An Efficient Neumann Series–Based Algorithm for Thermoacoustic and Photoacoustic Tomography with Variable Sound Speed*

Jianliang Qian[†], Plamen Stefanov[‡], Gunther Uhlmann[§], and Hongkai Zhao[¶]

- **Abstract.** We present an efficient algorithm for reconstructing an unknown source in thermoacoustic and photoacoustic tomography based on the recent advances in understanding the theoretical nature of the problem. We work with variable sound speeds that also might be discontinuous across some surface. The latter problem arises in brain imaging. The algorithmic development is based on an explicit formula in the form of a Neumann series. We present numerical examples with nontrapping, trapping, and piecewise smooth speeds, as well as examples with data on a part of the boundary. These numerical examples demonstrate the robust performance of the Neumann series–based algorithm.
- Key words. thermoacoustic tomography, photoacoustic tomography, inverse problems, Neumann series, variable sound speed

AMS subject classifications. 65Y10, 47G30

DOI. 10.1137/100817280

1. Introduction. Thermoacoustic tomography (TAT) and photoacoustic (PAT) tomography are emerging medical imaging modalities [35, 33]. These are hybrid medical imaging methods that combine the high resolution of acoustic waves with the large contrast of optical waves. TAT and PAT have been developed to overcome the limitations of both conventional ultrasound and microwave imaging. The physical principle underlying TAT and PAT is the photoacoustic effect, which can be roughly described as follows. A short impulse of electromagnetic microwaves or light is sent through a patient's body. The tissue heats up slightly, and the heat expansion generates weak acoustic waves. These waves are measured away from the patient's body, and one tries to recover the acoustic source that gives information about the rate of absorption at each point in the body, thus creating an image. One of the potential applications for TAT is early breast cancer detection. The American Cancer Society reports that breast cancer is the second overall leading cause of death among women in the United States. The mortality rate from breast cancer has declined in recent years due to progress in both

^{*}Received by the editors December 7, 2010; accepted for publication (in revised form) April 25, 2011; published electronically September 15, 2011.

http://www.siam.org/journals/siims/4-3/81728.html

[†]Department of Mathematics, Michigan State University, East Lansing, MI 48824 (qian@math.msu.edu). This author's work was partially supported by NSF 0810104, NSF 0830161, and NSF 1115363.

[‡]Mathematics Department, Purdue University, West Lafayette, IN 47907 (stefanov@math.purdue.edu). This author's work was partially supported by NSF grant DMS-0800428 and a Simons Research Professorship.

[§]Department of Mathematics, University of Washington, Seattle, WA 98195, and University of California, Irvine, CA 92697-3875 (gunther@math.washington.edu). This author's work was partially supported by NSF, a Chancellor Professorship at UC Berkeley, and a Senior Clay Award.

[¶]Department of Mathematics, University of California, Irvine, CA 92697-3875 (zhao@math.uci.edu). This author's work was partially supported by NSF 0811254.

early detection and more effective treatment. Better detection techniques, however, are still needed. The significance of TAT and PAT is that they yield images of high electromagnetic contrast at high ultrasonic resolution in relatively large volumes of biological tissues [35, 33].

A first step in TAT and PAT is to reconstruct the amount of deposited energy from timedependent boundary measurement of acoustic signals. To start with, we describe here the widely accepted mathematical model of TAT and PAT [35, 33, 8]. Let $\Omega \subset \mathbb{R}^n$ be an open set with a smooth strictly convex boundary. Let c(x) > 0 be the sound speed that is either smooth or piecewise smooth. Assume that c(x) = 1 outside Ω . The acoustic pressure u(t, x)then solves the wave equation

(1)
$$\begin{cases} (\partial_t^2 - c^2 \Delta) u = 0 & \text{in } (0, T) \times \mathbb{R}^n, \\ u|_{t=0} = f, \\ \partial_t u|_{t=0} = 0, \end{cases}$$

where T > 0 is fixed and f(x) is a source that we want to recover supported in Ω . The measurements are modeled by the operator

(2)
$$\Lambda f := u|_{[0,T] \times \partial \Omega}.$$

Given Λf , the problem is to reconstruct the unknown f that is related to the absorbing properties of the body at any point.

There have been significant advances in both mathematical theories and medical applications; see [1, 8, 9, 11, 12, 14, 15, 18, 19, 21, 25, 30, 31, 35, 38, 39, 16] and the references therein. Theoretically, one is interested in uniqueness and stability of the solution for the inverse problem; numerically, one is interested in designing efficient numerical algorithms to recover the solution of the inverse problem. Naturally, the above two aspects have been well studied in the case of the sound speed being constant. In fact, if the sound speed is constant and the observation surface $\partial\Omega$ is of some special geometry, such as a planar, spherical, or cylindrical surface, there are explicit closed-form inversion formulas; see [8, 34, 11, 12, 7] and the references therein. In practice there are many cases when the constant sound speed model is inaccurate [35, 16, 37, 20]. For instance, in breast imaging, the different components of the breast, such as the glandular tissues, stromal tissues, cancerous tissues, and other fatty issues, have different acoustic properties. The variations between their acoustic speeds can be as great as 10% [16].

To tackle variable sound speeds in TAT and PAT, the time-reversal method has been suggested in [8] and used in [36, 10, 14, 15]. However, only when the dimension is odd and the sound speed is constant, does the time-reversal method give an exact reconstruction for Tlarge enough by the Huygens principle. When the dimension is even or the sound speed is not constant, the time-reversal method yields only an approximate reconstruction as $T \gg 1$. A natural question arises immediately: is there any exact reconstruction formula which is able to handle both a variable sound speed and an irregular observation geometry for a fixed T? It turns out that the work [30] summarized below does provide such a formula and this formula is the foundation for our algorithmic development. Some results about variable sound speeds are contained in the references above, but a more complete analysis in this case has been given in [30]. This analysis includes if-and-only-if conditions for uniqueness and stability, including the case with observations on a part of the boundary. An explicit recovery formula of a type of convergent Neumann series is derived in [30] when Λf is known on the whole boundary $\partial\Omega$ and T is greater than the stability threshold. This formula is the base for our numerical reconstruction, and we also apply it if the speed is trapping.

There is another situation related to a variable discontinuous sound speed which arises in brain imaging [38, 22, 6]. The skull has a discontinuous sound speed which is piecewise smooth with jump-discontinuities across the boundary of the skull. A typical situation is that the speed in the skull is about twice as fast as that in the soft tissue; see, e.g., [38, 39]. Such speeds drastically change the way that singularities propagate; see Figure 1. In [31], the second and third authors studied this model and proved that the Neumann series expansion works as well, provided that all singularities issued from $\sup f$ have a path of nondiffractive segments that reaches $\partial \Omega$ until time T. Based on the work in [31], we develop an efficient numerical algorithm as well to handle the situation of variable discontinuous sound speeds in PAT and TAT.

The rest of the paper is organized as follows. In section 2, we define the energy spaces and provide some facts about propagation of singularities for the wave equation. In sections 3 and 4 we summarize the theoretical results for smooth sound speed in [30] for cases of both complete and partial data and explain the consequences of the theory. Section 5 describes the theoretical results in [31] when the sound speed is discontinuous. In section 6 we make comparisons between smooth and nonsmooth sound speeds in terms of uniqueness and stability. The iterative algorithm for constructing the Neumann series so that the inversion formula in [30] can be implemented is presented in section 7. In sections 8, 9, and 10, we give extensive examples to demonstrate the robustness of the Neumann series—based algorithm for PAT and TAT.

2. Preliminaries. Assume for now that c > 0 is smooth. The speed c defines a Riemannian metric $c^{-2}dx^2$. For any piecewise smooth curve $\gamma : t \in [a, b] \mapsto \gamma(t) \in \mathbb{R}^n$, the length of γ in that metric is given by

length(
$$\gamma$$
) = $\int_{a}^{b} \frac{|\dot{\gamma}(t)|}{c(\gamma(t))} dt$.

The so-defined length is independent of the parameterization of γ . The distance function $\operatorname{dist}(x, y)$ is then defined as the infimum of the lengths of all such curves connecting x and y.

For any $(x, \theta) \in \mathbb{R}^n \times S^{n-1}$ we denote by $\gamma_{x,\theta}(t)$ the unit speed (i.e., $|\dot{\gamma}| = c(\gamma)$) geodesics issued at x in the direction θ .

Similar to the settings in [30, 31], the energy of u(t, x) in a domain $U \subset \mathbb{R}^n$ is given by

$$E(u(t)) = \int_{U} \left(|\nabla_x u|^2 + c^{-2} |u_t|^2 \right) \mathrm{d}x,$$

where $u(t) = u(t, \cdot)$. The energy of any Cauchy data (f, g) for (1) is given by

$$E(f,g) = \int_{U} \left(|\nabla_x f|^2 + c^{-2} |g|^2 \right) \mathrm{d}x.$$

The energy norm is defined as the square root of the energy. In particular, the energy of (f, 0) in U is given by the square of the Dirichlet norm

$$||f||_{H_D(U)}^2 := \int_U |\nabla_x f|^2 \, \mathrm{d}x,$$

853

where the Hilbert space $H_D(U)$ is the completion of $C_0^{\infty}(U)$ under the above Dirichlet norm. We always assume below that the initial condition $f \in H_D(\Omega)$. We will denote by $\|\cdot\|$ the norm in $H_D(\Omega)$, and in the same way we denote the operator norm in that space.

