

Nr. 29
07. December 2016

Leopold-Franzens-Universität Innsbruck



Preprint-Series: Department of Mathematics - Applied Mathematics

Iterative Methods for Photoacoustic Tomography with Variable Sound Speed

Markus Haltmeier, Linh V. Nguyen



Technikerstraße 13 - 6020 Innsbruck - Austria
Tel.: +43 512 507 53803 Fax: +43 512 507 53898
<https://applied-math.uibk.ac.at>

Iterative Methods for Photoacoustic Tomography with Variable Sound Speed

Markus Haltmeier[†] Linh V. Nguyen[‡]

[†]Department of Mathematics, University of Innsbruck
Technikerstrasse 13, A-6020 Innsbruck, Austria
`Markus.Haltmeier@uibk.ac.at`

Abstract

In this article, we revisit iterative methods for solving the inverse problem of photoacoustic tomography in free space. Recently, there have been interesting developments on explicit formulations of the adjoint operator, demonstrating that iterative methods could be an attractive choice for photoacoustic image reconstruction. In this work, we propose several modifications of current formulations of the adjoint operator which help speed up the convergence and yield improved error estimates. We establish a stability analysis and show that, with our choices of the adjoint operator, Landweber's and the CG methods can achieve a linear rate of convergence either in L^2 or H^1 norm under the visibility condition. In addition, we analyze the normal operator from the microlocal analysis point of view. This helps us to have more insight into the convergence speed of the iterative methods as well as choosing proper weights for the mapping spaces. Finally, we present numerical results using various iterative reconstruction methods for trapping as well as non-trapping sound speed. Our results demonstrate that Nesterov's fast gradient and the CG methods converge faster than Landweber's and iterative time reversal methods in the visible as well as the invisible case.

Keywords: Photoacoustic tomography, variable sound speed, iterative regularization, adjoint operator, Landweber's method, Nesterov's method, CG method, visibility condition, invisibility condition, image reconstruction

AMS Subject Classification: 35R30; 92C55; 65F10; 35A18; 74J05.

1 Introduction

Photoacoustic tomography (PAT) is an emerging hybrid method of imaging. It combines the high contrast of optical imaging with the good resolution of ultrasound tomography. The biological object of interest is scanned with a laser light pulse. The photoelastic effect produces an ultrasound pressure propagating in the space. One measures the ultrasonic pressure on an observation surface. The aim of PAT is to recover the initial pressure inside the tissue from the measured data. This quantity contains helpful internal information of the object and is the image to be reconstructed.

The mathematical model for PAT is the acoustic wave equation

$$\begin{cases} c^{-2}(x) p_{tt}(x, t) - \Delta p(x, t) = 0, & (x, t) \in \mathbb{R}^d \times \mathbb{R}_+, \\ p(x, 0) = f(x), \quad p_t(x, 0) = 0, & x \in \mathbb{R}^d, \end{cases} \quad (1.1)$$

[†]Department of Mathematics, University of Idaho, 875 Perimeter Dr, Moscow, ID 83844, US. E-Mail: `lnguyen@uidaho.edu`

where $c: \mathbb{R}^d \rightarrow \mathbb{R}$ is the sound speed and $f: \mathbb{R}^d \rightarrow \mathbb{R}$ the initial pressure. Let us denote by S the observation surface. We will assume that S is a closed subset of $\partial\Omega$ with nonempty interior $\text{Int}(S)$. Here, Ω is an open subset of \mathbb{R}^d that contains the support of f . The mathematical problem of PAT is to invert the map $\mathbf{L}: f \rightarrow g := p|_{S \times (0, T)}$. We will refer to this problem as the inverse problem of PAT.

For later conveniences, let us fix several geometric conventions. We will always assume that the sound speed c is smooth and bounded from below by a positive constant. The space \mathbb{R}^d is considered as a Riemannian manifold with the metric $c^{-2}(x) dx^2$ and Ω is assumed to be strictly convex with respect to this metric. Then, all the geodesic rays originating inside Ω intersect the boundary $\partial\Omega$ at most once. We say that the speed c is nontrapping if all such geodesic rays intersect with $\partial\Omega$; otherwise, the speed c is called trapping. We will denote by $\mathbb{T}^*\Omega$ the cotangent bundle of Ω . It can be identified with $\Omega \times \mathbb{R}^d$. Also, $\mathbb{T}^*\Omega \setminus 0$ is the cotangent bundle of Ω minus the zero section, which can be identified with $\Omega \times (\mathbb{R}^d \setminus \{0\})$. A set $V \subset \mathbb{T}^*\Omega \setminus 0$ is said to be conic if $(x, \xi) \in V$ implies $(x, \alpha\xi) \in V$ for all $\alpha > 0$.

Let us assume that $\text{supp}(f) \subset \Omega_0$, where $\Omega_0 \Subset \Omega$. Then, \mathbf{L} is a well-defined linear bounded operator from $H^{s_1}(\Omega_0)$ to $H^{s_2}(S \times [0, T])$ for all $s_1 \geq s_2$ (see [54]). In this article, we will identify the correct mapping spaces for \mathbf{L} in order to stabilize the inverse problem of PAT and design proper algorithms. For the sake of simplicity, we will assume that \mathbf{L} is injective. The necessary and sufficient condition for this assumption to hold can be found in [54].

An essential feature of an inverse problem is its well-posedness or ill-posedness (see [13] and Section 2). The inversion of a linear operator \mathbf{T} is called well-posed if the ratio $\|\mathbf{T}x\|/\|x\|$ is bounded from below by a positive constant and ill-posed otherwise. The inverse problem of PAT can be either well-posed or ill-posed, as can be seen in the following two scenarios:

- (i) There is a closed subset $S_0 \subset \partial\Omega$ such that $S_0 \subset \text{Int}(S)$ and the following condition holds: for any element $(x, \xi) \in \mathbb{T}^*\Omega_0 \setminus 0$, at least one of the unit speed geodesic rays originating from x at time $t = 0$ along the direction of $\pm\xi$ intersects with S_0 at a time $t < T$. This is the so-called **visibility condition** [38, 61, 29, 47, 54].
- (ii) There is an open conic set $V \subset \mathbb{T}^*\Omega_0 \setminus 0$ such that for all $(x, \xi) \in V$ none of the unit speed geodesic rays originating from x at time $t = 0$ along the direction of $\pm\xi$ intersects with S at a time $t \leq T$. This is called the **invisibility condition**.

The visibility and invisibility conditions are almost, but not exactly, complementary. Under the visibility condition, the inversion of $\mathbf{L}: H^s(\Omega_0) \rightarrow H^s(S \times [0, T])$ is well-posed for any $s \geq 0$ (see [54]). On the other hand, when the invisibility condition holds, the inversion of $\mathbf{L}: H^{s_1}(\Omega_0) \rightarrow H^{s_2}(S \times [0, T])$ is ill-posed for all s_1, s_2 (see [47]). In this article, we solve the inverse problem of PAT for both well-posed and ill-posed settings by iterative (regularization) methods. They include Landweber's, Nesterov's and the conjugate gradient (CG) methods. These iterative methods are theoretically convergent to the exact solution in the absence of noise. However, the convergence speed as well as error estimates with respect to the noise level are different for the ill-posed and the well-posed settings.

There exist several methods to solve the problem of PAT such as explicit inversion formulas [16, 35, 15, 35, 46, 20, 21, 43, 50], series solutions [36, 2], time reversal [16, 29, 28, 54], and quasi-reversibility [9]. Reviews on these methods can be found in [29, 33, 34, 53].* To the best of our knowledge, the series solution, time reversal method, and quasi-reversibility method only apply to

*Other setups of PAT that use integrating detectors have been studied in, e.g., [22, 51, 7, 62]. However, we do not consider these setups in the present article.

the case that S is a closed surface. The inversion formula works for certain closed or flat observation surfaces. The algebraic/iterative methods have been commonly used in the engineering community [12, 30, 41, 60]. It seems that these works do not employ the explicit form of the adjoint operator, which has been described only recently [6, 3]. The goal of this article is to have a systematic study on the iterative methods to see their advantages as well as limitations.

Our approach is most closely related to [6], where Landweber's method was proposed to solve the inverse problem for PAT. However, we make several changes in order to preserve the well-posedness (when it holds) of the problem and speed up the convergence speed. For example, instead of considering \mathbf{L} as a mapping from $H_0^1(\Omega_0)$ to $L^2(S \times [0, T])$, we consider $\mathbf{L}: H_i(\Omega_0) \rightarrow H_i(\chi)$, $i = 0, 1$ (see Section 3). This change not only preserves the well-posedness of the inverse problem under the visibility condition but also makes it simpler to compute the adjoint operator and further speed up the convergence. Additionally, by introducing the weighted norm on $H_i(\chi)$, we supply the flexibility to the iterative method. This change is related to [3], where acoustic measurements on an open set are considered (see also Section 3.2, where we derive theoretical results for this setup).

Let us note that our established Landweber and GC algorithms converge linearly for the partial data problem under the visibility condition. This convergence rate, to the best of our knowledge, has not been obtained by any previous method. Comparable results for a different setup of PAT, where the acoustic wave is contained in a bounded domain, have recently been obtained in [1, 49, 56].

The article is organized as follows. In Section 2, we introduce several iterative methods that will be used for solving the inverse problem of PAT. We will see in that section that the knowledge of the adjoint operator is crucial for the iterative methods. In Section 3 we derive and analyze the adjoint operator \mathbf{L}^* of \mathbf{L} . We will revisit the PAT with open observation domain in Subsection 3.2. In Section 4, we describe our numerical implementations and present results in various scenarios including full and partial data as well as non-trapping and trapping sound speed.

2 Iterative methods for solving linear equations in Hilbert spaces

In this section, we review several main iterative methods for solving linear equations in Hilbert spaces. Let $\mathbf{T}: X \rightarrow Y$ be a linear operator mapping between two Hilbert spaces X and Y . We denote by $\mathcal{R}(\mathbf{T})$ its range. Assuming that \mathbf{T} is injective, we are interested in inverting \mathbf{T} . That is, we want to solve the following problem:

Problem 2.1. *Given $y \in \mathcal{R}(\mathbf{T})$, find the solution $x \in X$ of the equation $\mathbf{T}x = y$.*

Problem 2.1 is said to be well-posed if the inverse $\mathbf{T}^{-1}: \mathcal{R}(\mathbf{T}) \rightarrow X$ is bounded and ill-posed otherwise. It can be seen that Problem 2.1 is well-posed if and only if \mathbf{T} is bounded from below, i.e.,

$$b := \inf_{x \neq 0} \frac{\|\mathbf{T}x\|_Y}{\|x\|_X} = \sqrt{\inf_{x \neq 0} \frac{\langle \mathbf{T}^* \mathbf{T}x, x \rangle_X}{\|x\|_X^2}} > 0.$$

Let us now consider the inverse problem with noise:

Problem 2.2. *Given $y^\delta \in Y$ such that $\|y^\delta - \mathbf{T}x\|_Y \leq \delta$, find an approximation x^δ of x .*

We notice that Problem 2.1 is a special case of Problem 2.2 where $\delta = 0$. To solve Problem 2.2, it is helpful to understand the Moore-Penrose generalized inverse \mathbf{T}^\dagger of \mathbf{T} . It is defined for all $y = y' + y'' \in \mathcal{R}(\mathbf{T}) \oplus \mathcal{R}(\mathbf{T})^\perp =: \mathcal{D}(\mathbf{T}^\dagger)$ by the formula $\mathbf{T}^\dagger y = \mathbf{T}^{-1}y'$, which is the unique minimizer of the residual functional $\frac{1}{2}\|\mathbf{T}x - y\|_Y^2$. A reasonable solution for Problem 2.2 is $x^\delta = \mathbf{T}^\dagger y^\delta$ when

$y^\delta \in \mathcal{D}(\mathbf{T}^\dagger)$. Of course, some care is needed when $y^\delta \notin \mathcal{D}(\mathbf{T}^\dagger)$. We will discuss this issue in detail below.

There are several methods to solve Problems 2.1 and 2.2. In the rest of this section, we discuss three methods: Landweber's, Nesterov's, and the conjugate gradient (CG) methods.

2.1 Landweber's method

We first consider the exact data situation (see Problem 2.1). In this case, Landweber's method is simply the gradient descent method for minimizing the residual functional $\frac{1}{2}\|\mathbf{T}x - y\|^2$. It reads as follows:

$$x_{k+1} = x_k - \gamma \mathbf{T}^*(\mathbf{T}x_k - y), \quad k \geq 0. \quad (2.1)$$

In order for the algorithm to converge, the step size (or relaxation parameter) γ needs to be properly chosen. From (2.1), we obtain

$$x_{k+1} - x = (\mathbf{I} - \gamma \mathbf{T}^* \mathbf{T})(x_k - x). \quad (2.2)$$

Our goal is to choose γ such that $\|x_k - x\|$ decreases with respect to k . To that end, we observe that the spectrum of $\mathbf{I} - \gamma \mathbf{T}^* \mathbf{T}$ is contained in $[1 - \gamma \|\mathbf{T}\|^2, 1 - \gamma b^2]$ and therefore $\|\mathbf{I} - \gamma \mathbf{T}^* \mathbf{T}\| = \max\{|1 - \gamma b^2|, |1 - \gamma \|\mathbf{T}\|^2|\}$. Choosing a fixed step size $0 < \gamma < 2/\|\mathbf{T}\|^2$, we obtain

$$\|\mathbf{I} - \gamma \mathbf{T}^* \mathbf{T}\| \leq 1.$$

The equality, in fact, may still hold if $b = 0$ (i.e., in the case of the ill-posed problem). From equation (2.2) we obtain

$$\|x_{k+1} - x\|_X \leq \|x_k - x\|_X.$$

Keeping in mind that \mathbf{T} is injective, one can prove that the strict inequality holds. Moreover, $\|x_k - x\| \rightarrow 0$ as $k \rightarrow \infty$ (see, e.g., [13, 31]). The convergence speed, however, can be arbitrarily slow for the ill-posed problem. On the other hand, for the well-posed problem $\|\mathbf{I} - \gamma \mathbf{T}^* \mathbf{T}\| < 1$. Therefore, $x_k \rightarrow x$ linearly since $\|x_{k+1} - x\|_X \leq \|\mathbf{I} - \gamma \mathbf{T}^* \mathbf{T}\| \|x_k - x\|_X$. This implies

$$\|x_k - x\|_X \leq \|\mathbf{I} - \gamma \mathbf{T}^* \mathbf{T}\|^k \|x_0 - x\|_X. \quad (2.3)$$

In particular for the choice $\gamma = \frac{2}{b^2 + \|\mathbf{T}\|^2}$, we obtain the optimal convergence estimate

$$\|x_k - x\|_X \leq \left(\frac{\|\mathbf{T}\|^2 - b^2}{\|\mathbf{T}\| + b^2} \right)^k \|x_0 - x\|_X.$$

This convergence however is slow when the condition number $\kappa = \|\mathbf{T}\|/b$ is large.

Now consider the noisy problem (c.f., Problem 2.2). We then repeat the iteration (2.1) where x_k and y are replaced by x_k^δ and y^δ , respectively. If $y^\delta \in \mathcal{D}(\mathbf{T}^\dagger)$ and $x^\delta = \mathbf{T}^\dagger y^\delta$, then $x_k^\delta \rightarrow x^\delta$. However, if $y^\delta \notin \mathcal{D}(\mathbf{T}^\dagger)$, then $x_k \rightarrow \infty$ (see, e.g., [13]). One needs to propose a stopping criteria, for which we follow Morozov [42]. Let $\tau > 1$ be a fixed constant and k_δ is the smallest integer such that

$$\|\mathbf{T}x_{k_\delta}^\delta - y^\delta\|_Y < \tau\delta.$$

We then define $x_*^\delta := x_{k_\delta}^\delta$. The stopping index k_δ exists and satisfies $k_\delta = O(\delta^{-2})$. Moreover, $x_*^\delta \rightarrow x$ as $\delta \rightarrow 0$ (see [13, Proposition 6.4]). In the well posed case, Morozov's discrepancy principle obviously yields the optimal error estimate $\|x_*^\delta - x\|_X = O(\delta)$ for all $x \in X$. In the ill-posed case, however, this does not hold and error estimates can only be derived under appropriate smoothness assumptions on the unknown x .

