

Efficient Calculation of Jacobian and Adjoint Vector Products in Wave Propagational Inverse Problem Using Automatic Differentiation*

Thomas F. Coleman[†] Fadil Santosa[‡] Arun Verma[§]

Abstract

Wave propagational inverse problems arise in several applications including medical imaging and geophysical exploration. In these problems, one is interested in obtaining the parameters describing the medium from its response to excitations. The problems are characterized by their large size, and by the hyperbolic equation which models the physical phenomena. The inverse problems are often posed as a nonlinear data-fitting where the unknown parameters are found by minimizing the misfit between the predicted data and the actual data. In order to solve the problem numerically using a gradient-type approach, one must calculate the action of the Jacobian and its adjoint on a given vector. In this paper, we explore the use of automatic differentiation (AD) to develop codes that perform these calculations.

We show that by exploiting structure at 2 scales, we can arrive at a very efficient code whose main components are produced by AD. In the first scale we exploit the time-stepping nature of the hyperbolic solver by using the “Extended Jacobian” framework. In the second (finer) scale, we exploit the finite difference stencil in order to make explicit use of the sparsity in the dependence of the output variables to the input variables. The main ideas in this work are illustrated with a simpler, one-dimensional version of the problem. Numerical results are given for both one- and two- dimensional problems. We present computational templates that can be used in conjunction with optimization packages to solve the inverse problem.

1 Introduction

In the type of wave propagational inverse problems under consideration, the goal is to determine parameters, such as soundspeed distribution and density distribution, from measured

*This research is sponsored in part by the Applied Mathematical Sciences Research Program (KC-04-02) of the Office of Energy Research of the US Department of Energy under grants DE-FG02-97ER25013 and DE-FG02-94ER25225, and also by the National Science Foundation grant DMS 9503114, and by the Air Force Office of Scientific Research grant F49620-95-I-0305

[†]Computer Science Department and Center for Applied Mathematics, Cornell University, Ithaca NY 14850.

[‡]Minnesota Center for Industrial Mathematics, School of Mathematics, University of Minnesota, Minneapolis, MN 55455.

[§]Computer Science Department, Cornell University, Ithaca NY 14850.

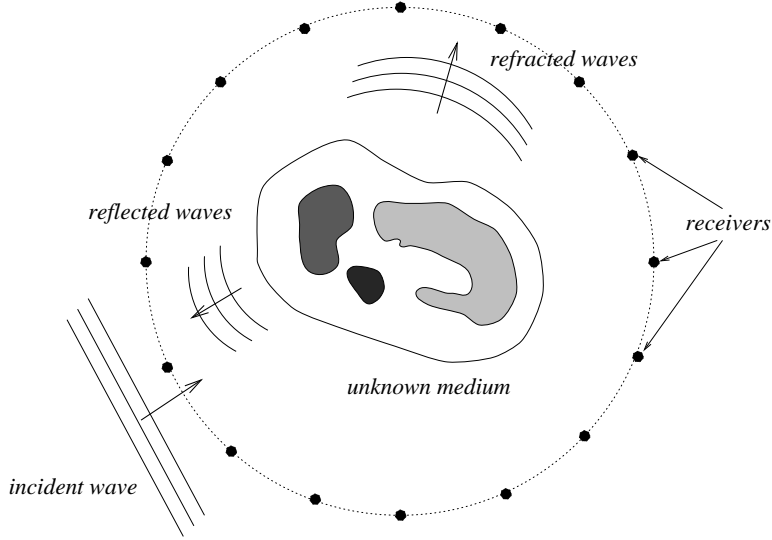


Figure 1: In this figure, the problem is to identify the unknown medium. An incident wave is generated, and as it travels into the medium being probed, reflected and refracted signals are generated. These are captured at the receivers. Several such experiments are carried out for a set of incident disturbances. The inverse problem is to find the properties of the unknown medium from the collected data.

data, which are collected at a set of receivers. Figure 1 explains the situation. An incident disturbance is generated, as it travels in the unknown medium and produces reflections and refractions. This information is collected at receivers placed at a set of locations. Several such experiments are carried out for a set of incident disturbances. The inverse problem is to determine properties of the unknown medium from the set of measured response.

Problems of this type arise in several applications including geophysical exploration and medical imaging. A common feature in these applications is that the problem is very large. Typically, the number of unknowns and equations could be in the range of 10^3 to 10^6 . Often, the most convenient way to solve this type of inverse is to pose it as an optimization, either using nonlinear leastsquares [16, 12] or other approach specialized to take advantage of the properties afforded by the particular application [14, 1]. In any event, what one will need for computation is derivative information of the relation between medium parameters and data. Because of the size of the problem, we cannot compute and store the entire Jacobian of the function, but rather, we must find ways of computing the action of the Jacobian and its transpose on a given vector, or the so called direct and adjoint products.

The goal of this work is to show that efficient calculation of direct and adjoint product is possible. The approach we take is to use automatic differentiation (AD) while exploiting structure to the extent possible. We emphasize that without taking advantage of structure, a direct application of current AD technology to the codes simulation the wave phenomena will lead to memory problems.

As we will show in the next section, the wave propagation can be modeled effectively

using time-stepping finite difference schemes. The time-stepping nature of the scheme can be exploited using the general Extended Jacobian framework [3, 4]. The spatial discretization by finite differences reveal further structure. Each finite difference stencil encodes the dependence of a computed intermediate variable on other variables. In particular, it shows that there is an inherent sparsity in the Jacobian. A combination of these structure exploitations allows us to overcome the problem posed by size, and its consequence on memory requirements.

In our implementation, we apply AD on the finite difference stencils and use the resulting codes to assemble a procedure for computing the Jacobian and adjoint vector products. The resulting code is as efficient as those that are obtained by directly performing summations-by-parts calculation on the simulation program. The advantage here is that we have avoided the error-prone and tedious procedure [13]. Instead, we can view the code writing process at a higher level, leaving the most difficult parts to AD.

The plan of this article is as follows. We proceed with a short introduction to inverse problem for acoustic waves. The model for the physics and its numerical discretization are described in Section 2. In Section 3, we review the Extended Jacobian framework and show how it can be used for our problem. We will provide templates for calculating the adjoint-vector product. The stencil approach its implementation is presented in Section 4. We will also show how the stencil can be described at a higher level as projections. Templates for calculating Jacobian and adjoint vector products that uses stencils are given. Section 5 summarizes our experience with this method of computation. A final section contains concluding remarks.

We acknowledge helpful discussions with William Symes, who has a similar on-going effort on automatic differentiation as ours [15]. Some of the ideas in this work were inspired by his presentation at the Institute for Mathematics and its Applications, Minnesota, in July 1997.

2 Inverse Problems and Numerical Modeling

2.1 One-dimensional problem

Consider a bar or string of unit length whose sound speed is location dependent. Let $u(x, t)$ represent a measure of the disturbance at time t and location x . Then u satisfies the wave equation

$$u_{tt} = c^2(x) u_{xx} \quad \text{for } 0 < x < L, \quad (1a)$$

where $c(x)$ is the sound speed of the medium. We assume that the medium is quiescent at $t = 0$,

$$u(x, 0) = 0, \quad \text{and} \quad u_t(x, 0) = 0, \quad 0 \leq x \leq L. \quad (1b)$$

Disturbance is introduced at the boundary $x = 0$ as a Neumann boundary condition

$$u_x(0, t) = f(t), \quad \text{for } t > 0. \quad (1c)$$

We will assume that $f(t)$ is compactly supported away from $t = 0$. On the right end, we assume a radiation boundary condition

$$[u_t - c(x)u_x]_{x=L} = 0. \quad (1d)$$

We are given $u(0, t) = g(t)$ for $0 < t < T$. The problem is to find the unknown $c(x)$.

