

## Research Article

# Multiparameter Inversion: Cramer's Rule for Pseudodifferential Operators

Rami Nammour and William W. Symes

Department of Computational and Applied Mathematics, Rice University, 6100 Main Street, Houston, TX 77005, USA

Correspondence should be addressed to William W. Symes, symes@caam.rice.edu

Received 16 January 2011; Accepted 31 May 2011

Academic Editor: Sergey Fomel

Copyright © 2011 R. Nammour and W. W. Symes. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Linearized multiparameter inversion is a model-driven variant of amplitude-versus-offset analysis, which seeks to separately account for the influences of several model parameters on the seismic response. Previous approaches to this class of problems have included geometric optics-based (Kirchhoff, GRT) inversion and iterative methods suitable for large linear systems. In this paper, we suggest an approach based on the mathematical nature of the normal operator of linearized inversion—it is a scaling operator in phase space—and on a very old idea from linear algebra, namely, Cramer's rule for computing the inverse of a matrix. The approximate solution of the linearized multiparameter problem so produced involves no ray theory computations. It may be sufficiently accurate for some purposes; for others, it can serve as a preconditioner to enhance the convergence of standard iterative methods.

## 1. Introduction

The linearized inverse problem in reflection seismology aims at recovering model perturbations from data perturbations, assuming known reference or background model. Both reference model and model perturbations consist of material parameter fields (functions of position  $x$ ) that characterize wave propagation. As the background model is considered known, we will suppress it from the notation and use the word "model" and the symbol  $m$  and similar symbols to refer to model perturbations. The linearized scattering operator  $F$  maps a model perturbation  $m$  to the corresponding perturbation  $Fm$  of the predicted data.

With these conventions, the inverse problem may be stated as follows: given observed data  $d$ , find a model perturbation  $m$  so that

$$Fm \approx d. \quad (1)$$

Interpreting (1) in a least squares sense yields the *normal equations*

$$F^*Fm = F^*d. \quad (2)$$

$F^*$  is the adjoint (transpose) of  $F$ : it is a migration operator (to be precise, the result of reverse time migration, properly construed). The migration operator maps the data space to model space. We will refer to the migration output  $m_{\text{mig}} = F^*d$  as the migrated image, acknowledging its typical role in seismic processing.  $N := F^*F$  is the normal operator, or Hessian (we will use these terms interchangeably).

An example of this setup is linear acoustics; we will treat this example explicitly in this paper. The model  $m$  consists of the velocity and density (or impedance and density, or any other equivalent combination) perturbation fields  $m = (\delta v, \delta \rho)$ , for instance. The forward map samples the solution of the linearized (Born approximation) acoustic wave equation for a source (right-hand side) indexed by source position, at receiver positions, possibly with additional filtering to emulate antenna patterns induced by source and receiver arrays.

In simulating seismic data generation with models such as linear acoustics, waves typically propagate over hundreds of wavelengths, and model fields must resolve features on the wavelength scale. The normal equations thus represent millions, or billions, of equations in the same number of unknowns, ruling out the possibility of solving them

by means of direct matrix methods such as Gaussian elimination. This paper presumes that the computations implementing the application of the normal operator are carried out by “wave equation” methods, that is, finite difference or finite element simulation. Thus each application is an expensive, large-scale computational procedure, a fact which places practical limits on the number of steps taken in an iterative scheme to solve the normal equations (2). For example, Krylov subspace methods (such as the conjugate gradient method) require at least one application of the normal operator per iteration.

We present an efficient method to approximate the solution of the normal equations that requires a few applications of the normal operator (one for one-parameter inversion, two for two-parameter inversion, and six for three-parameter inversion). This method leverages the properties of the normal operator: under some conditions (background model parameter fields slowly varying on the wavelength scale, diving wave energy eliminated from data, data polarized by propagating phases), it is a special type of matrix-valued spatially varying filter [1–9]. This particular type of spatially varying filter, called a *pseudodifferential operator* in the mathematical literature [10], preserves the discontinuities (events) in the model  $m$  to which it is applied while effecting a dip- and frequency-dependent scaling of the amplitudes. Its application also mixes events when several model parameters (density,  $p$ - and  $s$ -wave velocities, etc.) are present: thus a density event will show up in the velocity component of acoustic migration output, and viceversa, for example.

In this work, we shall show how to separate the events corresponding to various model parameters by means of several applications of the normal operator to permuted image vectors, so that the result differs from an inversion of the data by an overall spatially dependent filter, common to all components. We have previously solved the problem of estimating and correcting for such a filter [11, 12]: it is the same problem as occurs in the solution of a single-parameter linearized inverse problem. Combining the two techniques, we recover an accurate approximate inversion for all parameters, within constraints which we shall illustrate.

We review our approach to single-parameter inversion (based on constant density acoustic modeling, for instance) in Appendix B. It belongs to the genre of *scaling methods* [13–16], which approximate the Hessian or its inverse by estimating a filter of some sort from a single application of the normal operator to a migrated image or some other test image. In previous work on single-parameter inversion, the second author pointed out that the pseudodifferential nature of the Hessian makes possible, a very simple, and accurate approximation, provided that the reflectivity dip is unique at every imaging location [16]. Our approach to single-parameter inversion without assuming uniqueness of dip (Appendix B) carries this idea further; it is similar to that described by Guilton [15], but with different constraints on filter structure motivated by the theory of the pseudodifferential normal operator cited above. The core of our method is the Pseudodifferential Operator Algorithm [17], an efficient algorithm for application of filters of the required type, which we review in Appendix A. Several other authors

have explored similar pseudodifferential scaling methods based on other fast application algorithms [18, 19].

Note that data-adaptive scaling methods such as those cited in the previous paragraph differ in an essential way from scaling by an approximate diagonal of the Hessian [20, 21]. As noted by Rickett [14], a dip-independent scaling cannot well approximate an inverse Hessian in general, as the appropriate amplitude correction (from migrated image to inversion) depends on dip. On the other hand, these methods also differ from the so-called Kirchhoff or Generalized Radon Transform or Ray-Born inversion (e.g., [22–25]): both are asymptotic and approximate (either implicitly or explicitly) the normal operator as a pseudodifferential operator, but Generalized Radon Transform inversion involves explicit geometric optics computations, while scaling methods do not. Migration deconvolution methods [26] also approximate the Hessian, or its inverse, but differ from scaling methods in a similar way: migration deconvolution filters are constructed from approximate (layered) Green’s function computations, rather than by data-adaptive fitting of an *a priori* limited filter class. Several authors have explored construction of localized filter representation of the Hessian or inverse Hessian using a more general class of Green’s functions [27, 28]; again, scaling methods do not require precomputation of a Green’s function database.

It has long been recognized that the offset (or scattering angle) dependence of a reflection event encodes material parameter changes across a reflector. Approximations to this relation motivated by the analysis of reflection and transmission in layered media (the Zoeppritz equations and linearization thereof, [29, 30]) lead to amplitude-versus-offset (“AVO”) analysis techniques [31]. Linearized inversion for one or more material parameters might be viewed as a quantitative counterpart, that is, as AVO inversion, and attacked via iterative linear system solvers [32–34]. Herrmann [18] uses a single-parameter inversion approach resembling ours to precondition iterative inversion for faster convergence; we envision a similar use of our techniques. Note that migration deconvolution has also been used as an aid to quantitative AVO [35]. As already mentioned, migration deconvolution relies on layered modeling to obtain approximate Green’s functions. Pseudodifferential scaling is limited in other ways but does not require modeled wave dynamics used to construct the approximate Hessian to be close to those of layered media (and does not require the construction of Green’s functions).

Several studies have analyzed the conditioning, or error propagation properties, of multiparameter inversion [34, 36–38]. We will not discuss this important issue except in that characteristic conditioning behavior will be evident in our examples: for linearized acoustic inversion from surface data, velocity or impedance perturbations are considerably better resolved than are density perturbations.

We have used only synthetic data in the work reported here and rather simple synthetic data at that. In fact, the very few published examples of inversion (in the sense of noise-level data fit) of exploration-scale data [39–41], strongly suggest that the success of linearized multiparameter inversion is sensitive to other aspects of the inverse problem: source pulse

and radiation pattern estimation, background field approximation, modeling the physics accurately (including elasticity, attenuation, etc.), and to very careful data preprocessing. In this work, we chose to avoid these issues and focus only on the mathematical/computational issue of multiparameter inversion, given model-consistent data.

