GEOPHYSICS, VOL. 59, NO. 4 (APRIL 1994); P. 577–590, 11 FIGS., 1 TABLE.

Nonlinear velocity inversion by a two-step Monte Carlo method

Side Jin* and Raul Madariaga‡

ABSTRACT

Seismic reflection data contain information on small-scale impedance variations and a smooth reference velocity model. Given a reference velocity model, the reflectors can be obtained by linearized migration-inversion. If the reference velocity is incorrect, the reflectors obtained by inverting different subsets of the data will be incoherent. We propose to use the coherency of these images to invert for the background velocity distribution. We have developed a two-step iterative inversion method in which we separate the retrieval of small-scale variations of the seismic velocity from the longer-period reference velocity model. Given an initial background velocity model, we use a waveform misfit-functional for the inversion of small-scale velocity variations. For this linear step we use the linearized migration-inversion method based on ray theory that we have recently developed with Lambaré and Virieux. The reference velocity model is then updated by a Monte Carlo inversion method. For the nonlinear inversion of the velocity background, we introduce an objective functional that measures the coherency of the short wavelength components obtained by inverting different common shot gathers at the same locations. The nonlinear functional is calculated directly in migrated data space to avoid expensive numerical forward modeling by finite differences or ray theory. Our method is somewhat similar to an iterative migration velocity analysis, but we do an automatic search for relatively large-scale 1-D reference velocity models. We apply the nonlinear inversion method to a marine data set from the North Sea and also show that nonlinear inversion can be applied to realistic scale data sets to obtain a laterally heterogeneous velocity model with a reasonable amount of computer time.

INTRODUCTION

In conventional seismic processing, a very accurate estimate of the background velocity is necessary for obtaining good images; but, unfortunately, the derivation of an optimal velocity model is a difficult problem. Traditionally, background velocity is obtained from normal moveout (NMO) velocity analysis, based on the assumption of vertical velocity stratification with flat, horizontal reflectors. Obviously, this assumption makes the velocity analysis very restrictive. In fact, Lynn and Claerbout (1982) showed that the process of converting NMO and/or stacking velocity into interval velocity is unstable for layers with lateral velocity variation. To obtain velocity information for media with lateral velocity variation, kinematic methods such as traveltime tomography can be used (Bishop et al., 1985; Chiu et al., 1986). However, in practice, traveltime tomography suffers from a serious shortcoming: it requires accurate prestack reflection traveltime picking.

To be successful, the process of stacking and migration requires good background velocity information. Migrated stacks can conversely be used to estimate the background velocity model as has been proposed by many authors. Yilmaz and Chambers (1984) and Faye and Jeannot (1986), for example, proposed migration velocity analysis schemes based upon wavefront focusing. Kim and Gonzalez (1991) presented an implementation of the Kirchhoff integral which makes this approach more practical. Al-Yahya (1989) proposed another approach based on the principle that after migration with a correct velocity model, images in a common-receiver gather (CRG) should be aligned horizontally, regardless of structure. A CRG consists of seismic traces that have the same receiver coordinate. A detailed compar-

Manuscript received by the Editor March 19, 1992; revised manuscript received February 26, 1993.
*Elf GRC, 4th floor, 114A Cromwell Road, London SW7 4EU, United Kingdom.
‡Laboratoire de Sismologie, Institut de Physique du Globe de Paris et Université Paris 7, 4 Place Jussieu, Tour 14, F-75252 Paris Cedex 05, France.

ison of several migration velocity analysis methods can be found in Versteeg (1991). The drawback of migration velocity analysis is that it is usually interactive, requiring human intervention at each iteration step. The reasons for this are twofold: first, prestack migration is expensive; and, second, the criteria used for comparing the fit obtained with different velocity models are qualitative and nonobjective.

To avoid human intervention in velocity estimation, several approaches to nonlinear velocity inversion have been proposed in the literature. Recently, Snieder et al. (1989) and Cao et al. (1990) proposed to invert simultaneously both the short-wavelength impedance variations and the smooth background velocity by minimizing a least-squares misfit functional for the waveforms. If the velocity structure is known, the short wavelength impedance can be obtained by solving the inversion problem with a gradient algorithm. Gradient methods, however, are slow because they tend to converge on secondary minima. To find the reference velocity, Snieder et al. (1989) and Cao et al. (1990) used a relaxation method to update the impedance and velocity models in successive steps. The relaxation method was successfully used for inverting 1-D velocity backgrounds. In practical situations, however, the relaxation method is computationally expensive because the computation of the waveform misfit-functional requires numerous forward seismic modelings by finite differences or other numerical methods. More recently, to avoid the severe convergence difficulties associated with nonlinear least-squares inversion, Symes and Carrazone (1991) proposed a differential semblance optimization method (DSO) to invert reflection seismograms. This method exploits both moveout and amplitude characteristics of reflections. Although this method is based on the ideas underlying ordinary velocity analysis, it cannot avoid performing forward seismic modelings. So it is not clear if DSO can yield a computationally tractable approach for 2-D problems.

In many nonlinear inversion problems (for instance, simultaneous depth and moment tensor inversion from surface waves, Romanowicz, 1982), it is convenient to separate the linear from the nonlinear inversion stages. The reason is that the nonlinear inversion can be formulated with a relatively small number of parameters, while the linear part may have a huge number of degrees of freedom. In this paper, we separate linearized migration-inversion for small-scale impedance contrasts from fully nonlinear velocity analysis. It turns out that this two-step inversion method is very natural for seismic reflection profiles which have been processed separately for velocity analysis and migration. We design the two-step inversion method using different objective functionals to determine the short wavelength variations and the background velocity. For a given background velocity model, impedance contrasts are obtained by linear inversion optimizing a waveform-fit functional. Since we use a very fast algorithm for the linear step, we can use a fully nonlinear Monte Carlo inversion method for the retrieval of the background velocity model. The optimization criterion for the background velocity is based on a measure of the semblance of the spatial short wavelength component of the model obtained by inverting different subsets of the reflection data. Because our measure of the coherency is calculated in model space instead of data space, we do not need expensive forward seismic modeling to calculate coherency.

In view of the adopted optimization criterion, our method is similar to migration velocity analysis, but it explores the background velocity space automatically with a Monte Carlo technique. By an application to data from a marine seismic profile, we show that our method is practical, in the sense that it can be applied to field data sets and produce reasonable results with present-day computers.

THE PRINCIPLE OF NONLINEAR VELOCITY INVERSION

We study the nonlinear inversion of the acoustic parameters of a 2-D earth model with constant density. The only heterogeneity allowed in this model is the variation of the compressional-velocity $c(x, z)$, where x and z are the spatial coordinates of the model. We assume as usual that the velocity field can be adequately decomposed into two parts:

$$c(x, z) = c_0(x, z) + \delta c(x, z), \tag{1}$$

where $c_0(x, z)$ is a smooth *large-scale velocity* field and $\delta c(x, z)$ is a *small-scale perturbation* superimposed on this background. In standard seismic processing $c_0(x, z)$ is determined by velocity analysis and $\delta c(x, z)$ is determined by migration or linearized inversion. The information about $\delta c(\mathbf{x})$ is contained in the scattered waveforms, whereas the background velocity $c_0(\mathbf{x})$ manifests itself mainly in the traveltime of reflected waves and in the curvature of the reflection hyperbolas.

The separation of the velocity field into a smooth and a small-scale perturbation may seem somewhat arbitrary. As discussed by Claerbout (1985), this separation occurs naturally because of the finite frequency band of the sources used in applied geophysics. For wavelengths longer than the dominating frequency of the source, the heterogeneities appear to be smooth and their main effect is to modify the propagation velocity. Shorter wavelength heterogeneities, on the other hand, scatter the seismic waves and produce reflections from laterally coherent inhomogeneities. A complete discussion of this separation can be found in Jannane et al. (1989) who studied the separation by finite difference modeling. In this paper we will exploit the separation between short- and long-wavelength heterogeneities to formulate a two-step nonlinear inversion method.

Given a certain reference velocity model, an image of the reflectors defined by the velocity perturbation field $\delta c(\mathbf{x})$ can be constructed by linearized inversion, for instance, by the asymptotic inversion method proposed by Jin et al. (1992). In linearized inversion, the strength, or reflectivity, of a reflector is determined by the waveforms, and location is determined by the traveltime of the waves in the current background velocity model. Because of the redundancy of reflection data, if the background velocity $c_0(x, z)$ is good, perturbation models $\delta c(\mathbf{x})$ obtained from different subsets of the seismic data (for example, different shot gathers) should be kinematically the same, a property we have verified experimentally on seismic data from the North Sea (Lambaré et al., 1992). Conversely, if the background is incorrect, the same reflector will be mapped to different positions by different subsets of the data. Thus, the location of a reflector inverted from different data subsets contains

the information needed to obtain the background velocity field. As proposed by Al-Yahya (1989), by optimization of the coherency of the small-scale models $\delta c(\mathbf{x})$ obtained from different subsets, it should be possible to obtain the background velocity model. This principle has in fact always been used to perform velocity analysis; but in the classical procedures the background velocity is usually assumed to be homogeneous and the reflectors to be horizontal and infinitely extended laterally.

In the following we shall assume that the data set is redundant enough so that several short-wavelength models $\delta c(x, z)$ can be determined from selected subsets of the data. To apply a nonlinear inversion technique based on the coherency of the small-scale velocity models inverted from different data subsets, we perform inversion in two steps. In the first step, given an initial background velocity model, we use linear inversion to calculate δc from several data subsets for background velocity. In the second step, we perform nonlinear inversion for the background velocity model by a Monte Carlo technique. For the second step, we need an objective criterion to quantify the coherency of the models obtained from different data subsets.

