

Inversion of the star transform

Fan Zhao¹, John C Schotland² and Vadim A Markel³

¹ Department of Mathematics, University of Pennsylvania, Philadelphia, PA 19104, USA

² Department of Mathematics and Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA

³ Departments of Radiology and Bioengineering and Graduate Group in Applied Mathematics and Computational Science, University of Pennsylvania, Philadelphia, PA 19104, USA

E-mail: vmarkel@mail.med.upenn.edu

Received 29 January 2014, revised 16 April 2014

Accepted for publication 6 May 2014

Published 17 September 2014

Abstract

We define the star transform as a generalization of the broken ray transform introduced by us in previous work. The advantages of using the star transform include the possibility to reconstruct the absorption and the scattering coefficients of the medium separately and simultaneously (from the same data) and the possibility to utilize scattered radiation which, in the case of conventional x-ray tomography, is discarded. In this paper, we derive the star transform from physical principles, discuss its mathematical properties and analyze numerical stability of inversion. In particular, it is shown that stable inversion of the star transform can be obtained only for configurations involving odd number of rays. Several computationally-efficient inversion algorithms are derived and tested numerically.

Keywords: single-scattering tomography, broken ray transform, x-ray

(Some figures may appear in colour only in the online journal)

1. Introduction

Image reconstruction techniques based on inversion of the Radon transform are well-established. These methods generally neglect the phenomenon of scattering and are based on the laws of geometrical optics and Beer's law, which describes attenuation of rays upon straight-line propagation. However, x-rays do experience significant scattering in tissues and, at high energies, attenuation of x-rays is predominantly explained by Compton scattering. Just like ballistic photons, scattered photons carry information about the medium through which they propagate. Utilization of such photons can be advantageous. In the case of x-rays, account of scattering is simplified because, for the physical parameters encountered in typical

applications, the single-scattering approximation can be used safely. All this has stimulated interest in using single-scattered photons for tomographic imaging [1–6]. We also note that, under suitable conditions, single-scattering regime can be applicable to optical imaging as well, e.g., in the mesoscopic scattering regime [7].

When scattering is taken into consideration, the photon trajectories are no longer straight lines. Different curved or piecewise linear trajectories have been explored, the type of trajectory depending on energy- and angle-selectivity of sources and detectors and on the choice of contrast mechanism. Thus, a series of papers have explored a circular-arc transform, which arises when the signal is generated by first-order Compton scattering [1, 3, 8, 9]. In the imaging modality proposed in these references, the contrast mechanism is related to the spatially-varying efficiency of Compton scattering while attenuation of the scattered rays by the medium is neglected, although an approximate correction to account for the attenuation can be introduced [1]. Mathematically, it was shown that the generalized Radon transform on co-planar circles whose centers are restricted to a bounded domain is invertible [10]. Radon transforms on other smooth curves have also been considered [11–13]. In other applications, the trajectories associated with detection of single-scattered photons are piecewise linear [2, 4, 5, 14]. In other cases, the arising transform involves area integrals rather than integrals over well-defined trajectories [6].

In this paper, we focus on an imaging modality introduced by us in [15–17], wherein angularly-resolved sources and detectors are employed but no energy resolution or sensitivity is assumed. The corresponding integral transform of the medium is referred to as the broken ray Radon transform or, in some cases, as the V-line transform [18, 19], since the photon trajectories of interest resemble the letter V. A single broken ray consists of a vertex and two rays originating from the vertex, which are translated without rotations when the vertex is scanned. The imaging modality based on inverting the broken ray transform does not require multiple projections in the traditional sense. It turns out that it is sufficient to scan several angularly-resolved sources and detectors on one or both sides of a long strip. Another useful feature of this modality is the ability to reconstruct the attenuation and the scattering coefficients of the medium separately. However, in the simplest geometry involving a single broken ray (V-line) whose vertex is scanned over a two-dimensional area [15–17], the inverse problem is mildly ill-posed and this results in various image artifacts.

Imaging methods that utilize broken rays have attracted considerable recent attention [19–22]. An important result was obtained by Katsevich and Krylov [20] who have demonstrated that a linear combination of several broken ray measurements can be used to derive a purely local reconstruction algorithm that involves only first-order derivatives of the data. In this paper, we also explore an approach that utilizes linear combinations of broken ray measurements but take a different approach to image reconstruction. Namely, we describe a reconstruction algorithm in spatial Fourier domain where the inverse solution depends on the data nonlocally. The motivation behind this approach is three-fold. First, it allows one to use classical Tikhonov regularization. Second, within the spatial Fourier method, measurements of ballistic (nonscattered) rays can be easily combined with measurements of broken rays in the same image reconstruction algorithm. Third, the method developed by us is more flexible with respect to choosing the ray geometry.

In what follows, we show that the broken ray transform is a particular case of the star transform (introduced in this paper), which involves line integrals taken with different weights over several rays originating from the same vertex. As in the case of a single broken ray, the rays comprising the star are translated without rotations when the vertex is scanned. The local method of Katsevich and Krylov [20] and other similar methods can be obtained by taking linear combinations of the ray integrals with vector coefficients (this involves a definition of a vector data function). We will discuss some mathematical properties of the star transform,

derive several computationally efficient methods for its inversion in spatial Fourier domain, analyze stability of the inversion algorithms and illustrate the results with numerical examples. In particular, we show that stable inversion of a scalar star transform can be obtained only if the number of rays is odd. More detailed stability conditions are also obtained.

The paper is organized as follows. The star transform is introduced in section 2. In section 3, we explain how the star transform can be related to the physical measurements taken by angularly-resolved source-detector pairs. We also explain how the star transform can be constructed to facilitate separate reconstruction of the attenuation and the scattering coefficients. In section 4, we make a connection between the star transform and the local reconstruction method of Katsevich and Krylov [20], and describe a general framework for obtaining similar methods. In section 5, we obtain the Fourier-space representation of the star transform. In section 6, we explain how the star transform can be combined with traditional projection measurements without significant modifications of the reconstruction algorithm. In section 7, we derive several fast computational algorithms for inverting the star transform in Fourier domain. The results of this section are, in fact, more general because they apply to any matrix that is given by a sum of a term whose inverse or pseudo-inverse is known and a finite number of separable terms. In section 8 we analyze the numerical stability of inverting the star transform. We have obtained a few simple necessary conditions for stability. One such condition is that the star transform must contain an odd number of ray integrals. The results are illustrated with numerical examples of section 9. Finally, section 10 contains a discussion and a summary.

2. Imaging geometry and definition of the star transform

The imaging modality described in this paper reconstructs the properties of a three-dimensional medium slice-by-slice, similarly to the conventional methods of x-ray tomography (obviously, this analogy does not apply to the helical Radon transform and similar generalizations). In what follows, we assume that a slice $x = \text{const}$ of a three-dimensional medium has been selected and will focus on reconstructing the attenuation coefficient $\mu(y, z) = \mu_s(y, z) + \mu_a(y, z)$ in that slice. Here μ_s and μ_a are the scattering and absorption coefficients. Several angularly-resolved sources and detectors are scanned along the lines $z = 0$ and $z = L$. We seek to reconstruct the function $\mu(y, z)$ between these two lines in the open strip $\mathbb{S} = \{0 < z < L\}$. We will also show how the scattering coefficient $\mu_s(y, z)$ can be separately recovered.

The transform whose inversion we study in this paper is of the following form:

$$\Phi(\mathbf{R}) = \sum_{k=1}^K s_k I_k(\mathbf{R}), \quad \mathbf{R} \equiv (Y, Z) \in \bar{\mathbb{S}} = \{0 \leq z \leq L\}; \quad (1a)$$

$$I_k(\mathbf{R}) = \int_0^{\ell_k(Z)} \mu(\mathbf{R} + \hat{\mathbf{u}}_k \ell) d\ell. \quad (1b)$$

Here $\Phi(\mathbf{R})$ is the data function for a K -ray imaging geometry, which is illustrated in figure 1(a) for the particular case $K = 3$. It is defined in the closure of \mathbb{S} (denoted here by $\bar{\mathbb{S}}$). Also, $\hat{\mathbf{u}}_k = (u_{ky}, u_{kz})$ is a set of K unit vectors with nonzero projections onto the Z -axis (that is, $u_{ky}^2 + u_{kz}^2 = 1$ and $u_{kz} \neq 0$); $\ell_k(Z)$ is the distance (defined for each ray) from the vertex to the boundary of \mathbb{S} ; finally, $s_k \neq 0$ is a set of known coefficients. The data function $\Phi(\mathbf{R})$ is assumed to be known (measured).

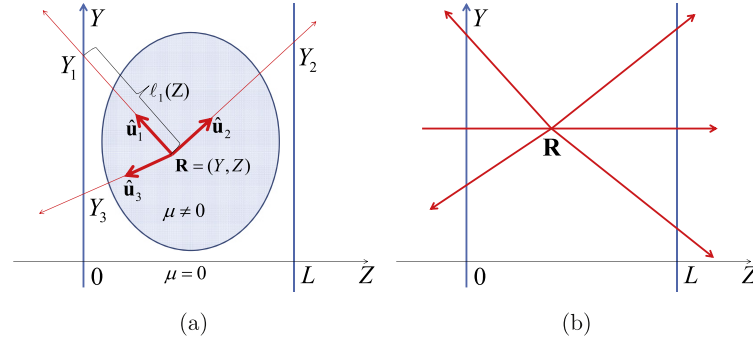


Figure 1. (a) Sketch of the imaging geometry for the case $N = 3$ (the distances ℓ_2 and ℓ_3 are not shown). (b) Imaging geometry in which simultaneous measurements of the ballistic and single-scattered rays (due to the same source) are utilized.

Since formation of broken rays depends on scattering, it is essential that $\mu_s(\mathbf{r}) > \text{const} > 0$ for $\mathbf{r} \in \mathbb{S}$ and this implies that $\mu(\mathbf{r}) > \text{const} > 0$ in \mathbb{S} . This is problematic for our purposes because $\mu(\mathbf{r})$ does not have in this case a Fourier transform. Of course, the scattering medium is never infinite in practice and one can introduce physical boundaries to alleviate this problem. Then the data function is no longer defined in an infinite strip. However, there exists a simpler approach, which we will follow in this paper. Namely, it is physically reasonable to expect that the medium properties are constant and known when $|y| \rightarrow \infty$. More specifically, let

$$\mu(y, z) = \bar{\mu} = \bar{\mu}_s + \bar{\mu}_a \quad \text{for } |y| > y_{\max} > 0, \quad (2)$$

where bar is used to denote known background values of all coefficients. This is true, in particular, if the inhomogeneities $\mu(\mathbf{r})$ are represented by a set of finite objects, as is often the case in practice. Then we can write $\mu(\mathbf{r}) = \bar{\mu} + \delta\mu(\mathbf{r})$, where $\delta\mu(\mathbf{r})$ is compactly supported in a subset of \mathbb{S} . This situation is schematically illustrated in figure 1(a). We emphasize that the support of $\delta\mu$ or the value of y_{\max} are not known *a priori*. However, we can use (2) to rewrite (1) as follows:

$$\delta\Phi(\mathbf{R}) \equiv \Phi(\mathbf{R}) - \bar{\mu} \sum_{k=1}^K s_k \ell_k(Z) = \sum_{k=1}^K s_k \delta I_k(\mathbf{R}), \quad \mathbf{R} \equiv (Y, Z) \in \mathbb{S}; \quad (3a)$$

$$\delta I_k(\mathbf{R}) = \int_0^{\ell_k(Z)} \delta\mu(\mathbf{R} + \hat{\mathbf{u}}_k \ell) d\ell. \quad (3b)$$

We notice that (3) is exactly of the same form as (1) except that some quantities have been redefined. Since the functions $\ell_k(Z)$ are known, the left-hand side of (3a) is also known and the expression for $\delta\Phi$ in terms of Φ and $\ell_k(Z)$ serves as a definition of the new data function. In what follows, we will write for simplicity μ instead of $\delta\mu$, Φ instead of $\delta\Phi$, I_k instead of δI_k and consider inversion of the transform (1). However, we should keep in mind that, in this formulation, μ is really the fluctuating part of the attenuation coefficient and, therefore, it can be negative. With these definitions accepted, it is obvious that, as long as $|u_{kz}| > \text{const} > 0$, we have $\Phi(Y, Z) = 0$ for $|Y| > Y_{\max} > 0$. Note that $Y_{\max} \neq y_{\max}$, where the latter quantity appears in (2), and the relation between Y_{\max} and y_{\max} depends on the ray geometry.

Image reconstruction methods described in this paper allow one to combine measurements of the type illustrated in figure 1(a), which employ single scattering at the vertex \mathbf{R} , with ballistic transmission measurements. This is schematically illustrated in figure 1(b). A family of ballistic rays yields a single plane-parallel projection in a CT scan. Since ballistic rays carry more photons than scattered rays, measurements of the former are less affected by noise. Measurements of ballistic rays that are perpendicular to the strip, as shown in figure 1(b), provide sufficient information to determine the integral $\int_0^L \mu(y, z) dz$, which is related by Fourier transform to the function $\mu_0(q)$ defined below in section 5. Ballistic rays that enter the strip at the angle θ to the normal provide information on the Fourier coefficients of a more general form, namely, $\mu_n(\kappa_n \text{ctg}\theta)$. Independent measurements of these coefficients can be used in image reconstruction as is described in section 6.

3. Physical principles and derivation of the star transform

We now explain what type of physical measurements are required to obtain a transform of the type (1) and also discuss how the scattering coefficient $\mu_s(\mathbf{r})$ can be separately recovered.

