Waveform inversion through MBTT formulation

Guy Chavent, François Clément

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Waveform inversion through MBTT formulation

Guy CHAVENT
François CLÉMENT

N° 1839
Janvier 1993

PROGRAMME 6

Calcul Scientifique,
Modélisation et
Logiciels numériques

Waveform inversion through MBTT formulation[1]

Inversion sismique par reformulation en temps

Guy CHAVENT[2]
François CLÉMENT[3]

[1]Work performed as part of the IFP-INRIA Ψ Consortium Project.

[2]CEREMADE, Université de Paris-Dauphine, pl. du M^{al} De Lattre de Tassigny, F-75775 Paris Cédex 16 - INRIA-Rocquencourt, Domaine de Voluceau, B.P. 105, F-78153 Le Chesnay Cédex.

[3]INRIA-Rocquencourt, Domaine de Voluceau, B.P. 105, F-78153 Le Chesnay Cédex.

ABSTRACT

This work deals with automatic determination of background velocity from full waveform data. For this purpose we have designed, in an earlier work, an original approach based on the MBTT (Migration Based Travel Time) reformulation of the waveform acoustic inversion problem. The goal of this paper is to give the present status of the development of this approach. We first describe how, with this approach, we can handle the multishot data which are required to obtain low frequency information on the background velocities.

We present then in detail four steps towards the implementation of this approach: separation of propagator/reflector unknowns via multiscale analysis, separate imaging of impedance and velocity reflectivity, amplitude preserving migration for the proper imaging of deep reflectors, determination of velocities and impedance from their background values and the two corresponding reflectivities. Each step is illustrated on a synthetic test example.

Finally we give numerical evidence that, already in the case of a single shot, the MBTT approach yields a dramatic improvement of the objective function, which becomes convex and looses all parasitic local minima, and we describe the calculation of the gradient of the new objective function.

Key words: inverse problem, parameter estimation, seismic inversion, background velocities determination, travel-time formulation, migration, multiscale analysis.

RÉSUMÉ

L'objet de ces travaux est la détermination automatique de modèles de vitesses à partir de données sismiques complètes. Dans ce but, nous avons développé, dans des travaux antérieurs, une approche originale reposant sur une reformulation en temps de parcours par migration (MBTT) du problème inverse sismique. Le présent rapport montre l'état d'avancement de cette approche. Nous décrivons d'abord comment nous pouvons prendre en compte des données multi-tirs, celles-ci étant nécessaires quant à l'obtention des composantes basses-fréquences des vitesses.

Nous présentons ensuite en détails 4 étapes de la mise en œuvre de cette approche : séparation des inconnues propagateur/reflecteurs par analyse multi-échelle, représentation séparée des réflectivités dûes aux vitesses et à l'impédance, migration préservant l'amplitude pour une représentation correcte des réflecteurs profonds et détermination des vitesses et de l'impédance à partir de leur modèle et des deux réflectivités correspondantes. Chaque étape est illustrée sur un exemple synthétique.

Finalement, nous montrons numériquement que, déjà dans le cas d'un seul tir, l'approche MBTT apporte une amélioration spectaculaire de la fonction objectif, qui devient convexe et perd tous ces minima locaux parasites. Nous décrivons, enfin, le calcul du gradient de cette nouvelle fonction objectif.

Mots clefs : problème inverse, estimation de paramètres, inversion sismique, détermination de modeles de vitesses, reformulation en temps de parcours, migration, analyse multi-échelle.

1 INTRODUCTION

Let us denote by

$$\varphi_n : (v, \sigma) \mapsto c_n, \quad \text{for } n = 1, \ldots, \text{NSHOT} \tag{1.1}$$

the mapping which associates, for the n-th shot, to any velocity and impedance depth sections v and σ, the corresponding acoustic synthetic shot gather c_n, given by

$$c_n = W_t G\, p_n \tag{1.2}$$

where p_n is the wavefield obtained from the acoustic equation (*cf.* Clément, 1991)

$$C(v, \sigma) p_n = f_n. \tag{1.3}$$

G is a matrix selecting the values at the geophone locations, and W_t is the time amplification factor for the compensation of the geometric attenuation and the muting off of the direct arrivals (see section 4). Then the usual output least-square error function is defined by

$$\mathcal{J}(v, \sigma) = \sum_{n=1}^{\text{NSHOT}} \mathcal{J}_n(v, \sigma) \tag{1.4}$$

$$\mathcal{J}_n(v, \sigma) = \frac{1}{2} \| d_n - \varphi_n(v, \sigma) \|^2, \qquad \text{for } n = 1, \ldots, \text{NSHOT} \tag{1.5}$$

where $\| \ \|$ denotes a norm on the seismic sections and d_n the recorded data corrected for the geometric attenuation.

As it is well-known, when v is accurately known, minimizing $\mathcal{J}$ with respect to σ yields a good estimation of the reflectivity, which takes into account all multiple reflections. However it is also well-known that, whenever the initial guess for the velocity is not good enough, the minimization of $\mathcal{J}$ with respect to v and σ fails, the gradient algorithm stopping in a local minimum which is far from the true value and far from being coherent with the data. This failure can be explained by two facts:

- for each shot n, the error function $\mathcal{J}_n$ is well-behaved with respect to σ (almost quadratic in the seismic band of σ), but extremely poorly behaved with respect to v: changes in v cause time shifts on the synthetic section $\varphi_n(v, \sigma)$, which create local minima for the error function. It will thus be important to *reformulate the problem* in order to have an objective function which is better behaved with respect to velocities.

- The multiplicity of shots, which could be expected to enhance the behavior of $\mathcal{J}$ through the redundancy of information, is in fact going in the opposite direction: the sole depth section σ is in charge of explaining all events appearing in the data for all shots. When the velocity section v is correct, this is all right, but when v is wrong, the kinematic would require to locate the reflectors in σ at different places for different shots, so that no single σ can correctly explain the events in all shots. It will thus be important to *duplicate the reflectivity unknowns* by having one reflectivity section per shot. Then the coherence or similarity or semblance of these sections comes naturally as a second objective to be satisfied, together with the consistency with the data.

The first point was addressed by the MBTT formulation (Clément, 1991) where the impedance depth section unknown σ is replaced by a time section reflectivity unknown s—similar to a seismogram without multiples. In this earlier approach—referred to as "classical MBTT" in the sequel—, the modified error function $\mathcal{MJ}$ is the same least-square error as before (*cf.* (1.5)), but considered as function of a propagator unknown v_s and of the time reflectivity unknown s. And time shifts on the synthetics vanish when we change the propagator v_s for fixed time reflectivity s. As the depth reflectivity σ is still an intermediate variable (depending on the propagator v_s and on the time reflectivity s via depth-migration), minimize the least-square error still permit us to evaluate a model $(\bar{v}, \bar{\sigma})$ which matches at best the data.

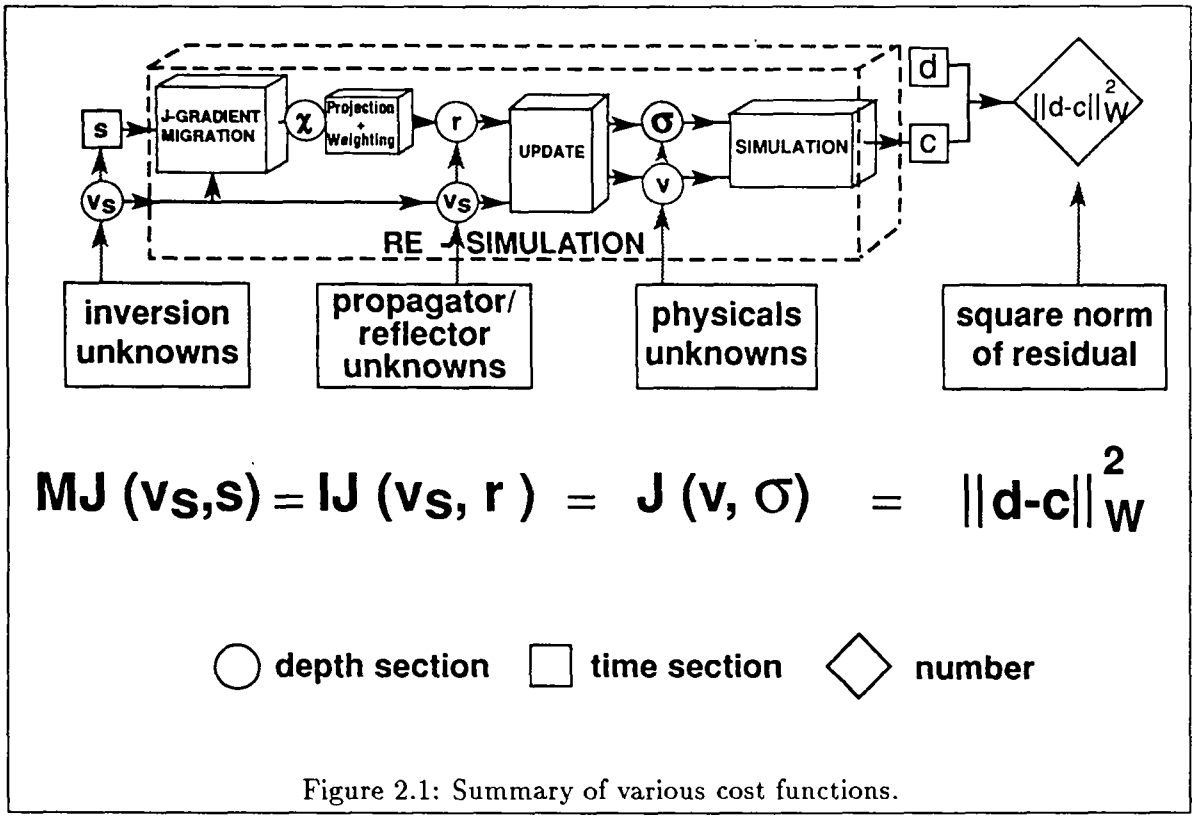
The second point was addressed by the various coherency or semblance formulations (Kleyn, 1983; Al-Yahya, 1989; Symes, 1988; Symes and Carrazone, 1991; Chavent and Jacewitz, 1990) which aimed at extracting information on the velocity by requiring that the depth section reflectivities for each shot look similar. The main difference with the Differential Semblance Optimization (DSO) of Symes and Carrazone (1991) lies on the necessity for DSO to perform complete minimization with respect to the depth reflectivity r as we just have to perform one gradient step, which is much costless.

We present in this paper a further development of the MBTT approach, designed to alleviate the limitations of the first version: better separation of propagator and reflector unknowns (section 3) and taking into account of v- and σ-reflectivity which solve the problem of the limitation in offset of the classical MBTT (Clément, 1991), better imaging of deep reflectors (section 4) which is necessary for the least-square error function to be sensitive to deep values of the propagator, and to allow the handling of multishot data by incorporating a coherency treatment of the depth reflectivities (complete algorithm presented in section 2). Numerical evidences of the convexity of the modified error function $\mathcal{MJ}$ are shown in section 5 on a synthetic example.

2 THE MBTT FORMULATION FOR MULTISHOT DATA

We describe in this section how to calculate the modified error function $\mathcal{MJ}$ as a function of the new unknowns, *i.e.* one propagator depth section v_s and NSHOT time reflectivity sections $s_1, \ldots, s_{\text{NSHOT}}$. The Figure 2.1 shows the different parts (boxes) of the complete algorithm from the inversion unknowns (v_s, s) to the least-square error (drawn for single shot data).

The preliminary step deals with the separation of the physical variables (v, σ) into propagator and reflector unknowns (v_s, r) (detailed in section 3). The first step corresponds to the first box of Figure 2.1 *J-gradient migration.* It is only the generalization of the classical MBTT for multishot data. The second step gathers the second and third boxes *projection + weighting* and *update* and performs the "gradient" step. This step has been improved by the introduction of weight and normalization (see section 4). We present two possibilities for the implementation of the coherency. The third step corresponds to the last box *simulation* and ends the big box *re-simulation.* In the fourth step we calculate the least-square error function which may involve a penalization term for the handling of the coherency.



$$MJ(v_s, s) = IJ(v_s, r) = J(v, \sigma) = \|d-c\|^2_W$$

Figure 2.1: Summary of various cost functions.

Preliminary step: definition of the propagator unknown v_s.
A velocity depth section v will be decomposed into a "smooth" part v_s and a "rough" part v_r (see section 3)

$$1/v = 1/v_s + 1/v_r, \tag{2.1}$$

(the reason for decomposing the slowness $1/v$ rather than v itself will be explained in paragraph 3.3) and similarly for the impedance section

$$\sigma = \sigma_s + \sigma_r. \tag{2.2}$$

This will allow to replace the physical unknowns v and σ by:

- one propagator depth section: $1/v_s$,
- two reflectivity depth sections: $1/v_r$ (proportional to v-reflectivity) and σ_r (proportional to σ-reflectivity).

The smooth impedance σ_s is always taken as given, its actual value having little influence on the seismograms. As we shall see in paragraph 3.3, the subspaces V_s (for $1/v_s$ and σ_s) and V_r (for $1/v_r$ and σ_r) have to be orthogonal, and we shall denote by P_s and P_r the corresponding projection operators.

