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, P.O. Box 1892, Houston TX 77251-1892

Summary

Seismic inversion produces model estimates which are at most unique in an *average* sense. The model resolution matrix quantifies how close an estimate is locally to the true model which generated the data. The resolution matrix is traditionally defined from the singular value decomposition of the discretized forward problem. Unfortunately, computing the singular value decomposition of the forward matrix is computationally prohibitive for inverse problems of realistic size. Inversion requires the solution of large normal matrix systems which are best tackled by iterative techniques 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. As an example of this method, I performed two linear viscoelastic inversion experiments on $p - \tau$ marine data from the Gulf of Mexico. The two sets of elastic parameter reflectivities were estimated from the same data using two different seismic sources. The two reflectivity/source models may be judged for accuracy based on how the estimates compare to measured well logs. The approximate model resolution matrix agrees with this ranking of the two models for this example and provides us with a way to compare different model estimates when, for example, such well-log measurements are not available.

Short Title: Approximating Resolution for Large Inverse Problems
Keywords: Seismic inverse problem, model resolution

1 Introduction

It is well known that seismic inverse problems are ill-posed. They generally do not have unique solutions. When we solve a linear inverse problem, we cannot rely on the model estimates to be accurate at each point. How accurate the model estimates are depends on how well the model reflects the subsurface phenomena and how the inversion procedure estimates the model from the data. The goal of resolution is to quantify which parts of the model parameters are well determined insofar as one can regard that part of the model estimate as being close to the true model.

An explanation of the general linear inverse problem and model resolution may be found in the book by Menke [MENKE, 1989] and in the now classic paper by Backus and Gilbert [BACKUS and GILBERT, 1968]. Backus and Gilbert describe how solutions of seismic inverse problems are most often not pointwise unique but may still be viewed as having unique average behavior. They describe how to determine the shortest length scale which the given data can resolve at a particular depth. (Related references include papers by Franklin [FRANKLIN, 1970], Jackson [JACKSON, 1979], and Kennett and Nolet [KENNETT and NOLET, 1978].) The model resolution matrix provides one way to quantify how close the model estimates are to the true model as a function of the independent variable (such as depth). Wiggins [WIGGINS, 1972] defines this matrix in terms of the singular value decomposition (SVD) and analyzes resolution for a very small problem (less than ten unknowns) using observed data from surface waves and free oscillations. He comments in the abstract to the paper 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." Forming the resolution matrix via the SVD has not become common practice, however, because it is computationally prohibitive to implement the idea directly on large prob-

lems. Papers by Martinez and McMechan [MARTINEZ and MCMECHAN, 1991], Assous and Collino [ASSOUS and COLLINO, 1990], Bishop et al. [BISHOP et al., 1985], and Ory and Pratt [ORY and PRATT, 1995] are attempts to use the SVD to compute the well and poorly-resolved parameters. All of these attempts are restricted by the cost of computing the SVD. Martinez and McMechan use a viscoelastic simulator to invert for velocities, density, and quality factors. They 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. estimate reflector depth and seismic velocities from tomography data. They analyze the determination of components of a simplified model analytically and substantiate their conclusions with a small numerical experiment performed on real data. Assous and Collino determine resolution for a more complicated problem by storing the full normal matrix and calling an IMSL generalized eigenvalue problem solver. Finally, Ory and Pratt examine the effect of different regularizing operators on resolution. They consider a synthetic experiment of determining three one-dimensional anisotropic velocity parameters from traveltime data. At the end of the paper (p. 420) they state:

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

Rather than computing the SVD of the forward operator, we approximate the eigenvalues and eigenvectors of the normal matrix via the Lanczos algorithm. Having solved the inverse problem via the closely-related conjugate gradient technique, we obtain this *approximate resolution matrix* at very little additional cost over solving the problem. The size of the problem does not restrict

our ability to compute the resolution. The technique may be applied without modification to any type of inverse problem where resolution should be analyzed.

In Section 2 of this paper, we describe model resolution for seismic inverse problems. Section 3 provides a discussion of some of the basic ideas behind the well-known conjugate gradient and Lanczos algorithms and lists the steps we take to determine our approximation to the model resolution matrix. A pseudocode version of the actual algorithm is also given in this section. In Section 4, we apply this approximate model resolution idea to a real marine dataset from the Gulf of Mexico. Using a viscoelastic model for wave propagation in the earth, we invert for the three elastic parameter reflectivities which linearly influence the data. Two inversion experiments are performed which differ only in the seismic sources used. Resolution diagrams are provided 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.

2 The Model Resolution Idea

Most geophysical inverse problems are of "mixed determined" type; i.e., these problems are ones in which some model parameters are well determined and others are not. When we perform an inversion, we should not think of the resulting model parameters as being accurate at each discretized point. It is more reasonable to think of the solution vector m as an *average* of the true model parameters $\langle m \rangle = a^t m$. The easiest such models to interpret are ones which have an averaging vector a which is only nonzero over a small interval in the domain of the model

parameters. Then, if the model parameters are discretely parametrized in a physically meaningful way, one can interpret this average. In the experiments I will describe, the model parameters depend on one spatial coordinate only, depth z, and the averaging vectors correspond to averages over depths.

We wish to quantify our ability to resolve the individual model parameters from the forward model for the seismogram G. We start by devising a generalized inverse or *estimator* G^{-g} which will act on the data and return an estimate of the model parameters. Assuming that there is a set of model parameters $m^{true} \in \Re^n$ which satisfies the equation $Gm^{true} = d^{obs}$, for the forward modeler $G \in \Re^{m \times n}$ and observed data $d^{obs} \in \Re^m$, we would like to be able to measure how close our estimated model m^{est} is to the (unknown) true model m^{true}.

$$m^{est} = G^{-g} d^{obs} = G^{-g} G m^{true} \equiv R m^{true} \tag{2.1}$$

where $R \in \Re^{n \times n}$ is the *model resolution matrix.* If $R = I$ then $m^{est} = m^{true}$ (or any other model which fits the data as well as m^{true} according to our forward modeler), 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 $m_i^{est} = a^t m^{true}$. Note that this averaging vector a corresponds to a column of the resolution matrix R.

As an example, assume G has full column rank (i.e., $\text{rank}(G) = n$). Then the least squares estimator G^{-g} comes from solving the normal equations: $G^{-g} = (G^t G)^{-1} G^t$. The resolution matrix corresponding to this estimator is defined to be

$$R = G^{-g} G = (G^t G)^{-1} G^t G = I.$$

The model parameters computed using exact least squares inversion are perfectly resolved.

The singular value decomposition is the computational tool one would like to apply to understand how close the estimated model is to the true model which generated the data. The singular value decomposition is defined in [GOLUB and LOAN, 1989], page 71, by the following theorem:

Theorem 2.1 *If $A \in \Re^{m\times n}$ matrix, then there exist orthogonal matrices $U = [u_1, \ldots, u_m] \in \Re^{m\times m}$ and $V = [v_1, \ldots, v_n] \in \Re^{n\times n}$ such that $U^t AV = diag(\sigma_1, \ldots, \sigma_p) \in \Re^{m\times n}$, $p = \min\{m, n\}$, where $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p \geq 0$.*

The columns of U are eigenvectors of AA^t (the left singular vectors of A), and the columns of V are the eigenvectors of A^tA (the right singular vectors of A). Finally, the p singular values on the diagonal of $\Sigma = diag(\sigma_1, \ldots, \sigma_p)$ are the square roots of the nonzero eigenvalues of AA^t and A^tA (the singular values).

The forward operator G, can be defined in terms of the singular value decomposition as $G = U\Sigma V^t$. In other words, the matrix U has columns which are the eigenvectors spanning the data space D, and V's columns are the eigenvectors spanning the model space M. The singular value matrix Σ can be partitioned into a square submatrix Σ_p of nonzero singular values and three zero matrices:

$$\begin{pmatrix} \Sigma_p & 0 \\ 0 & 0 \end{pmatrix}$$

We write $G = U\Sigma V^t = U_p\Sigma_p V_p^t$ (the truncated SVD) where U_p and V_p^t consist of the first p columns of U and V respectively. We note that although $V_p^t V_p = I$, in general since the p right singular vectors do not span the whole space, $V_p V_p^t \neq I$.

A version of the estimator G^{-g} may now be expressed using the truncated singular value

decomposition as

$$G^{-g} = V_p \Sigma_p^{-1} U_p^t$$

and the model resolution matrix (from expression 2.1) is simply

$$R = G^{-g} G = (V_p \Sigma_p^{-1} U_p^t)(U_p \Sigma_p V_p^t) = V_p V_p^t \tag{2.2}$$

In our implementation, we do not explicitly compute the SVD. Instead we 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.

3 Discussion of the Conjugate Gradient and Lanczos Algorithms

Due to the large size of the normal matrices which are typical for seismic inverse problems, we prefer to use iterative methods to approximately solve these linear systems. Since the normal operator is symmetric and positive definite, an obvious technique to use to solve this system is the conjugate gradient algorithm. Moreover, the conjugate gradient algorithm automatically generates the parameters needed to construct the tridiagonal matrix which arises in the Lanczos iteration. 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 also 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 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 [HESTENES, 1980], Hestenes and Stiefel [HESTENES and STIEFEL, 1952], and Lanczos [LANCZOS, 1950]. Golub and O'Leary give a bibliographical history of the two methods in [GOLUB and O'LEARY, 1989]. The following general description of the two algorithms may be found in expanded form in the book by Golub and Van Loan [GOLUB and LOAN, 1989], pp. 476–480 and 516–520.

3.1 The Conjugate Gradient Idea

The Hestenes-Stiefel conjugate gradient algorithm may be understood in the context of minimizing the function $\phi(x)$ defined by

$$\phi(x) = \frac{1}{2}x^t Ax - x^t b \tag{3.1}$$

where $b \in \Re^n$, and the matrix $A \in \Re^{n\times n}$ 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 obvious 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 largest to smallest eigenvalues) is large. The conjugate gradient algorithm, therefore, chooses to minimize ϕ in a set of directions $\{p_1, p_2, ...\}$ which do not necessarily correspond to the residual directions. One approach with obvious benefits is to choose linearly independent directions p_i so that each x_j solves

$$\min_{x \in \text{span}\{p_1, \ldots, p_j\}} \phi(x) \tag{3.2}$$

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 one-dimensional minimization problem

$$\min_{\alpha} \phi(x_{j-1} + \alpha p_j) \tag{3.3}$$

we also solve the j-dimensional problem 3.2. 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 $P_{j-1}^t A p_j = 0$. These requirements can be satisfied and an algorithmic implementation is described in Subsection 3.4 below.

3.2 The Lanczos Idea and Connection to the Conjugate Gradient Algorithm

Estimates of the eigenvalues and eigenvectors of the normal operator are very useful for analyzing inversion results. The Lanczos algorithm, when applied to a symmetric matrix $A \in \Re^{n \times n}$, generates a sequence of tridiagonal matrices $T_j \in \Re^{j \times j}$ 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 $\Re^{n \times j}$ with orthonormal columns, we define the scalars M_j and m_j by

$$M_j = \max_{y \neq 0} \frac{y^t (Q_j^t A Q_j) y}{y^t y} \leq \lambda_1(A) \tag{3.4}$$

$$m_j = \min_{y \neq 0} \frac{y^t (Q_j^t A Q_j) y}{y^t y} \geq \lambda_n(A) \tag{3.5}$$

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 $span\{x, Ax\}$. For this reason, the Lanczos vectors $\{q_i\}_1^j$ are chosen to be an orthonormal basis for the Krylov subspace

$$\kappa(A, q_1, j) \equiv \text{span}\{q_1, Aq_1, ..., A^{j-1}q_1\} = \text{span}\{q_1, ..., q_j\} \tag{3.6}$$

At the jth iteration of the Lanczos algorithm we have a matrix $Q_j \in \Re^{n \times j}$ (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 $T \in \Re^{j \times j}$. In fact, the Lanczos matrix "tridiagonalizes" the matrix A up to an error matrix.

$$AQ_j = Q_j T_j + r_j e_j^t. \tag{3.7}$$

The entries in T are combinations of the parameters generated in the conjugate gradient iteration (for details see the algorithm in Subsection 3.4).

3.3 Definition of the Approximate Model Resolution Matrix

As stated in Section 2, we are solving the problem $Gm = d$ where the forward operator $G \in \Re^{m \times n}$. Assuming that $rank(G) = n$, we may instead solve the normal system

$$G^t G m = G^t d. \tag{3.8}$$

Let $A \equiv G^t G$ and $b \equiv G^t d$. Calling the model solution x (rather than m), equation 3.8 may be written in more standard notation as $Ax = b$. Rather than calculating the full model resolution matrix defined in Section 2 by the SVD (computationally prohibitive), we approximate this matrix via the following steps:

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 $\{q_1, Aq_1, ..., A^{j-1}q_1\}$. Similarly, the j eigenvalue/vector estimates which come out of the Lanczos process are optimal for this subspace.