The mathematical concept of broken rays appears naturally when transport of waves or particles through a medium is considered within the first-order scattering approximation. Intensity carried by a broken ray can be measured by source-detector pairs that are not on axis. As is discussed in detail in [15–17], a broken ray is defined when an angularly-resolved source and an angularly-resolved detector are aligned so that the two rays drawn from the source and the detector intersect in the slice of interest. The broken ray consists of a vertex $\mathbf{R} = (Y, Z)$ and two rays originating from the vertex and pointing to the source and the detector (see figure 1(a) for an illustration). Each source and detector is characterized by the directionality vector $\hat{\mathbf{u}}_k$ and each can be scanned along the Y axis. Consider a source-detector pair characterized by the directionality vectors $\hat{\mathbf{u}}_j$ and $\hat{\mathbf{u}}_k$. The projections of the source and detector positions onto the Y and Z axes, Y_j, Z_j and Y_k, Z_k ($Z_j, Z_k = 0, L$), determine the broken ray geometry. Not all possible combinations of $Y_j, Z_j, \hat{\mathbf{u}}_j; Y_k, Z_k, \hat{\mathbf{u}}_k$ have a vertex in \mathcal{S} . Source-detector arrangements without a vertex or with a vertex outside of \mathcal{S} do not generate useful data. If $\mathbf{R} \in \mathcal{S}$, then the power measured by the j th detector due to the k th source, W_{jk} , is given by

$$W_{jk}(\mathbf{R}) = W_0 S_{jk} \mu_s(\mathbf{R}) \exp \left\{ - \left[I_j(\mathbf{R}) + I_k(\mathbf{R}) \right] \right\}, \quad (4)$$

where W_0 is the incident power generated by the source, which is assumed to be the same for all sources, $I_j(\mathbf{R})$ and $I_k(\mathbf{R})$ are the integrals defined in (1b) and S_{jk} is a geometrical factor, which depends on $\cos \theta_{jk} = \hat{\mathbf{u}}_j \cdot \hat{\mathbf{u}}_k$ but not on \mathbf{R} . The factors S_{jk} can be model-dependent, but we assume here that they are known. For example, if the power measured by the detector is described by the single-energy radiative transport equation, an explicit expression for S_{jk} is given in [15].

We now define data functions $\phi_{jk}(\mathbf{R})$ according to

$$\phi_{jk}(\mathbf{R}) = \ln \left[\frac{W_{jk}(\mathbf{R})}{W_0 S_{jk} \bar{\mu}_s} \right], \quad (5)$$

where $\bar{\mu}_s$ is the background scattering coefficient defined in (2). It is important to note that, when a function $\phi_{jk}(\mathbf{R})$ is measured, one of the end-points of the corresponding broken ray is a source and the other is a detector. Physically, these are different devices. However, all data functions ϕ_{jk} do not need to be acquired simultaneously. For example, in the case $K = 3$, the data functions ϕ_{12} and ϕ_{13} can be acquired simultaneously using one source at Y_1, Z_1 and two detectors at Y_2, Z_2 and Y_3, Z_3 . Acquisition of the function ϕ_{23} requires the use of a source at Y_2, Z_2 and a detector at Y_3, Z_3 or *vice versa*, and can be performed separately.

Substituting (4) into (5), we obtain the equation

$$\phi_{jk}(\mathbf{R}) = [I_j(\mathbf{R}) + I_k(\mathbf{R}) + \eta(\mathbf{R})](1 - \delta_{jk}), \quad (6)$$

where $\eta(\mathbf{R}) = \ln[\mu_s(\mathbf{R})/\bar{\mu}_s]$. The diagonal terms ϕ_{kk} are not measurable; for generality, we have defined these terms to be zero by including the factor $(1 - \delta_{jk})$ in (6). Since $\phi_{jk} = \phi_{kj}$, it is sufficient to consider only the pairs of indices with $j < k$. If K distinct directionality vectors are used, then the number of independent functions $\phi_{jk}(\mathbf{R})$ is $K(K - 1)/2$.

Consider first the case when the scattering coefficient is constant and equal to $\bar{\mu}$, so that $\eta(\mathbf{R}) = 0$. It is then easy to see that

$$\Phi \equiv \frac{1}{2(K - 1)} \sum_{k=1}^K \sum_{j=1}^K \phi_{jk} = \sum_{k=1}^K I_k. \quad (7)$$

This equation is of the form (1a) with $s_k = 1$. Therefore, if transform (1a) defines the attenuation coefficient uniquely, then the set $\phi_{jk}(\mathbf{R})$ necessarily contains redundant data points. Another way to obtain the star transform (1a) with $s_k = 1$ is to use cyclic summation. Let $j(k) = k + 1$ if $k < K$ and $j(k) = 1$ if $k = K$. Then

$$\Phi \equiv \frac{1}{2} \sum_{k=1}^K \phi_{kj(k)} = \sum_{k=1}^K I_k, \quad (8)$$

which is of the same form as (7). We note that there are many other linear combinations that result in the star transform with $s_k \neq 1$.

We now turn to the case when $\mu_s(\mathbf{R})$ can vary and, correspondingly, $\eta(\mathbf{R}) \neq 0$. Unlike in the case of uniform scattering coefficient, the data functions $\phi_{jk}(\mathbf{R})$ depends now on two unknown functions, $\mu(\mathbf{R})$ and $\eta(\mathbf{R})$. It is, in principle, possible to reconstruct both given a sufficient number of degrees of freedom in the data. However, we wish to simplify the inverse problem and exclude $\eta(\mathbf{R})$ from the equations analytically. In this case, introduction of the star transform is a necessity rather than a choice. The approach of [16, 17, 20] was to make linear combinations of the equations in (6) in such a way as to exclude $\eta(\mathbf{R})$; then, once the total attenuation function is reconstructed, $\eta(\mathbf{R})$ can be recovered from any of the equations in (6). The linear combinations were defined so that one of the ray integrals was excluded from the resulting equations. For example, in a three-ray geometry, we have defined [16, 17] the data function as $\Phi \equiv \phi_{12} - \phi_{13} = I_2 - I_3$. Thus, the scattering contrast function $\eta(\mathbf{R})$ was excluded but so was the ray integral I_1 . We note that the resultant two-ray geometry is *a priori* ill-posed, as is shown below in section 8.2.

Here we use a similar but somewhat more general approach. We note that the problem of excluding η while not excluding any of the ray integrals from (6) can be solved if we find a set of coefficients c_{jk} such that (i) $\sum_{jk} c_{jk} = 0$, (ii) $c_{kk} = 0$, (iii) $c_{jk} = c_{kj}$, and (iv) $s_k = \sum_j c_{jk} \neq 0$.

Then

$$\Phi \equiv \frac{1}{2} \sum_{j=1}^K \sum_{k=1}^K c_{jk} \phi_{jk} = \sum_{k=1}^K s_k I_k, \quad (9)$$

where ϕ_{jk} satisfy (6). This is again a transform of the form (1a) in which, by construction, $s_k \neq 0$. Generally, there exist many different sets of c_{jk} that satisfy the above conditions. Some obvious coefficient examples for the cases $K = 3$ and $K = 4$ are shown below:

$$\begin{array}{c|c} 0 & 1 & 1 & -1 & 1 \\ 1 & 0 & -2 & -1 & -1 \\ \hline 1 & -2 & 0 & -1 & 1 \\ 2 & -1 & -1 & 0 & -1 \end{array} \quad \begin{array}{c|c} 0 & 1 & 1 & -1 & 1 \\ 1 & 0 & -1 & -1 & -1 \\ 1 & -1 & 0 & 1 & 1 \\ \hline -1 & -1 & 1 & 0 & -1 \\ 1 & -1 & 1 & -1 & 0 \end{array}$$

In the second example, all coefficients c_{jk} and s_k are equal to ± 1 . Note that such arrangements are only possible if $K(K-1)/2$ is even ($K = 4, 5, 8, 9, \dots$).

4. Star transform and the method of Katsevich and Krylov

If direct measurement of a single term $I_k(\mathbf{R})$ was possible (that is, if the data function with $K = 1$ could be measured), then a simple purely local image reconstruction algorithm involving one first-order derivative could be obtained. Indeed, we obviously have

$$-(\hat{\mathbf{u}}_k \cdot \nabla_{\mathbf{R}}) I_k(\mathbf{R}) = \mu(\mathbf{R}), \quad \mathbf{R} \in \mathbb{S}. \quad (10)$$

Unfortunately, direct measurement of $I_k(\mathbf{R})$ is physically impossible. In the case of a spatially-uniform scattering coefficient $\mu_s(\mathbf{R}) = \text{const}$, we can define a star transform by taking linear combinations of individual measurements in such a way as to obtain an equation containing a single term of the form I_k . For example, in the case $K = 3$, we can take

$$\Phi \equiv \frac{1}{2} (\phi_{13} + \phi_{23} - \phi_{12}) = I_3. \quad (11)$$

We can invert (11) using the local formula (10). Here the coefficients of the star transform are $c_{13} = c_{23} = -c_{12} = 1$.

In the more general and more practically-important case of spatially-varying $\mu_s(\mathbf{r})$, a single term $I_k(\mathbf{R})$ can not be mathematically 'isolated' with the use of scalar coefficients c_{jk} . However, we can define *vector* coefficients \mathbf{c}_{jk} to obtain a local reconstruction algorithm involving only first-order derivatives. In this case the data function $\Phi(\mathbf{R})$ is also a vector and the reconstruction formula is of the form

$$\mu(\mathbf{R}) = -\frac{1}{\zeta} \nabla \cdot \Phi(\mathbf{R}), \quad (12)$$

where ζ is a coefficient. We will now outline a general approach to obtaining such reconstruction formulas and provide a few examples.

Let a set of two-dimensional vector coefficients \mathbf{c}_{jk} satisfy the following conditions: (i) $\sum_{jk} \mathbf{c}_{jk} = 0$, (ii) $\mathbf{c}_{kk} = 0$, (iii) $\mathbf{c}_{jk} = \mathbf{c}_{kj}$, and (iv) $\mathbf{s}_k = \sum_j \mathbf{c}_{jk} = \sigma_k \hat{\mathbf{u}}_k$, where σ_k is an arbitrary scalar. Here the conditions (i)-(iii) are quite analogous to the similarly numbered conditions of the previous section (for scalar coefficients) while condition (iv) is new: it requires that \mathbf{s}_k be

collinear to the unit vector $\hat{\mathbf{u}}_k$. The conditions (i) and (iv) are consistent if $\sum_k \sigma_k \hat{\mathbf{u}}_k = 0$. We can find the coefficients σ_k that satisfy this condition if $K \geq 3$ (assuming all vectors $\hat{\mathbf{u}}_k$ are different). Then define the vector data function according to

$$\Phi \equiv \frac{1}{2} \sum_{j=1}^K \sum_{k=1}^K c_{jk} \phi_{jk}, \tag{13}$$

It can be seen that this data function satisfies the equation

$$\Phi = \sum_{k=1}^K \sigma_k \hat{\mathbf{u}}_k I_k, \tag{14}$$

which can be inverted according to (12) with $\zeta = \sum_k \sigma_k$, assuming that the latter quantity is not zero.

As an example, consider the case $K = 3$. The matrix of coefficients that satisfy all the conditions stated above is

$$\begin{array}{ccc|c} 0 & \sigma_1 \hat{\mathbf{u}}_1 + \sigma_2 \hat{\mathbf{u}}_2 & \sigma_1 \hat{\mathbf{u}}_1 + \sigma_3 \hat{\mathbf{u}}_3 & \sigma_1 \hat{\mathbf{u}}_1 \\ \sigma_1 \hat{\mathbf{u}}_1 + \sigma_2 \hat{\mathbf{u}}_2 & 0 & \sigma_2 \hat{\mathbf{u}}_2 + \sigma_3 \hat{\mathbf{u}}_3 & \sigma_2 \hat{\mathbf{u}}_2 \\ \sigma_1 \hat{\mathbf{u}}_1 + \sigma_3 \hat{\mathbf{u}}_3 & \sigma_2 \hat{\mathbf{u}}_2 + \sigma_3 \hat{\mathbf{u}}_3 & 0 & \sigma_3 \hat{\mathbf{u}}_3 \\ \hline \sigma_1 \hat{\mathbf{u}}_1 & \sigma_2 \hat{\mathbf{u}}_2 & \sigma_3 \hat{\mathbf{u}}_3 & 0 \end{array}$$

where the coefficients σ_k are assumed to satisfy $\sigma_1 \hat{\mathbf{u}}_1 + \sigma_2 \hat{\mathbf{u}}_2 + \sigma_3 \hat{\mathbf{u}}_3 = 0$. Application of this scheme results in the equation

$$\begin{aligned} \mu &= -\frac{1}{\zeta} \nabla \cdot \left[\sigma_1 \hat{\mathbf{u}}_1 (\phi_{12} + \phi_{13}) + \sigma_2 \hat{\mathbf{u}}_2 (\phi_{12} + \phi_{23}) + \sigma_3 \hat{\mathbf{u}}_3 (\phi_{13} + \phi_{23}) \right] \\ &= \frac{1}{\zeta} \nabla \cdot \left[\sigma_1 \hat{\mathbf{u}}_1 (\phi_{23} - \phi_{12}) + \sigma_2 \hat{\mathbf{u}}_2 (\phi_{12} - \phi_{13}) \right]. \end{aligned} \tag{15}$$

