

# Transmission traveltimes tomography based on paraxial Liouville equations and level set formulations

Shingyu Leung<sup>1</sup> and Jianliang Qian<sup>2</sup>

<sup>1</sup> Department of Mathematics, UCLA, Los Angeles, CA 90095-1555, USA

<sup>2</sup> Department of Mathematics and Statistics, Wichita State University, Wichita, KS 67260-0033, USA

E-mail: [syleung@math.ucla.edu](mailto:syleung@math.ucla.edu) and [qian@math.wichita.edu](mailto:qian@math.wichita.edu)

Received 12 March 2006, in final form 13 February 2007

Published 13 March 2007

Online at [stacks.iop.org/IP/23/799](http://stacks.iop.org/IP/23/799)

## Abstract

We propose a new formulation for traveltimes tomography based on paraxial Liouville equations and level set formulations. This new formulation allows us to account for multivalued traveltimes and multipathing systematically in the tomography problem. To obtain efficient implementations, we use the adjoint state technique and the method of gradient descent. Starting from some initial guess, we minimize a nonlinear energy functional by a Newton-type method. The required gradient is computed by solving one forward and one adjoint problem of the paraxial Liouville equations. Then the velocity model is updated iteratively by solving a Helmholtz equation with the computed gradient as the right-hand side. Numerical examples with and without added noise demonstrate that the new formulation is effective and accurate. To our knowledge, this is the first Eulerian approach to take into account all arrivals systematically in traveltimes tomography.

(Some figures in this article are in colour only in the electronic version)

## 1. Introduction

Traveltimes tomography [5, 6, 11, 14, 34, 35, 51] is an important class of inverse problems. Theoretically, it amounts to determining uniquely a Riemannian metric by knowing the length of geodesics joining points of the boundary of a two-dimensional or three-dimensional compact domain; those geodesics solve a family of Hamilton–Jacobi (HJ) equations. Computationally, it amounts to designing fast, high resolution methods to invert given data for the unknown velocities or other material parameters [2–4, 22, 23, 44, 45, 49, 50]. For a general account of inverse problems for PDEs, see [24].

Geophysical traveltimes tomography is closely related to the x-ray computerized tomography used in medical diagnosis [15]. The so-called transmission traveltimes tomography

uses traveltimes data between boreholes to invert velocity as a discretized field [6, 13, 36, 51]. Seismic tomography is usually formulated as a minimization problem that produces a velocity model which minimizes the difference between traveltimes generated by tracing rays through the model and traveltimes measured from the data [2–6, 22, 23, 51]. The ray tracing in the above methods is based on Fermat's principle, and the resulting ray path is computed from an explicit discretization of a nonlinear ray-path integral; as such, the methods inherit some shortcomings from typical ray tracing methods, such as shadow zones due to non-uniform coverage of the computational domain.

On the other hand, the work in [44, 45] has shown that it is possible to formulate the transmission tomography problem in a purely Eulerian framework based on finite difference eikonal solvers; Sei and Symes [44, 45] have also derived the linearized eikonal equation and the traveltimes gradient based on the adjoint state method borrowed from optimal control theory.

However, all the above cited methods are based on first-arrival traveltimes only. As far as we know, no attempt has been made to formulate the transmission tomography problem by using all arrivals, including multivalued traveltimes, in a Eulerian framework. The works presented in [11, 14] have used multivalued traveltimes from ray tracing methods, but these works are on reflection traveltimes tomography which is different from transmission tomography in that rays start at the surface, reflect off interfaces whose depths are to be determined and return to the surface. A recent new method presented in [10] is able to take into account multiple arrivals in phase space framework, but it is based on a Lagrangian approach.

Because multivalued traveltimes and resulting multipathings are common in complex velocity structures, it is necessary to take into account all the arrivals systematically; an example of this is high resolution seismic imaging via integral transform in the presence of strong seismic refraction [20, 32, 37]. To achieve this purpose, we propose formulating transmission tomography using the Liouville equation and level-set-based PDE framework in phase space so that multivalued traveltimes can be used systematically.

The Liouville equation is a linear hyperbolic equation derived from a Hamiltonian system, which in turn is obtained from a HJ equation by the method of characteristics; therefore, the Liouville equation shares the same characteristics as the original nonlinear HJ equation, but it is linear at the price of doubling the number of independent variables; see [1, 8, 9, 12, 17, 19, 28, 38, 39]. In particular, inspired by Engquist *et al* [18, 38], in [31, 40, 41] we proposed paraxial formulations of the Liouville equation for multivalued geometrical optics; the resulting Eulerian framework for computing geometrical optics-related multivalued quantities has been shown to be computationally efficient and accurate, since to some extent this formulation has reduced the burden of doubling the number of independent variables in the Liouville framework.

Encouraged by the success of the formulations using the paraxial Liouville equations, we apply such techniques to seismic traveltimes transmission tomography problems so that multivalued traveltimes and resulting multipathings can be utilized systematically. To that end, we minimize a new energy functional which consists of mismatching terms for multivalued traveltimes and source locations. To minimize the functional, we derive its gradient using the adjoint state technique. Starting from some initial guess, we minimize the nonlinear energy functional by a Newton-type method. The required gradient is computed by solving one forward and one adjoint problem of the paraxial Liouville equations. Then the velocity model is updated iteratively by solving a Helmholtz equation with the computed gradient as the right-hand side. Therefore, the functional can be efficiently minimized by using the simple method of gradient descent, which is a simplified version of the Newton method. An abstract of some preliminary results in this paper has appeared in [29].

In section 2, we will first briefly review the approach in [31, 40, 41], in which we developed a Eulerian method based on level sets and Liouville equations to compute the multivalued solution of the eikonal equation. The corresponding inverse problem is formulated in section 3. Section 4 gives an algorithm and related implementation details. Section 5 details some practical issues associated with data measurements. Numerical examples are given in sections 6 and 7.

## 2. Level set method for paraxial geometrical optics

In [31, 40, 41] we have developed a level-set-based paraxial Liouville formulation for computing multivalued solutions of eikonal equations, and the resulting methods provide efficient ways for determining multivalued traveltimes of waves in inhomogeneous media. Here we briefly review the formulation for the two-dimensional case; see [40, 41] for more details.

Denote

$$\Omega_p = \{(x, z) : x_{\min} \leq x \leq x_{\max}, 0 \leq z \leq z_f\} \quad (1)$$

and assume that the source is located on the surface:  $x_{\min} \leq x_s \leq x_{\max}$  and  $z_s = 0$ .

Consider the paraxial eikonal equation for traveltime  $\tau = \tau(z, x)$  along the  $z$ -direction [47],

$$\frac{\partial \tau}{\partial z} = H\left(x, z, \frac{\partial \tau}{\partial x}\right) = \sqrt{\max\left(\frac{1}{c^2} - \left(\frac{\partial \tau}{\partial x}\right)^2, \frac{\cos^2 \theta_{\max}}{c^2}\right)}, \quad (2)$$

with a point source condition

$$\tau(z_s, x_s) = 0, \quad (3)$$

where  $0 \leq \theta_{\max} < 0.5\pi$ , and  $c = c(z, x) > 0$  is the positive velocity; the reciprocal of  $c$  is the slowness field. This equation comes from enforcing the so-called sub-horizontal condition in the eikonal equation,

$$|\nabla_{\mathbf{x}} \tau(\mathbf{x}, \mathbf{x}_s)| = \frac{1}{c(\mathbf{x})}; \quad (4)$$

see [47] for further details. By the paraxial assumption, we only allow downgoing rays and do not allow upgoing rays; however, the formulation still accommodates velocity models with variations in both transversal and vertical directions, as long as the downgoing assumption holds. However, as pointed out in [17], the paraxial approximation in geometrical optics signifies another simplification which can be made when there is one preferred wave propagation direction so that a stationary eikonal equation can be rewritten as an evolution equation in one of the spatial variables [21, 42, 43]; as long as a version of the so-called sub-horizontal condition holds, there is no approximation involved. The paraxial eikonal equation that we use first appeared in [21] and was rigorously justified in [47].

Applying the method of characteristics to the paraxial eikonal equation, we have a ray tracing system,

$$\frac{dX}{dz} = \tan \Theta, \quad X(0) = x_s, \quad (5)$$

$$\frac{d\Theta}{dz} = \frac{1}{c} \left( \frac{\partial c}{\partial z} \tan \Theta - \frac{\partial c}{\partial x} \right), \quad \Theta(0) = \theta_s, \quad (6)$$

$$\frac{dT}{dz} = \frac{1}{c \cos \Theta}, \quad T(0) = 0, \quad (7)$$

where  $x_s$  varies from  $x_{\min}$  to  $x_{\max}$ , and  $\theta_s$  varies from  $-\theta_{\max} \leq \theta \leq \theta_{\max} < \frac{\pi}{2}$ . This ray tracing system (5)–(7) is a Lagrangian formulation defined in the reduced phase space  $(z; x, \theta)$ .

Denote

$$\Omega = \{(x, \theta) : x_{\min} \leq x \leq x_{\max}, -\theta_{\max} \leq \theta \leq \theta_{\max}\}. \quad (8)$$

We introduce a function,

$$\phi = \phi(z, x, \theta) : [0, z_f] \times \Omega \rightarrow [x_{\min}, x_{\max}],$$

such that, for any  $x_s \in [x_{\min}, x_{\max}]$  and  $z \in [0, z_f]$ , the set

$$\Gamma(z; x_s) = \{(X(z), \Theta(z)) : \phi(z, X(z), \Theta(z)) = x_s\}$$

gives the location of the reduced bi-characteristic strip  $(X(z), \Theta(z))$  initiated from the source  $x_s$  with take-off angles  $-\theta_{\max} \leq \theta_s \leq \theta_{\max}$ .

As an initial condition, we let

$$\phi(0, x, \theta) = x, \quad (x, \theta) \in \Omega. \quad (9)$$

Admittedly, this initial condition seems to be unusual. For a fixed  $x_s \in [x_{\min}, x_{\max}]$ , the location where  $\phi(0, x, \theta) = x_s$  holds is

$$\Gamma(0; x_s) = \{(x, \theta) : x = x_s, -\theta_{\max} \leq \theta \leq \theta_{\max}\},$$

which states that the initial take-off angle varies from  $-\theta_{\max}$  to  $\theta_{\max}$  at the source location  $x_s$ .

Next, we differentiate the level set identity with respect to  $z$  to obtain

$$\phi_z + u\phi_x + w\phi_\theta = 0, \quad (10)$$

where

$$u(z, x, \theta) = \tan \theta, \quad w = \frac{1}{c}(c_z \tan \theta - c_x) \quad (11)$$

are functions defined in the space  $[0, z_f] \times \Omega$ ; we point out that  $u$  and  $w$  are obtained by respectively viewing  $\frac{dX}{dz}$  and  $\frac{d\Theta}{dz}$  as vector fields.

By the initial condition (9), we transport many source locations simultaneously. Consequently, evolving equations (10) and (9) will transport source locations to any  $z$  according to the vector fields  $u$  and  $w$ .

