

Wave-equation travelttime inversion

Y. Luo* and G. T. Schuster*

ABSTRACT

This paper presents a new travelttime inversion method based on the wave equation. In this new method, designated as wave-equation travelttime inversion (WT), seismograms are computed by any full-wave forward modeling method (we use a finite-difference method). The velocity model is perturbed until the travelttimes from the synthetic seismograms are best fitted to the observed travelttimes in a least squares sense. A gradient optimization method is used and the formula for the Frechét derivative (perturbation of travelttimes with respect to velocity) is derived directly from the wave equation. No travelttime picking or ray tracing is necessary, and there are no high frequency assumptions about the data. Body wave, diffraction, reflection and head wave travelttimes can be incorporated into the inversion. In the high-

frequency limit, WT inversion reduces to ray-based travelttime tomography. It can also be shown that WT inversion is approximately equivalent to full-wave inversion when the starting velocity model is "close" to the actual model.

Numerical simulations show that WT inversion succeeds for models with up to 80 percent velocity contrasts compared to the failure of full-wave inversion for some models with no more than 10 percent velocity contrast. We also show that the WT method succeeds in inverting a layered velocity model where a shooting ray-tracing method fails to compute the correct first arrival times. The disadvantage of the WT method is that it appears to provide less model resolution compared to full-wave inversion, but this problem can be remedied by a hybrid travelttime + full-wave inversion method (Luo and Schuster, 1989).

INTRODUCTION

Seismic inversion algorithms span the range between two extremes: travelttime inversion (Dines and Lytle, 1979; Paulsson et al., 1985; Ivansson, 1985; Bishop et al., 1985; Lines, 1988; Justice et al., 1989; and many others) and full-wave inversion (Tarantola, 1987; Johnson and Tracy, 1983; and others). Travelttime inversion typically uses ray tracing to compute both the travelttimes and Frechét derivative (perturbations of travelttimes with respect to velocities). While computationally efficient, travelttime inversion is subject to a high-frequency assumption about the data and can therefore fail when the earth's velocity variations are characterized by the same wavelength as in the source wavelet. On the other hand, we can show that the misfit function to be minimized (sum of the squared errors between observed and calculated travelttimes) can be *quasi-linear* with respect to the relative change between the assumed and actual velocity models. A gradient optimization algorithm

(e.g., conjugate gradients) can thus make rapid progress in searching for the correct velocity model and successful inversion can be achieved even if the starting model is far from the actual model.

Attempts to bridge the gap between the extremes of travelttime inversion and full-wave inversion include Born inversion (Clayton and Stolt, 1981; Weglein, 1982; Keys and Weglein, 1983; Bleistein and Gray, 1985; Carrion and Foster, 1985; and others) and other amplitude methods which are subject to restrictive assumptions about the data. These intermediate methods can be very successful for some data sets but usually not for data with strong contrasts in impedance. Intermediate methods also include surface-wave inversion (Wattus, 1989) and diffraction tomography (Lo et al., 1988)

Full-wave inversion overcomes limitations imposed by the high-frequency restrictions in travelttime inversion and the weak scattering approximation of Born methods by perturbing the velocity model until the synthetic seismograms match

Manuscript received by the Editor February 20, 1990; revised manuscript received September 13, 1990.

*Department of Geology and Geophysics, 717 W. C. Browning Bldg., University of Utah, Salt Lake City, UT 84112-1183.

© 1991 Society of Exploration Geophysicists. All rights reserved.

the observed seismograms. No approximations are necessary and the synthetic seismograms are usually computed by a finite-difference solution to the wave equation. In addition, the Frechét derivatives are elegantly computed by reverse time migration of the seismogram residuals. The problem with full-wave inversion, however, is that the misfit function (normed difference between observed and synthetic seismograms) can be highly *nonlinear* with respect to the velocity models. Gauthier et al. (1986) showed that full-wave inversion can fail for a velocity model with no more than 10 percent velocity contrast. A reason for this failure is that the misfit function is highly nonlinear with respect to velocity perturbations in the model. In this case a gradient method will tend to get stuck in local minima if the starting model is moderately far from the actual model.

Can one borrow the best characteristics of traveltime inversion (quasi-linear misfit function and robust convergence properties) and full-wave inversion (no approximations to the data) to create an inversion method free from approximations, robust in the presence of data noise, and quickly convergent for starting models far from the actual model? Traveltime inversion might achieve this goal if the wave equation, rather than the approximate method of ray tracing, is used to compute traveltimes and Frechét derivatives. This paper describes the derivation of a new velocity inversion method, wave-equation traveltime inversion (WT), which minimizes traveltime residuals using traveltimes and Frechét derivatives computed from solutions to the wave equation. The merits of WT inversion are that it can invert for some velocity models with more than 80 percent contrast in impedance, its misfit function is roughly independent of realistic density variations, it can invert for complicated velocity models where shooting ray-tracing methods fail, no high-frequency assumptions about the data are necessary, and traveltime picking and event identification may sometimes be unnecessary. The disadvantage is that the WT method is characterized by less model resolution compared to that associated with a full-wave inversion method. We first present a derivation of the WT method, and then present results from synthetic and real data tests.

THEORY

This section presents the derivation of the wave-equation traveltime inversion (WT) method. The key steps are to (1) define a connective (Luo and Schuster, 1991) function that connects the traveltime residual with the pressure seismograms (this step allows for the derivation of the Frechét derivative), (2) define a traveltime misfit function (the summed squared difference between observed and synthetic traveltimes), and (3) derive the perturbation of the misfit function with respect to velocity using the wave equation.

The following analysis assumes that the propagation of seismic waves honors the 2-D acoustic wave equation. Let $p(x_r, t; x_s)_{\text{obs}}$ denote the observed pressure seismograms measured at receiver location x_r due to a line source excited at time $t = 0$ and at location x_s . For a given velocity model, $p(x_r, t; x_s)_{\text{cal}}$ denotes the computed seismograms which satisfy the acoustic wave equation

$$\frac{1}{c^2(x)} \frac{\partial^2 p(x, t; x_s)}{\partial t^2} - \rho(x) \nabla \cdot \left[\frac{1}{\rho(x)} \nabla p(x, t; x_s) \right] = s(t; x), \quad (1)$$

where $\rho(x)$ is the density, $s(t; x)$ is the source function, and $c(x)$ is the wave speed.

