

Adjoint modeling for acoustic inversion^{a)}

Paul Hursky^{b)} and Michael B. Porter

Science Applications International Corporation, 10260 Campus Point Drive, San Diego, California 92121

B. D. Cornuelle, W. S. Hodgkiss, and W. A. Kuperman

Scripps Institution of Oceanography, University of California in San Diego, 9500 Gilman Drive, La Jolla, California 92093-0238

(Received 12 September 2002; accepted for publication 4 November 2003)

The use of adjoint modeling for acoustic inversion is investigated. An adjoint model is derived from a linearized forward propagation model to propagate data-model misfit at the observation points back through the medium to the medium perturbations not being accounted for in the model. This adjoint model can be used to aid in inverting for these unaccounted medium perturbations. Adjoint methods are being applied to a variety of inversion problems, but have not drawn much attention from the underwater acoustic community. This paper presents an application of adjoint methods to acoustic inversion. Inversions are demonstrated in simulation for both range-independent and range-dependent sound speed profiles using the adjoint of a parabolic equation model. Sensitivity and error analyses are discussed showing how the adjoint model enables calculations to be performed in the space of observations, rather than the often much larger space of model parameters. Using an adjoint model enables directions of steepest descent in the model parameters (what we invert for) to be calculated using far fewer modeling runs than if a forward model only were used. © 2004 Acoustical Society of America. [DOI: 10.1121/1.1636760]

PACS numbers: 43.30.Pe, 43.60.Pt, 43.60.Rw [WMS]

Pages: 607–619

I. INTRODUCTION

This paper investigates the use of adjoint modeling for acoustic inversion. Typically, in inversion problems a forward model driven by presumed medium properties is used to propagate a field from the source to the observation points, where the field predicted by the forward model is compared with the measured field there. The actual medium is not known and is typically a perturbation of the presumed medium, to be calculated by an inversion process so that subsequent forward modeling can be improved. If the properties of the medium used to drive the forward model match the actual medium properties, then the data-model misfit at the observation points is small. If not, these presumed medium properties are adjusted and the forward modeling repeated until the data-model misfit has been reduced. Note that the emphasis in this process is on the mapping from the unknown medium properties to the observations, when it is the medium properties that we want to estimate. What we really want is some way to propagate the data-model misfit from our observation points directly back to the points in the medium where we need to make adjustments. This is what an adjoint model does. It back-propagates an adjoint field, initialized to the data-model misfit at the observation points,

back to the points in the medium where a correction to the presumed medium properties can be made to reduce the data-model misfit.

An adjoint model has particular relevance to inversion problems in which the unknown parameter space is much greater than the observation space. Methods based on adjoint models have been used in geophysical inversion,¹ electromagnetic tomography² and oceanography.^{3–5} An adjoint approach to acoustic inverse scattering problems has also been explored, on a theoretical level, with emphasis on variational methods.⁶ Other recent work in acoustics includes using adjoint methods to invert for bottom boundary conditions in the ocean waveguide⁷ and using an adjoint approach with normal modes to compute derivatives with respect to environmental parameters.⁸ In this paper we show how the technique can be used to invert for medium properties in underwater acoustic problems, demonstrating an adjoint-based inversion for range-independent and range-dependent sound speed profiles.

To understand the benefits of the approach for acoustic inversion, it is useful to review the commonly used alternatives. A standard approach is to minimize a cost function that measures the difference between forward solutions of a model and observations. The most direct, brute force method is sometimes called “iterated solution of the forward problem:” one minimizes the cost function through an exhaustive search. While this method can be used for multi-parameter inversions the computational burden can be quite large. For instance, if there are 10 parameter values each sampled at 20 values, then the search involves 10^{20} runs of the forward problem. Hence, nonlinear methods involving a more strategic search such as simulated annealing,⁹ genetic algorithms¹⁰ and others¹¹ have emerged as more practical alternatives. In

^{a)}Portions of this work were presented in “Adjoint-assisted inversion for shallow water environmental parameters,” Proceedings of the Conference on Acoustic Variability, 16–20 September 2002, Lerici, La Spezia, Italy, Impact of Littoral Environmental Variability on Acoustic Predictions and Sonar Performance, edited by N. G. Pace and F. B. Jensen (Kluwer, Dordrecht, 2002). Portions of this work were also presented at the 142nd Acoustical Society Meeting, 3–7 December 2001.

^{b)}Electronic address: paul.hursky@saic.com

all of these forward solution techniques, the cost of running the forward model becomes the central factor. If the problem is only mildly nonlinear, a first attempt at decreasing this cost is to consider perturbations from known solutions so that the full cost of a massive forward search is avoided. Linear approximations to the sound pressure equation (as a function of medium properties) have been used in acoustic tomography.^{12,13} Even with this approach, computing the perturbations over the full multi-parameter space remains the dominant, limiting factor.

Solutions to inverse problems in ocean acoustics usually include *a priori* information, such as gross knowledge of the ocean environment. For example, the sound speed is typically close to 1500 m/s. In some cases the starting guess is close enough to the true solution that we are within the multi-dimensional bowl of the cost function that contains the global minimum (unfortunately, it is rarely close enough for a linear inverse to work). When this occurs, it is useful to know the gradient of the cost function with respect to the large number of unknown parameters so as to guide us to the minimum. This gradient can be obtained by operating on the data-model misfit by the adjoint of the forward problem operator.^{1,6} When the number of unknown parameters is much greater than the number of observations, the adjoint approach can be a less costly method of calculating the gradient than running the forward model for each variation in the unknown parameters.

For completeness, we mention that the derivative of a function mapping one normed vector space to another (e.g., sound speed versus range and depth to complex pressure amplitude across an array at a given frequency) is called a Fréchet derivative,¹⁴ often mentioned as the key component of the inverse procedure. The adjoint method provides a procedure to compute the Fréchet derivative.^{1,6,15} Adjoint methods also have an extensive earlier history in control theory, where they are derived from the Pontryagin principle.^{14,16–18} Although we pursue a different approach in this paper, the adjoint operator can be derived by applying variational techniques to minimize an objective function that includes the data-model mismatch and the differential equations that describe the forward model being used, using Lagrange multipliers to set the constraint that the dynamical model is perfect. This produces the Euler–Lagrange equations, which are integrated by parts to reveal the adjoint operator and the boundary conditions that it must satisfy.^{1,6,7,19} Adjoint methods have also been compared to the filtered back-propagation technique in diffraction tomography.²⁰

The adjoint operator is formulated for a linearized forward model, so its application is limited to mildly nonlinear problems. This scope of problems is similar to those to which the Born approximation can be applied: a zeroth-order forward modeling calculation is used to calculate first-order perturbations to the medium, so the actual perturbations (being solved for) should not be so great or many that they drastically change the character of the zeroth-order field. To a certain extent this limitation can be overcome by iterating the approach, as we will show in Sec. III B.

This paper demonstrates inversion for sound speed fluctuations in the water column. Medium fluctuations and the

variability they cause is a topic that is of much current interest, since variability introduces uncertainty into predictions of naval system performance. There is much ongoing work^{21–24} addressing how medium fluctuations impact our ability to predict various processes in the ocean.

Although adjoint modeling is a concept that can be applied to all the standard acoustic models, the parabolic equation has a form that best matches the development of adjoint techniques in physical oceanography. The reason is that the PE marching solution in range is analogous to the marching solution in time of ocean circulation models. In addition, the parabolic equation can model range-dependent environments, where adjoint techniques provide the most payoff (many more unknowns than measurements). Section II reviews the parabolic equation, including its tangent linear and adjoint models, and derives an iterative procedure for solving inverse problems (Sec. II C contrasts the generalized inverse solution with solutions based on the adjoint). Section III presents results of testing the inversion procedure on simulated data based on sound speed profiles measured during the INTIMATE 96 experiment, during which internal tides were observed. Section III discusses the linearity of the PE in the particular experiment configuration used, demonstrates both range-independent and range-dependent inversions, and shows how the adjoint model can be used to calculate confidence bounds for estimated parameters.

II. THE PARABOLIC EQUATION FORWARD MODEL, ITS TANGENT LINEAR MODEL, AND ITS ADJOINT

The oceanographic community is quite actively developing adjoint techniques, so the oceanographic literature provides a useful starting point for formulating an acoustic adjoint.^{4,5,19} The common oceanographic problem involves a marching forward in time of ocean temperature, currents, etc. through an ocean circulation model. One then observes the ocean conditions and seeks to understand how errors in initializing the ocean model, or in forcing it, caused those observation errors. The adjoint model directly back-propagates the data-model misfits, assuming there are no modeling errors, to provide corrections to the initial conditions and/or forcing.

In part, seeking a good acoustic analog of this process, we have in our initial work used a parabolic equation (PE) model which marches a starter acoustic pressure field forward in range from the source. The ocean and subbottom sound speed profile may be viewed as a forcing function and we will develop an adjoint model that back-propagates data-model misfits in the pressure field, thereby providing corrections to the sound speed profile that reduce these errors.

In using the PE to model acoustic propagation, modeling error is of much less concern than in ocean circulation models (which remain a work in progress), because the PE model is known to produce very accurate pressure predictions, whose error is very small compared to the changes in pressure caused by the sound speed profile changes that we are inverting for. As a result, we do not explicitly address modeling error in our inversion cost functions.

A. Perturbation of parabolic equation to produce tangent linear model

The first necessary ingredient for an adjoint model is the so-called tangent linear model. This is simply a linearized version of the PE that estimates (linearly) how perturbations in the sound speed profile map into perturbations in the pressure field. The standard homogeneous PE²⁵ (with no source in the medium) is

$$2ik_0 \frac{\partial p}{\partial r} + \frac{\partial^2 p}{\partial z^2} + k_0^2 [\gamma(r, z) - 1] p = 0, \quad (1)$$

where $\gamma(r, z)$ is the index of refraction squared, $c_0^2/c^2(r, z)$. Expanding this equation in terms of perturbations in pressure p and index of refraction squared γ to first order in ε ,

$$\begin{aligned} \gamma &= \gamma + \varepsilon \tilde{\gamma}, \\ p &= p + \varepsilon \tilde{p} + \dots, \end{aligned} \quad (2)$$

yields

$$2ik_0 \frac{\partial \tilde{p}}{\partial r} + \frac{\partial^2 \tilde{p}}{\partial z^2} + k_0^2 [\gamma(r, z) - 1] \tilde{p} = -k_0^2 \tilde{\gamma}(r, z) p.$$

We recognize that typical acoustic inversion problems will rarely be strictly linear. Therefore, in subsequent sections we will show an iterative method that relinearizes about its last estimate and can find its way to the solution despite mild nonlinearities.

