Practical aspects and applications of 2D stereotomography

Frederic Billette[*], Soazig Le Bégat[‡], Pascal Podvin[‡], and Gilles Lambaré[‡]

ABSTRACT

Stereotomography is a new velocity estimation method. This tomographic approach aims at retrieving subsurface velocities from prestack seismic data. In addition to traveltimes, the slope of locally coherent events are picked simultaneously in common offset, common source, common receiver, and common midpoint gathers. As the picking is realized on locally coherent events, they do not need to be interpreted in terms of reflection on given interfaces, but may represent diffractions or reflections from anywhere in the image. In the high-frequency approximation, each one of these events corresponds to a ray trajectory in the subsurface. Stereotomography consists of picking and analyzing these events to update both the associated ray paths and velocity model. In this paper, we describe the implementation of two critical features needed to put stereotomography into practice: an automatic picking tool and a robust multiscale iterative inversion technique. Applications to 2D reflection seismic are presented on synthetic data and on a 2D line extracted from a 3D towed streamer survey shot in West Africa for TotalFinaElf. The examples demonstrate that the method requires only minor human intervention and rapidly converges to a geologically plausible velocity model in these two very different and complex velocity regimes. The quality of the velocity models is verified by prestack depth migration results.

INTRODUCTION

Prestack depth migration is now accepted as the appropriate imaging technique for geologically complex structures. Although both migration and subsequent common-reflection-point (CRP) gather analysis technologies are in constant evolution toward better algorithms [e.g., multiarrival Kirchhoff (Operto et al., 2000; Xu et al., 2000) or wave-equation migration (Etgen and Biondi, 2001), followed by amplitude-variation-with-offset (AVO)/amplitude-variation-with-angle (AVA) analysis], they still require the estimation of an accurate velocity macromodel. In the case of quasi-horizontal sedimentary structures, conventional processing using simple normal-moveout (NMO) based velocity estimations has proven to work well. However, it has been demonstrated numerous times that imaging complex geological structures (e.g., faulted or tilted blocks, salt domes, gas sags, and significant lateral velocity contrasts in general) requires more sophisticated velocity estimation techniques because the 1D assumption is then violated.

Many advanced velocity-estimation methods are based on comparing traveltimes obtained from an initial velocity model with traveltimes extracted from seismic data. Standard "traveltime tomography" (Bishop et al., 1985) requires traveltimes to be picked on continuous horizons over the entire offset range. This approach requires a picking step where continuous events are interpreted and correlated in a 3D sense (inline, crossline, and offset dimensions). Furthermore, two-point ray tracing is involved, which significantly increases the computing time, especially when multiarrivals are taken into account (Clarke et al., 2001). Applications of traveltime tomography to 3D problems have been documented (Guiziou et al., 1996; Broto and Ehinger, 1998; Le Stunff and Grenier, 1998), but because it is rather time-consuming, this approach has not been routinely applied in exploration projects.

An alternative approach, migration velocity analysis (MVA) exploits seismic data redundancy to obtain a velocity model. Errors in the velocity model result in the appearance of residual moveout in constant image-point sections of the prestack migrated image cube. Migration-based velocity estimation aims at translating this residual moveout into velocity updates, the workflow being a loop between migration, analysis, and velocity updates. The initial concept (Yilmaz and Chambers, 1984; Al-Yahya, 1989) led to a wide range of developments including (1) differential semblance optimization (DSO) without picking (Symes and Carazzone, 1991), (2) picking of either locally coherent (Chauris et al., 2002) or continuous (Liu, 1997) events, (3) "gamma" residual analysis along a fine grid following

horizons (Etgen, 1991; Billette et al., 2002) or on a sparse grid of common depth points (CDPs) (Audebert et al., 1998; Woodward et al., 1998), and (4) a residual-migration wave-equation tomographic technique (Biondi and Sava, 1999). The efficiency and accuracy of these various migration-based methods strongly depends on the density of points where the analysis is carried out. Practically, the analysis is typically done on a 500 m × 500 m grid due to the high cost of the migration. Due to computational limitations of tomographic inversions, only long-wavelength anomalies can generally be determined.

In this context, we believe there is a need for a tomographic method that can estimate lateral velocity variations with only minimal human intervention. An original approach, the controlled directional reception (CDR) method, was introduced in Russia some time ago (Riabinkin, 1957) and updated using more recent tools at Stanford University (Sword, 1987). The basic concept of CDR is to pick and use both the traveltime and its spatial derivatives, which are the slope of locally coherent arrivals in the prestack time domain. Using this additional piece of information helps to better constrain the problem of velocity determination. As implemented by its authors and as applied by Whiting (1998), the CDR method can only handle applications to simple structures where rays are not bending too sharply and where no caustics occur. Billette (1998) and Billette and Lambaré (1998) recast slope-tomography approaches in a general formulation, of which stereotomography appears to be the most robust choice for velocity estimation in complex structures [the name "stereotomography" was already introduced in a very different context by Dor and Labossière (1984)]. They presented the theoretical aspects and gave preliminary results without discussing practical issues such as data picking. In this paper, we develop the practical aspects of the method described in Billette (1998) and Billette and Lambaré (1998), including data picking and the model updating procedure. The effective applicability of stereotomography to various 2D structures will be illustrated by two applications: "Marmousoft," a simplified version of the Marmousi data set, and a 2D line extracted from a 3D towed-streamer survey provided to us by TotalFinaElf.

STEREOTOMOGRAPHY: PRINCIPLES

Background

As introduced by Riabinkin (1957) and Sword (1987), slope tomographic methods are characterized by the fact that, in addition to the traveltime, the slopes of locally coherent events are also used by the tomographic procedure. A slope is the tangent of a coherent event measured from prestack seismic trace gathers. In a 2D common shot gather, one can measure the slope p_{Rx} of a given coherent event at a given receiver location (Figure 1a). In a 2D common receiver gather, one can measure the slope p_{Sx} at a given source location for the same event. In a 2D common offset gather, $p_{Sx} + p_{Rx}$ is measured, whereas we get $(p_{Sx} - p_{Rx})/2$ in a 2D CDP gather. Two types of prestack data gathers are necessary to obtain both p_{Rx} and p_{Sx} in the horizontal direction.

Given a high-frequency approximation of wave propagation, any reflection or diffraction can be modeled by a pair of ray segments joining three points together: the source S, the reflecting/diffracting point X in the subsurface, and the receiver R. The two rays emerge at the surface with a particular direction given by the horizontal components of the two associated slowness vectors. The slopes measured in the 2D prestack gathers (p_{Rx} and p_{Sx}) are equivalent to the horizontal component of the slowness vectors emerging in S and R, respectively (Figure 1b).

The CDR method

In the CDR method, ray couples are propagated downwards in a given initial model. The surface shooting angles are obtained through slope measurements of the data. One-way traveltimes are integrated along each ray with a constant depth step until their sum equals the two-way traveltime picked in the
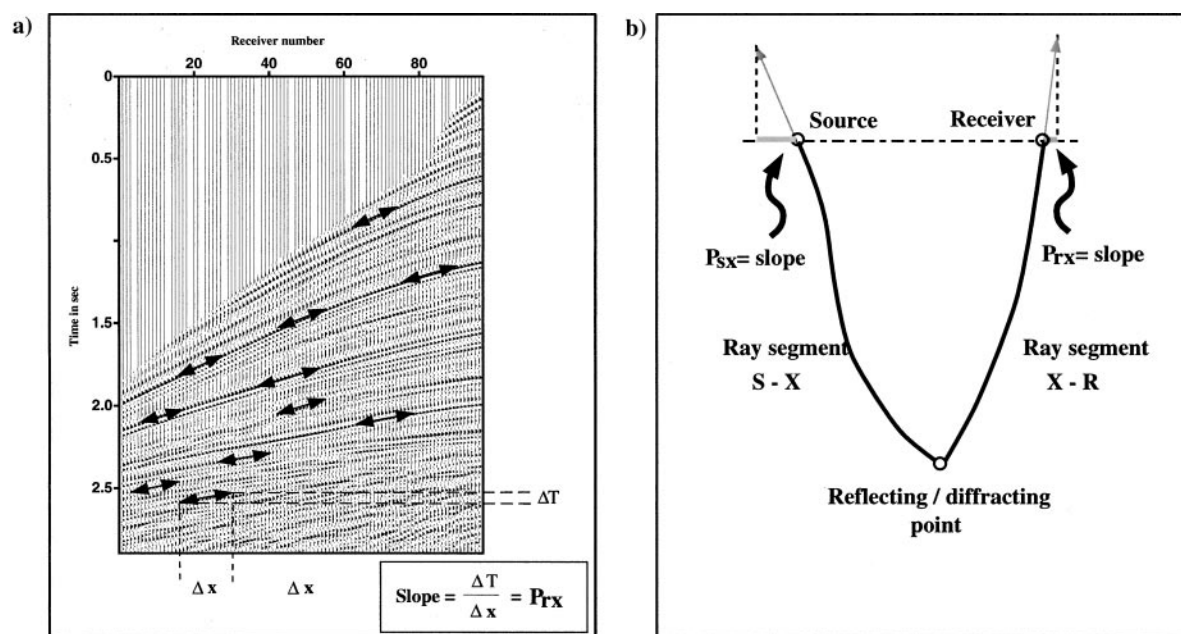

FIG. 1. The slope is the tangent of a locally coherent event in the data space (left) and also the horizontal component of the slowness vector in the model space (right).

data. Because the initial velocity model is incorrect, the distance between the final points of the two rays may not be zero. This misfit is used to update the velocity model until the two rays terminate at a unique X location. Note that no simplifying hypothesis has been introduced concerning subsurface velocities or the lateral continuity of the coherent event picked in the data; as long as all of the required slopes are picked, CDR is a fully 3D method [see Chalard et al. (2000) for a discussion of the number of slope components required in 3D]. On the other hand, the CDR method does not allow uncertainties about data measurement to be introduced. This is unfortunate, because inversion results are very sensitive to errors of the slopes measured in the data. In CDR, it is easily understood how this happens, because slope errors result in significant modification of ray trajectories. Furthermore, the convergence of the two rays to a single point may break down in the case of highly diverging rays or turning rays, which have to be eliminated. In fact, the weak points of CDR become amplified as soon as the quality of the data deteriorates (causing loss of slope-picking precision) as well as when the subsurface becomes somewhat complex (resulting in complicated ray paths).