There are two main geometric quantities that are crucial for the results below. First we set

(3)
$$T_0 := \max\{\operatorname{dist}(x, \partial\Omega) : x \in \overline{\Omega}\},\$$

where $\operatorname{dist}(x, \partial \Omega)$ is the distance in the given Riemannian metric $c^{-2} dx^2$. Let $T_1 \leq \infty$ be the supremum of the lengths of all maximal geodesics lying in $\overline{\Omega}$. Clearly, $T_0 < T_1$; however, while the first number is always finite, the second can be infinite. It actually can be shown that

$$(4) T_0 \le T_1/2$$

One of the main ingredients of our approach in [30, 31] was to understand the microlocal nature of the problem. We recall the definition of a wave front set of a function, or, more generally, a distribution; see [13]. The definition is based on the known property of the Fourier transform: one can tell whether a compactly supported function f is smooth by looking at the decay of the Fourier transform $\hat{f}(\xi)$ as $|\xi| \to \infty$: $f \in C_0^{\infty}$ if and only if $|\hat{f}(\xi)| \leq C_N (1+|\xi|)^{-N}$ for any N. The idea behind the wave front set is to localize this near a fixed x_0 and in a conic neighborhood of a fixed $\xi_0 \neq 0$. Conic neighborhoods are defined as open conic sets, i.e., sets of the type $\Gamma = \{r\theta; r > 0, \theta \in V\}$, where V is an open subset of S^{n-1} . We say that $(x_0, \xi_0) \notin WF(f), \xi_0 \neq 0$, if there exist $\phi \in C_0^{\infty}$ with $\phi(x_0) \neq 0$ and a conical neighborhood Γ of ξ_0 , so that

$$|\widehat{\phi f}(\xi)| \le C_N (1+|\xi|)^{-N} \quad \forall \xi \in \Gamma, \ \forall N.$$

If $(x,\xi) \in WF(f)$, we say that (x,ξ) is a singularity of f, or that f is singular at (x,ξ) . Since singularities are defined by conic sets, we can restrict ξ to unit vectors. For example, the Dirac Delta function $\delta(x)$ has a wave front set at x = 0 and all directions, i.e., $WF(\delta) =$ $\{(0,\xi); \xi \neq 0\}$, and a piecewise smooth function f that has a jump across some smooth surface S (and nowhere else) is singular at all points of S in (co)normal directions, i.e., $WF(f) = \{(x,\xi); x \in S, 0 \neq \xi \perp S \text{ at } x\}$.

The propagation of singularities for the wave equation (1) can be described as follows. If $(x,\theta) \in WF(f)$, then at time t, both $(\gamma_{x,\theta}(t), \dot{\gamma}_{x,\theta}(t))$ and $(\gamma_{x,\theta}(-t), \dot{\gamma}_{x,\theta}(-t))$ are in WF $(u(t, \cdot))$, where u is the solution of (1). This is due to the fact that the symbol of the wave operator is $\tau^2 - c^2 |\xi|^2$, which has two characteristic roots, $\tau = \pm c(x)|\xi|$, and that the initial velocity in (1) is zero; therefore each singularity splits into two equal parts starting to propagate in opposite directions. While this is a classical result in the linear PDE theory, we refer the reader to [30] for more details in this specific case.

3. Smooth speed and data on the whole $\partial \Omega$. We will describe below the theoretical results in [30].

3.1. Uniqueness. If $T \gg 1$, Λf recovers f uniquely. We have the following sharp result based on the unique continuation theorem by Tataru [32].

Theorem 1. Let $\Lambda f = 0$. Then f(x) = 0 for dist $(x, \partial \Omega) \leq T$. Moreover, f can be arbitrary in the set dist $(x, \partial \Omega) > T$ if the latter set is nonempty.

Corollary 1. Λ is injective on $H_D(\Omega)$ if and only if $T \ge T_0$. We refer to [30] for proofs.

3.2. Stability. We showed in [30] that Λf recovers f in a stable way if each singularity (x,ξ) , i.e., each element of the wave front set WF(f), reaches $\partial\Omega$ for time t (positive or negative) such that |t| < T. In other words, if functions f are a priori supported in a fixed compact $\mathcal{K} \subset \overline{\Omega}$, then we have the following equivalent statement:

(5) For any $(x, \theta) \in \mathcal{K} \times S^{n-1}$, the unit speed geodesic through (x, θ) at t = 0reaches $\partial \Omega$ at time |t| < T.

Moreover, the following condition is sufficient regardless of the choice of \mathcal{K} :

(6)
$$T_1/2 < T_1/2$$

This condition is equivalent to (5) if $\mathcal{K} = \overline{\Omega}$.

Furthermore, to formulate a condition equivalent to (5) by taking \mathcal{K} into account, we define $T_1 = T_1(\mathcal{K})$ as above, which is explicitly related to geodesics that pass through \mathcal{K} .

If $T_1 = \infty$, then the sound speed c is called trapping (in Ω). In this case, there is no stability regardless of the choice of T.

We summarize the above discussion in the following.

Theorem 2. Let $\mathcal{K} \subset \Omega$ be compact.

(a) Let $T > T_1(\mathcal{K})/2$. Then there exists a constant C > 0 so that

$$||f|| \le C ||\Lambda f||_{H^1([0,T] \times \partial \Omega)}.$$

(b) Let $T < T_1(\mathcal{K})/2$. Then for any C > 0, s_1 , and s_2 , there is $f \in C^{\infty}$ supported in \mathcal{K} so that

$$||f||_{H^{s_1}} \ge C ||\Lambda f||_{H^{s_2}([0,T] \times \partial \Omega)}.$$

In other words, $T > T_1(\mathcal{K})/2$ is a sufficient and necessary condition for stability for functions f supported in \mathcal{K} up to replacing the < sign by \leq .

Visible and invisible singularities. Condition (5) can be explained in the following way. As a general principle, for a stable recovery in such a linear inverse problem, we need to detect all singularities. We refer the reader to [29] for more details. As explained above, each singularity starts to travel in a positive and a negative direction because $u_t = 0$ at t = 0, so it can leave two traces on $\partial\Omega$. It is enough to detect one of them for stability. On the other hand, if we can detect both, we can expect better numerical results. Condition (5) then says that all singularities are visible at the boundary.

We want to emphasize that T_1 can be much larger than $\operatorname{diam}(\Omega) := \max\{\operatorname{dist}(x, y); (x, y) \in \partial\Omega \times \partial\Omega\}$. If there are no conjugate points in $\overline{\Omega}$, then those two quantities coincide. If there are conjugate points, however, those two quantities do not necessarily coincide. If there are closed geodesics, then $T_1 = \infty$, while $\operatorname{diam}(\Omega)$ is finite. In this case, although we always have uniqueness for some $T \gg 1$, we never have stability. Finally, we notice that while T_0 and

 $\operatorname{diam}(\Omega)$ can always be estimated analytically or numerically, T_1 is much harder to estimate or even to tell whether it is finite or not. In any case,

(7)
$$\operatorname{diam}(\Omega) \leq T_1.$$

If $\gamma_{x,\theta}$ does not hit the boundary at time $|t| \leq T$, we call (x, θ) an invisible (possible) singularity. It is easy to show that if one such pair exists, then there is a nonempty open set of invisible singularities.

3.3. Reconstruction. The reconstruction method in [30] is based on the following ideas. If we knew the Cauchy data (u, u_t) on $\{T\} \times \Omega$, we could just solve a mixed problem like the one below with that Cauchy data on t = T and boundary data given by Λf . Although we do not know (u, u_t) on $\{T\} \times \Omega$, we do know the boundary values of u on $\partial\Omega$ for t = T, i.e., u on $\{T\} \times \partial\Omega$. Assuming that [f, 0] is in the energy space, i.e., $f \in H_D(\Omega)$, we can only say that $u_t(T, \cdot)$ is in $L^2(\Omega)$, and its boundary values on $\{T\} \times \Omega$, we choose the one that minimizes the energy norm $\|\cdot\|_{H_D(\Omega)}$. By the Dirichlet principle, the function is given by the harmonic extension of $u(T, \cdot)|_{\partial\Omega}$.