B. Discretized PE and its tangent linear model

At this point, Eq. (2) can be discretized to implement a tangent linear model for the PE, from which an adjoint model can be constructed. However, it may be prudent to first derive a discretized PE—deriving an analytical adjoint and then discretizing it can introduce errors that break adjoint symmetry and result in suboptimal inversions.¹⁹ Therefore, in this section, we start with a discretized PE (the implicit finite differences PE²⁵) and derive a tangent linear model from its first-order terms.

1. Finite difference PE: Zeroth order

A finite difference version of Eq. (1) is

$$2ik_0 \frac{\mathbf{p}^{n+1} - \mathbf{p}^n}{\Delta r} + \mathbf{A} \frac{\mathbf{p}^{n+1} + \mathbf{p}^n}{2} = 0, \quad (3)$$

where \mathbf{p}^n is the vector of pressures sampled in depth at range n , Δr is the size of the range step, and matrix \mathbf{A} is the depth operator corresponding to the terms $\partial^2/\partial z^2 + k_0^2 [\gamma(r, z) - 1] p$ in Eq. (1). Matrix \mathbf{A} is

$$\begin{bmatrix} d_1 & e_2 & & & & \\ e_2 & d_2 & e_3 & & & \\ & e_3 & d_3 & & & \\ & & & \ddots & & \\ & & & & d_{N-2} & e_{N-1} \\ & & & & e_{N-1} & d_{N-1} & e_N \\ & & & & & e_N & d_N \end{bmatrix}, \quad (4)$$

$$\begin{aligned} d_j &= -\frac{2}{h^2} + k_0^2 [\gamma(r, z_j) - 1], \\ e_j &= \frac{1}{h^2}, \end{aligned} \quad (5)$$

where h is the depth increment (we have used a constant depth increment). Collecting terms

$$\left(\frac{2ik_0}{\Delta r} \mathbf{I} + \frac{\mathbf{A}}{2} \right) \mathbf{p}_{n+1} = \left(\frac{2ik_0}{\Delta r} \mathbf{I} - \frac{\mathbf{A}}{2} \right) \mathbf{p}_n, \quad (6)$$

which can be rewritten as

$$\mathbf{p}_{n+1} = \mathbf{F} \mathbf{p}_n, \quad (7)$$

where

$$\mathbf{F} = \left(\frac{2ik_0}{\Delta r} \mathbf{I} + \frac{\mathbf{A}}{2} \right)^{-1} \left(\frac{2ik_0}{\Delta r} \mathbf{I} - \frac{\mathbf{A}}{2} \right) = \mathbf{B}^{-1} \mathbf{C}. \quad (8)$$

This is the familiar PE marching solution, where \mathbf{F} contains the range-dependent index of refraction squared $\gamma(r, z)$.

2. PE tangent linear model from first-order finite difference PE

To construct the adjoint of the PE model, we must first construct a tangent linear model for it. A first-order perturbation analysis of Eq. (8) yields

$$(\mathbf{B} + \varepsilon \tilde{\mathbf{B}})(\mathbf{p}_{n+1} + \varepsilon \tilde{\mathbf{p}}_{n+1}) = (\mathbf{C} + \varepsilon \tilde{\mathbf{C}})(\mathbf{p}_n + \varepsilon \tilde{\mathbf{p}}_n), \quad (9)$$

where $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{C}}$ are of opposite sign, and are both diagonal, with the i th diagonal element having magnitude $(k_0^2/2) \tilde{\gamma}(r, z_i)$. The $\tilde{\gamma}(r, z_i)$ are the first-order medium properties we are inverting for, and are indexed on depth z_i and range r .

The terms of order ε produce

$$\mathbf{B} \tilde{\mathbf{p}}_{n+1} = \mathbf{C} \tilde{\mathbf{p}}_n - \tilde{\mathbf{B}} \mathbf{p}_{n+1} + \tilde{\mathbf{C}} \mathbf{p}_n. \quad (10)$$

Replacing matrices $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{C}}$ with diagonal matrix \mathbf{D} and setting the appropriate factors yields

$$\mathbf{B} \tilde{\mathbf{p}}_{n+1} = \mathbf{C} \tilde{\mathbf{p}}_n - \frac{k_0^2}{2} \mathbf{D} (\mathbf{p}_{n+1} + \mathbf{p}_n). \quad (11)$$

Swapping the diagonal elements of \mathbf{D} [i.e., $\tilde{\gamma}(r, z_i)$] with the elements of the vector that \mathbf{D} is multiplying (i.e., $\mathbf{p}_{n+1} + \mathbf{p}_n$),

$$\mathbf{B} \tilde{\mathbf{p}}_{n+1} = \mathbf{C} \tilde{\mathbf{p}}_n - \Gamma \tilde{\gamma}(\mathbf{z}), \quad (12)$$

with Γ a diagonal matrix whose diagonal elements consist of the vector $(k_0^2/2) (\mathbf{p}_{n+1} + \mathbf{p}_n)$. Multiplying both sides by \mathbf{B}^{-1} ,

$$\tilde{\mathbf{p}}_{n+1} = \mathbf{F} \tilde{\mathbf{p}}_n - \mathbf{G} \tilde{\mathbf{u}}_n, \quad (13)$$

$$\mathbf{G} = \mathbf{B}^{-1} \Gamma, \quad (14)$$

where $\tilde{\mathbf{u}}_n$ is the vector of $\tilde{\gamma}(r_n, z)$ values sampled in depth at range r_n [i.e., the i th element of $\tilde{\mathbf{u}}_n$ is $\tilde{u}_n(i) = \tilde{\gamma}(r_n, z_i)$, the first-order perturbation to the index of refraction squared at range n and depth i]. Matrix \mathbf{F} in Eq. (13) is identical to the \mathbf{F} in Eq. (8) for the zeroth-order pressures. Equation (13) propagates the first-order pressure field $\tilde{\mathbf{p}}$ (the perturbation to

the zeroth-order pressure field \mathbf{p}), with the first-order medium perturbations $\tilde{\mathbf{u}}_n$ [i.e., $\tilde{\gamma}(r,z)$] acting as a forcing function. The perturbations $\tilde{\gamma}(z)$, and matrices \mathbf{B} and \mathbf{C} , by which \mathbf{F} and \mathbf{G} are represented, can be range dependent. Equation (13) is the tangent linear model for our (otherwise nonlinear) PE model. It is important to keep in mind that the tangent linear model and the adjoint model derived from it in Sec. IID propagate quantities that are first-order perturbations.

C. How adjoint model arises in our PE-based inversion

In this section, we combine the sequence of PE marching steps into a single linear system, and show how the adjoint of this system back-propagates data-model misfit at the observation points back to the points in the medium where adjustments should be made. The purpose of this section is to illustrate what an adjoint model does in a very concrete way. We will actually use a more efficient adjoint formulation presented in the next section to solve inverse problems in later sections.

To simplify notation, we drop the tilde notation that was used in Sec. IIB to indicate a first-order perturbation and assume all references to \mathbf{p} and \mathbf{u} below are to their first-order terms.

Equation (13) generates the following equations, one for each range step (the subscripts indicate the index of the range step):

$$\begin{aligned} \mathbf{p}_1 &= \mathbf{F}_0 \mathbf{p}_0 - \mathbf{G}_0 \mathbf{u}_0, \\ \mathbf{p}_2 &= \mathbf{F}_1 \mathbf{F}_0 \mathbf{p}_0 - \mathbf{F}_1 \mathbf{G}_0 \mathbf{u}_0 - \mathbf{G}_1 \mathbf{u}_1, \\ &\vdots \\ \mathbf{p}_N &= \mathbf{F}_{N-1} \mathbf{F}_{N-2} \cdots \mathbf{F}_1 \mathbf{F}_0 \mathbf{p}_0 \\ &\quad - \sum_{n=0}^{N-1} \mathbf{F}_{N-1} \mathbf{F}_{N-2} \cdots \mathbf{F}_{n+1} \mathbf{G}_n \mathbf{u}_n. \end{aligned} \quad (15)$$

We include vector \mathbf{p}_0 , the correction to the presumed starter field, as an unknown here for completeness. We will only solve for medium properties in the sections that follow. Each pressure \mathbf{p}_n is a function of all the medium property vectors \mathbf{u}_i , $i=1, \dots, n$, leading up to it. A matrix representation of this summation is

$$[\mathbf{A} \quad \mathbf{B}_0 \quad \mathbf{B}_1 \quad \cdots \quad \mathbf{B}_{N-1}] \begin{bmatrix} \mathbf{p}_0 \\ \mathbf{u}_0 \\ \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_{N-1} \end{bmatrix} = \mathbf{p}_N, \quad (16)$$

where the submatrix \mathbf{A} is

$$\mathbf{A} = \mathbf{F}_{N-1} \mathbf{F}_{N-2} \cdots \mathbf{F}_0, \quad (17)$$

each submatrix \mathbf{B}_n is

$$\mathbf{B}_n = -\mathbf{F}_{N-1} \mathbf{F}_{N-2} \cdots \mathbf{F}_{n+1} \mathbf{G}_n, \quad (18)$$

and \mathbf{p}_0 and all the individual \mathbf{u}_n vectors have been stacked one on top of each other in Eq. (16) to form the super vector

\mathbf{u} . Equation (16) is a representation of the inverse problem in a familiar linear system form:

$$\mathbf{F} \mathbf{u} = \mathbf{p}_N, \quad (19)$$

where \mathbf{F} and \mathbf{u} are composite matrices.