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

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 $V \in \Re^{n \times k}$.

4. Form the approximate resolution matrix $\tilde{R} \equiv VV^t$. $\tilde{R} \in \Re^{n \times n}$.

It must be stressed that this resolution matrix is only an approximation and is not the matrix one would get from computing $R = G^{-g}G$ as described in Section 2. In the marine data experiments described in Section 4, we 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 us with a good estimate of the solution without requiring one to perform nearly this many iterations, however. In fact, we perform only 30 iterations of the combined conjugate gradient/Lanczos algorithm. Further, we dropped the 6 or 7 smallest eigenvectors from the matrix V because of approximation error.

The conjugate gradient algorithm attempts to solve the system $Ax = b$ by finding the best solution x_j where $x_j \in \kappa(A, q_1, j) \equiv \text{span}\{q_1, Aq_1, ..., A^{j-1}q_1\}$. This *Lanczos basis* $\{q_1, Aq_1, ..., A^{j-1}q_1\}$ is necessarily data dependent. The conjugate gradient algorithm chooses as its initial starting direction the steepest descent or residual direction, which is related to the data (or is the data if the initial guess for the solution x_0 is zero). The version of the Lanczos algorithm we use 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.

3.4 Algorithm

We present here a pseudocode version of the two algorithms (conjugate gradient and Lanczos) which we have implemented in our inversion code.

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 $A \in \Re^{n \times n}$ is symmetric and positive definite and $b \in \Re^n$ then the following algorithm computes $x \in \Re^n$ 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 $j = 1$

$\beta = 0$

$p = r$

$rtr_c = \langle r, r \rangle$

else

$\beta = rtr_c/rtr_p$

$p = r + \beta p$

end if

$ap = Ap$

$ptap = \langle p, ap \rangle$

if $ptap < tol$

break

end if

$\alpha_c = rtr_c/ptap$

$x = x + \alpha_c p$

$r = r - \alpha_c ap$

if eigenvalue flag = true

if $j = 1$

$T(j,j) = 1/\alpha_c$

$T(j,j-1) = 0$

$T(j-1,j) = T(j,j-1)$

else

$T(j,j) = rtr_c/(rtr_p\alpha_p) + 1/\alpha_c$

$T(j,j-1) = -\sqrt{rtr_c/rtr_p}/\alpha_p$

$T(j-1,j) = T(j,j-1)$

end if

Call LAPACK routine SSTEQR to get eigenvalues/vectors (Z) of T.

Compute error in approximate eigenvalue for normal operator.

end if

$rtr_p = rtr_c$

$rtr_c = \langle r, r \rangle$

$\alpha_p = \alpha_c$

if $\sqrt{rtr_c} < tol$

break

end if

if eigenvector flag = true

$QZ = X$

end if

end

4 Application of Resolution to $p - \tau$ Marine Data Experiments

4.1 Modeling, Inversion, and the Data

In this section we apply the conjugate gradient and Lanczos algorithms just described to the linear inverse problem of estimating three elastic parameter reflectivities in the viscoelastic model for wave propagation in the earth. We 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 both 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 modeling and inversion procedures are described in Appendix A. The geometry of the physical marine experiment as well as data preparation are given in Appendix B. (For a reference to a related inversion study of the same marine data, see [MINKOFF and SYMES, 1994].) 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 and Richards [AKI and 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*:

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

In the above expression, d^{pred} is the predicted seismic data, f the 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 relativized 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 *O*utput *L*east *S*quares (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

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

where $d^{obs}(t,p)$ is the "observed" $p-\tau$ data and $\|\cdot\|$ is the L^2 norm.

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 method. While the choice of source wavelet influenced the rate of convergence, 30 iterations of the conjugate gradient algorithm was generally sufficient to reduce the normal residual to close to 1% of its starting value. In both experiments, the initial estimates of the reflectivities were zero. The background P-wave velocity was estimated via nonlinear Differential Semblance Optimization (see [KERN and SYMES, 1994]) and fixed in these experiments (see Figure B.3).

Air gun modeling software gave an a priori estimate of the source signature and radiation

pattern. Figure B.4 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 Figure B.11.

The data used in this work was derived from a marine survey in the Gulf of Mexico. This area of the Gulf contains a strong gas-sand-related direct hydrocarbon indicator at about 2.3s (see Figure B.1). The data was Radon transformed, respecting 3D cylindrical symmetry, to yield 48 plane-wave traces per midpoint gather. Slowness values range from p_{min}=.1158ms/m to p_{max}=.36468ms/m.

The numerical experiments described in this paper were performed on a single midpoint gather located near a logged well (Figure B.2). For the well near this midpoint gather, we obtained block sonic and density (gamma ray) logs in the 1.4–2.6s (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.

4.2 Experimental Results

In this section the two linear inversion experiments are compared and more detailed results are given. As stated in the previous subsection, in Experiment 1 the seismic source (gotten from an air gun model) was assumed to be accurate and not updated in the inversion. The three elastic parameter reflectivities were estimated. In Experiment 2 both the reflectivities and the seismic source were estimated.

In Subsection 3.3, a list of the steps required to approximate the resolution matrix are given. In step 1 the first level of approximation error is introduced. In exact arithmetic n steps of the conjugate gradient algorithm would be sufficient to determine $x \in \Re^n$ 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 many 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 are optimal in a subspace spanned by only j basis vectors (generally $j \ll n$). The second level of approximation error occurs in steps 2 and 3 where 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 $\tilde{R}$). The number (k) of eigenvectors used to form V is determined by specifying an upper bound on acceptable error and by using the following theorem (from [GOLUB and LOAN, 1989], p. 479):

Theorem 4.1 *Suppose that j steps of the Lanczos algorithm have been performed and that $S_j^t T_j S_j = diag(\theta_1, \ldots, \theta_j)$ is the Schur decomposition of the tridiagonal matrix T_j. If $Y_j = [y_1, \ldots, y_j] = Q_j S_j \in \Re^{j \times j}$ 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 above, 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; 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 $\tilde{R}$ for each experiment, we could not, therefore, rely on all the computed eigenvectors of the normal operator (columns of V). Figures B.7, and B.14 show the approximate eigenvalues graphed with the associated errors. Figures B.8, and B.15 show the relative error in the eigenvalues plotted against eigenvalue number (where "1" is always the smallest computed eigenvalue.)

The columns of the matrix V are the eigenvectors corresponding to the largest eigenvalues (the best determined in these experiments). These eigenvectors (and corresponding eigenvalues) were chosen because they meet the (arbitrary) tolerance criterion, namely having an approximation

error $\leq 30\%$.

Another well-known source of error in the Lanczos process comes from roundoff 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 (see [PARLETT, 1980]). The approach we implemented does not take advantage of these newer methods and is, therefore, subject to the problems of loss of orthogonality of the Lanczos vectors after a large number of iterations of the conjugate gradient procedure have been performed. We ran numerous experiments, therefore, to try to maximize the accuracy of the eigenvalue approximations while minimizing the loss of orthogonality of the Lanczos vectors. Trial and error indicated that about 30 iterations of the conjugate gradient algorithm is optimal for this particular problem. If we let $Q \in \Re^{n \times j}$ be the Lanczos matrix gotten after j iterations of the conjugate gradient algorithm, then ideally we should have that $Q^t * Q = I$ with I the $j \times j$ identity matrix. Plots of $Q^t * Q$ for the two experiments are shown in Figures B.9 and B.16.

One of the main difficulties one has with trying to understand the resolution matrix R is processing the information. The model resolution matrix $R \in \Re^{n \times n}$ may be quite large. The closer this matrix is to the identity $I \in \Re^{n \times n}$, the better the model resolution. 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 Figure B.18. The last pair of graphs corresponds to an area near the target for each experiment. However, 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).

However, it is clear that none of the columns shown is very close to a unit spike. We are after all only estimating 30 eigenvalues out of the total of 626 possible (for the P-wave impedance) in these inversion experiments. Wiggins [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.

Rather than plotting arbitrary columns of the large resolution matrix, one could calculate a function of the matrix termed the resolution spread (see [BACKUS and GILBERT, 1968]). One example is the Backus-Gilbert spread function which returns a single number for each resolution matrix:

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

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 noise in the data. Another choice is the vector-valued spread function:

$$Sp_i = \frac{\sum_{j=1}^{n}(i-j)^2 R_{ij}^2}{\sum_{j=1}^{n} R_{ij}^2}. \tag{4.1}$$

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. More work needs to be done to understand the meaning of the size of the elements in the spread vector for these experiments.

We show various comparisons of resolution spread. Figure B.10 overplots the resolution spread for each of the three reflectivities estimated in Experiment 1. These curves give little information (increase rapidly) at the water bottom (approximately 500m depth, where there is no information

in the data) and at depths below the target which is located in the interval of 2100-2200m (due to noise). Nonetheless, one can use this graph to compare relative resolution for the three different parameters estimated in one inversion experiment. The normalization of the spread function prevents the P-wave impedance resolution (the part of the resolution matrix with biggest norm) from overpowering the other two parameters' resolution information. Figure B.17 plots the same spread functions for the three reflectivities estimated in Experiment 2. Figure B.19 compares the resolution spread for the P-wave impedance reflectivity for Experiment 1 (dashed line) and Experiment 2 (solid line). Figure B.20 is a graph of the same resolution comparison for the S-wave velocity reflectivity. Finally, B.21 compares the resolution spread for the two experiments for the P-wave velocity divided by density reflectivity. One notes that in all three of these graphs, the interval of interest (1000–2300m depth) is better resolved in Experiment 2 (the reflectivity estimation job using the inversion-estimated source) than in Experiment 1 (the 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 Figures B.6 and B.13. 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.

5 Conclusion

Model resolution for seismic inverse problems was first described in classic papers by Backus and Gilbert 25 years ago. The limited data obtained about the subsurface due to measuring tools and computational discretization prevents the inversion process from being able to return model estimates that are the true parameters which generated the data. The singular value decomposition

of the forward (seismogram) model and generalized inverse allow us to quantify to what extent the model estimates are localized averages of the true model 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 1960's and early 1970's when it seemed a promising tool. However, computing the resolution matrix for inverse problems is still not common practice because it is prohibitive to calculate the SVD for realistic experiments. We present here an inexpensive way to approximate the resolution matrix. We 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 resolution matrix for the reflectivities. We illustrate this process on two reflectivity experiments with different sources. The resulting resolution spread curves agree with the 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.

Acknowlegement: The author would like to thank Dr. William Symes of Rice University for many encouraging and insightful suggestions. Thanks also go to Exxon Production Research Company and James 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 are 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.

A Appendix

A.1 Viscoelastic Modeling

We model the earth as a viscoelastic medium. The viscoelastic simulator we used was built on the solution approximation given in [AKI and RICHARDS, 1980] pp.153–155. 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.

Assume that the earth is a layered medium. Thus the parameters in our model vary only with depth, $z \equiv x_3$. By applying the Radon integral transform (or plane-wave decomposition) to the solution of the viscoelastic wave equation (and to common midpoint gathers of the data) we reduce the three-dimensional model to a family of one-dimensional models (see [TREITEL et al., 1982]). In effect, we have synthesized incident plane wave "shot" records parametrized by slowness, p, and by midpoint. Use of 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.

Neither mode conversion nor multiply reflected energy appear to be important in the data set used for the experiments (most likely 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 (see [BEYLKIN, 1985]),

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

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 reflectivity $\tilde{r}$ may be written

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

The geometric optics reflectivity amplitude A_P is

$$A_P(z,t,p) = \int_{-\infty}^{\infty} d\omega F_P(z,p) \exp\left[i\omega(t - 2\tau(z,p)) - |\omega| \left(1 + \frac{2i}{\pi} \ln\left|\frac{\omega}{\omega_0}\right|\right) \alpha(z,p)\right].$$

The vertical travel time of plane waves with slowness, p, is

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

and

$$\alpha(z,p) = \int_{z_0}^{z} dy \frac{1}{q_P(y)v_P(y)(1-(pv_P(y))^2)^{1/2}}$$

is the corresponding attenuation factor. An algebraic combination of background parameters $v_P(z)$, $v_S(z)$, and $\rho(z)$, and slowness p gives $F_P(z,p)$. The temporal frequency variable is denoted ω; z_0 is the source depth, and q_P is the P-wave attenuation factor. Similar expressions define the geometric

optics amplitudes A_S and A_d. The reference frequency ω_0 calibrates the (frequency-dependent) velocity. Waves at frequency ω_0 move with the P-wave velocity v_P.