Data and model in stereotomography

Billette and Lambaré (1998) introduced a general formulation of slope tomography. Each locally coherent event picked on a given trace is described by source and receiver positions (S, R), a two-way traveltime T_{SR}, and two local slopes $(P_S = \partial T_{SR}/\partial S, P_R = \partial T_{SR}/\partial R)$. The event is modeled as the result of single diffraction (or, alternatively a specular reflection) at some location in depth. Stereotomography characterizes this event by a pair of ray segments terminating at the same point in depth. The theory shows that any perturbation of the parameters describing the picked events can be interpreted in terms of perturbations of both the velocity model and the ray segment pair. This approach is thus somehow "orthogonal" to CDR, where the distance between ray-segment final points in depth was the only criterion used to update velocities, whereas other ray path characteristics (surface locations, associated slopes, and traveltimes) were determined by picked data and considered perfectly known and accurate. In stereotomography, in contrast, the ray segments are joined in depth by definition (no experimental uncertainty is attached to this), but all of their boundary conditions are relaxed (Billette and Lambaré, 1998), that is, tested against (possibly noisy) picked data. In practice, two rays are propagated upwards in an a priori velocity model from a common a priori location in depth, with a priori shooting angles, and with two a priori one-way traveltimes. The ending points of the rays correspond to calculated source and receiver locations. The horizontal components of the ray parameters at these points provide calculated slopes p_{Rx} and p_{Sx}, and the sum of the two a priori one-way traveltimes directly provides the event's calculated two-way traveltime. If the a priori velocity model is not correct, these calculated positions, slopes and traveltime will not fit the corresponding values measured in the data. In this case, the misfits are jointly interpreted in terms of velocity and ray-segment errors. In this formulation, we try to fit calculated data to observed data (i.e., source and receiver positions, slopes, and two-way traveltimes) only within the error margins defined by measurement uncertainties. In practice, we consider an a priori covariance matrix

in the data space that is not the Identity matrix but a diagonal one.

In stereotomography, the model space $\mathbf{M}$ is defined by a model vector $\mathbf{m}$ that includes a velocity field C and N pairs of ray segments (see Figure 2):

$$\mathbf{m} = \left[[(X, \beta_S, \beta_R, T_S, T_R)_i]_{i=1}^N, [C_j]_{j=1}^J \right]. \quad (1)$$

The velocity model is described by J cardinal cubic B-splines $B_j(X)$, where $C(X) = \sum_{j=1}^J C_j B_j(X)$, B_j are the B-spline functions and C_j are the weights. This particular choice guarantees second-order continuity of the velocity field and allows for an efficient ray-tracing implementation. Moreover, this parameterization produces velocity models that are directly usable for ray-based migration without requiring any ad hoc smoothing, as would be needed with blocky velocity macromodels, for example. A smooth description of the velocity macromodel has many numerical advantages (Chapman, 1985; Lailly and Sinoquet, 1996; Lucio et al., 1996) and does not significantly alter imaging quality (Versteeg, 1993; Mispel and Hanitzsch, 1996; Gray, 2000). In equation (1), a pair of ray segments is described by a common starting point X, two takeoff angles β_S and β_R, and two one-way traveltimes T_S and T_R. It is important to point out that the number of parameters $\mathbf{m}$ required to describe the model increases with the number of picked events N. For large-size applications, this must be accounted for in the implementation of the method.

Inversion

Because we want to take into account measurement uncertainties, we develop stereotomography in the stochastic formulation of inverse problem theory as proposed by Tarantola (1987). We want to estimate the model parameters $\mathbf{m}$ that best explains the observed data $\mathbf{d}^r$. Data and model are linked by a relationship g that contains the physics of wave propagation. The misfits between the picked data $\mathbf{d}^r$ and the calculated data
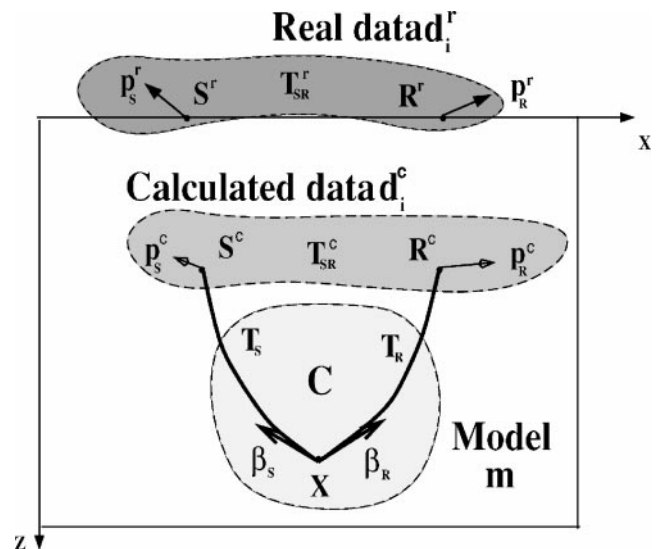

FIG. 2. Calculated data $\mathbf{d}_i^c = [(S^c, R^c, P_S^c, P_R^c, T_{SR}^c)_i]_{i=1}^N$ are from an initial model $\mathbf{m} = [[(X, \beta_S, \beta_R, T_S, T_R)_i]_{i=1}^N, [C_j]_{j=1}^J]$. They do not match real data $\mathbf{d}_i^r = [(S^r, R^r, P_S^r, P_R^r, T_{SR}^r)_i]_{i=1}^N$. The misfits will be used to update the model.

$\mathbf{d}^c = g(\mathbf{m})$ are minimized in the ℓ^2 sense. The cost function S(**m**) can be expressed as

$$S(\mathbf{m}) = \frac{1}{2}\big[(g(\mathbf{m}) - \mathbf{d}^r)^T \mathbf{C}_\mathbf{D}^{-1}(g(\mathbf{m}) - \mathbf{d}^r)$$
$$+ (\mathbf{m} - \mathbf{m_{prior}})^T \mathbf{C}_\mathbf{M}^{-1}(\mathbf{m} - \mathbf{m_{prior}})\big], \quad (2)$$

where $\mathbf{m_{prior}}$ is the a priori model, $g(\mathbf{m}) = \mathbf{d}^c$, **D** is the data space, **M** is the model space, and the superscript ^{T} denotes the adjoint operator. $\mathbf{C_M}$ and $\mathbf{C_D}$ are the covariance matrices in the model space and data space, respectively. We detail how they are built in the "A priori information" subsection below. Finding the minimum value of S(**m**) is a nonlinear problem. Since fully nonlinear approaches (Monte Carlo or simulated annealing techniques) are too costly to succeed for problems with high-dimensional data and model spaces, a common practice consists of converging iteratively through successive locally linearized solutions. The linearized solution of equation (2) can be expressed as

$$\mathbf{m}_{k+1} = \mathbf{m}_k - \left(\frac{\partial^2 S}{\partial \mathbf{m}^2}(\mathbf{m}_k)\right)^{-1} \frac{\partial \mathbf{S}}{\partial \mathbf{m}}(\mathbf{m}_k). \quad (3)$$

By introducing **G**, the matrix that contains the Fréchet derivatives of the data with respect to the parameters of the model **m** at iteration k,

$$\mathbf{G}_k = \left[\frac{\partial g^D}{\partial \mathbf{m}^\mathbf{M}}\right]_{\mathbf{m}_k}, \quad (4)$$

the iterative scheme of equation (3) can finally be written in the form

$$\mathbf{m}_{k+1} = \mathbf{m}_k - \big(\mathbf{G}_k^T \mathbf{C}_\mathbf{D}^{-1}\big(g(\mathbf{m}_k) - \mathbf{d}^r\big)\mathbf{G}_k + \mathbf{C}_\mathbf{M}^{-1}\big)^{-1}$$
$$\times \mathbf{G}_k^T \mathbf{C}_\mathbf{D}^{-1}\big(g(\mathbf{m}_k) - \mathbf{d}^r\big) + \mathbf{C}_\mathbf{m}^{-1}(\mathbf{m}_k - \mathbf{m_{prior}}). \quad (5)$$

STEREOTOMOGRAPHY: PRACTICAL ASPECTS

Stereotomography consists of picking locally coherent events in the data $\mathbf{d}^r$, calculating synthetic picks $\mathbf{d}^c$ in an a priori model, and then using the misfits between the two to update the model parameters.