Equation (19) can be solved using either a generalized inverse of \mathbf{F} ,

$$\hat{\mathbf{u}} = \mathbf{F}^+ \mathbf{p}_N, \quad (20)$$

or an iterative steepest descent technique,

$$\hat{\mathbf{u}}_i = \hat{\mathbf{u}}_{i-1} + \alpha \mathbf{F}^* (\mathbf{F} \hat{\mathbf{u}}_{i-1} - \mathbf{p}_N), \quad (21)$$

in which the second term on the right hand side is just the gradient of the squared data-model misfit, scaled by a step size parameter α . Note how the gradient is formed by the adjoint \mathbf{F}^* operating on the data-model misfit $\mathbf{F} \hat{\mathbf{u}}_{i-1} - \mathbf{p}_N$.

From Eq. (16), the conjugate transpose of \mathbf{F} (the adjoint of our linearized PE) is

$$\mathbf{F}^* = \begin{bmatrix} \mathbf{A}^* \\ \mathbf{B}_0^* \\ \vdots \\ \mathbf{B}_{N-1}^* \end{bmatrix}, \quad (22)$$

where the individual matrix

$$\mathbf{B}_n^* = -\mathbf{G}_n^* \mathbf{F}_{n+1}^* \cdots \mathbf{F}_{N-2}^* \mathbf{F}_{N-1}^* \quad (23)$$

can be interpreted as the transfer matrix which back-propagates the data-model misfit (often identified as the adjoint field) from the receiver at range N to points back in the medium at range n . Matrix \mathbf{F}_{N-1}^* steps from range N to $N-1$. Matrix \mathbf{F}_{N-2}^* steps from range $N-1$ to $N-2$, and so on, until at range n , \mathbf{G}_n^* transforms the adjoint field to a medium property correction at range step n (as we will show later). The adjoint of our forward model propagates information from the observations at the receiver array back to all the medium properties that we are inverting for (i.e., the \mathbf{u}_n at all ranges n). Note that matrix \mathbf{A}^* similarly propagates information all the way back to the source, where it can be used to correct the starter field (if this is treated as an unknown).

D. Fixed-point iteration to invert for medium properties

In this section, we present an alternative derivation of the adjoint model, using Lagrange multipliers. This yields a fixed-point iteration, which we relate to a steepest-descent iteration. To simplify the presentation, as in Sec. IIC, we drop the tilde notation that was used in Sec. IIB to indicate first-order perturbations and assume all references to \mathbf{p} and \mathbf{u} are to their first-order terms.

The first-order parabolic equation is

$$\mathbf{p}_{n+1} = \mathbf{F}_n \mathbf{p}_n - \mathbf{G}_n \mathbf{u}_n, \quad (24)$$

where the unknown medium properties \mathbf{u}_n act as forcing functions at each range step n of the marching solution [Eq. (24) is the same as Eq. (13), without the tilda notation]. Matrices \mathbf{F}_n and \mathbf{G}_n are functions only of the zeroth-order pressures and medium properties, and not their first-order perturbations, \mathbf{p}_n and \mathbf{u}_n .

To solve for \mathbf{u}_n , we formulate a cost function $J(\mathbf{p}, \mathbf{u}, \lambda)$ to be minimized:

$$J(\mathbf{p}, \mathbf{u}, \lambda) = \frac{1}{2} (\mathbf{p}_N - \mathbf{m}_N)^* (\mathbf{p}_N - \mathbf{m}_N) + \sum_{n=1}^N \lambda_n^* (\mathbf{p}_n - \mathbf{F}_{n-1} \mathbf{p}_{n-1} + \mathbf{G}_{n-1} \mathbf{u}_{n-1}) + \frac{\beta}{2} \sum_{n=0}^{N-1} \mathbf{u}_n^* \mathbf{u}_n. \quad (25)$$

The first term in J seeks to minimize the mismatch between the measured \mathbf{m}_N and modeled \mathbf{p}_N pressure increments, both at range index N . The data-model misfit in the first term will also contain measurement noise in \mathbf{m}_N , which we ignore, since we focus on the underdetermined case (which provides no redundant measurements that can be used to smooth the solution). Since we are dealing with first-order terms, \mathbf{m}_N is the difference between the measured pressure and the zeroth-order pressure prediction. The modeled pressure \mathbf{p}_N is calculated by propagating \mathbf{p}_n from the source to the receiver using Eq. (24) [i.e., our tangent linear model $\{\mathbf{F}_n, \mathbf{G}_n\}$ with the estimated environmental parameters \mathbf{u}_n as driving functions]. Given a solution for \mathbf{u}_n , the zeroth-order pressure plus the pressure correction \mathbf{p}_N calculated using \mathbf{u}_n should reproduce the measured pressure. The second term uses Lagrange multipliers λ_n (vectors at each range, sampled in depth) to enforce the hard constraint that the model is perfect (i.e., \mathbf{p}_n and \mathbf{u}_n are completely governed by the linearized forward model $\{\mathbf{F}_n, \mathbf{G}_n\}$). The third term is a regularizing term to minimize the amplitude of the medium property perturbations \mathbf{u}_n , with β a weighting term to regulate how much we allow the \mathbf{u}_n to stray from our assumed zeroth-order values (later, we will set β to the ratio of the variances of the pressure measurements and the medium properties).

Admittedly, this is an unusual way to formulate an inverse problem. We have set up a large number of unknowns in all the intermediate \mathbf{p}_n , in addition to the already large number of unknowns \mathbf{u}_n . We will show that minimizing the cost function in Eq. (25) is a more direct way of inverting for the \mathbf{u}_n than repeatedly running the forward model to explore the surface J as a function of \mathbf{u}_n .

Note that J is a function of \mathbf{u}_n at all ranges and depths, so its minimization can be used to resolve range-dependent features.

The partial derivatives of $J(\mathbf{p}, \mathbf{u}, \lambda)$ are

$$\frac{\partial J}{\partial \mathbf{p}_N} = \mathbf{p}_N - \mathbf{m}_N + \lambda_N, \quad (26)$$

$$\frac{\partial J}{\partial \mathbf{p}_n} = \lambda_n - \mathbf{F}_n^* \lambda_{n+1}, \quad (27)$$

$$\frac{\partial J}{\partial \mathbf{u}_n} = \beta \mathbf{u}_n + \mathbf{G}_n^* \lambda_{n+1}, \quad (28)$$

$$\frac{\partial J}{\partial \lambda_n} = \mathbf{p}_n - \mathbf{F}_{n-1} \mathbf{p}_{n-1} + \mathbf{G}_{n-1} \mathbf{u}_{n-1}. \quad (29)$$

Setting these partial derivatives to zero yields a system of equations which are solved to invert for the sound speed profile. Much of the adjoint literature describes iterative processes to solve these equations. It should be emphasized that the iteration is not something that directly follows in logical sequence but is simply a proposed approach to solving this large system. Actually, it is just a so-called fixed-point or functional iteration (also called a Picard iteration).

To explain the iteration in those terms, the derivatives in Eqs. (26)–(28) are set to zero and rearranged to yield

$$\mathbf{u}_n = f(\mathbf{u}_n) \quad (30)$$

with

$$f(\mathbf{u}_n) = -\frac{1}{\beta} \mathbf{G}_n^* \lambda_{n+1}. \quad (31)$$

The term $\mathbf{G}_n^* \lambda_{n+1}$ comes from the back-propagation of the λ_N to λ_n , with λ_N a function of \mathbf{p}_N , and \mathbf{p}_N itself calculated from \mathbf{u}_n . The idea of a fixed-point iteration is to pick a starting guess, \mathbf{u}^0 , and apply a simple recursion:

$$\mathbf{u}^{i+1} = f(\mathbf{u}^i). \quad (32)$$

Whether the fixed-point iteration converges or not is often unclear. However, if it does, then we have solved for the solution of our original equation. In fact, part of the art of constructing a good fixed-point iteration is to rearrange the original equation into a form where the iteration converges as rapidly as possible. For instance, we can obtain a broader class of iterations by multiplying Eq. (30) by α and adding \mathbf{u}_n on both sides,

$$\mathbf{u}_n + \alpha \mathbf{u}_n = \mathbf{u}_n + \alpha f(\mathbf{u}_n), \quad (33)$$

which is rearranged as

$$\mathbf{u}_n = \mathbf{u}_n - \alpha (\mathbf{u}_n - f(\mathbf{u}_n)). \quad (34)$$

Now, α provides a parameter that we can adjust to optimize the rate of convergence. To summarize the fixed-point iteration in terms of the variables in our problem, we have the following recursion:

$$\mathbf{p}_{n+1}^{i+1} = \mathbf{F}_n \mathbf{p}_n^{i+1} - \mathbf{G}_n \mathbf{u}_n^i, \quad (35)$$

$$\lambda_N^{i+1} = \mathbf{m}_N - \mathbf{p}_N^{i+1}, \quad (36)$$

$$\lambda_n^{i+1} = \mathbf{F}_n^* \lambda_{n+1}^i, \quad (37)$$

$$\mathbf{u}_n^{i+1} = \mathbf{u}_n^i - \alpha \left(\mathbf{u}_n^i + \frac{\mathbf{G}_n^* \lambda_{n+1}^{i+1}}{\beta} \right). \quad (38)$$

In words, we calculate \mathbf{p}_n^{i+1} using our forward model to propagate pressures from the source to the receiver based on our current estimate of \mathbf{u}_n^i in Eq. (35), initialize λ_N^{i+1} at the receiver to the data-model misfit [Eq. (36)], propagate the λ_n^{i+1} from the receiver to the source in Eq. (37), and calculate the updated first-order medium properties \mathbf{u}_n^{i+1} from the updated λ_n^{i+1} at all range indexes n in Eq. (38).

Equation (37) represents the *adjoint model*. The λ_n 's form the *adjoint field* propagated by the adjoint model \mathbf{F}_n^* . Equation (36) initializes the adjoint field λ_N to the data-model misfit at the receiver. Note the complementary relationship between the pressures \mathbf{p}_n , being propagated from

the source to the receiver by our forward model \mathbf{F}_n , and the Lagrange multipliers λ_n , being propagated by the adjoint model \mathbf{F}_n^* backward from the receiver to the source.