Consequently, given h defined on $[0,T] \times \partial \Omega$ which eventually will be replaced by Λf , we first solve the elliptic boundary value problem

(8)
$$\Delta \phi = 0, \quad \phi|_{\partial\Omega} = h(T, \cdot),$$

and we introduce the notation P_{Ω} for the Poisson operator of harmonic extension: $P_{\Omega}h(T, \cdot) := \phi$. In fact, the equation is $c^2\Delta u = 0$, but c^2 cancels out. Then we perform the modified back-projection

(9)
$$\begin{cases} (\partial_t^2 - c^2 \Delta) v = 0 & \text{in } (0, T) \times \Omega, \\ v|_{[0,T] \times \partial \Omega} = h, \\ v|_{t=T} = P_{\Omega} h(T, \cdot), \\ v_t|_{t=T} = 0. \end{cases}$$

Note that the initial data at t = T satisfy compatibility conditions of first order (no jump at $\{T\} \times \partial \Omega$). Then we define the following left pseudoinverse of Λ :

(10)
$$Ah := v(0, \cdot) \quad \text{in } \bar{\Omega}.$$

The operator A is not an actual inverse (unless n is odd, c is constant, and T is greater than the diameter), and we have

$$A\Lambda = I - K,$$

where K is an "error" operator. We showed in [30] that

(11)
$$\|Kf\|_{H_D(\Omega)} \le \|f\|_{H_D(\Omega)} \quad \forall f \in H_D(\Omega),$$

for any smooth speed, trapping or not, and for any time T > 0. If $T > T_0$, the inequality is strict; see [30]. To show that K is a contraction requires some assumptions, however.

Theorem 3. (a) Let c be nontrapping, and let $T > T_1/2$. Then $A\Lambda = I - K$, where $\|K\|_{H_D(\Omega) \to H_D(\Omega)} < 1$. In particular, I - K is invertible on $H_D(\Omega)$, and the inverse thermoacoustic problem has an explicit solution of the form

(12)
$$f = \sum_{m=0}^{\infty} K^m Ah, \quad h := \Lambda f.$$

(b) Let $T > T_1$. Then in addition to the conclusions above, K is compact in $H_D(\Omega)$.

For a proof, see [30, Theorem 1] and [31, Remark 2.2].

In other words, under condition (6), we have not only stability but an explicit solution in the form of a convergent Neumann series as well. On the other hand, if (6) fails, i.e., if $T < T_1/2$, there is no stability by Theorem 2. This does not mean that the series (12) will not converge in this case. If it does, it will converge to f for $T > T_0$. Indeed, then it is easy to see that the limit g solves (I - K)(g - f) = 0, and since (11) is strict in this case, f = g.

Based on this theorem and its proof, we can expect good convergence when $T > T_1$, and the first term would already be a good approximation of the high frequency part of f.

If $T_1/2 < T < T_1$, we can expect the first term in the series to recover only a fraction of the high frequency part of f and the successive terms to improve this gradually; the series (12) would still converge, but that convergence would be slower.

If $T < T_1/2$, then ||K|| = 1. Indeed, if we assume that ||K|| < 1, we would get uniform convergence of the Neumann series and stability; on the other hand, there is no stability in this case.

3.4. Summary: Dependence on T.

(i) $T < T_0$.

 Λf does not recover f uniquely; see Theorem 1. Then ||K|| = 1, and for any f supported in the inaccessible region, Kf = f.

(ii) $T_0 < T < T_1/2$.

This can happen only if there is a strict inequality in (4). Then we have uniqueness but not stability. In this case, ||K|| = 1, ||Kf|| < ||f||, and we do not know if the Neumann series (12) converges. If it does, it converges to f.

(iii) $T_1/2 < T < T_1$.

This assumes that Ω is nontrapping for c. The Neumann series (12) converges exponentially but maybe not as fast as in the next case. There is stability, and ||K|| < 1. (iv) $T_1 < T$.

This also assumes that Ω is nontrapping for c. The Neumann series (12) converges exponentially. There is stability, ||K|| < 1, and K is compact.

3.5. Recovery of singularities. In cases (iii) and (iv), we explicitly recover the whole f, including its singularities. If the goal is to recover only the singularities of f (the wave front set WF(f)) with less computation, then one can perform the classical back-projection (time reversal) as follows; see [36, 10, 14, 15, 21].

Let $\chi \in C_0^{\infty}(\mathbb{R})$ be such that $\chi(t) = 1$ for $t \in [0, T]$, and $\chi(t) = 0$ near t = T', where T' > T is "close enough" to T. Set

(13)
$$Rf = A\chi\Lambda f,$$

where A is as above, but the time reversal is done from t = T' to t = 0. Since $\chi \Lambda = 0$ near t = T', the harmonic extension is actually zero in this case, and we have the classical time reversal. In general, Rf is not close to f unless $T \to \infty$; see [14].

In case (iv), R is a parametrix of infinite order; see [30]. Therefore, it correctly recovers all singularities, including jumps across smooth surfaces—it will correctly recover the location and the size of the jump.

In case (iii), R is elliptic but not a parametrix itself; see [30, Theorem 3]. Then Rf will have the singularities at the right places, but the amplitudes will be in general between 1/2 and 1.

In cases (i) and (ii), only singularities "close enough to the boundary" will be recovered with amplitudes between 1/2 and 1.

Finally, if the speed is trapping in Ω , i.e., if $T_1 = \infty$, then Rf recovers the visible singularities up to time T by choosing χ appropriately.

These comments are just another way to formulate [30, Theorem 3].

3.6. Comparison with the time-reversal method. Let $T_1 < \infty$ first; i.e., assume that c is nontrapping. Let the time-reversal approximation be defined as in (13) with $\chi = 1$ in a neighborhood of $[0, T_1]$. Then R is a parametrix of infinite order, i.e., Rf = f - Qf, where $Qf \in C^{\infty}$ for any f. On the other hand, ||Q|| is not necessarily small for any fixed T. Assuming that χ is properly chosen, as $T \to \infty$, this norm gets smaller at a rate dictated by the local energy decay for the wave equation: $||Q|| = O(t^{1-n})$ for n even and $||Q|| = O(e^{-Ct})$ for n odd; see [14]. Therefore, for $T \gg 1$ so that ||Q|| < 1, one can write

$$Rf = (\mathbf{I} - Q)f,$$

and one can solve this equation by a Neumann series as above. However, it is not straightforward to impose sharp conditions on χ and T to guarantee ||Q|| < 1. On the other hand, for the method that we propose, that condition is $T > T_1$ and it is sharp for a stable inversion. Also, the proposed method minimizes the norm of the "error" operator, and when both Neumann expansions converge uniformly, the one in Theorem 3 will converge faster in the uniform topology. Numerical experiments not shown here confirm this.

If $T_1 = \infty$ (*c* is trapping), then the error in the time reversal-method decays like $O(1/\log T)$ if $f \in H_0^2(\Omega)$, and the error decays even slower if $f \in H_D(\Omega)$ only; see [4]. The first term of the Neumann series inversion has an error no greater than this, and numerically the error improves with a few more terms. We do not know whether the Neumann series converges or not, however.

4. Smooth speed and data on a part of $\partial \Omega$. Let $\Gamma \subset \partial \Omega$ be a relatively open set of $\partial \Omega$, and assume that we only have data available on $[0,T] \times \Gamma$. We will suppose that f is supported in some compact set $\mathcal{K} \subset \Omega$.

4.1. Uniqueness. As in (3), set

(14)
$$T_0 := T_0(\mathcal{K}, \Gamma) = \max\{\operatorname{dist}(x, \Gamma) : x \in \mathcal{K}\}.$$

Theorem 1 has the following analogue in this case.

Theorem 4. Let $\Lambda f = 0$ on $[0,T] \times \Gamma$. If $T \ge T_0$, then f = 0. If $T < T_0$, then f = 0 on $\mathcal{K} \cap \{x; \operatorname{dist}(x,\Gamma) < T\}$ and can be arbitrary in the complement of this set in \mathcal{K} .

The proof of this theorem is not easy. It combines Tataru's uniqueness theorem with arguments that first appeared in [8] in the case of constant speeds and were extended later in [30] to variable speeds.

As a corollary, $\Lambda f|_{[0,T]\times\Gamma}$ uniquely determines f if and only if $T \geq T_0$.

4.2. Stability. The following condition guarantees that we can detect all singularities originating from \mathcal{K} as singularities of our data:

(15) For any
$$(x, \theta) \in \mathcal{K} \times S^{n-1}$$
, the unit speed geodesic through (x, θ) at $t = 0$
reaches Γ at time $|t| < T$.

Let $T_1/2 \leq \infty$ be the maximum of all such times. The following condition is then equivalent to (15):

(16) $T_1/2 < T.$

Theorem 5. Let $\mathcal{K} \subset \Omega$ be compact.

(a) Let $T > T_1/2$. Then there exists a constant C > 0 so that

$$||f|| \le C ||\Lambda f||_{H^1([0,T] \times \Gamma)}.$$

(b) Let $T < T_1/2$. Then for any C > 0, s_1 , and s_2 , there is $f \in C^{\infty}$ supported in \mathcal{K} so that

 $||f||_{H^{s_1}} \ge C ||\Lambda f||_{H^{s_2}([0,T] \times \Gamma)}.$

In other words, $T > T_1/2$ is a sufficient and necessary condition for stability for functions f supported in \mathcal{K} up to replacing the < sign by \leq . Part (a) is proved in [30], while part (b) follows from [28].

