WEIGHTED ANGLE RADON TRANSFORM: CONVERGENCE RATES AND EFFICIENT ESTIMATION

Daniel Hohmann and Hajo Holzmann

Philipps-University Marburg

Abstract: In the statistics literature, recovering a signal observed under the Radon transform is considered a mildly ill-posed inverse problem. In this paper, we argue that several statistical models that involve the Radon transform lead to an observational design which strongly influences its degree of ill-posedness, and that the Radon transform can actually become severely ill-posed. The main ingredient here is a weight function λ on the angle. Extending results for the limited angle situation, we compute the singular value decomposition of the Radon transform as an operator between suitably weighted L_2-spaces, and show how the singular values relate to the eigenvalues of the sequence of Toeplitz matrices of λ. Further, in the associated white noise sequence model, we give upper and lower bounds on the rate of convergence, and in several special situations even obtain optimal rates with precise minimax constants. For the severely ill-posed limited angle problem, a simple projection estimator is adaptive in the exact minimax sense.

Key words and phrases: Efficient estimation, limited angle problem, minimax estimation, nonparametric estimation, Radon transform.

1. Introduction

Recovering images (functions) observed under the Radon transform is one of the most important and common inverse problems, with fundamental applications in tomography and other fields, see e.g., Natterer (1986) for an overview. The statistics literature has devoted a significant amount of effort to the issue, and considers the problem to be only mildly ill-posed. In this paper we show that the ill-posedness of the Radon transform strongly depends on the observational design, and that observational designs which lead to significantly more severe ill-posedness arise naturally in statistical models involving the Radon transform. We restrict attention to the two-dimensional case, in which the Radon transform is said to be only mildly ill-posed of degree 1/2.

Let $B_1(0) = \{x \in \mathbb{R}^2 : \|x\| \le 1\}$ be the unit disc in $\mathbb{R}^2$ and let $f : B_1(0) \to \mathbb{R}$ be integrable. Then its Radon transform is defined (for almost all (φ, s)) as

$$\mathrm{R}f(\varphi, s) = \int_{|t| \le \sqrt{1-s^2}} f(s\cos\varphi - t\sin\varphi, s\sin\varphi + t\cos\varphi)\, dt,$$

$$(\varphi, s) \in [-\frac{\pi}{2}, \frac{\pi}{2}] \times [-1, 1].$$

We follow Johnstone and Silverman (1990) and call the domain $[-\pi/2, \pi/2] \times [-1, 1]$ of $\mathrm{R}f$ the detector space, and $B_1(0)$ brain space. The aim is to estimate f from noisy data on its Radon transform.

We argue that due to the observational design, the Radon transform needs to be studied as an operator between weighted L_2-spaces

$$
\begin{aligned}
&\mathrm{R} : L_2(B_1(0); \mu_2) \longrightarrow L_2([-\frac{\pi}{2}, \frac{\pi}{2}] \times [-1, 1]; \mu_1), \\
&d\mu_2(x, y) = w_2(x, y)\, dx\, dy, \quad d\mu_1(\varphi, s) = \lambda(\varphi)w_1(s)\, d\varphi\, ds.
\end{aligned}
\tag{1.1}
$$

Here, the most striking feature is the weight function $\lambda : [-\pi/2, \pi/2] \to [0, \infty)$ on the angle in detector space. The case when λ has support $[-\eta, \eta]$ for some $\eta < \pi/2$ is called the limited angle Radon transform (cf. Davison (1983)). However, it turns out that even if λ only has two zeros at the boundary points $\pm\pi/2$, the degree of ill-posedness of R depends on λ. For the weight functions w_1 and w_2, we consider the parametric families in $\gamma > -1/2$,

$$
\begin{aligned}
w_1(s) &= \frac{\sqrt{\pi}\Gamma(\gamma + 1/2)}{\gamma\Gamma(\gamma)} (1 - s^2)^{1/2-\gamma}, \qquad -1 \le s \le 1, \\
w_2(x, y) &= \frac{\pi}{\gamma} (1 - x^2 - y^2)^{1-\gamma}, \qquad (x, y) \in B_1(0).
\end{aligned}
\tag{1.2}
$$

The weight function w_1 in detector space also corresponds to the measurement design, the most important cases being $\gamma = 1$ (fan beam design) and $\gamma = 1/2$ (parallel beam design). The weight w_2 with corresponding γ is then required, for technical reasons, to make the singular value decomposition (SVD) of R analytically tractable. In particular, in the parallel beam design $\gamma = 1/2$, the estimation error in brain space is measured with a weighted L_2-norm. In the following we discuss statistical models that involve the weight function λ, and also indicate the appropriate values of the parameter γ in w_1 and w_2.

Gaussian white noise

This is an idealized statistical model, in which we conduct our convergence analysis. We observe

$$
dY(\varphi, s) = \big(\mathrm{R}f\big)(\varphi, s)\, d\mu_1(\varphi, s) + \varepsilon\, dW(\varphi, s),
\tag{1.3}
$$

which means that for any $h(\varphi, s) \in L_2([-\pi/2, \pi/2] \times [-1, 1]; \mu_1)$, we may observe

$$
\begin{aligned}
Y(h) &= \int_{-\pi/2}^{\pi/2}\!\!\int_{-1}^{1} \mathrm{R}f(\varphi, s)h(\varphi, s)\lambda(\varphi)w_1(s)d\varphi ds + \varepsilon \int_{-\pi/2}^{\pi/2}\!\!\int_{-1}^{1} h(\varphi, s)dW(\varphi, s) \\
&= \langle (\mathrm{R}f), h \rangle_{\mu_1} + \varepsilon\, W(h),
\end{aligned}
\tag{1.4}
$$

where $W(h)$ is a Gaussian field with mean zero and covariance

$$
E\big(W(h_1)W(h_2)\big) = \langle h_1, h_2 \rangle_{\mu_1}, \qquad h_1, h_2 \in L_2([-\frac{\pi}{2}, \frac{\pi}{2}] \times [-1, 1]; \mu_1).
$$
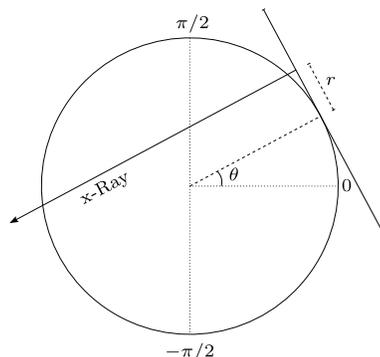
Figure 1. Parametrization of the measurement design in computerized to-
mography: Measurements are performed uniformly distributed on $[-\pi/2,$
$\pi/2] \times [-1, 1]$ in parallel beam design.

For direct observations, regression as well as density estimation problems are
asymptotically equivalent to white noise models under fairly general conditions,
see Nussbaum (1996) and Reiß (2008). While no corresponding results are yet
available for our indirect models, the analysis in the technically less complicated
white noise model still provides valuable insight into the difficulty of the estima-
tion problem.