## 2. Theory and Methods

Recall that the aim of this paper is to solve the normal equations

$$Nm = m_{\text{mig}} = F^* d \quad (3)$$

for a model vector  $m = (m_1, \dots, m_n)$  consisting of  $n$  component material parameter perturbation fields  $m_i$ . The migration output  $m_{\text{mig}}$  is a vector of the same type as the model, that is, an  $n$ -vector of components, which we shall call images, as they typically contain visual evidence of reflectors.

Since both input and output of the normal operator consist of  $n$ -vectors of material parameter fields, it is natural to represent  $N$  as an  $n \times n$  block matrix of operators  $N_{ij}$  mapping material parameter fields of type  $i$  to those of type  $j$ . Since material parameter fields are positive, it is possible to use relative perturbations to parametrize the model (thus,  $\delta\rho/\rho$  for instance), in which, case  $N$  and its blocks  $N_{ij}$  are nondimensional. We chose not to nondimensionalize the problem as the techniques to be described below automatically produce results with appropriate units.

The Hessian may be treated as a matrix of spatially varying filters. The filter coefficient  $A_{ij}$  defining the  $(i, j)$  component filter  $N_{ij}$  depends on both the spatial position  $x$  and the wavevector  $k$ ,

$$N_{ij}m_j(x) = \int A_{ij}(x, k) \widehat{m}_j(k) e^{ix \cdot k} dk, \quad (4)$$

in which  $\widehat{m}_j$  denotes the spatial Fourier transform of the  $j$ th model component  $m_j$ . Denote by  $N = \text{op}(A)$  the operator defined by the matrix  $A = (A_{ij})$  of filter coefficients via (4).

Under certain conditions, the normal operator is a matrix of spatially varying filters of a special type, known as pseudodifferential, to be described below. By Beylkin [1] and Rakesh [2] established the first results of this type, which were systematically extended by Nolan and Symes [4], Ten Kroode et al. [6], De Hoop and Bleistein [5], Stolk [42], Stolk and De Hoop [43] and others. We summarize the theory as follows: the Hessian is well approximated by a pseudodifferential operator provided

- (i) the material parameters in the background model vary smoothly on the scale of a wavelength (since the theory is asymptotic in frequency, the technical assumption is that they are smooth, that is, infinitely differentiable, but the practical meaning is as stated here);
- (ii) diving wave energy is not present in the data or has been muted or dip-filtered out;
- (iii) the data has been polarized into propagating phase components.

These conditions are likely to be essential: either major reflectors in the background model or nonpolarized multiphase data lead to Hessians which produce nonphysical reflector images at some distance from their sources, hence cannot be well-approximated by their near-diagonal behavior. Since pseudodifferential operators are nearly local, in a sense to be made precise below, Hessians producing major reflector shifts cannot be approximated by them. For that matter, no near-diagonal approximation to the Hessian could be accurate in such cases, so these same limitations would seem to apply to all scaling methods.

Pseudodifferential operators are distinguished from other types of spatially varying filters by strong constraints on the filter coefficients  $A_{ij}$ , which must exhibit polynomial-like growth in the wavenumber  $k = \sqrt{k \cdot k}$ , as  $k \rightarrow \infty$ . more precisely, require that a real  $\mu$  exist so that the mixed partial derivative of order  $\alpha = \alpha_1 + \alpha_2 + \alpha_3$  grows like  $k^{\mu - \alpha}$  as  $k \rightarrow \infty$ :

$$\left| \frac{\partial^\alpha}{\partial k_1^{\alpha_1} \partial k_2^{\alpha_2} \partial k_3^{\alpha_3}} A_{ij}(x, k) \right| \leq \text{const.} \times k^{\mu - \alpha}. \quad (5)$$

That is,  $A_{ij}$  grows like  $k^\mu$  for large  $k$ , and every mixed partial derivative in  $k$  decreases the order of growth by the order of the derivative. It is also required that spatial derivatives do not increase the order of growth of the coefficient—we will not write down precisely how this additional constraint is imposed, referring the interested reader to Taylor [10] for a more complete account.

Condition (5) also describes the growth behavior of polynomials in  $k$ , which are the filter coefficients of differential operators. Other functions of  $k$ , for instance  $k$  itself, also satisfy these conditions, so the type of filter described by (5) is more general than differential operators. These considerations led to the name “pseudodifferential operator” being applied to this type of filter. In fact, it is possible to approximate any pseudodifferential operator arbitrarily well by the composition of a differential operator and a power of the Laplacian (i.e., a filter of the form  $k^\mu$ ). In this sense, pseudodifferential operators are nearly local, that is, diagonal.

By convention, filter coefficients obeying the growth rules described here are called *symbols*, and we shall use this terminology as well to distinguish these special filter coefficients from the Fourier representation of general filters. The number  $\mu$  figuring in the constraint (5) is the *order* of the operator; it determines the extent to which application of the operator changes the growth (or decay) rate for in the Fourier domain, for large wavenumber. Note that order is inclusive, rather than precise: if an operator is of order  $\mu$ , it is of order  $\nu$  for any  $\nu \geq \mu$ . However, order does differentiate pseudodifferential operators by the size of their effect on oscillatory data, just as is the case for differential operators. The work cited at the beginning of this section shows how the order of the Hessian depends on the dimension of the model and the source-receiver geometry. For example, for constant density acoustics in 2D, the order is 1; for 3D and a full range (i.e.,  $\pi$ ) of azimuths, the order is 2.

Equation (5) is a very strict constraint on the structure of the pseudodifferential class of spatially varying filters, so

it should not be a surprise that the behavior of such filters is also highly constrained. Very important in the sequel are the composition properties: if  $A$  is a  $(n \times n)$  matrix pseudodifferential operator of order  $\mu$ ,  $B$  is another of order  $\nu$ , then

- (1) the product (in other words, composition)  $AB$  is again a pseudodifferential operator of order  $\mu + \nu$ , and the symbol of the product differs from the product of the symbols by a symbol of order  $\mu + \nu - 1$ ;
- (2) consequently, if  $n = 1$  (scalar operators), then the commutator  $AB - BA$  is of order  $\mu + \nu - 1$ .

That is, *to leading order in wavenumber, the product of pseudodifferential operators is simply obtained by multiplying their symbols, and scalar operators commute*. These properties single out pseudodifferential operators amongst general spatially varying filters. Proofs of these and other facts about pseudodifferential operators may be found in Taylor [10], for example.

The theory cited above also showed that the normal operator is “partly invertible”, that is, at each point in the subsurface, the Hessian scales spatial Fourier components with a certain range of dips by a *positive* multiple of  $k^u$ . It follows from point (2) above that an operator whose symbol is the reciprocal of that of the normal operator for this range of dips, and simply set to zero with suitable tapering outside this range will function as an approximate inverse Hessian. This construction gives an approximate inversion for the single-parameter inverse Hessian (e.g., for constant-density acoustics), illustrates the utility of these very special filters, and is explained briefly in Appendix B, and in more detail in [11, 12].

Multiparameter scattering results in a matrix Hessian, mixing influences between various parameters. In some cases, the Hessian is a matrix of pseudodifferential operators. For example, the Hessian for acoustic scattering or for polarized elastic scattering ( $P$ - $P$ ,  $S$ - $S$ ) has this property. Acoustic scattering has only one ( $p$ -wave) phase, so it has this property [44], as does elastic scattering with suitably polarized data [43]. The close relation between the matrix  $A = (A_{ij})$  of symbols and the normal operator  $N$  which it defines again allows us to calculate an approximate inverse via symbol computations.

To this end, recall the definition of the *adjugate*, or classical adjoint, of an  $n \times n$  matrix  $A$ , denoted by  $\text{Adj}(A)$ : it is the transpose of the matrix of cofactors of  $A$ . Explicitly,

$$\text{Adj}(A)_{ij} = (-1)^{i+j} \det M_{ji} \quad (6)$$

in which the  $(n-1) \times (n-1)$  minor  $M_{ij}$  is obtained from  $A$  by dropping the  $i$ th row and  $j$ th column. The  $ji$  cofactor is the number on the right-hand side of (6). In our case, the transpose of the cofactor matrix may be ignored, as the matrix  $A$  and thus its adjugate are symmetric (since the Hessian  $N$  is symmetric).