4.3. Reconstruction. An explicit formula of the type (12) is not available in this case, but one can show that the recovery is reduced to a Fredholm equation if (16) holds.

Let $\chi \in C^{\infty}([0,T] \times \Gamma)$ be a cutoff function supported in $[0,T] \times \Gamma$ so that $\chi = 1$ on a slightly smaller set $[0,T'] \times \Gamma'$ that still satisfies the condition. Assume also that $0 \le \chi \le 1$. Then we know $\chi \Lambda f$. Apply the time-reversal operator A to this to get $A\chi \Lambda f$. Since $\chi(\cdot,T) = 0$, the harmonic extension is zero, so this is the classical time reversal. Let K be the "error operator"

$$K = I - A\chi\Lambda.$$

To find f, we need to solve

(17)
$$(\mathbf{I} - K)f = h, \quad h := A\chi\Lambda f$$

with h determined by the data. By [30, Theorem 3], $A\chi\Lambda$ is a pseudodifferential operator that is elliptic under the stability condition (15). Its principal symbol is given by

$$\frac{1}{2}\chi(\gamma_{x,\xi}(\tau_+(x,\xi))) + \frac{1}{2}\chi(\gamma_{x,\xi}(\tau_-(x,\xi))),$$

where $\tau_{\pm}(x,\xi)$ is the positive/negative time of the (unit speed) geodesic $\gamma_{x,\xi}$ through (x,ξ) to reach $\partial\Omega$. Then K is also a pseudodifferential operator with the principal symbol

(18)
$$\sigma_p(K) = 1 - \frac{1}{2}\chi(\gamma_{x,\xi}(\tau_+(x,\xi))) - \frac{1}{2}\chi(\gamma_{x,\xi}(\tau_-(x,\xi))).$$

Since $0 \le \chi \le 1$, and by the stability condition, at least one of the χ terms above is equal to 1/2. Therefore,

$$0 \le \sigma_p(K) \le \frac{1}{2}.$$

Away from the set $\{0 < \chi < 1\}$, $\sigma_p(K)$ is either 0 or 1/2. This set is relatively small in applications because we do not want to lose data but not too small in order to keep $|\partial_t \chi| + |\partial_x \chi|$ not too large. By the Garding inequality, one can express K in the form

$$K = K_1 + K_2, \quad ||K_1|| \le \frac{1}{2}, \quad K_2 \text{ compact},$$

and if $\Gamma \neq \emptyset$, one can show that actually $||K_1|| = 1/2$. This does not allow us to claim that the Neumann series (12) converges. However, by [24], since the distance from K to the compact operators is less than one, for a fixed f,

(19)
$$\sum_{m=0}^{\infty} K^m A \chi \Lambda f$$

converges if and only if $K^m A \chi \Lambda f \to 0$ as $m \to \infty$. It is trivial to see that the limit g solves

(20)
$$A\chi\Lambda g = A\chi\Lambda f.$$

While we know that $\chi\Lambda$ is injective for $T > T_0$, this is not true for A, so we cannot draw the conclusion that g = f. The original equation $\chi\Lambda f = h$ with h in the range of Λ is Fredholm with a trivial kernel, but (20) for g is Fredholm which may have a nontrivial kernel.

On the other hand, the partial sums

(21)
$$g_N := \sum_{m=0}^N K^m A \chi \Lambda f$$

used in the reconstruction below recover singularities of f asymptotically as $N \to \infty$. This follows from the standard pseudodifferential calculus.

For the purpose of the numerical reconstruction, we take χ to be a function of $x \in \partial \Omega$ only, independent of t. In particular, we no longer have $\chi = 0$ near t = T. Then we define A as before, but this time the harmonic extension of $\chi \Lambda f(T, \cdot)$ is not trivial in general. Then we use partial sums as in (21).

If the stability condition (16) is not satisfied, then inverting $\chi \Lambda$ is not equivalent to solving a Fredholm equation anymore and it is unstable. One can show that K is the sum of an operator with norm not exceeding 1 and a compact operator. Of course, this does not guarantee convergence of the Neumann series, and the partial sums asymptotically recover the visible singularities only. In the numerical example below, we check the norm of each successive term in the partial sum (21) and stop when that norm starts to increase.



Figure 1. Rays in the case of a discontinuous sound speed. Dotted lines represent negative times.

5. Discontinuous sound speed: Modeling brain imaging. We will briefly review the results in [31]. Let c(x) be piecewise smooth with a nonzero jump across one or several smooth closed nonintersecting surfaces that we call S in Ω . We still assume that c = 1 outside Ω . The formulation of the problem is still the same, and classical solutions u(t, x) are assumed to be in C^1 . This implies the following transmission conditions across S where c(x) jumps: the limits of u(t, x) and its normal derivative match as x approaches S from either side.

5.1. Propagation of singularities. It is well known that this case is quite different from the previous one from the viewpoint of propagation of singularities. Let us see what happens when a "ray" (a geodesic) approaches S from one of its sides, which we call an interior one. Let c_{-} be the limit of c on S from inside, and let c_{+} be the limit from outside.

If $c_{-} > c_{+}$, this ray splits into two parts when hitting S. One reflects according to the usual laws of reflection and goes back to the interior. The other one goes into the exterior and refracts by changing its angle with respect to the boundary. The incoming and outgoing angles α_{\pm} with the normal satisfy Snell's law:

(22)
$$\frac{\sin \alpha_{-}}{\sin \alpha_{+}} = \frac{c_{-}}{c_{+}}$$

Assume now that $c_{-} < c_{+}$. Then there is a critical angle $0 < \alpha_{0} < \pi$ with the normal at any point so that if $\alpha_{-} < \alpha_{0}$, there are still a reflected and a transmitted (refracted) ray as above satisfying Snell's law. If $\alpha_{-} > \alpha_{0}$, then there is no refracted ray, while the reflected one still exists. This is known as a full internal reflection. This critical angle can be computed from Snell's law: it is the value of α_{-} that forces $\sin \alpha_{+}$ to be equal to 1, and therefore to be greater than 1 when $\alpha_{-} > \alpha_{0}$. Therefore, $\alpha_{0} = \sin^{-1}(c_{-}/c_{+})$. Propagation of singularities when $\alpha_{-} = \alpha_{0}$ is more delicate and will not be analyzed here.

After splitting or not into two rays, each segment may split into two, etc. Assuming that we never get rays tangent to S, reflection and refraction occur in the same way as above. Figure 1 shows a possible time evolution of a single singularity, both for positive and negative time. The sound speed in the "skull" is higher than that on either side.

We define the distance function dist(x, y) as the infimum of the length (in the metric $c^{-2}dx^2$) of all piecewise smooth curves connecting x and y that intersect S transversely.

5.2. Uniqueness. Theorem 1 and Corollary 1 still hold in this case; see [31, Proposition 5.1] and its proof. The definition of T_0 in this case is the same.

5.3. Stability. Visible and invisible singularities. The general principle that we should be able to detect each singularity for stable recovery still applies. However, in the case of full internal reflection, there might be singularities that never reach $\partial\Omega$. For example, let Ω be the ball B(0, R) (centered at 0 with radius R), and let S be the sphere $|x| = R_1, R_1 < R$. Assume also that inside the sphere S, the speed is a constant less than that outside S, where it is also a constant. Then all rays coming from the interior of S, hitting S at an angle smaller than the critical one, will reflect and give rise to no transmitted rays. By the rotational symmetry, those reflected rays will hit S again at the same angle, fully reflect, etc. Such singularities will be invisible, no matter what T we choose. This is similar to the trapping situation in the case of a smooth sound speed. Of course, we may also have rays that are trapped but never reach the interface S. Therefore the singularities that are certain to be visible up to time T consist of the following set:

(23)
$$\mathcal{U} = \{ (x,\theta) \in (\Omega \setminus S) \times S^{n-1}; \text{ there is a path of the "geodesic" issued from either} \\ (x,\theta) \text{ or } (x,-\theta) \text{ at } t = 0 \text{ never tangent to } S, \text{ that is outside } \bar{\Omega} \text{ at time } t = T \}.$$

If we restrict our attention to f a priori supported in some compact $\mathcal{K} \subset \Omega \setminus S$, then the stability condition can be formulated as follows:

(24)
$$\forall (x,\theta) \in \mathcal{K} \times S^{n-1}$$
, there is a path of the "geodesic" issued from either (x,θ) or $(x,-\theta)$ at $t=0$ never tangent to S , that is outside $\overline{\Omega}$ at time $t=T$.

The reconstruction formula below shows that this condition is sufficient not only for stability but also for an explicit convergent formula of the type (12).

Similar to the smooth case, we can define $T_1/2$ (instead of defining T_1 directly) to be the infimum of all T satisfying (24) when $\mathcal{K} = \overline{\Omega}$. Then (6) remains a necessary and sufficient condition (up to replacing $\langle by \leq \rangle$) for stability in this case as well.

