

# **A Finite-Difference Algorithm for Full Waveform Teleseismic Tomography**

S. Roecker, B. Baker, and J. McLaughlin

*Rensselaer Polytechnic Institute, Troy, NY 12180, USA*

## **SUMMARY**

We adapt a 2D spectral domain finite difference waveform tomography algorithm previously used in active source seismological imaging to the case of a plane wave propagating through a 2.5D viscoelastic medium in order to recover P and S wavespeed variations from body waves recorded at teleseismic distances. A transferable efficacy that permits recovery of arbitrarily heterogenous models on moderately sized computers provides the primary motivation for choosing this algorithm. Synthetic waveforms can be generated either by specifying an analytic solution for a background plane wave in a 1D model and solving for the source distribution that would produce it, or by solving for a scattered field excited by a plane wave source and then adding the background wavefield to it. Because the former approach typically involves a concentration of sources at the free surface, the latter tends to be more stable numerically. We adapt a gradient approach to solve the inverse problem to maintain tractability; calculating the gradient does not require much more computational effort than does the forward problem. The waveform tomography algorithm can be applied in a straightforward way to perform receiver function migration and travel time inversion.

**Key words:** Body waves – Computational seismology – Seismic tomography – Wave scattering and diffraction.

## 1 INTRODUCTION

Analyses of seismograms recorded by regional seismic networks of body waves generated by sources at teleseismic distances have made tremendous contributions to our understanding of the Earth's interior. Two techniques in particular, travel time tomography and receiver function analysis, have evolved from their introduction some 30 years ago (Aki et al., 1977; Vinnik, 1977; Langston, 1979) to become standard tools for visualizing properties of the medium directly beneath the network. Travel time tomography techniques attempt to unravel an integrative effect of slowness along a ray path or Fresnel zone, while receiver function analysis focuses more on the location of discrete scatterers, often in the form of continuous interfaces, that generate coherent mode conversions. Each approach has its particular advantages and shortcomings. Travel time tomography images are generally robust and useful for illuminating large scale features, but the resolution of the images is often low and features such as isolated zones of low wavespeed may be missed entirely. Receiver functions provide an excellent means for detecting localized contrasts in medium properties, but the locations of those contrasts in space and their implications for changes in the medium are often poorly defined. For example, the illumination of the Mohorovičić discontinuity is now a routine result of P receiver function analysis, and the Lithosphere-Asthenosphere boundary, until recently considered an exotic target, has become a fairly robust result of S receiver function analysis (e.g., Oreshin et al., 2002; Kumar et al., 2005). However, an accurate determination of crustal and lithospheric thickness is difficult without ancillary information on background wavespeeds, for example from surface wave dispersion analysis (e.g., Herrmann et al., 2001). Even with a reasonable background model, quantitative interpretation of mode conversions can be difficult if the medium is laterally heterogeneous, and as a result this type of analysis tends to fall into one of two end-member categories: the “modeling” end that presumes that the Earth beneath a network is, at least locally, approximately one-dimensional, and a “migration” end that is interested primarily in locating the sources of the mode conversions with at most a cursory explanation for the polarity and amplitude of the scattered phase. A number of investigators at both ends have helped to reduce this gap. For example, on the modeling end, Vinnik et al (2004) compensate for lateral heterogeneity by assuming piecewise homogeneity and including complementary information from travel time tomography. Advances on the migration side include relatively straightforward incorporation of tomographically derived wavespeeds (e.g., Wilson and Aster, 2003) and a variety of wave equation based algorithms that locate scatterers using a Born (single scatterer) approximation (e.g., Bostok et al, 2001; Poppeliers and Pavlis, 2003a; 2003b; Levander et al., 2005) Each of these approaches has the advantage of requiring modest computational resources, but still leaves a substantial gap between modelers and migraters.

In this paper we attempt to reduce this gap further by introducing a full waveform tomographic

technique for quantitative modeling of teleseismic body waves, both primary and multiply scattered, that can be applied to 2.5D viscoelastic media. The full waveform approach simultaneously models and migrates receiver functions and implicitly includes a form of travel time tomography as well.

## 2 TRACTABLE WAVEFORM TOMOGRAPHY

The term *waveform tomography* refers to a class of algorithms that attempts to recover information about medium properties from the entire seismic waveform. The basic concepts of waveform tomography were introduced through the seminal work of Lailly (1983), Tarantola (1984), Mora (1987) and Woodward (1992), and, with a few exceptions (receiver function modeling being one) have for the most part been used in the active source environment. An attractive advantage of waveform tomographic techniques over traditional travel time tomography is the potential for significantly improved image resolution: there simply is much more information in a waveform than there is in an arrival time. For example, in a side by side comparison of the two techniques in a simulated cross hole environment, Pratt (2003) demonstrated that while travel time resolution was limited by the size of the Fresnel zone, waveform tomography resolution was on the order of the dominant wavelength in the signal.

The algorithm we discuss here is an adaptation of waveform tomography techniques developed over the past 18 years by R. G. Pratt and coworkers (mostly G. Hicks, L. Sirgue, Z. M. Song, I. Štekl, and P. R. Williamson; for convenience these techniques are referred to collectively below as the RGP techniques). They are robust enough to have been successfully applied in analysis of real data (e.g., Brittan et al., 1997; Bleibinhaus et al., 2007). Equally as important as their track record, however, is their demonstrated ability to generate reasonably accurate images in heterogeneous media without resorting to computing on a massive scale. For example, they were used to image a hypothetical and highly heterogeneous structure in a competition promoted by a Commission on Controlled-Source Seismology (CCSS) workshop organized in 2003 (Brenders and Pratt, 2003; 2007; Hole et al., 2005). The image generated by these techniques was easily as good as those generated by higher powered algorithms but had the advantage of being produced using a single Linux box.

While massively parallel computing has been a boon to seismology (e.g., Acelik et al., 2002; 2003), it can be a rather inconvenient and expensive proposition. While some investigators have had success in generating images with 3D algorithms (e.g., Chen et al., 2007; Tape et al., 2009), existing codes that can handle general 3D media typically require large scale computing simply to do the forward problem. Hence, inverse modeling with these approaches can be prohibitively time consuming and expensive. Our aim is to develop a tool that investigators with modest resources will find useful. An attraction of the RGP techniques is that they are geared towards smaller machines.

Our starting point is the acoustic inverse problem described in Pratt et al. (1990, 1998, 1999, 2003) and Song et al. (1995a, 1995b) and an extension of the forward problem to the elastic case by Pratt (1990) and Štekl and Pratt (1998). For clarity we summarize below the key concepts of those studies that have direct bearing on the 2.5D teleseismic case.

### 3 WAVEFORM COMPUTATION

We define the forward problem for modeling teleseismic body waves as the response of a locally heterogeneous, viscoelastic medium to an excitation by plane waves generated by sources external to that medium. We start with the Cartesian form of the three dimensional elastic wave equation for displacement  $\mathbf{u} = (u, v, w)$  in an isotropic, heterogeneous medium in the frequency domain

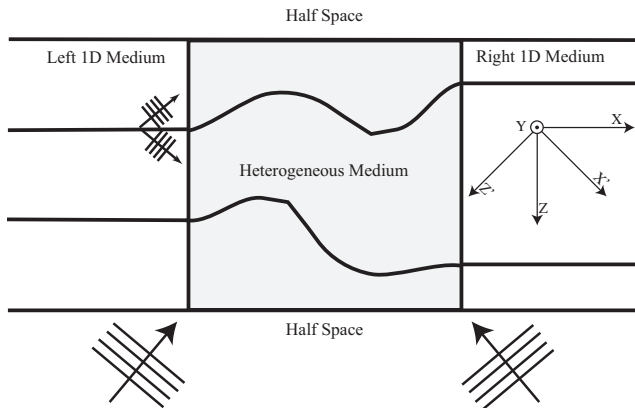
$$\begin{aligned}\omega^2 \rho u + \frac{\partial}{\partial x} \left[ (\lambda + 2\mu) \frac{\partial u}{\partial x} + \lambda \left( \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} \right) \right] + \frac{\partial}{\partial y} \left[ \mu \left( \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right) \right] + \frac{\partial}{\partial z} \left[ \mu \left( \frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} \right) \right] &= -f_x \\ \omega^2 \rho v + \frac{\partial}{\partial y} \left[ (\lambda + 2\mu) \frac{\partial v}{\partial y} + \lambda \left( \frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} \right) \right] + \frac{\partial}{\partial x} \left[ \mu \left( \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right) \right] + \frac{\partial}{\partial z} \left[ \mu \left( \frac{\partial v}{\partial z} + \frac{\partial w}{\partial y} \right) \right] &= -f_y \\ \omega^2 \rho w + \frac{\partial}{\partial z} \left[ (\lambda + 2\mu) \frac{\partial w}{\partial z} + \lambda \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) \right] + \frac{\partial}{\partial x} \left[ \mu \left( \frac{\partial w}{\partial x} + \frac{\partial u}{\partial z} \right) \right] + \frac{\partial}{\partial y} \left[ \mu \left( \frac{\partial w}{\partial y} + \frac{\partial v}{\partial z} \right) \right] &= -f_z\end{aligned}$$

where  $\omega$  is the radial frequency and  $\mathbf{f} = (f_x, f_y, f_z)$  is a source term. Conceding that a full 3D approach is beyond the capabilities of moderately sized computers, we reduce the above equations to 2.5D by recognizing a plane wave invariance in the  $y$  direction (Takenaka and Kennett, 1996) expressed in the frequency domain as

$$\mathbf{u}(x, y, z, \omega) = \mathbf{A}(\omega) \exp [i(\omega t - \omega p_y y - k_x x - k_z z)] \quad (1)$$

where  $p_y$  is the  $y$  component of the slowness vector. Recognizing that  $\partial \mathbf{u} / \partial y = -i\omega p_y \mathbf{u}$ ,  $\partial^2 \mathbf{u} / \partial y^2 = -\omega^2 p_y^2 \mathbf{u}$ ,  $\partial \lambda / \partial y = 0$ , and  $\partial \mu / \partial y = 0$ , the wave equations in 2.5D become

$$\begin{aligned}(\rho - p_y^2 \mu) \omega^2 u + \frac{\partial}{\partial x} \left[ (\lambda + 2\mu) \frac{\partial u}{\partial x} + \lambda \frac{\partial w}{\partial z} \right] + \frac{\partial}{\partial z} \left[ \mu \left( \frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} \right) \right] - i\omega p_y \left( \frac{\partial}{\partial x} \lambda v + \mu \frac{\partial v}{\partial x} \right) &= -f_x \\ (\rho - p_y^2 (\lambda + 2\mu)) \omega^2 v + \frac{\partial}{\partial x} \left( \mu \frac{\partial v}{\partial x} \right) + \frac{\partial}{\partial z} \left( \mu \frac{\partial v}{\partial z} \right) - i\omega p_y \left[ \lambda \left( \frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} \right) + \frac{\partial}{\partial x} \mu u + \frac{\partial}{\partial z} \mu w \right] &= -f_y \\ (\rho - p_y^2 \mu) \omega^2 w + \frac{\partial}{\partial z} \left[ (\lambda + 2\mu) \frac{\partial w}{\partial z} + \lambda \frac{\partial u}{\partial x} \right] + \frac{\partial}{\partial x} \left[ \mu \left( \frac{\partial w}{\partial x} + \frac{\partial u}{\partial z} \right) \right] - i\omega p_y \left( \frac{\partial}{\partial z} \lambda v + \mu \frac{\partial v}{\partial z} \right) &= -f_z\end{aligned}$$



**Figure 1.** Schematic of the model. We consider a heterogeneous medium (grey rectangle) bounded by half spaces on the top and bottom and vertically stratified media on either side. The upper boundary is usually, but not necessarily, a free surface. Plane waves enter the medium from below, but converted waves may also enter the heterogeneous region from the side. The 1D media on the left and right are used for waves entering from the left and right, respectively, and are not necessarily the same. To prevent diffractions, the wavespeeds in the heterogeneous medium match those in the 1D media at the boundaries. Orientations of the “non-rotated”  $(x, y, z)$  and  $45^\circ$  rotated  $(x', y, z')$  axes are indicated at the right.

The advantage of 2.5D is that we can model a wide variety of environments of geologic interest that can be expected to be approximately invariant along a particular axis (e.g., mountain belts, rift zones, subduction zones, basins) without significantly increasing the computational demands of the 2D case, but at the same time relaxing the 2D restriction that all sources and receivers be in the same plane.

We compute waveforms by solving the 2.5D system of equations above using finite differences on a 2D  $(x, z)$  grid. The medium is parameterized by specifying a density and either P- and S- wavespeeds or the Lamé parameters  $\lambda$  and  $\mu$  at each node. The grid is bounded by a vertically dipping rectangle imbedded in background 1D media (Figure 1). The media on either vertical side of the rectangle can be vertically stratified but are not necessarily identical. Our method requires only that phases scattered inside the heterogenous region not reenter the rectangle after they propagate outside of it. Hence the media above and below the rectangle must be homogeneous half spaces; typically the top of the rectangle is a free surface. The half space bounding the bottom of the heterogeneous medium ensures that the time-reduced waveform incident along this boundary is everywhere the same.

A principal concern in solving wave equations using finite differences is the potential for introducing numerically generated dispersion and anisotropy. These artifacts derive from the differences in P and S wavespeeds specified in the continuous medium and our finite difference representations of them. The size of these artifacts decreases as the grid density increases, but of course at a cost of increased memory and computational time. Following Štekl and Pratt (1998) and Jo et al. (1996)

we mitigate these artifacts by combining two Cartesian systems related by a 45° rotation about the  $y$  (invariant) axis (Figure 1). In the “non-rotated” system,  $x$  is horizontal and  $z$  is positive down. The rotated  $(x', z')$  system is defined by

$$x' = \frac{z + x}{\sqrt{2}} \text{ and } z' = \frac{z - x}{\sqrt{2}}$$

The  $y$  axis is the same for both systems. Positive directions for  $x$  and  $y$  are arbitrary but conform to  $\hat{\mathbf{x}} \times \hat{\mathbf{y}} = \hat{\mathbf{z}}$ .

To combine systems, we create an additional set of equations in the  $(x', z')$  45° rotated system by substituting

$$\frac{\partial}{\partial x} = \frac{1}{\sqrt{2}} \left( \frac{\partial}{\partial x'} - \frac{\partial}{\partial z'} \right) \text{ and } \frac{\partial}{\partial z} = \frac{1}{\sqrt{2}} \left( \frac{\partial}{\partial x'} + \frac{\partial}{\partial z'} \right)$$

into the 2.5D version of the wave equation. We then solve this system of equations using finite differences on a weighted sum of the contributions from the two (primed and unprimed) frames:

$$\begin{aligned} (\rho - p_y^2 \mu) \omega^2 u + aA + (1 - a)A' &= -f_x \\ [\rho - p_y^2 (\lambda + 2\mu)] \omega^2 v + aB + (1 - a)B' &= -f_y \\ (\rho - p_y^2 \mu) \omega^2 w + aC + (1 - a)C' &= -f_z \end{aligned} \quad (2)$$

For added stability we include a lumped mass contribution (Marfurt, 1983; Štekl and Pratt, 1998) for the central node of a 5-point differencing star:

$$\rho' \mathbf{u}' = b\rho \mathbf{u} + \frac{(1 - b)}{4} (\rho^+ \mathbf{u}^+ + \rho^- \mathbf{u}^- + \rho_+ \mathbf{u}_+ + \rho_- \mathbf{u}_-)$$

where we have used the differencing notation described in Appendix A. The weights  $a$  and  $b$  are determined by an optimization procedure, discussed in detail in Appendix B, that minimizes numerical dispersion over all directions of propagation. When  $p_y = 0$  the weights reduce to the forms found in Štekl and Pratt (1998) for the 2D elastic case\* for P-SV and to those in Jo et al. (1996) for the acoustic case for SH. The off-axis case is somewhat more complicated as it involves complex terms (specifically the  $ip_y \omega$  terms that come from the  $\partial/\partial y$  partial derivatives). Nevertheless, we find that with appropriate weighting numerical dispersion can be suppressed to less than a few percent with grid spacing as sparse as 4 points per wavelength.