The significance of the adjugate is this: when the matrix  $A$  is invertible, then the adjugate is proportional to the inverse,

$$A^{-1} := \frac{1}{\det(A)} \text{Adj}(A). \quad (7)$$

This relation is *Cramer’s rule* and is the critical observation leading to our approximation method. Note that

$$\text{Adj}(A)A = A\text{Adj}(A) = \det(A)I, \quad (8)$$

where  $I$  is the identity matrix. If we define the adjugate of  $N$  to be  $\text{Adj}(N) = \text{op}(\text{Adj}(A))$  (i.e., the matrix pseudodifferential operator with symbol matrix  $\text{Adj}(A)$ ), with slight abuse of notation, then the multiplicative order properties of pseudodifferential operators described above imply that

$$\text{Adj}(N)N \approx N\text{Adj}(N) \approx \det(N)I. \quad (9)$$

The equation above features another abuse of notation, with  $\det(N) := \text{op}(\det(A))$ . The power of (8) is revealed when applied to (3)

$$\text{Adj}(N)m_{\text{mig}} = \text{Adj}(N)Nm \approx \det(N)m. \quad (10)$$

Equation (10) reduces the inversion of  $N$  to two steps: (i) application of  $\text{Adj}(N)$ , followed by (ii) inversion of the scalar operator  $\det(N)$ , after the application of the adjugate.

Note also that while  $m_{\text{mig}}$  may be dimensionally heterogeneous,  $\text{Adj}(N)m_{\text{mig}}$  has redimensionalized all components so that their units all bear the same relation to the units of the components of  $m$ . Inversion of  $\det(N)$  removes this common unit factor and correctly dimensionalizes the model estimate.

For inversion of  $\det(N)$ , we resort to a method similar to the one we previously developed for  $n = 1$  (reviewed in Appendix B). First apply the normal operator again to the left-hand side of (10), to form

$$\begin{aligned} N\text{Adj}(N)m_{\text{mig}} &= N\det(N)m \approx \det(N)Nm \\ &= \det(N)m_{\text{mig}}. \end{aligned} \quad (11)$$

Here, we have used the fact that scalar pseudodifferential operators approximately commute with matrices of pseudodifferential operators to commute  $N$  and  $\det(N)$ .

Equation (11) shows how to find a pseudodifferential approximate inverse  $c$  to  $\det N$ , by minimizing an objective function, all parts of which are readily computable

$$\left\| m_{\text{mig}} - cN\text{Adj}(N)m_{\text{mig}} \right\|^2, \quad (12)$$

over a suitable class of pseudodifferential operators  $c$ . From (9), a solution  $c$  will approximately invert  $\det(N)$  and is scalar, since  $\det(N)$  is. Since  $c$  removes from the left-hand side of (10) a scalar filter, it plays a similar role to the scaling factors encountered in the various scaling methods mentioned in the introduction, so we will also refer to  $c$  as a scaling factor, although it is a pseudodifferential operator.

The reader will recall that  $N$  is only partially invertible for a range of model wavevectors  $k$ , the symbol determinant  $\det(A)$  is asymptotic to a positive multiple of  $k^u$  and to a lower power outside of this “illuminated” cone of model wavenumbers. The method explained in Appendix B constructs  $c$  as an approximation to the reciprocal of the symbol of  $\det(N)$  in the illuminated wavenumber cone, tapered to zero outside of this cone.

The algorithm explained in Appendix B is iterative and requires at least one application of the trial scaling factor  $c$  at each iteration. An efficient method for numerical application of pseudodifferential operators is therefore required; we use an algorithm introduced by Bao and Symes [17], which we review in Appendix A.

Having constructed  $c$ , we approximate the solution of normal equations by

$$m_{\text{inv}} = c \text{Adj}(N) m_{\text{mig}}. \quad (13)$$

Equations (12) and (13) show that computation of the adjugate action is key to this approach, so we examine it more carefully. We will address the two parameter cases ( $n = 2$ ) explicitly, both for inspiration and because it is this case for which we present examples in the next section. Writing

$$N = \begin{pmatrix} N_{11} & N_{12} \\ N_{12} & N_{22} \end{pmatrix}, \quad (14)$$

$\text{Adj}(N)$  is

$$\text{Adj}(N) = \begin{pmatrix} N_{22} & -N_{12} \\ -N_{12} & N_{11} \end{pmatrix}, \quad (15)$$

$$\text{Adj}(N)m = \begin{pmatrix} N_{22}m_1 & -N_{12}m_2 \\ -N_{12}m_1 & +N_{11}m_2 \end{pmatrix}.$$

It appears to be necessary to compute the various components of  $N$  separately: for example, if the underlying theory is acoustic scattering, then  $N_{22}$  might represent input and output density perturbations, and so on, so that each  $N_{ij}$  would appear to require a complete modeling/migration cycle.

In fact, it is possible to compute  $\text{Adj}(N)m$  using *only* applications of  $N$ , thus both reducing the cost of the application and eliminating the need to write special component-to-component migration codes. In the  $2 \times 2$  case, only one application of  $N$  is required! In fact,

$$\text{Adj}(N) = J^T N J, \quad (16)$$

here,  $J$  is the so-called symplectic matrix,

$$J = \begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}, \quad \text{so } J \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} = \begin{pmatrix} -m_2 \\ m_1 \end{pmatrix}. \quad (17)$$

Note that units are implicitly changed (as they are in the application of the adjugate operator) and that in practice it is presumed that the discrete representations of the two parameters are the same, so that, for example, density can be swapped for velocity and vice versa.

Equation (16) implies that the application of the adjugate on the migrated image requires one application of  $N$ , rather than the four that a naive implementation would suggest. Thus, approximating the inverse for  $n = 2$  requires two applications of the normal operator, plus a small amount of additional data manipulation.

The situation is more complicated for  $n = 3$ , but a similar result holds: it is possible to compute the application of the adjugate using five applications of the normal operator, a considerable improvement over the twenty-four that the naive algorithm based on the definition (6) would appear to require; note that separate application of every component  $N_{ij}$  by wave-equation methods is just as expensive as application of the entire operator  $N$ . See Nammour [38] for details. We conjecture that a similar drastic reduction is possible for general  $n$ .

As a final note, we can show explicitly for  $n = 2$  the sense in which equation (9) is an approximation, while equation (8) is exact. Denoting the commutator of two operators by  $[a, b] = ab - ba$ ,

$$\begin{aligned} \text{Adj}(N)N &= (N_{22}N_{11} - N_{12}^2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &+ \begin{pmatrix} 0 & [N_{22}, N_{12}] \\ [N_{11}, N_{12}] & [N_{11}, N_{22}] \end{pmatrix} \\ &= \det(N)I + \begin{pmatrix} 0 & [N_{22}, N_{12}] \\ [N_{11}, N_{12}] & [N_{11}, N_{22}] \end{pmatrix}. \end{aligned} \quad (18)$$

The theory of pseudodifferential operators shows that the commutator of two pseudodifferential operators is of lower order than their composition, that is, relatively negligible for highly oscillatory input fields. Using  $\approx$  to mean “differs by a lower order error”,

$$\text{Adj}(N)N \approx N \text{Adj}(N) \approx \det(N). \quad (19)$$

### 3. Numerical Examples: Layered Variable Density Acoustics

As a first application to the two parameter inversion, we construct a variable density acoustics model perturbation consisting of a thin oscillatory velocity layer and a thin oscillatory density layer in a different place (see Figures 1 and 2). The background model is homogeneous with  $v_p = 2$  km/s and  $dn = 2000$  kg/m<sup>3</sup>. The two thin layers are treated as perturbations.

We simulated reflection data for this model using the IWAVE software developed by The Rice Inversion project in linearized (Born) modeling mode [45]. The model extends around 1.7 km in depth and 6.5 km horizontally. One source is put in the middle, and receivers are laid out to create an offset ranging  $[-2.7$  km,  $2.7$  km], spaced 20 m apart and 40 m below the (absorbing) surface. The isotropic point source wavelet was a 2.5–5–15–20 Hz trapezoidal zero-phase bandpass filter. 3 seconds of data were recorded at 151 receiver positions. The boundary conditions were absorbing on all sides of the domain, in particular, free surface effects were not modeled. Since the model is layered, only a single source gather need be modeled.

All (reverse-time) migrations were also carried out with IWAVE [46]. The output is a layered model (i.e., two depth traces, one for each velocity and density perturbations), which we shall display as layered fields.

