Interval velocity estimation from beam-stacked data — 2-D gradient operator

Biondo Biondi

ABSTRACT

The theory that I described in SEP–57 (Biondi, 1988) can be applied to compute the gradient of energy in beam-stacks with respect to the velocity model when the model is two-dimensional (2-D). The computation of the gradient uses a linear operator that relates perturbations in the velocity model to resulting changes of beam-stack's parameters. To evaluate this operator I developed a new ray-tracing algorithm that computes the gradient of raypaths and traveltimes with respect to the velocity model. The resulting back-projection operator is spatially localized around the down-going and up-going rays. The operator should resolve well local velocity anomalies. The 2-D inversion is untested, but the 1-D method successfully reconstructed the velocity above a dipping reflector.

INTRODUCTION

Various methods for estimating velocity in geologically complex areas have been proposed. A leading approach uses a tomographic fitting of the velocity model to traveltimes of reflections. The traveltimes can be directly picked from the data (Bishop et al., 1985) or picked from the data transformed by use of local slant stacks (Sword, 1987). An alternative approach, which avoids the picking of traveltimes, maximizes energy in stacking-velocity domain (Toldi, 1985) or in migration-velocity domain (Fowler, 1988). The method that I described in a previous report (Biondi, 1988) combines some elements of Sword's method with Toldi's idea of maximizing energy in a transformed domain. I use energy in beam-stacked data as an objective function to be maximized by the best choice of the velocity model. A beam stack is a linear transformation on the data similar to a local slant stack; it differs from a slant stack in that its stacking-trajectory is hyperbolic instead of linear. (Kostov and Biondi, 1987).

In the previous paper I derived the general theory but I tested it only in the particular case of a horizontally-layered medium. The implementation of the method in the general two-dimensional (2-D) case requires a sophisticated ray-tracing algorithm; I present this algorithm in the Appendix. Using the new ray-tracing algorithm I evaluate the linear operator that relates perturbations of the velocity model to the resulting changes of beam-stack's parameters. The transpose of this operator is the back-projection operator that is used to compute the gradient of the objective function with respect to the velocity model. The properties of the back-projection operator determine the resolving power of the velocity estimation.

In the next section I review the theory of my estimation method and derive the formula I use to evaluate the gradient operator. In the section on the back-projection operator I display the operator and show that in a special case my operator is similar to Toldi's and Fowler's operators. In conclusion I present some preliminary inversion results from synthetic data.

INTERVAL VELOCITY ESTIMATION METHOD

In this section I review the velocity estimation method that I presented in my last report (Biondi, 1988).

The estimation starts from beam-stacked data. Beam stacks are functions of five variables: midpoint y, half-offset h, midpoint ray parameter p_y, offset ray parameter p_h and traveltime t. The amplitude of $Beam(y, h, t, p_y, p_h)$ is proportional to the energy of the reflected waves recorded with observed horizontal ray parameters p_y and p_h, at midpoint y, offset h, and traveltime t. The amplitude of the beam stack can be simply the square of the sum of the amplitudes along the stacking trajectory or the value of a coherency function, such as semblance, computed on the stacking trajectory.

Estimating the interval velocity in the original beam stack domain can be problematic (Biondi, 1987). Therefore in my last report I introduced the following transformations of coordinates;

$$\tau = t - p_h h, \tag{1a}$$

$$\xi = h - \frac{p_h V_0^2}{4} t, \tag{1b}$$

where V_0 is a constant average velocity. These transformations do not have a specific physical meaning, but they are convenient for the inversion.

Defining an objective function

The velocity estimation is a tomographic fitting to beam-stacked data of traveltimes and surface locations predicted by the velocity model. The transformed offset ξ, as a function of transformed traveltime τ and midpoint y, is computed by use of

ray tracing for each reflector point R_j and for all ray parameters p_h and p_y. The function $\xi = \xi(y, \tau; \mathbf{m})$ defines a manifold in the data space. The goal of inversion is to maximize the energy in beam-stacked data on the manifold $\xi = \xi(y, \tau; \mathbf{m})$. The inversion is formulated as the solution of the non-quadratic optimization problem of finding the maximum with respect to the slowness model of the total energy

$$E(\mathbf{m}) = \sum_{p_y} \sum_{p_h} \sum_{j} \overline{Beam}(y_j, \xi(y_j, \tau_j; \mathbf{m}), \tau_j, p_y, p_h), \tag{2}$$

or in more compact notation

$$E(\mathbf{m}) = \sum_{i} B_i(\xi_i(\mathbf{m})), \tag{3}$$

where i is the index of the data points used, including all reflector locations R_j and all ray parameters p_y and p_h.

Computing the gradient of the objective function

The maximum of $E(\mathbf{m})$ can be found with a gradient algorithm. Implementation of a gradient algorithm requires that the gradient of the objective function be computed with respect to the model. The gradient can be expressed as

$$\nabla E_{\mathbf{m}} = \sum_{i} \frac{\partial B_i(\xi_i(\mathbf{m}))}{\partial \mathbf{m}} = \sum_{i} \frac{\partial \xi_i}{\partial \mathbf{m}} \frac{\partial B_i(\xi_i)}{\partial \xi_i} = \mathbf{G}^T \mathbf{D}_1, \tag{4}$$

where the derivatives are computed at fixed y, τ, p_y and p_h.

The vector $\mathbf{D}_1$ is easily computed from beam-stacked data with a finite-difference approximation of the derivative operator. This vector represents the interaction of the inversion algorithm with the actual data.

To evaluate the Frechet derivatives the following relation is used:

$$\left.\frac{\partial \xi}{\partial \mathbf{m}}\right|_{(\bar{y},\bar{\tau})} = \left.\frac{\delta \xi}{\delta \mathbf{m}}\right|_{(\bar{R},\bar{p}_y,\bar{p}_h)} - \left.\frac{\delta y}{\delta \mathbf{m}}\right|_{(\bar{R},\bar{p}_y,\bar{p}_h)} \left.\frac{\partial \xi}{\partial y}\right|_{(\bar{y},\bar{\tau})} - \left.\frac{\delta \tau}{\delta \mathbf{m}}\right|_{(\bar{R},\bar{p}_y,\bar{p}_h)} \left.\frac{\partial \xi}{\partial \tau}\right|_{(\bar{y},\bar{\tau})}. \tag{5}$$

Partial derivatives $\partial \xi / \partial y$ and $\partial \xi / \partial \tau$ are evaluated by use of finite difference on the manifold defined by $\xi = \xi(y, \tau; \mathbf{m})$, at constant transformed traveltime $\bar{\tau}$ and midpoint $\bar{y}$. Total derivatives with respect to the velocity model $\delta \xi / \delta \mathbf{m}$, $\delta y / \delta \mathbf{m}$ and $\delta \tau / \delta \mathbf{m}$, are computed by use of the ray-tracing method presented in Appendix A, for fixed ray parameters $\bar{p}_y$, $\bar{p}_h$, and reflector point $\bar{R}$.

BACK-PROJECTION OPERATOR

The linear operator $\mathbf{G}$ relates slowness perturbations to the resulting perturbations of the transformed offset ξ. The transpose operator $\mathbf{G}^T$ is used for computing

the gradient. The operator $\mathbf{G}^T$ back-projects derivatives of the objective function with respect to the transformed offset ξ into the slowness model.