After replacing the continuous derivatives with the finite difference stencils of Appendix A, the forward problem can be written as the solution to a system of linear equations:

$$S\mathbf{u} = -\mathbf{f} \quad (3)$$

If the grid has  $n_x \times n_z$  nodes, then  $\mathbf{u}$  is a  $3n_x n_z$  length vector containing the  $(u, v, w)$  displacement

\* With the exception of an error in their derivation discussed in Appendix B

field at each point in the model,  $\mathbf{f}$  is a  $3n_x n_z$  length vector containing the sources at each point ( $\mathbf{f} = (f_x^i, f_y^i, f_z^i)$ ), and  $S$  is a banded, structurally symmetric  $3n_x n_z \times 3n_x n_z$  matrix containing the finite difference terms. Following Pratt (1990), we solve for the wavefield

$$\mathbf{u} = -S^{-1} \mathbf{f} \quad (4)$$

using LU decomposition. The number of nonzero terms (i.e., “fill”) in the decomposition is minimized by using the nested decomposition algorithm of George and Liu (1970) as implemented in the *Metis* package of Karypis and Kumar (1998).

### 3.1 Using a Background Wavefield to Specify Forces

In the case of teleseismic body waves the forces in equation (3) are plane waves generated by distant events external to the modeled medium. Specification of such forces in (3) may seem inappropriate because in the absence of internal sources

$$S\mathbf{u} = 0 \quad (5)$$

and the solution to the resulting homogeneous equation (5) is non-unique. Nevertheless, one can ask if there is there a distribution of sources  $\mathbf{f}_o$  that will generate a desired  $\mathbf{u} = \mathbf{u}_o$  in a medium described by a finite difference matrix  $S_o$ . Evidently, an appropriate force distribution can be computed directly from

$$S_o \mathbf{u}_o = \mathbf{f}_o \quad (6)$$

Hence if we have a means to specify a field  $\mathbf{u}_o$  in a medium described by  $S_o$ , we can use the force distribution above to compute the response  $\mathbf{u}$  of another medium described by  $S$  to the same excitation by solving  $\mathbf{u} = S^{-1} \mathbf{f}_o$ . Equation (6) shows that in order to generate an appropriate background wavefield we must specify forces at every point in the medium. However, the differencing nature of  $S_o$  means that most of the energy in  $\mathbf{f}_o$  will be located where the gradient in the displacement field is largest, and this generally will be at the boundaries of the grid. In particular, the forces at the free (top) surface can be several orders of magnitude greater than anywhere else. While not an issue conceptually, such a concentration of source energy all along the free surface can be numerically unstable as the wavefield will necessarily depend on choices we make on how to compute the response of this surface. For this reason we will describe an alternate approach that is theoretically identical but numerically more stable.

We can describe the total displacement field  $\mathbf{u}_t$  as the sum of a background wavefield  $\mathbf{u}_o$  and a scattered wave field  $\mathbf{u}_h$  generated by perturbations to the background medium

$$\mathbf{u}_t = (\mathbf{u}_o + \mathbf{u}_h) \quad (7)$$

We can similarly separate our differencing matrix  $S$  into a component  $S_o$  that produces the background wavefield and a component  $S_h$  that produces the scattered field

$$S = S_o + S_h \quad (8)$$

Recalling that in the absence of internal sources

$$S\mathbf{u}_t = (S_o + S_h)(\mathbf{u}_o + \mathbf{u}_h) = S_o\mathbf{u}_o + S_h\mathbf{u}_o + S\mathbf{u}_h = 0 \quad (9)$$

The background plane wave also has no internal sources, hence  $S_o\mathbf{u}_o = 0$  and

$$S\mathbf{u}_h = -S_h\mathbf{u}_o = -(S - S_o)\mathbf{u}_o \quad (10)$$

Using (10), we solve for the scattered wavefield and then add the background to obtain the total field

$$\mathbf{u}_t = \mathbf{u}_o - S^{-1}[(S - S_o)\mathbf{u}_o] \quad (11)$$

In this case the force distribution is  $\mathbf{f} = -(S - S_o)\mathbf{u}_o$ , meaning that forces appear only at perturbations to the background wavespeed, which generally will not include any of the boundaries. Hence in this case we avoid having large amplitude sources distributed all along the free surface. To show that the two approaches are conceptually equivalent, recall that in the total field approach (6) we specified a distribution of forces  $\mathbf{f}_o$  such that

$$S\mathbf{u}_t = S_o\mathbf{u}_o = \mathbf{f}_o \quad (12)$$

and so from (9)

$$S\mathbf{u}_t = (S_o + S_h)(\mathbf{u}_o + \mathbf{u}_h) = \mathbf{f}_o + S_h\mathbf{u}_o + S\mathbf{u}_h = \mathbf{f}_o \quad (13)$$

and finally

$$S\mathbf{u}_h = -S_h\mathbf{u}_o \quad (14)$$

which is equation (10). Hence we can solve for the total wavefield directly as a response to a distribution of sources that create a wavefield in the background medium, or solve for the scattered wavefield by computing the response a distribution of sources located at the heterogeneities in the medium. We generally choose to use the scattered approach in (10) as it seems to be more stable numerically.

### 3.2 The Plane Wave Source

In implementing (10) we could specify any background wavefield  $\mathbf{u}_o$  corresponding to a given  $S_o$ . For example,  $\mathbf{u}_o$  could be the wavefield computed from a previous iteration of the inverse problem for which we already have stored an appropriate  $S_o$ . Generally, however, wavefields for laterally varying media of interest are not known a priori, and so we begin with simpler media for which we have analytical solutions. In the case of teleseismic body waves a suitable  $\mathbf{u}_o$  should allow for surface reflections

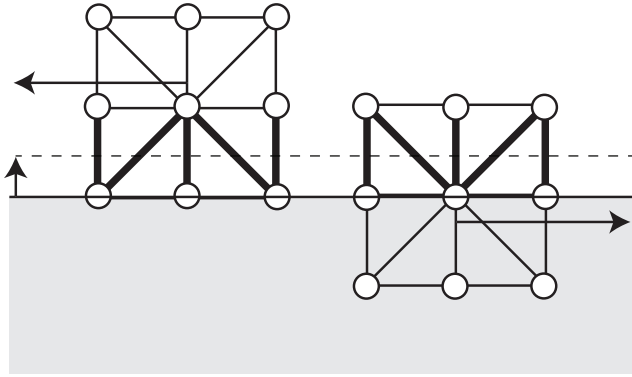
outside the heterogenous (“modeled”) region. It should also allow for both continuous variations in medium properties with depth and relatively abrupt changes, such as the Moho or sediment/basement interfaces, that would generate large amplitude mode conversions both inside and outside of the heterogenous region.

Fortunately there are several techniques available to compute the response of a 1D (vertically stratified) medium to plane wave excitation. For our purposes we employ an algorithm based on the Thomson-Haskell approach (Haskell, 1953; 1960; 1962), mostly because it is a frequency domain approach that has been used extensively in modeling 1D P- and S- receiver functions (e.g., Kosarev et al., 1993; Oreshin et al., 2002). We adapt it here by computing the response at the free surface to an impulse at the base of the model as described in Haskell (1960; 1962) and then recovering the response at any point in the medium by accumulating the propagator matrices in reverse order and applying an appropriate phase shift ( $\theta(\delta x) = e^{i\omega p_x \delta x}$  where  $p_x$  is the horizontal slowness) along the base of the model. The computation of the wavefield in this manner is highly efficient.

One can surmise from this approach that it will be desirable to match the external 1D model to the internal model along the vertical boundaries, otherwise the change in wavespeed could generate unwanted diffractions. At the same time, it is not necessary to assign the same 1D model to both sides of the box. The incident wave is never exactly vertical and so is influenced only by the model on the same side of the heterogenous medium. Hence we may allow discontinuities to change depth within the medium without having to return to an original depth on the opposite side (Figure 1).

In applying the analytical form for  $u_o$  one must be aware of potential artifacts caused by differences between an analytical (i.e., continuous) representation of the medium and our finite difference approximation of it. The problem of numerical dispersion was discussed above and a considerable amount of effort was spent minimizing those effects. We note that some investigations in the electromagnetic literature recommend an additional dispersion correction when using analytical plane wave solutions in the time domain (e.g., Schneider, 2004; Tan and Potter, 2006). We attempted to emulate this recommendation by computing phase velocity corrections in the manner described in Appendix B and applying them to the analytical model prior to computing  $u_o$ . However, we found that for the current algorithm such corrections were insignificant at best, and in some cases actually degraded the computed wavefield. Hence we do not apply any analogous correction here.

A more subtle artifact created by the mismatch between continuous and discrete models comes from the appearance of nonphysical or “ghost” phases caused when an interface in the background 1D models is displaced (or otherwise removed) from its original depth within the heterogenous medium. To understand the cause of this artifact, consider the sources created by application of equation (6) across an interface separating media with different properties (Figure 2). Forces of the form



**Figure 2.** Illustration of the origin of ghost phases generated when an interface aligns with the grid. The two nine-point stars average the medium properties between points connected by the thick black lines. The differences in the forces resulting from equation (6) are nearly equal but opposite in sign, leading to the creation of a force couple (indicated by horizontal arrows) separated by one grid space at the interface. The waves generated by this force couple can be attenuated either by replacing the interface with a gradual gradient in medium properties or by displacing the interface upwards (to the dashed line) an amount that equates travel times in the discrete and continuous representations of the medium between the grid points.

$\mu_{-1/2}(u_+ - u)$  will be created at points just above the interface, and of the form  $\mu_{+1/2}(u - u_-)$  at points along the interface. The net effect is a quasi-couple across the boundary with nearly equal forces in opposite directions. This force couple creates waves that compensate for the mismatch between the analytical and finite difference representations of the medium (as discussed below, waves generated using equation (6) in a 1D medium match the analytical form very well), but if the source of that mismatch (i.e., the interface) is removed these compensating waves still exist but no longer play the role of primary phase modifiers. Instead they appear as small amplitude waves that emanate from an invisible interface.

This effect is difficult to eliminate entirely, but because it is caused by the mismatch between continuous and discrete representations of wavespeed variations between the grid points near the interface, they may be reduced considerably by decreasing this mismatch. One way to do this is to replace the infinitesimal interface by a gradient in wavespeed over several grid points. Alternatively, one can displace the interface a fraction of a grid spacing upwards from the first grid point at or below the interface. An appropriate displacement can be estimated by equating the travel times across the grid space  $h$ ; for mixed mode waves we find that an average transit time for both phases works well. Specifically we calculate an upward displacement  $\delta z$  from:

$$\delta z = 0.5hV_{inc}^b \left[ 1.0/(V_{inc}^b + V_{inc}^a) + 1.0/(V_{inc}^b + V_{con}^a) \right] \quad (15)$$

where the superscripts  $a$  and  $b$  refer to the wavespeeds above below and above the interface, respec-

tively, and the subscripts *inc* and *con* refer the modes (P or S) of the incident and converted waves, respectively.

### 3.3 Boundary Conditions

The top side of the heterogenous medium generally is taken to be the free surface and we apply the usual traction-free conditions:

$$\begin{aligned}\sigma_{zx} &= \mu \left( \frac{\partial u}{\partial z} + \frac{\partial w}{\partial x} \right) = 0 \\ \sigma_{zy} &= \mu \left( \frac{\partial v}{\partial z} + \frac{\partial w}{\partial y} \right) = 0 \\ \sigma_{zz} &= \lambda \left( \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) + (\lambda + 2\mu) \frac{\partial w}{\partial z} = 0\end{aligned}$$

We represent these equations in finite difference form using the templates summarized in Appendix A, and in the 2.5D case we evaluate the  $y$  derivatives above as  $\partial/\partial y = -i\omega p_y$ . The remaining sides of the medium are fixed displacement (Dirichlet conditions) with some number (nominally 10, although as few as 5 give satisfactory results) of grid points forming a Perfectly Matched Layer (PML) to eliminate reflections (e.g., Collino and Tsogka, 2001). Adaptations of PML to the second order elastic wave equation are discussed in Komatitsch and Tromp (2003) and Zheng and Huang (2002). The two treatments are essentially equivalent although we found the direct use of complex coordinate stretching in Zheng and Huang (2002) easier to adapt to combinations of rotated coordinate systems. The incorporation of a PML in the determination of the scattered field involves solving the 2.5D wave equations using stretched coordinates, but a slightly different approach should be taken when the background wavefield is required to decay into the PML as well. Following Zheng and Huang (2002), we stretch the  $x_i$  coordinate by defining a complex quantity  $\epsilon_i$  such that

$$\hat{x}_i = x_i^o + \int_{x_i^o}^{x_i} \epsilon_i(x') dx' \quad (16)$$

where  $x_i^o$  marks the edge of the PML and  $x_i < x_i^o$  is the normal heterogeneous region. It is straightforward to show that  $\epsilon_i$  is related to the  $d(x)$  term of Komatitsch and Tromp (2003) and Collino and Tsogka (2001) by

$$\epsilon_i(x) = 1 - \frac{i}{\omega} d(x) \quad (17)$$

Collino and Tsogka (2001) suggest the following form for  $d(x)$ :

$$d(x) = \frac{3\alpha}{2\delta} \left( \frac{x}{\delta} \right)^2 \log_{10} \left( \frac{1}{R} \right) = Dx^2 \quad (18)$$

where  $\alpha$  is the P wavespeed,  $\delta$  is the width of the PML, and  $R$  is a reflection coefficient generally taken to be 0.001. To apply this concept to the background wavefield we evaluate the integral in (16) to obtain

$$\hat{x}_i = x_i + \frac{D}{i\omega} \frac{x_i^3 + x_o^3}{3} \quad (19)$$

Without loss of generality, we take  $x_o = 0$  so that  $x_i$  represents the distance into the PML:

$$\hat{x}_i = x_i \left( 1 - i \frac{D}{\omega} \frac{x_i^2}{3} \right) \quad (20)$$

We then compute  $\mathbf{u}(\omega)$  within the PML ( $\mathbf{u}_p(\omega)$ ) from

$$\begin{aligned} \mathbf{u}_p(\omega) &= \mathbf{A}(\omega) \exp [i(\omega t - \omega p_y y - k_x \hat{x} - k_z \hat{z})] \\ &= \mathbf{u}(\omega) \exp \left[ -\frac{D(p_x x^3 + p_z z^3)}{3} \right] \end{aligned} \quad (21)$$

which causes the background wavefield to attenuate exponentially within the PML in a manner consistent with the scattered wavefield.

### 3.4 The Inverse Problem

Once a forward problem is implemented there are a number of potential procedures for constructing an inverse problem. Several approaches germane to waveform tomography were investigated for the acoustic case by Pratt et al. (1998). We make use of the gradient approach here partly because of its intuitive character but more because it does not require significantly more computation than the forward problem. We will review the basics of this approach here to illustrate how it can be adapted to the 2.5D teleseismic case.