Data picking

In stereotomography, the data space **D** consists of a set of N locally coherent events **d** (Figure 2):

$$\mathbf{d} = [(S, R, P_S, P_R, T_{SR})_i]_{i=1}^N, \quad (6)$$

where S and R are the source and receiver positions; P_S and P_R are the slopes in common receiver and in common source gathers, respectively; and T_{SR} is the two-way traveltime. Picking is a major challenge for stereotomography applications. While picking traveltimes is a well-known and routinely used procedure, picking local slopes is less mature. A local slope at a given time is obtained from a data gather by forming local slant stacks (i.e., weighted summation over a group of traces located near a central reference trace). The local stack will be optimal for the local slope value. In a seismic line, the procedure is repeated for a sparse set of reference traces. In each trace, weighted slant stacks are calculated at each time step for a set of slope values. The weighting function we use is a Hamming window $[H_L(x) = 0.5 + 0.5\cos(\pi(x/L))$ for $x \in [-L, L]$ and $H_L(x) = 0$ outside]. Local coherency is measured by a normalized parameter, whose value is one for a perfectly coherent event (perfect crosscorrelation and perfectly linear moveout in the entire gather), and zero if no coherency exists at all. This parameter is a normalized envelope of trace semblance (see Figure 3). We use semblances because this has proven to be effective for separating interfering events (Taner et al., 1979) . The coherency parameter is

$$A(p, t; x_{ref})$$
$$= \frac{\left|\sum\limits_{traces} (H_L(x - x_{ref})\hat{u}(x, t - p(x - x_{ref})))\right|^2}{\sum\limits_{traces} (H_L(x - x_{ref})|\hat{u}(x, t - p(x - x_{ref}))|^2)}, \quad (7)$$
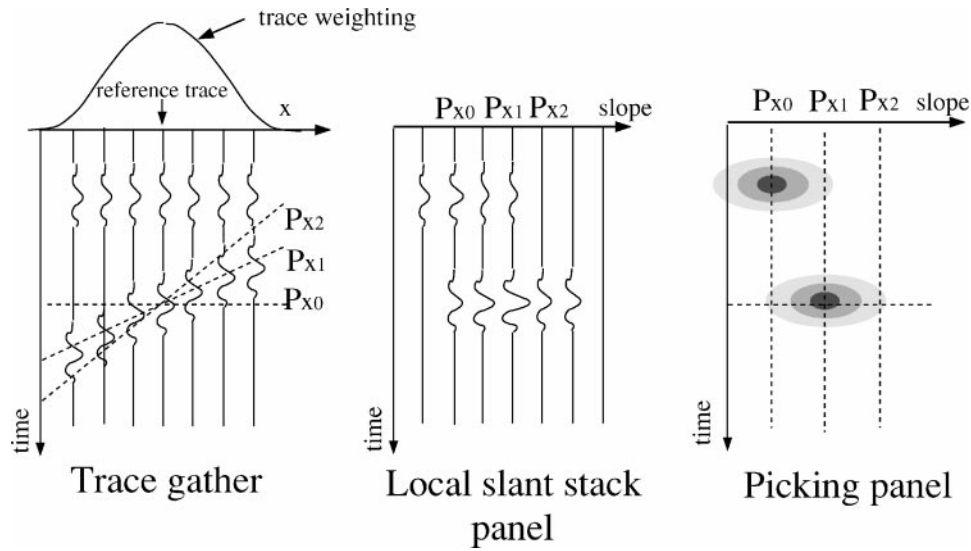

FIG. 3. Computation of the slope picking panels. From an initial trace gather (left), we compute a local slant stack panel (center) weighted around the reference trace using a Hamming window. The final panel on which we pick our data (right) corresponds to the normalized squared envelope of the local slant-stack panel.

where x denotes the position of the trace in the gather, x_{ref} the position of the reference trace, p the investigated slope, $H_L(x)$ the weight of the normalized Hamming term, and $\hat{u}(x,t) = (1 + i\,\text{Hilb})\,u(x,t)$ is the analytical trace. Indeed, this operation consists of a transformation from a trace gather with (distance-time) dimensions (Figure 4 left) to a semblance panel with (slope-time) dimensions for a given reference trace (Figure 4 right). It has to be performed on two data gathers to
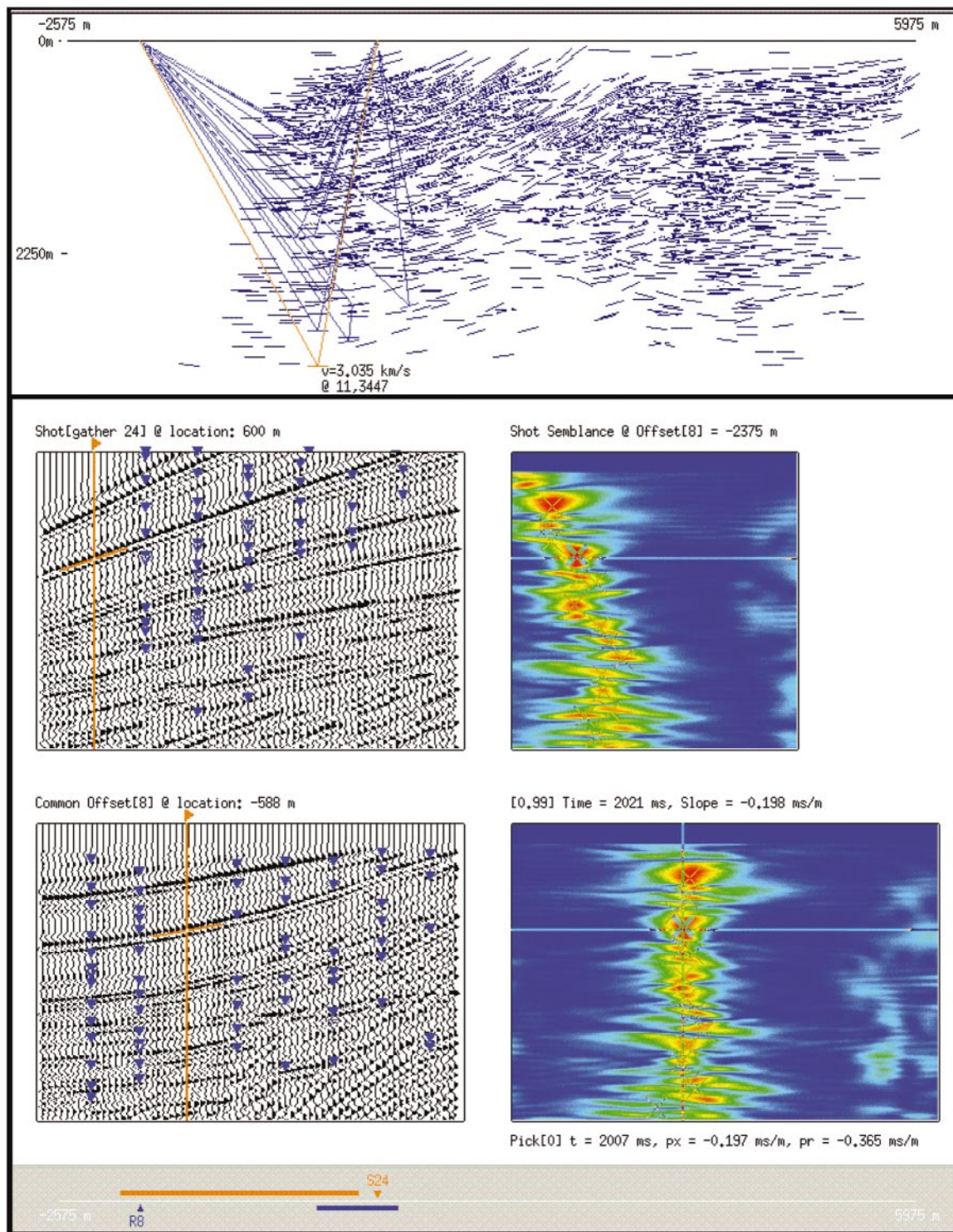

FIG. 4. The slope picking tool is composed of trace gathers (left), associated with semblance panels (right), and an equivalent model QC window (top). Each pick is the combination of two maxima in the semblance panels for the same traveltime. It corresponds to one traveltime and two local slopes for a reference trace (noted with a flag) and can be represented by an equivalent local dip in the top window.

obtain the slopes at both source and receiver locations. Picks are collected in the semblance panels, and slopes are visualized on the traces themselves for quality control (QC). Each pick is a set of two points, with maximum amplitude values picked simultaneously in the two semblance panels for the traveltime along the reference trace. An interactive slope-picking tool was built to manually collect traveltimes and slopes (Figure 4). In the case presented in Figure 4, the two trace collections chosen are common shot (center left) and common offset gathers (lower left). For a given reference trace (indicated by a flag), two associated semblance panels are computed on the fly (right panels). At the very top of Figure 4, an additional window interactively displays a QC panel for the location of the reflection/diffraction point and the average velocity in a homogeneous velocity approximation (see Figure 5). This equivalent model $[X_{eq}, V_{eq}]$ is built using basic geometric rules for a given pick. It gives a rough idea of the a priori spatial distribution of the reflection/diffraction points, but also allows us to filter out abnormal $[X_{eq}, V_{eq}]$ pairs (e.g., a reflection at a depth of 4 km with an average velocity less than 1500 m/s). Furthermore, the equivalent model may be used as an initial guess for updating the ray paths. More details, including analytic expressions, are given in Appendix A.