The back-projection operator is a function of the location of the slowness anomaly, the reflector point, the midpoint position, and the ray parameters. The next series of figures display some interesting *slices* of this multi-dimensional operator. Figure 1 shows the amplitude of the operator as a function of the location of the anomaly, at fixed reflector point, midpoint, and ray parameters. The operator is computed by application of equation (5) and ray tracing through a background model with a constant velocity of 2 km/s. The midpoint ray parameter is zero, and the offset ray parameter is .06 s/km. The operator is smooth because the model is parametrized by use of B-spline functions. The operator is non-zero in a band around the down-going and up-going rays. Above the rays the operator is positive (light in Figure 1) and below it is negative (dark in Figure 1). A slowness anomaly above the ray makes the ray travel more horizontally, and therefore increases the offset; below the ray an anomaly makes the ray travel more vertically and thus decreases the offset. Figure 2 shows a cross-section of Figure 1 taken at constant depth of 180 m. The operator is symmetric and looks like a smoothed *first derivative* operator.

Figure 3 shows the back-projection operator obtained when the midpoint ray parameter is .4 s/km and the offset ray parameter is .06 s/km. Figure 4 shows a cross-section of Figure 3 taken at constant depth of 180m. The operator is asymmetric.

Figure 5 shows the operator obtained when the background velocity is linearly increasing with depth at a rate of 1 s^{-1}. The midpoint ray parameter is .4 s/km and the offset ray parameter is .06 s/km. The rays are no longer straight lines, and the operator follows the bending of the rays. Figure 6 shows a cross-section of Figure 5 taken at constant depth of 180m.

Figure 7 shows the operator obtained when the background velocity has a constant lateral gradient of 1 s^{-1}. The midpoint ray parameter is zero and the offset ray parameter is .06 s/km. The apparent dip of the reflector is the effect of the ray-bending caused by the lateral gradient in velocity. Figure 8 shows a cross-section of Figure 7 taken at constant depth of 180m.

The previous figures show the gradient for one beam. The total gradient is the combination of the contributions of all the beams. The set of beams having the same reflector position and the same midpoint ray parameter, but varying offset ray parameter, corresponds to reflections from the same reflector segment (that is, to the same hyperbolic reflection in the data). The combination of the operators corresponding to this set of beams generates the gradient for the whole hyperbolic reflection, as recorded by all the geophones in the cable. Figure 9 shows the result of summing, with equal weights, the 12 operators corresponding to offset ray parameter from .005 s/km to .06 s/km. The background velocity is a constant of 2 km/s and the midpoint ray parameter is zero. Figure 10 shows a cross-section of Figure 9, taken at a depth of 180m.

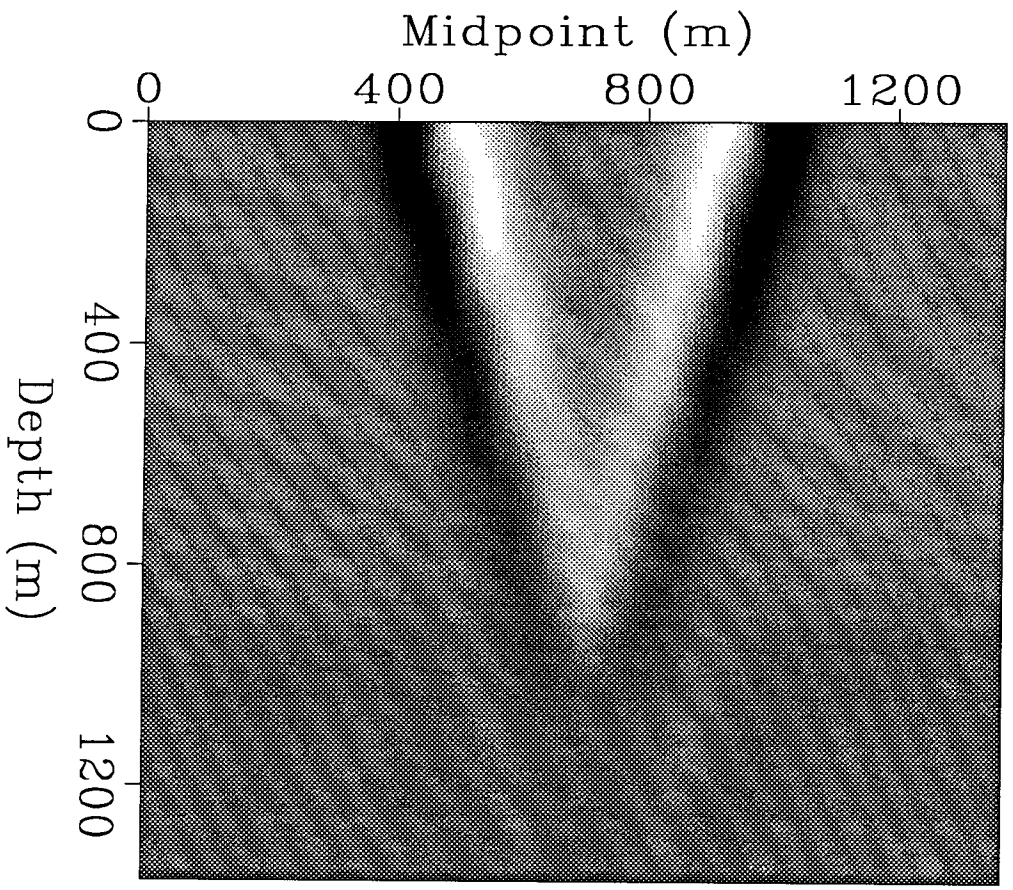

FIG. 1. The back-projection operator $\mathbf{G}^T$ as a function of the location of the slowness anomaly at fixed reflector point, midpoint and ray parameters. The midpoint ray parameter is zero and the offset ray parameter is .06 s/km. The operator is spatially localized around the rays. Light areas indicate positive amplitude and dark areas indicate negative amplitude.

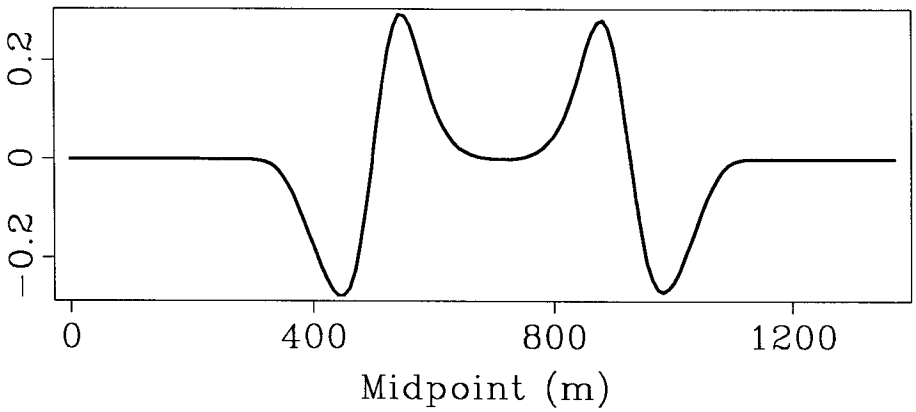

FIG. 2. A cross-section of the back-projection operator shown above taken at a depth of 180 m. The operator is symmetric and looks like a smoothed *first derivative* operator.