We begin by defining a residual vector  $\delta \mathbf{d} = \mathbf{u} - \mathbf{d}$  where  $\mathbf{u}$  and  $\mathbf{d}$  are the calculated and observed wavefields, respectively. We use  $\delta \mathbf{d}$  to create an objective function

$$E(\mathbf{m}) = \frac{1}{2} \delta \mathbf{d}^{*T} \delta \mathbf{d} \quad (22)$$

where the vector  $\mathbf{m}$  describes the model. The inverse problem then consists of solving for a perturbation  $\delta \mathbf{m}$  to  $\mathbf{m}$  that minimizes the objective function. The gradient method estimates  $\delta \mathbf{m}$  by taking a step in the direction of the gradient of  $E$ :

$$\delta \mathbf{m} = -\gamma \nabla_{\mathbf{m}} E(\mathbf{m}) \quad (23)$$

where  $\gamma$  is a scalar. It is not difficult to show that

$$\nabla_{\mathbf{m}} E(\mathbf{m}) = \frac{1}{2} \nabla_{\mathbf{m}} \delta \mathbf{d}^{*T} \delta \mathbf{d} = \text{Re} \sum_{i=1}^n \frac{\partial u_i}{\partial m} \delta d_i^* = \text{Re} (J^T \delta \mathbf{d}^*) \quad (24)$$

where  $J$  is the Jacobian. Recalling that the forward problem is  $S\mathbf{u} = -\mathbf{f}$ , a derivative with respect to a particular model element  $m_i$  gives

$$S \frac{\partial \mathbf{u}}{\partial m_i} + \frac{\partial S}{\partial m_i} \mathbf{u} = 0 \quad (25)$$

and so

$$\frac{\partial \mathbf{u}}{\partial m_i} = -S^{-1} \frac{\partial S}{\partial m_i} \mathbf{u} = -S^{-1} \mathbf{f}^{(i)} \quad (26)$$

where we define  $F = \mathbf{f}^{(i)}$  as a set of virtual forces. This leads to

$$\begin{aligned} \nabla_m E(\mathbf{m}) &= \text{Re} (J^T \delta \mathbf{d}^*) = \text{Re} \left[ (S^{-1} F)^T \delta \mathbf{d}^* \right] \\ &= \text{Re} \left[ F^T (S^{-1})^T \delta \mathbf{d}^* \right] = \text{Re} (F^T \mathbf{v}) \end{aligned} \quad (27)$$

Pratt et al. (1998) argue that as long as reciprocity holds it should be the case that

$$\mathbf{v} = (S^{-1})^T \delta \mathbf{d}^* = S^{-1} \delta \mathbf{d}^* \quad (28)$$

so that no additional factorization of  $S$  is required to do the inverse problem. When true, this identity significantly reduces the computational effort (LU factorization is the bottleneck in these algorithms), and is why the gradient method was chosen for this application: the inverse problem in this case requires only slightly more computation than the forward problem.

We note that in general if  $S$  is full rank then a singular value decomposition of a complex valued  $S = U\Lambda V^H$  can be manipulated to show that

$$(S^{-1})^T = (S^T)^{-1} \quad (29)$$

so that (28) is true as long as  $S$  is symmetric. Inspection of the templates that make up  $S$  (Appendix A) shows that in the 2D case it is indeed symmetric everywhere *except* within the PML, the reason being that direction of propagation is consequential within these boundary layers (and so reciprocity no longer applies). It is not difficult to compute  $(S^T)^{-1}$ , in fact we can do this with the same LU factorization of  $S$  (one simply solves the system of equations using  $S^T = U^T L^T$ ). However, we argue that because the PML is an artificial construct, and because  $S$  would be symmetric in an infinite grid with no artificial boundaries, it is in fact more correct to use  $S^{-1}$  in (28).

In the 2.5D case ( $p_y \neq 0$ ), the part of  $S$  outside the PML is Hermitian and thus no longer strictly symmetric (i.e.,  $S^T \neq S^H$ ). In this case we can write  $S = U\Lambda U^H$  and

$$(S^{-1})^T = U^{HT} \Lambda^{-1} U^T = U^* \Lambda^{-1} U^T = (S^*)^{-1}$$

where  $U^*$  refers to the complex conjugate of each element of the eigenvector matrix  $U$  without transposing, while  $U^H$  is the complex conjugate transpose of  $U$ . For a Hermitian matrix,  $S^* = S^T$ , so (29) still applies, but it is no longer the case that we can simply use the inverse of the  $S$  itself in (28). We

should instead use

$$\mathbf{v} = S^{*-1} \delta \mathbf{d}^* \quad (30)$$

in (28) but recall that  $S^*$  refers only to the part of  $S$  *outside* the PML. Because terms in the PML are complex, we form  $S^*$  simply by reversing the sign on  $p_y$ .

Pratt et al. (1998) identify  $\mathbf{v}$  in (28) as a wavefield generated by back propagating time reversed scattered waves from the receivers using the same propagator matrix as the forward problem. We retain the same back propagation aspect in (30), and in this case reversing the sign on  $p_y$  is equivalent to sending the wave back in the opposite  $y$  direction from whence it arrived.

The  $F$  term in (27) is the forward propagated wavefield modified by a scattering factor  $\frac{\partial S}{\partial m_i}$ . In the acoustic case this factor is a single scalar that is dependent on frequency and medium slowness. As Pratt et al. (1998) point out, because a product of a function with the complex conjugate of another function in the frequency domain is a correlation of these functions in the time domain, the gradient in (27) is simply a weighted correlation between predicted and observed scattered wavefields at every point in the medium. When the correlation is large, the corresponding perturbation to the model will be large as well. A similar intuitive interpretation applies in the 2.5D elastic case, although as we have formulated it here, each  $m_i$  appears in several elements of  $S$ . As a result, instead of a single scalar multiplying the forward wavefield, each  $\frac{\partial S}{\partial m_i}$  is represented by a 27 x 27 submatrix with as many as 369 nonzero elements (Appendix C). This much larger number of elements makes a simple generalization of their effects less straightforward but at the same time does not entail a significant increase in computational effort.

### 3.5 Source and Receiver Statics

We consider the plane wave impinging on the heterogenous medium from below to be associated with a particular source time function (STF) which in general includes all of the effects on the incident wave prior to its arrival at the model base. These would include the history of energy release at the source and the interaction of the waves (reflections and mode conversions) that take place near the source region. The STF usually is considered to be an unresolvable “nuisance” parameter in imaging algorithms. In travel time tomography it takes the form of an unknown origin time that we remove by substituting relative for absolute wavespeeds in the inverse problem. In receiver function analysis the STF typically is eliminated by deconvolving one component of motion from another. For example, in P-receiver function modeling we represent the radial ( $r$ ) and vertical ( $z$ ) components of observed

motion as

$$\begin{aligned} z(\omega) &= s(\omega)G_z(\omega)i_z(\omega) \\ r(\omega) &= s(\omega)G_r(\omega)i_r(\omega) \end{aligned} \quad (31)$$

where  $s(\omega)$  is the source time function,  $i_r$  and  $i_z$  are instrument responses that generally are known and removed prior to analysis, and  $G_z$  and  $G_r$  are the vertical and radial components of the model Green's functions. We then employ one of several algorithms in the time or frequency domain to deconvolve the  $z$  component from the  $r$  component to arrive at a form resembling

$$\frac{r(\omega)}{z(\omega)} = \frac{G_r(\omega)}{G_z(\omega)} \quad (32)$$

The STF is a nuisance parameter in teleseismic waveform tomography as well, but if we have an array of stations recording the same event we can adapt a technique suggested by Pratt (1990) for removing the source contribution from explosions in the active source environment. As stated in equation (31), the wavefield is a convolution of the STF with the medium Green's functions, and we can write (6) as

$$S\mathbf{u} = S[G(\omega)s(\omega)] = \mathbf{f}_o \quad (33)$$

We estimate the STF by adjusting  $s(\omega)$  to minimize the difference between observed and calculated wavefields. That is, if the set  $\mathbf{d}$  of observations from a given event satisfies

$$S\mathbf{d} = \mathbf{f}_s \quad (34)$$

then we seek a STF  $\hat{s}(\omega)$  such that

$$S[G(\omega)\hat{s}(\omega)] = \mathbf{f}_s \quad (35)$$

To do so, we modify the STF at any particular frequency by a complex scalar  $c$  so that

$$\begin{aligned} S[G(\omega)\hat{s}(\omega)] &= S[G(\omega)(c\hat{s}(\omega))] = S[c\mathbf{u}] = cS\mathbf{u} \\ &= c\mathbf{f}_o = \mathbf{f}_s \end{aligned}$$

Hence, to approach (34) we minimize the difference between  $\mathbf{d}$  and  $c\mathbf{u}$ . If we pose this as the minimization of an objective function

$$E(c) = \frac{1}{2} \sum_{i=1}^n (\mathbf{d} - c\mathbf{u})_i^* (\mathbf{d} - c\mathbf{u})_i$$

then setting  $\frac{\partial E(c)}{\partial c} = 0$  leads to

$$c = \frac{\mathbf{u}^* \mathbf{d}}{\mathbf{u}^* \mathbf{u}} \quad (36)$$

which is identical to the source correction term derived by Pratt (1990). Intuitively, (36) shows that  $c$

is determined by a correlation of observed ( $\mathbf{d}$ ) and predicted ( $\mathbf{u} = S^{-1}\mathbf{f}$ ) wavefields at each station (normalized by the predicted wavefield). It is similar to stacking, but in this case the stack is weighted by the expected medium response. Hence we solve for  $c$  each time we perturb the medium rather than simply determine an average waveform at the outset (note that if  $\mathbf{u}$  is a delta function, then  $c$  is a simple stack of observed waveforms). It is also analogous to eliminating origin times in travel time tomography, in that parts of the waveform that appear to be in some sense “common” to all stations are assigned to the source time function. Equation (36) allows for timing adjustments in the form of phase shifts, but these shifts are frequency dependent and an investigator needs to be mindful of the potential for cycle skips. Rather than depend solely on time adjustments calculated from (36) one would be well advised to generate a travel time tomography image to use as a starting model.

An advantage of this approach over traditional teleseismic receiver functions is that the usual initial step of source deconvolution, with all its attendant assumptions and potential for noise enhancement, is not required. Instead, a source time function can be determined for each event, and the response of the medium to this function can be calculated explicitly. The computation is economical: we compute the Green’s functions in  $S^{-1}$  only once for a given  $p_y$  and loop through all sources  $\mathbf{f}$  to generate the wavefields  $\mathbf{u}$  from which we can determine  $c$ .

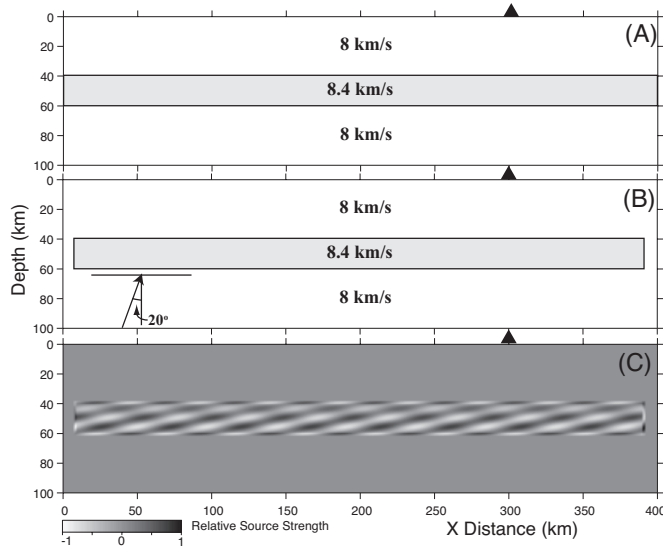
A common concern in teleseismic travel time tomography is timing offsets at stations caused by local wavespeed anomalies, and a standard method for compensating for these offsets is to subtract a station-based average residual (i.e., a “station correction”). In the waveform tomography approach an algorithm similar to that described above for the STF can be applied to compute receiver statics. In this case we gather the residual wavefields recorded at a given station as a vector  $\delta\mathbf{d} = \mathbf{d} - \mathbf{u}$  and minimize the energy in these residuals by scaling the observables to obtain  $\hat{\mathbf{d}} = r\mathbf{d}$ . Using the same procedure as above, we compute the value of  $r$  that minimizes the energy in  $\delta\hat{\mathbf{d}} = r\mathbf{d} - \mathbf{u}$  and find that

$$r = \frac{\mathbf{u}\mathbf{d}^*}{\mathbf{d}\mathbf{d}^*} \quad (37)$$

Note that a station correction determined in this fashion can correct not only for time delays (again being cognizant of potential cycle skips) but also for frequency dependent modifications to the waveforms such as resonances at stations located in basins.

#### **4 RELATING WAVEFORM TOMOGRAPHY TO RECEIVER FUNCTION MIGRATION AND TRAVEL TIME TOMOGRAPHY**

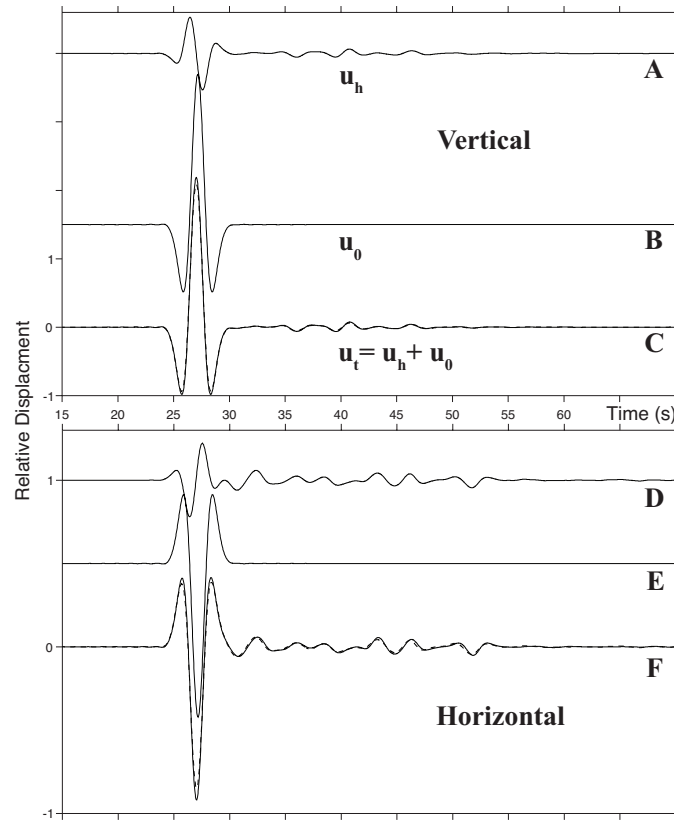
The waveform tomography approach described here can be expected to improve images of the subsurface by extracting more information from seismograms, but there is a trade off in an elevated sensitivity to noise and a nonlinearity that can make reaching a correct minimum through a series of linear steps



**Figure 3.** Model used to verify the computation of the response of heterogeneous media to plane wave excitation using equation (10). The trial medium (panel B) is a high wavespeed (8.4 km/s) slab embedded in a constant wavespeed (8 km/s) background. A scattered field is computed using sources located within the slab, shown in panel C for the real part of the source functions at 0.05 Hz. At points located far from the edges of the slab, the total wavefield should approximate the case of the 1D layer in panel A, for which analytical solutions are known. The trial plane wave arrives at a  $20^\circ$  angle of incidence, as shown in the inset in panel B. Closed triangles locate the point where the seismograms shown in Figure 4 are computed.