Since the signal-to-noise ratio is improved after stacking the gathers and no structural interpretation is needed, the slope picking operation can be fully automated. Several control parameters allow tuning the number and the reliability of the automatically collected picks: the distance between analyzed traces, the time sampling increment, the width of the Hamming window, a threshold on the semblance value, and a threshold on the product of the semblances from common shot, common receiver, and common offset gathers. Due to the uncorrelated nature of the local events, the automatic picking procedure is quite fast. Typically, the operation requires less than an hour for a marine 2D line on one workstation. An interesting usage of the equivalent model QC is that it can indicate under- and over-picked areas. In a second run, the automatic picker can add or decimate data in a given delineated area (e.g., while tuning the semblance threshold for given equivalent model boundaries). As any other method that requires data picking, stereotomography is sensitive to coherent noise in the data, such as surface-related multiples. It is important to point out that no robust antimultiple criterion has been implemented in the automatic picker. We assume that both water-bottom and inter-bed multiple reflections have been taken care of at an earlier stage. On the other hand, diffractions and trip-

licated arrivals can be picked as well as any simple primary reflection.

Model parameter estimation

In stereotomography, we estimate the subsurface velocity field and the ray paths modeling the wave propagation for a set of identified reflected/diffracted events. All of these parameters need to be initialized and will be updated.

Initialization.—The initial velocity field is usually described by a minimum number of nodes, four in each direction for cubic B-splines. In most cases, the initial velocity model is homogeneous, each C_j weight being set to an average value of the velocity at the surface. In general, any velocity grid after smoothing can be described by a set of B-spline weights and then used as an initial guess.

The locations of the initial diffraction/reflection points X in depth are given by the equivalent models calculated in the picking process. The angles are initialized through simple geometrical considerations in a homogeneous approximation. Each of the two one-way traveltime pairs is set to half of the corresponding two-way traveltime picked from the data. The initial ray paths do not usually reach the surface and are very often far from explaining the picked data. It is possible to improve the initial ray paths by solving a simplified version of the inverse problem for a fixed velocity model. In fact, we use a local iterative optimization of each pair of ray segments individually, trying to fit the data while solving the ray paths only, assuming the initial velocity model is correct.

The Fréchet derivative matrix is small (seven parameters in the data space for six parameters in the model space) and well-conditioned. The inverse problem is solved using the Gauss-Newton approach with a singular value decomposition (SVD) of the Fréchet derivative matrix constrained by a damping factor. This optimization always converges very quickly toward a solution that already reduces the cost function considerably. Rays are traced upwards using a second-order Runge-Kutta algorithm. Because rays can exit the model boundaries, virtual extensions of the velocity model are introduced.

A priori information.—A priori information in the data space is introduced as a diagonal covariance matrix $\mathbf{C_D}$ that contains typical measurement uncertainties for each class of picked data. These quantities might be chosen regarding the specificity of a given data set, although generic values (5 m for the positions, 5 ms for the traveltimes, and 0.01 s/km for the slopes) were used in the examples. This diagonal matrix is composed of a series of constant values for each class of picked data. A priori information in the model space is introduced as a covariance matrix $\mathbf{C_M}$ in the form of computational regularization terms. These include a damping factor ϵ for all of the parameters and a smoothing term Δ on the velocity model parameters, defined as the centered finite difference approximation of the 2D spatial Laplacian operator:

$$\Delta = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \tag{8}$$

There are two tuning parameters that control the inverse problem. A weight μ applied on the operator Δ will control the smoothness of the updated model. The damping factor will set
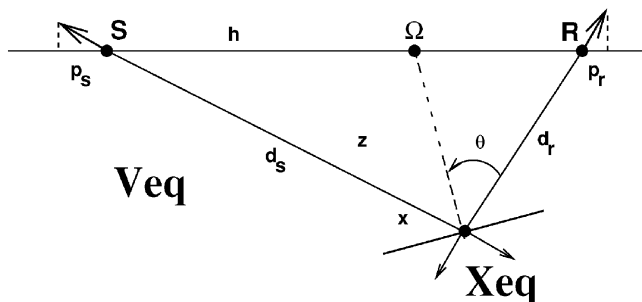

FIG. 5. The equivalent model in a homogeneous approximation is defined by a set of [position X_{eq}, velocity V_{eq}] pairs obtained from picked data $[S, R, P_S, P_R, T_{SR}]$.

how far from the initial model we want to investigate the model space. The choice of an adequate μ value is always a problem. Pratt and Chapman (1992) proposed a consistent stochastic inverse method to select the appropriate degree of regularization. It consists of investigating a range of μ values to produce a plot of data residuals as a function of the velocity roughness (both in a rms sense). Such curves (Figure 6) typically show a knee point that corresponds to a good estimate of the optimal level of regularization (large μ values produce excessive residuals, and small μ values produce excessive roughness). A correct level of smoothing will significantly compensate for both low ray coverage and edge effects.

Fréchet derivatives.—The Fréchet derivatives contained in $\mathbf{G}_k$ [see equation (5)] are calculated using paraxial ray tracing (Chapman, 1985; Farra and Madariaga, 1987). Paraxial ray theory is a powerful tool that gives access to first-order perturbations of a ray path while perturbing its initial conditions or the background velocity. It requires the second-order derivatives of the velocity model to be continuous, a condition that is honored by our parameterization choice (cubic B-splines). Both position and slowness vector perturbations are calculated on the fly in our forward modeling, using the propagator matrix (Aki and Richards, 1980). All of the details of Fréchet derivative calculations are given in Billette and Lambaré (1998).

Inversion.—Since the ray parameters describing the ray paths and the velocity model are coupled in the inversion (Billette and Lambaré, 1998), a joint inversion of all of them is the proper approach. Consequently, the inverse problem to be solved in stereotomography is large, sparse, and ill-conditioned. Thus, the solution of each linearized inversion is obtained using the conjugate-gradient method LSQR (Paige and Saunders, 1982) that is well adapted to this type of problem because it allows us to skip computing the null contributions. We compress the Fréchet derivative matrix by storing only non-null values. This has a significant impact on memory requirements for real-size applications. As recommended by Paige and Saunders (1982), we apply a column normalization to the augmented Fréchet derivatives matrix (energy of columns is equalized) before LSQR. This insures that each class of model parameters contributes significantly to the solution. LSQR is an iterative algorithm. Various tests have shown that a limited number of iterations, corresponding to a small conditioning number, are sufficient because several nonlinear iterations are also done. This will be illustrated in our applications.
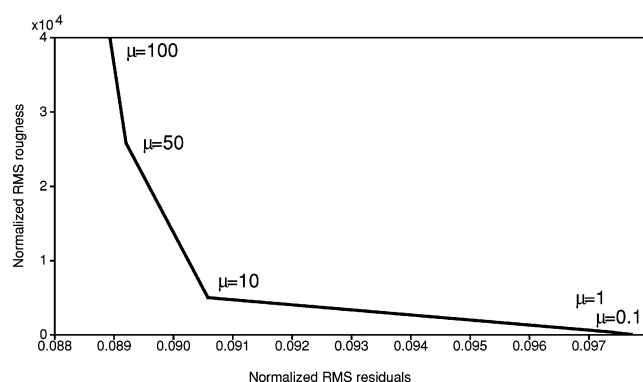

FIG. 6. Example of roughness versus residuals for a real-data case.

Multiscale optimization.— In their first applications, Billette et al. (1998) used a three-step optimization procedure: (1) invert for a homogeneous velocity $V(x, y, z) = V_0$, (2) invert for a vertical gradient $V(x, y, z) = V_0 + V_1.z$, and finally (3) compute $V(x, y, z) = V_0 + V_1.z + \Delta V_{\text{B-splines}}(x, y, z)$. This strategy was intended for compaction-driven shallow velocity regimes. It may fail for the case of complex overburden where the second step may not converge to a reasonable solution. In order to obtain a robust, fast, and automatic optimization scheme, we use a multiscale approach for the estimation of the velocity field (Le Begat al., 2000). The initial B-splines grid is resampled when the cost function stops, significantly decreasing with each new iteration. Additional velocity nodes are introduced in the model, alternately in vertical and horizontal directions. The philosophy of the multiscale optimization is to search for progressively shorter and shorter velocity wavelengths as longer ones are resolved, hopefully thus converging to the correct solution. In the next section, we demonstrate this method on a synthetic data set.

APPLICATION TO THE "MARMOUSOFT" SYNTHETIC DATA SET

Data

The 2D Marmousi model and data set have been considerably documented since their release by the Institut Français du Pétrole (IFP) (Bourgeois et al., 1991) and have been used to test several different velocity analysis and migration techniques. None of the methods that have been presented since then were able to image the target zone properly. The main difficulties are the model's structural complexity, coupled with multiples that are hard to attenuate. Since we do not want to challenge both of these problems together, we created an adapted 2D Marmousi synthetic data set using ray+Born modeling (Lambaré et al., 1992) that provides us with a data set free of multiple and refracted arrivals. The ray+Born modeling requires a velocity macromodel and a short-wavelength perturbation model. Our velocity macro model consists of a low-pass filtered version [to 480 m or $\tau = 240$ m as defined by Operto et al. (2000)] of the exact Marmousi model (Figure 7 top). The filtered model has been smoothed so that the ray field only has single arrivals. The short wavelength perturbation was a high-pass filtered version of the exact Marmousi velocity and density models [to $\tau = 75$ m as defined by Operto et al. (2000)]. We called the new model "Marmousoft" as it is a simplified version of the exact Marmousi model. In fact, Marmousoft is the combination of a smooth velocity field and a very complex reflectivity, and thus it is an ideal data set for testing the practical aspects of our method. The first-arrival ray+Born preserved amplitude prestack depth migration using the exact Marmousoft velocity model is presented in Figure 8 (top).

