Geophys. J. Int. (1996) **126,** 345–359

A computationally feasible approximate resolution matrix for seismic inverse problems

Susan E. Minkoff*

The Rice Inversion Project, Department of Computational and Applied Mathematics, Rice University, PO Box 1892, *Houston* TX 77251-1892, *USA*

Accepted 1996 February 29. Received 1996 February 13; in original form 1995 October 7

SUMMARY

Seismic inversion produces model estimates which are at most unique in an *average* sense. The model resolution matrix quantifies the spatial extent over which the estimate averages the true model. Although the resolution matrix has traditionally been defined in terms of the singular value decomposition of the discretized forward problem, this computation is prohibitive for inverse problems of realistic size. Inversion requires one to solve a large normal matrix system which is best tackled by an iterative technique such as the conjugate gradient method. The close connection between the conjugate gradient and Lanczos algorithms allows us to construct an extremely inexpensive *approximation* to the model resolution matrix. Synthetic experiments indicate the data dependence of this particular approximation. The approximation is very good in the vicinity of large events in the data. Two large linear viscoelastic inversion experiments on p–τ marine data from the Gulf of Mexico provide estimates of the elastic parameter reflectivities corresponding to two different seismic sources. Traditionally, one evaluates the accuracy of the two reflectivity estimates by comparing them with measured well logs. The approximate model resolution matrices agree with the well-log ranking of the two models and provide us with a way to compare different model estimates when, for example, such well-log measurements are not available.

Key words: inversion, seismic resolution, source time functions, viscoelasticity.

INTRODUCTION

In a classic paper, Backus & Gilbert (1968) gave a description of the general linear inverse problem and defined model resolution. They explained that solutions of seismic inverse problems are generally not pointwise unique but may still have unique average behaviour. Resolution quantifies how close the model estimates are to the true model as a function of spatial location (such as depth). It indicates the shortest length-scale which the given data can resolve. [References include papers by Franklin (1970), Jackson (1979), Kennett & Nolet (1978), and the book by Menke (1989).] Wiggins (1972) defines this matrix in terms of the singular value decomposition (SVD) and analyses resolution for a very small inversion problem (approximately 10 unknowns) using observed data from surface waves and free oscillations. He comments that, 'computation of parameter and information resolution is such a simple extension of any inversion procedure based on perturbation parameters that such inversion studies are incomplete without considering resolution.' Unfortunately, forming the resolution matrix via the SVD has not become common practice because it is computationally prohibitive for large problems.

There are a number of papers in which SVD resolution is computed for small inverse problems. For example, using waveform inversion and a viscoelastic simulator, Martinez & McMechan (1991) examine model resolution for a simple geometry experiment of a target layer located between two homogeneous half-spaces. For field data experiments they do not compute resolution. Bishop *et al.* (1985) analyse the resolution of components of a simplified tomography model analytically, and substantiate their conclusions with a small numerical experiment performed on real data.

There have also been attempts to compute the full SVD for realistic (large) inverse problems [see, for example, Assous & Collino (1990)]. Ory & Pratt (1995) examine the effect of different regularizing operators on resolution. They consider a synthetic experiment of determining three 1-D anisotropic velocity parameters from traveltime data, and state (p. 420):

One practical problem that arises is the cost of computing the model resolution matrix, either because of too large a number of data, or too large a number of parameters ... or both. It took about 3 hours CPU time ... to compute the matrix in our

*Now at: Texas Institute for Computational and Applied Mathematics, Taylor Hall 2.316, University of Texas, Austin, TX 78712, USA.

© 1996 RAS

Downloaded from https://academic.oup.com/gji/article-abstract/126/2/345/623352 by University of Texas at Dallas - McDermott Library user on 22 November 2019

one-dimensional anisotropic parameter problem (576 parameters and 1420 data). For larger problems, this computation becomes prohibitive.

Vasco, Pulliam & Johnson (1993) used a massively parallel computer to determine a model of *P*-wave velocity variations of the Earth's mantle from traveltime data. They formed the full resolution and covariance matrices. [See also the paper by Pulliam & Stark (1994) and the reference lists located in both papers.]

Recently, a number of authors have suggested computing partial (approximate) resolution information. Scales (1989) outlined a method for constructing the matrices needed for the Lanczos eigenvalue estimation algorithm from the conjugate gradient linear system solver. I use the same connection in this paper. He did not, however, form the model resolution matrix from the resulting approximate eigen information. Nolet & Snieder (1990) gave a theoretical overview of the application of Paige & Saunders' (1982) least-squares QR decomposition (LSQR) algorithm to the continuous inverse problem, but did not give any numerical examples. Berryman described both a Lanczos method (1994a) and an LSQR-based method (1994b) for computing resolution. The Lanczos technique is very similar to the method discussed in this paper. Although he gave a small (4×4) seismic tomography example of LSQR, he did not give a numerical example of the Lanczos method. Zhang & McMechan (1995) applied LSQR to a larger set of synthetic tomography problems.

In this paper, we approximately solve the inverse problem using the conjugate gradient technique and simultaneously estimate some of the eigenvalues and eigenvectors of the normal matrix via the closely related Lanczos algorithm. This subset of eigenvectors allows us to form a partial resolution matrix at almost no additional cost over solving the inverse problem. The size of the problem does not restrict our ability to compute this approximation to resolution. One can apply this technique, without modification, to any type of inverse problem where resolution should be analysed.

In the next sections, I give a brief description of model resolution, the conjugate gradient, and the Lanczos algorithms. This discussion includes a list of the steps taken to form the model resolution approximation. I apply the technique to synthetic and field data full waveform inversion tests of p–τ marine reflection data. The small acoustic synthetic example illustrates that, at least in some instances, this approximation converges extremely quickly to the true resolution matrix. The rate of convergence of the approximation improves near large events in the data. This hypothesis is reasonable because the Lanczos search directions depend on the data.

In the field data experiments I use a viscoelastic model for wave propagation in the Earth, and invert for the three elastic parameter reflectivities that linearly influence the data. The two inversion experiments differ only in the seismic sources used. I include resolution diagrams for the three parameters estimated in the two experiments. Detailed descriptions of the forward model, inversion procedure, and data are given in the Appendix. The resolution study of these two experiments agrees with independent measurements (for example, well-log comparisons to inversion results), indicating that one of the two sources (the inversion-estimated source) provides more accurate elastic parameter estimates than the other (air-gun model source).

THE MODEL RESOLUTION DATA

We wish to quantify our ability to resolve the individual model parameters from the forward model for the seismogram G. It is necessary to devise a generalized inverse or *estimator* G^{-g} which will act on the data and return an estimate of the model parameters. The simplest way to define the model resolution matrix is to assume that there is a set of model parameters $m^{\text{true}} \in \mathscr{R}^n$ which satisfies the equation $Gm^{\text{true}} = d^{\text{obs}}$, for the forward modeller $G \in \mathscr{R}^{m \times n}$ and observed data $d^{\text{obs}} \in \mathscr{R}^m$. The resolution matrix (Menke 1989) measures the averaging of the model m^{true} which enters into the model estimate m^{est} in the inversion process:

$$m^{\text{est}} = G^{-g} d^{\text{obs}} = G^{-g} G m^{\text{true}} \equiv R m^{\text{true}}, \tag{1}$$

where $R \in \mathscr{R}^{n \times n}$ is the *model resolution matrix*. If $R = I$ then $m^{\text{est}} = m^{\text{true}}$, and the model parameters are perfectly resolved in the inversion. In general, $R \neq I$, and then the model estimates m^{est} are weighted averages of the true parameters $m_i^{\text{est}} = a^{\mathrm{T}} m^{\text{true}}$. Note that this averaging vector a corresponds to a column of the resolution matrix R. In the experiments I will describe, the model parameters depend on one spatial coordinate only, namely depth (z), and the averaging vectors correspond to averages over depths.

The singular value decomposition allows us to decompose the forward operator into the matrix product $G = U\Sigma V^{\mathrm{T}}$ (Golub & Van Loan 1989, page 71). The columns of U are eigenvectors of GG^{T} (the left singular vectors of G), and the columns of V are the eigenvectors of $G^{\mathrm{T}}G$ (the right singular vectors of G). Finally, the p singular values on the diagonal of $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_p)$ are the square roots of the non-zero eigenvalues of GG^{T} and $G^{\mathrm{T}}G$ (the singular values). Of importance to us in this paper is the fact that V's columns are eigenvectors which span the model space M.

Only the p non-zero singular values contribute to the model resolution matrix, so we work instead with the truncated SVD defined by $G = U_p \Sigma_p V_p^{\mathrm{T}}$ where U_p and V_p^{T} consist of the first p columns of U and V respectively, and Σ_p is a $p \times p$ diagonal matrix. We note that, although $V_p^{\mathrm{T}} V_p = I$, in general, since the p right singular vectors do not span the whole space, $V_p V_p^{\mathrm{T}} \neq I$.

A version of the estimator G^{-g} can now be expressed in terms of the truncated singular value decomposition, $G^{-g} = V_p \Sigma_p^{-1} U_p^{\mathrm{T}}$, and the model resolution matrix [from expression (1)] is simply

$$R = G^{-g} G = (V_p \Sigma_p^{-1} U_p^{\mathrm{T}})(U_p \Sigma_p V_p^{\mathrm{T}}) = V_p V_p^{\mathrm{T}}. \tag{2}$$

In my implementation, I do not explicitly compute the SVD. Instead, I calculate an approximate resolution matrix by estimating some of the eigenvalues and eigenvectors of the normal matrix through the Lanczos iterative procedure, described in the next section.

DISCUSSION OF THE CONJUGATE GRADIENT AND LANCZOS ALGORITHMS

The normal matrices typical of seismic inverse problems are large, so we prefer to use iterative methods to solve these linear systems approximately. If the normal operator were symmetric and positive-definite, an appropriate technique for solving this system would be the conjugate gradient algorithm. [In reality, these inverse problems often produce normal matrices with large nullspaces (see Fig. 4 for the synthetic

Downloaded from https://academic.oup.com/gji/article-abstract/126/2/345/623352 by University of Texas at Dallas - McDermott Library user on 22 November 2019

example)]. Therefore, applying an iterative technique such as conjugate gradients (which we terminate after a small number of iterations) regularizes the problem. See, for example, Hanke (1995) and the excellent reference list therein.

The conjugate gradient algorithm provides the parameters needed to construct the tridiagonal matrix for the Lanczos iteration (Scales 1989). The extreme eigenvalues of this tridiagonal matrix approximate those of the original matrix (in our case, the normal matrix). The corresponding eigenvectors of the normal matrix can be found from the eigenvectors of the tridiagonal matrix by multiplying by the Lanczos matrix (changing bases). The eigenvalues provide information about the condition of the normal matrix. The eigenvectors of the normal matrix (right singular vectors of the forward operator) allow us to measure the model resolution.

The next two subsections describe the conjugate gradient and Lanczos algorithms in general terms. The classic references for these two algorithms are papers by Hestenes (1980), Hestenes & Stiefel (1952) and Lanczos (1950). Golub & O'Leary (1989) give a bibliographical reference history of the two methods.

The conjugate gradient idea

[The following general description of the conjugate gradient algorithm is taken from Golub & Van Loan (1989, pp. 516–520).] The Hestenes–Stiefel conjugate gradient algorithm can be understood in the context of minimizing the function $\phi(x)$ defined by

$$\phi(x) = \frac{1}{2}x^{\mathrm{T}}Ax - x^{\mathrm{T}}b, \tag{3}$$

where $b \in \mathscr{R}^n$, and the matrix $A \in \mathscr{R}^{n\times n}$ is assumed to be positive-definite and symmetric. The minimizer of ϕ is $x = A^{-1}b$. So, minimizing the function ϕ and solving the linear system $Ax = b$ are seen to be equivalent problems.

