

ADJOINT-ASSISTED INVERSION FOR SHALLOW WATER ENVIRONMENT PARAMETERS

PAUL HURSKY AND MICHAEL B. PORTER

*Science Applications International Corporation
1299 Prospect Street, Suite 305, La Jolla, CA 92037, USA
E-mail: paul.hursky@saic.com, michael.b.porter@saic.com*

BRUCE D. CORNUELLE, W. S. HODGKISS AND W. A. KUPERMAN

*Scripps Institution of Oceanography
University of California in San Diego, La Jolla, CA 92093-0701, USA
E-mail: bdc@ucsd.edu, wsh@mpl.ucsd.edu, wak@mpl.ucsd.edu*

The adjoint of a forward model can back-propagate mismatch between observations and their predictions and produce the corrections to the forward model inputs that caused the mismatch. As an example of this process, the adjoint of a parabolic equation propagation model is used to invert errors in pressure predictions at a receiver for sound speed perturbations due to internal tides.

1 Introduction

Using the adjoint of a forward model has the potential to sharply reduce the number of modeling runs usually needed to achieve an inversion. Typically, an inversion process varies the parameters of a forward model, running the forward model for many candidate sets of parameter values until the forward model matches the data. Unfortunately, this often requires many runs to adequately search the space of unknown parameters. We present an alternative technique based on the adjoint of the forward model. The adjoint model back-propagates a mismatch between model predictions and measured observations, producing corrections to model input parameters along the trajectory of the forward model. A single run of the adjoint model thus duplicates many forward modeling runs. In this paper, we will use the adjoint of a parabolic equation propagation model to invert for sound speed perturbations due to internal tides.

Adjoint methods have been used in many fields. Ref. [1] suggested adjoint methods for tomography. Refs. [2] and [3] present how adjoint methods can be used to assimilate data into oceanographic models. Ref. [6] derives the adjoint of the Helmholtz equation in terms of continuous variables and discusses the connection between adjoint techniques and time reversal. 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. The adjoint model back-propagates (time-reverses) the perturbed field to the unknown medium perturbations (viewed as sources of diffraction). Ref. [7] shows how adjoints can be used to calculate Fréchet derivatives used to solve inverse problems of the sort we address. Refs. [4] and [5] present how adjoints arise in optimal control theory, where their use is

known as the Pontryagin Principle.

In Section 2, we derive a tangent linear model for the parabolic equation. In Section 3, we show how the adjoint of this model can be used to solve acoustic inverse problems. In Section 4, we show in simulation how these models can be used to estimate the sound speed perturbations caused by internal tides.

2 Tangent linear model for the parabolic equation

The standard homogeneous PE equation (with no source in the medium) is

$$2ik_0 \frac{\partial p}{\partial r} + \frac{\partial^2 p}{\partial z^2} + k_0^2 (n^2 - 1) p = 0. \quad (1)$$

Expanding this equation in terms of perturbations in pressure p and index of refraction squared n^2 to first order in ε ,

$$n^2 = n_0^2 + \varepsilon n_1^2,$$

$$p = p_0 + \varepsilon p_1 + \dots,$$

yields

$$2ik_0 \frac{\partial p_1}{\partial r} + \frac{\partial^2 p_1}{\partial z^2} + k_0^2 [n_0^2(r, z) - 1] p_1 = -k_0^2 n_1^2(r, z) p_0. \quad (2)$$

We will use a discrete formulation of Eqs. 1 and 2 based on the implicit finite differences scheme described in Section 6.6 of Ref. [8]. A finite difference approximation to Eq. 1 (the unperturbed problem, with $\varepsilon = 0$) produces a marching solution for the zeroth-order pressures \mathbf{p}_0 ,

$$\mathbf{p}_0(r + \delta r) = \mathbf{F}(r) \mathbf{p}_0(r), \quad (3)$$

where \mathbf{p}_0 is a vector sampled in depth. Matrix $\mathbf{F}(r)$ is a symmetric, tri-diagonal matrix with diagonal elements

$$-\frac{2}{h^2} + k_0^2 (n_0^2(r, z) - 1),$$

and super-diagonal and sub-diagonal elements $\frac{1}{h^2}$. The diagonal elements contain $n_0^2(r, z)$, the zeroth-order index of refraction squared, which varies with range and depth. Matrix $\mathbf{F}(r)$ is used to propagate pressure vectors (sampled in depth) one range step at a time, given an initial pressure (or starter field). Equation 2 (the first-order perturbation terms, of order ε) generates a marching solution for the first-order pressure vectors \mathbf{p}_1 ,

$$\mathbf{p}_1(r + \delta r) = \mathbf{F}(r) \mathbf{p}_1(r) + \mathbf{G}(r) \mathbf{u}(r). \quad (4)$$