Regression

Suppose that we observe random variables (Y, Θ, S) from the model

$$Y = (\mathrm{R}f)(\Theta, S) + \epsilon, \qquad \mathrm{E}(\epsilon|\Theta, S) = 0.$$

If (Θ, S) is distributed according to μ_1, then given $h(\varphi, s)$, $\mathrm{E}\big(Yh(\Theta, S)\big) =$
$\langle (\mathrm{R}f), h \rangle_{\mu_1}$, which may be estimated unbiasedly from a sample of observations;
compare to (1.4) in the white noise model for analogy.

This model is the statistical framework for computerized tomography (Nat-
terer, 1986), and the measure μ_1 is determined by the measurement design. The
case $\gamma = 1$ corresponds to the fan beam design, the case $\gamma = 1/2$ to the parallel
beam design, see Figures 1 and 2.

For the fan beam design, most statistical literature uses SVD or derived
methods (such as needlets), see Cavalier and Tsybakov (2002) or Klemelä and
Mammen (2010). In case of parallel beam, Cavalier (1998) uses estimates based
on the filtered back-projection algorithm.

No paper in the statistics literature seems to take into account a weight
function λ on the angle, which arises most naturally in the parallel beam design
in form of a limited angle, e.g., in digital breast tomosynthesis, dental tomography
or electron microscopy, where the measurement device may only be rotated over
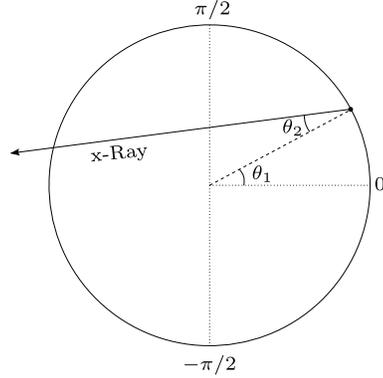
Figure 2. Parametrization of the measurement design in computerized tomography: Measurements are performed uniformly distributed on $[-\pi/2, \pi/2]^2$ in fan beam design.

a limited range. See Frikel (2013) for further references and also for a discussion of the bias of the filtered back-projection algorithm in case of limited angle.

Density estimation

Johnstone and Silverman (1990) propose a model of Positron emission tomography in which the emission density $f(x_1, x_2)$ on $B_1(0)$ needs to be estimated from data (Θ, S) distributed according to $\mathrm{R}f$. Here $\mathrm{E}\big(g(\Theta, S)\big) = \langle (\mathrm{R}f), g \rangle_{\mu_1}$ without weight functions ($\gamma = 1/2$ and $\lambda = 1$). In order to take advantage of the simpler form of the singular value decomposition in case $\gamma = 1$, they insert the weight w_1 with $\gamma = 1$ into $\mathrm{E}\big(g(\Theta, S)w_1(\Theta)\big)$. As a consequence, the variance term in the risk is difficult to handle, and therefore they resort to a surrogate mean integrated squared error in order to measure the precision of their estimators.

Nonparametric random coefficient regression models

Nonparametric estimation in random coefficient regression models was first studied in Beran, Feuerverger, and Hall (1996). These models have recently become quite popular in econometrics, see Hoderlein, Klemelä, and Mammen (2010). Suppose that we observe (Y, X) from the model $Y = X^T \beta$. Here $X, \beta \in \mathbb{R}^2$ are independent random vectors, and the unobserved β has a Lebesgue density f_β supported in $B_1(0)$. The aim is to estimate f_β. If we standardize $Z = Y/\|X\|$, $X/\|X\| = (\cos(\Phi), \sin(\Phi))$, then $f_{Z|\Phi=\varphi}(z) = (\mathrm{R}f_\beta)(\varphi, z)$. Given $h(\varphi, s)$, if Φ has a Lebesgue density f_Φ we have

$$\mathrm{E}\big(h(\Phi, Z)\big) = \int_{-\pi/2}^{\pi/2} \int_{-1}^{1} h(\varphi, z) f_{Z|\Phi=\varphi}(z) \, f_\Phi(\varphi) \, dz \, d\varphi$$

$$= \int_{-\pi/2}^{\pi/2} \int_{-1}^{1} (\mathrm{R}f_\beta)(\varphi, z) \, h(\varphi, z) \, d\mu(\varphi, z),$$
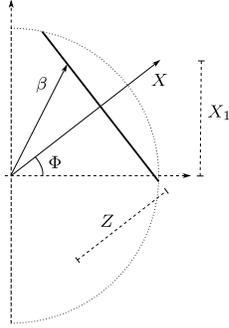
Figure 3. Parametrization in the random coefficient model. The bold line is the set of all $\beta \in B_1(0)$ for which $\beta' X/\|X\| = Z$.

where $d\mu(\varphi, z) = f_\Phi(\varphi)dz\, d\varphi = 2^{-1}\, d\mu_1(\varphi, z)$ with $\lambda(\varphi) = f_\Phi(\varphi)$ and $\gamma = 1/2$. Thus,

$$\mathrm{E}\big(h(\Phi, Z)\big) = 2^{-1}\, \langle (\mathrm{R}f_\beta), h\rangle_{\mu_1}$$

which, in analogy to the white noise model (1.4) may be unbiasedly estimated from a sample in this model.

If $X = (1, X_1)$ includes an intercept as well as an additional covariate, the support of X_1 determines the support of Φ, and in case of full support of X_1 with density f_{X_1}, the tails of X_1 determine the rate of decay of f_Φ at $\pm\pi/2$ since $f_\Phi(\varphi) = f_{X_1}(\tan\varphi)(1 + (\tan\varphi)^2)$. Thus, only for quite heavy tails of X_1 (Cauchy-type tails) is f_Φ bounded away from 0. See Figure 3 for an illustration. Our results show that for lighter tails, the Radon transform R on the weighted L_2-spaces is in fact more ill-posed.

Main results and structure of the paper

As a first main result, we determine the singular value decomposition of the Radon transform R as an operator between the weighted L_2-spaces in (1.1), and show how the singular values relate to the eigenvalues of certain Toeplitz matrices associated with the function λ. If we evaluate the white noise model (1.3) at the singular functions, we obtain a doubly-indexed sequence model. As a next major result, for the limited angle situation, $\lambda = \mathbf{1}_{[-\eta, \eta]}$ for an $\eta < \pi/2$, we show that the optimal rate of estimation over ellipsoidal smoothness classes is only logarithmic, and that a simple projection estimator achieves the optimal rate together with the optimal constant. If the weight function has an isolated zero, we give polynomial upper and lower bounds on the rate of convergence, the order of which depends on the degree of the zero. Finally, for functions λ with finite Fourier expansion, we even obtain optimal rates with precise minimax constants in case of the fan-beam design.

The paper is structured as follows. We start Section 2 by reviewing efficient estimation in general white noise sequence models, and in Section 2.2 we introduce the doubly indexed sequence model for the Radon transform. We discuss ellipsoidal smoothness assumptions, and how the Pinsker estimator applies in this model. In Section 2.3 we present the singular values, while the full derivation of the SVD of the weighted angle Radon transform, together with explicit expressions for the singular functions, is given in the supplementary Appendix B.1. Section 3 turns to nonparametric estimation in the sequence model for the Radon transform. We start in Section 3.1 with the severely ill-posed limited angle problem, in which a simple projection estimator is even sharp minimax adaptive. In Section 3.2 we give upper and lower bounds on the rate of convergence in case of polynomial decay of the singular values, and in Section 3.3 we obtain precise rates with asymptotic minimax constants for the fan-beam design ($\gamma = 1$). Section 4 concludes. Proofs are deferred to the supplementary Appendix A in Hohmann and Holzmann (2015). The derivation of the SVD, discussion of the ellipsoidal smoothness assumptions, as well as some further results can be found in the supplementary Appendix B in Hohmann and Holzmann (2015).