Connective function

We now use a crosscorrelation function to define a connective function that connects the traveltimes with the pressure field. The degree to which the synthetic and observed seismograms match each other can be estimated by the crosscorrelation function

$$f(x_r, \tau; x_s) = \int dt \frac{p(x_r, t + \tau; x_s)_{\text{obs}}}{A(x_r; x_s)_{\text{obs}}} p(x_r, t; x_s)_{\text{cal}},$$

where $A(x_r; x_s)_{\text{obs}}$ is the maximum amplitude of $p(x_r, t; x_s)_{\text{obs}}$ and τ is the shift time between synthetic and real seismograms. The divisor $A(x_r; x_s)_{\text{obs}}$ normalizes the observed seismograms to a maximum amplitude of 1 and eliminates amplitude problems due to inconsistent coupling of the geophones or source to the earth.

We seek a τ that shifts a synthetic seismogram so that it “best” matches the observed seismogram. The criterion for “best” match is defined as the traveltime residual $\Delta\tau$ that maximizes the crosscorrelation function $f(x_r, \tau; x_s)$, i.e.,

$$f(x_r, \Delta\tau; x_s) = \max \{f(x_r, \tau; x_s) | \tau \in [-T, T]\}, \quad (2)$$

where T is the estimated maximum traveltime difference between the observed and calculated seismograms. For the examples in this paper, we use only the transmitted waveforms by windowing out all other arrivals so that the $\Delta\tau$ corresponds to the traveltime difference between the observed and calculated transmitted arrivals. Note $\Delta\tau = 0$ indicates that the correct velocity model has been found which generates a transmitted wave arriving at the same time as the observed transmitted wave.

The derivative of $f(x_r, \tau; x_s)$ with respect to τ should be zero at $\Delta\tau$ unless its maximum is at an end point $\Delta\tau = T$ or $\Delta\tau = -T$:

$$\begin{aligned} \dot{f}_{\Delta\tau} &= \left[\frac{\partial f(x_r, \tau; x_s)}{\partial \tau} \right]_{\tau = \Delta\tau} \\ &= \int dt \frac{\dot{p}(x_r, t + \Delta\tau; x_s)_{\text{obs}}}{A(x_r; x_s)_{\text{obs}}} p(x_r, t; x_s)_{\text{cal}} = 0, \end{aligned} \quad (3)$$

where $\dot{p} = \partial p(x, t; x_s) / \partial t$. Equation (3) is the connective function which will be used to compute the Frechét derivative.

Misfit function

The WT method attempts to determine a velocity model $c(x)$ which predicts seismograms $p(x_r, t; x_s)_{\text{cal}}$ that minimize the following misfit function:

$$S = \frac{1}{2} \sum_s \sum_r \Delta\tau(x_r, x_s)^2, \quad (4)$$

where $\Delta\tau$ is defined by equation (2) and the factor $1/2$ is introduced for subsequent simplifications. This criterion can be generalized to account for the estimated observation errors or a priori information in model space.

A gradient method can be used to find the velocity model that minimizes equation (4). For simplicity, we discuss the steepest descent method although the conjugate gradient method can also be used (Tarantola, 1987). To update the velocity model, the steepest descent method gives

$$c(x)_{k+1} = c(x)_k + \alpha_k \gamma(x)_k, \quad (5)$$

where $\gamma(x)_k$ is the steepest descent direction of the misfit function S and α_k is the step length (see Appendix B) for the k th iteration. The central problem is how to compute $\gamma(x)$ using the wave equation. To obtain $\gamma(x)$, take the derivative of S with respect to the velocity model $c(x)$:

$$\gamma(x) = -\frac{\partial S}{\partial c(x)} = -\sum_s \sum_r \frac{\partial(\Delta\tau)}{\partial c(x)} \Delta\tau(x_r, x_s). \quad (6)$$

Using equation (3) and the rule for an implicit function derivative, we get

$$\begin{aligned} \frac{\partial(\Delta\tau)}{\partial c(x)} &= \frac{\left[\frac{\partial(\dot{f}_{\Delta\tau})}{\partial c(x)} \right]}{\left[\frac{\partial(\dot{f}_{\Delta\tau})}{\partial(\Delta\tau)} \right]} \\ &= \frac{1}{E} \int dt \dot{p}(x_r, t + \Delta\tau; x_s)_{\text{obs}} \frac{\partial p(x_r, t; x_s)_{\text{cal}}}{\partial c(x)}, \end{aligned} \quad (7a)$$

where

$$\begin{aligned} E &= - \int dt \ddot{p}(x_r, t + \Delta\tau; x_s)_{\text{obs}} p(x_r, t; x_s)_{\text{cal}} \\ &= \int dt \dot{p}(x_r, t + \Delta\tau; x_s)_{\text{obs}} \dot{p}(x_r, t; x_s)_{\text{cal}}. \end{aligned}$$

In Appendix A, we show that the Fréchet derivative of the pressure field $p(x_r, t; x_s)_{\text{cal}}$ is

$$\frac{\partial p(x_r, t; x_s)_{\text{cal}}}{\partial c(x)} = \frac{2}{c^3(x)} \dot{g}(x, t; x_r, 0) * \dot{p}(x, t; x_s), \quad (7c)$$

where $g(x, t; x', t')$ is the Green's function for equation (1); that is, the pressure field at point x and time t due to the impulse source $\delta(x - x')\delta(t - t')$. The asterisk represents time convolution. Substitution of equation (7c) into equation (7b) gives

$$\begin{aligned} \frac{\partial(\Delta\tau)}{\partial c(x)} &= \frac{2}{c^3(x)} \int dt \dot{g}(x, t; x_r, 0) * \dot{p}(x, t; x_s) \\ &\quad \times \frac{\dot{p}(x_r, t + \Delta\tau; x_s)_{\text{obs}}}{E} \end{aligned} \quad (7d)$$

and substituting equation (7d) into equation (6) gives

$$\gamma = \frac{1}{c^3(x)} \sum_s \sum_r \int dt \dot{g}(x, t; x_r, 0) * \dot{p}(x, t; x_s) \delta\tau(x_r, t; x_s), \quad (8a)$$

where $\delta\tau$ is the pseudotraveltime residual,

$$\delta\tau(x_r, t; x_s) = -\frac{2}{E} \dot{p}(x_r, t + \Delta\tau; x_s)_{\text{obs}} \Delta\tau(x_r, x_s). \quad (8b)$$