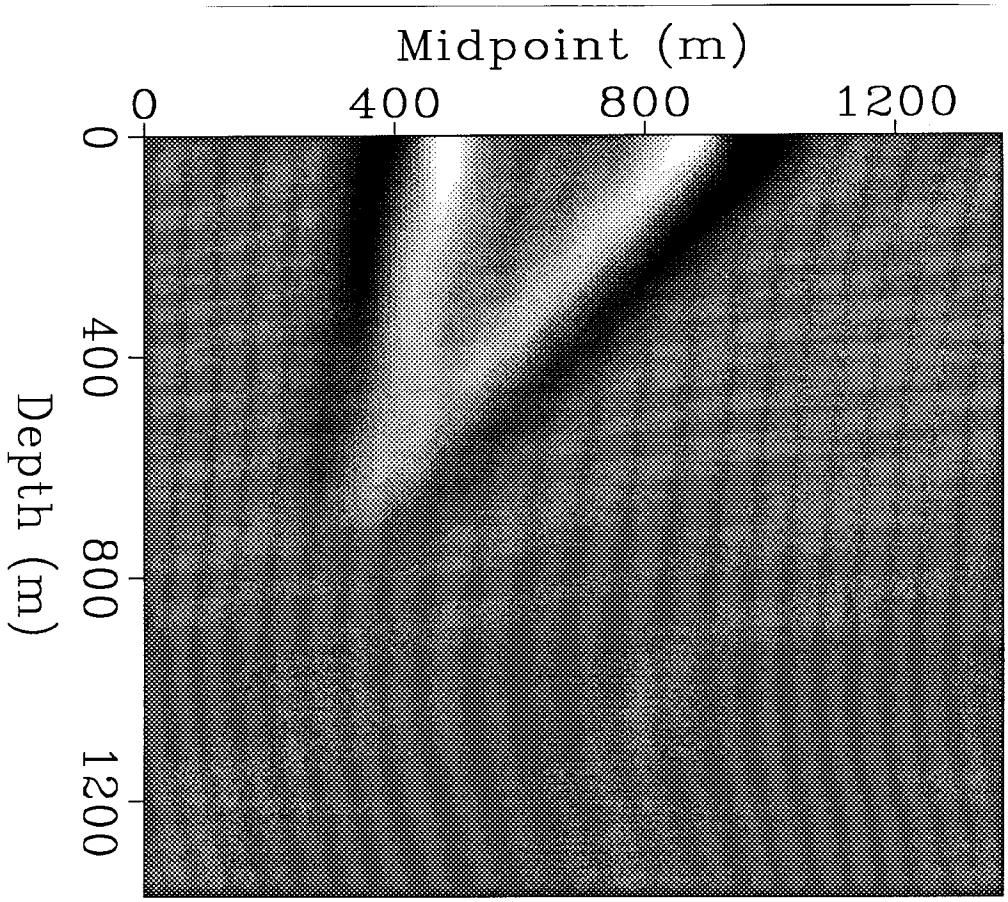

FIG. 3. The back-projection operator $\mathbf{G}^T$ as a function of the location of the slowness anomaly at fixed reflector point, midpoint and ray parameters. The midpoint ray parameter is .4 s/km and the offset ray parameter is .06 s/km. Light areas indicate positive amplitude and dark areas indicate negative amplitude.

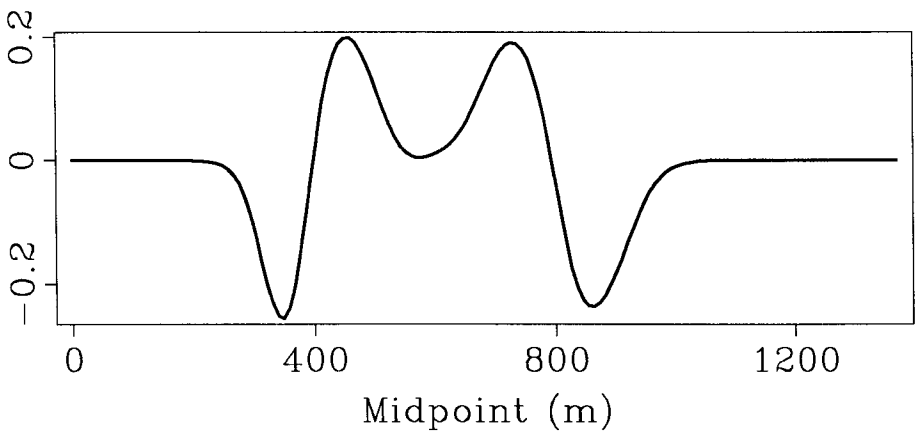

FIG. 4. A cross-section of the back-projection operator shown above taken at a depth of 180 m. The operator is asymmetric.

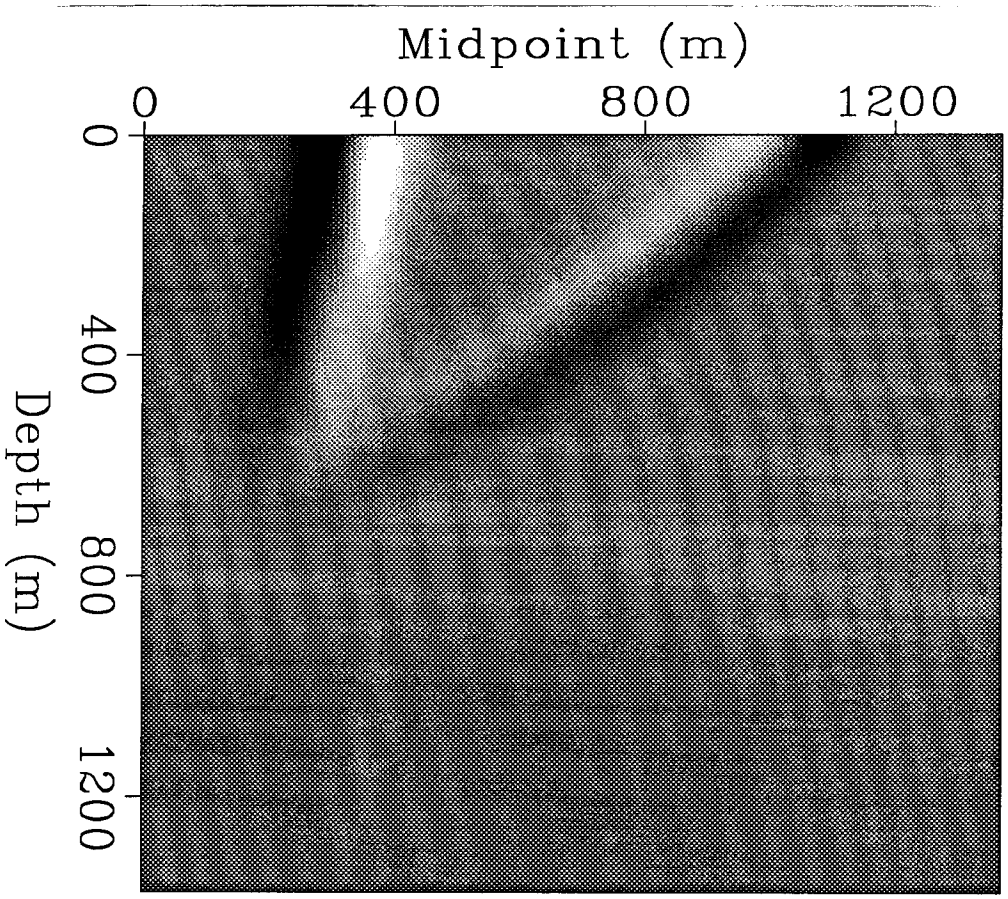

FIG. 5. The back-projection operator $\mathbf{G}^T$ as a function of the location of the slowness anomaly at fixed reflector point, midpoint and ray parameters. The midpoint ray parameter is .4 s/km and the offset ray parameter is .06 s/km. The background velocity is linearly increasing with depth at a rate of 1 s^{-1}. Light areas indicate positive amplitude and dark areas indicate negative amplitude.