In particular, if $\hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_2 + \hat{\mathbf{u}}_3 = 0$, we can use $\sigma_1 = \sigma_2 = \sigma_3 = 1$ and the above equation takes the simple form

$$\mu = \frac{1}{3} \nabla \cdot \left[\hat{\mathbf{u}}_1 (\phi_{23} - \phi_{12}) + \hat{\mathbf{u}}_2 (\phi_{12} - \phi_{13}) \right].$$

The method of Katsevich and Krylov [20] can also be derived using the mathematical formalism of this section. The particular implementation of [20] utilizes four rays, $K = 4$, but one of the scalar coefficients, say σ_1 , is zero, so that the corresponding ray does not enter the star transform. The coefficient matrix utilized by Katsevich and Krylov is

$$\begin{array}{cccc|c} 0 & \sigma_2 \hat{\mathbf{u}}_2 & \sigma_3 \hat{\mathbf{u}}_3 & \sigma_4 \hat{\mathbf{u}}_4 & 0 \\ \sigma_2 \hat{\mathbf{u}}_2 & 0 & 0 & 0 & \sigma_2 \hat{\mathbf{u}}_2 \\ \sigma_3 \hat{\mathbf{u}}_3 & 0 & 0 & 0 & \sigma_3 \hat{\mathbf{u}}_3 \\ \sigma_4 \hat{\mathbf{u}}_4 & 0 & 0 & 0 & \sigma_4 \hat{\mathbf{u}}_4 \\ \hline 0 & \sigma_2 \hat{\mathbf{u}}_2 & \sigma_3 \hat{\mathbf{u}}_3 & \sigma_4 \hat{\mathbf{u}}_4 & 0 \end{array} \tag{16}$$

where $\sigma_1 = 0$ and the remainder of the coefficients satisfy $\sigma_2 \hat{\mathbf{u}}_2 + \sigma_3 \hat{\mathbf{u}}_3 + \sigma_4 \hat{\mathbf{u}}_4 = 0$. The corresponding inversion formula is

$$\begin{aligned}\mu &= -\frac{1}{\zeta} \nabla \cdot (\sigma_2 \hat{\mathbf{u}}_2 \phi_{12} + \sigma_3 \hat{\mathbf{u}}_3 \phi_{13} + \sigma_4 \hat{\mathbf{u}}_4 \phi_{14}) \\ &= \frac{1}{\zeta} \nabla \cdot [\sigma_2 \hat{\mathbf{u}}_2 (\phi_{14} - \phi_{12}) + \sigma_3 \hat{\mathbf{u}}_3 (\phi_{14} - \phi_{13})].\end{aligned}\quad (17)$$

The difference between formulas (15) and (17) is that these formulas utilize different ‘individual measurements’ ϕ_{jk} and, therefore, a different physical arrangement of sources and detectors.

5. Fourier basis representation

We define the Fourier transform of $\mu(y, z)$ as follows:

$$\mu(y, z) = \int_{-\infty}^{\infty} \frac{dq}{2\pi} e^{iqy} \bar{\mu}(q, z) = \int_{-\infty}^{\infty} \frac{dq}{2\pi} e^{iqy} \frac{1}{L} \sum_{n=-\infty}^{\infty} \mu_n(q) e^{i\kappa_n z}, \quad (18a)$$

$$\mu_n(q) = \int_{-\infty}^{\infty} dy e^{-iqy} \int_0^L dz e^{-i\kappa_n z} \mu(y, z), \quad (18b)$$

$$\kappa_n = \frac{2\pi n}{L}. \quad (18c)$$

The same convention can be used for the data function $\Phi(Y, Z)$. We assume that all functions are sufficiently ‘nice’ so that the Fourier transforms exist. It should be noted that, if (1) is defined on an infinite plane, then the Fourier transform of $\Phi(Y, Z)$ does not exist even if we include generalized functions into consideration. When (1) is defined in \mathcal{S} and all rays cross the strip boundaries, this problem is removed. However, we should keep in mind that $\Phi(Y, 0) \neq \Phi(Y, L)$. We define

$$\Delta(Y) \equiv \frac{1}{2} [\Phi(Y, 0) + \Phi(Y, L)], \quad (19)$$

Then

$$\sum_{n=-\infty}^{\infty} \bar{\Phi}_n(q) = L \bar{\Delta}(q) = L \int_{-\infty}^{\infty} \Delta(Y) e^{-iqY} dY. \quad (20)$$

We start by Fourier-transforming (1) along the Y -direction:

$$\begin{aligned}\bar{\Phi}(q, Z) &= \sum_{k=1}^K s_k \int_0^{\ell_k(Z)} d\ell \int_{-\infty}^{\infty} \mu(Y + u_{ky} \ell, Z + u_{kz} \ell) e^{-iqY} dY \\ &= \sum_{k=1}^K s_k \int_0^{\ell_k(Z)} e^{iq u_{ky} \ell} \bar{\mu}(q, Z + u_{kz} \ell) d\ell \\ &= \sum_{k=1}^K \frac{s_k}{u_{kz}} e^{-i\beta_k(q)Z} \int_Z^{\bar{z}_k} e^{i\beta_k(q)z} \bar{\mu}(q, z) dz \\ &= \sum_{k=1}^K \frac{s_k}{u_{kz}} e^{-i\beta_k(q)Z} \sum_{n=-\infty}^{\infty} \frac{e^{i[\beta_k(q)+\kappa_n] \bar{z}_k} - e^{i[\beta_k(q)+\kappa_n]Z}}{i[\beta_k(q) + \kappa_n]} \mu_n(q).\end{aligned}\quad (21)$$

In the above derivation, we have taken advantage of the fact that the upper limit of integration over ℓ , $\ell_k(Z)$, is independent of Y and introduced the notations

$$\beta_k(q) = q \frac{u_{ky}}{u_{kz}}, \quad \xi_k = \begin{cases} L, & \text{if } u_{kz} > 0 \\ 0, & \text{if } u_{kz} < 0. \end{cases} \quad (22)$$

Here ξ_k is the Z -coordinate of the k th ray intersection with the boundary of \mathcal{S} . Since we assume that all rays intersect the strip boundaries, the quantities in (22) are well defined. It is useful to keep in mind that $\exp(i\xi_k \kappa_n) = 1$ for all values of indices.

Equation (21) is parameterized by q . To shorten the notations, we will omit the explicit dependence on q below by writing β_k, μ_n instead of $\beta_k(q), \mu_n(q)$, etc., except in a few special cases. We then fix q and take the Fourier transform of (21) with respect to Z , which results in the following infinite system of linear equations:

$$\Phi_n = \mu_n \sum_{k=1}^K \frac{is_k}{u_{kz}(\beta_k + \kappa_n)} + \sum_{k=1}^K \frac{s_k e^{i\beta_k \xi_k} (e^{-i\beta_k L} - 1)}{Lu_{kz}(\beta_k + \kappa_n)} \sum_{m=-\infty}^{\infty} \frac{\mu_m}{\beta_k + \kappa_m}. \quad (23)$$

Introducing the notations

$$d_n = \sum_{k=1}^K \frac{is_k}{u_{kz}(\beta_k + \kappa_n)} = \sum_{k=1}^K \frac{is_k}{\hat{\mathbf{u}}_k \cdot (q, \kappa_n)}, \quad \alpha_k = \frac{e^{i\beta_k \xi_k} (e^{-i\beta_k L} - 1)}{Lu_{kz}}, \quad (24)$$

we can rewrite (23) in the form

$$\Phi_n = d_n \mu_n + \sum_{k=1}^K \frac{s_k \alpha_k}{\beta_k + \kappa_n} \sum_{m=-\infty}^{\infty} \frac{\mu_m}{\beta_k + \kappa_m}. \quad (25)$$

In (24), (q, κ_n) is the two-dimensional Fourier vector. It can be seen that (25) is an infinite set of algebraic equations whose matrix is a sum of one diagonal matrix and K separable matrices. It will also prove useful to introduce Dirac notations. Let

$$|\mu\rangle = (\dots, \mu_{-1}, \mu_0, \mu_1, \dots)^T, \quad (26a)$$

$$|\Phi\rangle = (\dots, \Phi_{-1}, \Phi_0, \Phi_1, \dots)^T, \quad (26b)$$

$$|a_k\rangle = \left(\dots, \frac{1}{\beta_k + \kappa_{-1}}, \frac{1}{\beta_k + \kappa_0}, \frac{1}{\beta_k + \kappa_1}, \dots \right)^T. \quad (26c)$$

Then (25) can be written as

$$|\Phi\rangle = D |\mu\rangle + \sum_{k=1}^K s_k \alpha_k |a_k\rangle \langle a_k | \mu \rangle \equiv A |\mu\rangle. \quad (27)$$

Here the diagonal matrix D has the elements $D_{nm} = d_n \delta_{nm}$ and the second equality defines the matrix A .

6. Star transform combined with projection measurements

The function $\mu_0(q)$ can be measured with relatively high precision by utilizing ballistic (nonscattered) rays as is illustrated in figure 1(b). Measurements of ballistic rays and of the single-scattered rays can be performed simultaneously, without the need to employ additional sources of radiation. Therefore, information on $\mu_0(q)$ can be obtained as long as at least one of the sources is oriented perpendicularly to the strip, as shown in the figure. One motivation for developing the imaging modality described here is to reduce the radiation dose received by a

patient by reducing the number of projections employed. In this respect, discarding ballistic photons is not an efficient approach. Mathematically, the knowledge of $\mu_0(q)$ does not determine the function $\mu(y, z)$ [to this end, the whole set of coefficients $\mu_n(q)$ is required], but it does improve the conditioning of the inverse problem. Therefore, it is reasonable to assume that the projection measurements illustrated in figure 1(b) have been performed and that $\mu_0(q)$ is known. Then we can rewrite (25) as follows:

$$\Psi_0 = \sum_{k=1}^K \frac{s_k \alpha_k}{\beta_k} \sum_{m \neq 0} \frac{\mu_m}{\beta_k + \kappa_m}, \quad n = 0, \quad (28a)$$

$$\Psi_n = d_n \mu_n + \sum_{k=1}^K \frac{s_k \alpha_k}{\beta_k + \kappa_n} \sum_{m \neq 0} \frac{\mu_m}{\beta_k + \kappa_m}, \quad n \neq 0, \quad (28b)$$

where

$$\Psi_0 \equiv \Phi_0 - \mu_0 \left(d_0 + \sum_{k=1}^K \frac{s_k \alpha_k}{\beta_k^2} \right), \quad n = 0, \quad (29a)$$

$$\Psi_n \equiv \Phi_n - \mu_0 \sum_{k=1}^K \frac{s_k \alpha_k}{\beta_k (\beta_k + \kappa_n)}, \quad n \neq 0. \quad (29b)$$

The right-hand side of (28) depends only on μ_n with $n \neq 0$; therefore, this set (when truncated so that $|n| \leq n_{\max}$; see more detail on truncation in section 7.1) has more equations than unknowns. One possible approach is to disregard equation (28a). In an ideal setting, e.g., if the data are generated using inverse crime and contain no noise or systematic errors, all equations in (28) must be consistent and the disregard of (28a) does not affect the solution. However, if experimental measurements are used, then (28a) provides an additional useful constraint and it might be advantageous to seek the pseudo-inverse of the overdetermined system, as is described in section 7.2 below.

We finally note that measurement of ballistic rays that enter the slab at the angle θ can also be utilized in similar manner. Such rays yield independent measurements of the Fourier coefficients $\mu_n(q = \kappa_n \text{ctg}\theta)$.

7. Methods of solution

We now discuss two efficient algorithms for solving equations of the type (27). However, we will consider in this section inversion algorithms for a more general matrix A , which is of the form

$$A = D + V, \quad V = \sum_{k=1}^K |b_k\rangle \langle a_k|, \quad (30)$$

where $|a_k\rangle$ and $|b_k\rangle$ are not necessarily collinear. The special case of the star transform is obtained if we take $|b_k\rangle = s_k \alpha_k |a_k\rangle$ and assume that D is the diagonal matrix defined by (24). The size of A can be either finite or infinite. In particular, A can be a finite-size rectangular or square matrix. In the case of star transform, these details depend on the type of truncation of the infinite set of equations (27). Truncation is discussed in section 7.1. In section 7.2, we describe an algorithm for fast iterative computation of the Tikhonov-regularized pseudo-inverse of a finite-size rectangular A . An interesting feature of this algorithm is that it does not

require the knowledge or computation of the singular vectors and singular values of A . Finally, in section 7.3 we describe a method for direct inversion of A . In this method, A can be either finite and square or infinite. The power of this method is that it allows us to invert an infinite-dimensional matrix, essentially, without any approximations. However, it is not possible to measure all components of the infinite-dimensional vector of data $|\Phi\rangle$ and the components that are not measurable are approximated by zeros.

7.1. Truncation

Here we discuss various ways to truncate the infinite set of equations (27).

To perform image reconstruction on a finite grid, it is sufficient to know the coefficients μ_n for a finite range of n . For example, if the image is discretized in the Z -direction using $2n_{\max} + 1$ points $z_n = hn$, $n = 0, 1, \dots, 2n_{\max}$, $h = L/2n_{\max}$, then we only need to know μ_n for $-n_{\max} \leq n \leq n_{\max}$. If we sample the data function on the same grid, we have access to the quantities Φ_n with n in the same range. Other components of $|\Phi\rangle$ are in this case not measurable. If the matrix A in (27) were block-diagonal, with one block encompassing all indexes n that satisfy the above inequality, then we could have used the available measurements of Φ_n to reconstruct the required μ_n without any approximations. In this case, truncation of (27) would have been trivial.