A compelling interpretation of the adjoint is that the observed field is a superposition of a baseline field due to the presumed medium and a perturbed field due to the unknown medium perturbations (which we invert for). The adjoint model back-propagates (time-reverses) the perturbed field to the unknown medium perturbations, which are viewed as sources of diffraction.¹

An important insight about this fixed-point iteration is that it is also equivalent to a descent technique. To see this, observe that the second term in Eq. (34) [also Eq. (38)] is proportional to the gradient of J with respect to \mathbf{u} as given in Eq. (28). Thus, our fixed-point iteration is equivalent to an iteration

$$\mathbf{u}^{i+1} = \mathbf{u}^i - \frac{\alpha}{\beta} \frac{\partial J}{\partial \mathbf{u}}. \quad (39)$$

In this form, we see that the parameter α plays the usual role in a steepest descent method of controlling the step size in the descent direction. It is natural in this descent framework to consider all the standard extensions such as conjugate gradient techniques.²⁶ We do not discuss these options further here, but simply note they are part of the standard set of techniques used to improve the convergence of the adjoint iteration.

The fixed-point iteration equations derived above are in terms of \mathbf{F}_n and \mathbf{G}_n which are functions of the zeroth-order medium properties and pressures. As a result, the iterative process can periodically be relinearized about the current estimates for \mathbf{u}_n in order to update the tangent linear model. This enables the process to handle mildly nonlinear problems.

III. INVERTING FOR INTIMATE 96 INTERNAL TIDES (SIMULATED RESULTS)

In this section, the adjoint method described in Sec. II D is tested on synthetic acoustic data calculated for a series of sound speed profiles measured during the INTIMATE 96 experiment²⁷ during which internal tides were observed. Figure 1 shows the entire sequence of measured sound speed profiles, in the order that they were measured (the sampling in time was not uniform, so indexes rather than time units are shown on the horizontal axis). Figure 2 shows the entire set of measured sound speed profiles and the source-receiver configuration used in our simulation. These profiles were measured over roughly 60 h. The experiment configuration was fixed-fixed. The line joining the source and receiver was perpendicular to the movement of internal tides during part of the actual experiment. Such a configuration would suggest a *range-independent* sound speed field in the plane joining the source and receiver.

An important step when applying adjoint methods is to ensure that the problem has been linearized, so Sec. III A compares the pressures calculated by the original PE and the linearized PE models, for the configuration to be used to demonstrate the inversion. Section III B applies the iterative

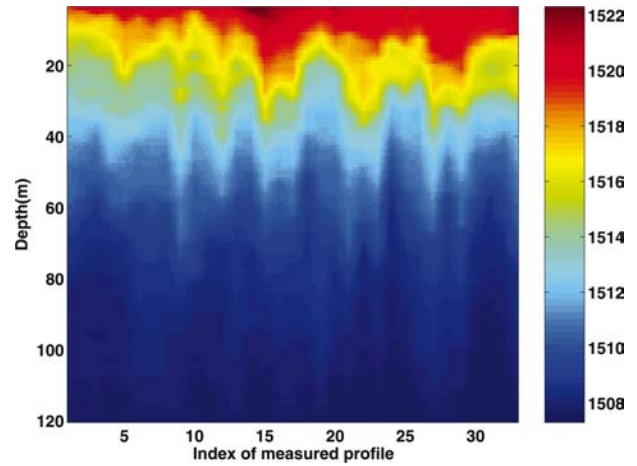


FIG. 1. Sound speed profiles measured during INTIMATE 96 over a time interval of roughly 3 days. The sound speed profiles were measured at only approximately regular time intervals, so the horizontal axis shows the index of the measurements. The image scale shows sound speed measured in m/s.

adjoint technique described in Sec. II to a series of *range-independent* sound speed fields. Section III C shows that our adjoint implementation reproduces the tangent linear model (using far fewer modeling runs than if the forward model alone were used). Section III D shows how a generalized inverse of the linearized forward model can be used as an alternative to our iterative adjoint inversion, and also how to calculate error bars for these inversion results. Section III E applies the iterative adjoint inversion to a range-dependent inversion problem.

The PE model was used to synthesize acoustic pressures to serve as “measured” data in our simulations. Pressures were calculated at 400 Hz in Sec. III B for the range-independent case and for a range of frequencies from 100 to 400 Hz in Sec. III E for the range-dependent cases. The receiver was a fully spanning vertical line array at a 2-km distance from the source. The source depth was at 50-m depth in Sec. III B for the range-independent case. Sources spanned 5 to 90 m for the range-dependent cases in Sec.

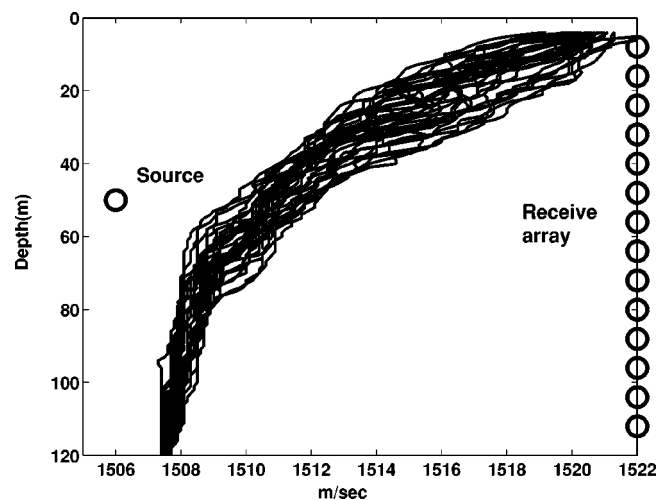


FIG. 2. Sound speed profiles measured during INTIMATE 96 and experimental configuration for which acoustic data was synthesized. All inversions were performed at a frequency of 400 Hz and a range of 2 km.

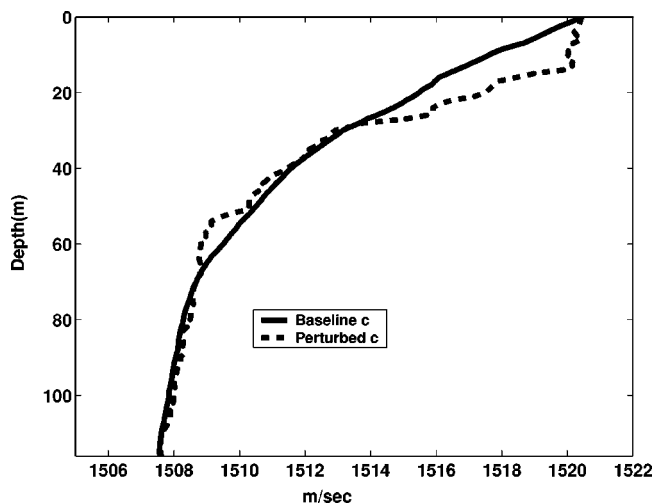


FIG. 3. Perturbed and baseline sound speed profiles used to test linearity of INTIMATE 96 configuration.

III E, with each source an independent omni-directional point source and not part of a transmitting array.

A. Testing our tangent linear model to verify we are in linear regime

Prior to using the adjoint method for inversion it is important to assess whether the tangent linear model and its adjoint are reasonable approximations to the full PE model in this experimental configuration. Using the same experimental configuration as in our inversions, we compare perturbations produced by our tangent linear model with perturbations predicted by the original (nonlinear) PE. Figure 3 shows two sound speed profiles, $c(z)$ to be used as a baseline, and $c(z) + \delta c(z)$ as a perturbed profile. The baseline profile is the mean profile calculated from the entire set of INTIMATE 96 sound speed measurements. The perturbed profile is one chosen from the measured profiles for its large deviation from the mean profile. Figure 4 shows the pressure field produced by the original PE model operating on the baseline sound speed profile ($P(c)$, where P represents the

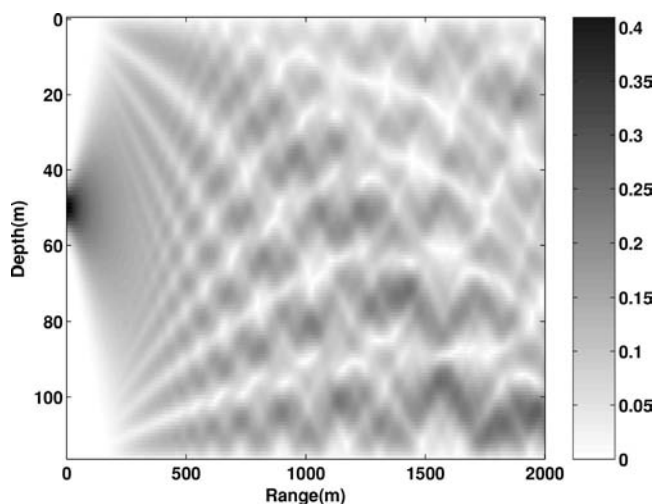


FIG. 4. Baseline pressure magnitude at 2 km, source depth of 50 m, and at 400 Hz.

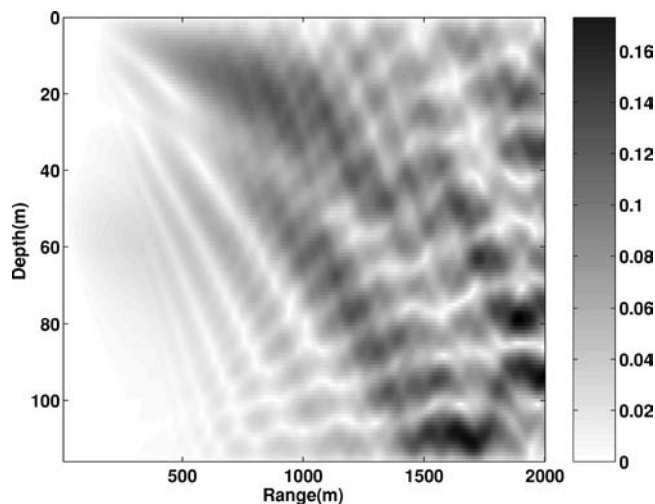


FIG. 5. Magnitude difference between perturbed and baseline pressures.

original PE). Figure 5 shows the magnitude difference between perturbed and baseline pressures calculated by the original PE [i.e., $P(c + \delta c) - P(c)$]. Figure 6 shows the magnitude difference between both sets of perturbed pressures, those calculated by the original PE and those calculated by its tangent linear model [i.e., $P(c + \delta c) - P(c) - \mathbf{F}(c)\delta c$, where $\mathbf{F}(c)$ represents the tangent linear model, linearized about c]. Comparing the images in Fig. 5 and Fig. 6 (whose color scales are the same) shows that the perturbation in pressures due to the change in medium properties (in Fig. 5) is always larger than the size of the error between the original PE model and its tangent linear model (in Fig. 6). This shows that the problem is not too nonlinear (our iterative process will relinearize after each iteration), and that our configuration is reasonable for testing our iterative adjoint process.