2. Gaussian White Noise Sequence Models

2.1. Review of general infinite white noise sequence models

We start by briefly reviewing some general facts about minimax estimation in infinite white noise sequence models from Cavalier and Tsybakov (2002). Consider observing

$$Y_k = \theta_k + \varepsilon \sigma_k^{-1} \xi_k, \qquad k = 0, 1, 2, \ldots, \tag{2.1}$$

with $(\xi_k)_k$ an i.i.d. Gaussian white noise, $\varepsilon > 0$ the noise level, and $(\sigma_k)_k$ a known sequence of strictly positive weights. The goal is to estimate the parameter $\theta = (\theta_0, \theta_1, \ldots)$ from the noisy observations Y_k. Certainly, estimating θ gets more involved the smaller the weights σ_k are. Asymptotics in this infinite sequence model are w.r.t. $\varepsilon \to 0$.

A *linear estimator* $\hat{\theta} = \hat{\theta}(h)$ of θ is defined as $\hat{\theta}_k = h_k Y_k$ for some given real sequence $h = (h_0, h_1, \ldots)$, not depending on the Y_k. The class of linear estimators thus corresponds to the class of real, countably infinite sequences h. The *mean squared risk* of an estimator $\hat{\theta}$ is defined as

$$R_\varepsilon(\hat{\theta}, \theta) = \mathrm{E}\|\hat{\theta} - \theta\|^2 = \sum_{k=0}^\infty \mathrm{E}\big[(\hat{\theta}_k - \theta_k)^2\big].$$

Define the *linear minimax risk* and the *minimax risk* on a class Θ by

$$r_\varepsilon^L(\Theta) = \inf_{h \in \mathbb{R}^\mathbb{N}} \sup_{\theta \in \Theta} R_\varepsilon(\hat{\theta}(h), \theta), \qquad r_\varepsilon(\Theta) = \inf_{\hat{\theta}} \sup_{\theta \in \Theta} R_\varepsilon(\hat{\theta}, \theta),$$

where $\inf_{\hat{\theta}}$ is the infimum over all possible estimators. An estimator $\hat{\theta}$ is said to be *rate optimal on Θ* if

$$\sup_{\theta \in \Theta} R_{\varepsilon}(\hat{\theta}, \theta) \asymp r_{\varepsilon}(\Theta) \qquad \text{as } \varepsilon \to 0.$$

It is said to be *asymptotically minimax* or *asymptotically efficient on Θ* if

$$\sup_{\theta \in \Theta} R_{\varepsilon}(\hat{\theta}, \theta) \sim r_{\varepsilon}(\Theta) \qquad \text{as } \varepsilon \to 0.$$

The class Θ is typically chosen to be an l_2-ellipsoid: given a constant $L > 0$ and a sequence $a = (a_0, a_1, \ldots)$ of real ellipsoid weights, set

$$\Theta = \Theta(a, L) = \left\{ \theta : \sum_{k=0}^{\infty} a_k^2 \theta_k^2 \leq L \right\}. \tag{2.2}$$

Now let $\Theta = \Theta(a, L)$ be an ellipsoid according to (2.2), and assume that for all $\varepsilon > 0$ there exists a solution c_{ε} to the equation

$$\varepsilon^2 \sum_{k=0}^{\infty} \sigma_k^{-2} a_k (1 - c_{\varepsilon} a_k)_+ = c_{\varepsilon} L, \tag{2.3}$$

where the subscript $+$ denotes positive part, $x_+ = \max\{x, 0\}$. Then, the *Pinsker estimator* is defined as the linear estimator $\hat{\theta}(h^*)$ with weights $h_k^* = (1 - c_{\varepsilon} a_k)_+$, $k = 0, 1, \ldots$.

Theorem 1 (Pinsker (1980)).
a. *The Pinsker estimator $\hat{\theta}(h^*)$ is linear minimax on $\Theta(a, L)$, $\sup_{\theta \in \Theta} R_{\varepsilon}(\hat{\theta}(h^*), \theta)$ $= r_{\varepsilon}^L(\Theta)$ for all $\varepsilon > 0$, where the linear minimax risk is given by*

$$r_{\varepsilon}^L(\Theta) = \varepsilon^2 \sum_{k=0}^{\infty} \sigma_k^{-2} (1 - c_{\varepsilon} a_k)_+. \tag{2.4}$$

b. *If*

$$\frac{\max_{k : a_k < T} \sigma_k^{-2}}{\sum_{k : a_k < T} \sigma_k^{-2}} \longrightarrow 0 \tag{2.5}$$

as $T \to \infty$, then $r_{\varepsilon}(\Theta) \sim r_{\varepsilon}^L(\Theta)$ as $\varepsilon \to 0$.

The condition (2.5) is from Cavalier and Tsybakov (2002). As we shall see below, the Pinsker estimator may also be efficient if this condition is not satisfied.

Remark 1. If the sequence a is monotonically non-decreasing, then there always exists a solution c_{ε} to (2.3) so that the Pinsker estimator is well-defined and

Theorem 1 applies. Even more, in this case c_ε is unique and known to be given by

$$c_\varepsilon = \frac{\sum_{k=0}^{N_\varepsilon} \sigma_k^{-2} a_k}{L/\varepsilon^2 + \sum_{k=0}^{N_\varepsilon} \sigma_k^{-2} a_k^2},$$

where

$$N_\varepsilon = \max\{k : a_k \leq c_\varepsilon^{-1}\} = \max\left\{n : \varepsilon^2 \sum_{k=0}^{n} \sigma_k^{-2} a_k(a_n - a_k) \leq L\right\}, \qquad (2.6)$$

and the minimax risk is attained at $(\hat{\theta}(h^*), \theta^*)$ with

$$\theta_k^* = \frac{\varepsilon}{\sigma_k} \sqrt{\frac{(1 - c_\varepsilon a_k)_+}{c_\varepsilon a_k}}. \qquad (2.7)$$

2.2. The sequence model for the Radon transform

Suppose now that we observe the Radon transform Rf of a function $f \in L_2(B_1(0); \mu_2)$ in the white noise model (1.3).

We require the singular value decomposition of the operator R in (1.1). It consists of triples

$$\left\{\Psi_{m,l}, \Phi_{m,l}, \sigma_{m,l}\right\}_{m \geq l \geq 0}, \qquad (2.8)$$

where the $(\Psi_{m,l})_{m \geq l \geq 0}$ form an orthonormal basis of $L_2(B_1(0); \mu_2)$, the $\{\Phi_{m,l}\}_{m \geq l \geq 0}$ are orthonormal in $L_2([-\pi/2, \pi/2] \times [-1, 1]; \mu_1)$ and complete in range(R), $\sigma_{m,l} > 0$ for all $m \geq l \geq 0$ and $R\Psi_{m,l} = \sigma_{m,l} \Phi_{m,l}$ and $R^*\Phi_{m,l} = \sigma_{m,l} \Psi_{m,l}$, where R^* is the adjoint operator of R, see Proposition B.2. The $(\Psi_{m,l})_{m \geq l \geq 0}$ and $(\Phi_{m,l})_{m \geq l \geq 0}$ are called the singular functions, the $(\sigma_{m,l})_{m \geq l \geq 0}$ the singular values. The singular values are presented in the next section, while the derivation of the SVD together with explicit forms of the singular functions in terms of orthogonal polynomials, is given in the supplementary Appendix B.1.