However, A is not block-diagonal. As a result, the coefficients μ_n with $-n_{\max} \leq n \leq n_{\max}$ depend on *all* measurements Φ_n . Therefore, there exist two different methods to truncate (27). The first method is to set $\mu_n = 0$ for $|n| > n_{\max}$ and to disregard all equations with $|n| > n_{\max}$. In this approach, *all* equations in (27) are changed, not only the ones that have been disregarded. The second approach is to keep all equations in (27) but to substitute the unavailable measurements $\Phi_n(q)$ with zeros. This second approach does not modify the matrix A but it makes an approximation of the data.

Normally, this second approach would not be practically feasible because infinite sets of equations can not be handled numerically. For the case at hand, however, we can utilize the known algebraic structure of A , as is described in section 7.3 below.

7.2. Iterative computation of the pseudo-inverse

Here we consider the case when A is finite and either square and singular or non-square. In both cases, inverse of A may not exist. Non-square A is encountered, for example, if projection measurements are used to compute some of the coefficients $\mu_n(q)$, as is described in section 6. We therefore wish to derive an inversion algorithm that is suitably regularized and contains no numerical instabilities. We note that none of the terms in the right-hand side of (30) can be expected to commute and, therefore, it is not possible to find analytically the singular value decomposition of A even in the simplest case of one separable term. The proposed method computes the pseudo-inverse of A rather than its singular-value decomposition, the latter being a more computationally demanding, yet unnecessary, task.

The main idea of this iterative computation of the pseudo-inverse is the following. Let us define the ‘forward’ recursion as the set of equations

$$D_k = D_{k-1} + |b_k\rangle\langle a_k|, \quad k = 1, 2, \dots, K, \quad (31)$$

where $D_0 = D$. It can be seen that $D_K = A$, where A is defined by (30) and contains K separable terms. Now let us assume that we know the pseudo-inverse of D_{k-1} , denoted here by

D_{k-1}^+ . Then we can compute the pseudo-inverse of the right-hand side of (31) analytically using the formulas given below. This will yield an expression for the pseudo-inverse of D_k in terms of the pseudo-inverse of D_{k-1} . This rule for pseudo-inverses can be referred to as the ‘inverse’ recursion. The inverse recursion can be started easily because the pseudo-inverse of $D_0 = D$ (a diagonal matrix) is known. Then the inverse recursion can be followed for K steps to compute $A^+ = D_K^+$.

Let A and all D_k be finite $N \times M$ matrices. The Tikhonov-regularized pseudo-inverse of D_k can be defined as follows:

$$D_k^+ = D_k^* S_{N,k} = S_{M,k} D_k^*, \quad (32)$$

where

$$S_{N,k} = (D_k D_k^* + \lambda^2 \mathbb{I}_N)^{-1}, \quad S_{M,k} = (D_k^* D_k + \lambda^2 \mathbb{I}_M)^{-1}, \quad (33)$$

Here \mathbb{I}_N and \mathbb{I}_M are the identity matrices of the size $N \times N$ and $M \times M$ respectively, and λ is the regularization parameter. We note that the inverses in (33) exist in the usual sense as long as $\lambda > 0$.

The recursion starts with computing $S_{N,0}$ and $S_{M,0}$ according to (33), where $D_0 = D$. In the case of star transform, $S_{N,0}$ and $S_{M,0}$ are diagonal matrices with the elements

$$s_m = \frac{1}{|d_m|^2 + \lambda^2}. \quad (34)$$

In the case $N > M$ (an overdetermined problem), the matrix $S_{N,0}$ has s_m in the first M diagonal positions and $1/\lambda^2$ in the positions $m = M + 1, \dots, N$, while $S_{M,0}$ is the $M \times M$ minor of $S_{N,0}$.

The generic inverse iteration step requires a formula for updating $S_{N,k+1}$, $S_{M,k+1}$ and D_{k+1}^+ in terms of $S_{N,k}$, $S_{M,k}$, and D_k^+ . This recursive rule is of the form

$$S_{N,k+1} = S_{N,k} - S_{N,k} T_{N,k} S_{N,k}, \quad (35a)$$

$$S_{M,k+1} = S_{M,k} - S_{M,k} T_{M,k} S_{M,k}, \quad (35b)$$

$$D_{k+1}^+ = D_{k+1}^* S_{N,k+1} = S_{M,k+1} D_{k+1}^*. \quad (35c)$$

In the last formula, any of the two equivalent expressions for D_{k+1}^+ can be used. Note that we have encountered two distinct (but related by a permutation of notations) matrices $T_{N,k}$ and $T_{M,k}$, which can be referred to as the pseudo-T matrices. Unlike the true T-matrix, the pseudo-T matrices are guaranteed to exist as long as $\lambda > 0$. Analytical expressions for the pseudo-T matrices can be obtained by tedious but straightforward algebraic calculation. Here we adduce the final result:

$$T_{M,k} = \frac{1}{\mathcal{D}_k} \left[\gamma_k |a_{k+1}\rangle \langle b_{k+1}| D_k + \text{H.c.} \right. \\ \left. + \lambda^2 P_k |a_{k+1}\rangle \langle a_{k+1}| - Q_k D_k^* |b_{k+1}\rangle \langle b_{k+1}| D_k \right], \quad (36a)$$

$$T_{N,k} = \frac{1}{\mathcal{D}_k} \left[\gamma_k D_k |a_{k+1}\rangle \langle b_{k+1}| + \text{H.c.} \right. \\ \left. + \lambda^2 Q_k |b_{k+1}\rangle \langle b_{k+1}| - P_k D_k |a_{k+1}\rangle \langle a_{k+1}| D_k^* \right], \quad (36b)$$

where ‘H.c.’ stands for Hermitian conjugate of the preceding term and

$$\gamma_k = 1 + \langle a_{k+1} | D_k^+ | b_{k+1} \rangle, \quad (37a)$$

$$P_k = \langle b_{k+1} | S_{N,k} | b_{k+1} \rangle > 0, \quad Q_k = \langle a_{k+1} | S_{M,k} | a_{k+1} \rangle > 0, \quad (37b)$$

$$\mathcal{D}_k = |\gamma_k|^2 + \lambda^2 P_k Q_k > 0. \quad (37c)$$

As can be seen, the quantities P_k , Q_k and \mathcal{D}_k are guaranteed to be positive. Correspondingly, the main iteration step (35) is always well-defined. The iterations are formally terminated at $k = K$ and the final result is obtained as $A^+ = D_K^+$.

Note that the algorithm described here involves numerical operations on matrices and vectors whose elements are not known analytically. Therefore, this algorithm can not be used in the infinite-dimensional case.

7.3. Solution by direct matrix inversion

Here we assume that A is of infinite size, although the algorithm is also applicable if A is finite and square.

We wish to solve the equation $A|\mu\rangle = |\Phi\rangle$, where A is of the form (30). Let $x_k = \langle a_k | \mu \rangle$, $k = 1, \dots, K$. Then multiply the equation by $\langle a_j | D^{-1}$ from the left. This yields a set of K linear equations

$$x_j + \sum_{k=1}^K M_{jk} x_k = R_j, \quad (38)$$

where

$$M_{jk} = \langle a_j | D^{-1} | b_k \rangle, \quad R_j = \langle a_j | D^{-1} | \Phi \rangle, \quad (39)$$

and we have tacitly assumed that D^{-1} exists. Once the unknown quantities x_k are found by solving the finite-dimensional set (38) numerically, the vector $|\mu\rangle$ can be easily computed according to

$$|\mu\rangle = D^{-1} \left(|\Phi\rangle - \sum_{k=1}^K x_k |b_k\rangle \right). \quad (40)$$

It can be seen that the T-matrix is given by

$$T = \sum_{j,k=1}^K |b_j\rangle M_{jk}^{-1} \langle a_k|. \quad (41)$$

Note that (40) can be used to compute μ_n with arbitrary n . In this sense, (40) is truly the solution to the infinite-dimensional system of equations. The coefficients M_{jk} and R_j can be computed numerically or analytically, depending on the problem. In what follows, we discuss computation of these quantities for the particular case of star transform.

As was discussed in section 7.1, there exist two different ways to truncate (27), and here we use the particular truncation in which A is infinite (not truncated) while the components Φ_n with $|n| > n_{\max}$ are set to zero. Also, the matrix D in the case of star transform is diagonal. Correspondingly, the expression for R_j (39) contains only finite sums, which can be computed without any approximations. We also recall that the particular case of the star transform is obtained if $|b_k\rangle = s_k \alpha_k |a_k\rangle$. Consequently, the expression for M_{jk} takes the form

$M_{jk} = \langle a_j | D^{-1} | a_k \rangle s_k \alpha_k$. This expression involves infinite series in which all terms are known analytically. The series can be easily summed numerically¹. Convergence can be accelerated by utilizing the known large- n asymptote of the terms. Indeed, let us write the series for the matrix element $\langle a_j | D^{-1} | a_k \rangle$ explicitly:

$$\langle a_j | D^{-1} | a_k \rangle = \sum_{n=-\infty}^{\infty} \frac{1}{\beta_j + \kappa_n} \frac{1}{d_n} \frac{1}{\beta_k + \kappa_n} = \frac{1}{\beta_j \beta_k d_0} + \sum_{n=1}^{\infty} t_{jk,n}, \quad (42a)$$

$$t_{jk,n} = \frac{1}{\beta_j + \kappa_n} \frac{1}{d_n} \frac{1}{\beta_k + \kappa_n} + \frac{1}{\beta_j - \kappa_n} \frac{1}{d_{-n}} \frac{1}{\beta_k - \kappa_n}, \quad (42b)$$

where β_j are defined in (22) and we have accounted for $\kappa_{-n} = -\kappa_n$. We note that the factor

$$\beta_j \beta_k d_0 = iq \frac{u_{jy} u_{ky}}{u_{jz} u_{kz}} \sum_{l=1}^K \frac{s_l}{u_{ly}}$$

turns to zero when $q = 0$, but this case can be considered separately as is described in section 8.1 below. We can now consider the large- n asymptote of $t_{jk,n}$. To this end we use the following expansion of d_n :

$$d_n = i \left(\Sigma_1 \frac{1}{\kappa_n} + \Sigma_2 \frac{q}{\kappa_n^2} + \dots \right), \quad |n| \rightarrow \infty, \quad (43)$$

where

$$\Sigma_0 = \sum_{k=1}^K \frac{s_k}{|u_{kz}|}, \quad \Sigma_1 = \sum_{k=1}^K \frac{s_k}{u_{kz}}, \quad \Sigma_2 = -\sum_{k=1}^K \frac{s_k u_{ky}}{u_{kz}^2}. \quad (44)$$

The quantity Σ_0 is not used in this section but will be needed below in section 8.1. From this, we obtain the following expansion for $t_{jk,n}$:

$$t_{jk,n} = \tau_{jk} \frac{1}{\kappa_n^2} + O\left(\frac{1}{\kappa_n^4}\right), \quad n \rightarrow \infty, \quad (45)$$

where

$$\tau_{jk} = \frac{2i \left[(\beta_j + \beta_k) \Sigma_1 + q \Sigma_2 \right]}{\Sigma_1^2} = \frac{2iq \sum_{l=1}^K s_l}{\Sigma_1^2} \left(\frac{u_{jy}}{u_{jz}} + \frac{u_{ky}}{u_{kz}} - \frac{u_{ly}}{u_{lz}} \right). \quad (46)$$

Note that the above derivations assume that $\Sigma_1 \neq 0$. If $\Sigma_1 = 0$, the diagonal elements of D decay with n as $1/n^2$ or faster, and the infinite system of equations (27) with the right-hand side coefficients Φ_n truncated to zero for $|n| > n_{\max}$ does not have a solution. In section 8.1 below, we also show that the star transform with $\Sigma_1 = 0$ is not invertible at $q = 0$.

We can now use the result (46) to rewrite (42a) as

$$\langle a_j | D^{-1} | a_k \rangle = \frac{1}{\beta_j \beta_k d_0} + \frac{L^2 \tau_{jk}}{24} + \sum_{n=1}^{\infty} \left(t_{jk,n} - \frac{\tau_{jk}}{\kappa_n^2} \right). \quad (47)$$

¹ In the case $K = 2$, the series can be summed analytically. We do not discuss this result because M_{jk} can be computed numerically without any noticeable loss of precision for general K .

The terms $(t_{jk,n} - \tau_{jk}/\kappa_n^2)$ decay with n as $1/n^4$ and, therefore, fast numerical convergence of (47) can be expected.

The disadvantage of the method presented here is that it is not regularized and can become numerically unstable. However, $A^*A + \lambda^2\mathbf{I}$, similarly to A , contains one diagonal and a finite number of separable terms. We can use this fact to generalize the method for computation of the Tikhonov-regularized pseudo-inverse.

7.4. Computational complexity

Each iteration of the methods of section 7.2 involves either one or a few matrix-vector products. Assuming for simplicity $N \times N$ square A , we obtain the computational complexity of $O(N^2)$ per iteration and the total computational complexity of $O(KN^2)$. It should be emphasized that the matrix-vector products involve a diagonal matrix at the first iteration but for the subsequent iteration the diagonality is lost. This is why we estimate the computational complexity to be $O(N^2)$ per iteration rather than $O(N)$. We should also keep in mind that A is parameterized by the Fourier variable q and that numerical inversion of $A(q)$ must be performed for every value of q used. Consider an example of reconstructing a rasterized $N \times N$ -pixel image. In this case, the number of discrete values of q to be used is also N . Then the cumulative computational complexity of reconstructing the image by the iterative methods is $O(KN^3)$. For comparison, a method based on pixelization of the image and representing the star transform (1) as a set of N^2 linear equations will result in the computational complexity of $O(N^6)$ for direct methods or $O(MN^4)$ for iterative methods such as the conjugate gradient descent, where M is the number of iterations needed for convergence. Therefore, Fourier-space representation of the star transform and iterative inversion of the resulting equations (parameterized by q) can result in a very significant computational advantage.