A convenient way to view the problem is to define the forward map as one that associates a given $c(x)$ with a boundary data $u(0, t)$. Let

$$A[c](t) := u(0, t), \quad 0 < t \leq T,$$

where it is understood that the evaluation of $A[c](\cdot)$ is through the initial-boundary value problem (IBVP) in (1). A least-squares formulation of this problem is to solve the minimization

$$\min_{c(x)} \int_0^T |A[c](t) - g(t)|^2 dt. \quad (2)$$

A common discretization for this problem is to use finite difference methods. Let

$$u_i^k \approx u(x = x_i, t = t_k) \quad \text{where} \quad x_i = i\Delta x, \quad i = 0 : n, \quad t_k = k\Delta t, \quad k = 0 : m,$$

and $\Delta x = L/n$, and $\Delta t = \lambda\Delta x$ for some $\lambda > 0$. A second order finite difference is chosen. The partial differential equation in (1a) is replaced by

$$u_i^{k+1} = 2(1 - \lambda^2 c_i^2)u_i^{k-1} - u_i^k + \lambda^2 c_i^2 (u_{i+1}^k + u_{i-1}^k), \quad \text{for } i = 1 : n-1, \quad k \geq 0. \quad (3a)$$

We use the initial conditions

$$u_i^{-1} = u_i^0 = 0. \quad (3b)$$

We discretize the boundary conditions as

$$u_0^{k+1} = 2(1 - \lambda^2 c_0^2)u_0^k - u_0^{k-1} + 2\lambda^2 c_0^2 u_1^k - 2f^k \lambda^2 c_0^2 \Delta x. \quad (3c)$$

for the inhomogeneous Neumann condition on the left end, and

$$u_n^{k+1} = u_n^k - c_n \frac{\Delta t}{\Delta x} (u_n^k - u_{n-1}^k) \quad (3d)$$

for the radiation boundary condition on the right. The discrete version of the forward map is obtained by running the finite difference forward in time and recording the left end value for u_i^k , that is,

$$A[c]^k := u_0^k$$

A way to describe the function evaluation is through a vector notation. Let us write the vectors $\mathbf{u}^k = [u_0^k, u_1^k, \dots, u_n^k]^T$ and $\mathbf{c} = [c_0, c_1, \dots, c_n]^T$. Then the finite difference scheme amounts to

$$\mathbf{u}^{k+1} = F(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1}), \quad \text{with } \mathbf{u}^{-1} = \mathbf{u}^0 = \mathbf{0}. \quad (4)$$

The forward map from \mathbf{c} to $A[\mathbf{c}]$ is given by, letting $\mathbf{e}_1 = [1, 0, \dots, 0]^T$

$$A[\mathbf{c}]^k = \mathbf{e}_1^T \mathbf{u}^k, \quad \text{for } k = 1 : m. \quad (5)$$

The inverse problem is to solve for \mathbf{c} in

$$\min_{\mathbf{c}} \|A[\mathbf{c}] - \mathbf{g}\|^2 \quad (6)$$

where \mathbf{g} is a data vector corresponding to a measurement.

2.2 Two-dimensional problem

The two-dimensional problem is motivated by a problem in acoustic imaging of human tissues. The geometry of the problem has been described in the previous section, and elsewhere [11]. Here we give a mathematical model of the physics.

Because any computational domain is necessarily finite, we will consider a box $\Omega := [-a, a] \times [-a, a]$. Letting $u(x, y, t)$ represent the excess pressure, a model for acoustics is given by the partial differential equation

$$u_{tt} = c(x, y)^2 \Delta u + f(x, y, t) \quad \text{in } \Omega, t > 0. \quad (7a)$$

Here $c(x, y)$ represents the unknown soundspeed distribution, while $f(x, y, t)$ is a known acoustic source. Initially, the system is at rest, hence

$$u(x, y, 0) = u_t(x, y, 0) = 0. \quad (7b)$$

We need to simulate an unbounded medium with a bounded domain. In the unbounded medium, we would have a boundary condition for $|x^2 + y^2|$ large that amounts to saying that waves which are sufficiently far away from the origin and travelling outward will be radiated to infinity. To simulate the unbounded medium, we assume that c is constant near the boundary of Ω and apply the Engquist-Majda boundary conditions [6] along the flat parts of $\partial\Omega$ (and a modification of Engquist-Majda at the corners of $\partial\Omega$). For points away from the corners, the boundary condition is given by

$$cu_{xt} - u_{tt} + \frac{c^2}{2}u_{yy} = 0 \quad \text{for } \{x = \pm a; |y| < a\}, \quad (7c)$$

$$cu_{yt} - u_{tt} + \frac{c^2}{2}u_{xx} = 0 \quad \text{for } \{y = \pm a; |x| < a\}. \quad (7d)$$

Let R represent the collection of coordinate points where receivers have been placed to record u . Thus,

$$R = \{(x_r, y_r) = (\rho \cos \theta_r, \rho \sin \theta_r), \quad r = 1 : p\}$$

for some $\rho > 0$. The forward map is given by

$$A[c; f]_r := u(x_r, y_r, t)$$

The source term $f(x, y, t)$ is assumed to be null for $t = 0$. We will view the forward map $A[\cdot]$ as dependent on c and parameterized by f .

The nonlinear least-squares formulation is given by

$$\min_{\mathbf{c}} \sum_l \sum_{r=1}^p \int_0^T |A[\mathbf{c}; f_l]_r(t) - g_{rl}(t)|^2$$

where $g_{rl}(t)$ is the measured response at location (x_r, y_r) for the source $f_l(x, y, t)$.

Discretization of (7) is quite straight-forward. The only tricky part comes in discretizing the Enquist-Majda boundary condition. Letting $(x_i, y_j) = (i\Delta, j\Delta)$, $-n \leq i \leq n$ and $-n \leq j \leq n$, we discretized the domain Ω by a regular mesh of size $\Delta = a/n$. Time is discretized as in the 1-D case: $t_k = k\Delta t$ for $k = 0 : m$.

Let the $(2n + 1)^2$ vector \mathbf{u}^k represent the value for $u(x, y, t)$ at the node points at time t_k . The finite difference scheme can be written in short hand as

$$\mathbf{u}^{k+1} = F(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1}), \quad \text{for } k = 1 : m,$$

with $\mathbf{u}^{-1} = \mathbf{u}^0 = \mathbf{0}$. The discrete forward map evaluates \mathbf{u} at each receiver, thus

$$A[\mathbf{c}; f]^k = T \mathbf{u}^k,$$

where T is a matrix of size p -by- $(2n + 1)^2$ and its function is to ‘grab’ values of \mathbf{u} at time step k at the receivers. In place of the integration in the nonlinear leastsquares, we have

$$\min_{\mathbf{c}} \sum_l \sum_{r=1}^p \sum_{k=1}^m |A[\mathbf{c}; f_l]_r^k - g_{rl}^k|^2 \quad (8)$$

Here g_{rl}^k is the measured response at receiver r at time step k when the excitation is f_l .

