Mathematical problems in computed tomography

Adel Faridani *†

1 Introduction

Computed tomography (CT) entails the reconstruction of a function f from line integrals of f. This mathematical problem is encountered in a growing number of diverse settings in medicine, science, and technology, ranging from the famous application in diagnostic radiology to research in quantum optics. As a consequence, many aspects of CT have been extensively studied and are now well understood, thus providing an interesting model case for the study of other inverse problems. Other aspects, notably three-dimensional reconstructions, still provide numerous open problems.

This article is based on a series of talks given by the author at the 1999 Mathematical Geophysics Summer School (MGSS) held at Stanford University. Its purpose is to give an introduction to the topic, treat some aspects in more detail, and to point out references for further study. The reader interested in a broader overview of the field, its relation to various branches of pure and applied mathematics, and its development over the years may wish to consult the monographs [6, 30, 31, 35, 61, 75], the volumes [20, 21, 27, 32, 33, 74], and review articles [41, 48, 55, 57, 64, 65, 81, 86].

In practice only integrals over finitely many lines can be measured, and the distribution of these lines is sometimes restricted. The following presentation is centered around the question: *What features of f can be stably recovered from a given collection of line integrals of f?* For example, we may ask what resolution can be achieved with the available data. If a full reconstruction of f is not possible, we may try to detect the location of boundaries (jump discontinuities of f), or also the sizes of the jumps.

The exposition is divided into three parts. The first part is concerned with general theory. Its main themes are questions of uniqueness, stability and inversion for the x-ray transform, as well as detection of singularities. The second part is devoted to local tomography and is centered around a discussion of recently developed methods for computing jumps of a function from local tomographic data. The third part treats optimal sampling and has at its core a detailed error analysis of the parallel-beam filtered

*Dept. of Mathematics, Oregon State University, Corvallis, OR 97331.
Email: faridani@math.orst.edu; Homepage: http://ucs.orst.edu/~faridana

†This work was supported by NSF grant DMS-9803352 and NIH grant R01 RR 11800-4.

backprojection algorithm. Appendices on wavelets and Matlab source code for the filtered backprojection algorithm conclude the manuscript.

Part I: Some General Aspects

2 The x-ray and Radon transforms

We begin by introducing some notation and background material. $\mathbb{R}^n$ consists of n-tuples of real numbers, usually designated by single letters, $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$, etc. The inner product and absolute value are defined by $< x, y >= \sum_1^n x_i y_i$ and $|x| = \sqrt{< x, x >}$. The unit sphere S^{n-1} consists of the points with absolute value 1. $C_0^\infty(\mathbb{R}^n)$ denotes the set of infinitely differentiable functions on $\mathbb{R}^n$ with compact support. A continuous linear functional on C_0^∞ is called a distribution. If X is a set, X° denotes its interior, $\overline{X}$ its closure, and X^c its complement. χ_X and χ_n denote the characteristic functions (indicator functions) of X, and of the unit ball in $\mathbb{R}^n$, respectively. I.e., $\chi_X(x) = 1$ if $x \in X$, and $\chi_X(x) = 0$ if $x \notin X$. $|X|$ denotes the n-dimensional Lebesgue measure of $X \subset \mathbb{R}^n$. However, when it is clear that X should be treated as a set of dimension $m < n$, $|X|$ is the m-dimensional area measure. Thus

$$|S^{k-1}| = 2\pi^{k/2}/\Gamma(k/2)$$

is the $(k-1)$-dimensional area of the $(k-1)$-dimensional sphere.

The convolution of two functions is given by

$$f * g(x) = \int_{\mathbb{R}^n} f(x-y)g(y)dy.$$

The Fourier transform is defined by

$$\hat{f}(\xi) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x)e^{-i<x,\xi>}\, dx$$

for integrable functions f, and is extended to larger classes of functions or distributions by continuity or duality. For square-integrable functions f, g we have

$$f * g(x) = \int_{\mathbb{R}^n} \hat{f}(\xi)\hat{g}(\xi)e^{i<x,\xi>}\, d\xi. \tag{1}$$

The integral transforms most relevant for tomography are the x-ray transform and the Radon transform.

Definition 2.1 *Let $\theta \in S^{n-1}$ and $\Theta^\perp$ the hyperplane through the origin orthogonal to θ. We parametrize a line $l(\theta, y)$ in $\mathbb{R}^n$ by specifying its direction $\theta \in S^{n-1}$ and the point y where the line intersects the hyperplane $\Theta^\perp$.*

The x-ray transform of a function $f \in L_1(\mathbb{R}^n)$ is given by

$$Pf(\theta, y) = P_\theta f(y) = \int_{\mathbb{R}} f(y + t\theta)dt, \quad y \in \Theta^\perp. \tag{2}$$

The Radon transform of f is defined by

$$Rf(\theta, s) = R_\theta f(s) = \int_{\Theta^\perp} f(x + s\theta)dx, \quad s \in \mathbb{R}. \tag{3}$$

We see that $Pf(\theta, x)$ is the integral of f over the line $l(\theta, y)$ parallel to θ which passes through $y \in \Theta^\perp$, and that $Rf(\theta, s)$ is the integral of f over the hyperplane orthogonal to θ with signed distance s from the origin. In the following we will be mostly concerned with the x-ray transform. In two dimensions the two transforms coincide apart from the parameterization: We parametrize $\theta \in S^1$ by its polar angle φ and define a vector $\theta^\perp$ orthogonal to θ such that

$$\theta = (\cos\varphi, \sin\varphi), \quad \theta^\perp = (-\sin\varphi, \cos\varphi). \tag{4}$$

Then the points in the subspace $\Theta^\perp$ are given by $\Theta^\perp = \{s\theta^\perp, \ s \in \mathbb{R}\}$ and we have the relation $Pf(\theta, s\theta^\perp) = Rf(\theta^\perp, s)$. Also, when working in two dimensions, we will often use the simplified notation $Pf(\theta, s)$ or $P_\theta f(s)$ instead of $Pf(\theta, s\theta^\perp)$. Occasionally we will also replace θ by the polar angle φ according to (4) and write $Pf(\varphi, s)$.

Let us consider two examples. Let G be the Gaussian function $G(x) = e^{-<x,x>/2}$. Then

$$PG(\theta, y) = e^{-<y,y>/2} \int_{\mathbb{R}} e^{-<t\theta,t\theta>/2}dt = (2\pi)^{1/2}e^{-<y,y>/2}, \quad y \in \Theta^\perp. \tag{5}$$

For χ_n, the characteristic function of the unit ball in $\mathbb{R}^n$, we can use a geometrical argument. We obtain $P\chi_n(\theta, y) = 0$ for $|y| > 1$ since then the line $l(\theta, y)$ does not intersect the unit ball. For $|y| \leq 1$ observe that the intersection of the line $l(\theta, y)$ with the unit ball in $\mathbb{R}^n$ is a line segment of length $2\sqrt{1 - |y|^2}$ and that $P\chi_n(\theta, y)$ is equal to this length.

The following relation between the Fourier transforms of $P_\theta f$ and f will prove to be useful:

Theorem 2.2 *Under the hypotheses of Definition 2.1,*

$$(P_\theta f)^\wedge(\eta) = (2\pi)^{1/2}\hat{f}(\eta), \quad \eta \in \Theta^\perp \tag{6}$$

$$(R_\theta f)^\wedge(\sigma) = (2\pi)^{(n-1)/2}\hat{f}(\sigma\theta), \quad \sigma \in \mathbb{R} \tag{7}$$

Proof: This is a straightforward computation. We demonstrate it for the x-ray transform. Let $\eta \in \Theta^\perp$. Then

$$(Pf)^\wedge(\theta, \eta)$$

$$= (2\pi)^{(1-n)/2} \int_{\Theta^\perp} Pf(\theta, x) e^{-i<x,\eta>} \, dx$$

$$= (2\pi)^{(1-n)/2} \int_{\Theta^\perp} \int_{\mathbb{R}} f(x + s\theta) \, ds \; e^{-i<x,\eta>} dx$$

$$= (2\pi)^{(1-n)/2} \int_{\mathbb{R}^n} f(y) e^{-i<y,\eta>} \, dy$$

$$= \sqrt{2\pi} \, \hat{f}(\eta), \quad \eta \in \Theta^\perp. \square$$

As we will see below, Theorem 2.2 can be used to explore questions of uniqueness, nonuniqueness, stability, and inversion.

Current medical scanners employ an x-ray source which moves around the patient. To describe this type of data collection, the parameterization of lines by $\theta \in S^{n-1}$ and $y \in \Theta^\perp$ is less convenient. It is more suitable to introduce the *divergent beam x-ray transform*

$$Df(a, \theta) = D_a f(\theta) = \int_0^\infty f(a + t\theta) dt, \quad \theta \in S^{n-1}, \tag{8}$$

which gives the integral of f over the ray with direction θ emanating from the source point a.

The x-ray and Radon transforms are special cases of the general k-plane transform, which maps a function into its integrals over k-dimensional affine subspaces; see, e.g., [41].

3 Uniqueness and nonuniqueness

Theorem 3.1 *([86, 41]) Let $f \in L_2(\mathbb{R}^n)$ have compact support, and let $Pf(\theta, \cdot) \equiv 0$ for infinitely many θ. Then $f \equiv 0$.*

Proof: The Fourier transform $\hat{f}$ is analytic and $\hat{f}(\eta) = \widehat{P_\theta f}(\eta) = 0$ on the hyperplanes $< \eta, \theta >= 0$. Since no nontrivial entire function can vanish on an infinite set of hyperplanes through the origin, we must have $\hat{f} \equiv 0$. $\square$

As an application, consider the so-called limited angle problem. Let $Pf(\theta, \cdot)$ be given for infinitely many θ concentrated in a cone C. Then f is uniquely determined, even if C is very small. However, if $C \neq S^{n-1}$, the reconstruction is not stable. Indeed, the proof of the above theorem shows that reconstructing f is equivalent to analytic continuation of $\hat{f}$, and analytic continuation is known to be extremely unstable.

The uniqueness result requires an infinite number of directions, while in practice only a finite number can be measured. It was already recognized by the pioneers of CT that this entails the loss of uniqueness; see the example given in [3]. The next theorem shows that the nonuniqueness is quite extensive, i.e., given $Pf(\theta_j, \cdot)$ for finitely many directions

θ_j, there are null functions which can be prescribed arbitrarily on a large portion of their domain.

Theorem 3.2 *([86]) Let $\theta_1, \ldots, \theta_p \in S^{n-1}$, $K \subset \mathbb{R}^n$ compact, and $f \in C_0^\infty(K)$. Let $K_0 \subset U \subset K$ with U open and K_0 compact. Then there is $f_0 \in C_0^\infty(K)$, $f_0 = f$ on K_0, and $Pf_0(\theta_k, \cdot) \equiv 0$, $k = 1, \ldots, p$.*

While this result makes it seem difficult to obtain reliable reconstructions in practice, it is not the end of the story. It turns out that the null functions for the x-ray transform are high-frequency functions [50, 51, 52, 59] , and that it is possible to suppress such functions in practical reconstructions.

Theorem 3.3 *[50, 51, 52] Let $f_0 \in L_2(\mathbb{R}^2)$ with support contained in the unit disk. If $Pf_0(\theta_k, \cdot) \equiv 0$ for $k = 1, \ldots, p$, then*

$$\hat{f}_0(\sigma\theta) = \sum_{m>p} i^m \sigma^{-1} J_{m+1}(\sigma) q_m(\theta),$$

where $\sigma \in \mathbb{R}$, $\theta \in S^1$, J_{m+1} the order $m+1$ Bessel function of the first kind and q_m a polynomial of degree m.

Since $J_l(t)$ is very small for $l > t$, it follows that if $P_\theta f$ vanishes for p distinct directions θ_j, then $\hat{f}(\xi)$ is almost entirely concentrated in the set $\{\xi \in \mathbb{R}^n : |\xi| > p\}$ [59]. This means that measuring $P_{\theta_j} f$ determines $\hat{f}(\xi)$ reliably for $|\xi| < p$. However, the reconstruction problem may still be severely unstable, e.g., when the directions are concentrated in a narrow range. In cases where sufficient stability is present, a low-pass filtered version of f may be recovered. A loose application of Shannon's sampling theorem yields that the reconstruction will resolve details of size $2\pi/p$ or greater.

Remark 3.4 *It follows that the influence of nonuniqueness may be avoided in practice under the following conditions:*

a) A-priori information that $|\hat{f}(\xi)|$ is small for $|\xi| > b$ is available.

b) Data $P_{\theta_j} f$ for $p > b$ directions θ_j are measured.

c) The reconstruction method used produces a function f_R with $|\hat{f}_R(\xi)|$ small for $|\xi| > b$.

Nonuniqueness theorems for the divergent beam x-ray transform have been proved in [47, 90]. A generalization to the general k-plane transform has been given in [41].

4 Inversion and Ill-posedness

Calderón's operator Λ is defined in terms of Fourier transforms by

$$(\Lambda\varphi)^\wedge(\xi) = |\xi|\hat{\varphi}(\xi), \quad \varphi \in C_0^\infty(\mathbb{R}^n).$$

It is extended by duality to the class of functions f for which $(1+|x|)^{-1-n}f$ is integrable [14]. Note that

$$\Lambda^2 = -\Delta, \quad \Delta = \text{Laplacian}. \tag{9}$$

For $n \geq 2$, Λ^{-1}, the inverse of Λ, is given by convolution with the Riesz kernel R_1:

$$\Lambda^{-1}f = R_1 * f, \quad R_1(x) = (\pi|S^{n-2}|)^{-1}|x|^{1-n}. \tag{10}$$

In dimension $n = 1$ we have $\Lambda f = \mathcal{H}\partial f$, where ∂f denotes the derivative of f and $\mathcal{H}$ denotes the Hilbert transform

$$\mathcal{H}f(s) = \frac{1}{\pi}\int_{\mathbb{R}} \frac{f(t)}{s-t}\, dt \tag{11}$$

where the integral is understood as a principal value.

We can formally derive an inversion formula for Pf by combining Theorem 2.2 and the inverse Fourier transform. For simplicity we first consider dimension $n = 2$. Using the Fourier inversion formula, Theorem 2.2, the relation (4) and changing to polar coordinates we obtain

$$\begin{aligned}
f(x) &= (2\pi)^{-1}\int_{\mathbb{R}^2} \hat{f}(\xi)e^{i<x,\xi>}d\xi \\
&= (2\pi)^{-1}\int_0^{2\pi}\int_0^\infty \sigma\hat{f}(\sigma\theta^\perp)e^{i<x,\sigma\theta^\perp>}d\sigma d\varphi \\
&= (4\pi)^{-1}\int_0^{2\pi}\int_{-\infty}^\infty |\sigma|\hat{f}(\sigma\theta^\perp)e^{i<x,\sigma\theta^\perp>}d\sigma d\varphi \\
&= \frac{1}{2}(2\pi)^{-3/2}\int_0^{2\pi}\int_{-\infty}^\infty |\sigma|\widehat{P_\theta f}(\sigma)e^{i\sigma<x,\theta^\perp>}d\sigma d\varphi && (12) \\
&= \frac{1}{2}(2\pi)^{-3/2}\int_0^{2\pi}\int_{\mathbb{R}} (\Lambda P_\theta f)^\wedge(\sigma)e^{i\sigma<x,\theta^\perp>}d\sigma d\varphi \\
&= (4\pi)^{-1}\int_0^{2\pi} \Lambda P_\theta f(<x,\theta^\perp>)d\varphi && (13) \\
&= \frac{1}{4\pi^2}\int_0^{2\pi}\int_{\mathbb{R}} \frac{\partial P_\theta f(s)}{<x,\theta^\perp> - s}\, ds d\varphi && (14)
\end{aligned}$$

In the last step we made use of the relation $\Lambda g = \mathcal{H}\partial g$ mentioned above.

For general dimension n one uses the change of variables

$$\int_{\mathbb{R}^n} g(\xi)d\xi = |S^{n-2}|^{-1}\int_{S^{n-1}}\int_{\Theta^\perp}|\eta|g(\eta)d\eta d\theta \tag{15}$$

([86, Formula (9.2')]), and obtains

$$f(x) = \left(2\pi|S^{n-2}|\right)^{-1}\int_{S^{n-1}}\Lambda P_\theta f(E_{\Theta^\perp}x)d\theta \tag{16}$$

where $E_{\Theta^\perp}x$ denotes the orthogonal projection of x onto the subspace $\Theta^\perp$.

If we use the *backprojection* operator $P^\sharp$ defined by

$$P^\sharp g(x) = \int_{S^{n-1}} g(\theta, E_{\Theta^\perp}x)d\theta, \tag{17}$$

then (16) assumes the compact form

$$f(x) = \left(2\pi|S^{n-2}|\right)^{-1} P^\sharp\Lambda Pf(x). \tag{18}$$

An inversion formula for the Radon transform can be derived in a similar way. For other inversion formulas see [61, §II.2].

¿From equation (14) we see that computation of $f(x)$ requires integrals over lines far from x, because the Hilbert transform kernel has unbounded support. Note that $P_\theta f(< x, \theta^\perp >)$ is the integral over the line with direction θ which passes through x. Hence the inversion formula is not "local". A local inversion formula would utilize only values $P_\theta f(s)$ with s close to $< x, \theta^\perp >$. We will discuss what can be done with local formulas in a later section.