5.4. Reconstruction. In [31], we also analyzed the energy at high frequencies that is carried by the reflected and the transmitted rays. We showed that if none of the incident, reflected, and transmitted rays are tangent to S, then a positive fraction of the energy reflects and transmits at high frequencies. Therefore, even though under condition (24) the corresponding singularity is visible at $\partial\Omega$, only a fraction of it would be measured. This fraction could be quite small, depending on the speed, the number of reflections and refractions before the ray hits $\partial\Omega$, and the angles there. So we should not expect the first term in the series (12), even if it converges, to be a good approximation even at high frequencies. Theorem 6 proved in [31], shows that an analogue of (12) still converges but needs to be modified first.

The needed modification is connected to the following problem. When applying A to Λf with supp $f \subset \mathcal{K}$, we get $A\Lambda f$ that may be supported everywhere in Ω . The successive terms in (12) would then force $A\Lambda$ to be applied to the result. To "restrict" $A\Lambda f$ to \mathcal{K} , we cannot

just restrict in the usual sense, because this may take the results out of the energy space (no zero trace on $\partial \mathcal{K}$). For this reason, we project orthogonally on $H_D(\mathcal{K})$. It turns out that the orthogonal projection is given by $\Pi_{\mathcal{K}} f := f - P_{\mathcal{K}}(f|\partial \mathcal{K})$, where $P_{\mathcal{K}}$ is the Poisson operator of harmonic extension from $\partial \mathcal{K}$ to \mathcal{K} defined in (8).

Theorem 6. Let \mathcal{K} satisfy (24). Then $\Pi_{\mathcal{K}}A\Lambda = I - K$ in $H_D(\mathcal{K})$, with $\|K\|_{H_D(\mathcal{K})} < 1$. In particular, I - K is invertible on $H_D(\mathcal{K})$, and Λ restricted to $H_D(\mathcal{K})$ has an explicit left inverse of the form

(25)
$$f = \sum_{m=0}^{\infty} K^m \Pi_{\mathcal{K}} Ah, \quad h = \Lambda f.$$

Note that the theorem does not say how to reconstruct f when $\operatorname{supp} f$ is not in a set \mathcal{K} satisfying (24). Estimate (11) still holds, but we do not know if I - K can be inverted by a Neumann series. In the numerical examples, however, we work with f supported everywhere in Ω , and then $\Pi_{\mathcal{K}} = \Pi_{\overline{\Omega}} = I$.

6. Comparison between the smooth and the nonsmooth case. If the sound speed c is smooth, then there is always uniqueness for large enough T because $T_0 < \infty$ and $T > T_0/2$ suffices for uniqueness. On the other hand, there is stability and the Neumann series converges to a solution only when c is nontrapping and $T > T_1/2$. In the trapping case, one still has stability and an explicit solution of the type in Theorem 3 under the a priori assumptions that $\operatorname{supp} f \subset \mathcal{K}$ and all singularities with base points $x \in \mathcal{K}$ are visible. The assumption $\operatorname{supp} f \subset \mathcal{K}$ means that we need to know f outside \mathcal{K} ; then we can subtract that part from fand apply the reconstruction procedure.

When c is of jump type, we may still have rays that are trapped even if they never reach S. On the other hand, S may split some rays. Under the stability condition (24) which is equivalent to (6) with the modified definition of T_1 , we still have stability and an explicit solution for $\operatorname{supp} f \subset \mathcal{K}$. We still need to know f outside \mathcal{K} for a full reconstruction. The essential difference is that even though all singularities except those (of measure zero) contributing to tangent rays on S will be detected in the best possible case, $T_1 < \infty$ and $T > T_1/2$, they might be detected with a loss of energy even when $T \gg 1$. In the example presented in Figure 1, the singularity that exits at the top has been split twice before that, and each such event takes away a fraction of the energy. In contrast, in the case of a smooth sound speed, such a singularity will exit for t > 0 with half of the energy of f, while the other half will exit for a negative time; as such, for example, we will detect both in case (iv). Going back to the nonsmooth case, the Neumann series (25) will gradually restore the right strength of the singularity, and the whole f, actually, but we can expect this to be more sensitive to noise and computational errors.

7. Algorithmic formulations. The main issue in implementing the above Neumann series formulation is how to compute the operator K. In the case of data available on the whole $\partial\Omega$, by definition, $K = I - A\Lambda$. Thus, given a function ψ ,

$$K\psi = \psi - A\Lambda\psi.$$

Thus we need an efficient computational algorithm to carry out the actions of time-reversal operator A and the measurement operator Λ . Since the measurement operator Λ can be simulated by forward modeling, we will detail the implementation of forward modeling first.

7.1. Complex scaling/perfectly matched layers for the acoustic wave equation. Since the forward model equation (1) is formulated as a pure initial value problem, we have to truncate the computational domain to be finite. The truncated domain has to be large enough to enclose the domain where the measurements are taken along its boundary. On the other hand, we have to impose some artificial boundary conditions on the boundary of the truncated domain. To perform accurate long-time simulation, we will adopt the complex scaling method (see [27] and the references therein), also known as perfectly matched layers (PML) as an absorbing boundary condition for acoustic waves [2, 5, 23]. This is done so that waves will not be reflected into the computational domain.

To derive the PML for the equation, we follow [23]. We rewrite the second-order equation as a first-order system by introducing the two-component velocity vector $\mathbf{v}(x, y, t) = (v_x(x, y, t), v_y(x, y, t))$:

(26)
$$\frac{\partial \mathbf{v}}{\partial t} = -\nabla u$$

(27)
$$\frac{\partial u}{\partial t} = -c^2 \nabla \cdot \mathbf{v}.$$

By using the complex coordinate stretching [2, 5, 23], we have the following equations:

(28)
$$\frac{\partial v_{\eta}}{\partial t} + \omega_{\eta} v_{\eta} = -\frac{\partial u}{\partial \eta},$$

(29)
$$\frac{\partial u^{(\eta)}}{\partial t} + \omega_{\eta} u^{(\eta)} = -c^2 \frac{\partial v_{\eta}}{\partial \eta},$$

where $\eta = x, y, u = u^{(x)} + u^{(y)}$, and ω_{η} represents a loss in the PML and is zero in a regular non-PML region.

In our simulation, we will take the following $\omega_{\eta}: [0,1] \to R^+$:

$$\omega_{\eta}(s) = \begin{cases} \frac{b}{\sigma} \left(\frac{s-\sigma}{\sigma}\right)^2, & s \in [0,\sigma], \\ 0, & s \in [\sigma, 1-\sigma], \\ \frac{b}{\sigma} \left(\frac{s-1+\sigma}{\sigma}\right)^2, & s \in [1-\sigma, 1], \end{cases}$$

where $\sigma > 0$ and b > 0 are some appropriate constants. Depending on the computational domain, this function will be rescaled accordingly.

Equations (28) and (29) are discretized by a staggered finite-difference scheme, and the details are omitted; see [23].

7.2. T_0 and the time reversal operator A. To compute the critical time T_0 , we first solve the eikonal equation

- (30) $c|\nabla \tilde{T}| = 1,$
- (31) $\tilde{T}|_{\Gamma} = 0$

by using the fast sweeping finite-difference scheme as designed in [40, 17, 26], where Γ can be the whole boundary $\partial\Omega$ or a part of $\partial\Omega$. Then $T_0 = \max\{\tilde{T}\}$.

The Poisson equation defining the harmonic extension is solved by a V-cycle multigrid method [3], which converges with the optimal rate independent of the mesh size.

The time-reversal wave equation (9) with terminal and boundary conditions is solved by a standard second-order finite-difference time-domain scheme.

7.3. Partial data. When the data are available only on a part of the boundary, we impose vanishing boundary conditions on the inaccessible boundary. To have overall compatible boundary conditions, we apply a smooth taping function (varying from 1 to 0) to transition from available data to artificial vanishing data.

7.4. Discontinuous sound speeds. In our current formulation, we use finite-difference methods on Cartesian meshes to simplify our implementation. As such, we will design our meshes so that interfaces of discontinuous sound speed are located on one of the Cartesian mesh lines. Therefore, in our current work, the discontinuous sound speeds shown in our numerical examples allow only some special geometry.

Nevertheless, let us emphasize that the methodology presented in this work can be equally applied to discontinuous sound speeds with irregular geometry because we easily replace a finite-difference method with a finite-element method for wave equations which allows us to treat arbitrary geometry. Such a line of research is under way, and we will report it in an upcoming paper.

8. Numerical results: Smooth speed, data on the whole $\partial \Omega$. In this section we show two-dimensional numerical examples to validate our algorithms.