2.2 Nesterov's fast gradient method

Let us now discuss the Nesterov's fast gradient method (see [44]), which is an acceleration of the gradient descent method. We state the algorithm here for minimizing the residual functional $\frac{1}{2}\|\mathbf{T}x - y\|^2$ in the exact data case (see Problem 2.1):

1. initialization: $x_0 = z_0$ and $t_0 = 1$.
2. while (not stop) do

$$\begin{aligned} x_{k+1} &= z_k - \gamma \mathbf{T}^*(\mathbf{T}z_k - y) \\ t_{k+1} &= \frac{1 + \sqrt{1 + 4t_k^2}}{2} \\ z_{k+1} &= x_{k+1} + \frac{t_k - 1}{t_{k+1}}(x_{k+1} - x_k) \end{aligned}$$

end do

Nesterov's method is known to satisfy $\|\mathbf{T}x_k - y\|^2 = \mathcal{O}(k^{-2})$. The algorithm for Problem 2.2 is similar. We also apply the Morozov's stopping criteria for this algorithm. Then, the stopping index satisfies $k_\delta = \mathcal{O}(\delta^{-1})$. It is not clear to see whether the Nesterov's method converges (at least) linearly for the well-posed case of Problem 2.1. However, as demonstrated in later in our numerical simulations, Nesterov's method converges much faster than Landweber's method for both Problem 2.1 and Problem 2.2.

2.3 The conjugate gradient (CG) method

The CG algorithm was originally proposed in [26] for solving linear equations $\mathbf{A}x = y$, where $\mathbf{A}: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is positive definite self adjoint operator. The basic idea of CG method is to find a decomposition of the solution x into the form

$$x = x_0 + \sum_k \alpha_k d_k,$$

where $(d_k)_k$ is a family of directions that are pairwise orthogonal with respect to the energy product $\langle d, d' \rangle_{\mathbf{A}} := \langle d, \mathbf{A}d' \rangle$. The CG algorithm has been generalized to the Hilbert space setting in [24, 11, 32, 23, 17, 45, 14, 4, 19, 39]. In this article, we will consider an operator \mathbf{T} which is not self-adjoint (indeed, it maps between different spaces X and Y , which will be specified in Section 3). In such a situation we can use the CG method to solve the normal equation

$$\mathbf{T}^* \mathbf{T}x = \mathbf{T}^* y.$$

This resulting method is called conjugate gradient method for the normal equation (CGNE). It reads as follows (see, e.g., [23, Algorithm 2.3]):

1. initialization: $k = 0$, $r_0 = y - \mathbf{T}x_0$, $d_0 = \mathbf{T}^* r_0$.
2. while (not stop) do

- i) $\alpha_k = \|\mathbf{T}^* r_k\|_X^2 / \|\mathbf{T}d_k\|_Y^2$
- ii) $x_{k+1} = x_k + \alpha_k d_k$
- iii) $r_{k+1} = r_k - \alpha_k \mathbf{T}d_k$

- iv) $\beta_k = \|\mathbf{T}^* r_{k+1}\|_X^2 / \|\mathbf{T}^* r_k\|_X^2$
- v) $d_{k+1} = \mathbf{T}^* r_{k+1} + \beta_k d_k$
- vi) $k = k + 1$

end while

In the following we refer to the CGNE algorithm simply as the CG method. The iterates x_k converges to the solution x of Problem 2.1. When the problem is well-posed, the CG method converges linearly in the energy norm (see, e.g., [11]):

$$\|\mathbf{T}(x_k - x)\|_Y \leq 2 \left(\frac{\|\mathbf{T}\| - b}{\|\mathbf{T}\| + b} \right)^k \|\mathbf{T}(x_0 - x)\|_Y,$$

and hence

$$\|x_k - x\|_X \leq 2 \frac{\|\mathbf{T}\|}{b} \left(\frac{\|\mathbf{T}\| - b}{\|\mathbf{T}\| + b} \right)^k \|x_0 - x\|_X.$$

This convergence rate for the CG method is better than for Landweber's method stated in (2.3). We also notice that for CG, we do not have to determine a step size parameter. Moreover, assume that $\mathbf{T}^* \mathbf{T} = \alpha \mathbf{I} + \mathbf{K}$ where \mathbf{K} is a compact operator. Then, the CG method converges superlinearly (see, e.g., [25]).

For the case of noisy data (see Problem 2.2), we use the same iteration in combination with Morozov's stopping criteria. Then, the stopping index exists and, in practice, it is much smaller than that of Landweber's method.

2.4 Iterative methods for PAT

In this article, we will study the above three iterative methods for the inverse problem of PAT. To that end, we need to establish the proper form of the adjoint operator (or equivalently, the mapping spaces for \mathbf{L}), which is done in Section 3. Our goal for the adjoint operator is two-fold: it should be relatively simple to implement and to speed up the convergence. In particular, with our choice of mapping spaces for \mathbf{L} , the inverse problem of PAT is well-posed under the visibility condition. Therefore, the linear convergence for the Landweber's and CG method is guaranteed for the exact problem. This convergence rate, for partial data problem of PAT, has not been obtained before for any other methods.

In Section 4, we will implement the iterative methods for PAT. We will also compare them with the iterative time reversal method proposed in [54] (see also [52] for the numerical discussion of the method).

3 The adjoint operator for PAT

Let us recall that $\mathbf{L}: f \rightarrow g := p|_{S \times (0, T)}$, where p is defined by the acoustic wave equation (1.1) and S is a closed subset of $\partial\Omega$. Our goal is to invert \mathbf{L} using the iterative methods introduced in the previous section. It is crucial to analyze the adjoint operator \mathbf{L}^* of \mathbf{L} . To that end, we first need to identify the correct mapping spaces for \mathbf{L} . We, indeed, will consider two realizations, \mathbf{L}_0 and \mathbf{L}_1 , of \mathbf{L} corresponding to two different choices of the mapping spaces.

We first recall our assumption $\text{supp}(f) \subset \Omega_0$, where $\Omega_0 \Subset \Omega$. For the spaces of f , let us denote

$$\begin{aligned} H_0(\Omega_0) &:= \{f \in L^2(\mathbb{R}^d) : \text{supp}(f) \subset \overline{\Omega_0}\}, \\ H_1(\Omega_0) &:= \{f \in H^1(\mathbb{R}^d) : \text{supp}(f) \subset \overline{\Omega_0}\}. \end{aligned}$$

Then, $H_0(\Omega_0)$ and $H_1(\Omega_0)$ are Hilbert spaces with the respective norms

$$\begin{aligned}\|f\|_{H_0(\Omega_0)} &= \|c^{-1}f\|_{L^2(\Omega_0)}, \\ \|f\|_{H_1(\Omega_0)} &= \|\nabla f\|_{L^2(\Omega_0)}.\end{aligned}$$

We note that $H_0(\Omega_0) \cong L^2(\Omega_0)$ and $H_1(\Omega_0) \cong H_0^1(\Omega_0)$. The above chosen norms are convenient for our later purposes.

For the spaces of g , we fix a nonnegative function $\chi \in L^\infty(\partial\Omega \times [0, T])$ such that $\text{supp}(\chi) = \Gamma := S \times [0, T]$. Let us denote:

$$\begin{aligned}H_0(\chi) &= \{g: \|g\|_{H_0(\chi)} := \|\sqrt{\chi}g\|_{L^2(\Gamma)} < \infty\}, \\ H_1(\chi) &= \{g: g(\cdot, 0) \equiv 0, \|g\|_{H_1(\chi)} := \|g_t\|_{H_0(\chi)} < \infty\}.\end{aligned}$$

We define

$$\mathbf{L}_i = \mathbf{L}|_{H_i(\Omega_0)}: (H_i(\Omega_0), \|\cdot\|_{H_i(\Omega_0)}) \rightarrow (H_i(\chi), \|\cdot\|_{H_i(\chi)}) \quad \text{for } i = 0, 1.$$

Let $H^i(\Gamma)$ be the standard Sobolev space of order i on Γ . Noticing that \mathbf{L} is a bounded map from $H_i(\Omega_0) \rightarrow H^i(\Gamma)$ (see [54]) and $H^i(\Gamma) \subset H_i(\chi)$, we obtain:

Theorem 3.1. *For $i = 0, 1$, \mathbf{L}_i is a bounded map from $H_i(\Omega_0)$ to $H_i(\chi)$.*

We now consider χg as a function on $\partial\Omega \times [0, T]$ which vanishes on $(\partial\Omega \setminus S) \times [0, T]$. The following theorem gives us an explicit formulation of \mathbf{L}_i^* :

Theorem 3.2. *Let $g \in C^\infty(\Gamma)$.*

(a) *Consider the wave equation*

$$\begin{cases} c^{-2}(x)q_{tt}(x, t) - \Delta q(x, t) = 0, & (x, t) \in (\mathbb{R}^d \setminus \partial\Omega) \times (0, T), \\ q(x, T) = 0, \quad q_t(x, T) = 0, & x \in \mathbb{R}^d, \\ [q](y, t) = 0, \quad \left[\frac{\partial q}{\partial \nu}\right](y, t) = \chi(y, t)g(y, t), & (y, t) \in \partial\Omega \times [0, T]. \end{cases} \quad (3.1)$$

Then

$$\mathbf{L}_0^*g = q_t(\cdot, 0)|_{\Omega_0}.$$

(b) *Assume further that χ is independent of t (i.e., $\chi(y, t) = \chi(y)$). We define*

$$\bar{g}(x, t) = g(x, t) - g(x, T),$$

and consider the wave equation

$$\begin{cases} c^{-2}(x)\bar{q}_{tt}(x, t) - \Delta \bar{q}(x, t) = 0, & (x, t) \in (\mathbb{R}^d \setminus \partial\Omega) \times (0, T), \\ \bar{q}(x, T) = 0, \quad \bar{q}_t(x, T) = 0, & x \in \mathbb{R}^d, \\ [\bar{q}](y, t) = 0, \quad \left[\frac{\partial \bar{q}}{\partial \nu}\right](y, t) = \chi(y, t)\bar{g}(y, t), & (y, t) \in \partial\Omega \times [0, T]. \end{cases} \quad (3.2)$$

Then,

$$\mathbf{L}_1^*g = \Pi[\bar{q}_t(\cdot, 0)].$$

Here, Π is the projection on the space $H_1(\Omega_0) \cong H_0^1(\Omega_0)$, given by

$$\Pi(f) = f - \phi,$$

where ϕ is the harmonic extension of $f|_{\partial\Omega_0}$ to $\bar{\Omega}_0$.

We note that the solution $q(x, t)$ in (3.1) is defined by the variational formulation (see [6])

$$\int_0^T \int_{\mathbb{R}^d} c^{-2}(x) q_{tt}(x, t) v(x, t) dx dt + \int_0^T \int_{\mathbb{R}^d} \nabla q(x, t) \nabla v(x, t) dx dt = - \int_0^T \int_{\partial\Omega} \chi(y, t) g(y, t) v(y, t) dy dt, \quad (3.3)$$

for any test function $v \in C^\infty(\mathbb{R}^d \times [0, T])$. Its existence and uniqueness can be found in [6]. The above variational form implies that (3.1) can be formally rewritten as the nonhomogeneous wave problem

$$\begin{cases} c^{-2}(x) q_{tt}(x, t) - \Delta q(x, t) = -\delta_{\partial\Omega}(x) \chi(x, t) g(x, t), & (x, t) \in \mathbb{R}^d \times (0, T), \\ q(x, T) = 0, \quad q_t(x, T) = 0, & x \in \mathbb{R}^d. \end{cases}$$

This formulation will be helpful later when we solve the problem numerically in Section 4.

Let us now proceed to prove Theorem 3.2.

Proof. Let $f \in C_0^\infty(\Omega_0)$ and p be the solution of (1.1). Choosing $v = p$ in (3.3) and taking integration by parts for the left hand side, we obtain

$$\begin{aligned} - \int_{\Omega_0} c^{-2}(x) q_t(x, 0) p(x, 0) dx + \int_0^T \int_{\mathbb{R}^d} q(x, t) [c^{-2}(x) p_{tt}(x, t) - \Delta p(x, t)] dx \\ = - \int_0^T \int_{\partial\Omega} \chi(y, t) g(y, t) \mathbf{L}(f)(y, t) dy dt. \end{aligned}$$

That is,

$$\int_{\Omega_0} c^{-2}(x) q_t(x, 0) f(x) dx = \int_0^T \int_{\partial\Omega} \chi(y, t) g(y, t) \mathbf{L}(f)(y, t) dy dt, \quad (3.4)$$

or

$$\langle q_t(\cdot, 0), f \rangle_{H_0(\Omega_0)} = \langle g, \mathbf{L}(f) \rangle_{H_0(\chi)}.$$

Since this is true for all $f \in C_0^\infty(\Omega_0)$, we obtain

$$\mathbf{L}_0^* g = q_t(\cdot, 0)|_{\Omega_0}.$$

This finishes the proof of (a). Let us now proceed to prove (b). Let \bar{q} be the solution of (3.2). That is, \bar{q} satisfies (3.3) with g being replaced by \bar{g} . In that identity, we pick $v = p_{tt}$. Then, taking integration by parts for the left hand side, we obtain

$$\int_{\Omega_0} c^{-2}(x) \bar{q}_t(x, 0) p_{tt}(x, 0) dx = \int_0^T \int_{\partial\Omega} \chi(y, t) \bar{g}(y, t) \partial_t^2 \mathbf{L}(f)(y, t) dy dt.$$

Noting that $c^{-2}(x) p_{tt}(x, 0) = \Delta f(x)$, we arrive to

$$\int_{\Omega_0} \bar{q}_t(x, 0) \Delta f(x) dx = \int_0^T \int_{\partial\Omega} \chi(y, t) \bar{g}(y, t) \partial_t^2 \mathbf{L}(f)(y, t) dy dt.$$

Let us now consider $\chi(y, t) = \chi(y)$. Taking integration by parts for the left hand side with respect to x and right hand side with respect to t , we get

$$\int_{\Omega_0} \nabla[\bar{q}_t(x, 0)] \nabla f(x) dx = \int_0^T \int_{\partial\Omega} \chi(y) \bar{g}_t(y, t) \partial_t \mathbf{L}(f)(y, t) dy dt.$$

Here, we have used $\bar{g}(\cdot, T) = \partial_t \mathbf{L}(f)(\cdot, 0) \equiv 0$. Let ϕ be the harmonic extension to $\bar{\Omega}_0$ of $q_t(\cdot, 0)|_{\partial\Omega_0}$. Since

$$\int_{\Omega_0} \nabla \phi(x) \nabla f(x) dx = - \int_{\Omega} \Delta \phi(x) f(x) dx = 0,$$

and $\bar{g}_t = g_t$, we obtain

$$\int_{\Omega_0} \nabla [\bar{q}_t(x, 0) - \phi(x)] \nabla f(x) dx = \int_0^T \int_{\partial\Omega} \chi(y) g_t(y, t) \partial_t \mathbf{L}(f)(y, t) dy dt. \quad (3.5)$$

We conclude

$$\langle \Pi[\bar{q}_t(\cdot, 0)], f \rangle_{H_1(\Omega_0)} = \langle g, \mathbf{L}f \rangle_{H_1(\chi)}.$$

This proves $\mathbf{L}_1^* g = \Pi[\bar{q}_t(\cdot, 0)]$. □

Remark 3.3. *Let us make the following observations:*

- (a) *Since $C^\infty(\Gamma)$ is dense in both $H_0(\chi)$ and $H_1(\chi)$, the adjoint operators \mathbf{L}_0^* and \mathbf{L}_1^* are uniquely determined from the formulas in Theorem 3.2.*
- (b) *Compared to \mathbf{L}_0^* , \mathbf{L}_1^* involves an extra projection operator. In our numerical experiments, we will only use \mathbf{L}_0^* since it is simpler to implement. However, the knowledge of \mathbf{L}_1^* is helpful in designing iterative algorithms that converge in the $H^1(\Omega_0)$ norm.*
- (c) *If, instead of $\mathbf{L} = \mathbf{L}_0 : H_0(\Omega_0) \rightarrow H_0(\chi)$, we consider $\mathbf{L} : L^2(\Omega_0) \rightarrow H_0(\chi)$, then $\mathbf{L}^* g = \frac{1}{2} q_t(\cdot, 0)$. Our formulations of \mathbf{L}^* are different from that in [6], which is $\mathbf{L}^* g = -\Delta^{-1}(\frac{1}{2} q_t(\cdot, 0))$. Our formulations make the inverse problem of PAT well-posed under the visibility condition (see Theorem 3.4 below).*

Let us recall the visibility condition described in the introduction:

Visibility condition. *There is a closed subset $S_0 \subset \partial\Omega$ such that $S_0 \subset \text{Int}(S)$ and the following condition holds: for any element $(x, \xi) \in \mathbb{T}^* \Omega_0 \setminus 0$, at least one of the unit speed geodesic rays originating from x at time $t = 0$ along the directions $\pm \xi$ intersects with S_0 at a time $t < T$.*

Let us prove that with our choices of mapping spaces, the inverse problem of PAT is well-posed.