Equation (12) gives us valuable information about the stability of the inversion. The factor $|\sigma|$ in the inverse Fourier integral will become arbitrarily large. This means that the inversion is unstable. In practice measurement and discretization errors will prevent accurate computation of $\widehat{P_\theta f}(\sigma)$ for large $|\sigma|$, and these errors are then amplified by multiplication with $|\sigma|$. In other words, due to the integration in P, Pf is smoother than f itself. The inversion has to reverse this smoothing and this makes it unstable. The extent of this instability will depend on the amount of smoothing inherent in P. This can be quantified using Sobolev norms. For functions f with compact support we define

$$\begin{aligned}
\| f \|_{H_0^\alpha} &= \left(\int_{\mathbb{R}^n}(1+|\xi|^2)^\alpha\,|\hat{f}(\xi)|^2\,d\xi\right)^{1/2} \\
\| Pf \|_\alpha &= \left(\int_{S^{n-1}} d\theta\int_{\Theta^\perp} d\eta\;(1+|\eta|^2)^\alpha\,|\widehat{P_\theta f}(\eta)|^2\right)^{1/2}.
\end{aligned}$$

Then we have

Theorem 4.1 *[61, p. 42] If $f \in C_0^\infty$ is supported in the unit ball, then there are constants $c(\alpha, n), C(\alpha, n)$ such that*

$$c(\alpha, n) \parallel f \parallel_{H_0^\alpha} \;\leq\; \parallel Pf \parallel_{\alpha+\frac{1}{2}} \;\leq\; C(\alpha, n) \parallel f \parallel_{H_0^\alpha} .$$

Hence the operator P smoothes by an order $1/2$ measured in a Sobolev scale. In order to see what the instability might mean in practice we assume that we have measured data g^ϵ such that $\parallel Pf - g^\epsilon \parallel_{L_2} \leq \epsilon$, and a-priori information about f of the form $\parallel f \parallel_{H_0^\beta} \leq \rho$. For $\beta > 0$ this excludes highly oscillatory functions, so this condition corresponds to condition a) in Remark 3.4. Let f_1, f_2 be two candidate functions for reconstruction, i.e., f_1, f_2 both satisfy the a-priori condition and $\parallel Pf_i - g^\epsilon \parallel_{L_2} \leq \epsilon$. We are interested to know by how much f_1 and f_2 can differ. Since $\parallel P(f_1 - f_2) \parallel_{L_2} \leq 2\epsilon$ and $\parallel f_1 - f_2 \parallel_{H_0^\beta} \leq 2\rho$, we have the worst case error

$$\parallel f_1 - f_2 \parallel_{L^2} \;\leq\; d(\epsilon, \rho, \beta), \text{ with}$$

$$d(\epsilon, \rho, \beta) \;=\; \sup \left\{ \parallel f \parallel_{L_2} : \; \parallel Pf \parallel_{L_2} \leq 2\epsilon, \;\; \parallel f \parallel_{H_0^\beta} \leq 2\rho \right\} .$$

A natural choice for β is such that functions which are smooth except for jump discontinuities along smooth boundaries are in H_0^β. This leads to the condition $\beta < 1/2$ [61, p. 92]. For the limiting case $\beta = 1/2$ the worst case error satisfies ([61, p.94])

$$d(\epsilon, \rho, \frac{1}{2}) \;\leq\; c(n)\sqrt{\epsilon\rho}.$$

This means that the reconstruction problem is moderately ill-posed. We expect a gain of $2k$ digits in data accuracy to yield k additional accurate digits in the reconstruction. In other words, the instability in the reconstruction causes a loss of half the number of accurate digits.

Another approach to quantify the degree of ill-posedness is provided by the singular value decomposition of P [59]. Here one looks at how fast the singular values converge to zero. Again, the assessment of moderate ill-posedness is confirmed.

In order to use the inversion formula in practice we have to stabilize it. This involves a well-known trade-off between stability and accuracy of the reconstruction. Here we give up the goal of recovering the function f itself, and aim instead at reconstructing an approximation $e * f$, where e is an approximate delta function. As the computation below shows, stabilization requires the Fourier transform $\hat{e}(\xi)$ to decay sufficiently fast for large $|\xi|$. The price to pay for the stabilization is limited resolution, so e must be chosen carefully, depending on the amount and accuracy of the available measurements. Note also that a proper choice of e helps to satisfy the condition c) for avoiding the influence of nonuniqueness given in Remark 3.4.

As we will see later, it is sometimes advantageous to reconstruct $\Lambda^m f$ instead of f, with $m > -1$ an integer. The case $m = 0$ of course yields an approximation to the function f itself. Using the convolution theorem (1) for the Fourier transform we obtain in a similar way as above

$$
\begin{aligned}
& e * \Lambda^m f(x) \\
= \;& \int_{\mathbb{R}^n} \hat{e}(\xi)|\xi|^m \hat{f}(\xi) e^{i<x,\xi>} d\xi \\
= \;& |S^{n-2}|^{-1} \int_{S^{n-1}} \int_{\Theta^\perp} |\eta|^{m+1} \hat{e}(\eta) \hat{f}(\eta) e^{i<x,\eta>} d\eta d\theta \\
= \;& (2\pi)^{-1} |S^{n-2}|^{-1} \int_{S^{n-1}} \int_{\Theta^\perp} |\eta|^{m+1} (P_\theta e)^\wedge(\eta) (P_\theta f)^\wedge(\eta) e^{i<E_{\Theta^\perp} x,\eta>} d\eta d\theta \\
= \;& \int_{S^{n-1}} (k * P_\theta f)(E_{\Theta^\perp} x) d\theta, \quad m \geq -1, \qquad (19)
\end{aligned}
$$

with the convolution kernel

$$k(y) = (2\pi |S^{n-2}|)^{-1} \Lambda^{m+1} P_\theta e(y), \quad y \in \Theta^\perp. \qquad (20)$$

If e is a radial function, then $P_\theta e$ and the convolution kernel k are independent of θ.

A corresponding formula for the Radon transform can be derived by using polar coordinates in $\mathbb{R}^n$ instead of (15). For rigorous proofs and general conditions on e and f for which (19) is valid see [47], [87] and [58]. Of greatest interest are the case $m = 0$, which gives the formulas for reconstructing the function f itself, and the cases $m = \pm 1$. Letting $e \to \delta$ yields the exact inversion formula

$$\Lambda^m f(x) = (2\pi |S^{n-2}|)^{-1} \int_{S^{n-1}} \Lambda^{m+1} P_\theta f(E_{\Theta^\perp} x) d\theta. \qquad (21)$$

A desirable property would be the possibility of local reconstruction, i.e., reconstruction at a point should require only lines passing through a small neighborhood of that point. Since the parameters θ and $y \in \Theta^\perp$ of a line passing through a point x must satisfy the equation $E_{\Theta^\perp} x = y$, reconstruction according to (19) will be local if the kernel k is supported in a small neighborhood of the origin. However, for m even and $\int_{\mathbb{R}^n} e(x)dx \neq 0$, $\hat{k}$ is not analytic, so k cannot have compact support. Hence ordinary tomography is global, not local. On the other hand, it follows from (20) and (9) that k has compact support if $m \geq -1$ is odd and e has compact support. This explains the interest in the cases $m = \pm 1$. Computing $\Lambda^{-1} f(x)$ consists of taking the average of all integrals over lines passing through x. This was done in early imaging techniques preceding CT. However, the result is a very blurry image of f which by itself is of limited usefulness; see the bottom left picture in Fig. 1. Current local tomography, reviewed below, avoids this disadvantage by computing a linear combination of Λf and $\Lambda^{-1} f$.

If f is supported in the unit ball, and the source points a lie on a sphere A with center in the origin and radius $R > 1$, then the approximate inversion formula for the divergent beam x-ray transform reads [87]

$$e * \Lambda^m f(x) = R^{-1} \int_A \int_{S^{n-1}} D_a f(\theta) \mid < a, \theta > \mid k(E_{\Theta^\perp}(x-a)) \, d\theta da, \qquad (22)$$

with $m \geq -1$ and k as in (20).

We conclude this section with a few remarks on reconstruction algorithms. The *filtered backprojection algorithm* is the most popular reconstruction method. It is a computer implementation of the approximate reconstruction formulas (19) and (22) for parallel-beam and fan-beam sampling, respectively. We will discuss it in detail in a later section. For references on the filtered backprojection algorithm see, e.g., [48].

The *Fourier reconstruction algorithm* uses the Fast Fourier transform to compute

$$\widehat{P_{\theta_j} f}(\eta) \;=\; \sqrt{2\pi} \hat{f}(\eta), \quad \eta \in \Theta_j^\perp, \;\; j = 0, ..., P-1.$$

In 2D this gives values of $\hat{f}$ on a polar grid. These are now interpolated onto a rectangular grid and a 2D inverse FFT is used to obtain f. This is much faster than filtered backprojection, but the interpolation is problematic, i.e., prone to cause artifacts in the reconstructed image. For further discussion and references on methods to overcome these drawbacks see [65].

Algebraic methods do not discretize an inversion formula or use the projection slice theorem, but start from an ansatz $f(x) = \sum_{i=1}^N c_i \; \psi_i(x)$ and then solve the linear system

$$\sum_{i=1}^N c_i \; P_{\theta_j} \psi_i(y_k) \;=\; g_{jk}, \quad j,k = 1,2,\ldots$$

for the unknown coefficients c_i. Here $g_{jk} = P_{\theta_j} f(y_k)$ are the measured data. Often the basis functions are the characteristic functions of pixels or voxels, but this is of course not the only choice. Indeed, the advantage of such methods lies in their flexibility, e.g., in incorporating irregular sampling geometries or available a-priori information on f. The resulting linear systems are large and sparse and require special (usually iterative) algorithms for sufficiently fast solution. Stabilization can be achieved by limiting resolution or by stopping the iteration before convergence is achieved (see, e.g., Fig. V.12 in [61]).

Numerous other reconstruction algorithms have been developed. For a survey see, e.g., [61, Ch. V] and [64, 65].

5 Incomplete data problems and detection of singularities

Incomplete data problems arise when measurements of $P_\theta f(y)$ are unavailable for a certain range of arguments (θ, y). In dimension 2 the most common examples are the limited angle problem, the exterior problem, and the interior problem. Assume that f has compact support contained in the unit disk. In the limited angle problem, measurements $Pf(\varphi, s)$ are available only in an angular range $\varphi \in [\varphi_1, \varphi_2]$ with $|\varphi_1 - \varphi_2| < \pi$. Note that because of $Pf(\varphi, s) = Pf(\varphi+\pi, -s)$, an angular range of π is sufficient for complete data. It follows from Theorem 3.1 that f is uniquely determined by the limited angle data. The problem is lack of stability. We see from Theorem 2.2 that the data determine the Fourier transform $\hat{f}(\xi)$ in the cone $\{\xi = \sigma(\cos\varphi, \sin\varphi) : \varphi \in [\varphi_1+\pi/2, \varphi_2+\pi/2],\ \sigma \in \mathbb{R}\}$. Reconstructing f is therefore equivalent to accomplishing an analytic continuation of $\hat{f}$, and analytic continuation is severely ill-posed. A more detailed picture emerges from the singular value decomposition of the limited angle transform [53]. The severe ill-posedness is reflected in exponentially decaying singular values. However, the spectrum splits into two parts, one with singular values close to the singular values in the full-range case, and the other with singular values close to 0. The components of f corresponding to singular functions in the first part are therefore recoverable. The characterization of the unrecoverable singular functions in [53] allows to predict and recognize typical reconstruction artifacts.

In the exterior problem only line integrals $P_\theta f(s)$ with $|s| > a > 0$ are available. Uniqueness holds in the measured region but stability is missing. The singular value decomposition was given in [68, 71], and used to develop a reconstruction algorithm [72, 73].

Finally, the interior problem is characterized by measurements in the range $|s| \leq a < 1$. Uniqueness does not hold, not even inside the disk $|x| < a$ where for each point x all integrals over lines passing through a small neighborhood of x are measured. A singular value decomposition has been derived in [60]. Promising new methods for the interior problem also include the wavelet-based approaches of [2, 77] and pseudolocal tomography [40]. The wavelet localization method presented in [66] requires additional integrals over a small number of lines not intersecting the disk $|x| < a$.

None of these problems provides both uniqueness and stability. It is now natural to ask that if the function f itself cannot be recovered stably, what features of f can? One answer to this question is provided by the singular value decompositions, which tell us that components of f corresponding to singular functions with large singular values are recoverable. Another approach is to ask which *singularities* of f can be stably recovered. In many applications f can be considered to be approximately piecewise constant with jump discontinuities along the boundaries between different features. Identifying the singularities of f thus allows to determine the shape of such features. A general answer has been given in [73] based on the correspondence between the wavefront sets of f and Pf. In the special case of f being smooth except for jump discontinuities along a smooth curve Γ, a singularity at a point x is detected stably if and only if integrals over lines in a neighborhood of the tangent line to Γ at x are available. Applying this rule to the incomplete data problems mentioned above yields that in the limited angle and exterior problems not all jumps can be stably detected, since for any point x there are lines passing through x for which the data are not available. On the other hand, in case of the interior problem one can stably determine all singularities inside the disk $|x| < a$. It is thus interesting to note that the interior problem is the worst behaved of the three with respect to uniqueness, but is the best behaved with respect to detecting singularities.

For general reconstruction methods where the reconstruction preserves the stable singularities see [44]. Several methods have been suggested to detect singularities directly from the line integrals without first performing a reconstruction [38, 73, 76].

A problem of great practical interest which still poses many open problems is three-dimensional cone-beam reconstruction with sources on a curve. See, e.g., [94] for an inversion formula, [18] for a general stability result, [73] for conditions to detect singularities, and [8, 17, 22, 56, 63] for reconstruction algorithms and other developments. The approximate inversion formula (19) is very useful in two dimensions, but not so in three dimensions. It needs integrals over all lines, but in three dimensions the lines form a four parameter family, so (19) requires far more data than should be needed to determine a function of three variables. In practical 3D tomography an x-ray source moves on a curve, so only integrals over lines intersecting the curve are measured. This situation is modelled by the divergent beam x-ray transform $D_a f(\theta)$, where a runs through the curve Γ. The conditions on the source curve Γ for stable inversion are restrictive, so that in most practical situations one has an incomplete data problem. Based on the exposition in [73], we now state the precise definitions for the microlocal concepts mentioned above and apply them to this situation. The reader interested in a deeper treatment may wish to first read [73] and [26], and then proceed to articles such as [1, 23, 24, 25, 70].

The following concept of a wavefront set uses the fact that the Fourier transform of a C_0^∞ function decays rapidly. A local version of this fact can be obtained by first multiplying f with a C_0^∞ cut-off function Φ with small support, and seeing if the Fourier transform of the product Φf decays rapidly. The wavefront set gives even more specific, so-called microlocal information, inasmuch as it identifies the directions in which the Fourier transform of Φf does not decrease rapidly.

Definition 5.1 *Let f be a distribution and let $x_0, \xi_0 \in \mathbb{R}^n$, $\xi_0 \neq 0$. Then (x_0, ξ_0) is in the wavefront set of f if and only if for each cut-off function Φ in C_0^∞ with $\Phi(x_0) \neq 0$, the Fourier transform of Φf does not decrease rapidly in any conic neighborhood of the ray $\{t\xi_0, t > 0\}$.*

Loosely speaking, we say that a singularity of f can be stably detected from available x-ray data, if there exists a corresponding singularity of comparable strength in the data. The strength of a singularity can be quantified microlocally using Sobolev space concepts:

Definition 5.2 *A distribution f is in the Sobolev space H^s microlocally near (x_0, ξ_0) if and only if there is a cut-off function $\Phi \in C_0^\infty(\mathbb{R}^n)$ with $\Phi(x_0) \neq 0$ and function $u(\xi)$ homogeneous of degree zero and smooth on $\mathbb{R}^n \backslash \{0\}$ and with $u(\xi_0) \neq 0$ such that $u(\xi)(\widehat{\Phi f})(\xi) \in L^2(\mathbb{R}^n, (1 + |\xi|^2)^s)$.*

First, one localizes near x_0 by multiplying f by Φ, then one microlocalizes near ξ_0 by forming $u\widehat{\Phi f}$. and sees how rapidly $\widehat{\Phi f}$ decays at infinity.

For 3D tomography with sources on a curve we have the following result:

Theorem 5.3 *(cf. [73, Theorem 4.1], and [1, 23]) Let Γ be a smooth curve in $\mathbb{R}^3$ and f a distribution whose support is compact and disjoint from Γ. Then any wavefront set of f at (x_0, ξ_0) is stably detected from divergent beam x-ray data Df with sources on Γ if and only if*

the plane $\mathcal{P}$ through x_0 and orthogonal to ξ_0, intersects Γ transversally.

If data are taken over an open set of rays with sources on Γ, then a ray in $\mathcal{P}$ from Γ to x_0 must be in the data set for stable detection to apply. In these cases f is in H^s microlocally near (x_0, ξ_0) if and only if the corresponding singularity of Df is in $H^{s+1/2}$.

We see that the corresponding singularities of Df are weaker by 1/2 Sobolev order, but this is still strong enough to allow stable detection in practice.

Theorem 5.3 allows to analyze singularity detection in 3D tomography in the same way as described above in the two-dimensional case.

It is now interesting to ask if the available numerical algorithms can actually reconstruct all the stable singularities. The results for a general class of restricted x-ray transforms obtained in [23, 24, 25] show that microlocal analysis is also a powerful tool to answer such a question. For an introduction to these results see [26]. Explicit calculations analysing an algorithm for contour reconstruction proposed in [56] and some closely related methods have recently been given in [37, 46].

The algorithm of [56] aims to reconstruct the function

$$f_R = -\Delta D^* D f, \tag{23}$$

with

$$D^* g(x) = \int_\Gamma \| x - a \|^{-1} g\left(a, \frac{x-a}{\| x - a \|}\right) da.$$

The results in [23, 37, 46] show that the wavefront set of f_R consists of two parts. The first part contains those wavefronts (x, ξ) of f for which the plane through x and normal to ξ intersects Γ. The second part may introduce new singularities, namely on the line from a source point $a \in \Gamma$ to x, the location of the original singularity in f. This will happen if the plane through x and normal to ξ contains a and the tangent vector to Γ at a is orthogonal to ξ, i.e., the plane touches Γ but does not intersect Γ transversally. In addition, the acceleration vector of the curve at a should not be orthogonal to ξ. The Sobolev strength of these additional singularities is the same as the reconstructed part of the original wavefront set [24, 25, 37], and they appear as artifacts in numerical simulations [34, 37].

An advantage of the formula (23) is that reconstruction of f_R is local, i.e., reconstruction at a point x requires only integrals over lines close to x. In [56] it is shown that f_R approximates Λf in certain cases. Another, and apparently the historically first method for 3D local tomography is an adaptation of the algorithm by Feldkamp et al. [17] developed by P.J. Thomas at the Mayo Clinic. While the details of this algorithm have not been published, it has been used in various papers, e.g., [91, 14].

Part II: Local Tomography

6 Local tomography

Often only part of an object needs to be imaged. In this case it would be preferable if only integrals over lines which intersect the region of interest (ROI) are needed. We know from the discussion of the interior problem above that we don't have uniqueness. However, it turns out that the null functions are nearly constant inside the ROI, and we know already that all singularities inside the ROI are stably determined. Several approaches have been developed in the literature. For example, the wavelet based method of [66] exploits the fact that the error contains mostly low frequencies, and that these can be recovered by supplementing the data with relatively few measurements outside the ROI. The method of [77] which will be discussed below, extrapolates the missing data and aims at reconstruction of f up to a constant error. Another method using extrapolation of the missing data is described in [61, §VI.4].