3 The Extended Jacobian Framework

We restrict our discussion to the 1-D problem for clarity of presentation. The prescription for computing Jacobian vector and adjoint vector products for the more complex 2-D problem follows the same lines as for the 1-D problem. An algorithm for the forward map for the 1-D case is

$\begin{aligned} &\mathbf{u}^{-1} = \mathbf{u}^0 = \mathbf{0} \\ &\text{for } k = 0 : m - 1 \\ &\quad \mathbf{u}^{k+1} = F(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1}) \\ &\quad h_{k+1} = \mathbf{e}_1^T \mathbf{u}^{k+1} \\ &\text{end} \end{aligned} \quad (9)$
--

We use the notation $h_k = A[\mathbf{c}]_k$. Thus the function in question is the mapping from \mathbf{c} to $\mathbf{h} = (h_1, h_2, \dots, h_m)^T$.

We can give an alternate description of this mapping by enumerating through the loop

$\begin{aligned} \mathbf{u}^1 &= F(\mathbf{c}, \mathbf{u}^0, \mathbf{u}^{-1}) \\ \mathbf{u}^2 &= F(\mathbf{c}, \mathbf{u}^1, \mathbf{u}^0) \\ &\vdots \\ \mathbf{u}^m &= F(\mathbf{c}, \mathbf{u}^{m-1}, \mathbf{u}^{m-2}) \\ \mathbf{h} &= \mathbf{e}_1 \mathbf{e}_1^T \mathbf{u}^1 + \mathbf{e}_2 \mathbf{e}_1^T \mathbf{u}^2 + \dots + \mathbf{e}_m \mathbf{e}_1^T \mathbf{u}^m \end{aligned}$

We call this the extended function. The extended function allows for an easy way to compute the Jacobian and its transpose. Formally, the directional derivative of \mathbf{h} in the direction $d\mathbf{c}$, i.e., the Jacobian-vector product, is given by the following calculation

$$\begin{array}{l}
 d\mathbf{u}^{-1} = d\mathbf{u}^0 = 0 \\
 \text{for } k = 0 : m - 1 \\
 \quad d\mathbf{u}^{k+1} = F_1(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1})d\mathbf{c} + F_2(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1})d\mathbf{u}^k + F_3(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1})d\mathbf{u}^{k-1} \\
 \quad dh_{k+1} = \mathbf{e}_1^T d\mathbf{u}^{k+1} \\
 \text{end}
 \end{array} \tag{10}$$

The matrices F_1 , F_2 , and F_3 are Jacobians of the function F with respect to the first, second and third variables. Therefore, they are $(n+1)$ -by- $(n+1)$ matrices. In a computer program, we would simply define $F(\mathbf{c}, \cdot, \cdot)$ and use AD to either compute these matrices or produce subprograms that calculates the action of these matrices on given vectors. The desired directional derivative (Jacobian times vector $d\mathbf{c}$) is $d\mathbf{h} = (dh_1, dh_2, \dots, dh_m)^T$.

3.1 Adjoint computation via linear algebra

The above calculation can be defined as a set of matrix equations through the use of the extended Jacobian framework [3, 4] (see also [5]). Let

$$dU = \begin{bmatrix} d\mathbf{u}^1 \\ d\mathbf{u}^2 \\ \vdots \\ d\mathbf{u}^m \end{bmatrix}$$

Define the $m(n+1) \times m(n+1)$ matrix

$$M = \begin{bmatrix} -I & 0 & \ddots & \ddots & 0 & 0 \\ F_2(\mathbf{c}, \mathbf{u}^1, \mathbf{u}^0) & -I & 0 & \ddots & \ddots & 0 \\ F_3(\mathbf{c}, \mathbf{u}^2, \mathbf{u}^1) & F_2(\mathbf{c}, \mathbf{u}^2, \mathbf{u}^1) & -I & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & -I & 0 \\ 0 & \ddots & \ddots & F_3(\mathbf{c}, \mathbf{u}^{m-1}, \mathbf{u}^{m-2}) & F_2(\mathbf{c}, \mathbf{u}^{m-1}, \mathbf{u}^{m-2}) & -I \end{bmatrix}$$

and $m(n+1) \times (n+1)$ matrix

$$B = \begin{bmatrix} F_1(\mathbf{c}, \mathbf{u}^0, \mathbf{u}^{-1}) \\ F_1(\mathbf{c}, \mathbf{u}^1, \mathbf{u}^0) \\ \vdots \\ F_1(\mathbf{c}, \mathbf{u}^{m-1}, \mathbf{u}^{m-2}) \end{bmatrix}$$

and $m \times m(n+1)$ matrix

$$T = [\mathbf{e}_1 \mathbf{e}_1^T \quad \mathbf{e}_2 \mathbf{e}_1^T \quad \cdots \quad \mathbf{e}_m \mathbf{e}_1^T]$$

Then

$$-M dU = B d\mathbf{c}, \quad \text{and} \quad d\mathbf{h} = T dU.$$

From the above, we can solve for dU and write

$$d\mathbf{h} = -TM^{-1}Bd\mathbf{c}, \quad (11)$$

which encapsulates the Jacobian-vector product calculation in (10).

To obtain a formula for the adjoint-vector product calculation, we start by formally taking the adjoint of (11). Let \mathbf{p} be the result of multiplying vector \mathbf{q} by the adjoint of the Jacobian. Then from (11)

$$\mathbf{p} = -B^T M^{-T} T^T \mathbf{q}. \quad (12)$$

We do not advocate computing the matrices B , M and T , but rather, use the formalism to generate an efficient algorithm for finding \mathbf{p} given \mathbf{q} .

Let \mathbf{q} be an m -vector. Then $Q = T^T \mathbf{q}$ is an $m(n+1)$ -vector. From (12) if we let $Y = -M^{-T} T^T \mathbf{q}$, then

$$-M^T Y = Q \quad \text{and} \quad \mathbf{p} = B^T Y. \quad (13)$$

By exploiting the structures of M and B , we can come up with an efficient algorithm to find \mathbf{p} for a given \mathbf{q} . Because of the lower-triangular structure of M , we never need to invert any matrices. The algorithm starts by chopping up Q into m separate pieces

$$Q = \begin{bmatrix} \mathbf{q}^1 \\ \mathbf{q}^2 \\ \vdots \\ \mathbf{q}^m \end{bmatrix}$$

and similarly for Y . Then, according to (13), we can calculate \mathbf{p} by

$\begin{aligned} \mathbf{y}^m &= \mathbf{q}^m; \quad \mathbf{p} = F_1(\mathbf{c}, \mathbf{u}^{m-1}, \mathbf{u}^{m-2})^T \mathbf{y}^m \\ \mathbf{y}^{m-1} &= \mathbf{q}^{m-1} + F_2(\mathbf{c}, \mathbf{u}^{m-1}, \mathbf{u}^{m-2})^T \mathbf{y}^m \\ \mathbf{p} &= \mathbf{p} + F_1(\mathbf{c}, \mathbf{u}^{m-2}, \mathbf{u}^{m-3})^T \mathbf{y}^{m-1} \\ \text{for } k &= m-2 : -1 : 1 \\ \quad \mathbf{y}^k &= \mathbf{q}^k + F_2(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1})^T \mathbf{y}^{k+1} + F_3(\mathbf{c}, \mathbf{u}^{k+1}, \mathbf{u}^k)^T \mathbf{y}^{k+2} \\ \quad \mathbf{p} &= \mathbf{p} + F_1(\mathbf{c}, \mathbf{u}^{k-1}, \mathbf{u}^{k-2})^T \mathbf{y}^k \\ \text{end} \end{aligned}$	(14)
---	------

Note that we have adjoints/transpose of $F_1(\mathbf{c}, \cdot, \cdot)$, $F_2(\mathbf{c}, \cdot, \cdot)$, and $F_3(\mathbf{c}, \cdot, \cdot)$. These adjoints can be computed explicitly if we have matrices F_1 , F_2 , and F_3 or we can resort to AD to produce subprograms that computes their action on given vectors.