Theorem 3.4. *Assume that the visible condition holds and $\chi \equiv 1$ on $S_1 \times [0, T]$, where S_1 is a closed subset of $\partial\Omega$ such that $S_0 \subset \text{Int}(S_1)$ and $S_1 \subset \text{Int}(S)$. For $i = 0, 1$, there is a constant $C > 0$ such that for any $g = \mathbf{L}f$, we have*

$$\|f\|_{H_i(\Omega_0)} \leq C \|g\|_{H_i(\chi)}.$$

Proof. Following the lines of [54], we obtain for $i = 0, 1$ [†]

$$\|f\|_{H_i(\Omega_0)} \leq C \|g\|_{H^i(S_1 \times [0, T])}. \quad (3.6)$$

Here and elsewhere, $H^i(S_1 \times [0, T])$ is the standard Sobolev space of order i on $S_1 \times [0, T]$ and C is a generic constant which may be different in one place from another.

[†]The result for $i = 1$ is obtained in that reference. The result for $i = 0$ is obtained similarly, one only needs to invoke [37, Theorem 2.3] instead of [37, Theorem 2.1].

Let us consider $i = 0$. Noticing that $\|g\|_{H^0(S_1 \times [0, T])} \leq \|g\|_{H_0(\chi)}$, we obtain

$$\|f\|_{H_0(\Omega_0)} \leq C \|g\|_{H_0(\chi)}.$$

Let us now consider $i = 1$ and compare $\|g\|_{H_1(\chi)}$ with $\|g\|_{H^1(S_1 \times [0, T])}$. To this end, we recall the hyperbolic zone in $\mathbb{T}^*(\partial\Omega \times (0, T))$

$$\mathcal{H} = \{(x, t, \eta, \tau) \in \mathbb{T}^*(\partial\Omega \times (0, T)) : c(x) |\eta| < \tau\}.$$

It is well-known (see, e.g., [54, 47]) that $\text{WF}(g) \subset \mathcal{H}$. Since ∂_t is an elliptic operator in \mathcal{H} ,

$$\|g\|_{H^1(S_1 \times [0, T])} \leq C(\|\partial_t g\|_{L^2(S_1 \times [0, T])} + \|g\|_{L^2(S_1 \times [0, T])}).$$

Keeping in mind that $g(\cdot, 0) \equiv 0$, we obtain

$$\|g\|_{H^1(S_1 \times [0, T])} \leq C \|\partial_t g\|_{L^2(S_1 \times [0, T])}.$$

Hence,

$$\|g\|_{H^1(S_1 \times [0, T])} \leq C \|g\|_{H_1(\chi)}.$$

This, together with (3.6), gives

$$\|f\|_{H_1(\Omega_0)} \leq C \|g\|_{H_1(\chi)},$$

which finishes our proof. \square

Remark 3.5. *Let us recall that (see Section 2) when the linear inverse problem is well-posed, Landweber's and the CG methods have a linear rate of convergence. Theorem 3.4 shows that with our choices of mapping spaces, the inverse problem of PAT is well-posed under the visibility condition. Therefore, Landweber's and the CG methods converge linearly in either L^2 -norm or H^1 -norm, depending on our choice of the adjoint operator in Theorem 3.2, if the visibility condition holds. This convergence rate has not been obtained before by any method.*

3.1 Microlocal analysis for the normal operator $\mathbf{L}^* \mathbf{L}$

To better understand the nature of $\mathbf{L}^* \mathbf{L}$, we will analyze it from the microlocal analysis point of view. Let us recall that $r_{\pm}(x, \xi)$ is the (unit speed) geodesic rays originating from x along direction of $\pm\xi$. We assume that $r_{\pm}(x, \xi)$ intersects the boundary $\partial\Omega$ at a unique point $x_{\pm} = x_{\pm}(x, \xi)$. We denote by θ_{\pm} the angle between $r_{\pm}(x, \xi)$ and the normal vector of $\partial\Omega$ at x_{\pm} . Our main result is the following theorem.

Theorem 3.6. *Assume that $\chi \in C^{\infty}(\partial\Omega \times [0, T])$. Then, the normal operator $\mathbf{N} = \mathbf{L}^* \mathbf{L}$ is a pseudo-differential operator of order zero, whose principal symbol is*

$$\sigma_0(x, \xi) = \frac{1}{4} \left(\frac{c(x_+) \chi(x_+, t_+)}{\cos(\theta_+)} + \frac{c(x_-) \chi(x_-, t_-)}{\cos(\theta_-)} \right). \quad (3.7)$$

Here, t_{\pm} is the geodesics distance between x and x_{\pm} .

Let us note that in the case of trapping speed, it may happen that one (or both) of the geodesic rays $r_{\pm}(x, \xi)$ does not intersect $\partial\Omega$. In that case, Theorem 3.6 still holds if we replace $\chi(x_{\pm}, t_{\pm})$ by 0. The above formula is true for both realizations \mathbf{L}_0 and \mathbf{L}_1 of \mathbf{L} since \mathbf{L}_0^* and \mathbf{L}_1^* only differ by a smooth term. In the below proof, we will consider $\mathbf{L} = \mathbf{L}_0$.

Proof. We first intuitively describe the effect of $\mathbf{L}^*\mathbf{L}$ on the wave front set of a function f supported inside $\overline{\Omega}_0$. For simplicity, we assume that f is microlocally supported near an element $(x, \xi) \in \mathbb{T}^*\Omega \setminus 0$. Let us analyze the effect of \mathbf{L} to f by considering the wave equation (1.1). At time $t = 0$, the singularity of f at (x, ξ) breaks into two equal parts (see [54]). They induce the singularities of p on the bicharacteristic rays $\mathcal{C}_\pm(x, \xi)$ originating at $(x, 0, \xi, \tau = c(x)|\xi|)$ and $(x, 0, -\xi, \tau = c(x)|\xi|)$ (see [27]). The projection of each bicharacteristic ray $\mathcal{C}_\pm(x, \xi)$ on the spatial domain \mathbb{R}^d is the geodesic ray $r_\pm(x, \xi)$ on \mathbb{R}^d (recalling that \mathbb{R}^d is equipped with the metric $c^{-2}(x) dx^2$). Each of the geodesic ray hits the boundary $\partial\Omega$ at a unique point x_\pm and time t_\pm . The corresponding singularity of p at (x_\pm, t_\pm) is denoted by $(x_\pm, t_\pm, \xi_\pm, \tau_\pm)$. Its projection on $\mathbb{T}^*_{(x_\pm, t_\pm)}(\partial\Omega \times [0, T])$ induces a singularity of g at $(x_\pm, t_\pm, \eta_\pm, \tau_\pm)$. Now, consider the adjoint equation (3.1) which defines \mathbf{L}^* . The singularity of g at $(x_\pm, t_\pm, \eta_\pm, \tau_\pm)$ then induces two singularities of q at $(x_\pm, t_\pm, \xi', \tau_\pm)$. Here, $\xi' = \eta_\pm \pm \sqrt{c^{-2}(x_\pm)\tau^2 - |\eta_\pm|^2} \nu$ where ν is the normal vector of $\partial\Omega$ at x_\pm (note that one of the such ξ' equals ξ_\pm). These two singularities propagate along two opposite directions when going backward in time, one into the domain Ω (along the ray $\mathcal{C}_\pm(x, \xi)$ but in the negative direction) and one away from Ω . At $t = 0$ the first one lands back to (x, ξ) and the other one lands outside of Ω . This shows the pseudo-locality of $\mathbf{L}^*\mathbf{L}: f \rightarrow q_t(\cdot, 0)|_{\Omega_0}$ and heuristically explains that $\mathbf{L}^*\mathbf{L}$ is a pseudo-differential operator. Our rigorous argument follows below.

Let us recall that up to a smooth term (e.g., [59]):

$$p(x, t) = \frac{1}{(2\pi)^d} \sum_{\sigma=\pm} \int e^{i\phi_\sigma(x, t, \xi)} a_\sigma(x, t, \xi) \hat{f}(\xi) d\xi = p_+(x, t) + p_-(x, t).$$

The phase function ϕ_σ satisfy the eikonal equation

$$\partial_t \phi_\sigma(x, t, \xi) + \sigma |\nabla_x \phi_\sigma(x, t, \xi)| = 0, \quad \phi_\sigma(x, 0, \xi) = x \cdot \xi.$$

The amplitude function a_σ satisfies

$$a(x, t, \xi) \sim \sum_{m=0}^{\infty} a_{-m}(x, t, \xi),$$

where a_{-m} is homogeneous of order $-m$ in ξ . The leading term $a_0 = a(x, t, \xi)$ satisfies the transport equation[‡]

$$(\partial_t \phi_\sigma \partial_t - c^2(x) \nabla_x \phi_\sigma \cdot \nabla_x) a_0(x, t, \xi) = 0,$$

with the initial condition $a_0(x, \xi, 0) = 1/2$.

Then, up to a smooth term, we obtain $g = (p_+ + p_-)|_{\partial\Omega} = g_+ + g_-$. Solving the adjoint problem (3.1), we obtain, up to a smooth term, $\mathbf{L}^*\mathbf{L}f = \partial_t q_+(\cdot, 0) + \partial_t q_-(\cdot, 0)$. Here, q_σ (for $\sigma = \pm$) is defined by

$$\begin{cases} c^{-2}(x) q_{\sigma, tt}(x, t) - \Delta q_\sigma(x, t) = 0, & (x, t) \in (\mathbb{R}^d \setminus \partial\Omega) \times (0, T), \\ q_\sigma(x, T) = 0, \quad q_{\sigma, t}(x, T) = 0, & x \in \mathbb{R}^d, \\ [q_\sigma](y, t) = 0, \quad \left[\frac{\partial q_\sigma}{\partial \nu} \right](y, t) = \chi(y, t) g_\sigma(y, t), & (y, t) \in \partial\Omega \times (0, T). \end{cases} \quad (3.8)$$

Let us show that $f \rightarrow f_+ := \partial_t q_+(\cdot, 0)$ is a pseudo-differential operator with the principal symbol

$$\sigma_+(x, \xi) = \frac{1}{4} \frac{\chi(x_+, t_+)}{\cos(\theta_+)}.$$

[‡]In several references, the equation contains a zero order term. However, that term turns out to be zero.

We recall that x_+ is the intersection of the positive geodesic ray $r_+(x, \xi)$ and $\partial\Omega$, and t_+ is the time to travel along the geodesic from x to x_+ . Let $(x_+, t_+, \xi_+, \tau_+)$ be the corresponding element on the bicharacteristic and $(x_+, t_+, \eta_+, \tau_+)$ its projection on $\mathbb{T}_{(x_+, t_+)}^*(\partial\Omega \times [0, T])$. Then, $(x_+, t_+, \eta_+, \tau_+)$ is in the hyperbolic zone, that is $\tau_+ > c(x_+) |\eta_+|$. Let us show that the mapping $q_+|_{\partial\Omega \times [0, T]} \rightarrow [\partial_\nu q_+]|_{\partial\Omega \times [0, T]}$ is an elliptic pseudo-differential operator near $(x_+, t_+, \eta_+, \tau_+)$.

Indeed, for simplicity, we assume that locally near x_+ , $\partial\Omega$ is flat, and $h = q_+|_{\partial\Omega \times [0, T]}$ is supported near (x_+, t_+) . We then can write $y = (y', 0)$ for all $y \in \partial\Omega$ and assume that $\Omega \subset \{x \in \mathbb{R}^d : x_n < 0\}$. The parametrix q_{in} (respectively q_{out}) for q_+ in Ω (respectively Ω^c) near (x_+, t_+) is of the form

$$q_{\text{in/out}}(x, t) = \frac{1}{(2\pi)^d} \sum_{s=F, B} \int_{\mathbb{R}} \int_{\mathbb{R}^{d-1}} e^{i\psi_s(x, t, \eta, \tau)} d_s(x, t, \eta, \tau) \hat{h}(\eta, \tau) d\eta d\tau. \quad (3.9)$$

Here,

$$d_F(y, t, \eta, \tau) + d_B(y, t, \eta, \tau) = 1, \quad \psi_s(y, t, \eta, \tau) = y' \cdot \eta - t\tau, \quad y \in \partial\Omega,$$

and

$$\hat{h}(\eta, \tau) = \int_{\mathbb{R}} \int_{\mathbb{R}^{d-1}} h(y', 0, t) e^{i(-\eta \cdot y' + t\tau)} dy' dt.$$

Similarly to ϕ_+ , the phase function $\psi = \psi_{F, B}$ satisfies the eikonal equation

$$|\partial_t \psi(x, t, \eta, \tau)| = c(x) |\nabla_x \psi(x, t, \eta, \tau)|.$$

In particular, we obtain

$$\partial_{x_n} \psi_F(y, t, \eta, \tau) = \sqrt{c^{-2}(y) \tau^2 - \eta^2}, \quad \partial_{x_n} \psi_B(y, t, \eta, \tau) = -\sqrt{c^{-2}(y) \tau^2 - \eta^2}, \quad y \in \partial\Omega.$$

Roughly speaking, the phase function ψ_F transmits (forward) wave from left to right (along the x_n direction) and ψ_B transmits (backward) wave to the opposite direction. Since $q(x, T) = q_t(x, T) = 0$ for all $x \in \mathbb{R}^d$, we obtain that there is no backward wave inside Ω and no forward wave outside Ω . That is,

$$q_{\text{in}}(x, t) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}} \int_{\mathbb{R}^{d-1}} e^{i\psi_F(x, t, \xi)} d_{\text{in}}(x, t, \eta, \tau) \hat{h}(\eta, \tau) d\eta d\tau,$$

and

$$q_{\text{out}}(x, t) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}} \int_{\mathbb{R}^{d-1}} e^{i\psi_B(x, t, \xi)} d_{\text{out}}(x, t, \eta, \tau) \hat{h}(\eta, \tau) d\eta d\tau.$$

Moreover,

$$d_{\text{in}}(y, t, \eta, \tau) = d_{\text{out}}(y, t, \eta, \tau) = 1, \quad y \in \partial\Omega.$$

Up to lower order terms, we obtain microlocally near the hyperbolic element $(x_+, t_+, \eta_+, \tau_+)$

$$[\partial_{x_n} q_+](y, t) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}} \int_{\mathbb{R}^{d-1}} e^{i(y' \cdot \eta - t\tau)} (-2i) \sqrt{c^{-2}(y) \tau^2 - \eta^2} \hat{h}(\eta, \tau) d\eta d\tau.$$

That is, the mapping $q_+|_{\partial\Omega} \rightarrow [\partial_\nu q_+]$ is an elliptic pseudo-differential operator at the element $(x_+, t_+, \eta_+, \tau_+)$ with principal symbol $(-2i) \sqrt{c^{-2}(x_+) \tau_+^2 - \eta_+^2}$. Therefore, the mapping $[\partial_\nu q_+]|_{\partial\Omega \times [0, T]} \rightarrow q_+|_{\partial\Omega \times [0, T]}$ is also a pseudo-differential operator near $(x_+, t_+, \eta_+, \tau_+)$ with the principal symbol $\frac{1}{(-2i) \sqrt{c^{-2}(x_+) \tau_+^2 - |\eta_+|^2}}$.

Noting that $f \rightarrow \chi g_+$ and $q_+|_{\partial\Omega \times [0, T]} \rightarrow q_+(\cdot, t)|_\Omega$ are FIOs, we obtain $f \rightarrow q_+(\cdot, t)|_\Omega$ is also an FIO. We, hence, can write the parametrix for q_+ in Ω in the form

$$q_+(x, t) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\phi_+(x, t, \xi)} b(x, t, \xi) \hat{f}(\xi) d\xi.$$

In particular,

$$q_+(y, t) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\phi_+(y, t, \xi)} b(y, t, \xi) \hat{f}(\xi) d\xi, \quad y \in \partial\Omega. \quad (3.10)$$

On the other hand,

$$\chi(y, t) g_+(y, t) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{i\phi_+(y, t, \xi)} \chi(y, t) a(y, t, \xi) \hat{f}(\xi) d\xi, \quad y \in \partial\Omega.$$

Since $\chi g_+ = [\partial_\nu q_+] \rightarrow q_+$ is a pseudo-differential with principal symbol $\frac{1}{(-2i)\sqrt{c^{-2}(x_+) \tau_+^2 - |\eta_+|^2}}$ at $(x_+, t_+, \eta_+, \tau_+)$, the principal part b_p of b satisfies

$$b_p(x_+, t_+, \xi_+) = -\frac{\chi(x_+, t_+)}{2i\sqrt{c^{-2}(x_+) \tau_+^2 - |\eta_+|^2}} a_0(x_+, \xi_+, t_+).$$

Since b_{in} and a_0 satisfy the same transport equation on the geodesic ray $r_+(x, \xi)$, the above equation implies

$$b_p(x, 0, \xi) = -\frac{\chi(x_+, t_+)}{2i\sqrt{c^{-2}\tau_+^2 - |\eta_+|^2}} a_0(x, 0, \xi) = -\frac{\chi(x_+, t_+)}{4i\sqrt{c^{-2}\tau_+^2 - |\xi_+|^2}}.$$

Noting that

$$\partial_t \phi_+(x, 0, \xi) = -c(x) |\nabla_x \phi_+(x, 0, \xi)| = -c(x) |\xi|,$$

we obtain, from (3.10), up to lower order terms,

$$f_+(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{ix \cdot \xi} \frac{c(x) |\xi| \chi(x_+, t_+)}{4\sqrt{c^{-2}(x_+) \tau_+^2 - |\xi_+|^2}} \hat{f}(\xi) d\xi.$$

Noting that $c(x) |\xi| = c(x_+) |\xi_+|$,[§] we obtain the mapping $f \rightarrow f_+$ is a pseudo-differential operator with principal symbol

$$\frac{c(x) |\xi| \chi(x_+, t_+)}{4\sqrt{c^{-2}(x_+) \tau_+^2 - |\eta_+|^2}} = \frac{c(x_+) |\xi_+| \chi(x_+, t_+)}{4\sqrt{|\xi_+|^2 - |\eta_+|^2}} = \frac{c(x_+) \chi(x_+, t_+)}{4 \cos(\theta_+)}.$$

Repeating the above argument for $f \rightarrow f_-$, we finish the proof. \square

Remark 3.7. *Let us make the following observations:*

- (a) *The calculus of symbols can be explained more intuitively by considering the current set up as the limit of the open set measurement. This will be discussed in Section 3.2.*

[§]This comes from the fact that $c(x) |\xi| = \tau$ and τ is constant on the bicharacteristic rays (see, e.g., [47]).