Lambda tomography, the main topic of this section, was introduced independently in [95] and [87]. It does not attempt to reconstruct the function f itself but instead produces the related function $Lf = \Lambda f + \mu\Lambda^{-1}f$. This has the advantage that the reconstruction is strictly local in the sense that computation of $Lf(x)$ requires only integrals over lines passing arbitrarily close to x. Local tomography has found applications in medical imaging [91], nondestructive testing [82, 96], and microtomography [15, 16, 80, 83]. Extensions to more general settings have been presented in [36, 44]. Other approaches include [2] and [40].

Intelligent use of Lambda tomography requires knowledge of what kind of useful information about f is retained in Lf. Let us consider an example. The upper left of Fig. 1 shows an ordinary, global reconstruction of the density function f of a calibration object used by the Siemens company. The data come from an old generation Siemens hospital scanner. Units are such that the radius of the global reconstruction circle is one. The figure displays the reconstruction inside the rectangle $[-.5, .5]^2$. The scanning geometry is a fan-beam geometry (53) with source radius $R = 2.868$, $p = 720$ source positions, and $2q = 512$ rays per source. The upper right of Fig. 1 shows a reconstruction of Λf. Reconstructions of $\Lambda^{-1}f$ and $Lf = \Lambda f + 46\Lambda^{-1}f$ are shown in the lower left and lower right, respectively. The similarity between the images of f and Λf is at first glance surprising. We expect that a good local reconstruction method should detect the singularities of f, since these are stably determined by the data. Indeed, since Λ is an invertible elliptic pseudo-differential operator, f and Λf have precisely the same singular set. However, we see that Λf is cupped where f is constant, and that the singularities are amplified in Λf. The image of $\Lambda^{-1}f$ by itself seems less useful, but it provides a countercup for the cup in Λf. Thus, the image of Lf shows less cupping and looks even more similar to f than the image of Λf. For example, the image of Lf indicates that the density just inside the boundary of the object is larger than the density outside the object, while this can not be clearly seen from the image of Λf. To achieve this effect, a good selection of μ is necessary. A prescription for selecting μ can be found in [15].

A more detailed understanding of images of Λf or Lf is obtained from studying quantitative relations between Λf, $\Lambda^{-1} f$ and f [14, 15]. Some of the results for Λf are as follows. For corresponding results on Λ^{-1} see [14].

Theorem 6.1 *([14]) Let X and Y be measurable sets, $n \geq 2$, and let $(1+|x|)^{-1-n} f$ be integrable.*

(a) If $f_r(x) = f(x/r)$, then $\Lambda f_r(x) = r^{-1} \Lambda f(x/r)$.

(b) $\Lambda \chi_X(x) > 0$ on X°, and < 0 on $X^{c\circ}$; $\Lambda \chi_{X^c} = -\Lambda \chi_X$.

(c) $\Lambda \chi_X$ is subharmonic (Laplacian ≥ 0) on X°, and superharmonic on $X^{c\circ}$. This implies that $\Lambda \chi_X$ cannot have a local maximum in X°, nor a local minimum in $X^{c\circ}$.

(d) If x is outside the support of f, then

$$\Lambda f(x) = \frac{1-n}{\pi |S^{n-2}|} \int_{\mathbb{R}^n} |x-y|^{-1-n} f(y) dy.$$

(e) Near ∂X, $|\Lambda \chi_X(y)| \sim \frac{1}{d(y, \partial X)}$, where $d(x, \partial X)$ denotes the distance of x to ∂X.

Remark 6.2 *The results for $\Lambda \chi_X$ are of practical interest, since in many applications the function f can be modeled as a linear combination of characteristic functions.*

- *As a consequence of (a), small features are amplified in images of Λf. This is beneficial for the detection of small, low contrast details. For example, in Fig. 1 the small holes in the rectangular pieces are more clearly visible in the image of Λf than in the image of f.*

- *Part (b) indicates that the jumps of Λf at discontinuities of f have the same direction as those of f.*

- *Part (c) explains why there are no oscillations which could be mistaken for actual details in images of Λf.*

- *Part (d) shows that if f has compact support, then Λf cannot. This means that there are global effects in images of Λf in the sense that the value of $\Lambda f(x_0)$ depends on the values of f everywhere. However, Part d) implies that $\Lambda f(x)$ will decay at least as $O(|x|^{-1-n})$ for $|x| \to \infty$. More refined estimates are derived in [15].*

- *Part (e) shows that a finite jump in f causes an infinite jump in Λf. In a neighborhood of ∂X, Λf is not a function but a principal value distribution [14].*

While Lf retains the signs of jumps in density, it does not give direct information about the size of these jumps. However, such information about density differences may be extracted in certain cases. In the following we will describe several methods. We assume that f is a linear combination of a smooth function and of characteristic functions of sets:

$$f = f_0 + \sum c_i \chi_{X_i}, \quad f_0 \in C_0^\infty, \quad |\partial X_i| = 0, \quad X_i = \overline{X_i^\circ}, \quad X_i^\circ \cap X_j^\circ = \emptyset \text{ if } i \neq j. \quad (24)$$

We are interested in estimating $c_j - c_i$ when X_j, X_i have a common nontrivial boundary Γ,

$$\Gamma = \partial X_i \cap \partial X_j \cap W \neq \emptyset, \quad W = (X_i \cup X_j)^\circ. \quad (25)$$

We first discuss the method developed in [15]. It is based on Theorem 6.3 below. The theorem expresses the fact that for x sufficiently close to Γ,

$$c_j - c_i \;=\; \frac{\Lambda f(x)}{\Lambda \chi_{X_j}(x)} \;+\; O(d), \qquad \text{and}$$

$$|c_j - c_i| \;=\; \frac{|\nabla \Lambda f(x)|}{|\nabla \Lambda \chi_{X_j}(x)|} \;+\; O(d^2),$$

where d is the distance from x to Γ.

Recall that a set Y has curvature $\leq 1/r$ along a subset Y_0 of ∂Y if for each point $\bar{y} \in Y_0$ there are open balls $B \subset Y$ and $B' \subset Y^c$ of radius r with $\bar{y} \in \bar{B} \cap \overline{B'}$. The distance of a point x to a set Y is denoted by $d(x, Y)$.

Theorem 6.3 *([15]) Let f be as in (24). Fix i, j, let $W = (X_i \cup X_j)^\circ$ and assume that*

$$\Gamma = \partial X_i \cap \partial X_j \cap W \neq \emptyset.$$

Let X_j have curvature $\leq 1/r$, $r > 0$, along a closed subset Γ_0 of Γ. Let $x \in W \backslash \Gamma$ be such that $d(x, \partial X_j) = d(x, \Gamma_0) = d$. Then

$$\left| \frac{\Lambda f(x)}{\Lambda \chi_{X_j}(x)} - (c_j - c_i) \right|$$

$$\leq \; F_1(d/r) \left(\max |\Lambda f_0| \;+\; C_1 \frac{(\max_{k \neq j} |c_k|)}{d(x, \partial W)} \right) d \quad (26)$$

$$\left| \frac{|\nabla \Lambda f(x)|}{|\nabla \Lambda \chi_{X_j}(x)|} \;-\; |c_j - c_i| \right|$$

$$\leq \; F_2(d/r) \left(\max |\nabla \Lambda f_0| \;+\; C_2 \frac{(\max_{k \neq j} |c_k|)}{d(x, \partial W)^2} \right) d^2 \quad (27)$$

The constants C_1 and C_2 and the functions F_1, F_2 can be given explicitly. E.g., for $n = 2$, $C_1 = 2$, and $C_2 = 3$. Furthermore,

$$\lim_{t \to 0^+} F_1(t) \;=\; \lim_{t \to 0^+} F_2(t) \;=\; \pi.$$

The error terms on the right-hand sides of (26) and (27) indicate that in general the estimate (27) should be more accurate than (26) when d is small. The terms involving $d(x, \partial W)$ come from the influence of other boundaries than Γ.

Numerical implementation of (26) or (27) requires computation of reconstructions of Λf and $\Lambda \chi_{X_j}$ inside a region of interest R. In the following let $\bar{\Lambda} f$ and $\bar{\Lambda} \chi_{X_j}$ denote these reconstructions, rather than the functions Λf and $\Lambda \chi_{X_j}$ themselves. It is also assumed that f has the form (24) with sets X_i such that $X_i \subset R$ or $X_i \cap R = \emptyset$. This entails no loss of generality since any set X_i violating this condition can be replaced by the two sets $X_i \cap R$ and $X_i \cap R^c$. $\bar{\Lambda} \chi_{X_j}$ is computed using simulated x-ray data, after ∂X_j has been found from $\bar{\Lambda} f$. In principle, either (26) or (27) can be used, but as discussed above the method based on (27) is likely to be more accurate. This gives only $|c_j - c_i|$, but since the sign of $c_j - c_i$ is preserved in Λf, this is all that is needed.

The method consists of the following steps:

1. Compute $\bar{\Lambda} f$ from local data inside a region of interest R.
2. Determine X_j by finding ∂X_j from $\bar{\Lambda} f$.
3. Compute $\bar{\Lambda} \chi_{X_j}$ inside the region of interest from simulated x-ray data, using the same sampling geometry as for the original data.
4. If $x \in \partial X_j$, take the ratio $|\nabla \bar{\Lambda} f(x)| / |\nabla \bar{\Lambda} \chi_{X_j}(x)|$ as an estimate for the magnitude of the density jump. It is advisable to use suitable averages of the gradients over points near the boundary of X_j instead of the gradient at a single point x. This reduces effects due to measurement noise.

Following [12], we demonstrate the method with x-ray data from a medical scanner. Additional applications of this method are reported in [15, 12, 80].

The top left picture of Fig. 2 shows again the global reconstruction of the calibration object. The region of interest R is indicated by the box. The picture in the upper right shows the local reconstruction $\bar{\Lambda} f$ inside the region of interest. The goal is to determine the density difference between the small hole and its surroundings. Let X_j be the characteristic function of the hole.

Finding ∂X_j involves edge detection. This is currently done by the user of the method, who specifies the vertices of a polygon approximating ∂X_j. MATLAB's image processing toolbox allows to do this selection conveniently. Our software gives the user the option to use either the reconstruction $\bar{\Lambda} f$ itself, an image of $|\nabla \bar{\Lambda} f|$, or the result of a standard automatic edge detector, for specifying the polygon. Which image is most convenient differs from application to application. In [15], where the method was applied to projection data from a human pelvis, the gradient image was most convenient. Here the result of the automatic edge detector (MATLAB's edge command) applied to the reconstruction $\bar{\Lambda} f$ is satisfactory, as can be seen from the lower right of Fig. 2.

The top left of Fig. 3 shows an image of $|\nabla \bar{\Lambda} f|$ inside the region of interest. The box indicates the subregion R' containing the part of the boundary which will be used to estimate the density jump. Here we average over the whole boundary of the small hole. MATLAB's 'imcrop' command allows convenient selection of R' by the user. The corresponding image for $|\nabla \bar{\Lambda} \chi_{X_j}|$ is shown in the top right part of the figure. This reconstruction was computed from simulated x-ray data using the same scanning geometry as in the reconstruction from real data. Having no specific information on the detectors, the effect of the positive detector width was modeled by averaging line integrals over the angular distance between two adjacent detectors.

The following averaging procedure was used to estimate the density difference. Let M be the maximum of $|\nabla \bar{\Lambda} f|$ in R'. Take the average of $|\nabla \bar{\Lambda} f(x)|$ over all points x in R' such that $|\nabla \bar{\Lambda} f(x)| > tM$ for some $t \in (0, 1)$. The same averaging procedure is applied to $|\nabla \bar{\Lambda} \chi_{X_j}(x)|$, with M replaced by the maximum of $|\nabla \bar{\Lambda} \chi_{X_j}|$ in R'. The ratio of the two averages is the estimate $d(t)$ for the density difference. This estimate depends on the choice of t. If t is too close to 1, the average is taken over very few points, while a small t will include points too far from the boundary. So t should be chosen small enough to have sufficiently many points for averaging, but large enough so that only points close to the boundary contribute to the averages. The graph in the bottom left of Fig. 3 displays the estimated density differences $d(t)$ for $.5 \leq t \leq .96$. The bottom right shows the numbers $N(t)$ of points contributing to the averages of $|\nabla \bar{\Lambda} f|$ (solid line), and of $|\nabla \bar{\Lambda} \chi_{X_j}|$ (dotted line). If $t \geq .9$ very few points contribute to the average of $|\nabla \bar{\Lambda} f(x)|$. The corresponding estimates are therefore likely to be unreliable. On the other hand, for $t < .6$ points away from the boundary begin to contribute to the average. The binary images in Fig. 4 show the location of the points considered for the averages in the case of $t = .6$, and $t = .9$, respectively. Hence reasonable estimates for the density difference are the values of $d(t)$ for $.6 \leq t \leq .9$. These values lie between 1782 and 1898. The global reconstruction indicates that the true density difference is approximately 1854. Hence all of the acceptable estimates lie between 96 and 102 per cent of the true value. Since the x-ray data have been scaled by an unknown factor, the reconstructed values do not correspond to Hounsfield numbers.

When implementing the method described above a few parameters have to be chosen judiciously and a few comments on how to do this are in order. If the filtered backprojection algorithm is used the reconstruction $\bar{\Lambda} f$ will, apart from discretization errors, be equal to $e * \Lambda f$. The point-spread function e is assumed to be of the form $e = e_r$ defined in (47) and (50) below, usually with $\alpha = 11.4174$; cf. the appendix of [14]. Choosing the point spread radius r entails the usual tradeoff between stability (larger r) and high resolution (smaller r) and will depend on the number of measured line integrals as well as on the accuracy of these measurements. In the example above $r = .0225$, which means that the minimum of the convolution kernel falls on the second detector; cf. [14, §9].

The other important parameter is the spacing h of the grid of points where $|\nabla\bar{\Lambda}f| \simeq |\nabla(e * \Lambda f)|$ is computed. $|\nabla(e * \Lambda f)|$ varies rapidly near a boundary and h has to be sufficiently small so that the maximum of the gradient at the gridpoints is close to the overall maximum. The special case with f the characteristic function of a halfspace seems to give sufficient guidance for practical purposes. If f is the characteristic function χ_H of a halfspace H then both Λf and $e * \Lambda f$ can be computed as follows. For $x \notin \partial H$ one has ([14, Theorem 4.5])

$$\Lambda\chi_H(x) = (\pi\tilde{d}(x))^{-1},$$

where $\tilde{d}(x)$ is the signed distance of x from ∂H, i.e., $\tilde{d}(x) = d(x, \partial H)$ for $x \in H$, and $\tilde{d}(x) = -d(x, \partial H)$ for $x \notin H$. Computing $e * \Lambda\chi_H$ involves the Radon transform (3) of e. Since e is radial, $R_\theta e$ does not depend on θ. Therefore the subscript θ will be suppressed and $Re(s)$ viewed as a function of the one variable s. It now follows that

$$e * \Lambda\chi_H(x) = \mathcal{H}Re(\tilde{d}(x)), \tag{28}$$

where $\mathcal{H}$ denotes the Hilbert transform as defined in (11). Observing that for functions f of one variable $\Lambda f(t) = \frac{d}{dt}\mathcal{H}f(t)$ gives

$$|\nabla(e * \Lambda\chi_H(x))| \;=\; |\Lambda Re(\tilde{d}(x))|. \tag{29}$$

Inspection of the graph of ΛRe for e as in (50) and $\alpha = 11.4174$ now yields that the width of the interval where $|\Lambda Re(t)| > 0.98(\max_{s\in\mathbb{R}} |\Lambda Re(s)|)$ is approximately $r/20$. Hence a rule of thumb for choosing h would be to set $h = r/20$.

The method described above can be simplified by making a priori assumptions about the unknown boundary ∂X_j, so that the polygonal approximations and the reconstruction from simulated data are avoided. For example, X_j could be assumed to be a halfspace H. Replacing Λf and $\Lambda\chi_{X_j}$ in (26) and (27) by $e * \Lambda f$ and $e * \Lambda\chi_H$, and using (28) and (29) gives the approximate formulas

$$c_j - c_i \simeq \frac{e * \Lambda f(x)}{\mathcal{H}Re(\tilde{d}(x))}, \tag{30}$$

$$|c_j - c_i| \simeq \frac{|\nabla(e * \Lambda f(x))|}{|\Lambda Re(\tilde{d}(x))|}. \tag{31}$$

These two formulas are the basis of two of the algorithms proposed in [39, 75] for dimension $n = 2$, cf. formulas (2.17) and (2.21) in [39]. The derivation in [39, 75] is different and employs an asymptotic expansion for Λf, where f is smooth except for jumps across smooth boundaries.

Another method to compute jumps of a function from essentially local data is *pseudolocal tomography* [40, 75]. We follow the presentation given in [4] which allows to understand the numerical implementation of this method in the framework of (30) and (31).

The starting point for pseudolocal tomography is the two-dimensional inversion formula (14) which we repeat here:

$$\begin{aligned} f(x) &= \frac{1}{4\pi}\int_{S^1} \mathcal{H}\partial P_\theta f(< x, \theta^\perp >) d\theta \\ &= \frac{1}{4\pi^2}\int_0^{2\pi}\int_{\mathbb{R}} \frac{\frac{d}{ds}P_\theta f(s)}{< x, \theta^\perp > -s}\, ds\, d\varphi. \end{aligned}$$

Now truncate the Hilbert transform integral and define

$$f_d(x) = \frac{1}{4\pi^2}\int_0^{2\pi}\int_{< x,\theta^\perp > -d}^{< x,\theta^\perp > +d} \frac{\frac{d}{ds}P_\theta f(s)}{< x, \theta^\perp > -s}\, ds\, d\varphi. \tag{32}$$

It was shown in [40] that $f - f_d$ is continuous, hence f_d has the same jumps as f. Recalling that $P_\theta f(< x, \theta^\perp >)$ is the integral over the line in direction θ which passes through x, we see that computation of $f_d(x)$ requires only integrals over lines with distance at most d from x ("pseudo-local" reconstruction.)

In practice one has to use an approximate inversion formula and computes

$$\begin{aligned} f_{d,r}(x) &= e_r * f_d(x) = \int_0^{2\pi}\int_{\mathbb{R}} \tilde{k}_{d,r}(< x, \theta^\perp > -s) P_\theta f(s)\, ds\, d\varphi \\ \tilde{k}_{d,r}(t) &= \frac{1}{4\pi^2}\int_{t-d}^{t+d} \frac{\frac{d}{ds}P_\theta e_r(s)}{t-s}\, ds \qquad (33) \\ & \qquad (34) \end{aligned}$$

where e_r is a *radial* function satisfying

$$e_r(x) = r^{-2} e_1(x/r), \quad e_1(x) = 0 \text{ for } |x| > 1, \quad \int_{\mathbb{R}^2} e_1\, dx = 1.$$

Note that $\tilde{k}_{d,r}(t) = 0$ for $|t| > d + r$, i.e., computation of $f_{d,r}(x)$ requires integrals over lines with distance at most $d + r$ from x. Furthermore, $\lim_{d\to\infty} \tilde{k}_{d,r}(t) = (4\pi)^{-1}\mathcal{H}\partial P_\theta e_r(t)$. Hence (19) gives that $\lim_{d\to\infty} f_{d,r}(x) = e_r * f(x)$. Indeed, the convolution kernel $k_{d,r}$ can be obtained from the kernel k in (20) by letting $m = 0$ and truncating the Hilbert transform integral. The relation $f_{d,r} = e_r * f_d$ was shown in [40].