The integrals described above are approximated by the trapezoidal rule in the modeling 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.

A.2 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 nonlinear in the P-wave velocity v_p. Although we did perform an initial inversion for the long-wavelength portion of the P-wave velocity, we will not discuss that experiment in this paper. (This inversion result is the fixed background velocity used in both experiments described in this paper. It is shown in Figure B.3.) 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 *O*utput *L*east *S*quares (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

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

where $d^{obs}(t,p)$ is the "observed" $p-\tau$ data, and $\gamma(t,p)$ is a *conditioning weight* factor. The

conditioning factor $\gamma(t,p)$ enhances the resolution of deeper events. It is a model-based gain. 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 ($Q_p \longrightarrow \infty$) no conditioning is necessary.

The elastic reflectivities and source parameters together have a nonlinear (bilinear) influence, so quadratic minimization algorithms cannot be used directly in Experiment 2. Instead, we used a method known as coordinate search or alternation. We will hereafter refer to a "round of alternation" to be 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 solution of simple linear least-squares problems. Obviously, quasi-Newton methods could be applied to the problem and would likely 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 2 rounds of alternation.

(For a more thorough description of joint source-reflectivity inversion for a simplified model see [MINKOFF and SYMES, 1995].)

B Appendix

B.1 Seismic Field Data — Geometry of the Geophysical Experiment and Preprocessing

The data used in this work was 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 15m with a minimum source-receiver separation of 148m. The shot interval was 22.5m. Each group contained 17 equally spaced and equally weighted hydrophones. The data was recorded without a low-cut filter. A 110 Hz high-cut filter was applied. The sampling rate was 2ms, and the total record length was 5s.

This area of the Gulf contains a strong gas-sand-related direct hydrocarbon indicator at about 2.3s. The stack shows this target horizon to be embedded in a sequence of nearly horizontal strata, beginning at roughly 1.5s (see Figure B.1). Therefore, layered medium modeling 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 was Radon transformed, respecting 3D cylindrical symmetry, to yield 48 plane-wave traces per midpoint gather. Slowness values range from p_{min}=.1158ms/m to p_{max}=.36468ms/m. To remove diffraction artifacts originating in the shallow subsurface, the plane-wave data were

time migrated in the midpoint dip domain, then modeled to $p_{mid} = 0$. This last step collapses diffractions while still preserving layered reflection amplitudes. While the diffracting structures are still visible in the upper 1.5s, 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 prestack $p - \tau$ data (see Figure B.2).

In order to lighten the computational burden of inversion, we performed our calculations on a low-pass filtered version of the $p - \tau$ 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 Figure B.2) located near a logged well.

References

[AKI and RICHARDS, 1980] AKI, K. and RICHARDS, P. (1980). *Quantitative Seismology: Theory and Methods.* Freeman, San Francisco.

[ASSOUS and COLLINO, 1990] ASSOUS, F. and 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 and GILBERT, 1968] BACKUS, G. and GILBERT, F. (1968). The resolving power of gross earth data. *Geophys. J. R. Astr. Soc.*, 16:169–205.

[BEYLKIN, 1985] 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 et al., 1985] BISHOP, T., BUBE, K., CUTLER, R., LANGAN, R., LOVE, P., RESNICK, J., SHUEY, R., SPINDLER, D., and WYLD, H. (1985). Tomographic determination of velocity and depth in laterally varying media. *Geophysics*, 50:903–923.

[FRANKLIN, 1970] FRANKLIN, J. (1970). Well-posed stochastic extensions of ill-posed linear problems. *Journal of Mathematical Analysis and Applications*, 31:682–716.

[GOLUB and LOAN, 1989] GOLUB, G. and LOAN, C. V. (1989). *Matrix Computations.* The Johns Hopkins University Press, Baltimore.

[GOLUB and O'LEARY, 1989] GOLUB, G. and O'LEARY, D. (1989). Some history of the conjugate gradient and lanzos algorithms: 1948-1976. *SIAM Review*, 31:50–102.

[HESTENES, 1980] HESTENES, M. (1980). *Conjugate Direction Methods in Optimization.* Springer-Verlag, Berlin.

[HESTENES and STIEFEL, 1952] HESTENES, M. and STIEFEL, E. (1952). Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Stand.*, 49:409–436.

[JACKSON, 1979] 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 and NOLET, 1978] KENNETT, B. and NOLET, G. (1978). Resolution analysis for discrete systems. *Geophys. J. R. astr. Soc.*, 53:413–425.

[KERN and SYMES, 1994] KERN, M. and SYMES, W. (1994). Inversion of reflection seismograms by differential semblance analysis: algorithm structure and synthetic examples. *Geophysical Prospecting*, 42:565–614.

[LANCZOS, 1950] 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 and MCMECHAN, 1991] MARTINEZ, R. and MCMECHAN, G. (1991). $\tau - p$ seismic data for viscoelastic media – part 2: linearized inversion. *Geophysical Prospecting*, 39:157–181.

[MENKE, 1989] MENKE, W. (1989). *Geophysical Data Analysis: Discrete Inverse Theory*. Academic Press, Inc., San Diego.

[MINKOFF and SYMES, 1994] MINKOFF, S. and SYMES, W. (1994). Viscoelastic modeling and inversion of a marine data set. In *Proc. 64th Annual International Meeting*, pages 1016–1019, Los Angeles, CA., U.S.A. Society of Exploration Geophysicists. Expanded abstract.

[MINKOFF and SYMES, 1995] MINKOFF, S. and SYMES, W. (1995). Estimating the energy source and reflectivity by seismic inversion. *Inverse Problems*, 11:383–395.

[ORY and PRATT, 1995] ORY, J. and PRATT, R. (1995). Are our parameter estimators biased? the significance of finite-difference regularization operators. *Inverse Problems*, 11:397–424.

[PARLETT, 1980] PARLETT, B. (1980). *The Symmetric Eigenvalue Problem*. Prentice-Hall.

[TREITEL et al., 1982] TREITEL, S., GUTOWSKI, P., and WAGNER, D. (1982). Plane-wave decomposition of seismograms. *Geophysics*, 47:1375–1401.

[WIGGINS, 1972] WIGGINS, R. A. (1972). The general linear inverse problem: implication of surface waves and free oscillations for earth structure. *Reviews of Geophysics and Space Physics*, 10(1):251–285.

Figure B.1: The stacked section of marine data. The diffracted energy originating in the shallow subsurface was suppressed by prestack time migration in the offset-midpoint slowness domain. Modeling 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 2.3s, is embedded in a flat-lying sequence.

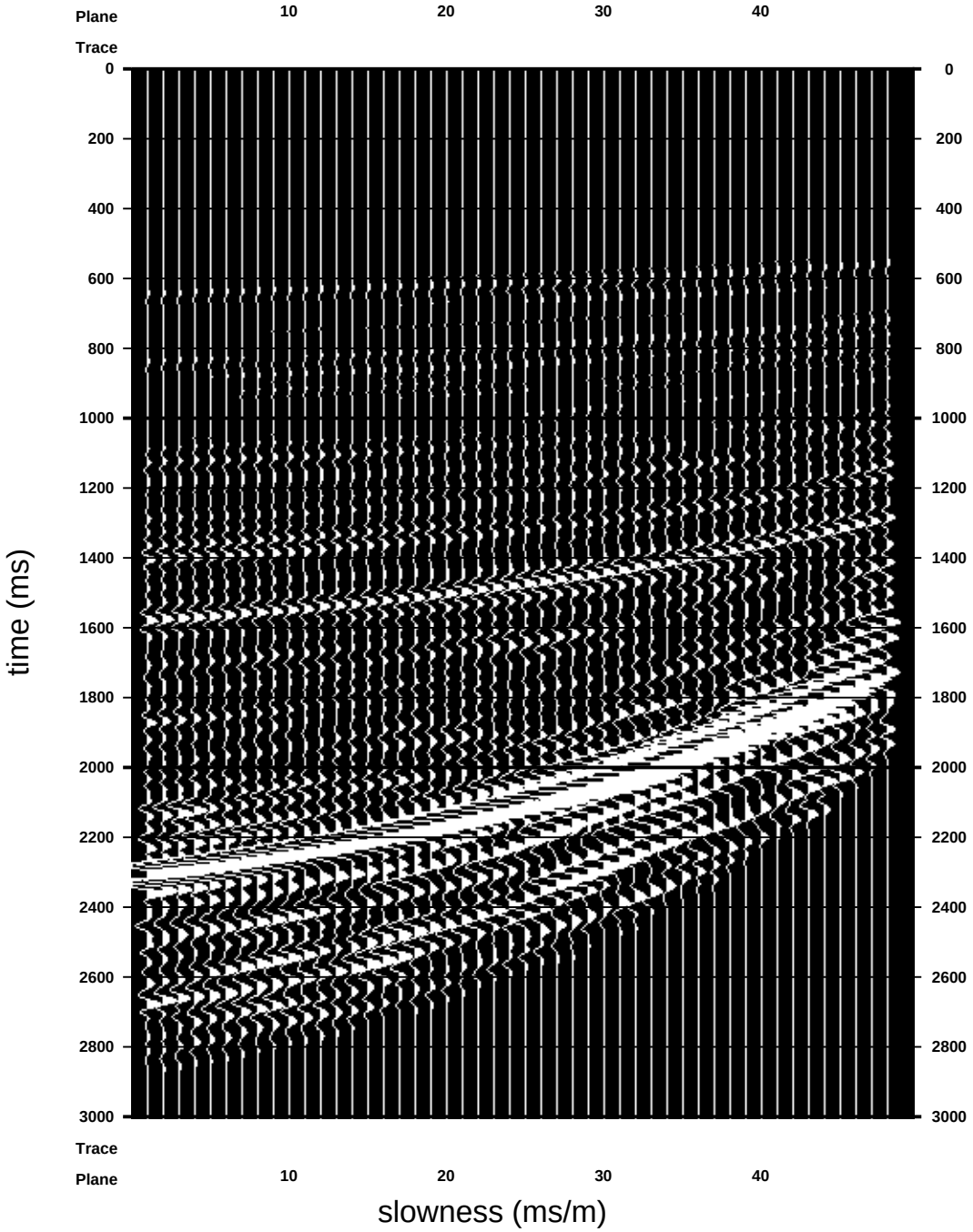

Figure B.2: The common midpoint data gather (CMP6) used for the two experiments contrasted in this paper.

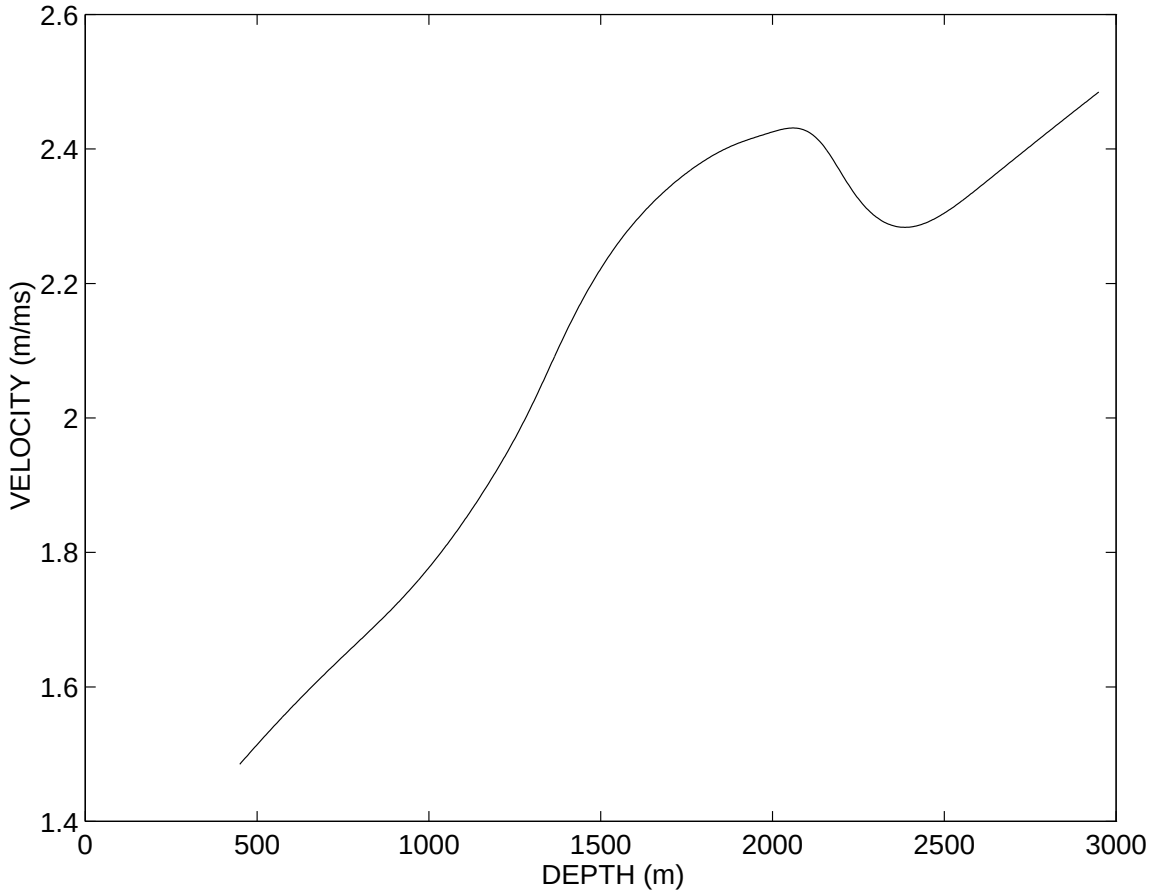

Figure B.3: The inversion-estimated P-wave background velocity used for both inversion experiments.