First step: for each shot n, determine the velocity- and impedance-gradient depth sections $\chi_{1/v,n}$ and $\chi_{\sigma,n}$ from the time reflectivities $s_1, \ldots, s_{\text{NSHOT}}$ using a gradient-type migration with

velocity background v_s.
For each shot n:

1. forward propagate the source using the smooth velocity and impedance

$$C(v_s, \sigma_s) p_{s,n} = f_n, \tag{2.3}$$

2. backward propagate the time reflectivity s_n using the smooth velocity and impedance

$$C(v_s, \sigma_s)^T q_{s,n} = G^T W_t s_n, \tag{2.4}$$

3. correlate $p_{s,n}$ and $q_{s,n}$

$$\chi_{1/v,n}(p_{s,n}, q_{s,n}; v_s, \sigma_s) = \\ -\frac{1}{\sigma_s} < p_{s,n}{}^{(1+)}, q_{s,n}{}^{(1+)} >_t - \frac{v_s^2}{\sigma_s} < \nabla p_{s,n}, \nabla q_{s,n} >_t, \tag{2.5}$$

$$\chi_{\sigma,n}(p_{s,n}, q_{s,n}; v_s, \sigma_s) = \\ \frac{1}{v_s \sigma_s^2} < p_{s,n}{}^{(1+)}, q_{s,n}{}^{(1+)} >_t - \frac{v_s}{\sigma_s^2} < \nabla p_{s,n}, \nabla q_{s,n} >_t, \tag{2.6}$$

(where $u^{(1+)}$ denotes the first discrete derivative in time of u, see Appendix C). Formula (2.5) and (2.6) are obtained by requiring that $\chi_{1/v,n}$ and $\chi_{\sigma,n}$ are gradient migrations (see Lailly, 1983), *i.e.*

$$\chi_{1/v,n} = \nabla_{1/v} \mathcal{J}_n(v_s, \sigma_s) \tag{2.7}$$

$$\chi_{\sigma,n} = \nabla_{\sigma} \mathcal{J}_n(v_s, \sigma_s), \tag{2.8}$$

as soon as the time reflectivity unknown s_n is chosen as the reflectivity part of the data, given by the residual

$$s_n = d_n - \varphi_n(v_s, \sigma_s). \tag{2.9}$$

Second step: for each shot n return to the material properties v_n and σ_n using the gradient-type migrations $\chi_{1/v,n}$ and $\chi_{\sigma,n}$ and the smooth backgrounds v_s and σ_s.
We shall distinguish two cases, depending on the way one implements the coherency.
Case 1: coherency through penalization. Each shot will be treated completely independently, thus leading to one couple of material data (v_n, σ_n) per shot. The coherency will then be taken into account in the fourth step by adding to the least-square error a term measuring the dissemblance of the depth reflectivity sections corresponding to successive shots (*cf.* Symes and Carrazone, 1991).
For each shot n:

1. correct the gradient images for amplitude and project on the reflectivity space V_r in order to obtain the migrated depth sections

$$m_{1/v,n} = -P_r W_z P_r \chi_{1/v,n}(p_{s,n}, q_{s,n}; v_s, \sigma_s), \tag{2.10}$$

$$m_{\sigma,n} = -P_r W_z P_r \chi_{\sigma,n}(p_{s,n}, q_{s,n}; v_s, \sigma_s), \tag{2.11}$$

(W_z is a weight defined in section 4)

2. normalize in order to obtain dimensionless v- and σ-reflectivity depth sections

$$r_{1/v,n} = \frac{m_{1/v,n}}{\|m_{1/v,n}\|_1}, \tag{2.12}$$

$$r_{\sigma,n} = \frac{m_{\sigma,n}}{\|m_{\sigma,n}\|_1}, \tag{2.13}$$

3. update $1/v_s$ and σ_s using fixed dimensionless steps $\alpha_{1/v}$ and α_σ(see section 4)

$$1/v_n = 1/v_s + \alpha_{1/v}\|1/v_s\|_1 r_{1/v,n}, \tag{2.14}$$

$$\sigma_n = \sigma_s + \alpha_\sigma\|\sigma_s\|_1 r_{\sigma,n}. \tag{2.15}$$

Case 2: coherency through stack. We use here a formulation inspired from Chavent and Jacewitz (1990). The stacks $\chi_{1/v}$ and χ_σ of the correlation sections $\chi_{1/v,n}$ and $\chi_{\sigma,n}$ of each shot are, when (2.9) holds, nothing but the gradients $\nabla_{1/v}\mathcal{J}(v_s,\sigma_s)$ and $\nabla_\sigma\mathcal{J}(v_s,\sigma_s)$ (as $\mathcal{J}$ is the sum of the $\mathcal{J}_n$'s), which tell us how to modify $1/v_s$ and σ_s to decrease $\mathcal{J}$, when only one set of material properties is used to model the data. If the velocity model is good, the events in the migrated sections of each shot will add constructively from one shot to the next, and the stacked section will be able to explain correctly the data from all shots. If the velocity model is wrong, the events will interfere destructively, and the stacked section will be much less able to match correctly the data of all shots. So one way of measuring the adequacy of the velocity model is to re-simulate the data of each shot, using the reflectivity of the stacked sections: the least-square error itself is then a measure of quality of the velocity model.

0. stack the gradient images $\chi_{1/v,n}$ and $\chi_{\sigma,n}$ of each shot

$$\chi_{1/v} = \sum_{n=1}^{\text{NSHOT}} \chi_{1/v,n}(p_{s,n},q_{s,n};v_s,\sigma_s), \tag{2.16}$$

$$\chi_\sigma = \sum_{n=1}^{\text{NSHOT}} \chi_{\sigma,n}(p_{s,n},q_{s,n};v_s,\sigma_s), \tag{2.17}$$

1. correct the stacked gradient images for amplitude and project on the reflectivity space V_r in order to obtain the migrated sections

$$m_{1/v} = -P_r W_z P_r \chi_{1/v}, \tag{2.18}$$

$$m_\sigma = -P_r W_z P_r \chi_\sigma, \tag{2.19}$$

(W_z is a weight defined in section 4)

2. normalize in order to obtain dimensionless v- and σ-reflectivity depth sections

$$r_{1/v} = \frac{m_{1/v}}{\|m_{1/v}\|_1}, \tag{2.20}$$

$$r_\sigma = \frac{m_\sigma}{\|m_\sigma\|_1}, \tag{2.21}$$

3. update $1/v_s$ and σ_s using dimensionless steps $\alpha_{1/v}$ and α_σ(see section 4)

$$1/v = 1/v_s + \alpha_{1/v}\|1/v_s\|_1 r_{1/v}, \tag{2.22}$$

$$\sigma = \sigma_s + \alpha_\sigma\|\sigma_s\|_1 r_\sigma. \tag{2.23}$$

Hence the material properties (v_n, σ_n) associated with each shot are all the same

$$v_n = v, \qquad \text{for all } n = 1, \ldots, \text{NSHOT}, \tag{2.24}$$

$$\sigma_n = \sigma, \qquad \text{for all } n = 1, \ldots, \text{NSHOT}. \tag{2.25}$$

Third step: for each shot, re-simulate the data using the material properties v_n and σ_n. For each shot n:

1. re-propagate the source using the full velocity and impedance associated with the shot

$$C(v_n, \sigma_n) p_n = f_n, \tag{2.26}$$

2. compute the wavefield at the geophones, correct for geometric attenuation and mute off the direct arrivals

$$c_n = W_t G\, p_n = \varphi_n(v_n, \sigma_n) \tag{2.27}$$

(this defines the "modified parameter to synthetics" mapping, or "re-simulation" mapping $\mathcal{M}\varphi_n : (v_s, s_1, \ldots, s_{\text{NSHOT}}) \mapsto c_n$).

Fourth step: compute the least-square error and the coherency term.
For each shot n, compute

$$\tilde{\mathcal{J}}_n(v_n, \sigma_n) = \frac{1}{2} \|\tilde{c}_n - \tilde{d}_n\|_2^{\,2} \tag{2.28}$$

where the $\tilde{}$ denotes l^1-normalized sections (see Appendix C). The MBTT-modified error functoin for the n-th shot is then

$$\mathcal{M}\mathcal{J}_n(v_s, s_1, \ldots, s_{\text{NSHOT}}) = \tilde{\mathcal{J}}_n(v_n, \sigma_n), \tag{2.29}$$

and the final objective function for all shots is

$$\begin{aligned} \mathcal{M}\mathcal{J}(v_s, s_1, \ldots, s_{\text{NSHOT}}) &= \sum_{n=1}^{\text{NSHOT}} \mathcal{M}\mathcal{J}_n(v_s, s_1, \ldots, s_{\text{NSHOT}}) \\ &+ w_v \sum_{n=2}^{\text{NSHOT}} \|v_n - v_{n-1}\|^2 + w_\sigma \sum_{n=2}^{\text{NSHOT}} \|\sigma_n - \sigma_{n-1}\|^2 \end{aligned} \tag{2.30}$$

where w_v and w_σ are the penalization weights associated with the coherency of the velocity and impedance reflectivities respectively. Notice that, when the "coherency through stack" approach is used, the coherency term in (2.30) disappears, as the material properties v_n and σ_n associated with all shots are the same.

We have now precisely defined, through steps 1 to 4, how to calculate the modified objective function $\mathcal{M}\mathcal{J}$ associated through the MBTT formulation with a given smooth velocity background v_s and NSHOT given time reflectivities $s_1, \ldots, s_{\text{NSHOT}}$. In order to perform a full waveform inversion, we have to solve the optimization problem

find $\hat{v}_s, \hat{s}_1, \ldots, \hat{s}_{\text{NSHOT}}$ such that

$$\mathcal{M}\mathcal{J}(\hat{v}_s, \hat{s}_1, \ldots, \hat{s}_{\text{NSHOT}}) \le \mathcal{M}\mathcal{J}(v_s, s_1, \ldots, s_{\text{NSHOT}}) \tag{2.31}$$

for all $v_s, s_1, \ldots, s_{\text{NSHOT}}$.

The first advantage which can be expected from this new formulation is that $\mathcal{MJ}$, for fixed reflectivities s_n ($n = 1, \ldots, \text{NSHOT}$), will be a smooth function of v_s without local minima. This is numerically demonstrated in the case of one shot in Clément (1991) and in section 5 of this paper. In the multishot case, this will be true for each of the $\mathcal{MJ}_n$'s defined in (2.29), and hence for their sum $\mathcal{MJ}$.

The second advantage is that one knows in fact a very good estimate of the optimal time reflectivity $\hat{s}_n$, namely the residual s_n^0 associated with the seismic data d_n, obtained by subtracting—or muting off—the direct arrivals in the d_n section (see Appendix A). It is thus hoped that a first estimate v_s^1 of the velocity background v_s can be obtained by minimizing $\mathcal{MJ}(v_s, s_1^0, \ldots, s_{\text{NSHOT}}^0)$ with respect to v_s. Then the time reflectivity s_n^0 of each shot can be enhanced (subtracting multiples,...) by minimizing $\mathcal{MJ}(v_s^1, s_1, \ldots, s_{\text{NSHOT}})$ with respect to $s_1, \ldots, s_{\text{NSHOT}}$. And so on...

3 SMOOTH/ROUGH DECOMPOSITION OF THE PHYSICAL PARAMETERS

We show in this section how we can perform the decomposition of the physical parameters of the preliminary step which will enable us to take into account all the reflectors and then suppress the limitation in offset which was necessary in classical MBTT (Clément, 1991).

We recall first where the information on velocities lies in the data: because of the lack of continuous and low-frequency components in the source wavelet, low-frequency components of the velocity and impedance do not produce sizable reflections. It is only through the kinematic of arrival times of reflections generated by high-frequency components of v and σ that the low-frequency part of the velocity plays a role in the data.

It is hence important to decouple the unknowns which governs the kinematics—*propagator unknowns*—from those who describe the reflectivity—*reflector unknowns.* In other terms the wave equation, when solved with any value of the propagator unknown and a zero reflectivity value of the reflector unknowns should produce a synthetic seismogram with no sizable reflection beside the direct arrivals.

A first choice would be to take v as the propagator unknown and σ as the only reflector unknown: the impedance σ is clearly a pure reflector unknown (value of σ do not interfere on the propagation), but the velocity v, beside being of course a propagator unknown, is also a reflector unknown as strong variations of v generate reflections for waves at non-normal incidences. It was the choice made in Clément (1991) where we concluded that the MBTT formulation was efficient if we restricted the maximum offset of the observation, *i.e.* considered only data where v-reflectors are negligible.

We shall take advantage of the band-limited nature of the source wavelet—and particularly of its lack of continuous and low-frequency components—to perform an *approximate decoupling of the propagator and reflector unknowns.*

Let V denote the function space for the material properties distribution. We split it into

the direct sum of two subspaces

$$V = V_s \oplus V_r \tag{3.1}$$

where V_s means "V-smooth" and contains "slowly varying functions", and V_r stands for "V-rough" which contains "rapidly oscillating functions". We shall denote by P_s and P_r the respective projections on each subspace parallel to the other one and decompose any velocity and impedance distribution into their smooth and rough parts (as we shall see section 3.3, it is rather the slowness $u = 1/v$ which has to be decomposed into a smooth and a rough part)

$$1/v = 1/v_s + 1/v_r = P_s(1/v) + P_r(1/v), \tag{3.2}$$

$$\sigma = \sigma_s + \sigma_r = P_s\sigma + P_r\sigma. \tag{3.3}$$

We first give one example of such a decomposition in the case where the depth section is covered by a fine rectangular grid used for the numerical simulation of the wave equation. Then V is made of functions defined by bilinear interpolation from their values at the nodes of the grid—*i.e.* V is the so-called Q_1-finite element space—and a natural choice for V_s is the space of functions defined also by the same bilinear interpolation but on a coarse subgrid. It is then easy to obtain a (non orthogonal) supplementary subspace V_r of V_s by simply complementing the canonical basis of V_s with elements of the canonical basis of V. In this example, we see that choosing V_s amounts simply to choose a coarse subgrid of the fine numerical grid.

We state now the conditions which have to be satisfied by V_s and V_r to be correct propagator and reflector spaces:

1. low reflectivity of V_s: the synthetic seismograms computed with $1/v_s \in V_s$ and $\sigma_s \in V_s$ contain no sizable reflections beside the direct arrivals,

2. sufficient kinematic content of V_s: replacing $1/v \in V$ by $1/v_s = P_s(1/v) \in V_s$ does not deteriorate too much the quality of the migrated images,

3. low kinematic content of V_r: replacing $1/v_s \in V_s$ by $1/v_s + 1/v_r$ for any $1/v_r \in V_r$ changes as little as possible the travel times between distant enough points.

A natural choice for the propagator and reflector unknowns is then:

- one propagator unknown: $1/v_s \in V_s$,
- two reflector unknowns: $1/v_r$ and $\sigma_r \in V_r$.

As it is well-known that the smooth part σ_s of the impedance plays no role in the seismic data, we can choose for σ_s any reasonable smooth function, *e.g.* a constant.

Conditions 1 and 2 tend to conflict: at one end, when V_s is made of constant functions, the low reflectivity condition is fully satisfied, but the kinematic content of the constant is definitely not sufficient, and at the other end, when $V_s = V$ the situation is exchanged. This conflict can be resolved only because of the band-limited nature of the source wavelet, and more particularly of its lack of low-frequency components. *We discuss now this issue in the case where V_s is constructed by Q_1-interpolation on a coarse grid* as in the above example.

3.1 SATISFYING THE LOW REFLECTIVITY CONDITION FOR V_s

If a coarse grid of size H is used, a transition from a given slowness $1/v_1$ to another given slowness $1/v_2$ has to be "stretched" linearly over a layer of thickness H when it is to be represented in the space V_s associated with this coarse grid. In order to analyze the effect of this "stretching" on the reflectivity, we use the fact that the corresponding velocity transition is close to be linear (see Figure 3.1) and define:

- r_s = reflection coefficient of the smoothed velocity transition for a normal incidence plane wave illumination at frequency f,
- $r = (v_2 - v_1)/(v_2 + v_1)$ = usual reflection coefficient corresponding to an abrupt transition.

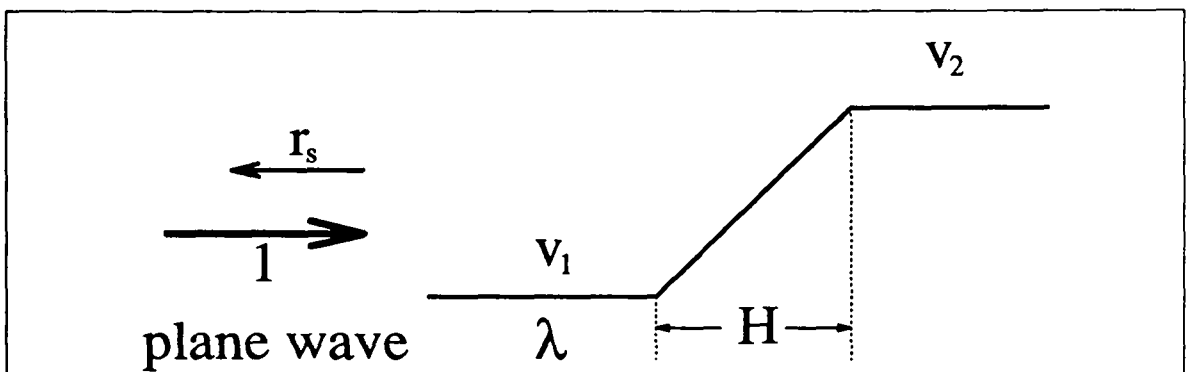


Figure 3.1: Velocity transition for the analysis of the low reflectivity condition.

