GEOPHYSICS, VOL. 69, NO. 2 (MARCH-APRIL 2004); P. 533–546, 27 FIGS.
10.1190/1.1707073

Incorporating geologic information into reflection tomography

Robert G. Clapp*, Biondo Biondi*, and Jon F. Claerbout*

ABSTRACT

In areas of complex geology, prestack depth migration is often necessary if we are to produce an accurate image of the subsurface. Prestack depth migration requires an accurate interval velocity model. With few exceptions, the subsurface velocities are not known beforehand and should be estimated. When the velocity structure is complex, with significant lateral variations, reflection-tomography methods are often an effective tool for improving the velocity estimate. Unfortunately, reflection tomography often converges slowly, to a model that is geologically unreasonable, or it does not converge at all. The large null space of reflection-tomography problems often forces us to add a sparse parameterization of the model and/or regularization criteria to the estimation. Standard tomography schemes tend to create isotropic features in velocity models that are inconsistent with geology. These isotropic features result, in large part, from using symmetric regularization operators or from choosing a poor model parameterization. If we replace the symmetric operators with nonstationary operators that tend to spread information along structural dips, the tomography will produce velocity models that are geologically more reasonable. In addition, by forming the operators in helical 1D space and performing polynomial division, we apply the inverse of these space-varying anisotropic operators. The inverse operators can be used as a preconditioner to a standard tomography problem, thereby significantly improving the speed of convergence compared with the typical regularized inversion problem. Results from 2D synthetic and 2D field data are shown. In each case, the velocity obtained improves the focusing of the migrated image.

INTRODUCTION

Obtaining an accurate velocity model is an essential part of imaging complex strucutures. In complex environments, reflection tomography is often an effective tool for improving the velocity estimate. The challenge is that reflection tomography is generally an underdetermined problem (Stork and Clayton, 1992). To obtain a pleasing result, the general solution is to impose some type of regularization criteria, thereby limiting inversion solutions to large singular values (Rowbatham and Pratt, 1997) or characterizing the model through a limited number of coefficients (Ji, 1997). These methods all create velocity models that can correctly model the recorded travel times but are often geologically unrealistic.

To create more geologically feasible velocity models and to speed up convergence of the tomography problem, Michelena and Harris (1991) suggested using varying-size grid cells. Unfortunately, such a parameterization is prone to error when the wrong-size blocks are chosen (Delprat-Jannaud and Lailly, 1992). Other authors have suggested locally clustering grid cells (Carrion, 1991) or characterizing the velocity model as a series of layers (Kosloff et al., 1996). These methods are also susceptible to errors when the wrong parameterization is chosen.

An attractive alternative approach is to add a model-regularization term to the objective function (van Trier, 1990). In theory, this regularization term should be the inverse model covariance matrix (Tarantola, 1987) obtained from some a priori information sources. For tomography, a geologist's structural model of the area, well-log information, or preliminary stack or migration results all could provide useful information. Incorporating these varied information sources into the objective function is problematic. Geostatisticians have successfully combined these mixed types of information (Hirsche et al., 1997). However, geostatistics solves many local least-squares problems, whereas tomography involves finding a global solution. Therefore, standard geostatistical methods are ill-suited for velocity estimation. More-promising approaches were presented by Delprat-Jannaud and Lailly (1992) and Kaipio et al. (1999). Delprat-Jannaud and Lailly (1992) incorporated into the objective function a term encouraging the velocity gradient to follow reflector dip. Kaipio et al., (1999) suggested using a priori structural information to create conditional covariance matrices.

In this paper, we take a different approach to adding geologic dip information to velocity estimation. We start from the

Manuscript received by the Editor May 31, 2002; revised manuscript received May 24, 2003.
*Stanford University, Department of Geophysics, Stanford Exploration Project, Mitchell Building, Stanford, California. E-mail: bob@sep.stanford.edu; biondo@sep.stanford.edu; jon@sep.stanford.edu.
© 2004 Society of Exploration Geophysicists. All rights reserved.

same basic assumption that Delprat-Jannaud and Lailly (1992) used, that velocity follows structural dip or some other known trend. Rather than minimizing the velocity gradient directly in the objective function, we approximate a single dip by creating a small plane-wave annihilation filter (Claerbout, 1992a). By adjusting the shape and coefficients within this filter (which we refer to as a *dip-penalty filter*), we can approximate a wide range of covariance responses. By building a space-varying regularization operator out of these filters (a *steering filter*), we can approximate a model covariance that is space-variant. To speed up convergence, we reformulate the regularization problem as a preconditioned problem (Claerbout, 1998a), using the helix transform and polynomial division (Claerbout, 1998b).

To demonstrate the effect of the steering filter, we begin by introducing a simple missing-data problem. We show how the geophysicist and geostatistician find different ways to characterize and incorporate the model covariance function into an inverse problem. We then introduce another way to characterize the model covariance function—dip-penalty filters—that combine the geostatistician's ability to use disparate and irregular information sources and the geophysicist's ability to solve complex inversion problems. We go on to describe how to find dip-penalty-filter directions and how to build a steering filter. We show how to speed up convergence by reformulating problems in terms of preconditioning rather than regularization. We show that a steering filter improves the tomography estimate on a simple-anticline synthetic model. We conclude by applying the technique on a marine North Sea data set.

WHY REGULARIZE?

Many geophysical inverse problems are ill-posed. A classic example of this is the missing-data problem (Claerbout, 1998a; Isaaks and Srivastava, 1989). The goal of the missing-data problem is to interpolate intelligently between a sparse set of known points. For example, let us take a synthetic velocity model with an upper horizontal reflector, an anticline between two unconformities, and an updipping layer at the bottom of the model, as shown in the left panel of Figure 1. Suppose we have velocity measurements at several wells and would like to interpolate them onto a regular 2D mesh (right panel of Figure 1).

The geophysicist might follow the approach described by Claerbout (1998a). First, interpolate the irregular data, **d**, onto a regular mesh by applying some type of binning operator; **B**, then define a fitting goal that requires the model, **m**, to fit the data exactly at the known points, **J**. We can then define an objective function Q as

$$Q(\mathbf{m}) = \|\mathbf{JBd} - \mathbf{Jm}\|^2. \tag{1}$$

In this objective function we are trying to minimize the residual **r**,

$$minimize\{\mathbf{r} = \mathbf{JBd} - \mathbf{Jm}\}. \tag{2}$$

A shorthand form of equation (2) is

$$\mathbf{JBd} \approx \mathbf{Jm}. \tag{3}$$

In equation (4), **0** is an array of zeros. In this shorthand form, the $\approx$ indicates that after minimization, the left-hand side is approximately equal to the right-hand side. Equation (3) is referred to as a *fitting goal* and makes representing complex objective functions more intuitive. We will use this shorthand form throughout the rest of the paper. At model locations where there are no data values, we want the model to be "smooth"; therefore, we will use Tikhonov regularization (Tikhonov and Arsenin, 1997) to minimize the output of a roughening operator **A** applied to the model,

$$\mathbf{0} \approx \mathbf{Am}. \tag{4}$$

In equation (4), **0** is an array of zeros. For the remainder of the paper we will refer to equation (4) as the model styling goal. The issue of determining the appropriate relative weights between the fitting goal and the model styling goal is complex and therefore beyond the scope of this paper.

Having no additional knowledge about the model, we might logically choose an isotropic operator like the Laplacian for **A**. If we apply the fitting goals implied by equations (3) and (4) for 200 iterations using the Laplacian for **A**, we get the model depicted in Figure 2. By spreading information isotropically, the model goes smoothly from the known points to some local average. We see little to no continuation of layers, which is generally an unsatisfactory result.

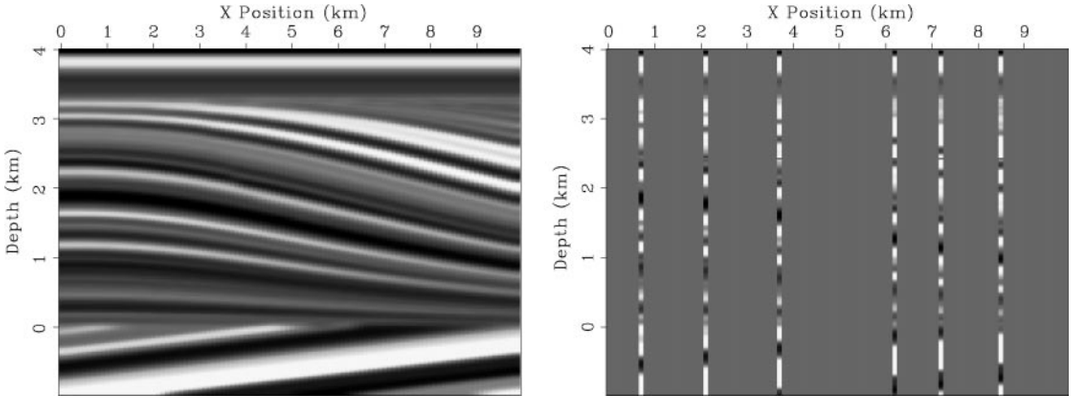

Figure 1. Left panel shows a synthetic velocity model, right panel shows a subset of those data chosen to simulate well-log data.

APPROXIMATING THE COVARIANCE MATRIX

Can we do better than the Laplacian operator? Bayes theory indicates that we should be using the inverse model covariance for the regularization operator (Tarantola, 1987). Unfortunately, we usually have no way to estimate the model covariance matrix. If the model covariance matrix is unreasonable to estimate and use, what is a more reasonable goal?

