CWP-453

Global wave-equation reflection tomography

Maarten V. de Hoop* & Robert D. van der Hilst†

*Center for Wave Phenomena, Colorado School of Mines, Golden, CO 80401-1887

†Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge MA 02139-4307

ABSTRACT

In tomography it has been recognized that the finiteness of the frequency content of seismic data leads to interference effects in the process of medium reconstruction, that need to be accounted for. Various ways of looking at these effects in the framework of *transmission* tomography can be found in the literature. Here, we consider single-scattered body waves and develop from our earlier work on inverse scattering a method of wave-equation *reflection* tomography – which in exploration seismics is identified as a method of wave-equation migration velocity analysis – admitting bandlimited data. In the transition from transmission tomography to reflection tomography the usual cross correlation between modelled and observed data is replaced by modelled annihilators applied to the observed data. Using the generalized screen expansion for one-way wave propagation, we derive the Fréchet or sensitivity kernels and explain how they can be evaluated with an adjoint state method.

INTRODUCTION

In tomography the finiteness of the bandwidth of seismic data leads to interference effects in the process of medium reconstruction that need to be accounted for. Various ways of investigating these effects in the framework of *transmitted* body waves can be found in the literature (Luo & Schuster, 1991; Woodward, 1992; Dahlen *et al.*, 2000). The implied approaches fall in the category of wave-equation tomography: We distinguish the one based on the idea of backprojecting phase residuals over a Fresnel-like volume rather than backprojecting residual traveltimes along rays, and the one based on the wave equation combined with the Born approximation formulated as an adjoint state method (Vasco *et al.*, 1995). Here, we consider single-*scattered* body waves and from our earlier work on imaging-inversion develop a method of wave-equation reflection tomography – which in exploration seismics would be a method of wave-equation migration velocity analysis – admitting finite-frequency data. In the transition from transmission tomography to reflection tomography the usual cross correlation between modelled and observed data is replaced by modelled annihilators applied to the observed data.

In the context of optimization, the key quantity to be found is the Fréchet kernel derived from the error criterion. In wave-equation tomography, the wavefields in the kernel can be computed directly from the time-domain wave equation, using a normal mode summation (Zhao & Jordan, 1998), or using the frequency-domain one-way wave equation, which is the method exploited here.

As with the high-frequency, ray-based approaches to reflection tomography and migration velocity analysis, the idea is to utilize scattered phases in conjunction with the inherent 'redundancy' in their observation. (In the absence of caustics, the redundancy is arises from observations at multiple offsets, as in exploration seismology, or at multiple angular epicentral distances, as in global seismology, and azimuths.) In the framework of reflection seismology in sedimentary basins, the waves scattered from the many reflectors and faults in the subsurface can be used for this purpose, while in the framework of global seismology in the study of the mantle, scattered phases such as PcP and ScS can be identified and used for this purpose. The idea of using scattered phases in global tomography is certainly not new: The ScS travel times, for example, were used by Grand (1994) to build a mantle model. Scattering includes the possibility of mode conversions.

The components of the approach presented in this paper are as follows. The data are downward continued (Claerbout, 1985; Clayton, 1978) and subjected to a wave-equation angle transform (De Bruin *et al.*, 1990; Sava *et al.*, 1999). Such a transform generates multiple images of the same part of Earth's reflecting structure for waves scattered over a range of angles (Stolk & De Hoop, 2003); these images are without artifacts (false reflectors) even in the presence of caustics unlike their counterparts generated by a generalized Radon transform (Brandsberg-Dahl *et al.*, 2003a). However, if the velocity model is incorrect, the images will differ for different scattering angles and azimuths. From the angle transform, annihilators of the data are derived. Whether the velocity model is acceptable is based upon whether the data are in the range of, i.e. can be predicted by, the modelling operator underlying our approach; *annihilators* detect this particular property of the data (Stolk & De Hoop, 2002). (In conventional tomography, one would detect whether the travel times in the data are in the range of the modelled travel times by *differencing*.) Tomography is then formulated as the problem of minimizing the action of these annihilators on the data. We develop the theory here up the expression for the Fréchet kernel following this approach. The evaluation of this kernel is formulated as an adjoint state method (see, for example (Tarantola, 1987)). The key assumption invoked in the application of one-way wave theory is that the rays connected to data points are nowhere horizontal in the subsurface (Stolk & De Hoop, 2003). This does not exclude the formation of caustics associated with multipathing.

Downward continuation and its use in seismic imaging dates back to Claerbout (1985). The principle of characterizing the range of operators of the type encountered in the modelling underlying the presented procedure can be found in Guillemin and Uhlmann (1981) and has been connected to seismic body-wave scattering by De Hoop and Uhlmann (2003). Our annihilator-based approach to velocity analysis is of the differential semblance type. Systematic methods for updating velocity models by optimization using the differential semblance criterion have been introduced by Symes and Carazzone (1991), and further developed and applied by Chauris and Noble (2001), Mulder and Ten Kroode (2002), and others. The wave-equation angle transform occurs in the work of De Bruin *et al.* (1990) and of Sava *et al.* (1999) for the purpose of generating and analyzing common-image-point gathers.