Note that in the algorithm (14), we need to have available values of the fields \mathbf{u}^k for all indices k . Depending on the size of the problem, it may be more efficient to store only values of \mathbf{u}^k for some indices $k \in K$, and use (9) to generate the field for other indices $k \notin K$. An efficient method to do this is discussed in [8].

3.2 Adjoint computation via adjoint variables

We give an alternate derivation of the algorithm in (14) which is based on using adjoint variables. Consider a simple calculation involving the following 3 steps. The input variable is c and the output variable is u_3 ; u_0 is a parameter. The steps are:

$$\begin{aligned} u_1 &= f(c, u_0, 0) \\ u_2 &= f(c, u_1, u_0) \\ u_3 &= f(c, u_2, u_1) \end{aligned}$$

We can view this as an extended function. The Jacobian calculation is

$$\begin{aligned} du_1 &= f_1(c, u_0, 0)dc \\ du_2 &= f_1(c, u_1, u_0)dc + f_2(c, u_1, u_0)du_1 \\ du_3 &= f_1(c, u_2, u_1)dc + f_2(c, u_2, u_1)du_2 + f_3(c, u_2, u_1)du_1 \end{aligned}$$

Therefore, the Jacobian (in this case, derivative) can be indentified as J from the output $du_3 = Jdc$. This is a forward mode computation.

Let the adjoint variables be p and v_3 so that we formally have $p = J^T v_3$. If we view du_1 and du_2 as intermediate variables, then we can associate to them adjoint variables v_1 and v_2 . From the third equation in the adjoint calculation, we can formally write

$$\begin{bmatrix} p \\ v_2 \\ v_1 \end{bmatrix} = \begin{bmatrix} f_1(c, u_2, u_1) \\ f_2(c, u_2, u_1) \\ f_3(c, u_2, u_1) \end{bmatrix} v_3$$

and from the second,

$$\begin{bmatrix} p \\ v_1 \end{bmatrix} = \begin{bmatrix} f_1(c, u_1, u_0) \\ f_2(c, u_1, u_0) \end{bmatrix} v_2$$

and from the first,

$$p = f_1(c, u_0, 0)v_1.$$

The contributions to each of the adjoint variables are summed over each operation, hence

$$\begin{aligned} v_2 &= f_2(c, u_2, u_1)v_3 \\ v_1 &= f_3(c, u_2, u_1)v_3 + f_2(c, u_1, u_0)v_2 \\ p &= f_1(c, u_2, u_1)v_3 + f_1(c, u_1, u_0)v_2 + f_1(c, u_0, 0)v_1 \end{aligned}$$

This is the reverse computation [9].

We can generalize this concept to the 1-D wave propagation problem. In (10), we identify adjoint variables \mathbf{p} with $d\mathbf{c}$ for the input, and \mathbf{q} with $d\mathbf{h}$ for the output. To the intermediate variables $d\mathbf{u}^k$, we associate adjoint variables \mathbf{v}^k . Performing the reverse mode calculation, we must start at index $k = m - 1$. Let q_k , for $k = 1 : m$ be the elements of \mathbf{q} .

The adjoint-times-vector algorithm is

<pre> set $\mathbf{v}^m = 0$ for $k = m - 1 : 0$ $\mathbf{v}^{k+1} = \mathbf{v}^{k+1} + \mathbf{e}_1 q_{k+1}$ $\mathbf{v}^k = \mathbf{v}^k + F_2(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1})^T \mathbf{v}^{k+1}$ $\mathbf{v}^{k-1} = \mathbf{v}^{k-1} + F_3(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1})^T \mathbf{v}^{k+1}$ $\mathbf{p} = \mathbf{p} + F_1(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1})^T \mathbf{v}^{k+1}$ end </pre>	(15)
---	------

At the end of the calculation, we can identify $\mathbf{p} = J^T \mathbf{q}$.

3.3 Remark

The foregoing discussion, while limited to the 1-D problem, can be adapted to solve the more complicated 2-D problem. What we wish to emphasize here is the conciseness of the extended Jacobian framework, and how to exploit the underlying problem structure. The algorithms in (10), (14) and (15) can be viewed as code templates for Jacobian and adjoint vector product calculations. AD is deployed in computing the Jacobian and adjoint of the subproblem described by the time stepping process (4).

We recall that the adjoint (reverse product) $F_1(\mathbf{c}, \cdot, \cdot)^T y$, etc., can be computed using the adjoint (reverse) mode of an AD tool. For large problems like this, computing the adjoint product of the timestep routine (4) can be very expensive, since the size of \mathbf{c} , \mathbf{u}^k , and \mathbf{u}^{k-1} can be large. An AD tool would by default assume that every element of $F_1(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1})$ depend on every element of \mathbf{c} , \mathbf{u}^k , and \mathbf{u}^{k-1} . This assumption on dependence generates a ‘table’ which is used in computing intermediate values in the reverse product mode. For example, ADOL-C [10] implements this lookup by creating a tape, which it will write on the disk if the problem size is large. When it does this, it becomes unacceptably inefficient.

This concern brings us to the main idea of this paper, i.e., that of AD applied to the finite difference stencil. Our approach is to use AD on the smallest component of the calculation – a kind of ‘microscopic’ structure exploitation. We discuss how this is done in the next section.

In principle, what we are exploiting is specific sparsity structure that is inherent in the finite difference scheme. A general approach for exploiting sparsity in AD is described in [2].

4 Exploiting the stencil structure

The finite difference method that we used in the 1-D can be written as indicated in (4) which we rewrite here

$$\mathbf{u}^{k+1} = F(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1}), \quad \text{with } \mathbf{u}^{-1} = \mathbf{u}^0 = \mathbf{0}.$$

This shorthand notation does not reveal the stencil structure given by the explicit formulas in (3). For the j th component of \mathbf{u}^{k+1} , j not equal to 0 or n , from (3a) we can write

$$u_j^{k+1} = f(c_j, u_{j-1}^k, u_j^k, u_{j-1}^k, u_j^{k-1}). \quad (16)$$

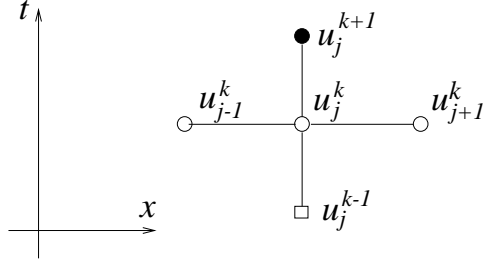


Figure 2: The stencil for the 1-D problem for $j \neq 0, n$. Boundary nodes are slightly different and require separate treatment.