One choice for decreasing the function ϕ is to travel in the negative gradient direction, $-\nabla\phi(x_c) = b - Ax_c$, from the current point x_c. One notices that the negative gradient direction is the residual direction r_c of the system at the current point. Unfortunately, as is well known, this method (steepest descent) may converge extremely slowly if the condition of the system (or ratio of the largest to smallest eigenvalues) is large. The conjugate gradient algorithm, therefore, chooses to minimize ϕ in a set of directions $\{p_1, p_2, \ldots\}$ which do not necessarily correspond to the residual directions. One approach with clear benefits is to choose linearly independent directions p_i so that each x_j solves

$$\min_{x\in \mathrm{span}\{p_1,\ldots,p_j\}} \phi(x). \tag{4}$$

This choice of search directions ensures finite termination of the algorithm in at most n steps. We would like a vector p_j such that, when we solve the 1-D minimization problem

$$\min_{\alpha} \phi(x_{j-1} + \alpha p_j), \tag{5}$$

we also solve the j-dimensional problem (4). Luckily, such a solution is possible if we require the directions p_j to be *A-conjugate* to the previous directions $p_1, \ldots, p_{j-1}$. The vectors $p_1, \ldots, p_j$ are A-conjugate if $P_{j-1}^{\mathrm{T}}Ap_j = 0$. The search directions satisfy the following theorem (Golub & Van Loan 1989, p. 521):

Theorem: After k iterations of the conjugate gradient algorithm, we have

$$\mathrm{span}\{p_1, \ldots, p_j\} = \mathrm{span}\{r_0, \ldots, r_{j-1}\} = \mathrm{span}\{b, Ab, \ldots, A^{j-1}b\}.$$

The Lanczos idea and connection to the conjugate gradient algorithm

[This description of the Lanczos algorithm can be found in more detail in Golub & Van Loan (1989, pp. 476–480).] The Lanczos algorithm, when applied to a symmetric matrix $A \in \mathscr{R}^{n\times n}$, generates a sequence of tridiagonal matrices $T_j \in \mathscr{R}^{j\times j}$ with extreme eigenvalues which are progressively better estimates of the extreme eigenvalues of A.

One way to motivate the Lanczos idea is to recall the Rayleigh quotient which can be used to approximate the eigenvalues of a matrix A. Let λ_1 be the largest eigenvalue of A and λ_n the smallest. For $Q_j = [q_1, \ldots, q_j]$, a matrix in $\mathscr{R}^{n\times j}$ with orthonormal columns, we define the scalars M_j and m_j by

$$M_j = \max_{y\neq 0} \frac{y^{\mathrm{T}}(Q_j^{\mathrm{T}}AQ_j)y}{y^{\mathrm{T}}y} \leq \lambda_1(A), \tag{6}$$

$$m_j = \min_{y\neq 0} \frac{y^{\mathrm{T}}(Q_j^{\mathrm{T}}AQ_j)y}{y^{\mathrm{T}}y} \geq \lambda_n(A). \tag{7}$$

The Lanczos algorithm provides a way to compute the q_j so that the scalars M_j and m_j are better and better estimates of $\lambda_1(A)$ and $\lambda_n(A)$. Let $x = Q_j y$. Then the Rayleigh quotient changes most rapidly in the direction of its gradient, which is a vector contained in $\mathrm{span}\{x, Ax\}$. For this reason, the Lanczos vectors $\{q_i\}_1^j$ are chosen to be an orthonormal basis for the Krylov subspace (Golub & Van Loan 1989, p. 477):

$$\kappa(A, q_1, j) \equiv \mathrm{span}\{q_1, Aq_1, \ldots, A^{j-1}q_1\} = \mathrm{span}\{q_1, \ldots, q_j\}. \tag{8}$$

At the jth iteration of the Lanczos algorithm we have a matrix $Q_j \in \mathscr{R}^{n\times j}$ (the Lanczos matrix) whose columns are the normalized residuals resulting from the conjugate gradient algorithm (which can be shown to be orthonormal) and a symmetric, tridiagonal matrix $T \in \mathscr{R}^{j\times j}$. In fact, the Lanczos matrix 'tridiagonalizes' the matrix A up to an error matrix (Golub & Van Loan 1989, pp. 477–478):

$$AQ_j = Q_jT_j + r_je_j^{\mathrm{T}}. \tag{9}$$

The entries in T are combinations of the parameters generated in the conjugate gradient iteration (for details see the Algorithm subsection below).

Definition of the approximate model resolution matrix

As stated in the section on model resolution, we are solving the problem $Gm = d$ where the forward operator $G \in \mathscr{R}^{m\times n}$. Assuming that $\mathrm{rank}(G) = n$, we may instead solve the normal system

$$G^{\mathrm{T}}Gm = G^{\mathrm{T}}d. \tag{10}$$

Let $A \equiv G^{\mathrm{T}}G$ and $b \equiv G^{\mathrm{T}}d$. Calling the model solution x (rather than m), eq. (10) can be written in more standard notation as $Ax = b$. Rather than calculating the full model resolution matrix defined by the SVD (computationally prohibitive), we approximate this matrix via the following steps.

Downloaded from https://academic.oup.com/gji/article-abstract/126/2/345/623352 by University of Texas at Dallas - McDermott Library user on 22 November 2019

(1) Solve the system $Ax = b$ approximately by performing j steps of the combined conjugate gradient/Lanczos algorithms. The solution x_j is the one which comes closest to solving the system $Ax = b$ in the Krylov subspace spanned by $\{b, Ab, \ldots, A^{j-1}b\}$. Similarly, the j eigenvalue estimates from the Lanczos process are optimal for this subspace.

(2) Decide on an acceptable tolerance level for the error in the eigenvalue approximations. From the j approximate eigenvectors, determine the k consecutive eigenvectors that correspond to the eigenvalues with approximation error $\leq$ tolerance.

(3) Form the matrix V whose columns are these k approximate eigenvectors. The remaining $j - k$ eigenvectors are deemed too erroneous to be included in V. The matrix V has dimension $V \in \mathscr{R}^{n \times k}$.

(4) Form the approximate resolution matrix $R_{\text{lanc}} \equiv VV^{\text{T}}$. $R_{\text{lanc}} \in \mathscr{R}^{n \times n}$.

It must be stressed that this resolution matrix is not the (traditional) resolution matrix $R = G^{-g}G$. In the examples described in the section on plane-wave marine field data experiments, I invert for three reflectivity parameters each of which is described by 626 points, so the model m consists of nearly 2000 points. The conjugate gradient algorithm provides a good estimate of the solution with many fewer iterations, however. In fact, I performed only 30 iterations of the combined conjugate gradient/Lanczos algorithm. Furthermore, I dropped the six or seven smallest eigenvectors from the matrix V because of approximation error.

Algorithm

A pseudocode version of the two algorithms (conjugate gradient and Lanczos) follows.

Variables used

A: normal operator
b: data
r: residual
x_0: starting solution
x: approximate solution
p: conjugate gradient direction
β: parameter used in computation of new direction p
α_c step length in current direction p
α_p step length in previous direction
rtr_c: inner product of current residual with itself
rtr_p: inner product of previous residual with itself
tol: relative residual tolerance used for determining algorithm convergence
Q: Lanczos matrix
T: tridiagonal matrix resulting from Lanczos process with eigenvalues approximating those of A
Z: matrix of eigenvectors of the tridiagonal matrix T
X: matrix of approximate eigenvectors of the original matrix A.

Algorithm (conjugate gradient/Lanczos)

If $A \in \mathscr{R}^{n \times n}$ is symmetric and positive-definite and $b \in \mathscr{R}^n$, then the following algorithm computes $x \in \mathscr{R}^n$ so that $Ax = b$. The algorithm also optionally approximates some of the eigenvalues and eigenvectors of the matrix A. The notation $\|\cdot\|$ indicates norm and $\langle \cdot, \cdot \rangle$ the associated inner product for that space. The algorithm is not restricted to models in L^2.

```
initialize:
  r = b − Ax_0
  x = x_0
for j = 1 :iteration limit
  if eigenvector flag = true
    Q(:, j) = r/‖r‖
  end if
  if j = 1
    β = 0
    p = r
    rtr_c = ⟨r, r⟩
  else
    β = rtr_c/rtr_p
    p = r + βp
  end if
  ap = Ap
  ptap = ⟨p, ap⟩
  if ptap < tol
    break
  end if
  α_c = rtr_c/ptap
  x = x + α_c p
  r = r − α_c ap
  if eigenvalue flag = true
    if j = 1
      T(j, j) = 1/α_c
      T(j, j − 1) = 0
      T(j − 1, j) = T(j, j − 1)
    else
      T(j, j) = rtr_c/(rtr_p α_p) + 1/α_c
      T(j, j − 1) = −√(rtr_c/rtr_p)/α_p
      T(j − 1, j) = T(j, j − 1)
    end if
    Call LAPACK routine to get eigenvalues/vectors (Z) of
    T (Anderson et al. 1992).
    Compute error in approximate eigenvalue for normal
    operator.
  end if
  rtr_p = rtr_c
  rtr_c = ⟨r, r⟩
  α_p = α_c
  if √rtr_c < tol
    break
  end if
  if eigenvector flag = true
    QZ = X
  end if
end
```

A QUALITATIVE LOOK AT APPROXIMATION ERROR—A SYNTHETIC EXPERIMENT

I used the same layered medium, primaries only, plane-wave forward simulator for both the synthetic and field data experiments. However, for the synthetic experiment I modified the simulator input to represent constant-density acoustic rather than viscoelastic wave propagation (which was used for the

Downloaded from https://academic.oup.com/gji/article-abstract/126/2/345/623352 by University of Texas at Dallas - McDermott Library user on 22 November 2019

field data experiments). This simpler model suffices to indicate when the approximation is accurate and when it is not.

Let $c(z)$ be the depth-dependent sound velocity. The reflectivity is then the short-scale heterogeneity in the velocity, or $r(z) = \delta c(z)/c(z)$. The layered-medium assumption (that the velocity depends only on depth) allows us to apply the Radon integral transform (or plane-wave decomposition) to the normal displacement obtained from the solution to the acoustic wave equation. Thus one can reduce the 3-D problem to a family of 1-D equations (Treitel, Gutowski & Wagner 1982). These equations are parametrized by slowness, p. By assuming a primaries-only or single-scattering approximation and by using high-frequency asymptotics, we can write the convolutional equation for the seismogram:

$$d^{\text{pred}}(t, p) = f(t) * \tilde{r}(t, p).$$

Here, '*' denotes convolution in time t, and f is the isotropic source (assumed to have point support). The expression for the reflectivity as a function of time, $\tilde{r}$ (or perturbation of the Green's function for the acoustic wave equation), is given by

$$\tilde{r}(t, p) \approx \int dz[A(z)r(z, p)]\delta(t - 2\tau),$$

where A is the reflectivity amplitude from geometric optics. The vertical (plane-wave) velocity changes form when we Radon transform, becoming

$$v(z, p) = c(z)/\sqrt{1 - c^2(z)p^2}.$$

Similarly, the reflectivity $r(z, p)$ is now the relative perturbation in this transformed velocity. The traveltime function is

$$\tau(z, p) = \int_0^z d\zeta \frac{1}{v(\zeta, p)}.$$

The plane-wave data used in the experiments (see Fig. 1) were generated from the isotropic source shown in Fig. 2 and the reflectivity shown in Fig. 3, dashed line. These two model parameters are based on field data parameters used in the next section, namely the first trace of the anisotropic air-gun model source (Fig. 11) and an unscaled version of the estimated reflectivity (Fig. 20, solid line). The reflectivity has been truncated to zero outside of the depth interval of 2000–2400 m. The model is discretized using 100 grid-points. The data has 13 plane-wave traces with slowness values ranging from 0.116 ms m^{-1} to 0.365 ms m^{-1}.