It turns out that for small d (i.e., local data), f_d is significantly different from zero only in a narrow region near a boundary (cf. [40, Fig. 3]), and that the convolution with the point spread function e_r alters these values so much that the jumps cannot just be simply read off the reconstructed image $f_{d,r}$. We need an algorithm to obtain information about the jumps of f. The methods developed by Katsevich and Ramm [40, 75] can be understood in the framework developed for Lambda tomography. According to (30) and (31) we have for x close to Γ

$$c_j - c_i \simeq \frac{E * \Lambda f(x)}{\mathcal{H}RE(\tilde{d}(x))} \tag{35}$$

$$|c_j - c_i| \simeq \frac{|\nabla E * \Lambda f(x)|}{|\Lambda \mathrm{R} E(\tilde{d}(x))|} \tag{36}$$

The task now is to find $E_{d,r}$ such that $E_{d,r} * \Lambda f = f_{d,r} = e_r * f_d$.

Proposition 6.4 *([75, 4]) Define $E_{d,r}$ by*

$$P_\theta E_{d,r} = (P_\theta e_r) * M_d$$

with

$$M_d(s) = -\frac{1}{\pi} \ln(|s/d|)\, \chi_{[-d,d]}(s).$$

Then

$$f_{d,r}(x) = E_{d,r} * \Lambda f(x).$$

With this result (35) and (36) give

$$c_j - c_i \simeq \frac{f_{d,r}(x)}{\mathcal{H}RE_{d,r}(\tilde{d}(x))} \tag{37}$$

$$|c_j - c_i| \simeq \frac{|\nabla f_{d,r}(x)|}{|\Lambda \mathrm{R} E_{d,r}(\tilde{d}(x))|} \tag{38}$$

and we can apply the same algorithms for recovering the jumps as in Lambda tomography.

Some remarks are in order.

1. Note that because $E_{d,r}$ is radial, $\mathcal{H}RE_{d,r}(0) = 0$, so $f_{d,r}(x) \simeq 0$ for $x \in \Gamma$. This makes it difficult to use the relation (37) in practice, since finding $\tilde{d}(x)$ is not easy, cf. the algorithm given [40] and further discussed in [4]. However, since $|\nabla f_{d,r}|$ is maximal for $x \in \Gamma$ one can find the points $x \in \Gamma$ by looking for the local maxima of $|\nabla f_{d,r}|$ and then estimate the jump by

$$|c_j - c_i| \simeq \frac{|\nabla f_{d,r}(x)|}{|\Lambda RE_{d,r}(0)|}, \quad x \in \Gamma.$$

   This approach has essentially been used in [75] for pseudolocal tomography and in [39] for Lambda tomography.

2. The property that f_d has the same jumps as f is not used in the algorithm.
3. $E_{d,r}(x) = 0$ for $|x| > d + r$. Hence our derivation of the algorithm is only justified for $d + r$ sufficiently small. In practice the method seems to work also for much larger values of $d + r$.

Another method which can be used for region of interest tomography is the *wavelet-based multiresolution local tomography* of [77]. It illustrates the possible uses of wavelets to 'localize' the x-ray transform, or, more precisely, to separate the features which are well determined by local data from those who are not. For readers unfamiliar with wavelets we have collected some basic facts in Appendix A.

Consider a (two-dimensional) multiresolution analysis of nested subspaces V_j, $j \in \mathbb{Z}$ of $L_2(\mathbb{R}^2)$. We assume a dilation matrix $M = 2I$ (cf. Definition 8.2 below), where I is the identity matrix, and use the notation

$$f_{j,k}(x) = 2f(2^j x - k), \quad j \in \mathbb{Z}, \quad k \in \Gamma = \mathbb{Z}^2, \quad x \in \mathbb{R}^2,$$

cf. (60). Let Φ be the scaling function and Ψ^μ, $\mu = 1, 2, 3$ the associated wavelets. Since the $\Phi_{j+1,k}$, $k \in \mathbb{Z}^2$ are a Riesz basis of the subspace V_{j+1}, a function $f \in V_{j+1}$ can be written as

$$f(x) = \sum_{k \in \mathbb{Z}^2} \tilde{A}_{j+1,k} \Phi_{j+1,k}(x).$$

The so-called *approximation coefficients* $\tilde{A}_{j,k}$ are given by

$$\tilde{A}_{j,k} =< f, \tilde{\Phi}_{j,k} >$$

where $<, >$ denotes the inner product in L_2 and $\tilde{\Phi}$ is the biorthogonal scaling function (cf. Definition 8.4). Alternatively we can use the relation $V_{j+1} = V_j + W_j$ and obtain the expansion

$$f(x) = \sum_{k \in \mathbb{Z}^2} \tilde{A}_{j,k} \Phi_{j,k}(x) \; + \; \sum_{\mu=1}^{3} \sum_{k \in \mathbb{Z}^2} \tilde{D}^\mu_{j,k} \Psi^\mu_{j,k}(x).$$

We can interpret the first sum as an approximation to f in $V_j \subset V_{j+1}$, i.e., at a lower resolution. The second sum supplies the missing detail information. Therefore the coefficients

$$\tilde{D}^\mu_{j,k} =< f, \tilde{\Psi}^\mu_{j,k} >$$

are called *detail coefficients*. The Fast Wavelet Transform and its inverse (see Theorem 8.6) allow efficient computation of the $\tilde{A}_{j,k}$ and $\tilde{D}^\mu_{j,k}$, $k \in \mathbb{Z}^2$ from the $\tilde{A}_{j+1,k}$, $k \in \mathbb{Z}^2$, and vice versa.

We now observe that the approximation and detail coefficients can be computed directly from the x-ray data. Let $f^\vee(x) = f(-x)$. Then

$$\tilde{A}_{j,k} =< f, \tilde{\Phi}_{j,k} >= (f * \overline{\tilde{\Phi}^\vee_{j,0}})(2^{-j}k) \tag{39}$$

Similarly,

$$\widetilde{D}^{\mu}_{j,k} =< f, \widetilde{\psi}^{\mu}_{j,k} >= (f * \overline{\widetilde{\Psi}^{\mu}_{j,0}{}^{\vee}})(2^{-j}k) \tag{40}$$

Hence we can use the approximate inversion formula (19) with $e(x) = \overline{\widetilde{\phi}^{\vee}_{j,0}}(x)$ and reconstruction on the grid $x = 2^{-j}k$, $k \in \mathbb{Z}^2$, to obtain the approximation coefficients directly from the x-ray data. For the detail coefficients we let $e = \overline{\widetilde{\Psi}^{\mu}_{j,0}{}^{\vee}}$. Alternatively one could first compute the approximation coefficients $\tilde{A}_{j+1,k}$ by letting $e(x) = \overline{\widetilde{\phi}^{\vee}_{j+1,0}}(x)$ and choosing the finer grid $x = 2^{-j-1}k$, $k \in \mathbb{Z}^2$, and then use the Fast Wavelet Transform to obtain the approximation and detail coefficients at level j. Since the additional computational burden of applying the Fast Wavelet Transform is negligible compared to the effort required for the reconstruction from the x-ray data, this alternative method seems preferable, since only one point-spread function and corresponding convolution kernel need to be used. However, if not all coefficients on level j are needed, the first method will be more efficient.

The next question is how this approach allows to 'localize' the x-ray transform, i.e., to separate features which are determined by local data from those which are not. It was observed in [66] that the detail coefficients for sufficiently large j should be well determined by local data, if the wavelets Ψ^{μ} have vanishing moments. Let us see why.

Definition 6.5 *A function f of n variables has vanishing moments of order up to N, if*

$$\int_{\mathbb{R}^n} x^{\alpha} f(x)dx = 0$$

for all multiindices $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $|\alpha| = \sum \alpha_i \leq N$. Recall that the α_i are nonnegative integers and that $x^{\alpha} = x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n}$.

The nonlocality in the approximate inversion formula comes from the convolution kernel k in (20) in case of $m = 0$. In two dimensions this is caused by the presence of the Hilbert transform in the formula $k = (4\pi)^{-1}\Lambda P_{\theta} e = (4\pi)^{-1}\mathcal{H}\partial P_{\theta} e$. The key observation now is that the Hilbert transform of a function with vanishing moments decays fast.

Lemma 6.6 *([77, p. 1418]) Let $f(t) \in L_2(\mathbb{R})$ vanish for $|t| > A$ and have vanishing moments of order up to N. Then, for $|s| > A$,*

$$|\mathcal{H}f(s)| \leq \frac{1}{\pi |s-A|^{N+2}} \int_{-A}^{A} |f(t)t^{N+1}|dt$$

It is well known how to construct wavelets with vanishing moments, and it turns out that the functions $\partial P_{\theta} \overline{\widetilde{\Psi}^{\mu}_{j,0}{}^{\vee}}$. inherit the vanishing moments from the $\widetilde{\Psi}^{\mu}$. Therefore the convolution kernels $k = (4\pi)^{-1}\mathcal{H}\partial P_{\theta} \overline{\widetilde{\Psi}^{\mu}_{j,0}{}^{\vee}}$ will decay rapidly outside the support of $P_{\theta}\widetilde{\Psi}^{\mu}_{j,0}{}^{\vee}$.

So we see that the detail coefficients for large j, when $\widetilde{\Psi}^{\mu}_{j,0}$ has small support, are well determined by local data. This is intuitively plausible since these coefficients contain high-frequency information, and we know already from Lambda tomography that high-frequency information is well-determined. So the nonlocality shows its greatest impact in the approximation coefficients. Since the scaling function satisfies $\int \widetilde{\Phi}(x)dx = 1$, its zero order moment does not vanish. One could still choose $\widetilde{\Phi}$ so that the moments of order 1 through N vanish. It is shown in [77, p. 1419] that in such a case the resulting convolution kernel k satisfies

$$|k(s)| = O(s^{-2}) + O(s^{-N-3}).$$

It seems that this does not achieve much, since we cannot remove the leading $O(s^{-2})$ term. Nevertheless, the authors of [77] found that some scaling functions having vanishing moments lead to convolution kernels with sufficiently rapid decay for practical purposes. These scaling functions where found from one-dimensional scaling functions by the method of Definition 8.8. In their reconstructions the authors of [77] also extrapolated the missing data by constant values, thus reducing cupping artifacts. While it is suggested in [77] to first compute the approximation and detail coefficients at level j and then use an inverse fast wavelet transform to obtain the approximation coefficients at level $j+1$, numerical tests in [84] indicated that the simpler approach of directly computing the approximation coefficients at level $j+1$ yields equivalent results. We observe that this can be done without using wavelet theory, namely just by specifying the particular point spread function $e = \overline{\widetilde{\Phi}^{\vee}_{j+1,0}}$ in (19).

Part III: Sampling, Resolution and Computational Issues

7 Sampling the 2D x-ray transform

We first consider the parallel-beam geometry in two dimensions. Our analysis of sampling and resolution will use techniques from Fourier analysis. These require both the domain of Pf as well as the sampling sets to have a group structure. In 2D we parameterize $\theta \in S^1$ by $\theta = (\cos\varphi, \sin\varphi)$, and let $\theta^\perp = (-\sin\varphi, \cos\varphi)$. We write $Pf(\varphi, s)$ for $Pf(\theta, s\theta^\perp)$ and consider Pf to be a function on the group $\mathbb{T} \times \mathbb{R}$, where $\mathbb{T}$ denotes the circle group. We take the interval $[0, 2\pi)$ with addition modulo 2π as a model for $\mathbb{T}$.

Recall that for fixed φ, the values of $Pf(\varphi, s)$ for different s correspond to integrals over a collection of parallel lines. We first consider the case where Pf is measured at points

$$(\varphi_j, s_{jl}), \quad j = 0, \ldots, P-1, \quad l \in \mathbb{Z}.$$

Since for each angle φ_j we measure integrals over a collection of parallel lines $l(\varphi_j, s_{jl})$, such an arrangement is called a parallel-beam geometry. We would like the set of points (φ_j, s_{jl}) to be a discrete subgroup of $\mathbb{T} \times \mathbb{R}$, and for practical reasons we require that more than one measurement is taken for each occurring angle φ_j. We call a sampling set which satisfies these requirements an *admissible sampling lattice* (ASL). There are several ways to parameterize such lattices [11, 13, 16]. Here we use the parametrization given in [16]. If $\mathbf{L}$ is an ASL, then there are $d > 0$ and integers N, P, such that $0 \le N < P$ and

$$\mathbf{L} = \mathbf{L}(d, N, P) = \{(\varphi_j, s_{jl}) : \varphi_j = 2\pi j/P,\ s_{jl} = d(l + jN/P), j = 0, \ldots, P-1;\ l \in \mathbb{Z}\}. \tag{41}$$

We see that P is the number of angles (views). For each view angle φ_j the values s_{jl}, $l \in \mathbb{Z}$, are equispaced with spacing d, hence d is the detector spacing. The parameter N characterizes an angle dependent shift of the detector array. We also see that there are P different lattices for given parameters d and P.

The most important lattices are the standard lattice

$$\mathbf{L}_S = \{(\varphi_j, s_l) : \varphi_j = 2\pi j/P,\ s_l = d\,l,\ j = 0, \ldots, P-1,\ l \in \mathbb{Z}\}$$

which is obtained by letting $N = 0$, and the interlaced lattice

$$\mathbf{L}_I = \{(\varphi_j, s_{jl}) : \varphi_j = 2\pi j/P,\ s_{jl} = d(l + j/2),\ j = 0, \ldots, P-1,\ l \in \mathbb{Z}\}.$$

where P is even and $N = P/2$. We see that for the standard lattice the detector positions s_l do not change with the angle of view. For the interlaced lattice the detector array is shifted by one-half of a detector spacing when going from one angle of view to the next.

In practice one chooses $P = 2p$ for both lattices, and for the interlaced lattice one lets p be even. Then, because of the symmetry relation

$$Pf(\varphi, s) = Pf(\varphi + \pi, -s), \tag{42}$$

only the angles $\varphi_j \in [0,\pi)$ need to be measured. It can be shown that among all admissible lattices the standard and interlaced lattices are the only ones which fully exploit this symmetry [16].

We now describe the implementation of the filtered backprojection algorithm, which is based on discretizing the approximate inversion formula (19) with the trapezoidal rule. In two dimensions (19) reads

$$\begin{aligned} e * \Lambda^m f(x) &= (\Lambda^m e) * f(x) \\ &= \int_0^{2\pi} \int_{\mathbb{R}} k(< x, \theta^\perp > -s) Pf(\varphi, s) ds\, d\varphi, \end{aligned} \tag{43}$$

$$k(s) = \frac{1}{4\pi} \Lambda^{m+1} P_\theta e(s). \tag{44}$$

We assume that we have sampled Pf on an admissible lattice. Discretizing (43) with the trapezoidal rule gives

$$\begin{aligned} e * f(x) &\simeq \frac{2\pi}{P} \sum_{j=0}^{P-1} Q_j(< x, \theta_j^\perp >), \\ Q_j(t) &= d \sum_l k(t - s_{jl}) Pf(\varphi_j, s_{jl}), \end{aligned}$$

with φ_j, s_{jl} as in (41), and $\theta_j^\perp = (-\sin\varphi_j, \cos\varphi_j)$. We assume that f is supported in the unit disk. Hence the sum in the discrete convolution is finite. The reconstruction is usually computed for values of x on a rectangular grid $x_{m_1 m_2} = (m_1/M_1, m_2/M_2)$, $|m_i| \le M_i$. Since computing the discrete convolution $Q_j(< x, \theta_j^\perp >)$ for each occurring value of $< x, \theta_j^\perp >$ would take too long, one first computes $Q_j(iH)$, $|i| \le 1/H$, and then obtains an approximation $I_H Q_j(< x, \theta_j^\perp >)$ for $Q_j(< x, \theta_j^\perp >)$ by linear interpolation with stepsize H. We assume that

$$H = d/(N'm), \text{ with } 0 < m, N' \in \mathbb{Z}, \text{ and } N'N/P \in \mathbb{Z}. \tag{45}$$

This gives $H = d/m$ for the standard lattice ($N' = 1$) and $H = d/(2m)$ for the interlaced lattice ($N' = 2$). Then the effect of interpolating the convolution is the same as replacing the kernel k with the piecewise linear function $I_H k$ which interpolates k at the points Hl, $l \in \mathbb{Z}$; see, e.g., [11, p.84]. Hence the algorithm computes the function

$$\begin{aligned} f_R(x) &= \frac{2\pi}{P} \sum_{j=0}^{P-1} I_H Q_j(< x, \theta_j^\perp >) \\ &= \frac{2\pi d}{P} \sum_{j=0}^{P-1} \sum_{l \in \mathbb{Z}} I_H k\left(< x, \theta_j^\perp > -s_{jl}\right) Pf\left(\varphi_j, s_{jl}\right). \end{aligned} \tag{46}$$

If e is not radial, then k and $I_H k$ will depend on θ, which is not explicitly reflected in our notation. MATLAB source code implementing (46) for the standard lattice is provided for illustrative purposes in Appendix B.

In practice one chooses a basic point spread function e_1 and then controls the resolution by using a scaled version

$$e(x) = e_r(x) = r^{-2} e_1(x/r). \tag{47}$$

The corresponding kernels scale as

$$k_r(s) = r^{-2-m} k_1(s/r). \tag{48}$$

Examples for point spread functions and kernels are as follows. A popular choice for global tomography ($m = 0$) is the Shepp-Logan kernel (cf. [61, p. 111])

$$k_1(s) = \frac{1}{2\pi^3} \frac{(\pi/2) - s \sin s}{(\pi/2)^2 - s^2}. \tag{49}$$

This kernel is bandlimited with bandwidth 1, i.e., the Fourier transform

$$\widehat{k_1}(\sigma) = \frac{1}{2}(2\pi)^{-3/2} |\sigma| \frac{\sin(\pi\sigma)}{\pi\sigma} \chi_{[-1,1]}(\sigma)$$

vanishes for $|\sigma| > 1$. It follows that the scaled kernel $k_r(s) = r^{-2} k_1(s/r)$ is bandlimited with bandwidth $b = 1/r$, and the same is true for the corresponding point-spread function e_r and hence for $e_r * f(x)$.