Using the identities (Tarantola, 1987)

$$\begin{aligned} \int dt [f(t) * g(t)]h(t) &= \int dt g(t)[f(-t) * h(t)], \\ g(x, -t; x', 0) &= g(x, 0; x', t), \end{aligned}$$

we can rewrite equation (8a) as

$$\gamma = \frac{1}{c^3(x)} \sum_s \int dt \dot{p}(x, t; x_s) \dot{p}'(x, t; x_s), \quad (9)$$

where

$$\dot{p}'(x, t; x_s) = \sum_r g(x, -t; x_r, 0) * \delta\tau(x_r, t; x_s),$$

$p(x, t; x_s)$ is the pressure field calculated for the current velocity model $c(x)$ and $p'(x, t; x_s)$ is the field computed by reverse time propagation of the pseudoresidual $\delta\tau(x_r, t; x_s)$ acting as a source at receiver location x_r . This result is the same as that of full-wave inversion except $\delta\tau$ is used instead of

$$\delta p = p(x_r, t; x_s)_{\text{obs}} - p(x_r, t; x_s)_{\text{cal}}. \quad (10)$$

Combining equations (9) and (5) yields an iterative method to invert for a velocity model $c(x)$ from traveltime residuals. In Appendix B, we describe the computer implementation of this theory.

RELATIONSHIP OF WT INVERSION TO RAY-TRACING TRAVELTIME TOMOGRAPHY AND FULL-WAVE INVERSION

We show that in the high-frequency limit and under a linear perturbation assumption the WT method reduces to ray-tracing traveltime tomography. We also show that the WT method is approximately equivalent to full-wave inversion if the starting model is close to the true model.

According to 3-D asymptotic ray theory,

$$p(x_r, t; x_s)_{\text{cal}} = A(x_r, x_s)_{\text{cal}} \delta[t - \tau(x_r, x_s)_{\text{cal}}], \quad (11)$$

where $\tau(x_r, x_s)_{\text{cal}}$ is the traveltime computed along rays for a given velocity model and $A(x_r, x_s)_{\text{cal}}$ is the amplitude factor which accounts for spherical spreading losses (Bleistein, 1984). Substituting equation (11) into (7b) yields

$$\begin{aligned} \frac{\partial(\Delta\tau)}{\partial c(x)} &= \frac{1}{E} \int dt \dot{p}(x_r, t + \Delta\tau; x_s)_{\text{obs}} \\ &\quad \times \left\{ \frac{\partial A(x_r, x_s)_{\text{cal}}}{\partial c(x)} \delta[t - \tau(x_r, x_s)_{\text{obs}}] \right. \\ &\quad \left. - A(x_r, x_s)_{\text{cal}} \delta[t - \tau(x_r, x_s)_{\text{cal}}] \frac{\partial \tau(x_r, x_s)_{\text{cal}}}{\partial c(x)} \right\}, \end{aligned} \quad (12a)$$

where

$$E = \int dt \dot{p}(x_r, t + \Delta\tau; x_s)_{\text{obs}} A(x_r; x_s)_{\text{cal}} \delta[t - \tau(x_r; x_s)_{\text{cal}}]. \quad (12b)$$

Ray-tracing tomography employs a linear perturbation assumption, i.e., when the velocity $c(x)$ is perturbed, the raypath remains unchanged for first-order variations in velocity and traveltimes. An unperturbed raypath leads to an unperturbed amplitude with a constant ray tube. Therefore, setting the derivative of $A(x_r; x_s)_{\text{cal}}$ with respect to $c(x)$ to 0 in equation (12a) and substituting equation (12b) into equation (12a), we get

$$\frac{\partial(\Delta\tau)}{\partial c(x)} = -\frac{\partial\tau(x_r; x_s)_{\text{cal}}}{\partial c(x)} = \frac{\partial(\Delta\tau')}{\partial c(x)}, \quad (12c)$$

where $\Delta\tau' = \tau(x_r; x_s)_{\text{obs}} - \tau(x_r; x_s)_{\text{cal}}$ and $\tau(x_r; x_s)_{\text{obs}}$ is the traveltimes picked from real data. This result shows that the Jacobian matrix in ray-tracing tomography is a special case of our Jacobian operator in equation (7) for the high-frequency and linear perturbation assumptions.

To establish the relationship between WT and full-wave inversion, assume $p(x_r, t; x_s)_{\text{obs}}$ is the seismogram for the true velocity model $c(x)$. If the current velocity model is $c(x) + \delta c(x)$ and $\delta c(x)$ is small, then the calculated seismo-

gram $p(x_r, t; x_s)_{\text{cal}}$ differs by only a time shift from $p(x_r, t; x_s)_{\text{obs}}$,

$$p(x_r, t; x_s)_{\text{cal}} \approx p(x_r, t + \Delta\tau; x_s)_{\text{obs}}$$

This assumes that the amplitude differences are negligible. Therefore,

$$\begin{aligned} \dot{p}(x_r, t + \Delta\tau; x_s)_{\text{obs}} &\approx \frac{p(x_r, t + \Delta\tau; x_s)_{\text{obs}} - p(x_r, t; x_s)_{\text{obs}}}{\Delta\tau} \\ &\approx \frac{p(x_r, t; x_s)_{\text{cal}} - p(x_r, t; x_s)_{\text{obs}}}{\Delta\tau}. \end{aligned}$$

Substituting the above equation into equation (8b) yields

$$\delta\tau(x_r, t; x_s) \approx \frac{2}{E} [p(x_r, t; x_s)_{\text{obs}} - p(x_r, t; x_s)_{\text{cal}}] = \frac{2}{E} \delta p. \quad (13)$$

This result shows that $\delta\tau$ equals δp except for a factor of $2/E$; the WT inversion method is similar to full-wave inversion when the current velocity model is close to the true model. In practice, however, it is difficult to determine $\Delta\tau$ by crosscorrelation [equation (2)] when $\Delta\tau$ is small. This results in less velocity resolution compared to that from the full-wave inversion method. An optimal method might be to use the WT method to reconstruct a moderately coarse velocity model and

