.96e-03)**	2^4	**2.13e-01 (2.35e-02)**
$\boldsymbol{\epsilon = 16}$					
L^1	1.03e+00 (4.55e-03)	2^6	**2.25e-01 (2.28e-03)**	2^4	**2.19e-01 (1.53e-03)**
L^2	9.42e-01 (2.99e-03)	2^6	2.05e-01 (1.51e-03)	2^4	**1.99e-01 (8.66e-04)**
L^3	9.49e-01 (3.49e-03)	2^5	**1.94e-01 (3.78e-05)**	2^4	1.99e-01 (1.07e-03)
L^4	9.88e-01 (5.73e-03)	2^6	2.14e-01 (2.86e-03)	2^4	**1.91e-01 (1.53e-03)**
L^6	1.14e+00 (1.67e-02)	2^5	2.40e-01 (1.73e-03)	2^4	**1.93e-01 (3.00e-03)**
L^{10}	1.49e+00 (5.48e-02)	2^5	**2.42e-01 (2.77e-03)**	2^4	**2.22e-01 (1.12e-02)**
L^∞	2.78e+00 (2.53e-01)	2^5	**2.65e-01 (9.50e-03)**	2^4	3.72e-01 (4.31e-02)
$\boldsymbol{\epsilon = 24}$					
L^1	1.54e+00 (6.82e-03)	2^5	**2.56e-01 (1.63e-03)**	2^4	2.65e-01 (2.51e-03)
L^2	1.41e+00 (4.48e-03)	2^5	**2.44e-01 (1.04e-03)**	2^4	**2.39e-01 (1.61e-03)**
L^3	1.42e+00 (5.23e-03)	2^5	**2.34e-01 (4.01e-05)**	2^4	2.39e-01 (1.85e-03)
L^4	1.48e+00 (8.59e-03)	2^5	2.47e-01 (1.76e-03)	2^4	**2.33e-01 (2.54e-03)**
L^6	1.71e+00 (2.51e-02)	2^5	**2.47e-01 (2.80e-03)**	2^4	**2.51e-01 (6.05e-03)**
L^{10}	2.24e+00 (8.22e-02)	2^5	**2.50e-01 (5.09e-03)**	2^4	3.17e-01 (1.81e-02)
L^∞	4.17e+00 (3.79e-01)	2^5	**3.00e-01 (3.24e-02)**	2^3	**3.95e-01 (6.28e-02)**
$\boldsymbol{\epsilon = 32}$					
L^1	2.05e+00 (9.09e-03)	2^5	**2.81e-01 (2.27e-03)**	2^4	3.17e-01 (3.56e-03)
L^2	1.88e+00 (5.98e-03)	2^5	**2.62e-01 (1.55e-03)**	2^4	2.86e-01 (2.40e-03)
L^3	1.90e+00 (6.97e-03)	2^5	**2.82e-01 (6.42e-05)**	2^4	**2.86e-01 (2.68e-03)**
L^4	1.98e+00 (1.15e-02)	2^5	**2.61e-01 (2.47e-03)**	2^4	2.86e-01 (3.66e-03)
L^6	2.28e+00 (3.35e-02)	2^5	**2.61e-01 (4.30e-03)**	2^4	3.22e-01 (9.05e-03)
L^{10}	2.98e+00 (1.10e-01)	2^5	**2.71e-01 (1.10e-02)**	2^3	3.82e-01 (4.00e-03)
L^∞	5.56e+00 (5.06e-01)	2^5	**3.85e-01 (5.17e-02)**	2^3	**4.02e-01 (8.37e-02)**

Table 2: Smoothed Logan : Linear Needlet and linear SVD relative estimation error in norm L^1, L^2, L^3, L^4, L^6, L^{10} and L^∞ for several noise level ϵ.