Since the kernel (49) does not have compact support, it is not useful for local tomography. There we start with a point spread function

$$e_1(x) = \begin{cases} C(1 - |x|^2)^{\alpha+1/2} & \text{for} \quad |x| < 1 \\ 0 & \text{for} \quad |x| \geq 1 \end{cases} \tag{50}$$

$$C = \Gamma(\alpha + 5/2)/(\pi \Gamma(\alpha + 3/2)),$$

cf. [14, p. 482]. The corresponding kernel for computing Λf (i.e., $m = 1$) is given by

$$K_1(s) \;=\; \frac{-1}{4\pi} \frac{d^2}{ds^2} P_\theta e_1(s) \;=\; \frac{2\Gamma(\alpha + 5/2)}{\sqrt{\pi}\,\Gamma(\alpha + 1)} (1 - s^2)^{\alpha - 1} (1 - (2\alpha + 1) s^2), \quad |s| < 1,$$

and $K_1(s) = 0$ for $|s| \geq 1$. Here we used that in one dimension $\Lambda^2 = -d^2/ds^2$. Now the kernel is no longer bandlimited, but has compact support. The scaled kernels $K_r(s) = r^{-3} K_1(s/r)$ vanish for $|s| > r$. The kernel for global tomography generated by the point spread function (50) has a complicated analytic expression but a quickly convergent series expansion [85].

Discretization of (22) yields the filtered backprojection algorithm for the fan-beam sampling geometry. Recall that f is supported in the unit disk. Let $R > 1$, $a =$

$R(\cos\alpha, \sin\alpha)$, $\theta = -(\cos(\alpha-\beta), \sin(\alpha-\beta))$, and $x-a = -|x-a|(\cos(\alpha-\gamma), \sin(\alpha-\gamma))$. Writing $Df(\alpha,\beta)$ for $D_a f(\theta)$, (22) becomes

$$e_r * \Lambda^m f(x) = R\int_0^{2\pi}\int_{-\pi/2}^{\pi/2} Df(\alpha,\beta)\cos(\beta)k_r(|x-a|\sin(\gamma-\beta))d\beta d\alpha. \qquad (51)$$

In order to evaluate the inner integral efficiently, a 'homogeneous approximation' [45] is needed. It follows from (20) and (48) that

$$k_r(|x-a|\sin(\gamma-\beta)) = |x-a|^{-2-m}k_c(\sin(\gamma-\beta)), \quad c = r/|x-a|.$$

The approximation consists in replacing $c = r/|x-a|$ by a constant independent of x and a. This gives

$$e_r * \Lambda^m f(x) \simeq R\int_0^{2\pi} |x-a|^{-2-m}\int_{-\pi/2}^{\pi/2} Df(\alpha,\beta)\cos(\beta)k_c(\sin(\gamma-\beta))d\beta d\alpha. \qquad (52)$$

¿From here we can proceed as before by discretizing with the trapezoidal rule and inserting an interpolation step. The standard sampling lattice for the fan-beam geometry has the form

$$\mathbf{L}_{SF} = \{(\alpha_j,\beta_l):\ \alpha_j = 2\pi j/p, \beta_l = l\arcsin(1/R)/q,\ j = 0,\dots,p-1,\ l = -q,\dots,q-1\}. \qquad (53)$$

The reconstruction of $\Lambda^{-1}f$ is not unstable, so convolution with e_r is not needed. One can directly discretize the formula

$$\Lambda^{-1}f(x) = (R/4\pi)\int_0^{2\pi} |x-a|^{-1}Df(\alpha,\gamma)\cos\gamma\, d\alpha, \qquad (54)$$

which comes from letting $e_b \to \delta$ in (51).

In order to further analyze the parallel-beam algorithm we use Shannon sampling theory. We begin with some definitions. We define the Fourier transform of a function g with domain $\mathbb{T}\times\mathbb{R}$ by

$$\hat{g}(k,\sigma) = \frac{1}{2\pi}\int_0^{2\pi}\int_{\mathbb{R}} g(\varphi,s)e^{-i(k\varphi+\sigma s)}\, ds\, d\varphi, \quad (k,\sigma)\in\mathbb{Z}\times\mathbb{R}.$$

The corresponding inverse Fourier transform is given by

$$\begin{aligned}\tilde{G}(\varphi,s) &= \frac{1}{2\pi}\int_{\mathbb{Z}\times\mathbb{R}} G(\zeta)e^{i<z,\zeta>}\, d\zeta, \quad z = (\varphi,s)\\ &= \frac{1}{2\pi}\sum_{k\in\mathbb{Z}}\int_{\mathbb{R}} G(k,\sigma)e^{-i(k\varphi+\sigma s)}\, d\sigma. \qquad (55)\end{aligned}$$

The reciprocal lattice $\mathbf{L}^\perp \subset \mathbb{Z} \times \mathbb{R}$ is defined as

$$\mathbf{L}^\perp(d, N, P) \;=\; \begin{pmatrix} P & -N \\ 0 & 2\pi/d \end{pmatrix} \mathbb{Z}^2.$$

For $g \in C_0^\infty(\mathbb{T} \times \mathbb{R})$, $\mathbf{L} = \mathbf{L}(d, N, P)$ an ASL, $K \subset \mathbb{Z} \times \mathbb{R}$ compact define

$$Sg(z) = \frac{d}{P} \sum_{y \in \mathbf{L}} g(y) \tilde{\chi}_K(z - y), \quad z \in \mathbb{T} \times \mathbb{R},$$

where $\tilde{\chi}_K$ is the inverse Fourier transform of the characteristic function of K. Sg may be viewed as an approximation of g computed from sampled values of g on the lattice $\mathbf{L}$. The following classical sampling theorem gives an error estimate for this approximation:

Theorem 7.1 *Let $g \in C_0^\infty(\mathbb{T} \times \mathbb{R})$, $\mathbf{L}$ an ASL and K be a compact subset of $\mathbb{Z} \times \mathbb{R}$ such that the translates $K + \eta, \eta \in \mathbf{L}^\perp$ are disjoint. Then*

$$|g(z) - Sg(z)| \;\leq\; \pi^{-1} \int_{(\mathbb{Z} \times \mathbb{R}) \setminus K} |\hat{g}(\zeta)| d\zeta.$$

This result is an adapted version of the multidimensional sampling theorem of Petersen and Middleton [67]. For a proof see, e.g., [61, p. 62] or [11, Theorem 2.2].

If $\mathrm{supp}(\hat{g}) \subseteq K$, then $g = Sg$, i.e., g can be recovered exactly from its samples on the lattice $\mathbf{L}$.

In order for Sg to be close to g, the set K should be such that $\hat{g}$ is concentrated in K. The 'sampling condition' that the translates $K + \eta$, $\eta \in \mathbf{L}^\perp$ be disjoint requires the reciprocal lattice $\mathbf{L}^\perp$ to be sufficiently sparse, and therefore the sampling lattice $\mathbf{L}$ to be sufficiently dense.

A suitable set K for sampling the 2D x-ray transform was given by Natterer [61] based on results by Lindgren and Rattey [78]:

Theorem 7.2 *[61, p. 71] For $b > 0$ and $0 < \vartheta < 1$ let*

$$K_0(\vartheta, b) = \left\{ (k, \sigma) \in \mathbb{Z} \times \mathbb{R} : \; |\sigma| < b, \; |k| < \vartheta^{-1} \max(|\sigma|, (1 - \vartheta) b) \right\}. \tag{56}$$

Let $f \in C_0^\infty(\Omega)$. Then

$$\int_{(\mathbb{Z} \times \mathbb{R}) \setminus K_0} |\widehat{Pf}(\zeta)| \, d\zeta \;\leq\; \frac{8}{\pi^2 \vartheta} \int_{|\xi| > b} |\hat{f}(\xi)| d\xi + \; \| f \|_{L_1} \eta(\vartheta, b), \tag{57}$$

where $\eta(\vartheta, b)$ decreases exponentially with b, satisfying an estimate

$$0 \leq \eta(\vartheta, b) \leq C(\vartheta) e^{-\lambda(\vartheta) b}$$

with constants $C(\vartheta), \lambda(\vartheta) > 0$.

Usually the parameter ϑ is chosen close to 1. The parameter b plays the role of a cut-off frequency. If $|\hat{f}(\xi)|$ is sufficiently small for $|\xi| > b$ then the right-hand side of (57) will be small. In this case Theorem 7.1 imposes the condition that the sampling lattice $\mathbf{L}$ be such that the translated sets $K_0(\vartheta, b) + \eta$, $\eta \in \mathbf{L}^\perp$ are disjoint. In terms of the lattice parameters d, N and P these conditions are as follows:

For the standard lattice $N = 0$, and the reciprocal lattice $\mathbf{L}^\perp$ equals

$$\mathbf{L}^\perp = \{(Pk_1,\ 2\pi k_2/d),\ k_1, k_2 \in \mathbb{Z}\}.$$

For reasons of efficiency as discussed above we let $P = 2p$ be even. The translated sets $K_0(\vartheta, b) + \eta$, $\eta \in \mathbf{L}^\perp$ are disjoint if and only if

$$d < \pi/b, \quad p > b/\vartheta,\ P = 2p. \tag{58}$$

For the interlaced lattice we again let $K = K_0(\vartheta, b)$ as in (56). For this lattice $P = 2p$ and $N = P/2 = p$. We always let p be even, so that because of the symmetry relation (42) only the angles $\varphi_j \in [0, \pi)$ need to be measured. The reciprocal lattice is $\mathbf{L}^\perp = \{(p(2k_1 - k_2), 2\pi k_2/d),\ k_1, k_2 \in \mathbb{Z}\}$. It turns out [61, 11] that the sets $K_0(\vartheta, b) + \eta$, $\eta \in \mathbf{L}^\perp$ are disjoint if either the conditions (58) are satisfied, or if

$$\frac{\pi}{b} < d \leq \frac{2\pi}{b},\ \ p > \max\left(\frac{2\pi}{\vartheta d}, \frac{(2-\vartheta)b}{\vartheta}\right),\ p \text{ even },P = 2p. \tag{59}$$

We see that the interlaced lattice allows for a maximal detector spacing of $d = 2\pi/b$ which is twice as large as the maximum allowed for standard lattice, with only a moderate increase in p. Sampling conditions for a general admissible sampling lattice have been given in [13]. We see that both (58) and (59) require p to be greater than the cut-off frequency b, which corresponds to condition b) in Remark 3.4 on avoiding the effects of non-uniqueness.

It remains to investigate if the theoretically superior resolution of the interlaced lattice can be realized in practice. In principle there are two obvious approaches: One could first interpolate the missing data to a denser lattice and then use any reconstruction algorithm. This approach has been successfully tried in [10], but we will not discuss it here. The second approach would be to reconstruct directly from interlaced data. It turns out that the filtered backprojection algorithm is very suitable for this purpose [43, 11, 16]. We have the following error estimate, which extends the results of [43] and [11].

Theorem 7.3 *([16]) Let e be radial and sufficiently smooth, $f \in C_0^\infty(\mathbb{R}^2)$ be supported in the unit disk and the sets $K_0(\vartheta, b) + \eta$, $\eta \in \mathbf{L}^\perp$ be disjoint. Then*

$$f_R(x) \quad = \quad G_H * e * \Lambda^m f(x) + \sum_{i=1}^{4} E_i(x),$$

$$
\begin{aligned}
\widehat{G}_H(\xi) &= (2\pi)^{-1}\mathrm{sinc}^2(H|\xi|/2)\chi_1(|\xi|/b) \\
|E_1(x)| &\leq c \parallel \hat{k} \parallel_\infty \int_{|\xi|>b} |\hat{f}(\xi)|\, d\xi \\
|E_2(x)| &\leq c \parallel f \parallel_\infty \int_{|\sigma|>b} |\hat{k}(\sigma)|\, d\sigma \\
|E_3(x)| &\leq (2\pi)^{3/2} \sup_\theta \int_{-b}^{b} \Big(1 - \mathrm{sinc}^2(H\sigma/2)\Big)\, |\hat{k}(\sigma)| \sum_{l\in\mathbb{Z}} \left| \hat{f}\left(\left(\sigma + \frac{2\pi l}{d} \right) \theta \right) \right| d\sigma \\
|E_4(x)| &\leq c \parallel f \parallel_\infty \parallel \hat{k} \parallel_\infty \eta(\vartheta, b)
\end{aligned}
$$

Here $\mathrm{sinc}(x) = \sin(x)/x$. The proof ([16]) is somewhat technical and will not be given here. However, it is worthwhile to note that apart from the interpolation step this is an estimate for the error of numerical integration by the trapezoidal rule. The estimate for this error is based on the Poisson summation formula for $\mathbb{T} \times \mathbb{R}$. This approach was first applied in the present context by Kruse [43].

We will discuss the origin and importance of the four error terms. The error E_1 is the so-called aliasing error stemming from the fact that f is not bandlimited, since it has compact support. If the cut-off frequency b is chosen sufficiently large, E_1 will be small. The sampling conditions then require that the number of data available is commensurate with b. The error E_2 is present when k is not bandlimited with bandwidth b. In global tomography, i.e., when $m = 0$, one can chose e and k to be b-bandlimited, so that E_2 vanishes. In local tomography one wishes k to have compact support, so k cannot be bandlimited.

The error E_3 is caused by the interpolation step and usually not a concern when using the standard lattice. This can be explained as follows ([11]): Consider the common parameter choice $d = H = \pi/b$. Since $\hat{f}(\xi)$ is assumed to be small for $|\xi| > b$, only the term with $l = 0$ in the sum will be significant, i.e., we have for $|\sigma| \leq b$

$$
\sum_{l\in\mathbb{Z}} \left| \hat{f}((\sigma + 2\pi l/d)\theta) \right| = \sum_{l\in\mathbb{Z}} \left| \hat{f}((\sigma + 2bl)\theta) \right| \simeq |\hat{f}(\sigma\theta)|.
$$

Usually the density function f is non-negative so that $|\hat{f}(\sigma\theta)|$ has a sharply peaked maximum at $\sigma = 0$ and is very small for $|\sigma|$ close to b. In such a case the error E_2 will be small since the factor $1 - \mathrm{sinc}^2(H\sigma/2)$ is small exactly where $|\hat{f}(\sigma\theta)|$ is large.

The error E_3 is of much greater concern when the interlaced lattice is used. Consider the choice of parameters $d = 2\pi/b$, $H = \pi/b$. Now the sum over l in the estimate for E_3 may have 3 significant terms for $|\sigma| < b$:

$$
\sum_{l\in\mathbb{Z}} \left| \hat{f}((\sigma + 2\pi l/d)\theta) \right| = \sum_{l\in\mathbb{Z}} |\hat{f}((\sigma + bl)\theta)| \simeq |\hat{f}((\sigma - b)\theta)| + |\hat{f}(\sigma\theta)| + |\hat{f}((\sigma + b)\theta)|.
$$

As discussed before the contribution of the term $\hat{f}(\sigma\theta)$ is largely cancelled by the factor $(1 - \mathrm{sinc}^2(H\sigma/2))$. However, this is not the case for the other two terms. E.g., let σ be close to b. Then, assuming again that $\hat{f}$ is large near the origin, $|\hat{f}((\sigma - b)\theta)|$ will be large and is not attenuated by the factor $(1 - \mathrm{sinc}^2(H\sigma/2))$ which will be close to 1. Therefore we expect considerable reconstruction errors for this choice of parameters. That this is indeed the case has been demonstrated for global tomography in [43, 11]. Hence when using the interlaced lattice one should choose $H \ll b$, so that $(1 - \mathrm{sinc}^2(H\sigma/2))$ is small for $|\sigma| < b$. Typical choices in practice are $H = \pi/(16b)$ or smaller. Choosing $H < \pi/b$ has also a cosmetic side effect. If a b-bandlimited convolution kernel is used whose Fourier transform has a jump discontinuity at $|\sigma| = b$ (e.g., a scaled version of the Shepp-Logan kernel (49)), then ringing artifacts are caused by this discontinuity. In case of the standard lattice with the parameter choice $d = H = \pi/b$ these artifacts are practically removed by the additional smoothing from the interpolation. For smaller H this effect is lost. In this case the Fourier transform of k should taper off continuously to zero if smooth images are desired [11].

Another effect of the interpolation is the additional filtering with G_H. Since this alters only the higher frequencies, it is usually not a concern. In any case, the effect can be eliminated by choosing very small H.

The last error E_4 decreases exponentially with b, as indicated by the notation $\eta(\vartheta, b)$. Explicit estimates are as follows [16]. Let

$$\beta = \left(1 - \vartheta^2 |x|^2\right)^{3/2}.$$

For the standard lattice we have

$$|E_4(x)| \le c \parallel f \parallel_\infty \parallel \hat{k} \parallel_\infty b \frac{e^{-\beta b/\vartheta}}{1 - e^{-\beta}}.$$

For the interlaced lattice we let $b' = (1 - \vartheta)b/\vartheta$, and obtain with β as above

$$|E_4(x)| \le c \parallel f \parallel_\infty \parallel \hat{k} \parallel_\infty b\vartheta e^{-\beta b'} \left(\frac{1 + b'}{1 - e^{-\beta}} + \frac{e^{-\beta}}{(1 - e^{-\beta})^2} \right).$$

In both cases the error decays exponetially with b, but in case of $\vartheta|x|$ close to 1, when β is close to zero, the above estimates indicate that the error should be larger in case of the interlaced lattice due to the term involving $(1 - e^{-\beta})^{-2}$. This effect has been observed in [11]. It causes a thin ring artifact in the region $|x| \simeq 1$, i.e., at the boundary of the reconstruction region. It can be eliminated by choosing a smaller value for ϑ, i.e., by increasing the number of views p, cf. (58), (59).

Numerical experiments for both global and local tomography, with simulated as well as real data [43, 11, 10, 12, 16] show that the higher efficiency of the interlaced lattice can at least be partly realized in practice. However, there are also some drawbacks. There is a somewhat reduced stability because of inaccurate convolutions. In case of the interlaced lattice the stepsize $d = 2\pi/b$ is not small enough to allow accurate computation of the convolutions. Because of the truly two-dimensional nature of the numerical integration, these errors cancel out during the discrete backprojection step. The sensitivity with regard to the interpolation stepsize H can be understood as coming from a disturbance of these cancellations by the additional interpolation. A second drawback is a requirement that the sampling condition with respect to the number of views P has to be strictly observed. The aliasing caused by violating this condition is usually quite moderate in case of the standard lattice but much more severe for the interlaced lattice. This can be easily seen from the pattern in which the translated sets $K_0 + \eta$ begin to overlap when the sampling condition for p is violated. Hence the interlaced lattice seems to be most useful when the detector spacing is the main factor limiting resolution.