A 2D towed streamer acquisition was simulated, consisting of 261 shots recorded on 96 traces with a 25-m spacing. The offsets range is 100–2450-m. The time sampling interval is 4 ms. A mute was applied to the shot gathers to filter out the direct arrival.

Stereotomography

A total of 3222 locally coherent events were automatically picked using a 125-m-long summation window *L*

(corresponding to 10 traces). Picking was done every 125 m in the data cube (common source, common receiver, common midpoint, and common offset trace gathers). The semblance threshold was set to 0.8. We introduced measurement errors into the covariance matrix of 5 m, 5 ms, and 0.01 s/km for positions, traveltimes, and slopes, respectively. Figure 9a shows the initial spatial distribution of the picked reflected/diffracted events positioned in depth using the homogeneous approximation of the equivalent model. As expected on these synthetic data, the quality of the model coverage was good.

In order not to bias our results by the known exact solution, the initial model (Figure 9a) was set to a constant 2 km/s velocity field, with a 1 km × 1 km B-spline node spacing. In general, we believe it is better to consider a too-slow rather than a too-fast initial velocity model, because (1) it keeps the ray paths inside the model boundaries and (2) the shallow velocities are already close to the solution and will be updated promptly, limiting the propagation of errors and allowing a layer-stripping type of convergence. Initial ray-segment paths were directly obtained from picked data, using the equivalent model information (see Appendix A). In this example, the cost function was reduced by three orders of magnitude in the first five iterations where we updated the ray paths only (Figure 10).

Next, joint inversion of the entire stereotomography model vector (ray paths and velocity nodes) was applied. The tuning parameter μ remained fixed for all resamplings of the velocity grid. For each nonlinear iteration, 100 LSQR iterations (corresponding to a conditioning number of about 1000) were sufficient for the cost function to converge. The evolution of the cost function as a function of iterations and velocity resampling (Figure 10) shows together with Figures 9a–9f that many iterations were required to estimate the long wavelengths features

of the velocity, and that very few iterations were necessary to add shorter wavelength variations. In this application, our multiscale optimization scheme allowed us to significantly decrease the computing time.

Results

A practical way of visualizing the ray segments is to represent them as dip bars obtained from optimized X positions and takeoff angles, and to overlay them on the velocity model. In Figures 9a–9f, we present the model updates obtained with stereotomography. The velocity resampling was stopped when we reached 250-m B-spline node spacing in both lateral and vertical directions (Figure 9f). The comparison between the updated model and the exact velocity model (Figure 7) shows great similarity. Some local areas have not been correctly retrieved, especially in the deep parts of the model, because very few reliable picks originated from there. In general, we can see that our local approach was able to align picks along continuous horizons and even faults when needed. Note that by overlaying the updated dip bars on the velocity model, we can see that the dip bars form a "skeleton" of the migrated image, analogous
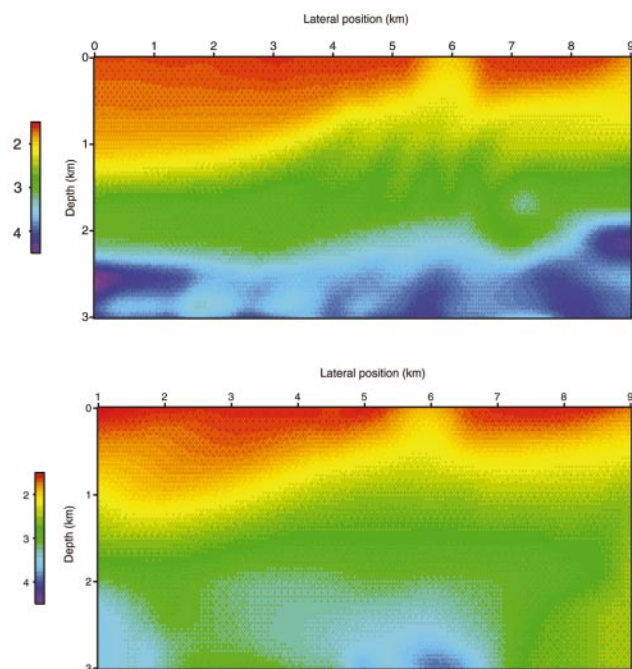

FIG. 7. Marmousoft synthetic example: (Top) exact velocity model, (bottom) final velocity model obtained from stereotomography.
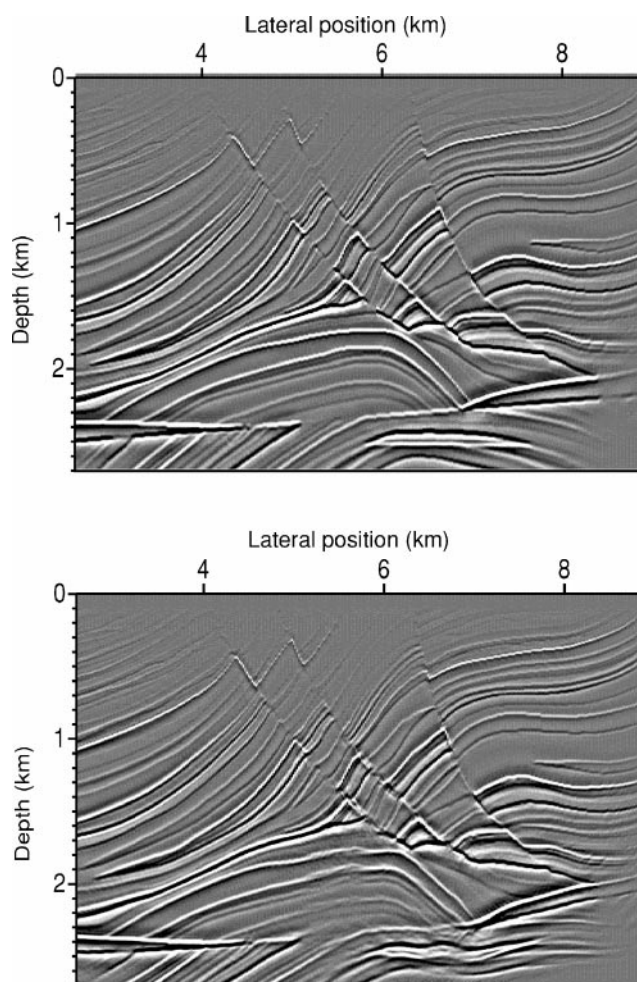

FIG. 8. Marmousoft synthetic example: prestack depth migration using (top) the exact velocity model, (bottom) the final velocity model obtained using stereotomography.

to old-style migration of reflectors by hand. Using the updated velocity model, we applied a first-arrival preserved-amplitude prestack depth migration (Thierry et al., 1999) to the data. Since the data only contains "first" arrivals, this should ide-