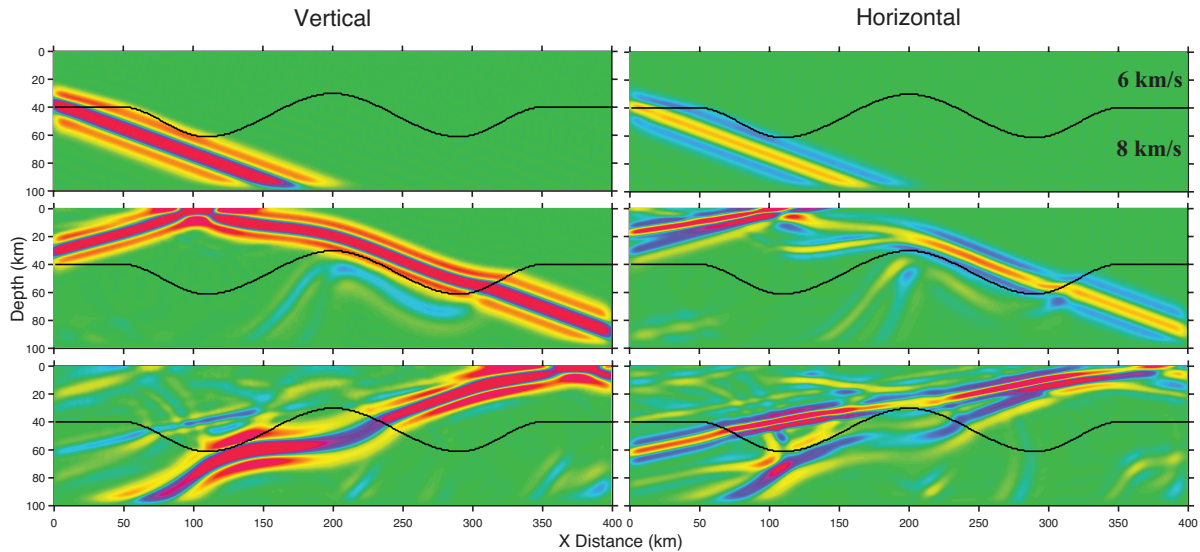
difficult. As mentioned above, a classic nonlinearity in spectral analysis is cycle skips; if we are more than  $\pi$  radians out of phase, most algorithms will shift the phase in the wrong direction. For this reason it is generally advisable to have a reasonable starting model using whatever information can be gleaned from other techniques. As mentioned at the outset, the images that come from travel time tomography and receiver function migration may be fuzzy or beset by ambiguities, but have an advantage in being relatively simple and robust so that investigators still can manage to recover some useful first order information about the medium.

In the event that background information or noise levels are inadequate to allow a full waveform tomography approach, one can still use parts of the algorithm described above to recover information about the medium in a manner analogous to travel time inversion and receiver function migration, but in ways that are arguably superior to either of these standard techniques. For example, one can generate a finite frequency travel time tomography image by normalizing amplitudes and solving equation (27) for phase shifts (i.e., time delays) after windowing the primary (P or S) phase. An analogous approach was employed by Maggi et al. (2009) and Tape et al. (2009). If a priori information about the subsurface is lacking, cycle skips can be reduced by fitting the low frequency end of the spectrum first before attempting a broad band inversion. We can also emulate receiver function migration by treating



**Figure 4.** Displacement seismograms computed at the free surface of the slab model (Figure 3) at  $x = 300$  km due to a plane wave entering the medium at a  $20^\circ$  angle of incidence. Trace “A” is the vertical component of ground motion ( $u_h$ ) due to the scattering sources located within the slab (panel C in Figure 3), while “B” is the displacement due to the constant wavespeed halfspace ( $u_0$ ). Traces “C” show the sum of the upper two seismograms ( $u_t = u_h + u_0$ ) plotted as a dashed line on a solid line representing the analytical solution for the 1D medium shown in panel A of Figure 3. Traces D, E, and F are the analogous results for the horizontal component of motion. Amplitudes of the horizontal components are magnified as indicated by the scales shown on the vertical axes.

the deconvolved waveform in (32) as the residual field ( $\delta d^*$ ) in (27) and (28), and backpropagating it into the medium from the stations. In this way, migration is equivalent to calculating the gradient in (27) without the additional step of solving for  $\delta m$ . An advantage of performing migration in this fashion is that we can more accurately account for the effects of lateral heterogeneity and multiple scattering than is possible with most standard techniques.



**Figure 5.** Snapshots of the wavefield generated by an incident plane P wave on a medium composed of a low wavespeed medium (6 km/s) over a high wavespeed medium (8 km/s) separated by an undulating surface indicated by the black line at about 40 km depth. The incident wave enters from the lower left corner at an angle of  $20^\circ$  from vertical. Left and right panels show, respectively, vertical and horizontal components of displacement. The top, middle, and bottom rows correspond to times of 11, 22, and 33 seconds, respectively, after the onset at the lower left corner at 100 km depth. Deep reds and blues correspond to large plus and minus amplitude and all displacements are plotted at the same scale.

## 5 EXAMPLES

In this section we discuss some simple examples that illustrate the basic concepts of teleseismic waveform tomography. The examples were generated using a FORTRAN code that implements the algorithms described above. Most of the computation was done on a Quad core AMD Opteron (2.2 GHz per core) with 16 Gb of RAM, but the examples were also run successfully on a MacBook Pro with a 2.5 GHz Intel Core Duo. The grid in the examples has 400 points in the  $x$  direction and 100 points in the  $z$  direction. The grid spacing is 1 km, which makes the model size about 400 km x 100 km and also sets the minimum wavelength that we can expect to model without significant numerical dispersion to about 4 km. A ten point (9 km) wide PML is included along the bottom and sides of the model. For reasons of stability, when we do the inversion we apply a mask to the PML and to the uppermost 2 grid points at the free surface, which keeps the wavespeeds near the boundaries constant.

Computation of the wavefield for a single frequency on the 400 x 100 grid required about 20 seconds on the Opteron and about 50 seconds on the MacBook Pro. Nominally a complete forward problem would require on the order of 40 frequencies, meaning that an entire suite of Green's functions can be determined for a moderately sized grid on a modest desktop computer in less than an hour.

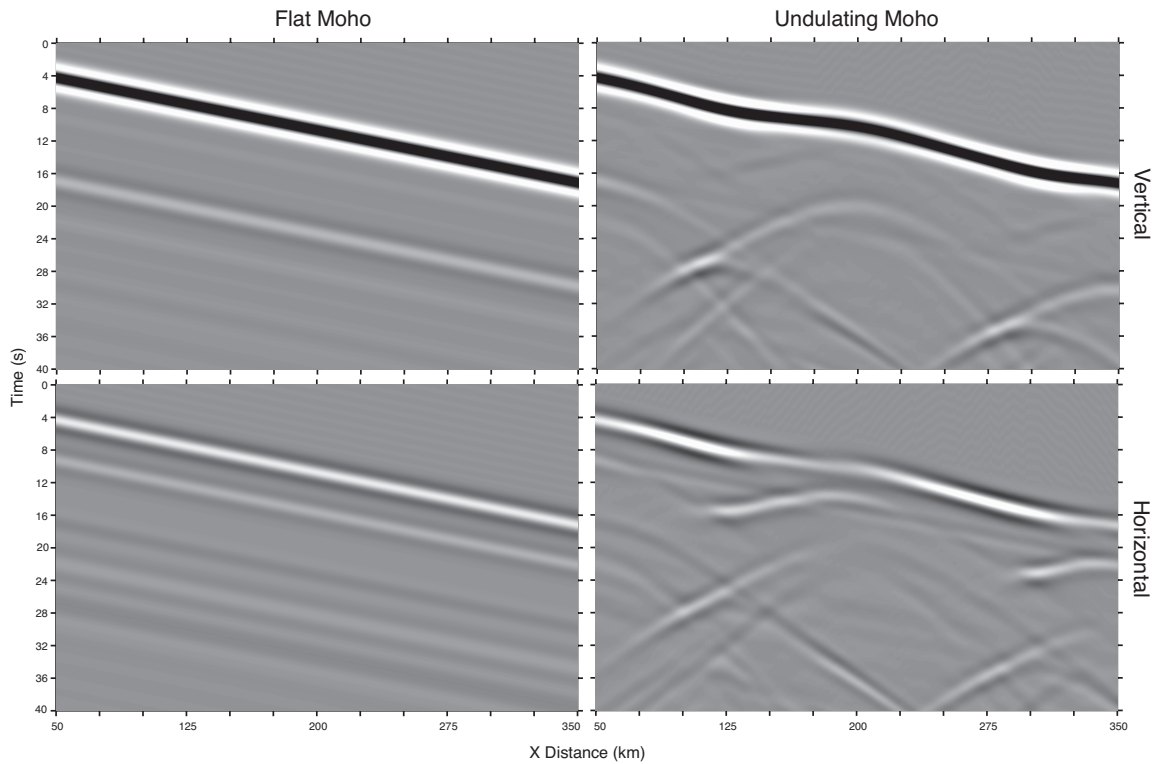
## 5.1 Waveform Synthetics

In the process of testing the code we generated waveforms for a wide range of trial models, and animations generated from several of these tests can be found online<sup>†</sup>. We review two examples here. The first tests the approach of equation (11) for computing the plane wave excitation of 2.5D heterogeneity. The heterogeneity in this case is a slab with wavespeeds 5% higher than those in a surrounding constant wavespeed background (Figure 3). While no analytical solution to the wave equation exists for this case, it should happen that the response at stations far from the edges of the slab should closely approximate the response of a 1D medium (i.e., an infinite slab), for which the response is known. We use equation (10) to compute a scattered wavefield ( $u_h$ ) from a distribution of sources in the slab (panel C in Figure 3) that can be added to the (known) half space response to recover an approximation to the layered medium response. The analytic and numerical waveforms (traces C and F in Figure 4) are quite similar, the main difference being a slight underestimation of the amplitude of the main phase caused by approximating a 1D layer by a bounded slab, which ignores the small but nonzero contributions of scattering sources that would extend out beyond the edges of the slab.

The second example demonstrates how the algorithm works in a model that has large wavespeed gradients in both horizontal and vertical directions. The choice of model dimension for this test was intended to represent the Tien Shan region of central Asia, where a significant amount of effort has already gone into generating travel time and receiver function images (e.g., Roecker et al., 1993; Vinnik et al., 2002; 2004; Li et al., 2009). One of the intriguing results of those investigations was evidence for an undulating Moho with remarkably large excursions from a mean depth of about 45 km (e.g., Oreshin et al., 2002). We reproduce this environment in a simplified manner here by specifying a constant wavespeed crust (6 km/s) over a constant wavespeed mantle (8 km/s) separated by a thin interface (Figure 5).

Wavefield snapshots (Figure 5) illustrate the response of this crustal/upper mantle model to an incident P wave approaching from the lower left at a 20° angle of incidence. Clearly visible in the panels are reflections from the curved parts of the Moho and focussing and defocussing caused by wavespeed changes and refractions on the boundary. SV mode conversions show up clearly on the horizontal components of motion (panels B and D in Figure 5). Multiply reflected phases entering from the stratified medium to the left of the modeled region are clearly seen at the left side of panels C and D. Record sections (Figure 6) computed at points at the surface illustrate the phenomenon of “ghost” phases described above, as the later arriving phases generated in the 1D background model (left side panels of Figure 6) appear as faint, but still discernible, phases in the 2.5D model. As discussed above,

<sup>†</sup> <http://gretchen.geo.rpi.edu/roecker/WT/ElasticFD.html>

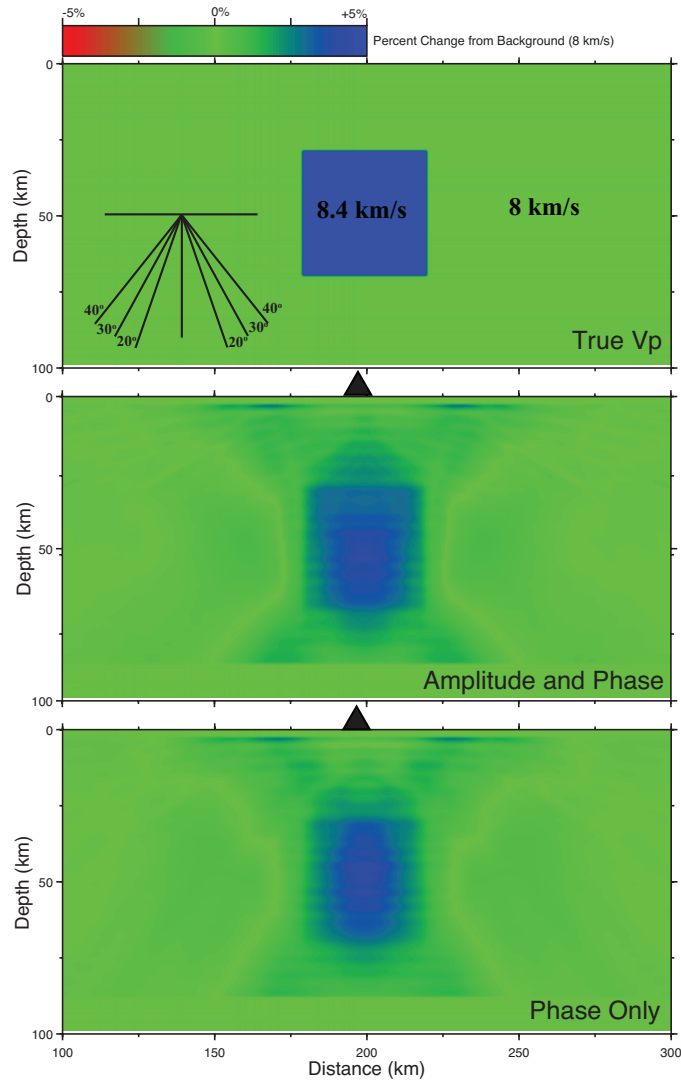


**Figure 6.** Record sections of displacement seismograms for the example shown in Figure 5. The panels on the left show the plane wave excitation for the case of a horizontal Moho at 40 km depth, which is also the background medium used to compute  $u_o$ . Panels on the right are the total field for the undulating Moho show in Figure 5. The upper and lower rows correspond to vertical and horizontal components of motion, respectively. Note the reduced, but still discernable, presence of the 1D background mode conversions (“ghost” phases) in the sections on the right.

it is difficult to eliminate these artifacts entirely, but we can significantly reduce their amplitude by displacing the 1D interface upwards by about half a grid spacing.