Given  $z \in [0, z_f]$  and  $x_s \in [x_{\min}, x_{\max}]$ , the set  $\Gamma(z; x_s)$  is a curve in  $\Omega$ , which defines an implicit function between  $X$  and  $\Theta$ . When  $z = 0$ ,  $\Gamma(0; x_s)$  is a vertical line in  $\Omega$ , indicating that the rays with take-off angles from  $-\theta_{\max}$  to  $\theta_{\max}$  emanate from the source location  $x_s$ . When  $z \neq 0$ ,  $\Gamma(z; x_s)$  being a curve indicates that for some  $X = x_r$  there are more than one  $\Theta = \theta_r^a$  such that  $\phi(z, x_r, \theta_r^a) = x_s$ , implying that more than one ray emanating from the source  $x_s$  reach the physical location  $(z, x_r)$  with arrival angles  $\theta_r^a$  (see figure 1).

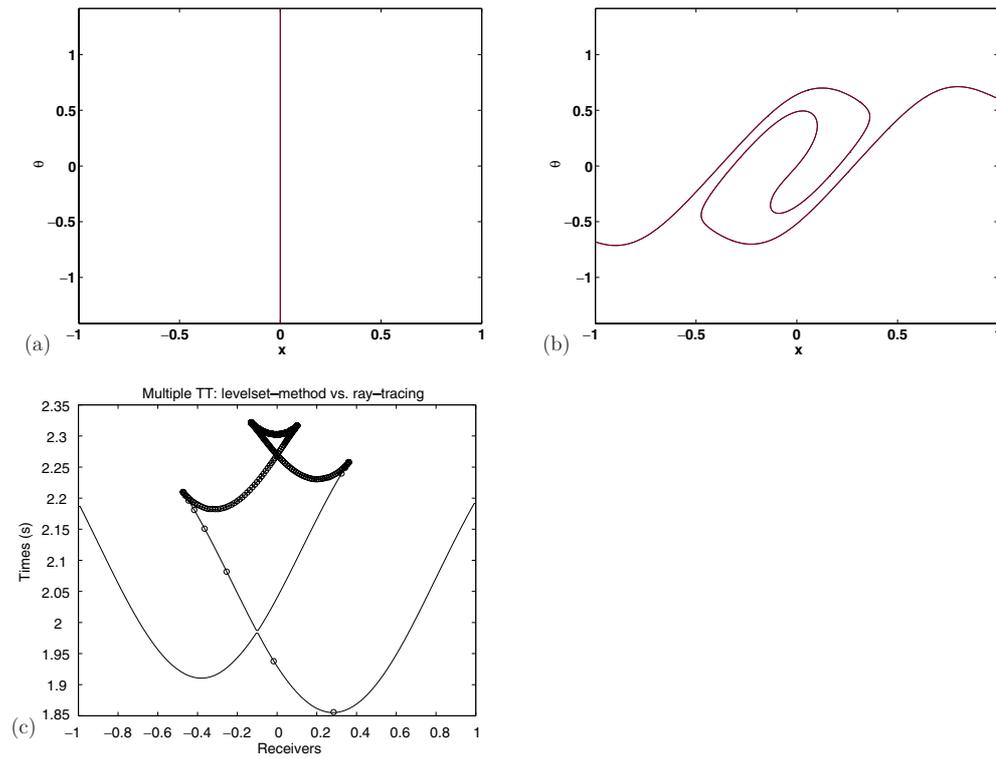
Next we derive an equation to determine the traveltime for those multiple rays reaching the physical location  $(z, x)$ . Consider  $T = T(z) = T(z, X(z), \theta(z))$  along a ray. Differentiating with respect to  $z$  in equation (7), we have

$$T_z + uT_x + wT_\theta = \frac{1}{c \cos \theta}, \quad (12)$$

which is equipped with the initial condition

$$T(0, x, \theta) = 0. \quad (13)$$

By solving the level set equation (10), we have the locations of reduced bi-characteristics; by solving the traveltime equation (12), we have the traveltime values at the corresponding



**Figure 1.** A sinusoidal model. The evolution of level sets and traveltimes for the source location  $x_s = 0$ . (a) The  $x_s$ -level set at  $z = 0$ , (b) the  $x_s$ -level set at  $z = 2.0$ , where the horizontal axis is the  $x$ -axis and the vertical is the  $\theta$ -axis: arrival angles and arrival locations are uniformly sampled in the  $x$ - $\theta$  space, implying that the corresponding take-off angles are not uniformly sampled; (c) the traveltimes are uniformly sampled in physical space; such a multivalued traveltime curve can be partitioned into monovalued branches at caustics.

locations. Therefore, the multivalued traveltime at a specific location in the physical space can be obtained by first computing the locations of the reduced bi-characteristics and then interpolating the traveltime at that location from gridded traveltimes. These equations give us a Eulerian formulation for computing multivalued traveltimes at any physical location for many sources simultaneously.

We also remark that since in the current linear Liouville setting all the characteristics are not intersecting with each other, we need not worry about splitting and merging as in many traditional applications of level set methods. On the other hand, we also remark that by adding one dimension to the level set formulation in the two-dimensional case we are able to parameterize multivalued solutions and deal with many sources simultaneously.

### 3. Tomography based on paraxial Liouville equations: inverse problems

In the direct problem, we assume that the velocity  $c$  in the physical domain  $\Omega_p$  is known; this velocity can then be used to compute the corresponding multivalued arrival times. The inverse problem is to determine the velocity inside  $\Omega_p$  using the multivalued arrival times measured on the physical boundary.

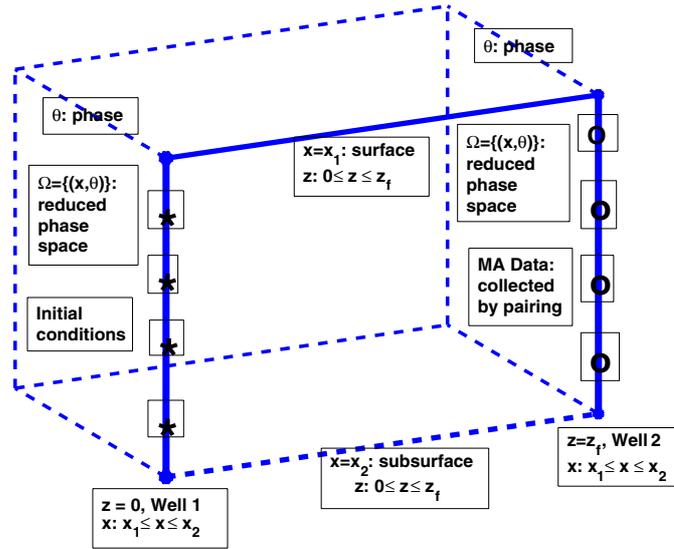


Figure 2. Use multi-arrivals from received time series via Liouville in phase space.

Denote  $\tilde{\Omega} = \Omega \times (0, z_f)$  and  $z_0 = 0$ . According to the forward problem given in the above section, we have the following state equations:

$$\phi_z + u\phi_x + v\phi_\theta = 0, \quad T_z + uT_x + vT_\theta = \frac{1}{c \cos \theta}, \quad (14)$$

where  $\mathbf{u} = (u, v) = (\tan \theta, m_z \tan \theta - m_x)$  and  $m = m(c) = \log c$ . The initial and boundary conditions of this system are given by

$$\phi(z_0, x, \theta) = x, \quad T(z_0, x, \theta) = 0, \quad (15)$$

and

$$\begin{aligned} \phi(z, x, \theta)|_{(x, \theta) \in \partial\Omega} &= \begin{cases} \phi^* & \text{if } (\mathbf{u} \cdot \mathbf{n}) < 0, \\ \text{no b.c. needed} & \text{if } (\mathbf{u} \cdot \mathbf{n}) \geq 0; \end{cases} \\ T(z, x, \theta)|_{(x, \theta) \in \partial\Omega} &= \begin{cases} T^* & \text{if } (\mathbf{u} \cdot \mathbf{n}) < 0, \\ \text{no b.c. needed} & \text{if } (\mathbf{u} \cdot \mathbf{n}) \geq 0, \end{cases} \end{aligned} \quad (16)$$

where  $\mathbf{n}$  is the outward normal vector of  $\partial\Omega$ , and  $*$  denotes the conditions on the inflow boundary of  $\partial\Omega$ , where  $\mathbf{u} \cdot \mathbf{n} < 0$ .

We note that in the above formulation we use all the possible source locations as initialized in equation (15) rather than a particular single source. In particular, we point out that in the above equation,  $\phi(z_0, x, \theta) = x$ ,  $\phi(z_0, x, \theta)$  represents the source location  $(z_0, x)$  in the phase space  $(z, x, \theta)$ .

In the following, without abusing the notation, we also use  $*$  to denote the measured values on the outflow boundary of  $\partial\Omega$ , where  $\mathbf{u} \cdot \mathbf{n} \geq 0$ , and the measured values on the level  $z = z_f$ .

Assume that one can measure data,  $\phi(z, x, \theta)|_{(x, \theta) \in \partial\Omega}$ , on the outflow boundary,  $\phi(z_f, x, \theta)$  on the level  $z = z_f$ ,  $T(z, x, \theta)|_{(x, \theta) \in \partial\Omega}$  on the outflow boundary,  $T(z_f, x, \theta)$  on the level  $z = z_f$  and  $m|_{\partial\Omega_p}$  (see figure 2).

We remark that (i) the boundary measurements,  $\phi(z, x, \theta)|_{(x, \theta) \in \partial\Omega}$  and  $T(z, x, \theta)|_{(x, \theta) \in \partial\Omega}$ , are available when the interior measurement is possible such as in transient elastography [26, 33], where careful treatment of multivalued traveltimes is needed as well; (ii) in transmission traveltome tomography the outflow boundary measurements are typically not

available, but we can avoid using these data by taking advantage of finite speed of wave propagation; see section 5 for more details. Nevertheless, to make the presentation transparent we make the above assumptions.

We also remark that data  $\phi(z_f, x, \theta)$   $((x, \theta) \in \Omega)$  and  $T(z_f, x, \theta)$   $((x, \theta) \in \Omega)$  are nothing but the arriving ray information at the receiver  $x$  (equivalent to the geophone ‘g’ in [14]) with the arriving angle  $\theta$  (equivalent to the ray parameter ‘p’ in [14]):  $\phi(z_f, x, \theta)$  being the initial source location which is transported to  $z = z_f$  (equivalent to the shot location ‘s’ in [14]) and  $T(z_f, x, \theta)$  being the transmission traveltime of the arriving ray (analogous to the reflection traveltime ‘t’ in [14]). Consequently, such data can be picked from seismic data by plain preprocessing as in [14]; see section 5 for more details. Although the information that we use is almost the same as that in [14], our formulation is a Eulerian partial differential equation formulation while the one in [14] is a Lagrangian ordinary differential equation formulation.