Because I chose to invert for a reflectivity defined over a small depth interval, I was able to compute the full SVD model resolution matrix and to compare this matrix with the approximation. The conjugate gradient algorithm solved the normal equations (with a relative error of 0.5 per cent) in eight iterations. Fig. 3 shows the target reflectivity (dashed curve) and the inversion result after eight iterations (solid curve). The normal operator singular values are shown in Fig. 4. (Note the very large nullspace.) Figs 5, 6 and 7 show three columns of the 100-column model resolution matrix after eight combined conjugate gradient–Lanczos iterations. Fig. 5 is the column of the resolution matrix corresponding to a depth of 2120 m. Fig. 6 is the column corresponding to 2180 m depth, and Fig. 7 corresponds to 2240 m depth. Each graph has three curves. The dashed line is the appropriate column of the full SVD resolution matrix. The solid line is the same column of the Lanczos-estimated resolution matrix formed from the equation $R_{\text{lanc}} = VV^{\text{T}}$, where V has eight columns each of

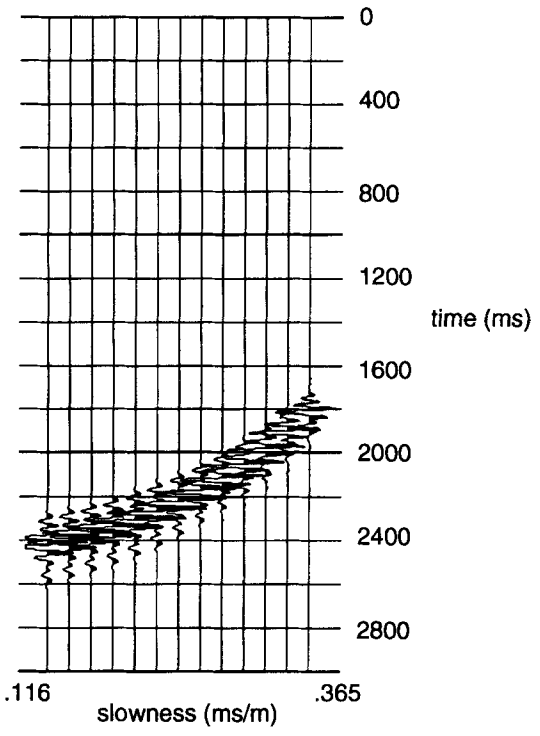

Figure 1. The common-midpoint data gather used for the synthetic experiment (13 plane-wave traces generated from the isotropic source function shown in Fig. 2 and the acoustic impedance shown in Fig. 3 dashed line).

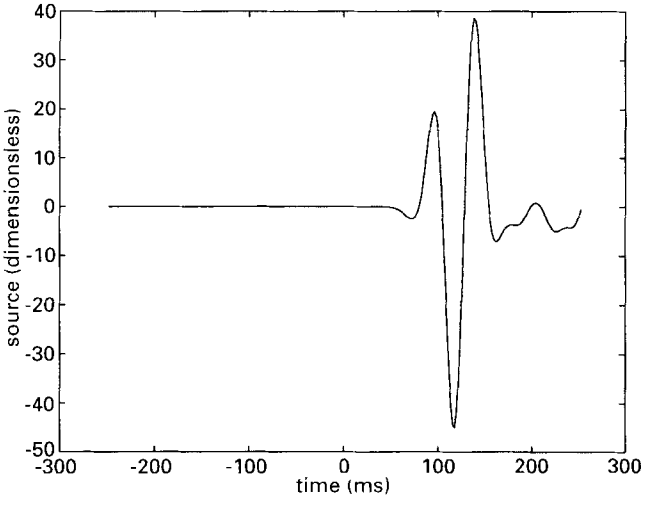

Figure 2. The isotropic (15 Hz peak frequency) source which generated the data shown in Fig. 1.

length 100 (the eight normal operator eigenvectors estimated from the Lanczos process). The crossed line is the same column of a second partial resolution matrix R_{psvd} formed from the eight eigenvectors corresponding to the eight largest eigenvalues of the full SVD. Although I at first believed the resolution matrix approximation R_{lanc} would be close to the partial SVD resolution matrix R_{pavd}, there are marked differences between the two approximations. At the target location of 2180 m depth (Fig. 6), the Lanczos approximation (R_{lanc}) converges to the full SVD resolution kernel (R) more quickly than does the partial resolution matrix (R_{psvd}) formed from the eight largest

Downloaded from https://academic.oup.com/gji/article-abstract/126/2/345/623352 by University of Texas at Dallas - McDermott Library user on 22 November 2019

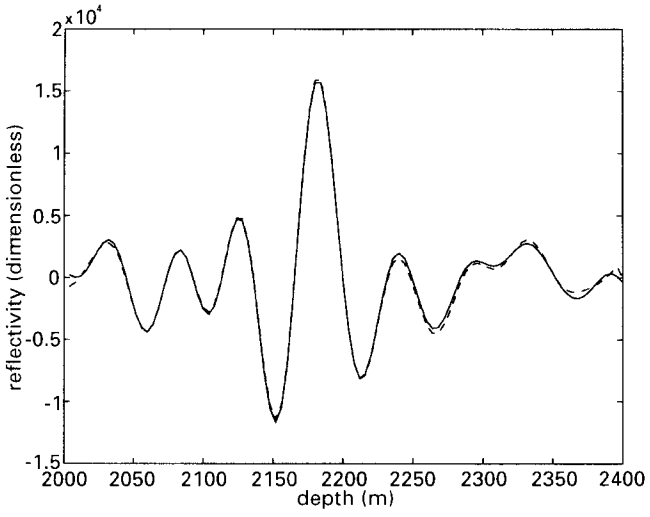

Figure 3. Solid curve: acoustic impedance inversion result after eight iterations. Dashed curve: acoustic impedance target which generated the data shown in Fig. 1.

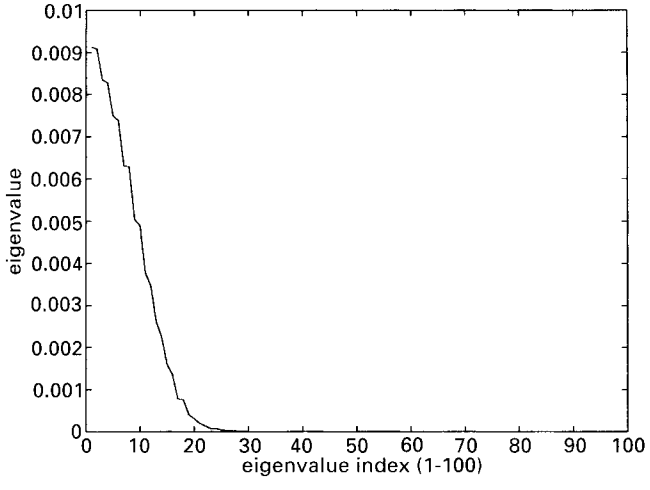

Figure 4. The complete eigenvalue spectrum of the normal operator for the synthetic experiment.

exact eigenvectors. Away from the target location, the Lanczos estimate does slightly worse. I observed the same behaviour after only five iterations of the conjugate gradient–Lanczos algorithm (stopping tolerance of 1 per cent), as well as in other synthetic experiments.

The conjugate gradient solution x_j is optimal for $x_j \in \kappa(A, b, j) \equiv \text{span}\{b, Ab, \ldots, A^{j-1}b\}$. The version of the Lanczos algorithm I used is extremely inexpensive because the necessary inputs to the algorithm come from the conjugate gradient algorithm directly. Thus, the eigen estimates must be data-dependent as well.

PLANE-WAVE MARINE FIELD DATA EXPERIMENTS

Modelling, inversion and the data

In this section I apply the conjugate gradient and Lanczos algorithms just described to the linear inverse problem of estimating three elastic parameter reflectivities in the visco-

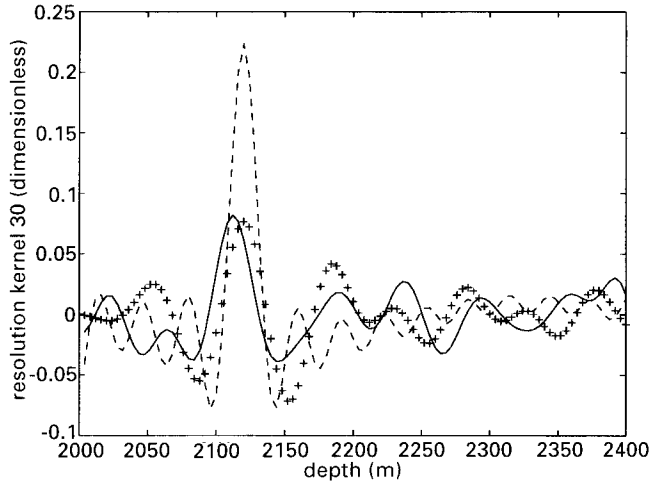

Figure 5. Dashed curve: column 30 (corresponding to a depth of 2120 m) of the full SVD model resolution matrix. Solid curve: Lanczos estimate of the same column using eight iterations of the combined conjugate gradient–Lanczos procedure (eight eigenvectors). Crossed line: SVD estimate of the same column using the eigenvectors corresponding to the eight largest eigenvalues from the full SVD of the normal matrix.

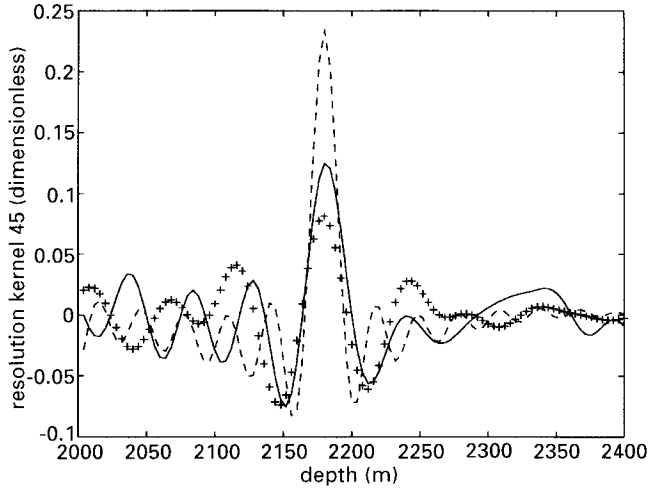

Figure 6. Dashed curve: column 45 (corresponding to a depth of 2180 m) of the full SVD model resolution matrix. Solid curve: Lanczos estimate of the same column using eight iterations of the combined conjugate gradient–Lanczos procedure (eight eigenvectors). Crossed line: SVD estimate of the same column using the eigenvectors corresponding to the eight largest eigenvalues from the full SVD of the normal matrix.

elastic model for wave propagation in the Earth. I will compare the resolution obtained for reflectivities from two different experiments. These experiments were performed on a real marine common-midpoint data gather close to a logged well. In Experiment 1, an air-gun model anisotropic source was used to estimate the three reflectivities. In Experiment 2, the reflectivities and an anisotropic source were estimated via linearized inversion. All other inputs to the model were fixed, having been estimated by inversion or some other means prior to these two experiments.

Most of the details of the forward modelling and inversion procedures are described in Appendices A and B. The geometry

Downloaded from https://academic.oup.com/gji/article-abstract/126/2/345/623352 by University of Texas at Dallas - McDermott Library user on 22 November 2019

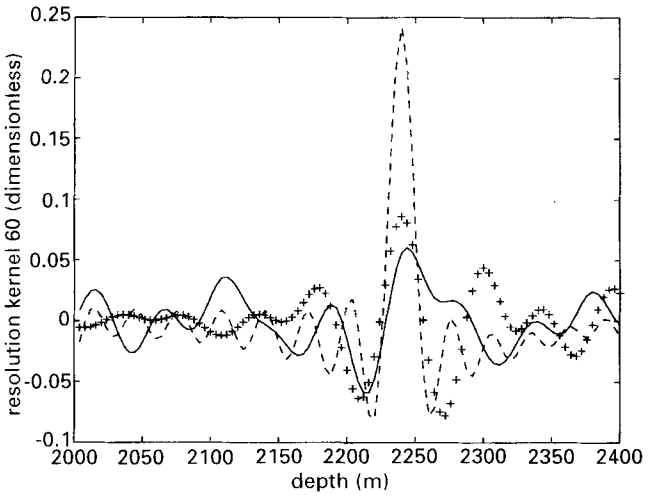

Figure 7. Dashed curve: column 60 (corresponding to a depth of 2240 m) of the full SVD model resolution matrix. Solid curve: Lanczos estimate of the same column using eight iterations of the combined conjugate gradient–Lanczos procedure (eight eigenvectors). Crossed line: SVD estimate of the same column using the eigenvectors corresponding to the eight largest eigenvalues from the full SVD of the normal matrix.

of the physical marine experiment as well as data preparation are given in Appendix C. [For a reference to a related inversion study of the same marine data, see Minkoff & Symes (1996).] In summary, we model the Earth as a viscoelastic medium. The viscoelastic simulator we used was built on the solution approximation given in the book by Aki & Richards (1980, pp. 153–155). We make three assumptions; namely, that the Earth is a layered medium, that we will only model primary reflections, and that the source is high frequency relative to the frequency content of the seismic data. These assumptions lead to the viscoelastic convolutional model for the seismogram:

$$d^{\text{pred}}(t, p) = f(t, p) * \tilde{r}(t, p).$$

In the above expression, d^{pred} is the predicted seismic data, f the anisotropic source wavelet, p denotes slowness, and t time. The '*' symbol is convolution in time. The time-dependent reflectivity $\tilde{r}$ is approximated by the expression

$$\tilde{r}(t, p) \approx \int dz [A_P(z, t, p) r_P(z) + A_S(z, t, p) r_S(z) + A_D(z, t, p) r_D(z)].$$

The depth-dependent P-wave velocity, S-wave velocity, and density reflectivities are the high-frequency quantities $r_P = \delta v_P / v_P$, $r_S = \delta v_S / v_S$, and $r_D = \delta \rho / \rho$ respectively. These reflectivity terms are normalized with respect to the corresponding background or low-frequency quantities. The geometric optics reflectivity amplitudes A_P, A_S and A_D are defined in Appendix A.