(b) We notice that the function χ plays the role of preconditioning for the inverse problem of PAT. Formula (3.7) may give us some hint on how to make a good choice of χ . Indeed, let us consider the case $c = 1$, S is the sphere of radius R , and $\chi(y, t) = t$. Since the geodesics are straight lines, we observe that $\theta_+ = \theta_- = \theta$ and hence

$$\sigma_0(x, \xi) = \frac{t_+ + t_-}{4 \cos \theta}.$$

We notice that $t_+ + t_-$ is the length of the line segment connecting x_+ and x_- . A simple geometric observation then gives:

$$\sigma_0(x, \xi) = \frac{R}{2}.$$

We obtain

$$\mathbf{L}^* \mathbf{L} = \frac{R}{2} \mathbf{I} + \mathbf{K}, \quad (3.11)$$

where \mathbf{K} is a compact operator. Therefore, in such a situation, the CG method for the inverse problem of PAT converges superlinearly (see the discussion of the CG method in Subsection 2.3). We note that (3.11) can be derived from the results in [16, 15]. Indeed, for odd d , [16] even gives $\mathbf{L}^* \mathbf{L} = \frac{R}{2} \mathbf{I}$.

The above discussion also suggests the choice of $\chi(y, t) = t$ when the speed is almost constant. The in-depth discussion on the preconditioning, however, is beyond the scope of this article.

Let us recall the time-reversal technique for PAT (see, e.g., [16, 29, 54]). Consider the time reversal wave equation

$$\begin{cases} c^{-2}(x) q_{tt}(x, t) - \Delta q(x, t) = 0, & (x, t) \in \Omega \times [0, T], \\ q(x, T) = \phi(x), \quad q_t(x, T) = 0, & x \in \Omega, \\ q(x, t) = \chi(x, t) g(x, t), & (x, t) \in \partial\Omega \times [0, T]. \end{cases}$$

Here, ϕ is the harmonic extension of $\chi(x, T) g(x, T)|_{x \in \partial\Omega}$ to $\bar{\Omega}$. The time-reversal operator is defined by $\Lambda g = q(\cdot, 0)$. It is proved in [54] that, if $\chi \equiv 1$,

$$\|\mathbf{I} - \Lambda \mathbf{L}\| < 1.$$

This suggest that $\Lambda \mathbf{L}$ can be used as the first step for iterative method (see [54]), which we now call iterative time reversal or Neumann series solution (see also [52] for the thorough numerical discussion and [55] for nonsmooth sound speed). It is shown in [54] that $\Lambda \mathbf{L}$ is a pseudodifferential operator of order zero with the principal symbol

$$\sigma_0(x, \xi) = \frac{1}{2} (\chi(x_+, t_+) + \chi(x_+, t_-)).$$

This is different from the symbol of $\mathbf{L}^* \mathbf{L}$ shown in Theorem 3.6. We, in particular, conclude that the adjoint operator \mathbf{L}^* is fundamentally different from the time reversal operator Λ . We also note that no proof for the convergence of iterative time reversal method is available for limited data problem, even under the visibility condition. However, numerically it works reasonably well in this situation (as demonstrated in [52] and also Section 4).

3.2 Open domain observations revisited

Let us consider the setup used in [3]. Namely, let us consider the operator \mathbf{L}_ω defined by

$$(\mathbf{L}_\omega f)(x, t) = \omega(x, t) p(x, t).$$

Here, p is the solution of (1.1) with initial pressure f , and $0 \leq \omega \in C^\infty(\mathbb{R}^d \times [0, T])$ is the window function, whose support determines the accessible region for the data. We assume that $\text{supp}(\omega) = \mathbf{D} \times [0, T]$, where $\mathbf{D} \subset \mathbb{R}^d$ is an open band. That is, $\mathbb{R}^d \setminus \overline{\mathbf{D}} = \Omega \cup \Omega'$, where $\Omega \cap \Omega' = \emptyset$, Ω is bounded and Ω' is unbounded. We again, assume that f is supported in $\overline{\Omega}_0$ where $\Omega_0 \Subset \Omega$, and are interested in the problem of finding f given $\mathbf{L}_\omega f$. It can be solved by the iterative methods described in Section 2, which we do not elaborate further in this article. We, instead, focus on analyzing the adjoint operator in this setup.

We propose the following method to compute the adjoint of \mathbf{L}_ω . Consider the time-reversed problem

$$\begin{cases} c^{-2}(x) q_{tt}(x, t) - \Delta q(x, t) = -\omega(x, t) h(x, t), & (x, t) \in \mathbb{R}^d \times (0, T), \\ q(x, T) = 0, \quad q_t(x, T) = 0, & x \in \mathbb{R}^d. \end{cases} \quad (3.12)$$

We define

$$\mathbf{L}_\omega^*(h) = \partial_t q(\cdot, 0).$$

Let $\bar{h}(\cdot, t) = h(\cdot, t) - h(\cdot, T)$ and \bar{q} be the solution of (3.12) with h being replaced by \bar{h} . We define (recalling that Π denotes the projection from $H_1(\Omega_0)$ onto $H_0^1(\Omega_0)$)

$$\bar{\mathbf{L}}_\omega^*(h) = \Pi[\bar{q}_t(\cdot, 0)],$$

Let us recall the space $H_i(\Omega_0)$ defined at the beginning of Section 3. Similarly to the spaces $H_i(\chi)$, we define

$$\begin{aligned} H_0(\omega) &:= \{h: \|h\|_{H_0(\omega)} := \|\sqrt{\omega} h\|_{L^2(\mathbf{D} \times [0, T])} < \infty\}, \\ H_1(\omega) &:= \{h: h(\cdot, 0) \equiv 0 \text{ in } \mathbf{D}, \|h\|_{H_1(\omega)} := \|h_t\|_{H_0(\omega)} < \infty\}. \end{aligned}$$

The following lemma shows that \mathbf{L}_ω^* and $\bar{\mathbf{L}}_\omega^*$ are the adjoints of \mathbf{L}_ω , given the correct mapping spaces.

Theorem 3.8. *We have*

(a) For all $f \in H_0(\Omega)$ and $h \in H_0(\omega)$,

$$\langle \mathbf{L}_\omega f, h \rangle_{H_0(\omega)} = \langle f, \mathbf{L}_\omega^* h \rangle_{H_0(\Omega_0)}.$$

That is, \mathbf{L}_ω^* is the adjoint of $\mathbf{L}_\omega: H_0(\Omega_0) \rightarrow H_0(\omega)$.

(b) Assume that ω is independent of t , then for all $f \in H_1(\Omega)$ and $h \in H_1(\omega)$,

$$\langle \mathbf{L} f, h \rangle_{H_1(\omega)} = \langle f, \bar{\mathbf{L}}_\omega^* h \rangle_{H_1(\Omega_0)}.$$

That is, $\bar{\mathbf{L}}_\omega^*$ is the adjoint of $\mathbf{L}_\omega: H_1(\Omega_0) \rightarrow H_1(\omega)$.

The proof of Theorem 3.8 is similar to that of Theorem 3.2. We skip it for the sake of brevity. Let us notice that our definition of \mathbf{L}_ω^* is slightly different from [3]. It is motivated by case of the observation on a surface discussed the previous section. Our definition matches with that in [3] if $\omega(\cdot, T)h(\cdot, T) \equiv 0$.

The following theorem gives us a microlocal characterization of the normal operator $\mathbf{L}_\omega^* \mathbf{L}_\omega$.

Theorem 3.9. *The operator $\mathbf{L}_\omega^* \mathbf{L}_\omega$ is a pseudo-differential operator of order zero whose principal symbol is*

$$\sigma_0(x, \xi) = \frac{1}{4} \left(\int_0^T c^2(x_+(t)) \omega(x_+(t), t) dt + \int_0^T c^2(x_-(t)) \omega(x_-(t), t) dt \right).$$

Here, $x_\pm(t) = r_\pm(x, \xi)(t)$ is unit speed geodesic ray originated from x at time $t = 0$ along the direction of $\pm\xi$.

Proof. Let us consider the solution q of the time reversed problem (3.12) with $h = p|_{\mathbb{D} \times [0, T]}$. Applying the Duhamel's principle, we can write

$$q(\cdot, t) = \int_t^T q(\cdot, t; s) ds,$$

where $q(x, t; s)$ satisfies

$$\begin{cases} c^{-2}(x) q_{tt}(x, t; s) - \Delta q(x, t; s) = 0, & (x, t) \in \mathbb{R}^d \times (0, s), \\ q(x, s; s) = 0, \quad q_t(x, s; s) = c^2(x) \omega(x, s) p(x, s), & x \in \mathbb{R}^d. \end{cases}$$

Therefore, denoting $q'(\cdot, t; s) = q_t(\cdot, t; s)$,

$$\mathbf{L}_\omega^*(p) = q_t(\cdot, 0) = \int_0^T q'(\cdot, 0; s) ds - q(\cdot, 0; 0) = \int_0^T q'(\cdot, 0; s) ds = \int_0^T \mathbf{L}(s)(p) ds.$$

Let us show that $f \rightarrow \mathbf{L}(s)(p) := q'(\cdot, 0; s)$ is a pseudo-differential operator with the principal symbol

$$\sigma_0(x, \xi; s) = \frac{1}{4} (c^2(x_+(s)) \omega(x_+(s), s) + c^2(x_-(s)) \omega(x_-(s), s)).$$

Indeed, we note that $q'(x, t; s)$ satisfies

$$\begin{cases} c^{-2}(x) q'_{tt}(x, t; s) - \Delta q'(x, t; s) = 0, & (x, t) \in \mathbb{R}^d \times (0, s), \\ q'(x, s; s) = c^2(x) \omega(x, s) p(x, s), \quad q'_t(x, s; s) = 0, & x \in \mathbb{R}^d. \end{cases} \quad (3.13)$$

Assume that $(x, \xi) \in (\mathbb{T}^* \Omega \setminus 0) \cap \text{WF}(f)$ and consider the propagation of p , governed by the wave equation (1.1). The singularity of f at (x, ξ) is broken into two equal parts propagating along the geodesic rays $r_\pm(x, \xi)$. Let us consider the propagation along $r_+(x, \xi)$. The projection of the propagated singularity at $t = s$ to $\mathbb{T}^* \mathbb{R}^d$ produces a corresponding singularity of $p(\cdot, s)$. Let us consider the propagation of that singularity due to the equation (3.13). Firstly, due to the end time condition at $t = s$, it is multiplied by $c^2(x) \omega(x, s)$. Then, it is broken into two equal parts propagating along two opposite directions (in reversed time). One of them hits back to (x, ξ) at $t = 0$ (this travel along $r_+(x, \xi)$ but in negative direction) and the other one lands outside of Ω . Therefore, the recovered singularity at (x, ξ) , as just described, is of $\frac{1}{2} c^2(x_+(s)) \omega(x_+(s), s)$ of the part of original singularity at (x, ξ) propagating along the ray $r_+(x, \xi)$. Similar argument for the negative

ray $r_-(x, \xi)$ gives us the second recovered singularity with the magnitude $\frac{1}{2}c^2(x_+(s))\omega(x_+(s), s)$ of the part of original singularity at (x, ξ) propagating along the negative ray $r_-(x, \xi)$. Since each part (propagating on each direction) is half of the original singularity, we obtain the recovered singularity is $\frac{1}{4}(c^2(x_+(s))\omega(x_+(s), s) + c^2(x_-(s))\omega(x_-(s), s))$ of the original singularity. This intuitively, shows that $\mathbf{L}(s)\mathbf{L}_\omega$ is a pseudo-differential operator of order zero with the principal symbol

$$\frac{1}{4}\left(c^2(x_+(s))\omega(x_+(s), s) + c^2(x_-(s))\omega(x_-(s), s)\right).$$

A more rigorous proof can be done by writing the corresponding form of the parametrix for the wave equations (1.1) and (3.13). However, we skip it for the sake of simplicity.

Now, since $\mathbf{L}_\omega^*\mathbf{L}_\omega = \int_0^T \mathbf{L}(s)\mathbf{L}_\omega ds$, we obtain that $\mathbf{L}_\omega^*\mathbf{L}_\omega$ is a pseudo-differential operator of order zero with the symbol

$$\frac{1}{4}\int_0^T \left(c^2(x_+(s))\omega(x_+(s), s) + c^2(x_-(s))\omega(x_-(s), s)\right) ds.$$

This finishes the proof of the theorem. □

Remark 3.10. *Let us consider $\omega = \omega_\epsilon$ to be a family of smooth function that approximate the $\chi(x, t)\delta_{\partial\Omega}(x)$. Then, the setup for the observation on the surface $\partial\Omega$ is just the limit as $\epsilon \rightarrow 0$. We note that*

$$\lim_{\epsilon \rightarrow 0} \int_0^T c^2(x_\pm(s))\omega(x_\pm(s), s) ds = \frac{c(x_\pm)\chi(x_\pm, t_\pm)}{\cos\theta_\pm}.$$

Therefore,

$$\lim_{\epsilon \rightarrow 0} \sigma_0(x, \xi) = \frac{1}{4} \left(\frac{c(x_+)\chi(x_+, t_+)}{\cos(\theta_+)} + \frac{c(x_-)\chi(x_-, t_-)}{\cos(\theta_-)} \right),$$

which is the symbol $\mathbf{L}^*\mathbf{L}$ in Theorem 3.6.

4 Numerical experiments

In this section we implement the iterative methods presented in Section 2 for PAT for the observation on a surface $\partial\Omega$. We will employ the explicit formulation of $\mathbf{L}^* = \mathbf{L}_0^*$ presented in Section 3. We will chose the weight function χ to be independent of the time variable t .[¶] We note that under the visibility condition, Landweber's and the CG methods have linear convergence in $H_0(\Omega_0) = L^2(\Omega_0)$, since by Theorem 3.4 the inversion of $\mathbf{L}f = g$ is well-posed in this situation. If using $\mathbf{L}^* = \mathbf{L}_1^*$, we would obtain the linear rate of convergence in $H_1(\Omega_0) \simeq H^1(\Omega_0)$. However, we will refrain from that choice.

We only consider the two dimensional space and Ω to be a disc centered at the origin: $\Omega = B_R(0) = \{x \in \mathbb{R}^2: |x| = R\}$. All presented results assume non-constant sound speed. We consider the following test cases

(T1) Non-trapping sound speed, complete data;

(T2) Non-trapping sound speed, partial data, visible phantom (i.e., the visibility condition holds);

[¶]Other choices of χ may result in better conditioning of the problem. However, studying optimal preconditioning is beyond the scope of this article.

(T3) Non-trapping sound speed, partial data, invisible phantom (i.e., the invisibility condition holds);

(T4) Trapping sound speed, complete data.

The reader is referred to the introduction (also, see, [29]) for the description of the trapping and non-trapping sound speed. We will compare the results for the Landweber’s method, Nesterov’s method, the CG method (as proposed in the present paper) as well as the iterative time reversal algorithm proposed in [52]. Thereby we investigate the numerical speed of convergence as well as stability and accuracy of all these algorithms. For Landweber’s and Nesterov’s method we have taken the step size equal to $\gamma = 1$, which worked well in all our numerical simulations.