Based on the above assumptions, the traveltime tomography problem is to determine  $m$ , and therefore  $c$ , such that the predicted data from the solutions for the system (14)–(16) are as close to these measurements as possible. To achieve this, we propose to minimize the following energy:

$$E(m) = \frac{1}{2} \int_{\Omega} (\phi - \phi^*)^2|_{z=z_f} + \frac{1}{2} \int_z \int_{\partial\Omega} (\mathbf{u} \cdot \mathbf{n})(\phi - \phi^*)^2 + \frac{\beta}{2} \int_{\Omega} (T - T^*)^2|_{z=z_f} + \frac{\beta}{2} \int_z \int_{\partial\Omega} (\mathbf{u} \cdot \mathbf{n})(T - T^*)^2. \quad (17)$$

The proposed energy is always positive. Although it is possible that  $(\mathbf{u} \cdot \mathbf{n}) < 0$  on  $\partial\Omega$ , in that case  $\phi - \phi^* = T - T^* = 0$ . This implies that these negative values of  $(\mathbf{u} \cdot \mathbf{n})$  will have no contribution to the overall energy. In other words, the proposed energy measures the difference between the computed solution and the data on the *outflow* boundary. On the *inflow* boundary, the numerical solution automatically matches the measurements according to the conditions in (16). The parameter  $\beta$  balances the mismatching of the traveltimes and the source locations.

In [14] a similar mismatching functional was proposed for discrete data sets in a Lagrangian formulation. Comparing the energy functional (17) with the objective function, formula (6) on page 705 in [14], we conclude that the two formulations encode similar information in phase space. Although our formulation is based on continuum data sets, we eventually have to discretize it, which will be addressed in detail later on.

To compute the minimizer of the energy (17), one can use the method of gradient descent. To do that we need to compute the gradient of this functional by linearization. Here we give a formal derivation of the gradient. Therefore, we perturb  $m$  by  $\epsilon\tilde{m}$ ; the corresponding changes in  $\phi$  and  $T$ , denoted by  $\epsilon\tilde{\phi}$  and  $\epsilon\tilde{T}$ , respectively, satisfy

$$\tilde{\phi}_z + u\tilde{\phi}_x + v\tilde{\phi}_\theta = [\tilde{m}_x - \tilde{m}_z \tan \theta]\phi_\theta, \quad (18)$$

$$\tilde{T}_z + u\tilde{T}_x + v\tilde{T}_\theta = [\tilde{m}_x - \tilde{m}_z \tan \theta]T_\theta - \frac{\tilde{m}}{c \cos \theta}. \quad (19)$$

Define  $\tilde{\mathbf{u}} = (0, \tilde{v}) = (0, \tilde{m}_z \tan \theta - \tilde{m}_x)$ . The corresponding change in the energy is given by

$$\begin{aligned} \delta E = E(m + \epsilon\tilde{m}) - E(m) = & \epsilon \left[ \int_{\Omega} \tilde{\phi}(\phi - \phi^*)|_{z=z_f} + \int_z \int_{\partial\Omega} (\mathbf{u} \cdot \mathbf{n})\tilde{\phi}(\phi - \phi^*) \right. \\ & + \frac{1}{2} \int_z \int_{\partial\Omega} (\tilde{\mathbf{u}} \cdot \mathbf{n})(\phi - \phi^*)^2 + \beta \int_{\Omega} \tilde{T}(T - T^*)|_{z=z_f} \\ & \left. + \beta \int_z \int_{\partial\Omega} (\mathbf{u} \cdot \mathbf{n})\tilde{T}(T - T^*) + \frac{\beta}{2} \int_z \int_{\partial\Omega} (\tilde{\mathbf{u}} \cdot \mathbf{n})(T - T^*)^2 \right] + O(\epsilon^2). \quad (20) \end{aligned}$$

However, it is not clear how to choose the perturbation  $\tilde{m}$  so that the energy is decreased. The reason is that  $\delta E$  depends not only on  $\tilde{m}$ , but also on the perturbations in  $\tilde{\phi}$  and  $\tilde{T}$ , and the latter two variations depend on  $\tilde{m}$  implicitly according to (18) and (19).

To make  $\tilde{m}$  explicit in the above energy perturbation, we use integration by parts and introduce adjoint equations. First, introducing a Lagrange multiplier  $\lambda_1$  for equation (18), multiplying the equation by  $\lambda_1$  and integrating over  $\tilde{\Omega}$ , we have

$$\begin{aligned} 0 &= \int_{\tilde{\Omega}} (\lambda_1, u\lambda_1, v\lambda_1) \cdot \nabla \tilde{\phi} + \int_{\tilde{\Omega}} \lambda_1 \tilde{v}\phi_\theta \\ &= \int_{\Omega} \lambda_1 \tilde{\phi}|_{z=z_f} + \int_z \int_{\partial\Omega} (\mathbf{u} \cdot \mathbf{n}) \lambda_1 \tilde{\phi} - \int_{\tilde{\Omega}} \nabla \cdot (\lambda_1, u\lambda_1, v\lambda_1) \tilde{\phi} + \int_{\tilde{\Omega}} \lambda_1 \tilde{v}\phi_\theta. \end{aligned} \quad (21)$$

Similarly, we introduce another Lagrange multiplier  $\lambda_2$  for equation (19). This gives

$$0 = \int_{\Omega} \lambda_2 \tilde{T}|_{z=z_f} + \int_z \int_{\partial\Omega} (\mathbf{u} \cdot \mathbf{n}) \lambda_2 \tilde{T} - \int_{\tilde{\Omega}} \nabla \cdot (\lambda_2, u\lambda_2, v\lambda_2) \tilde{T} + \int_{\tilde{\Omega}} \lambda_2 \left\{ \tilde{v}T_\theta + \frac{\tilde{m}}{c \cos \theta} \right\}. \quad (22)$$

To eliminate  $\tilde{\phi}$  and  $\tilde{T}$  in  $\delta E$ , we first multiply equations (21) and (22) by  $\epsilon$  and  $\beta\epsilon$ , respectively, and then add them to equation (20). This gives

$$\begin{aligned} \delta E &= \epsilon \left[ \int_{\Omega} \tilde{\phi}(\phi - \phi^*)|_{z=z_f} + \int_z \int_{\partial\Omega} (\mathbf{u} \cdot \mathbf{n}) \tilde{\phi}(\phi - \phi^*) + \frac{1}{2} \int_z \int_{\partial\Omega} (\tilde{\mathbf{u}} \cdot \mathbf{n})(\phi - \phi^*)^2 \right. \\ &\quad + \beta \int_{\Omega} \tilde{T}(T - T^*)|_{z=z_f} + \beta \int_z \int_{\partial\Omega} (\mathbf{u} \cdot \mathbf{n}) \tilde{T}(T - T^*) + \frac{\beta}{2} \int_z \int_{\partial\Omega} (\tilde{\mathbf{u}} \cdot \mathbf{n})(T - T^*)^2 \\ &\quad + \int_{\Omega} \lambda_1 \tilde{\phi}|_{z=z_f} + \int_z \int_{\partial\Omega} (\mathbf{u} \cdot \mathbf{n}) \lambda_1 \tilde{\phi} - \int_{\tilde{\Omega}} \nabla \cdot (\lambda_1, u\lambda_1, v\lambda_1) \tilde{\phi} + \int_{\tilde{\Omega}} \lambda_1 \tilde{v}\phi_\theta \\ &\quad + \beta \int_{\Omega} \lambda_2 \tilde{T}|_{z=z_f} + \beta \int_z \int_{\partial\Omega} (\mathbf{u} \cdot \mathbf{n}) \lambda_2 \tilde{T} - \beta \int_{\tilde{\Omega}} \nabla \cdot (\lambda_2, u\lambda_2, v\lambda_2) \tilde{T} \\ &\quad \left. + \beta \int_{\tilde{\Omega}} \lambda_2 \left\{ \tilde{v}T_\theta + \frac{\tilde{m}}{c \cos \theta} \right\} \right] + O(\epsilon^2). \end{aligned} \quad (23)$$

Next, we choose  $\lambda_1$  and  $\lambda_2$  satisfying

$$(\lambda_1)_z + (u\lambda_1)_x + (v\lambda_1)_\theta = 0, \quad (\lambda_2)_z + (u\lambda_2)_x + (v\lambda_2)_\theta = 0, \quad (24)$$

with the initial conditions on  $z = z_f$ ,

$$\lambda_1(z = z_f) = \phi^* - \phi \quad \text{and} \quad \lambda_2(z = z_f) = T^* - T, \quad (25)$$

and the boundary conditions,

$$\begin{aligned} \lambda_1|_{\partial\Omega} &= \begin{cases} \phi^* - \phi & \text{if } (\mathbf{u} \cdot \mathbf{n}) > 0, \\ \text{no b.c. needed} & \text{if } (\mathbf{u} \cdot \mathbf{n}) \leq 0; \end{cases} \\ \lambda_2|_{\partial\Omega} &= \begin{cases} T^* - T & \text{if } (\mathbf{u} \cdot \mathbf{n}) > 0, \\ \text{no b.c. needed} & \text{if } (\mathbf{u} \cdot \mathbf{n}) \leq 0. \end{cases} \end{aligned} \quad (26)$$

Ignoring terms of higher than linear order in the above equation for  $\delta E$ , we have the leading order term in the energy perturbation which is given by

$$\begin{aligned} \delta E &= \epsilon \int_{\tilde{\Omega}} \left[ \tilde{v}(\lambda_1\phi_\theta + \beta\lambda_2T_\theta) + \frac{\beta\lambda_2\tilde{m}}{c \cos \theta} \right] + \frac{\epsilon}{2} \int_{\Omega_p} \tilde{v}(\phi - \phi^*)^2 \Big|_{\theta=\theta_{\min}}^{\theta=\theta_{\max}} \\ &\quad + \frac{\epsilon\beta}{2} \int_{\Omega_p} \tilde{v}(T - T^*)^2 \Big|_{\theta=\theta_{\min}}^{\theta=\theta_{\max}} \end{aligned}$$

$$\begin{aligned}
 &= -\epsilon \int_{\Omega_p} \left\{ \tilde{m}_x \left[ \int_{\theta} \lambda_1 \phi_{\theta} + \beta \lambda_2 T_{\theta} \right] - \tilde{m}_z \left[ \int_{\theta} \tan \theta (\lambda_1 \phi_{\theta} + \beta \lambda_2 T_{\theta}) \right] \right. \\
 &\quad \left. - \frac{\tilde{m} \beta}{c} \left[ \int_{\theta} \frac{\lambda_2}{\cos \theta} \right] - \frac{1}{2} \int_{\Omega_p} \tilde{v} [(\phi - \phi^*)^2 + \beta(T - T^*)^2] \Big|_{\theta=\theta_{\min}}^{\theta=\theta_{\max}} \right\} \\
 &= \epsilon \int_{\Omega_p} \tilde{m} g,
 \end{aligned} \tag{27}$$

where

$$\begin{aligned}
 g(x, z) &= [f_1(x, z)]_x - [f_2(x, z)]_z + \frac{\beta}{c} f_3(x, z) + f_4(x, z), \\
 f_1(x, z) &= \int_{\theta} \lambda_1 \phi_{\theta} + \beta \lambda_2 T_{\theta}, \quad f_2(x, z) = \int_{\theta} \tan \theta (\lambda_1 \phi_{\theta} + \beta \lambda_2 T_{\theta}), \\
 f_3(x, z) &= \int_{\theta} \frac{\lambda_2}{\cos \theta}, \\
 f_4(x, z) &= \frac{1}{2} \left\{ \frac{\partial}{\partial x} [(\phi - \phi^*)^2 + \beta(T - T^*)^2] - \tan \theta \frac{\partial}{\partial z} [(\phi - \phi^*)^2 + \beta(T - T^*)^2] \right\} \Big|_{\theta=\theta_{\min}}^{\theta=\theta_{\max}}.
 \end{aligned} \tag{28}$$

To minimize the energy using the method of gradient descent, one could choose the perturbation  $\tilde{m} = -g$ . This implies

$$\delta E = -\epsilon \int_{\Omega_p} \tilde{m}^2 \leq 0, \tag{29}$$

and the equality holds when  $\|\tilde{m}\| = 0$ . However, it is not straightforward how one can guarantee the following two properties:

- (1)  $\tilde{m}|_{\partial\Omega_p} = 0$ ,
- (2)  $m^{k+1} = m^k + \epsilon \tilde{m}^k$  smooth.