For simplicity of formulation, we consider here *P* waves, possibly in transversely isotropic media with vertical symmetry axis described, approximately, by a scalar wave equation (see, e.g., Schoenberg and De Hoop (2000)). We also make the flattened Earth assumption to avoid contributions from the metric on a manifold in the equations.

The approach presented here is a synthesis of various developments: The generalized Bremmer coupling series to model seismic reflection data (De Hoop, 1996), inverse scattering based upon the first-order term of this series and the wave-equation angle transform (Stolk & De Hoop, 2003), the relation between wavefield reciprocity and optimization (De Hoop & De Hoop, 2000), and the generalized screen expansion for one-way wave propagation (Le Rousseau & De Hoop, 2001), which constitutes the downward continuation.

1 DIRECTIONAL WAVEFIELD DECOMPOSITION

We identify the z coordinate with depth, which coincides with the direction of continuation. The remaining lateral coordinates are collected in x. Time is indicated by t. The partial derivatives are denoted by

$$D_{x_{1,\ldots,n-1}} = -\mathrm{i}\partial_{x_{1,\ldots,n-1}} \ , \ D_z = -\mathrm{i}\partial_z \ , \ D_t = -\mathrm{i}\partial_t$$

so that their Fourier domain counterparts become multiplications by ξ, ζ (the wave vector) and ω. Here, $n = 2$ for 2D seismics while $n = 3$ for 3D seismics.

We consider the scalar wave equation for pressure u rewritten as a first-order system

$$\frac{\partial}{\partial z}\begin{pmatrix} u \\ \frac{\partial u}{\partial z}\end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -A(x,z,D_x,D_t) & 0 \end{pmatrix}\begin{pmatrix} u \\ \frac{\partial u}{\partial z}\end{pmatrix} + \begin{pmatrix} 0 \\ f \end{pmatrix}, \quad (1)$$

where $A(x,z,\xi,\omega) = c_0(x,z)^{-2}\omega^2 - \|\xi\|^2$ is the principal ('high-frequency') part of the symbol of A. In a stratified medium that is translationally invariant in the horizontal directions, the principal symbol equals the full symbol; in more general media, a symbol calculus develops the subprincipal part of the symbol. In fact, (x,z,t,ξ,ζ,ω) are coordinates on phase space; the symbol of A is defined on this phase space. Localization in phase space is called microlocalization in the analysis literature.

1.1 The system of one-way wave equations

Microlocally, away from the zeroes of $A(x,z,\xi,\omega)$, system (1) can be transformed into diagonal form modulo a smoothing operator. A family of pseudodifferential operator matrices $\mathsf{Q}(z) = \mathsf{Q}(x,z,D_x,D_t)$ exists such that

$$\begin{pmatrix} u_+ \\ u_- \end{pmatrix} = \mathsf{Q}(z)\begin{pmatrix} u \\ \frac{\partial u}{\partial z}\end{pmatrix}, \qquad \begin{pmatrix} f_+ \\ f_- \end{pmatrix} = \mathsf{Q}(z)\begin{pmatrix} 0 \\ f \end{pmatrix},$$

satisfy the one-way wave or single-square-root equations

$$\left(\frac{\partial}{\partial z} \pm \mathrm{i}B_\pm(x,z,D_x,D_t)\right) u_\pm = f_\pm \ . \quad (2)$$

Any coupling between $+$ and $-$ constituents is absorbed in $f_\pm$, but is of higher order. The principal symbol b of the $B_\pm$ is given by $b(x,z,\xi,\omega) = \sqrt{A(x,z,\xi,\omega)} = \omega\sqrt{\frac{1}{c_0(x,z)^2} - \omega^{-2}\|\xi\|^2}$; $\omega^{-1}b$ has the appearance of vertical wave slowness. For (x,t,ξ,ω) such that the symbol $B_\pm$ is real, the equation is of hyperbolic type, corresponding microlocally to propagating waves.

The subprincipal part of the $B_\pm$ depends on the normalization of $\mathsf{Q}(z)$. We choose this normalization such that (2) is selfadjoint microlocally where the symbol of $B_\pm$ is real; then

$$u = Q_+^* u_+ + Q_-^* u_-, \quad (3)$$

$$f_\pm = \pm\tfrac{1}{2}\mathrm{i}Q_\pm f \ , \quad (4)$$

where $Q_\pm = Q_\pm(z) = Q_\pm(x,z,D_x,D_t)$ are z-families of pseudodifferential operators with principal symbols $\omega^{-1/2}\left(\frac{1}{c_0(x,z)^2} - \omega^{-2}\|\xi\|^2\right)^{-1/4}$. The physical meaning of this choice of $Q_\pm$ is that the down- and upgoing fields are normalized in vertical-acoustic-power flux. The operators $Q_\pm$ form the first column of Q. We observe that $Q_-^* u_-$ represents the upgoing and $Q_+^* u_+$ represents the downgoing constituent of the wavefield u.