$\mathbf{F}(r)$ is the same as in Eq. 3. $\mathbf{G}(r)$ is a diagonal matrix with values $-k_0^2 \mathbf{p}_0(r)$ (i.e. the zeroth-order pressures sampled in depth at a particular range). Vector $\mathbf{u}(r)$ contains $n_1^2(r, z)$ sampled in depth. Equation 4 has a forcing function proportional to the zeroth-order pressure $\mathbf{p}_0(r)$ and the first-order perturbation to the index of refraction squared

$n_1^2(r, z)$. Matrices $\mathbf{F}(r)$ and $\mathbf{G}(r)$ define our tangent linear model and also form the basis of our adjoint model in the following sections.

3 Using an adjoint model to solve acoustic inverse problems

We will use Eq. 4 to formulate an adjoint method in terms of first-order perturbations to the pressures and the environmental parameters. We will no longer use subscripts 0 and 1 to indicate the order of the perturbation. All \mathbf{p}_r and \mathbf{u}_r will be first-order perturbations, or corrections to zeroth-order quantities calculated using our initial guess at the environmental parameters we are inverting for. These \mathbf{p}_r and \mathbf{u}_r will be vectors, sampled in depth, with subscripts r that indicate range indexes. Equation 4 in terms of this notation is

$$\mathbf{p}_{r+1} = \mathbf{F}_r \mathbf{p}_r + \mathbf{G}_r \mathbf{u}_r.$$

Note both \mathbf{F}_r and \mathbf{G}_r are functions of range. Matrix \mathbf{F}_r propagates the pressure correction vector \mathbf{p}_r at range index r one range step to \mathbf{p}_{r+1} . The environmental parameter correction vector \mathbf{u}_r influences the propagation of the pressure via a known matrix \mathbf{G}_r . The i th element of vector \mathbf{p}_r contains the pressure correction at range index r at the i th sampled depth. Similarly for vector \mathbf{u}_r . Matrices \mathbf{F} and \mathbf{G} , derived in Section 2, form a tangent linear model of the original, non-linear PE model. \mathbf{F} and \mathbf{G} are functions of the zeroth-order environmental parameters and the zeroth-order pressures (calculated using the original non-linear propagation model with the zeroth-order environmental parameters as inputs). The vectors \mathbf{p}_r (sampled in depth) are the pressure increments calculated by the tangent linear model \mathbf{F} and \mathbf{G} as corrections to the zeroth-order pressures due to the environmental correction vectors \mathbf{u}_r (also sampled in depth).

To solve for \mathbf{u}_r , we formulate an objective function $J(\mathbf{p}, \mathbf{u}, \lambda)$ to be minimized:

$$J(\mathbf{p}, \mathbf{u}, \lambda) = \frac{1}{2}(\mathbf{p}_N - \mathbf{m}_N)^2 + \sum_{r=1}^N \lambda_r^T (\mathbf{p}_r - \mathbf{F}_{r-1} \mathbf{p}_{r-1} - \mathbf{G}_{r-1} \mathbf{u}_{r-1}) + \frac{1}{2} \sum_{r=0}^{N-1} \mathbf{u}_r^2. \quad (5)$$

The first term in J seeks to minimize the mismatch between the measured pressure increment \mathbf{m}_N and the modeled pressure increment \mathbf{p}_N , both at range index N . 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}_r from the source to the receiver using our tangent linear model $\{\mathbf{F}_r, \mathbf{G}_r\}$ with the environmental parameter corrections \mathbf{u}_r as driving functions. Given a solution for \mathbf{u}_r , the zeroth-order pressure plus the pressure correction \mathbf{p}_N calculated using \mathbf{u}_r should reproduce the measured pressure. The second term uses Lagrange multipliers λ_r (vectors at each range, sampled in depth) to enforce the hard constraint that the \mathbf{p}_r and \mathbf{u}_r must be consistent with the model $\{\mathbf{F}_r, \mathbf{G}_r\}$. The third term is a regularizing term to minimize the amplitude of the environmental perturbations.