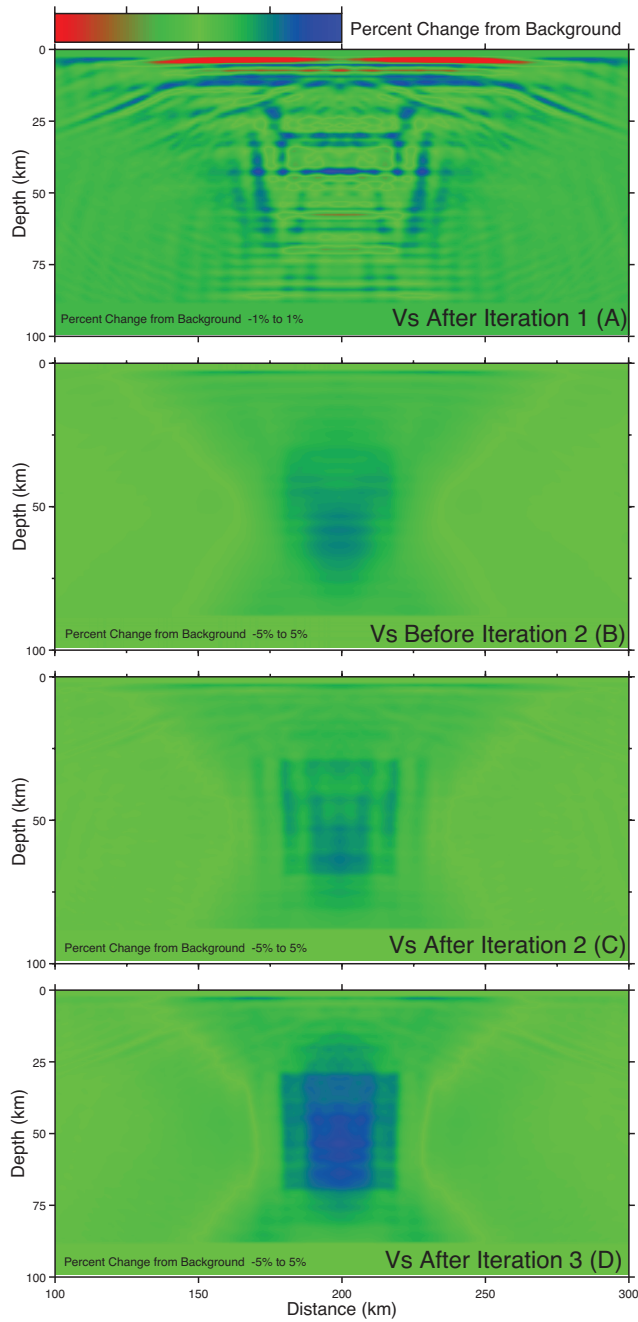
## 5.2 Tomography Example

To illustrate some of the key concepts of teleseismic waveform tomography we consider a simple model of a half space composed of a Poisson solid with constant 8 km/s compressional wavespeed in which a block 40 km on a side with wavespeeds 5% higher than the background is embedded (Figure 7). Synthetic seismograms are generated by calculating the responses to incident plane P waves from 6 events. Three of these are at a back azimuth of  $0^\circ$  (i.e., waves approaching from the right in Figure 7) and three are from  $180^\circ$  (waves approaching from the left) and with angles of incidence of  $20^\circ$ ,  $30^\circ$ , and  $40^\circ$  with respect to vertical. The source time function of the incident wave is a Küpper wavelet



**Figure 7.** Results of waveform tomography inversion for P wavespeed after three iterations. Only the middle 200 km of the total 400 km width of the model is shown here to enhance details in the image. The remainder of the model stays at the background wavespeed. *Top.* The target image. This is a 40 km x 40 km block with wavespeeds 5% higher than the 8 km/s (green) background. The diagram to the left of the block shows the angles of incidence for the 6 plane waves used to generate a synthetic data set. *Middle.* Reconstruction of the target using both amplitude and phase. *Bottom.* Reconstruction using only the phase. Receivers are located along the surface at a 1 km interval between  $x = 50$  and 350 km. The closed triangle locates the receiver at  $x = 190$  km where the seismograms in Figure 9 are calculated.

(Küpper, 1958) with a dominant frequency of 0.5 Hz and a bandwidth from 0.02 Hz to 1 Hz. The wavefield is recorded by 300 receivers placed 1 km apart along the surface between  $x = 50$  and 350 km. The starting model in the inversion is the constant background, and three iterations are performed to produce the final results shown in Figures 7 and 8.



**Figure 8.** Results of a waveform tomography inversion for S wavespeed. The target image is the same as that shown in Figure 7. Scaling of the palette is indicated at the lower left of each image. (A). Reconstruction after the first iteration starting with a constant wavespeed background. (B). Revised starting model for second iteration based on the P wavespeed model obtained after the first iteration. (C) & (D). Reconstruction after the second and third iterations.

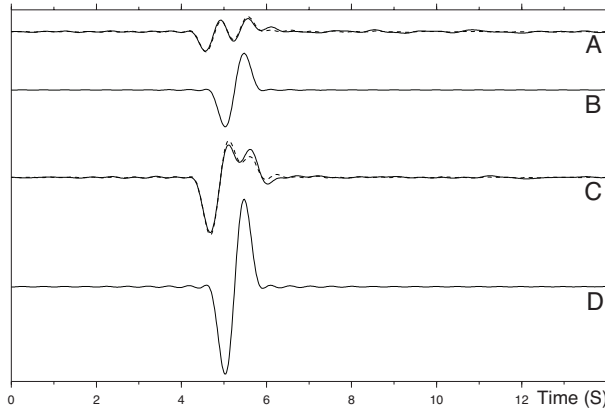
Besides being a common type of observation in body wave analysis, modeling an incident P wave and the scattered waves in its coda provides a good example of how waveform tomography incorporates aspects of both travel time tomography and receiver function migration in creating an image. Because the primary P arrival is much more energetic than the scattered phases, those parts of the Jacobian in (23) related to P wavespeeds will be sensitive primarily to the phase shifts caused by the P wavespeed anomaly. To see that this is the case, we compare an image reconstructed using the full waveform with one using only phase shifts. The results in Figure 7 show that while the edges on the full waveform image are somewhat better defined, the phase-only inversion does nearly as well in reconstructing the image.

In contrast to the P wavespeeds, there is very little information about S wavespeeds in the P coda, and the Ps conversions from the sides of the box are the dominant residual shear waves that the tomography algorithm migrates back into the medium. As shown in panel A in Figure 8, these residual waves converge in the general vicinity of the edges of the box, but focussing is poor due to the inaccuracy of the model used to perform the migration. Specifically, the sides of the box appear to dip inwards because the background is 5% slower than it should be. Note also that because *all* of the residuals are migrated, the misfit to the primary P phase shows up as large amplitude signals near the top of the model. If we tie the S wavespeeds to the imaged P wavespeeds as shown in Figure 8b, the migrated image is much better focussed, and on the third iteration the edges of the box are clearly seen even though the wavespeeds within the box are still not well resolved.

We note that while this test reconstructs what appears to be a fairly well resolved image with much less data (6 events and 1800 arrivals) than that usually available for tomographic imaging, we compensated for this lack of information by fixing the source time function and not contaminating the data with any but a small amount of numerical noise. The value of tests that quantify the influence of noise and geometry on an image usually depend strongly on the particular data being analyzed and so will wait for an investigation involving an actual dataset.

## 6 EXTENSIONS

This paper outlines a basic 2.5D full waveform teleseismic tomography algorithm, and the discussion has focussed on the case of heterogeneous, elastic media. There are additional techniques that can be adapted from the acoustic waveform tomography approach; for example, an extension from elastic to viscoelastic (i.e., finite Q) media can be achieved by using complex wavespeeds. Simple forms of vertical anisotropy can be included by stretching the coordinate system in the  $z$  direction. More sophisticated methodology like the receiver function approach of Levin and Park (1997) would be possible to employ in computing a background wavefield, but modeling propagation in a general heterogeneous,



**Figure 9.** Displacement seismograms for the station located at the position  $x = 190$  km in Figure 3 for the full wave inversion test. The plane wave source in this case has a  $20^\circ$  angle of incidence and is propagating in the  $+x$  direction. The uppermost traces (A) show the horizontal component of the recovered and true waveforms plotted as dashed and solid lines, respectively. The “recovered” component is computed using the model determined after three iterations (shown in figure 8d). The “true” waveform corresponds to the “true” model shown at the top of Figure 7. Trace (B) is the waveform in the starting (half space) model. The lower two sets of seismograms (C and D) show the equivalent waveforms for the vertical component. All amplitudes are scaled equally.

anisotropic media would most likely require significant computing resources. Topography or submarine environments that involve media with a solid-liquid interface can be modeled using a combination of elastic and acoustic equations when forming the difference matrix  $S$ ; methods for computing background fields in these cases are well known. For example, topography may be accounted for using a modified version of the transformed grid technique of Ruud and Hestholm (2001). Finally, teleseismic body waves could be combined with other types of waveforms in a joint inversion. An active source line parallel to the  $x$  axis would be the easiest to incorporate as the original RGP algorithm was designed for this type of application. Surface waves could also be included; as with body waves one need only be able to specify a known background wavefield. Local earthquakes could be included provided both sources and receivers are located within the  $(x, z)$  plane because the wavefield in this case is no longer invariant along the  $y$  direction. Nevertheless, one might argue that 2D sections of the 2.5D model still allow reasonable approximations of the wavefield.

Our intention in this study was to provide one example of a tractable inverse problem using gradients, but there are several other approaches that one could consider. In particular, Pratt et al. (1998) discuss several alternatives for the acoustic case that could be easily adapted to the visco-elastic teleseismic case. Also, where the location of specific boundaries is an important part of the inverse problem (for example, when focussing on receiver function migration) one might consider inverting for

interface topography instead of wavspeeds at individual grid points, as suggested for example by Dahlen (2005).

## ACKNOWLEDGMENTS

This work was supported by the National Science Foundation grants CMG-0327634 and EAR-0838384. We would like to acknowledge Carl Tape and an anonymous reviewer for providing comments that significantly improved the content of this paper. We benefitted from the availability of the source code for the acoustic wave algorithms *OMEGA* and *FULLWV* provided by R. G. Pratt, which greatly facilitated the development of our own code. We acknowledge and appreciate the many years of labor by a number of contributors that went into the development of those codes.

## REFERENCES

- Akçelik, V., G. Biros, and O. Ghattas, 2002. Parallel multiscale Gauss-Newton-Krylov methods for inverse wave propagation, *Proc. ACM/IEEE Supercomputing Conference (SC02)*, doi: 10.1109/SC.2002.10002, published on CD-ROM and at [www.sc-conference.org/sc2002](http://www.sc-conference.org/sc2002).
- Akçelik, V., G. Biros, and O. Ghattas, 2003. High resolution forward and inverse earthquake modeling on terascale computers, *Proc. ACM/IEEE Supercomputing Conference (SC03)*, doi: 10.1109/SC.2003.10056, published on CD-ROM and at [www.sc-conference.org/sc2003](http://www.sc-conference.org/sc2003).
- Aki, K., A. Christoffersson, and E. S. Husebye, 1977. Determination of the three-dimensional seismic structure of the lithosphere, *J. Geophys. Res.*, **82**, 277-296.
- Bleibinhaus, F., J. A. Hole, T. Ryberg, and G. S. Fuis, 2007. Structure of the California Coast Ranges and San Andreas Fault at SAFOD from seismic waveform inversion and reflection imaging. *J. Geophys. Res.*, **112**, B06315, doi:10.1029/2006JB004611.
- Bostock, M.G., Rondenay, S. and Shragge, J., 2001. Multiparameter 2-D inversion of scattered teleseismic body-waves. I. Theory for oblique incidence, *J. Geophys. Res.*, **106**, 30,771-30,782.
- Brenders, A., and Pratt, R. G., 2003. Waveform Tomography of a 2-D Full Wavefield, Elastic, Wide Angle, Synthetic Dataset , *12th Annual Commission on Controlled Source Seismology, Mountain Lake, Virginia*.
- Brenders, A., and Pratt, R. G., 2007. Full waveform tomography for lithospheric imaging; results from a blind test in a realistic crustal model, *Geophysical Journal International*, **168**, 133-151.
- Brittan, J., Forgues, E., Pratt, R., Morgan, J., Warner, M., Macintyre, H., and Marin, L., 1997. Wavefield inversion across the edge of the Chicxulub impact structure, *Journal of Conference Abstracts*, **1**, 65-72.
- Chen, P., L. Zhao, and T. H. Jordan, 2007. Full 3D tomography for the crustal structure of the Los Angeles region, *Bull. Seis. Soc. Am.*, **97** (4), 1094-1120.
- Collino, F. and C. Tsogka, 2001. Application of the PML absorbing layer model to the linear elastodynamic problem in anisotropic heterogeneous media, *Geophysics*, **66**, 294-307.

- Dahlen, F. A., 2005. Finite-frequency sensitivity kernels for boundary topography perturbations, *Geophys. J. Int.*, **162**, 525-540.
- George, A., and Lui, J. W. H., 1981. *Computer solutions of large positive definite systems*, Prentice-Hall Inc., Englewood Cliffs, NJ.
- Haskell, N. A., 1953. The dispersion of surface waves on multilayered media, *Bulletin of the Seismological Society of America*, **43** (1), 17-34.
- Haskell, N. A., 1960. Crustal reflection of plane SH waves, *Journal of Geophysical Research*, **65** (12), 4147-4150.
- Haskell, N. A., 1962. Crustal Reflection of Plane P and SV Waves, *Journal of Geophysical Research*, **67** (12), 4751-4767.
- Herrmann, R. B., C. J. Ammon, and J. Julia, 2001. Application of joint receiver-function surface-wave dispersion for local structure in Eurasia, *Proceedings of the 23rd seismic research review; worldwide monitoring of nuclear explosions*, **I**, 46-54.
- Hole, J. A., Zelt, C. A., and Pratt, R. G., 2005. Advances in controlled-source seismic imaging, *Eos Trans. Amer. Geophys. Union*, **86**, 177 - 181.
- Jo, C. H., Shin, C. S., and Suh, J. H., 1996. An optimal 9-point, finite-difference, frequency-space, 2-D scalar wave extrapolator, *Geophysics*, **61** (2), 529-537.
- Karypis, G., and V. Kumar, 1998. *METIS: A Software Package for Partitioning Unstructured Graphs, Partitioning Meshes, and Computing Fill-Reducing Orderings of Sparse Matrices, Version 4.0*, (Manual), University of Minnesota, Department of Computer Science / Army HPC Research Center, <http://www.cs.umn.edu/karypis>.
- Komatitsch, D., and J. Tromp, 2003. A perfectly matched layer absorbing boundary condition for the second-order seismic wave equation, *Geophysical Journal International*, **154** (1) , 146-153 doi:10.1046/j.1365-246X.2003.01950.x.
- Kosarev, G. L., N. V. Petersen, L. P. Vinnik, and S. W. Roecker, 1993. Receiver functions for the Tien Shan analog broadband network; contrasts in the evolution of structure across the Talasso-Fergano Fault *Journal of Geophysical Research*, **98**, 4437-4448.
- Kumar P., K. Wylegalla, C. Reigber, X. Yuan, I. Woelbern, P. Schwintzer, K. Fleming, T. Dahl-Jensen, T. B. Larsen, J. Schweitzer, K. Priestley, O. Gudmundsson, D. Wolf, R. Kind, W. Hanka, 2005. The lithosphere-aesthenosphere boundary in the North-West Atlantic region, *Earth and Planetary Science Letters*, **236**, 249-257.
- Küpper, F. J., Theoretische Untersuchung über die Mehrfachaufstellung von Geophonen, 1958. *Geophysical Prospecting*, DOI: 10.1111/j.1365-2478.1958.tb01652.x, 1958.
- Lailly, P., 1983. The seismic inverse problem as a sequence of before stack migrations, *SIAM Philadelphia, Conference on Inverse Scattering: Theory and Application*, Edited by J B Bednar, R Redner, E Robinson and A Weglein, 206-220.
- Langston, C. A., 1979. Structure under Mount Rainier, Washington, inferred from teleseismic body waves. *J. Geophys. Res.*, **84**, 4749-4762.