Evaluating (1.4) at the singular functions $\Phi_{m,l}$, we obtain the doubly indexed sequence of observations

$$Y(\Phi_{m,l}) = \langle Rf, \Phi_{m,l}\rangle_{\mu_1} + \epsilon W(\Phi_{m,l}) = \sigma_{m,l} \theta_{m,l} + \epsilon \xi_{m,l},$$

where $\theta_{m,l} = \langle f, \Psi_{m,l}\rangle_{\mu_2}$ are the Fourier coefficients of f w.r.t. the basis $(\Psi_{m,l})$, and $\xi_{m,l} = W(\Phi_{m,l})$ are independent standard-normal random variables. Now rescale $Y_{m,l} = \sigma_{m,l}^{-1} Y(\Phi_{m,l})$, so that

$$Y_{m,l} = \theta_{m,l} + \varepsilon \sigma_{m,l}^{-1} \xi_{m,l}, \qquad m \geq l \geq 0. \qquad (2.9)$$

Thus, in the doubly indexed sequence model (2.9), ellipsoidal smoothness assumptions on f correspond to the decay of the Fourier coefficients $\theta_{m,l}$ w.r.t. the

basis $(\Psi_{m,l})_{m \geq l \geq 0}$, while rates of convergence depend on the decay of the singular values $\sigma_{m,l}$.

We investigate estimation of θ over the ellipsoids

$$\Theta_1 = \Theta_1(\kappa, L) = \Big\{ \theta : \sum_{m \geq l \geq 0} (m+1)^{2\kappa} \theta_{m,l}^2 \leq L \Big\},$$

$$\Theta_2 = \Theta_2(\kappa, L) = \Big\{ \theta : \sum_{m \geq l \geq 0} (m-l+1)^{2\kappa}(l+1)^{2\kappa} \theta_{m,l}^2 \leq L \Big\}.$$

Compared to (2.2), where a is a full sequence of weights, here we use a slightly different notation in which the parameter κ determines the whole weighting sequence.

Since $m+1 \leq (m-l+1)(l+1) \leq (m+1)^2$ for any $0 \leq l \leq m$,

$$\Theta_1(2\kappa, L) \subset \Theta_2(\kappa, L) \subset \Theta_1(\kappa, L). \tag{2.10}$$

The ellipsoid Θ_2 was proposed by Johnstone and Silverman (1990) in the context of density estimation. Johnstone (1989) shows that in case of $\gamma = 1$ and $\lambda = 1$ it corresponds to a class of functions having 2κ weak derivatives in a weighted L_2-space, see also Proposition B.6 for a more general result. A simpler yet natural choice is the ellipsoid Θ_1.

Remark 2 (*Pinsker estimator for the Radon sequence model*)**.** In order to apply Pinsker's Theorem 1 to these ellipsoids in the doubly-indexed sequence model, we require total orderings $\prec_i$, $i = 1, 2$, of the index set $\{(m,l), \ m \geq l \geq 0\}$, for which the weights in Θ_i are non-decreasing: For Θ_1, we let $(m,l) \prec_1 (\tilde{m}, \tilde{l})$ if $m < \tilde{m}$ or if $m = \tilde{m}$ and $l < \tilde{l}$. Similarly, for Θ_2 we let $(m,l) \prec_2 (\tilde{m}, \tilde{l})$ if $(l+1)(m-l+1) < (\tilde{l}+1)(\tilde{m}-\tilde{l}+1)$ or if there is equality and $l < \tilde{l}$.

2.3. The singular values

In this subsection we present the singular values $\sigma_{m,l}$ in the SVD (2.8) of the Radon transform, see Section B in the supplement for the proofs. Let

$$C_m = \mathrm{diag}(c_{m,0}, \ldots, c_{m,m}), \ \ c_{m,j} = \binom{m}{j} \frac{\Gamma(2\gamma)\Gamma(j+\gamma)\Gamma(m-j+\gamma)}{\Gamma(m+2\gamma)\Gamma(\gamma)^2}, \tag{2.11}$$

$$A_m = \big(d_{j-k}\big)_{j,k=0,\ldots,m}, \qquad m = 0, 1, 2, \ldots, \tag{2.12}$$

which is the Toeplitz matrix determined by the sequence

$$d_z = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-iz\varphi'} \lambda(\frac{\varphi'}{2}) \, d\varphi' = \frac{1}{\pi} \int_{-\pi/2}^{\pi/2} e^{-i2z\varphi'} \lambda(\varphi') \, d\varphi', \qquad z \in \mathbb{Z}.$$

The Toeplitz matrix A_m is Hermitian and positive semidefinite, and it is well known that it is even positive definite whenever λ is not essentially zero (which we always assume), see for instance Tilli (2003) for universal lower bounds on the smallest eigenvalues of sequences of Toeplitz matrices. We denote its (positive) eigenvalues by $\alpha_{m,0} \geq \ldots \geq \alpha_{m,m} > 0$. The matrix $B_m := C_m A_m$ is then also diagonalizable, with strictly positive eigenvalues (see Section B), which we denote by $\beta_{m,0} \geq \ldots \geq \beta_{m,m} > 0$. The singular values of R are given by

$$\sigma_{m,l} = \sqrt{\pi \beta_{m,l}}, \quad m \geq l \geq 0. \tag{2.13}$$

The case $\gamma = 1$ (Fan beam design)

In this case the weights $c_{m,l}$ have the simple form $c_{m,l} = (m+1)^{-1}$ for all m, so, given the eigenvalues $\alpha_{m,l}$ of A_m, it follows that $\beta_{m,l} = \alpha_{m,l}/(m+1)$, and thus the singular values of the operator R are

$$\sigma_{m,l} = \sqrt{\frac{\pi \alpha_{m,l}}{m+1}}, \qquad m \geq l \geq 0. \tag{2.14}$$

Thus, for $\gamma = 1$ the decay of $\sigma_{m,l}$ is determined by the decay of the singular values of the the sequence of Toeplitz matrices A_m generated by the function λ. The asymptotic behavior of the eigenvalues of such sequences of Toeplitz matrices has been intensively studied in the literature. A famous result by Szegö, see Grenander and Szegö (1958), states that the averages of the eigenvalues of A_m tend to the normalized integral of $\lambda(\cdot/2)$. Further results mainly concern the extreme eigenvalues. We present the results that we require in Section 3.