Good reconstructions from the interlaced lattice can also be obtained by using the direct algebraic reconstruction algorithm [42], or by increasing the amount of data through interpolation according to the sampling theorem [10]. Results for the fan beam geometry can be found in [62, 63]. As we have seen, the interpolation step can introduce significant errors in certain cases. It was recently shown [63] that the interpolation can be avoided by chosing the points x where the reconstruction is computed on a polar grid rather than on a rectangular grid, and interchanging the order of the two summations. This algorithm should work well for the interlaced lattice [97] and is particularly beneficial in case of the fan-beam sampling geometry [63], since the method also avoids the homogeneous approximation, whose influence on the reconstruction is difficult to estimate.

8 Appendix A: Some basic facts about wavelets

A brief introduction will be given to the area of multi-dimensional biorthogonal wavelets. The following discussion is taken from [84] and is based on the presentation in [93] for multi-dimensional orthonormal wavelets. For more details on wavelets and filter banks, the reader is referred to the literature, e.g., [7] or [92].

Definition 8.1 *A* **lattice** Γ *is a discrete subgroup of* $\mathbf{R}^n$ *given by integral linear combinations of a vector space basis* $\{v_1, ..., v_n\}$ *of* $\mathbf{R}^n$.

Definition 8.2 *Let* Γ *be a lattice and* M *be an* $n \times n$ *matrix such that*
i) $M\Gamma \subset \Gamma$
ii) all eigenvalues λ *of* M *satisfy* $|\lambda| > 1$.
M *is called the dilation matrix. Let* $m = |det(M)|$.
A **multiresolution analysis** *with scaling function* ϕ, $\int \phi(t)dt = 1$, *is a sequence of subspaces* V_j *of* $L_2(\mathbf{R}^n)$, $j \in \mathbf{Z}$, *satisfying:*
1. $V_j \subset V_{j+1}$, $\bigcap V_j = \{0\}$ *and* $\overline{\bigcup V_j} = L^2(\mathbf{R}^n)$.
2. $f(t) \in V_j \iff f(Mt) \in V_{j+1}$.
3. $f(t) \in V_0 \iff f(t-k) \in V_0$, $k \in \Gamma$.
4. $\{\phi(t-k), k \in \Gamma\}$ *is a Riesz basis of* V_0.

Note: for notation purposes, the following convention will be used:

$$f_{j,k}(t) = m^{j/2} f(M^j t - k),\ \ k \in \Gamma,\ \ j \in \mathbf{Z} \tag{60}$$

It follows from the definition of a multiresolution analysis that $\{\phi_{j,k}(t), k \in \Gamma\}$ is a Riesz basis of V_j.

Definition 8.3 *Consider a multiresolution analysis with lattice* Γ *and dilation matrix* M*. For* $j \in \mathbf{Z}$*, let* W_0 *be such that* V_1 *is the direct sum of* V_0 *and* W_0*. Assume there are* $\psi^1, ..., \psi^{m-1} \in W_0$ *such that*
$\{\psi^{\mu}_{0,k}, \mu = 1, ..., m-1,\ k \in \Gamma\ \}$ *is a Riesz basis of* W_0*. The* ψ^{μ} *are called* **wavelets**.

For $j \in \mathbf{Z}$, let W_j be the subspace with Riesz basis $\{\Psi^{\mu}_{j,k}, k \in \Gamma, \mu = 1, ..., m-1\}$. It follows that $V_{j+1} = V_j \oplus W_j$, a direct sum but not necessarily orthogonal.

Definition 8.4 *Let* V_j, $\widetilde{V}_j$ *be two multiresolution analyses corresponding to the same lattice* Γ *and dilation matrix* M*. Let* ϕ, ψ^{μ}, $\mu = 1, ..., m-1$, *be the scaling function and wavelets corresponding to* V_j*. Let* $\tilde{\phi}$, $\tilde{\psi}^{\mu}$, $\mu = 1, ..., m-1$, *be the scaling function and wavelets corresponding to* $\widetilde{V}_j$*. The multiresolution analyses are called* **biorthogonal** *if the following conditions hold for* $j, j' \in \mathbf{Z}$, $\mu, \mu' = 1, ..., m-1$, *and* $k, k' \in \Gamma$

i) $< \psi^{\mu}_{j,k}, \widetilde{\psi}^{\mu'}_{j',k'} >= \delta(j,j')\delta(\mu,\mu')\delta(k,k')$

ii) $< \widetilde{\psi}^{\mu}_{j,k}, \phi_{j,k'} >=< \widetilde{\phi}_{j,k}, \psi^{\mu}_{j,k'} >= 0$

iii) $< \phi_{j,k}, \widetilde{\phi}_{j,k'} >= \delta(k,k')$

For $n = 1$, $M = 2$, $\Gamma = \mathbf{Z}$, and $\phi = \tilde{\phi}$, the multiresolution analysis becomes the familiar one-dimensional, orthonormal case.

Since $\phi \in V_0 \subset V_1$, and $\{\phi_{1,k}, k \in \Gamma\}$ is a basis for V_1, there are coefficients $F_0(k)$, $k \in \Gamma$ such that

$$\phi(t) = m^{1/2} \sum_{k \in \Gamma} F_0(k) \phi_{1,k}(t)$$

¿From condition *iii)* above, it follows that $F_0(k) = m^{-1/2} < \phi, \widetilde{\phi}_{1,k} >$. Similarly, since $\widetilde{\phi} \in \widetilde{V}_0 \subset \widetilde{V}_1$, and $\{\widetilde{\phi}_{1,k}, k \in \Gamma\}$ is a basis for $\widetilde{V}_1$,

$$\widetilde{\phi}(t) = m^{1/2} \sum_{k \in \Gamma} H_0(k) \widetilde{\phi}_{1,k}(t)$$

where $H_0(k) = m^{-1/2} < \widetilde{\phi}, \phi_{1,k} >$. The above equations are called dilation equations. Similarly, the wavelets ψ^{μ}, $\widetilde{\psi}^{\mu}$ must satisfy so-called wavelet equations:

$$\psi^{\mu}(t) = m^{1/2} \sum_{k \in \Gamma} F_{\mu}(k) \phi_{1,k}(t)$$

$$\widetilde{\psi}^{\mu}(t) = m^{1/2} \sum_{k\in\Gamma} H_{\mu}(k)\widetilde{\phi}_{1,k}(t)$$

where $\mu = 1, ..., m-1$, $F_{\mu}(k) = m^{-1/2} < \psi^{\mu}, \widetilde{\phi}_{1,k} >$ and $H_{\mu}(k) = m^{-1/2} < \widetilde{\psi}^{\mu}, \phi_{1,k} >$.

The following lemma and theorems show how to decompose a function f into its wavelet coefficients and how to reconstruct f if its wavelet coefficients are known.

Lemma 8.5 *For $j \in \mathbf{Z}$ and $\mu = 1, ..., m-1$,*

$$\phi_{j,l}(t) = m^{1/2} \sum_{k\in\Gamma} F_0(k)\phi_{j+1,Ml+k}(t) \tag{61}$$

$$\widetilde{\phi}_{j,l}(t) = m^{1/2} \sum_{k\in\Gamma} H_0(k)\widetilde{\phi}_{j+1,Ml+k}(t) \tag{62}$$

$$\psi^{\mu}_{j,l}(t) = m^{1/2} \sum_{k\in\Gamma} F_{\mu}(k)\phi_{j+1,Ml+k}(t) \tag{63}$$

$$\widetilde{\psi}^{\mu}_{j,l}(t) = m^{1/2} \sum_{k\in\Gamma} H_{\mu}(k)\widetilde{\phi}_{j+1,Ml+k}(t) \tag{64}$$

Proof: Follows directly from the dilation and wavelet equations. □

Theorem 8.6 (Fast Wavelet Transform) *Let $j \in \mathbf{Z}$ and $f \in V_{j+1}$. For $\mu = 1, ..., m-1$, and $k \in \Gamma$, define the approximation coefficients as*

$$\widetilde{A}_{j,k} =< f, \widetilde{\phi}_{j,k} >= (f * \overline{\widetilde{\phi}^{\vee}_{j,0}})(M^{-j}k) \tag{65}$$

and the detail coefficients as

$$\widetilde{D}^{\mu}_{j,k} =< f, \widetilde{\psi}^{\mu}_{j,k} >= (f * \overline{\widetilde{\psi}^{\mu}_{j,0}{}^{\vee}})(M^{-j}k) \tag{66}$$

Then

$$\widetilde{A}_{j,k} = m^{1/2} \sum_{l\in\Gamma} \overline{H_0(l-Mk)}\widetilde{A}_{j+1,l} \tag{67}$$

and

$$\widetilde{D}^{\mu}_{j,k} = m^{1/2} \sum_{l\in\Gamma} \overline{H_{\mu}(l-Mk)}\widetilde{A}_{j+1,l} \tag{68}$$

Proof: Consider the following expansion of f:

$$f(t) = \sum_{l\in\Gamma} \widetilde{A}_{j+1,l}\phi_{j+1,l}(t) \tag{69}$$

For $\widetilde{A}_{j,k}$, take an inner product of (69) with $\widetilde{\phi}_{j,k}$. Use (62) and biorthogonality to get (67).
For $\widetilde{D}^{\mu}_{j,k}$, take an inner product of (69) with $\widetilde{\psi}^{\mu}_{j,k}$. Use (64) and biorthogonality to get (68). □

Theorem 8.7 (Inverse Fast Wavelet Transform) *Under the hypothesis of Theorem 8.6*

$$\widetilde{A}_{j+1,l} = m^{1/2} \sum_{k\in\Gamma} \left(F_0(l - Mk)\widetilde{A}_{j,k} + \sum_{\mu=1}^{m-1} F_\mu(l - Mk)\widetilde{D}^{\mu}_{j,k} \right) \tag{70}$$

Proof: Consider the following expansions of f

$$\begin{aligned} f(t) &= \sum_{k\in\Gamma} \widetilde{A}_{j+1,k}\phi_{j+1,k}(t) \\ &= \sum_{k\in\Gamma} \widetilde{A}_{j,k}\phi_{j,k}(t) + \sum_{\mu=1}^{m-1} \sum_{k\in\Gamma} \widetilde{D}^{\mu}_{j,k}\psi^{\mu}_{j,k}(t) \end{aligned} \tag{71}$$

The second expansion comes from writing $f \in V_{j+1} = V_j \oplus W_j$ as a sum of elements of V_j and W_j. To get $\widetilde{A}_{j+1,l}$, take an inner product of (71) with $\widetilde{\phi}_{j+1,l}$. Use (61), (63) and biorthogonality to get (70). □

An easy way to obtain wavelets in $\mathbf{R}^n$ is to use a tensor product construction with the wavelets in $\mathbf{R}$. We will look specifically at the two-dimensional case. Define the lattice as $\Gamma = \mathbf{Z}^2$, and the dilation matrix $M = 2I$, where I denotes the identity matrix. Since $|\det M| = 4 = m$, one can expect 3 wavelets and 1 scaling function. Let the spaces V_j, W_j be the chosen one-dimensional multiresolution analysis with scaling function ϕ, and wavelet ψ. The coefficients for the dilation and wavelet equation are $F_\mu(k)$, $\mu = 0, 1$. Constructing the two-dimensional scaling function and wavelets by taking products of the one-dimensional functions leads to the following definition.

Definition 8.8 *¿From a one-dimensional scaling function $\phi(x)$ and its corresponding wavelet $\psi(x)$,* **two-dimensional separable wavelets** *are defined for $(x, y) \in \mathbf{R}^2$,*

$$\begin{aligned} \Phi(x, y) &= \phi(x)\phi(y) \\ \Psi^1(x, y) &= \phi(x)\psi(y) \\ \Psi^2(x, y) &= \psi(x)\phi(y) \\ \Psi^3(x, y) &= \psi(x)\psi(y) \end{aligned}$$

Consider a biorthogonal pair of one-dimensional multiresolution analyses. Recall for separable, two-dimensional wavelets, $M = 2I$, $m = 4$ and $\Gamma = \mathbf{Z}^2$. Let V_j and W_j be the multiresolution analysis with scaling function ϕ, wavelet ψ, and coefficients F_0, F_1 for the dilation and wavelet equations. Let $\widetilde{V}_j$ and $\widetilde{W}_j$ be the multiresolution analysis with scaling function $\widetilde{\phi}$ wavelet $\widetilde{\psi}$ and coefficients H_0, H_1 for the dilation and wavelet equations.

We want to rewrite the fast wavelet transform and inverse fast wavelet transform for the case of two-dimensional separable wavelets. Let $\mathbf{F}_\mu(\mathbf{k})$, $\mu = 0, 1, 2, 3$ be the coefficients in the dilation and wavelet equations for the separable wavelets. It is easy to verify that

$$\begin{aligned}
\mathbf{F}_0(\mathbf{k}) &= F_0(k_1)F_0(k_2) \\
\mathbf{F}_1(\mathbf{k}) &= F_0(k_1)F_1(k_2) \\
\mathbf{F}_2(\mathbf{k}) &= F_1(k_1)F_0(k_2) \\
\mathbf{F}_3(\mathbf{k}) &= F_1(k_1)F_1(k_2)
\end{aligned}$$

where $\mathbf{k} = (k_1, k_2)$. Similarly for the $\mathbf{H}_\mu(\mathbf{k})$, $\mu = 0, 1, 2, 3$. Thus, equation (67) becomes:

$$\widetilde{A}_{j,k} = 2 \sum_{l \in \Gamma} \overline{\mathbf{H}_0(l - 2k)} \widetilde{A}_{j+1,l} \tag{72}$$

and equation (68) becomes

$$\widetilde{D}^{\mu}_{j,k} = 2 \sum_{l \in \Gamma} \overline{\mathbf{H}_\mu(l - 2k)} \widetilde{A}_{j+1,l}, \text{ for } \mu = 1, 2, 3 \tag{73}$$

The inverse wavelet transform, equation (70), becomes:

$$\begin{aligned}
\widetilde{A}_{j+1,n} &= m^{1/2} \sum_{k \in \Gamma} \left(\mathbf{F}_0(n - Mk)\widetilde{A}_{j,k} + \sum_{\mu=1}^{3} \mathbf{F}_\mu(n - Mk)\widetilde{D}^{\mu}_{j,k} \right) \\
&= 2 \sum_{k_1, k_2 \in \mathbf{Z}} \Big(F_0(n_1 - 2k_1)F_0(n_2 - 2k_2)\widetilde{A}_{j,k} + F_0(n_1 - 2k_1)F_1(n_2 - 2k_2)\widetilde{D}^{1}_{j,k} \\
&\quad + F_1(n_1 - 2k_1)F_0(n_2 - 2k_2)\widetilde{D}^{2}_{j,k} + F_1(n_1 - 2k_1)F_1(n_2 - 2k_2)\widetilde{D}^{3}_{j,k} \Big)
\end{aligned} \tag{74}$$

Recall that the approximation coefficients are a convolution of f with $\overline{\widetilde{\phi}_{j,0}}$: $\widetilde{A}_{j,k} = (f * \overline{\widetilde{\phi}^{\vee}_{j,0}})(M^{-j}k)$. We would like to consider, for sufficiently large j, the approximation coefficients as an approximation for f at a particular point. Notice that for $M = 2I$, $M^j x = 2^j x$, and $m = 2^n$. From Real Analysis, the following lemma holds.

Lemma 8.9 *Let f be continuous, $g \in L^1(\mathbf{R}^n)$ with $\int g dx = 1$, g is bounded and has compact support. Then for all $x \in \mathbf{R}^n$*

$$f(x) = \lim_{j \to \infty} 2^{jn} \int f(x + y)\overline{g(2^j y)} dy = \lim_{j \to \infty} 2^{j/2} f * \overline{g^{\vee}_{j,0}}(x) \tag{75}$$

Proof: Let $t = 2^{-j}$. Then we have a special case of Theorem 8.15, page 235, in [19]. □