	Inversion		Needlet		SVD
	$\hat{f}^I = \hat{f}^S_{2^7}$	k^N_{opt}	$\hat{f}^N_{k^N_{\text{opt}}}$	k^S_{opt}	$\hat{f}^S_{k^S_{\text{opt}}}$
$\boldsymbol{\epsilon = 2}$					
L^1	1.29e-01 (5.93e-04)	2^5	**1.75e-02 (1.87e-04)**	2^4	**1.69e-02 (2.38e-04)**
L^2	1.40e-01 (4.24e-04)	2^5	**1.84e-02 (1.27e-04)**	2^4	**1.84e-02 (1.70e-04)**
L^3	1.58e-01 (4.75e-04)	2^5	2.11e-02 (1.63e-08)	2^4	**1.84e-02 (2.18e-04)**
L^4	1.80e-01 (6.83e-04)	2^5	**2.13e-02 (2.06e-04)**	2^4	2.48e-02 (3.91e-04)
L^6	2.34e-01 (1.79e-03)	2^5	**2.48e-02 (4.97e-04)**	2^4	3.44e-02 (9.77e-04)
L^{10}	3.36e-01 (6.82e-03)	2^5	**3.23e-02 (1.35e-03)**	2^4	5.00e-02 (1.96e-03)
L^∞	6.76e-01 (3.37e-02)	2^5	**5.62e-02 (3.82e-03)**	2^4	8.98e-02 (5.48e-03)
$\boldsymbol{\epsilon = 4}$					
L^1	2.58e-01 (1.19e-03)	2^5	**2.50e-02 (3.55e-04)**	2^4	3.30e-02 (4.80e-04)
L^2	2.81e-01 (8.48e-04)	2^5	**2.65e-02 (2.68e-04)**	2^4	3.60e-02 (3.48e-04)
L^3	3.16e-01 (9.51e-04)	2^5	4.11e-02 (7.73e-08)	2^4	**3.60e-02 (4.28e-04)**
L^4	3.61e-01 (1.37e-03)	2^5	**3.21e-02 (4.86e-04)**	2^4	4.78e-02 (6.99e-04)
L^6	4.69e-01 (3.58e-03)	2^5	**4.07e-02 (1.21e-03)**	2^4	6.49e-02 (1.68e-03)
L^{10}	6.72e-01 (1.36e-02)	2^5	**5.69e-02 (2.56e-03)**	2^4	9.40e-02 (3.52e-03)
L^∞	1.35e+00 (6.74e-02)	2^5	**1.01e-01 (7.06e-03)**	2^3	1.42e-01 (1.03e-02)
$\boldsymbol{\epsilon = 8}$					
L^1	5.16e-01 (2.37e-03)	2^5	**4.30e-02 (7.22e-04)**	2^4	6.56e-02 (9.63e-04)
L^2	5.61e-01 (1.70e-03)	2^5	**4.63e-02 (5.46e-04)**	2^4	7.17e-02 (7.02e-04)
L^3	6.33e-01 (1.90e-03)	2^5	8.16e-02 (5.11e-07)	2^4	**7.17e-02 (8.54e-04)**
L^4	7.22e-01 (2.73e-03)	2^5	**5.95e-02 (1.01e-03)**	2^4	9.48e-02 (1.36e-03)
L^6	9.38e-01 (7.16e-03)	2^5	**7.87e-02 (2.42e-03)**	2^3	1.06e-01 (6.22e-04)
L^{10}	1.34e+00 (2.73e-02)	2^5	**1.11e-01 (5.00e-03)**	2^3	**1.17e-01 (1.55e-03)**