All statistical measures have an underlying assumption of repetition. Generally, we deal with data from a single experiment, so we have a single value at each location. As a result, common practice is to use nearby points to simulate repetition. By using multiple points, we are making an assumption of stationarity; in other words, we assume that the statistics in a region do not change. Mathematically, stationarity means that the joint distribution of any two points does not depend on their location, just on the vector distance **x** between them.

If we accept the stationarity assumption, there are several related ways that we can characterize the relationship between nearby points. The covariance $C(\mathbf{x})$ is simply the autocorrelation with the mean subtracted,

$$C(\mathbf{x}) = \langle m_i - \hat{m}_i \rangle - \langle m_j - \hat{m}_j \rangle, \tag{5}$$

where $\langle\ \rangle$ is an ensemble average, m_i and m_j are points separated by the vector distance **x**, and $\hat{m}_i$ and $\hat{m}_j$ are the average values of the points.

Figure 3 shows covariance, calculated using equation (5), of the left panel of Figure 1. It is dominated by the dip from beneath the lowest unconformity, but other dips are also seen. In the next subsections, we will show how geophysicists and geostatisticians use these relative measurements to solve the missing-data problem.

Geophysical approach

One method to characterize the model covariance is through a Prediction Error Filter (PEF) (Claerbout, 1998a). The geophysicist notes that if we solve

$$\mathbf{Ma} \approx \mathbf{0}, \tag{6}$$

where **M** is convolution with a field that has the same properties as the model, and **a** is a multidimensional PEF (Figure 4); the output of this convolution is white [Claerbout, 1992a; see Claerbout and Robinson (1964) for discussion of this estimation procedure]. Therefore, **a** must have the inverse spectrum of the model. This is only true if we have chosen a sufficient shape and size for **a**. For example, in a 2D prediction problem, **a** must be of sufficient length that it can describe all of the dips in the model. The width of **a** should correspond to the number of dips in the model. If we construct **a** to find dips at all

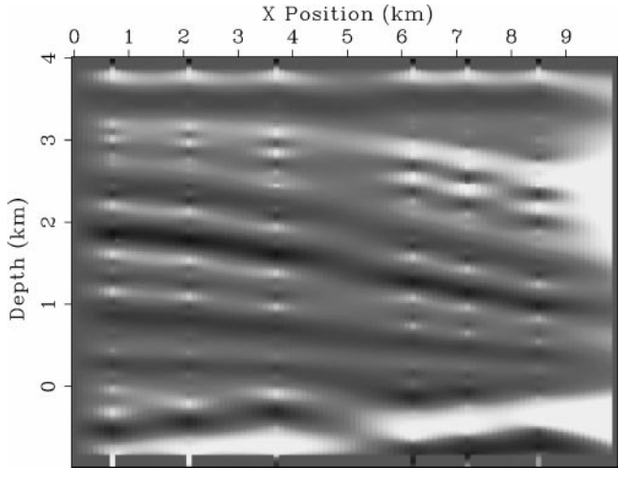

Figure 2. Interpolation result after 200 iterations using an inverse Laplacian regularization operator. Note the edge effects at the top and bottom of the model, resulting from using an internal convolution operator.

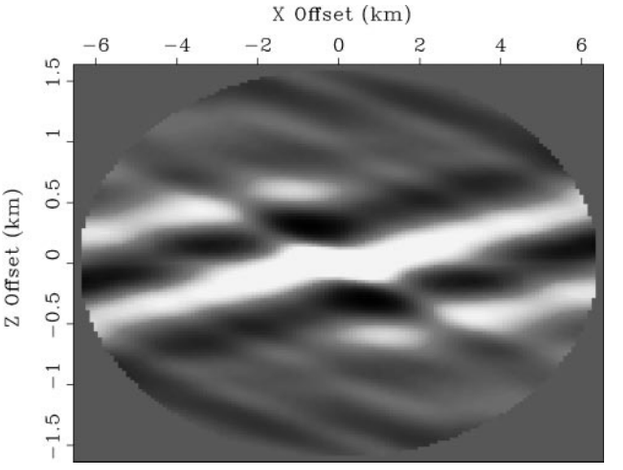

Figure 3. Spatial covariance matrix for the velocity model in Figure 1. Note that the dip below the lower unconformity dominates the covariance calculation.

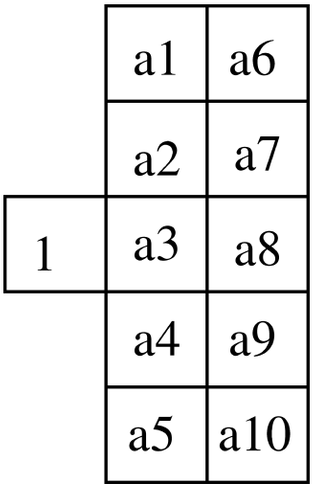

Figure 4. A sample 2D PEF. Coefficients $a1$–$a10$ are estimated by convolving the PEF with a known model. The number of columns in the PEF determine the number of dips that can be estimated; the number of rows determine the range of dips that the filter is sensitive to (Claerbout, 1998a).

possible angles and then apply equation (6), we get the impulse response of the filter (Figure 5).

The PEF captured the prominant dip going up at approximately 15° and also a minor dip going down at approximately 30°. The right panel of Figure 5 shows the result of applying the PEF to the missing-data problem. We have done a substantially better job of filling in the missing data, compared with the Laplacian result (Figure 2), but it is still far from ideal. The filter has introduced both dips at every location, and, as a result, we have a model that is unreasonable.

Geostatistical approach

The geostatistician takes another approach to characterizing the model covariance, called kriging (Deutsch and Journel, 1992). Instead of solving a global optimization problem, the geostatistician solves a series of small inversion problems. He assumes that the model point m is a linear combination of n nearby data points $d_1 \ldots d_n$. The data and model exist in the same vector space. We can define the model's location as $\mathbf{u_0}$ and the data-point locations as $\mathbf{u_1} \ldots \mathbf{u_n}$. The model is a linear combination of the data points,

$$m = \sum_{\alpha=1}^{n} \lambda_\alpha d_\alpha. \tag{7}$$

The weights λ_α are calculated to minimize the error variance and result in the set of equations,

$$\begin{bmatrix} C(\mathbf{u_1}-\mathbf{u_1}) & \ldots & C(\mathbf{u_n}-\mathbf{u_1}) & 1 \\ . & \ldots & . & . \\ . & \ldots & . & . \\ . & \ldots & . & . \\ C(\mathbf{u_1}-\mathbf{u_n}) & \ldots & C(\mathbf{u_n}-\mathbf{u_n}) & 1 \\ 1 & \ldots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ . \\ . \\ . \\ \lambda_n \\ \mu \end{bmatrix}$$

$$= \begin{bmatrix} C(\mathbf{u_0}-\mathbf{u_1}) \\ . \\ . \\ . \\ C(\mathbf{u_0}-\mathbf{u_n}) \\ 1 \end{bmatrix}, \tag{8}$$

where μ is the Lagrange parameter that makes equation (8) an unconstrained minimization problem (Isaaks and Srivastava, 1989). Each $C(\mathbf{x})$, where x is a vector distance, is taken from a predefined covariance estimate. An example of this can be seen in Figure 6. Here, the left side shows four points arranged in space. The right side shows which $\mathbf{C(x)}$ will be used

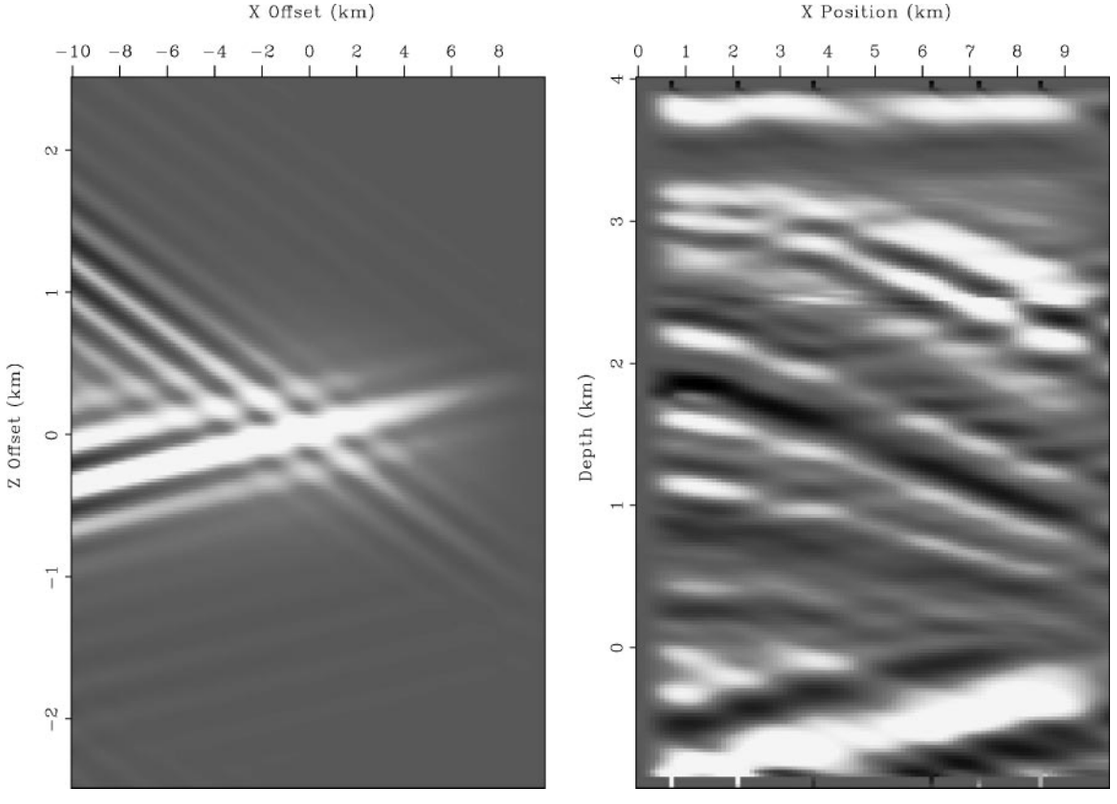

Figure 5. The left panel shows the impulse response of the PEF found from the velocity field (left panel of Figure 1). The right panel shows the result of applying the PEF to the missing-data problem (the input being the right side of Figure 1).