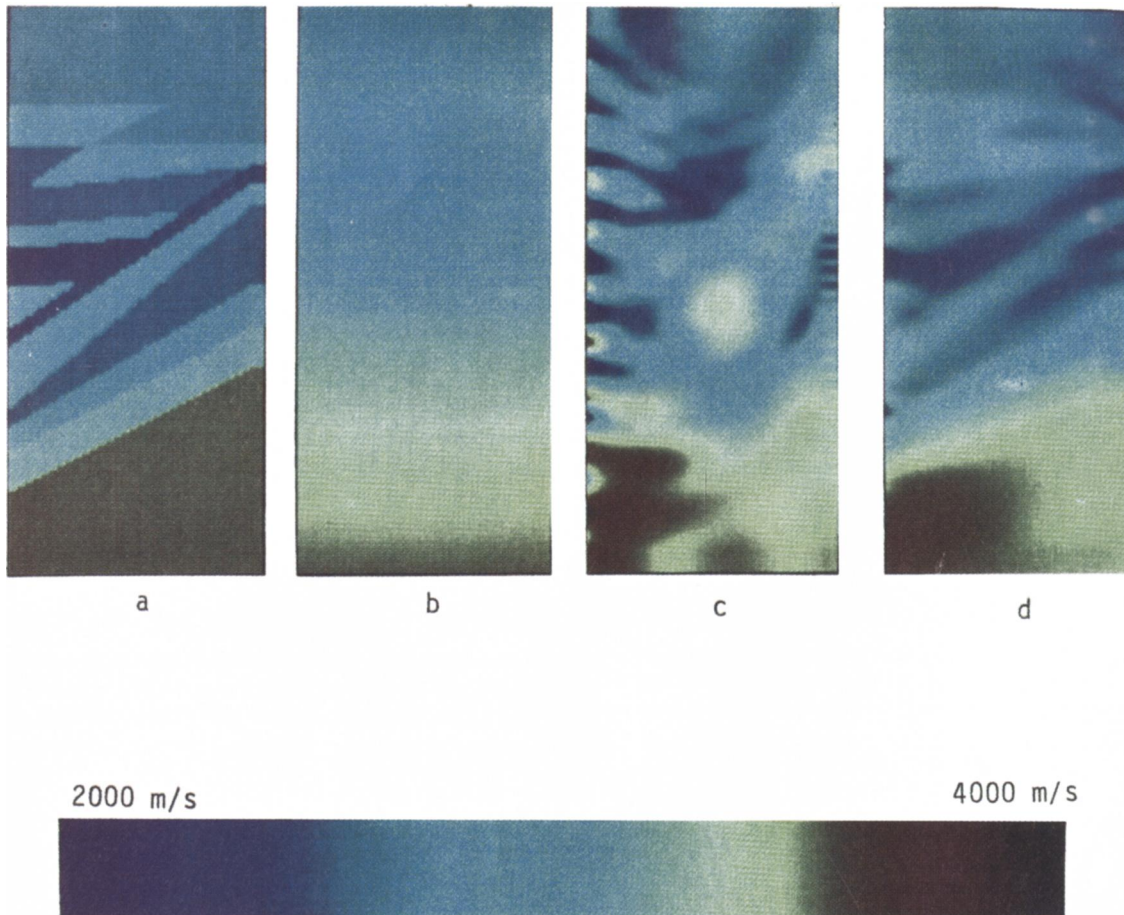


FIG. 1. (a) Dipping layer + fault model. (b) Initial starting model for both WT and full-wave inversion algorithms. (c) Velocity field reconstructed by full-wave inversion method after 20 iterations. Source and receiver parameters given in text. (d) Same as (c) except the WT method is used for the velocity reconstruction.

use full-wave inversion for the finer velocity details. This hybrid strategy is successfully exploited in Luo and Schuster (1989).

NUMERICAL EXAMPLES

The WT method is tested on three different crosswell data sets: synthetic crosswell data associated with a dipping layer and a fault model, synthetic crosswell data associated with an earth model (Langan velocity model) derived from a well log in Southern California (Langan et al., 1988), and real crosswell data collected by Exxon in Texas. The dipping layer + fault model is used to verify that WT inversion is more robust than full-wave inversion, and the Langan velocity model is used to show that WT inversion succeeds when ray tracing fails. The real data inversion is used to demonstrate that the WT method can successfully invert velocities from real data.

Crosswell fault model

Figure 1 illustrates a faulted crosswell velocity model where both the full-wave and WT methods are used to reconstruct the velocities from synthetic data. Figure 1a shows the true velocity model from which we calculate the synthetic seismograms by a finite-difference method. The vertical source and receiver wells are along the left and right margins of the model, the two wells are offset by 90 m, and the well depth is 210 m. The peak frequency of the Ricker source wavelet is 80 Hz, 21 sources are evenly spaced in the left well, and 36 geophones are evenly spaced in the right well. Figure 1b shows the initial velocity model where the velocity linearly increases with depth. Figure 1c shows the reconstructed velocity model after 20 iterations using a standard full-wave inversion (or nonlinear inversion) method. Figure 1d is the velocity model reconstructed by the WT method after 20 iterations. The average traveltime (≈ 1 ms) residual did not significantly decrease after 15 iterations. For this example, the full-wave method fails while the WT method provides an accurate velocity reconstruction.

Langan velocity model

Figure 2 (courtesy of R. Langan) is associated with an earth model derived from a sonic log in southern California. The crosswell configuration in Figure 2a consists of eight sources in the source well and 83 receivers spaced 4.8 m along the receiver well; the sonic log is given in Figure 2b. In this case, $\Delta x = \Delta z = 2.4$ m, $\Delta t = 0.5$ ms, and 700 time steps are calculated. Figure 2c shows the synthetic seismograms for the source at depth 168 m. To avoid aperture problems, the velocity is assumed to be known from the depths of 0–48 m and 350–400 m.

Langan et al. (1988) showed that a shooting ray-tracing method could not accurately compute the traveltimes in the shadow zones of the model, suggesting that a ray-tracing tomography algorithm may be inappropriate for a velocity reconstruction. Figure 3 shows the velocity profile reconstructed by the WT inversion method using a steepest descent method.

In the Langan velocity model, the density is kept constant ($4.0 \times 10^3 \text{ kg/m}^3$) for both forward modeling and inversion. In the Langan velocity-density model, the acquisition config-

uration and velocity are the same as that in the Langan velocity model (Figure 2b), except the density profile is that in Figure 4a. This density function was computed with a formula derived from well-log measurements (Gardner et al., 1974)

$$\frac{\rho}{\rho_0} = \left[\frac{c}{c_0} \right]^{1/4}, \quad (14)$$

where c is the velocity (Figure 2b), $c_0 = 2000$ m/s, and $1/\rho_0 = 2.5 \times 10^{-4} \text{ m}^3/\text{kg}$. In the inversion, the incorrect lightness profile in Figure 4b is used. Despite an incorrect assumption of homogeneous density, the WT method still achieves an accurate velocity reconstruction (Figure 4d) after 10 iterations.