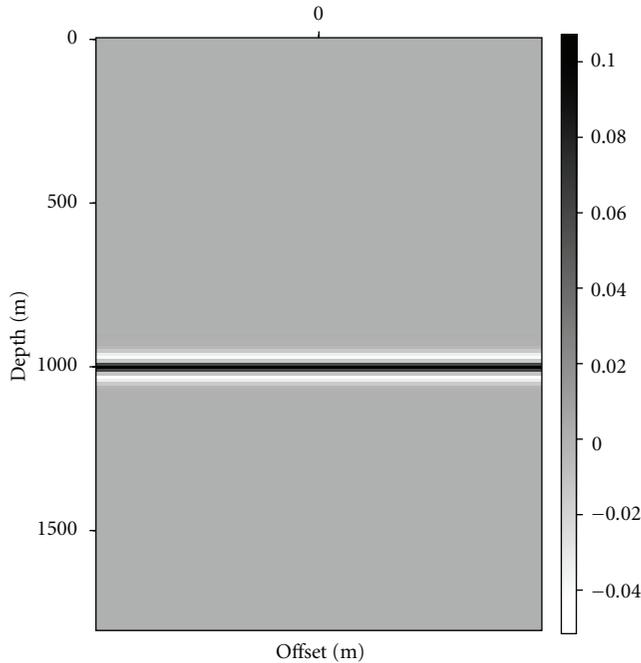


FIGURE 1: vp, velocity perturbation.

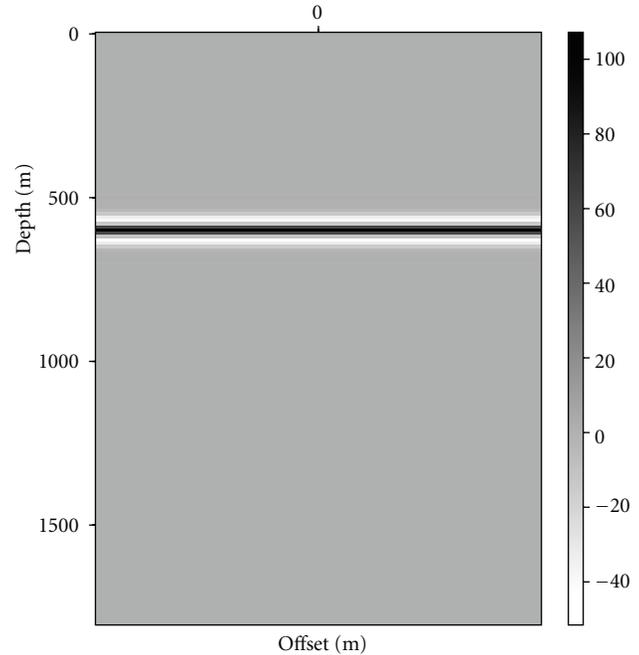


FIGURE 2: dn, density perturbation.

Migrating the Born data shows how migration mixes the effects of the two models in the two components of the migrated images (Figure 3). We shall refer to the migrated images as  $m_{\text{mig}_1}$  and  $m_{\text{mig}_2}$  to remain consistent with our notation where the vector of migrated images is  $m_{\text{mig}}$ . This example, albeit simple, stresses a new challenge of multiparameter inversion. For one parameter inversion, the events in the migrated image corresponded to events in the true model. In multiparameter inversion, events in the migrated images may correspond to an event in one or more of the components of the model. It is virtually impossible to tell that these migrated images correspond to a model with separate events for velocity and density without successful inversion.

Applying the scheme outlined above, we form

$$J^T N J m_{\text{mig}} \approx \det(N) m. \quad (20)$$

The result is shown in Figure 4 and shows how one application of the normal operator effectively separated the contributions of the velocity and density events. It remains to effect an amplitude correction by approximating an inverse to  $\det(N)$ . For this end, we are required to form  $N \det(N) m \approx \det(N) m_{\text{mig}}$ , shown in Figure 5.

The final step corrects the amplitudes of  $\det(N) m$  by undoing the effect of  $\det(N)$ , which yields an approximate inverse. This final step complements the separation we obtained earlier with an amplitude correction. Figure 6 shows that the approximate inverse compares favorably with true model. An interesting observation on this result is the fact that the velocity model is better recovered than the density model: traces of the velocity event in the density model are more apparent than those of the density event in the velocity model. This observation is in accordance with the

theoretical fact that the recovery of velocity in variable density acoustics is better conditioned than the recovery of density. The inversion is successful in that the inverted model succeeds in fitting 70% of the target data (see Figure 7).

#### 4. Discussion and Conclusion

We have presented a method to approximate the solution of the multiparameter linearized inverse problem an extension of Cramer's rule for matrices of pseudodifferential operators. The method consists of two steps:

- (1) reducing the multiparameter problem to a one parameter problem, which yields an amplitude scaling of the solution;
- (2) correcting the amplitudes of the result from the previous step to approximate the solution.

The application of the work flow to a layered variable density acoustics example shows how the effects of different material parameters, indistinguishable in the migrated images, are separated in the first step. The amplitude correction step successfully yields an approximate solution to the linearized inverse problem.

The work flow presented above applies without modification to any model. Results of application to models more complex than the layered model described above will be presented elsewhere. We point out that other aspects of the physics of seismic data generation will need to be accommodated if this (or any other) algorithm is to extract accurate results from field data. For example, as shown already by Minkoff and Symes [39], even when the Born approximation is adequate, the radiation pattern of the source has a first-order effect on the variation of amplitude

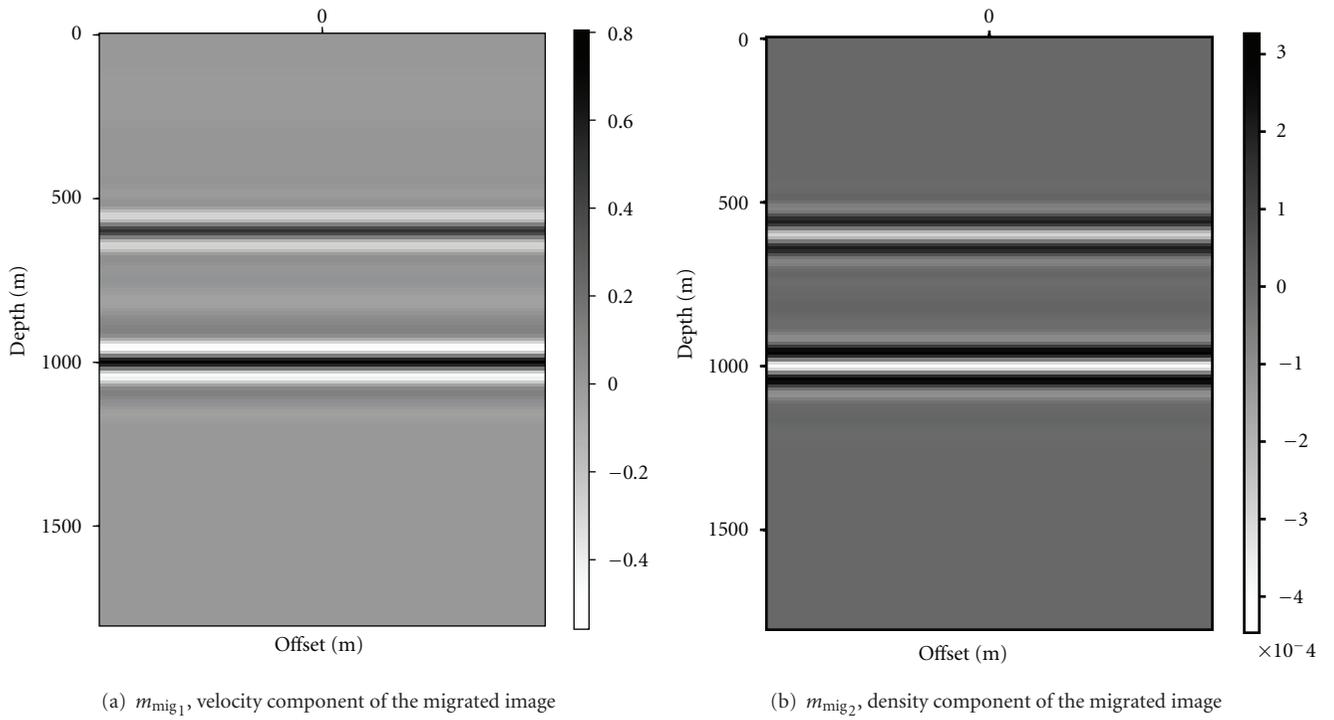


FIGURE 3: Migrated images mixing the contributions from density and velocity, and effecting a phase space scaling.