The predicted seismic data d^{pred} is linear in the seismic source f, and the elastic reflectivities r_P, r_S, r_D separately. The inversion technique used is Output Least Squares (or OLS). This method requires that we adjust the inversion parameters r_P, r_S, r_D and f to minimize the mean-squared error

$$J_{\text{OLS}} = \| d^{\text{pred}}(t, p) - d^{\text{obs}}(t, p) \|^2,$$

where $d^{\text{obs}}(t, p)$ is the 'observed' p–τ data and $\| \cdot \|$ is the L^2 norm.

The data used in this work were derived from a marine survey of the Gulf of Mexico. This area of the Gulf contains a strong gas-sand-related direct hydrocarbon indicator at about 2.3 s (see Fig. 8). The data were Radon transformed, respecting 3-D cylindrical symmetry, to yield 48 plane-wave traces per midpoint gather. Slowness values ranged from $p_{\min} = 0.116$ ms m^{-1} to $p_{\max} = 0.365$ ms m^{-1}.

The two numerical experiments described in this section were performed on a single midpoint gather (Fig. 9) located near a logged well. For the well near this midpoint gather, we obtained block sonic and density (gamma ray) logs in the 1.4–2.6 s (two-way time) interval. In the next subsection we describe how we used these logs in assessing the accuracy of the inversions and as a check on the resolution results.

As the number of reflectivity samples is moderately large (approximately 2000 total for the experiments described here), it is natural to use an iterative minimization scheme. While the choice of source wavelet influenced the rate of convergence, 30 iterations of the conjugate gradient algorithm were generally sufficient to reduce the normal residual to close to 1 per cent of its starting value. In both experiments, the initial estimates of the reflectivities were zero. The background P-wave velocity was estimated via non-linear differential semblance optimization (Kern & Symes 1994) and fixed in these experiments (see Fig. 10).

Air-gun modelling software gave an *a priori* estimate of the source signature and radiation pattern. Fig. 11 shows this air-gun model source over the range of slowness values in the data, with every fourth trace displayed for clarity. This source was used in Experiment 1. The inversion-estimated source determined in Experiment 2 is shown in Fig. 18. Note that all the Experiment 1 figures are shown grouped together first in the paper. Then the figures for Experiment 2 are shown.

Experimental results

In Experiment 1, the seismic source (an air-gun model) was not updated in the inversion. The three elastic parameter reflectivities

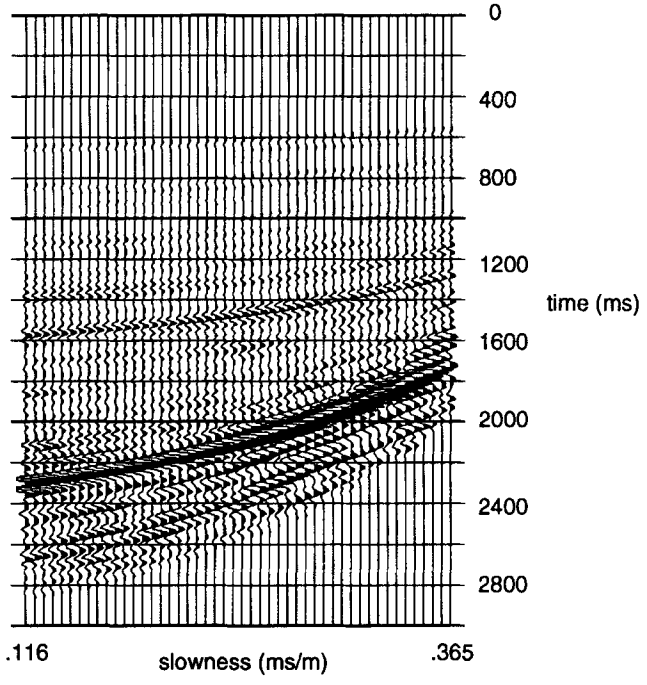

Figure 9. The common-midpoint data gather (CMP6) used for the two field data experiments contrasted in this paper.

Downloaded from https://academic.oup.com/gji/article-abstract/126/2/345/623352 by University of Texas at Dallas - McDermott Library user on 22 November 2019

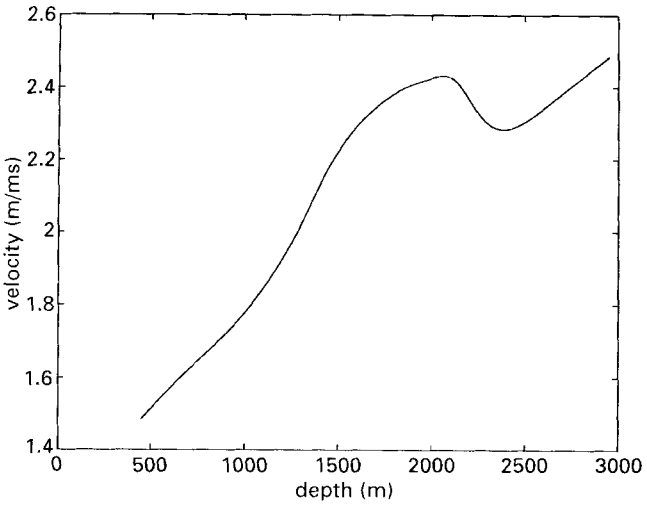

Figure 10. The inversion-estimated P-wave background velocity used for all the inversion experiments in this paper.

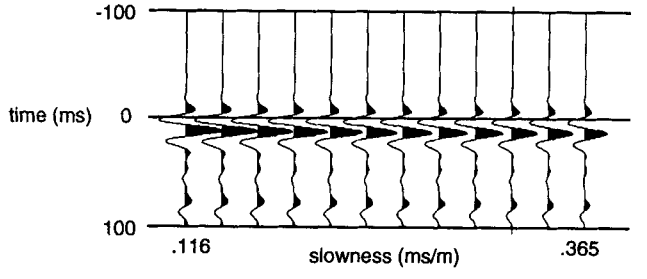

Figure 11. The air-gun model source estimate used in Experiment 1 with every fourth trace shown.

were estimated. In Experiment 2 both the reflectivities and the seismic source were estimated using inversion.

In an earlier subsection, a list of the steps required to approximate the resolution matrix was given. The first level of approximation error is introduced in step 1. In exact arithmetic, n steps of the conjugate gradient algorithm would be sufficient to determine the $x \in \mathscr{R}^n$ which solves $Ax = b$ exactly, since x would be found by searching the space spanned by n independent vectors. Similarly, the eigenvectors would be exactly determined. We actually perform far fewer than n iterations of the conjugate gradient algorithm, however, and so after j steps of the process we search for a solution to $Ax = b$ (and eigenvectors of A) which is optimal in a subspace spanned by only j basis vectors (generally $j \ll n$). A much less significant second level of approximation error occurs in steps 2 and 3 when not all of the approximate eigenvectors are deemed reliable enough to be used. Instead, we select k of the j eigenvectors ($k \leq j$) to form the matrix V (and therefore R_{lanc}). The number (k) of eigenvectors used to form V is determined by specifying an upper bound on acceptable error and applying the following result (Golub & Van Loan 1989, page 479):

Theorem: Suppose that j steps of the Lanczos algorithm have been performed and that $S_j^{\mathrm{T}} T_j S_j = \mathrm{diag}(\theta_1, \ldots, \theta_j)$ is the Schur decomposition of the tridiagonal matrix T_j. If $Y_j = [y_1, \ldots, y_j] = Q_j S_j \in \mathscr{R}^{j \times j}$ then for $i = 1{:}\ j$ we have $\|Ay_i - \theta_i y_i\|_{L^2} = |\beta_j| |s_{ji}|$ where $S_j = (s_{pq})$.

In the theorem, the matrix S contains the eigenvectors of the tridiagonal matrix T which comes from the Lanczos procedure. Q is the Lanczos matrix, Y is the matrix of approximate eigenvectors of the original matrix A, and θ_i are the eigenvalues of T and approximate eigenvalues of A. Finally, β_j is the last off-diagonal entry in the tridiagonal matrix T_j.

To compute our analogue of the SVD resolution matrix R_{lanc} for each experiment, we did not, therefore, rely on all the computed eigenvectors of the normal operator (columns of V). Figs 14 and 21 show the approximate eigenvalues graphed with the associated errors. Figs 15 and 22 show the relative error in the eigenvalues plotted against eigenvalue number (where '1' is always the smallest computed eigenvalue).

For these experiments, the columns of the matrix V are the eigenvectors that correspond most closely to the largest eigenvalues (the best determined in these experiments). However, this correspondence is not exact (see the synthetic experiment in the last section). These eigenvectors were chosen because they meet our (arbitrary) tolerance criterion, namely that the eigenvalues have an approximation error of less than 30 per cent.

Another well-known source of error in the Lanczos process comes from round-off and cancellation. These sources of error cause the Lanczos vectors to lose orthogonality. Error analysis done on this problem has been the motivation behind newer Lanczos procedures which attempt to minimize this loss of orthogonality (Parlett 1980). The approach I have implemented does not take advantage of these newer methods and is, therefore, subject to the problems caused by loss of orthogonality of the Lanczos vectors after a large number of iterations of the conjugate gradient procedure have been performed. I ran numerous experiments, therefore, to try to maximize the accuracy of the eigenvalue approximations while

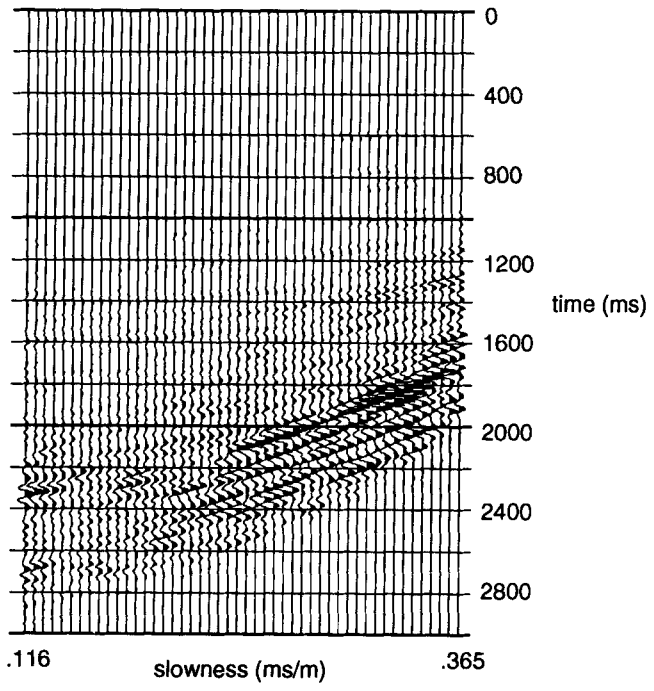

Figure 12. Difference between actual and predicted data which resulted from inverting for the reflectivities with the air-gun source fixed (Experiment 1). The misfit is plotted on the same scale as the actual data.