in equation (8). To guarantee that the matrix in equation (8) is positive definite, the geostatistician approximates the covariance function through a linear combination of a limited set of functions. Each function is described by *range*, the distance at which the covariance function has essentially gone to 0, *anisotropy*, the amount the covariance function depends on radial angle, and *orientation*, the major orientation axis of the covariance function. These parameters are used to describe a spherical, exponential, Gaussian, or power model that is guaranteed to produce a positive definite covariance matrix.

The left panel of Figure 7 shows the geostatistical approximation of the model covariance for the missing-data problem. This approximation does a good job of characterizing the primary dip of the covariance function but does not accurately describe the range of the covariance function. The right panel of Figure 7 shows the result of applying kriging to the same missing-data problem. The result is comparable to the geophysical result. Instead of adding a second dip at every location, we have imposed the primary dip. Because this approximation adds insufficient range for the covariance function, the answer returns to the local average between the two wells with the greatest separation. Overall, the result is as unsatisfactory as the geophysical approach to the problem.

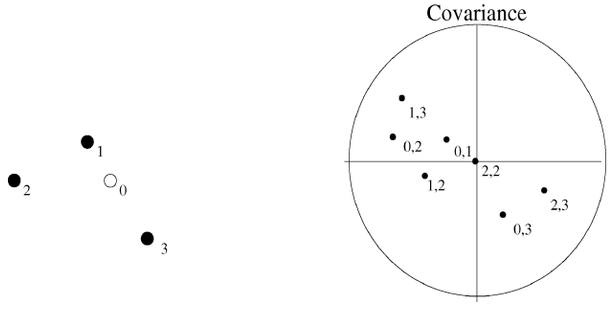

Figure 6. Definition of the terms in equation (8). A vector is drawn between two points. The covariance at the angle and distance describing the vector is then selected.

STATIONARITY

Both the geostatistical and geophysical approaches give poor results because the field we are characterizing is nonstationary, whereas both methods are built on a stationarity assumption. If we look at the covariance of four different regions in the data, as shown in Figure 8, we can see that covariance changes dramatically throughout the model. The top left panel shows the covariance taken from the flat structure at the top of the model and shows a strong horizontal trend. The top right panel represents the covariance along the upper portion of the anticline and has a slight dip down to the right. The bottom left panel is the covariance from the lower portion of the anticline and captures the sharper dip down to the right in the region. The final panel represents the area below the unconformity and captures the updipping structure.

Patching

A common solution to nonstationarity is to break the problem into patches (geophysics terminology) (Claerbout, 1992b) or "distinct subzones" (geostatistics terminology). We define regions where the assumption of stationarity is valid and apply the given technique in the region. We then recombine the subregions into the final model. Figure 9 shows the result of dividing the model space into the four different regions of Figure 8 and then applying the geophysical and geostatistical approaches. Note that the image is significantly improved over the single-covariance-function approach. The range description for the geostatistical approach is poor, and the images are still of significantly lower frequency than in the known model.

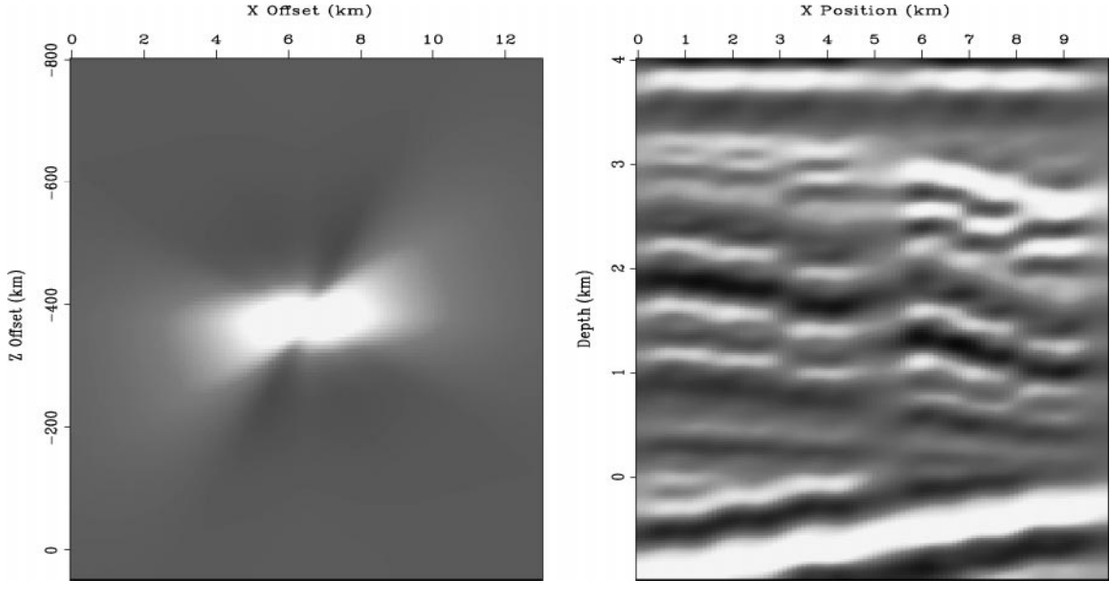

Figure 7. Left panel shows the approximated covariance matrix; right panel is the result of solving equation (8) at every unknown model point.

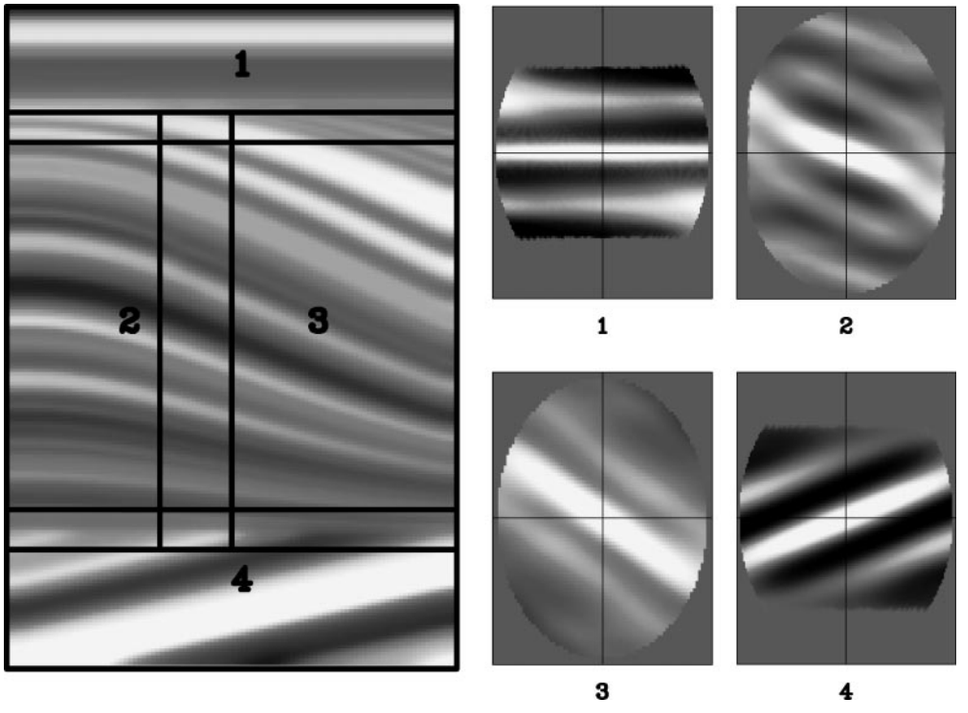

Figure 8. The covariance at four different regions of the model (left panel of Figure 1). The top left is above the upper unconformity; top right, the upper portion of the anticline; bottom left, lower portion of the anticline; and bottom right, below the lower anticline. In each covariance display, a line is drawn through zero offset.