- Levander, A., F. Niu, and W. W. Symes, 2005. Imaging teleseismic P to S scattered waves using the Kirchhoff integral, *Geophysical Monograph*, **157**, 149-169 Published by: American Geophysical Union : Washington, DC, United States.
- Levin V., and J. Park, 1997. P-SH conversions in a flat-layered medium with anisotropy of arbitrary orientation, *Geophysical Journal International*, **131**, 253-266.
- Li, Z., and S. Roecker, 2007. Tomographic Imaging of the Crust and Upper Mantle Beneath the Western Tien Shan, *Eos Trans. AGU*, **88**(52), Fall Meet. Suppl., Abstract T31B-0472.
- Mathews, J. H., and Fink, K. D., 2004. *Numerical Methods: Using Matlab, Fourth Edition*, Prentice-Hall Pub..
- Mora, P. R., 1987. Nonlinear two-dimensional elastic inversion of multi-offset seismic data, *Geophysics*, **52**, 1211-1228.
- Oreshin S., L. Vinnik, D. Peregoudov, S. Roecker, 2002. Lithosphere and asthenosphere of the Tien Shan imaged by S receiver functions, *Geophys. Res. Lett.*, **29**(8), doi: 1029/2001GLO14441.
- Poppeliers, C., and G. L. Pavlis, 2003a. Three-dimensional, prestack, plane wave migration of teleseismic P-to-S converted phases:1. Theory, *Journal of Geophysical Research*, **108**(B2), doi:10.1029/2001JB000216.
- Poppeliers, C., and G. L. Pavlis, 2003b. Three-dimensional, prestack, plane wave migration of teleseismic P-to-S converted phases; 2, Stacking multiple events , *Journal of Geophysical Research*, **108**(B5), doi:10.1029/2001JB001583.
- Pratt, R. G., and Worthington, M. H., 1990. Inverse theory applied to multi-source crosshole tomography. Part I: Acoustic wave-equation method, *Geophysical Prospecting*, **38**, 287-310.
- Pratt, R. G., Shin, C., and Hicks, G., 1998. Gauss-newton and full newton methods in frequency-space seismic waveform inversion, *Geophys. J. Int.*, **133**, 341-362.
- Pratt, R. G., 1990. Frequency domain elastic wave modeling by finite differences: A tool for cross-hole seismic imaging, *Geophysics*, **55**, 626-632.
- Pratt, R. G., 1999. Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model, *Geophysics*, **64**, 888-901.
- Pratt, R. G., 2003. *Waveform Tomography: theory and practice*, available at: <http://www.geophys.geos.vt.edu/hole/ccss/prattCCSS.pdf>.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992. *Numerical recipes in FORTRAN (2nd ed.): the art of scientific computing*, Cambridge University Press, 994 pp.
- Roecker, S.W., Sabitova, T. M., Vinnik, L. P., Burmakov, Y. A., Golvanov, M. I., Mamatkanova, R., and Munirova, L., 1993. Three-dimensional elastic wave velocity structure of the western and central Tien Shan, *J. Geophys. Res.*, **98**, 15779-15795.
- Ruud, B., and S. Hestholm, 2001. 2D surface topography boundary conditions in seismic wave modelling, *Geophysical Prospecting*, **49**, 445-460.
- Schneider, J. B., 2004. Plane Wave in FDTD Simulations and a Nearly Perfect Total-Field/Scattered-Field Boundary, *IEEE Trans. Antennas Propagat.*, **48**, 494-501.
- Song, Z. M., and Williamson, P. R., 1995. Frequency-domain acoustic wave modeling and inversion of cross-

- hole data: Part I - 2.5-D modeling method, *Geophysics*, **60**, 784-795.
- Song, Z. M., Williamson, P. R., and Pratt, R. G., 1995b. Frequency-domain acoustic wave modeling and inversion of cross-hole data: Part II - Inversion method, synthetic experiments and real data results, *Geophysics*, **60**, 796-809.
- Štekl, I., and R. G. Pratt, 1998. Accurate visco-elastic modeling by frequency-domain finite differences using rotated operators, *Geophysics*, **63**, 1179-1794.
- Takenaka, H., and B. L. N. Kennett, 1996. A 2.5-D time-domain elastodynamic equation for plane-wave incidence, *Geophysical Journal International*, **125**, F5-F9.
- Tan, T., and M. Potter, 2006. Perfect Plane Wave Source for Total-Field/Scattered-Field Formulation in FDTD using Time Domain Multipoint 1D Auxiliary Grid, *Antennas and Propagation Society International Symposium 2006, IEEE*, 3777-3780.
- Tape, C., Q. Liu, A. Maggi, and J. Tromp, 2009. Adjoint tomography of the southern California crust, *Science*, **325**, 988992.
- Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**, 1259-1266.
- Vinnik, L., 1977. Detection of waves converted from P to SV in the mantle. *Phys. Earth Planet. Inter.*, **15**, 39-45.
- Vinnik, L. P., D. Peregoudov, L. Makeyeva, S. Oreshin, and S. Roecker, 2002a. Towards 3-D fabric in the continental lithosphere and asthenosphere: the Tien Shan, *Geophys. Res. Lett.*, 10.1029/2001GL014588.
- Vinnik, L. P., S.W. Roecker, G. L. Kosarev, S.I. Oreshin, and I. Yu. Koulakov, 2002b. Crustal Structure and Dynamics of the Tien Shan, *Geophys. Res. Lett.*, **29**, 41-44.
- Vinnik, L. P., Ch. Reigber, I. M. Aleshin, G. L. Kosarev, M. K. Kaban, S. I. Oreshin, and S. Roecker, 2004. Receiver function tomography of the central Tien Shan, *Earth Planetary Science Letters*, **225**, 131-146.
- Williamson, P. W., and Pratt, R. G., 1995. A critical review of acoustic wave modelling procedures in 2.5 D, *Geophysics*, **60**, 591-595.
- Wilson, D., and R. Aster, 2003. Seismic structure of the lithosphere in the Southwestern United States using receiver functions, *Leading Edge*, **22**, 234-237.
- Woodward, M. J., 1992. Wave-equation tomography, *Geophysics*, **57**, 15-26.
- Yang, D. H., E. Liu, Z. J. Zhang, and J. Teng, 2002. Finite-difference modelling in two-dimensional anisotropic media using a flux-corrected transport technique, *Geophys. J. Int.*, **148**, 320-328.
- Yang, X., G. L., Pavlis, S. W. Roecker, and F. L. Vernon, 2003. Teleseismic tomographic images of the central Tien Shan, *Eos, Transactions, American Geophysical Union*, 2003, 84, **46**, Suppl., pp. 1057.
- Zheng, Y., and X. Huang, 2002. Anisotropic perfectly matched layers for elastic waves in Cartesian and curvilinear coordinates, in: *Earth Resources Laboratory 2002 Industry Consortium Meeting*, Dept. of Earth, Atmospheric, and Planetary sciences, Massachusetts Institute of Technology, Cambridge, MA, USA.

**APPENDIX A: FINITE DIFFERENCE FORMS**

Our finite difference forms are defined on a nine-point star shown diagrammatically as

$$\begin{array}{ccc} \mu^- & \mu_- & \mu_-^+ \\ \mu^- & \mu & \mu^+ \\ \mu_+^- & \mu_+ & \mu_+^+ \end{array}$$

We also define average medium properties between these points as (in this case for  $\mu$ )

$$\begin{array}{ccccc} \mu^- & \mu_-^{-1/2} & \mu_- & \mu_-^{+1/2} & \mu_-^+ \\ \mu_-^{-1/2} & \mu_-^{-1/2} & \mu_{-1/2} & \mu_{-1/2}^{+1/2} & \mu_{-1/2}^+ \\ \mu^- & \mu^{-1/2} & \mu & \mu^{+1/2} & \mu^+ \\ \mu_{+1/2}^- & \mu_{+1/2}^{-1/2} & \mu_{+1/2} & \mu_{+1/2}^{+1/2} & \mu_{+1/2}^+ \\ \mu_+^- & \mu_+^{-1/2} & \mu_+ & \mu_+^{+1/2} & \mu_+^+ \end{array}$$

where, for example,  $\mu_{+1/2}^{+1/2} = (\mu_+ + \mu_+^+) / 2$ . The general forms of the stencils used in the finite difference formulation in the non-rotated frame are

$$\begin{aligned} \mu \frac{\partial u}{\partial x} &= [\mu^{+1/2}(u^+ - u) + \mu^{-1/2}(u - u^-)] / 2\Delta \\ \mu \frac{\partial u}{\partial z} &= [\mu_{+1/2}(u_+ - u) + \mu_{-1/2}(u - u_-)] / 2\Delta \\ \frac{\partial}{\partial x} [\mu u] &= [\mu^{+1/2}(u^+ + u) - \mu^{-1/2}(u^- + u)] / 2\Delta \\ \frac{\partial}{\partial z} [\mu u] &= [\mu_{+1/2}(u_+ + u) - \mu_{-1/2}(u_- + u)] / 2\Delta \\ \frac{\partial}{\partial x} \left[ \mu \frac{\partial u}{\partial x} \right] &= [\mu^{+1/2}(u^+ - u) - \mu^{-1/2}(u - u^-)] / \Delta^2 \\ \frac{\partial}{\partial z} \left[ \mu \frac{\partial u}{\partial z} \right] &= [\mu_{+1/2}(u_+ - u) - \mu_{-1/2}(u - u_-)] / \Delta^2 \\ \frac{\partial}{\partial x} \left[ \mu \frac{\partial u}{\partial z} \right] &= [\mu^+(u_+^+ - u_-^+) - \mu^-(u_+^- - u_-^-)] / 4\Delta^2 \\ \frac{\partial}{\partial z} \left[ \mu \frac{\partial u}{\partial x} \right] &= [\mu_+(u_+^+ - u_+^-) - \mu_-(u_-^+ - u_-^-)] / 4\Delta^2 \end{aligned}$$

and in the 45° rotated frame

$$\begin{aligned}
 \mu \frac{\partial u}{\partial x'} &= \left[ \mu_{+1/2}^{+1/2} (u_+^+ - u) + \mu_{-1/2}^{-1/2} (u - u_-^-) \right] / 2\sqrt{2}\Delta \\
 \mu \frac{\partial u}{\partial z'} &= \left[ \mu_{+1/2}^{-1/2} (u_+^- - u) + \mu_{-1/2}^{+1/2} (u - u_-^+) \right] / 2\sqrt{2}\Delta \\
 \frac{\partial}{\partial x'} [\mu u] &= \left[ \mu_{+1/2}^{+1/2} (u_+^+ + u) - \mu_{-1/2}^{-1/2} (u_-^- + u) \right] / 2\sqrt{2}\Delta \\
 \frac{\partial}{\partial z'} [\mu u] &= \left[ \mu_{+1/2}^{-1/2} (u_+^- + u) - \mu_{-1/2}^{+1/2} (u_-^+ + u) \right] / 2\sqrt{2}\Delta \\
 \frac{\partial}{\partial x'} \left[ \mu \frac{\partial u}{\partial x'} \right] &= \left[ \mu_{+1/2}^{+1/2} (u_+^+ - u) - \mu_{-1/2}^{-1/2} (u - u_-^-) \right] / 2\Delta^2 \\
 \frac{\partial}{\partial z'} \left[ \mu \frac{\partial u}{\partial z'} \right] &= \left[ \mu_{+1/2}^{-1/2} (u_+^- - u) - \mu_{-1/2}^{+1/2} (u - u_-^+) \right] / 2\Delta^2 \\
 \frac{\partial}{\partial x'} \left[ \mu \frac{\partial u}{\partial z'} \right] &= \left[ \mu_{+1/2}^{+1/2} (u_+^+ - u_-^+) - \mu_{-1/2}^{-1/2} (u_-^- - u_+^-) \right] / 2\Delta^2 \\
 \frac{\partial}{\partial z'} \left[ \mu \frac{\partial u}{\partial x'} \right] &= \left[ \mu_{+1/2}^{-1/2} (u_+^- - u_-^-) - \mu_{-1/2}^{+1/2} (u_-^+ - u_+^+) \right] / 2\Delta^2
 \end{aligned}$$

When computing the optimizing weights of the sum of the two frames to minimize numerical dispersion (Appendix B), we use homogeneous versions for the non-rotated frame

$$\begin{aligned}
 \frac{\partial u}{\partial x} &= \frac{u^+ - u^-}{2\Delta} \\
 \frac{\partial u}{\partial z} &= \frac{u_+ - u_-}{2\Delta} \\
 \frac{\partial^2 u}{\partial x^2} &= \frac{u^+ - 2u + u^-}{\Delta^2} \\
 \frac{\partial^2 u}{\partial z^2} &= \frac{u_+ - 2u + u_-}{\Delta^2} \\
 \frac{\partial^2 u}{\partial x \partial z} &= \frac{u_+^+ - u_-^+ - u_+^- + u_-^-}{4\Delta^2}
 \end{aligned}$$

and in the 45° rotated frame

$$\begin{aligned}
 \frac{\partial u}{\partial x'} &= \frac{u_+^+ - u_-^-}{2\sqrt{2}\Delta} \\
 \frac{\partial u}{\partial z'} &= \frac{u_+^- - u_-^+}{2\sqrt{2}\Delta} \\
 \frac{\partial^2 u}{\partial x'^2} &= \frac{u_+^+ - 2u + u_-^-}{2\Delta^2} \\
 \frac{\partial^2 u}{\partial z'^2} &= \frac{u_+^- - 2u + u_-^+}{2\Delta^2} \\
 \frac{\partial^2 u}{\partial x' \partial z'} &= \frac{u_+^+ - u_-^+ - u_+^- + u_-^-}{2\Delta^2}
 \end{aligned}$$

## APPENDIX B: OPTIMIZATION OF GRIDS FOR NUMERICAL DISPERSION IN THE 2.5D ELASTIC CASE

Following Jo et al. (1996) and Štekl and Pratt (1998) we use a weighted sum of rotated grids to minimize the effects of numerical dispersion. We quantify these effects by determining the difference between true compressional ( $\alpha$ ) and shear ( $\beta$ ) velocities in the medium and the associated phase ( $\frac{\omega}{k}$ ) and group ( $\frac{d\omega}{dk}$ ) velocities predicted by the finite difference forms. We define a set of functionals

$$E = \max \left[ 1 - \frac{v_{grid}}{v_{true}} \right]^2 \quad (\text{B.1})$$

and adjust the weights to minimize an integral of these functionals

$$\frac{1}{2\pi} \int_0^{2\pi/G} \int_0^{\pi/4} \max\{E_{ph}^C, E_{ph}^{S_1}, E_{ph}^{S_2}, E_{gr}^C, E_{gr}^{S_1}, E_{gr}^{S_2}\} d\theta dk \quad (\text{B.2})$$

where  $k^* = \frac{k\Delta}{2\pi} = \frac{1}{G}$ ,  $\Delta$  is the grid spacing,  $\theta$  is the rotation angle, and  $E_{ph}^P$ , etc. refers to the functionals defined above evaluated using the phase (ph) or group (gr) velocities for three phases, one compressional ( $C$ ) and two shear ( $S_1$  and  $S_2$ ), that in the 2D case correspond to P, SV and SH. We start with the 3D elastic wave equation in homogeneous media

$$-\omega^2 \rho u_i = (\lambda + \mu) \frac{\partial}{\partial x_i} \frac{\partial u_j}{\partial x_j} + \mu \frac{\partial^2 u_i}{\partial^2 x_j} \quad (\text{B.3})$$

which in 2.5D becomes

$$\omega^2 (\beta^2 p_y^2 - 1) u = \alpha^2 \frac{\partial^2 u}{\partial x^2} + \beta^2 \frac{\partial^2 u}{\partial z^2} + (\alpha^2 - \beta^2) \frac{\partial^2 w}{\partial x \partial z} - i\omega p_y (\alpha^2 - \beta^2) \frac{\partial v}{\partial x} \quad (\text{B.4})$$

$$\omega^2 (\alpha^2 p_y^2 - 1) v = \beta^2 \left( \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial z^2} \right) - i\omega p_y (\alpha^2 - \beta^2) \frac{\partial u}{\partial x} - i\omega p_y (\alpha^2 - \beta^2) \frac{\partial w}{\partial z} \quad (\text{B.5})$$

$$\omega^2 (\beta^2 p_y^2 - 1) w = \alpha^2 \frac{\partial^2 w}{\partial z^2} + \beta^2 \frac{\partial^2 w}{\partial x^2} + (\alpha^2 - \beta^2) \frac{\partial^2 u}{\partial x \partial z} - i\omega p_y (\alpha^2 - \beta^2) \frac{\partial v}{\partial z} \quad (\text{B.6})$$