The general case

In the general case, the eigenvalues of B_m cannot be expressed in terms of those of A_m, but it is possible to derive certain bounds. First, concerning the $c_{m,l}$, and using $\Gamma(x+\delta)/\Gamma(x) \sim x^\delta$ as $x \to \infty$ for all $\delta \in \mathbb{R}$, it is easily seen that the inner weights, those for which l grows as pm for some $p \in (0,1)$, decay according to

$$c_{m,l} \sim \frac{\Gamma(2\gamma)}{\Gamma(\gamma)^2}(p(1-p))^{\gamma-1}(m+1)^{-1},$$

while the outer weights with l (or $m-l$) fixed behave like

$$c_{m,l} \sim \frac{\Gamma(2\gamma)}{\Gamma(\gamma)^2}\frac{\Gamma(l+\gamma)}{\Gamma(l+1)}(m+1)^{-\gamma},$$

both as $m \to \infty$. In particular, for $\gamma \leq 1$, the extreme weights satisfy

$$\min_{l=0,\ldots,m} c_{m,l} \sim \frac{\Gamma(2\gamma)}{\Gamma(\gamma)^2 4^{\gamma-1}}(m+1)^{-1}, \qquad \max_{l=0,\ldots,m} c_{m,l} \sim \frac{\Gamma(2\gamma)}{\Gamma(\gamma)}(m+1)^{-\gamma}. \tag{2.15}$$

For $\gamma > 1$ the roles of min and max are reversed.

From these estimates as well as general bounds on the eigenvalues of products of positive definite Hermitian matrices, see for instance Wang and Zhang (1992) and Zhang and Zhang (2006), we obtain the bounds

$$
\frac{\Gamma(2\gamma)}{\Gamma(\gamma)^2 4^{\gamma-1}} \frac{\alpha_{m,m}}{m+1}(1+o(1)) \overset{(\geq)}{\leq} \beta_{m,m} \overset{(\geq)}{\leq} \frac{\Gamma(2\gamma)}{\Gamma(\gamma)} \frac{\alpha_{m,m}}{(m+1)^\gamma}(1+o(1)), \overset{(\gamma>1)}{-\frac{1}{2} < \gamma \leq 1} .
$$
(2.16)

3. Minimax and Efficient Estimation for the Radon Transform

3.1. Limited angle Radon transform

We start with estimation in the limited angle case, where $\lambda = \mathbf{1}_{[-\eta,\eta]}$ for an $\eta < \pi/2$. The Toeplitz matrices A_m generated by this λ are given by

$$
A_m = \left(\frac{\sin(2(j-k)\eta)}{\pi(j-k)} \right)_{j,k=0,\ldots,m}
$$

where, for $j = k$, this expression is understood as the continuous continuation with value $2\eta/\pi$. It is well known that the small eigenvalues of A_m decay to zero exponentially fast, see Slepian (1978), and specifically that

$$
\alpha_{m,m} \sim Cm^{1/2}e^{-\xi m} \qquad \text{as} \quad m \to \infty,
$$
(3.1)

where the constants $C, \xi > 0$ only depend on the angle η, and where

$$
\xi = \log\left(1 + \frac{2\sqrt{1-\cos(\pi-2\eta)}}{\sqrt{2}-\sqrt{1-\cos(\pi-2\eta)}} \right).
$$
(3.2)

Slepian (1978) also discusses the behaviour of the other extreme as well as of the intermediate eigenvalues.

From (2.16) and (2.14) one has exponential decay of $\sigma_{m,m}$ as well, leading to a severely ill-posed inverse problem, see e.g. Mair and Ruymgaart (1996). In this case (2.5) fails to hold and therefore the second part of Pinsker's Theorem 1 as stated above does not apply. Since no general results are available, we start from scratch and give a specifically tailored result for minimax rates in severely ill-posed, doubly indexed sequence models, where in particular the rate of decay of $\sigma_{m,m}$ is only known up to a polynomial factor.

We define the projection estimator $\hat{\theta}(h^{Pr})$ with truncation level M_ϵ as the linear estimator with $h_{m,l} = 1$ for all $0 \leq l \leq m \leq M_\epsilon$, and $h_{m,l} = 0$ otherwise.

Theorem 2. *If there exist $\rho_1, \rho_2 \in \mathbb{R}$ and $\tau_1 \geq \tau_2 > 0$ such that the sequence of smallest singular values $\sigma_{m,m}$ satisfies*

$$
m^{\rho_1}e^{-\tau_1 m} \lesssim \sigma_{m,m} \lesssim m^{\rho_2}e^{-\tau_2 m} \qquad \text{as } m \to \infty,
$$
(3.3)

then $r_\varepsilon(\Theta_i(\kappa, L)) \log(1/\varepsilon)^{2\kappa}(L^{-1} + o(1)) \in [\tau_2^{2\kappa}, \tau_1^{2\kappa}]$ as $\varepsilon \to 0$, $i = 1, 2$. If in particular $\tau_1 = \tau_2 = \tau$, then any projection estimator $\hat{\theta}(h^{Pr})$ with truncation level $M_\varepsilon = \lfloor \tau^{-1} \log(1/\varepsilon)(1 - \log(1/\varepsilon)^{-\delta}) \rfloor$ for some $\delta \in (0, 1)$ is efficient on $\Theta_i(\kappa, L)$, $i = 1, 2$, and the corresponding minimax risk is given by

$$r_\varepsilon(\Theta_i(\kappa, L)) \sim \tau^{2\kappa} L \log(\frac{1}{\varepsilon})^{-2\kappa} \qquad as \ \varepsilon \to 0.$$

This result now provides the minimax rate for the limited angle tomography problem for any $\gamma > -1/2$. Indeed, in view of (3.1) as well as the bound given in (2.16),

$$m^{-1/4}e^{-\xi m/2} \overset{(\gtrsim)}{\sim} \sigma_{m,m} \overset{(\gtrsim)}{\sim} e^{-\xi m/2} m^{1/4-\gamma/2}, \qquad -1 \overset{(\gamma>1)}{<} \gamma \leq 1 .$$

Corollary 1. *For any $\gamma > -1/2$, the limited angle tomography problem with $\eta < \pi/2$ has minimax risk*

$$r_\varepsilon(\Theta_i(\kappa, L)) \sim (\frac{\xi}{2})^{2\kappa} L \log(\frac{1}{\varepsilon})^{-2\kappa} \qquad as \ \varepsilon \to 0, \quad i = 1, 2,$$

where ξ is given in (3.2).

Remark 3. In severely ill-posed problems, the variance is dominated by the bias, even when achieving the optimal constant. Therefore, there are several asymptotically efficient estimators, among them the simple projection estimator.

Remark 4. The projection estimator is asymptotically efficient and does not depend on the parameters κ and L of the smoothness class Θ_i, it is thus adaptive. Since the projection estimator is linear and the Pinsker estimator linear minimax (for fixed ϵ), the Pinsker estimator is also efficient.

Remark 5. Golubev and Khasminskii (1999) also investigate a single indexed sequence model, in which $\sigma_k^{-2} = e^{\alpha k}/k$ for an $\alpha > 0$. They show that the Pinsker estimator is even second order minimax, the second order term being of order $\sim \log \log \epsilon^{-2}/(\log \epsilon^{-2})^{2\kappa+1}$, where the parameter κ corresponds to the smoothness class. Analogous results in our model appear to be difficult, since the singular values are less precisely known.

Finally, we show that the logarithmic rate remains true for general λ (not necessarily an indicator function) which vanishes on an interval at the boundaries.