As described in Subsection 4.2 the proposed iterative schemes are implemented by numerical realizations of all involved operators. Thereby the most crucial steps are accurate discrete solvers for the forward and backward wave equation. For that purpose we implemented the k -space method (described in Subsection 4.1) that is an efficient FFT based numerical solution method that does not suffer from numerical dispersion that arises when solving the wave equation with standard finite difference or finite element methods.

4.1 The k -space method for numerically solving the wave equation

In this subsection we briefly describe the k -space method as we use it to numerically compute the solution of wave equation, which is required for evaluating the forward operator \mathbf{L} and its adjoint \mathbf{L}^* .

Consider the solution $p: \mathbb{R}^2 \times (0, T) \rightarrow \mathbb{R}$ of the two-dimensional wave equation

$$c^{-2}(x) p_{tt}(x, t) - \Delta p(x, t) = s(x, t) \quad \text{for } (x, t) \in \mathbb{R}^2 \times (0, T), \quad (4.1)$$

$$p(x, 0) = f(x) \quad \text{for } x \in \mathbb{R}^2, \quad (4.2)$$

$$p_t(x, 0) = 0 \quad \text{for } x \in \mathbb{R}^2, \quad (4.3)$$

where $s: \mathbb{R}^2 \times (0, T) \rightarrow \mathbb{R}$ is a given source term and $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ the given initial pressure. Several well investigated methods for numerically solving (4.1) (and analogously for the wave equation in higher dimensions) are available and have been used for photoacoustic tomography. This includes finite difference methods [8, 49, 56], finite element methods [6] as well as Fourier spectral and k -space methods [10, 30, 58]. In this paper we use a k -space method for numerically solving (4.1) because this method does not suffer from numerical dispersion. The k -space method is implemented in the freely available k -wave toolbox (see [58]); in order to be flexible in our implementations we have developed our own code as described below.

The k -space method makes the ansatz (see [10, 40, 57])

$$p(x, t) = w(x, t) - v(x, t) \quad \text{for } (x, t) \in \mathbb{R}^2 \times (0, T), \quad (4.4)$$

where the pressure p is written as linear combination of the auxiliary quantities $w := c_0^2/c^2 p$ and $v := (1 - c_0^2/c^2) p$. Here $c_0 > 0$ a suitable constant; we take $c_0 := \max \{c(x) : x \in \mathbb{R}^2\}$. One easily verifies that the wave (4.1) is equivalent to the following system of equations,

$$\begin{cases} w_{tt}(x, t) - c_0^2(x) \Delta w(x, t) = c_0^2 s(x, t) - c_0^2 \Delta v(x, t) & \text{for } (x, t) \in \mathbb{R}^2 \times (0, T), \\ v(x, t) = \frac{c_0^2 - c(x)^2}{c_0^2} w(x, t) & \text{for } (x, t) \in \mathbb{R}^2 \times (0, T). \end{cases} \quad (4.5)$$

Interpreting Δv as an additional source term, the first equation in (4.5) is a standard wave equation with constant sound speed. This suggests the time stepping formula

$$w(x, t + h_t) = 2w(x, t) + w(x, t - h_t) - 4\mathcal{F}_\xi^{-1} [\sin(c_0|\xi|h_t/2)^2 \mathcal{F}_x[w(x, t) - v(x, t)] - (c_0 h_t/2)^2 \text{sinc}(c_0|\xi|h_t/2)^2 \mathcal{F}_x[s(x, t)]] , \quad (4.6)$$

where \mathcal{F}_x and \mathcal{F}_ξ^{-1} denote the Fourier and inverse Fourier transform in the spatial variable x and the spatial frequency variable ξ , respectively, and $h_t > 0$ is a time stepping size. For constant sound speed we have $v = 0$, in which case the solution of (4.1) exactly satisfies (4.6) and vice versa. In the case of variable sound speed there is no such equivalence because v is itself dependent on w . Nevertheless, in any case (4.6) serves as the basis of an efficient and accurate iterative time stepping scheme for numerically computing the solution of the wave equation.

The resulting k -space method for solving (4.1) is summarized in Algorithm 1.

Algorithm 1 (The k -space method). *For given initial pressure $f(x)$ and source term $s(x, t)$ approximate the solution $p(x, t)$ of (4.1) as follows:*

- (1) Define initial conditions $w(x, -h_t) = w(x, 0) = c_0^2/c^2 f(x)$;
- (2) Set $t = 0$;
- (3) Compute $w(x, t + h_t)$ by evaluating (4.6);
- (4) Compute $v(x, t + h_t) := (1 - c(x)^2/c_0^2) w(x, t + h_t)$;
- (5) Compute $p(x, t + h_t) := w(x, t + h_t) - v(x, t + h_t)$;
- (6) Set $t \leftarrow t + h_t$ and go back to (3).

Algorithm 1 can directly be used to evaluate the forward operator $\mathbf{L}f$ by taking $s(x, t) = 0$ and restricting the solution to the measurement surface S_R , that is $\mathbf{L}f = p|_{S_R \times (0, T)}$. Recall that the adjoint operator is given by $\mathbf{L}^*g = q_t(\cdot, 0)$, where $q: \mathbb{R}^2 \times (0, T) \rightarrow \mathbb{R}$ satisfies the adjoint wave equation

$$c^{-2}(x) q_{tt}(x, t) - \Delta q(x, t) = -\delta_{S_R}(x) g(x, t) \quad \text{for } (x, t) \in \mathbb{R}^2 \times (0, T) \quad (4.7)$$

$$q_t(x, T) = q(x, T) = 0 \quad \text{for } x \in \mathbb{R}^2. \quad (4.8)$$

By substituting $t \leftarrow T - t$ and taking $s(x, t) = g(x, T - t) \delta_S(x)$ as source term in (4.1), Algorithm 1 can also be used to evaluate the \mathbf{L}^* . In the partial data case where measurements are made on a subset $S \subsetneq S_R$ only, the adjoint can be implemented by taking the source $s(x, t) = \chi(x, t) g(x, T - t) \delta_{S_R}(x)$ with an appropriate window function $\chi(x, t)$. In order to use all available data, in our implementations we take the window function to be equal to one on the observation part S and zero outside. This choice of the window function is known to create streak artifacts into the picture [18, 48, 5]. However, as we will see below, the artifacts fade away quickly after several iterations when the problem is well-posed.

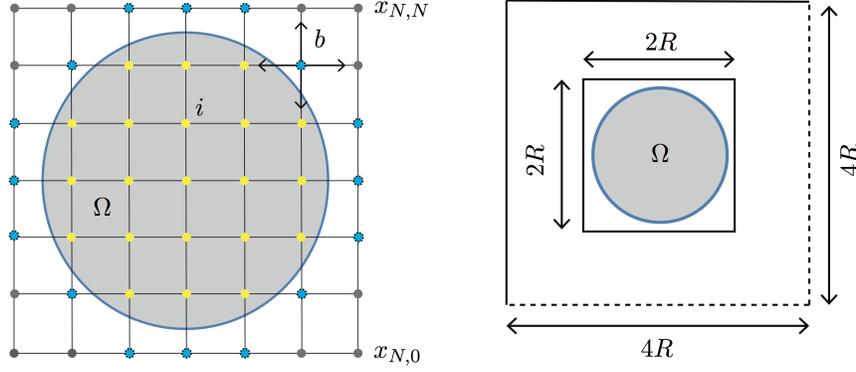


Figure 4.1: Left: The discrete domain Ω_N is defined as the set of all indices $i \in \{0, \dots, N\}^2$ with $x_i \in \Omega$. Right: The index b is contained in the discrete boundary $\partial\Omega_N$, because one of its neighbors is contained in Ω_N . Right: The domain $[-R, R]$ is embedded in a larger domain $[-2R, 2R]$ to avoid effects due to periodization.

4.2 Numerical realization

The iterative approaches for solving the equation $\mathbf{L}f = g$ are implemented with discrete counterparts of all operators introduced above. Thereby the function $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ is represented by a discrete vector

$$\mathbf{f} = (f(x_i))_{i_1, i_2=0}^N \in \mathbb{R}^{(N+1) \times (N+1)},$$

where $x_i = (-R, -R) + 2iR/N$ for $i = (i_1, i_2) \in \{0, \dots, N\}^2$ are equidistant grid points in the square $[-R, R]^2$. We define the discrete domain $\Omega_N \subset \{0, \dots, N\}^2$ as the set of all indices i with $x_i \in \Omega$. Further, the discrete measurements are made on parts of the discrete boundary $\partial\Omega_N$, that is defined as the set of all elements $b = (b_1, b_2) \in \{0, \dots, N\}^2 \setminus \Omega_N$ for which at least one of the discrete neighbors $(b_1 + 1, b_2)$, $(b_1 - 1, b_2)$, $(b_1, b_2 + 1)$, $(b_1, b_2 - 1)$ is contained in Ω_N , see the right image in Figure 4.1. All phantoms in our numerical simulations are chosen to have support in a compact subset of Ω . We will choose the discrete version of Ω_0 to be the set $\{x_i: i \in \Omega_N\}$.

The discrete forward operator can be written in the form

$$\mathbf{L}_{N,M}: \mathbb{R}^{(N+1) \times (N+1)} \rightarrow \mathbb{R}^{|\partial\Omega_N| \times (M+1)}: \mathbf{f} \mapsto (\mathbf{R}_{N,M} \circ \mathbf{W}_{N,M})\mathbf{f}. \quad (4.9)$$

Here $M + 1$ is the number of equidistant temporal sampling points in $[0, T]$, $\mathbf{W}_{N,M}$ is a discretization of the solution operator for the wave equation and $\mathbf{R}_{N,M}$ the linear operator that restricts the discrete pressure to spatial grid points restricted to $\partial\Omega_N \subset \{0, \dots, N\}^2$. The adjoint operator is then given by $\mathbf{L}_{N,M}^* = \mathbf{D}_0 \circ \mathbf{W}_{N,M}^* \circ \mathbf{R}_{N,M}^*$, where $\mathbf{R}_{N,M}^*$ is the embedding operator from $\mathbb{R}^{|\partial\Omega_N|}$ to $\mathbb{R}^{(N+1) \times (N+1)}$, $\mathbf{W}_{N,M}^*$ is the solution operator the adjoint wave equation (4.7), and \mathbf{D}_0 a discretization of the time derivative evaluated at $t = 0$ and restricted to Ω_0 .

For computing the solution operator $\mathbf{W}_{N,M}$ we use the k -space method described in Subsection 4.1. In the actual implementation of Algorithm 1, the Fourier transform of the function f is replaced by the FFT algorithm applied to \mathbf{f} (and likewise for the inverse Fourier transform). When applied directly to the given function values, the FFT algorithm causes the numerical solution to be $2R$ periodic. For the numerical solution of the wave equation we therefore embed the data vector $\mathbf{f} \in \mathbb{R}^{(N+1) \times (N+1)}$ in a larger vector in $\mathbb{R}^{(2N+1) \times (2N+1)}$, whose entries correspond to sampled values on an equidistant grid in $[-2R, 2R]^2$ (see the right image in Figure 4.1). As the sound speed is assumed to be equal to one outside of Ω , the numerical solution for times $t \leq 2R$ (which will

always be the case in our simulations) is free from periodization artifacts in the domain Ω and on the measurement surface.

Remark 4.1 (Numerical complexity of our iterative algorithms). *Using the FFT algorithm any time step in Algorithm 1 can be implemented using $\mathcal{O}(N^2 \log N)$ floating point operations (FLOPS). Performing $M \sim N$ times steps therefore this yields to $\mathcal{O}(N^3 \log N)$ algorithms for implementing the forward operator $\mathbf{L}_{N,M}$ and its adjoint $\mathbf{L}_{N,M}^*$. Consequently, performing one iterative step (for example using the CG or the Landweber iteration) is almost as fast as applying the filtered backprojection type algorithm (which requires $\mathcal{O}(N^3)$ FLOPS) for evaluating the adjoint or the inverse of \mathbf{L} . In three spatial dimensions the complexity of the k -space method scales to $\mathcal{O}(N^4 \log N)$. In this case one iterative step is already faster than filtered backprojection type algorithms (which in this case requires $\mathcal{O}(N^5)$ FLOPS). As we will see in the numerical results presented below, around 10 iterations with the CG method already gives very accurate reconstruction results. This shows that our iterative algorithms are a good option for PAT image reconstruction even in situations, where an explicit filtered backprojection type formula is available.*

4.3 Test case (T1): Non-trapping speed, complete data

We first consider the non-tapping sound speed (taken from [52])

$$c(x) = 1 + w(x) (0.1 \cos(2\pi x_1) + 0.05 \sin(2\pi x_2)) , \quad (4.10)$$

where $w: \mathbb{R}^2 \rightarrow [0, 1]$ is a smooth function that vanishes outside $B_1(0)$ and is equal to one on $B_{1/2}(0)$. The non-trapping sound speed c , the phantom f and the corresponding full data $\mathbf{L}f$ are illustrated in the top row in Figure 4.2. For the results presented in this section we use $R = 1$ and $N = 200$, which yields a spatial step size of $h_x = 2R/N = 1/100$. We further use a final time $T = 1.5$ and take $M = 800$ for the temporal discretization.

We performed iterative reconstructions using the following methods:

- (a) Iterative time reversal method
- (b) Landweber's method
- (c) Nesterov's method
- (d) CG method.

Figure 4.2 shows reconstruction results using these methods after 1, 10 and 200 iterations. One notices that all iterations converge quite rapidly to the original phantom. To investigate the convergence behavior more carefully, in Figure 4.3 we plot the logarithm of the squared discrete L^2 -reconstruction error and squared residual

$$\begin{aligned} \|\mathbf{f}_n - \mathbf{f}\|_2^2 &:= \sum_i |\mathbf{f}_n[i] - \mathbf{f}[i]|^2 h_x^2 \simeq \|f - f_n\|_{L^2}^2 , \\ \|\mathbf{L}_{N,M}\mathbf{f}_n - \mathbf{g}\|^2 &:= \sum_{b,j} |\mathbf{L}_{N,M}\mathbf{f}_n[b, j] - \mathbf{g}[b, j]|^2 h_x h_t \simeq \|\mathbf{L}f_n - \mathbf{g}\|_{L^2}^2 , \end{aligned}$$

respectively. One concludes from Figure 4.3, that all iterative schemes converge quite rapidly. In particular the CG method is the fastest.

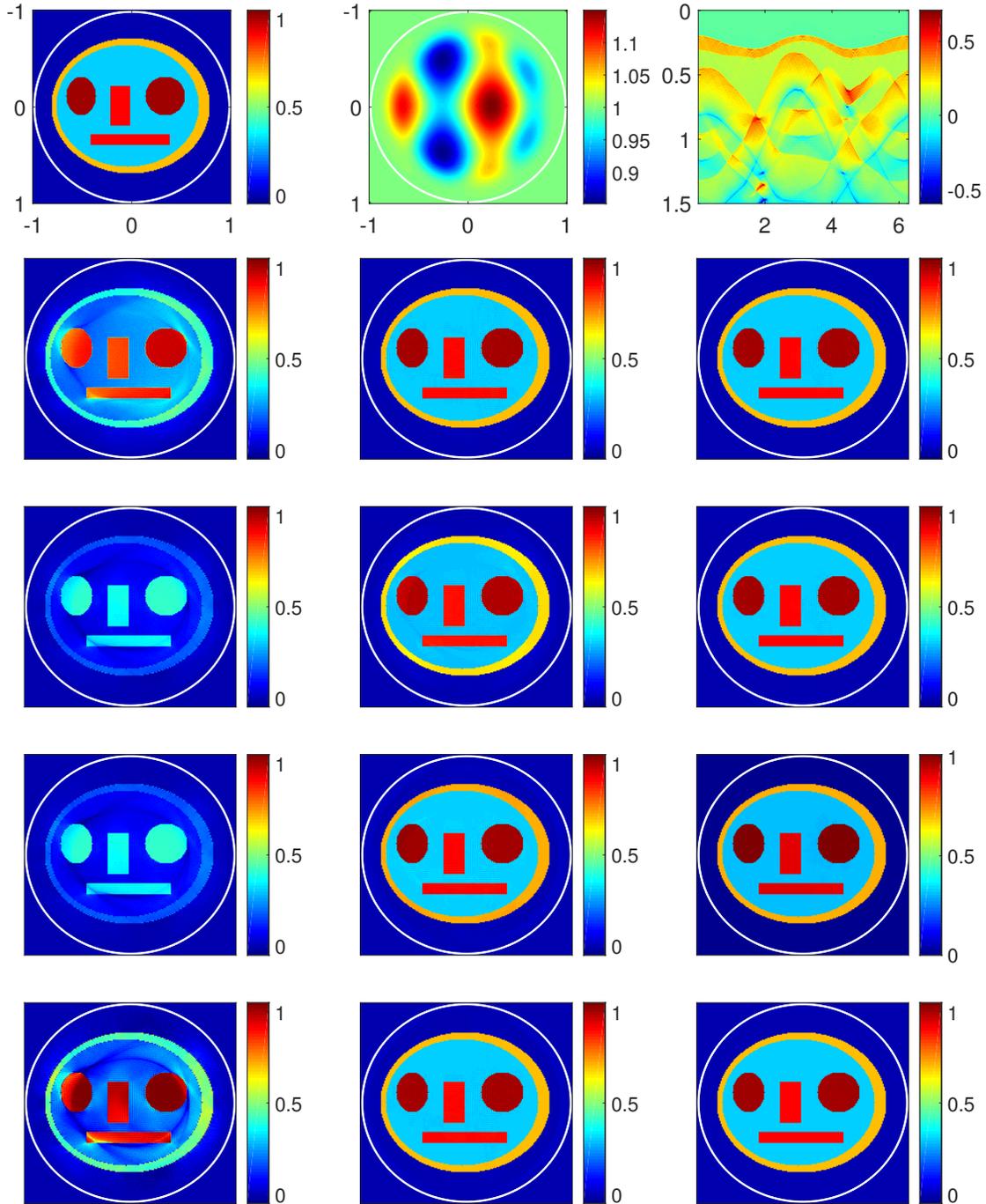


Figure 4.2: TEST CASE (T1), RECONSTRUCTIONS FROM EXACT DATA. Row 1: Initial pressure data f (left), non-trapping sound speed c (middle), and computed pressure data $\mathbf{L}f$ (right). Row 2: Iterative time reversal (after 1, 10 and 200 iterations). Row 3: Landweber's method (after 1, 10 and 200 iterations). Row 4: Nesterov's method (after 1, 10 and 200 iterations). Row 5: the CG method (after 1, 10 and 200 iterations). The white dots indicate the measurement curve.