ally give a perfect image. In Figure 8, we can compare the image obtained using the velocity model from stereotomography with the image migrated using the exact model. In the shallow part of the section, the quality of the migrated image
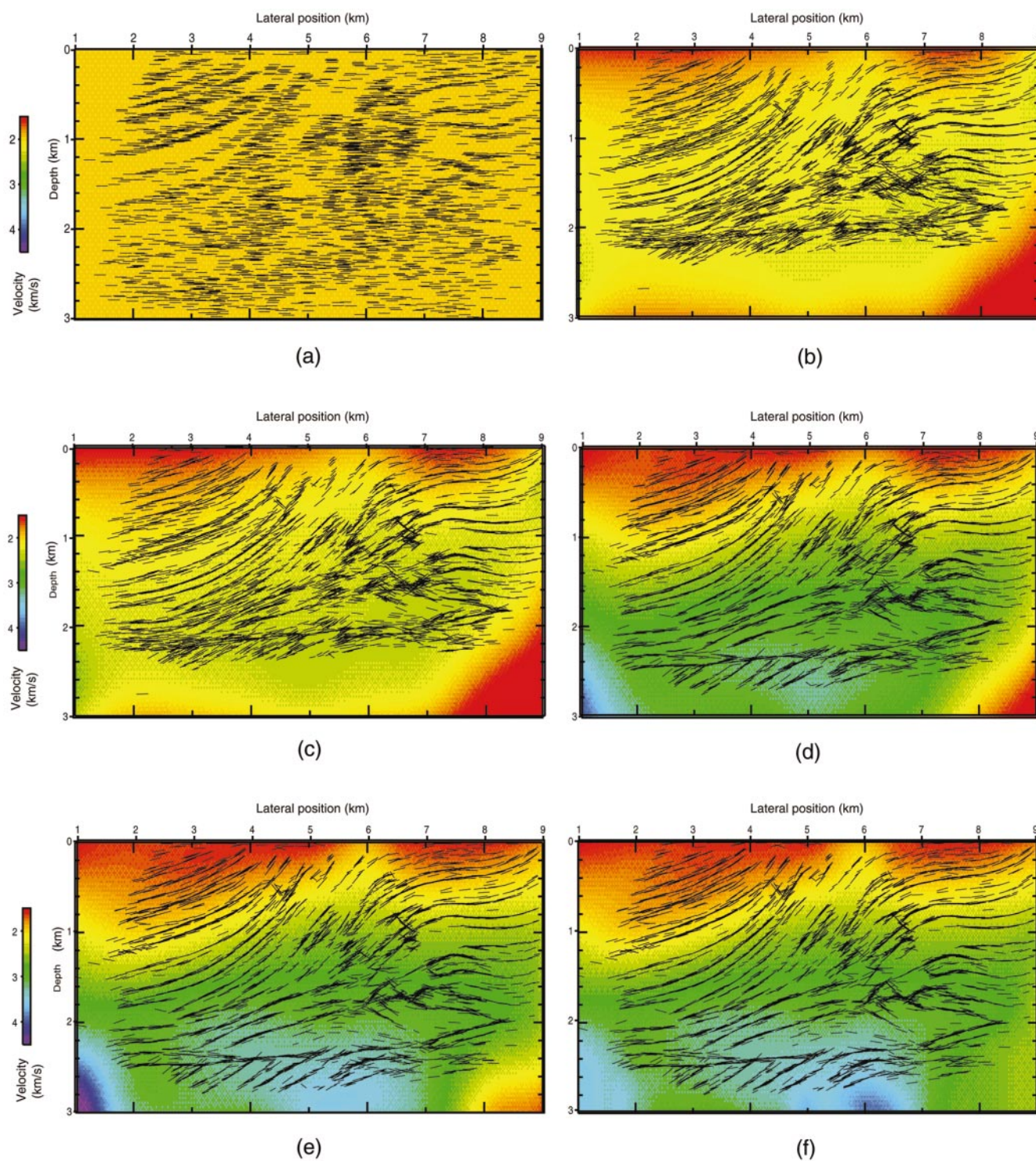

FIG. 9. Marmousoft synthetic example velocity models and dip bars (representing ray paths in the current velocity model) at various iterations of stereotomography: (a) initial model with a (1 km × 1 km) node spacing, (b) 5th iteration with a (1 km × 1 km) node spacing, (c) 10th iteration with a (1 km × 1 km) node spacing, (d) 50th iteration with a (1 km × 1 km) node spacing, (e) with a (500 m × 500 m) node spacing , (f) final model with a (250 m × 250 m) node spacing.

obtained using stereotomography is good, both in terms of focusing and reflector continuity. In particular, the faults are clearly defined and well-positioned. As expected, the areas that were not well-covered by the picking operation are not perfectly imaged. The target zone (lateral position 7 km, depth 2.4 km) is well-positioned, but we did not manage to reproduce the flat reflectors with their correct geometry. In Figure 11, we present a selection of migrated CRP gathers that are all reasonably flat. The small columns represent the CRP gathers migrated with the exact velocity model. This comparison shows we imaged the major reflectors to almost a perfect depth in the left and center part of the model (to 5 km), but the results deteriorate beneath 1.5 km in the right part of the section, where stereotomography could not find a satisfactory answer due to poor illumination and limited offsets.
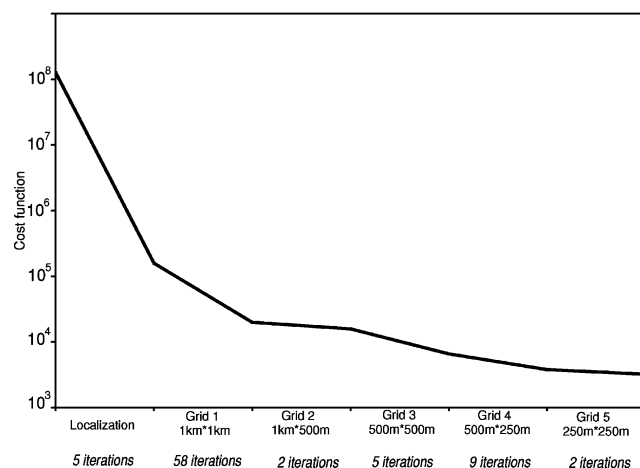

FIG. 10. Marmousoft synthetic example cost function as a function of iterations during multiscale optimization. Note that the first step (five iterations) corresponds to the initial localization (ray paths only).
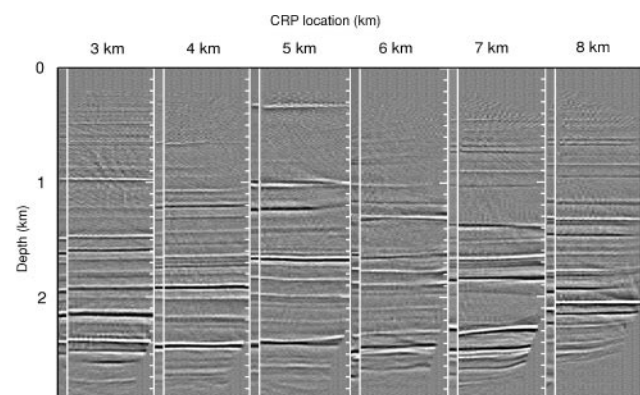

FIG. 11. Marmousoft synthetic example CRP migrated gathers for various CRP locations. The small columns correspond to the shortest offsets of the migration with the exact model. They are compared side by side to the full offset range of CRP gathers migrated using the model updated with stereotomography (large columns).

APPLICATION TO A 2D MARINE LINE

Data

We applied stereotomography on a 2D line extracted from a 3D towed-streamer marine survey shot in West Africa for TotalFinaElf. This data set was provided to us by TotalFinaElf after a crossline isotropic prestack time migration. It contains 1011 shots with 92 hydrophones recorded with a 4-ms time sampling interval. The offset range is 100–2380 m. We applied a [5, 20, 45, 65] Hz band pass filter to the data. The 500-m common offset section is shown in Figure 12 before further processing.

Stereotomography

The automatic picking tool collected 2860 events in less than one hour. The semblance threshold was set to 0.8. No particular
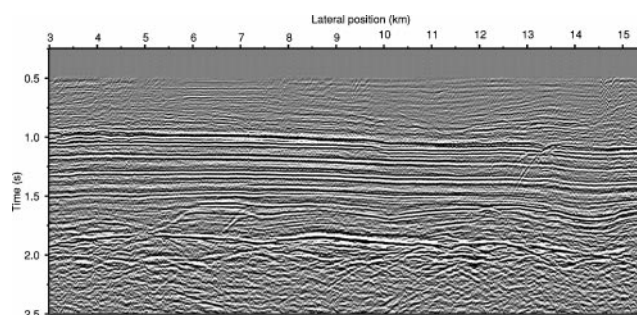

FIG. 12. Application to real data: the common-offset time-migrated section (500-m offset) shows horizontal layers on the top of a more complex structure.
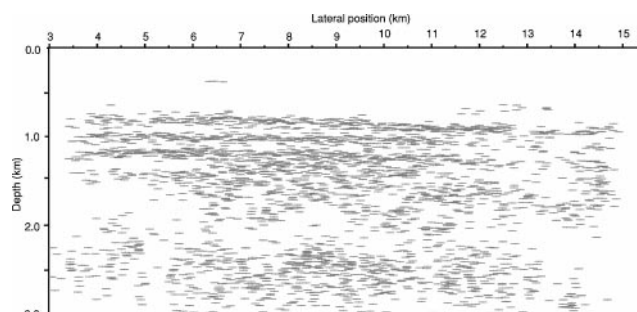

FIG. 13. Application to real data: initial picked events located in depth using the homogeneous approximation of the equivalent model.
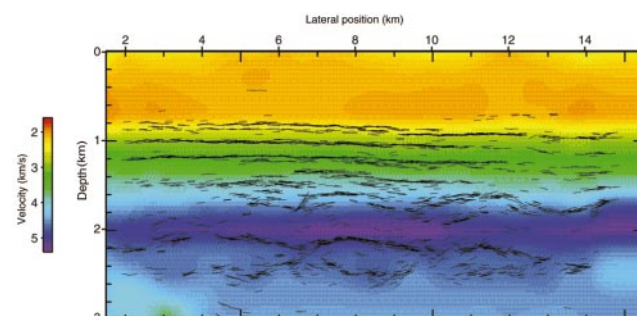

FIG. 14. Application to real data: final velocity model obtained with stereotomography.

knowledge of the area was available, so we proceeded as in our previous synthetic example: the initial model was composed of a homogeneous velocity field described by 60 nodes (2 km lateral spacing and 1 km vertical spacing), all set to the water velocity, and 2860 ray-segment pairs built using picked traveltime and slope values (Figure 13). The spatial distribution of the reflection/diffraction points was quite good except in the shallow part, where almost no picks were collected down to 800-m depth.