L^∞	2.71e+00 (1.35e-01)	2^4	**1.39e-01 (2.64e-03)**	2^3	1.67e-01 (8.80e-03)
$\boldsymbol{\epsilon = 16}$					
L^1	1.03e+00 (4.74e-03)	2^5	**8.19e-02 (1.48e-03)**	2^3	9.66e-02 (8.21e-04)
L^2	1.12e+00 (3.39e-03)	2^5	**8.91e-02 (1.09e-03)**	2^3	9.97e-02 (4.62e-04)
L^3	1.27e+00 (3.80e-03)	2^4	1.05e-01 (3.80e-06)	2^3	**9.97e-02 (5.37e-04)**
L^4	1.44e+00 (5.47e-03)	2^4	1.17e-01 (5.84e-04)	2^3	**1.11e-01 (7.99e-04)**
L^6	1.88e+00 (1.43e-02)	2^4	**1.21e-01 (8.44e-04)**	2^3	**1.23e-01 (1.89e-03)**
L^{10}	2.69e+00 (5.46e-02)	2^4	**1.27e-01 (1.54e-03)**	2^3	1.46e-01 (5.40e-03)
L^∞	5.41e+00 (2.70e-01)	2^4	**1.55e-01 (8.63e-03)**	2^3	2.33e-01 (1.66e-02)
$\boldsymbol{\epsilon = 24}$					
L^1	1.55e+00 (7.12e-03)	2^4	**1.15e-01 (1.04e-03)**	2^3	**1.11e-01 (1.43e-03)**
L^2	1.68e+00 (5.09e-03)	2^4	**1.16e-01 (7.17e-04)**	2^3	**1.15e-01 (8.95e-04)**
L^3	1.90e+00 (5.71e-03)	2^4	1.23e-01 (3.34e-06)	2^3	**1.15e-01 (1.02e-03)**
L^4	2.16e+00 (8.20e-03)	2^4	**1.24e-01 (9.93e-04)**	2^3	1.32e-01 (1.50e-03)
L^6	2.81e+00 (2.15e-02)	2^4	**1.30e-01 (1.59e-03)**	2^3	1.52e-01 (3.62e-03)
L^{10}	4.03e+00 (8.18e-02)	2^4	**1.40e-01 (3.75e-03)**	2^3	1.91e-01 (8.54e-03)
L^∞	8.12e+00 (4.05e-01)	2^4	**1.92e-01 (1.49e-02)**	2^3	3.10e-01 (2.32e-02)
$\boldsymbol{\epsilon = 32}$					
L^1	2.07e+00 (9.49e-03)	2^4	**1.22e-01 (1.48e-03)**	2^3	**1.27e-01 (2.08e-03)**
L^2	2.24e+00 (6.79e-03)	2^4	**1.24e-01 (1.07e-03)**	2^3	1.34e-01 (1.37e-03)
L^3	2.53e+00 (7.61e-03)	2^4	1.45e-01 (5.80e-06)	2^3	**1.34e-01 (1.55e-03)**
L^4	2.89e+00 (1.09e-02)	2^4	**1.34e-01 (1.52e-03)**	2^3	1.57e-01 (2.28e-03)
L^6	3.75e+00 (2.86e-02)	2^4	**1.44e-01 (2.78e-03)**	2^3	1.88e-01 (5.21e-03)
L^{10}	5.38e+00 (1.09e-01)	2^4	**1.63e-01 (6.44e-03)**	2^3	2.42e-01 (1.10e-02)
L^∞	1.08e+01 (5.39e-01)	2^4	**2.39e-01 (1.94e-02)**	2^2	3.61e-01 (2.97e-02)