Upon regularization (for a precise treatment, see (Stolk & De Hoop, 2003Appendix A)), there is a well defined solution operator $G_-(z,z_0)$, $z < z_0$, of the initial value problem for u_- given by (2) with $f_- = 0$; this operator describes propagation from z_0 to z, in the upward direction (of decreasing z). The *adjoint* $G_-(z,z_0)^*$ describes propagation from z to z_0 of (2) or from z_0 to z in reverse time. The operator G_- is a Fourier integral operator with complex phase

(Melin & Sjöstrand, 1975), (Hörmander, 1985chapter XXV), (Treves, 1980chapters X and XI). By Duhamel's principle, a solution operator for the inhomogeneous equation (2) is given by

$$u_-(\cdot, z) = \int_z^\infty G_-(z, z_0) f_-(\cdot, z_0)\,\mathrm{d}z_0 \ . \tag{5}$$

Note that the '$-$' or upgoing constituent of the original Green's function is generated by $-\frac{1}{2}\mathrm{i}Q_-^*(z)G_-(z,z_0)Q_-(z_0)$ (cf. (3)-(4)).

1.2 Generalized screen expansion

1.2.1 The single-square-root operator

The generalized screen expansion of the principal symbol of the single-square-root operator up to order N is of the form

$$b(x,z,\xi,\omega) = \sum_{j=0}^{N} A_j(\xi,\omega,z)\, U_j[c_0](x,z) \ , \tag{6}$$

with $U_0[c_0](x,z) = 1$ (cf. (Le Rousseau & De Hoop, 2001(16))) *. The factors A_j mostly control the shape (bending) of the local slowness surface while the factors U_j account for the change in the slowness surface due to the lateral medium fluctuations with respect to some average background $\bar{c}_0 = \bar{c}_0(z)$ varying with depth only. The factors A_j depend upon $\bar{c}_0$ but not upon the lateral medium fluctuations. This expansion is valid away from $\|\omega^{-1}\xi\| = \bar{c}_0^{-1}$ at each depth z.

Up to principal parts, interpreting the symbol as a right or dual symbol, the operator B_- acts on the one-way wavefield u_- as

$$B_- u_- \sim F^{-1}_{\omega\to t} \sum_{j=0}^{N} F^{-1}_{\xi\to x} A_j(\xi,\omega,z) F_{x'\to\xi} U_j[c_0](x',z) F_{t'\to\omega} u_- \ ,$$

rewritten in the time-Fourier or frequency (ω) domain as

$$\hat{B}_- \hat{u}_- \sim \sum_{j=0}^{N} F^{-1}_{\xi\to x} A_j(\xi,\omega,z) F_{x'\to\xi} U_j[c_0](x',z)\hat{u}_- \tag{7}$$

(cf. (Le Rousseau & De Hoop, 2001(31))), where F denotes the Fourier transform and $\hat{}$ indicates the time-Fourier domain. Thus the dependency of the operator on the laterally varying component of the background medium is completely contained in the factors U_j.

1.2.2 The perturbed single-square-root operator

Here, for the later application of tomography, we consider how the single-square-root operator is perturbed under a smooth perturbation $\delta c_0(x,z)$ of c_0 subject to the constraint that $\bar{c}_0(z)$ is kept fixed. (In fact, any perturbation in $\bar{c}_0(z)$ is absorbed in $\delta c_0(x,z)$.) In view of (7) we have

$$(\delta\hat{B}_-)\hat{u}_- \sim \sum_{j=0}^{N} F^{-1}_{\xi\to x} A_j(\xi,\omega,z) F_{x'\to\xi} \delta U_j(x',z)\hat{u}_- \ . \tag{8}$$

The perturbation δU_j in U_j is expressed in terms of Fréchet derivatives U_j' as

$$\delta U_j(x,z) = U_j'[c_0](x,z)\,\delta c_0(x,z) \ ; \tag{9}$$

this is a directional derivative, along a curve in c_0 space †. The curves follow a parametric representation of c_0, for example, in terms of cubic B-splines. The operation in (9) is multiplicative.

Substituting (9) into (8), we find an operator

$$\hat{B}_-'(\hat{u}_-) \sim \sum_{j=0}^{N} F^{-1}_{\xi\to x} A_j(\xi,\omega,z) F_{x'\to\xi} \underbrace{U_j'(x',z)\hat{u}_-}_{'\text{screen}'} \tag{10}$$

*The generalized screen expansion implies $A_j(\xi,\omega,z) = \omega a_j[\bar{c}_0(z)^{-2} - \omega^{-2}\|\xi\|^2]^{(-2j+1)/2}$ with $a_j = (-1)^{j+1}\frac{1\cdot 3\cdots(2j-1)}{j!2^j}$, and $U_j(x,z) = [c_0(x,z)^{-2} - \bar{c}_0(z)^{-2}]^j$.

†Here, $U_j'(x,z) = j\,[c_0(x,z)^{-2} - \bar{c}_0(z)^{-2}]^{j-1}(-2)c_0(x,z)^{-3}$, $j = 1, 2, \ldots$

such that

$$(\delta\hat{B}_-)\hat{u}_- \sim \hat{B}'_-(\hat{u}_-)\,\delta c_0\ . \tag{11}$$

Equation (10) indicates how the field is absorbed as a factor in the screen function.