Thus, for j sufficiently large, we get an approximation for f:

$$2^{j/2} f * \overline{\widetilde{\phi}^{\vee}_{j,0}}(2^{-j}k) \approx f(2^{-j}k)$$

or, with $n = 2$

$$2^{j/2}\widetilde{A}_{j,k} \approx f(2^{-j}k)$$

9 Appendix B: Matlab code for the filtered back-projection algorithm

Below we give a few MATLAB M-files which implement the parallel-beam filtered backprojection algorithm for the standard lattice. *This source code is provided for illustrative purposes only and comes without warranties of any kind.* The main file fbp.m is supplemented by three function M-files. Rad.m computes line integrals for a mathematical phantom consisting of ellipses, slkernel.m computes the discrete Shepp-Logan convolution kernel (cf. (49)), and window3.m allows to view the reconstructed image. It is automatically called at the end of the reconstruction. However, with the example phantom given in fbp.m the picture shown does not display the most interesting details. It is better to call window3 again with the parameters window3(-0.07,0.07,roi,P).

```
%fbp.m
%Parallel-beam filtered backprojection algorithm
% for the standard lattice

%specify input parameters here

p=200; %number of view angles between 0 and pi
q=64; %q=1/d, d = detector spacing
MX=128;  MY = 128; %matrix dimensions
roi=[-1  1 -1  1]; %roi=[xmin xmax ymin ymax]
                   %region of interest where
                   %reconstruction is computed
circle = 1; % If circle = 1 image computed only inside
            % circle inscribed in roi.


%Specify parameters of ellipses for mathematical phantom.
% xe = vector of x-coordinates of centers
% ye = vector of y-coordinates of centers
% ae = vector of first half axes
% be = vector of second half axes
% alpha = vector of rotation angles (degrees)
% rho = vector of densities
xe=[0 0 0.22 -0.22 0 0 0 -0.08 0 0.06 0.5538];
ye=[0 -0.0184 0 0 0.35 0.1 -0.1 -0.605 -0.605 -0.605 -0.3858];
ae=[0.69 0.6624 0.11 0.16 0.21 0.046 0.046 0.046 0.023 0.023 0.0333];
be=[0.92 0.874 0.31 0.41 0.25 0.046 0.046 0.023 0.023 0.046 0.206];
alpha=[0 0 -18 18 0 0 0 0 0 0 -18];
```

```
rho=  [1 -0.98 -0.02 -0.02 0.01 0.01 0.01 0.01 0.01 0.01 0.03];

%end of input section

b=pi*q; rps=1/b;
alpha = alpha*pi/180;

if MX > 1
 hx = (roi(2)-roi(1))/(MX-1);
 xrange = roi(1) + hx*[0:MX-1];
else
 hx = 0; xrange = roi(1);
end

if MY > 1
 hy = (roi(4)-roi(3))/(MY-1);
 yrange = flipud((roi(3) + hy*[0:MY-1])');
else
 hx = 0; yrange = roi(3);
end

center = [(roi(1)+roi(2)), (roi(3)+roi(4))]/2;
x1 = ones(MY,1)*xrange; %x-coordinate matrix
x2 = yrange*ones(1,MX); %y-coordinate matrix
if circle == 1
  re = min([roi(2)-roi(1),roi(4)-roi(3)])/2;
  chi = ((x1-center(1)).^2 + (x2-center(2)).^2 <= re^2); %char. fct of roi;
else
  chi = ones(size(x1));
end
x1 = x1(chi); x2 = x2(chi);
P = zeros(MY,MX);Pchi = P(chi);


h = 1/q;
s = h*[-q:q-1];
bs = [-2*q:2*q-1]/(q*rps);
wb = slkernel(bs)/(rps^2); %compute discrete convolution kernel.
for j = 1:p
  j
  phi = (pi*(j-1)/p);
```

```
  theta = [cos(phi);sin(phi)];
  RF = Rad(theta,phi,s,xe,ye,ae,be,alpha,rho); %compute line integrals

%    convolution
  C = conv(RF,wb);
  Q = h*C(2*q+1:4*q); Q(2*q+1)=0;


%    interpolation and backprojection

     Q = [real(Q)'; 0];
     t = theta(1)*x1 + theta(2)*x2;
     k1 = floor(t/h);
     u = (t/h-k1);
     k = max(1,k1+q+1); k = min(k,2*q);
     Pupdate =  ((1-u).*Q(k)+u.*Q(k+1));
     Pchi=Pchi+ Pupdate;

end   % j-loop

P(chi) = Pchi*(2*pi/p);

pmin = min(min(P));
pmax = max(max(P));
window3(pmin,pmax,roi,P); % view the computed image


% --- Cut here for Rad.m

function [RF] = Rad(theta,phi,s,x,y,u,v,alpha,rho)

%  This function computes the Radon transform of ellipses
% centered at (x,y) with major axis u, minor axis v,
% rotated through angle alpha, with weight rho.

RF = zeros(size(s));

for mu = 1:max(size(x))
   a = (u(mu)*cos(phi-alpha(mu)))^2+(v(mu)*sin(phi-alpha(mu)))^2;
   test = a-(s-[x(mu);y(mu)]'*theta).^2;
   ind = test>0;
```

```
   RF(ind) = RF(ind)+rho(mu)*(2*u(mu)*v(mu)*sqrt(test(ind)))/a;
end  % mu-loop


% --- Cut here for slkernel.m

function u = slkernel(t)

u = zeros(size(t));
i1 = abs(abs(t)-pi/2)<=1.e-6;
u(i1) = ones(size(u(i1)))/pi;
t1 = t(abs(abs(t)-pi/2)>1.e-6);
v = (pi/2 - t1.*sin(t1))./((pi/2)^2 - t1.^2);
u(abs(abs(t)-pi/2)>1.e-6) = v;
u = u/(2*pi^3);

% --- Cut here for window3.m

function pic1 = window3(mi,ma,roi,pic);
%function pic1 = window3(mi,ma,roi,pic);
% displays image pic with coordinates given by roi
% roi = [xmin xmax ymin ymax]
x = [roi(1), roi(2)]; y = [roi(3), roi(4)];
colors = 128; co = colors-1;
pic1 = pic - mi*ones(size(pic));
pic1 = (co/(ma-mi))*pic1;
P = (pic1 >= 0);
pic1 = pic1.*P;
P = (pic1 <= co);
pic1 = pic1.*P + co*(ones(size(pic1)) - P);
colormap(gray(colors));
image(x,fliplr(y),flipud(pic1));
axis('square');
```

References

[1] J. Boman and E.T. Quinto, *Support theorems for real-analytic Radon transforms on line complexes in three-space*, Trans. Amer. Math. Soc., 335(1993), pp. 877-890.

[2] C. Berenstein and D. Walnut, *Local inversion of the Radon transform in even dimensions using wavelets*, in [21, pp. 45-69].

[3] A. M. Cormack, *Representation of a function by its line integrals, with some radiological applications*, J. Appl. Phys., 34 (1963), pp. 2722-2727.

[4] K. Buglione, *Pseudolocal tomography*, M.S. paper, Dept. of Mathematics, Oregon State University, Corvallis, OR 97331, U.S.A., (1998).

[5] A. M. Cormack, *Sampling the Radon transform with beams of finite width*, Phys. Med. Biol., 23 (1978), pp. 1141-1148.

[6] S. R. Deans, *The Radon transform and some of its applications*, Wiley, 1983.

[7] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Capital City Press, VT, 1992

[8] M. Defrise and R. Clack, *A cone-beam reconstruction algorithm using shift–variant filtering and cone–beam backprojection*, IEEE Trans. Med. Imag., MI-13 (1994), pp. 186-195.

[9] L. Desbat, *Efficient sampling on coarse grids in tomography*, Inverse Problems, 9(1993), pp. 251-269.

[10] A. Faridani, *An application of a multidimensional sampling theorem to computed tomography*, in [27, pp. 65-80].

[11] A. Faridani, *Reconstructing from efficiently sampled data in parallel-beam computed tomography*, in G.F. Roach (ed.), Inverse problems and imaging, Pitman Research Notes in Mathematics Series, Vol. 245, Longman, 1991, pp. 68-102.

[12] A. Faridani, *Results, old and new, in computed tomography*, in: Inverse Problems in Wave Propagation, G. Chavent et al. (editors), The IMA Volumes in Mathematics and its Applications, Vol. 90, Springer Verlag, New York, 1997, pp. 167-193.

[13] A. Faridani, *Sampling in parallel-beam tomography*, in: Inverse Problems and Imaging, A.G. Ramm (editor), Plenum, New York, 1998, pp. 33-53.

[14] A. Faridani, E. L. Ritman, and K. T. Smith, *Local tomography*, SIAM J. Appl. Math., 52 (1992) , pp. 459-484. *Examples of local tomography*, SIAM J. Appl. Math., 52 (1992) , pp. 1193-1198.

[15] A. Faridani, D.V. Finch, E. L. Ritman, and K. T. Smith, *Local tomography II*, SIAM J. Appl. Math., 57 (1997), pp. 1095-1127.

[16] A. Faridani and E. L. Ritman, *High-resolution computed tomography*, Inverse Problems, 16(2000), pp. 635-650.

[17] L. A. Feldkamp, L. C. Davis, and J. W. Kress, *Practical cone–beam algorithm*, J. Opt. Soc. Am. A, 1 (1984), pp. 612-619.

[18] D. V. Finch, *Cone beam reconstruction with sources on a curve*, SIAM J. Appl. Math., 45(1985), pp. 665-673.

[19] G. Folland, *Real Analysis–Modern Techniques and their Applications* John Wiley and Sons, Inc, (1984)

[20] I.M. Gelfand and S. G. Gindikin (eds.), *Mathematical problems of tomography*, Translations of Mathematical Monographs, Vol. 81, Amer. Math. Soc., 1990.

[21] S. Gindikin and P. Michor (eds.), *75 Years of Radon Transform*, Conference Proceedings and Lecture Notes in Mathematical Physics, Vol. 4, International Press, Boston, 1994.

[22] P. Grangeat, *Mathematical framework of cone beam 3D reconstruction via the first derivative of the Radon transform*, in [33, pp. 66-97].

[23] A. Greenleaf and G. Uhlmann, *Nonlocal inversion formulas for the X-ray transform*, Duke Math. J., 58(1989), pp. 205-240.

[24] A. Greenleaf and G. Uhlmann, *Estimates for singular Radon transforms and pseudodifferential operators with singular symbols*, J. Funct. Anal., 89(1990), pp. 202-232.

[25] A. Greenleaf and G. Uhlmann, *Composition of some singular Fourier integral operators and estimates for restricted X-ray transforms.*, Ann. Inst. Fourier, 40(1990), pp. 443-466.

[26] A. Greenleaf and G. Uhlmann, *Microlocal techniques in integral geometry.*, in [27, pp. 121-135].

[27] E. Grinberg and E.T. Quinto (eds.), *Integral Geometry and Tomography*, Contemporary Mathematics, Vol. 113, Amer. Math. Soc., Providence, R.I., 1990.

[28] C. W. Groetsch, *Inverse problems in the mathematical sciences*, Vieweg, Braunschweig, 1993.

[29] S. Gutmann, J .H. B. Kemperman, J. A. Reeds, and L. A. Shepp, *Existence of probability measures with given marginals*, Annals of Probability, 19 (1991), pp. 1781-1797.

[30] S. Helgason, *The Radon transform*, Birkhäuser, 1980.

[31] G. T. Herman, *Image reconstruction from projections: the fundamentals of computerized tomography*, Academic Press, New York, 1980.

[32] G. T. Herman (ed.), *Image reconstruction from projections: implementation and applications*, Springer, 1979.

[33] G.T. Herman, A.K. Louis, and F. Natterer (eds.), *Mathematical Methods in Tomography*, Lecture Notes in Mathematics, Vol. 1497, Springer, 1991.

[34] O. D. Iancu, *Contour reconstruction in 3D x-ray computed tomography.* M.S. paper, Dept. of Mathematics, Oregon State University, Corvallis, OR 97331, U.S.A., (1999).

[35] A. C. Kak and M. Slaney, *Principles of computerized tomographic imaging*, IEEE Press, New York, 1988.

[36] A. I. Katsevich, *Local Tomography for the generalized Radon transform*, SIAM J. Appl. Math. 57(1997), pp. 1128-1162.

[37] A. Katsevich, *Cone beam local tomography*, SIAM J. Appl. Math, 59(1999), pp. 2224-2246.

[38] A.I. Katsevich and A. G. Ramm, *A method for finding discontinuities of functions from the tomographic data*, in [74, pp. 115-123].

[39] A.I. Katsevich and A. G. Ramm, *New methods for finding jumps of a function from its local tomographic data*, Inverse Problems, 11 (1995), pp. 1005-1023.

[40] A.I. Katsevich and A. G. Ramm, *Pseudolocal tomography*, SIAM J. Appl. Math., 56, (1996), pp. 167-191.

[41] F. Keinert, *Inversion of k-plane transforms and applications in computer tomography*, SIAM Review, 31(1989), pp. 273-298.

[42] W. Klaverkamp, *Tomographische Bildrekonstruktion mit direkten algebraischen Verfahren*, Ph.D. thesis, Fachbereich Mathematik der Universität Münster, Münster, Germany, 1991.

[43] H. Kruse, *Resolution of reconstruction methods in computerized tomography*, SIAM J. Sci. Stat. Comput.,10(1989), pp. 447-474.

[44] P. Kuchment, K. Lancaster and L. Mogilevskaya, *On local tomography*, Inverse Problems, 11 (1995), pp. 571-589.

[45] A. V. Lakshminarayanan, *Reconstruction from divergent x-ray data*, Suny Tech. Report 32, Comp. Sci. Dept., State University of New York, Buffalo, NY, 1975.

[46] I. Lan, *On an operator associated to a restricted x-ray transform*, Ph.D. thesis, Dept. of Mathematics, Oregon State University, Corvallis, OR 97331, U.S.A., (1999).

[47] J. V. Leahy, K. T. Smith, and D. C. Solmon, *Uniqueness, nonuniqueness and inversion in the x-ray and Radon problems*, submitted to Proc. Internat. Symp. on Ill-posed Problems, Univ. Delaware Newark, 1979. The proceedings did not appear. Some results of this article have been published in [90].

[48] R. M. Lewitt, *Reconstruction algorithms: transform methods*, Proc. IEEE, 71 (1983), pp. 390-408.

[49] R. M. Lewitt, R. H. T. Bates, and T. M. Peters, *Image reconstruction from projections: II: Modified back-projection methods*, Optik, 50 (1978), pp. 85-109.

[50] B. F. Logan, *The uncertainty principle in reconstructing functions from projections*, Duke Math. J., 42 (1975), pp. 661-706.

[51] A. K. Louis, *Nonuniqueness in inverse Radon problems: the frequency distribution of the ghosts*, Math. Z. 185 (1984), pp. 429-440.

[52] A. K. Louis, *Orthogonal function series expansions and the null space of the Radon transform*, SIAM J. Math. Anal., 15(1984), pp. 621-633.

[53] A. K. Louis, *Incomplete data problems in x-ray computerized tomography I. Singular value decomposition of the limited angle transform*, Numer. Math. 48(1986), pp. 251-262.

[54] A. K. Louis, *Inverse und schlecht gestellte Probleme*, B.G. Teubner, Stuttgart, 1989.

[55] A. K. Louis, *Medical imaging: state of the art and future development*, Inverse Problems 8(1992), pp. 709-738.

[56] A. K. Louis, and P. Maass, *Contour reconstruction in 3-D x-ray CT*, IEEE Trans. Med. Imag., MI-12 (1993), pp. 764–769.

[57] A. K. Louis and F. Natterer, *Mathematical problems in computerized tomography*, Proc. IEEE, 71 (1983), pp. 379-389.

[58] W. R. Madych, *Summability and approximate reconstruction from Radon transform data*, in [27, pp. 189-219].

[59] P. Maass, *The x-ray transform: singular value decomposition and resolution*, Inverse Problems, 3 (1987), pp. 729-741.

[60] P. Maass, *The interior Radon transform*, SIAM J. Appl. Math. 52(1992), pp. 710-724.

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

[62] F. Natterer, *Sampling in fan-beam tomography*, SIAM J. Appl. Math. 53(1993), pp. 358-380.

[63] F. Natterer, *Recent developments in x-ray tomography*, in [74, pp. 177-198].

[64] F. Natterer, *Algorithms in Tomography*, Preprint, (1997). Available online at wwwmath.uni-muenster.de/math/inst/num

[65] F. Natterer, *Numerical methods in tomography*, in Acta Numerica, Vol. 8, 1999, Cambridge University Press, New York, pp. 107-141.

[66] T. Olson and J. de Stefano, *Wavelet localization of the Radon transform*, IEEE Trans. Sig. Proc., 42 (1994), pp. 2055-2067 .

[67] D. P. Petersen and D. Middleton, *Sampling and reconstruction of wave-number-limited functions in N-dimensional euclidean space*, Inf. Control, 5(1962), pp. 279-323.

[68] R. M. Perry, *On reconstructing a function on the exterior of a disc from its Radon transform*, J. Math. Anal. Appl. 59 (1977), pp. 324-341.

[69] D. A. Popov, *On convergence of a class of algorithms for the inversion of the numerical Radon transform*, in [20, pp. 7-65].

[70] E.T. Quinto, *The dependence of the generalized Radon transform on defining measures*, Trans. Amer. Math. Soc. 257 (1980), pp. 331-346.

[71] E. T. Quinto, *Singular value decompositions and inversion methods for the exterior Radon transform and a spherical transform*, J. Math. Anal. Appl., 95 (1983), pp.437-448.

[72] E.T. Quinto, *Tomographic reconstructions from incomplete data – numerical inversion of the exterior Radon transform*, Inverse Problems, 4 (1988), pp. 867-876.

[73] E. T. Quinto, *Singularities of the x-ray transform and limited data tomography in* $\mathbb{R}^2$ *and* $\mathbb{R}^3$, SIAM J. Math. Anal., 24 (1993), pp. 1215-1225.

[74] E.T. Quinto, M. Cheney, and P. Kuchment (eds.), *Tomography, Impedance Imaging, and Integral Geometry*, Lectures in Applied Mathematics, Vol. 30, Amer. Math. Soc., 1994.

[75] A. G. Ramm and A. I. Katsevich, *The Radon Transform and Local Tomography*, CRC Press, Boca Raton, 1996.

[76] A. G. Ramm and A.I. Zaslavsky, *Reconstructing singularities of a function given its Radon transform*, Math. and Comput. Modelling, 18(1993), pp. 109-138.

[77] F. Rashid-Farrokhi, K. J. R. Liu, C. A. Berenstein, and D. Walnut, *Wavelet-based multiresolution local tomography*, IEEE Transactions on Image Processing, 6 (1997), pp. 1412-1430.

[78] P. A. Rattey and A. G. Lindgren, *Sampling the 2-D Radon transform*, IEEE Trans. Acoust. Speech Signal Processing, ASSP-29(1981), pp. 994-1002.

[79] M.G. Raymer, M. Beck, and D.F. McAllister, *Complex wave-field reconstruction using phase-space tomography*, Phys. Rev. Lett., 72(1994), pp. 1137-1140.

[80] E. L. Ritman, J. H. Dunsmuir, A. Faridani, D. V. Finch, K. T. Smith, and P. J. Thomas, *Local reconstruction applied to microtomography*, in: Inverse Problems in Wave Propagation, G. Chavent et al. (editors), The IMA Volumes in Mathematics and its Applications, Vol. 90, Springer Verlag, New York, 1997, pp. 443-452.

[81] L. A. Shepp and J. B. Kruskal, *Computerized tomography: the new medical x-ray technology*, Amer. Math. Monthly, 85 (1978), pp. 420-439.

[82] E. A. Sivers, D. L. Halloway, W. A. Ellingson, and J. Ling, *Development and application of local 3-D CT reconstruction software for imaging critical regions in large ceramic turbine rotors*, in Rev. Prog. Quant. Nondest. Eval.:, D.O. Thompson and D.E. Chimenti (eds.), Plenum, New York, 1993, pp. 357-364.

[83] E. A. Sivers, D. L. Halloway, W. A. Ellingson, *Obtaining high-resolution images of ceramics from 3-D x-ray microtomography by region-of-interest reconstruction*, Ceramic Eng. Sci. Proc., 14, no. 7-8, (1993), pp. 463-472.

[84] J. Skaggs, *Region of interest tomography using biorthogonal wavelets*, M.S. paper, Dept. of Mathematics, Oregon State University, Corvallis, OR 97331, U.S.A., (1997).

[85] K. T. Smith, *Reconstruction formulas in computed tomography*, Proc. Sympos. Appl. Math., No. 27, L. A. Shepp, ed., AMS, Providence, RI, (1983), pp. 7-23.

[86] K. T. Smith, D.C. Solmon, and S. L. Wagner, *Practical and mathematical aspects of the problem of reconstructing objects from radiographs*, Bull. Amer. Math. Soc., 83(1977), pp. 1227-1270. Addendum in Bull. Amer. Math. Soc., 84(1978), p. 691.

[87] K. T. Smith and F. Keinert, *Mathematical foundations of computed tomography*, Appl. Optics 24 (1985), pp. 3950-3957.

[88] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani, *Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne*

tomography: application to squeezed states and the vacuum, Phys. Rev. Lett., 70 (1993), pp. 1244-1247.