Using the results of Brekhovskikh (1960), we obtain (see Appendix B) that the reflectivity reduction factor r_s/r satisfies

$$\frac{r_s}{r} \leq \frac{1}{4n_\lambda}, \qquad \text{for } n_\lambda \geq 0.3 \text{ and } r \leq 0.33 \tag{3.4}$$

where n_λ is the thickness of the transition measured in terms of the wavelength λ in the incident medium

$$n_\lambda = \frac{H}{\lambda}, \qquad \text{with } \lambda = \frac{v_1}{f}. \tag{3.5}$$

In other words n_λ is the "number of wavelength per grid cell" associated with the coarse grid, the velocity of the medium and the frequency of the source. Using the exact formulas of Appendix B, we have plotted on Figure 3.2 the exact value of the reduction factor r_s/r as a function of n_λ. Both formula (3.4) and Figure 3.2 show that

> *reflectivity is reduced by a factor 10 at least as soon as the number of wavelength per grid cell is larger than 2.5.* (3.6)

Notice that this minoration of n_λ is rather pessimistic as we can see in Brekhovskikh (1960) that if the transition is a little smoother, the reflection coefficient is much smaller.

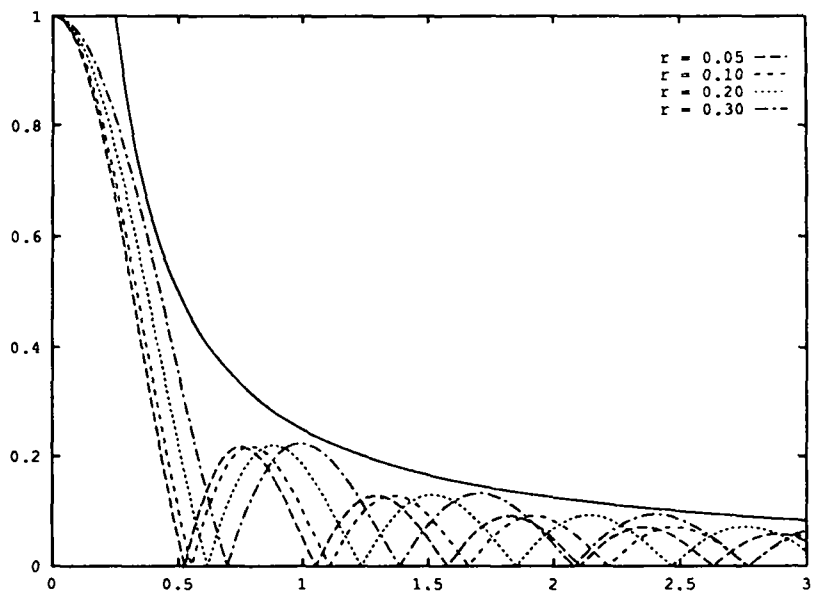


Figure 3.2: The reflectivity reduction factor r_s/r as a function of the number of wavelength per grid cell n_λ. Dashed lines: reflection coefficient r associated with the abrupt transition takes values 0.05, 0.10, 0.20 and 0.30. Solid line: $1/4n_\lambda$ majorating function.

3.2 SATISFYING THE SUFFICIENT KINEMATIC CONTENT CONDITION FOR V_s

It is difficult to analyze mathematically the decrease in the quality of the migrated images when a slowness model $1/v \in V$ is replaced by its best approximation on a given subspace V_s. But this problem has been recently considered by Versteeg *et al.* (1990, 1991), who performed migrations of the 2D synthetic data of the Marmousi model using first the exact, extremely detailed velocity model, and then more and more smoothed velocity models. Though they did not use the same smoothing procedure as the one we are discussing here, we believe that qualitative answers can still be drawn from their study for our case. Their conclusion was that the smoothing did not alterate the quality of the migrated images, provided that the smoothed velocity model retained from the original one all spatial wavelength larger than twice the dominant wavelength λ of the signal. If we consider, as a rule of thumb, that the subspace V_s associated with a coarse grid of size H contains all spatial wavelength larger than $4H$, then we can expect that

the quality of migrated images is maintained when one replaces background slownesses $1/v \in V$ *by smoothed slownesses* $1/v_s = P_s(1/v) \in V_s$ *provided that the number of wavelength per grid cell* n_λ *is smaller than 0.5.* (3.7)

3.3 SATISFYING THE LOW KINEMATIC CONTENT CONDITION FOR V_r

Let us denote by u the slowness of the model

$$u = \frac{1}{v},$$

which decomposes into

$$u = u_s + u_r = P_s u + P_r u.$$

Minimizing the kinematic content of V_r amounts to choose the projection P_s on V_s in such a way that replacing $u \in V$ by $u_s = P_s u \in V_s$ perturbs as little as possible the travel times between points at a distance large with respect to H. For example (see Figure 3.3) the travel time between two points M_1 and M_2 located on both sides of a velocity discontinuity should be basically unchanged when u is replaced by u_s, *i.e.* where the discontinuity has been replaced by a smooth slowness transition zone. This property will be satisfied as soon as

$$\int_{\Omega_h} (u - u_s) w_s = 0, \qquad \text{for any } w_s \in V_s \tag{3.8}$$

where Ω is the space domain. This can be seen by choosing in (3.8) for w_s a function with support approaching as close as possible the M_1M_2 segment as in Figure 3.3, and taking values close to one on this segment, in which case the left-hand side of (3.8) approaches the difference of travel time between M_1 and M_2.

Hence we see from (3.8) that V_r will have a low kinematic content as soon as P_s is chosen to be the orthogonal projection on V_s, which means that

the subspaces V_s and V_r have to be orthogonal. (3.9)

Remark: in all the numerical results we have used, for historical reasons, an orthogonal decomposition of the velocity instead of the slowness. As the error induced on the travel time by this choice was quite small, we did not change to orthogonal decomposition of slowness yet.

3.4 RESOLVING THE CONFLICT

We see that the conditions (3.6) and (3.7) for V_s are actually conflicting: the first one would require that $n_\lambda \geq 2.5$ (for a reduction of reflectivity of 10) and the second that $n_\lambda \leq 0.5$ (for a perfect ability to migrate). The reader has to keep in mind that the first inequality is quite pessimistic (see paragraph 3.1) and that the second lies on empirical results, so we do not have to worry too much about this contradiction. Once V_s has been chosen, V_r is ideally defined as the orthogonal supplement to V_s in order to satisfy at best the low kinematic content condition.

We illustrate on the SYNCLAY data set $(\bar{v}, \bar{\sigma})$ (see Appendix A) how this conflict can actually be resolved. The initial 192×192 discretization mesh is successively sampled by dichotomy to define seven scales of approximation ($192 = 3 \times 2^6$) from the coarsest grid 3×3 to the finest grid 192×192. At each scale N, we have

$$V_N = V_{N-1} \oplus W_N \tag{3.10}$$

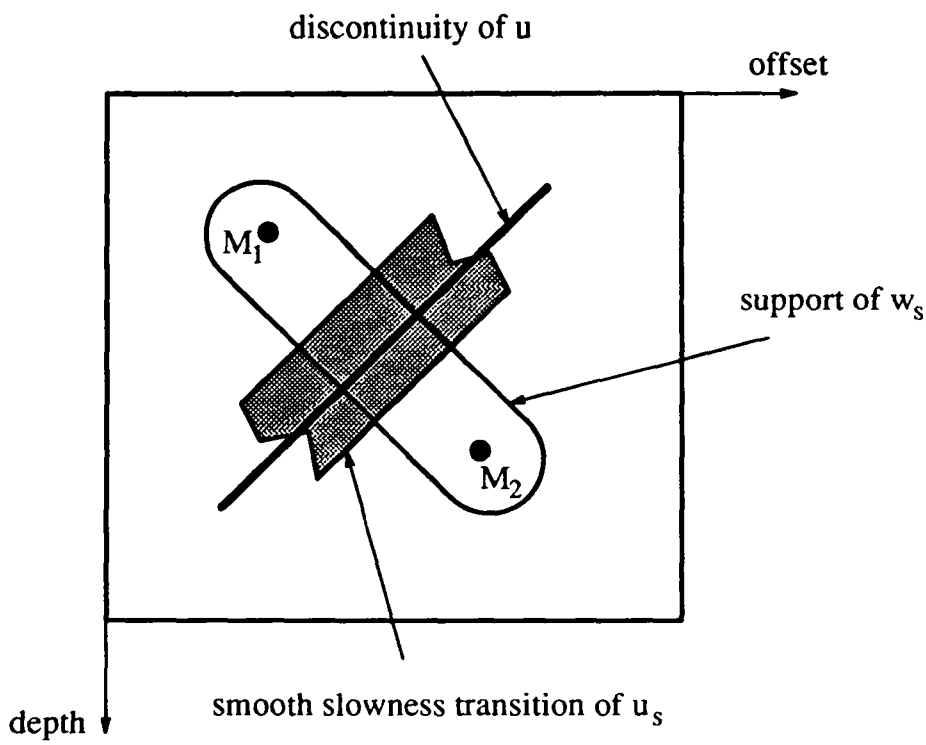


Figure 3.3: Notations for the analysis of the kinematic error induced by smoothing the slownesses.

where W_N is the orthogonal complement to V_{N-1} in V_N. Hence

$$V = V_6 = \underbrace{V_0 \oplus W_1 \oplus \cdots}_{V_s} \underbrace{\oplus W_6}_{V_r}. \tag{3.11}$$

The following set of numerical experiments was obtained by moving a cursor from scale 5 ($V_s = V_5, V_r = W_6$) down to scale 0 ($V_s = V_0, V_r = W_1 \oplus \cdots \oplus W_6$). In Figure 3.4 to 3.6 we have represented for each value of the cursor the synthetics and the v- and σ-reflectivity depth sections r_v and r_σ obtained using a constant background impedance and a background propagator $v_s = P_s \bar{v}$ for the projection P_s corresponding to the cursor position.

Accordingly to the previous discussion, one sees that, as the cursor moves down from 5 to 0, the seismogram weakens to the sole direct arrivals and the reflectivity images degenerate. It can be seen that, when the cursor position is 3, one has both a synthetic seismogram which contains only the direct arrival, and still accurate and correctly imaged depth reflectivity. So this experimentation suggests to choose, for the inversion of the SYNCLAY and MODIFIED SYNCLAY models (see Appendix A),

$$V_s = V_3 \tag{3.12}$$

which turns out to correspond to $n_\lambda = 1$ (see Tab. 3.1). It is in the range of the conflicting values $n_\lambda = 0.5$ and $n_\lambda = 2.5$. Thus in all the sequel we have used a value of 3 for the scale cursor: $V_s = V_3$ and $P_s = P_3$.

cursor	0	1	2	3	4	5	6
grid	3×3	6×6	12×12	24×24	48×48	96×96	192×192
H (meters)	333	167	83	42	21	10	5
n_λ	8	4	2	1	0.5	0.25	0.125

Table 3.1: Number of wavelength per grid cell n_λ at the different scales for peak frequency f = 36 Hertz, *i.e.* λ = 42 meters in the first layer (v_1 =1500 meter per second).

In Figure 3.7 we have illustrated the importance of the orthogonality on the near surface area of the migrated images: when V_r is not orthogonal to V_s (on the right), the v-reflectivity depth section still contains continuous components in its shallow part. Hence updating v_s with r_v will substantially change the travel time in the upper layer and introduce phase shifts in the corresponding synthetics.

4 RETURNING TO THE PHYSICAL PARAMETERS: AMPLITUDE PRESERVING MIGRATION

We describe in this section how to return in an efficient way to the physical parameters v and σ (second step of section 2) once the gradient images $\chi_{1/v}$ and χ_σ have been computed for a given background v_s, σ_s. For sake of simplicity we consider only one shot.

One of the key issue in the MBTT formulation (see last paragraph of section 2) is that one must be able to re-simulate late reflections, which requires a proper imaging of deep reflectors using as initial guess s^0 the reflectivity part of d. The gradient images $\chi_{1/v}$ and χ_σ of the v- and σ-reflectors have been computed in the first step by a correlation between the field p_s (forward propagation of the source) and the field q_s (backward propagation of the time reflectivity s). When it propagates, the signal attenuates for two reasons:

- geometrical spreading due to energy conservation of the wavefront generated by a point-source,
- reflection/transmission losses at each interface.

This attenuation comes in at two different places in the imaging process. First, the initial guess s^0 of the time reflectivity is deduced from the data themselves by muting off the direct arrivals. The data, which result from the propagation of a wavefield in the actual (unknown) medium, have been subject to both kind of attenuation. We classically correct for the geometrical attenuation all data d_n (we apply an amplification factor $\sqrt{t}$ to our synthetic 2D data) and use the corresponding weight $W_t = \sqrt{t}$ in the formula (1.2) for the computed seismograms c_n. The reflection/transmission losses for a signal which has been transmitted through reflectors $1, 2, \ldots, i-1$, reflected at reflector i and again transmitted through reflectors $i-1, \ldots, 2, 1$

is

$$(1-r_1^2)\ldots(1-r_{i-1}^2)r_i. \tag{4.1}$$

Correcting for this attenuation would require an estimation of the reflectivity encountered by each signal and has been left for later study.