Corollary 2. *Let the weight function $\lambda : [-\pi/2, \pi/2] \to [0, \infty)$ be Lebesgue measurable and bounded above. If there exist $0 < \eta_1 < \eta_2 < \pi/2$ such that $\inf_{|\varphi| \leq \eta_1} \lambda(\varphi) > 0$ and $\sup_{|\varphi| > \eta_2} \lambda(\varphi) = 0$, then*

$$r_\varepsilon(\Theta_i(\kappa, L)) \log(\frac{1}{\varepsilon})^{2\kappa}(2^{2\kappa} L^{-1} + o(1)) \in [\xi_2^{2\kappa}, \xi_1^{2\kappa}] \qquad as \ \varepsilon \to 0, \quad i = 1, 2,$$

for any $\gamma > -1/2$, where the ξ_j correspond to η_j according to (3.2).

3.2. Weight functions with isolated zeros

In case of a single root of λ (mod π, typically $\pi/2$), the extreme eigenvalues α_{mm} of the sequence of Toeplitz matrices A_m decay polynomially, with degree depending on the order of the root. More precisely, if $\lambda : \mathbb{R} \to \mathbb{R}_+$ is continuous and π-periodic, if there is a unique value φ_0 (mod π) such that $\lambda(\varphi_0) = 0$, and if there exists $\rho > 0$ such that, with $k = k(\rho) = \lfloor \rho/2 \rfloor$, $g(\varphi) = \lambda(\varphi)^{2k/\rho}$ has $2k$ continuous derivatives in some neighborhood of φ_0, and $g^{(2k)}$ is the first non-vanishing derivative of g at φ_0, then there exists $C > 0$ such that $\alpha_{m,m}^{-1} \sim Cm^\rho$, see Parter (1961). For example, for $\lambda = \cos^2$, $\alpha_{m,m}^{-1} \asymp m^2$. By (2.16), this implies polynomial decay of the singular values $\sigma_{m,l}$ as well.

Here is a general result.

Proposition 1.

a. *If there exists $\rho \geq 0$ such that $\beta_{m,m} \gtrsim m^{-\rho}$ as $m \to \infty$, then*

$$r_\varepsilon(\Theta_i(\kappa, L)) = O\big(\varepsilon^{4\kappa/(2\kappa+\rho+2)}\big) \qquad as\ \varepsilon \to 0, \quad i = 1, 2.$$

b. *Let $C > 0$ and $0 \leq \rho_1 \leq \rho < \rho_1 + 1$. If*

$$m^{-\rho} \lesssim \beta_{m,m} \lesssim m^{-\rho_1} \qquad as\ m \to \infty, \tag{3.4}$$

then the Pinsker estimator on $\Theta_i(\alpha, L)$ is asymptotically efficient, and

$$r_\varepsilon(\Theta_i(\alpha, L)) \gtrsim \varepsilon^{(4\kappa+2(\rho-\rho_1))/(2\kappa+\rho+1)} \qquad as\ \varepsilon \to 0, \quad i = 1, 2.$$

c. *If*

$$\beta_{m,m}^{-1} \sim Cm^\rho \qquad as\ m \to \infty, \tag{3.5}$$

$$r_\varepsilon(\Theta_i(\kappa, L)) \geq \tilde{C}\varepsilon^{4\kappa/(2\kappa+\rho+1)}(1 + o(1)) \qquad as\ \varepsilon \to 0, \quad i = 1, 2,$$

where

$$\tilde{C} = \tilde{C}(\kappa, \rho, L, C) = \Big(\frac{C\kappa}{\pi(\kappa + \rho + 1)}\Big)^{2\kappa/(2\kappa+\rho+1)} \frac{(L(2\kappa + \rho + 1))^{(\rho+1)/(2\kappa+\rho+1)}}{\rho + 1}.$$

Remark 6. If the minimal eigenvalue $\alpha_{m,m} \asymp m^{-\tilde{\rho}}$, then from the estimate in (2.16), the condition of a. is satisfied with $\rho = \tilde{\rho} + 1$ in case $-1/2 < \gamma \leq 1$, as well as $\rho = \tilde{\rho} + \gamma$ for $\gamma > 1$. Further, (3.4) is satisfied if $0 < \gamma < 2$, in which case ρ is as before and $\rho_1 = \tilde{\rho} + \gamma$ for $0 < \gamma < 1$, and $\rho_1 = \tilde{\rho} + 1$ otherwise. Finally, for condition (3.5) we require $\gamma = 1$.

3.3. Exact minimax rates and efficiency constants in case $\gamma = 1$

Next we intend to find minimax rates and efficiency constants in case where the minimal eigenvalue $\beta_{m,m}$ and hence the minimal singular value $\sigma_{m,m}$ decays

at a polynomial rate. We require quite precise asymptotics of all singular values $\sigma_{m,l}$, for which, in general, only bounds are available.

Therefore, in this section we restrict ourselves to the case $\gamma = 1$ (fan beam design), so that $\sigma_{m,l} = \sqrt{\pi\alpha_{m,l}/(m+1)}$ as given in (2.14). We impose assumptions on the eigenvalues $\alpha_{m,l}$ of the Toeplitz matrices A_m.

Assumption. There exist $C > 0$ and $\rho \geq 1$ such that

$$\sum_{l=0}^{m} \alpha_{m,l}^{-1} \sim Cm^{\rho-1} \qquad \text{as } m \to \infty. \tag{3.6}$$

Assumption. There exist $\rho \geq 2$, $\delta > 0$, and a positive, bounded sequence $c = (c_0, c_1, \ldots)$ such that

$$\alpha_{m,l}^{-1} = c_{m-l}l^{\rho-1} + O\big(((m-l+1)(l+1))^{\rho-1-\delta}\big), \qquad m \geq l \geq 0. \tag{3.7}$$

Remark 7. We use the exponent $\rho - 1$ instead of ρ since the parameter ρ then corresponds to that of Section 3.2.

First we show that the above conditions are satisfied in certain specific cases. We say that λ is banded if

$$\lambda(\varphi) = \sum_{k=-r}^{r} d_k e^{i2k\varphi}, \qquad r \in \mathbb{N}, \quad d_r \neq 0, \quad \bar{d}_k = d_{-k},$$

since, by construction, the Hermitian Toeplitz matrices A_m generated by λ are banded in this case, and in fact, the coefficients d_k are exactly the entries of A_m. In particular, the condition $\bar{d}_k = d_{-k}$ ensures that λ is real.