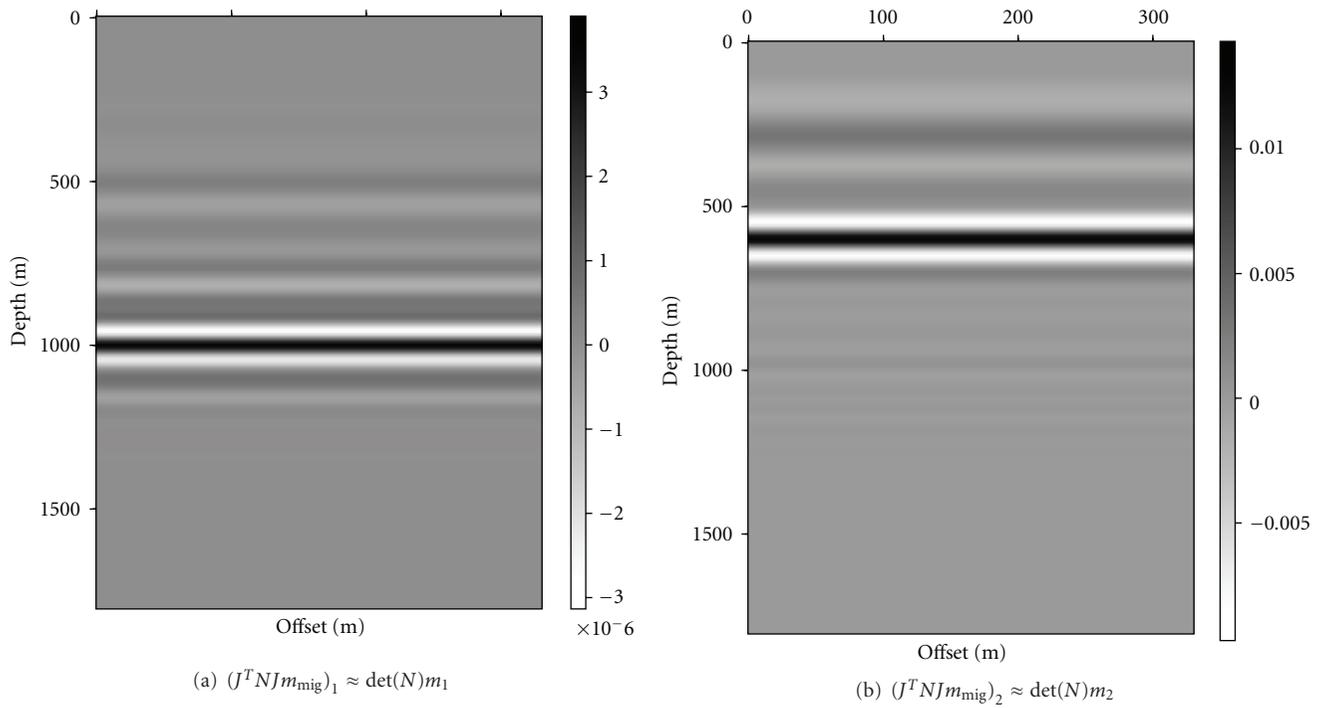


FIGURE 4: The application of the adjugate separates the velocity and density contributions. This result is a phase space scaling of the true model.

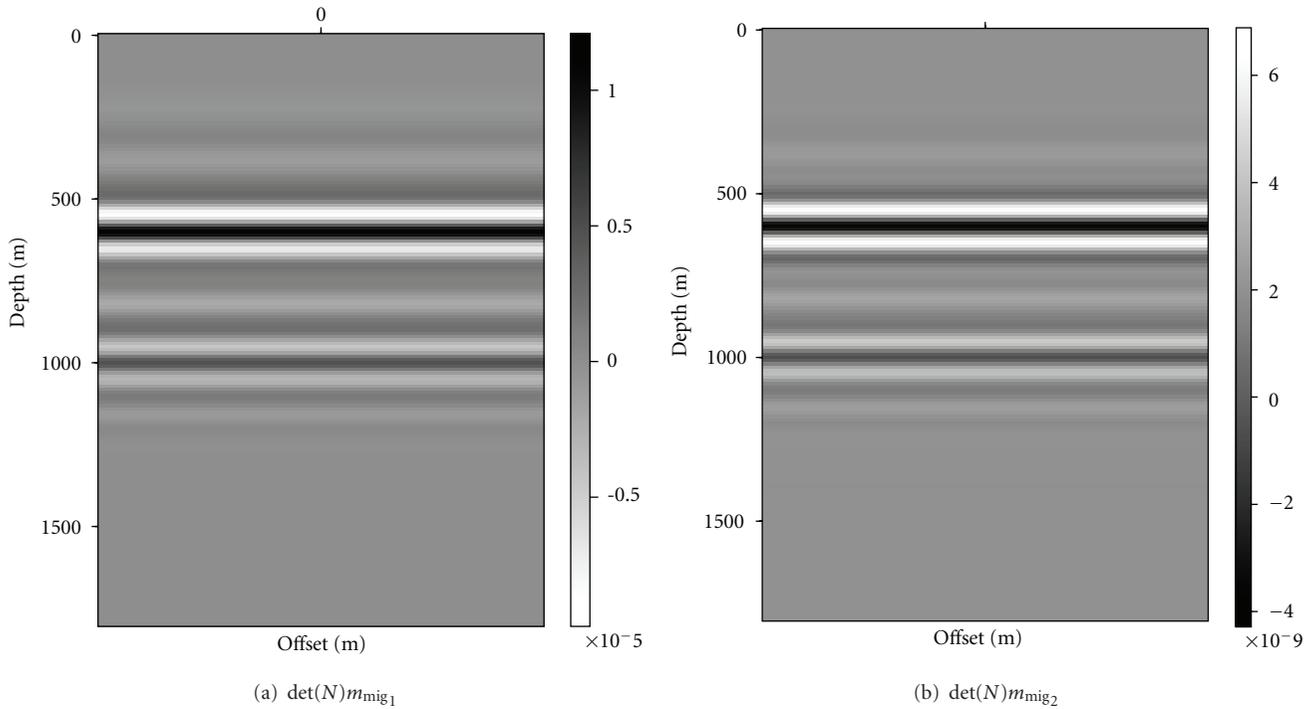


FIGURE 5: Scaling of the migrated images by  $\det(N)$ , used to undo the determinant.

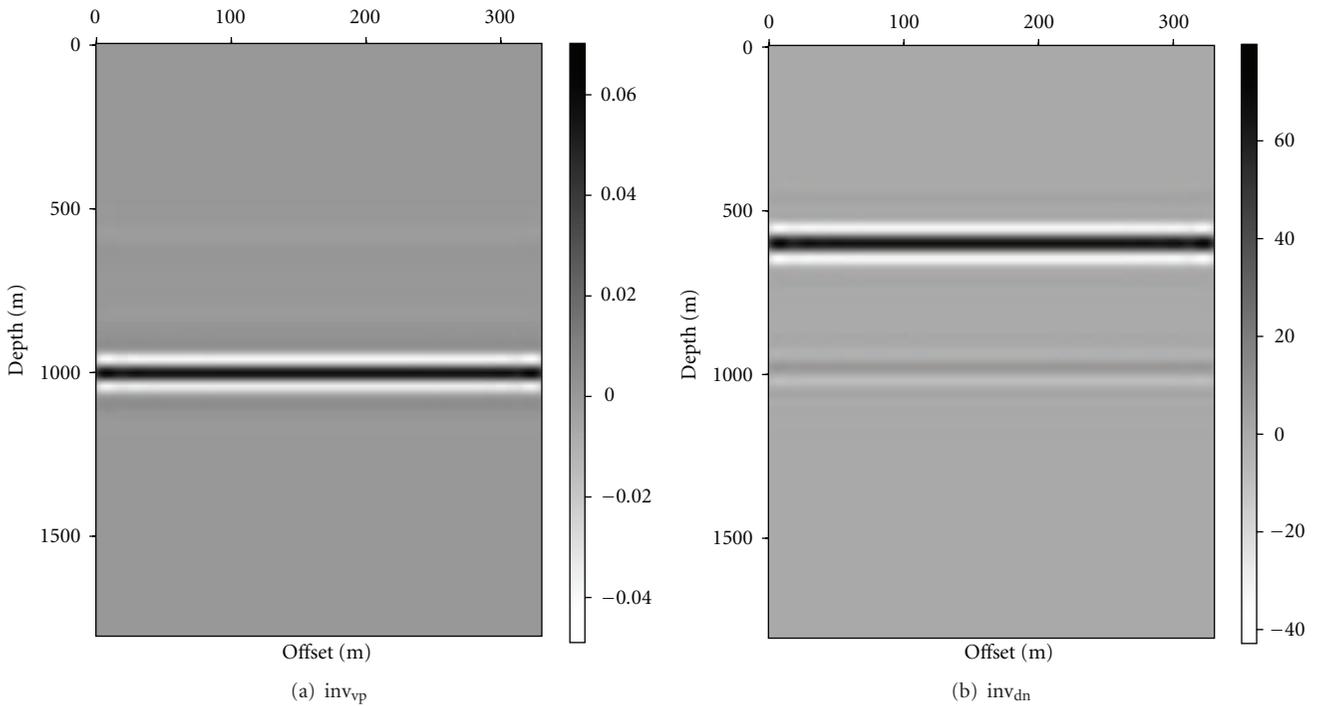


FIGURE 6: The approximate inverse. The contributions from velocity and density are separated, and the amplitudes are corrected.

with angle and must therefore be included in the parameters to be estimated.