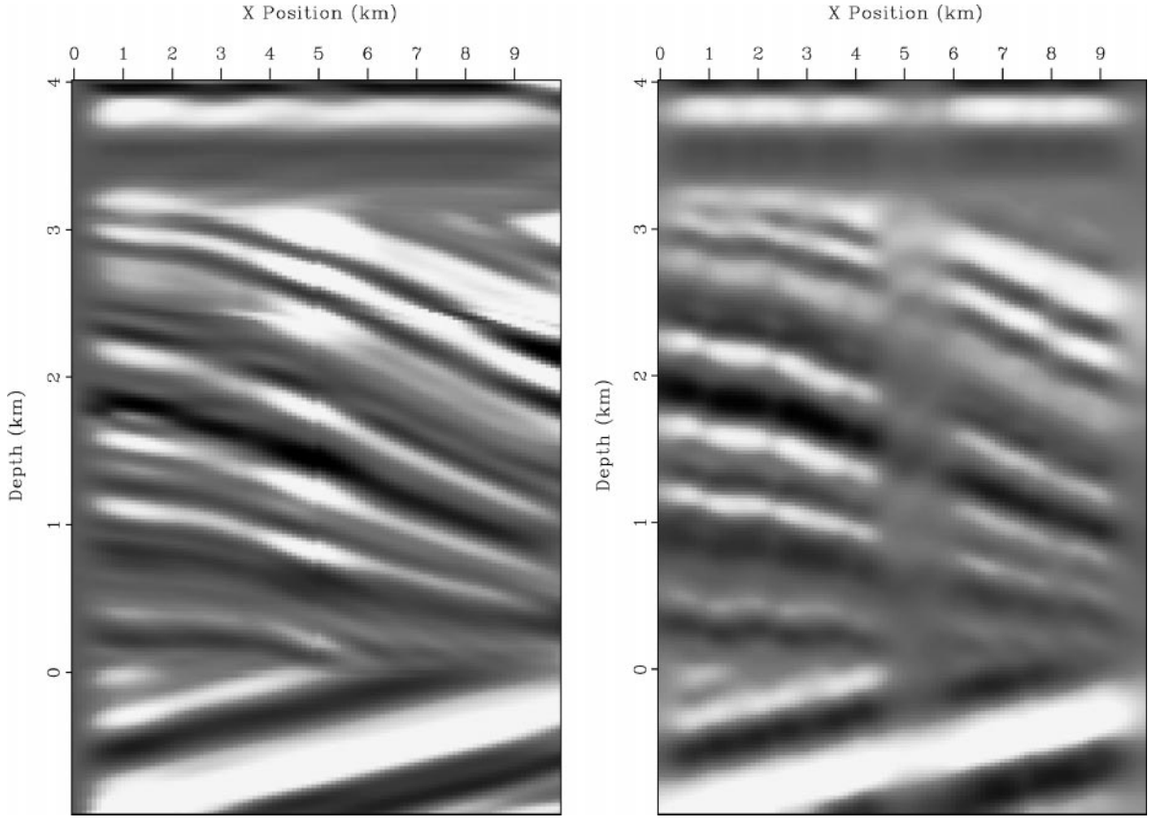

Figure 9. The result of breaking the problem into four patches and solving them independently. Left panel is the result of applying the geophysical method, the right panel is the result of applying geostatistics.

STEERING FILTER

It is apparent that the covariance function varies within at least two of the four patches (we can also see this in the covariance functions of patches 2 and 4 of Figure 8). Therefore, it follows that we should get a better image by making smaller and smaller patches. Crawley (1998) showed that this is true when one is solving a data-interpolation problem. Unfortunately, a patch size large enough to generate sufficient statistics to find the PEF or solve the kriging equations may still contain different covariance functions. In addition, the geophysical solution relies on having a field with the same statistical properties as does the model, which is often not the case. Often, what we have is what geostatisticians refer to as "soft data." Soft data have generally the same properties as does the variable of interest, but they are often in incompatible form. A classic example of this is tying well measurements to flow simulation results. Geostatistics is generally better suited for combining mixed and limited data. Kriging requires us to provide only a covariance function for each subzone/region. On the other hand, the geostatistical approach is not well suited for fitting into an iterative optimization problem. Kriging wants the physics of the problem to be describable by a known function that would then form the basis for a space-varying mean for the kriging problem (Isaaks and Srivastava, 1989). For problems like tomography, this is an unacceptable requirement.

In geophysics, attempts to combine different information sources (Stork, 1994) have met with limited success. Generally, the inversion problems are too large to use classical hard-constraint mechanisms (Polak, 1997). We need a method to construct a space-varying filter that does a good job of describing the model covariance but that can be obtained from limited and disparate information sources.

If we look at regions 1 and 4 in Figure 8, we can see that when the stationary assumption is valid, the covariance matrix is fairly simple. We have (1) a primary trend oriented along the dip of the velocity field that slowly dies out and (2) a ringing effect resulting from the sinusoidal nature of the model. We would like to come up with a way to emulate the primary trend of the covariance matrix through minimal information.

To do this, it is important to remember that the regularization operator should have the inverse spectrum of the covariance matrix. Therefore, if the covariance function is primarily a dipping event, the regularization operator should be destroying that dip. Claerbout (1990, 1992a) showed how to estimate the primary dip in a region and how to construct a filter that could destroy that dip. These small filters, which we refer to as dip-penalty filters, can be as simple as a two- or three-point filter, as in Figure 10. A dip-penalty filter consists of a fixed "1" and one or more coefficients in the next column. The location of the filter coefficients in the second column determines the dip that the filter will destroy. Figure 11 is the inverse impulse response of Figure 10. Note how the general orientation of the impulse response is approximately the same as the covariance function below the lower unconformity, but the anisotropy and range are not correct. In the next section, we will discuss how we can also control these parameters with dip-penalty filters.

Constructing a filter

When building a steering filter, we want to create a filter that destroys a given slope p. Further, we would like to control the bandwidth response of filters oriented at different slopes. We can achieve both of these goals by constructing a simple filter. The filter will have a 1 at the zero-lag location and the rest of its values one column away. For determining the non-zero-lag values, imagine constructing a triangle with an apex located on the zero-lag location. The center of the triangle will be located so that a line connecting it to the zero-lag apex has the desired slope p. The width w of the triangle determines the size of the filter (only coefficients within the triangle will be used) and the filter's level of anisotropy. The height of the triangle a determines the range over which the filter will operate.

For example, let's return to the missing-data problem. When we limit the a priori information to the assumption that velocity follows structure, and we have some guesses at reflector position, we can use this information to build a complex operator (which we refer to as a steering filter) composed of dip-penalty filters. For this problem, we will assume that we have the location of four reflectors: one above the top unconformity, two between the unconformities, and one below the lower unconformity (left panel of Figure 12). If we interpolate these dips to the entire model space, we have all we need to construct a steering-filter operator. If we use this operator as the regularizer, we get Figure 13 as the interpolation result. The steering filter did a significantly better job than did the patching approach. With more information about the model (more reflectors, an approximation of the level of anisotropy in different portions of the model, etc.), we could do even better by

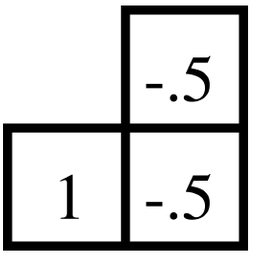

Figure 10. A dip-penalty filter designed to annihilate dips of approximately 22.5°.

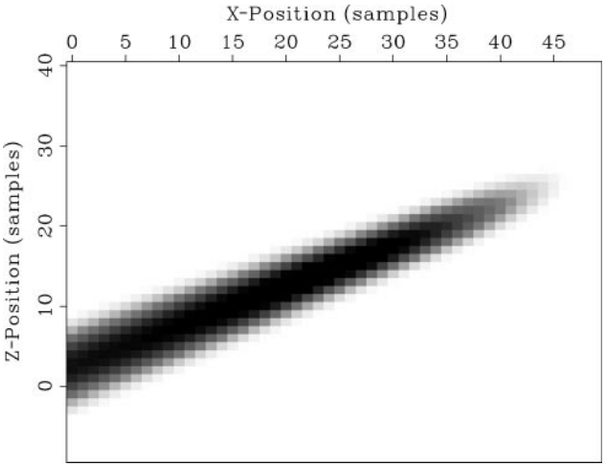

Figure 11. The result of applying $\frac{1}{\mathbf{A}\mathbf{A}'}$, where $\mathbf{A}$ is the filter in Figure 10.

using some of the other adjustable parameters available when constructing the steering-filter operator.

The three adjustable parameters in the filter construction—possibly different at every model point—can seem daunting, but they are what enables almost any covariance function to be described by steering filters. In certain regions of the model, one might feel that the covariance function is much more isotropic. In these regions, one could consider making the triangle bigger to smooth the filter coefficients over a wider angle range, while keeping it small in areas where the covariance is much more anisotropic.

REGULARIZING VS. PRECONDITIONING

An important consideration in many geophysical problems, including tomography, is speed of convergence. Tomography, even when it is ray based, is computationally intensive, so we should minimize the number of steps it takes to get to a reasonable solution. One reason for slow convergence is that the regularization operators, including steering filters, only respond to low frequencies. As a result, the condition number of the matrix we are attempting to invert is large. A classic solution is to reformulate the regularized inversion problem into a preconditioned problem in terms of some new variable **p** (Polak, 1997) as used in tomography by Harlan (1995). The goal is to replace the regularization operator **A** with a preconditioning operator **B** that smooths long distances with each iteration. Any smoother could be an effective preconditioning operator, but the ideal choice for **B** would be $\mathbf{A}^{-1}$, because if $\mathbf{B} = \mathbf{A}^{-1}$, the shorthand regularized fitting goals (3) and (4) would be equivalent to the new preconditioned fitting goals (Fomel et al., 1997;

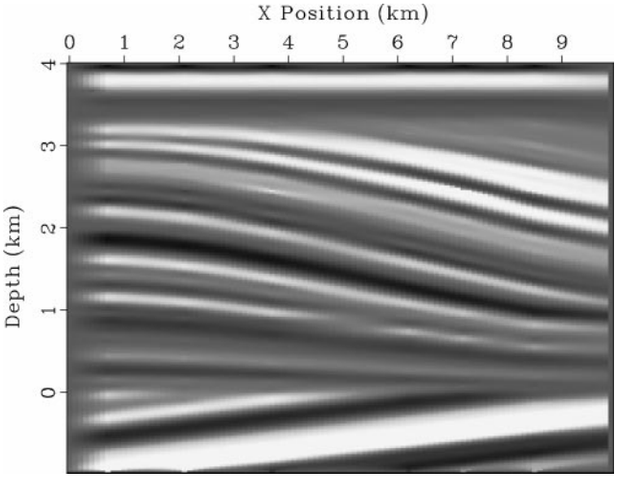

Figure 13. The result of using the steering filter operator as regularizer to the missing-data problem.