Our computational domain is taken to be $[-1.5, 1.5]^2$, and $\Omega = [-1.28, 1.28]^2$. We work with three sound speeds:

(32)
$$c_1(x,y) = 1.0 + 0.2\sin(2\pi x) + 0.1\cos(2\pi y),$$

(33)
$$c_2(x,y) = \frac{9(x^2+y^2)}{1+9(x^2+y^2)} + \exp\left(-90(x^2+y^2)\right) - 0.4\exp\left(-10(3\sqrt{(x^2+y^2)}-2)^2\right),$$

(34) $c_3(x,y) = 1.25 + \sin(2.0\pi x)\cos(2\pi y).$

See Figure 2, with some geodesics plotted. We also use a smooth cutoff for a smooth transition of those speeds to 1 near $\partial\Omega$ and outside Ω .

In the numerical examples presented in the following sections, we use mesh sizes $\Delta x = \Delta y = 0.01$ to discretize the domain into a 301×301 mesh. The PML are distributed around the ring $[-1.5, 1.5]^2 \setminus [-1.4, 1.4]^2$. To solve the wave equation, we impose the Courant–Friedrichs–Lewy condition to decide the time step according to given mesh sizes. In our examples, we take $\Delta t = 0.5\Delta x/\max\{c\}$, where c is the sound speed.

In all examples below, we use the abbreviation NS for the Neumann series, and TR for time reversal. We also mention that the (modified) TR corresponds to the first term (10) of the NS solution. We take a finite number, k + 1, of terms in the NS expansion by stopping when the error gets below 5% in the stable cases and somewhat higher in the nonstable ones. Taking more terms, especially in the cases where we have stability, improves the error several



Figure 2. Sound speed models. (a) The variable nontrapping sound speed c_1 . (b) The variable radial trapping sound speed c_2 . (c) The variable trapping speed c_3 .

times, but we do not show those examples. Next, we compute and comment on only the L^2 error since all test images model non- H^1 functions. We also show a diagram of the distance function, dist $(x, \partial \Omega)$.

In practice, we might not know the exact solution that we look for. To test for the convergence of the NS method, we will monitor the size of each term in the NS, and we will stop when the sizes of successive terms do not vary too much. We will study such a convergence criterion in a future work.

8.1. Example 1: The Shepp–Logan phantom.

8.1.1. Nontrapping speed c_1 , Figures 3–5. The sound speed is given by (32); see Figure 2(a). The original image as the source function f is shown in Figure 3(b). Numerically, we estimate T_0 to be $T_0 \approx 1.1767$. A rough estimate of T_1 is $3 < T_1 < 4$. We add four bright disks (f = 1 there) to the classical Shepp–Logan phantom.

Figure 3. $T = 2T_0 \approx 2.3535$. The time T is slightly above the stability threshold $T_1/2$ but below T_1 . The error of the NS solution with k = 8 (9 terms) of the series is 6.63% vs. error 37.76% for the TR solution. Since $T < T_1$, the TR solution does not recover the correct size of the jumps—they are recovered with amplitudes ranging from 1/2 to 1, and for many of them, it is just 1/2; this is clear from the slice diagrams. In contrast, the NS solution has the right amplitudes and would improve with more terms.

Figure 4. $T = 4T_0$. The time T is doubled, and it is greater than T_1 . The error of the NS solution with k = 8 (9 terms) of the series is 4.99% vs. error 7.07% for the TR solution. The TR



Figure 3. Example 1 with the nontrapping speed c_1 . Case 1: $T = 2T_0$. (a) The boundary distance map. (b) The exact initial condition. (c) The TR solution. (d) The NS solution. (e) x-slices of the TR solution (continuous line) and the exact solution (dashed line). (f) x-slices of the NS solution (continuous line) and the exact solution (dashed line). (g) y-slices of the TR solution (continuous line) and the exact solution (dashed line). (h) y-slices of the NS solution (continuous line) and the exact solution (dashed line).



Figure 4. Example 1 with the nontrapping speed c_1 . Case 2: $T = 4T_0$. (a) The boundary distance map. (b) The exact initial condition. (c) The TR solution. (d) The NS solution. (e) x-slices of the TR solution (continuous line) and the exact solution (dashed line). (f) x-slices of the NS solution (continuous line) and the exact solution (dashed line). (g) y-slices of the TR solution (continuous line) and the exact solution (dashed line). (h) y-slices of the NS solution (continuous line) and the exact solution (dashed line).

reconstruction gets better, as expected. While the singularities are recovered correctly, the low frequency part in the TR reconstruction is still not well recovered, as evidenced in the slice diagrams. For example, in Figures 4(e) and 4(g) for the TR reconstruction there are oscillations in some flat regions; in contrast, in Figures 4(f) and 4(h) for the NS reconstruction those oscillations are gone in the same flat regions.

Figure 5. $T = 4T_0$ with 10% noise. The time T is doubled, and it is greater than T_1 . The error of the NS solution with k = 8 (9 terms) of the series is 7.07% vs. error 9.71% for the TR one. The TR reconstruction gets better as expected. While the singularities are recovered correctly, the low frequency part in the TR reconstruction is still not well recovered, as evidenced in the slice diagrams. For example, in Figures 5(e) and 5(g) for the TR reconstruction there are oscillations in some flat regions; in contrast, in Figures 5(f) and 5(h) for the NS reconstruction those oscillations are gone in the same flat regions.

8.1.2. Trapping sound speed c_3 , Figure 6. The sound speed is c_3 given by (34) and Figure 2(c). Although it is hard to show that the speed is actually trapping, one can easily construct numerically geodesics of length at least 20; therefore, based on this numerical evidence, T_1 is at least 20 and possibly ∞ . The times T that we choose are much smaller than 20/2, so there is a considerable set of singularities that are not detected.

We have $T_0 \approx 1.25$. In Figure 6, we present reconstructions with $T = 4T_0$. The phantom in the middle of the NS image (error 18.11%, k = 8) can be separated from the background, while in the TR one (error 39.56%), this is much harder. The invisible singularities in this case are distributed in a chaotic way, and what appears as noise in both images is due to the fact that there are singularities there that cannot be resolved—even if the actual image is not singular there!

8.2. Example 2: Zebras, Figures 7–11.

8.2.1. Nontrapping speed c_1 . The sound speed is given by (32), and it is visualized in Figure 2(a). Here f is now represented by the zebras image that has more complex structure.

Figure 7. $T = 4T_0$. Both methods give a good reconstruction, and with k = 2 only, the NS one has error 4.6%, which is about half of the TR one. The error can be improved significantly with more terms.

Figure 8. $T = 4T_0$ with 10% noise. Here with k = 8, the NS has an error of 6.25%, which is about half of the TR one.

8.2.2. Trapping sound speed c_3 . The sound speed is given by (34). As we indicated above, $T_0 \approx 1.23$, $T_1 > 20$, and the geodesic flow is somewhat chaotic with invisible singularities that have also chaotic distribution.

Figure 9. $T = 4T_0$. The NS reconstruction with k = 20 is much cleaner with error 9.64%, which is a bit less than 1/2 of the TR error.

Figure 10. $T = 4T_0$ with noise. The noise slightly increases the errors in both cases.

8.2.3. Trapping speed c_2 .

Figure 11. $T = 4T_0$. The sound speed is given by (33), and we estimate T_0 to be $T_0 \approx 2.1547$. It is trapping with the circles of radii approximately 0.23 and 0.67 being stable trapped rays. The invisible singularities are much more structured now and are in some neighborhoods of those two circles. As a result, jumps across radial and close to radial lines



Figure 5. Example 1 with the nontrapping speed c_1 . Case 3: $T = 4T_0$ with 10% noise. (a) The boundary distance map. (b) The exact initial condition. (c) The TR solution. (d) The NS solution. (e) x-slices of the TR solution (continuous line) and the exact solution (dashed line). (f) x-slices of the NS solution (continuous line) and the exact solution (dashed line). (g) y-slices of the TR solution (continuous line) and the exact solution (continuous line) and the exact solution (dashed line). (h) y-slices of the NS solution (continuous line) and the exact solution (dashed line).



Figure 6. Example 1 with the trapping speed c_3 . $T = 4T_0$. (a) The boundary distance map. (b) The exact initial condition. (c) The TR solution. (d) The NS solution. (e) x-slices of the TR solution ("-") and the exact solution ("o"). (f) x-slices of the NS solution ("-") and the exact solution ("o"). (g) y-slices of the TR solution ("-") and the exact solution ("o"). (h) y-slices of the NS solution ("-") and the exact solution ("o").



Figure 7. Example 2 with the nontrapping speed c_1 . Case 1: $T = 4T_0$. (a) The boundary distance map. (b) The exact initial condition. (c) The TR solution. (d) The NS solution. (e) x-slices of the TR solution ("-") and the exact solution ("o"). (f) x-slices of the NS solution ("-") and the exact solution ("o"). (g) y-slices of the TR solution ("-") and the exact solution ("o"). (h) y-slices of the NS solution ("-") and the exact solution ("o").



Figure 8. Example 2 with the nontrapping speed c_1 . Case 2: $T = 4T_0$ with 10% noise. (a) The boundary distance map. (b) The exact initial condition. (c) The TR solution. (d) The NS solution. (e) x-slices of the TR solution ("-") and the exact solution ("o"). (f) x-slices of the NS solution ("-") and the exact solution ("o"). (g) y-slices of the TR solution ("-") and the exact solution ("o"). (h) y-slices of the NS solution ("-") and the exact solution ("o").