We have extended the Pseudodifferential Operator Algorithm (Appendix A) to three spatial dimensions (3D) which allows approximate inversion of 3D models. The symbol is

represented by a truncated spherical harmonics expansion. The formulation of the method in this case remains intact, though the cost of modeling and migration increases. This extension of the method is described in detail, and examples are given in Nammour [38].

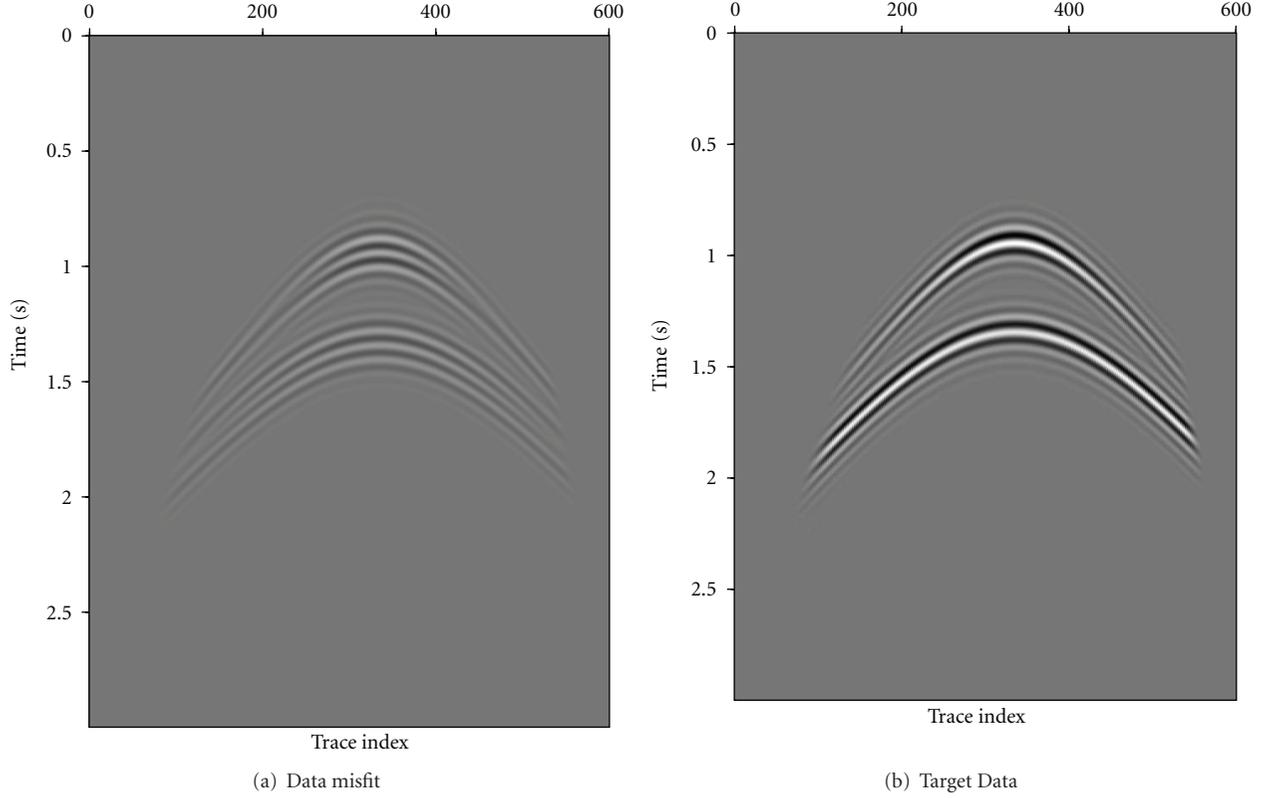


FIGURE 7: Data misfit versus target data. The inverted model fits 70% of the data.

The explicit derivation analogous to  $n = 2$  for more parameters (namely  $n = 3$ ) is similar to the one presented here but requires tedious algebraic manipulations. The case  $n = 3$  becomes relevant in the linearized inverse problem for linear elasticity, for example. Six applications of the normal operator are required to approximate the solution of the linearized inverse problem. See Nammour [38] for a detailed discussion.

We end by reminding the reader that the approach to approximate inversion pursued here relies on the validity of the pseudodifferential characterization of the Hessian. As mentioned in the theory section, two of the three conditions (background model smooth on the wavelength scale, polarized data) are essential for *any* near-diagonal Hessian approximation, adaptive or not, to be effective. These limitations seem likely to be acceptable in some exploration settings, for example, inversion of marine data for supersalt sediment properties with soft sea floor, but may seriously obstruct this type of technique in other settings.

## Appendices

### A. The Pseudodifferential Operator Algorithm

Bao and Symes [17] presents an algorithm to numerically apply pseudodifferential operators, which we use to represent the scaling factor  $c$  in (12). We will discuss this algorithm in

this section as it is central to the feasibility of the method presented here.

This discussion is restricted to 2D, so we may write  $\mathbf{x} = (x, z)$ . Recall that a pseudodifferential operator is characterized by its symbol and defined by

$$Q_m u(x, z) = \iint q_m(x, z, k_x, k_z) \hat{u}(k_x, k_z) e^{i(xk_x + zk_z)} dk_x dk_z, \quad (\text{A.1})$$

where  $q_m(x, z, k_x, k_z)$  is the principal symbol, homogeneous of degree  $m$ , and  $\hat{u} = \mathcal{F}[u]$  is the Fourier transform of  $u$ .

Thus, writing  $k_x = \omega \cos \theta$ ,  $k_z = \omega \sin \theta$  and using the homogeneity of  $q_m$ , we have

$$q_m(x, z, k_x, k_z) = \omega^m \tilde{q}_m(x, z, \theta). \quad (\text{A.2})$$

Notice that  $\tilde{q}_m(x, z, \theta) = q_m(x, z, \cos \theta, \sin \theta)$  is periodic and smooth in  $\theta$ , and hence it admits a rapidly converging Fourier expansion. We thus truncate the Fourier series, approximating the symbol by its first  $K + 1$  Fourier modes,

$$\tilde{q}_m(x, z, \theta) \approx \sum_{l=-K/2}^{l=K/2} c_l(x, z) e^{il\theta} = \sum_{l=-K/2}^{l=K/2} \omega^{-l} c_l(x, z) (k_x + ik_z)^l. \quad (\text{A.3})$$

Plugging (A.3) into (A.1), we obtain

$$Q_m u(x, z) \approx \sum_{l=-K/2}^{l=K/2} c_l(x, z) \mathcal{F}^{-1} \left[ \omega^{m-l} (k_x + ik_z)^l \hat{u}(k_x, k_z) \right]. \quad (\text{A.4})$$

Fourier transform theory identifies  $\omega^{m-l}$  as the symbol of  $(-\nabla)^{(m-l)/2}$ , and  $k_x$  and  $k_z$  are, respectively, the symbols of  $D_x = -i\partial_x$  and  $D_z = -i\partial_z$ .

Sampling the field  $u(x, z)$  and the symbol  $\tilde{q}_m(x, z, \theta)$ ,

$$\begin{aligned} U_{ij} &= u(x_0 + (i-1)\Delta x, z_0 + (j-1)\Delta z), \\ Q_{ijk} &= \tilde{q}_m(x_0 + (i-1)\Delta x, z_0 + (j-1)\Delta z, k\Delta\theta), \\ i &= 1, \dots, M, \quad j = 1, \dots, N, \quad k = -\frac{K}{2}, \dots, \frac{K}{2}. \end{aligned} \quad (\text{A.5})$$

Choosing  $\Delta k_x = 1/(M-1)\Delta x$  and  $\Delta k_z = 1/(N-1)\Delta z$  yields the unaliased discretizations of the symbols of the square root of the negative Laplacian,  $D_x$  and  $D_z$

$$\begin{aligned} \Omega_{\text{pr}} &= 2\pi\sqrt{(p\Delta k_x)^2 + (r\Delta k_z)^2}, \\ X_{\text{pr}} &= 2\pi p\Delta k_x, \\ Z_{\text{pr}} &= 2\pi r\Delta k_z, \\ p &= -\frac{M}{2}, \dots, \frac{M}{2}, \quad r = -\frac{N}{2}, \dots, \frac{N}{2}. \end{aligned} \quad (\text{A.6})$$

Equation (A.4) suggests the following algorithm to estimate  $Q_m u$  [17]. All Fourier transforms refer to a discrete Fourier transform.