Exxon crosswell data

Calnan and Schuster (1989) inverted the first arrival times from an Exxon crosswell data set using a ray tracing tomography algorithm. The crosswell geometry consisted of 96 evenly spaced downhole sources and receivers, 23 evenly spaced surface sources and receivers, the source and receiver well depths were 305.0 m, and the well offset was 183.0 m. The data and first arrival picks were of superb quality, partially due to the use of explosive sources. Figure 5 depicts a typical common shot point gather from the Exxon data set. Figure 6 shows the tomograms inverted from this crosswell data set.

Figure 6 compares the Calnan and Schuster ray-tracing tomogram on the right with the WT tomogram on the left. The ray-tracing and WT tomograms are quite similar, al-

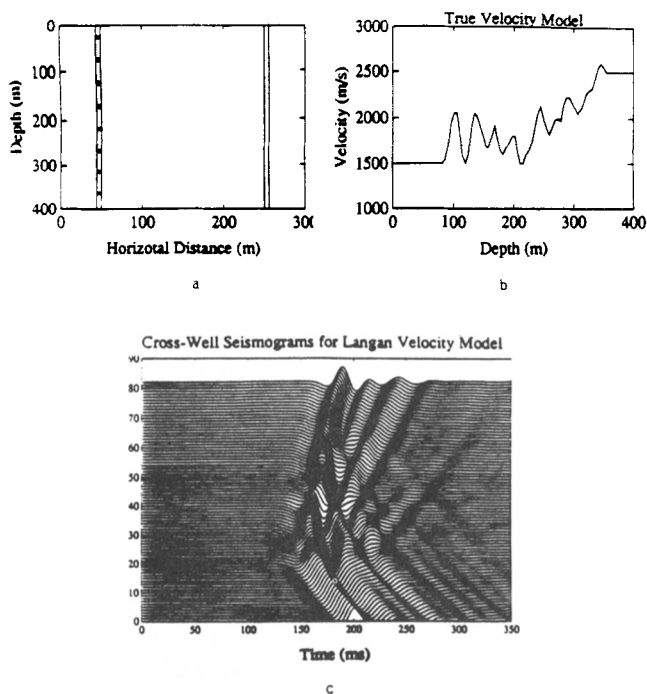


FIG. 2. (a) Crosswell geometry with eight sources evenly distributed along left well and 83 receivers evenly distributed along right well. (b) Sonic log, and (c) pressure seismogram for a source at depth 300 m.

though the WT tomogram appears to be smoother. The input data in each case consisted of the same picked traveltimes associated with the well-to-well (8104 traveltimes) and well-to-surface (4037 traveltimes) first arrivals. In these figures, darkest red corresponds to 2286.0 m/s, darkest blue corresponds to 1219.2 m/s, and the offset (depth) scale is 183.0 (305.0) m. The ray-tracing result required three iterations using a Gauss-Newton method with a conjugate gradient solver, while the WT method associated with Figure 6 used 9 iterations with a preconditioned steepest descent method. The ray-tracing algorithm used a slowness parameterization of a 51 by 31 grid of unknown slownesses, while the WT algorithm used a grid of 120 by 200 unknown velocities.

The associated WT traveltimes residuals after 0, 3, 6, and 8 iterations are shown in Figure 7. The CPU time per iteration is about 2 hours for each method, although the WT method used a 30 Mflop Stellar 2000 computer while the ray-tracing tomogram was calculated on a 20 Mflop Stellar 1000 com-

puter. Also, the ray-tracing code was considered to be inefficiently written.

CONCLUSION

A new seismic inversion method is presented which reconstructs velocities from traveltimes computed from solutions to the wave equation. No high-frequency assumptions about the data are needed, traveltimes picking and event identification are sometimes unnecessary, velocities are practically decoupled from densities, and the computer time is no more than that of full-wave inversion. Synthetic tests show that successful reconstructions can be achieved with models having large velocity contrasts. This is an improvement over standard full-wave inversion which can fail for velocity models with little more than 10 percent velocity contrast. Real data tests suggest that this method can be as accurate as ray-tracing tomography, and it can

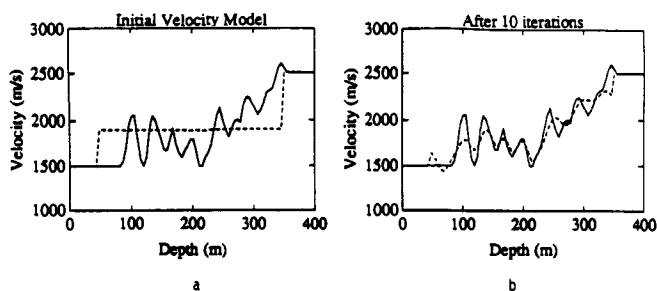


FIG. 3. (a) Actual velocity profile (solid) and initial velocity model (dashed line) for WT inversion. (b) Actual (solid line) and reconstructed (dashed line) velocity model.

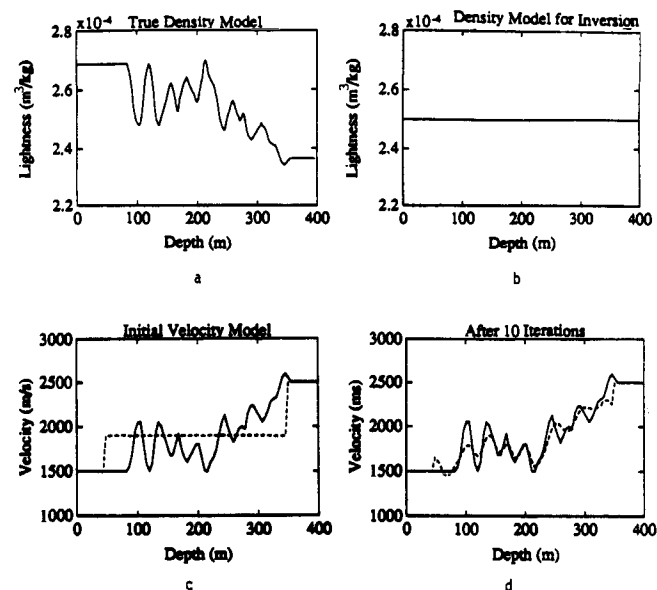


FIG. 4. (a) Density log computed from equation (14) and the velocity profile in Figure (2b). (b) Assumed density profile for WT inversion. (c) Actual (solid line) and reconstructed (dashed line) velocity profile. (d) Reconstructed velocity profile (dashed line) after 10 WT iterations compared to actual velocity profile (solid line).