An even larger computational advantage is gained by using the direct method of section 7.3. In this method the property of diagonality in the matrix-vector products is not lost and the computational complexity per one q is $O(\max(N, n_{\text{sum}}) + O(K^3))$, where n_{sum} is the maximum value of n needed for accurate approximation of the series of the type (47). Assuming that $N \gg n_{\text{sum}}, K^3$, the cumulative computational complexity of reconstructing an image is $O(N^2)$.

8. Numerical stability

In this section, we discuss the numerical stability of inverting the matrices $A(q)$. Stability is understood here in the algebraic sense, namely, we require the condition number of $A(q)$ (the ratio of its maximum and minimum singular values) to be within the numerical range that still allows for reliable numerical arithmetic. Simply stated, we say that inversion of $A(q)$ is stable if the inverse of $A(q)$ can be computed with numerical accuracy up to the usual round-off errors. This definition of stability is quite useful in practice, and several illustrative examples will be given below. It should not be confused with the stability of inverting the star transform in suitable function spaces, a topic which is deserving of additional research.

Further, we will analyze the stability of inverting $A(q)$ for two different cases: $qL \ll 1$ and $qL \gg 1$. Results for general q are not available at this point, although we have seen numerical evidence that $A(q)$ can be ill-posed at intermediate values of q (that is, for $qL \sim 1$) even if it is well posed in the two cases mentioned above. The instability at intermediate

values of q has occurred in numerical experiments where all rays involved crossed the same boundary. The corresponding artifacts were localized near the boundary of the strip and were not very severe. We do not show such examples in section 9 because the geometries with all rays crossing the same boundary are not very important in practice.

The main concern for image reconstruction is instability at large values of q . This instability results in high-frequency noise which is not localized and tends to obscure the whole image. Examples of such noise can be found, for example, in [17].

We finally note that the Fourier reconstruction formula involves the inverse of $A(q)$ inside an integral over q . Therefore, if inversion of $A(q)$ is unstable in a very narrow interval of q (that is, $A^{-1}(q)$ has a small integral weight), such instability is not of practical importance.

We now consider the two cases $qL \rightarrow 0$ and $qL \rightarrow \infty$ separately.

8.1. The case $qL \rightarrow 0$

The limit $q \rightarrow 0$ of (25) is not immediately obvious; the cases $n = 0$ and $n \neq 0$ must be considered separately and, in the first case, the function $\alpha_k(q)$ must be expanded to second order. The computations are, however, routine, and we adduce here the final result. For $q = 0$, (25) takes the following form:

$$\Phi_0 = \Sigma_0 \frac{\mu_0 L}{2} - i \Sigma_1 \sum_{m \neq 0} \frac{\mu_m}{\kappa_m}, \quad n = 0, \quad (48a)$$

$$\Phi_n = i \Sigma_1 \frac{\mu_n - \mu_0}{\kappa_n}, \quad n \neq 0, \quad (48b)$$

where Σ_0 , Σ_1 and Σ_2 are defined in (44).

In order for (48) to be invertible, the factors Σ_0 and Σ_1 defined above must be nonzero simultaneously. It is possible to select such ray geometries that this condition holds. For example, consider the $K = 4$ star with the coefficient matrix c_{jk} given in section 3. In this case $s_1 = s_3 = 1$ and $s_2 = s_4 = -1$. If we take $u_{1z} = u_{3z} = u/2$ and $u_{2z} = u_{4z} = -u$, where $0 < u < 1$, then $\Sigma_0 = 2/u$ and $\Sigma_1 = 6/u$.

If $\Sigma_0 \neq 0$ and $\Sigma_1 \neq 0$, we can invert (48) analytically. Using the equality $\sum_{m \neq 0} (1/\kappa_m) = 0$, we find that the solution to (48) is

$$\mu_0 = \frac{2}{L \Sigma_0} \sum_{m=-\infty}^{\infty} \Phi_m = \frac{2\tilde{\Delta}(0)}{\Sigma_0}, \quad n = 0, \quad (49a)$$

$$\mu_n = \mu_0 - i \frac{\kappa_n \Phi_n}{\Sigma_1}, \quad n \neq 0. \quad (49b)$$

where $\tilde{\Delta}(q)$ is defined in (19). We note that the coefficients μ_n defined by (49) can be obtained by a term-by-term differentiation of a Fourier series of a discontinuous function. The resultant series does not converge in the usual sense but rather yields two delta-functions centered at $z = 0$ and $z = L$. Therefore, the series should not be evaluated numerically too close to the boundaries of \mathbb{S} . However, sufficiently far from these boundaries, the series converges quite fast to the correct result, that is, to the function $\tilde{\mu}(0, z)$, as can be easily verified numerically using a number of examples. The region of bad convergence is small and can be reduced in practice to one or two pixels (on each side) of an image containing ~ 100 pixels in the Z -direction. We emphasize that in order to obtain good convergence, the coefficient μ_0 must be determined with sufficient precision from (49a). thus, (48) can be solved in two steps. In the

first step we compute μ_0 according to (49a). This step is well-defined as long as $\Sigma_0 \neq 0$. In the second step we use the previous result to compute μ_n for $n \neq 0$ according to (49b). This step is well-defined if $\Sigma_1 \neq 0$.

If μ_0 is known or can be measured independently, then the occurrence of $\Sigma_0 = 0$ does not pose a problem for image reconstruction because only (49b) needs to be used in this case. In section 6, we have explained how μ_0 can be obtained using measurements of ballistic (non-scattered) rays. However, the occurrence of $\Sigma_1 = 0$ is truly problematic. We will refer to the imaging geometries with $\Sigma_1 = 0$ as *symmetric*. We note that arrangements with $\Sigma_1 = 0$ have been inadvertently used by us before, e.g., in [16]. Image reconstruction is still possible in this case, but it is not possible to estimate correctly the integrals of the type $\int_{-\infty}^{\infty} \mu(y, z) dy$ ². This point is illustrated graphically in figure 2 where we show two star configurations with $K = 2$ and $\Sigma_1 = 0$. The (a) configuration has been used by us in [16, 17]. It can be seen that the long vertical inhomogeneity shown in the figures by a dark-shaded rectangle is difficult to reconstruct using the ray geometries shown in the figure. The data function in these two cases is invariant with respect to the shift of the inhomogeneity along the Z axis as long as the rays do not intersect or touch the upper or lower sides of the rectangle. The mathematical manifestation of this observation is the ill-posedness of (25) at $q = 0$.

8.2. The case $qL \rightarrow \infty$

Consider equation (40) in the limit $qL \rightarrow \infty$. It can be seen that the second term inside the brackets is larger than the first term by a factor $O(qL)$. Conversely, the second term in the right-hand side of equation (27) is smaller than the first term by a factor $O(1/qL)$. This is not surprising. The separable terms in the expression for A are due to the boundaries of \mathbb{S} . In the high spatial frequency limit, the boundaries are unimportant and the diagonal term in the expression for A dominates the separable terms. This fact can be used to analyze the stability of inverting $A(q)$ when q is large. We emphasize that the high-frequency instability is of primary concern for image reconstruction, as will be illustrated with several numerical examples below.

Since in the limit considered the diagonal term D dominates in (40), we will analyze the conditions under which the diagonal elements $d_n(q)$, are nonzero for all values of n .

Let us write $d_n(q)$ in the form

$$d_n(q) = \sum_{k=1}^K \frac{is_k}{\hat{\mathbf{u}}_k \cdot (q, \kappa_n)} = \frac{i}{|(q, \kappa_n)|} \sum_{k=1}^K \frac{s_k}{\hat{\mathbf{u}}_k \cdot \hat{\mathbf{v}}}, \quad (50)$$

where $\hat{\mathbf{v}} = (q, \kappa_n) / \sqrt{q^2 + \kappa_n^2}$ is the unit vector pointing in the direction of (q, κ_n) . Obviously, $\hat{\mathbf{v}}$ is defined only if $(q, \kappa_n) \neq 0$, which is the case here. We then define the function

$$f(\theta) = \sum_{k=1}^K \frac{s_k}{\cos(\theta - \theta_k)}, \quad (51)$$

where θ and θ_k are the polar angles of $\hat{\mathbf{v}}$ and $\hat{\mathbf{u}}_k$ in the YZ plane. It is clear that $d_n(q)$ cannot turn to zero if $f(\theta)$ does not have zeros for real θ . On the other hand, if $f(\theta)$ has zeros, then

² Note that if $\Sigma_1 = 0$ but $\Sigma_0 \neq 0$, it is still possible to estimate correctly $\mu_0(q)$ and, by Fourier transform, the quantity $\int_0^t \mu(y, z) dz$.

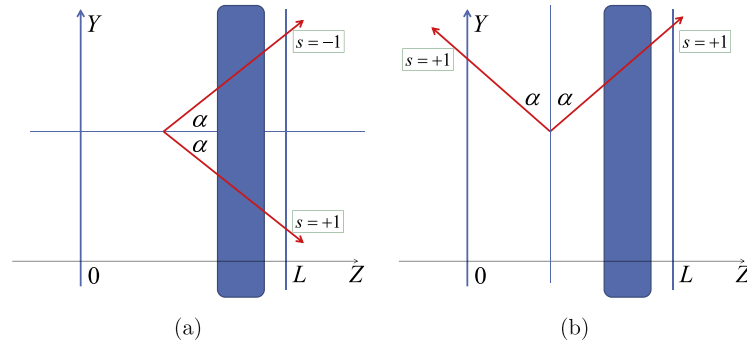


Figure 2. Two-ray symmetric stars with $\Sigma_1 = 0$. Here α denotes the angle. The long vertical inhomogeneity shown by the shaded rectangle is difficult to reconstruct using these geometries. We note that image reconstruction for the two cases shown in the figure is also ill-posed in the limit $qL \rightarrow \infty$ because the number of rays is even (see section 8.2 below).

we can find an arbitrarily small element $d_n(q)$. Therefore, the sufficient and necessary condition of stability of inverting the star transform is that the function $f(\theta)$ (51) does not have zeros on the real axis. We note that $f(\theta + \pi) = -f(\theta)$ and, therefore, it is sufficient to consider the interval $0 \leq \theta < \pi$.

Of course, for any given set of $\hat{\mathbf{u}}_k$ and s_k , it is a trivial matter to plot $f(\theta)$ and visually determine whether it has zeros or not. We will, however, show that $f(\theta)$ always has zeros if the number of rays K is even and, moreover, if K is odd, then $f(\theta)$ has zeros if the vectors $s_k \hat{\mathbf{u}}_k$ can be placed in the same half-plane (have simultaneously non-negative projections onto the same axis). Therefore, for inversion of the star transform to be well-posed for $qL \gg 1$, the following two necessary conditions must hold: (i) the number of rays K should be odd; (ii) the vectors $s_k \hat{\mathbf{u}}_k$ should not be contained in the same half-plane. This result will allow us to exclude star configurations that are ill-posed *a priori*.

We will now show how the above conditions have been obtained. We start with the observations that, when $\theta = \theta_k \pm \pi/2$ (that is, when $\hat{\mathbf{v}}$ is perpendicular to one of the unit vectors $\hat{\mathbf{u}}_k$), $f(\theta)$ diverges, and that $f(\theta_k - \epsilon)$ and $f(\theta_k + \epsilon)$ have different signs, where ϵ is an infinitesimal constant. Between the singular points, $f(\theta)$ is continuous. We therefore must determine whether $f(\theta)$ changes sign in at least one of the intervals where it is continuous. To this end, diagrams such as those shown in figure 3 can be useful. Consider a circle and draw each vector $s_k \hat{\mathbf{u}}_k$ as an arrow originating from the circle center. Then draw a perpendicular to each arrow, also through the circle center, as is shown by the dashed lines. Near each point where a dashed line intersects the circumference, draw a pair of signs, plus and minus, on each side of the line. These signs indicate the sign of $f(\theta)$ near the singularity. There are two singularities and four signs associated with each line. When placing the signs, a couple of obvious rules must be obeyed: different signs are placed on different sides of a singularity and similar signs are placed in each of the two half-planes created by a given line. Note that changing the sign of s_k corresponds to changing all signs associated with the corresponding line. Now, let us, starting from an arbitrary point, move along the circumference in any direction and make a complete revolution. Crossing the dashed lines corresponds to crossing the singularities of $f(\theta)$ while motion from one dashed line to the next corresponds to the

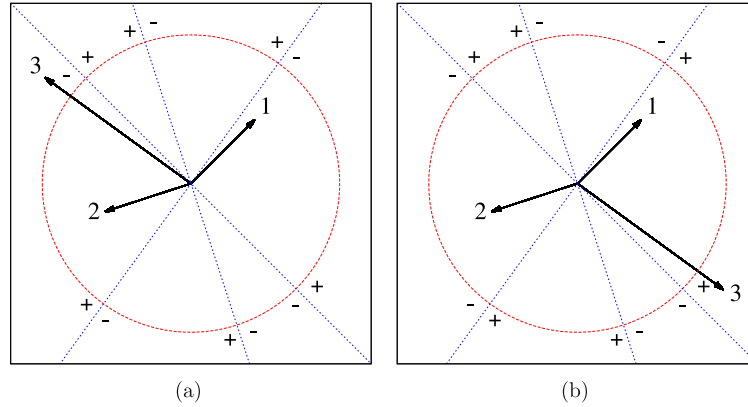


Figure 3. Sign diagrams for the function $f(\theta)$ in the case $K = 3$ and $s_1 = s_2 = 1$, $s_3 = -2$. In both cases (a) and (b), $\theta_1 = 0.25\pi$ and $\theta_2 = 1.1\pi$, and the third angle is $\theta_3 = -0.2\pi$ (a) and $\theta_3 = 0.8$ (b). Angles are measured from the positive direction of the Z-axis in the counter-clockwise direction. It can be seen that the drawing of the third line, which runs from the top right to the bottom left corner, removes the contradiction in case (b) but not in (a). Also, in (a) all vectors $s_k \hat{\mathbf{u}}_k$ can be placed in the same half-plane while in (b) the same is not true.

intervals of θ where $f(\theta)$ is continuous. If the motion over an interval of continuity connects two opposite signs, as is the case in the top and bottom segments of the diagram in figure 3(a), then $f(\theta)$ has at least one zero in that interval.