Figure 9. Example 2 with the trapping speed c_3 . Case 1: $T = 4T_0$. (a) The boundary distance map. (b) The exact initial condition. (c) The TR solution. (d) The NS solution.

near those circles are affected more. Notice that the NS solution with k = 20 is much cleaner in the smooth parts close to the boundary, and the TR one is brighter than the original close to the center. This shows that the NS solutions reconstruct the low frequency modes better.

9. Numerical results: Discontinuous sound speed. We present here numerical examples with two discontinuous sound speeds illustrated in Figure 12. The first one, which we denote by c_4 , is equal to 0.8 in the square $[-1, 1]^2$, then jumps by about a factor of 2 from the interior to the exterior, and then jumps to 1. The second one, c_5 , has similar jumps but inside the square $[-1, 1]^2$ is variable, equal to the speed (32).

If the speed jumps by a factor of 2 when going out of a square, all rays hitting the boundary at an angle less than 60 degrees are completely reflected. Those rays that hit the boundary at an angle greater than 30 degrees generate a reflected ray hitting the boundary again without a transmitted component, etc. Therefore, all rays hitting the boundary of the smaller square at angles between 30 and 60 degrees are completely trapped for all times.

9.1. Example 3: The Shepp–Logan phantom, Figures 13–14.

9.1.1. Piecewise constant discontinuous speed c_4 .



Figure 10. Example 2 with the trapping speed c_3 . Case 2: $T = 4T_0$ with 10% random noise. (a) The boundary distance map. (b) The exact initial condition. (c) The TR solution. (d) The NS solution.

Figure 13. The sound speed is c_4 , given by Figure 12(a). We have $T_0 \approx 1.50$ with $T = 4T_0$. The artifacts in the TR image are quite strong and can be explained by the way that singularities propagate in this case. In Figure 1, for example, the ray that exits on the top carries a fraction of the energy only. When we reverse the time, at the first contact with the outer boundary of the "skull," this ray will create a reflected one (not seen in the figure) together with the transmitted one, shown there. The reflected one is not part of the actual graph that we are trying to invert. It will reflect off $\partial\Omega$ and go back to the interior of Ω , creating more artificial rays, etc. If we had an infinite time T, then those artificial rays would be canceled by other such artificial rays, leading to an exact reconstruction, as $T \to \infty$ (at a very slow logarithmic rate for smooth f); this explains the artifacts in the TR image. The NS expansion with k = 16 creates very few artifacts, and the few that are seen are mostly due to singularities near or on the "skull."

9.1.2. Discontinuous speed c_5 .

Figure 14. The sound speed c_5 is given by Figure 12(b). Here, $T_0 \approx 1.36$, $T = 4T_0$. The speed is not so symmetric anymore, and the artifacts in the TR image are still strong but more random. The variable speed inside improves the NS image, where k = 16.



Figure 11. Example 2 with the trapping speed c_2 . $T = 4T_0$. (a) The boundary distance map. (b) The exact initial condition. (c) The TR solution. (d) The NS solution.



Figure 12. Sound speed models. (a) A discontinuous piecewise sound speed c_4 . (b) A nonpiecewise constant discontinuous sound speed c_5 .



Figure 13. Example 3 with the discontinuous sound speed c_4 . $T = 4T_0$. (a) The boundary distance map. (b) The exact initial condition. (c) The TR solution. (d) The NS solution.

9.2. Example 4: Zebras, Figure 15.

Figure 15. The sound speed c_5 is given by Figure 12(b). Here, $T_0 \approx 1.36$, and we take $T = 4T_0$. As in Figure 14, the TR reconstruction (error 21.83%) contains many wrong singularities, while the NS image (error 8%, k = 16) is very clean.

10. Numerical results: Partial data. We present here numerical examples with data given on three or two sides only. In the first case, we remove data on the right-hand side of the square; in the second, we remove data on the right-hand side and on the bottom of the square. We also use a smooth cutoff. We present examples with the nontrapping and the trapping speeds that we used before. Note that the notion of trapping changes with partial data since the stability depends on whether all singularities can reach Γ for times $\leq T$. Still, removing the data on some of the sides in addition to the much longer times that signals need to reach Γ produces worse reconstructions when the speed is trapping.

10.1. Partial data, nontrapping speed c_1 . The speed c_1 is given by (32). In all three cases, T = 4.7.



Figure 14. Example 3 with the discontinuous sound speed c_5 . (a) The boundary distance map. (b) The exact initial condition. (c) The TR solution. (d) The NS solution.

10.1.1. The Shepp-Logan phantom, Figures 16(a) and 16(b). We estimate T_0 to be $T_0 \approx 2.5$. It is greater than before because Γ is not the whole $\partial\Omega$ anymore. We have data on two adjacent sides. If the speed were constant, the singularities below the lower-left-to-upperright diagonal that would exit in the sides without measurements would be invisible. That would affect jumps across surfaces in that triangle with normals parallel or nearly parallel to that diagonal, roughly speaking. The speed is variable but not too far from constant, and we see the expected behavior. Note that this is an unstable case, and our analysis does not exclude an even exponential divergence (in the low frequency part), but the error gets smaller up to k = 8 when the computation is stopped.

10.1.2. Zebras, Figures 17(a) and 17(b). We have data on three sides. In this case, $T_1 \approx 1.35$, and we would have stability if the speed were constant. However, the speed is not constant, but in some sense it is not too far from a constant one. There might be a small set of geodesics that enter through the right-hand side and exit there as well, thus creating instability. The reconstruction is very good, however, with an error of 6.11%, k = 10. The chosen T is slightly greater than what would be the stability time, but the result is computed without the contribution of those geodesics.



Figure 15. Example 4 with the discontinuous speed c_5 : Figure 12(b). (a) The boundary distance map. (b) The exact initial condition. (c) The TR solution. (d) The NS solution.



Figure 16. Nontrapping speed c_1 with partial data on two adjacent sides. T = 4.7. (a) The boundary distance map, data on two sides. (b) The Shepp-Logan phantom reconstruction.



Figure 17. A modified example with the nontrapping speed c_1 , partial data. T = 4.7. (a) The boundary distance map, data on three sides. (b) The reconstructed "zebras" image. (c) The boundary distance map, data on two sides. (d) The reconstructed "zebras" image.

10.1.3. Zebras, Figures 17(c) and 17(d). We use data on two sides. This is a very unstable case, and one can see artifacts where the invisible singularities lie—in the lower right-hand triangle, with slopes close to 1, roughly speaking (jumps across curves with slopes in a neighborhood of -1). The error is 10.6%, and k = 10.

10.1.4. Zebras, Figures 18(a) and 18(b). The zebras image reconstructed in Figure 17(d) does not have so many invisible singularities in this particular setup (data on two sides), however. For this reason, we present another example, Figures 18(a) and 18(b), with a modified image that shows the expected behavior of the visible and invisible singularities. The error is 11.39%, and k = 11.

10.2. Partial data, trapping sound speeds c_3 and c_2 . The sound speed is c_3 (the first two examples) and c_2 (the third example).

10.2.1. Figures 19(a) and 19(b). Here, $T = 4.93 > T_0 \approx 1.3$, with data on three sides. The "chaotic" trapping speed c_3 makes the reconstruction worse than before. This is an unstable case because the trapping speed leaves many singularities invisible. As expected, the



Figure 18. Nontrapping speed c_1 examples, partial data on the left and the upper side. T = 4.7. (a) The exact initial condition. (b) The reconstructed "zebras" image.

worst part is near the side with no observations due to geodesics that enter and exit through that side. There are invisible singularities everywhere as well, due to the speed. The error is 19.25%, and k = 10.

10.2.2. Figures 19(c) and 19(d). Here, $T = 4.93 > T_0 \approx 2.1$, with data on two sides, and the speed is the same as above. As expected, the reconstruction is quite bad near the sides with no data. The error is 28.60%, and k = 16.

10.2.3. Figures 19(e) and 19(f). Here, $T = 8.61 > T_0 \approx 2.6$, with data on two sides, and the speed is c_2 . The time T is larger than above, but T_0 is slightly larger as well. The reconstruction is better due to the larger time and (probably) due to the fact that the trapping region of this speed is farther away from the sides where no observations are done. Experiments with times T closer to that in the two examples above (not shown) still yield a better reconstruction with this speed. The error is 16.17%, and k = 16.

11. Conclusions. We have implemented an efficient Neumann series-based algorithm for reconstructing an unknown source in TAT and PAT. The algorithm is built on the recent advances in understanding the theoretical nature of the problem. We worked with variable sound speeds that might also be discontinuous across some surface. The latter problem arises in brain imaging. Our algorithm is based on an explicit formula in the form of a Neumann series. We first modify the data at t = T using a harmonic extension, then do time reversal, and then iterate. The extensive numerical examples demonstrate the robust performance of the algorithm.