Second, the gradient images $\chi_{1/v}$ and χ_σ are obtained by correlations of the wavefields p_s and q_s, which are propagating in a smooth medium (velocity v_s, impedance σ_s) with no reflectivity (see section 3). Hence p_s and q_s are affected only by the geometrical attenuation. At a given point M of the offset×depth section, the forward field p_s has been attenuated proportionally to the distance $SM^{\frac{1}{2}}$ (where S denotes the source location) and, in the case of one receiver only, the backward field q_s has been attenuated proportionally to the distance $RM^{\frac{1}{2}}$ (where R denotes the receiver location). Hence the geometrical attenuation on p_s and q_s induces on the gradient images $\chi_{1/v}$ and χ_σ an attenuation proportional to $SM^{\frac{1}{2}} \times RM^{\frac{1}{2}}$ (where now R denotes a mean receiver location). A rough approximation (valid for the zero-offset case) of this attenuation factor is simply the depth z of the point M. So we decide to

> *amplify our gradient images $\chi_{1/v}$ and χ_σ by a depth factor $W_z = z$ in order to obtain the migrated sections $m_{1/v}$ and m_σ.* (4.2)

This amplification has to be performed in a way coherent with the fact that our depth reflectivity unknowns are only the rough part of $1/v$ and σ, and that we want that our migrated sections $m_{1/v}$ and m_σ allow us to update the smooth background $1/v_s$ and σ_s for the missing reflectivity part $1/v_r$ and σ_r. This updating has to be performed in such a way that, when the time reflectivity $s = d - \varphi(v_s, \sigma_s)$ (= reflectivity part of the data) has been used to generate the migrated sections $m_{1/v}$ and m_σ, the least-square error function $\mathcal{J}$ is decreased (*i.e.* the updated $1/v$ and σ generate seismograms close to the data d). This is achieved by *requiring that* (in the case of the impedance example) *the migrated section m_σ is a descent direction for $\mathcal{J}$ in the reflectivity space V_r, i.e.*

$$< \nabla_{\sigma_r}\mathcal{J}, m_\sigma > < 0. \tag{4.3}$$

It becomes now natural to define the migrated sections by

$$m_{1/v} = -P_r W_z P_r \chi_{1/v}, \tag{4.4}$$

$$m_\sigma = -P_r W_z P_r \chi_\sigma, \tag{4.5}$$

(*i.e.* (2.10) and (2.11) of section 2). This definition for $m_{1/v}$ and m_σ satisfies clearly (4.2). We check now that it satisfies also (4.3), for m_σ for example. As the spaces V_s and V_r are orthogonal, one has

$$\nabla_{\sigma_r}\mathcal{J}(v_s, \sigma_s) = P_r \nabla_\sigma \mathcal{J}(v_s, \sigma_s) \tag{4.6}$$

and, by definition of χ_σ as a gradient image, one has, when $s = d - \varphi(v_s, \sigma_s)$ (*cf.* equation (2.8))

$$\chi_\sigma = \nabla_\sigma \mathcal{J}(v_s, \sigma_s). \tag{4.7}$$

Hence we see that

$$\begin{aligned} < \nabla_{\sigma_r}\mathcal{J}, m_\sigma > &= < P_r \nabla_\sigma \mathcal{J}, -P_r W_z P_r \chi_\sigma > \\ &= - < P_r \chi_\sigma, P_r W_z P_r \chi_\sigma > \end{aligned}$$

i.e. , as $P_r^2 = P_r$

$$< \nabla_{\sigma_r} \mathcal{J}, m_\sigma >= - < P_r \chi_\sigma, W_z P_r \chi_\sigma > \tag{4.8}$$

which is negative as W_z is a positive weight. With the definition (4.4) (4.5) of the migrated sections, the update of the smooth backgrounds $1/v_s$ and σ_s by addition of a fraction of $m_{1/v}$ and m_σ amounts to perform one descent step on the usual least-square error function $\mathcal{J}$.

In practice, this involves a linear search for an optimal step, which is quite expensive, as getting close to the minimum of the error function requires many wave equation simulations. Moreover, as one can see in Tab. 4.1 (columns 4 and 5), the optimal steps are extremely dependent on the choice of the smooth background and on the data (see Appendix A for the description of the two data sets). In order to avoid this linear search, we normalize both the migrated sections (then we call them "reflectivity images" $r_{1/v}$ and r_σ, cf (2.12) (2.13)) and the length of the gradient steps (equations (2.14) (2.15)). Columns 6 and 7 of Tab. 4.1 show that the dimensionless steps α_v and α_σ are much less sensitive to the background and data. So we have chosen:

$$\alpha_v = 8\,10^{-4} \qquad \text{and} \qquad \alpha_\sigma = 2\,10^{-3} \tag{4.9}$$

as it is always better to underestimate the optimal step rather than overestimate it.

data set	background		not normalized		normalized	
	v_s	σ_s	$\alpha_v \times 10^3$	α_σ	$\alpha_v \times 10^3$	$\alpha_\sigma \times 10^3$
SYNCLAY	$P_3\bar{v}$	$1.5\ 10^6$	0.0	2.7	1.0	3.5
SYNCLAY	1500	$1.5\ 10^6$	2.3	1.3	0.9	1.5
MODIFIED SYNCLAY	$P_3\bar{v}$	$1.5\ 10^6$	0.3	0.5	0.6	3.0

Table 4.1: Normalization forces the optimal gradient steps to be almost independent of the initial guesses and of the data.

Finally, we illustrate in Figure 4.1 the improvement brought by the introduction of the weights W_t and W_z in the imaging of the deep reflectors of the v- and σ-depth reflectivities and in the re-simulation of late reflections. Since the late events are still significant in the corrected re-simulation, we can hope an actual dependency of the new error function on the deep values of the propagator unknown v_s.

5 BEHAVIOR OF THE MODIFIED OBJECTIVE FUNCTION

The first numerical results (Clément, 1991) concerned the case where the propagator v_s was taken to be the full velocity section itself v, so that the time reflectivity unknown s was in charge of explaining only the reflections caused by changes in the impedance. It was

then found that, in the case of one shot, the modified error function $\mathcal{MJ}$, for fixed time reflectivity s, had no local minima with respect to v on a large velocity interval (provided that the offset of the data was kept small enough in order to eliminate reflections caused by velocity discontinuities), whereas the usual error function $\mathcal{J}$, for fixed impedance σ, had many parasitic local minima.

We present here two tests for the validation of the MBTT approach of the 2D seismic inverse problem: it is shown that the new objective function reveals no local minima in a large range of velocities, whereas the usual one has many of them, as it is well-known.

In the first test, we compute the two least-square error functions $\mathcal{J}$ and $\mathcal{MJ}$ for a uniform perturbation of the propagator around its true value $v_{\text{MODIFIED SYNCLAY}}$ (for $\mathcal{J}$) and $\bar{v}_s = P_3\bar{v} = P_3 v_{\text{MODIFIED SYNCLAY}}$ (for $\mathcal{MJ}$), supposing that the true depth reflectivity $\sigma_{\text{MODIFIED SYNCLAY}}$ (for $\mathcal{J}$) and the true time reflectivity $s_{\text{MODIFIED SYNCLAY}}$ (for $\mathcal{MJ}$) are known. We use the MODIFIED SYNCLAY data set for which we effectively know all the true parameters. So for zero perturbation, $\mathcal{J} = \mathcal{J}(v_{\text{MODIFIED SYNCLAY}}, \sigma_{\text{MODIFIED SYNCLAY}}) = 0$ and $\mathcal{MJ} = \mathcal{MJ}(\bar{v}_s, s_{\text{MODIFIED SYNCLAY}}) = 0$. To enhance the difficulty of the test, we suppose that the first layer—which is water—is known, so that the first reflector will always be well-imaged. Hence by uniform perturbation we mean a function which is zero in the shallow layer and constant in the rest of the domain but has a smooth transition in the case of $\mathcal{MJ}$—as then, the propagator perturbation belongs to V_s. The size of the perturbation is ± 400 m/s. The velocity ranges from 1800 m/s in the second layer to 4000 m/s in the deeper one. We can see Figure 5.1 that the attraction domain of the global minimum for the new objective function is wide enough to allow feasible gradient-based minimization, whereas one must already have a 3% guess of the true propagator in its pocket to be able to pertain the global minimum of the usual error function.

The second test is similar to the previous one but with two degrees of freedom and no *a priori* guess of the exact value of the propagator (the global minima will not correspond to zero value of the objective functions). So we consider the SYNCLAY dataset. It points out the ability of the MBTT approach to extract information on the propagator at deep locations. We make the hypothesis that the propagator is horizontally stratified, therefore completely defined by its depth profile. The first layer is known again. And to reduce the number of unknowns, we make the further assumption that this profile is linear from depth 100 m down to depth 1000 m. Then we plot Figure 5.2 the two objective functions for propagator ranging from 1250 m/s to 2500 m/s at depth 100 m and from 1500 m/s to 4000 m/s at depth 1000 m and reflectivity fixed to $\sigma = cst$ (for $\mathcal{J}$) and $s =$ residual $= s_{\text{MODIFIED SYNCLAY}}$ (for $\mathcal{MJ}$). The convexity result is still obvious. Moreover, the optimal linear profile for the modified error function is far closer to the true propagator: this shows that the new objective function is much more sensitive to deep values of the propagator.

The success of these two tests show that it will actually be possible to perform the minimization of the new least-square error function $\mathcal{MJ}$ with respect to the new inversion unknowns propagator v_s and time reflectivity s using a gradient-type algorithm. We give in Appendix C detailed calculation of the two gradients $\nabla_{1/v_s}\mathcal{MJ}(v_s, s)$ and $\nabla_s\mathcal{MJ}(v_s, s)$. We are now testing the computation of these gradients.

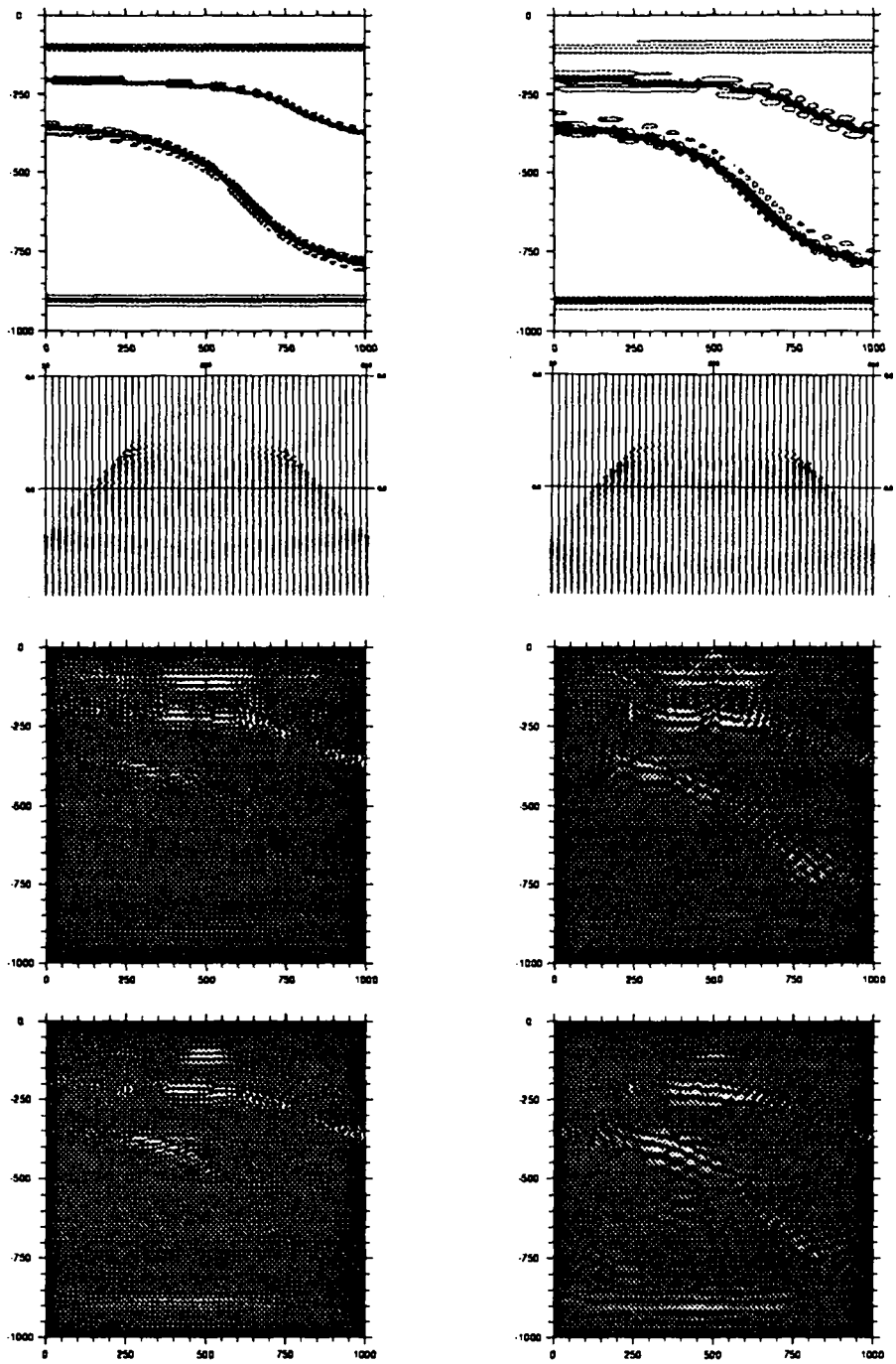


Figure 3.4: Moving down the scale cursor clears off the synthetics but degrades the migrated images (deepest reflector). Left: cursor = 5, right: cursor = 4. From top to bottom: propagator $v_s = P_s\bar{v}$, synthetic seismogram $c_s = \varphi(v_s, \sigma_s = cst)$ and reflectivity depth sections r_v and r_σ obtained from formula (2.3) to (2.13) with the same background v_s, σ_s.

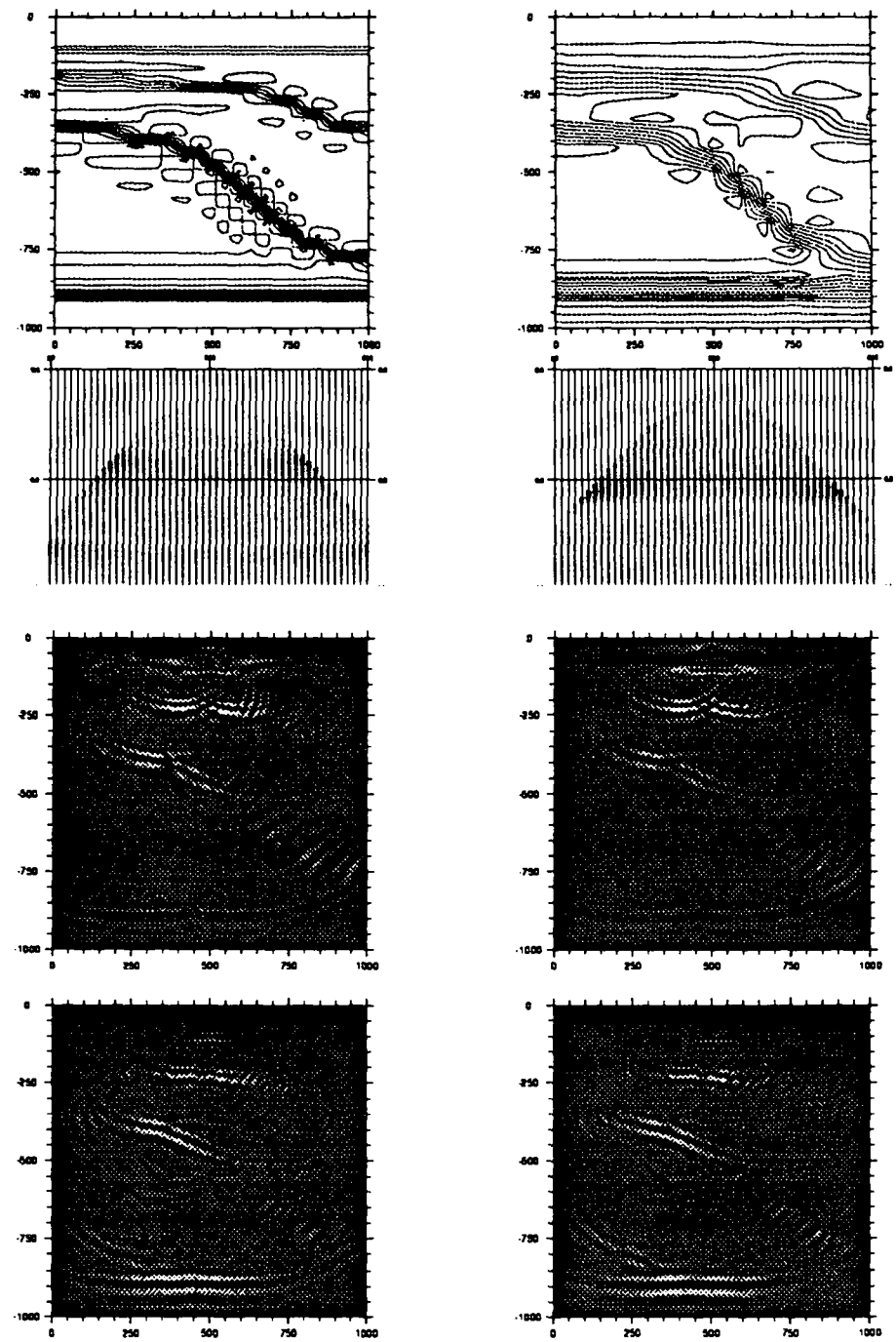


Figure 3.5: Moving down the scale cursor clears off the synthetics but degrades the migrated images (deepest reflector). Left: cursor = 3, right: cursor = 2. From top to bottom: propagator $v_s = P_s\bar{v}$, synthetic seismogram $c_s = \varphi(v_s, \sigma_s = cst)$ and reflectivity depth sections r_v and r_σ obtained from formula (2.3) to (2.13) with the same background v_s, σ_s.