As we have shown in Secs. II C and II D, the adjoint process is a local optimization, so if we are outside the solution's "basin of attraction," there is no mechanism for escaping a spurious local minimum. This can happen if the initial guess is too far from the correct solution, or if the linear approximation being used by the tangent linear model

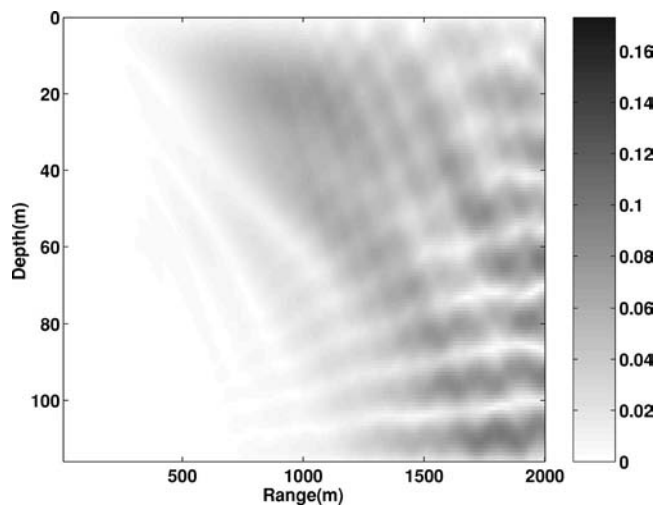


FIG. 6. Magnitude difference between perturbed and tangent linear pressures (would be zero if tangent linear model perfectly predicted the perturbed pressure). Compare to differences shown in Fig. 5.

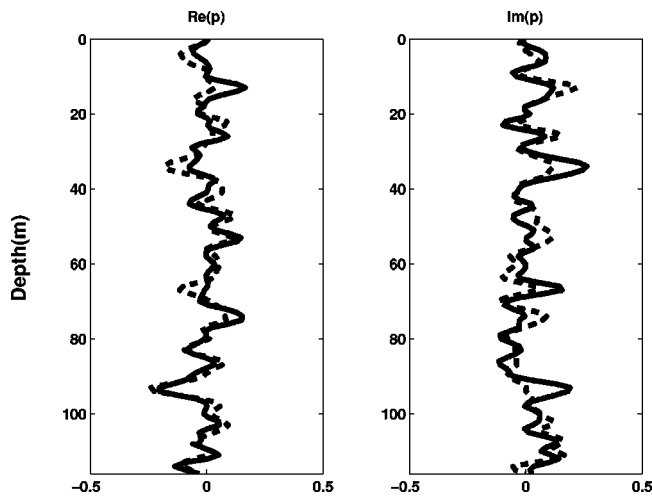


FIG. 7. “Measured” pressure was synthesized using PE with true sound speed profile (solid line). Modeled pressure was synthesized using the PE model with mean sound speed profile (dashed line). Real components are shown in left plot. Imaginary components are shown in right plot.

(and its adjoint) fails to reproduce the true cost function. Because it relinearizes about its last estimate and can follow a nonlinear basin of attraction if started reasonably close to the true solution, our iterative adjoint process can handle slightly nonlinear problems.

B. Adjoint inversions using fixed-point iteration (range-independent case)

In this section, the iterative adjoint inversion process discussed in Sec. IID is tested on *synthetic* acoustic data to assess the feasibility of inverting for sound speed profiles that indicate the internal tides present during the INTIMATE 96 experiment. For each measured profile from INTIMATE 96, an acoustic pressure vector is calculated by the PE model to serve as a “measured” pressure vector for our inversions. Each of these “measured” pressures is inverted to estimate the sound speed profile used to synthesize it. The mean of the entire set of profiles serves as the initial guess in each inversion.

Before showing the results on the entire sequence of profiles, we show the details of an inversion on a single profile, profile 28, selected for its comparatively large deviation from the mean. Figure 7 shows the “measured” pressure (calculated using profile 28) and the modeled pressure upon initialization (calculated for the mean profile, used to initialize our iterative process). A single iteration consists of using the forward model to calculate a modeled pressure, then using the adjoint model to calculate a corrected sound speed profile. Figure 8 shows cost function values [see Eq. (25)] at iterations 1–50. Figure 9 shows the modeled pressures and the measured pressures after 50 iterations. Figure 10 shows a comparison between the true and estimated sound speed profile perturbations (i.e., after 50 iterations).

The inversion shown in detail for profile 28 (in Figs. 7–10) was repeated at each of the profiles measured during the INTIMATE 96 experiment. Figure 11 shows the inversion results for the entire set of INTIMATE 96 sound speed profiles. The sound speed profiles were measured at only

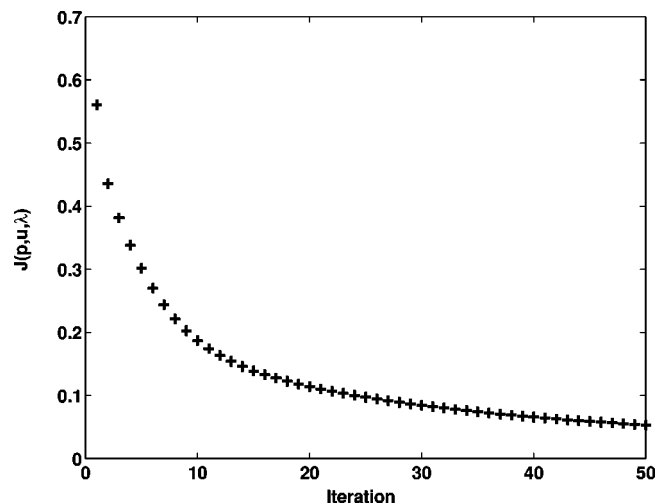


FIG. 8. Evolution of objective function produced by iterative adjoint process.

approximately regular time intervals, so the horizontal axes show the indexes of the measurements. The upper left plot shows the “measured” profiles (the true profiles, in our simulation). The lower left plot shows the estimated profiles after 50 iterations at each profile. The upper right plot shows the “true” profiles minus the mean profile. These are the first order quantities that are being estimated. The lower right plot shows the difference between the “true” and estimated sound speed profiles. Both the upper right and lower right plots have the same scale, so that the estimation errors can be compared with the quantities being estimated. The estimation errors are significantly smaller. This experiment was repeated with only four receiver elements (using the same configuration as the INTIMATE 96 experiment) and at a number of signal-to-noise ratios (additive white gaussian noise was added at each of the four receivers). Similar results to those shown in Fig. 11 were obtained at SNRs greater than 10 dB.

C. Using adjoint model to construct tangent linear model (range-independent case)

In this section, partly to assess the validity of our adjoint implementation, we calculate the tangent linear model using

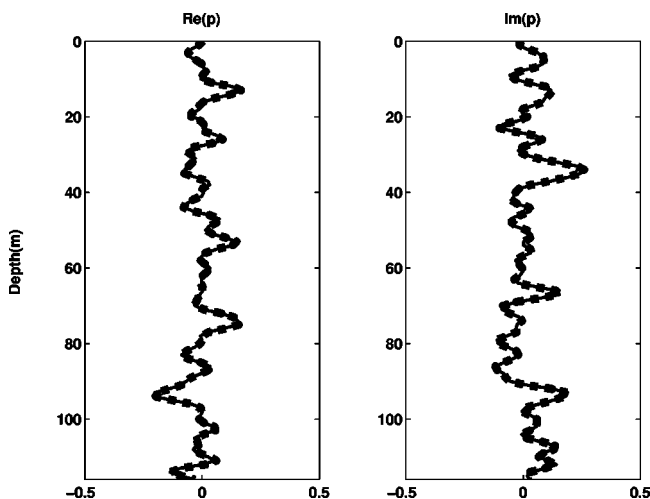


FIG. 9. Measured and final estimated pressure magnitudes.

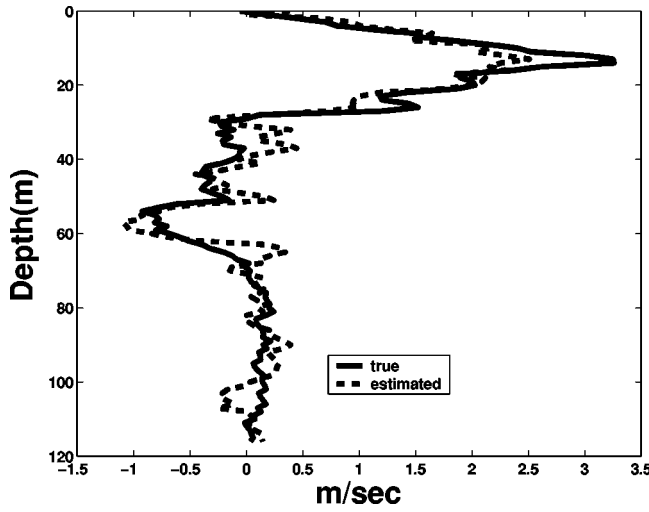


FIG. 10. True and final estimated sound speed profile perturbations.

our adjoint model and compare it with a tangent linear model calculated in a brute force fashion using the original PE. If our tangent linear model \mathbf{F} is \mathbf{M} by \mathbf{N} (\mathbf{M} observations, \mathbf{N} waveguide parameters), we will use \mathbf{M} runs of our adjoint model to calculate the \mathbf{M} rows of \mathbf{F} instead of using \mathbf{N} runs of our forward model to calculate the \mathbf{N} columns of \mathbf{F} . This comparison is made at one of the profiles estimated in Sec. III B (profile 28). With a much larger number of medium properties than measurements, the adjoint model provides a much less computationally demanding means of calculating the tangent linear model. As we will show in Sec. III D, such a tangent linear model can be used to form a generalized inverse solution to the inverse problem (an alternative to the iterative process we have demonstrated above) and to calculate error bars around the inverse problem solution.