First, let us consider the case of even K . Assume that $f(\theta)$ does not have zeros. Then take an arbitrary line and place the signs associated with it in accordance with the above two rules but otherwise arbitrarily, and start moving from the two singular points just ‘signed’ in, say, the clock-wise direction. The hypothesis that $f(\theta)$ does not have zeros forces a unique choice of signs for the next two singular points. Continue this process until only one ‘unsigned’ line remains. If K is even, then this last line can not be ‘signed’ without violating the underlying hypothesis. There will appear two symmetrically-situated intervals of θ (contained between two lines) where $f(\theta)$ is continuous and changes sign. Therefore, we have arrived at a contradiction and $f(\theta)$ must have zeros. For example, if $K = 2$, $f(\theta)$ has zeros when $\theta = \pm \theta [s_2 \hat{\mathbf{u}}_1 + s_1 \hat{\mathbf{u}}_2]$, where the last expression denotes the polar angle of the vector in the square brackets.

Consideration of odd K is somewhat more complicated. We can again make the hypothesis that $f(\theta)$ does not have zeros and start from $K - 1$ rays, where $K - 1$ is even. As discussed above, there will appear two intervals of θ where $f(\theta)$ is continuous and changes sign. Since these intervals are contained between the same two lines, each of them can be divided in two by drawing one additional, K th line. However, the contradiction is removed only for one particular choice of signs associated with this last line. This is illustrated in figure 3 for $K = 3$. Here we start with $K - 1 = 2$ and break the two intervals of θ that connect opposite signs with a third line. In the considered example, the choice of signs for this third line depends on the direction of $\hat{\mathbf{u}}_3$ (the coefficient s_3 is the same in cases (a) and (b)). In the case (b), the contradiction to the original hypothesis is removed but in the case (a) it is not. It can be further determined by inspection that in the case (a) all three vectors $s_k \hat{\mathbf{u}}_k$ (shown by

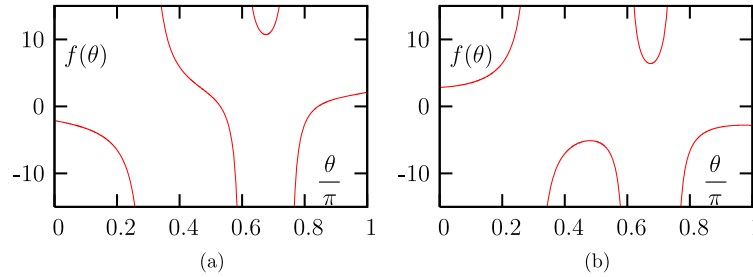


Figure 4. The function $f(\theta)$ for the two cases (a) and (b) shown in figure 3.

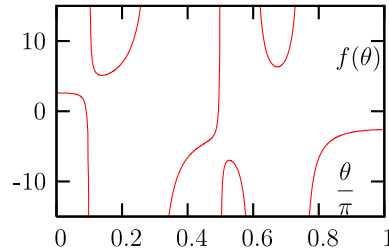


Figure 5. A counter-example to conditions (i)–(iii) of section 8.2. Function $f(\theta)$ for the same rays as in figure 3(b) plus two additional rays with $s_4 = -0.1$, $s_5 = 0.1$ and $\theta_4 = \pi$, $\theta_5 = 0.6\pi$. The total number of rays is $K = 5$. Conditions (i)–(iii) of section 8.2 are satisfied but $f(\theta)$ has zeros.

arrows) are contained in the same half-plane. More generally, we can follow similar considerations for an arbitrary K to show that the K th ray removes the contradiction only if $s_K \hat{\mathbf{u}}_K$ is not contained in the same half-plane as the vectors $s_1 \hat{\mathbf{u}}_1, \dots, s_{K-1} \hat{\mathbf{u}}_{K-1}$.

The function $f(\theta)$ for the two cases shown in figure 3 is plotted in figure 4 in the interval $0 \leq \theta \leq \pi$. Image reconstruction for the two star configurations of figure 3 are shown in figure 5.

We emphasize that the condition based on zeros of $f(\theta)$ is necessary and sufficient. However the conditions (i)–(ii) of this section are necessary only. We can easily find a case when (i)–(ii) are satisfied but $f(\theta)$ has zeros. A counter-example of this type for $K = 5$ is shown in (6). Here we have used the same three rays as in figure 3(b) plus two additional rays with $s_4 = -0.1$, $s_5 = 0.1$ and $\theta_4 = \pi$, $\theta_5 = 0.6\pi$. In this case the conditions (i)–(ii) are obviously satisfied and, additionally, we have preserved the condition $\sum_k s_k = 0$, which is required for simultaneous reconstruction of attenuation and scattering.

Counter-examples of this type can be generated by considering a star configuration with $K - 2$ rays, where $K - 2$ is odd, $\sum_{k=1}^{K-2} s_k = 0$, and in which $f(\theta)$ has no zeros. We then add to the configuration a pair of rays with the coefficients $s_{K-1} = -s_K$ and $|s_K| \ll 1$ to obtain a new configuration with K rays. Addition of such ray pairs with a ‘wrong’ sign of s_K can generate zeros in $f(\theta)$ without violating any of the conditions (i)–(ii) or the sum rule $\sum_k s_k = 0$. However, such counter-examples are of little practical concern. Indeed, in all such examples, the derivative of $f(\theta)$ near its zeros is very large. This can be clearly seen in figure 6. As was

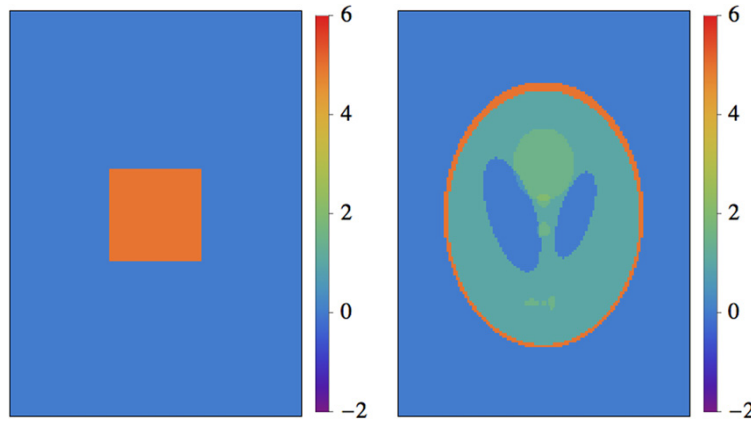


Figure 6. The phantoms used in the numerical experiments: square (left) and Shepp–Logan phantom (right). The quantity depicted by the color scale is the dimensionless parameter μL . Outside the inhomogeneities, this quantity is set to zero. The same color scale as in this figure is used in all reconstructed images shown below. Whenever the reconstructed values are outside the interval $-2 < \mu L < 6$, we use the following color clipping: the areas with $\mu L < -2$ are shown by black color and the areas with $\mu L > 6$ are shown by white (this clipping is not used in this rendering of phantoms but will appear in the reconstructed images below).

already mentioned, the inverse of $A(q)$ appears in an integral over q . Therefore, singularities of this sort have small integral weight and do not result in noticeable image distortions or artifacts. Of course in the limit $s_x \rightarrow 0$, the two additional rays do not influence image reconstruction at all, even though $f(\theta)$ still formally has zeros.

9. Simulations

In this section, we report several numerical experiments, which illustrate possible applications of the star transform to imaging. We use different star geometries, some of which are applicable to simultaneous reconstruction of scattering and attenuation coefficients. However, at this stage, we show only reconstructions of the attenuation. Physically, the reconstructed quantity is $\delta\mu(y, z)$, the deviation of the total attenuation coefficient from its background value (see the discussion around equation (3)). As such, the reconstructed quantity is zero outside of a few finite objects. Strictly speaking, the data function used is $\delta\Phi$ defined in (3). However, as elsewhere in the paper, we refer to the reconstructed quantity as to μ and to the data function as to Φ .

We use equation (1) to generate the data $\Phi(\mathbf{R})$ by computing the ray integrals analytically, which can be regarded as ‘inverse crime’ (that is, generating the data from the same model as is used for image reconstruction). We make our numerical tests more realistic by including Poissonian random noise in the data. The noise is added as follows. First, we recall that the data function Φ is constructed as a linear superposition of functions ϕ_{jk} , which, in turn, are related the physical measurements of intensity W_{jk} by (5). We therefore add noise directly to the functions ϕ_{jk} and then construct the data function of the star transform, Φ , according to (9). In this expression, the numerical coefficients c_{jk} are known precisely and determined as

Table 1. Weight coefficients s_k and ray angles θ_k for all cases considered in section 9. The angles are measured with respect to the positive direction of the Z -axis. Counter-clockwise rotation direction is assumed to be positive. Also shown for each case are the number of zeros NZ of the function $f(\theta)$ (51) in the interval $0 \leq \theta \leq \pi$ and the expansion coefficients Σ_0 and Σ_1 (44) rounded off to three significant figures. The cases 3a and 3b correspond to the cases (a) and (b) of figures 3 and 4.

Case	1a	1b	2a	2b	3a	3b
K	2		3		3	
s_1	1		1		1	
s_2	1		1		1	
s_3	N/A		1		-2	
θ_1/π	1	0.82	1		0.25	
θ_2/π	0.25	0.23	0.25		1.1	
θ_3/π	N/A		-0.25	-1/6	-0.2	0.8
NZ	1	1	0	0	2	0
Σ_0	2.41	2.52	3.83	3.57	-0.01	-0.01
Σ_1	0.41	0.15	1.83	1.57	-2.11	2.83

described in section 3 but the terms ϕ_{jk} contain noise. To add noise to ϕ_{jk} , we use the following procedure. For each data point ϕ_{jk} , we first compute an integer \overline{M} according to $\overline{M} = \text{nint}[\mathcal{N} \exp(-\phi_{jk})]$, where \mathcal{N} is a fixed integer number and $\text{nint}(x)$ denotes the nearest integer to x . We then choose randomly a number M from a Poissonian distribution whose average is \overline{M} . Finally, we compute the noise-affected data point ϕ'_{jk} as $\phi'_{jk} = -\log(M/\mathcal{N})$. It can be seen that $\phi'_{jk} \rightarrow \phi_{jk}$ in the limit $\mathcal{N} \rightarrow \infty$. This method of adding noise is somewhat *ad hoc*, but it has a well-controlled behavior and distorts randomly the reconstructed images when \mathcal{N} is not very large. Undoubtedly, more realistic noise models will be required in the future.

In all reconstructions, the data function and the reconstructed image were sampled on a square grid with the step h , where $h = L/126$ (L being the transverse width of the strip). Thus, there were $N=125$ real-space samples of $\Phi(Y, Z)$ in the Z -direction, and similarly in Y . This corresponds to the truncation of Fourier coefficients $\mu_n(q)$ such that $-n_{\max} \leq n \leq n_{\max}$, where $n_{\max} = (N - 1)/2 = 62$. The Fourier variable q was similarly sampled and truncated. Correspondingly, we considered a square matrix A of the size $N \times N$.

To generate forward data, we employ two different phantoms: a square and the Shepp–Logan phantom. The attenuation coefficient of the square phantom was $\mu = 5/L$ and in the case of Shepp–Logan phantom it varied from $\mu = 1/L$ to $\mu = 5/L$ (inside the inhomogeneities). In the background, we had $\mu = 0$. The phantoms are shown in figure 6.

Inversion of the star transform with two and three rays has been simulated. The exact parameters of the imaging geometries used are listed in table 1, where we also list the number of zeros of the function $f(\theta)$ in the interval $0 \leq \theta \leq \pi$ and the quantities Σ_0 , Σ_1 (44) for each geometry. As was discussed in the paper, these quantities can be used to ascertain the ill-posedness of the star transform. The imaging geometries are grouped into three cases. In Cases 1 and 2, all coefficients s_k are equal to unity. As was discussed in section 3,

these types of star transform are applicable to reconstruction of purely absorptive contrast in a uniform scattering background. The geometry of Case 3b satisfies all the conditions (i)–(iv) of section 3 and therefore can be used to reconstruct simultaneously attenuation and scattering contrast. We note that the star transform of Case 3b can be obtained from physical measurements as is suggested by the first coefficient table in the end of section 3.

9.1. Case 1

We start with image reconstructions for the ray geometries of Case 1 ($K = 2$; see table 1 for more detail). Case 1 is only applicable to reconstructing absorptive contrast in a medium with a spatially-uniform scattering coefficient. Results for four different levels of noise are shown in figure 7. The images were obtained by Tikhonov-regularized pseudo-inverse as described in section 7.2.