For the purpose of tomography, we will also need the *adjoint* $\hat{B}'_-(\hat{u}_-)^*$ of $\hat{B}'_-(\hat{u}_-)$. Since $(F_{x\to\xi})^* = F^{-1}_{\xi\to x}$ (up to factors of 2π), we have

$$\hat{B}'_-(\hat{u}_-)^* \sim \sum_{j=0}^{N} \hat{u}^*_- U'_j(x',z) F^{-1}_{\xi\to x'} A_j(\xi,\omega,z) F_{x\to\xi} \tag{12}$$

with the property that

$$\langle \hat{B}'_-(\hat{u}_-)\delta c_0, \hat{v}\rangle_{(x)} = \langle \delta c_0, \hat{B}'_-(\hat{u}_-)^*\hat{v}\rangle_{(x)} \ \text{ for given } z \text{ and } \omega$$

where $\langle .,.\rangle$ indicates the standard inner product in L^2.

2 WAVEFIELD CONTINUATION IN DEPTH

2.1 Modelling of seismic data

Let $u = u(s,r,t,z)$ now represent seismic data as a function of time, generated by sources (earthquakes) s and observed by receivers (stations) r if they were at the depth z below the surface. In the single-scattering approximation, such data satisfy an inhomogeneous double-square-root (DSR) equation,

$$(\tfrac{\partial}{\partial z} \underbrace{-\mathrm{i}B_-(s,z,D_s,D_t) - \mathrm{i}B_-(r,z,D_r,D_t)}_{-\mathrm{i}C(s,r,z,D_s,D_r,D_t)})u = g_-\ . \tag{13}$$

Here,

$$g_- = Q_{-,s}(z) Q_{-,r}(z) g\ ,$$

where $Q_{-,s}(z)$ is short for $Q_-(s,z,D_s,D_t)$ and $Q_{-,r}(z)$ is short for $Q_-(r,z,D_r,D_t)$, while g is given by

$$g(s,r,t,z) = \tfrac{1}{2} D_t^2 (I_2 I_1\, c_0^{-3}\delta c)(s,r,t,z)\ , \tag{14}$$

in which the maps I_1, I_2 are defined as

$$I_1 : c(x,z) \mapsto \delta(r-s) c(\tfrac{r+s}{2}, z)\ , \tag{15}$$

$$I_2 : c(s,r,z) \mapsto \delta(t) c(s,r,z)\ ; \tag{16}$$

the form of g is consistent with the Born approximation for scattered waves. For convience of notation, we have absorbed a factor $(-\frac{1}{2}\mathrm{i})^2$ in g. The perturbation δc contains the singular variations in medium wave speed; thus the total medium wave speed is composed as follows

$$c^{-2}(x,z) = \bar{c}_0^{-2}(z) + [c_0^{-2}(x,z) - \bar{c}_0^{-2}(z)] - 2c_0^{-3}(x,z)\,\delta c(x,z)\ .$$

It follows from the definition of $G_-(z,z_0)$ that the solution operator, $H(z,z_0)$ for $z > z_0$, for the Cauchy initial value problem for the regularized DSR equation can be written as

$$(H(z,z_0))(s,r,t,s_0,r_0,t_0) = \int_{\mathbb{R}} (G_-(z,z_0))(s,t-t_0-t',s_0,0)(G_-(z,z_0))(r,t',r_0,0)\,\mathrm{d}t'\ . \tag{17}$$

Here $(G_-(z,z_0))(r,t',r_0,0)$ denotes the distribution kernel of $G_-(z,z_0)$, and a likewise notation is employed for $H(z,z_0)$.

Using (17) in (13) it follows that the singular part of the data are modelled in the single-scattering (Born) approximation as

$$d(s,r,t) = \int_0^Z Q^*_{-,s}(0) Q^*_{-,r}(0) H(0,z_0) Q_{-,s}(z_0) Q_{-,r}(z_0)\, g(\cdot,z_0)\,\mathrm{d}z_0\ , \tag{18}$$

We assume that the support of δc is contained in the half space $\{(x,z) \mid z > 0\}$. For (18) to be valid we assume that the rays are nowhere horizontal. In equation (18) we recognize the time convolution of an incoming Green's function connecting the source at s (and zero depth) with the scattering point at $r_0 = s_0$ and depth z_0 with the outgoing Green's function connecting this scattering point with the receiver at r (and zero depth).

2.2 Downward continuation of seismic data

We assume that the original data, $d = d(s, r, t)$, are taken on the level surface $z_0 = 0$. (This may require a source depth correction.) The *adjoint* operator $H(0, z)^*$ is used to propagate the data backward in time. We consider

$$H(0, z)^* \ : \ d \mapsto H(0, z)^* \psi d = H(z, 0) \psi d \ , \tag{19}$$

where ψ is an appropriate tapered mute. The tapered mute is applied to guarantee that the seismic events corresponding to rays in the scattering process that do not satisfy the various conditions are eliminated from the data. $H(0, z)^* \psi d$ is a function of (s, r, t, z).