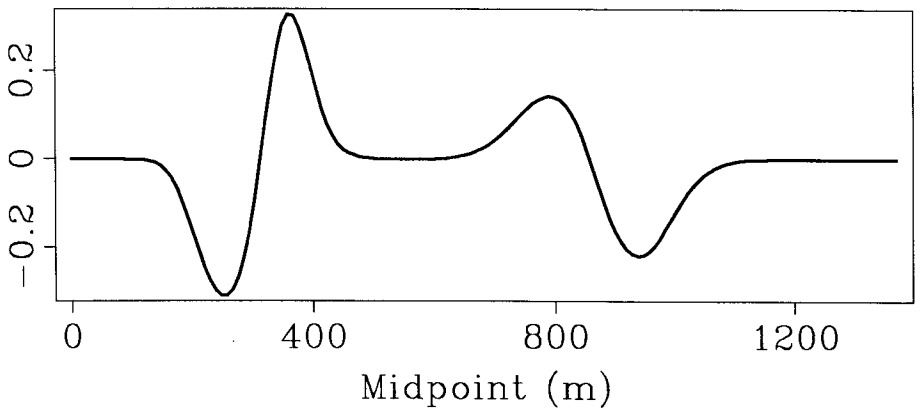

FIG. 6. A cross-section of the back-projection operator shown above taken at a depth of 180 m.

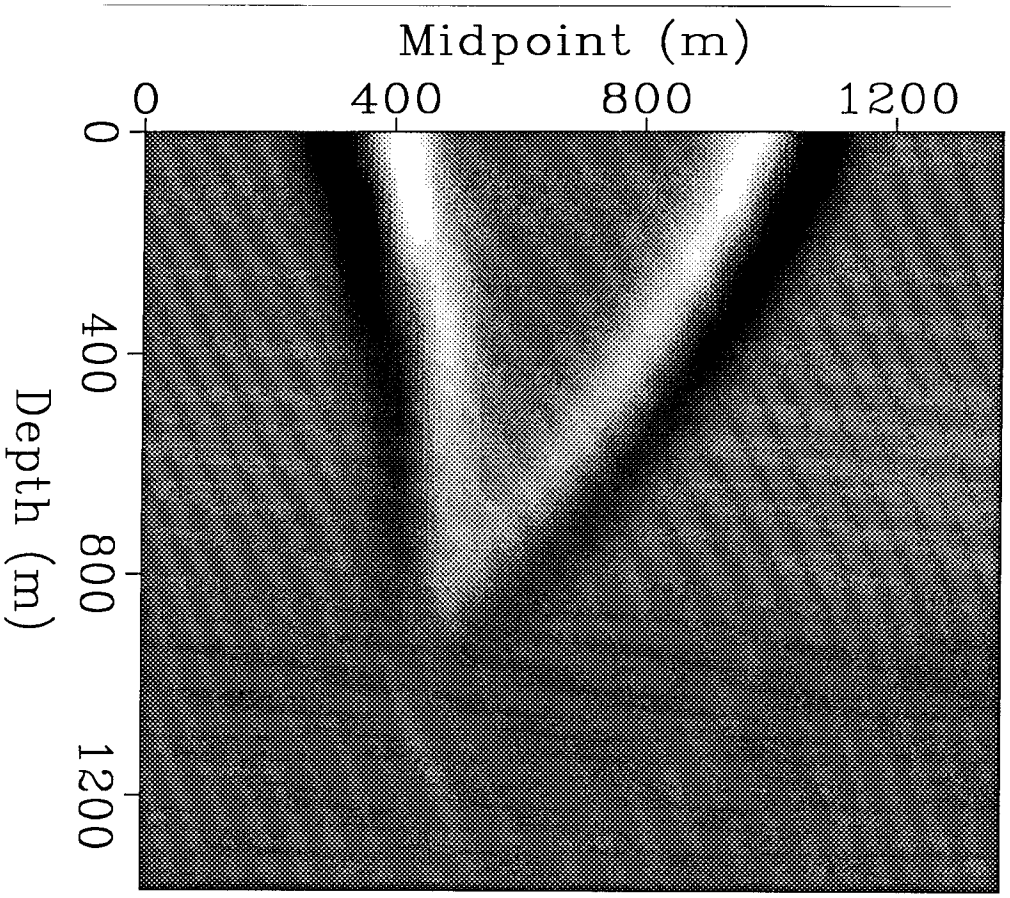

FIG. 7. The back-projection operator $\mathbf{G}^T$ as a function of the location of the slowness anomaly at fixed reflector point, midpoint and ray parameters. The midpoint ray parameter is zero and the offset ray parameter is .06 s/km. The background velocity has a constant lateral gradient of 1 s^{-1}. The apparent dip of the reflector is the effect of ray bending caused by the lateral gradient in velocity. Light areas indicate positive amplitude and dark areas indicate negative amplitude.

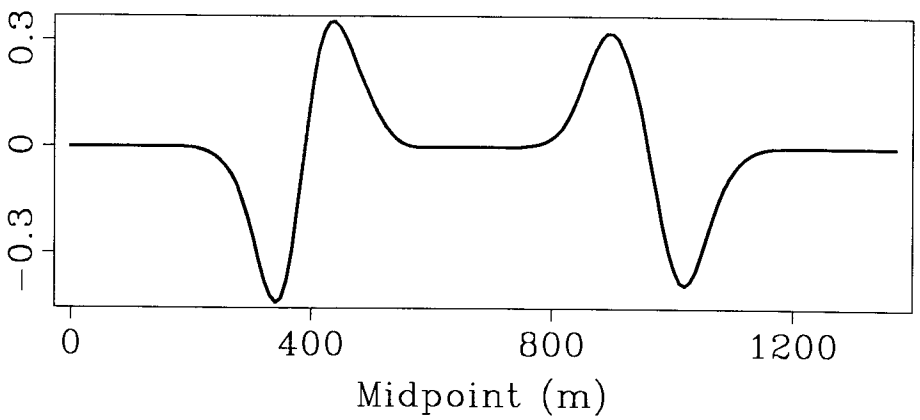

FIG. 8. A cross-section of the back-projection operator shown above taken at a depth of 180 m.