The above expression spells out clearly that the dependence of \mathbf{u}^{k+1} on \mathbf{c} , \mathbf{u}^k , and \mathbf{u}^{k-1} , is very *sparse*. This is best visualized by studying Figure 2. Thus, we need only to deal with f which is a function of only 5 variables. From (3c)-(3d), we have two more such functions which but they depend only on 4 and 3 variables respectively, and are given by

$$\begin{aligned} u_0^{k+1} &= f_L(c_0, u_0^k, u_1^k, u_0^{k-1}), \\ u_n^{k+1} &= f_R(c_n, u_{n-1}^k, u_n^k). \end{aligned}$$

The function $F(\cdot, \cdot, \cdot)$, representing a time-step, is now replaced with the pseudo-code

```

function  $\mathbf{u}^{k+1} = F(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1})$ 
 $u_0^{k+1} = f_L(c_0, u_0^k, u_1^k, u_0^{k-1})$ 
 $u_n^{k+1} = f_R(c_n, u_{n-1}^k, u_n^k)$ 
for  $j = 1 : n - 1$ 
     $u_j^{k+1} = f(c_j, u_{j-1}^k, u_j^k, u_{j-1}^k, u_j^{k-1})$ 
end

```

(17)

It is to these ‘small’ functions of a few variables that we want to apply automatic differentiation. The benefits are that we will have efficient codes which explicitly exploit the structure of the problem. The cost is that the derivative and adjoint codes will be slightly more complicated to assemble. We discuss this next.

4.1 Sparse Jacobian

Due to the sparsity of afforded by the stencil structure, it is feasible to calculate the full Jacobian (rather than the Jacobian vector product). To see this we introduce projection matrices. Let \mathbf{e}_j be the j th unit vector (We will let j run from 0 to n for convenience). Then (16) can be rewritten in terms of vectors \mathbf{c} , \mathbf{u}^k , and \mathbf{u}^{k-1} as

$$u_j^{k+1} = f(\mathbf{e}_j^T \mathbf{c}, \mathbf{e}_{j-1}^T \mathbf{u}^k, \mathbf{e}_j^T \mathbf{u}^k, \mathbf{e}_{j+1}^T \mathbf{u}^k, \mathbf{e}_j^T \mathbf{u}^{k-1}). \quad (18)$$

In computing the Jacobian, we needed the derivatives of $F(\cdot, \cdot, \cdot)$ with respect to the 3 variables. We next derive procedures to do this using the stencils.

The components of $F_1(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1})$ are

$$F_1 = \begin{bmatrix} (\nabla_{\mathbf{c}} u_0^{k+1})^T \\ (\nabla_{\mathbf{c}} u_1^{k+1})^T \\ \vdots \\ (\nabla_{\mathbf{c}} u_{n-1}^{k+1})^T \\ (\nabla_{\mathbf{c}} u_n^{k+1})^T \end{bmatrix}.$$

The gradients are easily obtained by differentiating (18) with respect to \mathbf{c} . We obtain, for $j \neq 0, n$,

$$\nabla_{\mathbf{c}} u_j^{k+1} = \partial_1 f(\mathbf{e}_j^T \mathbf{c}, \mathbf{e}_{j-1}^T \mathbf{u}^k, \mathbf{e}_j^T \mathbf{u}^k, \mathbf{e}_{j+1}^T \mathbf{u}^k, \mathbf{e}_j^T \mathbf{u}^{k-1}) \mathbf{e}_j,$$

which is an $(n+1)$ -vector with a single nonzero entry at j . Thus, it can be seen that $F_1(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1})$ is a *diagonal* matrix. This property is not apparent to state-of-the-arts automatic differentiation programs.

The Jacobian $F_3(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1})$ will also be diagonal for the same reason, and be computed by applying AD to (16). The Jacobian $F_2(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1})$ will be slightly more complicated. The components of the Jacobian are similar to those of $F_1(\cdot, \cdot, \cdot)$ except that the gradient will be with respect to \mathbf{u}^k . Directly differentiating (16) with respect to \mathbf{u}^k yields

$$\begin{aligned} \nabla_{\mathbf{u}^k} u_j^{k+1} &= \partial_2 f(\mathbf{e}_j^T \mathbf{c}, \mathbf{e}_{j-1}^T \mathbf{u}^k, \mathbf{e}_j^T \mathbf{u}^k, \mathbf{e}_{j+1}^T \mathbf{u}^k, \mathbf{e}_j^T \mathbf{u}^{k-1}) \mathbf{e}_{j-1} \\ &+ \partial_3 f(\mathbf{e}_j^T \mathbf{c}, \mathbf{e}_{j-1}^T \mathbf{u}^k, \mathbf{e}_j^T \mathbf{u}^k, \mathbf{e}_{j+1}^T \mathbf{u}^k, \mathbf{e}_j^T \mathbf{u}^{k-1}) \mathbf{e}_j \\ &+ \partial_4 f(\mathbf{e}_j^T \mathbf{c}, \mathbf{e}_{j-1}^T \mathbf{u}^k, \mathbf{e}_j^T \mathbf{u}^k, \mathbf{e}_{j+1}^T \mathbf{u}^k, \mathbf{e}_j^T \mathbf{u}^{k-1}) \mathbf{e}_{j+1}. \end{aligned}$$

Thus, the matrix $F_2(\mathbf{c}, \mathbf{u}^k, \mathbf{u}^{k-1})$ is a *tridiagonal* matrix.

We can summarize the steps in a MATLAB pseudo-code¹

```

F1 = [ $\partial_1 f_L \mathbf{e}_0^T$ ];
F2 = [ $(\partial_2 f_L \mathbf{e}_0^T + \partial_3 f_L \mathbf{e}_1^T)$ ];
F3 = [ $\partial_4 f_L \mathbf{e}_0^T$ ];
for j = 1 : n - 1
    F1 = [ $F_1; \partial_1 f \mathbf{e}_j^T$ ];
    F2 = [ $F_2; (\partial_2 f \mathbf{e}_{j-1}^T + \partial_3 f \mathbf{e}_j^T + \partial_4 f \mathbf{e}_{j+1}^T)$ ];
    F3 = [ $F_3; \partial_5 f \mathbf{e}_j^T$ ];
end
F1 = [ $F_1; \partial_1 f_R \mathbf{e}_n^T$ ];
F2 = [ $F_2; \partial_2 f_R \mathbf{e}_{n-1}^T + \partial_3 f_R \mathbf{e}_n^T$ ];
F3 = [ $F_3; \mathbf{zeros}(1, n + 1)$ ];
```

Once the matrices F_1 , F_2 and F_3 are obtained, we can use the code in (10) to compute the forward derivatives and the code in (14) to compute the adjoint. The codes for the partial derivatives of f , f_L and f_R are easily obtained using AD. These codes are expected

¹One would use `sparse` utilities to implement this in MATLAB.