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}_r , in addition to the already large number of unknowns \mathbf{u}_r . We will show how minimizing the objective function above leads to an iteration that seems to be a much more direct way of inverting for the \mathbf{u}_r than repeatedly running the forward model to explore the surface J as a function of \mathbf{u}_r .

Note that J is a function of \mathbf{u}_r at all ranges and depths, so its minimization has the potential to resolve range-dependent features. We will demonstrate inversions using measurements at a single frequency and a single source depth for range-independent features in Section 4. To resolve range-dependent features requires a richer set of measurements, using more sources and a wider band.

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

$$\frac{\partial J}{\partial \mathbf{p}_N} = \mathbf{p}_N - \mathbf{m}_N + \lambda_N, \quad (6)$$

$$\frac{\partial J}{\partial \mathbf{p}_r} = \lambda_r - \mathbf{F}_r^T \lambda_{r+1}, \quad (7)$$

$$\frac{\partial J}{\partial \mathbf{u}_r} = \mathbf{u}_r - \mathbf{G}_r^T \lambda_{r+1}. \quad (8)$$

Setting the partial with respect to \mathbf{p}_N to zero in Eq. 6 yields

$$\lambda_N = \mathbf{m}_N - \mathbf{p}_N, \quad (9)$$

initializing λ_N to the mismatch between measured and modeled pressures at range N . Setting the partials with respect to \mathbf{p}_r to zero in Eq. 7 produces a recursion relation,

$$\lambda_r = \mathbf{F}_r^T \lambda_{r+1}, \quad (10)$$

that enables us to propagate the Lagrange multipliers λ_r from the receiver to the source. The λ_r can be viewed as a field propagated by the adjoint model. The starter field of the adjoint model, given by Eq. 9, is the mismatch in our observations at the receiver (i.e. mismatch with our predictions, produced by our zeroth-order model with the zeroth-order environmental parameters as inputs). We are inverting for the corrections \mathbf{u}_r to these zeroth-order environmental parameters that will account for this mismatch. Setting the partials with respect to the \mathbf{u}_r to zero in Eq. 8 yields

$$\mathbf{u}_r = \mathbf{G}_r^T \lambda_{r+1}, \quad (11)$$

producing equations for \mathbf{u}_r at each range index r in terms of the Lagrange multipliers λ_{r+1} .

Equations 9, 10, and 11, in the order presented, can be used to calculate \mathbf{u}_r , the first-order corrections to the environmental parameters driving our forward model. However, in the non-linear problem we address, the gradients used to derive these equations may only be accurate in a small neighborhood about the zeroth-order pressure predictions. As a result, we used these equations as the basis for an iterative procedure which is able to follow the curvature of our objective surface J that will inevitably arise in some configurations. At each iteration, we calculate a new set of \mathbf{u}_r , given the current $\{\mathbf{F}_r, \mathbf{G}_r\}$. We use these \mathbf{u}_r to adjust the environmental parameters that are then used to re-calculate a new set of zeroth-order pressures at all ranges (either using the tangent linear model, or the

original non-linear model). If these calculated pressures match our measurements, we have found a solution to our problem. If not, the new zeroth-order pressures and the corrected environmental parameters are embedded in \mathbf{F} and \mathbf{G} , and the adjoint model is used to calculate another set of corrections to the environmental parameters.

The tangent linear model can be used to calculate the forward sensitivity in a problem (e.g. how sensitive \mathbf{p}_N is to perturbations \mathbf{p}_0 and \mathbf{u}_r). The adjoint model can be used to calculate the backward sensitivity in a problem (e.g. how sensitive \mathbf{p}_0 and \mathbf{u}_r are to perturbations \mathbf{p}_N).

4 Inverting for INTIMATE 96 internal tides (simulated results)

We will demonstrate the adjoint method described in Section 3 on an ensemble of sound speed profiles measured during the INTIMATE 96 experiment (see Ref. [9]) when the passage of internal tides was clearly visible (see Figure 2). We use our PE model to synthesize pressure measurements at 400 Hertz on a vertical line array at a range of 2 km from a source at depth of 50 meters, using each individual profile to generate a measured pressure vector. Each of these pressure vectors was inverted to estimate the sound speed profile which was used to synthesize it, using the iterative process outlined in Section 3 (using tangent linear and adjoint models derived for our PE model). The mean profile of the entire ensemble served as the initial guess for each inversion. We set up our inversions to solve for coefficients of empirical orthogonal functions (EOFs), averaged over the deviations from the mean profile.