3 WAVE-EQUATION ANGLE TRANSFORM AND ANNIHILATORS

We discuss here how redundancy in the (downward-continued) data can be understood and exploited for the purpose of tomography. The redundancy arises in the data because the wavefield in the subsurface scatters over different angles from the same reflection (bounce) point. The reflection point becomes the image point in the imaging process, and the redundancy in the data becomes manifest in the multiple images that are generated by extracting the different scattering angles and azimuths in this process. This is accomplished by the angle transform.

The test whether the velocity model is acceptable is based on the test whether the data are in the range of the single scattering modelling operator (18). This is established by applying annihilators derived from the angle transform to the data and introducing a norm as in the differential semblance criterion.

3.1 Angle transform

We define the wave-equation angle transform A by the following integral transform of the downward continued data:

$$(Ad)(x, z, p) = R \underbrace{H(0, z)^* \psi d}_{u(s,r,t,z)} , \tag{20}$$

with

$$\begin{aligned}(Ru)(x, z, p) &= \int_{\mathbb{R}^{n-1}} u(x - \tfrac{h}{2}, x + \tfrac{h}{2}, ph, z) \chi(x, z, h) \, \mathrm{d}h \\ &= \frac{1}{2\pi} \int \int_{\mathbb{R}^{n-1}} \hat{u}(x - \tfrac{h}{2}, x + \tfrac{h}{2}, \omega, z) \exp(-\mathrm{i}\omega p h) \, \chi(x, z, h) \, \mathrm{d}h \mathrm{d}\omega \ .\end{aligned} \tag{21}$$

This is basically an example of double beamforming with the downward continued data, in sources and receivers (for the surface counterpart, see Scherbaum *et al.* (1997)). Clearly

$$(Ad)(x, z, p) = \int_{\mathbb{R}^{n-1}} (H(0, z)^* \psi d)(x - \tfrac{h}{2}, x + \tfrac{h}{2}, ph) \chi(x, z, h) \, \mathrm{d}h \tag{22}$$

(cf. (Sava *et al.*, 1999; Stolk & De Hoop, 2003)), where $h \mapsto \chi(x, z, h)$ is a compactly supported cutoff function the support of which contains $h = 0$ such that $\chi(x, z, 0) = 1$. With an appropriate choice of support of χ, A is an invertible Fourier integral operator (Stolk & De Hoop, 2003). The notation in (21) is such ph is a multiplication for $n = 2$ and an inner product for $n \geq 3$. When the correct background medium c_0 is given, A maps data d to a p-family of reconstructions of δc in the form of pseudodifferential operators acting on δc.

The 'true-amplitude' counterpart of transform A follows from replacing

$$d \mapsto u = (H(0, z)^* \psi d)$$

by

$$d \mapsto u = 2 c_0(x, z)^3 \, \Xi(z) Q_{-,s}^*(z)^{-1} Q_{-,r}^*(z)^{-1} H(0, z)^* Q_{-,s}(0)^{-1} Q_{-,r}(0)^{-1} D_t^{-2} \psi d \ ,$$

where $\Xi(z)$ is a pseudodifferential operator with principal symbol

$$\begin{aligned}&\Xi(s, r, t, \sigma, \rho, \omega) \\ &= \frac{1}{c_0(s, z)^2} \left(\frac{1}{c_0(s, z)^2} - \omega^{-2} \|\sigma\|^2 \right)^{-1/2} + \frac{1}{c_0(r, z)^2} \left(\frac{1}{c_0(r, z)^2} - \omega^{-2} \|\rho\|^2 \right)^{-1/2}\end{aligned} \tag{23}$$

(Stolk & De Hoop, 2003(34),(41)). Using the true-amplitude version of the angle transform, the outcome is related to δc as an operator accounting for illumination effects only. We suppress the true-amplitude component operators in our formulas to keep presentation the simple.

3.2 Annihilators

Since the outcome $(Ad)(x,z,p)$ should be independent of p, annihilators of the data follow to be

$$W_i := \langle A^{-1} \rangle \frac{\partial}{\partial p_i} A\ ,\ i = 1, \ldots, n-1\ ,$$

where $\langle A^{-1} \rangle$ indicates a regularized inverse of A. Thus $W_i d = 0$. Rather than considering the annihilators in the data domain, we can consider the operators

$$\begin{aligned}(R_i' u)(x,z) &= \frac{\partial}{\partial p_i} A D_t^{-1} d \\ &= \frac{1}{2\pi} \int \int_{\mathbb{R}^{n-1}} \hat{u}(x - \tfrac{h}{2}, x + \tfrac{h}{2}, \omega, z) \exp(-\mathrm{i}\omega p h)\, h_i \chi(x,z,h)\, \mathrm{d}h \mathrm{d}\omega \end{aligned} \tag{24}$$

in the image domain. The annihilation of the data is replaced by an annihilation of the set of images $AD_t^{-1}d$. This approach was followed in the framework of the ray geometrical generalized Radon transform (or Kirchhoff-style) migration velocity analysis by Brandsberg-Dahl, Ursin and De Hoop (2003b).

