THE APPROXIMATE INVERSE IN ACTION II: CONVERGENCE AND STABILITY

ANDREAS RIEDER[†] AND THOMAS SCHUSTER[‡§]

September 10, 2001

Abstract. The approximate inverse is a scheme for constructing stable inversion formulæ for operator equations. Originally, it is defined on L^2-spaces. In the present article we extend the concept of approximate inverse to more general settings which allow us to investigate the discrete version of the approximate inverse which actually underlies numerical computations. Indeed, we show convergence if the discretization parameter tends to zero. Further, we prove stability, that is, we show the regularization property. Finally we apply the results to the filtered backprojection algorithm in 2D-tomography to obtain convergence rates.

Key words. Approximate inverse, mollification, Radon transform, filtered backprojection

AMS subject classifications. 65J10, 65R10

1. Setting the stage. The approximate inverse is a numerical scheme for solving operator equations of the first kind in Hilbert spaces. In this paper we further develop its analytic foundation relying on our results from [14]. Especially, we present an abstract convergence theory which we apply to classical X-ray tomography.

The concept of approximate inverse goes back to the article [9] by Louis and Maaß, see Louis [7, 8] for further information. Originally, it is confined to an L^2-setting. Suppose we want to solve the first kind operator equation

$$Af = g \tag{1.1}$$

where $A : L^2(\Omega) \to Y$ is a linear, injective, and bounded operator mapping square integrable functions over $\Omega \subset \mathbb{R}^\ell$ into the infinite dimensional Hilbert space Y. From a practical point of view we are only able to observe $g \in Y$ at finitely many instances. We model this fact by introducing an observation operator $\Psi_n : Y \to \mathbb{C}^n$. So we only have $g_n = \Psi_n g$ at our hands for solving (1.1). We therefore consider the semi-discrete equation

$$A_n f \;=\; g_n \tag{1.2}$$

with $A_n = \Psi_n A$ which we assume—for the time being—to be continuous.

Problem (1.2) is under-determined and we can only search for its minimum norm solution $f_n^\dagger$:

$$A_n^* A_n f_n^\dagger \;=\; A_n^* g_n \quad \text{ and } \quad f_n^\dagger \in \mathsf{N}(A_n)^\perp \tag{1.3}$$

where $\mathsf{N}(A_n)^\perp$ is the orthogonal complement of the null space of A_n and A_n^* is the adjoint of A_n. If (1.3) is a discrete ill-posed problem, which is very likely in case (1.1) is ill-posed, then the computation of $f_n^\dagger$ is instable. To avoid noise amplification

[†]Institut für Wissenschaftliches Rechnen und Mathematische Modellbildung (IWRMM), Universität Karlsruhe, 76128 Karlsruhe, Germany, email: `andreas.rieder@math.uni-karlsruhe.de`

[‡]Fachbereich Mathematik, Geb. 36, Universität des Saarlandes, 66041 Saarbrücken, Germany, email: `thomas.schuster@num.uni-sb.de`

[§]supported by Deutsche Forschungsgemeinschaft under grant Lo310/4-1

Louis and Maaß [9] suggested to find a smoothed solution f_γ to (1.3) defined by the moments

$$f_\gamma(x) \;=\; E_\gamma f_n^\dagger(x) \;:=\; \langle f_n^\dagger(\cdot), e_\gamma(x,\cdot)\rangle_{L^2(\Omega)}.$$

Here $\{e_\gamma(x,\cdot)\}_{\gamma>0} \subset L^2(\Omega)$, $x \in \Omega$, is a family of *mollifiers*, that is, a family of smooth functions with

$$\lim_{\gamma\to 0} \|E_\gamma w - w\|_{L^2(\Omega)} \;=\; 0 \quad \text{for any } \; w \in L^2(\Omega). \tag{1.4}$$

We can compute $f_\gamma(x)$ from the data g_n by the help of the *reconstruction kernel* $\upsilon_\gamma(x) \in \mathbb{C}^n$ defined as a solution of the normal equation

$$A_n A_n^* \upsilon_\gamma(x) \;=\; A_n e_\gamma(x,\cdot), \tag{1.5}$$

that is,

$$\|A_n^*\upsilon_\gamma(x) - e_\gamma(x,\cdot)\|_{L^2(\Omega)} \;=\; \min\big\{\, \|A_n^* u - e_\gamma(x,\cdot)\|_{L^2(\Omega)} \,\big|\, u \in \mathbb{C}^n \,\big\}.$$

If g_n is in $\mathsf{R}(A_n)$, the range of A_n, then

$$f_\gamma(x) \;=\; \langle g_n, \upsilon_\gamma(x)\rangle_{\mathbb{C}^n}\,,$$

see [14, Lemma 1.1]. The mapping $S_\gamma : g_n \mapsto \langle g_n, \upsilon_\gamma(\cdot)\rangle_{\mathbb{C}^n}$ is called *(semi-discrete) approximate inverse* of A_n. By choosing γ we are able to balance the approximation error and the data error. Hence, γ acts as a *regularization parameter*, see Louis [8].

In [14] we proposed and analyzed a technique for approximating the reconstruction kernel efficiently in case of

- large n, thus avoiding the solution of the densely populated and ill-conditioned system (1.5),
- an unbounded A_n, that is, the approximate inverse is not meaningfully defined.

Let $\upsilon_\gamma^n(x)$ be our substitution for the reconstruction kernel then we replaced $S_\gamma g_n$ by $S_{n,\gamma} g_n(\cdot) = \langle g_n, \upsilon_\gamma^n(\cdot)\rangle_{\mathbb{C}^n}$. Under meaningful assumptions we have been able to prove

$$\lim_{n\to\infty} S_{n,\gamma}\, g_n(x) \;=\; E_\gamma f(x) \qquad \text{for } \; g_n = A_n f \text{ and } \gamma > 0 \text{ fixed},$$

see Corollary 3.8 and Theorem 3.12 in [14].

In the present paper we go one step further. We take into consideration that, in practice, we never compute $f_\gamma(x)$ for all $x \in \Omega$. Instead we select a finite set of points $\{x_1, \ldots, x_d\} \subset \Omega$ in which we evaluate f_γ, that is, we compute the vector $(\Sigma_{n,\gamma} g_n)_i = \langle g_n, \upsilon_\gamma^n(x_i)\rangle_{\mathbb{C}^n}$, $i = 1, \ldots, d$. In the second step we apply an interpolation-like operator $\mathrm{I}_{n,d} : \mathbb{C}^d \to L^2(\Omega)$ to $\Sigma_{n,\gamma} g_n$ such that $\mathrm{I}_{n,d}\Sigma_{n,\gamma} A_n f \approx E_\gamma f$. Moreover, we will couple γ and d with n to establish the convergence

$$\lim_{n\to\infty} \big\| \mathrm{I}_{n,d_n} \Sigma_{n,\gamma_n} A_n f - f \big\|_{L^2(\Omega)} = \;0.$$

In our subsequent investigations we allow a more general setting replacing $L^2(\Omega)$ by an arbitrary real or complex Hilbert space X.

The structure of the paper is as follows. In the next section we set the stage by introducing the technical details. The interpolation-like operator I_{n,d_n} is precisely defined in Section 3 and examples are given. Section 4 contains our main results:

the proofs of the above mentioned convergence and of the regularization property of $\mathrm{I}_{n,d_n}\Sigma_{n,\gamma_n}$. To foster the understanding of the abstract concepts we apply our results to the reconstruction problem in 2D-tomography (Section 5). The approximate inverse then coincides with the filtered backprojection algorithm. Thus, we obtain L^2-convergence with rates and the regularization property of the filtered backprojection algorithm.

To our knowledge only pointwise convergence restricted to a small class of functions has been established before by Popov [12]. We comment in more detail on Popov's results in Remark 5.7 below.

2. The technical set-up. Throughout the paper X and Y denote real or complex infinite dimensional Hilbert spaces and the injective* operator A is in $\mathcal{L}(X,Y)$, the space of linear and bounded mappings from X to Y. Additionally, A is assumed to have the following mapping property

$$A : X_1 \to Y_1 \text{ is continuous} \tag{2.1}$$

where X_1 and Y_1 are Banach spaces such that the embeddings $X_1 \hookrightarrow X$ as well as $Y_1 \hookrightarrow Y$ are continuous, injective, and dense. One can consider X_1 and Y_1 subspaces of X and Y, respectively, which contain "smooth" elements. For instance, if X and Y are L^2-spaces, X_1 and Y_1 could be Sobolev spaces.