We would have similar formulas for $j = 0$ and $j = n$ with the difference that the vector of independent variables would be 4 and 3 dimensional, respectively. We can therefore assemble the du_j^{k+1} within an outer loop which corresponds to the time steps. The pseudo-code would take the form

<pre> du⁻¹ = du⁰ = 0 for $k = 0 : m - 1$ $X = [c_0, u_0^k, u_1^k, u_0^{k-1}]^T$; $dX = [dc_0, du_0^k, du_1^k, du_0^{k-1}]^T$ $du_0^{k+1} = \nabla f_L(X) \cdot dX$ $X = [c_n, u_{n-1}^k, u_n^k]^T$; $dX = [dc_n, du_{n-1}^k, du_n^k]^T$ $du_n^{k+1} = \nabla f_R(X) \cdot dX$ for $j = 1 : n - 1$ $X = [c_j, u_{j-1}^k, u_j^k, u_{j+1}^k, u_j^{k-1}]^T$; $dX = [dc_j, du_{j-1}^k, du_j^k, du_{j+1}^k, du_j^{k-1}]^T$ $du_j^{k+1} = \nabla f(X) \cdot dX$ end $dh_{k+1} = du_0^{k+1}$ end </pre>	(19)
---	------

The adjoint codes generated by AD on the stencil formula (16) would compute the following. Given a scalar v and a vector X , the adjoint code calculates a 5-vector

$$v \nabla f(X).$$

We have similar procedures for $f_L(\cdot)$ and $f_R(\cdot)$. In reverse mode, we want to perform a calculation similar to (15). We start with a vector $\mathbf{q} = [q_1, q_2, \dots, q_m]^T$, and we wish to compute $\mathbf{p} = J^T \mathbf{q}$. The pseudo-code for this is as follows

<pre> set $\mathbf{v}^m = 0$ for $k = m - 1 : 0$ $v_0^{k+1} = v_0^{k+1} + q_{k+1}$ $X = [c_0, u_0^k, u_1^k, u_0^{k-1}]^T$ $Y = [p_0, v_0^k, v_1^k, v_0^{k-1}]^T$ $Y = Y + dv_0^{k+1} \nabla f_L(X)$ for $j = 1 : n - 1$ $X = [c_j, u_{j-1}^k, u_j^k, u_{j+1}^k, u_j^{k-1}]^T$ $Y = [p_j, v_{j-1}^k, v_j^k, v_{j+1}^k, v_j^{k-1}]^T$ $Y = Y + dv_j^{k+1} \nabla f(X)$ end $X = [c_n, u_{n-1}^k, u_n^k]^T$ $Y = [p_n, v_{n-1}^k, v_n^k]^T$ $Y = Y + dv_n^{k+1} \nabla f_R(X)$ end </pre>	(20)
--	------

Algorithm (19) for the forward product calculation and algorithm (20) for the reverse product calculation will be extremely efficient because the codes produced by AD for calculating

the derivative of f and its adjoint will be nearly as short and simple as the function calculation. The number of independent variables is small, and there are no loops as can be seen in (3a).

In 2-D, the stencil is a bit more complex as already pointed out, but the general principle described here applies. Indeed, we have coded a version of algorithms (19) and (20) for the 2-D problem. We discuss the results of our numerical calculations next.

5 Numerical results

We present some results from our numerical computations. In both examples, we use TAMC [7] to obtain derivative and adjoint codes from fortran sources. All the fortran codes were ‘wrapped’ as MATLAB mex-files and used in conjunction with MATLAB codes.

Our goal in this paper is to demonstrate the use of extended Jacobian framework together with exploitation of stencil in Jacobian and adjoint calculations. In a subsequent work, we apply our approach to solve a 2-D inverse problem arising in acoustic imaging.

5.1 1-D problem

In our example, we choose $\Delta x = 1$ and $\Delta t = 0.8$. The domain is of length $L = (N - 1)\Delta x$. We will use several N in our calculations. The initial boundary value problem for the 1-D wave equation is discretized according to (3). The number of time steps is m , which will be varied as well. For excitation $f(t)$, we choose the derivative of the Gaussian. The graph of f is shown in Figure 5.1a. We take two sound speeds $c_1(x)$ and $c_2(x)$, shown in Figure 5.1b when $n = 100$. The resulting boundary data is $h(t) = u(0, t)$. When the medium is $c_1(x)$ the boundary data is $h_1(t)$, and $h_2(t)$ when the medium is $c_2(x)$. Let $q = h_1 - h_2$; the graph of $q(t)$ is displayed in Figure 5.1c for $m = 200$.

We will first compute the Jacobian at $c_1(x)$ times the difference $c_2(x) - c_1(x)$. The resulting output vector should be very close to $q(t)$. A comparison of $p(t)$ with $J(\mathbf{c}_1)(\mathbf{c}_2 - \mathbf{c}_1)$ is shown in 5.1c. Next we calculate the adjoint times $q(t)$; i.e.,

$$\mathbf{p} = J(\mathbf{c}_1)^T (\mathbf{h}_2 - \mathbf{h}_1).$$

The output of this calculation will be the steepest descent direction corresponding to the nonlinear leastsquares functional in (6). This direction would be similar to $c_2(x) - c_1(x)$. The graph of $p(t)$ is shown in Figure 5.1d. One can see that the 2 big signals, which are scaled versions of f , are reproduced near the places where $c_2(x) - c_1(x)$ take jumps. Unfortunately, the similarity ends there, the result shows that the inverse problem is not very well posed. However, we did check that the Jacobian and the adjoint are correctly computed by evaluating

$$\mathbf{q}^T J d\mathbf{c} \quad \text{and} \quad d\mathbf{c}^T J^T \mathbf{q} \tag{21}$$

and comparing their values for any choice of \mathbf{c} , \mathbf{q} , and $d\mathbf{c}$. The agreement is usually 14 digits. We show a typical run in Figure 5.1.

Computation time is linear in the number of x nodes for a fixed number of t nodes. There is no difficulty with memory as the stencil codes are very simple with small number of independent variables.

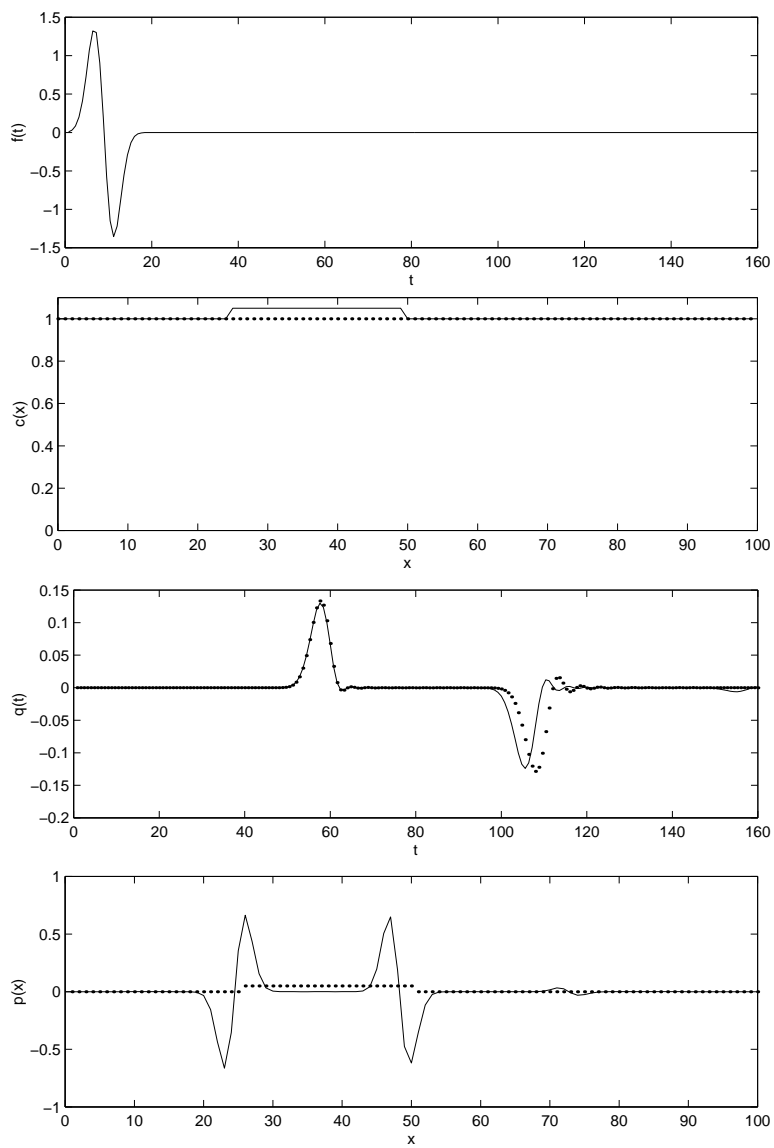


Figure 4: (a) The excitation used in the examples. (b) The two sound speed profiles, $c_1(x)$ in dots, $c_2(x)$ in solid. (c) The graphs of $(\mathbf{h}_2 - \mathbf{h}_1)$ and $J(\mathbf{c}_1)(\mathbf{c}_2 - \mathbf{c}_1)$. (d) The graph of $\mathbf{p} = J(\mathbf{c}_1)^T(\mathbf{h}_2(t) - \mathbf{h}_1)$, shown for comparison is the graph of $\mathbf{c}_2 - \mathbf{c}_1$.