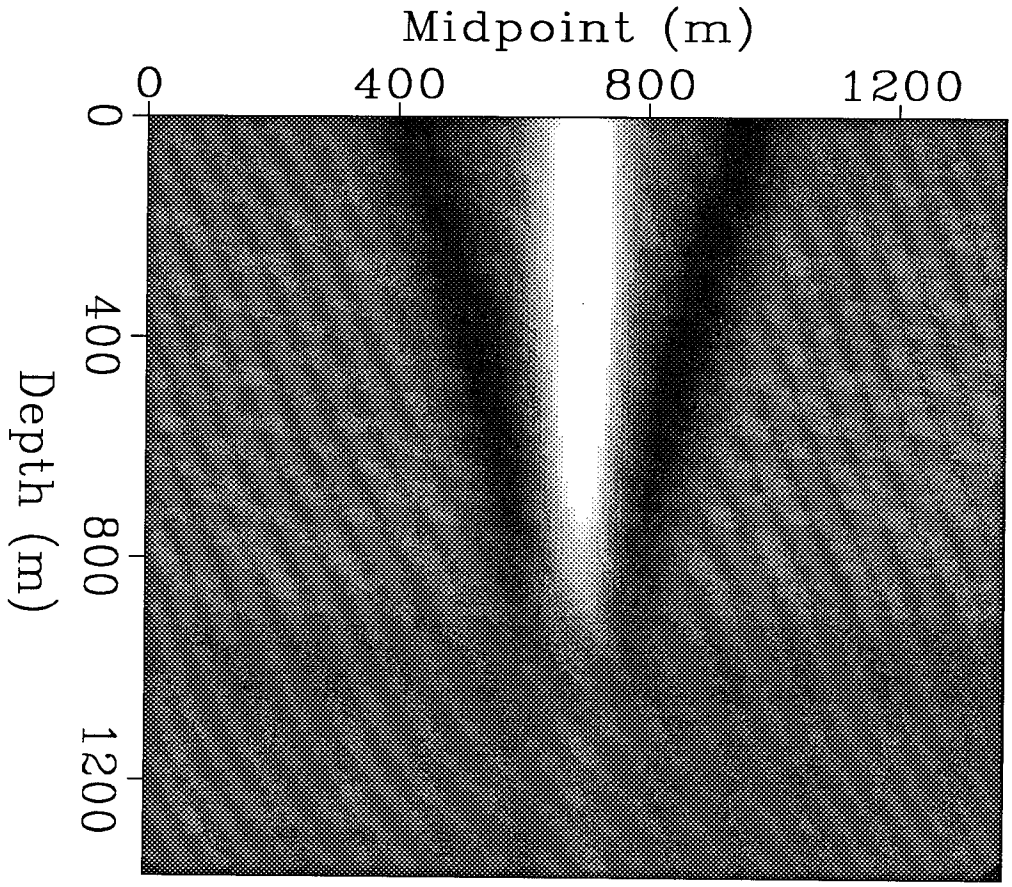

FIG. 9. The combination of the back-projection operators corresponding to offset ray parameters from .005 s/km to .06 s/km, but at fixed midpoint ray parameter, reflector point and midpoint location. The midpoint ray parameter is zero. This operator is similar to Toldi's operator for stacking slowness. Light areas indicate positive amplitude and dark areas indicate negative amplitude.

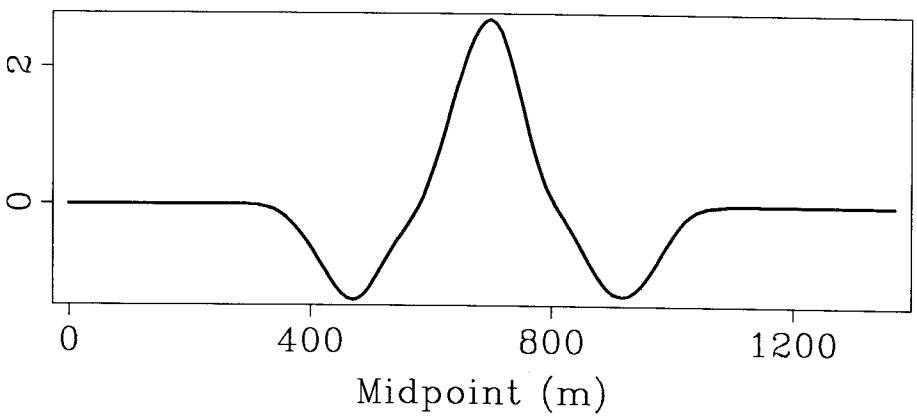

FIG. 10. A cross-section of the back-projection operator shown above taken at a depth of 180 m.

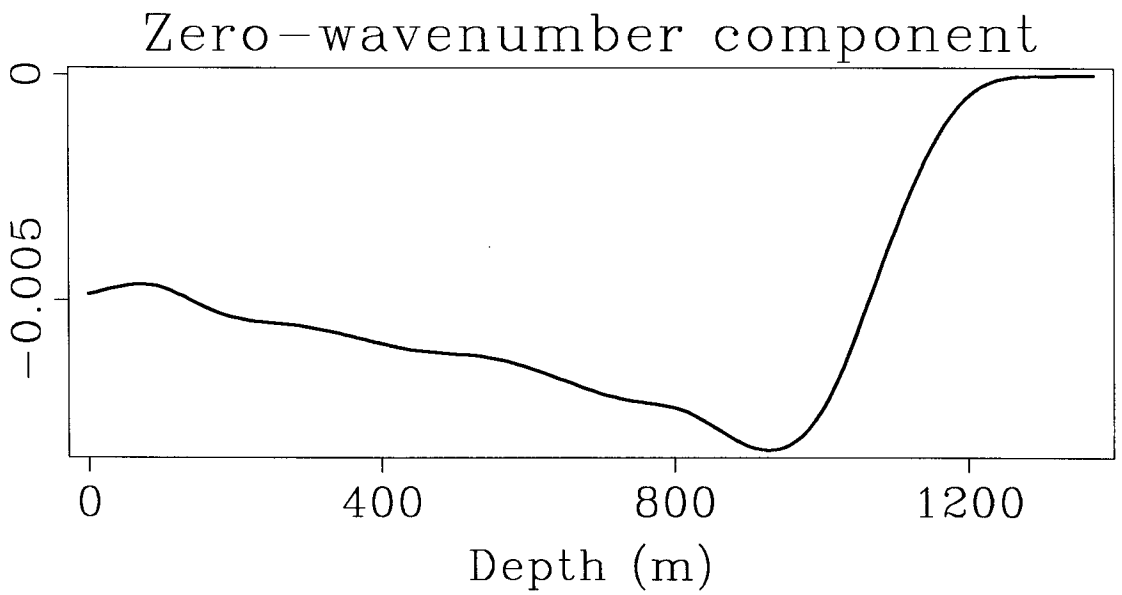

FIG. 11. The zero-wavenumber component of the back-projection operator as a function of depth. The function is negative because a constant increase in slowness causes a decrease in the transformed offset ξ. The amplitudes are low compared to the amplitudes of the operator itself (Figure 2).

The resulting operator is similar to a smoothed version of the operator that John Toldi (1985) derived for stacking velocity and Paul Fowler (1988) derived for migration velocity. Their operator is therefore functionally equivalent to a special case of mine, when the contributions of all the beams corresponding to the same reflector segment are equally weighted. I think that there is a trade-off between using one operator or the other. The gradient operator for the beam-stack is more local and has better spatial resolution than the operator for stacking velocity. On the other hand, the computations of stacking velocity, or migration velocity, use larger arrays than are used in computing beam-stack. The resulting estimation of stacking velocity, or migration velocity, is more robust than the estimation of beam-stack's parameters.

Figure 11 shows the result of integrating along the horizontal axis the operator shown in Figure 1. The plot shows the zero-wavenumber component of the back-projection operator as a function of depth. The zero-wavenumber component is negative because a constant increase in slowness causes a decrease in the transformed offset ξ. The amplitudes are low compared to the amplitudes of the operator itself (Figure 2). Therefore the determination of the low-wavenumber components of the slowness model using the 2-D operator is difficult. I think that one solution for estimating the low-wavenumber component is to begin the inversion procedure with a slowness model very smooth in the horizontal direction. At the limit the 1-D operator (Biondi, 1988) can be used. In the next section I show the result of

the inversion of a simple 2-D model in which the 1-D operator is used for a joint inversion of many midpoints.

RESULTS OF 1-D INVERSION OF SYNTHETIC DATA

I implemented an algorithm based on the 1-D operator for simultaneously inverting more than one common midpoint (CMP) gather. Although this algorithm is inaccurate when the velocity model is two-dimensional, I think that a joint 1-D inversion of many CMP can be a valuable tool for estimating the low-wavenumber components of the model. I show the result of inversion of a data-set containing a dipping reflector because in my previous report I tested the 1-D algorithm only in the case of a horizontally layered medium.

The synthetic data was generated using a finite difference modeling program, assuming the velocity function shown in Figure 12. For the inversion I used 12 CMP gathers, located from 300m to 480m. The maximum offset was 900m and the geophone spacing 15m. I computed beam stacks for two values of the midpoint ray parameter p_y; one for the flat reflector and one for the dipping one. In the offset direction I evaluated beam stacks for four offset ray parameters p_h. The starting model was a velocity function linearly increasing with depth and constant in the horizontal direction. Figure 13 shows the result of the final iteration of the inversion algorithm. The algorithm is limited by the 1-D assumption and therefore it back-projected in the wrong location the velocity information contained in the reflections originated by the dipping bed. Figure 14 shows the velocity functions at the horizontal location of 380m. The true model is drawn with a dashed line, the starting model with a dotted line, and the inversion result with a solid line. The inversion correctly estimated the value of velocity in the two upper layers.