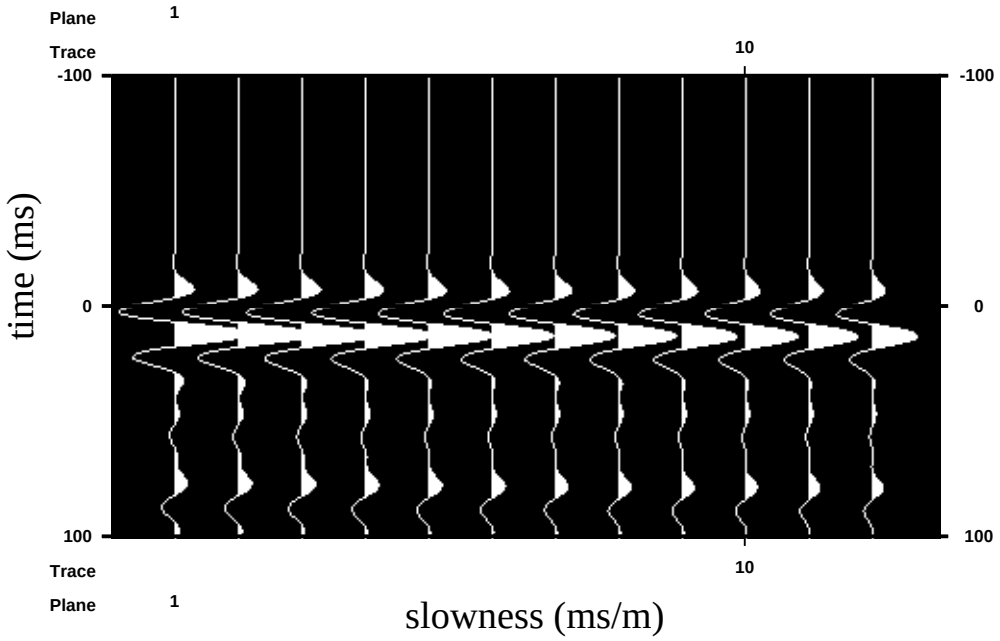

Figure B.4: The air gun model source estimate used in Experiment 1 with every fourth trace shown.

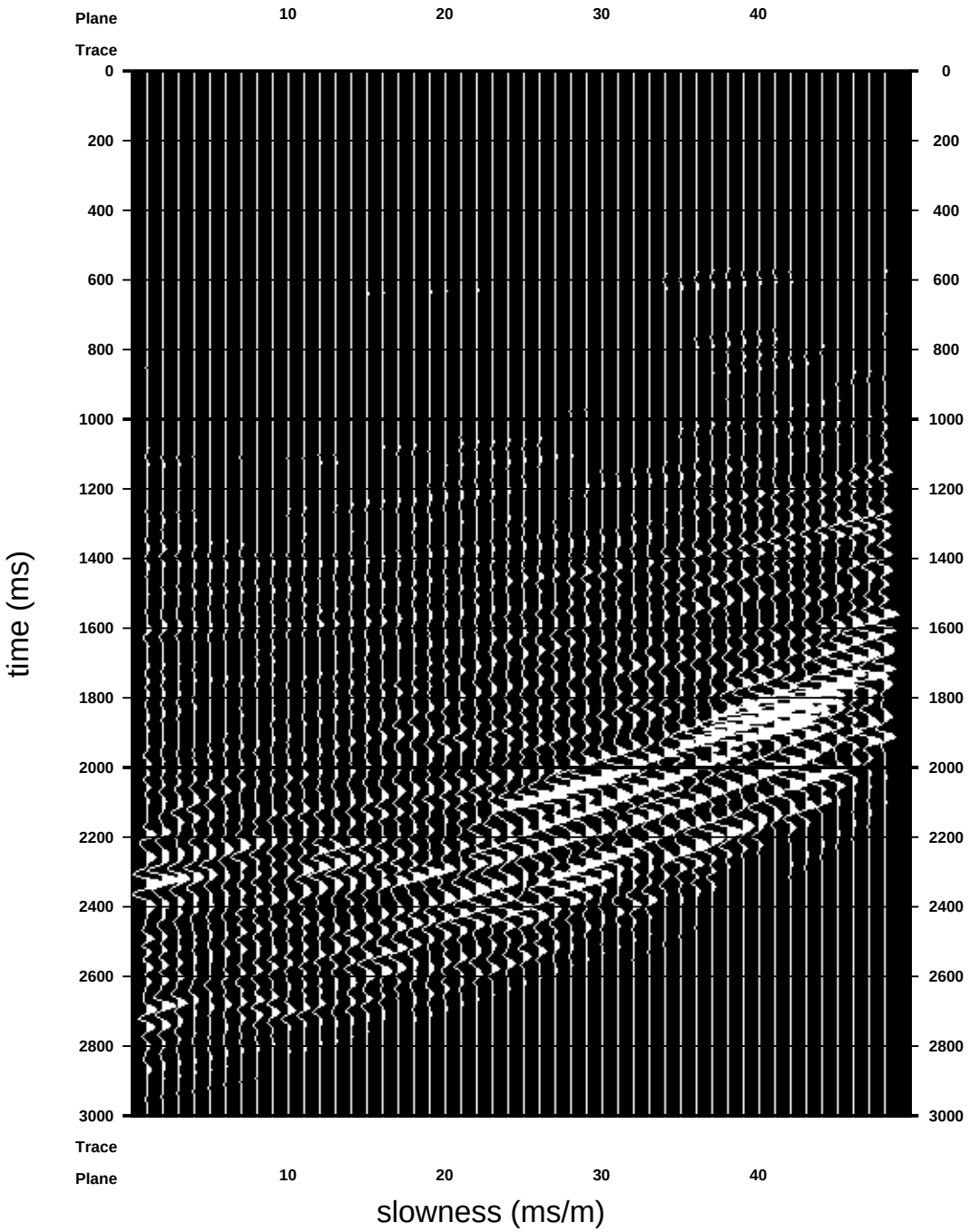

Figure B.5: Difference between actual and predicted data gotten from inverting the reflectivities with the air gun source fixed (Experiment 1). The misfit is plotted on the same scale as the actual data.

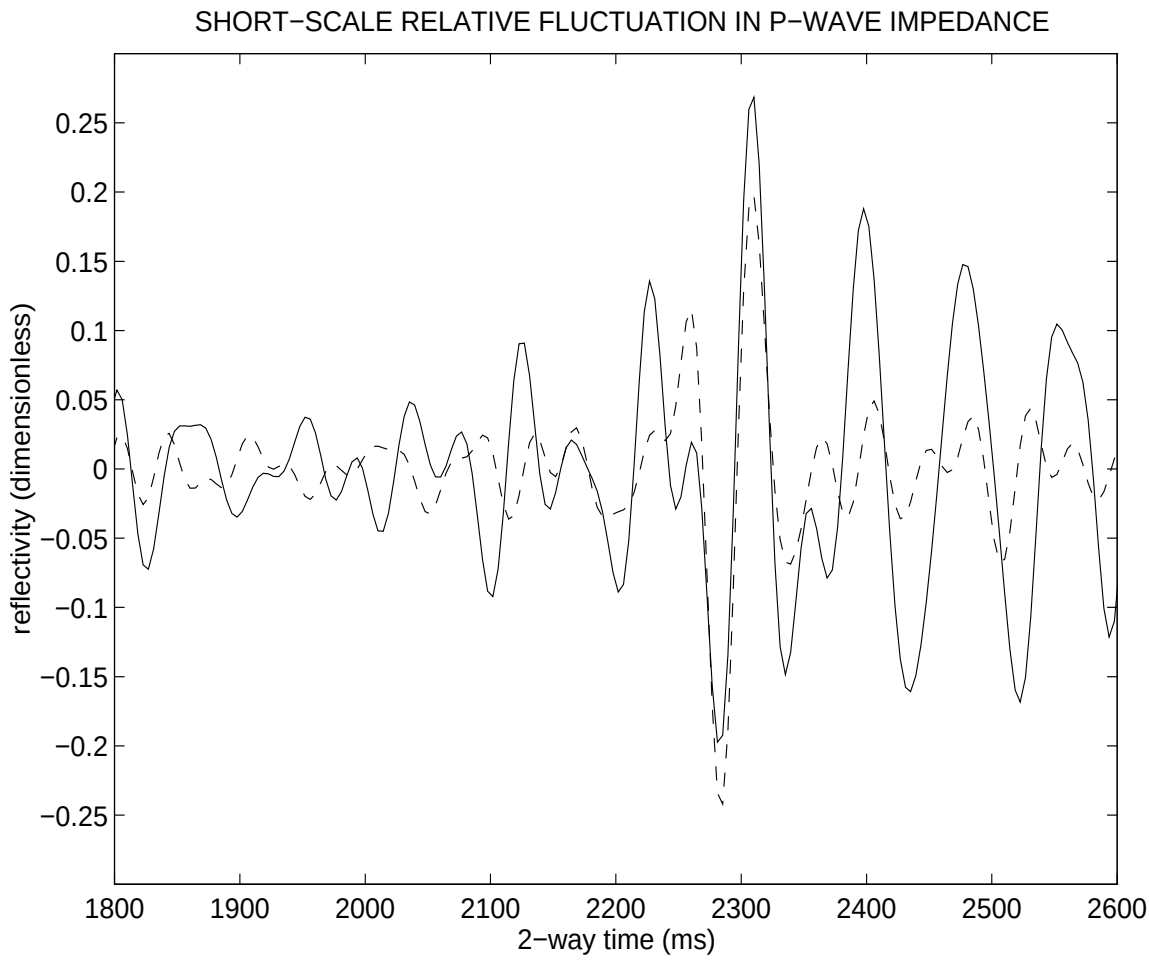

Figure B.6: Comparison of the independent well-log measurement of the relative short-scale fluctuation in the P-wave impedance with the result of inversion done on CMP6 using the air gun model source estimate (Experiment 1). The solid line shows the inversion result (scaled). The dashed line shows the 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.