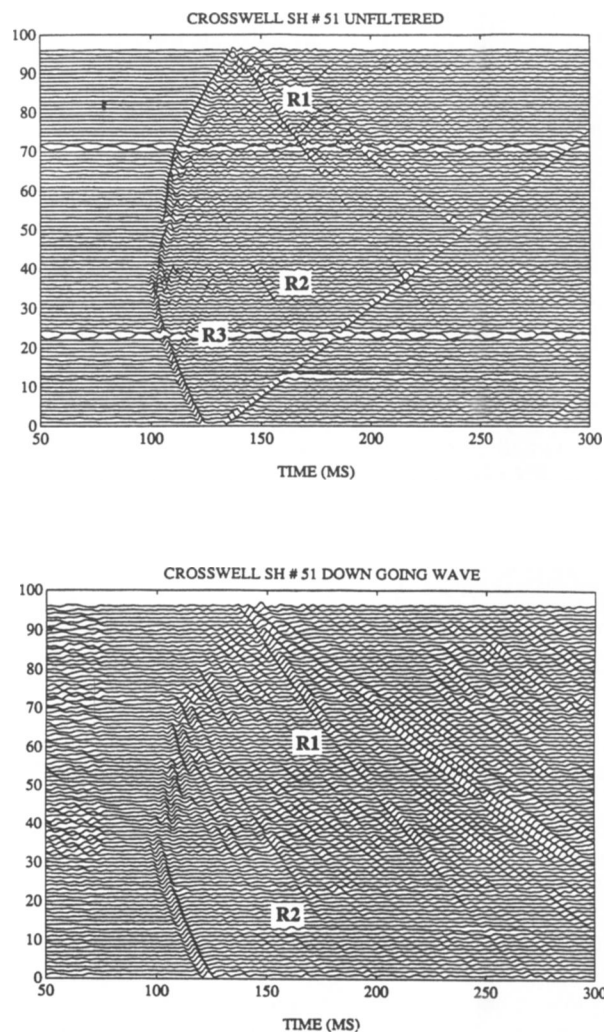


FIG. 5. Common shotpoint gather from shothole number 51 associated with the Exxon crosswell experiment described in the text. Top figure is unfiltered CSP gather and bottom figure is f - k filtered gather to highlight downgoing reflections. Reflections from interfaces are labeled R1, R2, and R3. Note the impulsive quality of the transmitted wavelets.

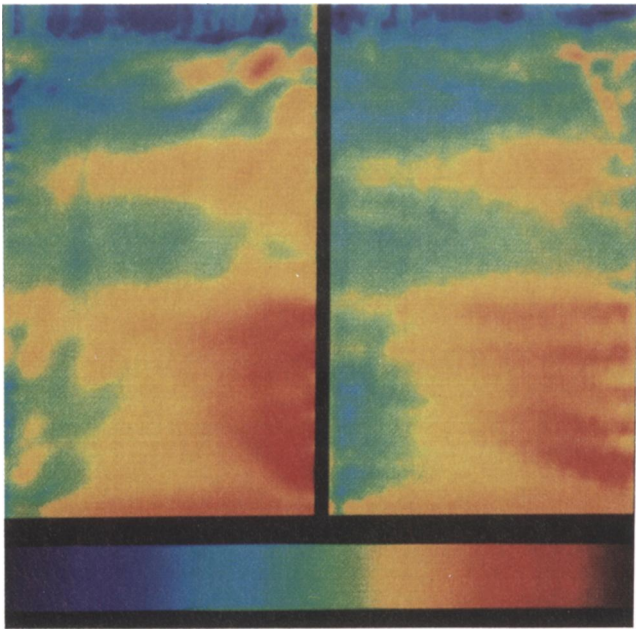


FIG. 6. WT tomogram on the left and ray-tracing tomogram on the right. Input data consist of over 12 000 traveltimes from the first arrivals of the Exxon data set.

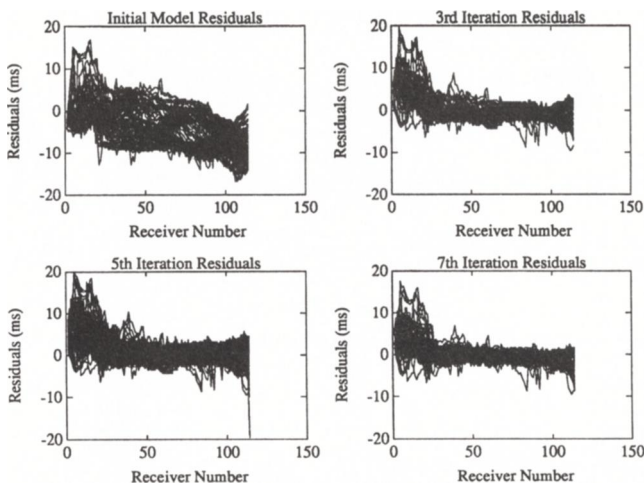


FIG. 7. WT traveltime residuals versus geophone number for 0, 3, 6, and 8 iterations. Note the variance of the residual decreases as iteration number increases.

offer the promise of incorporating reflection and diffraction information into the inversion.

ACKNOWLEDGMENTS

We thank the Gas Research Institute for sponsoring this research under contract 5089-215-1872. Neither GRI, members of GRI, nor any person acting on behalf of GRI assumes any liability in the use of information in this report. We are also grateful for the support of the 1989 University of Utah

tomography consortium members: Amoco, ARCO, British Petroleum, Chevron, Conoco, Exxon, GRI, Marathon, Mobil, Phillips, Texaco. Finally, we thank Sen Chen, Linda Zimmerman, and Exxon for providing the Exxon crosswell data set to us.