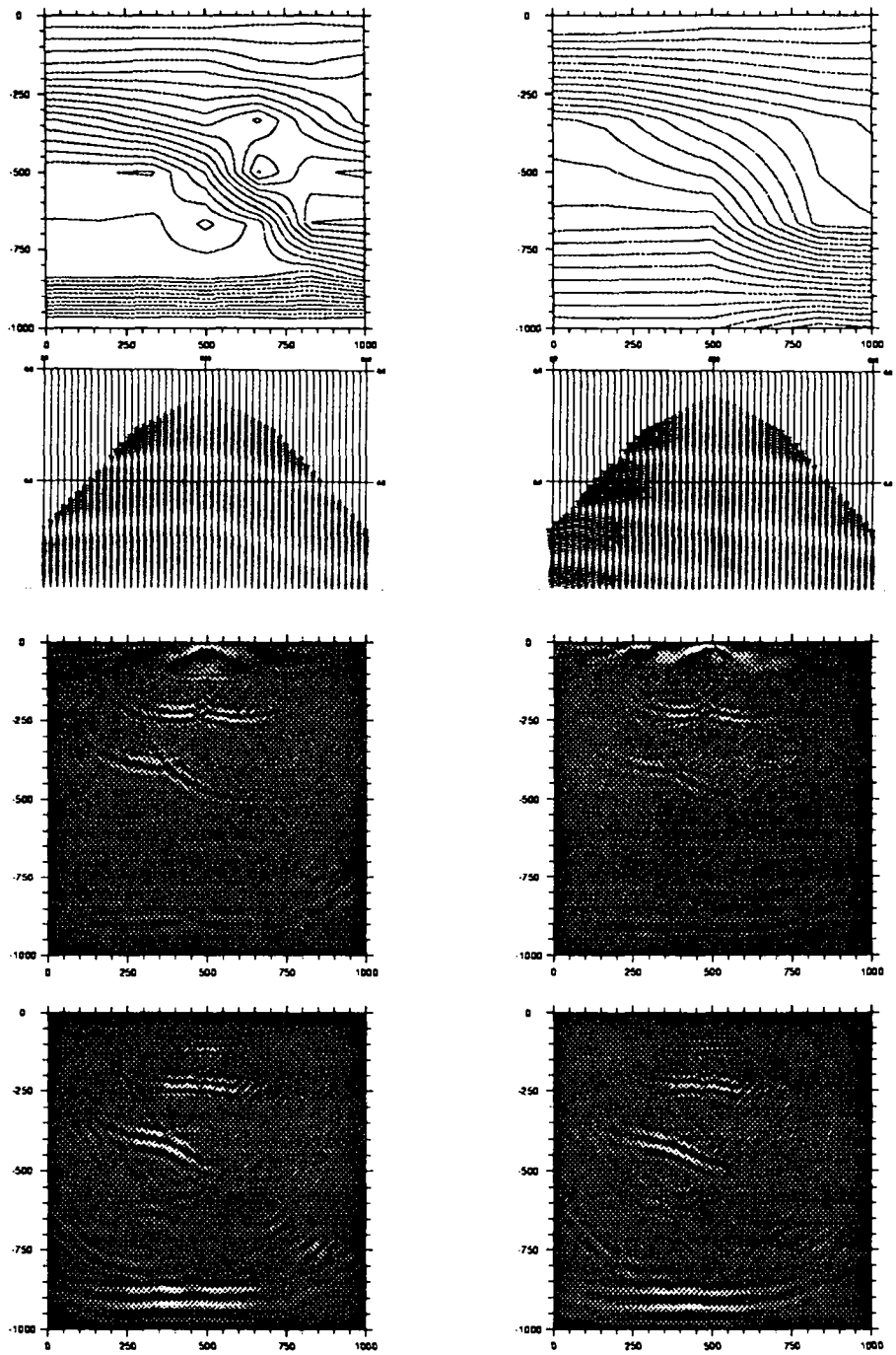


Figure 3.6: Moving down the scale cursor clears off the synthetics but degrades the migrated images (deepest reflector). Left: cursor = 1, right: cursor = 0. From top to bottom: propagator $v_s = P_s\bar{v}$, synthetic seismogram $c_s = \varphi(v_s, \sigma_s = cst)$ and reflectivity depth sections r_v and r_σ obtained from formula (2.3) to (2.13) with the same background v_s, σ_s.

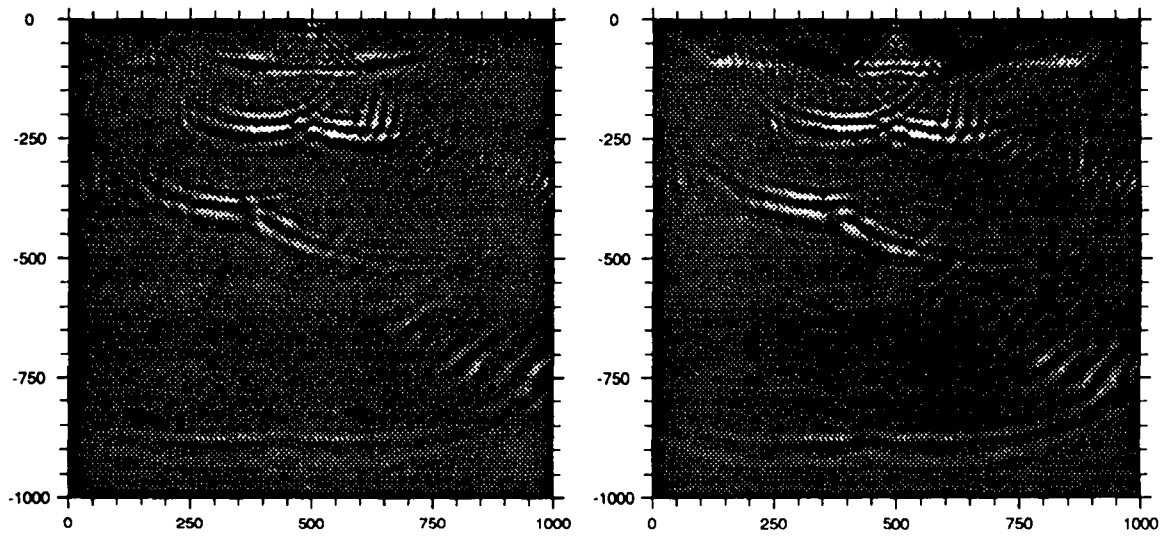


Figure 3.7: Influence of the orthogonalization of the subspaces $V_s = V_3$ and V_r on the shallow part of the v-reflectivity depth section r_v. Left: orthogonalized, right: non-orthogonalized.

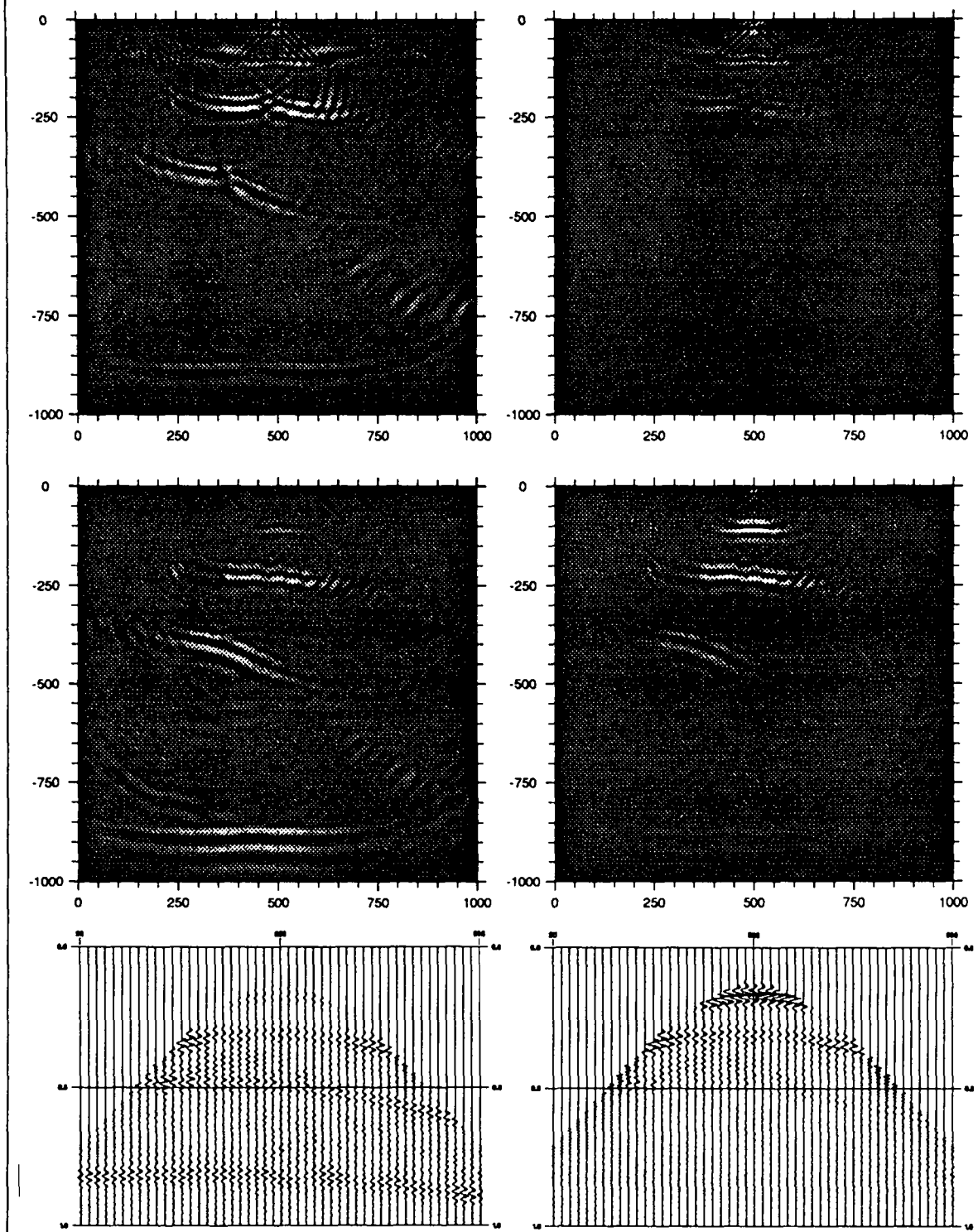


Figure 4.1: Correcting the synthetics (weight W_t) and the depth reflectivities (weight W_z) for geometrical spreading enhances the imaging of the deep reflectors and the quality the re-simulation. Left: amplified, right: not amplified. From top to bottom: the correction $\alpha_v \|v_s\| r_v$ to velocities, $\alpha_\sigma \|\sigma_s\| r_\sigma$ to impedances, and the re-simulated data. The displayed images are actually comparable in the sense that we use the same grey-map for the corresponding normalized optimal gradient updates $\alpha_v r_v$ and $\alpha_\sigma r_\sigma$ which is 25 equal grey levels in the range ± 1.6 % for v-reflectivity and ± 4 % for the σ-reflectivity.

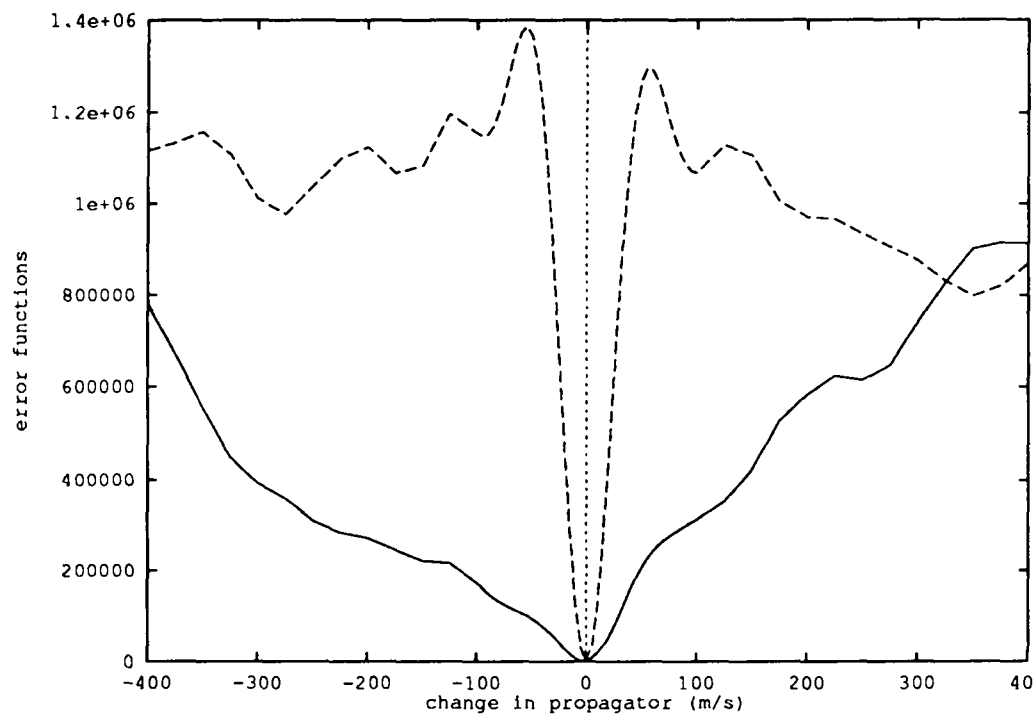


Figure 5.1: Effect of uniform perturbations of the true propagator with true reflectivity. MODIFIED SYNCLAY data set. The shallow water layer is known. Dashed line: usual least-square error function, solid line: MBTT modified least-square error function.