The tangent linear model is the matrix \mathbf{F} whose element F_{ij} is the i th observed pressure due to a unit perturbation of the j th index of refraction squared value. Thus, if we use the configuration of source and a vertical line receive array (as in our simulated example), a column of \mathbf{F} can be calculated by two runs of our unperturbed PE model:

$$\mathbf{f}_i = P(\mathbf{u} + \tilde{\mathbf{u}}_i) - P(\mathbf{u}), \quad (40)$$

where \mathbf{f}_i represents the i th column of \mathbf{F} , and $P(\mathbf{u})$ represented the pressures calculated at the receive array elements by the original PE model using medium properties \mathbf{u} . The perturbation in \mathbf{u} is $\tilde{\mathbf{u}}_i = \alpha \mathbf{e}_i$, where \mathbf{e}_i is the i th elementary basis vector (having a one at its i th element and zeros at all other elements), and α is a scaling factor small enough to keep the model linear. Each \mathbf{f}_i is a vector of pressure perturbations at the vertical line receive array elements due to the perturbation of \mathbf{u} at its i th element. The matrix \mathbf{F} is rescaled by dividing it by α , to compensate for the earlier scaling of the index of refraction perturbations $\tilde{\mathbf{u}}_i$. We will call the matrix calculated using the forward model $\mathbf{F}_u^{\text{forward}}$.

For the adjoint approach, we run the adjoint model to calculate the sound speed correction that would reproduce pressure perturbations at each individual receiver,

$$\tilde{\mathbf{u}}_j = \mathbf{F}^*(\tilde{\mathbf{p}}_j), \quad (41)$$

where \mathbf{F}^* is the adjoint model [see Eqs. (37) and (38)] and $\tilde{\mathbf{p}}_j = \alpha \mathbf{e}_j$. The conjugates of vectors $\tilde{\mathbf{u}}_j$ are put into the rows of a new matrix, $\mathbf{F}_u^{\text{adjoint}}$. The $\tilde{\mathbf{u}}_j$ are often called the representer adjoints, which must be calculated prior to calculating the representers themselves in a process which converts the inverse problem from a two-point boundary value problem to a linear algebraic problem in which the solution is expressed as a linear combination of representers.¹⁹

We compared the left eigenvectors and singular values (calculated by a singular value decomposition) of $\mathbf{F}_u^{\text{forward}}$ with $\mathbf{F}_u^{\text{adjoint}}$ to assess how well the adjoint model reproduced the tangent linear model calculated by the unperturbed PE. At the significant eigenvectors, the adjoint model was able to reproduce the eigenstructure of the $\mathbf{F}_u^{\text{forward}}$ matrix.²⁸

D. Generalized inverse solution and posterior covariances (range-independent case)

In this section, we use the representer adjoints calculated in the previous section to calculate a generalized inverse solution, as discussed in Sec. II C. We show this calculation as a final correction to the estimate produced by our iterative adjoint process in Sec. III B, but it could in principle replace the iterative adjoint-based inversion. We also show how the representer adjoints can be used to calculate the *posteriori* covariances of our sound speed estimates, given prior covariances for our pressure measurements and medium properties. All the calculations in this section (as in the previous section) are performed at profile 28, estimated in Sec. III B.

Prior knowledge of measurement and medium property error statistics is typically incorporated into the cost function we seek to minimize using covariance matrices \mathbf{C}_p (for pressure measurement errors \mathbf{p}) and \mathbf{C}_u (for medium properties \mathbf{u}). These covariance matrices map the data vector space and model vector space to their dual spaces, enabling a consistent formulation of the optimization in terms of a gradient (which is an element of the dual space).²⁹ Although incorporating more realistic covariances is usually necessary in practical applications, we will simplify the material that follows by assuming the covariances are scaled identity matrices:

$$\mathbf{C}_p = \sigma_p^2 \mathbf{I}, \quad \mathbf{C}_u = \sigma_u^2 \mathbf{I}, \quad (42)$$

and use $\beta = \sigma_p^2 / \sigma_u^2$ to represent our prior expectation of the relative accuracy of our measurements compared with that of our baseline medium properties. For high SNR, we set β low (σ_p small compared with σ_u), allowing our \mathbf{u} estimates to stray further from our prior expectations (the zeroth-order medium properties). Note that although we have not explicitly included measurement errors in Eq. (25), it is addressed implicitly by β , (and in general by \mathbf{C}_p and \mathbf{C}_u when they are not scaled identity matrices).

Defining a cost function $J(\mathbf{u})$ in terms of the data-model misfit and norm of \mathbf{u} (to penalize large \mathbf{u}),

$$J(\mathbf{u}) = \frac{1}{2} (\mathbf{F}\mathbf{u} - \mathbf{m}_N)^* \mathbf{C}_p^{-1} (\mathbf{F}\mathbf{u} - \mathbf{m}_N) + \frac{1}{2} \mathbf{u}^* \mathbf{C}_u^{-1} \mathbf{u}, \quad (43)$$

setting the gradient of $J(\mathbf{u})$ to zero, and solving for \mathbf{u} produces

$$\mathbf{u} = (\mathbf{F}^* \mathbf{C}_p^{-1} \mathbf{F} + \mathbf{C}_u^{-1})^{-1} \mathbf{F}^* \mathbf{C}_p^{-1} \mathbf{m}_N, \quad (44)$$

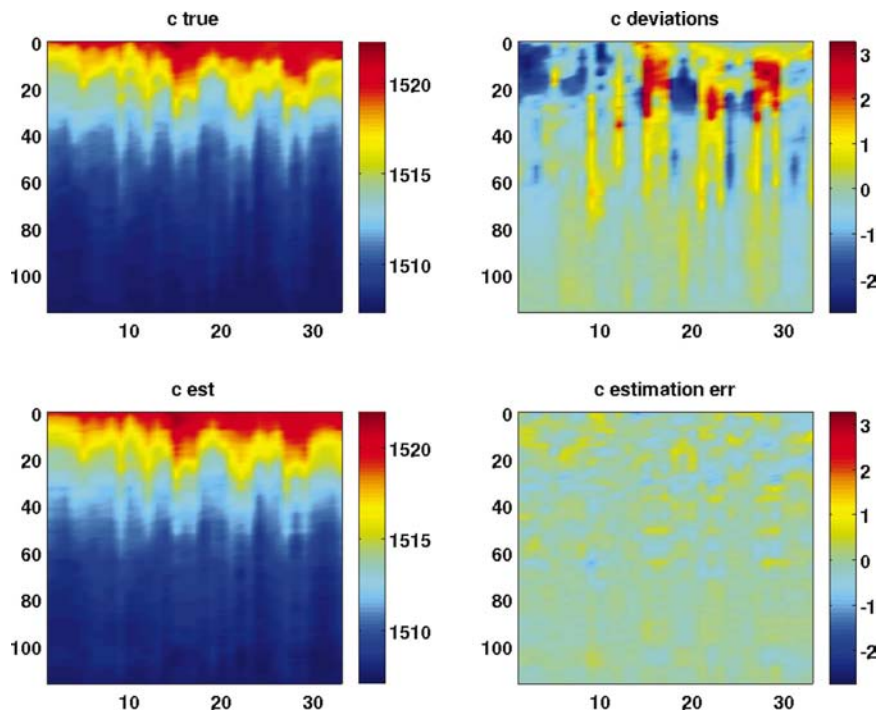


FIG. 11. Tracking internal tides over several diurnal cycles during INTIMATE 96 experiment. Horizontal axes are the sample numbers of the measured profiles. Note the different image scales in the left and right columns (the same image scale is used in each column so that the upper and lower plots can be compared). All image scales are sound speeds measured in m/s.

in which the matrix inverse is performed in the space of model parameters (analogous to a least-squares solution if the regularizing term $\mathbf{C}_u^{-1}\mathbf{u}$ is omitted). An alternative form,

$$\mathbf{u} = \mathbf{C}_u \mathbf{F}^* (\mathbf{F} \mathbf{C}_u \mathbf{F}^* + \mathbf{C}_p)^{-1} \mathbf{m}_N, \quad (45)$$

enables the matrix inverse to be performed in the data space (analogous to a minimum-norm solution). These solutions both involve the calculation of a potentially expensive generalized inverse—the choice can be made by which of the two spaces, model or data, is smaller.

The inversions in Sec. III B produced estimates of \mathbf{u} and residual errors $\partial\mathbf{p}$ (data-model misfit after 50 iterations). Setting \mathbf{C}_p and \mathbf{C}_u to be scaled identity matrices and β to 10^{-3} in Eq. (45), we substitute $\partial\mathbf{p}$ for \mathbf{m}_N and calculate $\partial\mathbf{u}$, a final

correction to our previous estimate \mathbf{u} . Figure 12 shows the adjoint solution from the previous section, the true solution, and the solution refined by the generalized inverse of \mathbf{F}_u . At almost all depths, the generalized inverse has improved the adjoint estimates.

To calculate error bars for our estimates of \mathbf{u} , we assume uncorrelated Gaussian errors and that our estimate is close to the true sound speed (so that a linear assumption holds). When this is the case, the *posterior* error covariance matrix (of the medium property estimation errors) is defined in terms of a prior pressure measurement covariance \mathbf{C}_p and a prior medium property covariance matrix \mathbf{C}_u as

$$\mathbf{R}_u = (\mathbf{F}_u^* \mathbf{C}_p^{-1} \mathbf{F}_u + \mathbf{C}_u^{-1})^{-1} \quad (46)$$

or, in the space of measurements,

$$\mathbf{R}_u = \mathbf{C}_u - \mathbf{C}_u \mathbf{F}_u^* (\mathbf{F}_u \mathbf{C}_u \mathbf{F}_u^* + \mathbf{C}_p)^{-1} \mathbf{F}_u \mathbf{C}_u. \quad (47)$$

Setting \mathbf{C}_p and \mathbf{C}_u to scaled identity matrices, setting σ_p^2 to the mean squared residual error (i.e., the data-mode misfit from the adjoint-based iterative inversion after 50 iterations), and setting σ_u^2 so that β was 25, we calculated \mathbf{R}_u using Eq. (46). Using the residual errors to set the prior covariance would not be justifiable with real data (this variance should represent *a priori* knowledge of the measurement system accuracy), but seemed a reasonable way in this illustrative simulated example to set a representative level for these variances.