REFERENCES

- Beydoun, W., and Mendes, M., 1989, Elastic ray Born I_2 -migration/inversion, *Geophys. J.*, **97**, 151–160.
- 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.
- Bleistein, N., 1984, *Mathematical methods for wave phenomena*: Academic Press.
- Bleistein, N., and Gray, S., 1985, An extension of the Born inversion method to a depth dependent reference profile: *Geophys. Prosp.*, **33**, 999–1022.
- Carrion, P., and Foster, D., 1985, Inversion of seismic data using precritical reflection and refraction data: *Geophysics*, **50**, 759–765.
- Calnan, C., and Schuster, G. T., 1989, Transmission + reflection crosswell tomography: Presented at the 59th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 908–911.
- Clayton, R., and Stolt, R., 1981, A Born WKBJ inversion method for acoustic reflection data: *Geophysics*, **46**, 1559–1567.
- Dines, K., and Lytle, R., 1979, Computerized geophysical tomography: *Proc. IEEE*, **67**, 1065–1072.
- Gardner, G. H. F., Gardner, L. W., and Gregory, A. R., 1974, Formation velocity and density—the diagnostic basis of stratigraphic traps: *Geophysics*, **39**, 770–780.
- Gauthier, O., Verieux, J., and Tarantola, A., 1986, Two-dimensional nonlinear inversion of seismic waveforms: numerical results: *Geophysics*, **51**, 1387–1403.
- Ivansson, S., 1985, A study of methods for tomographic velocity estimation in the presence of low-velocity zones: *Geophysics*, **50**, 969–988.
- Jannane, M., Beydoun, W., Crase, E., Cao, D., Koren, Z., Landa, E., Mendes, M., Pica, A., Noble, M., Roeth, G., Singh, S., Snieder, R., Tarantola, A., Trezeguet, D., and Zie, M., 1989, Wavelength of earth structure that can be resolved from seismic reflection data: *Geophysics*, **54**, 906–910.
- Johnson, S., and Tracy, M. L., 1983, Inverse scattering solutions by a sinc basis, multiple source, moment method; part I, Theory; *Ultrasonic Imaging* **5**, 361–375.
- Justice, J., Vassiliou, A., Singh, S., Logel, J., Hansen, P., Hall, B., Hutt, P., and Solankil, J., 1989, Acoustic tomography for enhancing oil recovery: *The Leading Edge*, **8**, 12–19.
- Keys, R., and Weglein, A., 1983, Generalized linear inversion and the first Born theory: *J. Math. Phys.*, **24**, 1444–1449.
- Langan, R., Paulsson, B., Stefani, J., Finnstrom, E., and Fairborn, J., 1988, Cross-well seismology: A new production tool: Presented at the Joint SEG/CPS Meeting at the Daqing oil field.
- Lines, L., 1988, Inversion of geophysical data: Soc. Expl. Geophys., Reprint series no. 9.
- Lo, T., Toksöz, N., Shao-hui, X., and Wu, R., 1988, Ultrasonic tests of geophysical tomographic reconstruction: *Geophysics*, **53**, 947–956.
- Luo, Y., and Schuster, G., 1989, A hybrid traveltime + full waveform inversion method: 1989 Univ. Utah Ann. Seismic Tomography report.
- 1991, Wave equation inversion of skeletalized geophysical data: *Geophys. J.*, in press.
- Paulsson, B., Cook, N., and McEvelly, T., 1985, Elastic wave velocities and attenuation in an underground repository of nuclear waste: *Geophysics*, **50**, 551–570.
- Pica, A., Diet, J. P., and Tarantola, A., 1988, Practice of nonlinear inversion of seismic reflection data in a laterally invariant medium: *Geophysics*, **55**, 284–292.
- Tarantola, A., 1987, *Inverse problem theory*: Elsevier Sci. Publ.
- Virieux, J., 1984, Shear-wave propagation in heterogeneous media. velocity-stress finite-difference method: *Geophysics*, **49**, 1933–1942.
- Wattus, N., 1989, Inversion of ground roll dispersion for near-surface shear-wave velocity variations: Presented at the 59th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 946–948.
- Weglein, A., 1982, Multidimensional seismic analysis: Migration and inversion: *Geophys.*, **20**, 47–60.

APPENDIX A
FRECHÉT DERIVATIVE

To obtain the Frechét derivative of the pressure field with respect to the velocity [equation (7c)], we introduce the pressure field $p(x, t; x_s)$ which satisfies the wave equation

$$\frac{1}{c(x)^2} \frac{\partial^2 p(x, t; x_s)}{\partial t^2} - \rho(x) \nabla \cdot \left[\frac{1}{\rho(x)} \nabla p(x, t; x_s) \right] = s(t; x_s), \quad (\text{A-1})$$

The corresponding Green's function obeys

$$\frac{1}{c(x)^2} \frac{\partial^2 g(x, t; x', t')}{\partial t^2} - \rho(x) \nabla \cdot \left[\frac{1}{\rho(x)} \nabla g(x, t; x', t') \right] = \delta(x - x') \delta(t - t') \quad (\text{A-2})$$

$$g(x, t; x', t') = 0 \quad \dot{g}(x, t; x', t') = 0 \quad \text{for } (t \leq t').$$

A perturbation of velocity $c(x) \rightarrow c(x) + \delta c(x)$ will produce a field $p(x, t; x_s) + \delta p(x, t; x_s)$ which obeys

$$\frac{1}{[c(x) + \delta c(x)]^2} \frac{\partial^2 [p(x, t; x_s) + \delta p(x, t; x_s)]}{\partial t^2} - \rho(x) \nabla \cdot \left[\frac{1}{\rho(x)} \nabla [p(x, t; x_s) + \delta p(x, t; x_s)] \right] = s(t; x) \quad (\text{A-3})$$

$$p(x, 0; x_s) + \delta p(x, 0; x_s) = 0$$

$$\dot{p}(x, 0; x_s) + \delta \dot{p}(x, 0; x_s) = 0.$$

Using

$$\frac{1}{[c(x) + \delta c(x)]^2} \approx \frac{1}{c(x)^2} - \frac{2\delta c(x)}{c(x)^3},$$

and subtracting equation (A-2) from equation (A-3) gives

$$\frac{1}{c(x)^2} \frac{\partial^2 \delta p(x, t; x_s)}{\partial t^2} - \rho(x) \nabla \cdot \left[\frac{1}{\rho(x)} \nabla \delta p(x, t; x_s) \right] = \frac{\partial^2 p(x, t; x_s)}{\partial t^2} \frac{2 \cdot \delta c(x)}{c(x)^3} + O(\delta c(x)^2), \quad (\text{A-4})$$

$$\delta p(x, 0; x_s) = 0 \quad \delta \dot{p}(x, 0; x_s) = 0.$$

Using the Green's function, the solution of equation (A-4) can be written

$$\delta p(x_r, t; x_s) = \int_v dv(x') g(x_r, t; x', 0) * \ddot{p}(x', t; x_s) \frac{2 \cdot \delta c(x')}{c(x')^3}, \quad (\text{A-5})$$

where the asterisk denotes time convolution. Since the perturbation occurs only at one point, set

$$\delta c(x') = \Delta c \delta(x' - x).$$

Then equation (A-5) becomes

$$\delta p(x_r, t; x_s) = g(x_r, t; x, 0) * \ddot{p}(x, t; x_s) \frac{2\Delta c}{c(x)^3}.$$

Dividing by Δc on both sides, we get equation (7c):

$$\frac{\partial p(x_r, t; x_s)_{\text{cal}}}{\partial c(x)} = \frac{2}{c(x)^3} \dot{g}(x, t; x_r, 0) * \dot{p}(x, t; x_s). \quad (\text{A-6})$$

We use reciprocity to allow the exchange $x \leftrightarrow x_r$.