For the purpose of tomography, we also need the *adjoint* $(R')^*$ of R'. Let $\mathcal{I}$ denote a test image as a function of (x,z,p), then the adjoint follows from

$$\begin{aligned}\langle R'u, \mathcal{I} \rangle_{(x,p)} &= \int_{\mathbb{R}^{2n-2}} \frac{1}{2\pi} \int \int_{\mathbb{R}^{n-1}} \\ &\qquad \hat{u}(x - \tfrac{h}{2}, x + \tfrac{h}{2}, \omega, z) \exp(-\mathrm{i}\omega p h)\, \chi(x,z,h)\, \mathrm{d}h \mathrm{d}\omega\, \mathcal{I}(x,z,p)^* \mathrm{d}x \mathrm{d}p \\ \frac{1}{2\pi} &\int \int_{\mathbb{R}^{2n-2}} \hat{u}(s,r,\omega,z) \left(\int_{\mathbb{R}^{n-1}} \right. \\ &\qquad \left. \exp[\mathrm{i}\omega p(r-s)]\, \mathcal{I}(\tfrac{1}{2}(s+r), z, p) \mathrm{d}p \right)^* \chi(\tfrac{1}{2}(s+r), z, r-s)\, \mathrm{d}s \mathrm{d}r \mathrm{d}\omega \\ &= \langle \hat{u}, (R')^* \mathcal{I} \rangle_{(s,r,\omega)} \ \text{ for given } z, \end{aligned} \tag{25}$$

from which we identify $(R')^*$.

4 AN OPTIMIZATION PROCEDURE FOR REFLECTION TOMOGRAPHY

The model estimation in our approach to reflection tomography is based upon minimizing the functional

$$\mathcal{J}[c_0] = \tfrac{1}{2} \int |(R'u)(x,z)|^2 \mathrm{d}(x,z) \mathrm{d}p\ . \tag{26}$$

Note that the measure of traveltime mismatch has been replaced by the operator R'; picking of traveltimes has been avoided. The integration is over all scattering or image points and angles.

The method for evaluating the Fréchet kernel or gradient of a functional derived from the solution of a partial (or pseudo)differential equation is known as the adjoint state method. It has been introduced and used in seismology by Tarantola (1987) and many others.

4.1 Perturbing the functional

Under a smooth perturbation δc_0 of c_0 subject to the constraint that $\bar{c}_0$ is kept the same, the perturbation of the downward-continued data u is δu while the perturbation of the functional follows to be

$$\delta \mathcal{J} = \int (R'u)\, (R' \delta u)\, \mathrm{d}(x,z) \mathrm{d}p\ , \tag{27}$$

because R' is independent of c_0. Hence,

$$\delta \mathcal{J} = \frac{1}{2\pi} \int \int \underbrace{((R')^* R' u)}_{\text{source}} (\overbrace{\delta \hat{u}}^{\text{field}})^*\, \mathrm{d}(s,r,z)\, \mathrm{d}\omega \tag{28}$$

by definition of adjoint in (25). In preparation of applying the reciprocity theorem of the time-correlation type in the framework of one-way wave theory, we interpret $(R')^* R' u$ as a source, viz., in the adjoint field equation for v, say,

$$(-\tfrac{\partial}{\partial z} \underbrace{+\mathrm{i}B_-(s,z,D_s,D_t)^* + \mathrm{i}B_-(r,z,D_r,D_t)^*}_{-(\mathrm{i}C(s,r,z,D_s,D_r,D_t))^*})v = \underbrace{(R')^* R' u}_{\text{penalty}} \ . \tag{29}$$

This equation is solved in the direction of decreasing z ('upward') with vanishing initial condition for some large z at the bottom of the model. Its right-hand side is interpreted as a penalty force, which vanishes if c_0 were an acceptable background medium. The field δu satisfies the scattered field equation

$$(\tfrac{\partial}{\partial z} \underbrace{-\mathrm{i}B_-(s,z,D_s,D_t) - \mathrm{i}B_-(r,z,D_r,D_t)}_{-\mathrm{i}C(s,r,z,D_s,D_r,D_t)})\delta u = \mathrm{i}\,\underbrace{\delta C(s,r,z,D_s,D_r,D_t)}_{\text{contrast}}\, u \tag{30}$$

in the single-scattering approximation, where δC is perturbation of operator C with δc_0, see (8) and (13). This equation is solved in the direction of increasing z (downward) with homogeneous initial conditions at $z = 0$. Its right-hand side is interpreted as a contrast source.

The reciprocity theorem of the time-correlation type now implies that we can write (28) in the equivalent form

$$\delta\mathcal{J} = \frac{1}{2\pi}\int\int \overbrace{\hat{v}}^{\text{field}} (\underbrace{\mathrm{i}\,\delta\hat{C}\hat{u}}_{\text{source}})^* \,\mathrm{d}(s,r,z)\,\mathrm{d}\omega \ . \tag{31}$$

4.2 The kernel

To arrive at the sensitivity kernel, we write the perturbation $\delta\hat{C}$ in expression (31) in the form $\delta\hat{C}\hat{u} = \hat{C}'(\hat{u})\,\delta c_0$, with

$$\hat{C}'(\hat{u}) \sim \sum_{j=0}^{N} F^{-1}_{\sigma\to x} A_j(\sigma,\omega,z) F_{s\to\sigma} U_j'(s,z)\hat{u}(s,r,\omega,z) + \sum_{j=0}^{N} F^{-1}_{\rho\to x} A_j(\rho,\omega,z) F_{r\to\rho} U_j'(r,z)\hat{u}(s,r,\omega,z) \ .$$