The first condition assumes that we can measure  $m$  on the boundary  $\partial\Omega_p$ , denoted by  $m^*|_{\partial\Omega_p}$ , which is a reasonable assumption. This means that the variation of the velocity function near the boundary should be zero.

The second condition is a regularity condition imposed on  $m^k$  because of the ill-posed nature of the transmission tomography problem, caused partly by the use of incomplete and imprecise data sets having limited ray-path coverage and aperture. Theoretically, the Tikhonov regularization can be used to bound the growth of solutions so as to overcome ill posedness and have a conditionally well-posed problem. Computationally, regularization will allow large step sizes,  $dz$ , in solving the advection equations (14) and the conservation laws (24). In general, one only needs  $m^k \in C^1$  to guarantee well posedness of state equations (14). However, assuming that one uses  $\tilde{m}^k = -g$  directly, it is not clear whether this function  $g$  would give us the desired regularity. Even if  $g \in C^1$ , the numerical solution of  $g$  could have jumps or spikes. These irregularities will force one to pick a very small step size,  $\epsilon^k$ , in the minimization process, and/or a very restrictive step size,  $dz$ , in solving the advection equations (14) and the conservation laws (24). Therefore, we regularize the problem at each iteration.

One way to satisfy the above two properties is to use the descent direction

$$\tilde{m} = -(I - \nu \Delta)^{-1} g, \tag{30}$$

where  $I$  is the identity operator,  $\Delta$  is the Laplacian operator and  $\nu \geq 0$  controls the amount of regularity that one wants. The homogeneous boundary condition is imposed when inverting

the operator  $(I - \nu\Delta)$ . With this  $\tilde{m}$ , we have

$$\delta E = -\epsilon \int_{\Omega_p} (|\tilde{m}|^2 + \nu|\nabla\tilde{m}|^2) \leq 0, \quad (31)$$

and the equality is achieved when  $\tilde{m} \equiv 0$ .

Here, we briefly justify the above regularization. Assume that  $g \in L^2(\Omega_p)$ . Consider the operator equation

$$-\tilde{m} = g \quad (32)$$

in  $H_0^1(\Omega_p)$ . Apparently this equation is ill posed. To use the Tikhonov regularization method [48, 16], we constrain  $\tilde{m}$  to be bounded in  $H_0^1(\Omega_p)$ ; then by the Poincare inequality, we only need to bound  $\nabla\tilde{m}$  in the  $L^2$ -norm.

We have the following theorem.

**Theorem 1.** Consider equation (32) in

$$\tilde{\mathcal{M}}_{ad} = \{\tilde{m} : \|\nabla\tilde{m}\|_{L^2} \leq B, \tilde{m} \in H_0^1(\Omega_p)\},$$

where  $B$  is a positive constant. The following functional:

$$J(\tilde{m}) = \|-\tilde{m} - g\|_{L^2}^2 + \nu\|\nabla\tilde{m}\|_{L^2}^2 \quad (33)$$

has a unique minimizer in  $\tilde{\mathcal{M}}_{ad}$  satisfying

$$-(I - \nu\Delta)\tilde{m} = g, \quad (34)$$

where  $\nu \geq 0$  is the so-called regularization parameter.

**Proof.** It is easy to verify that  $\tilde{\mathcal{M}}_{ad}$  is a closed, convex bounded set in  $H_0^1(\Omega_p)$ . The Tikhonov regularization reduces solving equation (32) in  $\tilde{\mathcal{M}}_{ad}$  to minimize the functional  $J(\tilde{m})$ . Then the variational principle applied to  $J(\tilde{m})$  yields equation (34). Because  $J(\tilde{m})$  is a quadratic functional, according to standard theory  $J(\tilde{m})$  has a unique minimizer given by condition (34).  $\square$

Determining an optimal regularization parameter is one of the crucial components in applying regularization methods. One simple way is to fix this parameter in all iterations. In most examples below, we determine this regularization parameter by trial and error. However, as  $m^k$  is getting closer to a minimizer, it is not necessary to have large regularization. Therefore, one may try to decrease the magnitude of  $\nu$  as  $k$  increases.

Some additional remarks are in order here. Theoretically, we may define some functional spaces and inner products, introduce some operator notations, derive the corresponding adjoint operators and gradient operators accordingly; however, here we prefer the above formal derivation which we think is much more transparent. Moreover, we have to justify the Frechet differentiability of the nonlinear functional, which in general is not an easy task. Finally, we may also carry out the sensitivity analysis of the underlying forward nonlinear operator. The above issues will be addressed in a forthcoming paper.

#### 4. An algorithm and some implementation details

According to the above formulation, we have the following algorithm.

Algorithm:

1. Initialize  $m = m^k$  for  $k = 0$ .

2. Obtain  $\phi(x, \theta, z)$  and  $T(x, \theta, z)$  by solving (14), with the velocity  $m = m^k$  using the initial and boundary conditions (16).
3. Obtain  $\lambda_1(x, \theta, z)$  and  $\lambda_2(x, \theta, z)$  by solving (24) with the velocity  $m = m^k$  using the initial condition (25) and the boundary condition (26).
4. Compute  $g$  using (28).
5. Obtain  $\tilde{m}^k(x, z)$  by solving (30).
6. Update

$$m^{k+1} = m^k + \epsilon^k \tilde{m}^k. \tag{35}$$

7. Go back to step 2 until  $\|\tilde{m}^k\| \leq \delta$  or  $k \geq k_{\max}$ , where  $\delta$  is a small positive number and  $k_{\max}$  is a large positive integer.

In the above algorithm, one may improve the computational efficiency by using, for example, the Armijo–Goldstein rule in picking  $\epsilon^k$ , or the BFGS method to replace the method of gradient descent. However, to simplify the presentation we use the method of gradient descent to demonstrate the effectiveness of the new formulation; thus we simply choose  $\epsilon^k = \epsilon$  to be a sufficiently small constant.

The initial guess for  $m$ , denoted by  $m^0$ , has to satisfy the conditions

- (1)  $m^0|_{\partial\Omega_p} = m^*|_{\partial\Omega_p}$ ,
- (2)  $m^0$  smooth.

One way to meet these conditions is to solve

$$(I - \nu\Delta)c^0 = 0, \tag{36}$$

with the boundary condition  $c^0|_{\partial\Omega_p} = c^*|_{\partial\Omega_p}$ , and set  $m^0 = \log c^0$ . Or,  $m^0$  can directly be obtained by solving  $(I - \nu\Delta)m^0 = 0$  with the boundary condition  $m^0|_{\partial\Omega_p} = \log c^*|_{\partial\Omega_p}$ .

Yet another possibility is to initialize it by solving a transmission tomography problem based on the first-arrival time only, for example, as in [44, 45]. By using this approach, the initial guess  $m^0$  should be closer to the minimizer of (17) than that from solving the above elliptic equation. See [30] for a detailed implementation based on fast sweeping methods [52] and adjoint state methods.

In step 2 and step 3 in the above algorithm, we need to solve two advection equations and two conservation laws. They are solved by the third-order weighted essentially non-oscillatory (WENO) scheme [27] in space and the second-order total variation diminishing (TVD) Runge–Kutta (RK) method in time [46]. The complexity of these schemes is  $O(M * N)$ , where  $N$  is the total number of mesh points in the  $(x, \theta)$ -space and  $M$  is the total number of steps in the  $z$ -direction, since these schemes are explicit. In step 5, we solve the Helmholtz equation (30) by FFT.

## 5. Practical issues and null spaces

### 5.1. Data measurement

On the final level  $z = z_f$ , we have assumed that data  $\phi(z_f, x, \theta)$  and  $T(z_f, x, \theta)$  for  $(x, \theta) \in \Omega$  are available in the reduced phase space. It seems that one cannot have such information since one can only measure data in physical space. Actually, this issue has been addressed successfully in [14] as mentioned earlier. To be self-contained, we give a short explanation of this.

We recall that data  $\phi(z_f, x, \theta)$  ( $(x, \theta) \in \Omega$ ) and  $T(z_f, x, \theta)$  ( $(x, \theta) \in \Omega$ ) consist of the arriving ray information at the receiver  $x$  (equivalent to the geophone ‘g’ in [14]) with

the arriving angle  $\theta$  (equivalent to the ray parameter ‘p’ in [14]):  $\phi(z_f, x, \theta)$  recording the initial source location (equivalent to the shot location ‘s’ in [14]) and  $T(z_f, x, \theta)$  recording the transmission traveltime of the arriving ray (analogous to the reflection traveltime ‘t’ in [14]).

How to obtain these data? For a given source location  $x_s$ , the multivalued traveltimes recorded at the receivers in between  $x_{\min}$  and  $x_{\max}$  can be collected as demonstrated in [14]. Thus, firstly we consider the (multivalued) traveltime data collected at the receivers on  $z = z_f$  due to a single point source at  $(x, z) = (x_s, 0)$ ; typically we have a multibranch curve as illustrated in figure 1(c). According to such a multivalued traveltime curve, we can partition the curve into single-valued branches at caustics [40, 41]. At each point  $x$  along each branch we can compute the partial derivative  $T_x(x)$ , and the arrival angles along that branch are computed by

$$\tan \theta^*(x) = \frac{T_x(x)}{\sqrt{c(x, z_f)^{-2} - T_x(x)^2}}, \quad (37)$$

where  $c$  is given along the boundary as we have assumed; next we assign values to the level set function  $\phi$  and the traveltime function  $T$  on  $z = z_f$  according to

$$\phi^*(z_f, x, \theta^*(x)) = x_s, \quad (38)$$

$$T^*(z_f, x, \theta^*(x)) = T^*(x); \quad (39)$$

the above process is repeated for each branch. In this way, the multivalued traveltimes associated with a single source  $x_s$  are assigned to functions  $\phi(z_f, x, \theta)$  and  $T(z_f, x, \theta)$  for  $(x, \theta) \in \Omega$  in the  $x$ - $\theta$  space, namely, the receiver arrival-angle space.

Secondly, we can do the same for other source locations as for the single point source  $(x, z) = (x_s, 0)$  because our formulation for the forward problem allows multiple point sources to be treated simultaneously, as seen in the initial condition (15).

Thirdly, we may apply similar procedures to receivers on  $x = x_{\min}$  and  $x = x_{\max}$ .

We note that the above collected data are not necessarily on the mesh points of the discretization of the  $x$ - $\theta$  space. However, since functions  $\phi$  and  $T$  are smooth in the  $x$ - $\theta$  space at any  $z$  if the given velocity model is smooth, we can always interpolate the numerical solution on a regular Cartesian mesh onto those irregular sampling points in the  $x$ - $\theta$  space. Mathematically, this just needs an additional interpolation operator; numerically, this will not introduce any difficulty in implementation.