Table 3: Very Smoothed Logan : Linear Needlet and linear SVD relative estimation error in norm L^1, L^2, L^3, L^4, L^6, L^{10} and L^∞ for several noise level ϵ.

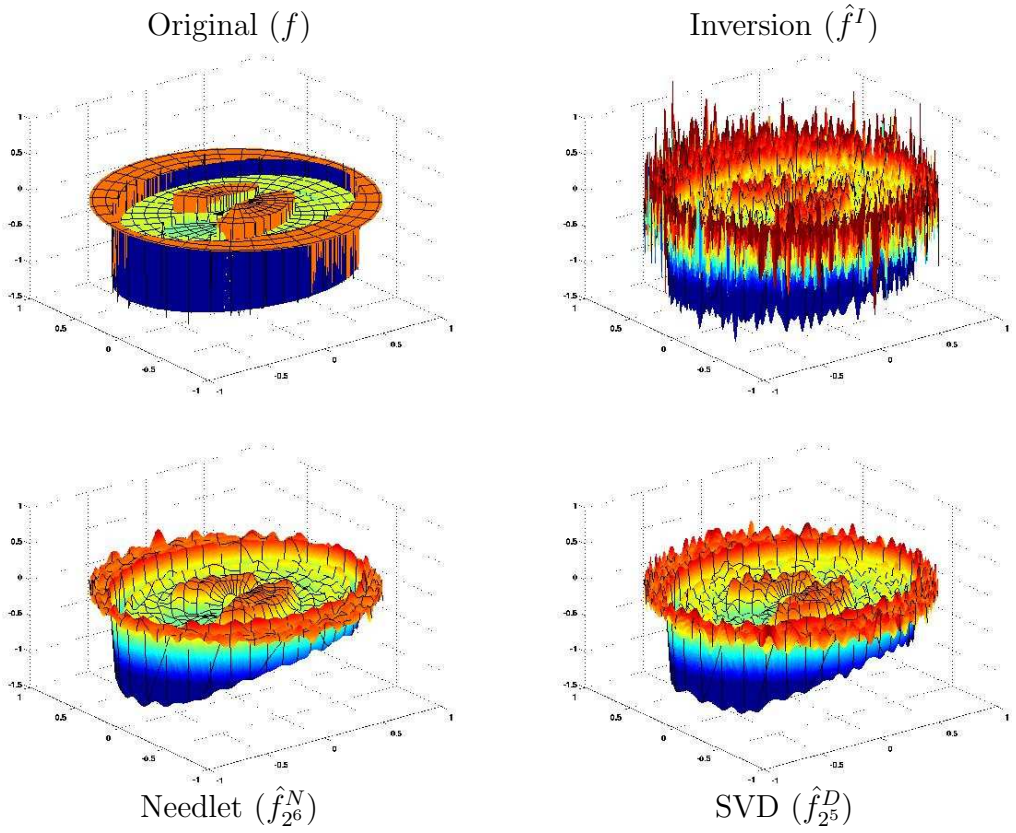

Figure 4: Visual comparison for the original Logan Shepp phantom with $\epsilon = 8$

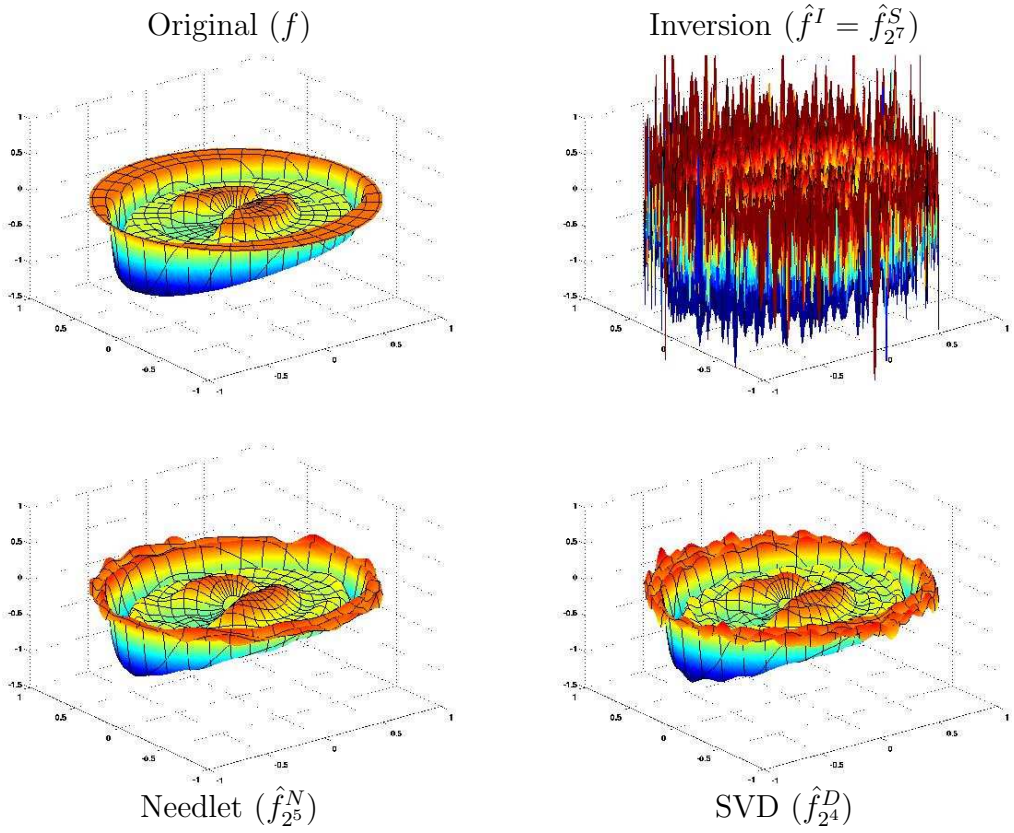

Figure 5: Visual comparison for the smoothed Logan Shepp phantom with $\epsilon = 16$