This approach also has the potential to improve the reconstruction when the possibly discontinuous speed is not well known by extracting most of the useful information for limited values of the observation time T; see [6]. Future applications include incorporating our algorithm into some practical experiments to see how the algorithm will behave on practical data. We will report on this in an upcoming paper.

lary distance map

-0.

0.5





Figure 19. Examples with the trapping speeds c_3 (the first two rows) and c_2 (the last row). (a) The boundary distance map, data on three sides. (b) The reconstructed "zebras" image, T = 4.92. (c) The boundary distance map, data on two sides. (d) The reconstructed "zebras" image, T = 4.92, k = 16. (e) The boundary distance map, data on two sides. (f) The reconstructed "zebras" image, T = 8.62, k = 16.

Acknowledgments. We thank MSRI for a semester-long program on inverse problems in a stimulating atmosphere, which initiated this project. We also thank the reviewers for their useful suggestions for improving the presentation of this paper. Qian would like to dedicate this paper to Professor Jiaqi Liu of Harbin Institute of Technology on the occasion of his 70th birthday.

REFERENCES

- M. AGRANOVSKY, P. KUCHMENT, AND L. KUNYANSKY, On reconstruction formulas and algorithms for the thermoacoustic tomography, in Photoacoustic Imaging and Spectroscopy, CRC Press, Boca Raton, FL, 2009, pp. 89–101.
- J.-P. BERENGER, A perfectly matched layer for the absorption of electromagnetic waves, J. Comput. Phys., 114 (1994), pp. 185–200.
- [3] A. BRANDT, Multi-level adaptive solutions to boundary-value problems, Math. Comp., 31 (1977), pp. 333– 390.
- [4] N. BURQ, Décroissance de l'énergie locale de l'équation des ondes pour le problème extérieur et absence de résonance au voisinage du réel, Acta Math., 180 (1998), pp. 1–29.
- [5] W. C. CHEW AND W. H. WEEDON, A 3D perfectly matched medium from modified Maxwell's equations with stretched coordinates, Microw. Opt. Tech. Lett., 7 (1994), pp. 599–604.
- B. COX AND B. TREEBY, Artifact trapping during time reversal photoacoustic imaging for acoustically heterogeneous media, IEEE Trans. Med. Imaging, 29 (2010), pp. 387–396.
- [7] D. FINCH, M. HALTMEIER, AND RAKESH, Inversion of spherical means and the wave equation in even dimensions, SIAM J. Appl. Math., 68 (2007), pp. 392–412.
- [8] D. FINCH, S. K. PATCH, AND RAKESH, Determining a function from its mean values over a family of spheres, SIAM J. Math. Anal., 35 (2004), pp. 1213–1240.
- D. FINCH AND RAKESH, Recovering a function from its spherical mean values in two and three dimensions, in Photoacoustic Imaging and Spectroscopy, CRC Press, Boca Raton, FL, 2009.
- [10] H. GRUN, C. HOFER, M. HALTMEIER, G. PALTAUFF, AND P. BURGHOLZER, Thermoacoustic imaging using time reversal, in Proceedings of the International Congress on Ultrasonics, Vienna, 2007, Paper ID 1542.
- [11] M. HALTMEIER, O. SCHERZER, P. BURGHOLZER, AND G. PALTAUF, Thermoacoustic computed tomography with large planar receivers, Inverse Problems, 20 (2004), pp. 1663–1673.
- [12] M. HALTMEIER, T. SCHUSTER, AND O. SCHERZER, Filtered backprojection for thermoacoustic computed tomography in spherical geometry, Math. Methods Appl. Sci., 28 (2005), pp. 1919–1937.
- [13] L. HÖRMANDER, Fourier integral operators. I, Acta Math., 127 (1971), pp. 79–183.
- [14] Y. HRISTOVA, Time reversal in thermoacoustic tomography—an error estimate, Inverse Problems, 25 (2009), 055008.
- [15] Y. HRISTOVA, P. KUCHMENT, AND L. NGUYEN, Reconstruction and time reversal in thermoacoustic tomography in acoustically homogeneous and inhomogeneous media, Inverse Problems, 24 (2008), 055006.
- [16] X. JIN AND L. V. WANG, Thermoacoustic tomography with correction for acoustic speed variations, Phys. Med. Biol., 51 (2006), pp. 6437–6448.
- [17] C. Y. KAO, S. J. OSHER, AND J. QIAN, Lax-Friedrichs sweeping schemes for static Hamilton-Jacobi equations, J. Comput. Phys., 196 (2004), pp. 367–391.
- [18] R. A. KRUGER, W. L. KISER, D. R. REINECKE, AND G. A. KRUGER, Thermoacoustic computed tomography using a conventional linear transducer array, Med. Phys., 30 (2003), pp. 856–860.
- [19] R. A. KRUGER, D. R. REINECKE, AND G. A. KRUGER, Thermoacoustic computed tomography—technical considerations, Med. Phys., 26 (1999), pp. 1832–1837.
- [20] G. KU, B. FORNAGE, X. JIN, M. XU, K. HUNT, AND L. V. WANG, Thermoacoustic and photoacoustic tomography of thick biological tissues toward breast imaging, Technol. Cancer Res. Treat., 4 (2005), pp. 559–565.
- [21] P. KUCHMENT AND L. KUNYANSKY, Mathematics of thermoacoustic tomography, European J. Appl. Math., 19 (2008), pp. 191–224.
- [22] M. LI, J. OH, X. XIE, G. KU, W. WANG, C. LI, G. LUNGU, G. STOICA, AND L. V. WANG, Simultaneous molecular and hypoxia imaging of brain tumors in vivo using spectroscopic photoacoustic tomography, Proc. IEEE, 96 (2008), pp. 481–489.

THERMO- AND PHOTOACOUSTIC TOMOGRAPHY WITH VARIABLE SOUND SPEED

- [23] Q. H. LIU AND J. TAO, The perfectly matched layer for acoustic waves in absorptive media, J. Acoust. Soc. Am., 102 (1997), pp. 2072–2082.
- [24] D. MEDKOVÁ, On the convergence of Neumann series for noncompact operators, Czechoslovak Math. J., 41 (1991), pp. 312–316.
- [25] S. K. PATCH, Thermoacoustic tomography—consistency conditions and the partial scan problem, Phys. Med. Biol., 49 (2004), pp. 2305–2315.
- [26] J. QIAN, Y. T.-ZHANG, AND H.-K. ZHAO, Fast sweeping methods for eikonal equations on triangular meshes, SIAM J. Numer. Anal., 45 (2007), pp. 83–107.
- [27] J. SJÖSTRAND AND M. ZWORSKI, Complex scaling and the distribution of scattering poles, J. Amer. Math. Soc., 4 (1991), pp. 729–769.
- [28] P. STEFANOV AND G. UHLMANN, Stability estimates for the hyperbolic Dirichlet to Neumann map in anisotropic media, J. Funct. Anal., 154 (1998), pp. 330–358.
- [29] P. STEFANOV AND G. UHLMANN, Linearizing non-linear inverse problems and an application to inverse backscattering, J. Funct. Anal., 256 (2009), pp. 2842–2866.
- [30] P. STEFANOV AND G. UHLMANN, Thermoacoustic tomography with variable sound speed, Inverse Problems, 25 (2009), 075011.
- [31] P. STEFANOV AND G. UHLMANN, Thermoacoustic tomography arising in brain imaging, Inverse Problems, 27 (2011), 045004.
- [32] D. TATARU, Unique continuation for operators with partially analytic coefficients, J. Math. Pures Appl. (9), 78 (1999), pp. 505–521.
- [33] L. V. WANG AND H.-I. WU, Biomedical Optics: Principles and Imaging, Wiley-Interscience, Hoboken, NJ, 2007.
- [34] M. XU AND L.-H. V. WANG, Universal back-projection algorithm for photoacoustic computed tomography, Phys. Rev. E., 71 (2005), 016706.
- [35] M. XU AND L. V. WANG, Photoacoustic imaging in biomedicine, Rev. Sci. Instrum., 77 (2006), 041101.
- [36] Y. XU AND L.-H. V. WANG, Time reversal and its application to tomography with diffracting sources, Phys. Rev. Lett., 92 (2004), 033902.
- [37] Y. XU AND L. V. WANG, Effects of acoustic heterogeneity in breast thermoacoustic tomography, IEEE Trans. Ultrason. Ferroelectr. Freq. Control, 50 (2003), pp. 1134–1146.
- [38] Y. XU AND L. V. WANG, Rhesus monkey brain imaging through intact skull with thermoacoustic tomography, IEEE Trans. Ultrason. Ferroelectr. Freq. Control, 53 (2006), pp. 542–548.
- [39] X. YANG AND L. V. WANG, Monkey brain cortex imaging by photoacoustic tomography, J. Biomed. Opt., 13 (2008), 044009.
- [40] H. K. ZHAO, A fast sweeping method for eikonal equations, Math. Comp., 74 (2005), pp. 603–627.