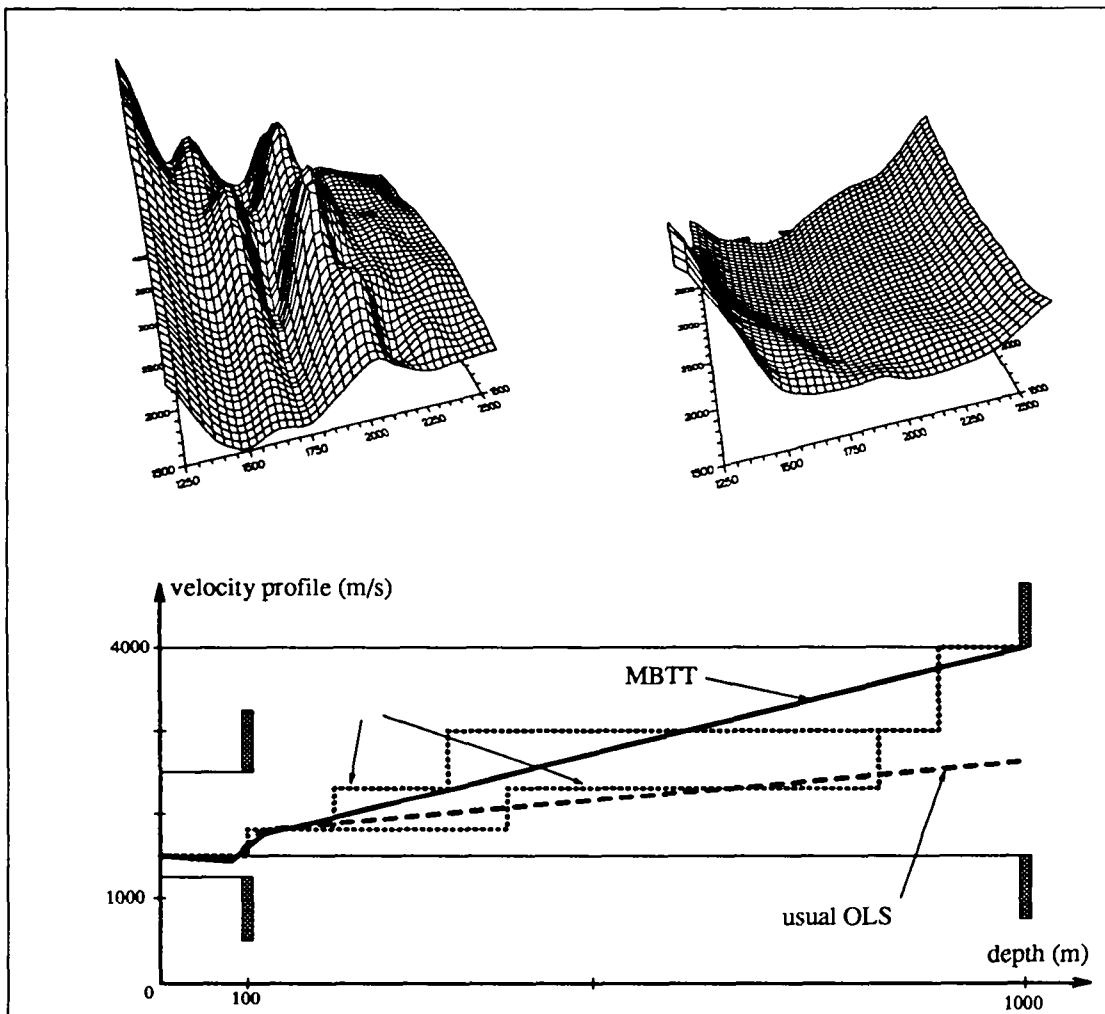


Figure 5.2: Search of the best linear profile for horizontally layered propagator. SYNCLAY data set. The shallow water layer is known. Top left: usual least-square error function, top right: MBTT modified least-square error function. First axis: velocity at depth 100 m (from 1250 m/s to 2500 m/s), second axis: velocity at depth 1000 m (from 1500 m/s to 4000 m/s). Bottom: corresponding optimal profiles for the usual error function (dashed line) and for the modified error function (solid line) plus, for reference, the target profiles of the SYNCLAY model from $x = 0$ m, to $x = 1000$ m, (dotted area).

6 CONCLUSION

The Migrated-Based Travel Time (MBTT) formulation of the waveform acoustic inversion problem consists in replacing the search for velocity and impedance depth section v and σ by the search of one propagator depth section v_s ("smooth velocity") and one reflectivity time section s ("demultiplied seismogram"). The objective of this change of unknowns is to render the usual least-square error, called $\mathcal{MJ}$ as a function of v_s and s, amenable to minimization by a local technique.

The findings are:

- the MBTT formulation can be adapted to handle the case of multiple shot data, which are necessary to obtain information on the low-frequency trends of the velocities through the coherency of the migrated shot images.
- A detailed analysis of the implementation of the MBTT formulation has been performed in the case of one shot:
  - a satisfying separation of propagator and reflector unknowns can be achieved by decomposing the slowness and impedance depth sections on two orthogonal subspaces V_s and V_r of smooth and rough functions;
  - fullwave reverse time migration operators have been defined for the separate imaging of the velocity and impedance reflectivity, which both i) give a good imaging of deep reflectors by incorporating proper time and depth weights, and ii) correspond to a descent direction in the space V_r for the least-square error;
  - updating the smooth backgrounds $(1/v_s, \sigma_s)$ with the migrated images $m_{1/v}$ and m_σ obtained from a given time reflectivity s can be performed in such a way that the updated parameters $(1/v, \sigma)$ generate a seismogram (the re-simulated reflectivity) close to s (up to the direct arrivals of course). Normalization makes the choice of the step size more independent of the data and the background parameters.
- Using the above implementation of the MBTT approach, the least-square error function $\mathcal{MJ}(v_s, s)$ has been shown, on two examples, to possess, for a fixed time reflectivity s, a nice behavior without parasitic local minima over a large range of the changes in the low-frequency trend of the propagator v_s. The behavior of the usual least-square error function $\mathcal{J}(v, \sigma)$ in the same situation was disastrous, as it exhibited many local minima.
- From a computational point of view, the calculation of $\mathcal{MJ}(v_s, s)$ requires one fullwave reverse time migration (*i.e.* two acoustic simulations with smooth backgrounds) plus one acoustical simulation (with full background). The computation of the two gradients of $\mathcal{MJ}$ with respect to v_s and s requires, as shown in Appendix C, one additional adjoint acoustical simulation (with full background) and one additional adjoint migration (*i.e.* two acoustic simulations with smooth backgrounds).

The above result seem extremely encouraging, and we are now proceeding to the implementation of the computation of the gradients of $\mathcal{MJ}$ and of its minimization.

Acknowledgments

This research was carried out as part of the Prestak Structural Interpretation consortium project (PSI) and was also supported by the FIRTECH Calcul Scientifique. The authors hereby acknowledge the support provided by the sponsors of these projects.

APPENDIX A: THE SYNCLAY SYNTHETIC TEST MODEL

The SYNCLAY model is a simple but realistic synthetic model: it is composed of five non-horizontal homogeneous layers leaning on a synclinal fold. The shallow layer is made of water. The Figure A.1 shows the material properties $(\bar{v}, \bar{\sigma})$ of the SYNCLAY model, the density is constant $\bar{\rho} = 1000$ kg/m^3. Hence velocity and impedance distributions have the same geometry. The grid size is $h = 5.2$ m, so our model of 1 km×1 km is discretized on a 192 × 192 FD grid.

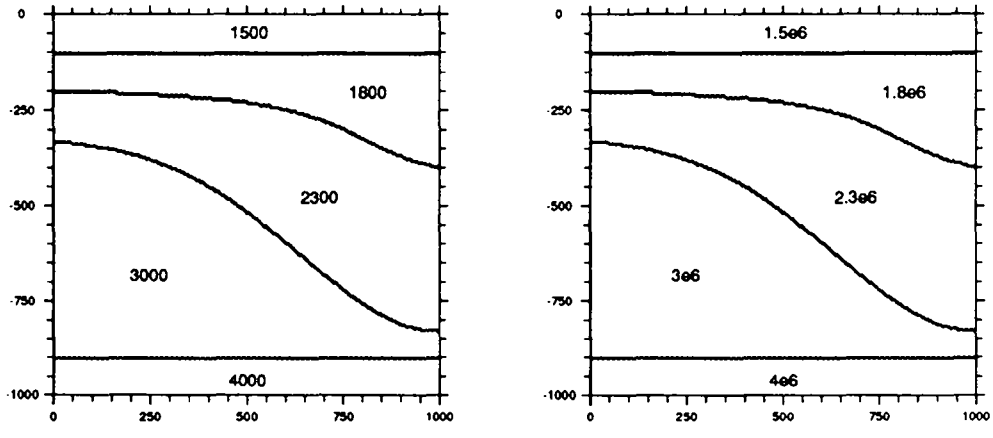


Figure A.1: Material properties of the SYNCLAY model. Left: velocity (m/s), right: impedance (kg/m^{2}s).

The Ricker point-source located in the middle at depth 10 m has peak frequency 36 Hz. The hydrophones are located all over the model at the same depth of 10 m and spaced by four grid meshes—*i.e.* 20 m—thus, there are 49 traces per recording. The recording duration is of 1 s sampled over 1250 time steps ($\Delta t = 0.8$ ms).

The SYNCLAY model, as defined above, allows fast computation of the wave equation using the numerical simulator describe in Clément (1990, 1991)—6 s cpu on a Cray-2—but is realistic enough to generate many local minima in the usual least-square error function $\mathcal{J}$. It is thus adequate to test the enhancement brought by the MBTT formulation. The corresponding (synthetic) data d_{SYNCLAY} are represented on Figure A.2. It is corrected by a factor $\sqrt{t}$ (as advocated in section 4) and the direct arrivals have been muted off.

The time reflectivity $\bar{s}$ which corresponds to the SYNCLAY data, *i.e.* such that

$$d_{\text{SYNCLAY}} = \varphi(\bar{v}, \bar{\sigma}) = \mathcal{M}\varphi(\bar{v}_s = P_3\bar{v}, \bar{s}). \qquad (A.1)$$

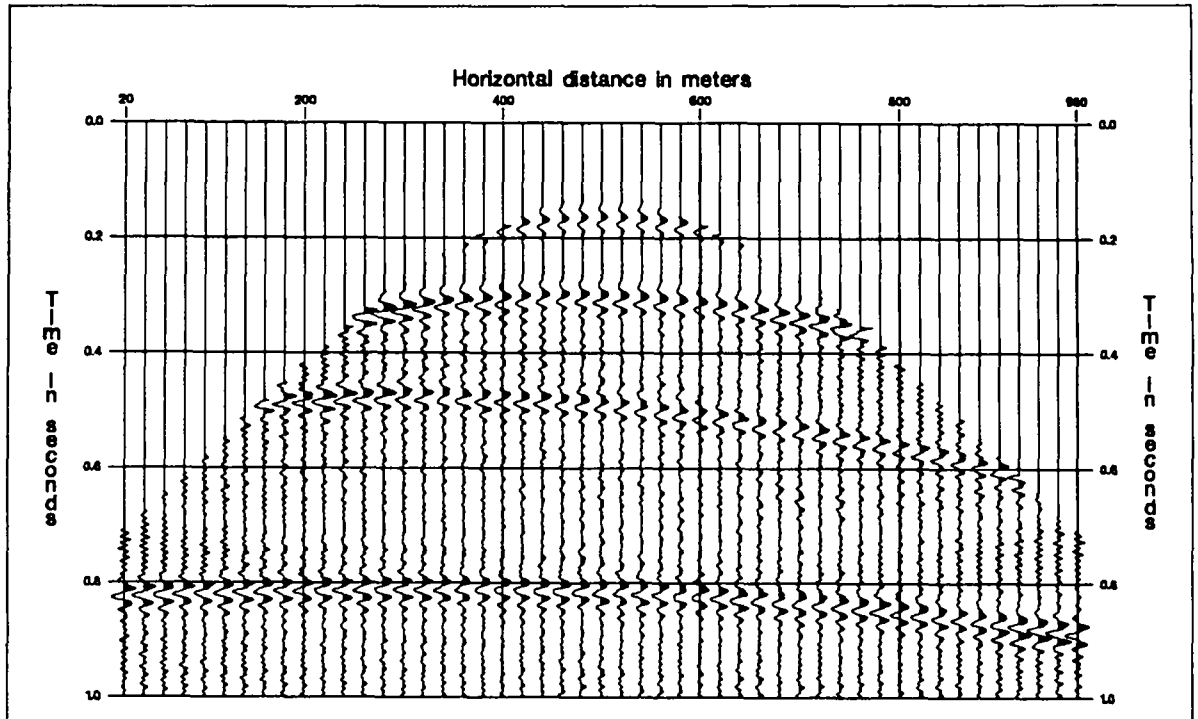


Figure A.2: The SYNCLAY data (corrected for geometrical spreading and muted off).

can be computed only by minimization of $\mathcal{MJ}$ with respect to s ($\mathcal{M}\phi$ is defined in 2.27). We could not perform this task, as we have not yet validated the computations for the gradient $\nabla_s \mathcal{MJ}$ described in appendix 6. But we have used a slight modification of the SYNCLAY model: the MODIFIED SYNCLAY model which has the same propagator $\bar{v}_s$ and a slightly different time reflectivity

$$s_{\text{MODIFIED SYNCLAY}} = d_{\text{SYNCLAY}} - \varphi(\bar{v}_s, \bar{\sigma}_s = cst) \tag{A.2}$$

The data corresponding to this model is then

$$\begin{aligned} d_{\text{MODIFIED SYNCLAY}} &= \mathcal{M}\varphi(\bar{v}_s, s_{\text{MODIFIED SYNCLAY}}) \\ &= \varphi(\underbrace{\bar{v}_s + \alpha_v \|\bar{v}_s\|_1 r_v}_{v_{\text{MODIFIED SYNCLAY}}}, \underbrace{\bar{\sigma}_s + \alpha_\sigma \|\bar{\sigma}_s\|_1 r_\sigma}_{\sigma_{\text{MODIFIED SYNCLAY}}}) \end{aligned} \tag{A.3}$$

where r_v and r_σ are the depth reflectivities obtained from $s_{\text{MODIFIED SYNCLAY}}$, *i.e.* they are the depth sections of Figure 3.5 for scale cursor = 3. We see Figure A.3 that the corresponding synthetic seismogram is very close to the previous one which shows that $s_{\text{MODIFIED SYNCLAY}}$ is a good approximation to the (yet unknown) $\bar{s}$.