CONCLUSIONS

The gradient of energy in beam-stacks with respect to the slowness model can be computed evaluating a linear operator relating perturbation in interval slowness to perturbations in beam-stack's parameters. The transpose of this linear operator is the back-projection operator of the tomographic inversion. For one beam the back-projection operator is spatially localized around the rays. An inversion algorithm using this back-projection operator should well resolve velocity anomalies.

ACKNOWLEDGMENTS

I thank Jos Van Trier for giving me his ray-tracing code and for suggestions on the implementation of my ray-tracing algorithm. I also thank John Etgen for many discussions on velocity analysis.

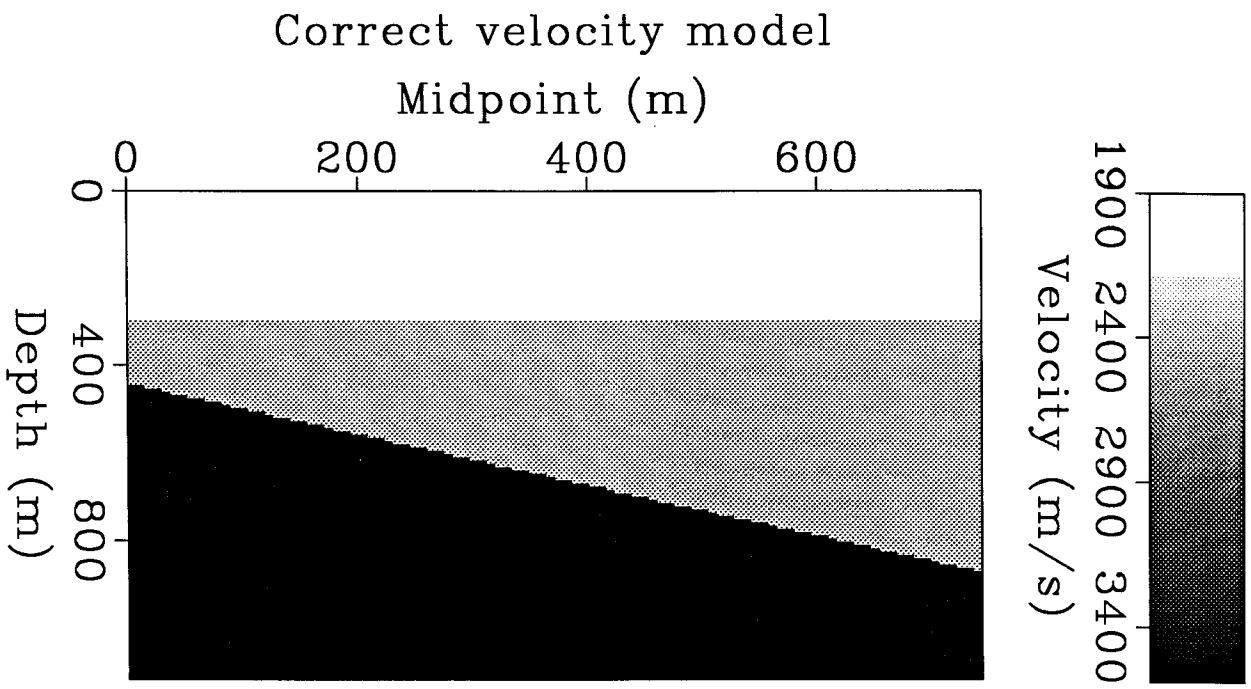

FIG. 12. The velocity model used to generated the synthetic data. The dipping reflector has a dip angle of 30°.

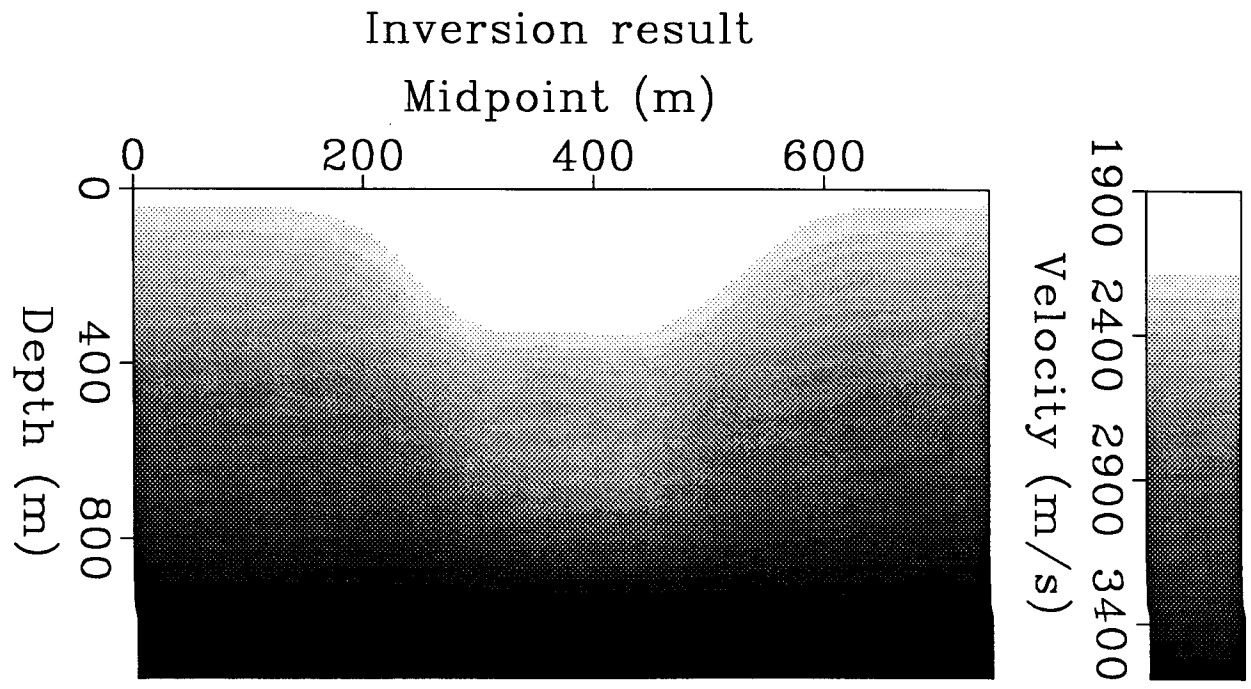

FIG. 13. The result of the last iteration of the inversion algorithm using 12 CMP gathers located from 300m to 480m. The algorithm is limited by the 1-D assumption and back-projected in the wrong location the velocity information contained in the reflections originated by the dipping bed.