## 5.2. Avoiding use of outflow boundary measurements

In the formulation for the inverse problem, we have assumed that the outflow boundary conditions are available. Depending on applications, we may or may not have such data available.

In transient elastography [26, 33], for example, the interior measurement of the multivalued traveltimes is possible so that the outflow boundary measurements,  $\phi(z, x, \theta)|_{(x, \theta) \in \partial\Omega}$  and  $T(z, x, \theta)|_{(x, \theta) \in \partial\Omega}$ , are available.

In transmission traveltime tomography the outflow boundary measurements are typically not available, but we can avoid using these data by taking advantage of finite speed of wave propagation. Consequently, we assume that we only have sources on  $z = 0$  and receivers on  $z = z_f$ . To tackle the inverse problem in this new setup, we propose the following strategy.

We denote the measurements on the final level  $z = z_f$  in the reduced phase space by the set

$$\mathcal{C}(z_f) = \{(x, \theta) : \text{both } \phi(z_f, x, \theta) \text{ and } T(z_f, x, \theta) \text{ are known}\}.$$

Next, we define  $\mathcal{C}(z)$  as the collection of all characteristic curves in phase space obtained by tracing back all arrival rays from the set  $\mathcal{C}(z_f)$  according to the Liouville equation. Because of the paraxial assumption, any characteristic from  $\mathcal{C}(z)$  has to satisfy

$$-\pi/2 < -\theta_{\max} < \theta(z) < \theta_{\max} < \pi/2$$

for all  $z \in [0, z_f]$ . This means that characteristics from  $\mathcal{C}(z)$  will never touch the surfaces  $\{(z, x, \theta) : 0 \leq z \leq z_f, x_{\min} \leq x \leq x_{\max} \text{ and } \theta = \pm\pi/2\}$ .

Because the Liouville equation is a linear hyperbolic equation, characteristics will never intersect. By finite speed of wave propagation for hyperbolic problems, we can use an enlarged numerical domain in computation so that boundary conditions (16) and (26) imposed on  $\phi, T, \lambda_1$  and  $\lambda_2$  will not interfere with solutions in  $\mathcal{C}(z)$  at all. In the implementation, one could, for example, numerically set  $(\mathbf{n} \cdot \nabla \phi)|_{\partial\Omega} = 0$  for the level set equation, if a boundary condition on  $\partial\Omega$  is necessary.

### 5.3. A modified energy

Next, we need to modify our energy. Because all characteristics from  $\mathcal{C}(z)$  will not touch  $\partial\Omega$ , our original energy reduces to

$$E_{\text{new}}(m) = \frac{1}{2} \int_{\Omega} (\phi - \phi^*)^2 \delta(\mathcal{C}(z))|_{z=z_f} + \frac{\beta}{2} \int_{\Omega} (T - T^*)^2 \delta(\mathcal{C}(z))|_{z=z_f}, \quad (40)$$

where  $\delta(\cdot)$  is Dirac's delta function.

We need to make a few changes to incorporate these modifications into our implementation. For the forward problem, we can still keep both paraxial Liouville equations (14) and their corresponding initial conditions. Furthermore, because we have enlarged our computational domain, the boundary conditions for these equations will not affect the energy that we are going to minimize. If necessary, as mentioned above, we could numerically impose Neumann boundary conditions.

The adjoint equations are still the same as in (24). The corresponding initial conditions (25), however, are modified to incorporate the measurements into the formulation,

$$\lambda_1(z = z_f) = (\phi^* - \phi)\delta(\mathcal{C}(z_f)) \quad \text{and} \quad \lambda_2(z = z_f) = (T^* - T)\delta(\mathcal{C}(z_f)). \quad (41)$$

The required boundary conditions are imposed in the same way as done for the state equations. Numerically, we could approximate the  $\delta$ -function by the following smoothed version as in the standard level set method:

$$\delta_{\epsilon}(x) = \frac{1}{4\epsilon} [1 - \text{sign}(|x| - \epsilon)] \left[ 1 + \cos\left(\frac{\pi x}{\epsilon}\right) \right], \quad (42)$$

for some small  $\epsilon > 0$ . Therefore, we are smearing the set of measurements  $\mathcal{C}(z_f)$ . The characteristics in  $\mathcal{C}(z)$  now become tubes with radius  $\epsilon$  in the reduced phase space.

The gradient of this new energy,  $g_{\text{new}}(x, z)$ , is also similar to the expression as before. Following the same approach as in section 3, we have

$$g_{\text{new}}(x, z) = [f_1(x, z)]_x - [f_2(x, z)]_z + \frac{\beta}{c} f_3(x, z), \quad (43)$$

where functions  $f_1, f_2$  and  $f_3$  are defined in (28).

In summary we have the following algorithm without using the outflow boundary conditions.

Practical Algorithm:

1. Initialize  $m = m^k$  for  $k = 0$ .
2. Obtain  $\phi(x, \theta, z)$  and  $T(x, \theta, z)$  by solving (14), with the velocity  $m = m^k$  using the initial conditions as in (16) and Neumann boundary conditions.
3. Obtain  $\lambda_1(x, \theta, z)$  and  $\lambda_2(x, \theta, z)$  by solving (24) with the velocity  $m = m^k$  using the initial conditions (41) and Neumann boundary conditions.
4. Compute  $g_{\text{new}}(x, z)$  using (43).
5. Obtain  $\tilde{m}^k(x, z)$  by solving (30).
6. Update

$$m^{k+1} = m^k + \epsilon^k \tilde{m}^k. \quad (44)$$

7. Go back to step 2 until  $\|\tilde{m}^k\| \leq \delta$  or  $k \geq k_{\text{max}}$ , where  $\delta$  is a small positive number and  $k_{\text{max}}$  is a large positive integer.

#### 5.4. Null spaces in transmission traveltome tomography

In cross-well transmission tomography, many features are poorly resolved due to limited ray-path coverage. To investigate the nature of this indeterminacy, it is vital to identify the null space for linearized tomography, namely, the class of model perturbation of a background medium which to first order causes no perturbation at all in the cross-well transmission (first-arrival and/or multi-arrival) traveltome data [7].

In the setting of first-arrival transmission traveltome tomography, one can easily construct different slowness models to give the same first-arrival traveltimes for all source–receiver combinations; see [25]. In a linearized continuum formulation of the transmission tomography problem, the difference between two such slowness models is a model perturbation in the null space of the derivative operator; since this operator maps small perturbations in the model slowness into the resulting small changes in traveltome, a model perturbation in the null space results in an identically zero (linearized) first-arrival traveltome perturbation [7]. In [7] by linearizing the forward problem of the ray-tracing-based first-arrival traveltome tomography about a homogeneous background medium, Bube and Meadows have identified polynomial slowness perturbations in the null space of the corresponding derivative operator and the ‘edge effect’ slowness perturbations in the null space if there are regions near the boundary of the model through which no rays pass. In practice, model discretization can remove some elements of the null space, but the slowness perturbations in the null space remain poorly resolved, even with perfect data recorded over wide apertures. Because the ray paths in a cross-well experiment are more horizontal than vertical, most of these poorly resolved features result in a degradation of the horizontal resolution relative to the vertical [7].

In the setting of multi-arrival-based traveltome tomography, to our knowledge no work has been done to identify the null space of the derivative operator of the corresponding forward problem. Although the edge effect slowness perturbations are in the null space if there are regions near the boundary of the model through which no rays pass, we need to analyse whether polynomial slowness perturbations are in the null space and, if they are, how large the subspace is of the set of all slowness perturbations.

## 6. Examples with ideal illumination

In the following examples, both the parameters  $\beta$  and  $\nu$  are set to be equal to 1 unless otherwise specified. This means that the regularization for the velocity comes from the operator  $(I - \nu\Delta)$ .

The physical domain  $\Omega_p$  is defined by  $(x_{\min}, x_{\max}) = (-1, 1)$  and  $(z_0, z_f) = (0, 2)$ . For the constant model, we use  $(\theta_{\min}, \theta_{\max}) = (-9\pi/20, 9\pi/20)$ . For the waveguide and Gaussian example, we use  $(\theta_{\min}, \theta_{\max}) = (-\pi/3, \pi/3)$ . The initial guess of the model  $m^0$  is determined by solving equation (36) with FFT.

The inflow boundary conditions for  $\phi$  and  $T$ , represented by  $\phi^*$  and  $T^*$ , respectively, are obtained by solving equations (14) using a semi-Lagrangian method, as in [31]; here the characteristic system is solved using the MATLAB function `ode45`.

The outflow boundary conditions  $\phi^*$  and  $T^*$  are obtained by solving the system (14) using the exact velocity  $c$  with the third-order WENO scheme in space and the second-order TVD-RK stepping in time.

Therefore, we assume that we have ideal illumination of the to-be-imaged domain.

### 6.1. Constant velocity model

The exact velocity model is given by  $c \equiv 1$ . In this case, the boundary measurements of  $m^*$  are simply given by

$$m^*|_{\partial\Omega_p} = 0. \quad (45)$$

The number of grid points in this test case is  $129 \times 129 \times 129$  in the  $x$ - $\theta$ - $z$  space.

Figure 3 shows the comparison of the numerical solution and the exact solution, the relative error in  $c^{500}$  and the convergence history of energy during the iteration. The results validate the effectiveness and the convergence of the algorithm.

We note that the large relative errors near  $z = 0$  in figure 3(c) are caused by the edge effect slowness perturbation in the null space because there are some regions near  $z = 0$  through which no rays pass due to the aperture limitation introduced by the paraxial formulation.

### 6.2. Waveguide velocity model

The exact velocity model is given by

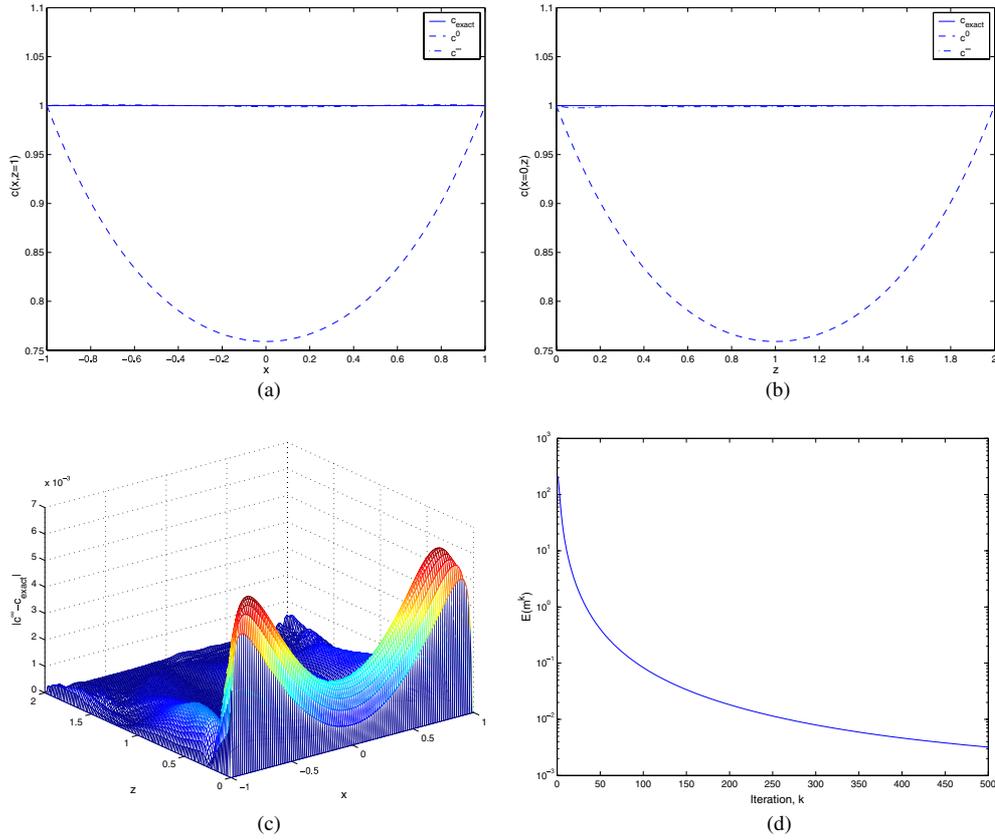
$$c(x, z) = 3.0 - 2.5 \exp\left(-\frac{x^2}{2}\right). \quad (46)$$

The number of grid points is  $129 \times 129 \times 129$  in the  $x$ - $\theta$ - $z$  space.