We write these equations in the 45° rotated frame as

$$\begin{aligned} \omega^2 (\beta^2 p_y^2 - 1) u &= \frac{(\alpha^2 + \beta^2)}{2} \left( \frac{\partial^2 u}{\partial x'^2} + \frac{\partial^2 u}{\partial z'^2} \right) \\ &- (\alpha^2 - \beta^2) \frac{\partial^2 u}{\partial x' \partial z'} + \frac{(\alpha^2 - \beta^2)}{2} \left( \frac{\partial^2 w}{\partial x'^2} - \frac{\partial^2 w}{\partial z'^2} \right) - i\omega p_y \frac{(\alpha^2 - \beta^2)}{\sqrt{2}} \left( \frac{\partial v}{\partial x'} - \frac{\partial v}{\partial z'} \right) \end{aligned} \quad (\text{B.7a})$$

$$\begin{aligned} \omega^2 (\alpha^2 p_y^2 - 1) v &= \beta^2 \left( \frac{\partial^2 v}{\partial x'^2} + \frac{\partial^2 v}{\partial z'^2} \right) \\ &- i\omega p_y \frac{(\alpha^2 - \beta^2)}{\sqrt{2}} \left( \frac{\partial u}{\partial x'} - \frac{\partial u}{\partial z'} \right) - i\omega p_y \frac{(\alpha^2 - \beta^2)}{\sqrt{2}} \left( \frac{\partial w}{\partial x'} + \frac{\partial w}{\partial z'} \right) \end{aligned} \quad (\text{B.7b})$$

$$\begin{aligned} \omega^2(\beta^2 p_y^2 - 1)w &= \frac{(\alpha^2 + \beta^2)}{2} \left( \frac{\partial^2 w}{\partial x'^2} + \frac{\partial^2 w}{\partial z'^2} \right) \\ &+ (\alpha^2 - \beta^2) \frac{\partial^2 w}{\partial x' \partial z'} + \frac{(\alpha^2 - \beta^2)}{2} \left( \frac{\partial^2 u}{\partial x'^2} - \frac{\partial^2 u}{\partial z'^2} \right) - i\omega p_y \frac{(\alpha^2 - \beta^2)}{\sqrt{2}} \left( \frac{\partial v}{\partial x'} + \frac{\partial v}{\partial z'} \right) \end{aligned} \quad (\text{B.7c})$$

and combine the two systems as a weighted sum

$$(R^2 p_y^2 - \frac{1}{\alpha^2})\omega^2 u = aA + (1-a)A' \quad (\text{B.8a})$$

$$(p_y^2 - \frac{1}{\alpha^2})\omega^2 v = aB + (1-a)B' \quad (\text{B.8b})$$

$$(R^2 p_y^2 - \frac{1}{\alpha^2})\omega^2 w = aC + (1-a)C' \quad (\text{B.8c})$$

where the primes refer to the 45° rotated frame and  $R^2 = \frac{\beta^2}{\alpha^2}$ . We also include mass lumping as

$$\hat{u}_i = bu_{i-}^+ + \frac{(1-b)}{4}(u_i^+ + u_i^- + u_{i+} + u_{i-}) \quad (\text{B.9})$$

where the superscripts and subscripts refer to the elements of the nine point star as discussed in Appendix A. The objective of the dispersion analysis is to determine the values of  $a$  in (B.8) and  $b$  in (B.9) that minimize the integral in (B.2).

We proceed with a von Neumann analysis by introducing a plane wave solution  $(u_0, v_0, w_0)e^{-i(k_x x + k_y y + k_z z)}$ , where we recall that for 2.5D  $k_y = \omega p_y$ . This plane wave is expressed in the finite difference stencils of Appendix A by, for example

$$f_{-}^+ = A(\omega)e^{i(\omega t - \omega p_y y - k_x(x+\Delta) - k_z(z-\Delta))} = f e^{-i(k_x - k_z)\Delta} \quad (\text{B.10})$$

and similarly for the remaining points in the nine point star. Then defining  $T_1 = (R^2 p_y^2 - \frac{1}{\alpha^2})$  and  $T_2 = (p_y^2 - \frac{1}{\alpha^2})$  and collecting the finite difference forms we can write the combined wave equation (B.8) as

$$\omega^2 \begin{bmatrix} T_1 M_1 & 0 & 0 \\ 0 & T_2 M_1 & 0 \\ 0 & 0 & T_1 M_1 \end{bmatrix} \begin{bmatrix} u_0 \\ v_0 \\ w_0 \end{bmatrix} = \frac{1}{\Delta^2} \begin{bmatrix} K_{11} & \omega p_y K_{21} \Delta & K_{31} \\ \omega p_y K_{21} \Delta & K_{22} & \omega p_y K_{32} \Delta \\ K_{31} & \omega p_y K_{32} \Delta & K_{33} \end{bmatrix} \begin{bmatrix} u_0 \\ v_0 \\ w_0 \end{bmatrix} \quad (\text{B.11})$$

where

$$\begin{aligned} M_1 &= b + \frac{1-b}{4}(2 \cos k_x \Delta + 2 \cos k_z \Delta) \\ &= b + \frac{(1-b)}{2}(\cos(k\Delta \cos \theta) + \cos(k\Delta \sin \theta)) \end{aligned}$$

$$\begin{aligned}
K_{11} &= a \left[ \frac{\partial^2 u}{\partial x^2} + R^2 \frac{\partial^2 u}{\partial z^2} \right] + (1-a) \left[ \frac{(1+R^2)}{2} \left( \frac{\partial^2 u}{\partial x'^2} + \frac{\partial^2 u}{\partial z'^2} \right) - (1-R^2) \frac{\partial^2 u}{\partial x' \partial z'} \right] \\
&= 2a [\cos(k\Delta \cos \theta) - 1 + R^2(\cos(k\Delta \sin \theta) - 1)] \\
&\quad + (1-a)[(1+R^2)(\cos(k\Delta \cos \theta) \cos(k\Delta \sin \theta) - 1) - (1-R^2)(\cos(k\Delta \sin \theta) - \cos(k\Delta \cos \theta))] \\
K_{21} &= -i(1-R^2) \left[ a \frac{dv}{dx} - \frac{(1-a)}{\sqrt{2}} \left( \frac{dv}{dx'} - \frac{dv}{dz'} \right) \right] \\
&= -(1-R^2) [a \sin(k\Delta \cos \theta) + (1-a) \sin(k\Delta \cos \theta) \cos(k\Delta \sin \theta)] \\
K_{31} &= a(1-R^2) \frac{\partial^2 w}{\partial x \partial z} + (1-a) \frac{1-R^2}{2} \left( \frac{\partial^2 w}{\partial x'} - \frac{\partial^2 w}{\partial z'^2} \right) \\
&= -(1-R^2) \sin(k\Delta \cos \theta) \sin(k\Delta \sin \theta) \\
K_{22} &= R^2 \left[ a \left( \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial z^2} \right) + (1-a) \left( \frac{\partial^2 v}{\partial x'^2} + \frac{\partial^2 v}{\partial z'^2} \right) \right] \\
&= 2R^2 [a(\cos(k\Delta \cos \theta) + \cos(k\Delta \sin \theta) - 2) + (1-a)(\cos(k\Delta \cos \theta) \cos(k\Delta \sin \theta) - 1)] \\
K_{32} &= -i(1-R^2) \left( a \frac{\partial w}{\partial z} + \frac{(1-a)}{\sqrt{2}} \left( \frac{\partial w}{\partial x'} + \frac{\partial w}{\partial z'} \right) \right) \\
&= -(1-R^2) [a \sin(k\Delta \sin \theta) - (1-a) \sin(k\Delta \sin \theta) \cos(k\Delta \cos \theta)] \\
K_{33} &= a \left( \frac{\partial^2 w}{\partial z^2} + R^2 \frac{\partial^2 w}{\partial x^2} \right) + (1-a) \left( \frac{(1+R^2)}{2} \left( \frac{\partial^2 w}{\partial x'^2} + \frac{\partial^2 w}{\partial z'^2} \right) + (1-R^2) \frac{\partial^2 w}{\partial x' \partial z'} \right) \\
&= 2a [\cos(k\Delta \sin \theta) - 1 + R^2(\cos(k\Delta \cos \theta) - 1)] \\
&\quad + (1-a)[(1+R^2)(\cos(k\Delta \cos \theta)(k\Delta \sin \theta) - 1) \\
&\quad + (1-R^2)(\cos(k\Delta \sin \theta) - \cos(k\Delta \cos \theta))]
\end{aligned}$$

Identifying the eigenvalue problem for  $\omega\Delta$  in (B.11) gives us the form

$$\begin{aligned}
&\omega^6 T_1^2 T_2 M_1^3 - \frac{\omega^4}{\Delta^2} T_1 T_2 M_1^2 (K_{11} + K_{33}) - \frac{\omega^4}{\Delta^2} T_1^2 M_1^2 K_{22} - \frac{\omega^4}{\Delta^2} p_y^2 T_1 M_1 (K_{21}^2 + K_{32}^2) \\
&\quad + \frac{\omega^2}{\Delta^4} T_1 M_1 (K_{11} K_{22} + K_{22} K_{33}) + \frac{\omega^2}{\Delta^4} T_2 M_1 (K_{11} K_{33} - K_{31}^2) \\
&\quad + \frac{\omega^2}{\Delta^4} p_y^2 (K_{32}^2 K_{11} + K_{21}^2 K_{33} - 2K_{31} K_{21} K_{32}) \\
&\quad + \frac{1}{\Delta^6} K_{31}^2 K_{22} - K_{11} K_{22} K_{33} = 0
\end{aligned}$$

or rearranging

$$\begin{aligned}
&(\omega\Delta)^6 - (\omega\Delta)^4 \left[ \frac{(K_{11} + K_{33})}{T_1 M_1} + \frac{K_{22}}{T_2 M_1} + p_y^2 \frac{(K_{21}^2 + K_{32}^2)}{T_1 T_2 M_1^2} \right] \\
&\quad + (\omega\Delta)^2 \left[ \frac{(K_{11} K_{22} + K_{22} K_{33})}{T_1 T_2 M_1^2} + \frac{(K_{11} K_{33} - K_{31}^2)}{T_1^2 M_1^2} + p_y^2 \frac{(K_{32}^2 K_{11} + K_{21}^2 K_{33} - 2K_{31} K_{21} K_{32})}{T_1^2 T_2 M_1^3} \right] \\
&\quad + \frac{K_{31}^2 K_{22} - K_{11} K_{22} K_{33}}{T_1^2 T_2 M_1^3} = 0
\end{aligned}$$

We now have a problem of the form  $x^3 + ax^2 + bx + c = 0$  where  $x = (\omega\Delta)^2$ . Following Press et al. (1992), to solve this cubic with real coefficients we first let

$$Q = \frac{a^2 - 3b}{9}; \quad R = \frac{2a^3 - 9ab + 27c}{54} \quad (\text{B.12})$$

Since  $Q$  and  $R$  will in this case always be real and  $R^2 < Q^3$  the three real roots in this case are given by<sup>‡</sup>

$$x_1 = -2\sqrt{Q} \cos\left(\frac{\theta}{3}\right) - \frac{a}{3} \quad (\text{B.13a})$$

$$x_2 = 2\sqrt{Q} \cos\left(\frac{\theta + \pi}{3}\right) - \frac{a}{3} \quad (\text{B.13b})$$

$$x_3 = 2\sqrt{Q} \cos\left(\frac{\theta - \pi}{3}\right) - \frac{a}{3} \quad (\text{B.13c})$$

where we have defined

$$\theta \equiv \cos^{-1}\left(\frac{R}{\sqrt{Q^3}}\right) \quad (\text{B.14})$$

Evaluation of these roots indicates that  $\theta$  is close to  $\pi$  which makes  $x_1$  and  $x_2$  approximately equal and always less than  $x_3$ . Hence we associate the two lower frequency components with shear wavespeeds ( $x_1 \Rightarrow S_1, x_2 \Rightarrow S_2$ ) and the higher frequency  $x_3$  with compressional wavespeed ( $x_3 \Rightarrow C$ ). In the 2D case,  $\theta = \pi$  and  $x_1 = x_2$ . The fact that they are different when  $p_y \neq 0$  implies a directional dependence, or numerical anisotropy, for shear waves. Recalling that  $x_i = (\omega_i\Delta)^2$ , the numerical phase velocities are

$$v_{ph_i} = \frac{\omega_i}{k} = \frac{\sqrt{x_i}}{k\Delta} \quad (\text{B.15})$$

and group velocities are

$$v_{gr_i} = \frac{d\omega_i}{dk} = \frac{1}{2\sqrt{x_i}\Delta} \Delta \frac{dx_i}{dk} = \frac{1}{2\sqrt{x_i}} \frac{dx_i}{dk} \quad (\text{B.16})$$

## B1 Numerical Integration

To approximate the integral in B.2 we use a 2D Simpson's rule described in Mathews and Fink (2004) as

$$\int \int_R f(x, y) dA = \int_a^b \int_c^d f(x, y) dy dx \approx \text{S2D}(f, h, k)$$

<sup>‡</sup> Note that these formulas differ from those prescribed by Press et al. (1992) by a minus sign and the  $\pi$  factor in the argument of the cosine functions for  $x_2$  and  $x_3$ . We discovered this error only after a significant amount of aspirin and hair loss.

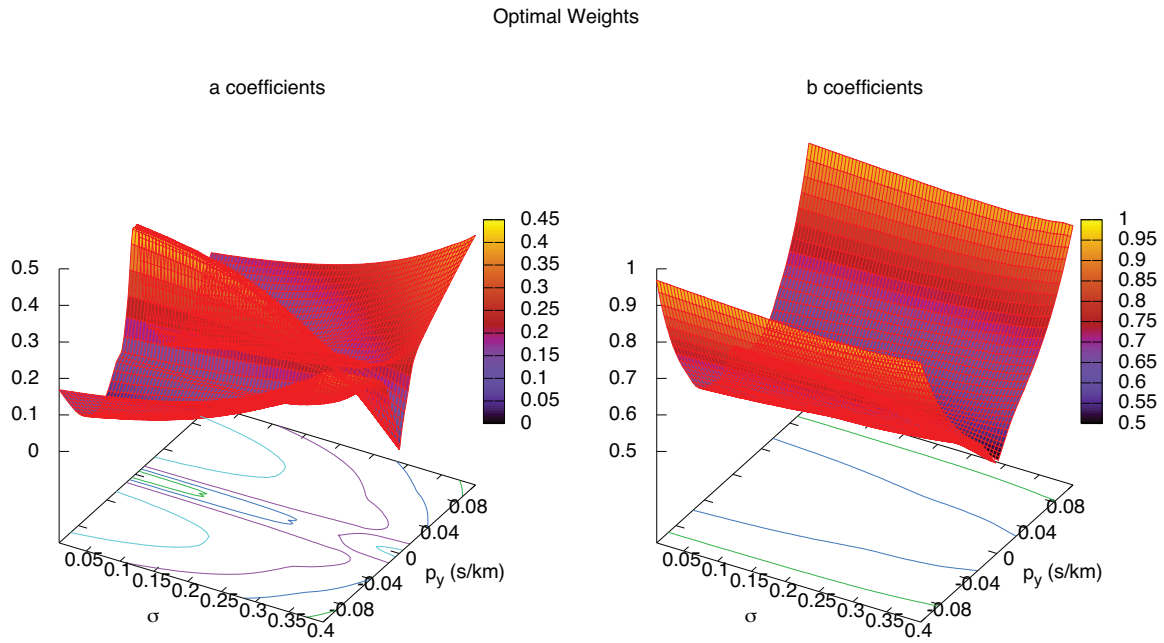
where

$$\begin{aligned}
\frac{9}{hk}S2D &= f(a, c) + f(a, d) + f(b, c) + f(b, d) \\
&+ 4 \sum_{j=1}^n f(a, y_{2j-1}) + 2 \sum_{j=1}^{n-1} f(a, y_{2j}) + 4 \sum_{j=1}^n f(b, y_{2j-1}) \\
&+ 2 \sum_{j=1}^{n-1} f(b, y_{2j}) + 4 \sum_{i=1}^m f(x_{2i-1}, c) + 2 \sum_{i=1}^{m-1} f(x_{2i}, c) \\
&+ 4 \sum_{i=1}^m f(x_{2i-1}, d) + 2 \sum_{i=1}^{m-1} f(x_{2i}, d) \\
&+ 16 \sum_{j=1}^n \left( \sum_{i=1}^m f(x_{2i-1}, y_{2j-1}) \right) + 8 \sum_{j=1}^{n-1} \left( \sum_{i=1}^m f(x_{2i-1}, y_{2j}) \right) \\
&+ 8 \sum_{j=1}^n \left( \sum_{i=1}^{m-1} f(x_{2i}, y_{2j-1}) \right) + 4 \sum_{j=1}^{n-1} \left( \sum_{i=1}^{m-1} f(x_{2i}, y_{2j}) \right)
\end{aligned}$$