Figure 13 shows the true sound speed, the adjoint estimate for the sound speed, and bounding curves three standard deviations above and below the adjoint estimate. The standard deviations are the diagonal elements of \mathbf{R}_u . The eigenvectors of \mathbf{R}_u are the principal components of the errors, and can be plotted to show the spatial correlation of the errors. Because \mathbf{C}_u is diagonal, its addition only serves to

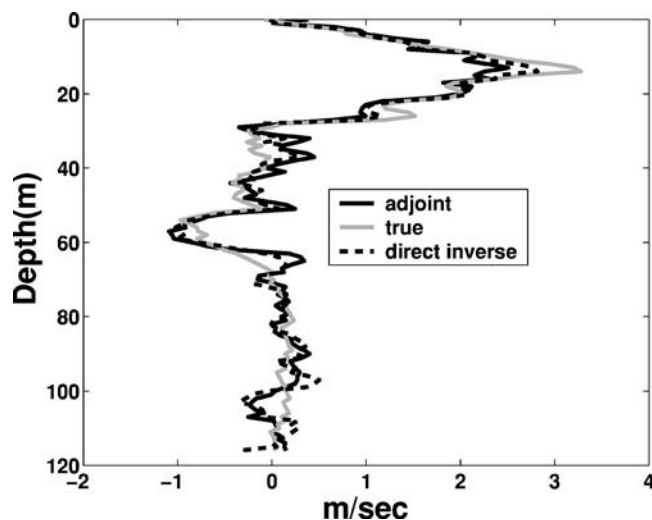


FIG. 12. Gray line shows sound speed estimate produced by iterative adjoint process. Black line shows the true sound speed. Dashed line shows the result of using generalized inverse to further reduce the residual errors from our iterative process.

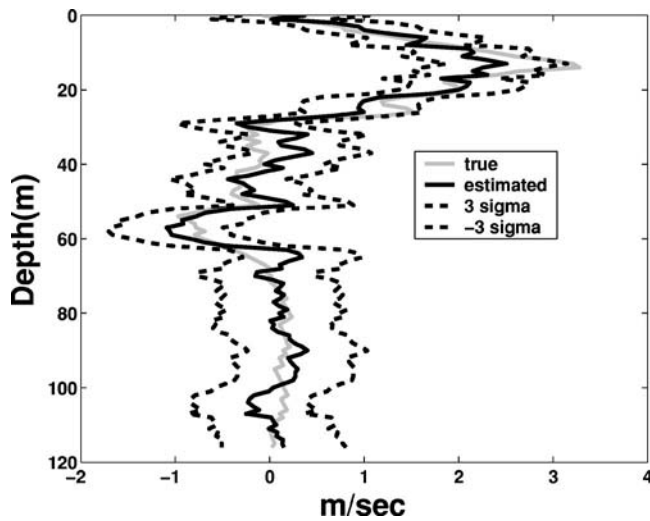


FIG. 13. Sound speed estimates with error bars (relative to baseline sound speed profile).

modify the eigenvalues of \mathbf{R}_u and not its eigenvectors, causing the eigenvalues of \mathbf{R}_u to increase or decrease according to changes in σ_u^2 .

E. Adjoint inversions (range-dependent case)

In Sec. III B, we inverted for range-independent sound speed perturbations. In this section, we invert for range-dependent sound speed perturbations. To construct a range-dependent test case, we treat the time sequence of INTIMATE 96 sound speed profiles as a sequence of profiles spanning a range interval. Figure 14 shows the range-dependent sound speed perturbation we will be imaging with our adjoint model (this is a perturbation from the mean profile). This 2D sound speed perturbation was generated in two steps. First, the sequence of INTIMATE 96 profiles was interpolated to cover 2 km in range with a 1-m spacing between profiles. Second, we isolated a short range interval where a large deviation from the mean profile occurred, by

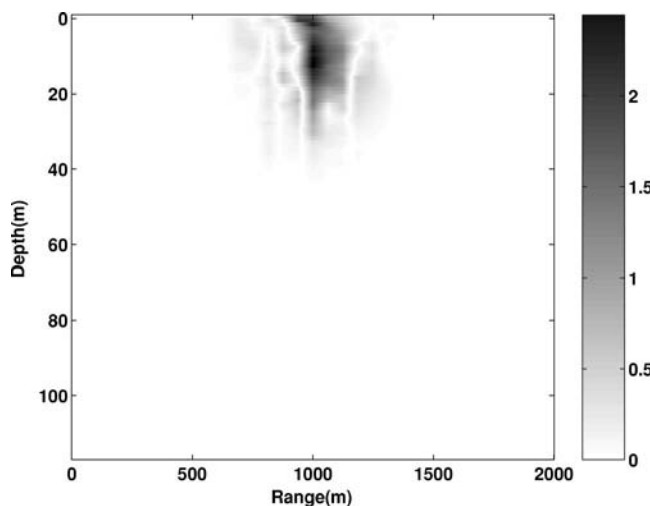


FIG. 14. Range-dependent sound speed perturbation being imaged by adjoint process. This figure shows the range-dependent feature being imaged at 1000 m. Image scale shows magnitude of perturbation in sound speed measured in m/s.

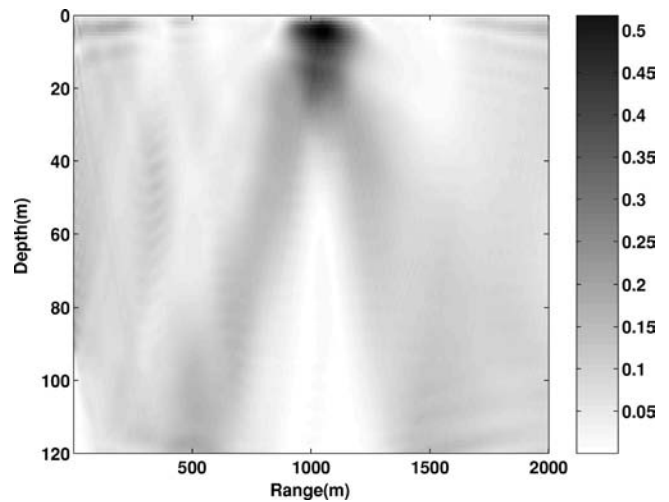


FIG. 15. Image of 2D feature at 1000 m after a single iteration of the adjoint process. Image scale shows magnitude of the correction to be applied to the sound speed field.

applying a Gaussian weighting centered at this interval. This interval covers profiles 16–18 in Fig. 11. Working with an isolated 2D feature enabled us to translate this structure to different ranges (500, 1000 and 1500 m), so that we could compare our ability to image it at these different ranges.

The parabolic equation model described in Sec. II was used to generate synthetic acoustic data on the receive array used in previous sections, for sources at depths from 5 to 90 m in 5-m steps at 100–400 Hz in 5-Hz steps. In contrast to the results in Sec. III B, where the calculated sound speed corrections were averaged over range (since range-independent sound speed fields were being inverted for), each source depth and frequency here produced a 2D map of sound speed corrections versus range and depth. Several combinations of source depths and source frequencies were used to calculate inversions for the range-dependent sound speed perturbation shown in Fig. 14. Using only a single source depth resolved the location of the range-dependent feature, but did not resolve the shape very well, producing only several highlights within the shape. The positions of these highlights changed with source depth. Using multiple frequencies did not improve the results as much as using multiple source depths, probably because it is more important to adequately sample the feature with acoustic paths than to use a wider bandwidth. Changing the frequency of the ranging signal does not appreciably change the acoustic paths visiting the 2D feature being imaged.

Figure 15 shows the summation of maps from all source depths (5–90 m in 5-m increments) and all frequencies (100–400 Hz in 5-Hz increments). Only a single iteration of the process applied in Sec. III B was used. Figures 14 and 15 show the adjoint model has imaged the 2D feature at 1000 m. Further iterations would better resolve the shape and produce an amplitude that better matches that shown in Fig. 14. This feature was similarly imaged when translated to 500 and 1500 m (although these results are not shown).

This shows that an adjoint model can produce a 2D map of large dimension using a single pass of the adjoint model. However, there is no guarantee that the map of corrections

produced by the adjoint model is unique. With too few measurements (e.g., too few source depths), it is conceivable that many 2D maps of corrections could be found to reproduce the same data-model misfit at the receive array. If the inverse problem has multiple solutions, the adjoint process provides no indication that a solution it has found is not unique. Analysis of the matrix used to linearize the problem [e.g., Eq. (19)] and its null space could be used to alert us to this ambiguity.

These results also rely upon the range-dependent sound speed perturbation being small enough that our initial guess (the mean profile, as defined in previous sections) is in some neighborhood of the objective function minimum. An initial guess too far away from the true profile could result in the inversion process getting trapped in a spurious local minimum.

IV. CONCLUSIONS

As we have shown above, an adjoint model propagates the data-model misfit (which is often identified as the adjoint field) back through the medium and calculates the perturbations to the forward model inputs needed to correct for this misfit. The adjoint model generates a sensitivity map of the model parameters \mathbf{u}_n to the observations \mathbf{m}_N , providing a direction for steepest descent in inverse problems. The alternative is to repeatedly run the forward model to test how each possible perturbation in the model parameters \mathbf{u}_n influences the observables \mathbf{m}_N . The adjoint model can thus provide a more economical way of exploring the search space (of model parameters \mathbf{u}_n) in aid of solving inverse problems. The computational savings of using an adjoint model increases with the number of unknowns in the problem being addressed, since each unknown in the problem presents a potentially distinct perturbation that must be assessed by the forward model. As a result, inversion processes using an adjoint model can accommodate large-dimensional problems, without demanding lower-dimensional representations, often required by alternative inverse methods in order to reduce the number of forward modeling runs needed to explore the search space.