The reconstructions contain fairly severe artifacts, especially in Case 1a. This is a consequence of the ill-posedness of the star transform at large values of q , in agreement with section 8.2 (K is even in Case 1). Similar artifacts have been observed by us previously in the purely numerical reconstructions utilizing two-ray geometries [15, 16]. The artifacts in Case 1b are less severe. We attribute this to the fact that the derivative of $f(\theta)$ near the point where it turns to zero is much larger in Case 1b than in Case 1a. Correspondingly, the singularity in Case 1a has a larger integral weight. The role of the derivative of $f(\theta)$ near its zeros was briefly discussed in the end of section 8.2. Note that in Case 1, regularization was required to obtain a recognizable image at all noise levels considered, including the case with no noise. An illustration of the effects of regularization is given in figure 8.

9.2. Case 2

We now turn to Case 2 ($K = 3$; see table 1 for more detail). Similarly to Case 1, Case 2 is only applicable to reconstructing absorptive contrast in a medium with a spatially-uniform scattering coefficient. Reconstructions are shown in figure 10. In Case 2, the function $f(\theta)$ does not have zeros and conditioning of the inverse problem is significantly improved. Correspondingly, we have obtained reasonable reconstructions at all noise levels considered without regularization, that is, by using $\lambda = 0$. Analytically, pseudo-inverse with $\lambda = 0$ is indistinguishable from the ordinary inverse, and we have verified this fact numerically.

Although figure 9 shows reasonable reconstructions obtained with $\lambda = 0$, the effects of regularization deserve additional discussion. We have verified that in Case 2 regularization does not yield noticeable improvement of image quality at all noise levels used. This result is expected when one is inverting a well-posed operator such as the forward operator of the discrete Fourier transform, which has a flat spectrum of singular values. In such cases, introduction of Tikhonov regularization is not justified. However, in the case considered here, the spectrum of singular values is not flat. Apparently, there exists at least one singular value, which is numerically small, yet not small enough to cause significant instability at $\lambda = 0$. Under the circumstances, regularization with an inappropriately chosen parameter λ can produce image artifacts, as is illustrated in figure 10. It can be seen that the images with $\lambda = 0$ and $\lambda = 2 \times 10^{-2}$ are comparable in quality, although in the first case the reconstructed boundary of the square inhomogeneity is sharper while in the second case the image appears to be less noisy. At the intermediate values of λ (e.g., $\lambda = 10^{-4}$) a severe artifact appears in the reconstructions. The artifact has the form of oscillations whose wave vector is aligned with

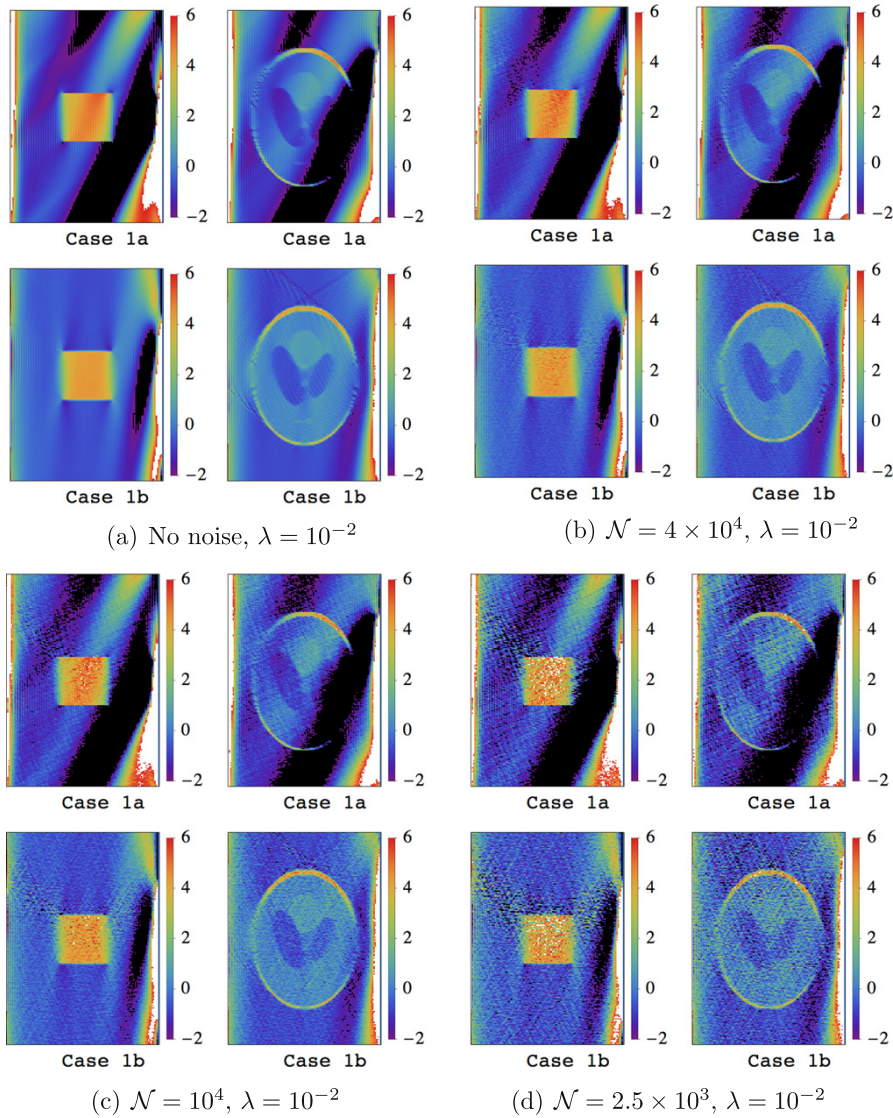


Figure 7. Case 1: image reconstructions for various levels of noise and regularization parameter λ , as labeled. Here \mathcal{N} is the integer parameter controlling the level of Poissonian noise in the data. the following clipping of the color scale has been used: Black color corresponds to numerical values $\mu L < -2$ and white color corresponds to $\mu L > 6$.

the Z-axis. We attribute this to the fact that the Fourier series expansion of the image does not converge in the usual sense at $q = 0$, as was discussed in section 8.1 (after equation (49)). Note, however, that in the more practically-important Case 3, this type of artifact is not present.

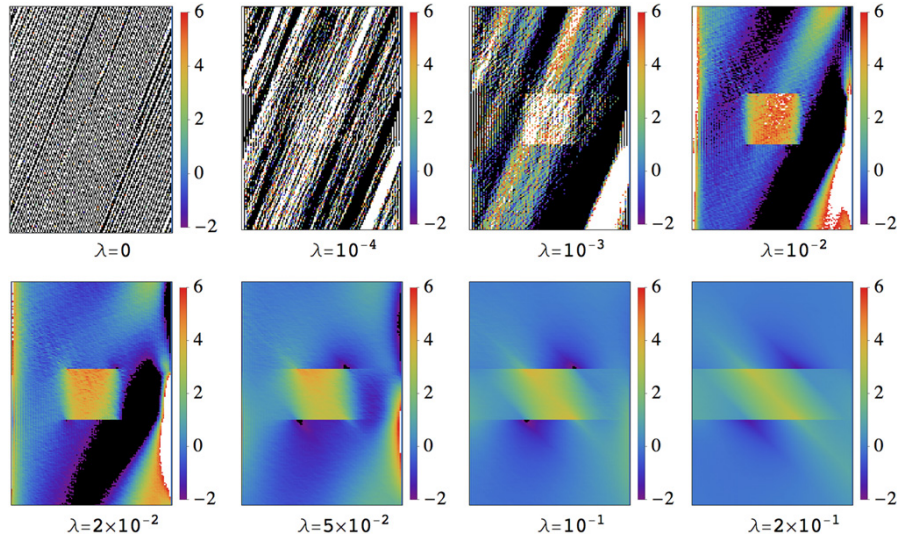


Figure 8. Case 1a: Effects of regularization for the noise level $\mathcal{N} = 10^4$ and different regularization parameters λ , as labeled.

9.3. Case 3

We now turn to Case 3 ($K = 3$; see table 1 for more detail). Reconstruction results for this geometry are shown in figure 11. Here the weight coefficients s_k satisfy the condition $\sum_k s_k = 0$, which makes possible simultaneous reconstruction of attenuation and scattering. Note that the Cases 3a and 3b correspond to the cases (a) and (b) illustrated in figures 3 and 4. The function $f(\theta)$ has zeros in Case 3a but not in Case 3b. Correspondingly, Case 3b allows for a stable inversion. Indeed, it can be seen that the reconstructions in Case 3b are much more stable in the presence of noise than in Case 3a. Without regularization, addition of noise to Case 3a results in noisy images that do not resemble the phantom. This can be alleviated by using Tikhonov regularization, as shown in figure 12. However, when compared at the same level of regularization, the image quality is always better in Case 3b. Note that the oscillating artifact that was seen in Case 2 at intermediate values of λ (figure 11) is not present in Case 3.

10. Discussion

In this paper, we have introduced the star transform as a generalization of the broken-ray transform, which was studied by us previously in [15–17]. We have shown that this generalization can be useful for improving the stability of inversion. Theoretical considerations demonstrate that a linear combination of individual measurements that leads to the star transform is required when the scattering coefficient of the medium is not spatially uniform and if one wishes to formulate the inverse problem with respect to only one unknown function, $\mu(\mathbf{R})$. The previously-developed approaches to inverting the broken ray transform in a medium with non-uniform scattering [16, 17] were, in fact, special cases of the more general framework introduced in this paper. This framework, based on constructing and inverting the star transform, can be used to avoid ill-posedness and thereby improve the image quality.

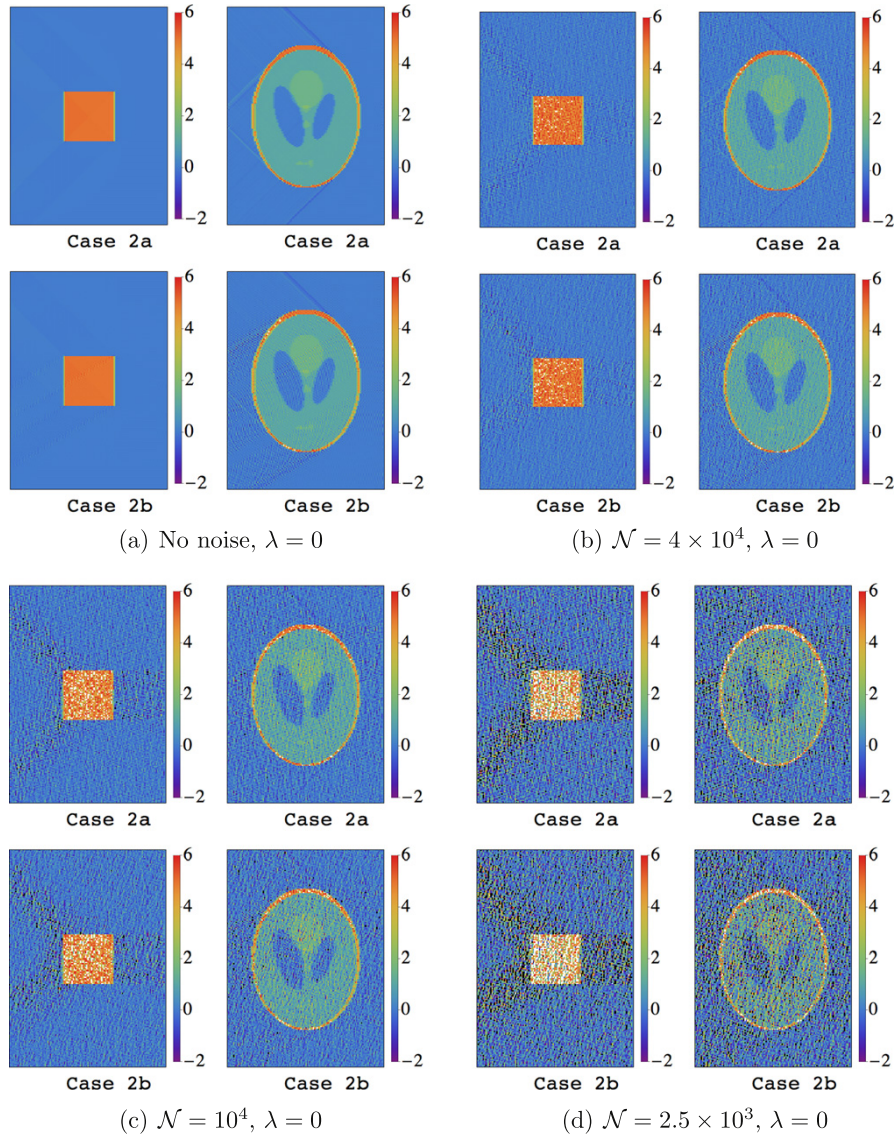


Figure 9. Same as in figure 8 but for Case 2; $\lambda = 0$ in all cases (regularization is not used).