In Section 4.1, because our inversion process is essentially a steepest descent method, we use our PE tangent linear model to assure ourselves that we are reasonably close to a solution. In Section 4.2, we show the results of our inversions. Note that we are using *synthetic* acoustic data to demonstrate the feasibility of inverting for a sequence of internal tides measured during the INTIMATE 96 experiment. The experiment configuration was fixed-fixed, with the line joining the source and receiver perpendicular to the passage of the internal tides, so the inversion was formulated to be *range-independent*. Each profile from the INTIMATE 96 sequence was estimated by a separate inversion.

4.1 Tangent linear modeling to verify we are in linear regime

To assess the linearity of our experimental configuration, we compare pressures produced by our fully non-linear PE model and its tangent linear version, given perturbations on the order of those actually observed during INTIMATE 96. We refer to the fully non-linear PE propagation model as $p = F(c)$. We refer to the tangent linear model as $\delta p = \delta F(c_0, \delta c)$, where we have written δF as a function of both c_0 , the baseline profile, and δc , its perturbation. Note the tangent linear model depends upon both c_0 and δc . The baseline profile is the mean profile calculated from the entire set of available sound speed profile measurements during the INTIMATE 96 experiment. We chose the profile that deviated the most from the mean profile as a test case. We wanted to see how closely the tangent linear model matched the original PE model in predicting pressure perturbations in the INTIMATE 96 configuration. We ran the original PE model on our baseline sound speed profile c_0 and on our perturbed sound speed profile $c_1 = c_0 + \delta c$, producing p_0 and p_1 . We ran our tangent linear model to produce δp , to see if it would reproduce the perturbation in pressure, $p_1 - p_0$, due to the perturbation in sound speed δc . Equations 12 through 15,

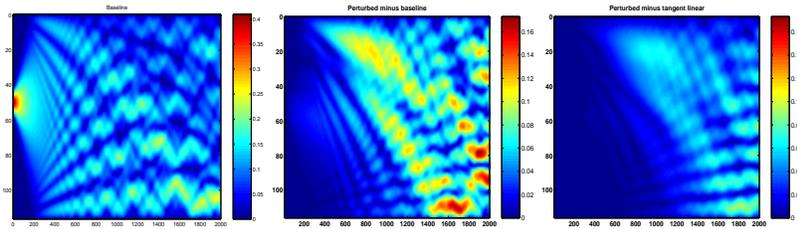


Figure 1. Baseline pressure (left image). Magnitude differences between perturbed and baseline pressures (center image) and between perturbed and tangent linear pressures (right image). Center and right images have the same color scale.

$$p_0 = F(c_0), \quad (12)$$

$$p_1 = F(c_0 + \delta c), \quad (13)$$

$$\delta p = \delta F(c_0, \delta c), \quad (14)$$

$$p_2 = p_0 + \delta p, \quad (15)$$

summarize how the three relevant pressures, p_0 , p_1 , and p_2 , were calculated. Figure 1, containing three images, shows several combinations of these pressures. The left image shows the baseline pressure magnitude, p_0 . The following two images have the same color scale, so that their values can be compared. The center and right images show magnitudes of pressure differences. The center image shows the perturbed minus baseline pressures, $|p_1 - p_0|$. The right image shows perturbed minus tangent linear pressures, $|p_1 - p_2|$. The significantly lower amplitudes in the right image indicate that the tangent linear model is able to match the predictions of the fully non-linear PE model reasonably well, at least given the size of the sound speed perturbations and the source-receiver range in our configuration. The quality of the tangent linear model degrades with increasing range, but as we will see in the next section, we are still close enough to a solution that our iterative process resolves the internal tides.

These results indicate that although the problem we are addressing is not strictly linear, it remains reasonable to attempt our iterative process, which we expect can tolerate slightly non-linear problems, because it presumably can follow a non-linear basin of attraction if started out close enough to the final solution point.