6 Appendix

6.1 Proof of identity (2.10)

From [17, p. 99] with some adjustment of notation, we have

$$\lambda_k^2 = \frac{|\mathbb{S}^{d-2}|\pi^{(d-1)/2}}{\Gamma\big(\frac{d+1}{2}\big)C_k^{d/2}(1)C_l^{(d-2)/2}(1)}\int_{-1}^{1} C_k^{d/2}(t)C_l^{(d-2)/2}(t)(1-t^2)^{(d-3)/2}dt,$$

where $0 \le l \le k$ and $l \equiv k \pmod 2$. As will be seen shortly λ_k is independent of l.

We will only consider the case $d > 2$ (the case $d = 2$ is simpler, see [17, p. 99]). To compute the above integral we will use the well known identity (cf. [22, (4.7.29)])

$$(n+\lambda)C_n^\lambda(t) = \lambda(C_n^{\lambda+1}(t) - C_{n-2}^{\lambda+1}(t)).$$

Summing up these identities (with indices $n, n-2, \dots$) and taking into account that $C_0^\lambda(t) = 1$, $C_1^\lambda(t) = 2\lambda(t)$, we get

$$C_n^{\lambda+1}(t) = \sum_{j=0}^{\lfloor n/2\rfloor} \frac{n-2j+\lambda}{\lambda} C_{n-2j}^\lambda(t). \tag{6.1}$$

This with $\lambda = (d-2)/2$ and the orthogonality of the polynomials $C_n^{(d-2)/2}(t)$, $n \ge 0$, yield

$$\begin{aligned}\int_{-1}^{1} C_k^{d/2}(t)C_l^{(d-2)/2}(t)(1-t^2)^{(d-3)/2}dt &= \frac{l+\lambda}{\lambda}\int_{-1}^{1}[C_l^{(d-2)/2}(t)]^2(1-t^2)^{(d-3)/2}dt \\ &= \frac{l+\lambda}{\lambda}h_l^{(\lambda)} = \frac{(l+\lambda)2^{1-2\lambda}\pi}{\lambda\Gamma(\lambda)^2}\frac{\Gamma(l+2\lambda)}{(l+\lambda)\Gamma(l+1)}.\end{aligned}$$

We use this and that $C_n^\lambda(1) = \frac{\Gamma(n+2\lambda)}{n!\Gamma(2\lambda)}$ (see §2.3.2) and $|\mathbb{S}^{d-2}| = \frac{2\pi^{(d-1)/2}}{\Gamma((d-1)/2)}$ to obtain

$$\begin{aligned}\lambda_k^2 &= \frac{2\pi^{d-1}}{\Gamma\big(\frac{d-1}{2}\big)\Gamma\big(\frac{d+1}{2}\big)}\frac{k!\Gamma(d)l!\Gamma(d-2)}{\Gamma(k+d)\Gamma(l+d-2)}\frac{2^{4-d}\pi}{(d-2)\Gamma\big(\frac{d-2}{2}\big)^2}\frac{\Gamma(l+d-2)}{\Gamma(l+1)} \\ &= \frac{2^{5-d}\pi^d}{\Gamma\big(\frac{d-1}{2}\big)\Gamma\big(\frac{d+1}{2}\big)}\frac{\Gamma(d)\Gamma(d-2)}{(d-2)\Gamma\big(\frac{d-2}{2}\big)^2}\frac{1}{(k+1)_{d-1}}.\end{aligned} \tag{6.2}$$

The doubling formula for Gamma-function says: $\Gamma(2z) = \frac{2^{2z-1}}{\sqrt{\pi}}\Gamma(z)\Gamma(z+1)$ (see e.g. [22]) and hence

$$\Gamma(d)\Gamma(d-2) = (d-1)(d-2)\Gamma(d-2)^2 = \frac{2^{2(d-3)}}{\pi}(d-1)(d-2)\Gamma\Big(\frac{d-2}{2}\Big)^2\Gamma\Big(\frac{d-1}{2}\Big)^2.$$