THE ASYMPTOTIC INVERSION METHOD

For the linearized inversion of the short-wavelength perturbations we use the method proposed by Jin et al. (1992) in the elastic approximation and by Lambaré et al. (1992) for the acoustic case. We will invert the small-scale velocity perturbation field from the common source gather (CSG) for every available source s.

We will briefly review the linearized inversion technique for reference. For simplicity, we start from the scalar wave equation for a line source. The extension to a point source in a medium infinitely extended in the direction perpendicular to the seismic profile (sometimes called the 2.5-D approximation) can be found in Lambaré et al. (1992).

$$\nabla^2 u(\mathbf{x}, \mathrm{s}, \omega) + \frac{\omega^2}{c^2(\mathbf{x})} u(\mathbf{x}, \mathrm{s}, \omega) = \delta(\mathbf{x} - \mathrm{s}), \tag{2}$$

where $u(\mathbf{x}, \mathrm{s}, \omega)$ is the acoustic pressure field, $\mathbf{x}$ is the current field point, s is the source position, ω is the frequency, and $c(\mathbf{x})$ is the variable acoustic velocity. Density is assumed to be constant in our inversions, because as has been extensively discussed in the literature, only velocity or impedance can be inverted for near-vertical reflection data. To invert for density simultaneously with velocity, we would need either wide-angle reflections or multicomponent data. As in the velocity decomposition (1), we split the pressure field into two parts

$$u(\mathbf{x}, \mathrm{s}, \omega) = u_0(\mathbf{x}, \mathrm{s}, \omega) + \delta u(\mathbf{x}, \mathrm{s}, \omega), \tag{3}$$

where $u_0(\mathbf{x}, \mathrm{s}, \omega)$ is the pressure field propagated in the background model $c_0(x, z)$ and $\delta u(\mathbf{x}, \mathrm{s}, \omega)$ is the acoustic field scattered by the small-scale perturbation $\delta c(x, z)$ of the model.

Within the first Born approximation, the scattered field $\delta u(\mathbf{x}_r, \mathrm{s}, \omega)$ at the receiver position $\mathbf{x}_r$ is given by

$$\delta u(\mathbf{x}_r, \mathrm{s}, \omega) = -\omega^2 \int_{\mathcal{M}} G(\mathbf{x}_r, \mathbf{y}, \omega) G(\mathbf{y}, \mathrm{s}, \omega) f(\mathbf{y})\, d\mathbf{y}, \tag{4}$$

where $G(\mathbf{x}, \mathrm{s}, \omega)$ is the Green's function for the background velocity model, i.e., the solution of equation (2) when $c(\mathbf{x})$ is replaced by $c_0(\mathbf{x})$. $\mathcal{M}$ is the model space. For convenience we write $\mathbf{x}_r$, $\mathbf{y}$, and ω as continuous variables. Their integrals can in all cases be replaced by a sum over sampled data points. For simplicity of notation in equation (4) we introduce the model perturbation function

$$f(\mathbf{y}) = \delta[c^{-2}(\mathbf{y})] = -2 \frac{\delta c(\mathbf{y})}{c^3(\mathbf{y})}. \tag{5}$$

Equation (4) is obtained by standard perturbation methods (Tarantola, 1984; Beylkin, 1985; Miller et al., 1987). Mathematically, $f(\mathbf{y})$ represents the short-wavelength part of the velocity field $c(\mathbf{y})$, so that only this part of the field can be obtained by the linearized inversion of the data. The low frequency part of the velocity field $c_0(\mathbf{y})$ affects mainly the traveltimes and amplitudes implicit in the calculation of G.

We assume, following Beylkin (1985) that the Green's function G in equation (4) may be accurately estimated by the first-order geometrical optics approximation. In 2-D, the ray approximation to G is given by

$$G(\mathbf{x}, \mathrm{s}, \omega) = (-i\omega)^{-1/2} A(\mathbf{x}, \mathrm{s}) e^{i\omega\tau(\mathbf{x}, \mathrm{s})}, \tag{6}$$

where the traveltime function τ satisfies the eikonal equation for the background velocity c_0 and the amplitude A satisfies the transport equation along the ray connecting the source at s to the current point $\mathbf{x}$.

In linearized inversion, given a certain reference velocity model c_0, we view expression (4) as a linear integral equation for the unknown function $f(\mathbf{x})$. This problem is very ill-posed unfortunately and can be solved directly only by making a number of approximations such as assuming a continuous distribution of sources and receivers (Beylkin, 1985; Beylkin and Burridge, 1990). We can avoid such assumptions using the generalized inversion methods introduced by Tarantola (1984). Following Jin et al. (1992), we pose the inverse problem as follows: given a fixed source point s, find the model perturbation $f(\mathbf{x})$ that minimizes the following objective functional:

$$J_1(f, \mathbf{x}|\mathrm{s}) = \frac{1}{2\pi} \int_r d\mathbf{x}_r \int d\omega q(\mathbf{x}, \mathrm{s}, \mathbf{x}_r, \omega) \times [\delta u_{obs}(\mathbf{x}_r, \mathrm{s}, \omega) - \delta u_{cal}(\mathbf{x}_r, \mathrm{s}, \omega)]^2, \tag{7}$$

where the integral extends over all available receivers $\mathbf{x}_r$ for the present source point s. $\delta u_{obs}(\mathbf{x}_r, \mathrm{s}, \omega)$ are the observed reflection data and $\delta u_{cal}(\mathbf{x}_r, \mathrm{s}, \omega)$ are the seismograms computed by equation (4) for the current perturbation model $f(\mathbf{x})$ and background velocity $c_0(\mathbf{x})$. The weight function $q(\mathbf{x}, \mathrm{s}, \mathbf{x}_r, \omega)$ was defined by Jin et al. (1992) as

$$q(\mathbf{x}, \mathrm{s}, \mathbf{x}_r, \omega) = \frac{|\mathbf{p}(\mathrm{s}, \mathbf{x}, \mathbf{x}_r)|^2 \mathcal{J}(\mathbf{p}; \mathbf{x}_r, \mathrm{s}, \mathbf{x})}{(2\pi) A^2(\mathrm{s}, \mathbf{x}, \mathbf{x}_r)\omega}, \tag{8}$$

where $\mathbf{p}$ is the total slowness vector or gradient $\mathbf{p}(\mathrm{s}, \mathbf{x}, \mathbf{x}_r) = \nabla_x \tau(\mathrm{s}, \mathbf{x}, \mathbf{x}_r)$ of the two way traveltime $\tau(\mathrm{s}, \mathbf{x}, \mathbf{x}_r) = \tau(\mathrm{s}, \mathbf{x})$

$+ \tau(\mathbf{x}, \mathbf{x}_r)$ from the source at s to the scatterer at x and then to the receiver located at $\mathbf{x}_r$. $A(\mathbf{s}, \mathbf{x}, \mathbf{x}_r) = A(\mathbf{s}, \mathbf{x})A(\mathbf{x}, \mathbf{x}_r)$ is the product of the amplitudes along these two ray segments. $\mathcal{J}(\mathbf{p}; \mathbf{x}_r, \mathbf{s}, \mathbf{x})$ is a Jacobian relating the receiver position $\mathbf{x}_r$, the source s and the point x where we want to estimate the model to the total slowness vector $\mathbf{p}(\mathbf{s}, \mathbf{x}, \mathbf{x}_r)$ (see Figure 1).

If we use the operator notation $\underset{\sim}{\mathbf{P}}$ to represent the forward operator defined by equation (4)—i.e., we rewrite it in the concise form $\delta\mathbf{u} = \underset{\sim}{\mathbf{P}}\mathbf{f}$—the least-squares solution of equation (7) satisfies the following normal equation:

$$\underset{\sim}{\mathbf{P}}^{\dagger} q \underset{\sim}{\mathbf{P}}\mathbf{f} = \underset{\sim}{\mathbf{P}}^{\dagger} q \delta\mathbf{u}_{\text{obs}}, \tag{9a}$$

where $\underset{\sim}{\mathbf{P}}^{\dagger}$ denotes the adjoint of the operator $\underset{\sim}{\mathbf{P}}$. If we define the Hessian operator $\underset{\sim}{\mathbf{H}} = \underset{\sim}{\mathbf{P}}^{\dagger} q \underset{\sim}{\mathbf{P}}$, then

$$\underset{\sim}{\mathbf{H}}\mathbf{f} = \underset{\sim}{\mathbf{P}}^{\dagger} q \delta\mathbf{u}_{\text{obs}}. \tag{9b}$$

In general, it is not possible to solve this equation exactly because of the huge number of variables in the Hessian. However, as shown by Jin et al. (1992), by a judicious choice of the weight function $q(\mathbf{x}, \mathbf{s}, \mathbf{r}, \omega)$, the Hessian $\underset{\sim}{\mathbf{H}}$ at the point x can be approximately diagonalized within the theory of asymptotic Fourier integrals introduced by Beylkin (1985). As shown in the appendix, within the frequency band of the source signal the Hessian $\underset{\sim}{\mathbf{H}}$ is approximately the identity operator

$$\underset{\sim}{\mathbf{H}} = \underset{\sim}{\mathbf{I}} \tag{10}$$

at x and the approximate solution of the inverse problem (7) is simply:

$$\mathbf{f} = \underset{\sim}{\mathbf{P}}^{\dagger} q \delta\mathbf{u}_{\text{obs}}, \tag{11}$$