Using the results of Böttcher, Grudsky, and Maksimenko (2010) on the uniform behavior of the eigenvalues of banded Toeplitz matrices A_m, we obtain

Proposition 2. *Suppose λ is banded and satisfies $\lambda(-\pi/2) = \lambda(\pi/2) = 0$. Further, assume that there is a unique maximizer φ_0 such that λ is strictly increasing on $(-\pi/2, \varphi_0)$ and strictly decreasing on $(\varphi_0, \pi/2)$, and the second derivatives of λ at $\pi/2$ and φ_0 are non-zero. Then the eigenvalues $\alpha_{m,l}$ satisfy (3.6) with $\rho = 3$ and $C = 4/(3\lambda''(\pi/2))$, as well as (3.7) with $\rho = 3$ and $c_j = (8/\lambda''(\pi/2)\pi^2)(j+1)^{-2}$.*

Linear Minimax risk on Θ_1 under (3.6)

Let $a_{m,l} = (m+1)^\kappa$ be the ellipsoid weights corresponding to $\Theta_1(\kappa, L)$. From (2.6) we have

$$(m,l)_\epsilon = \max\left\{(\tilde{m}, \tilde{l}) : \varepsilon^2 \sum_{(m,l)\prec_1(\tilde{m},\tilde{l})} \sigma_{m,l}^{-2} a_{m,l}(a_{\tilde{m},\tilde{l}} - a_{m,l}) \leq L\right\},$$

where the maximum is taken w.r.t. the total ordering $\prec_1$ defined at the end of Section 2. Since $a_{m,0} = \ldots = a_{m,m}$ for all m, we may include all l for the maximal value of m (since these do not increase the sum). Therefore, $(m,l)_\epsilon = (N_\epsilon, N_\epsilon)$, where

$$N_\varepsilon = \max\left\{n : \varepsilon^2 \sum_{m=0}^{n} \sum_{l=0}^{m} \sigma_{m,l}^{-2} a_{m,l}(a_{n,n} - a_{m,l}) \le L\right\}.$$

By (3.6) we have $\sum_{l=0}^{m} \sigma_{m,l}^{-2} \sim C\pi^{-1} m^\rho$, yielding

$$\sum_{m=0}^{n} \sum_{l=0}^{m} \sigma_{m,l}^{-2} a_{m,l}(a_{n,n} - a_{m,l}) \sim \frac{C}{\pi} \sum_{m=0}^{n} \left(n^\kappa m^{\kappa+\rho} - m^{2\kappa+\rho}\right)$$
$$\sim \frac{C}{\pi} \frac{\kappa}{(\kappa+\rho+1)(2\kappa+\rho+1)} n^{2\kappa+\rho+1}$$

as $n \to \infty$, and thus

$$N_\varepsilon \sim \left(\frac{\pi L(\kappa+\rho+1)(2\kappa+\rho+1)}{C\kappa\varepsilon^2}\right)^{1/(2\kappa+\rho+1)} \qquad \text{as } \varepsilon \to 0.$$

Since $c_\varepsilon \sim N_\varepsilon^{-\kappa}$ by (2.6), and minding that $(1 - c_\varepsilon a_{m,l})_+ = 0$ for $m > N_\varepsilon$, from Pinsker's theorem we obtain

$$\begin{aligned}
r_\varepsilon^L(\Theta_1(\kappa,L)) &\sim \varepsilon^2 \sum_{m=0}^{N_\varepsilon} \sum_{l=0}^{m} \sigma_{m,l}^{-2}(1 - N_\varepsilon^{-\kappa}(m+1)^\kappa) \\
&\sim \frac{C\varepsilon^2}{\pi} \sum_{m=0}^{N_\varepsilon} \left(m^\rho - N_\varepsilon^{-\kappa} m^{\kappa+\rho}\right) \\
&\sim \frac{C\varepsilon^2}{\pi} \frac{\kappa}{(\rho+1)(\kappa+\rho+1)} N_\varepsilon^{\rho+1} \\
&\sim C_1^* \varepsilon^{4\kappa/(2\kappa+\rho+1)}
\end{aligned} \tag{3.8}$$

with $C_1^* = C_1^*(\kappa, \rho, L, C)$ given in Theorem 3 below.

Linear Minimax risk on Θ_2 under (3.7)

In order to simplify calculations, note that the ellipsoid Θ_2 can be rewritten as

$$\Theta_2(\kappa, L) = \left\{\theta : \sum_{j,k\ge 0} (j+1)^{2\kappa}(k+1)^{2\kappa}\theta_{j+k,k}^2 \le L\right\},$$

corresponding to the sequence of ellipsoid weights $a_{j+k,k} = (j+1)^\kappa(k+1)^\kappa$, $j, k \ge 0$. (3.7) then reads

$$\alpha_{j+k,k}^{-1} = c_j k^{\rho-1} + O\left(((j+1)(k+1))^{\rho-1-\delta}\right), \qquad j,k \ge 0. \tag{3.9}$$

Define the totally ordered index sets $(n) = \{(j,k) \in \mathbb{N}_0^2 : (j+1)(k+1) \le n\}$, $n \in \mathbb{N}$. Similarly as above, for the parameter $(j,k)_\epsilon$ in (2.6) we have $\{(j,k) \prec_2 (j,k)_\epsilon\} \cup \{(j,k)_\epsilon\} = (N_\epsilon)$, where

$$N_\varepsilon = \max\Big\{n : \varepsilon^2 \sum_{(j,k)\in(n)} \sigma_{j+k,k}^{-2} a_{j+k,k}\big(n^\kappa - a_{j+k,k}\big) \le L\Big\}.$$

Since $\sigma_{j+k,k}^{-2} = (j+k+1)\pi^{-1}\alpha_{j+k,k}^{-1}$, Lemma A.6 in Section A.3 gives

$$\sum_{(j,k)\in(n)} \sigma_{j+k,k}^{-2} a_{j+k,k}\big(n^\kappa - a_{j+k,k}\big) \sim \frac{K(\rho,c)}{\pi} \frac{\kappa}{(\kappa + \rho + 1)(2\kappa + \rho + 1)} n^{2\kappa+\rho+1}$$

as $n \to \infty$, where

$$K(\rho,c) = \sum_{j=0}^{\infty} c_j(j+1)^{-(\rho+1)}. \tag{3.10}$$

Therefore,

$$N_\varepsilon \sim \Big(\frac{\pi L(\kappa + \rho + 1)(2\kappa + \rho + 1)}{K(\rho,c)\kappa\varepsilon^2}\Big)^{1/(2\kappa+\rho+1)} \qquad \text{as } \varepsilon \to 0,$$

so following the lines in (3.8) and using Lemma A.6, we find that

$$\begin{aligned} r_\varepsilon^L(\Theta_2(\kappa,L)) &= \varepsilon^2 \sum_{(j,k)\in(N_\varepsilon)} \sigma_{j+k,k}^{-2}\big(1 - N_\varepsilon^{-\kappa} a_{j+k,k}\big) \\ &\sim C_2^* \, \varepsilon^{4\kappa/(2\kappa+\rho+1)} \end{aligned} \tag{3.11}$$

with $C_2^* = C_2^*(\kappa,\rho,L,c)$ given in Theorem 3 below.