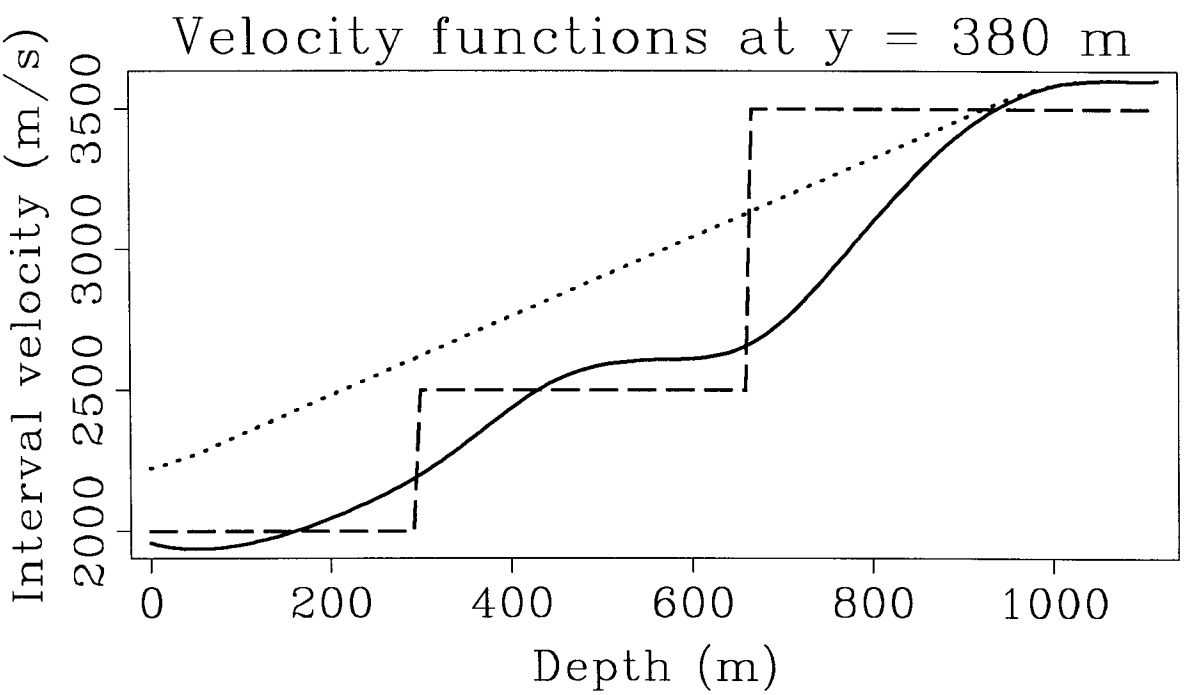

FIG. 14. The velocity functions at the horizontal location of 380m. The true model is drawn with a dashed line, the starting model with a dotted line, and the inversion result with a solid line. The inversion correctly estimated the value of velocity in the two upper layers.

REFERENCES

Biondi, B., 1987, Interval velocity estimation from beam-stacked data: SEP–**51**, 13–27.

Biondi, B., 1988, Interval velocity estimation from beam-stacked data — an improved method: SEP–**57**, 99–115.

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

Cerveny, V., 1987, Ray tracing algorithms in three-dimensional laterally varying layered structures: Seismic Tomography, edited by G. Nolet, Riedel Publishing Co., 99–134.

Diercks, P., 1975, An algorithm for smoothing, differentiation and integration of experimental data using splines functions. J. Comp. Appl. Maths. **1**, 165–185.

Fawcett, J., 1983, I. Three-dimensional ray-tracing and ray-inversion in layered media. II. Inverse scattering and curved ray tomography with applications to seismology: Ph.D. thesis, California Institute of Technology.

Fowler, P., 1988, Seismic velocity estimation using prestack time migration: Ph.D. thesis, Stanford University.

Kostov, C., and Biondi, B., 1987, Improved resolution of slant stacks using beam stacks: SEP–**51.**, 343–350.

Sword, C.H., 1987, Tomographic determination of interval velocities from reflection seismic data: The method of controlled directional reception: Ph.D. thesis, Stanford University.

Toldi, J.L., 1985, Velocity analysis without picking: Ph.D. thesis, Stanford University.

Van Trier, J., 1988, Vectorized initial-value ray tracing for arbitrary velocity models: SEP–**57**, 279–287.

APPENDIX A

In this appendix I present a method for computing the raypath and traveltime, and their derivatives with respect to the slowness model, of a ray traveling in a general 2-D medium. The raypath is computed by the solution of a ray-tracing system of first order ordinary differential equations, derived from the Eikonal equation (Cerveny, 1987). The gradients are computed by use of a continuation method (Fawcett, 1983).

Computing the raypath and traveltime

Rays are traced in a general 2-D medium with slowness function $S(z, y; \mathbf{m})$, where z is depth, y the horizontal coordinate, and $\mathbf{m}$ is the vector of slowness parameter. The slowness function is parametrized by use of B-spline functions(Diercks, 1975).

The rays are traced in depth solving the ray-tracing system

$$dy = \frac{p_y dz}{p_z} \tag{A-1a}$$

$$dp_z = \frac{S(z,y,\mathbf{m}) S_z(z,y,\mathbf{m}) dz}{p_z} \tag{A-1b}$$

$$dp_y = \frac{S(z,y,\mathbf{m}) S_y(z,y,\mathbf{m}) dz}{p_z}; \tag{A-1c}$$

and the traveltime is computed with the equation

$$dt = \frac{S^2(z,y,\mathbf{m}) dz}{p_z}. \tag{A-1d}$$

Here $S_z(z,y,\mathbf{m})$ and $S_y(z,y,\mathbf{m})$ are the partial derivatives of slowness with respect to the spatial coordinates, p_y is the horizontal ray parameter, and p_z is the vertical ray parameter.

The rays are traced down from the surface, thus the initial conditions are

$$z = 0$$

$$y = y_0$$

$$p_z = \sqrt{S^2(0, y_0, \mathbf{m}) - p_{obs}^2}$$

$$p_y = p_{obs}$$

and

$$t = 0,$$

where p_{obs} is the horizontal ray parameter measured at the surface by beam stacks.

The ray-tracing system can be numerically solved by use of a standard Runge-Kutta algorithm. The computation can be efficiently vectorized by the simultaneous tracing of a fan of rays (Van Trier, 1988).

Computing the gradients with respect to the slowness model

The velocity estimation method needs the gradient of the raypath with respect to the slowness model. This gradient is computed with a continuation method. The differential equations in the ray-tracing system can be approximated by the following system of difference equations:

$$Y^k(\mathbf{m}) = y^k - y^{k-1} - \frac{\bar{p}_y^k \Delta z}{\bar{p}_z^k} \tag{A-2a}$$

$$P_z^k(\mathbf{m}) = p_z^k - p_z^{k-1} - \frac{\bar{S}^k \bar{S}_z^k \Delta z}{\bar{p}_z^k} \tag{A-2b}$$

$$P_y^k(\mathbf{m}) = p_y^k - p_y^{k-1} - \frac{\bar{S}^k \bar{S}_y^k \Delta z}{\bar{p}_z^k} \tag{A-2c}$$

and

$$T^k(\mathbf{m}) = t^k - t^{k-1} - \frac{(\bar{S}^k)^2 \Delta z}{\bar{p}_z^k}, \tag{A-2d}$$

where $\Delta z = z^k - z^{k-1}$ is constant, $\bar{S}^k$, $\bar{S}_z^k$ and $\bar{S}_y^k$ are the slowness and its derivatives evaluated at the intermediate point $\left[\bar{z}^k = (z^k + z^{k-1})/2,\ \bar{y}^k = (y^k + y^{k-1})/2\right]$, and $\bar{p}_z^k$ and $\bar{p}_y^k$ are the averages $\bar{p}_z^k = (p_z^k + p_z^{k-1})/2$ and $\bar{p}_y^k = (p_y^k + p_y^{k-1})/2$.