Now we define the observation operator $\Psi_n : Y_1 \to \mathbb{C}^n$. Given n functionals $\{\psi_{n,k}\}_{1\le k\le n}$ in Y_1', the dual to Y_1, let

$$(\Psi_n v)_k := \langle \psi_{n,k}, v\rangle_{Y_1'\times Y_1}, \quad k = 1,\ldots,n, \tag{2.2}$$

where $\langle\cdot,\cdot\rangle_{Y_1'\times Y_1}$ is the duality pairing on $Y_1' \times Y_1$.

The semi-discrete operator $A_n := \Psi_n A : \mathsf{D}(A_n) \subset X \to \mathbb{C}^n$ with domain of definition $\mathsf{D}(A_n) \subset X_1$ might be bounded or unbounded. The first is the case if $X = X_1$ (topologically). Here A_n^* exists, hence, the approximate inverse is well defined. Typical examples are integral operators with smooth kernels.

Example 2.1. Let $A : L^2(0,1) \to L^2(0,1)$, $Af(x) := \int_0^1 \mathsf{k}(x,y)\, f(y)\,\mathrm{d}y$, where the kernel k is such that $A : L^2(0,1) \to H^{1/2+\varepsilon}(0,1)$ is bounded for an $\varepsilon > 0$. Thus, $X = X_1 = Y = L^2(0,1)$ and $Y_1 = H^{1/2+\varepsilon}(0,1)$. On the Sobolev space $H^{1/2+\varepsilon}$ point evaluations are continuous functionals, so $\Psi_n g = (g(x_1),\ldots,g(x_n))^t$, $x_i \in]0,1[$, is the right choice if we are able to observe Af at x_i. Thus, $A_n^* w(y) = \sum_i \overline{\mathsf{k}(x_i,y)}\, w_i$ and $(A_n A_n^*)_{i,j} = \int_0^1 \overline{\mathsf{k}(x_i,y)}\, \mathsf{k}(x_j,y)\,\mathrm{d}y$. ♣

It may happen that $A_n : \mathsf{D}(A_n) \subset X \to \mathbb{C}^n$ is unbounded and the concept of approximate inverse does not apply. A prominent example is the Radon transform where $\mathsf{D}(A_n^*) = \{0\}$, see [14, Theorem 5.1].

As already mentioned in §1 both, the bounded and unbounded, cases can be dealt with by extending the concept of approximate inverse. We will shortly describe our technique from [14]. The basic idea is to replace the reconstruction kernel (which may exist or not) by an approximation.

*The injectivity of the operator is not a crucial assumption. It only facilitates the presentation of the material.

Let $\{e_i\}_{1\le i\le d}$, $d \in \mathbb{N}$, be a set of functions in X which we call mollifiers. In our abstract framework e_i plays the role of $e_\gamma(x_i,\cdot)$, $x_i \in \Omega$, from the L^2-environment considered in §1. Due to the injectivity of A the range of A^* is dense in X. Therefore, to any $\varepsilon_i > 0$ we find a $v_i \in Y_1$ (Y_1 is dense in Y!) such that

$$\|e_i - A^* v_i\|_X \le \varepsilon_i, \quad i = 1,\dots,d. \tag{2.3}$$

In Section 3.2 of [14] we demonstrated how to obtain v_i from e_i knowing a singular value decomposition of A. With the v_i's we define a linear mapping $\Sigma_{n,d} : \mathbb{C}^n \to \mathbb{C}^d$ via

$$(\Sigma_{n,d} w)_i := \langle w, G_n \Psi_n v_i\rangle_{\mathbb{C}^n}, \quad i = 1,\dots,d, \tag{2.4}$$

where the $n\times n$-matrix G_n is related to Ψ_n and will be defined below. With the right choice of G_n we have

$$(\Sigma_{n,d} A_n f)_i \approx \langle f, e_i\rangle_X \tag{2.5}$$

(for the precise formulation see Theorem 2.2 below). Hence, we call $G_n\Psi_n v_i$ an *approximate reconstruction kernel* for A_n belonging to the mollifier e_i.

The matrix G_n in (2.4) is the Gramian relative to a family $\{\varphi_k\}_{1\le k\le n}$ in Y which is closely connected to Ψ_n by the operator $\Pi_n : Y_1 \to Y$,

$$\Pi_n v := \sum_{k=1}^{n} (\Psi_n v)_k\, \varphi_k = \sum_{k=1}^{n} \langle \psi_{n,k}, v\rangle_{Y_1'\times Y_1}\, \varphi_k. \tag{2.6}$$

The family $\{\varphi_k\}_{1\le k\le n}$ is required to build a *Riesz system*, that is,

$$\frac{1}{n}\sum_{k=1}^{n} |a_k|^2 \lesssim \Big\|\sum_{k=1}^{n} a_k\,\varphi_k\Big\|_Y^2 \lesssim \frac{1}{n}\sum_{k=1}^{n}|a_k|^2 \quad \text{for all } n \in \mathbb{N}. \tag{2.7}$$

Our notation $A \lesssim B$ indicates the existence of a generic constant $c > 0$ such that $A \le c\,B$. The constant c will not depend on the arguments of A and B. This means that the constants involved in (2.7) do not depend on n. By $A \simeq B$ we abbreviate two-sided inequalities like (2.7).

As a consequence of (2.7) we find a bound for the spectral norm of the Gramian:

$$\|G_n\| \lesssim n^{-1}. \tag{2.8}$$

Furthermore, Π_n is assumed to be *uniformly bounded* in n,

$$\|\Pi_n\|_{Y_1\to Y} \lesssim 1 \quad \text{as } n\to\infty, \tag{2.9}$$

Our second hypothesis on Π_n is the *approximation property* (2.10): let there be a sequence $\{\rho_n\}\subset[0,1]$ converging monotonically to zero such that

$$\|v - \Pi_n v\|_Y \lesssim \rho_n\, \|v\|_{Y_1} \quad \text{for all } v \in Y_1 \text{ as } n\to\infty. \tag{2.10}$$

Now we have all ingredients to give precise meaning to (2.5).

THEOREM 2.2. *Adopt the assumptions specified above in this section, especially, the v_i's are in Y_1. Let f be in X_1 ($X_1 = X$ is permitted). Then, for $1\le i\le d$,*

$$\big|(\Sigma_{n,d}A_n f)_i - \langle f, e_i\rangle_X\big| \lesssim \big(\rho_n\,\|v_i\|_{Y_1} + \varepsilon_i\big)\,\|f\|_{X_1} \quad \text{as } n\to\infty.$$

The constant implicitly involved in the above estimate does not *depend on i, d, and n.*

Proof. See Theorem 3.9 and Section 4 in [14]. ∎

3. The abstract mollifier property and approximate inverse. Having pairs (e_i, v_i), $1 \le i \le d$, satisfying (2.3) we are by now able to compute good approximations to the moments $\langle f, e_i\rangle_X$ of f from the data $g_n = A_n f$. In the remainder of this paper we will investigate how to reconstruct f from its moments. To this end we pose further conditions on the e_i's.

With $\{e_i\}_{1\le i\le d}$ we associate the family $\{b_i\}_{1\le i\le d} \subset X$ by defining the operator $E_d : X \to X$,

$$E_d f \;:=\; \sum_{i=1}^{d} \langle f, e_i\rangle_X \, b_i. \tag{3.1}$$

The crucial condition is the *mollifier property* (3.2) of E_d which establishes the interplay of $\{e_i\}_{1\le i\le d}$ and $\{b_i\}_{1\le i\le d}$:

$$\lim_{d\to\infty} \|E_d f \;-\; f\|_X \;=\; 0. \tag{3.2}$$

The mollifier property of E_d is the abstract analogue of the L^2-convergence (1.4).

The family $\{b_i\}_{1\le i\le d}$ is assumed to have the Riesz system property, that is, a two-sided estimate like (2.7) holds analogously for $\{b_i\}_{1\le i\le d}$ with respect to the X-norm.

Finally we define the *(fully-discrete) approximate inverse* $\widetilde{A}_{n,d} : \mathbb{C}^n \to X$ of A_n by

$$\widetilde{A}_{n,d} w \;:=\; \sum_{i=1}^{d} (\Sigma_{n,d} w)_i \;b_i \;=\; \sum_{i=1}^{d} \langle w, G_n \Psi_n v_i\rangle_{\mathbb{C}^n} \;b_i. \tag{3.3}$$

In view of Theorem 2.2 and (3.2) we expect that

$$\widetilde{A}_{n,d} A_n f \;\approx\; f \qquad \text{for all } f \in X_1 \tag{3.4}$$

which justifies the notation *approximate inverse* for $\widetilde{A}_{n,d}$. We emphasize that the definition of $\widetilde{A}_{n,d}$ depends on the triplets $\{(e_i, v_i, b_i)\}_{1\le i\le d} \subset X \times Y_1 \times X$ bound together by (2.3) and (3.2).