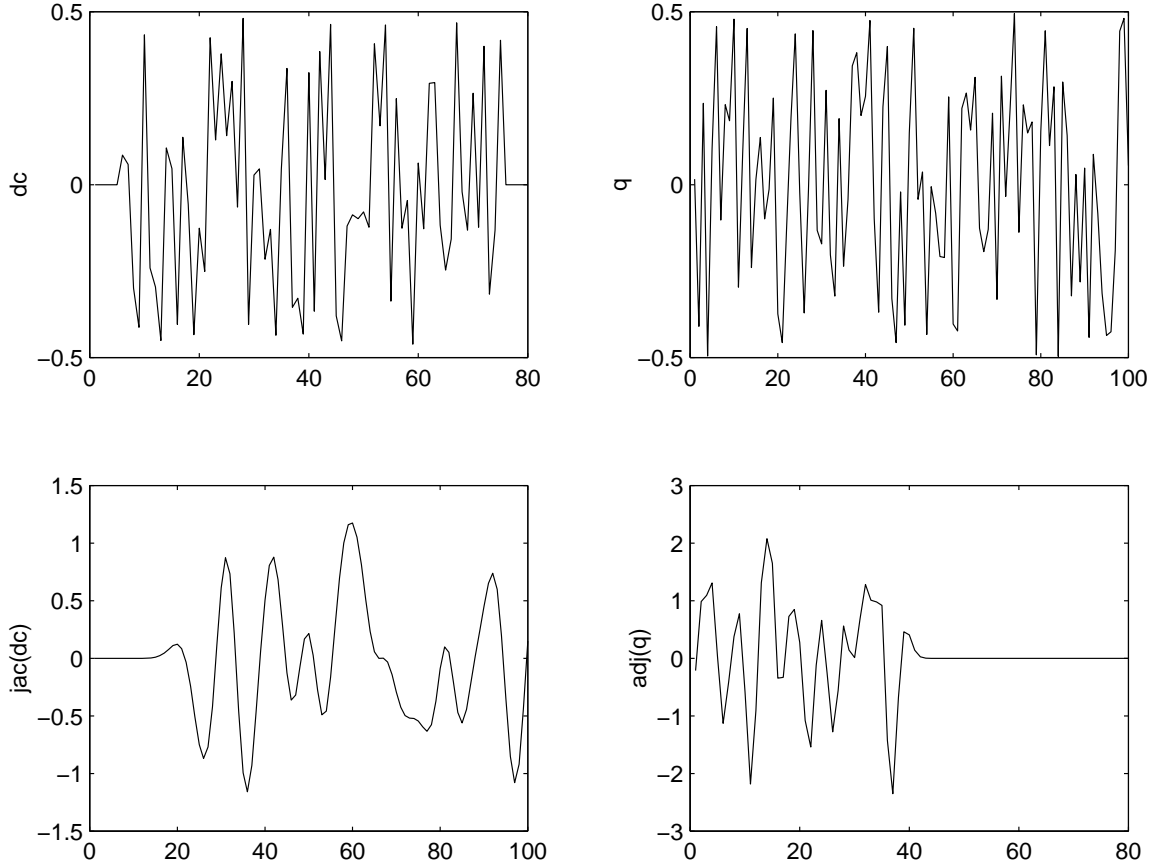


Figure 5: A test of the correctness of Jacobian and adjoint calculations. In this example, $N = 80$ and $m = 100$. We choose at random 2 vectors \mathbf{dc} and \mathbf{q} displayed on the first row. On the second row, we show $J \mathbf{dc}$ and $J^T \mathbf{q}$. The inner products $\mathbf{q}^T J \mathbf{dc}$ and $\mathbf{dc}^T J^T \mathbf{q}$ are evaluated. They agree to 14 digits.

5.2 2-D problem

In the 2-D problem, we set up a grid of 161-by-161 node points. The computational domain is $[-80, 80] \times [-80, 80]$, thus $\Delta = 1$. For interior nodes, we use a second order accurate discretization of the wave equation (7a). On the boundary nodes, we use a second order discretization of the Enquist-Majda boundary condition (7c-d). The corner nodes, and the 2 nodes adjacent to the corner on the boundary, require special stencils. The stencils are obtained by requiring that the discrete wave equation be satisfied at the node while at the same time satisfying the discrete version of the absorbing boundary condition.

For excitation, we choose a point source. To model this, if the source is at node (i_s, j_s) , i.e., located at position $(i_s \Delta, j_s \Delta)$, we assume that $f(x, y, t)$ is

$$f(x, y, t) = \begin{cases} \phi(t) & \text{at } (i_s \Delta, j_s \Delta) \\ 0 & \text{otherwise} \end{cases} .$$

The time-dependent function $\phi(t)$ is chosen to be a Gaussian, and will be sampled at the time increments $\Delta t = 0.55$, which is the time step chosen for the finite difference scheme.

Data will be collected at 64 stations located at node points. These points are nodes that lie close to a set of points distributed evenly at 64 places on the circumference of a circle of radius 72. We will take 381 time steps. A window of size $[-70, 70] \times [-70, 70]$ represents where $c(x, y)$ is allowed to vary. Thus the mapping from sound speed c to data at the receiver is $\mathbb{R}^{141 \times 141}$ to $\mathbb{R}^{64 \times 381}$.

In Figure 5.2a we display the sound speed distribution in the domain. The receivers are marked with circles; receiver 1 is at 0° from the positive x -axis. The source is located by a \star . Next, in Figure 5.2b, we display the receiver data when the medium has the two cylinders shown. The difference between the previous data and those when the domain is a homogeneous is shown in Figure 5.2c. In Figure 5.2d, we show the result of applying the adjoint on the difference data in Figure 5.2c. This process is often referred to as back-propagation, and corresponds to the steepest descent direction for the nonlinear least squares functional in (8). The resulting vector should resemble the image of the two cylinders. Indeed this is the case if one compares 5.2d and 5.2e, the latter displayed for comparison.

In numerous experiments with random vectors, we were able to get the inner products similar to (21) to agree 14 digits. The adjoint calculations take approximately 115 seconds on a 4-processor SGI Challenge L.

6 Conclusions

We have described an inverse problem arising in wave propagation, and how the need arises for efficient computation of Jacobian and adjoint products when the problem is posed as a least-squares problem. In this work, we describe the Extended Jacobian framework which gives a high level description of Jacobian and adjoint vector product calculation. The framework is particularly appropriate for functions whose evaluation involve some type of time stepping, such as those that arise in discretizing the wave equation.