The optimization of the ray segments converged in three iterations, which reduced the cost function from 10^8 to 4.6×10^5. A maximum of 150 LSQR iterations was imposed for each nonlinear iteration. Our final result using stereotomography was obtained after 47 nonlinear iterations. The final value of the cost function is 5100. The final velocity model was described by a total of 840 nodes, with a final sampling of 250 m laterally and 125 m vertically.
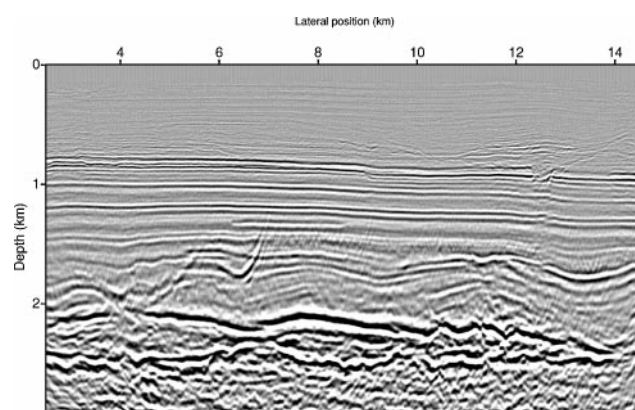

FIG. 15. Application to real data: prestack depth migration with the velocity model updated using stereotomography.

Results

The final velocity model interleaved with the updated dip bars is presented in Figure 14. The velocity regime is quasi-1D down to about 1.8-km depth, where we found a high-velocity zone (that may represent strongly compacted limestones). Underneath, the velocity slows down, and considerable lateral velocity contrasts are present. It should be pointed out that these sharp velocity contrasts were retrieved using a smooth velocity model description.

Using the updated velocity model, we applied a first-arrival preserved amplitude prestack depth migration (Thierry et al., 1999) to the data. The migrated image is shown in Figure 15, and is reasonably well-focused. A channel at 12.5-km lateral position and 800 m depth is well-imaged. Tilted blocks in the very high velocity zone are also clearly defined from 1.4-km to 2.1-km depth. Finally, the deeper structure indicated by the dip bars is seen to be continuous in the migrated stack, and several salt bodies can be delineated.

The CRP gathers (Figure 16) are fairly flat down to about 2-km depth. Below that, there is still some slight residual moveout. It seems that stereotomography underestimated the velocities in the limestones above the salt. This problem may come from a velocity sampling that was too sparse to capture some short wavelength anomalies. Stereotomography was successful for this application and provided an accurate velocity model in a very short time frame and in a very automated way.

CONCLUSIONS AND PERSPECTIVES

Stereotomography, as introduced by Billette and Lambaré (1998), was very promising from a theoretical point of view. We have shown that an advanced analysis of locally coherent
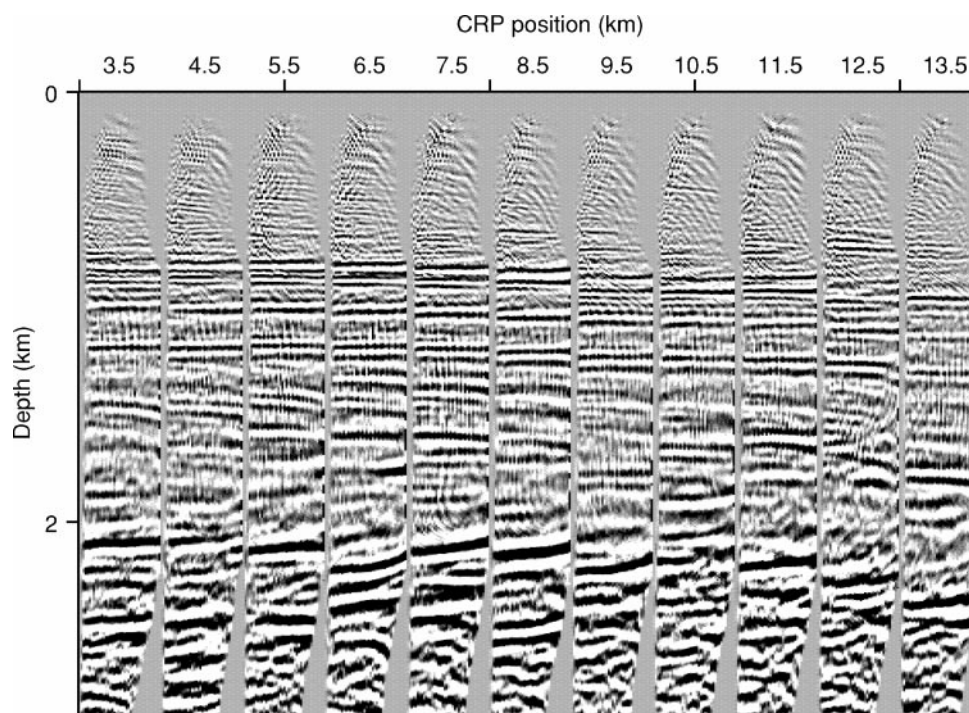

FIG. 16. Application to real data: CRP migrated gathers obtained using the model updated using stereotomography.

events in real seismic data can also be used as a basis of a practical velocity-estimation approach. In particular, we presented:

1) An efficient dedicated automatic picking tool that collects information from locally coherent events on semblance panels with visual QC capabilities.

2) A smooth velocity-model description that allows the use of ray-based prestack depth-imaging tools with an adaptive model representation.

3) A stable and efficient multiscale optimization approach.

The two applications of the method we presented were conducted in a very short time frame. Stereotomography was easy to use and efficient in both cases. Both applications were realized with a minimum of human intervention and both converged to a geologically plausible model that produced a good quality image.

Two extensions of this work on stereotomography are underway. The first developments and synthetic tests in 3D have been recently presented by Chalard et al. (2000), who demonstrated, in theory, the applicability of the method in 3D for conventional towed-streamer acquisitions. The first application with real multistreamer data is currently underway. Another very promising extension of stereotomography concerns the joint estimation of P and S velocity fields from multicomponent data (Granger et al., 2000; Alerini et al., 2001). Indeed, in contrast to how most layer-based velocity estimation methods are implemented, stereotomography does not require given events to be interpreted simultaneously in PP and PS sections, which can be a major challenge in building P and S velocity models.

ACKNOWLEDGMENTS

REFERENCES

Aki, K., and Richards, P. G., 1980, Quantitative seismology: Theory and methods: W. H. Freeman & Co.

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

Alerini, M., Le Bégat, S., Lambaré, G., and Boelle, J.-L., 2001, Preserved amplitude depth imaging: application to an OBC line (mahogany): 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 833–836.

Audebert, F., Guillaume, P., Zhang, X., and Jones, I., 1998, CRP scan—Solving 3D pre-SDM velocity analysis with zero offset tomography: 60th Ann. Mtg., Eur. Ass. Expl. Geophys, Expanded Abstracts, 1–5.

Billette, F., 1998, Estimation de macro-modéles de vitesse en sismique réflexion par stéréotomographie: Ph.D. diss., Université Paris VII.

Billette, F., Etgen, J., and Rietveld, W., 2002, The key practical aspects of 3D tomography: Data picking and model representation: 64th Ann. Mtg., Eur. Ass. Geosc. Eng., Expanded Abstracts, B006.

Billette, F., and Lambaré, G., 1998, Velocity macro-model estimation from seismic reflection data by stereotomography: Geophys. J. Internat., **135**, 671–680.

Billette, F., Podvin, P., and Lambaré, G., 1998, Stereotomography with automatic picking: Application to the Marmousi dataset: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1317–1320.

Biondi, B., and Sava, P., 1999, Wave-equation migration velocity analysis: 69th Ann. Internat, Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1723–1726.

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

Bourgeois, A., Bourget, M., Lailly, P., Poulet, M., Ricarte, P., and Versteeg, R., 1991, Marmousi, model and data: 52nd Ann. Mtg., Eur. Ass. Expl. Geophys, The Marmousi Experience, 5–16.

Broto, K., and Ehinger, A., 1998, How to access 3D prestack kinematic information in case of complex geologies?: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1401–1404.

Chalard, E., Podvin, P., Lambaré, G., and Audebert, F., 2000, Estimation of velocity macro-models by 3-D stereotomography: 70th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 2257–2260.

Chapman, C. H., 1985, Ray theory and its extensions: WKBJ and Maslov seismogram: J. Geophys., **58**, 27–43.

Chauris, H., Noble, M., Lambaré, G., and Podvin, P., 2002, Migration velocity analysis from locally coherent events in 2-D laterally heterogeneous media, Parts I and II: Geophysics, **67**, 1202–1224.

Clarke, R., Alazard, B., Pelle, L., Sinoquet, D., Lailly, P., Delprat-Jannaud, F., and Jannaud, L., 2001, 3D traveltime reflection tomography with multi-valued arrivals: 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1875–1878.

Dor, P., and Labossière, A., 1984, Stereotomography: Engineering Geology, **20**, 311–324.

Etgen, J. T., 1991, Residual prestack migration and interval-velocity estimation: Ph.D. diss., Stanford University.

Etgen, J., and Biondi, B., 2001, SEG workshop "Beyond kirchhoff imaging": 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Workshop Abstracts, in press.

Farra, V., and Madariaga, R., 1987, Seismic waveform modeling in heterogeneous media by ray perturbation theory: J. Geophys. Res., **92**, 2697–2712.

Granger, P., Herrenschmidt, A., Audebert, F., Gerea, C., Etienne, G., Stoppin, A., Bégat, S. L., Alerini, M., Lambaré, G., Berthet, P., Nebieridze, S., and Boelle, J., 2000, Comparison of different strategies for velocity model building and imaging of p and ps real data, *m* Recent advances in shear wave technology for reservoir characterization: A new beginning?: Soc. Expl. Geophys. Research Workshop Abstracts.