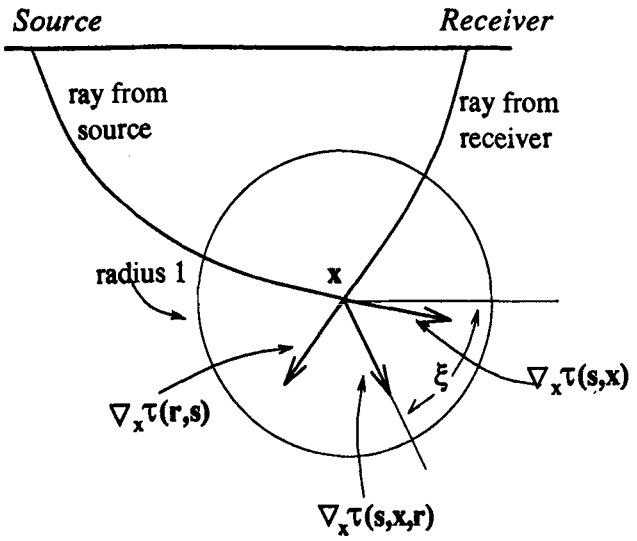

FIG. 1. Ray geometry in the vicinity of a diffraction point for the linearized inversion of a CSG. A ray from the fixed source arrives at this point and is scattered toward the receiver. The receiver position can be parameterized either by its coordinate along the surface or by the angle ξ of the direction of the total traveltime gradient $\nabla_x\tau(\mathbf{s}, \mathbf{x}, \mathbf{r})$.

which can be written explicitly in the time domain as:

$$f(\mathbf{x}|\mathbf{s}, c_0) = \frac{1}{2\pi} \int_r d\mathbf{x}_r \mathcal{J}(\mathbf{p}; \mathbf{x}_r, \mathbf{s}, \mathbf{x}) \frac{p^2(\mathbf{s}, \mathbf{x}, \mathbf{x}_r)}{A(\mathbf{s}, \mathbf{x}, \mathbf{x}_r)} \times \mathcal{H}[\delta u_{obs}(\mathbf{s}, \mathbf{x}_r, \tau(\mathbf{s}, \mathbf{x}, \mathbf{x}_r)], \tag{12}$$

where we use the notation $f(\mathbf{x}|\mathbf{s}, c_0)$ to emphasize that equation (12) computes a small-scale model for a fixed source position s and for a given background velocity model c_0. $\mathcal{H}$ denotes the Hilbert transform with respect to the time variable. Equation (12) is a modified prestack depth migration with a dynamic correction for amplitude and phase. A similar equation was found by Beylkin (1985) and Miller et al. (1987), assuming a continuous distribution of receivers r. As shown by Jin et al. (1992), the approximation of the Hessian (10) is not exact, so that expression (12) should be considered only as the first iteration in a quasi-Newton iterative procedure. However, for our purposes of estimating the background velocity, a single iteration is good enough. Once the background has been inverted, it is still possible to improve the perturbation model by further iterations. The traveltime τ, the Jacobian $\mathcal{J}$, the slowness p and the ray amplitude A that appear in equation (12) are calculated by ray tracing in the current background velocity model (see Lambaré et al., 1992).

THE COST FUNCTIONAL FOR NONLINEAR INVERSION

Nonlinear inversion is generally done by maximizing or minimizing an objective functional that measures the difference between a certain function of the observed data and the same function calculated from a model. A model is considered to be good if it can predict the data with the least possible error. The cost functional used in nonlinear velocity inversion by most authors (e.g., Snieder et al., 1989; Cao et al., 1990) is a least-squares waveform misfit-functional defined directly from the observed seismograms. A practical shortcoming of this function is that it requires computationally expensive forward seismic modeling to evaluate it. Another disadvantage is that the misfit-functional is very flat for bad models and it has deep, narrow minima near the good models. Most linearized iterative inversion methods converge extremely slowly, and convergence to local minima is frequent. Thus a vast amount of seismic modeling is unavoidable if this method is used to obtain a velocity model, so that velocity inversion using this misfit function is difficult to envisage in practical situations.

To avoid forward seismic modeling, we propose an objective function calculated directly in migrated data space. Inspired by the iterative migration velocity analysis of Al Yahya (1989), we define a cost functional that is computed from the data migrated by our asymptotic linearized inversion method. The objective functional we propose is

$$J_2(c_0) = \sum_x a(x) \sum_z \left| \sum_s f(x, z|\mathbf{s}, c_0) \right|, \tag{13}$$

where c_0 is the background velocity model used for the asymptotic linearized inversion-migration of each common shot gather (CSG). For each source s, we calculate a small-scale velocity model $f(x, z|\mathbf{s}, c_0)$ at a set of surface locations x and depths z. For each surface point x we can

construct an Iso-X gather, aligning the different models f obtained for each source point s. Several of these Iso-X gathers are shown in Figure 7. As suggested by the synthetic experiments of Al Yahya (1989) a correct background velocity model is one that makes the reflection events in each Iso-X gather flat. It is clear that the sum of an Iso-X gather over the sources s is a measure of the alignment of the reflection events and, therefore, of the validity of the background velocity model used to generate the Iso-X gather.

The objective functional J_2 that we propose is obtained summing each Iso-X gather over all sources and depths. Finally, we sum over x using a weighting coefficient $a(x)$. If the background velocity model is correct, $J_2(c_0)$ attains a global maximum, because for a surface location x, all $f(x, z|\mathbf{s}, c_0)$ are summed constructively. Thus the cost functional J_2 provides a measure of the coherency of the small-scale models f. The velocity model which maximizes the cost functional J_2 is considered to be the best model.

Because the number of reflectors may be different for different surface locations x, we use a weighting coefficient $a(x)$ in the objective function. $a(x)$ is chosen so that $a(x) \sum_z |\sum_s f(x, z|\mathbf{s}, c_0)|$ are nearly independent of x. In our implementation, we obtain this as follows: at the beginning, we set $a(x) = 1$. When we have obtained a "good enough" model (tenth accepted model, for example), we set $a(x) = 1./\sum_z |\sum_s f(x, z|s, c_0)|$. Finally we use this $a(x)$ to continue the velocity-optimization process. We can do this more than once in the course of the optimization if necessary.

Let us remark that unlike the norm $J_1(f, \mathbf{x}|\mathbf{s})$ defined in equation (7), which is an $\mathcal{L}^2$-norm (quadratic), J_2 is an $\mathcal{L}^1$ (absolute value) norm. As shown by Crase et al. (1990), $\mathcal{L}^1$ norms are usually more robust and can handle noisy data better than $\mathcal{L}^2$ norms. For the linear step we kept an $\mathcal{L}^2$-norm because it gives the simple migration-inversion algorithm (12). For the nonlinear step we are free to choose the more robust $\mathcal{L}^1$-norm, since the inversion is entirely done numerically.

MONTE CARLO NONLINEAR INVERSION OF BACKGROUND VELOCITY

The goal of the inversion method is to choose a smooth velocity model which maximizes the objective functional J_2 defined in equation (13). This goal is achieved by iteratively changing the background velocity model c_0 by a Monte Carlo random walk in model space and calculating $f(x, z|\mathbf{s}, c_0)$ from different CSGs. Then we calculate $J_2(c_0)$ and compare it to its previous values. The process is repeated until the Monte Carlo method cannot find a velocity model that increases the value of $J_2(c_0)$. We restrict the model space to realistic velocity models, and the more *a priori* information we have about the model we are looking for, the smaller the space of models we have to explore with the Monte Carlo method. For instance, if we know the velocity at some locations (at the earth surface, for example), we fix the velocities of these points to reduce the number of degrees of freedom of the problem. We chose the Monte Carlo method for background velocity inversion because, as suggested by Landa et al. (1989), this is an efficient method for the global optimization of a nonlinear functional. Monte Carlo techniques are of course very expensive if the forward model contains many parameters. Thanks to our separation of small- and large-scale velocity inversion, the number of model parameters used in the nonlinear inversion step can be reduced to a few B-spline coefficients. Combined with an efficient ray-tracing technique, we found that it is quite possible to use the Monte Carlo method to maximize the cost functional J_2. Since the computation of J_2 does not require human intervention, our process is automatic.

The flow diagram for nonlinear reference velocity inversion is depicted in Figure 2. Starting from an initial velocity model $c_0 = c_a$, we choose at random a direction $\mathbf{r}$ in model space and define a new model by

$$c_b = c_a + dc\ \mathbf{r}. \tag{14}$$

The step in velocity model space dc must be chosen sufficiently large so that the Monte Carlo method can rapidly explore all the interesting parts of the model space. Using the criterion J_2, we decide whether to keep or reject each successive model. The complete cycle is restarted again by choosing a new random direction $\mathbf{r}$ and using equation (14) to compute a new trial velocity model. The index n counts the

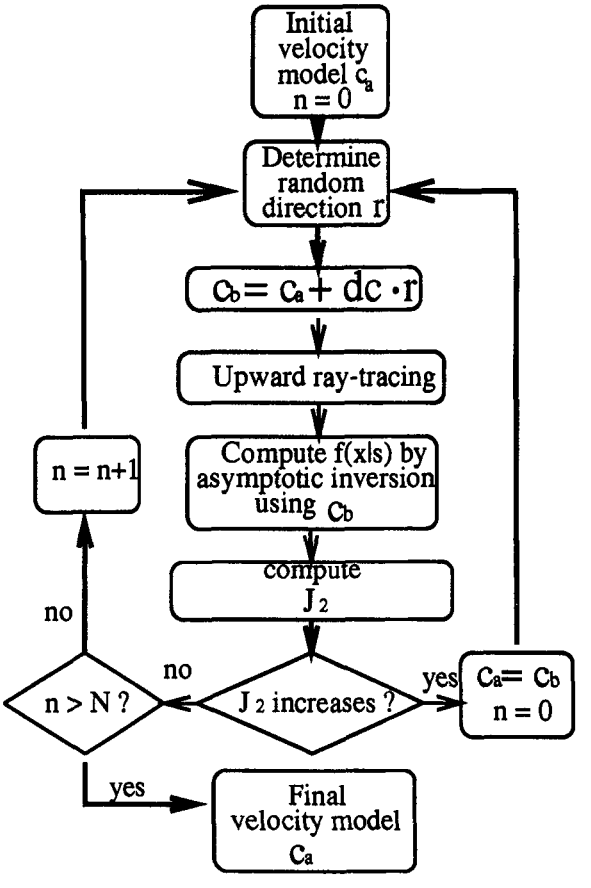

FIG. 2. Flow diagram for the nonlinear Monte Carlo inversion method. Starting from an initial background velocity model $\mathbf{c}_a$, we do a linear inversion for the small-scale velocity perturbation $f(\mathbf{x}|\mathbf{s}, \mathbf{c}_a)$. From this we calculate the cost functional J_2 and decide whether to keep or delete this model. The process starts again and continues until a number $n > N$ models have been tested without increase in the value of J_2. The maximum number of random steps N is chosen by the user.

number of trials that fail to find an acceptable model after the last acceptance. If this number is greater than a given number N, the random walk in background velocity space is terminated and the last accepted model—the one with maximum J_2—is taken as the best velocity model.