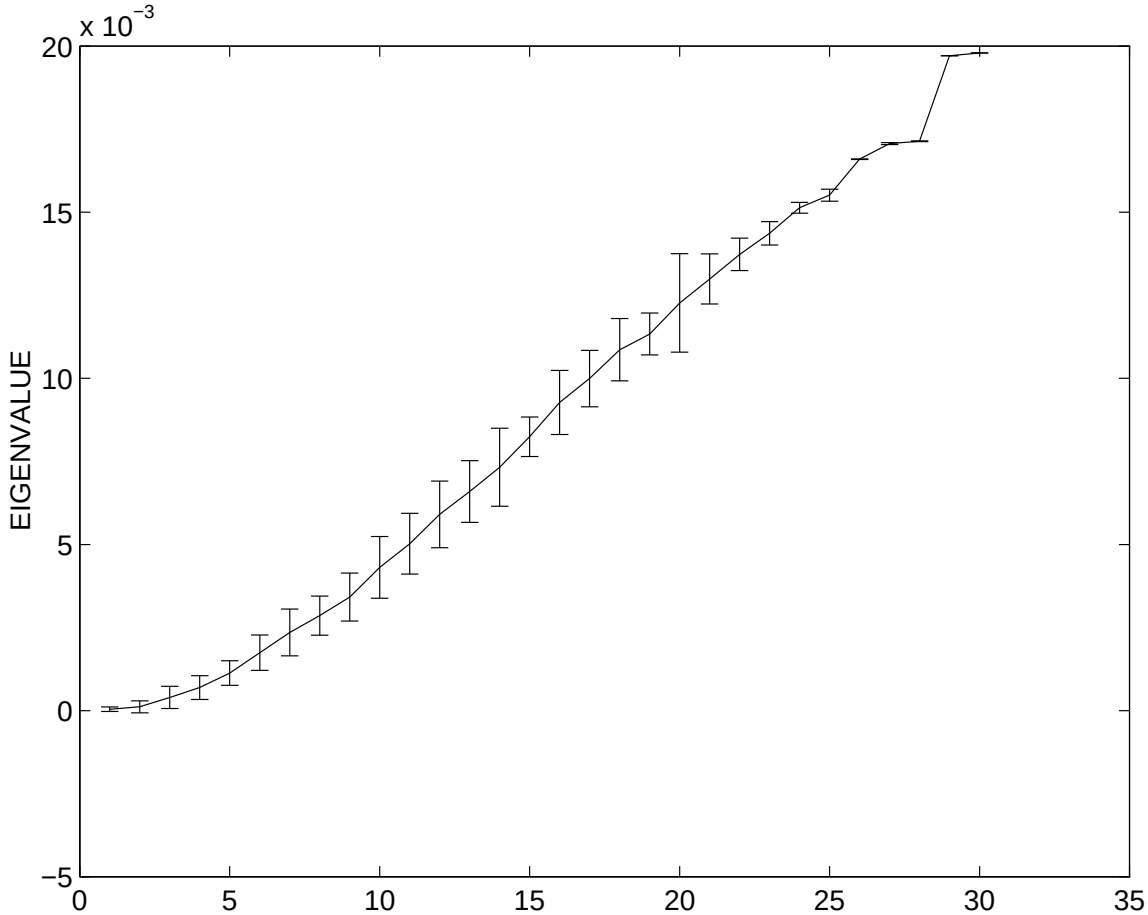

Figure B.7: The approximate eigenvalues of the normal operator for the reflectivity inversion with a fixed air gun model source (Experiment 1). The errors in the eigenvalues are shown as errorbars on the graph.

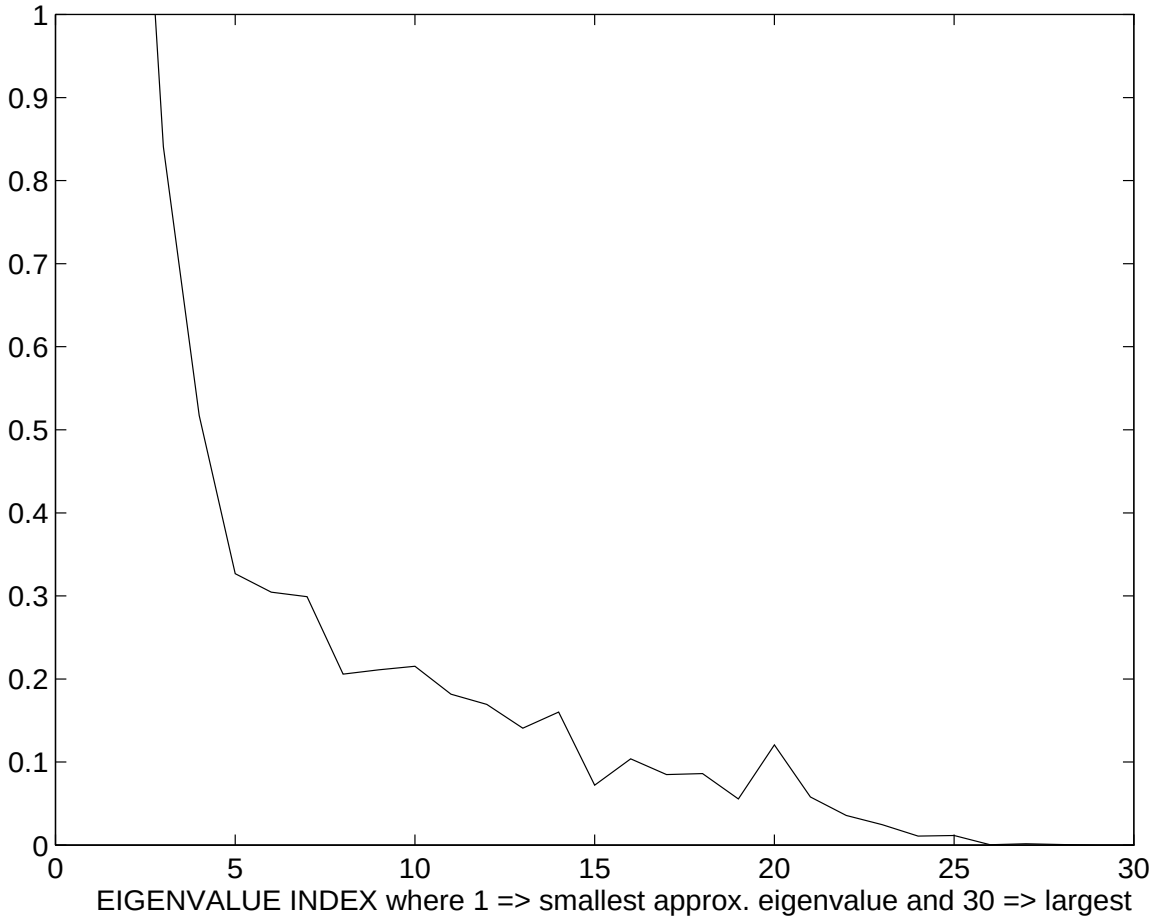

Figure B.8: The relative error in the approximate eigenvalues of the normal operator for the reflectivity inversion with a fixed air gun model source (Experiment 1).

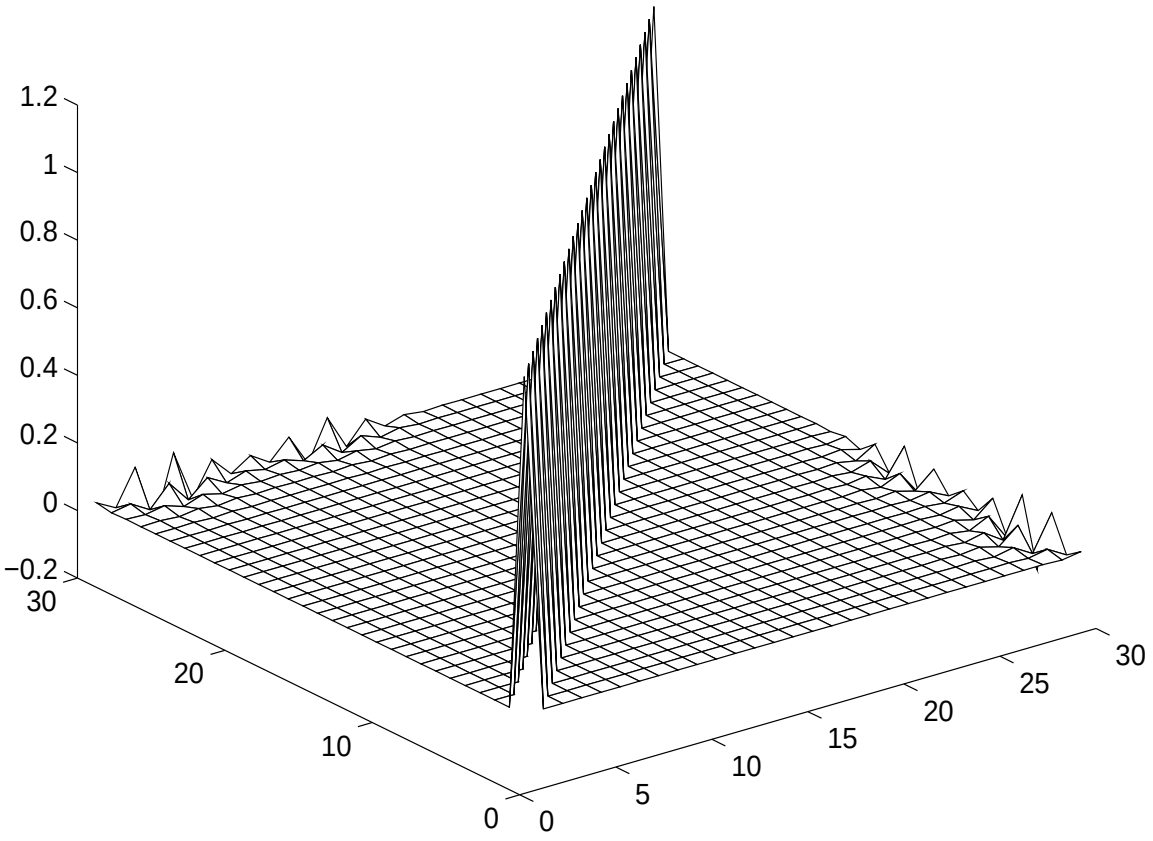

Figure B.9: The dot product of Lanczos matrix with itself for Experiment 1 after 30 iterations of the conjugate gradient algorithm.

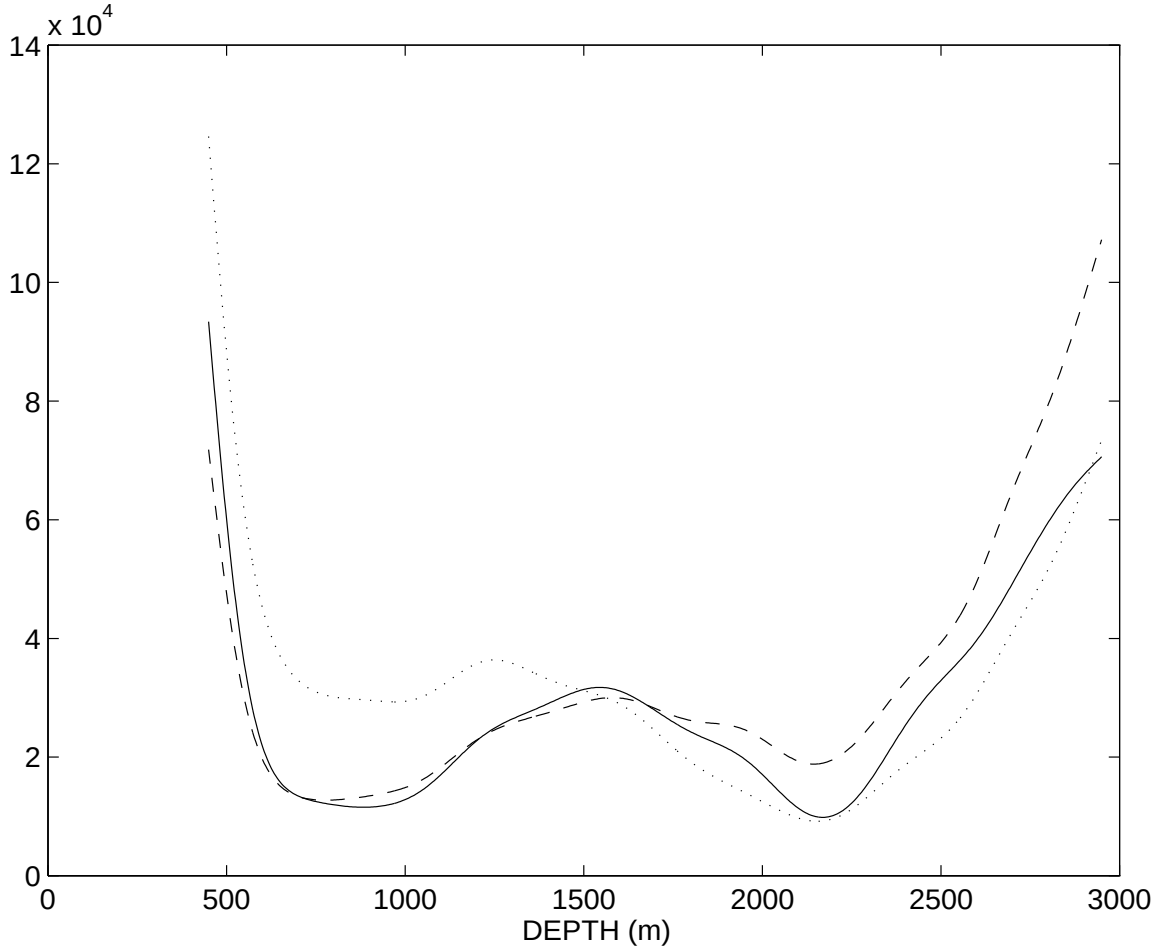

Figure B.10: Graph of the spread of the resolution matrix for Experiment 1. The solid line corresponds to the spread for the P-wave impedance reflectivity estimate. The dashed line describes the spread for the S-wave velocity reflectivity. The dotted line is the spread for the P-wave velocity divided by density reflectivity.