[89] D. C. Solmon, *The x-ray transform*, J. Math. Anal. Appl., 56(1976), pp. 61-83.

[90] D. C. Solmon, *Nonuniqueness and the null space of the divergent beam x-ray transform*, in [27, pp. 243-249].

[91] W. J. T. Spyra, A. Faridani, E. L. Ritman, and K. T. Smith, *Computed tomographic imaging of the coronary arterial tree - use of local tomography*, IEEE Trans. Med. Imag., 9 (1990), pp. 1-4.

[92] G. Strang, T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley, MA, (1996).

[93] R. Strichartz, *Construction of Orthonormal Wavelets*, in *Wavelets: Mathematics and Applications*, pg. 23-50, Editors: L. Benedetto, M. Frazier, CRC Press, Inc, Baton Rouge, FL, (1994).

[94] H. K. Tuy, *An inversion formula for cone beam reconstruction.* SIAM J. Appl. Math. 43(1983), pp. 546-552.

[95] É. I. Vainberg, I. A. Kazak, and V. P. Kurozaev, *Reconstruction of the internal three-dimensional structure of objects based on real-time internal projections*, Soviet J. Nondestructive Testing, 17 (1981), pp. 415-423.

[96] É. I. Vainberg, I. A. Kazak, and M. L. Faingoiz, *X-ray computerized back projection tomography with filtration by double differentiation. Procedure and information features*, Soviet J. Nondestructive Testing, 21 (1985), pp. 106-113.

[97] R. C. Vaughn, *The parallel-beam filtered backprojection algorithm with reconstruction on a polar grid*, M.S. paper, Dept. of Mathematics, Oregon State University, Corvallis, OR 97331, U.S.A., (1994).

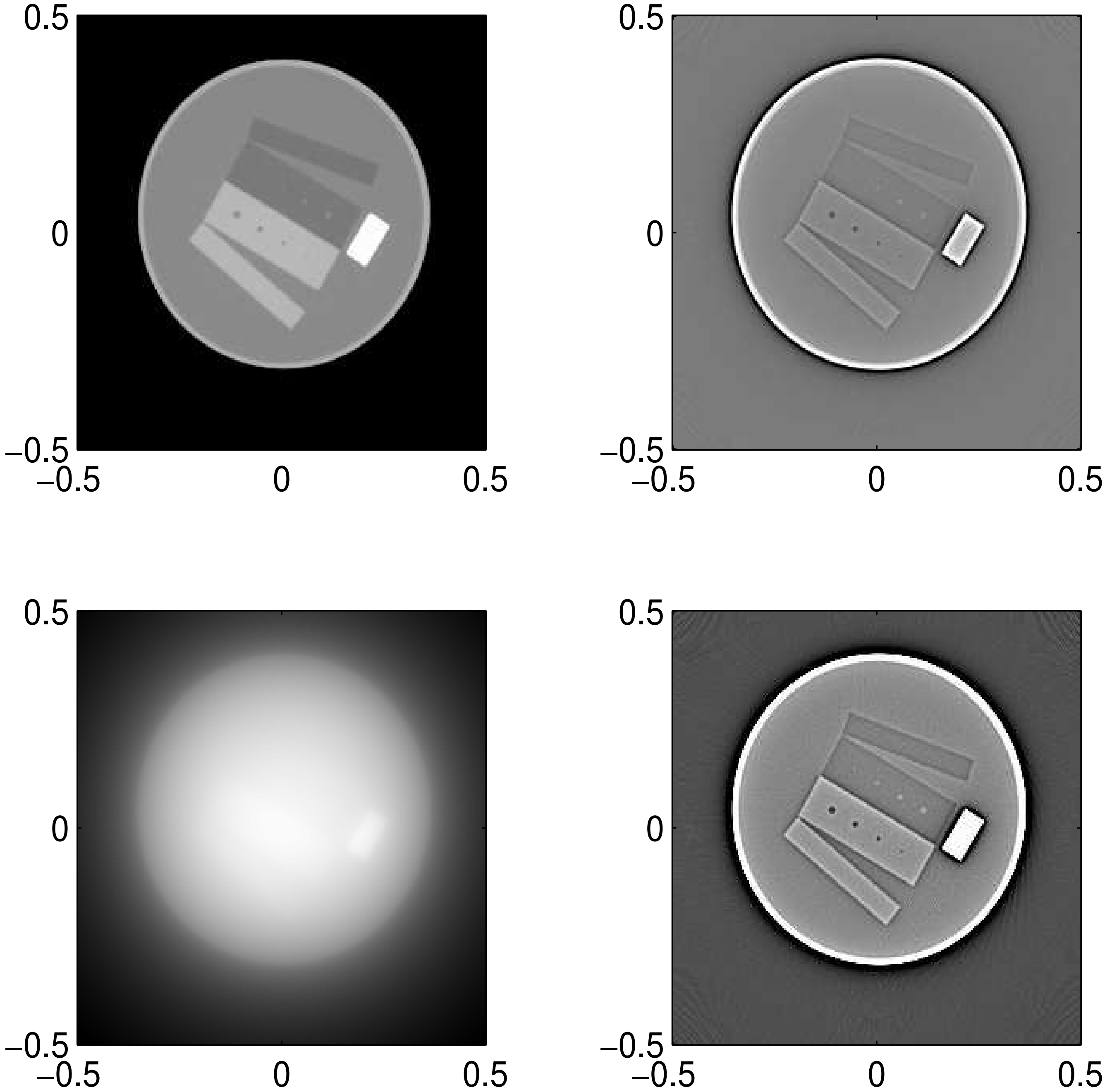

Figure 1: Top left: Global reconstruction of density $f(x)$ of calibration object. Top right: Reconstruction of Λf . Bottom left: Reconstruction of $\Lambda^{-1}f$. Bottom right: Reconstruction of $Lf = \Lambda f + \mu\Lambda^{-1}f$, $\mu = 46$.

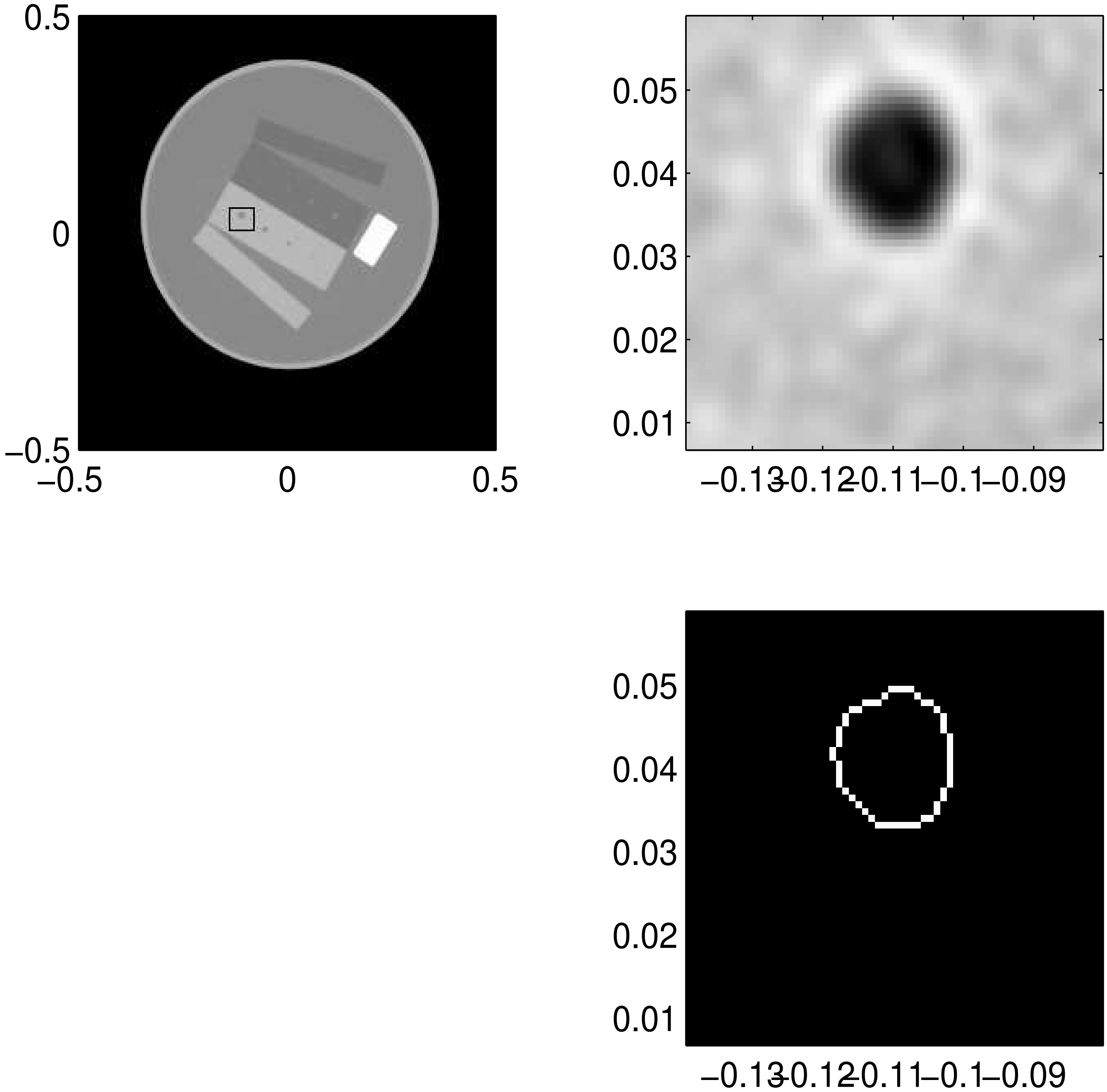

Figure 2: Top left: Global reconstruction of density $f(x)$ of calibration object. Box indicates region of interest R. Top right: Local reconstruction $\bar{\Lambda} f$ inside R. Bottom right: Result of automatic edge detector applied to the image of $\bar{\Lambda} f$ shown in top right. Pixels where an edge is detected are white.

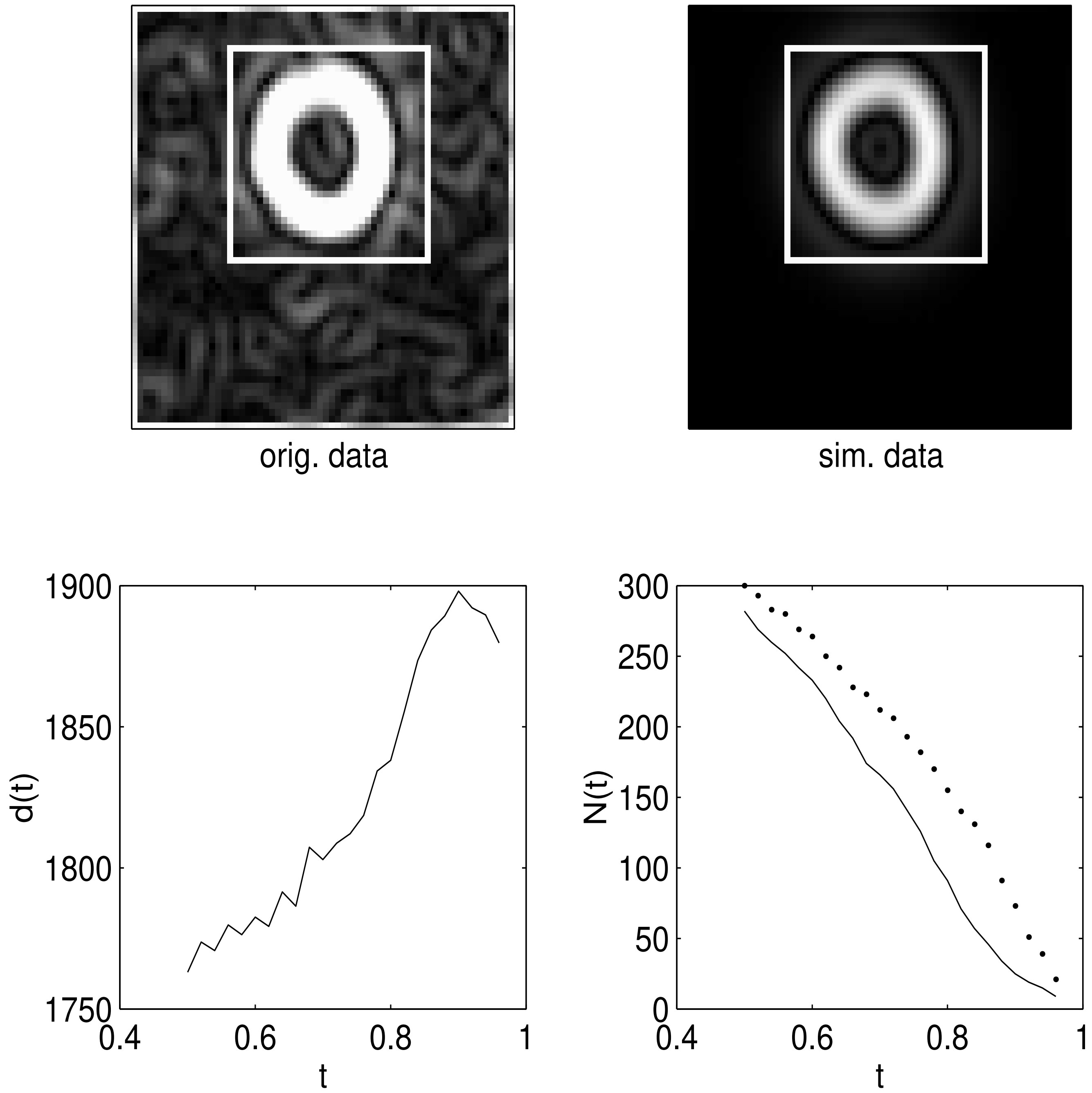

Figure 3: Top left: Image of $|\nabla\bar{\Lambda}f|$ inside R, f being the density function of the calibration object. Box indicates region R'. Top right: Image of $|\nabla\bar{\Lambda}\chi_{X_j}|$ inside R. Box indicates region R'. Bottom left: Graph of estimated density difference $d(t)$ for $0.5 \leq d(t) \leq 0.95$. Bottom right: Number $N(t)$ of points contributing to the averages of $|\nabla\bar{\Lambda}f|$ (solid line), and of $|\nabla\bar{\Lambda}\chi_{X_j}|$ (dotted line).

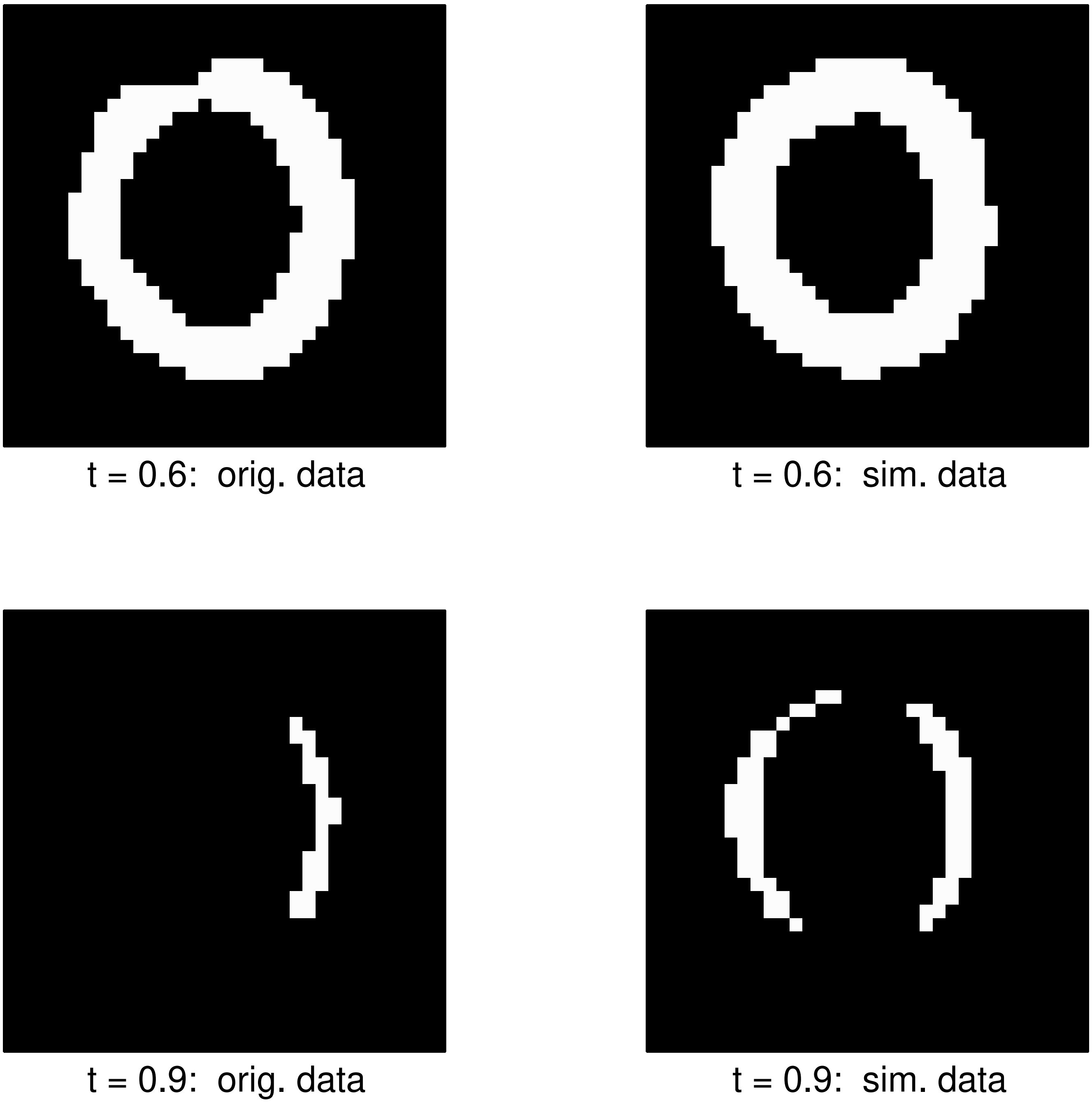

Figure 4: Left half: White points indicate where $|\nabla\bar{\Lambda}f(x)| > t\max_{y\in R'}|\nabla\bar{\Lambda}f(y)|$ for $t = 0.6$ (top left) and $t = 0.9$ (bottom left). Right half: White points indicate where $|\nabla\bar{\Lambda}\chi_{X_j}(x)| > t\max_{y\in R'}|\nabla\bar{\Lambda}\chi_{X_j}(y)|$ for $t = 0.6$ (top right) and $t = 0.9$ (bottom right).