A problem that arises when a trial background velocity is far away from the true earth velocity is that the first Born approximation used for computing $f(\mathbf{x}|\mathbf{s}, c_0)$ breaks down. However, this effect is just what we want to observe because the violation of the Born approximation reveals a bad velocity model which we will reject. The illegal use of the Born approximation for an inadequate background velocity model degrades the obtained images. The error in the images can be measured by J_2. In this sense, J_2 is a measure of the validity of the Born approximation. A good background velocity model has been obtained if and only if the Born approximation is applicable for this model.

MODEL PARAMETERIZATION

To accelerate the computation, we did not apply the Monte Carlo method directly for the inversion of the complete velocity model. Instead, we divided the velocity model into several interlaced layers from top to bottom and inversion was performed for each layer from the top down. This layer-stripping procedure reduces considerably the computation volume because the number of ray tracings needed for migration is reduced by the factor of the number of layers in the model. Let us remark, however, that the model should not be divided too thinly, because the top-down strategy has the obvious drawback of accumulating errors. To alleviate these errors, the layers must be interlaced and the thickness of each layer must be large enough to contain a sufficiently large number of reflectors so the inversion method does not converge to a local minimum. Of course, if one of the layers is not well-inverted, we cannot obtain a globally optimized velocity model. Ideally we would like to invert globally, but the cost of the Monte Carlo method increases very rapidly with the number of parameters retained in the inversion.

Because the Monte Carlo method is applied to successive vertical layers, the velocity model has to be parameterized using different vertical and horizontal interpolation functions. To decouple the velocity variation in a layer from that of the other layers, we must use a low order interpolation in the active layer. In our implementation we used a simple quasi-bidimensional interpolation function so that the values of the velocity c_0 at a point (x, z) are given by

$$c_0(x, z) = c_0(x, z_i) + \frac{z - z_i}{z_{i+1} - z_i} \times [c_0(x, z_{i+1}) - c_0(x, z_i)], \quad (15)$$

where z_i are the coordinates of the horizontal grid lines. The velocities $c_0(x, z_i)$ along these lines are obtained by cubic spline interpolation (de Boor, 1978). The quasi-bidimensional method is particularly adapted to our case where the knots are regularly spaced in the horizontal direction and at variable depth in the vertical direction. The quasi-bidimensional approach was reviewed by Gonzalez-Casanova and Alvarez (1985). The derivatives of $c_0(x, z)$ that are needed for ray and dynamic-ray tracing can be obtained by derivation of the spline functions. In our implementation, grid spacing in the vertical and horizontal directions were different. Vertical spacing is controlled by the number of significant reflectors that are included in each layer. The horizontal grid spacing, on the other hand, should be chosen as loose as possible to reduce computer costs.

Figure 3 shows an example of ray tracing where, for clarity's sake, we show only a few rays shot from a single depth horizon. In the actual computer implementation, rays were traced much more densely and shot from many depths. In this figure, we also show schematically the velocity grid. Ray quantities at diffraction points needed for the application of our linearized inversion scheme were obtained by high-order interpolation of the parameters calculated at the grid points.

AN EXPERIMENT WITH NORTH SEA REFLECTION DATA

To test the ability of our procedure to handle practical problems, we applied it to a reflection profile from the North Sea. The reflection profile consists of more than 800 shots with a total cable length of 2400 m for 96 groups of hydrophones; the group interval and the shot interval were both 25 m; the sampling rate was 0.004 ms; and total record length was 6 s. In the previous paper by Lambaré et al. (1992), we applied our linearized asymptotic inversion method to this data-set. In Figure 4 we show the result of linearized inversion using a 2-D background velocity model obtained by classical 1-D velocity analysis.

The only preprocessing that we performed on the data was to attenuate the multiples coming from the ocean bottom. Theoretically, to obtain the small-scale variations of the velocity field using equation (12), we should first deconvolve the source function from the field data. This is a difficult procedure, however, because the source functions are band-limited and they cannot be completely removed from the

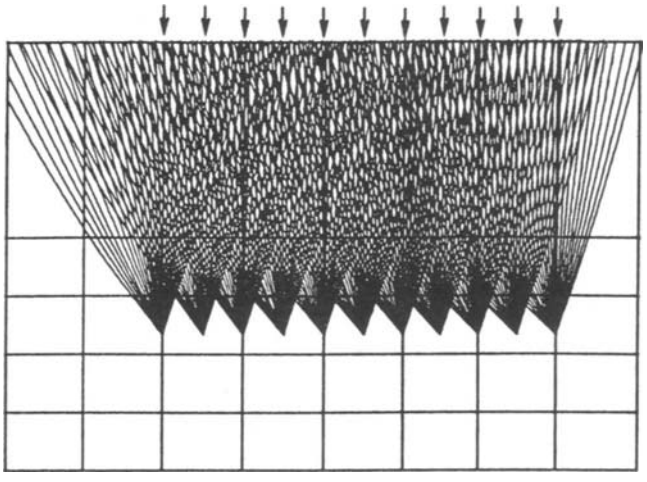

FIG. 3. Geometry of upward ray tracing. Rays are traced from the image point toward sources and receivers located on the earth surface. The arrows denote the Iso-X points. Superimposed on the rays we show schematically the typical size of the grid where the velocity model is defined. Velocities in each cell are represented by quasi-bidimensional spline interpolation function.

seismic data (Lambaré et al., 1992). The main effect of the source function is to limit the frequency band of the small-scale velocity model, inverted using equation (12). Deconvolution is needed only if we want the small-scale velocity model inverted from (12) to match as closely as possible the true short-wavelength components of the medium. In nonlinear inversion we are actually interested in the lateral coherency of the small-scale images obtained by linearized inversion of continuous Common Source Gathers (CSG), not on the exact value of $f(\mathbf{x}|\mathbf{s}, \mathbf{c}_0)$. For this reason we decided not to deconvolve the source signature from the seismic traces. An obvious danger of this procedure is that of cycle-skipping in the inversion which would cause the Monte Carlo technique to converge to the wrong cycle in the reflector event. As long as the source-time functions are nearly the same for all CSGs and do not remove the main features of the small-scale velocity model, the source wavelet should not influence the results of background velocity inversion. However, if the onset times and source functions vary significantly from CSG to CSG, deconvolution may be necessary to make the recorded CSGs more consistent. A final point to notice is that waves propagate in three dimensions. To compensate approximately for 3-D geometrical spreading, we replaced the weighting coefficient $a(\mathbf{x})$ by $a(\mathbf{x})\sqrt{z}$ where z is the depth of an image point.

To reduce computational costs and to save memory we applied nonlinear inversion to a subsection of the seismic profile. We chose a 7200-m section from shot numbers 260 to 548. To facilitate ray tracing and to take into account receivers located beyond the target zone, we extended the model 400 m on each side, giving a total model length of 8 km. Then, from this segment of the profile, we took only one of every two shots, so that we inverted a total number of 144 CSGs. A further reduction of the data set was obtained by using only 48 traces in each CSG, a reduction of 50 percent from the original data. This reduction has little influence on the inverted background velocity model, because our data set is very over-determined. With this reduction of available data, the shot and receiver intervals both became 50 m.

Since we are interested in inverting the large-scale part of the velocity model, we used a very coarse parameterization. The grid spacing was chosen as 1000 m horizontally and 300 m vertically. The velocity field was parameterized using the quasi-bidimensional scheme presented in the previous section. Kinematic and dynamic ray tracing were calculated on this interpolated model. For the computation of each Iso-X gather we inverted 24 CSGs. Since the intertrace interval is 50 m, the offset between the first and last trace in each Iso-X gather was 1.2 km. Iso-X panels were calculated every 500 m; and in the computation of $J_2(\mathbf{c}_0)$, 13 Iso-X gathers were summed for a total profile length of 6 km.