4.2 Adjoint iterative process results

Figure 2 shows the result of applying the process described in Section 3 to each of a sequence of profiles measured during the INTIMATE 96 experiment. The entire sequence of measured profiles is shown in the upper left hand plot, with the horizontal and vertical

ADJOINT-ASSISTED INVERSION

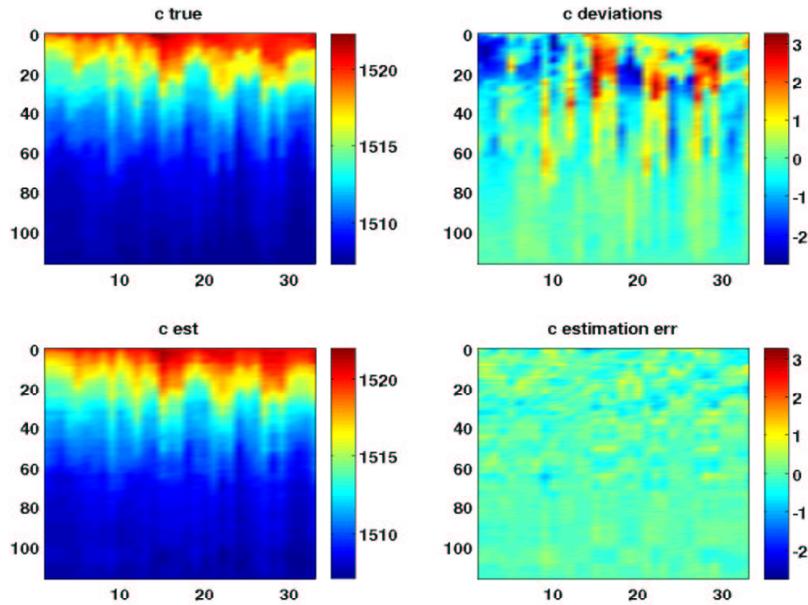


Figure 2. Inversion results for internal tides measured during INTIMATE 96 experiment (with simulated acoustic data). Upper left image shows the true sound speed profiles versus time. The lower left shows the estimated profiles versus time. The upper right image shows the sound speed profile deviations from the mean profile (what we solve for directly using the adjoint method). The lower right image shows the estimation errors (the difference between the upper left and lower left images).

axes corresponding to time and depth. Each profile was processed independently of the others. We ran our adjoint-based inversion for 50 iterations on each profile, using the mean profile as an initial guess. The resulting estimated profiles are shown in the lower left plot. Clearly, the coarse features have been resolved. For a more quantitative assessment, we show the two plots on the right. The upper right plot shows the deviations from the mean profile (the corrections we actually invert for). The lower right plot shows the estimation errors (i.e. the difference between the measured and estimated profiles). Both right-hand plots have the same color scale. The relatively smaller magnitudes of the estimation errors compared to the deviations from the mean indicate that our adjoint process has done a good job resolving the internal tides during this interval.

5 Conclusions

We have shown how the adjoint of a parabolic equation forward model can be used to invert pressure measurements for sound speed perturbations in the water column. The adjoint technique we have presented uses far fewer propagation model runs than techniques currently being used, in which the forward model is run for each candidate point in a high-dimensional search space.

Acknowledgments

We thank Aaron Thode for pointing out References [4], [6], and [7].

References

1. Walter Munk, Peter Worcester and Carl Wunsch, *Ocean Acoustic Tomography* (Cambridge University Press, 40 West 20th Street, New York, NY 10011-4211, 1995) pp. 318-319.
2. Carl Wunsch, *The Ocean Circulation Inverse Problem* (Cambridge University Press, 40 West 20th Street, New York, NY 10011-4211, 1996) pp. 362-391.
3. Andrew F. Bennett, *Inverse methods in physical oceanography*, Chapter 5, (Cambridge University Press, 40 West 20th Street, New York, NY 10011-4211, 1992) pp. 112-135.
4. Donald E. Kirk, *Optimal Control Theory* (Prentice-Hall, Englewood Cliffs, NJ, 07632, 1970) pp. 184-240.
5. I. M. Gelfand and S. V. Fomin, *Calculus of variations* (Prentice-Hall, Englewood Cliffs, NJ, 07632, 1963) pp. 218-225.
6. Albert Tarantola, Inversion of seismic reflection data in the acoustic approximation, *Geophysics* **49**, 1259–1266 (1984).
7. Stephen J. Norton, Iterative inverse scattering algorithms: Methods of computing Fréchet derivatives, *J. Acoust. Soc. Am.* **106**, 2653–2660 (1999).
8. F. B. Jensen, W. A. Kuperman, M. B. Porter and H. Schmidt, *Computational Ocean Acoustics* (AIP Press, 500 Sunnyside Boulevard, Woodbury, NY 11797-2999, 1994) pp. 366-375.
9. Y. 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*, ed. by A. Caiti, J. P. Hermand, S. M. Jesus, and M. B. Porter (Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000) pp. 19–38.