We insert this in (6.2) and then a little algebra shows that $\lambda_k^2 = \frac{2^d\pi^{d-1}}{(k+1)_{d-1}}$. □

6.2 Proof of Theorem 3.2

For the proof of estimates (3.23)-(3.24) we first note that by (3.3)

$$\sum_{\xi\in\chi_j}\|h_{j,\xi}\|_p^p \le c2^{jdp/2}2^{-jd}\sum_{\xi\in\chi_j}\frac{1}{\left(2^{-j}+\sqrt{1-|\xi|^2}\right)^{p/2-1}}$$

and we need an upper bound for $\Omega_r := 2^{-jd}\sum_{\xi\in\chi_j}\frac{1}{\left(2^{-j}+\sqrt{1-|\xi|^2}\right)^r}$. To this end, we will use the natural bijection between B^d and $\mathbb{S}^d_+$ considered in the remark in §3.2.2. Thus for $x\in B^d$ we write $\tilde{x}=(x,\sqrt{1-|x|^2})\in\mathbb{S}^d_+$. Let $\tilde{p}=(0,1)$ be the "north pole" of $\mathbb{S}^d$. For $\tilde{\xi}\in\chi_j$ we denote by $B_{\mathbb{S}^d}(\tilde{\xi},\rho)$ is the geodesic ball on $\mathbb{S}^d$ of radius ρ centered at $\tilde{\xi}$, i.e. $B_{\mathbb{S}^d}(\tilde{\xi},\rho):=\{\tilde{x}\in\mathbb{S}^d : d_{\mathbb{S}^d}(\tilde{x},\tilde{p})<\rho\}$, where $d_{\mathbb{S}^d}(\tilde{x},\tilde{p})=\text{Arccos }(\sqrt{1-|x|^2})=\text{Arccos }\langle\tilde{x},\tilde{p}\rangle$ is the geodesic distance between $\tilde{x}$, $\tilde{p}$. Using that $||u|-|\xi||\le|\langle\tilde{u},\tilde{p}\rangle-\langle\tilde{\xi},\tilde{p}\rangle|\le d_{\mathbb{S}^d}(\tilde{\xi},\tilde{u})\le\rho$ for $\tilde{u}\in B_{\mathbb{S}^d}(\tilde{\xi},\rho)$, and $\rho=\tau 2^{-j}\le 2^{-j}$ (see Proposition 3.1), it follows that

$$\frac{1}{2^{-j}+\sqrt{1-|\xi|^2}}\le\frac{2}{2^{-j}+\sqrt{1-|u|^2}}=\frac{2}{2^{-j}+\langle\tilde{u},\tilde{p}\rangle}\quad\forall\tilde{u}\in B_{\mathbb{S}^d}(\tilde{\xi},\rho).$$

On the other hand, we have $|B_{\mathbb{S}^d}(\tilde{\xi},\rho)|=|\mathbb{S}^{d-1}|\int_0^\rho(\sin\theta)^{d-1}d\theta\ge\rho^d|\mathbb{S}^{d-1}|\frac{2^{d-1}}{d\pi^{d-1}}$ with $|\mathbb{S}^{d-1}|=\frac{2\pi^{d/2}}{\Gamma(d/2)}$. We use the above and the fact that the balls $\{B_{\mathbb{S}^d}(\tilde{\xi},\rho)\}_{\xi\in\chi_j}$ are disjoint to obtain

$$\begin{aligned}\Omega_r &\le 2^{-jd}\sum_{\xi\in\chi_j}\frac{1}{|B_{\mathbb{S}^d}(\tilde{\xi},\rho)|}\int_{B_{\mathbb{S}^d}(\tilde{\xi},\rho)}\frac{1}{(2^{-j}+\langle\tilde{u},\tilde{p}\rangle)^r}d\sigma(\tilde{u})\\ &\le c\int_{\mathbb{S}^d_+}\frac{1}{(2^{-j}+\langle\tilde{u},\tilde{p}\rangle)^r}d\sigma(\tilde{u})\le c|\mathbb{S}^{d-1}|\int_0^{\pi/2}\frac{(\sin\theta)^{d-1}}{(2^{-j}+\cos\theta)^r}d\theta\\ &\le c\int_0^{\pi/2}\frac{\sin\theta}{(2^{-j}+\cos\theta)^r}d\theta=c\int_0^1\frac{1}{(2^{-j}+t)^r}dt\le c(d,\tau,r)\int_{2^{-j}}^2 t^{-r}dt.\end{aligned}$$

This yields estimates (3.23)-(3.24).