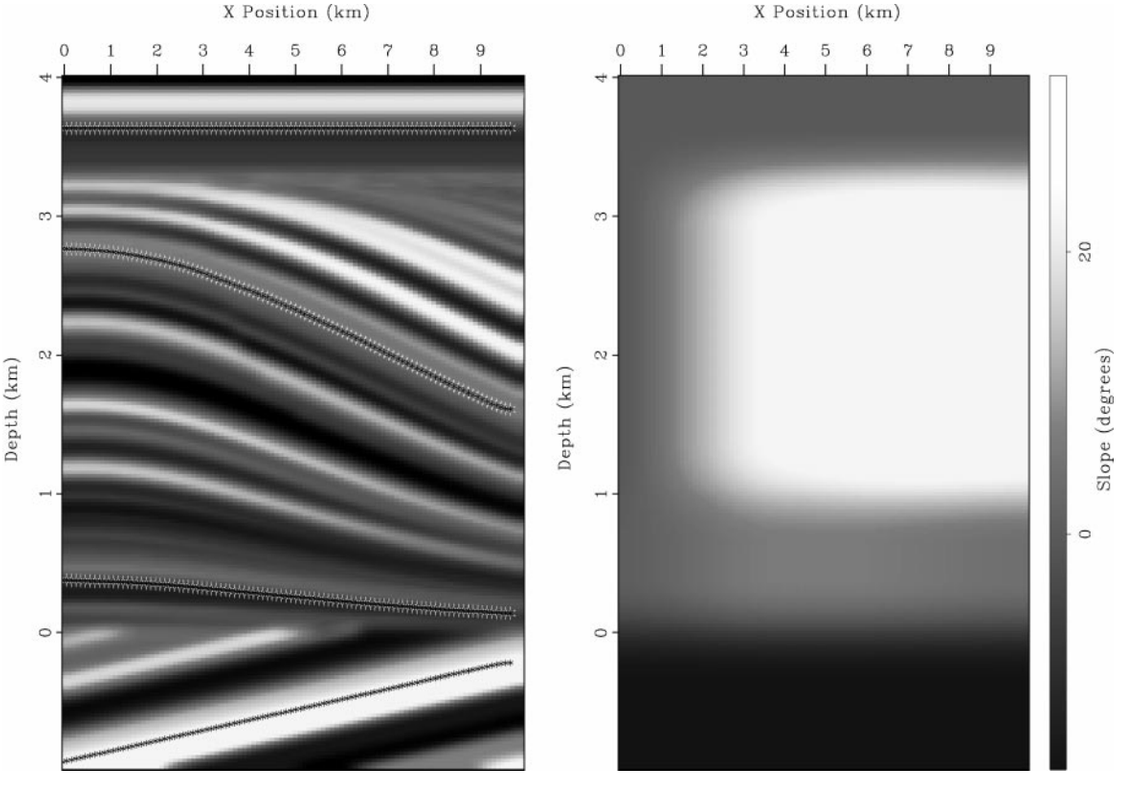

Figure 12. Left panel are four reflectors chosen to represent the a priori information. The right panel is the interpolated slope calculated from the reflectors that will form the basis of the dip-penalty filter.

Fomel, 2001), where **L** is a linear operator,

$$\mathbf{d} \approx \mathbf{L}\mathbf{A}^{-1}\mathbf{p}$$

$$\mathbf{0} \approx \mathbf{A}\mathbf{m} = \mathbf{A}\mathbf{A}^{-1}\mathbf{p} = \mathbf{I}\mathbf{p}. \quad (9)$$

Recall that in this shorthand form, the $\approx$ means that after minimization, the left-hand side is approximately equal to the right-hand side. The speed up is the result not only of the preconditioner spreading information long distances with every iteration but also of the regularization operator now being the identity matrix, **I**. The identity matrix is its own inverse, thereby reducing the condition number of the matrix we are trying to invert.

Helix transform

If our regularization operator were 1D, applying the inverse of the filter could easily be done through polynomial division (Claerbout, 1976). If the filter is greater than 1D, there is no general method to apply its inverse. The problem is that our regularization operator, the steering filter, is multidimensional. We can apply the inverse of the operator by taking advantage of the helix transform (Claerbout, 1998b). The general approach is to transform the problem into 1D space. Given a multidimensional filter, such as Figure 10, first map it onto the coordinate space of the data (panel *a* of Figure 14). Then imagine wrapping the data around a cylinder, with the end of column one connected to the beginning of column two. Finally, unwind the data into a single string of numbers, and the multidimensional filter is converted into a 1D filter. If this new, one-dimensional filter is causal and minimum phase, which the steering filter is, we can apply polynomial division (Claerbout, 1976).

With the helix, we can now solve the missing-data problem that took us 40 iterations as a regularized problem, as in Figure 13. Now we solve it in six iterations, at the same cost per iteration and with the same quality of result (Figure 15).

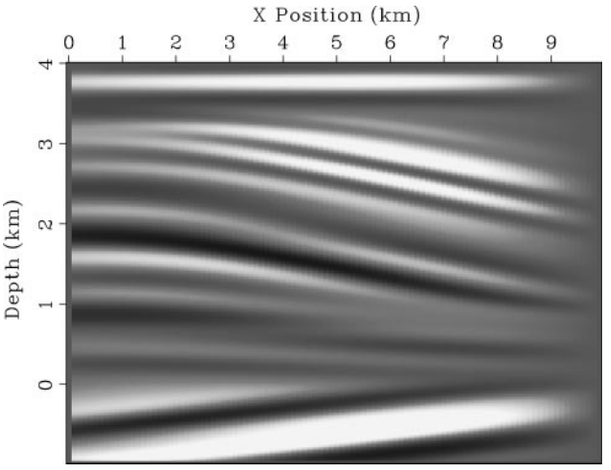

Figure 15. The interpolation result after six iterations using the preconditioned formulation (10) of the problem.

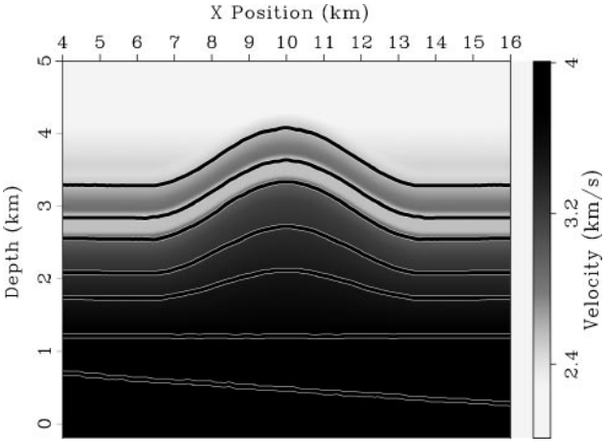

Figure 16. The correct velocity model with the reflector positions overlayed.

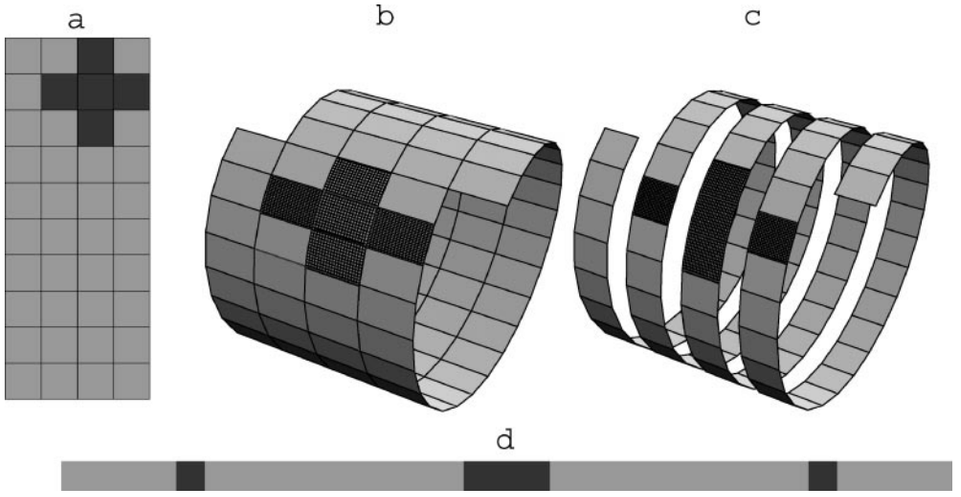

Figure 14. Filtering on a helix. The same filter coefficients overlie the same data values if the 2D coils are unwound into 1D strips. Figure courtesy of Sergey Fomel.