Before we give a precise meaning to (3.4) in the next section we look at an example for E_d possessing the mollifier property.

EXAMPLE 3.1. We first present systems $\{e_i\}_{1\le i\le d}$ and $\{b_i\}_{1\le i\le d}$ in $X = L^2(0,1)$ giving raise to (3.2). Then we discuss how to generalize the given example.

Let the even function $e \in L^2(\mathbb{R})$ have compact support in $[-1,1]$ and a normalized mean-value, that is, $\int e(x)\,\mathrm{d}x = 1$. For $h = 1/d$, $d \in \mathbb{N}$ and $d \ge 2$, we define

$$e_i(x) \;:=\; e(x/h - i)/h, \quad i = 1, \ldots, d-1. \tag{3.5}$$

The definition of the mollifiers e_0 and e_d located at the boundary need special attention. Let e^{b} be a function compactly supported in $[0,1]$ with $\int e^{\mathrm{b}}(x)\,\mathrm{d}x = 1$ and $\int x\, e^{\mathrm{b}}(x)\,\mathrm{d}x = 0$. Then,

$$e_0(x) \;:=\; e^{\mathrm{b}}(x/h)/h \quad \text{ and } \quad e_d(x) \;:=\; e^{\mathrm{b}}(d - x/h)/h. \tag{3.6}$$

Note that supp $e_i = [x_{i-1}, x_{i+1}]$, $i = 1, \ldots, d-1$, supp $e_0 = [0, x_1]$, and supp $e_d = [x_{d-1}, 1]$ where $x_i = h\,i$. Further, $\int e_i(x)\,\mathrm{d}x = 1$, $i = 0, \ldots, d$. See Figure 3.1 for an example of $\{e_i\}_{0\le i\le d}$.

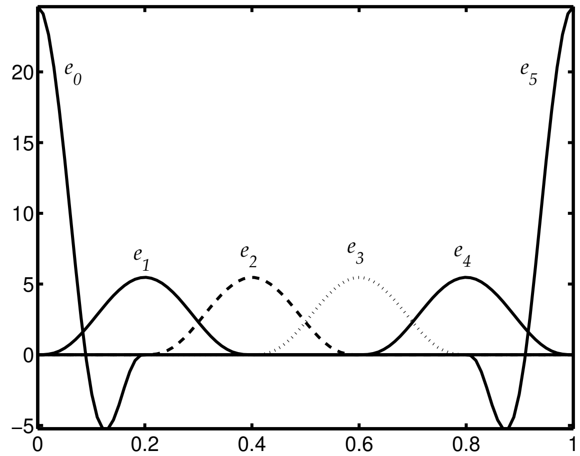

FIG. 3.1. *A set of six mollifiers as discussed in Example* 3.1. *Here,* $d = 5$*. The interior mollifiers* $e_1, \ldots, e_4$ *are obtained by* (3.5) *where* $e(x) = \frac{35}{32}\,(1-x^2)^3$*, for* $|x| \le 1$*, and* $e(x) = 0$*, otherwise. The two boundary mollifiers* e_0 *and* e_5 *are constructed from* $e^b(x) = \frac{1575}{64}\,(1-x^2)^3\,(\frac{1}{5}-x^2)$*, for* $0 \le x \le 1$*, and* $e(x) = 0$*, otherwise, see* (3.6)*.*

Analogously we define the b_i's starting out from the linear B-spline b given by

$$b(x) := \begin{cases} 1-|x| & : \quad |x| \le 1 \\ 0 & : \quad \text{otherwise} \end{cases}\,. \tag{3.7}$$

We set $b_i(x) := b(x/h - i)$, $i = 1, \ldots, d-1$, as well as

$$b_0(x) := \chi_{[0,h]}(x)\, b(x/h) \qquad \text{and} \qquad b_d(x) := \chi_{[1-h,1]}(x)\, b(x/h - d)$$

where χ_J is the indicator function of J. The function $E_d f := \sum_{i=0}^{d} \langle f, e_i\rangle_{L^2(0,1)}\, b_i$ is continuous and affine-linear when restricted to $[x_k, x_{k+1}]$, $k = 0, \ldots, d-1$. Further E_d interpolates the moments at x_k, that is,

$$E_d f(x_k) = \langle f, e_k\rangle_{L^2(0,1)}, \quad k = 0, \ldots, d.$$

In Appendix A we verify the mollifier property of E_d:

$$\lim_{d\to\infty} \|E_d f - f\|_{L^2(0,1)} = 0 \quad \text{for all } f \in L^2(0,1). \tag{3.8}$$

Furthermore we show that

$$\|E_d f - f\|_{L^2(0,1)} \lesssim d^{-\min\{2,s\}}\, \|f\|_{H^s(0,1)} \quad \text{as } d \to \infty \tag{3.9}$$

whenever f is in the Sobolev space $H^s(0,1)$ for $s > 0$. For the definition of $H^s = W_2^s$ see, e.g., Wloka [17].

A careful look at the proof given in Appendix A shows that this example can be extended easily to higher order B-splines. Let $\{b_0, \ldots, b_{N+d-2}\}$ be the B-spline basis of the polynomial spline space of order N with respect to the knot sequence $\{x_0, \ldots, x_d\}$, see, e.g., Schumaker [15, Chap. 4]. Let $E_d f := \sum_{i=0}^{N+d-2} \langle f, e_i\rangle_{L^2(0,1)}\, b_i$. Then, $E_d f \in \mathcal{C}^{N-2}(0,1)$ and $E_d f|_{[x_k,x_{k+1}]}$ is a polynomial of degree $N-1$ at most. If we choose the e_i's with compact support and normalized mean value then E_d reproduces constants. Hence, the mollifier property (3.8) holds. Moreover, if we can find e_i's such that E_d reproduces polynomial of degree up to $j \le N-1$ we even have

$$\|E_d f \;-\; f\|_{L^2(0,1)} \lesssim d^{-\min\{j+1,s\}} \,\|f\|_{H^s(0,1)} \quad \text{as } d \to \infty. \tag{3.10}$$

For instance, considering the B-spline as scaling function we can obtain the mollifiers from a dual scaling function as constructed by Cohen, Daubechies, and Feauveau [2] and by Dahmen, Kunoth, and Urban [3]. In the latter framework (3.10) holds with $j = N-1$. In the multivariate situation we may define E_d using tensor product B-splines. ♣

4. Analysis of convergence and stability. We will first estimate the reconstruction error of the approximate inverse to find a criterion on n and d such that

$$\lim_{\substack{n\to\infty\\ d\to\infty}} \|\widetilde{A}_{n,d} A_n f - f\|_X \;=\; 0. \tag{4.1}$$

Then we explore the stability of the approximate inverse in the presence of perturbed data, in particular, we show its regularization property. The regularization property of the continuous approximate inverse has already been investigated by Louis [8].

4.1. Convergence. We formulate a first result in the following theorem.

THEOREM 4.1. *Let A and Ψ_n be as specified in Section 2. Let the triplets $\{(e_i, v_i, b_i)\}_{1\le i\le d} \subset X \times Y_1 \times X$ such that (2.3) and (3.2) hold true. If $f \in X_1$ then*

$$\|\widetilde{A}_{n,d} A_n f - f\|_X \lesssim \|f - E_d f\|_X \;+\; \Big(\frac{1}{d}\sum_{i=1}^{d}\big(\rho_n^2\,\|v_i\|_{Y_1}^2 \;+\; \varepsilon_i^2\big)\Big)^{1/2} \,\|f\|_{X_1}. \tag{4.2}$$

Proof. We have that

$$\widetilde{A}_{n,d} A_n f - E_d f \;=\; \sum_{i=1}^{d}\big((\Sigma_{n,d} A_n f)_i \;-\; \langle f, e_i\rangle_X\big)\, b_i.$$

Using the Riesz system property of $\{b_i\}$ we estimate

$$\begin{aligned}
\|\widetilde{A}_{n,d} A_n f - E_d f\|_X^2 &\lesssim \frac{1}{d}\sum_{i=1}^{d}\big|(\Sigma_{n,d} A_n f)_i \;-\; \langle f, e_i\rangle_X\big|^2 \\
&\lesssim \frac{1}{d}\sum_{i=1}^{d}\big(\rho_n^2\,\|v_i\|_{Y_1}^2 \;+\; \varepsilon_i^2\big)\,\|f\|_{X_1}^2
\end{aligned}$$

where we used Theorem 2.2 in the last step. Since $\|\widetilde{A}_{n,d} A_n f - f\|_X \le \|\widetilde{A}_{n,d} A_n f - E_d f\|_X + \|f - E_d f\|_X$ we are done with the proof of Theorem 4.1. □

According to the above theorem we have the desired convergence (4.1) provided $d^{-1}\sum_{i=1}^{d}\varepsilon_i^2 \to 0$ as $d \to \infty$ and $\rho_n^2\, d^{-1}\sum_{i=1}^{d}\|v_i\|_{Y_1}^2 \to 0$ as $n, d \to \infty$. In the latter term we have a coupling of the number of data points with the number of reconstruction points. This was, of course, to be expected.