Figure 4 shows the comparison of the numerical solution and the exact solution, the relative error in  $c^{100}$ , and the convergence history of energy during the iteration. The results validate the effectiveness and the convergence of the algorithm.

Figure 5 shows the contour plots of the level set function  $\phi(z = 2, x, \theta)$  using  $c = c^0$  and the exact velocity model; the dashed line represents the zero level set. An essential part of the transmission tomography problem is to drive the level sets with  $c = c^0$  to match with the level sets with the exact velocity model  $c$ . Multivalued arrival angles in  $\theta$  are clearly seen, and this implies that the corresponding traveltimes are also multivalued. For example, for the exact velocity field  $c$ , if we look at rays from the source  $(x_s, z_s) = (0, 0)$ , then we have three arrivals at most locations on the level  $z = 2$ . Even though multivalued solutions are obtained from the initial guess, the shape of the level sets is quite different from that produced by the exact velocity model.

We note that the large relative errors near  $z = 0$  in figure 4(c) are caused by the edge effect slowness perturbation in the null space because there are some regions near  $z = 0$  through which no rays pass due to the aperture limitation introduced by the paraxial formulation.



**Figure 3.** (Constant model) (a) and (b) slices of the initial guess, the final approximated  $c$  and the exact  $c$ ; (c) the relative error in the solution; (d) the convergence history of energy in semi-log scale.

On the other hand, comparing figures 4(a) and (b), we note that the  $x$ -direction variations of the velocity model are better resolved than the  $z$ -direction variations; this is because the ray paths in the transmission experiment are more horizontal in the  $z$ -direction than vertical in the  $x$ -direction. Most of these poorly resolved features in the null space result in a degradation of the  $z$ -direction resolution, which is analogous to the first-arrival traveltimes tomography [7].

### 6.3. Gaussian model

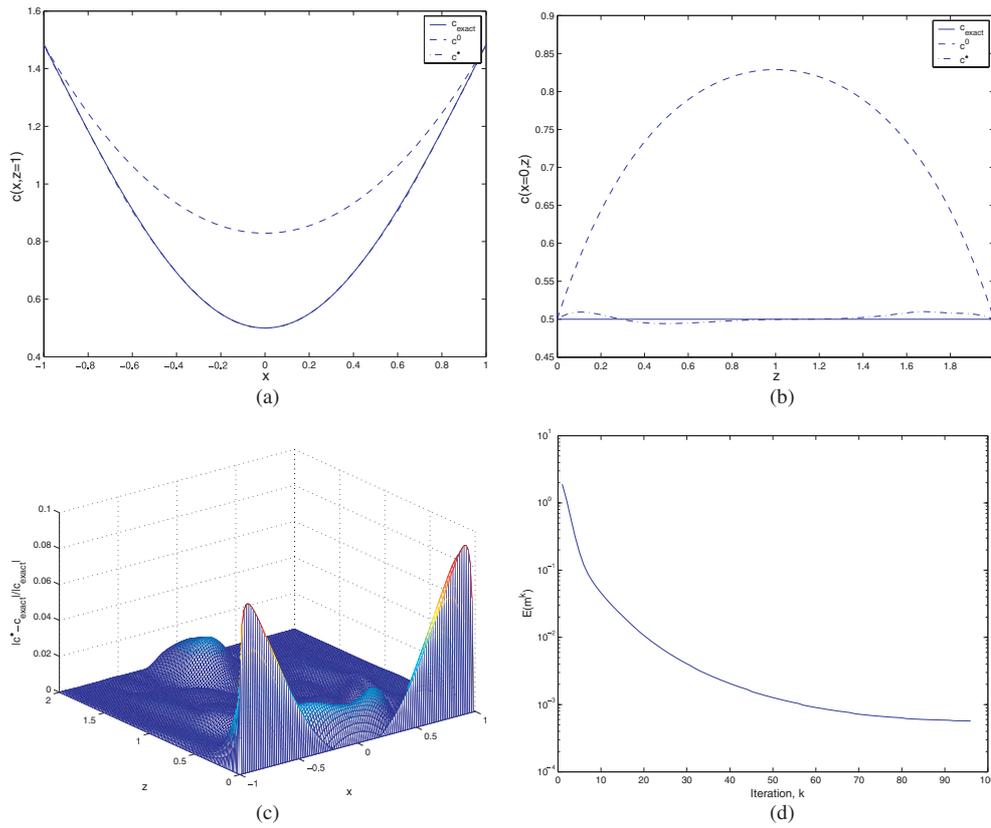
The exact velocity model is given by

$$c(x, z) = 3 - \frac{1}{2} \exp\left(-\frac{x^2 + (z - 0.5)^2}{0.5^2}\right) - \exp\left(-\frac{x^2 + (z - 1.25)^2}{0.5^2}\right). \quad (47)$$

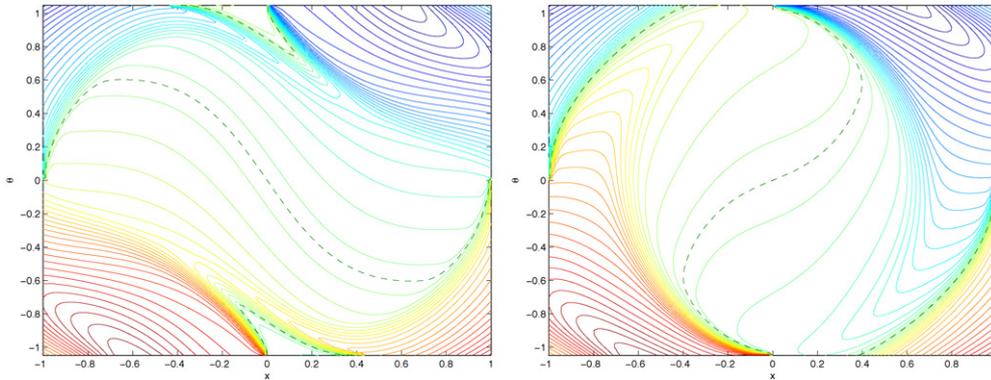
The number of grid points used is  $257 \times 129 \times 257$  in the  $x$ - $\theta$ - $z$  space.

Figure 6 shows the comparison of the numerical solution and the exact solution, the relative error in  $c^{300}$  and the convergence history of energy during the iteration. The results validate the effectiveness and the convergence of the algorithm.

Figure 7 shows the contour plots of the level set function  $\phi(z = 2, x, \theta)$  using  $c = c^0$  and the exact velocity model; the dashed line represents the zero level set.

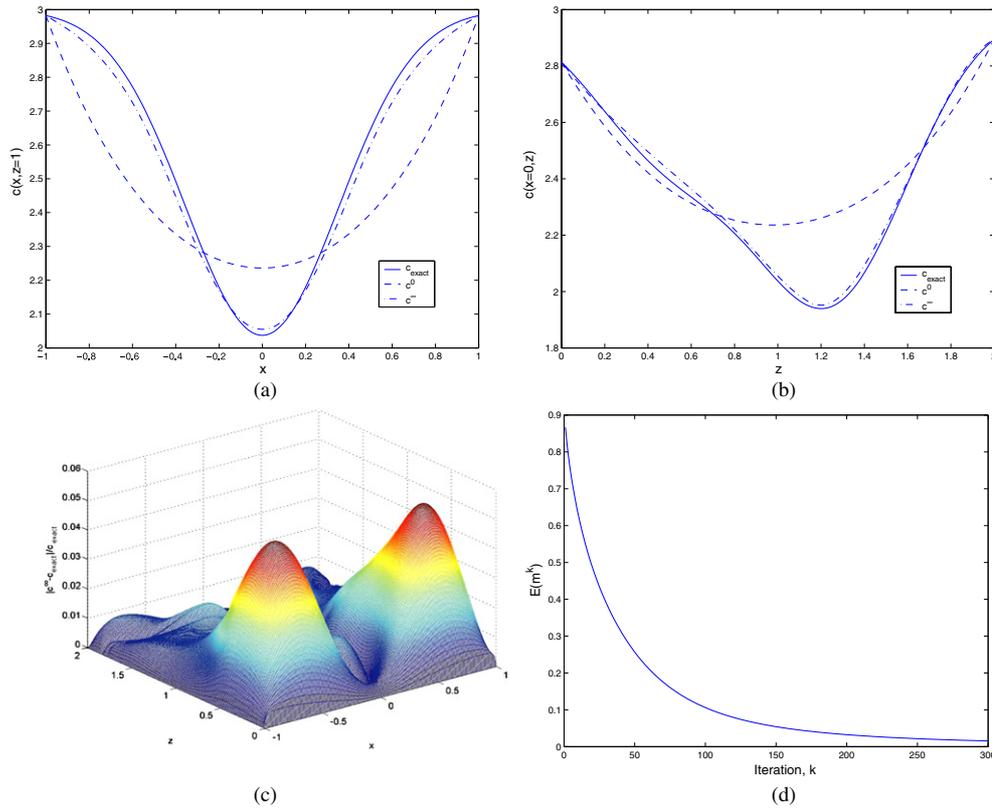


**Figure 4.** (Waveguide) (a) and (b) slices of the initial guess, the final approximated  $c$  and the exact  $c$ ; (c) the relative error in the solution; (d) the convergence history of energy in semi-log scale.