- (1) Compute  $\hat{U}_{\text{pr}} = \mathcal{F}[U_{ij}]$ .
- (2) For each  $i \in [1, M]$  and  $j \in [1, N]$ , compute  $\hat{Q}_{ij} = \left\{ \hat{Q}_{ijl} \right\}_{l=-K/2}^{K/2}$  the discrete Fourier transform of  $Q_{ij} = \left\{ Q_{ijk} \right\}_{k=-K/2}^{K/2}$ .
- (3) Initialize  $(QU)_{ij} = 0$ , for  $i \in [1, M]$ ,  $j \in [1, N]$ , For  $l = -K/2 : K/2$

(a) compute  $\{R_{ij}^l\}_{i=1, j=1}^{M, N} = \mathcal{F}^{-1}[\Omega_{\text{pr}}^{m-l}(X_{\text{pr}} + iZ_{\text{pr}})^l \hat{U}_{\text{pr}}]$  for  $p = -M/2, \dots, M/2$  and  $r = -N/2, \dots, N/2$ ,

(b) accumulate  $(QU)_{ij} = (QU)_{ij} + \hat{Q}_{ijl} R_{ij}^l$

End

A straightforward discretization of (A.1) has a computational complexity of  $O(N^4 \log(N))$ . The algorithm described above uses FFT (Fast Fourier Transform) and thus exhibits a complexity of  $O(KN^2(\log(N) + \log(K)))$ . The appeal of this approach is that  $K$  is independent of  $N$ . In fact, applications to reflection seismology require that the symbol be smooth and slowly varying in  $\theta$ , thus, may be captured accurately by a modest number of Fourier modes or, more explicitly, a small  $K$ .

The dependence on dip is captured in the angle variable  $\theta$ , and the method allows us to capture multiple dip events by increasing  $K > 1$ .

## B. The One-Parameter Case

In this appendix, we review the method developed in [11, 12] to construct an approximate inverse Hessian in the one-parameter case. The aim is to solve

$$Nm = m_{\text{mig}}, \quad (\text{B.1})$$

where  $m_{\text{mig}} = F^* d \in \text{Range}(N)$ . We seek a pseudo differential scaling factor and formulate its recovery as an optimization problem. Given the migrated image  $m_{\text{mig}}$  and the remigrated image  $Nm_{\text{mig}}$ ,

$$c = \underset{c \in \Psi\text{DO}}{\text{argmin}} \left\| m_{\text{mig}} - cNm_{\text{mig}} \right\|^2. \quad (\text{B.2})$$

The scaling factor  $c$  is chosen from a class of pseudodifferential operators represented numerically using the PsiDO algorithm (Appendix A). In this setting, the scaling factor approximates the action of the inverse of the normal operator on the migrated image  $m_{\text{mig}}$ . More specifically,

$$\begin{aligned} m &= N^{-1}m_{\text{mig}} \approx N^{-1}cNm_{\text{mig}} \approx cN^{-1}Nm_{\text{mig}} \\ &= c m_{\text{mig}} := m_{\text{inv}}. \end{aligned} \quad (\text{B.3})$$

The first of these equations expresses the true solution  $m$ , the second approximate equality follows from because pseudodifferential operators approximately commute. Defining  $m_{\text{inv}} := cm_{\text{mig}}$  thus yields an approximation to the true model  $m$ . Equation (B.3) shows that the scaling factor approximates the action of the inverse on the normal operator on the migrated image, and it is only in that sense that  $c$  approximates  $N^{-1}$ .

The scaling factor  $c$  is represented explicitly by the PsiDO algorithm,

$$\begin{aligned} cNm_{\text{mig}}(x, z) &\approx \sum_{l=-K/2}^{l=K/2} c_l(x, z) \mathcal{F}^{-1} \left[ \omega^{m-l}(k_x + ik_z)^l \widehat{Nm_{\text{mig}}}(k_x, k_z) \right]. \end{aligned} \quad (\text{B.4})$$

We enforce the continuity of  $c_l(x, z)$  using a parsimonious basis technique. Let  $\{\psi_j(x, z)\}_{j=1}^J$  be a set of smooth shape functions (cubic b-splines, for example). Write

$$c_l(x, z) = \sum_{j=1}^J c_l^j \psi_j(x, z), \quad (\text{B.5})$$

and plug into the objective function,

$$\begin{aligned} J &= \left\| m_{\text{mig}} - cNm_{\text{mig}} \right\|^2 \\ &= \sum_{x_i, z_k} \left| m_{\text{mig}}(x_i, z_k) - (cNm_{\text{mig}})(x_i, z_k) \right|^2. \end{aligned} \quad (\text{B.6})$$

We have denoted by  $J$  the objective function in (B.2), and used the fact that the images  $m_{\text{mig}}$  and  $Nm_{\text{mig}}$  are discretized on a grid  $\{x_i\}_{i=1}^N$ ,  $\{z_k\}_{k=1}^M$  to write out the norm explicitly.

We use limited memory BFGS (LBFGS) to minimize the objective function (B.2), which requires the gradient to be supplied by the user. The gradient is easily derived as

$$\begin{aligned} \frac{\partial J}{\partial c_j^l} = & 2 \sum_{x_i, z_k} \left\{ \left( m_{\text{mig}}(x_i, z_k) - (cN m_{\text{mig}})(x_i, z_k) \right) \psi_j(x_i, z_k) \right. \\ & \times \left( \mathcal{F}^{-1} \left[ \omega^{m-l} (k_x + ik_z)^l \widehat{N m_{\text{mig}}}(k_x, k_z) \right] \right) \\ & \left. \times (x_i, z_k) \right\}. \end{aligned} \quad (\text{B.7})$$

We have also explored enforcing the positivity of the scaling factor explicitly and other symmetries of the problem, which yields to slightly different representations of the scaling factor  $c$  and in turn different objective function gradient. For more details on this point and a discussion about the ability of the method to resolve multiple dip events please consult [12].

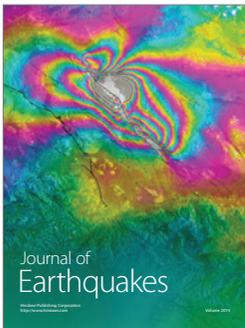
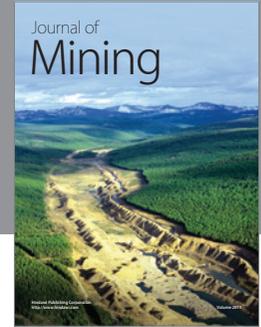
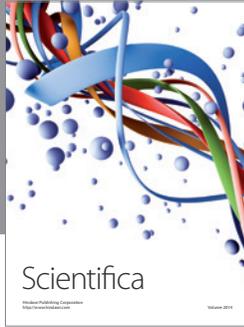
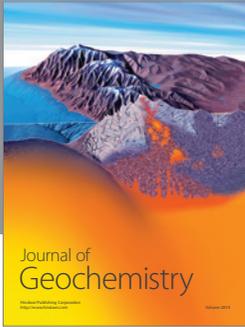
## Acknowledgments

This work was supported in part by the National Science Foundation under grant DMS 0620821, and by the sponsors of The Rice Inversion Project. The authors are grateful to an anonymous reviewer and to Guest Editor Sergey Fomel for thoughtful critiques.