REGULARIZING TOMOGRAPHY

To test the advantage of incorporating non-stationary-dip information into the tomography problem, we constructed a synthetic model (Figure 16). The model is an anticline with seven reflectors. The top five reflectors are within the anticline structure. The last two, a flat reflector and a dipping reflector, are below the anticline. The imaging challenge is to position the bottom reflectors correctly. For added difficulty there is a low-velocity layer between the second and third reflectors. The model was used to do acoustic wave modeling, with the resulting data set having 40-m common-midpoint (CMP) spacing and 60 offsets spaced 80 m apart. If the initial estimation of the slowness is the $s(z)$ function from outside the anticline, the migrated reflector positions are pulled up because of our using too low a velocity within the anticline.

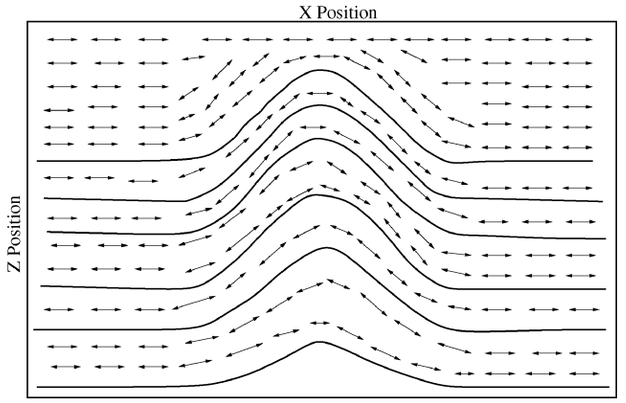

Figure 17. Steering filter directions as a function of geologic dip. The bold, solid lines are estimated reflector positions (taken from the initial migration); the remaining lines represent dip bars estimated by interpolating the dips between the reflectors.

Finding smoothing directions

To build the steering-filter operator, we need a dip field, the range, and the level of anisotropy throughout the model. In many cases, it is important to have all three of these parameters vary spatially. For this simple example, we can assign a constant range and anisotropy to the entire model. For the dip field estimate, we assume that velocity follows dip and use the position in which the reflectors image at zero offset. Figure 17 shows the dip directions based on the migrated reflector positions using the initial velocity estimate.

Fitting goals

For the tomography problem, we begin by linearizing around an initial slowness model $\mathbf{s_0}$. We then construct a ray-based tomography operator $\mathbf{T}$ (Stork and Clayton, 1991; Stork, 1992) that relates changes in migrated reflector position (and through local velocity changes in traveltime $\Delta\mathbf{t}$) to slowness perturbations $\Delta\mathbf{s}$.

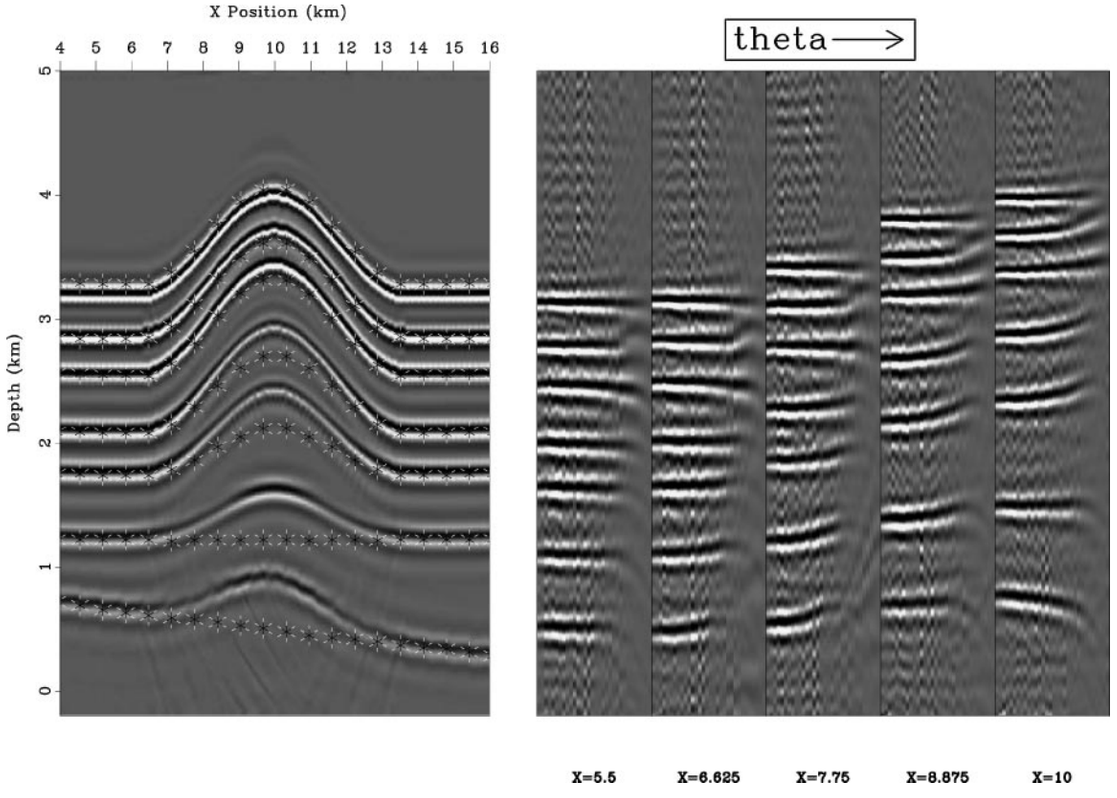

Figure 18. The initial migration result using an $s(z)$ velocity function from the edge of the model. Right are five (z, θ) planes at different x positions. Note that the CRP gather at 5.5 is flat, whereas the other four CRP gathers show significant moveout.

Following the same procedure that we used for the missing-data problem, we add the regularization goal to the tomography problem to get our shorthand fitting goals:

$$\begin{aligned} \Delta\mathbf{t} &\approx \mathbf{T}\Delta\mathbf{s} \\ \mathbf{0} &\approx \epsilon\mathbf{A}\Delta\mathbf{s}, \end{aligned} \tag{10}$$

where $\mathbf{A}$ is the steering filter operator and ϵ is a scalar parameter controlling the importance of the smoothing criteria.

However, these fitting goals don't accurately describe what we really want. The steering filters are based on the desired slowness $\mathbf{s}$, rather than on the change of slowness $\Delta\mathbf{s}$. With this fact in mind, we rewrite the model styling goal as

$$\mathbf{0} \approx \epsilon\mathbf{A}(\mathbf{s}_0 + \Delta\mathbf{s}). \tag{11}$$

Here we see the advantage of the helix. Without having the exact inverse of $\mathbf{A}$ as our preconditioner, we would not be able to do the same substitution. The left-hand side of the model styling goal is not equal to zero, but we can still do the same preconditioning substitution:

$$\begin{aligned} \Delta\mathbf{t} &\approx \mathbf{T}\mathbf{A}^{-1}\mathbf{p} \\ -\epsilon\mathbf{A}\mathbf{s}_0 &\approx \epsilon\mathbf{I}\mathbf{p}. \end{aligned} \tag{12}$$

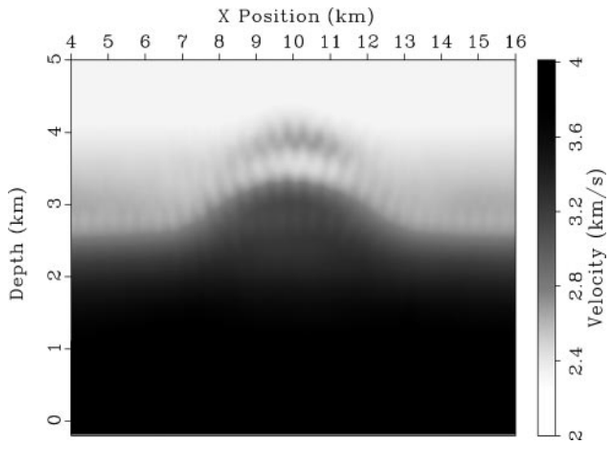

Figure 19. The velocity model using an isotropic regularization operator.

SYNTHETIC COMPARISONS

Now that we have set up the tomography problem, it is time to test the theory and compare the advantage of using an anisotropic steering filter versus using isotropic regularization. For this test case, we will use a wave-equation method to migrate the data and construct angle gathers.

Figure 18 is the result of migrating with an $s(z)$ slowness function (varying only as a function of depth) extracted from $x = 4$. The correct reflector positions are indicated with asterisks (*). As we expected, away from the anticline, the gathers are flat and we have correctly positioned the reflectors. Along the edge of the anticline, we see upward curvature in the CRP gathers, indicating that we have used too slow a velocity at this location. Below the center of the anticline, $x = 10$, the bottom reflector shows some reverse moveout, the well-known W pattern.