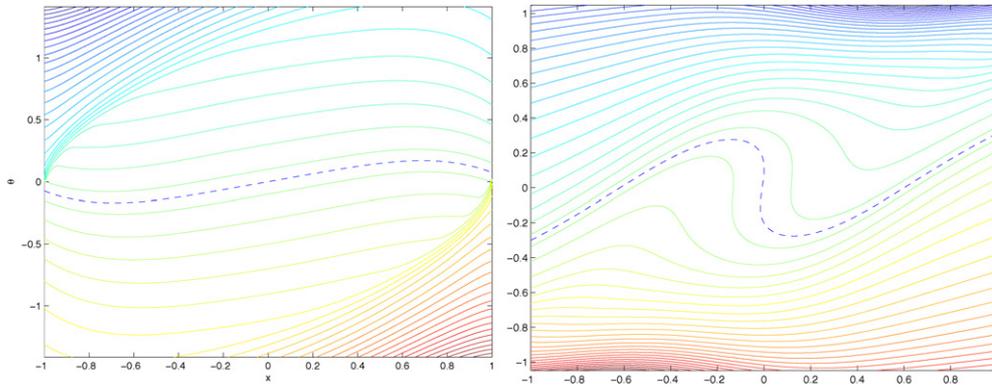


**Figure 5.** (Waveguide) The contour plot of  $\phi(z = 2, x, \theta)$  using  $c^0$  and the exact  $c$ , respectively.

To test the robustness of the new formulation, we have added the Gaussian noise up to 5% to the measured data and carried out the tomography process. The results are shown in figure 8; comparing with the results without noise, the new formulation is robust and accurate.

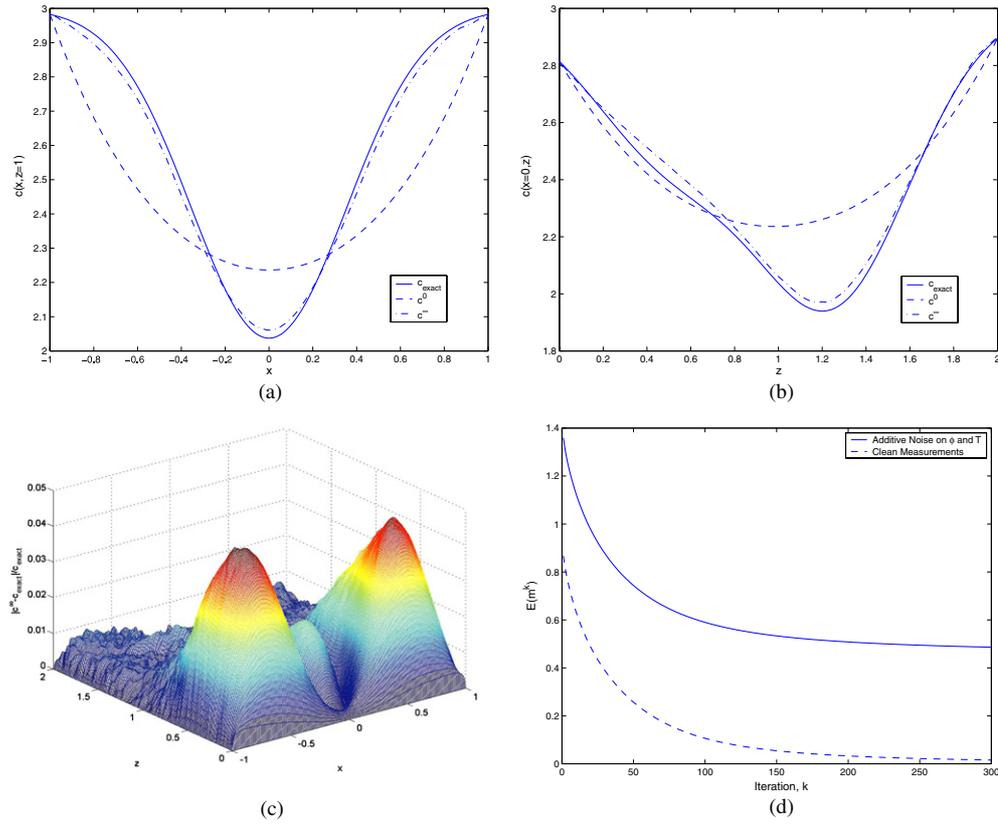


**Figure 6.** (Gaussian) (a) and (b) slices of the initial guess, the final approximated  $c$  and the exact  $c$ ; (c) the relative error in the solution; (d) the convergence history of energy in semi-log scale.



**Figure 7.** (Gaussian) The contour plot of  $\phi(z = 2, x, \theta)$  using  $c^0$  and the exact  $c$ , respectively.

Similar to constant and waveguide models, some analogous observations about the effects of the null space on the resolution of velocity models can be made about the Gaussian model.



**Figure 8.** (Gaussian model: additive Gaussian noise) (a) and (b) slices of the initial guess, the final approximated  $c$  and the exact  $c$ ; (c) the relative error in the solution; (d) the convergence history of energy in semi-log scale.

## 7. Examples with partial illumination

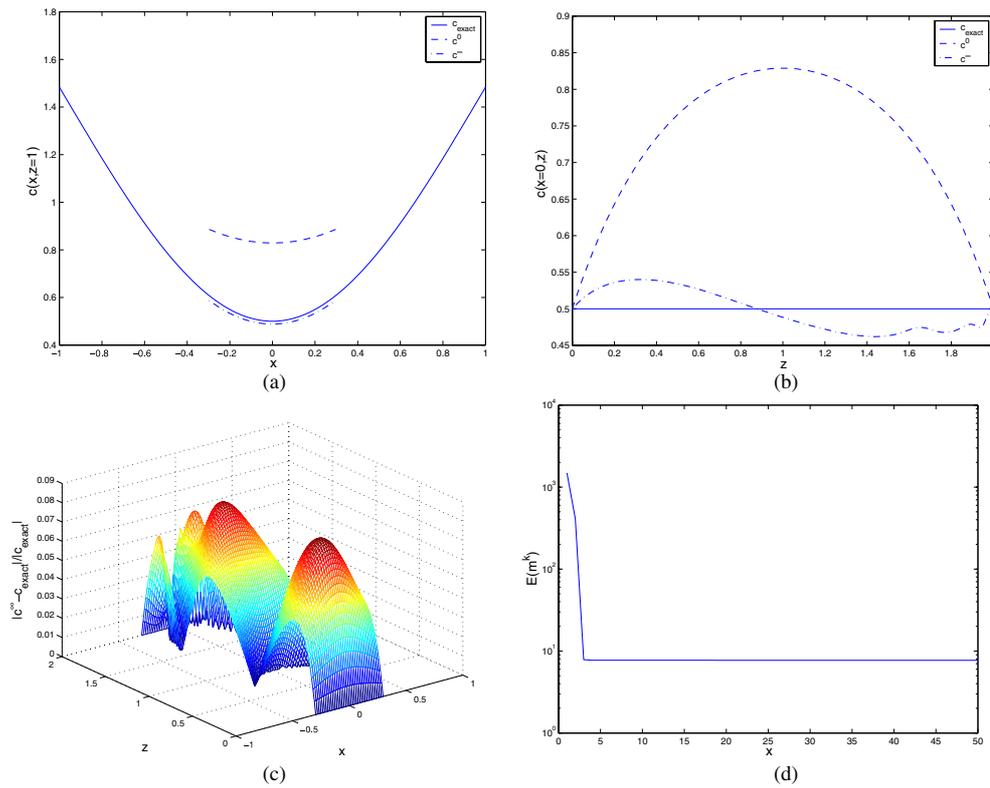
To demonstrate performance of the practical algorithm under partial illumination of the to-be-imaged domain, we set up the data collecting geometry in the following way.

Only one source is located at  $(x, z) = (0, 0)$  and five receivers are located at  $z = 2$  and  $x_r = -0.2, -0.1, 0.0, 0.1$  and  $0.2$ . The set  $\mathcal{C}(z_f)$  is obtained by solving (14) using a semi-Lagrangian method for all grid points on the level  $z = z_f$ . Then,

$$\mathcal{C}(z_f) = \left\{ (x, \theta) : \sum_{x_s, x_r} \delta_\epsilon(\phi(z_f, x, \theta) - x_s) \delta_\epsilon(x - x_r) \neq 0 \right\},$$

where  $\delta_\epsilon(\cdot)$  is defined by (42).

We use the above partial illumination to invert for the waveguide velocity model and the Gaussian velocity model. However, to solve the forward problem we use an enlarged numerical domain just as we explained in section 5. Here the enlarged numerical domain is taken to be  $(x_{\min}, x_{\max}) = (-1, 1)$ ,  $(z_0, z_f) = (0, 2)$  and  $(\theta_{\min}, \theta_{\max}) = (-\pi/3, \pi/3)$ .



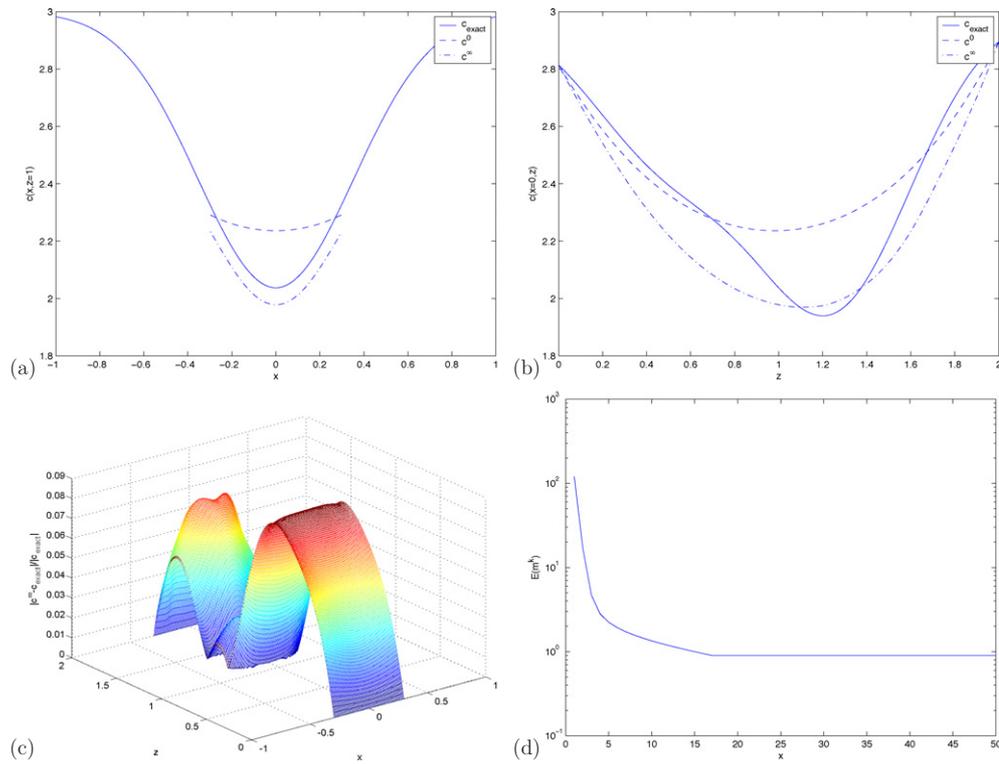
**Figure 9.** (Waveguide with partial illumination) (a) and (b) slices of the initial guess, the final approximated  $c$  and the exact  $c$ ; (c) the relative error in the solution; (d) the convergence history of energy in semi-log scale.

### 7.1. Waveguide velocity model

In this model, we discretize the enlarged domain in the reduced phase space to be an  $129 \times 129 \times 129$  mesh. To regularize the numerical gradient, we carry out the regularization step in the enlarged domain  $[-1, 1] \times [0, 2]$  with  $\nu = 1$ . However, we only extract the solution in the smaller domain,  $[-0.3, 0.3] \times [0, 2]$ . Note that all the receivers are contained in the interval  $[-0.3, 0.3]$ .