APPENDIX B: REFLECTION ON A TRANSITION LAYER

The reflection coefficient r_s for a transition layer as the one depicted in Figure 3.1 in a

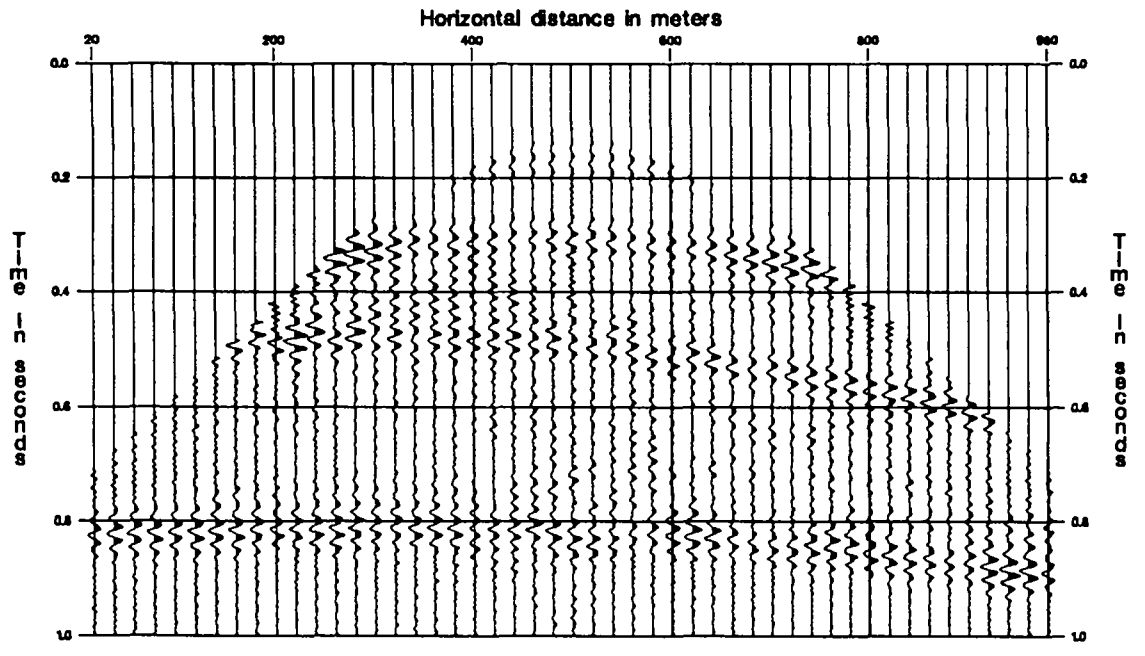


Figure A.3: The MODIFIED SYNCLAY data (corrected for geometrical spreading and muted off).

medium with constant density is given in Brekhovskikh (1960). Details about the calculation of the complex reflection coefficient are given in Berryman *et al.* (1958) in the general case of piecewise linear velocities.

Set $\zeta = 2\pi n_\lambda \frac{1-r}{r}$ (n_λ is the number of wavelength per grid cell and r is the reflection coefficient corresponding to an abrupt transition). We have to distinguish two cases:

• <u>for $0 < \zeta \leq 1$</u>,

$$|r_s| = \frac{\sinh \alpha_1}{(\sinh^2 \alpha_1 + 1 - \zeta^2)^{\frac{1}{2}}} \qquad (B.1)$$

$$\text{where} \qquad \alpha_1 = \frac{1}{2}(1-\zeta^2)^{\frac{1}{2}} \ln \frac{1+r}{1-r}.$$

• <u>for $\zeta > 1$</u>,

$$|r_s| = \frac{\sin \alpha_2}{(\sin^2 \alpha_2 + \zeta^2 - 1)^{\frac{1}{2}}} \qquad (B.2)$$

$$\text{where} \qquad \alpha_2 = \frac{1}{2}(\zeta^2-1)^{\frac{1}{2}} \ln \frac{1+r}{1-r}.$$

One checks easily that $|r_s| \to r$ when $n_\lambda \to 0$, *i.e.* when the transition layer becomes small in comparison to the wavelength. When $n_\lambda \to \frac{1}{2\pi}\frac{r}{1-r}$ then $|r_s|$ is continuous.

When $n_\lambda \to \infty$, we have the majoration

$$|r_s| \leq (\zeta^2 - 1)^{-\frac{1}{2}}.$$

Suppose that the reflection coefficient r is not too strong, *e.g.* $r \leq 0.33$, then

$$r^2(\zeta^2 - 1) \geq \frac{16}{9}\pi^2 n_\lambda^2 - \frac{1}{9}$$

and, with $n_\lambda \geq 0.3$, we have

$$\frac{16}{9}\pi^2 n_\lambda^2 - \frac{1}{9} \geq 16 n_\lambda^2.$$

Finally, with $r \leq 0.33$ and $n_\lambda \geq 0.3$, we have the following majoration of the reduction of the reflection coefficient

$$|\frac{r_s}{r}| \leq \frac{1}{4n_\lambda}. \qquad (B.3)$$

The curves of Figure 3.2 correspond to the formulas (B.1) and (B.2).

APPENDIX C: GRADIENTS OF THE MODIFIED COST FUNCTION $\mathcal{MJ}$

We present here the main steps required for the establishment of the formulas giving the exact gradients of the modified cost function $\mathcal{MJ}$. For sake of simplicity, we have considered only the case of a single shot. We introduce first some notations and state preliminary results.

C.1 PRELIMINARIES

Notations

• The letters "p" (in p_s, p, $\bar{p}_s$ or δp_s) and "q" (in q_s, q, $\bar{q}_s$ or δq_s) will respectively denote forward and backward wavefields with zero initial or final conditions:

$$p = (0, 0, p^2, \ldots, p^{NT}), \qquad q = (q^1, \ldots, q^{NT-1}, 0, 0),$$

where, at each time step k, p^k and q^k represent the values of the wavefields at the grid nodes M:

$$p^k = (p_M^k)_{M \in \Omega_{0h}}, \qquad q^k = (q_M^k)_{M \in \Omega_{0h}},$$

with $\Omega_{0h} = \Omega_h \backslash \Gamma_{0h}$ where Ω_h is the set of all grid nodes and Γ_{0h} the subset of all surface nodes (the free surface condition forces all wavefields to be zero at the surface nodes). The "s" subscript will denote a "smooth" wavefield, *i.e.* a wavefield computed using smooth velocity and impedance backgrounds, for example p_s which satisfies (2.3). The parameter fields v and σ are defined over the whole grid

$$v = (v_M)_{M \in \Omega_h}, \qquad \sigma = (\sigma_M)_{M \in \Omega_h}.$$

• The scalar products are entirely defined by their subscripts: "Ω_h", "Ω_{0h}", "G" or "t", respectively for a sum over all grid nodes, all grid nodes but surface nodes, all geophone locations or all time steps; or a space×time combination of them. For example, the usual least-square error function of (1.5) involves G and t:

$$\mathcal{J}_n = \frac{1}{2}\|d_n - c_n\|^2 = \frac{1}{2} < d_n - c_n, d_n - c_n >_{G,t} = \frac{1}{2}\sum_{k=2}^{NT}\sum_{G}(d_{n,G}^k - c_{n,G}^k)^2.$$

• We symbolize the first and second discrete time derivatives by

$$\begin{aligned}
u^{(1+)k} &= \frac{1}{\Delta t}(u^{k+1} - u^k), \\
u^{(1-)k} &= \frac{1}{\Delta t}(u^k - u^{k-1}) = u^{(1+)k-1}, \\
u^{(1)k} &= \frac{1}{2\Delta t}(u^{k+1} - u^{k-1}) = \frac{1}{2}(u^{(1+)k} + u^{(1-)k}), \\
u^{(2)k} &= \frac{1}{\Delta t^2}(u^{k+1} - 2u^k + u^{k-1}) = u^{(1+)(1-)k} = u^{(1-)(1+)k} = u^{(1+)(1+)k-1}.
\end{aligned}$$

• We introduce a specific notation for the l^1-normalization: an tilded letter will denote a field normalized by its own l^1-norm

$$\tilde{u} = \frac{u}{\|u\|_1}.$$

One will always be explicitly advised for other tilde occurrences, *e.g.* $\tilde{\mathcal{J}}$ will denote the least-square error between the l^1-normalized seismograms.

• Let f be a function of x and y, where x and y are fields defined on a grid Ω. When differentiating f, we will use the following notations

$$\delta f = \delta_x f + \delta_y f = < \nabla_x f, \delta x >_\Omega + < \nabla_y f, \delta y >_\Omega .$$

• The forward wave equation with second order absorbing conditions (Clément, 1990, 1991) will be denoted by

$$C(v,\sigma)p = f.$$

At all interior nodes $M \in \Omega_h \backslash \partial\Omega_h$ ($\partial\Omega_h$= surface + artificial boundary), this simplifies to

$$B\left(\frac{1}{v\sigma}\right)p^{(2)k} + A\left(\frac{v}{\sigma}\right)p^k = f^k \qquad \text{for } k = 1, \ldots, NT-1$$

where $B(\alpha)$ is the diagonal mass matrix defined by

$$< B(\alpha)p, q >_{\Omega_{0h}} = < \alpha p, q >_{\Omega_{0h}}$$

and $A(\alpha)$ the penta-diagonal generalized laplacian matrix defined by

$$< A(\alpha)p, q >_{\Omega_{0h}} = < p, A(\alpha)q >_{\Omega_{0h}} = < \alpha, \nabla p \nabla q >_{\Omega_h}$$

with the following discrete gradient correlation

$$\begin{aligned}
(\nabla p \nabla q)_M &= \frac{(p_M - p_R)(q_M - q_R) + (p_M - p_L)(q_M - q_L)}{2\Delta x^2} \\
&\quad + \frac{(p_M - p_U)(q_M - q_U) + (p_M - p_D)(q_M - q_D)}{2\Delta z^2}
\end{aligned}$$

(U (upper), R (right), D (down) and L (left) are the four neighbors of node M).

- The backward wave equation with second order absorbing conditions will be denoted by

$$C^T(v,\sigma)q = g.$$

At all interior nodes $M \in \Omega_h \backslash \partial\Omega_h$, this simplifies to

$$B\left(\frac{1}{v\sigma}\right) q^{(2)k} + A\left(\frac{v}{\sigma}\right) q^k = g^k \qquad \text{for } k = NT, \ldots, 2$$

with the same definition as above for B and A. It is important to notice that $C(v,\sigma)$ and $C^T(v,\sigma)$ are actually transposed one from the other (both are square matrices of (huge) size $(NT-1) \times |\Omega_{0h}|$)

$$< C(v,\sigma)p, q >_{\Omega_{0h},t} = < p, C^T(v,\sigma)q >_{\Omega_{0h},t}$$

for any wavefields p and q.

Differentiation of the l^1-normalization

Let u be a field defined over a grid Ω made of NB nodes. Let $\| \ \|_1$ be the following l^1-norm on this grid

$$\|u\|_1 = \frac{1}{NB} \sum_{M \in \Omega} |u_M|.$$

Then differentiating this norm and its inverse yields

$$\delta(\|u\|_1) = \frac{1}{NB} < \mathrm{sgn}(u), \delta u >_\Omega, \tag{C.1}$$

$$\delta\left(\frac{1}{\|u\|_1}\right) = -\frac{1}{NB\,\|u\|_1^{\,2}} < \mathrm{sgn}(u), \delta u >_\Omega \tag{C.2}$$

where $\mathrm{sgn}(u)_M$ denotes the sign of u_M. Hence

$$\delta(\tilde{u}) = \delta\left(\frac{u}{\|u\|_1}\right) = \frac{1}{\|u\|_1}\delta u + u\delta\left(\frac{1}{\|u\|_1}\right)$$

and for any field v defined over Ω, we have

$$\begin{aligned} < v, \delta\left(\frac{u}{\|u\|_1}\right) >_\Omega &= < \frac{v}{\|u\|_1}, \delta u >_\Omega + < v, -\frac{u}{NB\,\|u\|_1^{\,2}} < \mathrm{sgn}(u), \delta u >_\Omega >_\Omega, \\ &= < \frac{v}{\|u\|_1} - \frac{\mathrm{sgn}(u)}{NB} < \frac{u}{\|u\|_1}, \frac{v}{\|u\|_1} >_\Omega, \delta u >_\Omega . \end{aligned}$$

So, with $\tilde{v} = v/\|u\|_1$ and $\mathcal{M}v = \tilde{v} - \frac{\mathrm{sgn}(\tilde{u})}{NB} < \tilde{u}, \tilde{v} >_\Omega$, we have the following characterization of the action of the differential of the l^1-normalization on any test vector v

$$< v, \delta\tilde{u} >_\Omega = < \mathcal{M}v, \delta u >_\Omega . \tag{C.3}$$

Gradients of the usual least-square error functions

We shall need to compute the gradients, with respect to slowness $1/v$ and impedance σ, of the error functions

$$\mathcal{J}(v,\sigma) = \frac{1}{2}\|d - W_t G\, p\|_2{}^2 \qquad (cf.\ (1.5))$$

and

$$\tilde{\mathcal{J}}(v,\sigma) = \frac{1}{2}\|\frac{d}{\|d\|_1} - \frac{W_t G\, p}{\|W_t G\, p\|_1}\|_2{}^2 \qquad (cf.\ (2.28)).$$

They are two special cases of the more general problem of computing the gradients, with respect to $1/v$ and σ, of

$$j = j(p)$$

where j is a given functional over p, and where p is solution of the wave equation

$$C(v,\sigma)p = f.$$

These gradients can be computed using a technique similar to the one described in paragraph C.2 (but applied to a simpler case), as described in Clément (1990, appendix B):

$$\nabla_{1/v} j = -\frac{1}{\sigma} < p^{(1+)}, q^{(1+)} >_t - \frac{v^2}{\sigma} < \nabla p, \nabla q >_t = \chi_{1/v}(p,q;v,\sigma) \qquad (C.4)$$

$$\nabla_{\sigma} j = \frac{1}{v\sigma^2} < p^{(1+)}, q^{(1+)} >_t - \frac{v}{\sigma^2} < \nabla p, \nabla q >_t = \chi_{\sigma}(p,q;v,\sigma) \qquad (C.5)$$

where q is solution of the backward wave equation

$$C(v,\sigma)^T q = -\nabla_p j. \qquad (C.6)$$

The simple error function of equation (1.5) corresponds to

$$j(p) = \frac{1}{2}\|d - W_t G\, p\|^2 = \frac{1}{2}\|e\|^2 = \frac{1}{2} < e, e >_{G,t}$$

with $e = d - c$ and $c = W_t G\, p$. Differentiating j yields

$$\delta j = < e, \delta e >_{G,t} = - < G^T W_t\, e, \delta p >_{\Omega_{0h},t},$$

hence

$$\nabla_p j = -G^T W_t\, e. \qquad (C.7)$$

The normalized error function of equation (2.28) corresponds to

$$\tilde{j}(p) = \frac{1}{2}\|\tilde{d} - \tilde{c}\|^2 = \frac{1}{2}\|e\|^2 = \frac{1}{2} < e, e >_{G,t}$$

with $e = \tilde{d} - \tilde{c}$ and $c = W_t G\, p$. Differentiating $\tilde{j}$ yields again

$$\delta \tilde{j} = < e, \delta e >_{G,t} = < e, \delta \tilde{c} >_{G,t}\ .$$

Let $\tilde{e} = e/\|c\|_1$ and $\mathcal{M}e = \tilde{e} - \mathrm{sgn}(\tilde{c}) < \tilde{c}, \tilde{e} >_{G,t} /(NG \times (NT+1))$, where NG is the number of geophones and $NT+1$ the number of time steps, we obtain from equation (C.3) that

$$\delta \tilde{j} = - < \mathcal{M}e, \delta c >_{G,t},$$

hence

$$\nabla_p \tilde{j} = -G^T W_t\, \mathcal{M}e. \qquad (C.8)$$

C.2 LAGRANGIAN FOR THE CALCULATION OF THE GRADIENTS OF $\mathcal{MJ}$

The MBTT-modified error function

$$\mathcal{MJ} : (v_s, s) \longrightarrow \frac{1}{2}\|\tilde{d} - \tilde{c}\|_2^2 \qquad (C.9)$$

is defined through a complex sequence of operations (step 1 to 4) described in section 2. In order to handle the difficulties one after the other, we break this long chain of operations after step 1.2:

$$(v_s, s) \overset{\text{eqs. (2.3) and (2.4)}}{\longrightarrow} (p_s, q_s; v_s, \sigma_s) \overset{\text{eqs. (2.5) to (2.28)}}{\longrightarrow} \frac{1}{2}\|\tilde{d} - \tilde{c}\|_2^2 \qquad (C.10)$$

where p_s and q_s are the "smooth" wavefields used to migrate the time reflectivity s into a depth reflectivity, and where the function H is defined through the sequence of operations described through equations (2.5) to (2.28).