Gray, S., 2000, Velocity smoothing for depth migration: How much is too much?: 70th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1055–1058.

Guiziou, J. L., Mallet, J. L., and Madariaga, R., 1996, 3-D seismic reflection tomography on top of the GOCAD depth modeler: Geophysics, **61**, 1499–1510.

Lailly, P., and Sinoquet, D., 1996, Smooth velocity models in reflection tomography for imaging complex geological structures: Geophys. J. Internat., **124**, 349–362.

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

Le Bégat, S., Podvin, P., and Lambaré, G., 2000, Application of 2-D stereotomography to marine seismic reflection data: 70th Ann. Internat Mtg., Soc. Expl. Geophys., Expanded Abstracts, 2142–2145.

Le Stunff, Y., and Grenier, D., 1998, Taking into account a priori information in 3D tomography: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1875–1878.

Liu, Z., 1997, An analytical approach to migration velocity analysis: Geophysics, **62**, 1238–1249.

Lucio, P. S., Lambaré, G., and Hanyga, A., 1996, 3D multivalued travel time and amplitude maps: Pageoph, **148**, 449–479.

Mispel, J., and Hanitzsch, C., 1996, The use of layered or smoothed velocity models for prestack Kirchhoff depth migration: 66st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 519–521.

Operto, S., Xu, S., and Lambaré, G., 2000, Can we image quantitatively complex models with rays?: Geophysics, **65**, 1223–1238.

Paige, C., and Saunders, M. A., 1982, LSQR: Sparse linear equations and least squares problems, Part I and Part II: ACM Trans. Math. Soft., **8**, 43–71.

Pratt, R. G., and Chapman, C. H., 1992, Traveltime tomography in anisotropic media: Geophys. J. Internat., **109**.

Riabinkin, L. A., 1957, Fundamentals of resolving power of controlled directional reception (CDR) of seismic waves, in Slant stack processing: Soc. Expl. Geophys., **14**, 36–60 (translated and paraphrased from Prikladnaya, **16**, 3–36).

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

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

Taner, M. T., Koehler, F., and Sheriff, R., 1979, Complex seismic trace analysis: Geophysics, **44**, 1041–1063.

Tarantola, A., 1987, Inverse problem theory: Methods for data fitting and model parameter estimation: Elsevier.

Thierry, P., Operto, S., and Lambaré, G., 1999, Fast 2D ray-Born inversion/migration in complex media: Geophysics, **64**, 162–181.

Versteeg, R., 1993, Sensitivity of prestack depth migration to the velocity model: Geophysics, **58**, 873–882.

Whiting, P., 1998, Reflection tomography without picking: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1226–1229.

Woodward, M., Farmer, P., Nichols, D., and Charles, S., 1998, Automated 3D tomographic velocity analysis of residual moveout in prestack migrated common image point gathers: 68th Ann. Internat Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1218–1221.

Xu, S., Lambaré, G., and Calandra, H., 2000, 3D migration/inversion in complex media : Application to SEG/EAGE salt model: 70th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 886–889.

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

APPENDIX A

In stereotomography, the elementary datum is a locally coherent event picked in prestack seismic data and is characterized by five parameters: T_{SR}, P_S, P_R, x_S, x_R. T_{SR} is the event's two-way traveltime. S, R respectively represent the source and the receiver, x_S and x_R give their locations, and P_S and P_R are the event's slopes as measured in the data, defined by:

$$\begin{cases} P_S = \left(\dfrac{\partial T_{SR}}{\partial x_S}\right)_{x_R} \\ P_R = \left(\dfrac{\partial T_{SR}}{\partial x_R}\right)_{x_S} \end{cases}. \tag{A-1}$$

As already identified by Sword (1987), it is possible, in most cases, to attach such an event to a simple model where it is interpreted as the result of single diffraction of the downgoing wave by a diffracting point M in depth, at position (x_M, z_M), in a homogeneous medium with wave velocity v_{eq}. The triplet (v_{eq}, x_M, z_M) defines what we call the event's "equivalent medium." This model is of kinematic nature, and thus does not allow differentiating point diffraction from specular reflection. It is therefore just as legitimate to describe the equivalent model as a locally flat reflecting segment centered at point M, embedded in a uniform medium. The segment's dip straightforwardly derives from the geometry of the two rays SM, RM connecting the source and the receiver to the diffracting/reflecting point M in this medium.

The three parameters defining an event's equivalent medium (v_{eq}, x_{eq}, z_{eq}) can be computed from the picked event's attributes $(T_{SR}, P_S, P_R, x_S, x_R)$ with closed-form formulas, that can be derived from simple kinematic and geometric constraints.

Let d_S, d_R be the distances SM, RM (positive by construction) and let us suppose we know the equivalent slowness $s_{eq} = 1/v_{eq}$. There exist two independent ways of calculating the depth z_M:

$$z_M = d_S \frac{\sqrt{s_{eq}^2 - P_S^2}}{s_{eq}} = d_R \frac{\sqrt{s_{eq}^2 - P_R^2}}{s_{eq}}. \tag{A-2}$$

Meanwhile, the picked event's two-way traveltime is written as

$$s_{eq}(d_S + d_R) = T_{SR}. \tag{A-3}$$

These constraints form a linear system in d_S, d_R. Therefore,

$$d_S = \frac{T_{SR}}{s_{eq}} \frac{\sqrt{s_{eq}^2 - P_R^2}}{\sqrt{s_{eq}^2 - P_S^2} + \sqrt{s_{eq}^2 - P_R^2}}$$

and

$$d_R = \frac{T_{SR}}{s_{eq}} \frac{\sqrt{s_{eq}^2 - P_S^2}}{\sqrt{s_{eq}^2 - P_S^2} + \sqrt{s_{eq}^2 - P_R^2}}. \tag{A-4}$$

A supplementary constraint comes from acquisition geometry, simply stating that the two rays reach the surface at locations separated by the offset $2h = x_R - x_S$:

$$-\frac{P_S}{s_{eq}} d_S + \frac{P_R}{s_{eq}} d_R = 2h. \tag{A-5}$$

Equation (A-4) and (A-5) can be combined to eliminate d_S, d_R, leading to

$$(s_{eq}^2 - P_R^2)\left(P_S + \frac{2hs_{eq}^2}{T_{SR}}\right)^2 = (s_{eq}^2 - P_S^2)\left(P_R - \frac{2hs_{eq}^2}{T_{SR}}\right)^2, \tag{A-6}$$

from which the equivalent velocity can be derived:

$$v_{eq}^2 = \frac{4h}{T_{SR}} \frac{T_{SR} - h(P_R - P_S)}{T_{SR}(P_R - P_S) + 4hP_S P_R}. \tag{A-7}$$

We can now simply evaluate the vertical components of the parameters with

$$s_{eq}^2 - P_S^2 = \frac{(P_R - P_S)(T_{SR} + 2hP_S)^2}{4h(T_{SR} - h(P_R - P_S))}$$

and

$$s_{eq}^2 - P_R^2 = \frac{(P_R - P_S)(T_{SR} - 2hP_R)^2}{4h(T_{SR} - h(P_R - P_S))}. \tag{A-8}$$

Let x be the midpoint location $x = (x_S + x_R)/2$. The abscissa x_M of the diffracting point is obtained by

$$\begin{cases} h - (x_M - x) = -\dfrac{P_S}{s_{eq}} d_S \\ h + (x_M - x) = \dfrac{P_R}{s_{eq}} d_R \end{cases}, \tag{A-9}$$

giving $2(x_M - x) = (P_S d_S + P_R d_R)/s_{eq}$.
When all calculations are done, we have

$$x_M = x - h \frac{T_{SR}(P_S + P_R)}{T_{SR}(P_R - P_S) + 4hP_S P_R}. \tag{A-10}$$

Finally, the depth of the diffracting point is

$$z_M = d_S \frac{\sqrt{s_{eq}^2 - P_S^2}}{s_{eq}}, \tag{A-11}$$

giving

$$z_M = \frac{(T_{SR} + 2hP_S)(T_{SR} - 2hP_R)}{T_{SR}(P_R - P_S) + 4hP_S P_R} \sqrt{\frac{h(P_R - P_S)}{T_{SR} - h(P_R - P_S)}}. \tag{A-12}$$

From the equations above, it appears that a picked event may have no equivalent medium (e.g., because its equivalent velocity is imaginary). In our experience, most events picked in real or synthetic data do not create such mathematical exceptions (i.e., an equivalent medium can be computed successfully).

Events with no equivalent medium can only exist in cases where the (unknown true) velocity model exhibits strong lateral velocity variations. This is why these picks, when they exist, may bear precious information in the frame of velocity estimation.

This remark is better understood when noticing that there is a very close relationship between the computation of equivalent media and the normal moveout philosophy. In fact, the computations described above merely consist in seeking the best diffraction hyperbola locally tangent to the picked event. As a matter of fact, if the event does effectively result from specular reflection on a horizontal reflector in a 1D velocity model, then the equations degenerate because, in these conditions, the event's slopes P_S, P_R are equal but with opposite signs (by mere symmetry). In this case, it is easy to prove that the equivalent velocity is nothing else than the velocity that would be picked using standard NMO velocity analysis procedures.