We now turn to the proof of estimate (3.25). We will employ the maximal operator $\mathcal{M}_t$ ($t>0$), defined by

$$\mathcal{M}_t f(x):=\sup_{B\ni x}\left(\frac{1}{|B|}\int_B|f(y)|^t dy\right)^{1/t},\quad x\in B^d,\tag{6.3}$$

where the sup is over all balls $B\subset B^d$ with respect to the distance $d(\cdot,\cdot)$ from (3.9) containing x. It is easy to show that (see §2.3 in [13]) the Lebesgue measure on B^d is a doubling measure with respect to the distance $d(\cdot,\cdot)$. Hence the general theory

of maximal operators applies. In particular, the Fefferman-Stein vector-valued maximal inequality is valid: If $0 < p < \infty, 0 < q \le \infty$, and $0 < t < \min\{p, q\}$ then for any sequence of functions $\{f_\nu\}_\nu$ on B^d

$$\|(\sum_{\nu=1}^{\infty} |\mathcal{M}_t f_\nu(\cdot)|^q)^{1/q}\|_p \le c\|(\sum_{\nu=1}^{\infty} |f_\nu(\cdot)|^q)^{1/q}\|_p. \tag{6.4}$$

Denote by $B(\xi, r)$ the projection of $B_{\mathbb{S}^d}(\tilde{\xi}, r)$ onto B^d, i.e. $B(\xi, r) := \{x \in B^d : d(x, \xi) < r\}$. By [13, Lemma 2.5], we have

$$(\mathcal{M}_t \mathbb{1}_{B(\xi,r)})(x) \ge c\Big(1 + \frac{d(\xi, x)}{r}\Big)^{-(d+1)/t}, \quad \xi \in B^d,\ 0 < r \le \pi. \tag{6.5}$$

It is easy to see (cf. [13]) that

$$|B(\xi, \rho)| \sim 2^{-jd}(2^{-j} + \sqrt{1 - |\xi|^2}) \sim 2^{-jd} W_j(\xi), \quad \xi \in \chi_j. \tag{6.6}$$

Also, we let $\tilde{\mathbb{1}}_E := \frac{1}{|E|}\mathbb{1}_E$ denote the L^2-normalized characteristic function of $E \subset B^d$. Then (3.21) and (6.6) imply

$$\|h_{j,\xi}\|_p \sim \|\tilde{\mathbb{1}}_{B(\xi,\rho)}\|_p, \quad \xi \in \chi_j. \tag{6.7}$$

Now, pick $0 < t < 1$ and $M > (d + 1)/t$. From (3.21) and (6.5) it follows that

$$|h_{j,\xi}(x)| \le c(\mathcal{M}_t \tilde{\mathbb{1}}_{B(\xi,\rho)})(x), \quad x \in B^d. \tag{6.8}$$

Using this, the maximal inequality (6.4), and (6.7) we obtain

$$\begin{aligned} \|\sum_{\xi\in\chi_j} d_\xi h_{j,\xi}\|_p &\le c\|\sum_{\xi\in\chi_j} \mathcal{M}_t(d_\xi \tilde{\mathbb{1}}_{B(\xi,\rho)})\|_p \le c\|\sum_{\xi\in\chi_j} d_\xi \tilde{\mathbb{1}}_{B(\xi,\rho)}\|_p \\ &\le c\Big(\sum_{\xi\in\chi_j} \|d_\xi \tilde{\mathbb{1}}_{B(\xi,\rho)}\|_p^p\Big)^{1/p} \le c\Big(\sum_{\xi\in\chi_j} \|d_\xi h_{j,\xi}\|_p^p\Big)^{1/p}. \end{aligned}$$

This completes the proof of (3.25). □

References

[1] G. Andew, R. Askey, and R. Roy. *Special Functions.* Cambridge University Press.