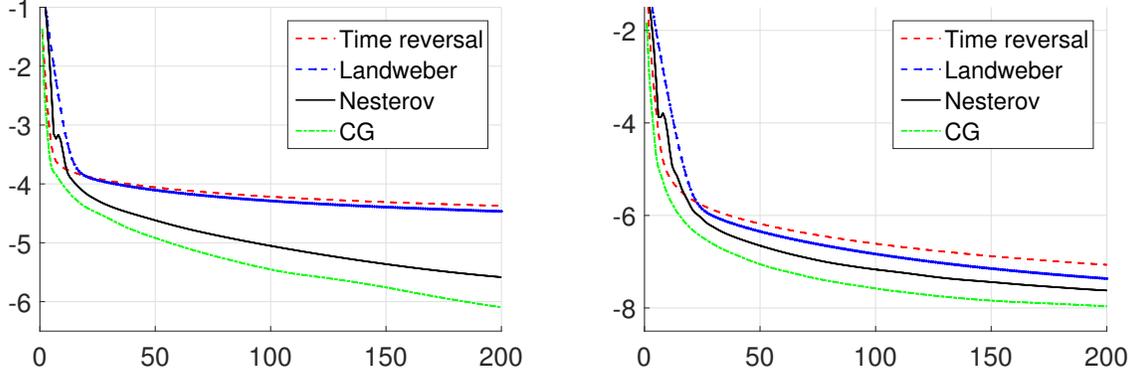


Figure 4.3: TEST CASE (T1), CONVERGENCE BEHAVIOR FOR EXACT DATA Left: Logarithm of squared reconstruction error $\|\mathbf{f}_n - \mathbf{f}\|_2^2$ in dependence of the iteration number. Left: Logarithm of residual $\|\mathbf{L}_{N,M}\mathbf{f}_n - \mathbf{g}\|_2^2$ in dependence of the iteration number.

In order to further investigate the behavior of the algorithms we repeated the computations with inexact data. To that end, we generated the data on a different grid, where we use $N = 350$ and $M = 1300$ (recall that the iterative algorithm uses $N = 200$ and $M = 800$). Further, we added Gaussian white noise to the data with a standard deviation equal to 5% of the L^2 -norm of $\mathbf{L}_{N,M}\mathbf{f}$. The total L^2 -error in the data is 0.049 compared to the L^2 -norm $\|\mathbf{g}\|_2 = 0.44$ of the exact data.

Figure 4.4 shows the reconstruction results from inexact data using iterative time reversal, Landweber’s, Nesterov’s, and the CG methods. The errors and the residuals again decrease quite rapidly in the first iterative steps. However after about 10 iterations the error as well as the residuals do not further decrease. Consequently, the iterations can be stopped at a certain iteration index n_* . This is due to the noise in the data which causes the data to be outside the range of $\mathbf{L}_{N,M}$. However, these results also reveal that we are in a stable situation, because the error does not significantly increase after reaching the stopping index n_* .

Note that we do not show results using the Landweber’s method proposed in [6]. Due to the smoothing operator $-\Delta^{-1}$ (which is the adjoint of the embedding $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$), that method is much slower than the Landweber’s method presented in the present article. On the other hand, the application of $-\Delta^{-1}$ may have the advantage of stabilizing the iteration.

4.4 Test case (T2): Non-trapping speed, partial data, visible phantom

As next test case we investigate the case of partial data where all singularities of the phantom are visible. As before we compare iterative time reversal, Landweber’s, Nesterov’s, and the CG methods using the sound speed given in (4.10). We again take $N = 200$, $R = 1$, $T = 1.5$ and $M = 800$. The detection curve together with the phantom, the non-trapping sound speed, and the simulated data are shown in the top row in Figure 4.5. One notices that the partial data have been collected on an arc with opening angle $4\pi/3$.

Rows 2 to 5 in Figure 4.5 show reconstruction results with iterative time reversal, Landweber, Nesterov’s, and the CG methods after 1, 10 and 200 iterations. Because the visibility condition holds, the linear convergence rate for the Landweber’s and CG methods is applicable. And indeed one observes fast convergence of all methods. As demonstrated by Figure 4.9 (which shows the reconstruction error and the residuals depending on the iteration index n) the CG iteration is again the most rapidly converging method.

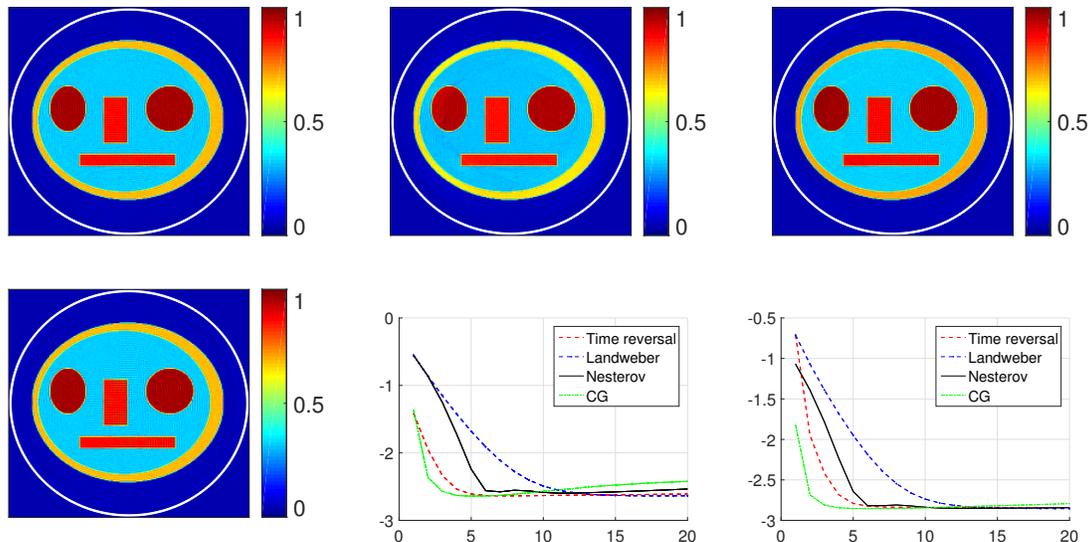


Figure 4.4: TEST CASE (T1), INEXACT DATA. Top left: Iterative time reversal, Top center: Landweber’s method, top right: Nesterov’s method. Bottom left: the CG method (all after 10 iterations). Bottom center: squared error. Bottom right: squared residuals.

To avoid inverse crimes and to investigate the behavior of the algorithms under real life scenario, we repeated the simulations with inexact data where we simulated the data on a different grid (using $N = 350$ and $M = 1300$) and further added Gaussian noise to the data (again with a standard deviation equal to 5% of the L^2 -norm of the exact data). The reconstruction results from inexact data are shown in Figure 4.10. They clearly demonstrate that all schemes provide good results. The CG method is again the fastest. The error $\|L_{N,M}f - g^\delta\|_2$ in the data is 0.0258. The residuals after 10 iterations are 0.2208 for the iterative time reversal, 0.0196 for the CG, 0.0199 for Nesterov’s, and 0.0222 for the Landweber’s methods. This in particular also shows that the discrepancy principle yields a well defined stopping index with an reconstruction error in the order of the data error.

4.5 Test case (T3): Non-trapping speed, partial data, invisible phantom

In this case we investigate the ill-posed problem, since the invisibility condition holds. We use again the sound speed given in (4.10) and take $N = 200$, $R = 1$, $T = 1.5$ and $M = 800$. The detection curve together with the phantom, the non-trapping sound speed, and the simulated data are shown in the top row in Figure 4.8. One notices that the partial data has been collected on an arc with opening angle $2\pi/3$ (the measurement curve). The invisibility condition holds in this setup.

Rows 2 to 5 in Figure 4.8 show reconstruction results with iterative time reversal, Landweber’s, Nesterov’s, and the CG methods after 1, 10 and 200 iterations. Because the inverse problem is ill-posed no convergence rate results for the Landweber, Nesterov’s, and CG methods are available. And indeed one observes that the reconstruction results are worse compared to the the case of a completely visible phantom. Again we investigated the convergence behavior more carefully. For that purpose Figure 4.9 shows the reconstruction error and the residuals depending on the iteration index n . While the residuals tend to zero quite fast, reconstruction error now decreases much slower than in the previous examples. Nevertheless also in this situation the CG iteration clearly yields the smallest reconstruction error for a given number of iterations.

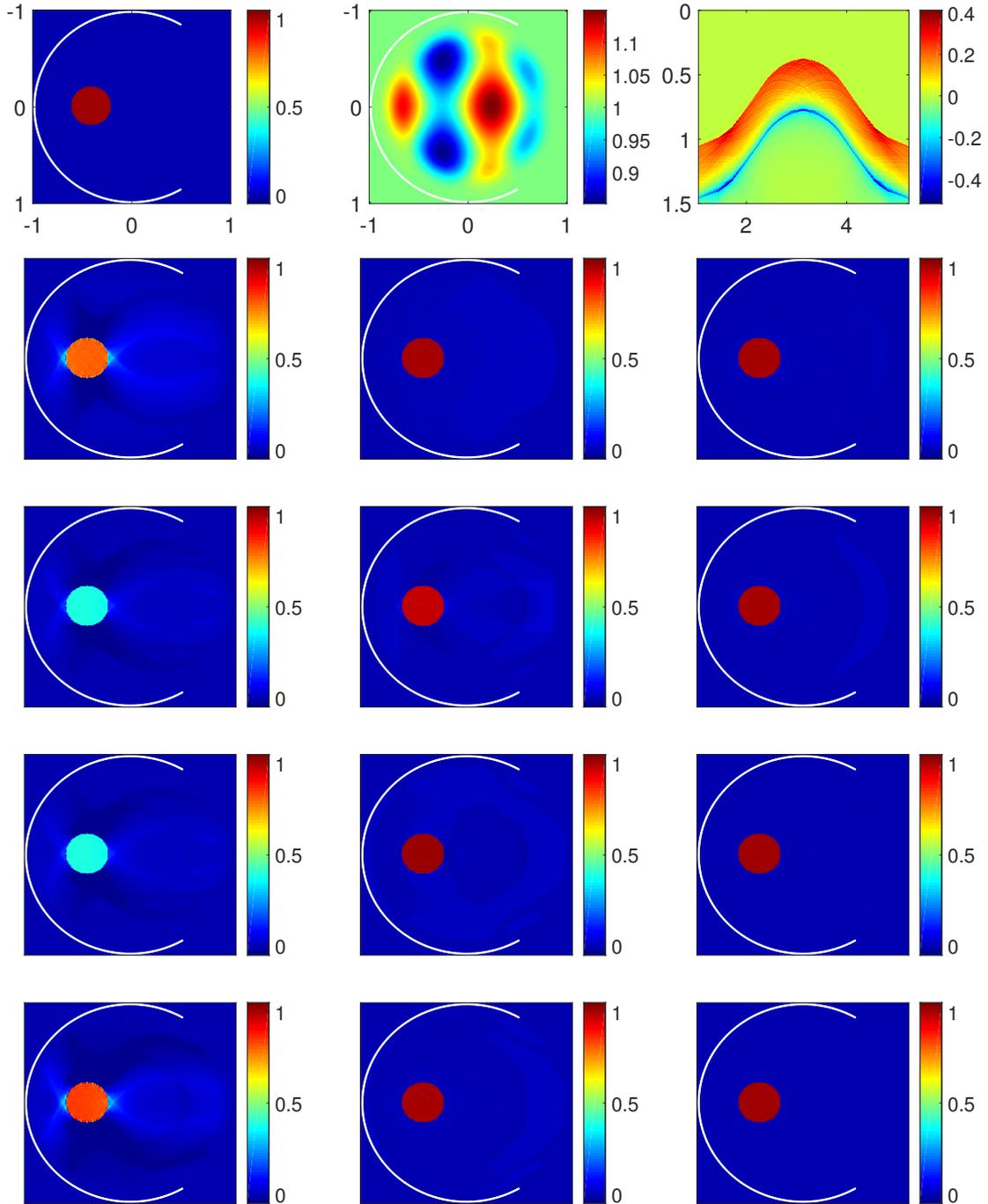


Figure 4.5: TEST CASE (T2): PARTIAL DATA, VISIBLE PHANTOM. Row 1: Initial pressure data f (left), non-trapping sound speed c (middle), and computed pressure data $\mathbf{L}f$ (right). Row 2: Iterative time reversal (after 1, 10 and 200 iterations). Row 3: Landweber’s method (after 1, 10 and 200 iterations). Row 4: Nesterov’s method (after 1, 10 and 200 iterations). Row 5: the CG method (after 1, 10 and 200 iterations). The white dots indicate the measurement curve.

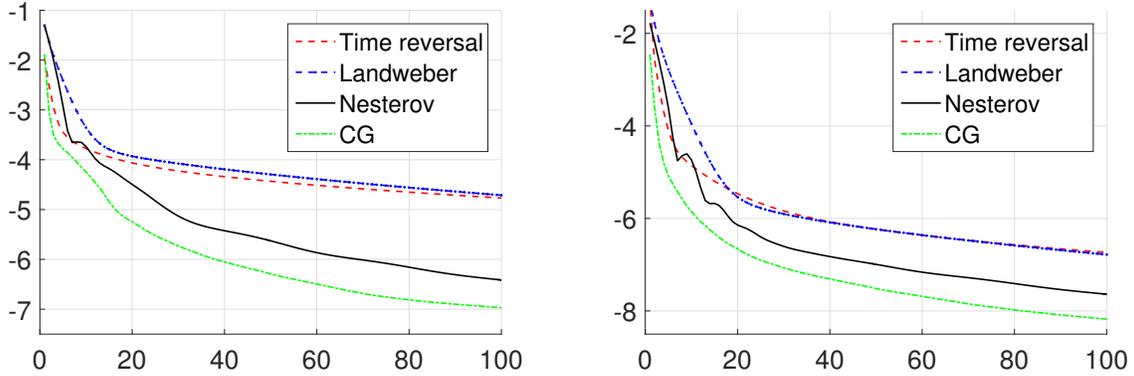


Figure 4.6: TEST CASE (T2), CONVERGENCE BEHAVIOR FOR EXACT DATA. Left: Logarithm of squared reconstruction error $\|\mathbf{f}_n - \mathbf{f}\|_2^2$ in dependence of the iteration number. Left: Logarithm of residual $\|\mathbf{L}_{N,M}\mathbf{f}_n - \mathbf{g}\|_2^2$ in dependence of the iteration number.