In [14] we proposed a construction scheme for the v_i's from the e_i's satisfying (2.3) arbitrarily accurate. Let us now investigate how Theorem 4.1 reads for these reconstruction kernels. To this end we shortly recall the construction principle. Let $A : X \to Y$ be a compact operator and let $\{(\mathrm{v}_k, u_k; \sigma_k) \mid k \in \mathbb{N}_0\} \subset X \times Y \times]0, \infty[$ be its *singular system*, that is,

$$Ax \;=\; \sum_{k=0}^{\infty} \sigma_k \;\langle x, \mathrm{v}_k\rangle_X \;u_k.$$

The sets of *singular functions* $\{\mathrm{v}_k\}$ and $\{u_k\}$ are orthonormal bases in X (A is injective) and $\overline{\mathsf{R}(A)}$, respectively. The positive numbers σ_k are the *singular values* of A satisfying $\lim_{k\to\infty} \sigma_k = 0$ (monotonically). We assume that all u_k's are in Y_1 and we define

$$v_{i,M} \;:=\; \sum_{k=0}^{M_i-1} \sigma_k^{-1} \;\langle e_i, \mathrm{v}_k\rangle_X \;u_k \tag{4.3}$$

yielding $\|e_i - A^* v_{i,M_i}\|_X \to 0$ as $M_i \to \infty$.

Theorem 4.2. *Let $A : X \to Y$ be compact with singular system $\{(\mathrm{v}_k, u_k; \sigma_k)\}$. Assume that $\sigma_k \simeq (k+1)^{-\gamma}$ for a $\gamma > 0$ as $k \to \infty$ and that $\|u_k\|_{Y_1} \lesssim \sigma_k^{-\beta}$ for a $\beta \geq 0$.*

Adopt the hypotheses of Theorem 4.1. Additionally, let e_i be in $\mathsf{D}\big((A^*A)^{-\alpha}\big)$, the domain of definition of $(A^*A)^{-\alpha}$. Further, let e_i and v_{i,M_i} be related by (4.3).

If $\alpha > (1+\beta)/2 + 1/(4\gamma)$ and $M_i = M_i(n) \gtrsim \rho_n^{-1/(\alpha\gamma)}$ (ρ_n from (2.10)) then

$$\|\widetilde{A}_{n,d} A_n f - f\|_X \;\lesssim\; \|f - E_d f\|_X \;+\; \rho_n \,\Big(\frac{1}{d}\,\sum_{i=1}^{d} \|(A^*A)^{-\alpha} e_i\|_X^2\Big)^{1/2} \;\|f\|_{X_1}.$$

Proof. Theorem 4.2 follows readily from Theorem 4.1 and Theorem 3.12 of [14]. ∎

4.2. Regularization property. Here we allow the data $g_n = \Psi_n A f$ to be contaminated by noise. More precisely we assume that any data point is accurate up to a relative error $\delta > 0$, that is,

$$(\Psi_n^\delta w)_i \;=\; (\Psi_n w)_i \;+\; \delta_i \;\|w\|_{Y_1}, \quad |\delta_i| \leq \delta, \quad i = 1, \ldots, n. \tag{4.4}$$

The following theorem tells us that the approximate inverse is a regularization scheme provided n is properly chosen depending on the noise level δ. Thus, even in the presence of data noise the approximate inverse delivers a reliable reconstruction. We refer, e.g., to Engl, Hanke, and Neubauer [4] and to Louis [6] for an introduction to the regularization of ill-posed problems.

Theorem 4.3. *Under the hypotheses of Theorem 4.1 assume that the triplets $\{(e_i, v_i, b_i)\}_{1\leq i\leq d} \subset X \times Y_1 \times X$ allow a coupling of d with n such that $d = d_n$ diverges to infinity with n and $\lim_{n\to\infty} \rho_n^2\, d_n^{-1} \sum_{i=1}^{d_n} \|v_i\|_{Y_1}^2 = 0$ as well as $\lim_{n\to\infty} d_n^{-1} \sum_{i=1}^{d_n} \varepsilon_i^2 = 0$. (The three conditions on d_n yield convergence of the approximate inverse for unperturbed data and $n \to \infty$, see (4.2).)*

If $n = n_\delta$ is such that n_δ diverges to infinity and $\delta/\rho_{n_\delta} = \mathrm{O}(1)$ as $\delta \to 0$ then

$$\lim_{\delta\to 0} \sup\big\{\|\widetilde{A}_{n_\delta, d_{n_\delta}} w - f\|_X \,\big|\, w = \Psi_{n_\delta}^\delta A f,\; \Psi_{n_\delta}^\delta \text{ satisfies (4.4)}\big\} \;=\; 0$$

for all $f \in X_1$.

Proof. We set $g_n = A_n f$ and $g_n^\delta = \Psi_n^\delta A f$ where Ψ_n^δ is a perturbation of Ψ_n according to (4.4). Using the Riesz system property of the b_i's we find

$$\|\widetilde{A}_{n,d}(g_n - g_n^\delta)\|_X \lesssim d^{-1/2} \, \|\Sigma_{n,d}(\Psi_n - \Psi_n^\delta)Af\|_{\mathbb{C}^n}.$$

Further, by (2.4), (2.8), (2.9), and (2.1),

$$\begin{aligned}\left|\left(\Sigma_{n,d}(\Psi_n - \Psi_n^\delta)Af\right)_i\right| &= \left|\left\langle G_n^{1/2}(\Psi_n - \Psi_n^\delta)Af,\, G_n^{1/2}\Psi_n v_i\right\rangle_{\mathbb{C}^n}\right| \\ &\lesssim n^{-1/2}\, \|(\Psi_n - \Psi_n^\delta)Af\|_{\mathbb{C}^n}\, \|\Pi_n v_i\|_Y \\ &\lesssim \delta\, \|f\|_{X_1}\, \|v_i\|_{Y_1}\end{aligned}$$

which yields

$$\|\widetilde{A}_{n,d}(g_n - g_n^\delta)\|_X \lesssim \delta\, \|f\|_{X_1} \left(\frac{1}{d}\sum_{i=1}^d \|v_i\|_{Y_1}^2\right)^{1/2}.$$

The latter estimate together with (4.2) implies that

$$\begin{aligned}\|\widetilde{A}_{n,d} g_n^\delta - f\|_X &\lesssim \|f - E_d f\|_X \\ &\quad + \left[(\delta + \rho_n)\left(\frac{1}{d}\sum_{i=1}^d \|v_i\|_{Y_1}^2\right)^{1/2} + \left(\frac{1}{d}\sum_{i=1}^d \varepsilon_i^2\right)^{1/2}\right] \|f\|_{X_1}. \quad (4.5)\end{aligned}$$

Replacing n by n_δ and d by d_{n_δ} the above right hand side tends to zero as δ decreases. Thus, Theorem 4.3 is proved. □

5. Convergence of filtered backprojection algorithm for 2D computerized tomography. Computer tomography entails the reconstruction of a density distribution from its integrals along straight lines. There exists a wide area of applications for computerized tomography – the most prominent one being medical imaging.

In the present section we apply our abstract convergence results from the former sections to the reconstruction problem in 2D-tomography. We will earn a rigorous convergence proof of the filtered backprojection algorithm for the parallel scanning geometry. To our knowledge there exists no other convergence proof in the literature.

5.1. Radon transform: definition and smoothing property. The mathematical model for 2D computerized tomography is the Radon transform mapping a function $f \in L^2(\Omega)$ to its line integrals. Here, Ω is the unit ball in $\mathbb{R}^2$ centered about the origin. More precisely,

$$\mathbf{R}f(s,\vartheta) := \int_{L(s,\vartheta)\cap\Omega} f(x)\, d\sigma(x).$$

The lines are parameterized by $L(s,\vartheta) = \{\tau\,\omega^\perp(\vartheta) + s\,\omega(\vartheta) \mid \tau \in \mathbb{R}\}$ where $s \in]-1,1[$, $\omega(\vartheta) = (\cos\vartheta, \sin\vartheta)^t$ and $\omega^\perp(\vartheta) = (-\sin\vartheta, \cos\vartheta)^t$ for $\vartheta \in]0,\pi[$. This parameterization of lines gives rise to the parallel scanning geometry.