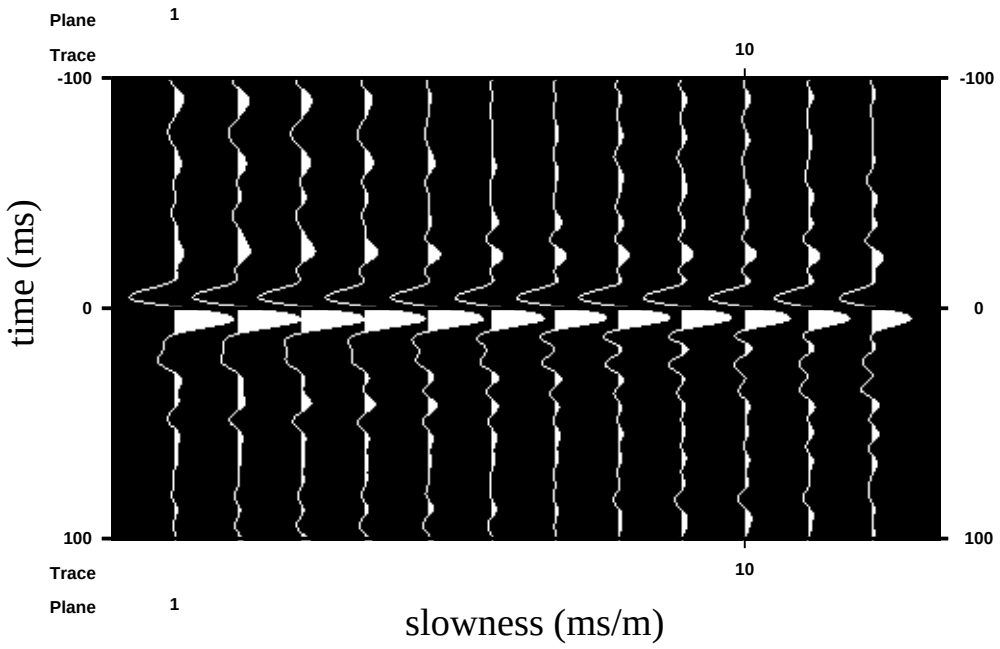

Figure B.11: The estimated anisotropic source from the linear source-reflectivity inversions where only every fourth trace is shown for clarity (Experiment 2).

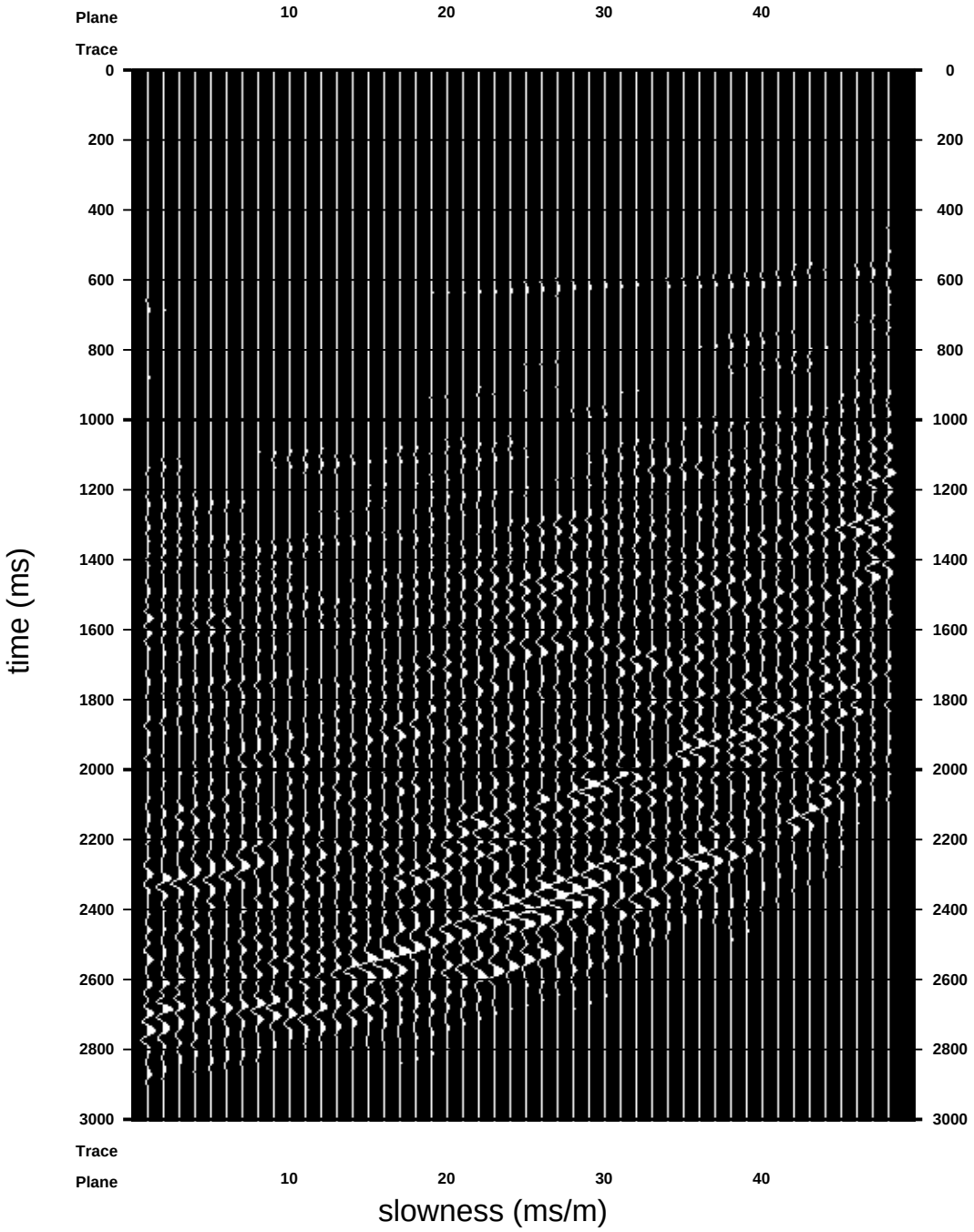

Figure B.12: Difference between actual and predicted data gotten from inverting for reflectivities and an anisotropic source. The misfit is plotted on the same scale as the actual data.

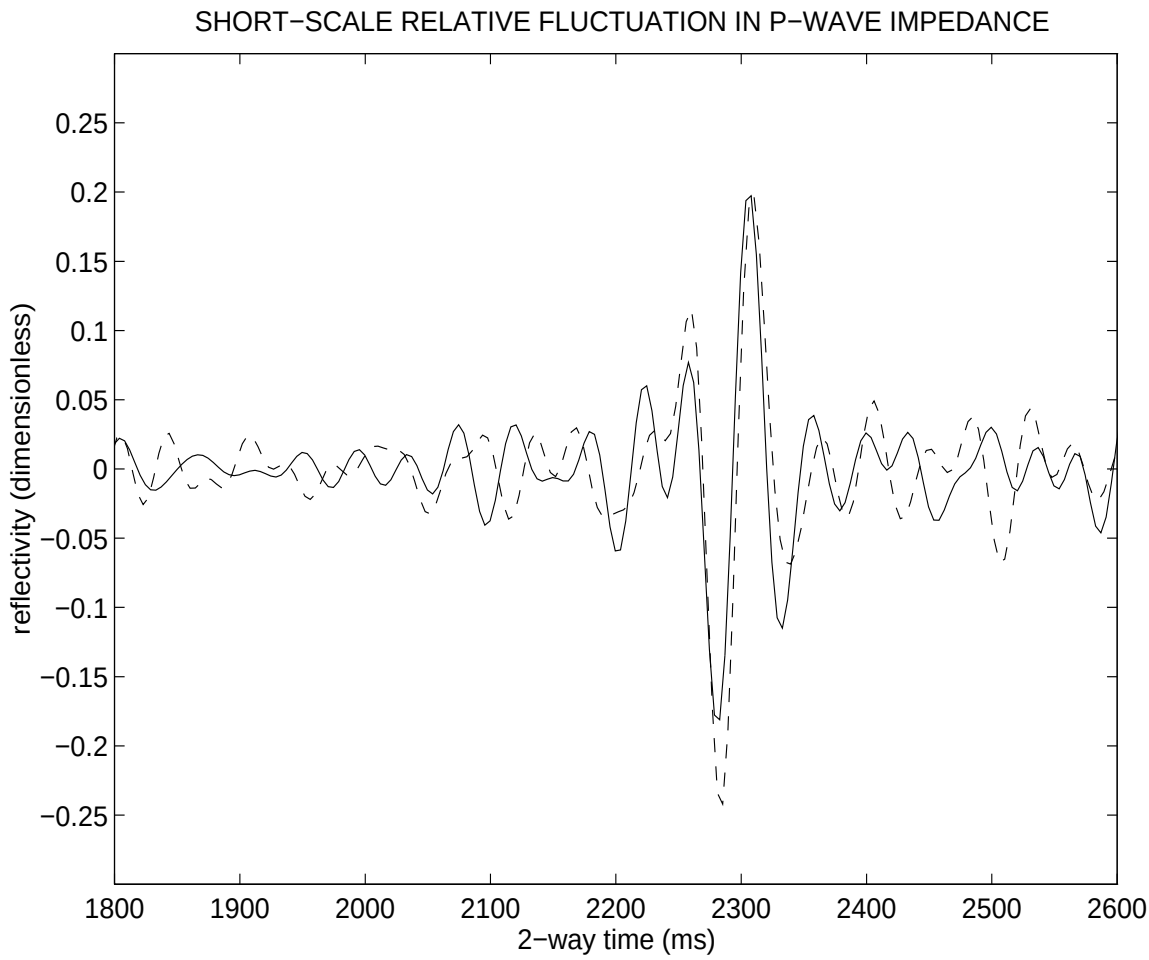

Figure B.13: Comparison of the independent well-log measurement of the relative short-scale fluctuation in the P-wave impedance with the result of inversion done on CMP6 using an anisotropic source estimate from inversion (Experiment 2). The solid line shows the inversion result (scaled and shifted left 44ms). 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.

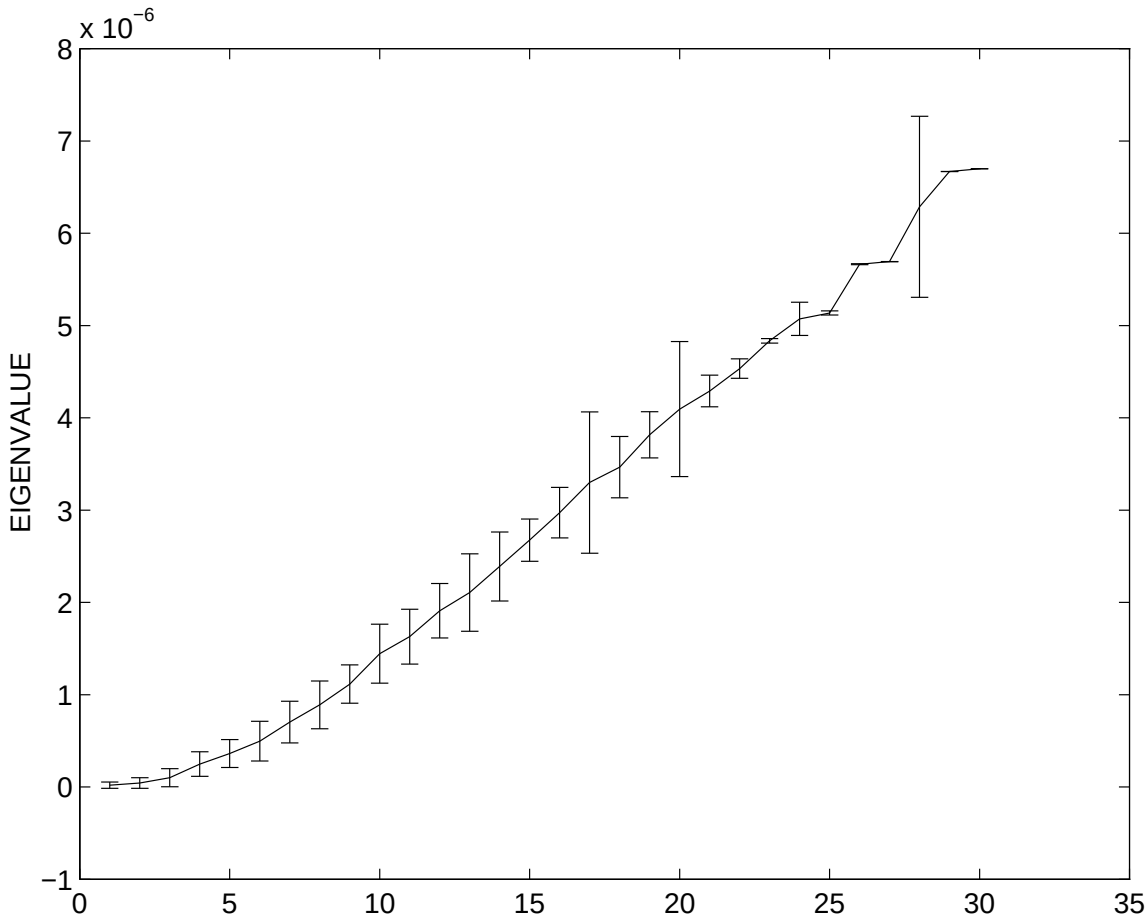

Figure B.14: The approximate eigenvalues of the normal operator for the reflectivity inversion with an inversion-estimated anisotropic source (Experiment 2). The errors in the eigenvalues are shown as errorbars.