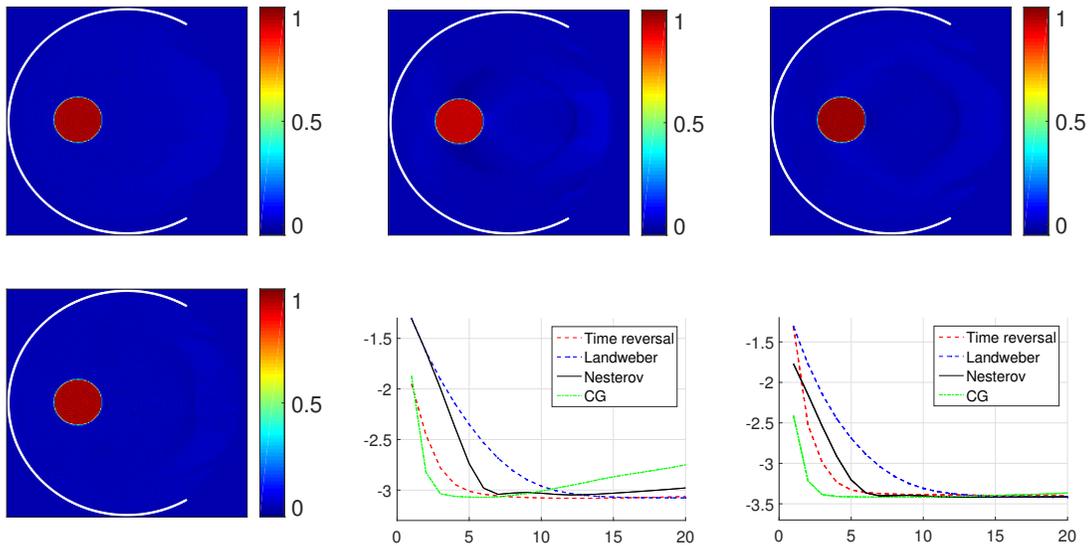


Figure 4.7: TEST CASE (T2), INEXACT DATA. Top left: Iterative time reversal, Top center: Landweber’s method, top right: Nesterov’s method. Bottom left: the CG method (all after 10 iterations). Bottom center: squared error. Bottom right: squared residuals.

Again we repeated the simulations with inexact data where we simulated the data on a different grid and further added Gaussian noise to the data. Due the ill-posedness of the problem we cannot expect complete convergence for noisy data. In fact, as can be seen in Figure 4.13 all iterations show the typical semi-convergence behavior: The error decreases until a certain optimal index n^* , after which the error starts to increase. Stopping the iteration (for example with Morozovs discrepancy principle) yields approximate but stable solutions. Incorporating additional regularization could further improve the results. Such investigations, however, are beyond the scope of this paper.

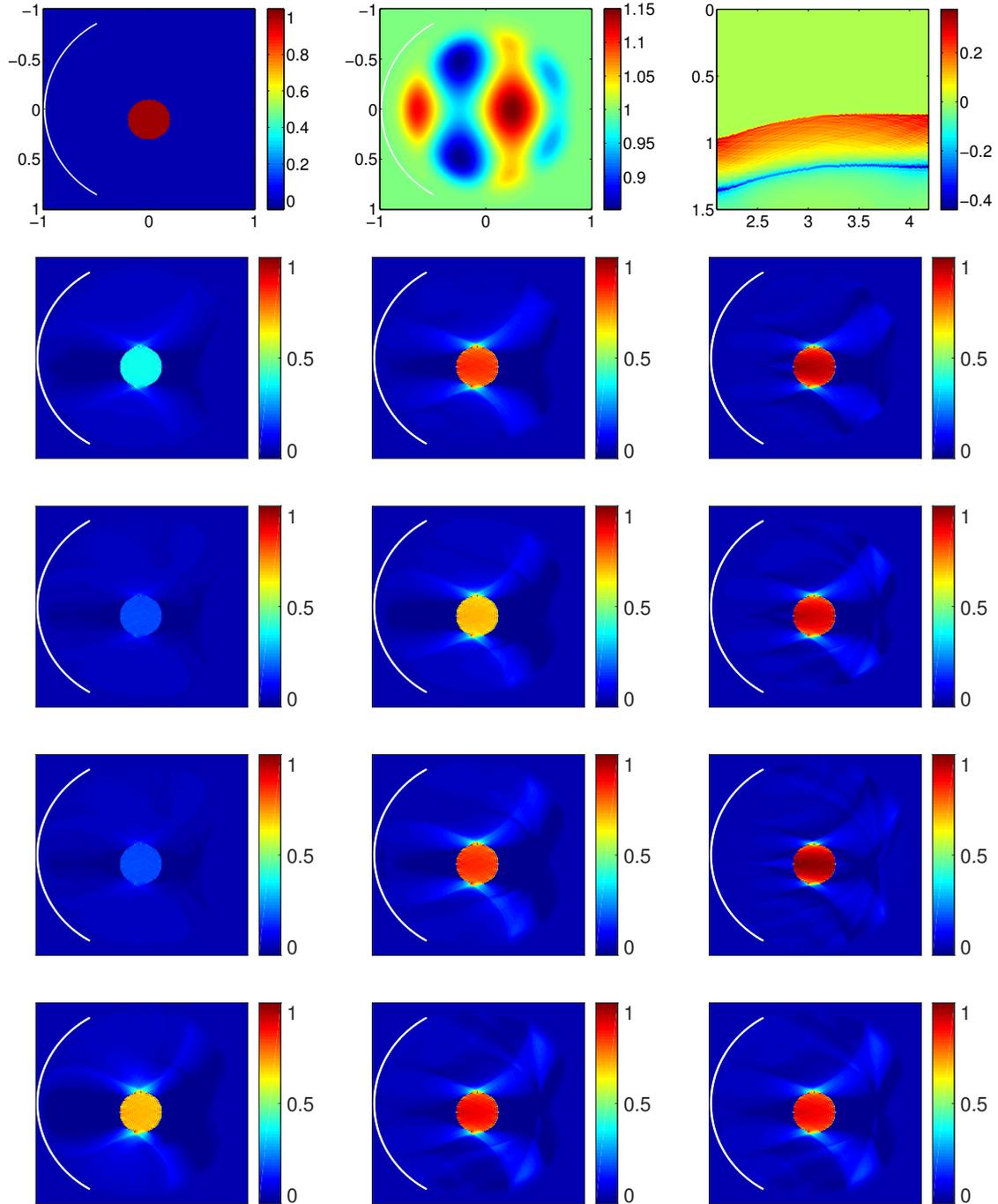


Figure 4.8: TEST CASE (T3): PARTIAL DATA, INVISIBLE PHANTOM Row 1: Non-trapping sound speed c (left), initial pressure data f (middle), and computed pressure data $\mathbf{L}f$ (right). The white dots indicate the measurement surface. Row 2: Iterative time reversal (after 1, 10 and 200 iterations). Row 3: Landweber's method (after 1, 10 and 200 iterations). Row 4: Nesterov's method (after 1, 10 and 200 iterations). Row 5: the CG method (after 1, 10 and 200 iterations).

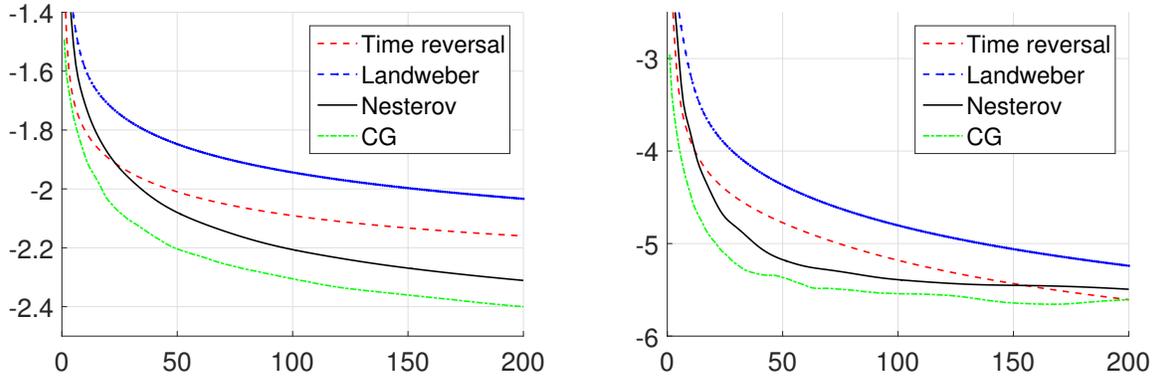


Figure 4.9: TEST CASE (T3), CONVERGENCE BEHAVIOR FOR EXACT DATA Left: Logarithm of squared reconstruction error $\|\mathbf{f}_n - \mathbf{f}\|_2^2$ in dependence of the iteration number. Left: Logarithm of residual $\|\mathbf{L}_{N,M}\mathbf{f}_n - \mathbf{g}\|_2^2$ in dependence of the iteration number.

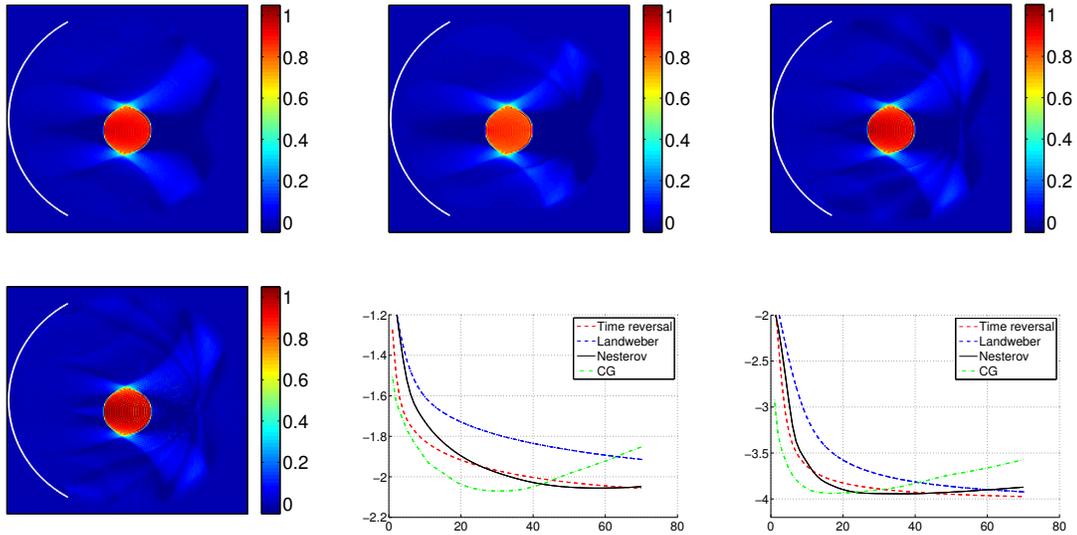


Figure 4.10: TEST CASE (T3), INEXACT DATA. Top left: Iterative time reversal, Top center: Landweber’s method, top right: Nesterov’s method. Bottom left: the CG method (all after 30 iterations). Bottom center: squared error. Bottom right: squared residuals.

4.6 Test case (T4): Trapping speed, complete data

Finally we consider the trapping sound speed (again taken from [52])

$$c(x) = 1 + 0.8w(x) \sin(2\pi x_1) \cos(2\pi x_2),$$

where $w: \mathbb{R}^2 \rightarrow [0, 1]$ is a smooth function that vanishes outside $B_1(0)$ and is equal to one on $B_{1/2}(0)$. The phantom f , the trapping sound speed c , and the corresponding data are illustrated in the top row in Figure 4.11. Reconstructions with the iterative time reversal, the Landweber’s, Nesterov’s, and the CG methods are shown in rows 2 to 5 in Figure 4.11. The squared iteration error and the squared residuals are shown in Figure 4.12. Again one notices that all methods

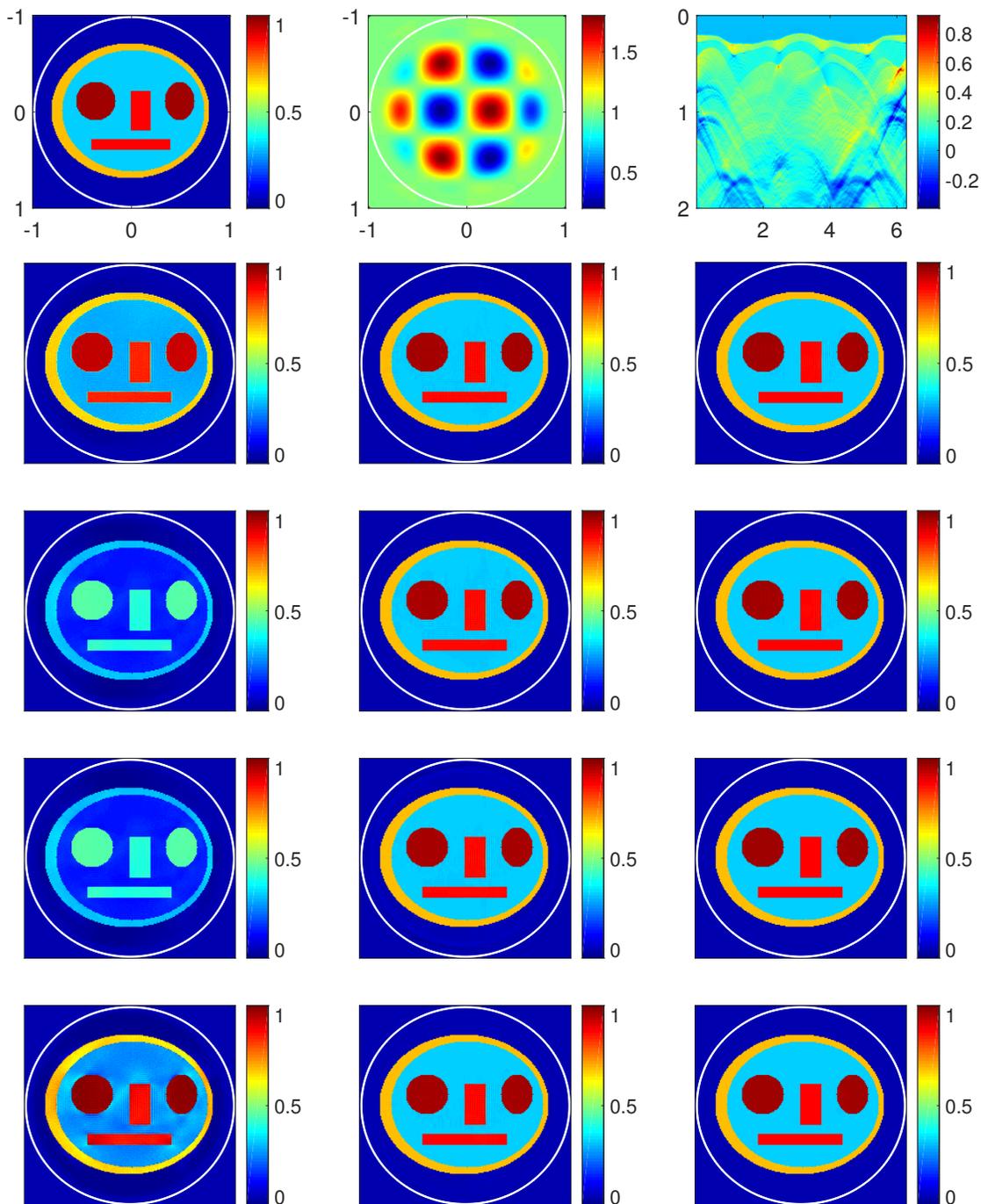


Figure 4.11: TEST CASE (T4): TRAPPING SPEED, COMPLETE DATA. Row 1: Trapping sound speed c (left), initial pressure data f (middle), and computed pressure data $\mathbf{L}f$ (right). The white dots indicate the measurement surface. Row 2: Iterative time reversal (after 1, 10 and 200 iterations). Row 3: Landweber's method (after 1, 10 and 200 iterations). Row 4: Nesterov's method (after 1, 10 and 200 iterations). Row 5: the CG (after 1, 10 and 200 iterations).

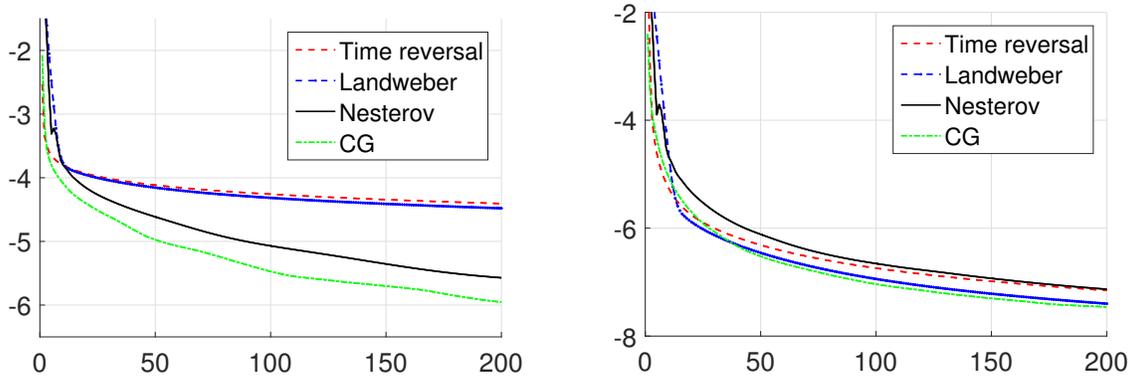


Figure 4.12: TEST CASE (T4), CONVERGENCE BEHAVIOR FOR EXACT DATA Left: Logarithm of squared reconstruction error $\|\mathbf{f}_n - \mathbf{f}\|_2^2$ in dependence of the iteration number. Left: Logarithm of residual $\|\mathbf{L}_{N,M}\mathbf{f}_n - \mathbf{g}\|_2^2$ in dependence of the iteration number.