Figure 9 shows the comparison of the numerical solution and the exact solution, the relative error in  $c^{50}$  and the convergence history of energy during the iteration. The results validate the effectiveness and the convergence of the algorithm as the practical algorithm is able to invert for the velocity inside the well-illuminated domain under partial illumination.

Under partial illumination, both ray-path coverage and apertures are more limited so that the null space corresponding to the derivative operator of the corresponding forward problem is even larger. In figures 9(b) and (c), we can see the effects of the null space on the resolution of the inverted velocity model, analogous to the ideal illumination case for the waveguide model.



**Figure 10.** (Gaussian with partial illumination) (a) and (b) slices of the initial guess, the final approximated  $c$  and the exact  $c$ ; (c) the relative error in the solution; (d) the convergence history of energy in semi-log scale.

### 7.2. Gaussian velocity model

In this model we discretize the enlarged domain in the reduced phase space to be an  $257 \times 129 \times 257$  mesh. To regularize the numerical gradient, we carry out the regularization step in the enlarged domain  $[-1, 1] \times [0, 2]$  with  $\nu = 1$ . However, we only extract the solution in the smaller domain,  $[-0.3, 0.3] \times [0, 2]$ . Note that all the receivers are contained in the interval  $[-0.3, 0.3]$ .

Figure 10 shows the comparison of the numerical solution and the exact solution, the relative error in  $c^{50}$  and the convergence history of energy during the iteration. The results validate the effectiveness and the convergence of the algorithm as the practical algorithm is able to invert for the velocity inside the well-illuminated domain under partial illumination.

## 8. Conclusion

We have developed a new formulation for traveltime transmission tomography based on the paraxial Liouville equations and level set formulations. This new formulation allows us to take into account multivalued traveltimes and multipathing systematically in the tomography problem. To obtain an efficient implementation, we have used adjoint equations and the method of gradient descent. Numerical examples indicate that the new formulation is effective.

Future work includes extending the formulation to deal with reflection tomography problems, carrying out a systematic comparison between first-arrival and multi-arrival-based traveltimes tomography, and identifying the null space of the derivative operator for the forward modelling problem proposed here.

## Acknowledgment

SL is supported by ONR grant #N00014-02-1-0720. JQ is supported by NSF DMS-0542174.

## References

- [1] Bal G, Keller J B, Papanicolaou G and Ryzhik L 1999 Transport theory for acoustic waves with reflection and transmission at interfaces *Wave Motion* **30** 303–27
- [2] Berryman J 1990 Stable iterative reconstruction algorithm for nonlinear traveltimes tomography *Inverse Problems* **6** 21–42
- [3] Berryman J 2000 Analysis of approximate inverses in tomography: I. Resolution analysis of common inverses *Optim. Eng.* **1** 87–115
- [4] Berryman J 2000 Analysis of approximate inverses in tomography: II. Iterative inverses *Optim. Eng.* **1** 437–73
- [5] Bishop T N, Bube K P, Cutler R T, Langan R T, Love P L, Resnick J R, Shuey R T, Spindler D A and Wyld H W 1985 Tomographic determination of velocity and depth in laterally varying media *Geophysics* **50** 903–23
- [6] Bube K P and Langan R T 1997 Hybrid  $l^1$ - $l^2$  minimization with applications to tomography *Geophysics* **62** 1183–95
- [7] Bube K P and Meadows M A 1998 Characterization of the null space of a generally anisotropic medium in linearized cross-well tomography *Geophys. J. Int.* **133** 65–84
- [8] Cheng L-T, Kang M, Osher S, Shim H and Tsai Y-H 2004 Reflection in a level set framework for geometric optics *CMES Comput. Methods Sci. Eng.* **5** 347–60
- [9] Cheng L-T, Liu H and Osher S J 2003 High frequency wave propagation in Schrodinger equations using the level set method *Commun. Math. Sci.* **1** 593–621
- [10] Chung E, Qian J, Uhlmann G and Zhao H K 2007 A new phase space method for recovering index of refraction from travel times *Inverse Problems* **23** 309–29
- [11] Clarke R A, Alazard B, Pelle L, Sinuquet D, Lailly P, Delprat-Jannaud F and Jannaud L 2001 3D traveltimes reflection tomography with multi-valued arrivals *71st Ann. Int. Mtg., Soc. Expl. Geophys., Expanded Abstracts (Tulsa, OK)* pp 1601–4
- [12] Cockburn B, Qian J, Reitich F and Wang J 2005 An accurate spectral/discontinuous finite-element formulation of a phase-space-based level set approach to geometrical optics *J. Comput. Phys.* **208** 75–195
- [13] Dahlen F A, Hung S-H and Nolet G 2000 Frechet kernels for finite-frequency traveltimes: I. Theory *Geophys. J. Int.* **141** 157–74
- [14] Delprat-Jannaud F and Lailly P 1995 Reflection tomography: how to handle multiple arrivals? *J. Geophys. Res.* **100** 703–15
- [15] Dines K A and Lytle R J 1979 Computerized geophysical tomography *Proc. IEEE* **67** 1065–73
- [16] Engl H, Hanke M and Neubauer A 1996 *Regularization of inverse problems* (Dordrecht: Kluwer–Academic)
- [17] Engquist B and Runborg O 2003 Computational high frequency wave propagation *Acta Numer.* **12** 181–266
- [18] Engquist B, Runborg O and Tornberg A-K 2002 High frequency wave propagation by the segment projection method *J. Comp. Phys.* **178** 373–90
- [19] Fomel S and Sethian J 2002 Fast phase space computation of multiple traveltimes *Proc. Natl Acad. Sci.* **99** 7329–34
- [20] Geoltrain S and Brac J 1993 Can we image complex structures with first-arrival traveltimes *Geophysics* **58** 564–75
- [21] Gray S and May W 1994 Kirchhoff migration using eikonal equation traveltimes *Geophysics* **59** 810–7
- [22] Harris J M, Nolen-Hoeksema R C, Langan R T, Van Schaack M, Lazaratos S K and Rector J W III 1995 High-resolution crosswell imaging of a west texas carbonate reservoir: I. Project summary and interpretation *Geophysics* **60** 667–81
- [23] Hyndman D W and Harris J M 1996 Traveltimes inversion for the geometry of aquifer lithologies *Geophysics* **61** 1728–37
- [24] Isakov V 1998 *Inverse problems for partial differential equations* (New York: Springer)

- [25] Ivansson S 1986 Some remarks concerning seismic reflection tomography and velocity analysis *Geophys. J. R. Astron. Soc.* **87** 539–57
- [26] Ji L, McLaughlin J R, Renzi D and Yoon J-R 2003 Interior elastodynamics inverse problems: shear wave speed reconstruction in transient elastography *Inverse Problems* **19** S1–S29
- [27] Jiang G S and Peng D 2000 Weighted ENO schemes for Hamilton–Jacobi equations *SIAM J. Sci. Comput.* **21** 2126–43
- [28] Jin S and Osher S 2003 A level set method for the computation of multivalued solutions to quasi-linear hyperbolic PDEs and Hamilton–Jacobi equations *Commun. Math. Sci.* **1** 575–91
- [29] Leung S and Qian J 2005 A transmission tomography problem based on multiple arrivals from paraxial Liouville equations *75th Ann. Int. Mtg., Soc. Expl. Geophys., Expanded Abstracts (Tulsa, OK)* pp 2589–92
- [30] Leung S and Qian J 2006 An adjoint state method for three-dimensional transmission traveltime tomography using first-arrivals *Commun. Math. Sci.* **4** 249–66
- [31] Leung S, Qian J and Osher S 2004 A level set method for three dimensional paraxial geometrical optics with multiple sources *Commun. Math. Sci.* **2** 657–86
- [32] Liu Z and Bleistein N 1995 Migration velocity analysis: theory and an iterative algorithm *Geophysics* **60** 142–53
- [33] McLaughlin J R and Yoon J-R 2004 Unique identifiability of elastic parameters from time-dependent interior displacement measurement *Inverse Problems* **20** 25–45
- [34] Minkoff S M 1996 A computationally feasible approximate resolution matrix for seismic inverse problems *Geophys. J. Int.* **126** 345–59
- [35] Montelli R, Nolet G, Dahlen F A, Masters G, Engdahl E R and Hung S-H 2004 Finite-frequency tomography reveals a variety of plumes in the mantle *Science* **303** 338–43
- [36] Montelli R, Nolet G, Masters G, Dahlen F A and Hung S-H 2004 Global P and PP traveltime tomography: rays versus waves *Geophys. J. Int.* **158** 637–54
- [37] Operto S, Xu S and Lambare G 2000 Can we image quantitatively complex models with rays *Geophysics* **65** 1223–38
- [38] Osher S, Cheng L-T, Kang M, Shim H and Tsai Y-H 2002 Geometrical optics in a phase space based level set and Eulerian framework *J. Comput. Phys.* **179** 622–48
- [39] Qian J, Cheng L-T and Osher S J 2003 A level set based Eulerian approach for anisotropic wave propagations *Wave Motion* **37** 365–79
- [40] Qian J and Leung S 2004 A level set method for paraxial multivalued traveltimes *J. Comp. Phys.* **197** 711–36
- [41] Qian J and Leung S 2006 A local level set method for paraxial multivalued geometric optics *SIAM J. Sci. Comput.* **28** 206–23
- [42] Qian J and Symes W W 2002 Adaptive finite difference method for traveltime and amplitude *Geophysics* **67** 167–76
- [43] Qian J and Symes W W 2002 Finite-difference quasi-P traveltimes for anisotropic media *Geophysics* **67** 147–55
- [44] Sei A and Symes W W 1994 Gradient calculation of the traveltime cost function without ray tracing *65th Ann. Int. Mtg., Soc. Expl. Geophys., Expanded Abstracts (Tulsa, OK)* pp 1351–4
- [45] Sei A and Symes W W 1995 Convergent finite-difference traveltime gradient for tomography *66th Ann. Int. Mtg., Soc. Expl. Geophys., Expanded Abstracts (Tulsa, OK)* pp 1258–61
- [46] Shu C W and Osher S J 1988 Efficient implementation of essentially non-oscillatory shock capturing schemes *J. Comput. Phys.* **77** 439–71
- [47] Symes W W and Qian J 2003 A slowness matching Eulerian method for multivalued solutions of eikonal equations *J. Sci. Comput.* **19** 501–26
- [48] Tikhonov A N and Arsenin V Y 1977 *On the solution of ill-posed problems* (New York: Wiley)
- [49] Uhlmann G 2001 Travel-time tomography *J. Korean Math. Soc.* **38** 711–22
- [50] Uhlmann G and Hansen S 2002 Recovering acoustic and elastic parameters from travel-times *International Mechanical Engineering Congress and Exhibition, Proc. ASME* pp 32149
- [51] Washbourne J K, Rector J W and Bube K P 2002 Crosswell traveltime tomography in three dimensions *Geophysics* **67** 853–71
- [52] Zhao H K 2005 Fast sweeping method for eikonal equations *Math. Comput.* **74** 603–27