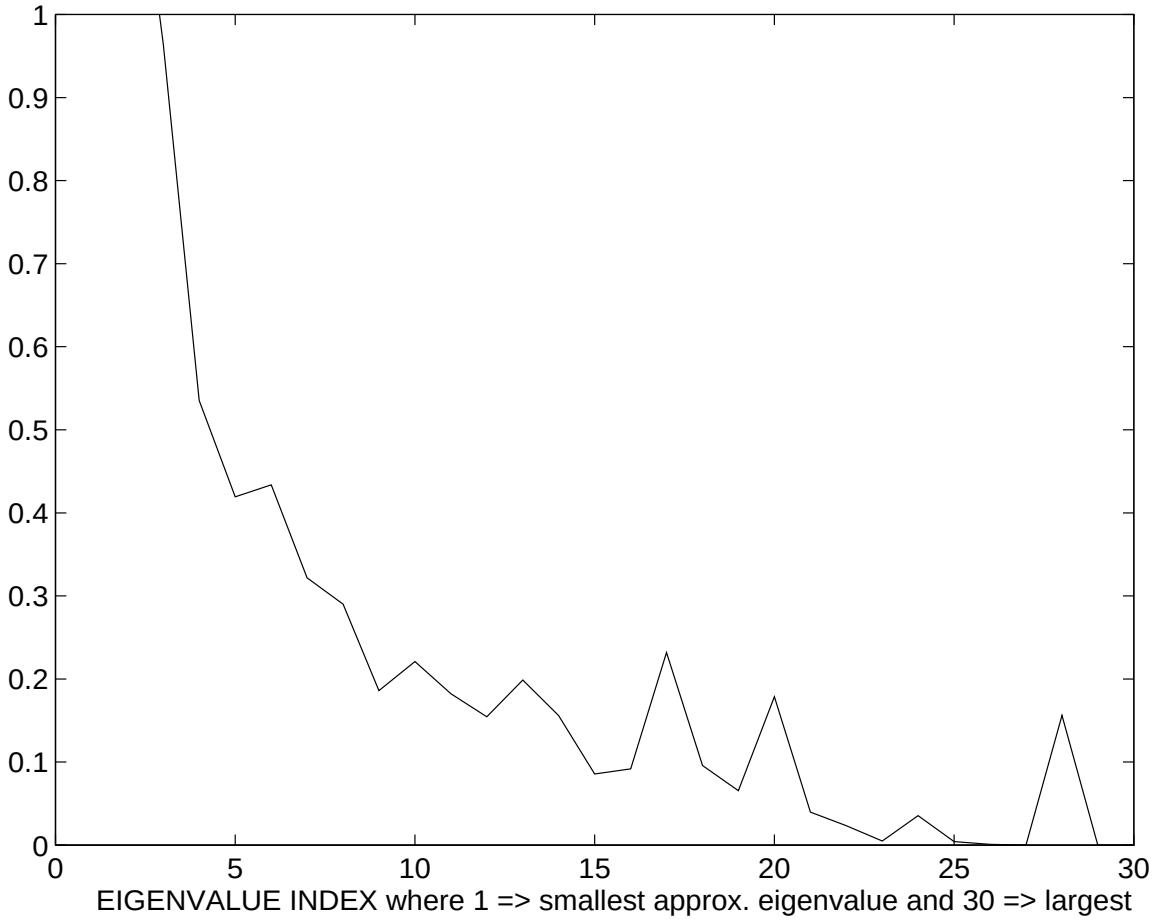

Figure B.15: The relative error in the approximate eigenvalues of the normal operator for the reflectivity inversion with an anisotropic inversion-estimated source (Experiment 2).

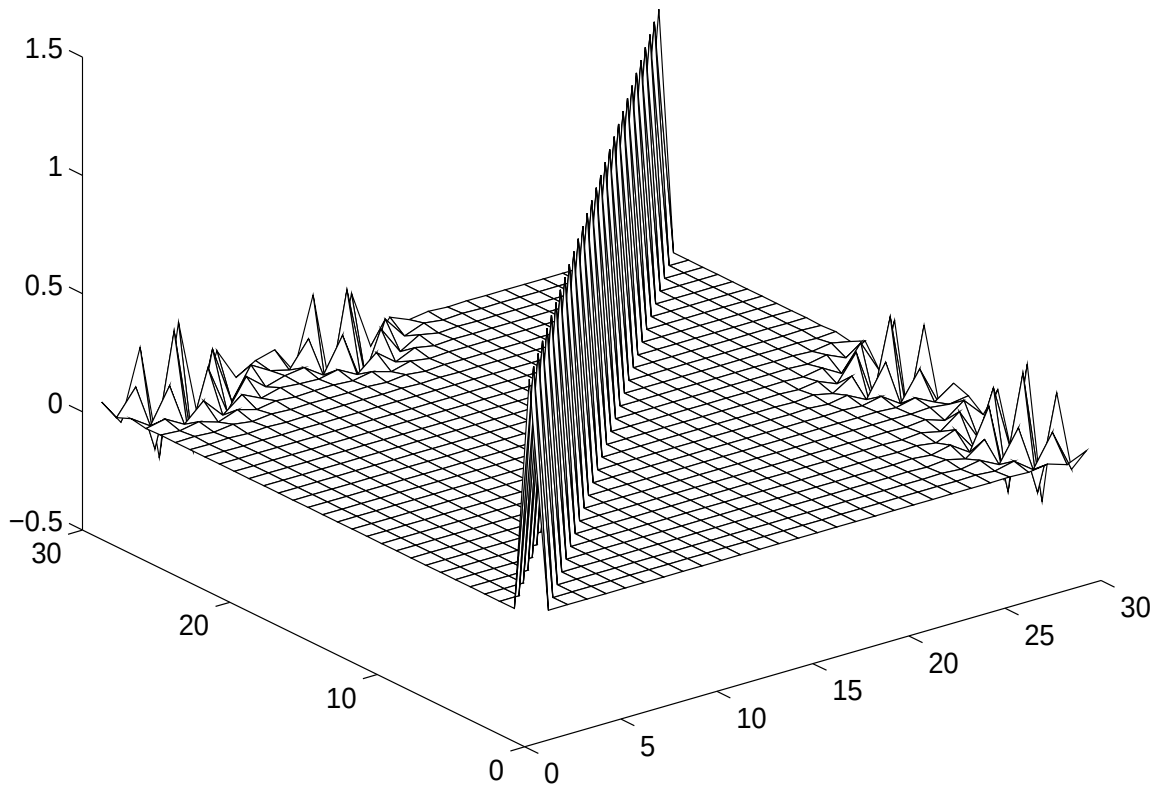

Figure B.16: The dot product of Lanczos matrix with itself for Experiment 2 after 30 iterations of the conjugate gradient algorithm.

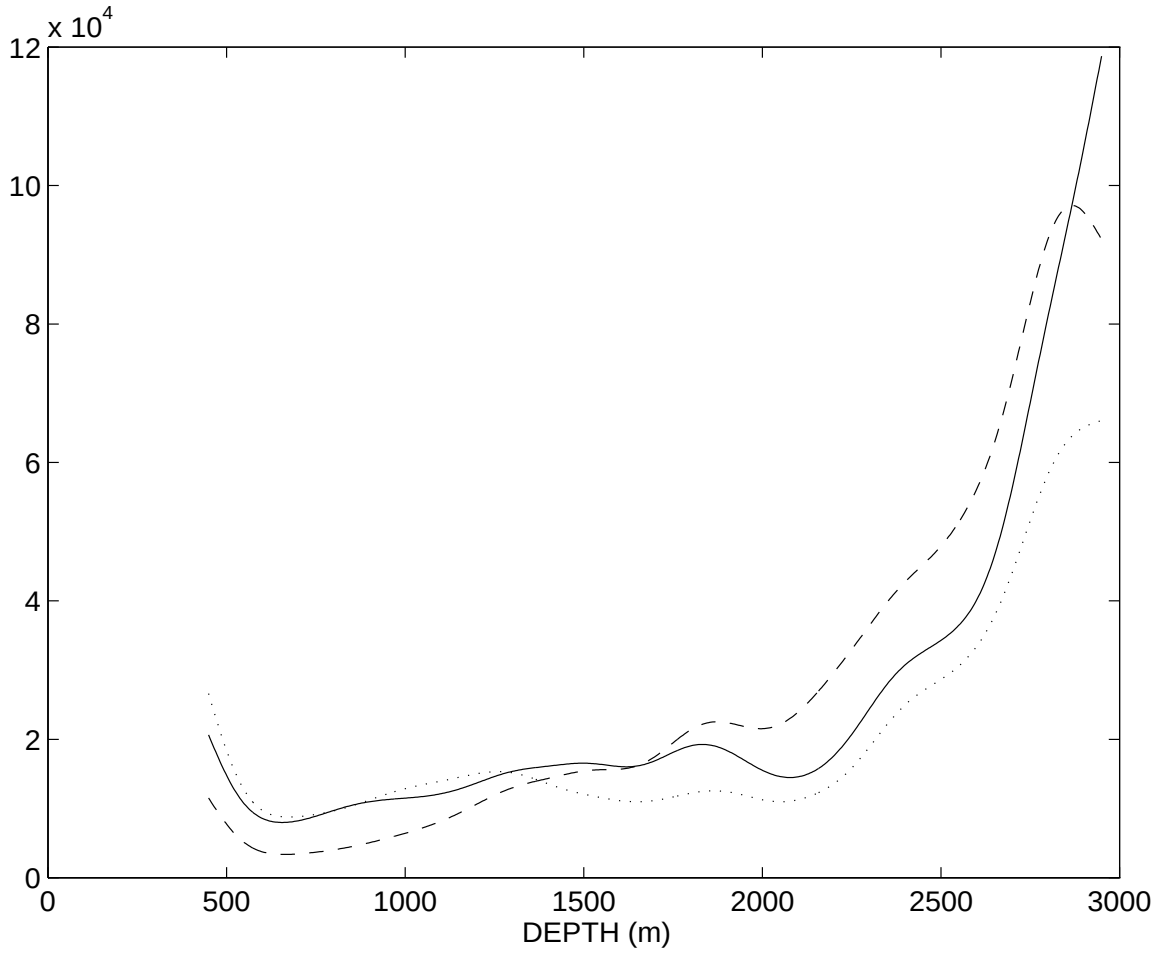

Figure B.17: Graph of the spread of the resolution matrix for Experiment 2. The solid line corresponds to the spread for the P-wave impedance reflectivity estimate. The dashed line describes the spread for the S-wave velocity reflectivity. The dotted line is the spread for the P-wave velocity divided by density reflectivity.

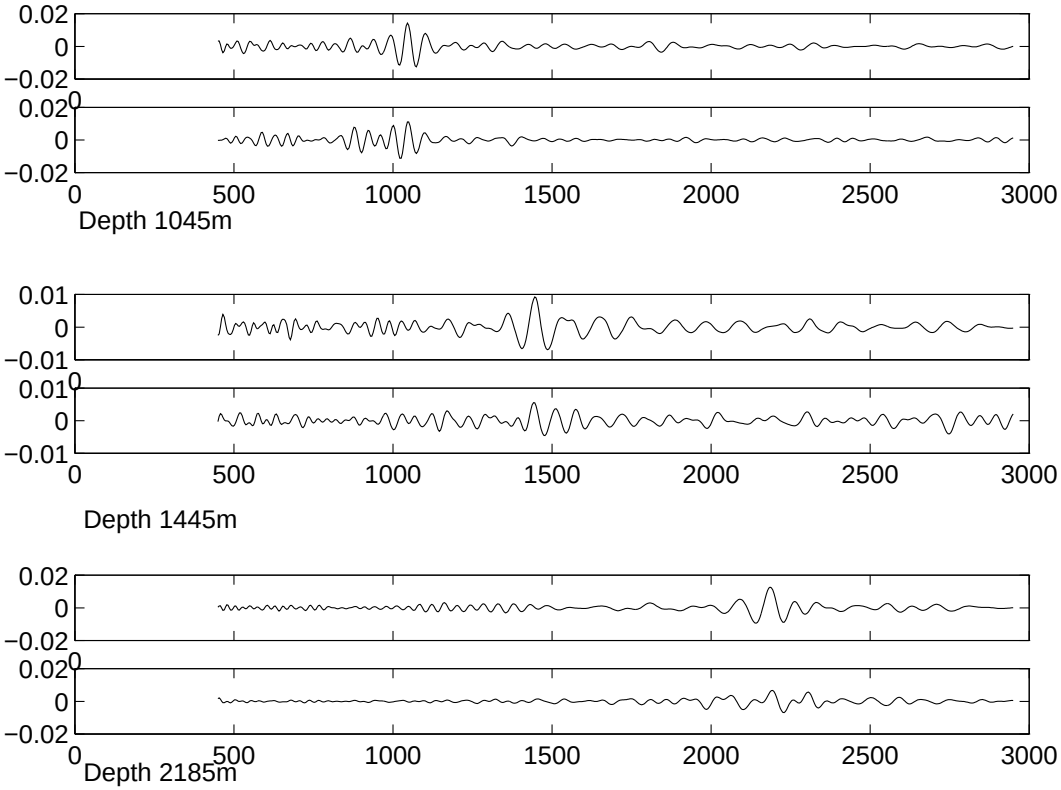

Figure B.18: Three selected columns of the resolution matrices for the two experiments. The top two graphs correspond to depth 1045m. The middle two graphs are the column of the resolution matrices corresponding to 1445m. The bottom two graphs correspond to depth 2185m (a region between where the two experiments place the target). 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).