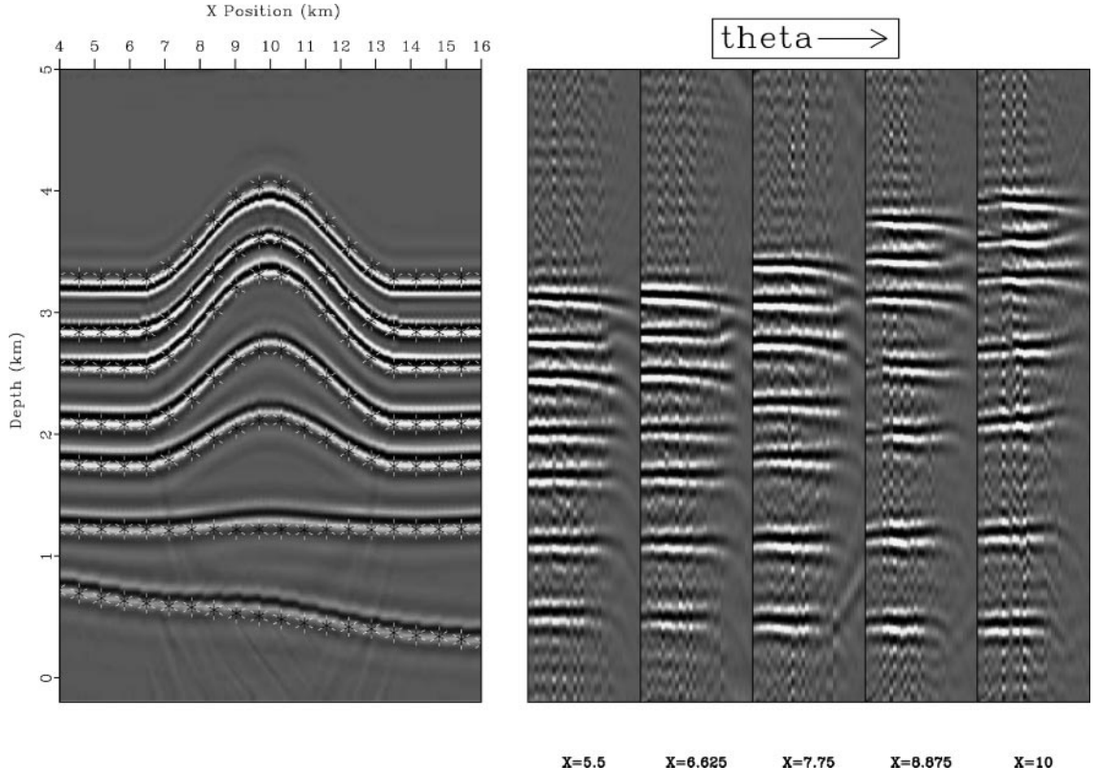

Figure 20. The migration result using the velocity of Figure 19.

For comparison, we performed one nonlinear iteration of tomography using different operators for **A** in equation (12). In the first case, we use an isotropic regularizer for **A**. Figure 19 shows the resulting velocity model. We have begun to recover the anticline shape but have significantly smeared out the low-velocity zone. If we migrate with this new velocity, as shown in Figure 20, we see that we have significantly improved the image. The CRP gathers are flatter but still show moveout errors. The reflectors are better positioned, but we still have some pull-up.

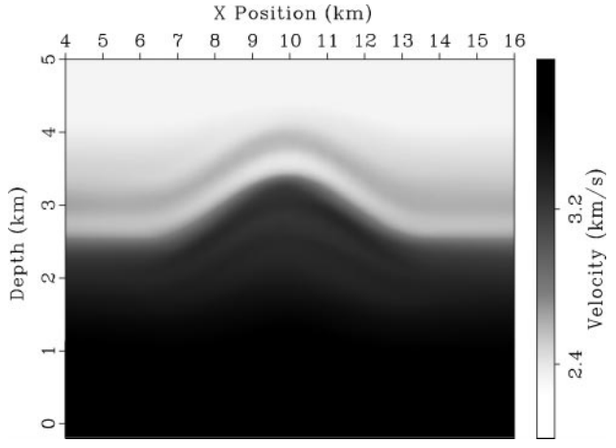

Figure 21. The velocity model using steering-filter regularization.

We get a substantially improved result using a steering-filter operator constructed from our initial migration positions. Figure 21 shows the resulting velocity model. Note how the low-velocity layer is much better defined in Figure 21 than it was in Figure 19. By migrating with this velocity, we get an excellent image, seen in Figure 22. The events are generally better focused, we have little moveout in the CRP gathers, and there is almost perfect placement of the reflectors.

FIELD DATA

To test the methodology on field data, we chose a data set over a salt dome in the North Sea. The data are very clean, with strong reflectors that are generally continuous. The data contain a chalk layer that causes a velocity inversion below it.

The data set is 3D marine, acquired using four cables with geophones every 25 m. In this paper, we will be dealing with a 2D subset of the 3D data set. The subset was created by partial stacking and then applying azimuth moveout (AMO) (Biondi et al., 1998) to the CMP gathers.

For our initial velocity, we started with Figure 23. Figure 24 is the result of migrating the data with the velocity of Figure 23. The initial migration shows little reflector coherency as we approach the left edge of the salt (A). Very little coherent energy is visible directly below the strong reflection at 1.7 km (B). To the right of the salt, we lose reflector coherency below 2.8 km (C). In addition, the left salt boundary (D) is nearly nonexistent.

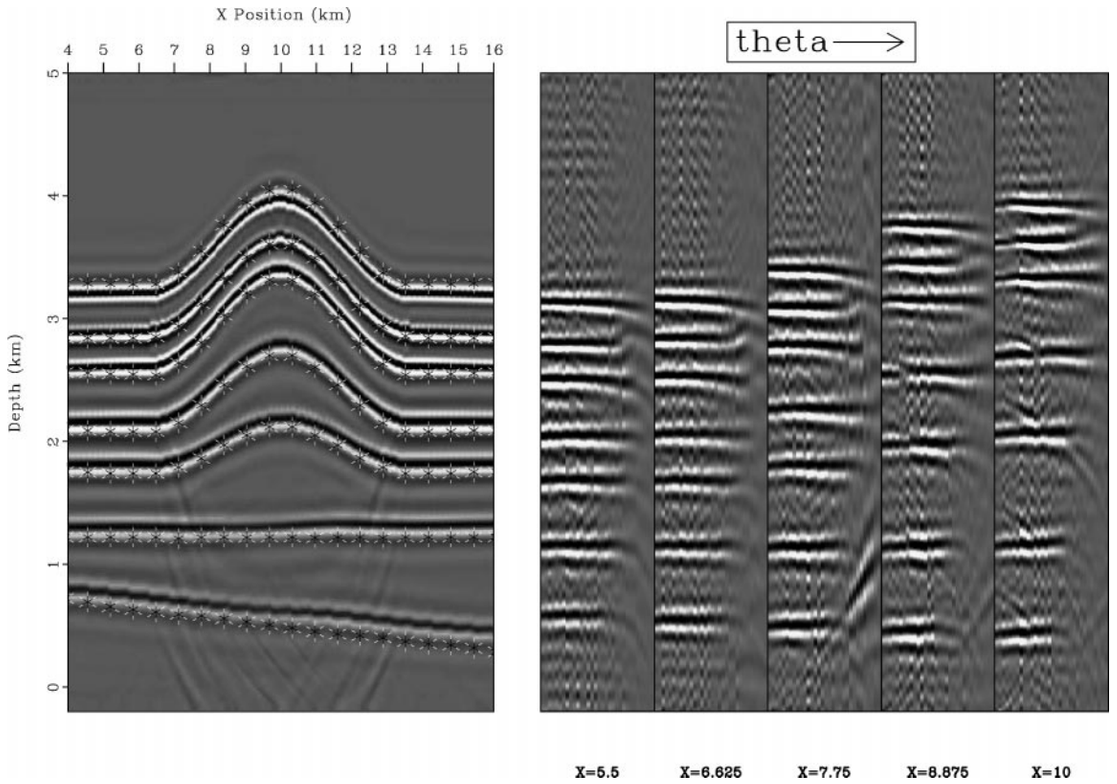

Figure 22. The migration result after one iteration of tomography using a steering filter as the regularization operator.

Using the initial migrated image, we chose 11 reflectors with which to perform tomography (Figure 25). To constrain the upper portion of the model, we chose the water-bottom reflection and two reflectors above the salt. We also picked the salt top and salt bottom and three reflectors on both sides of the salt body.

We performed two nonlinear iterations of tomography holding the salt velocity constant. We used a steering filter as our regularization operator and obtained the velocity in Figure 26. Note how the velocity generally follows structure. The velocity contours follow the top of the salt structure and the basin structure to the right of the salt. On the left side of the salt, velocity is not as consistent with the structural model.

The migration (Figure 27) using the new velocity (Figure 26) has also improved. Whereas little coherent energy was seen along the left the edge of the salt body (A) in the initial migration, we now see nice, continuous reflections to the salt edge. At B, the focusing of events has significantly improved. We also see coherent events to the right of the salt (C). The left salt edge (D), which was absent in the initial migration, is also visible.

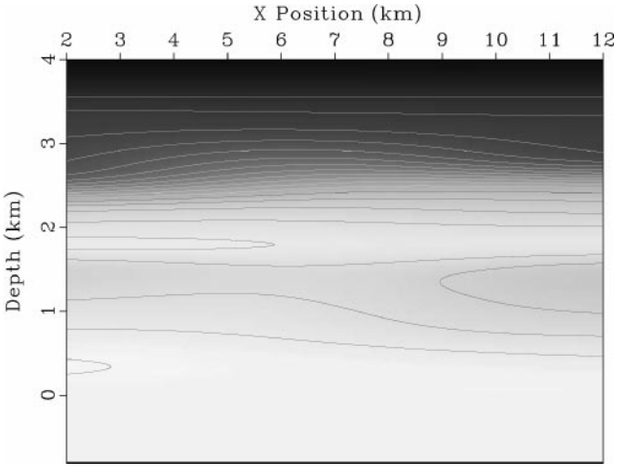

Figure 23. The initial velocity model for the 2D line.