Asymptotic efficiency on Θ_1 and Θ_2
   Given (3.8) and (3.11), we now easily arrive at

Theorem 3. *For $i = 1,2$, under* (3.6) *and* (3.7), *respectively,*

$$r_\varepsilon(\Theta_i(\kappa,L)) \sim C_i^* \, \varepsilon^{4\kappa/(2\kappa+\rho+1)} \qquad \text{as } \varepsilon \to 0,$$

where

$$C_i^* = \Big(\frac{\Xi_i\kappa}{\pi(\kappa + \rho + 1)}\Big)^{2\kappa/(2\kappa+\rho+1)} \frac{\big(L(2\kappa + \rho + 1)\big)^{(\rho+1)/(2\kappa+\rho+1)}}{\rho + 1},$$

$$\Xi_i = \begin{cases} C, & i = 1, \\ K(\rho,c), & i = 2. \end{cases} \tag{3.12}$$

Example 1. For the ordinary Radon transform, $\lambda = 1$, we have $\sum_{l=0}^{m} \alpha_{m,l}^{-1} = m+1$, whence (3.6) is satisfied for $C = 1$ and $\rho = 2$, leading to the minimax rate

$$r_\varepsilon(\Theta_1(\kappa, L)) \asymp \varepsilon^{4\kappa/(2\kappa+3)} \qquad \text{as } \varepsilon \to 0.$$

On the other hand, Cavalier and Tsybakov (2002) proved that in this case we have

$$r_\varepsilon(\Theta_2(\kappa, L)) \asymp \varepsilon^{4\kappa/(2\kappa+2)} \qquad \text{as } \varepsilon \to 0,$$

so we apparently improve by estimating within the smaller ellipsoid Θ_2. This is no longer true in general, however, when the inverse problem gets more ill-posed. For a banded weight function λ satisfying the assumptions of Proposition 2, Theorem 3 implies that

$$r_\varepsilon(\Theta_i(\kappa, L)) \asymp \varepsilon^{4\kappa/(2\kappa+4)} \qquad \text{as } \varepsilon \to 0$$

for both $i = 1$ and $i = 2$. A slight improvement can only be found for the efficiency constant. Here, $\Xi_1 = 4/(3\lambda''(\pi/2))$ and $\Xi_2 = 8/(\lambda''(\pi/2)\zeta(6)\pi^2)$, where ζ denotes the Riemann zeta function. Thus, $\Xi_1/\Xi_2 = \pi^2\zeta(6)/6 \approx 1.63$.

4. Concluding Remarks

We have shown how the design influences the degree of ill-posedness of the Radon transform in two dimensions, and that the whole range from mildly ill-posedness to severely ill-posedness may arise quite naturally. Without weight on the angle, the rate of convergence remains the same over $\Theta_1(\kappa, L)$ for all parameters $\gamma \in (0, 1]$ (which governs the weight function on the signed distance), see Section B.3 in the supplement, where we also derive the asymptotic minimax constants. In order to avoid computation of the spectral data, iterative methods might be an alternative, see Bissantz et al. (2007). The case of an unknown weight function, for which additional data is available (e.g. in the random coefficients model) leads to a problem with noisy operator, as studied in Hoffmann and Reiß (2008). Finally, we remark that in higher dimensions, injectivity of the limited angle Radon transform as well as the analytic form of its SVD seem not be established.

Supplementary Materials

The online supplement Hohmann and Holzmann (2015) contains the proofs as well as a detailed derivation of the singular value decomposition of the weighted angle Radon transform.

Acknowledgements

References

Beran, R., Feuerverger, A. and Hall, P. (1996). On nonparametric estimation of intercept and slope distributions in random coefficient regression. *Ann. Stat.* **24**, 2569-2592.

Bissantz, N., Hohage, T., Munk, A. and Ruymgaartm F. (2007). Convergence rates of general regularization methods for statistical inverse problems and applications. *SIAM J. Numer. Anal.* **45**, 2610–2636.

Böttcher, A., Grudsky, S. M. and Maksimenko, E. A. (2010). Inside the Eigenvalues of certain Hermitian Toeplitz Band Matrices. *J. Comput. Appl. Math.* **233**, 2245-2264.

Cavalier, L. (1998). Asymptotically efficient estimation in a problem related to tomography. *Math. Methods of Stat.* **7**, 445-456.

Cavalier, L. and Tsybakov, A. B. (2002). Sharp adaptation for inverse problems with random noise. *Probab. Theory Rel.* **123**, 323-354.

Davison, M. E. (1983). The ill-conditioned nature of the limited angle tomography problem. *Siam J. Appl. Math.* **43**, 428-448.

Frikel, J. (2013). Sparse regularization in limited angle tomography. *Appl. Comput. Harmon. Anal.* **34**, 117-141.

Golubev, G. K. and Khasminskii, R. Z. (1999). A statistical approach to some inverse problems for partial differential equations. *Probl. Inf. Transm.* **35**, 136-149.

Grenander, U. and Szegö, G. (1958). *Toeplitz Forms and Their Applications.* Univ. of California Press.

Hoderlein, S., Klemelä, J. and Mammen, E. (2010). Analyzing the Random Coefficient Model Nonparametrically. *Econometric Theory* **26**, 804-837.

Hoffmann, M. and Reiß, M. (2008). Nonlinear estimation for linear inverse problems with error in the operator. *Ann. Stat.* **36**, 310-336.

Hohmann, D. and Holzmann, H. (2015). Weighted angle Radon transform : Convergence rates and efficient estimation. *Statistica Sinica, Supplementary Materials.*

Johnstone, I. M. (1989). On singular value decompositions for the Radon transform and smoothness classes of functions. Dept. Statist., Stanford University, Technical Report.

Johnstone, I. M. and Silverman, B. (1990). Speed of estimation in positron emission tomography and related inverse problems. *Ann. Statist.* **18**, 251-280.

Klemelä, J. and Mammen, E. (2010). Empirical risk minimization in inverse problems. *Ann. Stat.* **38**, 482-511.

Mair, B. A. and Ruymgaart, F. H. (1996). Statistical inverse estimation in Hilbert scales. *SIAM J. Appl. Math.* **56**, 1424-1444.

Natterer, F. (1986). *The Mathematics of Computerized Tomography.* John Wiley & Sons.

Nussbaum, M. (1996). Asymptotic equivalence of density estimation and Gaussian white noise. *Ann. Stat.* **24**, 2399-2430.

Parter, S. V. (1961). On the extreme eigenvalues of toeplitz matrices. *Trans. Amer. Math. Soc.* **100**, 263-276.

Pinsker, M. S. (1980). Optimal filtering of square-integrable signals in Gaussian white noise. *Probl. Peredachi Inf.* **16**, 52-68.

Reiß, M. (2008). Asymptotic equivalence for nonparametric regression with multivariate and random design. *Ann. Stat.* **36**, 1957-1982.

Slepian, D. (1978). Prolate spheroidal wave functions, Fourier analysis, and uncertainty - V: The discrete case. *Bell System Technical J.* **57**, 1371-1430.

Tilli, P. (2003). Universal bounds on the convergence rate of extreme Toeplitz eigenvalues. *Linear Algebra Appl.* **366**, 403-416.

Wang, B. and Zhang, F. (1992). Some inequalities for the eigenvalues of the product of positive semidefinite Hermitian matrices. *Linear Algebra Appl.* **160**, 113-118.

Zhang, F. and Zhang, Q. (2006). Eigenvalue inequalities for matrix product. *IEEE Trans. Automat. Control* **51**, 1506-1509.

Philipps-University Marburg, Deparment of Mathematics and Computer Science, Hans- Meerweinstr., 35032 Marburg, Germany.

E-mail: hohmann@mathematik.uni-marburg.de

Philipps-University Marburg, Deparment of Mathematics and Computer Science, Hans- Meerweinstr., 35032 Marburg, Germany.

E-mail: holzmann@mathematik.uni-marburg.de