The Radon transform maps $L^2(\Omega)$ compactly to $L^2(Z)$, $Z =]-1,1[\times]0,\pi[$, see, e.g., Natterer [10, Chap. IV.3]. In [14, Theorem A.2] we verified the following mapping property:

$$\mathbf{R} : H_0^\alpha(\Omega) \to H^{\alpha+1/2}(Z) \quad \text{is continuous for any} \quad \alpha \ge 0. \quad (5.1)$$

By $H_0^\alpha(\Omega)$ we denote the Sobolev space being the closure of $\mathcal{C}_0^\infty(\Omega)$ with respect to the norm $\|f\|_\alpha^2 = \int_{\mathbb{R}^2} \left(1+\|\xi\|^2\right)^\alpha |\widehat{f}(\xi)|^2\, d\xi$. Here, $\widehat{f}$ is the Fourier transform of f. The space $H^\beta(Z) = W_2^\beta(Z)$ is an L^2-Sobolev space defined on the rectangular domain Z, see, e.g., Wloka [17].

Please note that the Radon transform $\mathbf{R} : L^2(\Omega) \to L^2(Z)$ will play the role of the operator $A : X \to Y$ from our abstract setting. Hence, (5.1) corresponds to the mapping property (2.1).

5.2. Approximate inverse for the Radon transform. In this subsection we provide all ingredients necessary to apply the approximate inverse to the reconstruction of density functions from discrete Radon data. These ingredients are mollifiers and reconstruction kernels, the observation operator Ψ_n, see (2.2), the interpolation-like operator Π_n, see (2.6), and the mollifier operator E_d, see (3.1).

First we introduce observation operators. Let $\ell \in \{1,2\}$ (with ℓ we distinguish two different scenarios) and define

$$s_i = i\, h_s, \qquad h_s = 1/q, \qquad i = -q,\dots,q_\ell, \tag{5.2a}$$

$$\vartheta_j = j\, h_\vartheta, \qquad h_\vartheta = \pi/p, \qquad j = 0,\dots,p_\ell, \tag{5.2b}$$

where $(q,p \in \mathbb{N})$

$$q_1 = q-1, \quad p_1 = p-1, \quad q_2 = q, \quad p_2 = p.$$

If $\alpha > 1/2$ then point evaluations are stable operations on $H^{\alpha+1/2}(Z)$. Therefore, we define the bounded operator

$$\Psi_{q,p}^{(\ell)} : H^{\alpha+1/2}(Z) \to \mathbb{R}^{n_\ell} \quad \text{by} \quad \left(\Psi_{q,p}^{(\ell)} y\right)_{i,j} := y(s_i,\vartheta_j) \tag{5.3}$$

where $n_\ell = (q+q_\ell+1)(p_\ell+1)$. The 2D-reconstruction problem now reads $(\alpha > 1/2)$:

$$\text{given } g_{q,p} \in \mathbb{R}^{n_\ell} \text{ find } f \in H_0^\alpha(\Omega) \text{ such that } \Psi_{q,p}^{(\ell)}\mathbf{R} f = g_{q,p}.$$

For applying the approximate inverse to the reconstruction problem we consider $\Psi_{q,p}^{(\ell)}\mathbf{R}$ in the natural L^2-topology in which $\mathbf{R}$ is bounded. Unfortunately the L^2-boundedness of $\mathbf{R}$ does not carry over to the semi-discrete Radon transform, see [14, Theorem 5.1]:

LEMMA 5.1. *The semi-discrete Radon transform*

$$\Psi_{q,p}^{(\ell)}\mathbf{R} : H_0^\alpha(\Omega) \subset L^2(\Omega) \to \mathbb{R}^{n_\ell}$$

is unbounded *for any* $\alpha > 1/2$.

Canonical candidates for the approximation spaces related to $\Psi_{q,p}^{(\ell)}$ are tensor product spline spaces $V_{q,p}^{(\ell)} = S_s^{(\ell)} \otimes S_\vartheta^{(\ell)}$. The univariate spaces $S_s^{(\ell)}$ and $S_\vartheta^{(\ell)}$ are the piecewise constant ($\ell = 1$) or linear ($\ell = 2$) spline spaces with respect to the knot sequences $\{s_i\}$ and $\{\vartheta_j\}$ respectively. We equip $V_{q,p}^{(\ell)}$ with the tensor product B-spline basis

$$\left\{B_{q,i}^{(\ell)} \otimes B_{p,j}^{(\ell)} \,\middle|\, i,j \text{ as in (5.2)}\right\} \tag{5.4}$$

For instance we have that

$$B_{p,j}^{(1)} = \chi_{[\vartheta_j,\vartheta_{j+1}[} \quad \text{and} \quad B_{q,i}^{(2)}(s_k) = \begin{cases} 1 & : \quad i = k \\ 0 & : \quad \text{otherwise} \end{cases}$$

where χ_I is the indicator function of the interval I. The other basis functions are defined accordingly. The interpolation operator $\Pi^{(\ell)}_{q,p} : H^\alpha(Z) \to V_{q,p}$, $\alpha > 1$,

$$\Pi^{(\ell)}_{q,p} y \;:=\; \sum_{i=-q}^{q_\ell} \sum_{j=0}^{p_\ell} y(s_i, \vartheta_j)\; B^{(\ell)}_{q,i} \otimes B^{(\ell)}_{p,j},$$

satisfies the uniform boundedness

$$\|\Pi^{(\ell)}_{q,p} y\|_{L^2(Z)} \;\lesssim\; \|y\|_{H^\alpha(Z)}$$

as well as the approximation property

$$\|\Pi^{(\ell)}_{q,p} y - y\|_{L^2(Z)} \;\lesssim\; h^{\min\{\alpha,\ell\}}\; \|y\|_{H^\alpha(Z)}, \quad h := \max\{h_s, h_\vartheta\}. \tag{5.5}$$

Both latter estimates may be proved along the lines presented by Schumaker [15, Chap. 12].

We define the mollifier operator $E_d : L^2(\Omega) \to L^2(\Omega)$ by

$$E_d f(x) \;:=\; \sum_{k \in \mathfrak{J}_d} \langle f, e_{d,k} \rangle_{L^2(\Omega)}\; B(dx - k) \tag{5.6}$$

where B is the tensor product linear B-spline: $B = b \otimes b$, see (3.7). Further, $\mathfrak{J}_d = \{k \in \mathbb{Z}^2 \,|\, \operatorname{supp} B(d\,\cdot - k) \cap \Omega \neq \emptyset\}$. The mollifiers used in defining E_d are scaled and translated versions of the mollifier e:

$$e_{d,k}(x) \;=\; \mathfrak{T}^{d,k}_1 e(x) \;:=\; d^2\; e(dx - k).$$

For e being radially symmetric with compact support in Ω we have the mollifier property (see Appendix B)

$$\lim_{d \to \infty} \|E_d f - f\|_{L^2(\Omega)} \;=\; 0 \quad \text{for all } \; f \in L^2(\Omega) \tag{5.7}$$

as well as the estimate

$$\|E_d f - f\|_{L^2(\Omega)} \lesssim d^{-\min\{2,\alpha\}}\; \|f\|_{H^\alpha(\Omega)} \quad \text{for all } \; f \in H^\alpha_0(\Omega), \; \alpha > 0. \tag{5.8}$$

Now let $v \in L^2(Z)$ be the solution of $\mathbf{R}^* v = e$. Since e is radially symmetric, v does not depend on the angle ϑ. For special mollifiers, v is well defined and explicitly known, see Example 5.2 below. Furthermore, v is a reconstruction kernel for $\mathbf{R}$ belonging to e, compare (2.3).