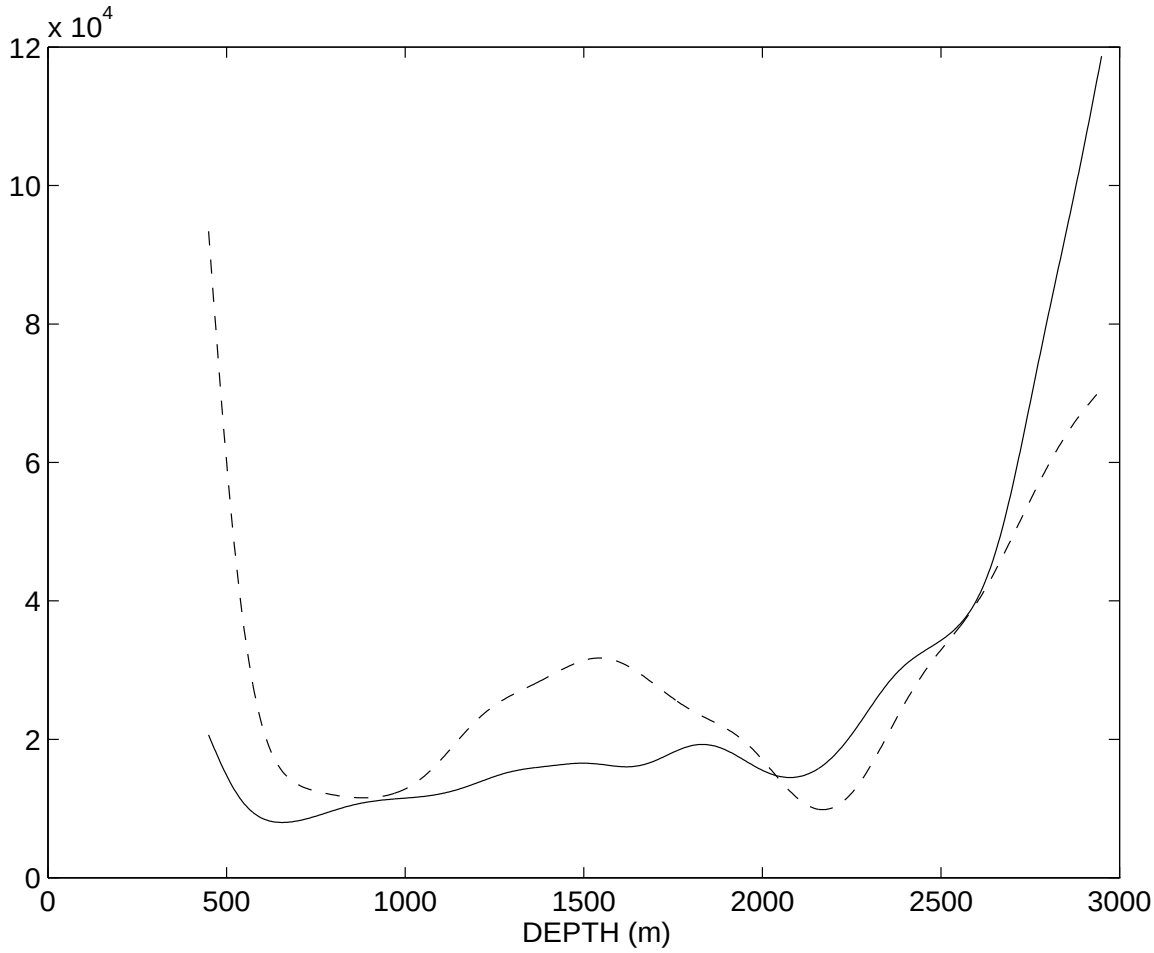

Figure B.19: The graph of the resolution spread for two estimates of the P-wave impedance reflectivity. The solid line corresponds to Experiment 2; the dashed line to Experiment 1.

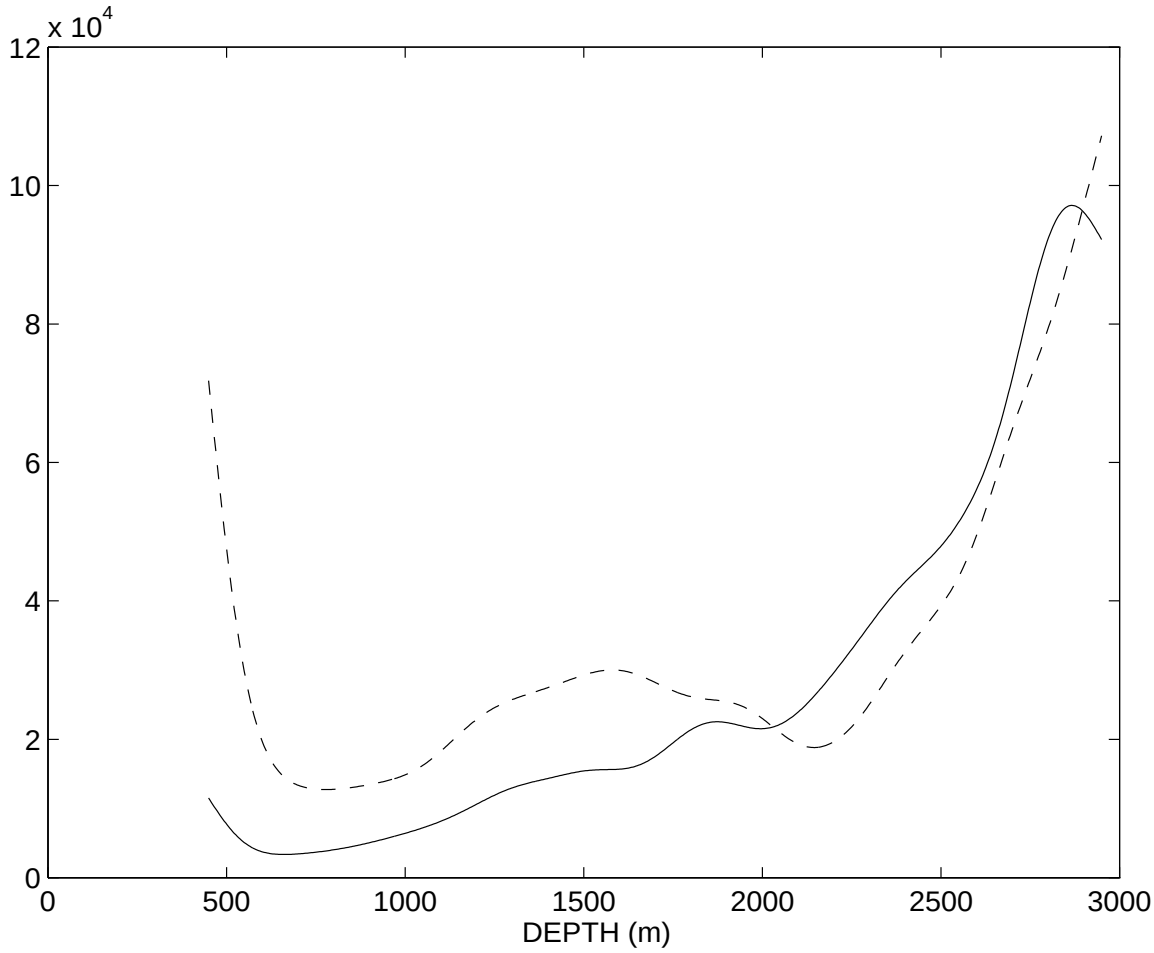

Figure B.20: The graph of the resolution spread for two estimates of the S-wave velocity reflectivity. The solid line corresponds to Experiment 2; the dashed line to Experiment 1.

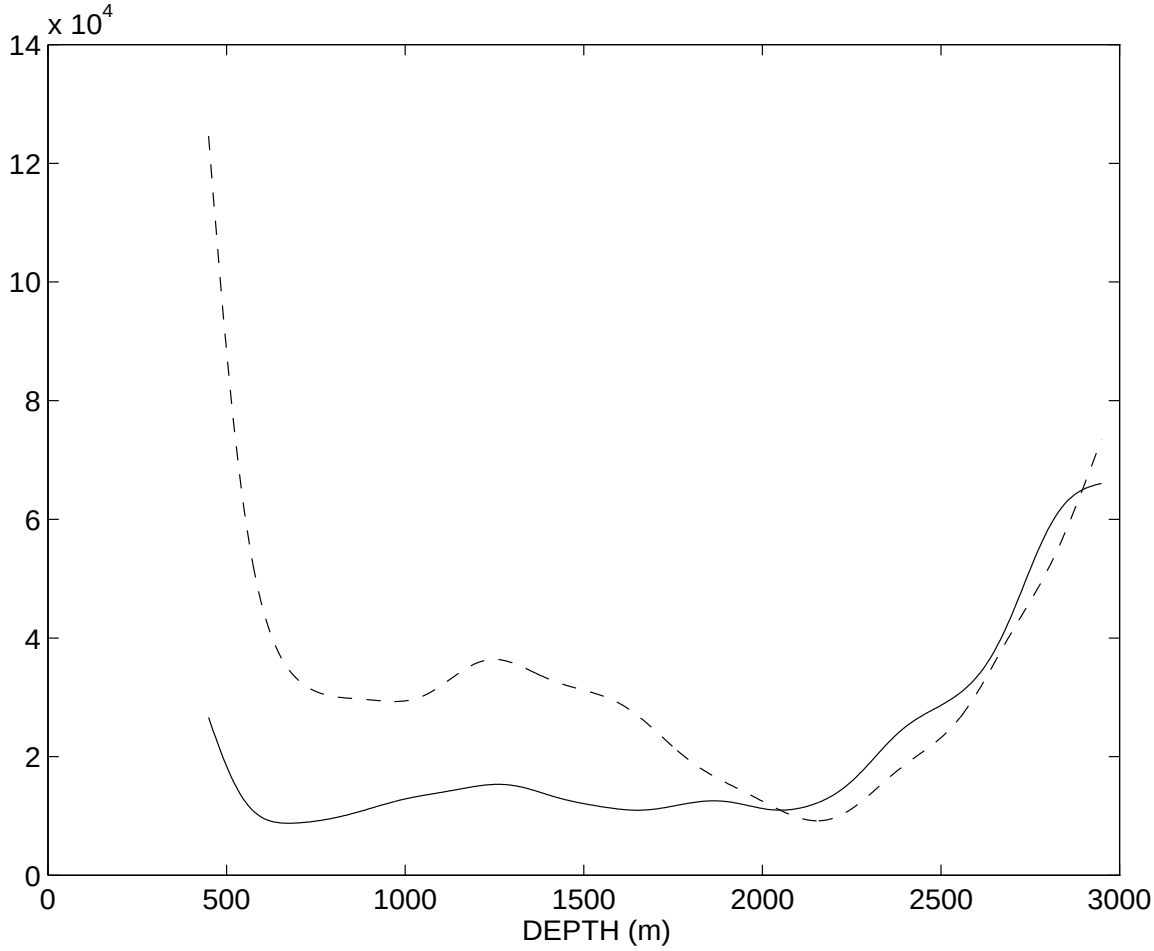

Figure B.21: The graph of the resolution spread for two estimates of the P-wave velocity divided by density reflectivity. The solid line corresponds to Experiment 2; the dashed line to Experiment 1.