For nonlinear inversion, we used the top-down strategy described in the previous section. The layers used in inversion, as well as the knots for the definition of the model, are displayed in Figure 5. We fixed the velocity of the top of the model at 1570 m/s, which is the acoustic velocity of the sea water. The model was divided into five layers from top to bottom. The thickness of the first layer was 1000 m and the others were 900 m thick. We did not invert for the velocity of the first layer (from 0–900 m) because of the presence of strong, complicated water-bottom multiples that appear at wide reflection angles. The velocity in this range was obtained by interpolation. As explained earlier, to reduce error accumulation, the layers were interlaced, so that the bottom of a layer is in the interior of the following one. Each layer contained 24 unknowns, eight of which were reinverted in the following step because they belonged to two consecutive layers. For each layer, nonlinear inversion was automatically terminated after about 800 iterations because of the inability to find a new acceptable model during the last 120

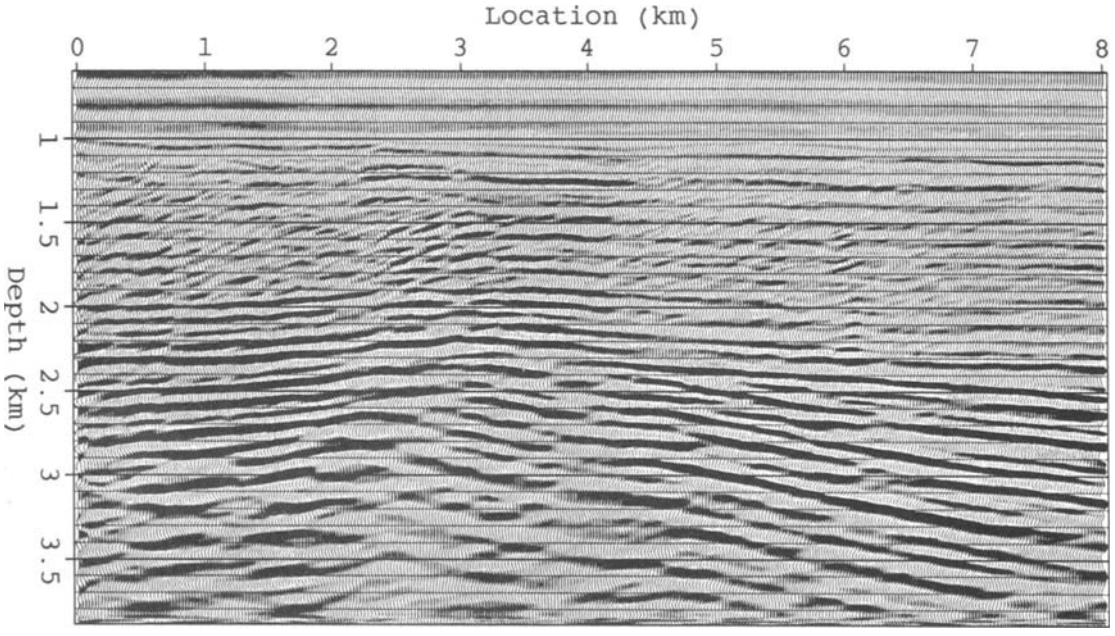

FIG. 4. Seismic profile from the North Sea. Result of linearized asymptotic inversion obtained by Lambaré et al. (1992) using the background velocity obtained by velocity analysis.

iterations. Since we were interested only in smooth velocity models, in the random walk we accepted only those models that made ray tracing regular. Therefore, in the implementation of our algorithm, we systematically rejected models that contained ray singularities that prevented the rays from reaching the surface. In the optimization process, the velocity variation range was limited to the interval from 1700 m/s to 5500 m/s. The maximum step dc (Figure 2) for the Monte Carlo random walk was chosen as 3000 m/s. This step is large enough to allow the optimization algorithm to escape from secondary maxima of the objective functional.

The final result of the inversion is shown in Figure 6. To facilitate ray tracing, the velocity model is extrapolated on both sides of the section. Since our objective function J_2 measures the lateral coherency of the small-scale models, Iso-X gathers provide a direct visual verification of the quality of the result of nonlinear inversion. In Figure 7 we present several Iso-X gathers calculated using the inverted background velocity model of Figure 6. From this figure, we see that maximizing $J_2(\mathbf{c}_0)$ is actually equivalent to aligning the reflectors in the different Iso-X gathers. Thus our best model is expected to be the same as the one that would be obtained using migration-velocity analysis. Some residual moveouts in Iso-X gathers of Figure 7 still exist, especially for the deeper layers. This is probably because of the accumulation of errors from the process of layer-stripping. It is likely that this problem would disappear if we had inverted for all depths simultaneously. But this requires the use of a much bigger computer than the workstation we had available.

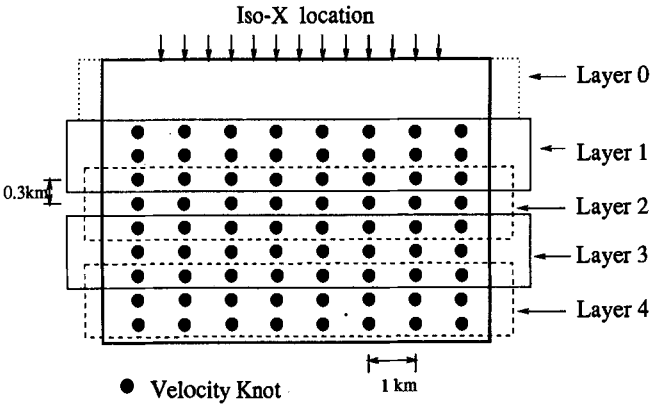

FIG. 5. Parameterization of the model for Monte Carlo inversion. The background velocity model for the profile is parameterized using 72 knots indicated by the filled circles. A Monte Carlo inversion following the flow diagram of Figure 2 was performed on each of the four interlaced layers shown in the figure. Twenty-four parameters were simultaneously inverted in each Monte Carlo step.

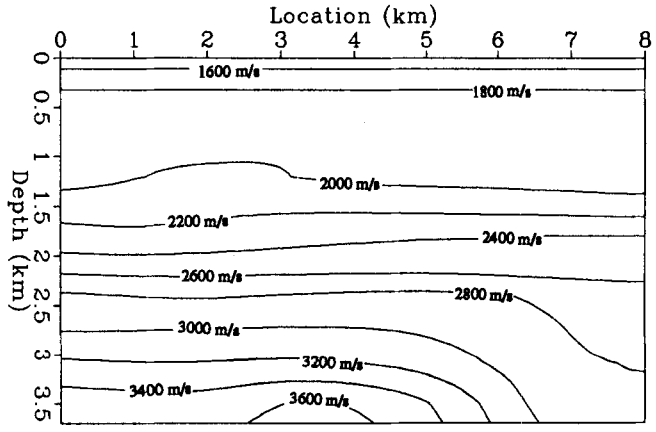

FIG. 6. Contour plot of the final background velocity model obtained by nonlinear Monte Carlo inversion.

To compare the result of nonlinear inversion with that obtained by conventional velocity analysis, we show in Figure 8 the velocity model obtained by classical procedures by Lambaré et al. (1992). This model was obtained by smoothing the interval velocities calculated by the Dix equation from the stacking velocities. Comparing with Figure 6, we observe that the result of the velocity analysis is different from our model obtained by the Monte Carlo method, especially below the depth of 2 km. The effect of the velocity model on the inverted small-scale profile can be observed by comparing the results shown in Figure 4 with that of Figure 9. Figure 4 was calculated with the background velocity model obtained by velocity analysis (Figure 8), while Figure 9 was calculated using the model obtained by nonlinear inversion (Figure 6). Using the background velocity inverted by our technique, linearized inversion has better resolution, while, using the velocity model obtained by velocity analysis, the inverted section (Figure 4) is less sharp and has many unrealistic lateral discontinuities below the depth of 2500 m, between surface locations x = 2000 m and x = 4000 m. To verify that the model obtained by Monte Carlo inversion was better than that determined by velocity analysis, we calculated the value of the contributions of each Iso-X to the objective functional J_2. The result presented in Table 1 shows that the model obtained by nonlinear inversion is systematically better than that of velocity analysis at all depths. J_2 is 2 percent larger for the inverted model. Since our inversion method is based on the lateral coherency of the Iso-X gathers, in Figure 10 we compare the Iso-X gathers calculated at x = 3200 m for the two velocity models. Although in both of them the events in the Iso-X gathers are visually well-aligned, they do not give the same migrated sections (cf. Figures 4 and 9). This is because in nonlinear inversion we assumed the background velocity model to be 2-D; while, for velocity analysis, velocity variation was assumed to vary only with depth. Although the subsurface we studied has only weak lateral variations, the assumption of a local 1-D earth considerably degrades the migrated results. The conventional velocity analysis is done independently at one surface location after another. In this manner it produces unrealistic strong lateral velocities gradients (Figure 8). In contrast, velocity analysis by the Monte Carlo method is done simultaneously at many surface points. The velocities under one surface point must maximize not only the stacking power of this point but also the stacking powers of the neighbor points. This means the velocities are subjected to more constraints than CMP stacking velocity analysis. In this way, we can reduce more nonuniqueness of the solutions than can conventional velocity analysis.

As a direct verification of our results, we compare in Figure 11 the large-scale velocity model obtained from

nonlinear inversion with a velocity-log from a 2500-m deep well located in the vicinity of shot 350, near x = 4 km. The well log has been low-pass filtered with a cut-off frequency similar to the maximum frequency contained in the seismic data. The agreement between the low frequency trend of the average sonic velocity with the result of inversion is quite satisfactory, although some low frequencies are missing, especially above 1.2 km depth. The main reasons for the missing frequencies are: first, the Monte Carlo method can find the solution only near the true one for finite iterations. Second, systematically low velocities above 1.2-km depth may be produced by the multiples generated at the sea bottom. Although comparing to the well log is an imperfect method, this result confirms that our nonlinear inversion method produces reasonably accurate velocity estimates.