APPENDIX B
IMPLEMENTATION OF WT INVERSION

Forward modeling

In principle, one can use any forward modeling scheme which simulates wave propagation; we use a staggered grid finite-difference scheme (Virieux, 1984). To use this scheme, we rewrite equation (1) as two first order equations

$$\begin{aligned} \frac{\partial p(x_r, t; x_s)}{\partial t} &= c(x)^2 \rho(x) \nabla \cdot (w(x_r, t; x_s)) \\ &\quad + c(x)^2 \bar{s}(t; x), \\ \frac{\partial w(x_r, t; x_s)}{\partial t} &= \frac{1}{\rho(x)} \nabla p(x_r, t; x_s), \end{aligned} \quad (\text{B-1})$$

where w is the particle velocity vector and the initial conditions are given as

$$p(x_r, 0; x_s) = 0; \quad w(x_r, 0; x_s) = 0. \quad \text{for } (t \leq 0)$$

Here \bar{s} is

$$\bar{s}(t; x) = \int_0^t dt s(t; x),$$

where $s(t; x)$ is the source term in the second-order wave equation.

From equation (B-1), we can get $p(x_r, t; x_s)_{\text{cal}}$ and $p(x, t; x_s)$ which will be used in the time correlation with the reverse time propagation field $p'(x_r, t; x_s)$. As for the time correlation in equation (9), we need to multiply the field $p(x, t; x_s)$ by $p'(x_r, t; x_s)$. We can either choose to store the entire history of field $p(x, t; x_s)$ in the computer memory or to recalculate it, backward in time, simultaneously with the

calculation of the field $p'(x_r, t; x_s)$ (Gauthier et al., 1986). We chose the latter option. For recalculation of the field $p(x, t; x_s)$, we need to store the history of the field of $p(x, t; x_s)$ at the boundaries, and, of course, the final two states of the field.

Backward Propagation

From equation (9), $p'(x_r, t; x_s)$ should satisfy

$$\frac{1}{c(x)^2} \frac{\partial^2 p'(x_r, t; x_s)}{\partial t^2} - \rho(x) \nabla \cdot \left[\frac{1}{\rho(x)} \nabla p'(x_r, t; x_s) \right] = \delta\tau(x_r, t; x_s).$$

In the time correlation of equation (9), the field $\dot{p}'(x_r, t; x_s)$ is used so that the time derivative can be taken on both sides of the above equation:

$$\frac{1}{c(x)^2} \frac{\partial^2 \dot{p}'(x_r, t; x_s)}{\partial t^2} - \rho(x) \nabla \cdot \left[\frac{1}{\rho(x)} \nabla \dot{p}'(x_r, t; x_s) \right] = \dot{\delta}\tau(x_r, t; x_s). \quad (\text{B-2})$$

Again, to use a staggered finite-difference scheme, rewrite equation (B-2) as

$$\frac{\partial \dot{p}'(x_r, t; x_s)}{\partial t} = c(x)^2 \rho(x) \nabla \cdot [\dot{w}'(x_r, t; x_s)] + c(x)^2 \delta\tau(x_r, t; x_s), \quad (\text{B-3})$$

$$\frac{\partial \dot{w}'(x_r, t; x_s)}{\partial t} = \frac{1}{\rho(x)} \nabla \dot{p}'(x_r, t; x_s),$$

with initial condition

$$\dot{p}'(x_r, T; x_s) = 0; \quad \dot{w}'(x_r, T + 1/2\Delta t; x_s) = 0,$$

where, Δt is the discretized time-step interval used in the finite-difference method and T is the total recording length. The pseudoresidual $\delta\tau$ is calculated from equation (8b) and the $\Delta\tau$ is obtained from equation (2). Since this initial condition is an approximation, we need to attenuate the amplitudes at the end of each trace to make this approximation more reasonable.

Direction of updating the model

Instead of using a steepest gradient direction, we can use some modified direction for updating the model. In general this update scheme can be expressed as

$$c(x)_{k+1} = c(x)_k + \alpha_k \cdot \phi_k, \quad (\text{B-4})$$

where the steepest descent method uses

$$\phi_k = \gamma(x)_k,$$

where $\gamma(x)$ is the negative gradient of the misfit function S given by equation (9). Another modification is to use a preconditioned gradient direction

$$\phi_k = \beta(x)_k = \gamma(x)_k \|x - x_s\|^{1/2} \cdot \|x - x_r\|^{1/2}.$$

This preconditioning compensates for geometrical expansion (Beydoun and Mendes, 1989). Of course, one can use the well known conjugate gradient direction

$$\phi_k = \beta(x)_k + \lambda(\phi_{k-1}), \quad k \geq 2$$

where

$$\lambda = \frac{[\beta_k]^t \cdot [\gamma_k]}{[\beta_{k-1}]^t \cdot [\gamma_{k-1}]}$$

and superscript t indicates matrix transpose.

Calculation of the step length

Pica et al. (1988) gives a formula for the estimation of step length α_k in equation (B-4). The final formula is

$$\alpha_k = \frac{[\phi_k]^t [\gamma(x)_k]}{[F\phi_k]^t [F\phi_k]}, \quad (\text{B-5})$$

where

$$[\phi_k]^t [\gamma(x)_k] = \sum_x [\phi_k \gamma(x)_k]$$

and

$$[F\phi_k] = \frac{g[c(x) + \epsilon\phi_k] - g[c(x)]}{\epsilon} = \frac{\delta p(x_r, t; x_s)}{\epsilon}.$$

$g[c(x)]$ implies forward modeling to get seismograms for the velocity model $c(x)$,

$$[F\phi_k]^t [F\phi_k] = \sum_r \int dt \left[\frac{\delta p(x_r, t; x_s)}{\epsilon} \right]^2,$$

where ϵ is estimated by

$$\max \{ \epsilon \cdot \phi_k \} \leq \frac{\max \{ c(x)_k \}}{100}.$$