We show further that the stencil structure of the finite difference scheme that provides the underlying function evaluation can be exploited. Automatic differentiation is applied at the stencil level, and the resulting subprograms fit nicely within the Extended Jacobian framework. The framework provides a guide for building highly efficient codes for Jacobian and adjoint vector product evaluations. The one drawback of the approach is that we have given up some ‘automation’ for efficiency. A small amount of handcoding is required to assemble the programs. Nevertheless, our approach provides a way to overcome memory problems associated with present AD technology.

Overall, the idea of exposing the stencil structure is very promising and can lead to an order of magnitude improvement in the adjoint code, as our numerical results show.

References

- [1] G. Chavent, F. Clement, and S. Gomez, Waveform inversion by MBTT formulation, in Mathematical and Numerical Aspects of Wave Propagation, Cohen et al eds., SIAM,

- Philadelphia, 1995, pp. 713-722.
- [2] T. Coleman and A. Verma, The efficient computation of sparse Jacobian matrices using automatic differentiation, *SIAM Journal on Scientific Computing* (1998), **19**, pp. 1210-1233.
 - [3] T. Coleman and A. Verma, Structure and efficient Hessian calculation, in *Advances in Nonlinear Programming*, Yuan, ed., Kluwer, Boston, 1996.
 - [4] T. Coleman and A. Verma, Structure and efficient Jacobian calculation, in *Computational Differentiation: Techniques, Applications, and Tools*, Berz, et al eds., SIAM, Philadelphia, 1996, pp. 149-159.
 - [5] T. Coleman, F. Santosa, and A. Verma, Semi-automatic differentiation, *Proceedings of Optimal Design and Control Workshop*, VPI, 1997.
 - [6] B. Enquist and A. Majda, Absorbing boundary conditions for the numerical simulation of waves, *Math. Comp.* (1977), **31**, pp. 629-651.
 - [7] R. Giering, *Tangent Linear and Adjoint Model Compiler, User Manual, TAMC Version 4.7*, 1997.
 - [8] A. Griewank, Achieving logarithmic growth of temporal and spatial complexity in reverse automatic differentiation, *Optimization Methods and Software* (1992), **1**, pp. 35-54.
 - [9] A. Griewank, Some bounds on the complexity of gradients, Jacobians, and Hessians, in *Complexity in Nonlinear Optimization*, Pardalos, ed., World Scientific Publishers, 1993.
 - [10] A. Griewank, D. Juedes, and J. Utke, ADOL-C, a package for the automatic differentiation of algorithms written in C/C++, *ACM Trans. Math. Software* (1996), **22**, 131-167.
 - [11] T. Mast, A. Nachman, and R. Waag, Focussing and imaging using eigenfunctions of the scattering operator, *J. Acous. Soc. Am.*, to appear.
 - [12] F. Santosa and W. Symes, *An Analysis of Least-Squares Velocity Inversion*, Society of Exploration Geophysicists, Tulsa, 1989.
 - [13] F. Santosa and W. Symes, Computation of the Hessian for least-squares solutions of inverse problems of reflection seismology, *Inverse Problem*, **4**(1988), pp. 211-233
 - [14] W. Symes, A differential semblance criterion for inversion of multioffset seismic reflection data, *J. Geophys. Res* (1993), **98**, pp. 2061-2073.
 - [15] W. Symes and C. Zhang, A finite difference time stepping class, Rice University TRIP report, 1997.
 - [16] A. Tarantola, *Inverse Problem Theory*, Elsevier, Amsterdam, 1987.

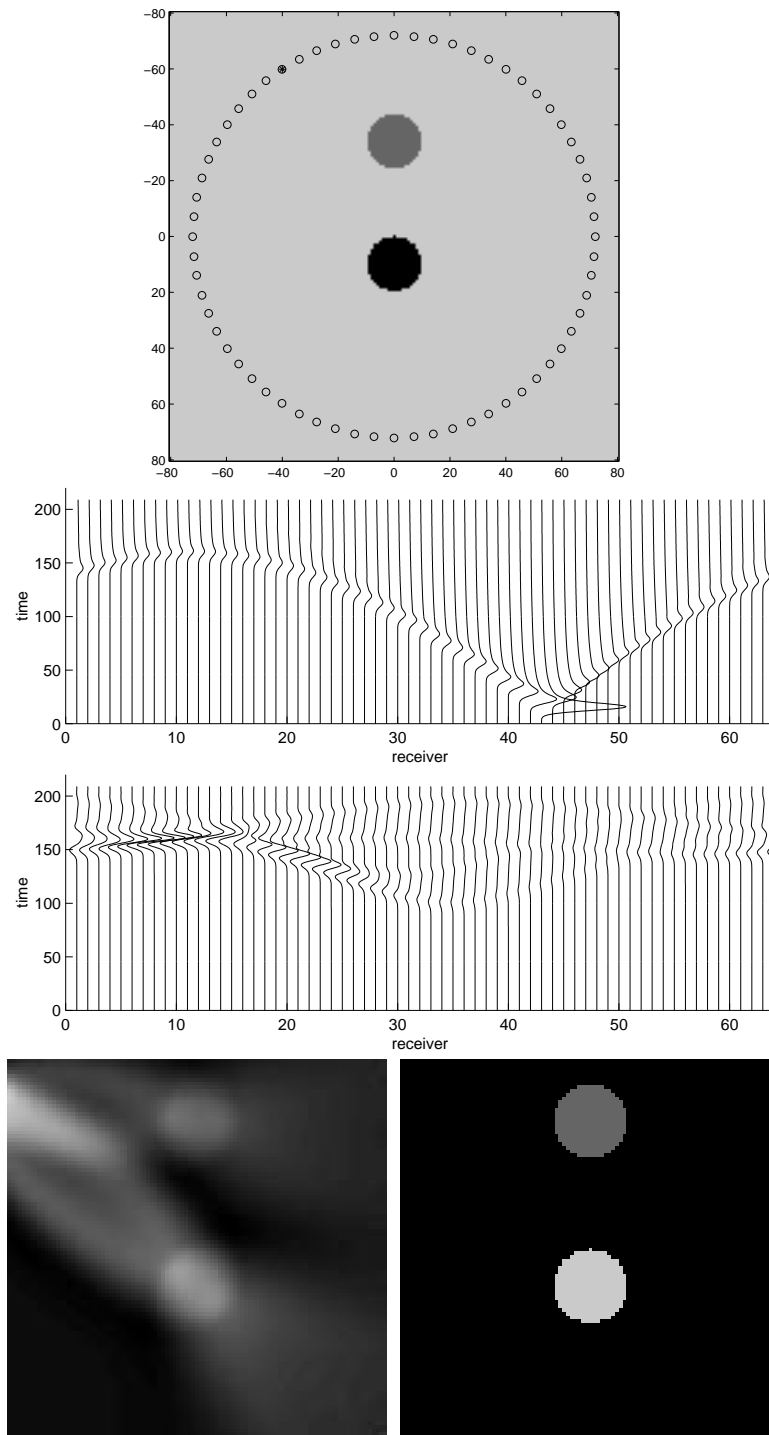


Figure 6: (a) The setup for the numerical experiment. The two cylinders represent sound speed anomalies. The darker cylinder is 2% faster while the lighter cylinder is 1% faster than the background medium. Shown in circles are the receiver locations. Star marks the location of the point source. (b) The receiver data when the two cylinders are present. (c) The difference between (b) and receiver data when the medium is constant. (d) The adjoint applied to (c). (e) The two cylinders plotted on the same scale as in (d) for comparison.