Finally, as an indication of the computer costs of our algorithm, we provide the following information. The computation of one Monte Carlo step and an evaluation of J_2 for one velocity model for the set of 144 shot gathers (a total of 6912 traces, each of which contains 1500 samples) took

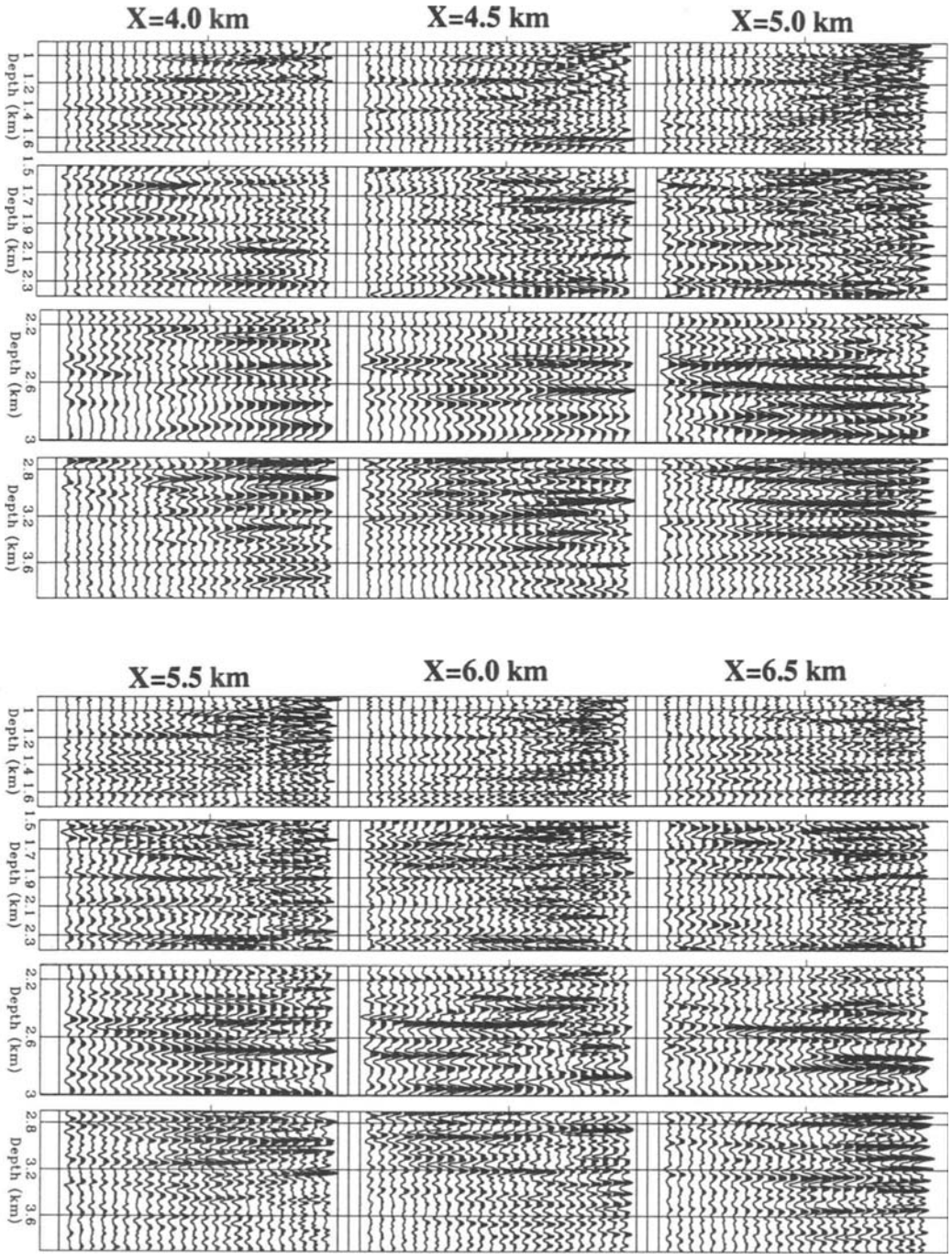

FIG. 7. Iso-X gathers calculated for each of the four layers of Figure 5 using the background velocity model inverted by the Monte Carlo method. The results show that maximizing our objective function is equivalent to aligning the reflectors of the Iso-X gathers.

about 2 minutes of CPU on a Sun Sparc-2 workstation. This CPU time includes the computation of inversion-migration for all the traces and the computation of J_2.

DISCUSSION

We have shown that inverting seismic profiles in two steps—a fast asymptotic linearized step for the small-scale velocity perturbations and a full nonlinear Monte Carlo inversion for the large-scale part—yields an effective algorithm that can be used in present-day computers. Two different objective functions were used for the velocity determination problem. The global minimum of the waveform misfit-function J_1 is obtained only if J_2 is close to its global maximum. The use of two different cost functionals is critical to the success of our method. Without this separation, optimization based only on the waveform misfit would be too slow because of the sheer number of degrees of freedom contained in the waveforms.

Although we did not make explicit use of traveltimes for the determination of the reference velocity, the cost functional J_2 belongs to the class of kinematically coherent criteria. A model that gives the correct kinematical travel-

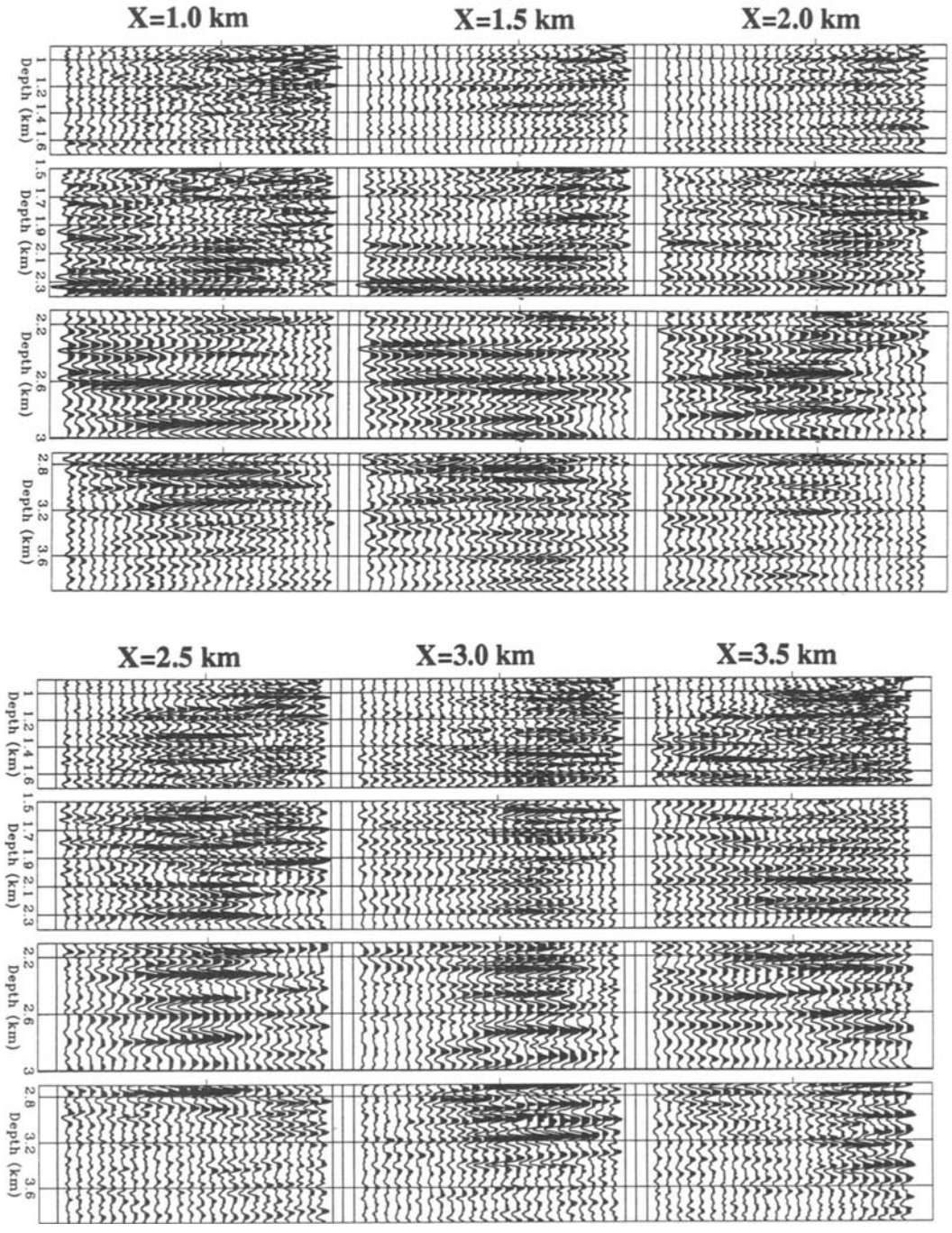

FIG. 7. CONTINUED

times gets a high value of J_2, so that it is preferred to those that give incorrect traveltimes under this criterion. Kinematically correct models are not unique, however, because there is always a group of velocity models that do not affect the traveltime data. For example, Claerbout (1985) and Jannane et al. (1989) found that, with standard reflection data recorded at the Earth's surface, it is possible to retrieve only the long-wavelength structure of the velocity and the short-wavelength structure of the impedance. A whole band of intermediate wavenumbers is undetermined by reflection profiles. Only transmitted waves seem to have information about these intermediate wavelengths. In practice this means that, under any kinematic criterion, the model obtained from inversion can be modified at certain intermediate-scale wavelengths without affecting the fit of the waveforms or the traveltimes. This basic indeterminacy has to be taken into account when comparing the results obtained by classical velocity analysis (Figure 8) with those obtained by inversion (Figure 6). This nonuniqueness also makes it very difficult to obtain estimates of the accuracy of the velocity model obtained by nonlinear inversion. This problem will be the subject of future work.

In this paper we used a Monte Carlo random search in the large-scale velocity model space. This is certainly not the only, or the best, nonlinear method for the exploration of solution space. Other algorithms that can accelerate convergence can be used; for instance, simulated annealing (Landa et al., 1989) or genetic algorithms (Stoffa and Sen, 1991). The

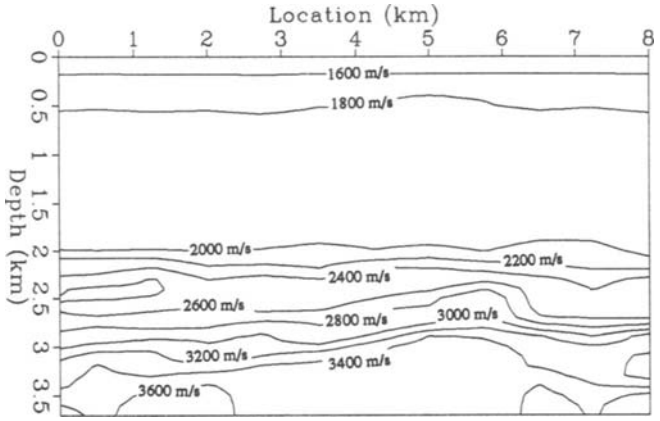

FIG. 8. Velocity model obtained by Lambaré et al. (1992) using conventional velocity analysis.