Example 5.2. We give concrete examples of mollifier/reconstruction kernel pairs for the Radon transform. These and other examples can be found in [13]. We define a family $\{e^n\}_{n>0}$ of radial mollifiers compactly supported in Ω:

$$e^n(x) \;:=\; \frac{n+1}{\pi} \begin{cases} (1 - \|x\|^2)^n & : \quad \|x\| \le 1 \\ 0 & : \quad \text{otherwise} \end{cases} \;\in\; H^s_0(\Omega) \quad \text{for any } \; s < n + 1/2.$$

Here, the equation $\mathbf{R}^* v^n = e^n$ is solved by

$$v^n(s) \;=\; \frac{1}{2\pi^2} \begin{cases} 2(n+1)\; {}_2\mathrm{F}_1(1,\, -n;\, 1/2;\, s^2) & : \quad |s| \le 1 \\ -\; {}_2\mathrm{F}_1(1,\, 3/2;\, n+2;\, 1/s^2)/s^2 & : \quad |s| > 1 \end{cases}$$

where ${}_2\mathrm{F}_1$ is the *hypergeometric series.* ♣

We have the invariance property

$$\mathcal{T}_1^{d,k}\mathbf{R}^* = \mathbf{R}^*\mathcal{T}_2^{d,k} \quad \text{for } k \in \mathbb{Z}^2 \text{ with } \|k\| \le d-1\,^\dagger$$

where the translation-dilation operator $\mathcal{T}_2^{d,k}$ is given by

$$\mathcal{T}_2^{d,k}w(s,\vartheta) := d^2\; w\big(d\,s - k^t\,\omega(\vartheta),\vartheta\big).$$

Due to the invariance property the reconstruction kernel $\upsilon_{d,k}$, which belongs to $e_{d,k}$, can be computed from υ by applying $\mathcal{T}_2^{d,k}$:

$$\upsilon_{d,k} = \mathcal{T}_2^{d,k}\upsilon \quad \Longrightarrow \quad \mathbf{R}^*\upsilon_{d,k} = e_{d,k} \;\text{ for }\; \|k\| \le d-1.$$

For later use we record a continuity result of $\mathcal{T}_2^{d,k}$. Its proof will be supplied in Appendix C.

LEMMA 5.3. *Let $Z^{d,k} = \big\{\big(d\,s - k^t\,\omega(\vartheta),\,\vartheta\big)\,\big|\,(s,\vartheta) \in Z\big\}$ and let $\|k\| \le d$. Then,*

$$\big\|\mathcal{T}_2^{d,k}g\big\|_{H^\kappa(Z)} \lesssim d^{\kappa+3/2}\;\|g\|_{H^\kappa(Z^{d,k})}$$

whenever the right hand side is defined for $\kappa \ge 0$.

After these preparations we are able to define the approximate inverse $\widetilde{\mathbf{R}}^{(\ell)}_{n_\ell,d}$: $\mathbb{R}^{n_\ell} \to L^2(\Omega)$ of $\Psi^{(\ell)}_{q,p}\mathbf{R}$, see (3.3):

$$\widetilde{\mathbf{R}}^{(\ell)}_{n_\ell,d}w(x) := \sum_{\substack{k\in\mathbb{Z}^2\\ \|k\|\le d-1}} \big(\Sigma^{(\ell)}_{n_\ell,d}w\big)_k\; B(d\,x-k), \quad x \in \Omega,$$

where

$$\big(\Sigma^{(\ell)}_{n_\ell,d}w\big)_k := \big\langle w,\, G^{(\ell)}_{q,p}\Psi^{(\ell)}_{q,p}\mathcal{T}_2^{d,k}\upsilon\big\rangle_{\mathbb{R}^{n_\ell}}$$

with $G^{(\ell)}_{q,p} \in \mathbb{R}^{n_\ell\times n_\ell}$ being the Gramian matrix with respect to the spline basis (5.4) ($G^{(1)}_{p,q}$ is a multiple of the identity matrix). Please note that $\Sigma^{(\ell)}_{n_\ell,d}w$ can be evaluated by an algorithm of filtered backprojection type, see, e.g., Natterer [10, Chap. V.1].

In the remainder of this section we will show a convergence result for $\widetilde{\mathbf{R}}^{(\ell)}_{n_\ell,d}$ which is analogous to Theorem 4.1, see Theorem 5.5 below. We start by presenting the version of Theorem 2.2 in the Radon transform framework.

LEMMA 5.4. *Let $\alpha > 1/2$ and let f be in $H_0^\alpha(\Omega)$. The radially symmetric mollifier e is assumed to be in $H_0^{\alpha+1}(\Omega)$. Then,*

$$\big|\big(\Sigma^{(\ell)}_{n_\ell,d}\Psi^{(\ell)}_{q,p}\mathbf{R}f\big)_k - \langle f,\,e_{d,k}\rangle_{L^2(\Omega)}\big| \lesssim h^{\min\{\alpha+1/2,\ell\}}\;d^{\alpha+2}\;\|f\|_{H^\alpha(\Omega)}\;\|\upsilon\|_{H^{\alpha+1/2}(\mathbb{R})}$$

for $k \in \mathbb{Z}^2$ with $\|k\| \le d-1$. Above, h is as in (5.5).

Proof. We apply Theorem 2.2 with $\varepsilon_i = 0$. Thus,

$$\big|\big(\Sigma^{(\ell)}_{n_\ell,d}\Psi^{(\ell)}_{q,p}\mathbf{R}f\big)_k - \langle f,\,e_{d,k}\rangle_{L^2(\Omega)}\big| \lesssim h^{\min\{\alpha+1/2,\ell\}}\;\|f\|_{H^\alpha(\Omega)}\;\|\mathcal{T}_2^{d,k}\upsilon\|_{H^{\alpha+1/2}(Z)}$$

†The restriction on $\|k\|$ guarantees that $\mathcal{T}_1^{d,k}e$ is compactly supported in $\overline{\Omega}$.

In view of Lemma 5.3 we are done provided $v \in H^{\alpha+1/2}(\mathbb{R})$. We have that $v = \Lambda \mathbf{R} e/(2\pi)$, see [13, Sec. 3], where the Λ-operator is defined by the Fourier transform via $\widehat{\Lambda f}(\xi) = |\xi|\,\widehat{f}(\xi)$. Since $\mathbf{R}e \in H_0^{\alpha+3/2}(-1,1)$, see (5.1), and since Λ maps $H^s(\mathbb{R})$ boundedly to $H^{s-1}(\mathbb{R})$ we are finally finished with the proof Lemma 5.4. $\square$

THEOREM 5.5. *Let $\alpha > 1/2$ and let f be in $H_0^\alpha(\Omega)$ with $\operatorname{supp} f \Subset \Omega$. The radially symmetric mollifier e is assumed to be in $H_0^{\alpha+1}(\Omega)$. Let $\widetilde{d} = \widetilde{d}(f)$ be the smallest positive integer such that $\operatorname{supp} f$ is contained in $\mathcal{B}_{1-1/\widetilde{d}}(0)$, the ball about the origin with radius $1 - 1/\widetilde{d}$. If $d \geq \widetilde{d}$ then*

$$\|\widetilde{\mathbf{R}}^{(\ell)}_{n_\ell,d}\Psi^{(\ell)}_{q,p}\mathbf{R}f - f\|_{L^2(\Omega)} \lesssim \left(d^{-\min\{2,\alpha\}} + h^{\min\{\ell,\alpha+1/2\}}\, d^{\alpha+2}\right)\|f\|_{H^\alpha(\Omega)}. \tag{5.9}$$

Proof. For $d \geq \widetilde{d}$ and $\|k\| \geq d$ the inner products $\langle f, e_{d,k}\rangle_{L^2(\Omega)}$ vanish. Thus,

$$E_d f(x) = \sum_{\substack{k\in\mathbb{N}^2\\ \|k\|\leq d-1}} \langle f, e_{d,k}\rangle_{L^2(\Omega)}\; B(dx-k).$$

Therefore, we may apply Theorem 4.1 together with (5.8) and Lemma 5.4. $\square$

Now we investigate how $d = d(q,p)$ has to be chosen to guarantee convergence when q and p, that is, the number of data points, grow to infinity. The Radon data provide optimal resolution if q and p are related by the sampling relation $p = \pi\, q$, see Natterer [10, Chap. III]. In the sequel we therefore assume that $p \simeq q$. The proof of the following corollary is a straight forward consequence of (5.9).

COROLLARY 5.6. *Adopt the assumptions of Theorem 5.5. Suppose that $p \simeq q$ and that $d \simeq q^\lambda$ with $0 < \lambda < \frac{\min\{\ell,\alpha+1/2\}}{\alpha+2}$. Then,*

$$\lim_{q\to\infty} \|\widetilde{\mathbf{R}}^{(\ell)}_{n_\ell,d}\Psi^{(\ell)}_{q,p}\mathbf{R}f - f\|_{L^2(\Omega)} = 0.$$

Moreover, if $\lambda = \lambda(\alpha) = \frac{\min\{\ell,\alpha+1/2\}}{\min\{2,\alpha\}+\alpha+2}$ then

$$\|\widetilde{\mathbf{R}}^{(\ell)}_{n_\ell,d}\Psi^{(\ell)}_{q,p}\mathbf{R}f - f\|_{L^2(\Omega)} \lesssim q^{-\lambda(\alpha)\,\min\{2,\alpha\}}\;\|f\|_{H^\alpha(\Omega)} \quad \textit{as } q\to\infty.$$

The best possible L^2-convergence rate for the reconstruction of $f \in H_0^\alpha(\Omega)$ from the discrete Radon data is $q^{-\alpha}\,\|f\|_{H^\alpha(\Omega)}$, see Natterer [10, Chap. IV, Theorem 2.2]. Since $\lambda(\alpha) \leq 2/5$ for $1/2 < \alpha \leq 3/2$ and $\ell \in \{1,2\}$ we could only derive a *suboptimal* convergence result for the filtered backprojection algorithm.