Easy-to-use necessary conditions for avoiding the ill-posedness have been formulated in section 8. However, these conditions are not sufficient. Conditions for numerical stability have been formulated separately in the limits $qL \ll 1$ and $qL \gg 1$. A necessary and sufficient condition of numerical stability for $qL \ll 1$ is $\Sigma_0 \neq 0$, $\Sigma_1 \neq 0$, where Σ_k are defined in (44). However, if measurements of ballistic rays are used, the condition $\Sigma_0 \neq 0$ is no longer

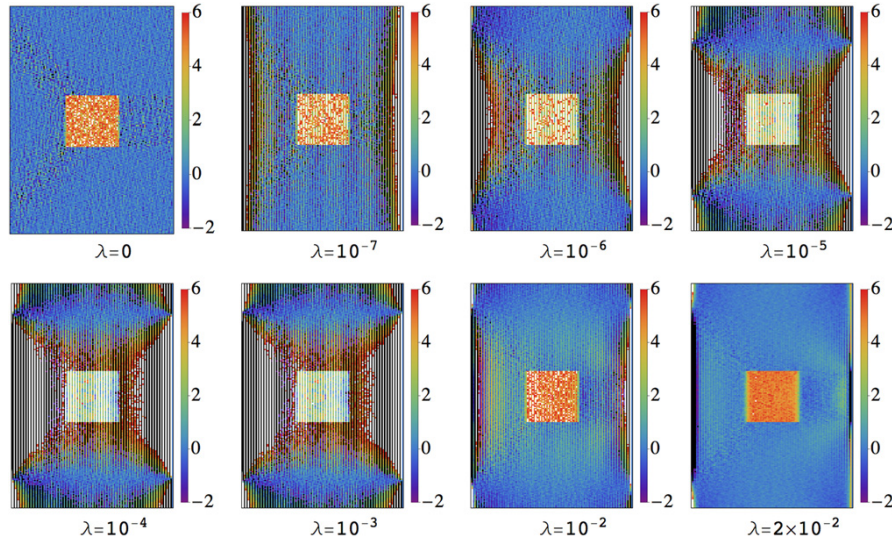


Figure 10. Case 2a: effects of regularization for the noise level $\mathcal{N} = 10^4$ and the regularization parameter λ , as labeled.

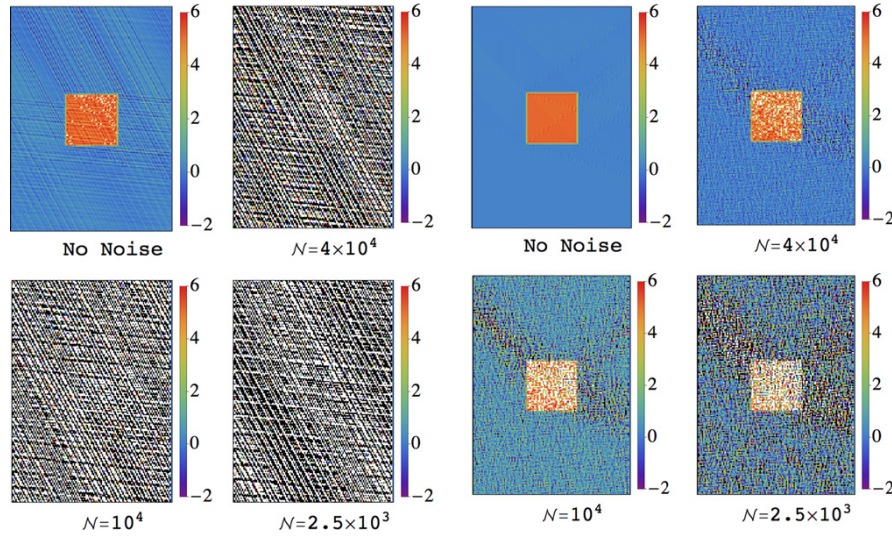


Figure 11. Case 3: Reconstructions of a square inhomogeneity for different levels of noise and $\lambda = 0$ (without regularization).

necessary. In the limit $qL \gg 1$, a necessary and sufficient condition for stability can be obtained by plotting the function $f(\theta)$ [defined in (51)] and visually determining whether it has zeros in the interval $0 \leq \theta \leq \pi$. For intermediate values of q , no analytical results concerning numerical stability have been obtained. However, we have seen numerical evidence

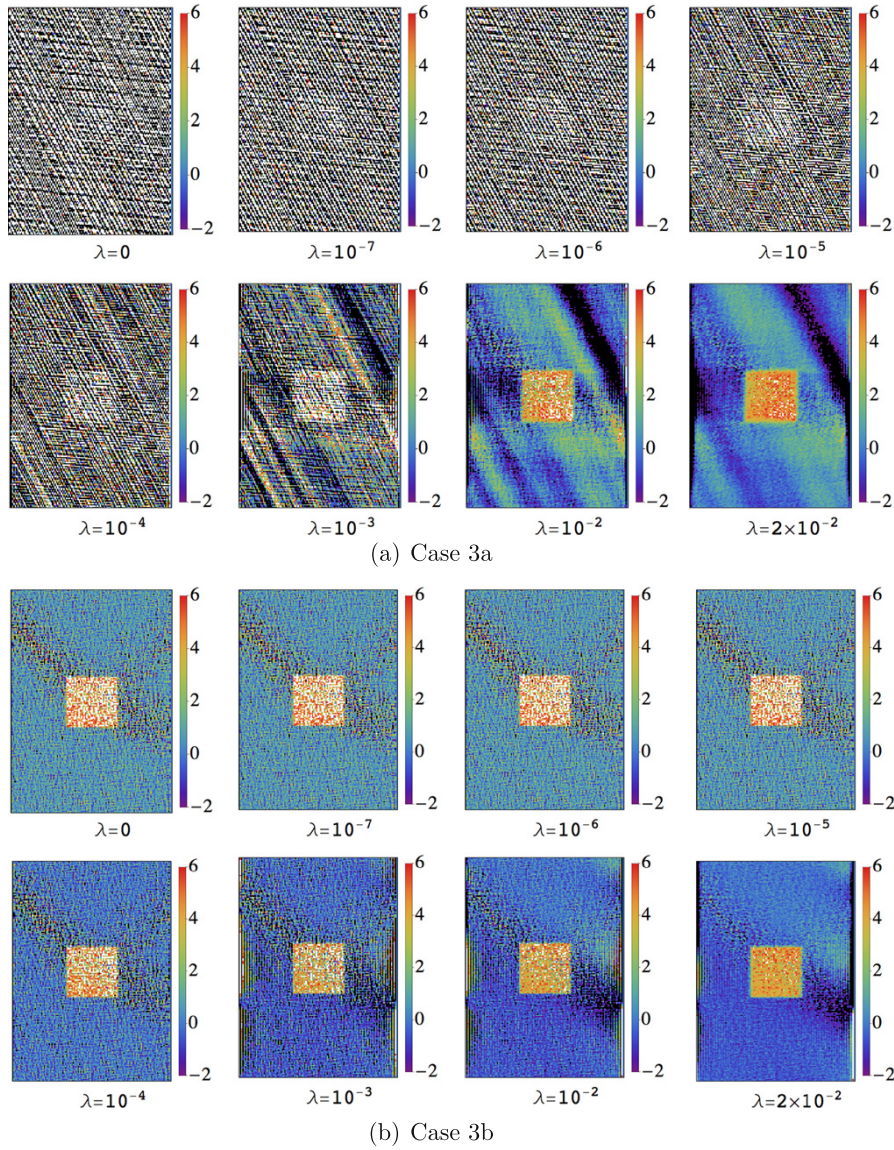


Figure 12. Case 3: effects of regularization for the noise level $\mathcal{N} = 10^4$ and the regularization parameter λ , as labeled.

that $A(q)$ can be ill-posed for $qL \sim 1$ in the imaging geometries where all rays involved crossed the same boundary.

We have also developed in section 7 several computationally efficient methods for inverting or pseudo-inverting a matrix of the form (30). Fourier-space inversion of the star transform is obtained as a special case. However, the results of section 7 may have a broader utility because matrices of the form (30) are commonly encountered in applications.

Finally, we have provided an initial numerical test of an imaging modality based on inverting the star transform. We have reconstructed both a square phantom and the Shepp–Logan phantom at various levels of Poissonian noise. Imaging geometries suitable for simultaneous reconstruction of the scattering and attenuation coefficients have been used, although in this paper only reconstruction of attenuation has been demonstrated.

Therefore, we have shown that the star transform is a feasible approach to imaging the medium with the account of single scattering. But is formation of the star transform necessary? As is described in detail in section 3, the star transform is obtained by taking certain linear combinations of individual ‘measurements’ $\phi_{jk}(\mathbf{R})$. Of course, one can not claim that taking these linear combinations provides additional useful information about the image. In fact, the complete set of measurements $\phi_{jk}(\mathbf{R})$ contains all experimentally-available information. The question is, therefore, whether one should use the functions $\phi_{jk}(\mathbf{R})$ directly or form the star transform.

The answer to this question depends on the numerical method used for reconstruction, available computational resources and statistical properties of the measurement noise. One possible approach is the following. Let us take all $K(K-1)/2$ mathematically-independent measurements $\phi_{jk}(\mathbf{R})$ and discretize the equations (6) on a square grid of the size $N \times N$. A discretization scheme relevant to ray integrals is described, for example, in [2], and we have used a similar discretization approach in [15]. This will result in a system of $K(K-1)N^2/2$ linear equations with respect to $2N^2$ unknowns μ_{lm} and η_{lm} , $l, m = 1, \dots, N$. In practice, we can (and should) assume *a priori* that the functions to be reconstructed are zero in the pixels adjacent to the strip boundaries; this will reduce somewhat the number of unknowns. This system of equations can be viewed as a problem of optimization, which can be solved by a variety of methods, including computation of the pseudo-inverse, TV-regularization or iterative optimization with nonlinear constraints. All these methods have considerable advantages and should be investigated in their own right. However, an important limiting factor is the computational complexity. Assume that we are aiming at reconstructing a megapixel image, i.e., a rasterized $N \times N$ image with $N \sim 10^3$. The computational complexity of the methods just described is prohibitively high in this case even for reconstructing a single slice of the medium. On the other hand, the methods based on constructing and inverting the star transform are characterized by a much smaller computational complexity because some of the steps necessary for image reconstruction are performed analytically rather than numerically.

Let us estimate the computational complexity of solving the image-reconstruction problem just considered. We assume that the number of data points is of the same order as the number of unknowns and will disregard factors of the order of unity. A direct method (matrix inversion or computing the regularized pseudo-inverse) requires $(N^2)^3 = N^6 \sim 10^{18}$ floating-point operations. This is clearly out of reach, even with the use of supercomputers. The only feasible option is in this case to use iterative optimization in which the computational cost of each iteration scales as $(N^2)^2 = N^4 \sim 10^{12}$. On an average modern computer with the peak performance of 10 Gflops, one iteration will cost about 100 s of computational time. This is acceptable as long as only a few iterations are required, but there is little hope that this would be generally the case, especially if optimization with nonlinear constraints is used. On the other hand, the iterative method of section 7.2 requires only $KN^3 \sim K \cdot 10^9$ floating-point operations per slice, where K is of the order of unity. This means that one slice can be

reconstructed in less than a second. The direct method of section 7.3 is even faster. The practical improvement of utilizing the star transform is therefore obvious.

In summary, construction and numerical inversion of the star transform as described in this paper is a computationally efficient approach for image reconstruction applicable to rays or particles undergoing predominantly single scattering. In the case of x-ray imaging, utilization of single-scattered photons is expected to reduce the total radiation dose received by a patient. Applications of the proposed methodology to mesoscopic optical imaging can also be envisioned.

Acknowledgments

This research was supported by the NSF under the grant DMS-1115616 to VAM and the grants DMS-1115574 and DMS-1108969 to JCS. The authors are grateful to Guillaume Bal and Alexander Katsevich for valuable discussions.

References

- [1] Norton S J 1994 *J. Appl. Phys.* **76** 2007–15
- [2] Yuasa T, Akiba M, Kazama M, Hoshino A, Watanabe Y, Hyodo K, Dilmanian F A, Akatsuka T and Itai Y 1997 *IEEE Trans. Nucl. Sci.* **44** 1760–9
- [3] Wang J, Chi Z and Wang Y 1999 *J. Appl. Phys.* **86** 1693–8
- [4] Busono P and Hussein E M A 1999 *Phys. Med. Biol.* **44** 1455–77
- [5] Cong W and Wang G 2011 *J. X-ray Sci. Technol.* **19** 219–27
- [6] Aviles J E A, Pistorius S, Gordon R and Elbakiri I A 2011 *J. X-ray Sci. Technol.* **19** 35–56
- [7] van Rossum M C W and Nieuwenhuizen T M 1999 *Rev. Mod. Phys.* **71** 313–71
- [8] Driol C, Nguyen M K and Truong T T 2008 *Simul. Model. Practice Theor.* **16** 1067–76
- [9] Nguyen M K and Truong T T 2010 *Inverse Problems* **26** 065005
- [10] Agranovsky M L and Quinto E T 1996 *J. Funct. Anal.* **139** 383–414
- [11] Maass P 1989 *Inverse Problems* **5** 849–57
- [12] Lissiano S and Ponomarev I 1997 *Inverse Problems* **13** 1053–62
- [13] Maeland E 1998 *Geophysics* **63** 1708–15
- [14] Ilmavirta J 2013 *Inverse Problems* **29** 035008
- [15] Florescu L, Schotland J C and Markel V A 2009 *Phys. Rev. E* **79** 036607
- [16] Florescu L, Schotland J C and Markel V A 2010 *Phys. Rev. E* **81** 016602
- [17] Florescu L, Markel V A and Schotland J C 2011 *Inverse Problems* **27** 025002
- [18] Morvidone M, Nguyen M K, Truong T T and Zaidi H 2010 *Int. J. Biomed. Imag.* **2010** 208179
- [19] Ambartsoumian G 2012 *Comput. Math. Appl.* **64** 260–5
- [20] Katsevich A and Krylov R 2013 *Inverse Problems* **29** 075008
- [21] Gouia-Zarrad R and Ambartsoumian G 2014 *Inverse Problems* **30** 045007
- [22] Haltmeier M 2014 *Inverse Problems* **30** 035001