Downloaded from https://academic.oup.com/gji/article-abstract/126/2/345/623352 by University of Texas at Dallas - McDermott Library user on 22 November 2019

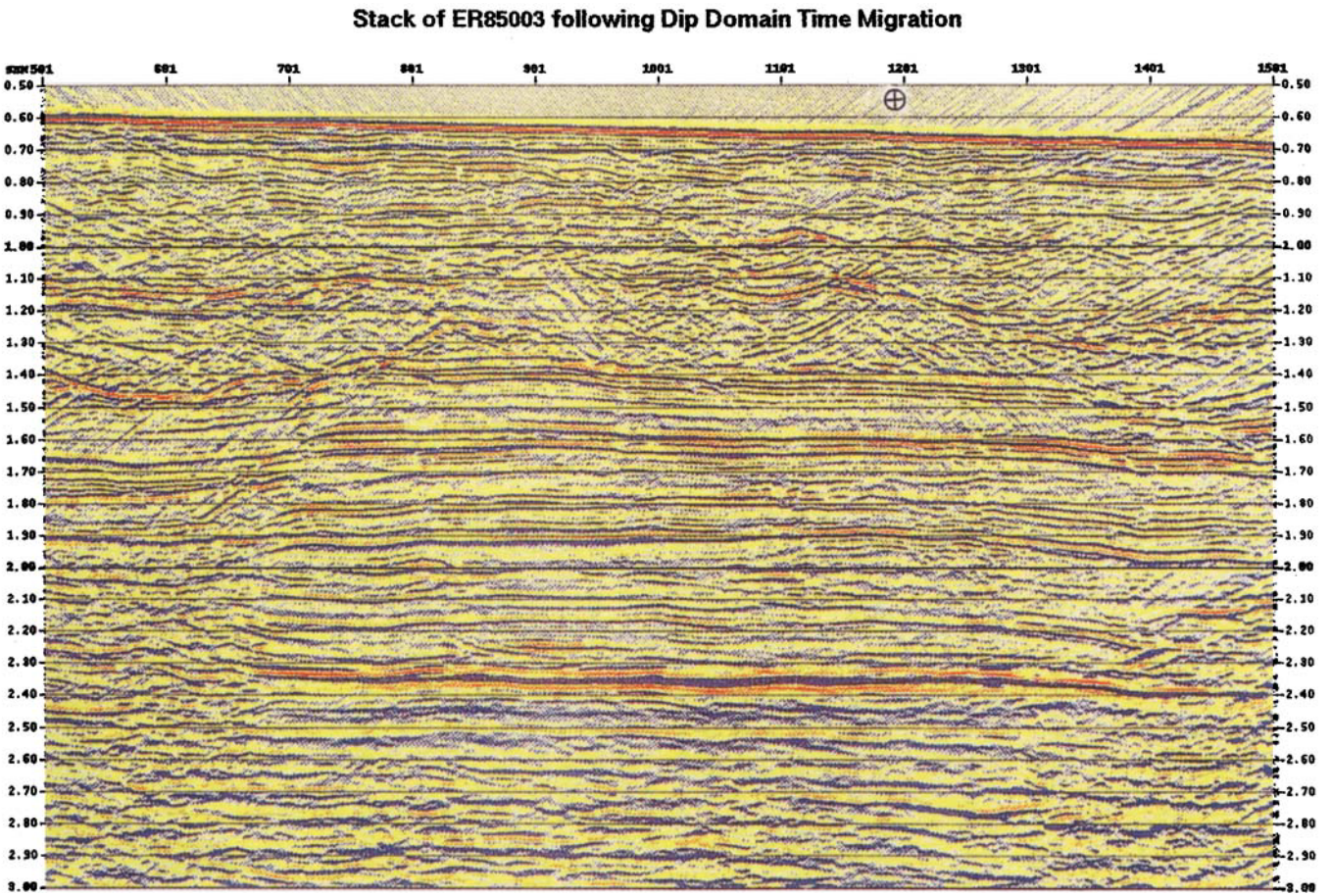

Figure 8. The stacked section of marine data. The diffracted energy originating in the shallow subsurface was suppressed by pre-stack time migration in the offset-midpoint slowness domain. Modelling recreated the data with zero midpoint slowness. The location of the logged well referenced in the text is marked. Note that the bright reflecting horizon, a gas sand at 2.3 s, is embedded in a flat-lying sequence.

Downloaded from https://academic.oup.com/gji/article-abstract/126/2/345/623352 by University of Texas at Dallas - McDermott Library user on 22 November 2019

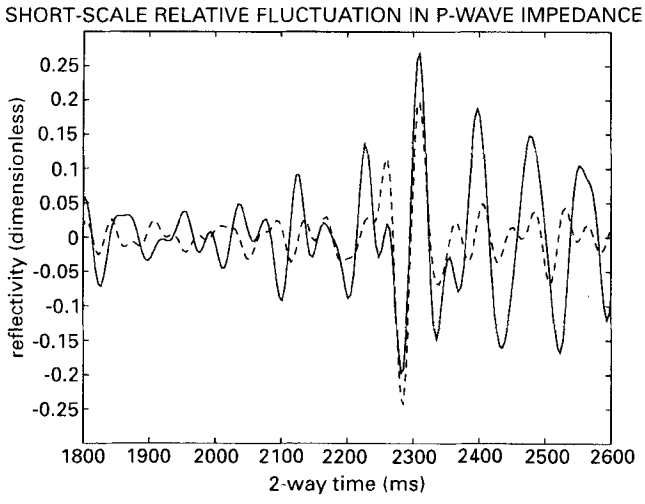

Figure 13. Comparison of the independent well-log measurement of the relative short-scale fluctuation in the *P*-wave impedance with the result of inversion done on CMP6 using the air-gun model source estimate (Experiment 1). The solid line shows the inversion result (scaled). The dashed line shows the detrended well log. Both graphs have been plotted as a function of two-way time and filtered to match the frequency content of the source.

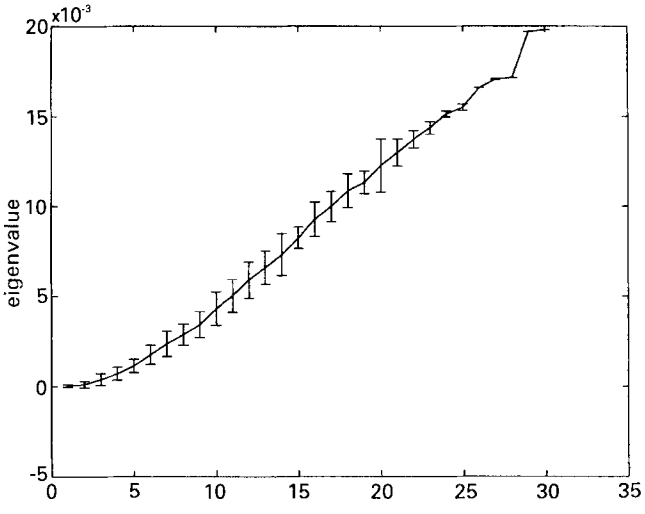

Figure 14. The approximate eigenvalues of the normal operator for the reflectivity inversion with a fixed air-gun model source (Experiment 1). The errors in the eigenvalues are shown as error bars on the graph.

minimizing the loss of orthogonality of the Lanczos vectors. Trial and error indicated that about 30 iterations of the conjugate gradient algorithm are optimal for this particular problem. If we let $Q \in \mathscr{R}^{n \times j}$ be the Lanczos matrix after j iterations of the conjugate gradient algorithm, then ideally we should have that $Q^{\mathrm{T}}Q = I$, with I the $j \times j$ identity matrix. Plots of $Q^{\mathrm{T}}Q$ for the two experiments are shown in Figs 16 and 23.

The model resolution matrix $R \in \mathscr{R}^{n \times n}$ may be quite large, which makes interpreting the data difficult. Various techniques have been devised for selecting interesting columns of R to examine. Three pairs of columns of the resolution matrix corresponding to the *P*-wave impedance reflectivity for the two experiments are shown in Fig. 25. The last pair of graphs

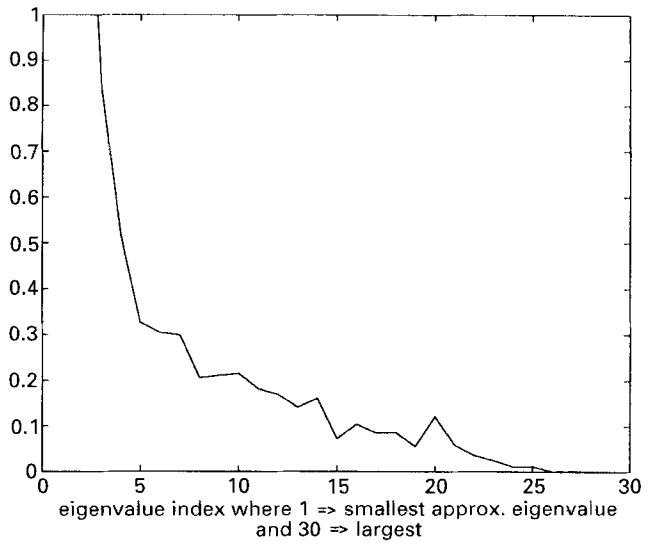

Figure 15. The relative error in the approximate eigenvalues of the normal operator for the reflectivity inversion with a fixed air-gun model source (Experiment 1).

corresponds to an area near the target for each experiment. The top two pairs were chosen at random. If the model parameter is perfectly resolved in the inversion at a certain depth, the plotted column should have a unit spike at the depth corresponding to that column of the resolution matrix and be zero elsewhere. For the three column comparisons shown, the top graph of the pairs (corresponding to the inversion-estimated source) is closer to a spike than the bottom graph (corresponding to the air-gun source). It is clear, however, that none of the columns shown is very close to a unit spike. I am, after all, only estimating 30 eigenvalues out of the total of 626 possible (for the *P*-wave impedance) in these inversion experiments. Wiggins (1972) advocates examining a more intelligent choice of columns of the resolution matrix, namely those with the largest diagonal elements in the resolution matrix (which he calls 'delta vectors'). These vectors should indicate the locations of well-determined components of the model (see also Bube & Langan 1995).

Rather than plotting arbitrary columns of the large resolution matrix, one could calculate a function of the matrix

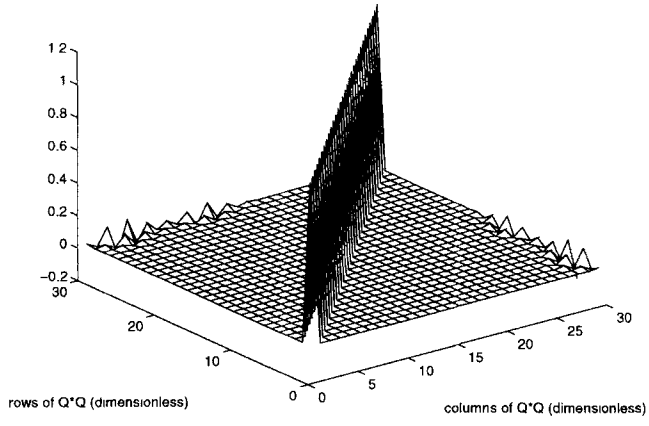

Figure 16. The dot product of the Lanczos matrix with itself for Experiment 1 after 30 iterations of the combined conjugate gradient–Lanczos algorithm.

Downloaded from https://academic.oup.com/gji/article-abstract/126/2/345/623352 by University of Texas at Dallas - McDermott Library user on 22 November 2019

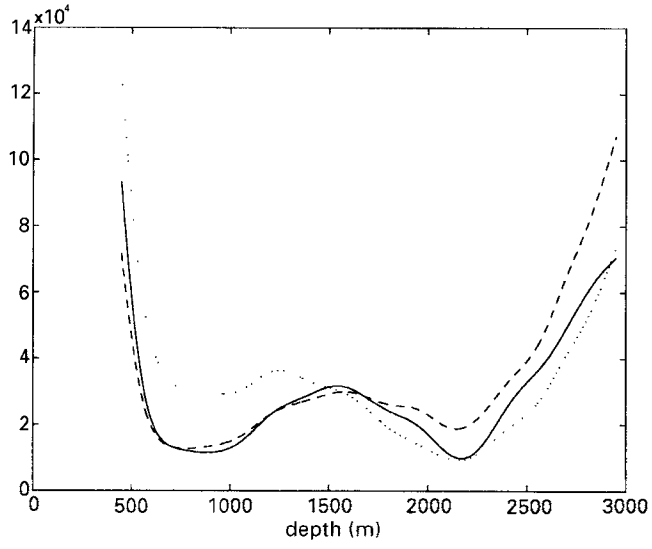

Figure 17. Graph of the spread of the resolution matrix for Experiment 1. The solid line corresponds to the spread for the *P*-wave impedance reflectivity estimate. The dashed line describes the spread for the *S*-wave velocity reflectivity. The dotted line is the spread for the *P*-wave velocity divided by density reflectivity.