REMARK 5.7. In [12] Popov showed *pointwise* convergence of the filtered backprojection algorithm. However, the results are restricted to compactly supported functions of the form $f(x) = \sum_{i=1}^N g_i(x)\,\chi_{D_i}(x)$ where $g_i \in \mathcal{C}^\infty$ and χ_{D_i} is the characteristic function of a region D_i with a smooth boundary. For instance, the often used Shepp-Logan head phantom [16] is of this form as a superposition of indicator functions of ellipses. Our analysis covers L^2-convergence for a different class of functions ($H_0^\alpha(\Omega)$ with $\alpha > 1/2$). Popov's functions are in $H_0^\alpha(\Omega)$ for any $\alpha < 1/2$ and therefore our analysis does 'just' not apply to them. On the other hand, Popov needs Radon data which are averaged (smoothed) with respect to the variable s.

Finally we formulate the regularization property of the filtered backprojection algorithm. Therefore, let $\Psi_{q,p}^{(\ell),\delta}$ be a noisy version of the observation operator (5.3) with a relative noise level $\delta > 0$, cf. (4.4):

$$\big|\big(\Psi_{q,p}^{(\ell),\delta} y\big)_{i,j} - \big(\Psi_{q,p}^{(\ell)} y\big)_{i,j}\big| \le \delta\, \|y\|_{H^{\alpha+1/2}(Z)}. \tag{5.10}$$

COROLLARY 5.8. *Adopt the assumptions of Theorem* 5.5. *Suppose that* $p \simeq q$ *and that* $d \simeq q^{\lambda}$ *with* $\lambda = \lambda(\alpha) = \frac{\min\{\ell,\alpha+1/2\}}{\min\{2,\alpha\}+\alpha+2}$. *Further, assume* (5.10).

If $q \simeq \delta^{-1/\min\{\ell,\alpha+1/2\}}$ *and* $\alpha \le 2$ *then*

$$\|\widetilde{\mathbf{R}}_{n_\ell,d}^{(\ell)} \Psi_{q,p}^{(\ell),\delta} \mathbf{R} f - f\|_{L^2(\Omega)} \lesssim \delta^{\frac{\alpha}{2(\alpha+1)}}\, \|f\|_{H^{\alpha}(\Omega)} \quad \text{as } \delta \to 0. \tag{5.11}$$

Proof. The abstract estimate (4.5) readily implies (5.11). □

A. Appendix: mollifier property – proof of (3.8) **and** (3.9)**.** The notation is from Example 3.1. We follow a standard procedure from approximation theory, see, e.g., Oswald [11, Chap. 2.3]: first we show that E_d reproduces affine-linear functions, then we bound $E_d : L^2(x_r, x_{r+1}) \to L^2(x_r, x_{r+1})$ uniformly in d, finally we apply the classical Lebesgue estimate.

We begin with verifying that $E_d p = p|_{[0,1]}$ for any affine-linear function p. Let $1 \le i \le d-1$ then

$$\langle p, e_i\rangle_{L^2(0,1)} = p(x_i)$$

where we used $\int e(x)\,\mathrm{d}x = 1$ and $\int x\, e(x)\,\mathrm{d}x = 0$ (e is even). Similarly, by using $\int e^{\mathrm{b}}(x)\,\mathrm{d}x = 1$ as well as $\int x\, e^{\mathrm{b}}(x)\,\mathrm{d}x = 0$,

$$\langle p, e_0\rangle_{L^2(0,1)} = p(0) \quad \text{and} \quad \langle p, e_d\rangle_{L^2(0,1)} = p(1).$$

Since $E_d p(x_k) = \langle p, e_k\rangle_{L^2(0,1)} = p(x_k)$, $k = 0,\ldots,d$ we have that

$$E_d p = p|_{[0,1]} \quad \text{for all affine-linear functions } p. \tag{A.1}$$

In the second step we prove the boundedness of $E_d : L^2(x_r, x_{r+1}) \to L^2(x_r, x_{r+1})$ uniformly in d and r. We have that

$$\|E_d f\|^2_{L^2(x_r,x_{r+1})} = \sum_{i,k=r}^{r+1} \langle f, e_i\rangle_{L^2(0,1)}\, \overline{\langle f, e_k\rangle}_{L^2(0,1)}\, \mathcal{B}_{i,k}, \quad \mathcal{B}_{i,k} := \langle b_i, b_k\rangle_{L^2(x_r,x_{r+1})}.$$

The symmetric 2×2-matrix $\mathcal{B}$ does not depend on r. By a Gershgorin argument we see that the spectral norm of $\mathcal{B}$ is bounded by a multiple of h : $\|\mathcal{B}\|_2 \lesssim h$. Thus,

$$\|E_d f\|^2_{L^2(x_r,x_{r+1})} \lesssim h \sum_{i=r}^{r+1} |\langle f, e_i\rangle_{L^2(0,1)}|^2 \lesssim h \sum_{i=r}^{r+1} \|f\|^2_{L^2(\operatorname{supp} e_i)}\, \|e_i\|^2_{L^2(0,1)}.$$

Since $\|e_i\|^2_{L^2} \lesssim h^{-1}$, $i = 0,\ldots,d$, we arrive at

$$\|E_d f\|^2_{L^2(x_r,x_{r+1})} \lesssim \sum_{i=r}^{r+1} \|f\|^2_{L^2(\operatorname{supp} e_i)} \lesssim \|f\|^2_{L^2(S_r)} \tag{A.2}$$

where $S_r = \operatorname{supp} e_r \cup \operatorname{supp} e_{r+1}$. We proceed with estimating $E_d f - f$. For $p \in \Pi_1$, the space of polynomials up to degree 1, we obtain

$$\|E_d f - f\|_{L^2(x_r,x_{r+1})} \leq \|f - p\|_{L^2(x_r,x_{r+1})} + \|E_d(f-p)\|_{L^2(x_r,x_{r+1})} \lesssim \|f-p\|_{L^2(S_r)}$$

where we used (A.1) as well as (A.2). Thus,

$$\|E_d f - f\|_{L^2(x_r,x_{r+1})} \lesssim \inf_{p \in \Pi_1} \|f-p\|_{L^2(S_r)}.$$

Now we apply a *Bramble-Hilbert* estimate [1] which yields

$$\|E_d f - f\|_{L^2(x_r,x_{r+1})} \lesssim h^\alpha \, |f|_{H^\alpha(S_r)} \quad \text{for } 0 \leq \alpha \leq 2.$$

Above, $|\cdot|_{H^\alpha(S_r)}$ denotes the H^α-semi-norm. Summing over r the square of both sides of the latter displayed inequality readily gives (3.9) which, in turn, implies (3.8) by a density argument.

B. Appendix: mollifier property – proof of (5.7) **and** (5.8)**.** We follow the general principle demonstrated in Appendix A. Therefore, we will be brief.

Let $e \in L^2(\mathbb{R}^n)$, $n \in \mathbb{N}$, be a radially symmetric mollifier with compact support. We consider the operator $E_d^{(n)} : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$,

$$E_d^{(n)} f(x) := \sum_{k \in \mathbb{Z}^n} \langle f, e_{d,k} \rangle_{L^2(\mathbb{R}^n)} \, B(d\,x - k),$$

being the n-dimensional generalization of E_d from (5.6), that is, $e_{d,k}(x) = d^n \, e(d\,x - k)$ and B is the n-fold tensor product of the linear B-spline b, see (3.7).

Let $p \in \Pi_1^{(n)}$, the space of polynomials up to degree 1 in n variables, then

$$E_d^{(n)} p(x) = p(x) \quad \text{for all } x \in \mathbb{R}^n.$$

The latter identity comes from $\int e(x)\,\mathrm{d}x = 1$ and $\int x_i \, e(x)\,\mathrm{d}x = 0$, $i = 1, \ldots, n$ (e is radially symmetric).