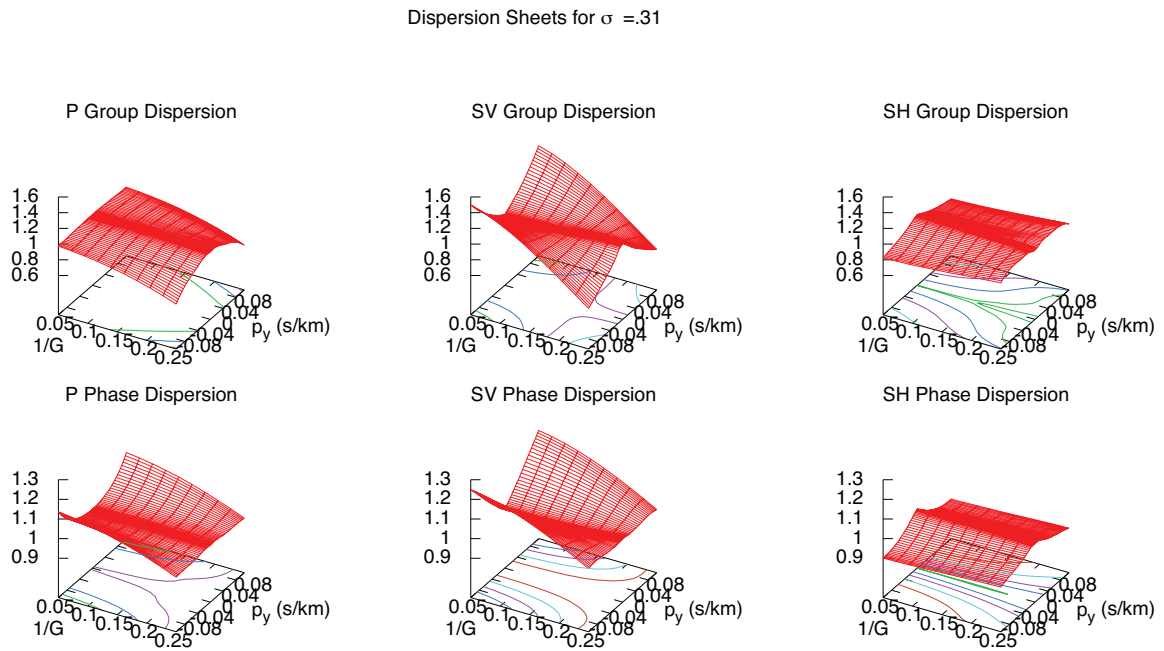
and the rectangle,  $R = \{(x, y) : a \leq x \leq b, c \leq y \leq d\}$  where  $x$  is divided into  $2m$  subintervals by  $h = \frac{b-a}{2m}$ ,  $i = 0, 1, \dots, 2m$  and  $y$  is divided into  $2n$  subintervals by  $k = \frac{d-c}{2n}$ ,  $j = 0, 1, \dots, 2n$ . We choose a Simpson's rule because the error can be shown to be of  $\mathbf{O}(h^4) + \mathbf{O}(k^4)$ . Other schemes, such as Gauss family integration rules, can work just as well provided the integral is "smooth" over the region. In our case, Simpson's rule has the advantage of ease of implementation.

## B2 Results

The optimal  $a$  and  $b$  coefficients determined from minimizing (B.2) depend upon three variables:  $p_y$ , the local P wavespeed ( $\alpha$ ), and Poisson's ratio ( $\sigma$ ). An exploration through typical ranges of these variables reveals a weak dependence of  $a$  and  $b$  on  $\alpha$ . An example of the dependence of  $a$  and  $b$  on  $p_y$  and  $\sigma$  is shown in Figure A1 for  $\alpha = 5.0$  km/s. The  $p_y$  range of  $-0.1224$  to  $0.1224$  corresponds approximately to an incident P wave with a variation in angle of incidence of about  $0^\circ$  to  $40^\circ$  and azimuths between  $-30^\circ$  and  $30^\circ$ , which is a representative range of directions for this application. These plots show that that the optimal values of  $a$  and  $b$  are much more dependent upon on  $\sigma$  than on  $p_y$ . The effect of different choices of  $\alpha$  is a slight a compression or stretching of this surface along the  $p_y$  axis. To a good approximation then, only  $\sigma$  and  $p_y$  need to be specified to recover the optimal values at any grid point. An exception is "fluid-like" media with small shear moduli, as the minimization procedure becomes unstable when Poisson's ratio is greater than about 0.40. As discussed in Štekl and Pratt (1998), in the fluid case one should use  $a = 0$  (rotated system only). Functional forms of the error curves (e.g., Figure A2) for a selection of propagation directions show that for nearly all cases the numerical errors are less that a few percent for  $1/G < 0.25$  or 4 grid points per wavelength. Somewhat



**Figure A1.** Optimal weighting coefficients for given  $p_y$  from  $-0.1224$  to  $0.1224$  and Poisson's ratio from  $0.005$  to  $0.4$



**Figure A2.** Examples of dispersive effects on phase and group velocities using the optimal coefficients in Figure A1 for  $\sigma = 0.31$  and a propagation angle of  $22.5^\circ$ . The  $p_y$  values range from  $-0.1224$  to  $0.1224$ , and the grid values from  $0$  to  $0.25$ .

surprisingly, the largest errors were found for horizontally propagating shear wave waves at large values of  $p_y$  implying that numerical anisotropy may be significant in this particular case. Horizontally propagating shear waves are not expected to play a significant role in teleseismic tomography, but should be considered if this technique is combined with surface waves. Even in this event, the typically longer wavelengths used in surface wave analysis should keep the total errors below a few percent.

### B3 Reduction to acoustic and 2D elastic cases

The 2.5D forms shown above can be reduced to 2D by setting  $p_y = 0$  in (B.11). In this case we arrive at a factored problem of the form

$$\left(\frac{\omega^2 \Delta^2}{\alpha^2} M_1 + K_{22}\right) \left[ \left(\frac{\omega^2 \Delta^2}{\alpha^2} M_1 + K_{11}\right) \left(\frac{\omega^2 \Delta^2}{\alpha^2} M_1 + K_{33}\right) - K_{31}^2 \right] = 0 \quad (\text{B.17})$$

and we find that  $M_1$  and  $K_{22}$  are equivalent to the acoustic forms given by Jo et al. (1996) and correspond to the SH phase in the elastic case.  $K_{11}$ ,  $K_{22}$ , and  $K_{31}$  are equivalent to the forms for P-SV in the 2D elastic case given by Štekl and Pratt (1998)<sup>§</sup>

If we then solve (B.17) where  $x = \frac{\omega \Delta}{\alpha}$ , and match up with the values calculated by the cubic equation, we can identify the roots as

$$x_1 = \frac{-\frac{K_{11}+K_{33}}{M_1} + \sqrt{\frac{(K_{11}+K_{33})^2}{M_1^2} + \frac{K_{11}K_{33}-K_{31}^2}{M_1}}}{2} \quad (\text{B.18a})$$

$$x_2 = \frac{-\frac{K_{11}+K_{33}}{M_1} - \sqrt{\frac{(K_{11}+K_{33})^2}{M_1^2} + \frac{K_{11}K_{33}-K_{31}^2}{M_1}}}{2} \quad (\text{B.18b})$$

$$x_3 = -\frac{K_{22}}{M_1} \quad (\text{B.18c})$$

where  $x_1$  and  $x_2$  are the P and SV coupled 2D elastic case from the quadratic, and  $x_3$  is the SH acoustic case.

## APPENDIX C: CALCULATING THE JACOBIAN FORCES

In equation (26) we defined a set of forces  $F = \mathbf{f}^{(i)}$  that generate the Jacobian:

$$J = \left[ \frac{\partial \mathbf{u}}{\partial m_1} \frac{\partial \mathbf{u}}{\partial m_2} \dots \frac{\partial \mathbf{u}}{\partial m_n} \right] = S^{-1} \left[ \mathbf{f}^{(1)} \mathbf{f}^{(2)} \dots \mathbf{f}^{(n)} \right] \quad (\text{C.1})$$

<sup>§</sup> The computations of Štekl and Pratt (1998) differ from 2D terms we derived independently as their finite difference forms of the cross derivatives in the rotated frame (e.g.  $\frac{\partial^2 u}{\partial x' \partial z'}$ ) appear to be incorrect.

where

$$\mathbf{f}^{(i)} = -\frac{\partial S}{\partial m_i} \mathbf{u} \quad (\text{C.2})$$

The displacement field  $u$  is determined using equation (11). Only a very small percentage of the matrix  $\partial S/\partial m_i$  is nonzero, and to locate these terms we first examine the structure of  $S$ .

### C1 Structure of the Finite Difference Matrix

Each point in the medium is associated with three components of displacement  $(u, v, w)$ , and in general each component is coupled to the other two. Hence the basic element of  $S$  is a 3 x 3 matrix  $A$  that we order as

$$A = \begin{bmatrix} A_{uu} & A_{uv} & A_{uw} \\ A_{vu} & A_{vv} & A_{vw} \\ A_{wu} & A_{wv} & A_{ww} \end{bmatrix}$$

where  $A_{ij}$  contains the finite difference contributions to displacement in direction  $i$  from partial derivatives of displacement in direction  $j$ . Because we use a 9 point star to compute differences, the displacements at any point in the medium are described by 9 of these  $A$  matrices ordered spatially as

$$\begin{array}{ccc} A_1 & A_4 & A_7 \\ A_2 & A_5 & A_8 \\ A_3 & A_6 & A_9 \end{array}$$

and arranged to form 3 rows in the  $S$  matrix as

$$\left[ A_1 \ A_2 \ A_3 \ \cdots \ A_4 \ A_5 \ A_6 \ \cdots \ A_7 \ A_8 \ A_9 \right]$$

For convenience we will refer to this set of 3 rows as a “meta-row”.

### C2 Computing derivatives of $S$

As discussed in Appendix A, medium properties are averaged between points in the 9 point star. Hence terms involving  $m_i$  will be found in each of the  $A$  matrices in the meta-row corresponding to that particular element. Additionally, averaging creates terms involving  $m_i$  that appear in the meta-rows associated with adjacent points in the medium, so that in general there will be 8 additional meta-rows with nonzero  $\partial S/\partial m_i$  elements. The locations of these elements can be summarized by first defining a blockette of  $A$  matrices in  $S$  as

$$B_n = \begin{bmatrix} A_{n-1} & A_n & A_{n+1} & 0 & 0 \\ 0 & A_{n-1} & A_n & A_{n+1} & 0 \\ 0 & 0 & A_{n-1} & A_n & A_{n+1} \end{bmatrix}$$

where the subscript on  $A$  above refers to the position of that matrix within the 9 point star associated with a particular meta-row (i.e.,  $A_n$  in a given meta-row is not necessarily equal to  $A_n$  in any other meta-row). That part of  $S$  involving  $m_i$  consists of 7 of these blockettes arranged as

$$\begin{bmatrix} B_5 & \cdots & B_8 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ B_2 & \cdots & B_5 & \cdots & B_8 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & B_2 & \cdots & B_5 \end{bmatrix} \quad (\text{C.3})$$

Typically we take the  $m_i$  to be the wavespeeds ( $\alpha_i$  and  $\beta_i$ ) specified at the grid points. Let  $m_j$  be a wavespeed at point  $j$  of the 9 point star and define  $A_{n,j}$  as the 3 x 3 matrix of derivatives of  $A_n$  with respect to  $m_j$ , i.e.

$$A_{n,j} = \frac{\partial A_n}{\partial m_j} \quad (\text{C.4})$$

Then for a given  $m_i$ , the nonzero elements of the banded  $\partial S/\partial m_i$  matrix in equation (C.3) can be written in a compact form as

$$\frac{\partial S}{\partial m_i} = \begin{bmatrix} A_{5,9} & A_{6,9} & 0 & A_{8,9} & A_{9,9} & 0 & 0 & 0 & 0 \\ 0 & A_{5,8} & 0 & A_{7,8} & A_{8,8} & A_{9,8} & 0 & 0 & 0 \\ 0 & A_{4,7} & A_{5,7} & 0 & A_{7,7} & A_{8,7} & 0 & 0 & 0 \\ 0 & A_{3,6} & 0 & A_{5,6} & A_{6,6} & 0 & 0 & A_{9,6} & 0 \\ A_{1,5} & A_{2,5} & A_{3,5} & A_{4,5} & A_{5,5} & A_{6,5} & A_{7,5} & A_{8,5} & A_{9,5} \\ 0 & A_{1,4} & 0 & 0 & A_{4,4} & A_{5,4} & 0 & A_{7,4} & 0 \\ 0 & 0 & 0 & A_{2,3} & A_{3,3} & 0 & A_{5,3} & A_{6,3} & 0 \\ 0 & 0 & 0 & A_{1,2} & A_{2,2} & A_{3,2} & 0 & A_{5,2} & 0 \\ 0 & 0 & 0 & 0 & A_{1,1} & A_{2,1} & 0 & A_{4,1} & A_{5,1} \end{bmatrix}$$

where we recall that  $A_{5,9}$  references  $m_i$  when it is located at position 9 in the nine point star,  $A_{5,8}$  when  $m_i$  is at position 8, and so on.

Let  $D_k$  represent the 3 x 1 subvector  $(u_k, v_k, w_k)$  of  $\mathbf{u}$  that is multiplied by the  $k^{th}$  meta-row of

$\partial S/\partial m_i$ . Then from equation (C.2)

$$\mathbf{f}^{(i)} = - \begin{bmatrix} A_{5,9}D_1 + A_{6,9}D_2 + A_{8,9}D_4 + A_{9,9}D_5 \\ A_{5,8}D_2 + A_{7,8}D_4 + A_{8,8}D_5 + A_{9,8}D_6 \\ A_{4,7}D_2 + A_{5,7}D_3 + A_{7,7}D_5 + A_{8,7}D_6 \\ A_{3,6}D_2 + A_{5,6}D_4 + A_{6,6}D_5 + A_{9,6}D_8 \\ \sum_{k=1}^9 A_{k,5}D_i \\ A_{1,4}D_2 + A_{4,4}D_5 + A_{5,4}D_6 + A_{7,4}D_8 \\ A_{2,3}D_4 + A_{3,3}D_5 + A_{5,3}D_7 + A_{6,3}D_8 \\ A_{1,2}D_4 + A_{2,2}D_5 + A_{3,2}D_6 + A_{5,2}D_8 \\ A_{1,1}D_5 + A_{2,1}D_6 + A_{4,1}D_8 + A_{5,1}D_9 \end{bmatrix} \quad (\text{C.5})$$

The 41  $A_{n,j}$  matrices in (C.5) require that we specify 369 medium dependent factors to compute the 27 elements of  $\mathbf{f}^{(i)}$  for each  $m_i$ . While somewhat more onerous than the acoustic case, where only one such factor is needed, the derivatives in (C.4) can be expressed analytically and nearly half of them are zero, so the extra computational effort is not overly burdensome.