Substituting this expression into (31) and taking the adjoint, we obtain

$$\delta\mathcal{J} = \frac{1}{2\pi}\int\int \hat{v}\,(\mathrm{i}\,\delta\hat{C}\hat{u})^*\,\mathrm{d}(s,r,z)\mathrm{d}\omega = \int (\mathrm{i}\,C'(u))^* v\,\delta c_0\,\mathrm{d}(x,z) \ . \tag{32}$$

To extract the kernel of the (directional) derivative of $\mathcal{J}$ out of expression (31), we make use of relation (12) with x' playing the role of s or r. Thus, the kernel attains the form

$$\begin{aligned} C'(u)^* v &= \frac{1}{2\pi}\int \mathrm{d}\omega \left[\int \hat{B}'_{-,s}(\hat{u})^*\hat{v}\,\mathrm{d}r + \int \hat{B}'_{-,r}(\hat{u})^*\hat{v}\,\mathrm{d}s\right] \\ &\sim \frac{1}{\pi}\,\mathrm{Re}\int_0^\infty \mathrm{d}\omega \sum_{j=0}^{N}\left[\int \hat{u}(x,r,\omega,z)^* U_j'(x,z) F^{-1}_{\sigma\to x} A_j(\sigma,\omega,z) F_{s\to\sigma}\hat{v}(.,r,\omega,z)\,\mathrm{d}r \right. \\ &\quad \left. + \int \hat{u}(s,x,\omega,z)^* U_j'(x,z) F^{-1}_{\rho\to x} A_j(\rho,\omega,z) F_{r\to\rho}\hat{v}(s,.,\omega,z)\,\mathrm{d}s\right] . \end{aligned} \tag{33}$$

We used the symmetry in frequency to restrict the evaluation to positive values. This expression is of the form of a time cross correlation of the downward continued data u (cf. (19)) and the adjoint field v excited by a penalty force nested in a generalized screen operation. It differs from the standard imaging condition in particular through the integrations over s and r.

4.3 Discretization in depth and generalized screen propagators

Discretization of the integral over z in (31) into a sum over a finite set of depths $\{z_n\}$ admits the incorporation of a DSR wave propagators. Let the depth step be $z_{n+1} - z_n = \Delta z$. The downward propagator, $h(z_{n+1}, z_n)$, is derived from (17) and satisfies the relation

$$u_{n+1} = h(z_{n+1}, z_n) u_n \ .$$

First, the contrast source $\mathrm{i}\,\delta C u$ evaluated at depth z_n in (31) is identified with the discretization of the left-hand side of (30),

$$(\Delta z)^{-1}[\delta u_{n+1} - (1 + \mathrm{i}\Delta z C(z_n))\,\delta u_n] \ ,$$

where the subscript $_n$ indicates the evaluation at $z = z_n$. Secondly, the operator $1 + \mathrm{i}\Delta z C(z_n)$ is viewed as an approximation to the propagator $h(z_{n+1}, z_n)$. Thirdly, substituting such a replacement in the expression above implies that upon discretization the contrast source $\mathrm{i}\,\delta\hat{C}\hat{u}$ inside the z-integral can be viewed as an approximation to

$$\mathrm{i}\hat{h}(z_{n+1}, z_n)(\delta\hat{C}(z_n)\hat{u}_n)$$

inside a summation over n. This expression is of the form of the difference of two residual propagators (Le Rousseau & De Hoop, 2001(27)) acting on $\hat{u}_n$.

5 DISCUSSION

We derived an explicit expression for the kernel driving wave-equation reflection tomography, using the generalized screen expansion for the downward continuation of the seismic data. Our criterion is based upon the condition that a velocity model is acceptable if the data are in the range of our modelling operator, here developed in the single scattering approximation. The evaluation of the kernel has a few unusual aspects: (i) the 'penalty' occurs and is evaluated in the subsurface rather than at the surface, (ii) the adjoint state method, derived from the reciprocity theorem of the time correlation type, leads to a procedure of propagating the adjoint field upwards and the data downwards and taking their cross correlation at the depths where the fields 'meet', (iii) a generalized screen operator is nested in the cross correlation in time. The penalty force can attain non-vanishing values only in regions that contain reflectors, where $\delta c \neq 0$.

The presented procedure admits bandlimited data. Any effects related to 'wavefront healing' are accounted for in the downward continuation restricted to the frequency band of the data irrespective of the formation of caustics. The full waveform is used, not the phase of the data.

The current method is developed for single scattering, i.e. primary reflections, and qP waves. Replacing the acoustic wavefield decomposition and generalized screen expansion by their elastic counterparts (Le Rousseau & De Hoop, 2003), in principal, the method can be extended to shear waves and mode conversions.