Next we demonstrate the local boundedness of $E_d^{(n)}$. Let $\square =]0,1[^n$ and $\square_{d,r} = d^{-n}(\square + r)$, $r \in \mathbb{Z}^n$. Using the tensor product structure of B we find that

$$\|E_d^{(n)} f\|^2_{L^2(\square_{d,r})} \lesssim d^{-n} \sum_{k \in \mathcal{F}_r} \|f\|^2_{L^2(\operatorname{supp} e_{d,k})} \, \|e_{d,k}\|^2_{L^2(\mathbb{R}^n)}$$

by the same arguments as in Appendix A. The index set $\mathcal{F}_r$ is given by $\mathcal{F}_r = \{k \in \mathbb{Z}^n \mid \operatorname{supp} B(\cdot - k) \cap \square_{1,r} \neq \emptyset\}$. Since $\|e_{d,k}\|^2_{L^2} \lesssim d^n$ we obtain

$$\|E_d^{(n)} f\|^2_{L^2(\square_{d,r})} \lesssim \|f\|^2_{L^2(S_{d,r})}$$

where $S_{d,r} = \bigcup_{k \in \mathcal{F}_r} \operatorname{supp} e_{d,k}$.

Since $E_d^{(n)}$ preserves polynomials of degree 1 and is locally bounded we arrive at

$$\|E_d^{(n)} f - f\|_{L^2(S_r)} \lesssim \inf_{p \in \Pi_1^{(n)}} \|f-p\|_{L^2(S_r)} \lesssim d^{-\alpha} \, |f|_{H^\alpha(S_r)}, \quad 0 \leq \alpha \leq 2,$$

where the last estimate follows from a Bramble-Hilbert argument. Thus,

$$\|E_d^{(n)} f - f\|_{L^2(\mathbb{R}^n)} \lesssim d^{-\min\{2,\alpha\}} \, |f|_{H^\alpha(\mathbb{R}^n)}.$$

Now let f be in $H_0^\alpha(\Omega) \subset H^\alpha(\mathbb{R}^2)$. Then,

$$\|E_d^{(2)} f - f\|_{L^2(\Omega)} \leq \|E_d^{(2)} f - f\|_{L^2(\mathbb{R}^2)} \lesssim d^{-\min\{2,\alpha\}} \, \|f\|_{H^\alpha(\Omega)}$$

which is (5.8). Finally, (5.8) implies the mollifier property (5.7) by a density argument.

C. Appendix: proof of Lemma 5.3. The transformation $\Phi(s,\vartheta) = (d\,s - k^t\omega(\vartheta),\vartheta)^t$ is $\mathcal{C}^\infty$-diffeomorphism between Z and $Z^{d,k}$. Further, $\mathcal{T}_2^{d,k} g = d^2\, g\circ\Phi$ and $\det J\Phi(s,\vartheta) = d$ where $J\Phi$ is the Jacobian of Φ.

Let κ be a non-negative integer. By induction we will establish that

$$|g\circ\Phi|_{H^\kappa(Z)} \lesssim d^{\kappa-1/2}\;|g|_{H^\kappa(Z^{d,k})} \tag{C.1}$$

for the semi-norm $|g|^2_{H^\kappa(Z^{d,k})} = \sum_{|\alpha|=\kappa}\|D^\alpha g\|^2_{L^2(Z^{d,k})}$ where α is in $\mathbb{N}_0^2$. Using the transformation result for integrals estimate (C.1) is easily verified for $\kappa = 0$. Now assume that (C.1) holds true for a $\kappa\in\mathbb{N}$. We have that

$$|g\circ\Phi|^2_{H^{\kappa+1}(Z)} \le |D^{(1,0)}(g\circ\Phi)|^2_{H^\kappa(Z)} + |D^{(0,1)}(g\circ\Phi)|^2_{H^\kappa(Z)}.$$

Further,

$$\begin{aligned} D^{(1,0)}(g\circ\Phi) &= d\;(D^{(1,0)}g)\circ\Phi \\ D^{(0,1)}(g\circ\Phi) &= k^t\omega(\vartheta)^\perp\;(D^{(1,0)}g)\circ\Phi + (D^{(0,1)}g)\circ\Phi. \end{aligned}$$

Therefore,

$$\begin{aligned} |D^{(1,0)}(g\circ\Phi)|^2_{H^\kappa(Z)} &= d^2\;|(D^{(1,0)}g)\circ\Phi|^2_{H^\kappa(Z)} \\ &\lesssim d^{2\kappa+1}\;|D^{(1,0)}g|^2_{H^\kappa(Z^{d,k})} \le d^{2\kappa+1}\;|g|^2_{H^{\kappa+1}(Z^{d,k})} \end{aligned}$$

by the inductive assumption. Taking into account that $|k^t\omega(\vartheta)^\perp| \le d$ (recall that $\|k\|\le d$) we also find by the arguments used above that

$$|D^{(0,1)}(g\circ\Phi)|^2_{H^\kappa(Z)} \lesssim d^{2\kappa+1}\;|g|^2_{H^{\kappa+1}(Z^{d,k})}$$

thereby finishing the inductive step to prove (C.1). Finally,

$$\left\|\mathcal{T}_2^{d,k} g\right\|^2_{H^\kappa(Z)} = d^4\,\|g\circ\Phi\|^2_{H^\kappa(Z)} = d^4\sum_{i=0}^{\kappa}|g\circ\Phi|^2_{H^i(Z)} \overset{\text{(C.1)}}{\lesssim} d^{2\kappa+3}\;\|g\|^2_{H^\kappa(Z^{d,k})}.$$

We have proved Lemma 5.3 for $\kappa\in\mathbb{N}_0$. Arbitrary $\kappa\ge 0$ can be dealt with, for instance, using arguments from interpolation theory for Sobolev spaces, see, e.g., Lions and Magenes [5, Chap. 5.1].

REFERENCES

[1] J. H. Bramble and S. R. Hilbert, *Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation*, SIAM J. Numer. Anal., 7 (1970), pp. 112–124.

[2] A. Cohen, I. Daubechies, and J.-C. Feauveau, *Biorthogonal bases of compactly supported wavelets*, Comm. Pure Appl. Math., 45 (1992), pp. 485–560.

[3] W. Dahmen, A. Kunoth, and K. Urban, *Biorthogonal spline wavelets on the interval – stability and moment conditions*, Appl. Comp. Harm. Anal., 6 (1999), pp. 132–196.

[4] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, vol. 375 of Mathematics and Its Applications, Kluwer Academic Publishers, Dordrecht, 1996.

[5] J. L. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications, Vol.* 1, Springer-Verlag, New York, 1972.

[6] A. K. Louis, *Inverse und schlecht gestellte Probleme*, Studienbücher Mathematik, B. G. Teubner, Stuttgart, Germany, 1989.

[7] ———, *Approximate inverse for linear and some nonlinear problems*, Inverse Problems, 12 (1996), pp. 175–190.

[8] ———, *A unified approach to regularization methods for linear ill-posed problems*, Inverse Problems, 15 (1999), pp. 489–498.

[9] A. K. Louis and P. Maass, *A mollifier method for linear operator equations of the first kind*, Inverse Problems, 6 (1990), pp. 427–440.

[10] F. Natterer, *The Mathematics of Computerized Tomography*, Wiley, Chichester, 1986.

[11] P. Oswald, *Multilevel Finite Element Approximation: Theory and Applications*, Teubner Skripten zur Numerik, B. G. Teubner, Stuttgart, Germany, 1994.

[12] D. A. Popov, *On convergence of a class of algorithms for the inversion of the numerical Radon transform*, in Mathematical Problems of Tomography, I. M. Gelfand and S. G. Gindikin, eds., vol. 81 of Translations of Mathematical Monographs, AMS, Providence, R.I., 1990, pp. 7–65.

[13] A. Rieder, *Principles of reconstruction filter design in 2D-computerized tomography*, in Radon Transforms and Tomography, T. Quinto, L. Ehrenpreis, A. Faridani, F. Gonzales, and E. Grinberg, eds., vol. 278 of Contemporary Mathematics, AMS, Providence, RI, 2001.

[14] A. Rieder and Th. Schuster, *The approximate inverse in action with an application to computerized tomography*, SIAM J. Numer. Anal., 37 (2000), pp. 1909–1929.

[15] L. L. Schumaker, *Spline Functions: Basic Theory*, Pure & Applied Mathematics, John Wiley & Sons, New York, 1981.

[16] L. A. Shepp and B. F. Logan, *The Fourier reconstruction of a head section*, IEEE Trans. Nuc. Sci., 21 (1974), pp. 21–43.

[17] J. Wloka, *Partial Differential Equations*, Cambridge University Press, Cambridge, U. K., 1987.