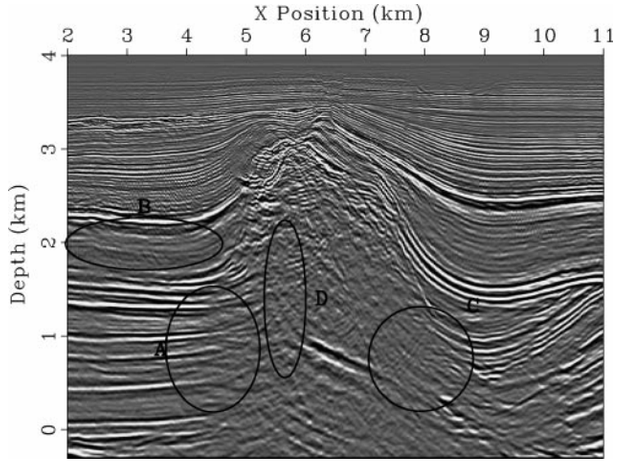

Figure 24. Migration result using the velocity from Figure 23.

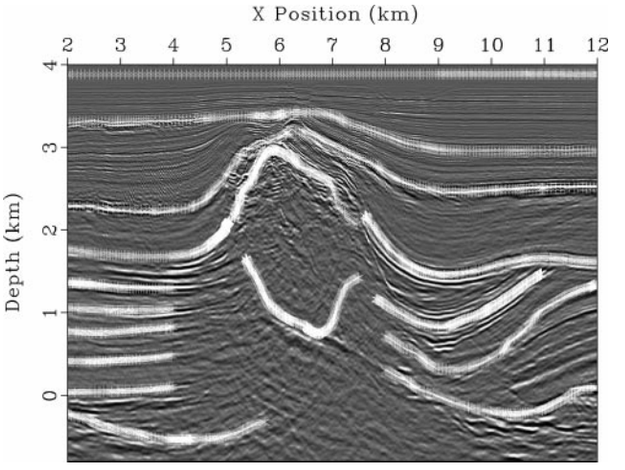

Figure 25. Initial migration with picked reflectors overlaid.

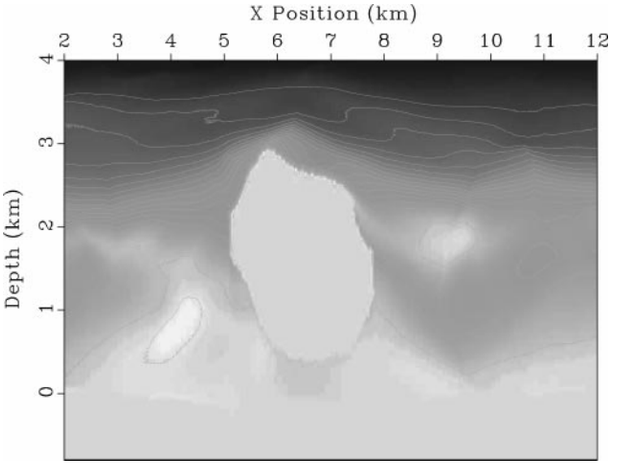

Figure 26. Final velocity. Note how the velocity generally follows structure. The velocity contours follow the top of the salt structure and the basin structure to the right of the salt. On the left side, the salt velocity is not as consistent with the structural model.

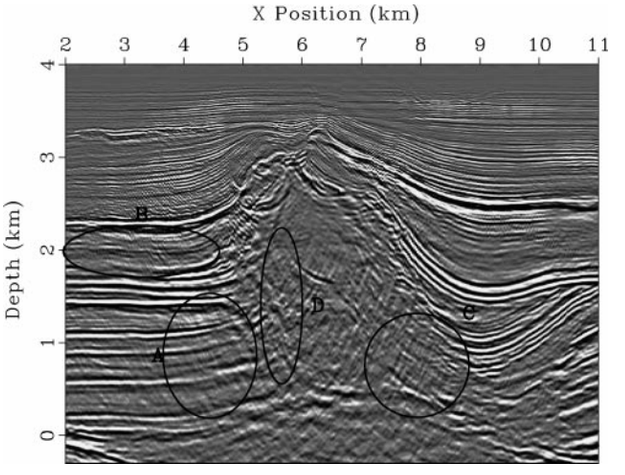

Figure 27. Final migrated image. Locations A–D show improvement, compared with Figure 24.

CONCLUSIONS

In this paper, we demonstrated that a poor choice of regularization operator can degrade inversion estimates. We showed that a standard regularization choice, a stationary isotropic operator, produced unrealistic features in our velocity model for a synthetic and real tomography example. We overcame that problem by using readily available a priori information—reflector dip estimates based on early migration results—to construct a nonstationary anisotropic regularization operator. By using this new regularization operator, we produced velocity models that were more geologically reasonable and produced better images. In the synthetic example, the image produced from the velocity model using the nonstationary operator better positioned the reflectors and flattened the gathers, compared with the model obtained using the isotropic regularization operator. In the field-data example, the image using the velocity model estimated using the nonstationary operator showed significantly better focusing and more reflector continuity.

ACKNOWLEDGMENTS

We thank TotalFinaElf for providing the field data set and the sponsors of Stanford Exploration Project for financial support for this work.

REFERENCES

Biondi, B., Fomel, S., and Chemingui, N., 1998, Azimuth moveout for 3-D prestack imaging: Geophysics, **63**, no. 2, 574–588.
Carrion, P., 1991, Dual tomography for imaging complex structures: Geophysics, **56**, 1395–1404.
Claerbout, J., 1976, Fundamentals of geophysical data processing: http:/sepwww.stanford.edu/sep/prof/.
Claerbout, J. F., 1990, Nonlinear problems: Stanford Exploration Project **SEP-65**, 229–240.
Claerbout, J. F., 1992a, Earth soundings analysis: Processing versus inversion: Blackwell Scientific Publications.
Claerbout, J. F., 1992b, Nonstationarity and conjugacy: Utilities for data patch work: Stanford Exploration Project **SEP-73**, 391–400.
Claerbout, J., 1998a, Geophysical estimation by example: Environmental soundings image enhancement: http://sepwww.stanford.edu/sep/prof/.
Claerbout, J. F., 1998b, Multi-dimensional recursive filtering via the helix: Geophysics, **63**, no. 5, 1532–1541.
Claerbout, J. F., and Robinson, E. A., 1964, The error in least-squares inverse filtering (short note): The error in least-squares inverse filtering (short note): Geophysics, **29**, 118–120.
Crawley, S., 1998, Shot interpolation for Radon multiple suppression: 68th Annual International Meeting, SEG, Expanded Abstracts, 1238–1241.
Delprat-Jannaud, F., and Lailly, P., 1992, What information on the earth model do reflection travel time provide?: Journal of Geophysical Research, **97**, 19827–19844.
Deutsch, C. V., and Journel, A., 1992, GSLIB: Geostatistical software library and user's guide: Oxford University Press.
Fomel, S., 2001, Three-dimensional seismic data regularization: Ph.D. thesis, Stanford University.
Fomel, S., Clapp, R., and Claerbout, J., 1997, Missing data interpolation by recursive filter preconditioning: Stanford Exploration Project **SEP-95**, 15–25.
Harlan, W. S., 1995, Regularization by model reparameterization: http:www.billharlan.com/pub/papers/regularization.pdf.
Hirsche, K., Porter-Hirsche, J., Mewhort, L., and Davis, R., 1997, The use and abuse of geostatistics: The Leading Edge, **16**, no. 3, 253–260.
Isaaks, E. H., and Srivastava, R. M., 1989, An Introduction to applied geostatistics: Oxford University Press.
Ji, J., 1997, Tomographic velocity estimation with plane-wave synthesis: Geophysics, **62**, no. 6, 1825–1838.
Kaipio, J. P., Kolehmainen, V., Vauhkonen, M., and Somersalo, E., 1999, Inverse problems with structural prior information: Inverse Problems, **15**, 713–729.
Kosloff, D., Sherwood, J., Koren, Z., Machet, E., and Falkovitz, Y., 1996, Velocity and interface depth determination by tomography of depth migrated gathers: Geophysics, **61**, no. 5, 1511–1523.
Michelena, R. J., and Harris, J. M., 1991, Tomographic traveltime inversion using natural pixels: Geophysics, **56**, no. 5, 635–653.
Polak, E., 1997, Optimization: Algorithms and consistent approximations: Springer-Verlag.
Rowbotham, P. S., and Pratt, R. G., 1997, Improved inversion through use of the null space: Geophysics, **62**, no. 3, 869–883.
Stork, C., 1992, Reflection tomography in the postmigrated domain: Geophysics, **57**, no. 5, 680–692.
Stork, C., 1994, Demonstration of mva tomography with controls and constraints for determining an accurate velocity model for prestack depth migration: 64th Annual International Meeting, SEG, Expanded Abstracts, 1338–1342.
Stork, C., and Clayton, R. W., 1991, An implementation of tomographic velocity analysis: Geophysics, **56**, no. 4, 483–495.
Stork, C., and Clayton, R. W., 1992, Using constraints to address the instabilities of automated prestack velocity analysis: Using constraints to address the instabilities of automated prestack velocity analysis: Geophysics, **57**, 404–419.
Tarantola, A., 1987, Inverse problem theory: Elsevier.
Tikhonov, A. N., and Arsenin, V. Y., 1977, Solution of ill-posed problems: John Wiley and Sons.
van Trier, J., 1990, Tomographic determination of structural velocities from depth migrated seismic data: Ph.D. thesis, Stanford University.