Equations (A-2) must be satisfied for all values of the slowness model $\mathbf{m}$, therefore the derivatives of $Y^k(\mathbf{m})$, $P_z^k(\mathbf{m})$, and $P_y^k(\mathbf{m})$ with respect to the slowness model are zero. These derivatives with respect to the model parameter m_i can be expanded with the chain rule as

$$\frac{\partial Y^k}{\partial m_i} = \frac{\partial Y^k}{\partial y^k}\frac{\partial y^k}{\partial m_i} + \frac{\partial Y^k}{\partial y^{k-1}}\frac{\partial y^{k-1}}{\partial m_i} + \frac{\partial Y^k}{\partial \bar{p}_z^k}\frac{\partial \bar{p}_z^k}{\partial m_i} + \frac{\partial Y^k}{\partial \bar{p}_y^k}\frac{\partial \bar{p}_y^k}{\partial m_i} = 0 \tag{A-3a}$$

$$\frac{\partial P_z{}^k}{\partial m_i} = \frac{\partial P_z{}^k}{\partial p_z^k}\frac{\partial p_z{}^k}{\partial m_i} + \frac{\partial P_z{}^k}{\partial p_z^{k-1}}\frac{\partial p_z{}^{k-1}}{\partial m_i} + \frac{\partial P_z{}^k}{\partial \bar{p}_z^k}\frac{\partial \bar{p}_z^k}{\partial m_i} +$$

$$\frac{\partial P_z{}^k}{\partial \bar{S}^k}\frac{\partial \bar{S}^k}{\partial m_i} + \frac{\partial P_z{}^k}{\partial \bar{S}_z^k}\frac{\partial \bar{S}_z^k}{\partial m_i} = 0 \tag{A-3b}$$

$$\frac{\partial P_y{}^k}{\partial m_i} = \frac{\partial P_y{}^k}{\partial p_y^k}\frac{\partial p_y{}^k}{\partial m_i} + \frac{\partial P_y{}^k}{\partial p_y^{k-1}}\frac{\partial p_y{}^{k-1}}{\partial m_i} + \frac{\partial P_y{}^k}{\partial \bar{p}_z^k}\frac{\partial \bar{p}_z^k}{\partial m_i} + \frac{\partial P_y{}^k}{\partial \bar{S}^k}\frac{\partial \bar{S}^k}{\partial m_i} + \frac{\partial P_y{}^k}{\partial \bar{S}_y^k}\frac{\partial \bar{S}_y^k}{\partial m_i} = 0. \tag{A-3c}$$

Evaluating all the partial derivatives in equations (A-3) it is possible to obtain a system that can be sequentially solved at each step k for the derivatives $\partial y^k/\partial m_i$, $\partial p_z{}^k/\partial m_i$ and $\partial p_y{}^k/\partial m_i$, when the derivatives at the previous step $k-1$ are known. Equations (A-3) are not independent, and therefore a system of three equations in three unknowns must be solved at each step. The three equations can be solved sequentially after these terms are dropped: those that depend on the second derivatives of slowness with respect to the spatial coordinates, and those terms that depend on the product of two first derivatives. It is possible to drop these terms because the slowness function is assumed to be smooth. After these approximations and some simplifications, equations (A-3) become

$$\frac{\partial y^k}{\partial m_i} = \frac{\partial y^{k-1}}{\partial m_i} + \frac{\Delta z}{2\bar{p}_z^k}\left(\frac{\partial p_y{}^k}{\partial m_i} + \frac{\partial p_y{}^{k-1}}{\partial m_i}\right) - \frac{\bar{p}_y^k \Delta z}{2(\bar{p}_z^k)^2}\left(\frac{\partial p_z{}^k}{\partial m_i} + \frac{\partial p_z{}^{k-1}}{\partial m_i}\right) \tag{A-4a}$$

$$\frac{\partial p_z{}^k}{\partial m_i}\left(1 + \frac{\bar{S}^k \bar{S}_z^k \Delta z}{2(\bar{p}_z^k)^2}\right) = \frac{\partial p_z{}^{k-1}}{\partial m_i}\left(1 - \frac{\bar{S}^k \bar{S}_z^k \Delta z}{2(\bar{p}_z^k)^2}\right) + \frac{\bar{S}^k \Delta z}{\bar{p}_z^k}\left.\frac{\partial S_z(z,y,\mathbf{m})}{\partial m_i}\right|_{(\bar{z},\bar{y})} + \frac{\bar{S}_z^k \Delta z}{\bar{p}_z^k}\left.\frac{\partial S(z,y,\mathbf{m})}{\partial m_i}\right|_{(\bar{z},\bar{y})} \tag{A-4b}$$

$$\frac{\partial p_y{}^k}{\partial m_i} = \frac{\partial p_y{}^{k-1}}{\partial m_i} + \frac{\bar{S}^k \Delta z}{\bar{p}_z^k}\left.\frac{\partial S_y(z,y,\mathbf{m})}{\partial m_i}\right|_{(\bar{z},\bar{y})} + \frac{\bar{S}_y^k \Delta z}{\bar{p}_z^k}\left.\frac{\partial S(z,y,\mathbf{m})}{\partial m_i}\right|_{(\bar{z},\bar{y})} - \frac{\bar{S}^k \bar{S}_y^k \Delta z}{2(\bar{p}_z^k)^2}\left(\frac{\partial p_z{}^k}{\partial m_i} + \frac{\partial p_z{}^{k-1}}{\partial m_i}\right). \tag{A-4c}$$

Equations (A-4) can be solved sequentially: start from equation (A-4b), follow with equation (A-4c), and then equation (A-4a). The initial conditions at the surface are:

$$\frac{\partial y^0}{\partial m_i} = 0 \tag{A-5a}$$

$$\frac{\partial p_z{}^0}{\partial m_i} = \frac{S(0,y_0,\mathbf{m})}{\sqrt{S^2(0,y_0,\mathbf{m}) - p_{obs}^2}}\left.\frac{\partial S(0,y_0,\mathbf{m})}{\partial m_i}\right|_{(0,y_0)} \tag{A-5b}$$

$$\frac{\partial {p_y}^0}{\partial m_i} = 0. \tag{A-5c}$$

Once the gradient of the raypath with respect to the slowness model is computed, it is easy to evaluate the gradient of traveltime. Differentiating equation (A-2d) I obtain the equation

$$\frac{\partial t^k}{\partial m_i} = \frac{\partial t^{k-1}}{\partial m_i} + \frac{2\bar{S}^k \Delta z}{\bar{p}_z^k} \left[\left. \frac{\partial S(z,y,\mathrm{m})}{\partial m_i} \right|_{(\bar{z},\bar{y})} + \frac{\bar{S}_y^k}{2} \left(\frac{\partial y^k}{\partial m_i} + \frac{\partial y^{k-1}}{\partial m_i} \right) \right] -$$

$$\frac{(\bar{S}^k)^2 \Delta z}{2(\bar{p}_z^k)^2} \left(\frac{\partial {p_z}^k}{\partial m_i} + \frac{\partial {p_z}^{k-1}}{\partial m_i} \right), \tag{A-6}$$

which can be used to evaluate sequentially the traveltime derivatives.

Tracing rays downward or upward ?

The theory presented at the beginning of this paper requires the rays be traced from each reflector point up to the surface. In practice it is more efficient to trace the rays downward and couple them at each depth level. The raypath is independent from the direction of propagation but the gradients are not. Equation (A-4b) and (A-4c) must be solved starting from the surface because the initial conditions are specified at the surface, where the horizontal ray parameter is measured. Equation (A-4a) can be solved either way. Between two specified depth points the solution have the same magnitude but opposite sign. Therefore the only term that is really affected by the direction of propagation is the gradient of traveltime. In particular the term

$$\frac{\bar{S}_y^k}{2} \left(\frac{\partial y^k}{\partial m_i} + \frac{\partial y^{k-1}}{\partial m_i} \right)$$

depends on the direction of propagation. I ran some practical test to study quantitatively the importance of this term; as a result I have decided to ignore it in the computation of the gradient of traveltime with respect to the slowness model.

Interactive Burg spectrum ala SEP-56 p263