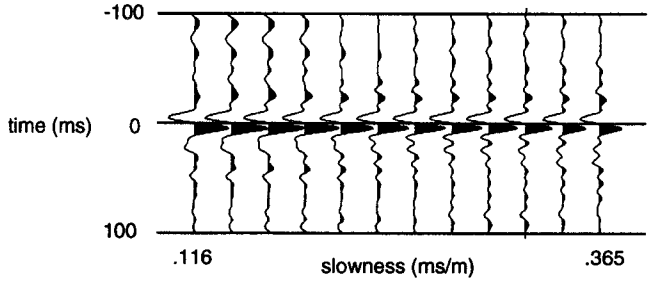

Figure 18. The estimated anisotropic source from the linear source reflectivity inversions, where only every fourth trace is shown for clarity (Experiment 2).

termed the resolution spread (Backus & Gilbert 1968). One example is the spread function:

$$\sum_{i=1}^{n}\sum_{j=1}^{n}(i-j)^2R_{ij}^2.$$

Unfortunately, this function is attempting to convey a large amount of information in a single number, and is easily corrupted by lack of information in some parts of the model domain or by noise in the data. Another choice is the vector-valued spread function:

$$Sp_i=\frac{\sum_{j=1}^{n}(i-j)^2R_{ij}^2}{\sum_{j=1}^{n}R_{ij}^2}. \qquad (11)$$

The idea behind both the scalar and vector-valued spread functions is to weight more heavily the parts of the model estimate derived from averaging the true model over wide intervals. The points on the resolution-spread curve closest to zero are the best-resolved depths.

I show various comparisons of resolution spread. Fig. 17 overplots the resolution spread for each of the three reflectivities

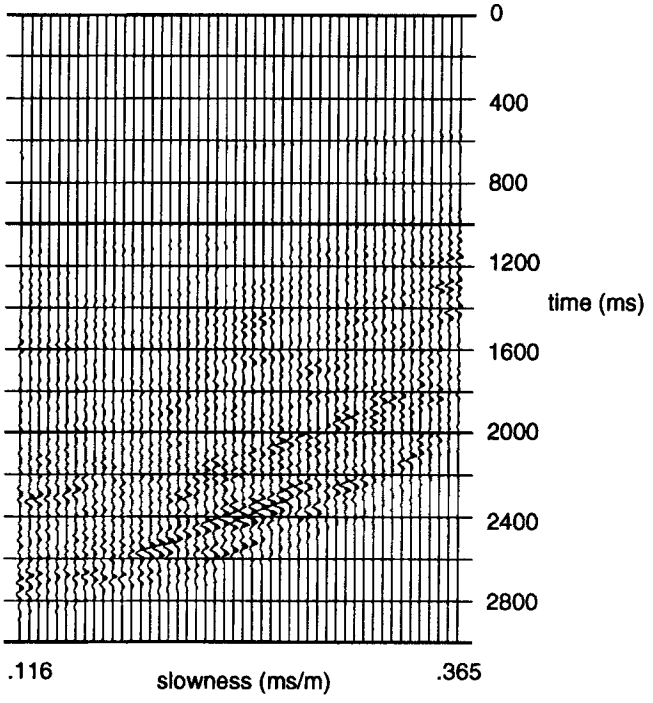

Figure 19. Difference between actual and predicted data obtained from inverting for reflectivities and an anisotropic source (Experiment 2). The misfit is plotted on the same scale as the actual data.

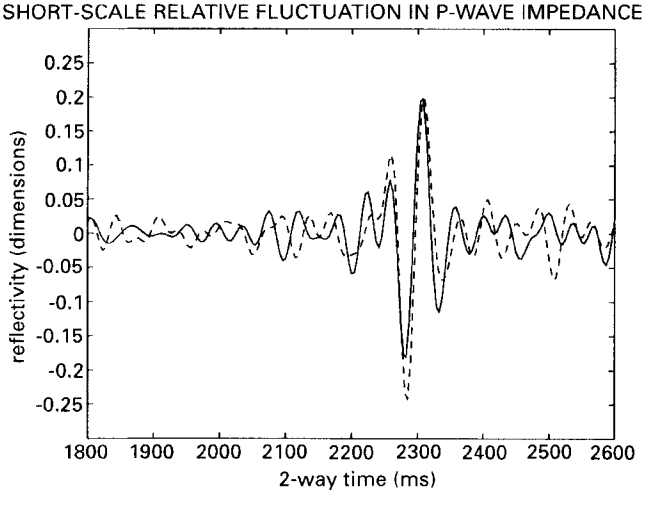

Figure 20. Comparison of the independent well-log measurement of the relative short-scale fluctuation in the *P*-wave impedance with the result of inversion done on CMP6 using an anisotropic source estimate from inversion (Experiment 2). The solid line shows the inversion result (scaled and shifted left 44 ms). The dashed line shows the detrended well log. Both graphs have been plotted as a function of two-way time and filtered to match the frequency content of the source.

estimated in Experiment 1. These curves give little information (increase rapidly) at the water bottom (500 m depth), where there is no information in the data. At depths below the target (2100–2200 m), noise in the data also renders the spread curves unreliable. Nonetheless, one can use this graph to compare the resolution for the three different parameters estimated in one inversion experiment. The normalization of the spread

Downloaded from https://academic.oup.com/gji/article-abstract/126/2/345/623352 by University of Texas at Dallas - McDermott Library user on 22 November 2019

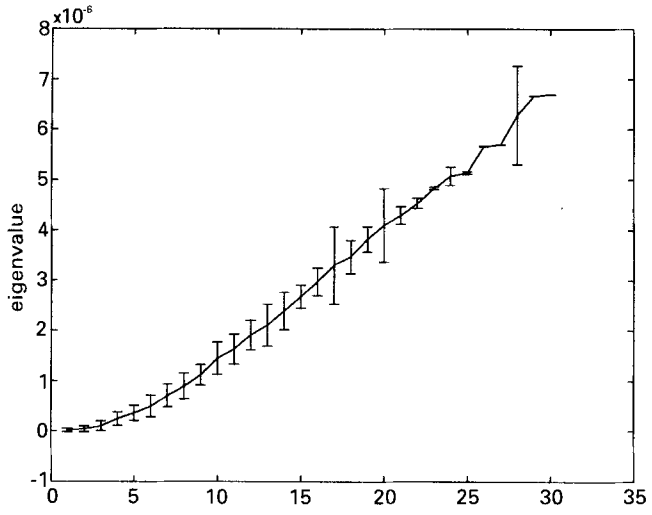

Figure 21. The approximate eigenvalues of the normal operator for the reflectivity inversion with an inversion-estimated anisotropic source (Experiment 2). The errors in the eigenvalues are shown as error bars.

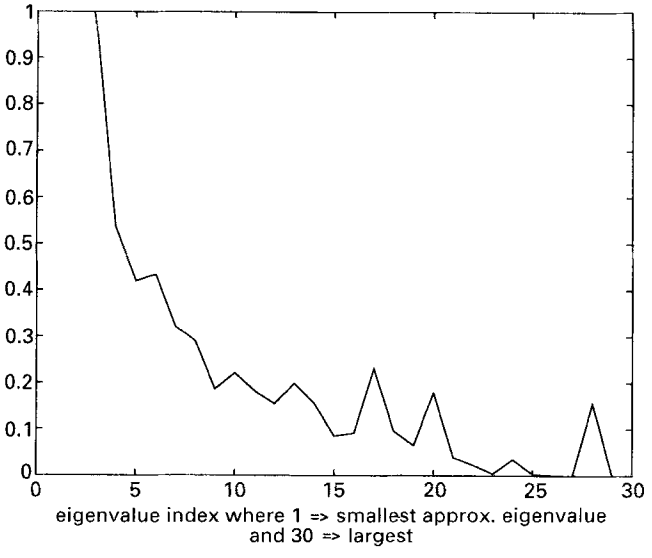

Figure 22. The relative error in the approximate eigenvalues of the normal operator for the reflectivity inversion with an anisotropic inversion-estimated source (Experiment 2).

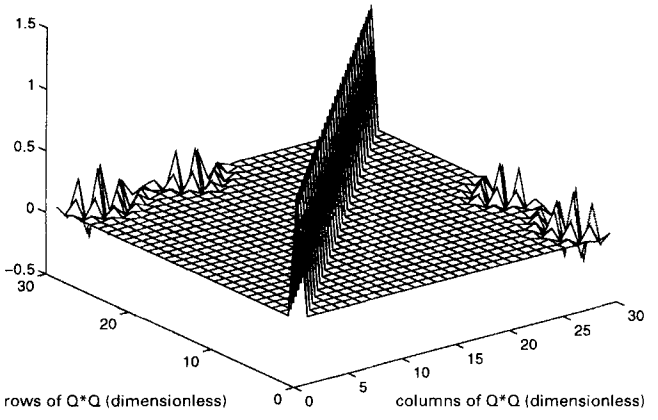

Figure 23. The dot product of the Lanczos matrix with itself for Experiment 2 after 30 iterations of the combined conjugate gradient–Lanczos algorithm.

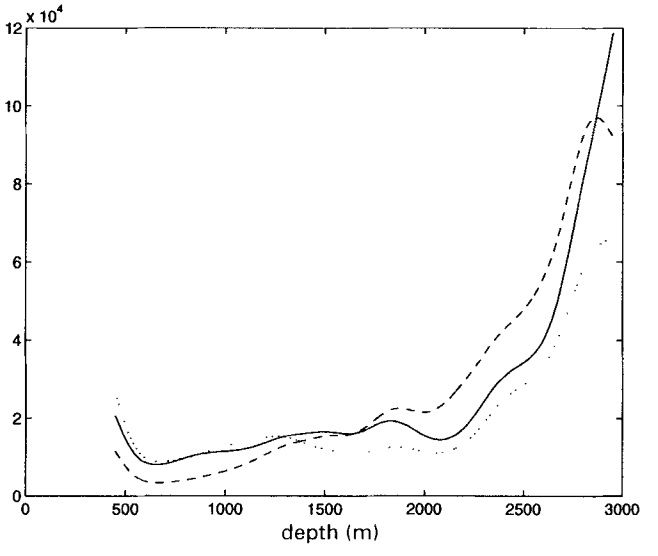

Figure 24. Graph of the spread of the resolution matrix for Experiment 2. The solid line corresponds to the spread for the *P*-wave impedance reflectivity estimate. The dashed line describes the spread for the *S*-wave velocity reflectivity. The dotted line is the spread for the *P*-wave velocity divided by density reflectivity.

function prevents the *P*-wave impedance resolution (the part of the resolution matrix with largest norm) from overpowering the resolution information from the other two parameters. Fig. 24 plots the same spread functions for the three reflectivities estimated in Experiment 2. Fig. 26 compares the resolution spread for the *P*-wave impedance reflectivity for Experiment 1 (dashed line) and Experiment 2 (solid line). Fig. 27 is a graph of the same resolution comparison for the *S*-wave velocity reflectivity. Finally, Fig. 28 compares the resolution spread for the third elastic parameter in the inversion (*P*-wave velocity divided by density reflectivity). One notes that in all three of these graphs, the interval of interest (1000–2300 m depth) is better resolved in Experiment 2 (reflectivity estimation using the inversion-estimated source) than in Experiment 1 (air-gun source experiment). At the target, however, the air-gun source does slightly better than the inversion-estimated source. These resolution-spread pictures agree with the well-log comparisons shown in Figs 13 and 20. The inversion-estimated source tends to do a better job of agreeing with the well logs over the depth domain of interest. However, at the target, the air-gun source does a reasonable job of matching the well-log. (For completeness, I included Figs 12 and 19, which show the misfit between the actual and predicted data for the two real data experiments.)