ACKNOWLEDGMENTS

The first author thanks the members of the Department of Earth, Atmospheric and Planetary Sciences at the Massachusetts Institute of Technology, and in particular Professor Prinn, for their hospitality, for financial support, and for providing a very stimulating environment during the Spring semester in 2003.

REFERENCES

Brandsberg-Dahl, S., De Hoop, M.V., & Ursin, B. 2003a. Focusing in dip and AVA compensation on scattering-angle/azimuth common image gathers. *Geoph.*, **68**, 232–254.

Brandsberg-Dahl, S., Ursin, B., & De Hoop, M.V. 2003b. Seismic velocity analysis in the scattering-angle/azimuth domain. *Geoph. Prosp.*, in print.

Chauris, H., & Noble, M. 2001. Two-dimensional velocity macro-model estimation from seismic reflection data by local differential semblance optimization: Applications to synthetic and real data sets. *Geophys. J. Int.*, **144**, 14–26.

Claerbout, J.F. 1985. *Imaging the earth's interior*. Oxford: Blackwell Scientific Publishing.

Clayton, R.W. 1978. Common midpoint migration. *Technical Report, Stanford University*, **SEP-14**.

Dahlen, F.A., Hung, S.-H., & Nolet, G. 2000. Fréchet kernels for finite-frequency traveltimes–I. Theory. *Geophys. J. Int.*, **141**, 157–174.

De Bruin, C.G.M., Wapenaar, C.P.A., & Berkhout, A.J. 1990. Angle-dependent reflectivity by means of prestack migration. *Geoph.*, **55**, 1223–1234.

De Hoop, M.V. 1996. Generalization of the Bremmer coupling series. *J. Math. Phys.*, **37**, 3246–3282.

De Hoop, M.V., & De Hoop, A.T. 2000. Wavefield reciprocity and optimization in remote sensing. *Proc. R. Soc. Lond. A (Mathematical, Physical and Engineering Sciences)*, **456**, 641–682.

De Hoop, M.V., & Uhlmann, G. 2003. in preparation.

Grand, S.P. 1994. Mantle shear structure beneath the Americas and surrounding oceans. *J. Geophys. Res.*, **99**, 11591–11621.

Guillemin, V., & Uhlmann, G. 1981. Oscillatory integrals with singular symbols. *Duke Math. Journal*, **48**, 251–267.

Hörmander, L. 1985. *The Analysis of Linear Partial Differential Operators*. Vol. 4. Berlin: Springer-Verlag.

Le Rousseau, J.H., & De Hoop, M.V. 2001. Modeling and imaging with the scalar generalized-screen algorithms in isotropic media. *Geoph.*, **66**, 1551–1568.

Le Rousseau, J.H., & De Hoop, M.V. 2003. Generalized-screen approximation and algorithm for the scattering of elastic waves. *Q. J. Mech. Appl. Math.*, **56**, 1–33.

Luo, Y., & Schuster, G.T. 1991. Wave-equation travel time inversion. *Geoph.*, **56**, 645–653.

Melin, A., & Sjöstrand, J. 1975. Fourier integral operators with complex-valued phase functions. *Pages 120–223 of: Fourier integral operators and partial differential equations*. Berlin: Springer-Verlag, for Lecture Notes in Math. Vol.459.

Mulder, W.A., & Ten Kroode, A.P.E. 2002. Automatic velocity analysis by differential semblance optimization. *Geoph.*, **67**, 1184–1191.

Sava, P., Biondi, B., & Fomel, S. 1999. Amplitude-preserved common image gathers by wave-equation migration. *Pages 824–827 of: Expanded Abstracts*. Society of Exploration Geophysicists.

Scherbaum, F., Krüger, F., & Weber, M. 1997. Double beam forming: Mapping lower mantle heterogeneities using combinations of source and receiver arrays. *J. Geophys. Res.*, **102**, 507–522.

Schoenberg, M., & De Hoop, M.V. 2000. Approximate dispersion relations for qP-qSV waves in transversely isotropic media. *Geoph.*, **65**, 919–933.

Stolk, C.C., & De Hoop, M.V. 2002. Microlocal analysis of seismic inverse scattering in anisotropic elastic media. *Comm. Pure Appl. Math.*, **55**, 261–301.

Stolk, C.C., & De Hoop, M.V. 2003. Seismic inverse scattering in the 'wave-equation' approach. *SIAM J. Appl. Math.*, submitted.

Symes, W.W., & Carazzone, J. 1991. Velocity inversion by differential semblance optimization. *Geoph.*, **56**, 654–663.

Tarantola, A. 1987. *Inverse problem theory*. Amsterdam: Elsevier.

Treves, F. 1980. *Introduction to Pseudodifferential and Fourier Integral Operators*. Vol. 2. New York: Plenum Press.

Vasco, D.W., Peterson, J.E., & Majer, E.L. 1995. Beyond ray tomography: Wavepaths and Fresnel volumes. *Geoph.*, **60**, 1790–1804.
Woodward, M. 1992. Wave equation tomography. *Geoph.*, **57**, 15–26.
Zhao, L., & Jordan, T.H. 1998. Sensitivity of frequency-dependent travel times to laterally heterogeneous, anisotropic Earth structure. *Geophys. J. Int.*, **133**, 683–704.