## References

- [1] G. Beylkin, "Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized Radon transform," *Journal of Mathematical Physics*, vol. 26, no. 1, pp. 99–108, 1985.
- [2] Rakesh, "A linearized inverse problem for the wave equation," *Communications on Partial Differential Equations*, vol. 13, pp. 573–601, 1988.
- [3] G. Beylkin and R. Burridge, "Linearized inverse scattering problems in acoustics and elasticity," *Wave Motion*, vol. 12, no. 1, pp. 15–52, 1990.
- [4] C. J. Nolan and W. W. Symes, "Global solution of a linearized inverse problem for the wave equation," *Communications in Partial Differential Equations*, vol. 22, no. 5-6, pp. 919–952, 1997.
- [5] M. V. de Hoop and N. Bleistein, "Generalized Radon transform inversions for reflectivity in anisotropic elastic media," *Inverse Problems*, vol. 13, no. 3, pp. 669–690, 1997.
- [6] A. P. E. Ten Kroode, D. J. Smit, and A. R. Verdel, "A microlocal analysis of migration," *Wave Motion*, vol. 28, no. 2, pp. 149–172, 1998.
- [7] R. Burridge, M. V. de Hoop, D. Miller, and C. Spencer, "Multi-parameter inversion in anisotropic elastic media," *Geophysical Journal International*, vol. 134, no. 3, pp. 757–777, 1998.
- [8] C. C. Stolk, *On the modeling and inversion of seismic data*, Ph.D. thesis, Universiteit Utrecht, 2000.
- [9] S.-K. Foss, B. Ursin, and M. V. de Hoop, "Depth-consistent P- and S-wave velocity reflection tomography using PP and PS seismic data," in *Proceedings of the 53rd SEG Annual International Meeting*, Expanded Abstracts, pp. 2363–2367, Society of Exploration Geophysicists, 2004.
- [10] M. Taylor, *Pseudodifferential Operators*, Princeton University Press, 1981.
- [11] R. Nammour and W. Symes, "Approximate constant density acoustic inverse scattering using dip-dependent scaling," in *Proceedings of the 80th SEG Annual International Meeting*, Expanded Abstracts SI 3.6, Society of Exploration Geophysicists, 2009.
- [12] R. Nammour, "Approximate inverse scattering using pseudo-differential scaling," Tech. Rep. 09-09, Department of Computational and Applied Mathematics, Rice University, Houston, Tex, USA, 2009.
- [13] J. F. Claerbout and D. Nichols, "Spectral preconditioning," Tech. Rep. 82, Stanford Exploration Project, Stanford University, Stanford, Calif, USA, 1994.
- [14] J. E. Rickett, "Illumination-based normalization for wave-equation depth migration," *Geophysics*, vol. 68, no. 4, pp. 1371–1379, 2003.
- [15] A. Guitton, "Amplitude and kinematic corrections of migrated images for nonunitary imaging operators," *Geophysics*, vol. 69, no. 4, pp. 1017–1024, 2004.
- [16] W. W. Symes, "Approximate linearized inversion by optimal scaling of prestack depth migration," *Geophysics*, vol. 73, no. 2, pp. R23–R35, 2008.
- [17] G. Bao and W. W. Symes, "Computation of pseudo-differential operators," *SIAM Journal on Scientific Computing*, vol. 17, no. 2, pp. 416–429, 1996.
- [18] F. J. Herrmann, P. Moghaddam, and C. C. Stolk, "Sparsity- and continuity-promoting seismic image recovery with curvelet frames," *Applied and Computational Harmonic Analysis*, vol. 24, no. 2, pp. 150–173, 2008.
- [19] L. Demanet, P. Létourneau, N. Boumal, H. Calandra, J. Chiu, and S. Snelson, "Matrix probing: a randomized preconditioner for the wave-equation Hessian," 2011, <http://math.mit.edu/icg/papers/matrix-probing-seismic.pdf>.
- [20] C. Shin, S. Jang, and D.-J. Min, "Improved amplitude preservation for prestack depth migration by inverse scattering theory," *Geophysical Prospecting*, vol. 49, no. 5, pp. 592–606, 2001.
- [21] U. Jang, D. J. Min, and C. Shin, "Comparison of scaling methods for waveform inversion," *Geophysical Prospecting*, vol. 57, no. 1, pp. 49–59, 2009.
- [22] M. F. Sullivan and J. K. Cohen, "Prestack kirchhoff inversion of common-offset data," *Geophysics*, vol. 52, no. 6, pp. 745–754, 1987.
- [23] N. Bleistein, "On the imaging of reflectors in the earth," *Geophysics*, vol. 52, no. 7, pp. 931–942, 1987.
- [24] M. S. Operto, S. Xu, and G. Lambare, "Can we quantitatively image complex structures with rays?" *Geophysics*, vol. 65, no. 4, pp. 1223–1238, 2000.
- [25] S.-K. Foss, M. V. de Hoop, and B. Ursin, "Linearized 2.5-dimensional parameter imaging inversion in anisotropic elastic media," *Geophysical Journal International*, vol. 161, no. 3, pp. 722–738, 2005.
- [26] J. Yu, J. Hu, G. T. Schuster, and R. Estill, "Prestack migration deconvolution," *Geophysics*, vol. 71, no. 2, pp. S53–S62, 2006.
- [27] G. Chavent and R.-E. Plessix, "An optimal true-amplitude least-squares prestack depth-migration operator," *Geophysics*, vol. 64, no. 2, pp. 508–515, 1999.
- [28] A. A. Valenciano, B. Biondi, and A. Guitton, "Target-oriented wave-equation inversion," *Geophysics*, vol. 71, no. 4, pp. A35–A38, 2006.
- [29] K. Aki and P. Richards, *Quantitative Seismology: Theory and Methods*, Freeman, 1980.
- [30] R. T. Shuey, "Simplification of the Zoeppritz equations," *Geophysics*, vol. 50, no. 4, pp. 609–614, 1985.

- [31] F. Hilterman, *Seismic Amplitude Interpretation*, Society of Exploration Geophysicists, 2001.
- [32] O. Gauthier, A. Tarantola, and J. Virieux, "Two-dimensional nonlinear inversion of seismic waveforms: numerical results," *Geophysics*, vol. 51, no. 7, pp. 1387–1403, 1986.
- [33] A. Bourgeois, B. F. Jiang, and P. Lailly, "Linearized inversion: a significant step beyond pre-stack migration," *Geophysical Journal International*, vol. 99, no. 2, pp. 435–445, 1989.
- [34] S. Jin, R. Madariaga, J. Virieux, and G. Lambare, "Two-dimensional asymptotic iterative elastic inversion," *Geophysical Journal International*, vol. 108, no. 2, pp. 575–588, 1992.
- [35] J. Yu, J. Hu, and M. V. de Hoop, "Prestack migration deconvolution+AVO inversion," in *Proceedings of the 73rd SEG Annual International Meeting*, Expanded Abstracts, pp. 1051–1054, Society of Exploration Geophysicists, 2003.
- [36] F. Santosa and W. W. Symes, "High-frequency perturbational analysis of the surface point-source response of a layered fluid," *Journal of Computational Physics*, vol. 74, no. 2, pp. 318–381, 1988.
- [37] G. J. M. Lörtzer and A. J. Berkhout, "Linear avo inversion of multicomponent seismic data," vol. 8, pp. 967–972, 1989, SEG Technical Program Expanded Abstracts.
- [38] R. Nammour, "Approximate multi-parameter inverse scattering using pseudodifferential scaling," Tech. Rep. 11-05, Department of Computational and Applied Mathematics, Rice University, Houston, Tex, USA, 2011.
- [39] S. E. Minkoff and W. W. Symes, "Full waveform inversion of marine reflection data in the plane-wave domain," *Geophysics*, vol. 62, no. 2, pp. 540–553, 1997.
- [40] M. Charara, C. Barnes, and A. Tarantola, "Constrained waveform inversion of seismic well data," in *Inverse Methods*, B. Jacobsen, K. Mosegaard, and P. Sibani, Eds., vol. 63 of *Lecture Notes in Earth Sciences*, pp. 98–112, Springer, Heidelberg, Germany, 1996.
- [41] M. Charara and C. Barnes, "Nonlinear inversion of seismic waveforms: a NorthSea offset VSP example," in *Proceedings of the 81st Annual International Meeting of Expanded Abstracts*, pp. 3966–3970, Society of Exploration Geophysicists, 2010.
- [42] C. C. Stolk, "Microlocal analysis of a seismic linearized inverse problem," *Wave Motion*, vol. 32, no. 3, pp. 267–290, 2000.
- [43] C. C. Stolk and M. V. de Hoop, "Microlocal analysis of seismic inverse scattering in anisotropic elastic media," *Communications on Pure and Applied Mathematics*, vol. 55, no. 3, pp. 261–301, 2002.
- [44] W. Symes, "Mathematics of reflection seismology," 1998, <http://www.trip.caam.rice.edu>.
- [45] D. Sun and W. W. Symes, "IWAVE implementation of Born simulation," Tech. Rep. 10-05, Department of Computational and Applied Mathematics, Rice University, Houston, Tex, USA, 2010.
- [46] D. Sun and W. W. Symes, "IWAVE implementation of adjoint state method," Tech. Rep. 10-06, Department of Computational and Applied Mathematics, Rice University, Houston, Tex, USA, 2010.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