Table 1. Comparison of the contribution to the objective functional J_2 of several different Iso-X sections for the two models obtained by velocity analysis and by the Monte Carlo technique. The objective functional J_2 is the sum of the columns.

Iso-X Location x	Velocity analysis	Monte Carlo
1.0 km	0.749	0.759
1.5 km	0.932	1.000
2.0 km	0.653	0.694
2.5 km	0.632	0.634
3.0 km	0.677	0.683
3.5 km	0.718	0.680
4.0 km	0.783	0.829
4.5 km	0.788	0.800
5.0 km	0.696	0.704
5.5 km	0.737	0.749
6.0 km	0.730	0.729
6.5 km	0.778	0.811
7.0 km	0.886	0.898
J_2	9.759	9.970

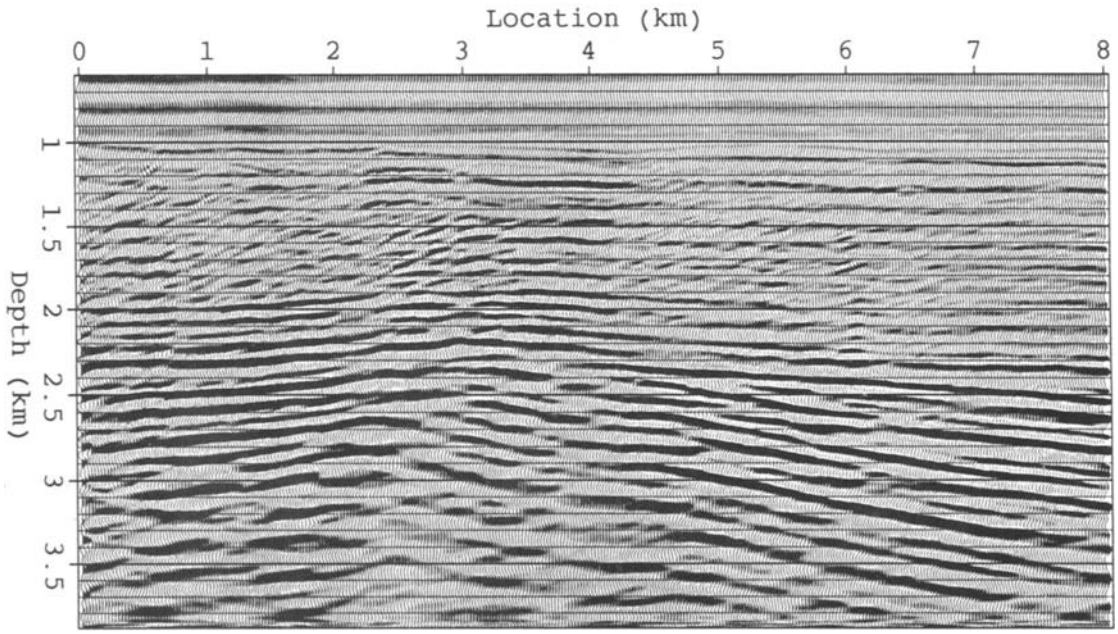

FIG. 9. Seismic profile obtained by linearized inversion-migration using the background velocity obtained by the Monte Carlo technique.

use of genetic algorithms for the inversion of background velocities was considered by us in a recent publication (Jin and Madariaga, 1993).

Let us remark that, in principle, any fast prestack migration or linearized inversion method can be used to determine the small-scale impedance model used for the evaluation of the cost functional $J_2(\mathbf{c}_0)$. However, our linearized asymptotic inversion based on ray theory has two advantages: first, since the Iso-X are sparsely sampled, we only need to trace a relatively small number of rays. Second, in contrast with finite difference migration-inversion we do not have to do a backward migration of surface data and then apply the imaging condition. The two steps are reduced to a single one using ray theory. This is the main reason why our nonlinear inversion is so fast.

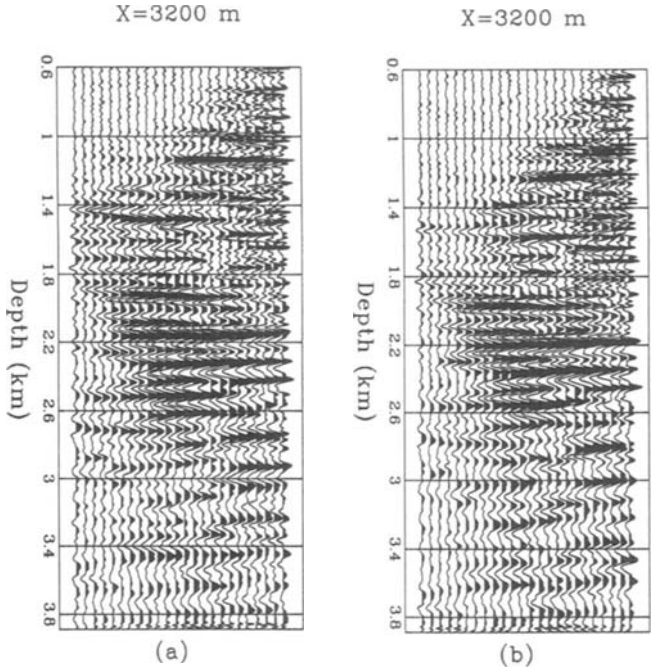

FIG. 10. Comparison of the Iso-X gathers calculated at x = 3200 m: (a) using the background velocity obtained by the Monte Carlo technique, (b) using the background velocity obtained by velocity analysis.

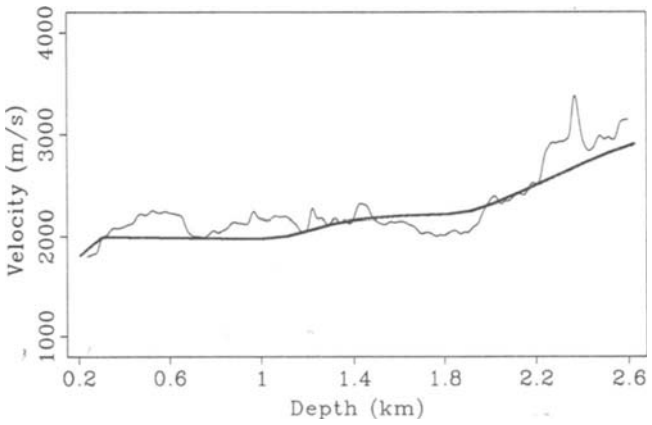

FIG. 11. Comparison of the background velocity model calculated by the Monte Carlo technique (smooth line) at shot position 350 (close to x = 4 km in Figure 9) with a filtered velocity log from a well located in the vicinity of this shot.

Like any method that analyzes reflections, nonlinear inversion is limited by its inability to resolve the velocities where there are no reflectors. Our method is also limited by the lack of redundancy of the data, by strong coherent shot noise that may degrade the resolution, and by water-reverberations and multiple arrivals that perturb the model retrieved for shallow depths. Generally, the influence of the multiples can be removed at greater depths because the velocity differences between primary and multiple events is sufficient to remove a significant amount of multiple energy by common-shot stacking in the cost functional $J_2(\mathbf{c}_0)$. Errors caused by waveform cycle-skipping in horizontally aligned Iso-X gathers occur only when the velocity model is correct. These errors will affect, of course, the small-scale model but they will not modify the results of velocity inversion. Once the velocity model has been sufficiently well-determined, a few steps of iteration of the linearized inversion method should be performed to eliminate the effects of cycle-skipping.

The precision of the obtained velocity is mainly controlled by the offset range of the data. Under the assumption that the medium is invariant with depth, Lynn and Claerbout (1982) studied the relation between the minimum obtainable lateral resolution and cable lengths. The relation between lateral velocity resolution and source-receiver separations was also discussed in Bickel (1990).

CONCLUSION

We have proposed a nonlinear Monte Carlo inversion method for the iterative determination of large- and small-scale components of the velocity model from surface seismic data. Realizing that the smooth background model requires fewer parameters than the short scale model, we separated the inversion into two parts: a linear inversion for the small-scale part, and a nonlinear inversion for the smooth, large-scale part of the model. The advantage of this separation is that nonlinear inversion can be formulated with a relatively small number of parameters, while linear inversion has a very large number of degrees of freedom.

For the inversion of the smooth background velocity model, we used the fact that the location of reflectors obtained by linearized inversion contains the required information to invert the background velocity. We showed that the coherency of the images reconstructed by inverting different data subsets can be used as a measure of the quality of the velocity model. In our inversion method we used the coherency of the models obtained by different CSGs. Using the fast linearized asymptotic inversion technique that we developed previously (Jin et al., 1992; Lambaré et al., 1992), the iterative process of velocity determination can be performed automatically by a standard Monte Carlo method. To accelerate nonlinear inversion, we do not invert the whole model in a single operation; instead we proceed by a series of interlaced layers and perform inversion one layer at a time. This reduction in the number of parameters for the nonlinear part renders the problem tractable in present-day computers for a seismic profile that is almost 7 km long. Much larger

models could, of course, be inverted in supercomputers and, in particular, it would be possible to avoid layer-stripping in the inversion.