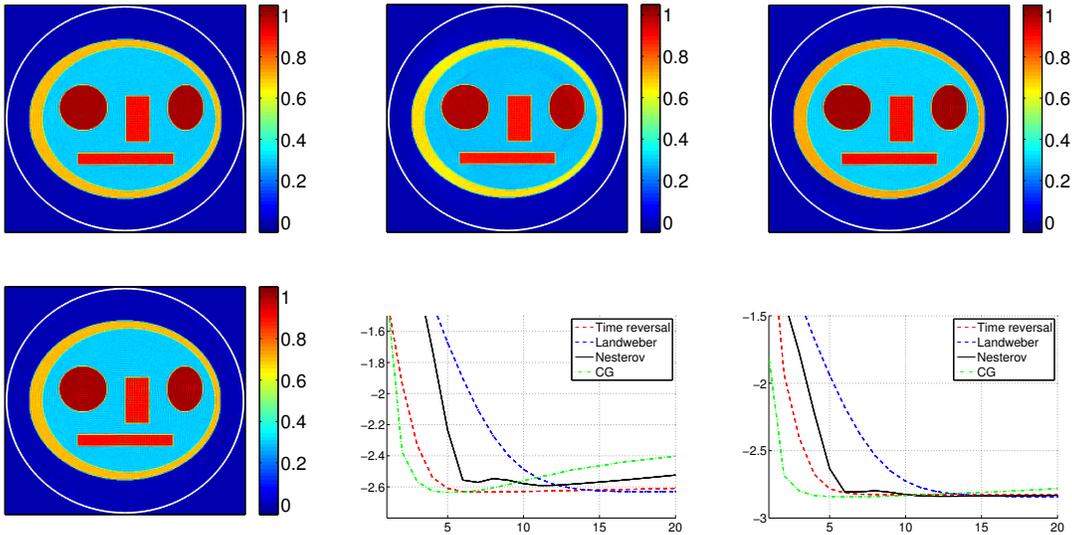


Figure 4.13: TEST CASE (T4), INEXACT DATA. Top left: Iterative time reversal, Top center: Landweber method, top right: Nesterov's method. Bottom left: CG iteration (all after 10 iterations). Bottom center: squared error. Bottom right: squared residuals.

converge to the original phantom quite fast. Again the CG method is the most rapidly converging method.

As in the previous cases we repeated the computations with inexact data \mathbf{g}^δ . Therefore we simulated the data on a different grid (using $N = 350$ and $M = 1300$) and additionally Gaussian white noise with a standard deviation equal to 5% the L^2 -norm of the exact data. The total error in the data is $\delta = 0.0490$ compared to the L^2 -norm $\|\mathbf{g}\|_2 = 0.4457$ of the exact data. The results for noisy data are shown in Figure 4.13. For any of the iterative methods, we observe a semi-convergence behavior. The minimal error reached by all iterative methods is approximately the same. The fastest method is again the CG method. Finally, we note that the residuals after 10 iterations are 0.0386 for iterative time reversal, 0.0433 for the Landweber's, 0.0386 for Nesterov's,

and 0.0384 for the CG method. For any method they are already below the noise level 0.0490, which shows that Morozov's discrepancy principle yields a well defined stopping index $n_* \leq 10$ for any considered iterative method.

5 Conclusion and outlook

In this paper we derived, analyzed, and implemented iterative algorithms for PAT with variable sound speed. We considered the full and partial data situation as well as non-trapping and trapping sound speed. In most of the cases under consideration, Landweber's method performs as well as the iterative time reversal method while Nesterov's and CG method converge faster. Landweber's and Nesterov's methods are convenient for regularization, which we will investigate in an upcoming work.

Acknowledgement

Linh Nguyen's research is partially supported by the NSF grants DMS 1212125 and DMS 1616904. He also thanks University of Innsbruck for financial support and hospitality during his visit.

References

- [1] S. ACOSTA AND C. MONTALTO, *Multiwave imaging in an enclosure with variable wave speed*, Inverse Problems, 31 (2015), p. 065009.
- [2] M. AGRANOVSKY AND P. KUCHMENT, *Uniqueness of reconstruction and an inversion procedure for thermoacoustic and photoacoustic tomography with variable sound speed*, Inverse Problems, 23 (2007), p. 2089.
- [3] S. R. ARRIDGE, M. M. BETCKE, B. T. COX, F. LUCKA, AND B. E. TREEBY, *On the adjoint operator in photoacoustic tomography*, Inverse Problems, 32 (2016), p. 115012 (19pp).
- [4] O. AXELSSON AND J. KARÁTON, *On the rate of convergence of the conjugate gradient method for linear operators in hilbert space*, Numerische Mathematik, 48 (2002), pp. 499–523.
- [5] L. L. BARANNYK, J. FRIKEL, AND L. V. NGUYEN, *On Artifacts in Limited Data Spherical Radon Transform: Curved Observation Surface*, Inverse Problems, 32 (2015).
- [6] Z. BELHACHMI, T. GLATZ, AND O. SCHERZER, *A direct method for photoacoustic tomography with inhomogeneous sound speed*, Inverse Problems, 32 (2016), p. 045005.
- [7] P. BURGHOLZER, C. HOFER, G. PALTAUF, G. MATT, M. HALTMEIER, AND O. SCHERZER, *Thermoacoustic tomography using a fiber based fabry-perot interferometer as an integrating line detector*, in Proc. of SPIE Vol, vol. 6086, 2006, pp. 60861N–1.
- [8] P. BURGHOLZER, G. J. MATT, M. HALTMEIER, AND G. PALTAUF, *Exact and approximate imaging methods for photoacoustic tomography using an arbitrary detection surface*, Physical Review E, 75 (2007), p. 046706.
- [9] C. CLASON AND M. V. KLIBANOV, *The quasi-reversibility method for thermoacoustic tomography in a heterogeneous medium*, SIAM Journal on Scientific Computing, 30 (2008), pp. 1–23.

- [10] B. COX, S. KARA, S. ARRIDGE, AND P. BEARD, *k-space propagation models for acoustically heterogeneous media: Application to biomedical photoacoustics*, The Journal of the Acoustical Society of America, 121 (2007), pp. 3453–3464.
- [11] J. W. DANIEL, *The conjugate gradient method for linear and nonlinear operator equations*, SIAM Journal on Numerical Analysis, 4 (1967), pp. 10–26.
- [12] X. L. DEAN-BEN, A. BUEHLER, V. NTZIACHRISTOS, AND D. RAZANSKY, *Accurate model-based reconstruction algorithm for three-dimensional optoacoustic tomography*, IEEE Transactions on Medical Imaging, 31 (2012), pp. 1922–1928.
- [13] H. W. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of inverse problems*, vol. 375, Springer Science & Business Media, 1996.
- [14] O. G. ERNST, *Minimal and orthogonal residual methods and their generalizations for solving linear operator equations*, 2000. Habilitation thesis, TU Bergakademie Freiberg.
- [15] D. FINCH, M. HALTMEIER, AND RAKESH, *Inversion of spherical means and the wave equation in even dimensions*, SIAM Journal on Applied Mathematics, 68 (2007), pp. 392–412.
- [16] D. FINCH, S. K. PATCH, AND RAKESH, *Determining a function from its mean values over a family of spheres*, SIAM Journal on Mathematical Analysis, 35 (2004), pp. 1213–1240 (electronic).
- [17] Z. FORTUNA, *Some convergence properties of the conjugate gradient method in hilbert space*, SIAM Journal on Numerical Analysis, 16 (1979), pp. 380–384.
- [18] J. FRIKEL AND E. T. QUINTO, *Artifacts in incomplete data tomography with applications to photoacoustic tomography and sonar*, SIAM Journal on Applied Mathematics, 75 (2015), pp. 703–725.
- [19] R. GLOWINSKI AND S. LAPIN, *Iterative solution of linear variational problems in hilbert spaces: some conjugate gradients success stories*, in *Conjugate Gradient Algorithms and Finite Element Methods*, Springer, 2004, pp. 223–245.
- [20] M. HALTMEIER, *Inversion of circular means and the wave equation on convex planar domains*, Computers & Mathematics with Applications. An International Journal, 65 (2013), pp. 1025–1036.
- [21] M. HALTMEIER, *Universal inversion formulas for recovering a function from spherical means*, SIAM Journal on Mathematical Analysis, 46 (2014), pp. 214–232.
- [22] M. HALTMEIER, O. SCHERZER, P. BURGHOLZER, AND G. PALTAUF, *Thermoacoustic computed tomography with large planar receivers*, Inverse Problems, 20 (2004), p. 1663.
- [23] M. HANKE, *Conjugate gradient type methods for ill-posed problems*, vol. 327, CRC Press, 1995.
- [24] R. M. HAYES, *Iterative methods of solving linear problems on hilbert space*, Bulletin of the American Mathematical Society, 58 (1952), pp. 653–653.
- [25] R. HERZOG AND E. SACHS, *Superlinear convergence of krylov subspace methods for self-adjoint problems in hilbert space*, SIAM Journal on Numerical Analysis, 53 (2015), pp. 1304–1324.

- [26] E. HESTENES, M. R. AND STIEFEL, *Methods of conjugate gradients for solving linear systems*, Journal of Research of the National Bureau of Standards, 49 (1952).
- [27] L. HÖRMANDER, *The analysis of linear partial differential operators. I*, Classics in Mathematics, Springer-Verlage, Berlin, 2003.
- [28] Y. HRISTOVA, *Time reversal in thermoacoustic tomography—an error estimate*, Inverse Problems, 25 (2009), pp. 055008, 14.
- [29] Y. HRISTOVA, P. KUCHMENT, AND L. NGUYEN, *Reconstruction and time reversal in thermoacoustic tomography in acoustically homogeneous and inhomogeneous media*, Inverse Problems, 24 (2008), pp. 055006, 25.
- [30] C. HUANG, K. WANG, L. NIE, AND M. A. WANG, L. V. AND ANASTASIO, *Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media*, IEEE Transactions on Medical Imaging, 32 (2013), pp. 1097–1110.
- [31] J. KAIPIO AND E. SOMERSALO, *Statistical and computational inverse problems*, vol. 160, Springer Science & Business Media, 2006.
- [32] W. J. KAMMERER AND M. Z. NASHED, *On the convergence of the conjugate gradient method for singular linear operator equations*, SIAM Journal on Numerical Analysis, 9 (1972), pp. 165–181.
- [33] P. KUCHMENT, *The Radon transform and medical imaging*, vol. 85, SIAM, 2014.
- [34] P. KUCHMENT AND L. KUNYANSKY, *Mathematics of thermoacoustic tomography*, European Journal of Applied Mathematics, 19 (2008), pp. 191–224.
- [35] L. A. KUNYANSKY, *Explicit inversion formulae for the spherical mean Radon transform*, Inverse Problems, 23 (2007), pp. 373–383.
- [36] L. A. KUNYANSKY, *A series solution and a fast algorithm for the inversion of the spherical mean radon transform*, Inverse Problems, 23 (2007), p. S11.
- [37] I. LASIECKA, J.-L. LIONS, AND R. TRIGGIANI, *Nonhomogeneous boundary value problems for second order hyperbolic operators*, J. Math. Pures Appl, 65 (1986), pp. 149–192.
- [38] A. K. LOUIS AND E. T. QUINTO, *Local tomographic methods in sonar*, in Surveys on solution methods for inverse problems, Springer, Vienna, 2000, pp. 147–154.
- [39] K.-A. MARDAL AND R. WINTHER, *Preconditioning discretizations of systems of partial differential equations*, Numerical Linear Algebra with Applications, 18 (2011), pp. 1–40.
- [40] T. D. MAST, L. P. SOURIAU, D. D. LIU, M. TABELI, A. I. NACHMAN, AND R. C. WAAG, *A k -space method for large-scale models of wave propagation in tissue*, IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 48 (2001), pp. 341–354.
- [41] D. MODGIL, M. ANASTASIO, AND P. LA RIVIÈRE, *Image reconstruction in photoacoustic tomography with variable speed of sound using a higher-order geometrical acoustics approximation*, Journal of Biomedical Optics, 15 (2010), p. 021308.

- [42] V. A. MOROZOV, *Methods for solving incorrectly posed problems*, Springer Science & Business Media, 2012.
- [43] F. NATTERER, *Photo-acoustic inversion in convex domains*, Inverse Problems Imaging, (2012).
- [44] Y. E. NESTEROV, *A method for solving the convex programming problem with convergence rate $o(1/k^2)$* , Doklady Akademii Nauk, 269 (1983), pp. 543–547. in Russian.
- [45] O. NEVANLINNA, *Convergence of iterations for linear equations*, Birkhäuser, 2012.
- [46] L. V. NGUYEN, *A family of inversion formulas in thermoacoustic tomography*, Inverse Probl. Imaging, 3 (2009), pp. 649–675.
- [47] L. V. NGUYEN, *On singularities and instability of reconstruction in thermoacoustic tomography*, Tomography and inverse transport theory, Contemporary Mathematics, 559 (2011), pp. 163–170.
- [48] L. V. NGUYEN, *On artifacts in limited data spherical radon transform: Flat observation surfaces*, SIAM J. Math. Analysis, 47 (2015), pp. 2984–3004.
- [49] L. V. NGUYEN AND L. A. KUNYANSKY, *A dissipative time reversal technique for photoacoustic tomography in a cavity*, SIAM Journal on Imaging Sciences, 9 (2016), pp. 748–769.
- [50] V. P. PALAMODOV, *A uniform reconstruction formula in integral geometry*, Inverse Probl., 28 (2012), p. 065014.
- [51] G. PALTAUF, P. BURGHOLZER, M. HALTMEIER, AND O. SCHERZER, *Thermoacoustic tomography using optical line detection*, in European Conference on Biomedical Optics, Optical Society of America, 2005.
- [52] J. QIAN, P. STEFANOV, G. UHLMANN, AND H. ZHAO, *An efficient neumann series-based algorithm for thermoacoustic and photoacoustic tomography with variable sound speed*, SIAM Journal on Imaging Sciences, 4 (2011), pp. 850–883.
- [53] A. ROSENTHAL, V. NTZIACHRISTOS, AND D. RAZANSKY, *Acoustic inversion in optoacoustic tomography: A review*, Current medical imaging reviews, 9 (2013), p. 318.
- [54] P. STEFANOV AND G. UHLMANN, *Thermoacoustic tomography with variable sound speed*, Inverse Problems, 25 (2009), pp. 075011, 16.
- [55] P. STEFANOV AND G. UHLMANN, *Thermoacoustic tomography arising in brain imaging*, Inverse Problems, 27 (2011), p. 045004.
- [56] P. STEFANOV AND Y. YANG, *Multiwave tomography with reflectors: Landweber’s iteration*, ArXiv e-prints, (2016).
- [57] M. TABELI, T. D. MAST, AND R. C. WAAG, *A k -space method for coupled first-order acoustic propagation equations*, Journal of the Acoustical Society of America, 111 (2002), pp. 53–63.
- [58] B. E. TREEBY AND B. T. COX, *k -wave: Matlab toolbox for the simulation and reconstruction of photoacoustic wave fields*, Journal of biomedical optics, 15 (2010), pp. 021314–021314.

- [59] F. TRÈVES, *Introduction to pseudodifferential and Fourier integral operators. Vol. 2*, Plenum Press, New York, 1980. Fourier integral operators, The University Series in Mathematics.
- [60] K. WANG, R. SU, A. A. ORAEVSKY, AND M. A. ANASTASIO, *Investigation of iterative image reconstruction in three-dimensional optoacoustic tomography*, Physics in medicine and biology, 57 (2012), p. 5399.
- [61] M. XU, L. V. WANG, G. AMBARTSOUMIAN, AND P. KUCHMENT, *Limited view thermoacoustic tomography*, in Photoacoustic imaging and spectroscopy, CRC Press, 2009, pp. pp. 61–73.
- [62] G. ZANGERL, O. SCHERZER, AND M. HALTMEIER, *Circular integrating detectors in photo and thermoacoustic tomography*, Inverse Probl. Sci. Eng, 17 (2009), pp. 133–142.