CONCLUSION

Since measuring tools and computational discretization provide limited data about the subsurface, the inversion process cannot generally return model estimates which are the true parameters that generated the data. The singular value decomposition of the forward (seismogram) model and generalized inverse (in theory) allow us to quantify to what extent the model estimates are localized averages of the true model

Downloaded from https://academic.oup.com/gji/article-abstract/126/2/345/623352 by University of Texas at Dallas - McDermott Library user on 22 November 2019

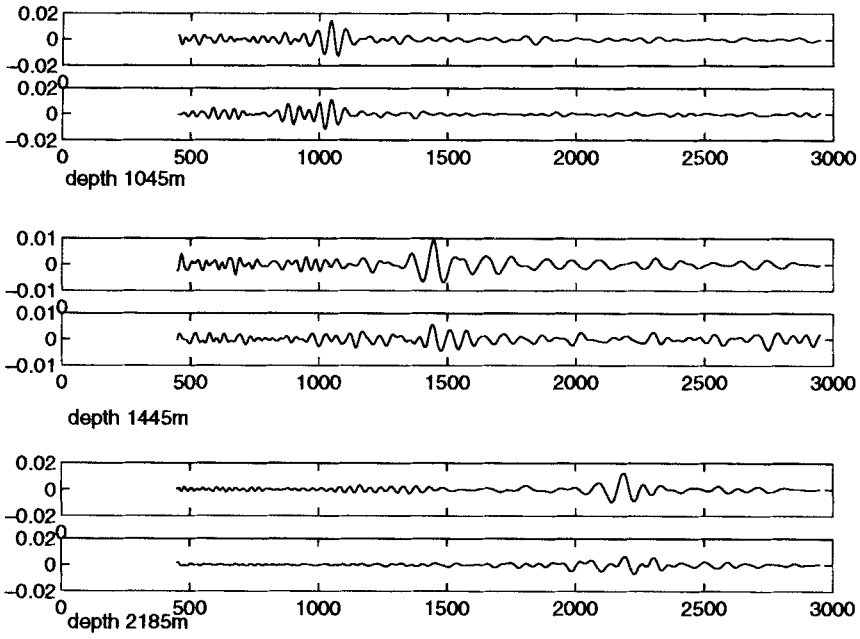

Figure 25. Three selected columns of the resolution matrices for the two experiments. The top two graphs correspond to a depth of 1045 m. The middle two graphs are the column of the resolution matrices corresponding to 1445 m. The bottom two graphs correspond to depth 2185 m (a region between where the two experiments place the targets). In each pair, the top graph corresponds to Experiment 2 (reflectivity estimation with the anisotropic inversion-estimated source). The bottom graph corresponds to Experiment 1 (reflectivity estimation with the air-gun model source).

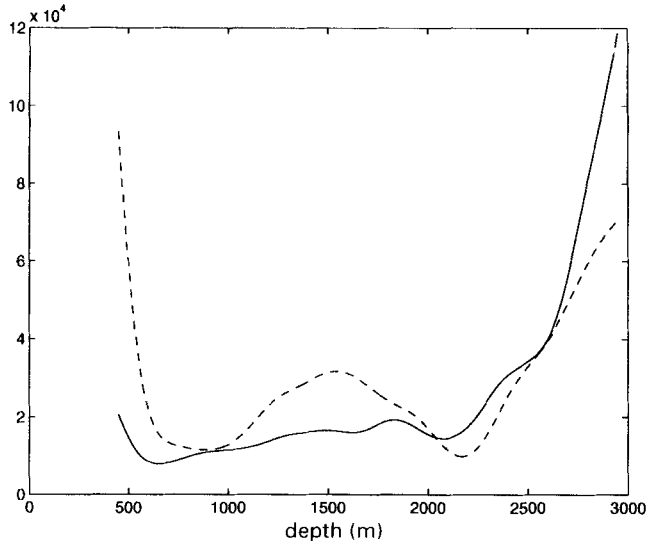

Figure 26. The graph of the resolution spread for two estimates of the *P*-wave impedance reflectivity. The solid line corresponds to Experiment 2; the dashed line, to Experiment 1.

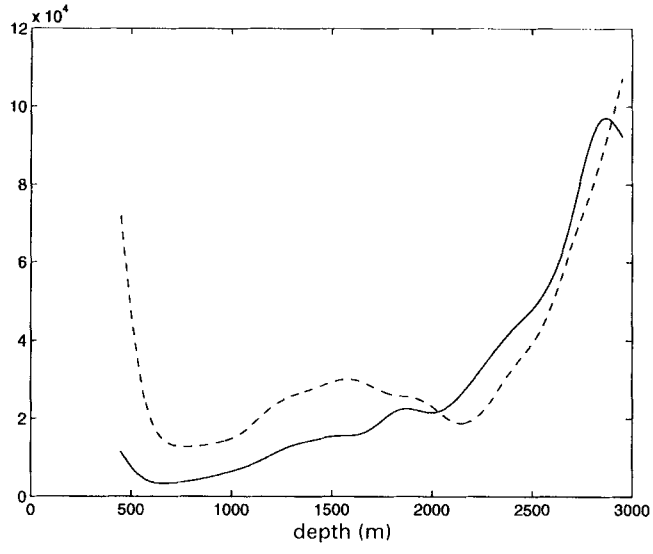

Figure 27. The graph of the resolution spread for two estimates of the *S*-wave velocity reflectivity. The solid line corresponds to Experiment 2; the dashed line, to Experiment 1.

parameters. The smaller the averaging interval (for example in depth) of the model estimates, the better the resolution. A number of papers were written on this subject in the late 1960s and early 1970s, when it seemed a promising tool. More recently, the prohibitive nature of these computations has prompted a number of authors to suggest ways to approximate the model resolution matrix for large (realistic) inverse problems. I present one of these methods in this paper: solve the least-squares inverse problem for the reflectivities using the conjugate gradient algorithm and simultaneously estimate some of the eigenvalues and eigenvectors of the normal matrix by the Lanczos procedure. The estimated eigenvectors are used to form the approximate resolution matrix. I applied this technique to a small synthetic test case which indicated that the approximation is most accurate near large data events. I then applied this process to two field-data reflectivity experiments with different sources. The resulting resolution-spread curves agree with well-log comparisons and imply that resolution defined in this way could be useful for ranking inversion-estimated models when well-log information is not available.

Downloaded from https://academic.oup.com/gji/article-abstract/126/2/345/623352 by University of Texas at Dallas - McDermott Library user on 22 November 2019

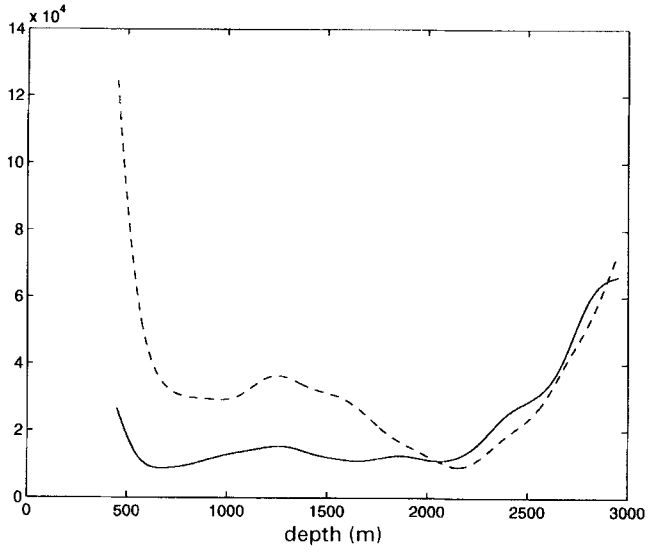

Figure 28. The graph of the resolution spread for two estimates of the *P*-wave velocity divided by density reflectivity. The solid line corresponds to Experiment 2; the dashed line, to Experiment 1.

ACKNOWLEGMENTS

The author would like to thank Bill Symes of Rice University for many encouraging and insightful suggestions. Thanks also go to Exxon Production Research Company and, specifically, to Jim Carazzone for providing the data and viscoelastic simulator used in this study. This work was partially supported by the National Science Foundation, the Office of Naval Research, the Texas Geophysical Parallel Computation Project, the Schlumberger Foundation, IBM, and The Rice Inversion Project. TRIP Sponsors for 1995 were Advance Geophysical, Amerada Hess, Amoco Production Co., Conoco Inc., Cray Research Inc., Discovery Bay, Exxon Production Research Co., Interactive Network Technologies, and Mobil Research and Development Corp.

REFERENCES

Abramowitz, M. & Stegun, I., eds, 1965. *Handbook of Mathematical Functions, with Formulas, Graphs, and Mathematical Tables,* Dover Publ. Inc., New York, NY.

Aki, K. & Richards, P., 1980. *Quantitative Seismology: Theory and Methods,* Freeman, San Francisco, CA.

Anderson, E., Bai, Z., Bischof, C., Demmel, J., Dongarra, J., Croz, J.D., Greenbaum, A., Hammarling, S., McKenney, A., Ostrouchov, S. & Sorensen, D., 1992. *LAPACK Users' Guide,* SIAM, Philadelphia.

Assous, F. & Collino, F., 1990. A numerical method for the exploration of sensitivity: the case of the identification of the 2d stratified elastic medium, *Inverse Problems,* **6,** 487–513.

Backus, G. & Gilbert, F., 1968. The resolving power of gross earth data, *Geophys. J. R. astr. Soc.,* **16,** 169–205.

Berryman, J., 1994a. Resolution of iterative inverses in seismic tomography, in *Proc. Cornelius Lanczos Int. Centenary Conf.* pp. 297-299, eds Brown, J., Chu, M., Ellison, D. & Plemmons, R., SIAM, Philadelphia.

Berryman, J., 1994b. Tomographic resolution without singular value decomposition, SPIE—The International Society for Optical Engineering, Mathematical Methods in Geophysical Imaging, **II,** 2–13.

Beylkin, G., 1985. Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized radon transform, *J. Math. Phys.,* **26,** 99–108.

Bishop, T., Bube, K., Cutler, R., Langan, R., Love, P., Resnick, J., Shuey, R., Spindler, D. & Wyld, H., 1985. Tomographic determination of velocity and depth in laterally varying media, *Geophysics,* **50,** 903–923.

Brysk, H. & McCowan, D., 1986. A slant-stack procedure for point-source data, *Geophysics,* **51,** 1370–1388.

Bube, K. & Langan, R., 1995. Resolution of crosswell tomography with transmission and reflection traveltimes, *65th Ann. Int. Mtg, Soc. Expl. Geophys., Expanded Abstracts,* 77–80.

Franklin, J., 1970. Well-posed stochastic extensions of ill-posed linear problems, *J. Math. Analysis and Applications,* **31,** 682–716.

Golub, G. & O'Leary, D., 1989. Some history of the conjugate gradient and Lanczos algorithms: 1948–1976, *SIAM Review,* **31,** 50–102.

Golub, G. & Van Loan, C., 1989. *Matrix Computations,* The Johns Hopkins University Press, Baltimore, MD.

Hanke, M., 1995. *Conjugate Gradient-type Methods for Ill-posed Problems,* John Wiley & Sons, Inc., New York, NY.

Hestenes, M., 1980. *Conjugate Direction Methods in Optimization,* Springer-Verlag, Berlin.

Hestenes, M. & Stiefel, E., 1952. Methods of conjugate gradients for solving linear systems, *J. Res. Nat. Bur. Stand.,* **49,** 409–436.

Jackson, D., 1979. The use of *a priori* data to resolve non-uniqueness in linear inversion, *Geophys. J. R. astr. Soc.,* **57,** 137–157.

Kennett, B. & Nolet, G., 1978. Resolution analysis for discrete systems, *Geophys. J. R. astr. Soc.,* **53,** 413–425.

Kern, M. & Symes, W., 1994. Inversion of reflection seismograms by differential semblance analysis: algorithm structure and synthetic examples, *Geophys. Prospect.,* **42,** 565–614.

Lanczos, C., 1950. An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Res. Nat. Bur. Stand.,* **45,** 255–282.

Martinez, R. & McMechan, G., 1991. τ–ρ seismic data for viscoelastic media—part 2: linearized inversion, *Geophys. Prospect.,* **39,** 157–181.

Menke, W., 1989. *Geophysical Data Analysis: Discrete Inverse Theory,* Academic Press Inc., San Diego, CA.

Minkoff, S. & Symes, W., 1995. Estimating the energy source and reflectivity by seismic inversion, *Inverse Problems,* **11,** 383–395.

Minkoff, S. & Symes, W., 1996. Full waveform inversion of marine reflection data in the plane-wave domain, *Geophysics,* in press.