Inversion of a profile from the North Sea shows that the Monte Carlo method can be used to retrieve a laterally heterogenous velocity model in a realistic situation.

ACKNOWLEDGMENTS

This work is a contribution of Groupement de Recherche Sismique sponsored by Centre National de la Recherche Scientifique, Elf Aquitaine Company, and Institut Français du Pétrole. The seismic data set was kindly provided by SNEA(P). We thank Jean Virieux for providing his ray-tracing code and helping us with many problems related to ray tracing. Gilles Lambaré prepared the data and provided many constructive ideas throughout this work.

REFERENCES

Al Yahya, K. M., 1989, Velocity analysis by iterative profile migration: Geophysics, **54**, 718–729.

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

Beylkin, G., and Burridge, R., 1990, Linearized inverse scattering problems in acoustics and elasticity: Wave Motion, **12**, 15–52.

Bickel, S. H., 1990, Velocity-depth ambiguity of reflection traveltimes: Geophysics, **55**, 266–276.

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

Cao, D., Beydoun, W. B., Singh, S. C., and Tarantola, A., 1990, A simultaneous inversion for background velocity and impedance maps: Geophysics, **55**, 458–469.

Chiu, S., Kanasewich, E., and Padke, S., 1986, Three-dimensional determination of structure and velocity by seismic tomography: Geophysics, **51**, 1559–1571.

Claerbout, J. F., 1985, Imaging the Earth's interior: Blackwell Scientific Publications, Inc.

Crase, E., Pica, A., Noble, M., McDonald, J., and Tarantola, A., 1990, Robust elastic nonlinear waveform inversion: Application to real data: Geophysics, **55**, 527–538.

De Boor, K., 1978, A practical guide to splines, Springer-Verlag, New York.

Faye, J. P., and Jeannot, J. P., 1986, Prestack migration velocities from focusing depth analysis: 56th Ann. Internat. Mtg., Soc. Expl. Geophysics, Expanded Abstracts, 438–440.

Gonzalez-Casanova, P., and Alvarez, R., 1985, Splines in Geophysics: Geophysics, **50**, 2831–2848.

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., Trézeguet, D., and Xie, M. Y., 1989, short note: Wavelengths of Earth structures that can be resolved from seismic reflection data: Geophysics, **54**, 906–907.

Jin, S., and Madariaga, R., 1993, Background velocity inversion with a genetic algorithm: Geophys. Res. Lett., **20**, 93–96.

Jin, S., Madariaga, R., Virieux, J., and Lambaré, G., 1992, Two-dimensional asymptotic iterative elastic inversion: Geophys. J. Int., **108**, 575–588.

Kim, Y. C., and Gonzalez, 1991, Migration velocity analysis with the Kirchhoff integral: Geophysics, **56**, 365–370.

Lambaré, G., Virieux, J., Madariaga, R., and Jin, S., 1992, Iterative asymptotic inversion of seismic profiles in the acoustic approximation: Geophysics, **57**, 1138–1154.

Landa, E., Beydoun, W., and Tarantola, A., 1989, Reference velocity model estimation from prestack waveforms: Coherency optimization by simulated annealing, Geophysics, **54**, 984–990.

Lynn, W. S., and Claerbout, J. F., 1982, Velocity estimation in laterally varying media: Geophysics, **47**, 884–897.

Miller, D., Oristaglio, M. L., and Beylkin, G., 1987, A new slant on seismic imaging: Migration and integral geometry: Geophysics, **52**, 943–964.

Romanowicz, B., 1982, Moment tensor inversion of long period Rayleigh waves: A new approach: J. Geophys. Res., **87**, 5395–5407.

Snieder, R., Xie, M. Y., Pica, A., and Tarantola, A., 1989, Retrieving both the impedance contrast and background velocity: A global strategy for the seismic reflection problem: Geophysics, **54**, 991–1000.

Stoffa, P. L., and Sen, M. K., 1991, Nonlinear multiparameter optimization using genetic algorithms: Inversion of plane-wave seismograms: Geophysics, **56**, 1794–1810.

Symes, W. W., and Carazzone, J. J., 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663.

Tarantola, A., 1984, Linearized inversion of seismic reflection data: Geophys. Prosp., **32**, 998–1015.

Versteeg, R., 1991, Analyse du probème de la détermination du modele de vitesse pour l'imagerie sismique: Ph.D. dissertation, Univ. of Paris 7, France.

Yilmaz, O., and Chambers, R. E., 1984, Migration velocity analysis by wavefield extrapolation: Geophysics, **49**, 1664–1674.

APPENDIX

ESTIMATION OF THE HESSIAN

We show very briefly the method used to approximate the Hessian $\underline{\mathbf{H}} = \underline{\mathbf{P}}^{\dagger}\underline{\mathbf{q}}\underline{\mathbf{P}}$ defined in equation (9). The linearized forward modeling operator $\underline{\mathbf{P}}$ was defined as the operator in expression (4). Therefore the Hessian is:

$$\underline{\mathbf{H}}(\mathbf{x}, \mathbf{y}|\mathbf{s}) = \frac{-\omega^4}{2\pi}\int_{\omega_{\min}}^{\omega_{\max}} d\,\omega \int_{\partial R} d\mathbf{x}_r\, q(\mathbf{x}, \mathbf{s}, \mathbf{x}_r, \omega) \times G^{\dagger}(\mathbf{x}_r, \mathbf{y}, \omega)G^{\dagger}(\mathbf{y}, \mathbf{s}, \omega)G(\mathbf{x}_r, \mathbf{x}, \omega)G(\mathbf{x}, \mathbf{s}, \omega), \quad \text{(A-1)}$$

where $\mathcal{M}$ is the support of $f(\mathbf{y})$, ∂R is the data acquisition line. $(\omega_{\min}, \omega_{\max})$ is the frequency band of the seismic source signal.

To estimate $\underline{\mathbf{H}}$, we replace the Green's functions in equation (A1) by their high-frequency asymptotic approximation (6). We consider the inversion of $\mathbf{f}$ in a small vicinity of the current point $\mathbf{x}$. We approximate the amplitude functions $A(\mathbf{x}_r, \mathbf{s}, \mathbf{y}) = A(\mathbf{y}, \mathbf{s})A(\mathbf{x}_r, \mathbf{y})$ near $\mathbf{x}$ by a constant amplitude $A(\mathbf{x}_r, \mathbf{s}, \mathbf{x})$, and we expand the two-way traveltime function $\tau(\mathbf{x}_r, \mathbf{s}, \mathbf{y}) = \tau(\mathbf{y}, \mathbf{s}) + \tau(\mathbf{x}_r, \mathbf{y})$ in a first-order Taylor series about the point $\mathbf{x}$:

$$\tau(\mathbf{x}_r, \mathbf{s}, \mathbf{y}) - \tau(\mathbf{x}_r, \mathbf{s}, \mathbf{x}) = \mathbf{p}(\mathbf{x}_r, \mathbf{s}, \mathbf{x}) \cdot (\mathbf{y} - \mathbf{x}). \quad \text{(A-2)}$$

We thus obtain from equation (A-1)

$$\underline{\mathbf{H}}(\mathbf{x}, \mathbf{y}|\mathbf{s}) = \frac{1}{(2\pi)^2}\int_{\omega_{\min}}^{\omega_{\max}} \omega\, d\omega \int_{\partial R} d\mathbf{x}_r p^2 \mathcal{J}(\theta, \xi; \mathbf{x}_r, \mathbf{s}, \mathbf{x}) \times e^{i\omega \mathbf{p}\cdot(\mathbf{y}-\mathbf{x})},$$

where $p = \|\mathbf{p}\|$ was defined in equation (8). For a given source point $\mathbf{s}$ there is a one-to-one mapping from the total slowness vector $\mathbf{p}$ to receiver position $\mathbf{x}_r$. We can therefore, following Figure 1, change variables from $\mathbf{x}_r$ to ξ. The Jacobian of this transformation is precisely $\mathcal{J}$. By this transformation, ∂R is mapped into a subset Ξ of a unit circle.

$$\underline{\mathbf{H}}(\mathbf{x}, \mathbf{y}|\mathbf{s}) = \frac{1}{(2\pi)^2} \int_{\omega_{\min}}^{\omega_{\max}} \omega \, d\omega \int_{\Xi} d\xi p^2 e^{i\omega|p|\xi \cdot (\mathbf{y}-\mathbf{x})}. \tag{A-3}$$

Finally, changing variables from ω and $\boldsymbol{\xi}$ to $\mathbf{k}$ defined by

$$\mathbf{k} = \omega \mathbf{p}, \tag{A-4}$$

we get, using $\kappa = \|\mathbf{k}\|$:

$$\underline{\mathbf{H}}(\mathbf{x}, \mathbf{y}|\mathbf{s}) = \frac{1}{(2\pi)^2} \int_{\mathcal{D}_k} \kappa \, d\kappa \, d\xi \, e^{i\mathbf{k} \cdot (\mathbf{y}-\mathbf{x})}, \tag{A-5}$$

where $\mathcal{D}_k$ is the domain of definition of equation (A-5) in the Fourier domain κ. This domain depends on the data acquisition geometry and the signal frequency band. It controls the spatial resolution of $\mathbf{f}(\mathbf{x}|\mathbf{s})$. Equation (A-5) is a band-limited approximation to a delta function. When $\mathcal{D}_\kappa$ is sufficiently large, $\underline{\mathbf{H}}$ approaches

$$\underline{\mathbf{H}}(\mathbf{x}, \mathbf{y}, \mathbf{s}) \simeq \delta(\mathbf{x} - \mathbf{y}), \tag{A-6}$$

which is the approximation we have used in deriving equation (11). Let us remark that the Hessian is unitary thanks to the choice of the weighting function $\mathbf{q}$ in the objective functional J_1. For further details refer to Jin et al. (1992).