[2] L. Cavalier, G. K. Golubev, D. Picard, and A. B. Tsybakov. Oracle inequalities for inverse problems. *Ann. Statist.*, 30(3):843–874, 2002.

[3] L. Cavalier and A. Tsybakov. Sharp adaptation for inverse problems with random noise. *Probab. Theory Related Fields*, 123(3):323–354, 2002.

[4] M. Davison. A singular value decomposition for the radon transform in n-dimensional eucledian space. *Numer. Func. Anal. and Optimiz.*, 3:321–340, 1981.

[5] V. Dicken and P. Maass. Wavelet-Galerkin methods for ill-posed problems. *J. Inverse Ill-Posed Probl.*, 4(3):203–221, 1996.

[6] C. Dunkl and Y. Xu. *Orthogonal polynomials of several variables*, volume 81 of *Cambridge University Press*. Cambridge University Press, New York, 2001.

[7] S. Efromovich and V. Koltchinskii. On inverse problems with unknown operators. *IEEE Trans. Inform. Theory*, 47(7):2876–2894, 2001.

[8] A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi. *Higher Transcendental Functions*, volume 2. McGraw-Hill, New York, 1953.

[9] A. Goldenshluger and S. Pereverzev. On adaptive inverse estimation of linear functionals in Hilbert scales. *Bernoulli*, 9(5):783–807, 2003.

[10] S. Helgason. *The Radon transform*. Birkhäuser, Basel and Boston, 1980.

[11] H. Johnen and K. Scherer. On the equivalence of the k functional and moduli of continuity and some applications. *L.N. Math.*, 571:119–140, 1976.

[12] G. Kerkyacharian, P. Petrushev, D. Picard, and T. Willer. Needlet algorithms for estimation in inverse problems. *Electron. J. Stat.*, 1:30–76 (electronic), 2007.

[13] G. Kyriazis, P. Petrushev, and Y. Xu. Decomposition of weighted triebel-lizorkin and besov spaces on the ball. *Proc. London Math. Soc.*, 2008. to appear
(arXiv:math/0703403).

[14] B. Logan and L. Shepp. Optimal reconstruction of a function from its projections. *Duke Math. J.*, 42(4):645–659, 1975.

[15] A. Louis. Orthogonal function series expansions and the null space of the radon transform. *SIAM J. Math. Anal.*, 15:621–633, 1984.

[16] P. Mathé and S. Pereverzev. Geometry of linear ill-posed problems in variable Hilbert scales. *Inverse Problems*, 19(3):789–803, 2003.

[17] F. Natterer. *The mathematics of computerized tomography*. Classical in Applied mathematics, Philadelphia, PA, 1981.

[18] P. Petrushev. Approximation by ridge functions and neural networks. *SIAM J. Math. Anal.*, 30:155–189, 1999.

[19] P. Petrushev and Y. Xu. Localized polynomials frames on the ball. *Constr. Approx.*, 27:121–148, 2008.

[20] L. A. Shepp and B. F. Logan. The fourier reconstruction of a head section. *IEEE Trans. Nucl. Sci.*, NS-21:21—43, 1974.

[21] E. Stein and G. Weiss. *Introduction to Fourier Analysis on Euclidian spaces*. Princeton University Press.

[22] G. Szegő. *Orthogonal polynomials*. American Mathematical Society, Providence, R.I., 1975.

[23] A. Tsybakov. On the best rate of adaptive estimation in some inverse problems. *C. R. Acad. Sci. Paris Sér. I Math.*, 330(9):835–840, 2000.

[24] Y. Xu. Orthogonal polynomials and cubature formulae on spheres and on balls. *SIAM J. Math. Anal.*, 29(3):779–793 (electronic), 1998.

[25] Y. Xu. Reconstruction from radon projections and orthogonal expansion on ball. 2007. (arXiv: 0705.1984v1).

[26] Y. Xu and O. Tischenko. Fast oped algorithm for reconstruction of images from radon data. 2007. (arXiv: math/0703617v1).

[27] Y. Xu, O. Tischenko, and C. Hoeschen. Image reconstruction by OPED algorithm with averaging. *Numer. Algorithms*, 45(1-4):179–193, 2007.