To show how an adjoint model arises in the familiar setting of solving a linear system, we have shown how first-order perturbations to the PE marching solution can be formulated as a linear system and solved using a steepest descent iteration, where the adjoint of the linear system is used to calculate the gradient. The steepest-descent iteration provides an alternative to calculating the inverse of a typically very large matrix. We have also presented an alternative formulation of the adjoint, showing how the adjoint model for the marching PE solution can be derived using Lagrange multipliers, which are the “adjoint field.” The adjoint model back-propagates its field from the measurement locations back to the medium perturbations that are causing data-model misfit. Because the “adjoint field” provides a mechanism for back-propagating information in this way, there is no need to construct the potentially very large linear system matrix, needed in the former derivation. Instead, we use two models, the original PE forward model and its adjoint, to propagate their respective fields in their respective direc-

tions. A fixed-point iteration results from this formulation, which we have shown is equivalent to a steepest-descent iteration based on a gradient, calculated from the “adjoint field” back-propagated by the adjoint model.

As a demonstration of these techniques, we have applied our fixed-point iteration to the series of sound speed profiles measured during the INTIMATE 96 experiment. We use acoustic “data” calculated by our PE model with the measured sound speed profiles provided as inputs. We invert the (synthetic) acoustic “data” for the sound speed profiles, to demonstrate our technique. We have shown successful inversions for both range-independent and range-dependent sound speed features. For the range-independent case we have also shown how to provide estimates of the accuracy of our results.

The adjoint-assisted inversion presented in this paper can be viewed as a wave-theoretic approach to tomography, and thus may provide an alternative to ray-based tomography and its shortcomings.³⁰ However, we recognize that relying upon pressure as the observable quantity is problematic, and hope in the future to extend the adjoint formulation to more robust observables.

Lest we leave the impression that adjoint modeling is a panacea for all inverse problems, we describe some difficulties that must be overcome in adjoint modeling. An adjoint model only works perfectly when the analogous forward model is linear. This has not inhibited the use of adjoint models in other fields, since mildly nonlinear problems can typically be linearized and more than slightly nonlinear problems can be attacked with iterative formulations^{14,31} in which an adjoint model remains a valuable component. Nevertheless, some problems will not lend themselves to adjoint modeling, because of their inherent nonlinearity. The adjoint method finds a local minimum, with no guarantee of it being globally optimal in problems whose objective function is not convex. Like other local methods, the adjoint iterative process may require a more sophisticated optimization strategy (e.g., conjugate gradients) if the simple steepest descent presented in the paper converges too slowly (the fixed point iteration may not converge at all). Finally, an adjoint model poses the burden of requiring additional implementations of both a tangent linear and an adjoint model, over and above the implementation of the forward model. Efforts are underway to automate the construction of such models from the original forward model code.³²

ACKNOWLEDGMENTS

The paper was much improved by the suggestions of the anonymous reviewers. This work was supported by the Office of Naval Research, Grant No. N00014-01-1-0768.

¹A. Tarantola, “Inversion of seismic reflection data in the acoustic approximation,” *Geophysics* **49**(8), 1259–1266 (1984).

²O. Dorn, E. L. Miller, and C. M. Rappaport, “A shape reconstruction method for electromagnetic tomography using adjoint fields and level sets,” *Inverse Probl.* **16**, 1119–1156 (2000).

³W. Munk, P. Worcester, and C. Wunsch, *Ocean Acoustic Tomography* (Cambridge U.P., New York, 1995), pp. 318–319.

⁴C. Wunsch, *The Ocean Circulation Inverse Problem* (Cambridge U.P., New York, 1996), pp. 362–391.

- ⁵A. F. Bennett, *Inverse Methods in Physical Oceanography* (Cambridge U.P., New York, 1992), pp. 112–135.
- ⁶S. J. Norton, “Iterative inverse scattering algorithms: Methods of computing Fréchet derivatives,” *J. Acoust. Soc. Am.* **106**, 2653–2660 (1999).
- ⁷M. Asch, J.-C. L. Gac, and P. Helluy, “An adjoint method for geoacoustic inversions,” in *Proceedings of the 2nd Conference on Inverse Problems, Control and Shape Optimization* (Carthage, Tunisia, 2002).
- ⁸A. Thode, “An adjoint/normal-mode approach for computing multiple-order derivatives of pressure with respect to environmental parameters and frequency: range-independent waveguide,” *J. Acoust. Soc. Am.*, (2003) (submitted).
- ⁹M. D. Collins and W. A. Kuperman, “Focalization: Environmental focusing and source localization,” *J. Acoust. Soc. Am.* **90**, 1410–1422 (1991).
- ¹⁰P. Gerstoft, “Inversion of seismoacoustic data using genetic algorithms and a posteriori probability distributions,” *J. Acoust. Soc. Am.* **95**, 770–782 (1994).
- ¹¹A. Tolstoy, N. Chapman, and G. Brooke, “Workshop ’97: Benchmarking for geoacoustic inversion in shallow water,” *J. Comput. Acoust.* **6**(1), 1–28 (1998).
- ¹²A. Tolstoy, “Linearization of the matched field processing approach to acoustic tomography,” *J. Acoust. Soc. Am.* **91**, 781–787 (1992).
- ¹³V. V. Goncharov and A. G. Voronovich, “An experiment on matched-field acoustic tomography with continuous wave signals in the Norway Sea,” *J. Acoust. Soc. Am.* **93**, 1873–1881 (1993).
- ¹⁴D. G. Luenberger, *Optimization by Vector Space Methods* (Wiley, New York, 1969).
- ¹⁵P. R. McGillivray and D. W. Oldenburg, “Methods for calculating Fréchet derivatives and sensitivities for the non-linear inverse problem: a comparative study,” *Geophys. Prospect.* **38**, 499–524 (1990).
- ¹⁶D. E. Kirk, *Optimal Control Theory* (Prentice–Hall, Englewood Cliffs, NJ, 1970), pp. 184–240.
- ¹⁷W. L. Brogan, *Modern Control Theory*, 3rd ed. (Prentice–Hall, Englewood Cliffs, NJ, 1991).
- ¹⁸I. M. Gelfand and S. V. Fomin, *Calculus of Variations* (Prentice–Hall, Englewood Cliffs, NJ, 1963), pp. 218–225.
- ¹⁹A. F. Bennett, *Inverse Modeling of the Ocean and Atmosphere* (Cambridge U.P., New York, 2002), pp. 18–29.
- ²⁰A. J. Devaney, “A filtered backpropagation algorithm for diffraction tomography,” *Ultrason. Imaging* **4**, 336–350 (1982).
- ²¹E. Coelho, “Mesoscale-small scale oceanic variability effects on underwater acoustic signal propagation,” in *Impact of Littoral Environmental Variability on Acoustic Predictions and Sonar Performance*, edited by N. G. Pace and F. B. Jensen (Kluwer Academic, Dordrecht, The Netherlands, 2002), pp. 49–54.
- ²²H. A. Deferrari, N. J. Williams, and H. B. Nguyen, “Variability, coherence and predictability of shallow water acoustic propagation in the straits of Florida,” in *Impact of Littoral Environmental Variability on Acoustic Predictions and Sonar Performance*, edited by N. G. Pace and F. B. Jensen (Kluwer Academic, Dordrecht, The Netherlands, 2002), pp. 245–254.
- ²³A. R. Robinson, P. Abbott, P. F. J. Lermusiaux, and L. Dillman, “Transfer of uncertainties through physical-acoustical-sonar end-to-end systems: a conceptual basis,” in *Impact of Littoral Environmental Variability on Acoustic Predictions and Sonar Performance*, edited by N. G. Pace and F. B. Jensen (Kluwer Academic, Dordrecht, The Netherlands, 2002), pp. 409–416.
- ²⁴A. Warn-Varnas, S. Chin-Bing, D. King, J. Hawkins, K. Lamb, and M. Teixeira, “Yellow Sea internal solitary wave variability,” in *Impact of Littoral Environmental Variability on Acoustic Predictions and Sonar Performance*, edited by N. G. Pace and F. B. Jensen (Kluwer Academic, Dordrecht, The Netherlands, 2002), pp. 409–416.
- ²⁵F. B. Jensen, W. A. Kuperman, M. B. Porter, and H. Schmidt, *Computational Ocean Acoustics* (AIP, Woodbury, NY, 1994), pp. 366–375.
- ²⁶J. C. Strikwerda, *Finite Difference Schemes and Partial Differential Equations* (Chapman and Hall, New York, 1989), pp. 323–350.
- ²⁷Y. Stéphan, X. Démoulin, T. Folégot, S. M. Jesus, M. B. Porter, and E. Coelho, “Acoustical effects of internal tides on shallow water propagation: an overview of the INTIMATE96 experiment,” in *Experimental Acoustic Inversion Methods for Exploration of the Shallow Water Environment*, edited by A. Caiti, J. P. Hernand, S. M. Jesus, and M. B. Porter (Kluwer Academic, Dordrecht, The Netherlands, 2000), pp. 19–38.
- ²⁸The singular values were within 1% of each other out to 110 of 117 of the eigenvectors. Similarly, the eigenvectors were correlated to within 1% out to 80 of 117 eigenvectors, and to within 5% out to 110 of 117 eigenvectors. The singular values for eigenvectors 80 and 110 were 60 and 80 dB down from the first eigenvector, respectively.
- ²⁹A. Tarantola, *Inverse Problem Theory* (Elsevier Science, New York, 1987), pp. 187–294.
- ³⁰G. A. Athanassoulis and E. K. Skarsoulis, “Arrival-time perturbations of broadband tomographic signals due to sound-speed disturbances. A wave-theoretic approach,” *J. Acoust. Soc. Am.* **97**, 3575–3588 (1995).
- ³¹P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization* (Academic, San Diego, CA, 1981).
- ³²R. Giering and T. Kaminsky, “Recipes for adjoint code construction,” *ACM Trans. Math. Softw.* **24**(4), 437–474 (1998).