Nolet, G. & Snieder, R., 1990. Solving large linear inverse problems by projection, *Geophys. J. Int.,* **103,** 565–568.

Ory, J. & Pratt, R., 1995. Are our parameter estimators biased? The significance of finite-difference regularization operators, *Inverse Problems,* **11,** 397–424.

Paige, C. & Saunders, M., 1982. LSQR: an algorithm for sparse linear equations and sparse least squares, *ACM Trans. Math. Softw.,* **8,** 43–71.

Parlett, B., 1980. *The Symmetric Eigenvalue Problem,* Prentice-Hall, Englewood Cliffs, NJ.

Pulliam, R. & Stark, P., 1994. Confidence regions for mantle heterogeneity, *J. geophys. Res.,* **99,** 6931–6943.

Scales, J., 1989. On the use of conjugate gradient to calculate the eigenvalues and singular values of large, sparse matrices, *Geophys. J.,* **97,** 179–183.

Treitel, S., Gutowski, P. & Wagner, D., 1982. Plane-wave decomposition of seismograms, *Geophysics,* **47,** 1375–1401.

Vasco, D., Pulliam, R. & Johnson, L., 1993. Formal inversion of ISC arrival times for mantle *P*-velocity structure, *Geophys. J. Int.,* **113,** 586–606.

Wiggins, R.A., 1972. The general linear inverse problem: implication of surface waves and free oscillations for earth structure, *Rev. Geophys. Space Phys.,* **10,** 251–285.

Zhang, J. & McMechan, G., 1995. Estimation of resolution and covariance for large matrix inversions, *Geophys. J. Int.,* **121,** 409–426.

Downloaded from https://academic.oup.com/gji/article-abstract/126/2/345/623352 by University of Texas at Dallas - McDermott Library user on 22 November 2019

APPENDIX A: VISCOELASTIC MODELLING

We model the Earth as a layered viscoelastic medium and use a convolutional approximation for pre-critical reflections of viscoelastic plane waves. We do not give the full derivation of the convolutional approximation here. It is closely related to the description of elastic plane waves in Aki & Richards (1980, pp. 153–155). For related experiments see Minkoff & Symes (1996).

The mechanical parameters in the model include the density, ρ, the shear and compressional wave velocities, v_S and v_P, and the shear and compressional quality factors, q_S and q_P. The time-dependent, anisotropic source is assumed to have, approximately, point support.

The parameters in our model vary only with depth, $z \equiv x_3$. Thus we can apply the Radon integral transform (or plane-wave decomposition) to the solution of the viscoelastic wave equation (and to common-midpoint gathers of the data) to reduce the 3-D model to a family of 1-D models (Treitel *et al.* 1982). In effect, we have synthesized incident plane-wave 'shot' records parametrized by slowness, p, and by midpoint. The plane-wave approach is justified by the fact that the reflection angles of interest for inversion were well sampled by the recording arrangement for the frequencies produced by the source. For typical marine acquisition geometry, Brysk & McCowen (1986) explained how the 3-D transform can be applied to common-midpoint gathers.

Neither mode conversion nor multiply reflected energy appears to be important in the data set used for the experiments (probably because of the soft water-bottom materials in this part of the subsurface). Thus we were able to assume a primaries-only, or single-scattering, approximation. The mechanical parameters are separated into the long-wavelength (smooth) background velocities and density, v_P, v_S, and ρ, and short-wavelength relative perturbations of these parameters (reflectivities) $r_P = \delta v_P / v_P$, $r_S = \delta v_S / v_S$, and $r_D = \delta\rho/\rho$. High-frequency asymptotics leads to the convolutional model prediction of the P-wave seismogram (Beylkin 1985):

$$d^{\text{pred}}(t, p) = f(t, p) * \tilde{r}(t, p). \tag{A1}$$

In the above expression, d^{pred} is the seismic data predicted by our model, f the source wavelet; p denotes slowness, and t time. The '$*$' symbol is convolution in time.

The source wavelet is anisotropic, i.e. slowness-dependent. A convenient representation is a Legendre series in p (which also regulates the extent of its anisotropy):

$$f(t, p) = \sum_{i=0}^{N} f_i(t) L_i(p), \tag{A2}$$

where L_i is the ith Legendre polynomial (Abramowitz & Stegun 1965).

The time-slowness-domain reflectivity $\tilde{r}$ is a linear combination of the elastic reflectivities r_P, r_S and r_D:

$$\tilde{r}(t, p) \approx \int dz [A_P(z, t, p) r_P(z) + A_S(z, t, p) r_S(z) + A_D(z, t, p) r_D(z)]. \tag{A3}$$

The geometric optics P–P reflectivity kernels A_P, A_S, A_D are (for $j = P$, S, or D)

$$A_j(z, t, p) = F_j(z, p) \int_{-\infty}^{\infty} d\omega \times \exp\left[i\omega(t - 2\tau(z, p)) - |\omega| \left(1 - \frac{2i}{\pi} \ln\left|\frac{\omega}{\omega_0}\right|\right) \alpha(z, p) \right], \tag{A4}$$

where

$$F_D(z, p) = K_P^2(z, p)[1 - 4p^2 v_S^2(z)], \tag{A5}$$

$$F_P(z, p) = \frac{K_P^2(z, p)}{[1 - p^2 v_P^2(z)]}, \tag{A6}$$

$$F_S(z, p) = -8K_P^2(z, p) p^2 v_S^2(z), \tag{A7}$$

and

$$K_P(z, p) = \frac{(1 - 2v_S^2(z)p^2)\sqrt{[\rho(z)/v_P(z)]\sqrt{1 - v_P^2(z)p^2}}}{\rho(z)\{[1 - 2v_S^2(z)p^2] + 4v_S^4(z)p^2\{p^2 + [1/(v_P(z)v_S(z))] \times \sqrt{1 - v_P^2(z)p^2}\sqrt{1 - v_S^2(z)p^2}\}\}}. \tag{A8}$$

The vertical traveltime of plane waves with slowness p is

$$\tau(z, p) = \int_{z_0}^{z} dy \frac{\sqrt{1 - [pv_P(y)]^2}}{v_P(y)}. \tag{A9}$$

The P-wave attenuation factor at slowness p is

$$\alpha(z, p) = \int_{z_0}^{z} dy \frac{1}{Q_P(y) v_P(y) \{1 - [pv_P(y)]^2\}^{1/2}}. \tag{A10}$$

The temporal frequency variable is denoted ω; z_0 is the source depth, and Q_P is the P-wave quality factor. The reference frequency ω_0 calibrates the (frequency-dependent) velocity. Waves at frequency ω_0 move with the P-wave velocity v_P. The form of the exponent in the integral expression (A4) follows Aki & Richards (1980, p. 182, display (5.88)).

The integrals described above are approximated by the trapezoidal rule in the modelling code. For use in the optimization, both the linearizations of the above expressions and their adjoints are required. These are computed by applying first-order perturbation theory to the discretized integral transforms.

APPENDIX B: INVERSION

The predicted seismic data d^{pred} is linear in each of the parameters f (the seismic source) and r_P, r_S, r_D (the elastic reflectivities). It is very non-linear in the P-wave velocity v_p. Although we did perform an initial inversion for the long-wavelength portion of the P-wave velocity (Fig. 10), we will not discuss that experiment in this paper (see instead Minkoff & Symes 1996). It is fixed in both the experiments described in this paper. The predicted seismogram also depends on the (background) S-wave velocity v_S and density ρ, and on the quality factors Q_P and Q_S. We have assumed here that v_S and ρ are known with sufficient accuracy from logs and regional relationships which hold on the average over long scales. The quality factors were estimated by roughly matching the rate of energy decay in the data with predictions from log-derived synthetics. Instead we focus on inverting for the source (Experiment 2) and three elastic parameter reflectivities (Experiments 1 and 2).

The basic inversion principle is Output Least Squares (or OLS). This method requires that we adjust the inversion parameters r_P, r_S, r_D and f to minimize the mean-squared

Downloaded from https://academic.oup.com/gji/article-abstract/126/2/345/623352 by University of Texas at Dallas - McDermott Library user on 22 November 2019

error

$$J_{\mathrm{OLS}} = \int_{t_{\min}}^{t_{\max}} dt \int_{p_{\min}}^{p_{\max}} dp \gamma(t,p) |d^{\mathrm{pred}}(t,p) - d^{\mathrm{obs}}(t,p)|^2 ,$$

where $d^{\mathrm{obs}}(t,p)$ is the 'observed' p–τ data, and $\gamma(t,p)$ is a *conditioning weight* factor. The conditioning factor $\gamma(t,p)$ enhances the resolution of deeper events. The desired outcome of conditioning permits considerable freedom in the design of the weight. We have used

$$\gamma(t,p) = \exp[|\omega_0|\alpha(z,p)] ,$$

where $t = \tau(z,p)$ in this formula. In the elastic limit where $Q_P \to \infty$, $\gamma = 1$. Thus γ is a model-based gain.

The elastic reflectivities and source parameters together have a non-linear (bilinear) influence, so quadratic minimization algorithms cannot be used directly in Experiment 2. Instead, we used a method known as coordinate search or alternation. This method of simultaneous inversion for source and reflectivities was introduced in the paper by Minkoff & Symes (1995). We will hereafter refer to a 'round of alternation' as one pass through the four steps of the following algorithm.

Repeat until convergence:

(1) Given the current source, f_c, and current reflectivity, r_c, invert for a new estimate of the reflectivity r_+.

(2) Replace r_c by r_+.

(3) Given the current source and reflectivity guesses, f_c, r_+, invert for a new estimate of the source f_+.

(4) Replace f_c by f_+.

Alternation, although notoriously inefficient, is attractive for initial experiments because it requires only successive solutions of simple linear least-squares problems. Obviously, quasi-Newton methods could be applied to the problem and would probably reduce the number of iterations dramatically. While further algorithmic development for source-reflectivity inversion is definitely needed, we were able to obtain a reasonable result in Experiment 2 with only two rounds of alternation.

APPENDIX C: SEISMIC FIELD DATA—GEOMETRY OF THE GEOPHYSICAL EXPERIMENT AND PREPROCESSING

The data used in this work were derived from a marine survey in the Gulf of Mexico. The survey line consisted of 511 shots recorded with 301 hydrophone groups. The group interval was 15 m with a minimum source–receiver separation of 148 m. The shot interval was 22.5 m. Each group contained 17 equally spaced and equally weighted hydrophones. The data were recorded without a low-cut filter. A 110 Hz high-cut filter was applied. The sampling rate was 2 ms, and the total record length was 5 s.

This area of the Gulf contains a strong gas-sand-related direct hydrocarbon indicator at about 2.3 s. The stack shows this target horizon to be embedded in a sequence of nearly horizontal strata, beginning at roughly 1.5 s (see Fig. 8). Therefore, layered medium modelling appeared to be a plausible tool for target-oriented inversion. Very little evidence of multiply-reflected energy appears above or near the target event, suggesting the viability of a primaries-only approach. Similarly, the apparent absence of mode-converted events justifies restriction of the propagation model to P waves.

The data were Radon-transformed, respecting 3-D cylindrical symmetry, to yield 48 plane-wave traces per midpoint gather. Slowness values range from $p_{\min} = 0.116$ ms m^{-1} to $p_{\max} = 0.365$ ms m^{-1}. To remove diffraction artefacts originating in the shallow subsurface, the plane-wave data were time-migrated in the midpoint dip domain, and then modelled to $p_{\mathrm{mid}} = 0$. This last step collapses diffractions while still preserving layered reflection amplitudes. While the diffracting structures are still visible in the upper 1.5 s, the diffraction tails are largely removed and no longer interfere with the lower, layered structure in the stack. The same is largely true of the pre-stack p–τ data (see Fig. 9).

In order to lighten the computational burden of inversion, we performed our calculations on a low-pass filtered version of the p–τ data, which resulted from convolving all the traces with a 15 Hz zero-phase Ricker wavelet. The numerical experiments described in this paper were performed on the single midpoint gather of filtered data (see Fig. 9) located near a logged well. We also filtered the corresponding air-gun model source.

Downloaded from https://academic.oup.com/gji/article-abstract/126/2/345/623352 by University of Texas at Dallas - McDermott Library user on 22 November 2019