In the present section we shall suppose that we know how to evaluate H and all its partial derivatives, and concentrate on the handling of the state equations (2.3), (2.4). In the next section we shall indicate how the partials of H with respect to p_s, q_s *etc...* can actually be computed.

As one can see by comparing (C.9) and (C.10), minimizing $\mathcal{MJ}$ with respect to v_s and s is equivalent to minimizing H with respect to p_s, q_s, v_s and σ_s under the constraints that the equations (2.3) and (2.4) are satisfied. We can write the Lagrangian function associated with this constrained optimization problem:

$$\begin{aligned} \mathcal{ML}(p_s, \bar{q}_s, q_s, \bar{p}_s; v_s, \sigma_s, s) \;=\; & H(p_s, q_s; v_s, \sigma_s) \\ & + < C(v_s, \sigma_s)p_s - f, \bar{q}_s >_{\Omega_{0h}, t} \\ & + < C(v_s, \sigma_s)^T q_s - G^T W_t \; s, \bar{p}_s >_{\Omega_{0h}, t} . \end{aligned} \qquad (C.11)$$

When p_s and q_s satisfy the constraints, then for any adjoint wavefields $\bar{q}_s$ and $\bar{p}_s$ and for any inversion parameters v_s and s, we have

$$\mathcal{MJ}(v_s, s) = \mathcal{ML}(p_s, \bar{q}_s, q_s, \bar{p}_s; v_s, \sigma_s, s).$$

Thus we can differentiate this equation with respect to $1/v_s$ and s for fixed (and yet unknown) $\bar{q}_s$ and $\bar{p}_s$:

$$\delta\mathcal{MJ} = \delta_{p_s}\mathcal{ML} + \delta_{q_s}\mathcal{ML} + \delta_{1/v_s}\mathcal{ML} + \delta_s\mathcal{ML}. \qquad (C.12)$$

If we require that the adjoint states satisfy

$$\delta_{p_s}\mathcal{ML} = < \nabla_{p_s}\mathcal{ML}, \delta p_s >_{\Omega_{0h}, t} = 0 \qquad \text{for all } \delta p_s \qquad (C.13)$$

$$\delta_{q_s}\mathcal{ML} = < \nabla_{q_s}\mathcal{ML}, \delta q_s >_{\Omega_{0h}, t} = 0 \qquad \text{for all } \delta q_s, \qquad (C.14)$$

then formula (C.12) simplifies to

$$\delta\mathcal{MJ} = < \nabla_{1/v_s}\mathcal{ML}, \delta 1/v_s >_{\Omega_h} + < \nabla_s\mathcal{ML}, \delta s >_{G,t} . \qquad (C.15)$$

But (C.13) is satisfied as soon as $\bar{q}_s$ is solution of the backward wave equation

$$C(v_s, \sigma_s)^T \bar{q}_s = -\nabla_{p_s} H \qquad (C.16)$$

and (C.14) is satisfied as soon as $\bar{p}_s$ is solution of the forward wave equation

$$C(v_s, \sigma_s)\bar{p}_s = -\nabla_{q_s} H. \qquad (C.17)$$

Then we see from (C.15) that the gradients of the modified least-square error function are simply given by

$$\nabla_{1/v_s} \mathcal{MJ} = \nabla_{1/v_s} \mathcal{ML} \qquad \text{and} \qquad \nabla_{\sigma_s} \mathcal{MJ} = \nabla_{\sigma_s} \mathcal{ML}$$

where p_s, q_s, $\bar{q}_s$ and $\bar{p}_s$ are defined by (2.3), (2.4), (C.16) and (C.17). This gives (see Clément, 1990, appendix B)

$$\nabla_{1/v_s} \mathcal{MJ} = \nabla_{1/v_s} H + P_s \left\{ \chi_{1/v}(p_s, \bar{q}_s; v_s, \sigma_s) + \chi_{1/v}(\bar{p}_s, q_s; v_s, \sigma_s) \right\}, \qquad (C.18)$$

$$\nabla_{\sigma_s} \mathcal{MJ} = W_t G \bar{p}_s. \qquad (C.19)$$

Notice that $\nabla_{\sigma_s} \mathcal{MJ}$ is the seismogram corresponding to the adjoint $\bar{p}_s$ of the backward propagation q_s of the time reflectivity s.

C.3 DIFFERENTIATION OF H

In order to obtain explicit formulas for the right-hand sides of the adjoint states and the first term in $\nabla_{1/v_s} \mathcal{MJ}$, we have to differentiate H with respect to its arguments (except σ_s which is supposed known).

The function H is defined through the sequence of operations from equation (2.5) to (2.28):

$$H : (p_s, q_s; v_s, \sigma_s) \overset{\text{eqs. (2.5) to (2.28)}}{\longrightarrow} \frac{1}{2} \| \tilde{d} - \tilde{c} \|_2^{\,2}. \qquad (C.20)$$

Once again, we break this chain of operations (after step 2) in order to ease the calculation of the partial derivatives of H:

$$(p_s, q_s; v_s, \sigma_s) \overset{\text{eqs. (2.5) to (2.25)}}{\longrightarrow} (v, \sigma) \overset{\text{eqs. (2.26) to (2.28)}}{\longrightarrow} \frac{1}{2} \| \tilde{d} - \tilde{c} \|_2^{\,2}. \qquad (C.21)$$

We have already calculated, in appendix C.1, the gradients of the usual error function $\tilde{\mathcal{J}}$(the last arrow of (C.21)) with respect to v and σ. It will then be easy, using the chain rule differentiation, to infer the partials with respect to p_s, q_s, v_s and σ_s as equations (2.5) to (2.25) are all made of explicit closed-form formulas.

From (C.20) and (C.21) we see that

$$\delta H = \delta \tilde{\mathcal{J}} = < \nabla_{1/v} \tilde{\mathcal{J}}, \delta(1/v) >_{\Omega_h} + < \nabla_{\sigma} \tilde{\mathcal{J}}, \delta\sigma >_{\Omega_h}. \qquad (C.22)$$

The gradients of $\tilde{\mathcal{J}}$ are computed through correlations (C.4) and (C.5) between the resimulated wavefield p of (2.27), and the adjoint state q of (C.6) with right-hand side given by (C.8). Then, from (2.22) and (2.23),

$$\delta(1/v) = \delta(1/v_s) + \alpha_{1/v} \delta(\|1/v_s\|_1) r_{1/v} + \alpha_{1/v} \|1/v_s\|_1 \delta r_{1/v} \qquad (C.23)$$

$$\delta\sigma = \alpha_\sigma \|\sigma_s\|_1 \delta r_\sigma \qquad (C.24)$$

From (C.1), we have

$$< \nabla_{1/v}\tilde{\mathcal{J}}, \alpha_{1/v}\delta(\|1/v_s\|_1)r_{1/v} >_{\Omega_h} = < \frac{\alpha_{1/v}}{NB} < \nabla_{1/v}\tilde{\mathcal{J}}, r_{1/v} >_{\Omega_h}, \delta(1/v_s) >_{\Omega_h} . \qquad (C.25)$$

We set successively

$$\begin{aligned} \gamma_{1/v} &= \alpha_{1/v}\|1/v_s\|_1 \nabla_{1/v}\tilde{\mathcal{J}}, \\ \mathcal{M}\gamma_{1/v} &= \tilde{\gamma}_{1/v} - \frac{\mathrm{sgn}(r_{1/v})}{NB} < r_{1/v}, \tilde{\gamma}_{1/v} >_{\Omega_h}, \\ g_{1/v} &= -P_r W_z P_r \mathcal{M}\gamma_{1/v} \end{aligned}$$

in order to write

$$\begin{aligned} < \nabla_{1/v}\tilde{\mathcal{J}}, \alpha_{1/v}\|1/v_s\|_1 \delta r_{1/v} >_{\Omega_h} &= < \gamma_{1/v}, \delta r_{1/v} >_{\Omega_h} \\ &= < \mathcal{M}\gamma_{1/v}, \delta m_{1/v} >_{\Omega_h} \\ &= < g_{1/v}, \delta\chi_{1/v} >_{\Omega_h} . \end{aligned} \qquad (C.26)$$

In the same way, we can write

$$< \nabla_\sigma \tilde{\mathcal{J}}, \alpha_\sigma \|\sigma_s\|_1 \delta r_\sigma >_{\Omega_h} = < g_\sigma, \delta\chi_\sigma >_{\Omega_h} . \qquad (C.27)$$

The correlations are linear functions of the wavefields, hence

$$\delta\chi_{1/v} = \chi_{1/v}(\delta p_s, q_s; v_s, \sigma_s) + \chi_{1/v}(p_s, \delta q_s; v_s, \sigma_s) + \delta_{1/v_s}\chi_{1/v}(p_s, q_s; v_s, \sigma_s) \qquad (C.28)$$

$$\delta\chi_\sigma = \chi_\sigma(\delta p_s, q_s; v_s, \sigma_s) + \chi_\sigma(p_s, \delta q_s; v_s, \sigma_s) + \delta_{1/v_s}\chi_\sigma(p_s, q_s; v_s, \sigma_s). \qquad (C.29)$$

We continue the calculation only for $\chi_{1/v}$, p_s and $1/v_s$, since it is similar for the other terms. After integration by parts in time, we can rewrite (C.4) as

$$\chi_{1/v}(p_s, q_s; v_s, \sigma_s) = \frac{1}{\sigma_s} < p_s, q_s^{(2)} >_t - \frac{v_s^2}{\sigma_s} < \nabla p_s, \nabla q_s >_t$$

hence, from the definition of A,

$$< g_{1/v}, \delta_{p_s}\chi_{1/v} >_{\Omega_h} = < \frac{g_{1/v}}{\sigma_s} q_s^{(2)} - A\left(\frac{g_{1/v}v_s^2}{\sigma_s}\right) q_s, \delta p_s >_{\Omega_{0h}, t}$$

and, after elimination of the second derivative in time with the wave equation (2.4),

$$\begin{aligned} < g_{1/v}, \delta_{p_s}\chi_{1/v} >_{\Omega_h} &= < g_{1/v} v_s G^T W_t s - g_{1/v} v_s A\left(\frac{v_s}{\sigma_s}\right) q_s \\ &\quad - A\left(\frac{g_{1/v}v_s^2}{\sigma_s}\right) q_s, \delta p_s >_{\Omega_{0h}, t} . \end{aligned} \qquad (C.30)$$

Differentiate (C.4) with respect to $1/v_s$ gives

$$\begin{aligned} \delta_{1/v_s}\chi_{1/v,M} &= < \nabla_{1/v_s}\chi_{1/v,M}, \delta(1/v_s) >_{\Omega_h} \\ &= \frac{\partial\chi_{1/v,M}}{\partial(1/v_{s,M})}\delta(1/v_{s,M}) \\ &= \frac{2v_{s,M}^3\delta(1/v_{s,M})}{\sigma_{s,M}}(< \nabla p_s, \nabla q_s >_t)_M \end{aligned}$$

hence

$$< g_{1/v}, \delta_{1/v_s} \chi_{1/v} >_{\Omega_h} = < g_{1/v}, \frac{2v_s^3 \delta(1/v_s)}{\sigma_s} < \nabla p_s, \nabla q_s >_t >_{\Omega_h}$$
$$= < \frac{2g_{1/v} v_s^3}{\sigma_s} < \nabla p_s, \nabla q_s >_t, \delta(1/v_s) >_{\Omega_h} . \qquad \text{(C.31)}$$

We proceed exactly in the same way to obtain the formulas for the "q_s" and "σ" terms. We set

$$\check{A} = \left(-g_{1/v} v_s + \frac{g_\sigma}{\sigma_s}\right) A\left(\frac{v_s}{\sigma_s}\right) - A\left(\frac{g_{1/v} v_s^2}{\sigma_s}\right) - A\left(\frac{g_\sigma v_s}{\sigma_s^2}\right).$$

And we find

$$\nabla_{p_s} H = \left(g_{1/v} v_s - \frac{g_\sigma}{\sigma_s}\right) G^T W_t s + \check{A} q_s, \qquad \text{(C.32)}$$

$$\nabla_{q_s} H = \left(g_{1/v} v_s - \frac{g_\sigma}{\sigma_s}\right) f + \check{A} p_s, \qquad \text{(C.33)}$$

$$\nabla_{1/v_s} H = P_s \left\{ \nabla_{1/v} \tilde{\mathcal{J}} + \frac{\alpha_{1/v}}{NB} < \nabla_{1/v} \tilde{\mathcal{J}}, r_{1/v} >_{\Omega_h} - \frac{g_\sigma}{v_s^2 \sigma_s^2} < p_s^{(1+)}, q_s^{(1+)} >_t \right.$$
$$\left. + \left(\frac{2g_{1/v} v_s^3}{\sigma_s} - \frac{g_\sigma}{\sigma_s^2}\right) < \nabla p_s, \nabla q_s >_t \right\}. \qquad \text{(C.34)}$$

REFERENCES

Al-Yahya, K., 1989, Velocity analysis by iterative profile migrations: Geophysics, 54(6), 718–729.

Berryman, L. H., Goupillaud, P. L., and Waters, K.H., 1958, Reflections from multiple transition layers, part one: Theoretical results: Geophysics, 23(2), 223–243.

Brekhovskikh, L. M., 1960, Waves in Layered Media, volume 6, Academic Press.

Chavent, G., and Jacewitz, C. A., 1990, Automatic determination of background velocities by multiple migration fitting: 60th Ann. Internat. Mtg., Soc. Expl. Geophys.

Clément, F., 1990, A migration-based travel-time formulation for the inversion of 2d seismic data: PSI 1990 Annual Report, IFP, Rueil Malmaison, France.

Clément, F., 1991, A migration-based travel-time formulation for the inversion of 2d seismic reflection data: Presented at 1st Internat. Conf. on Mathematical and Numerical Aspects of Wave Propagation Phenomena, INRIA and SIAM.

Clément, F., 1991, A migration-based travel-time formulation for the determination of 2d velocity distributions: PSI 1991 Annual Report, IFP, Rueil Malmaison, France.

Kleyn, A. H., 1983, Seismic Reflection Interpretation: Elsevier Applied Science.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Proc. of Conf. on Inverse Scattering, Theory and Applic., SIAM.

Symes, W. W., 1988, Velocity inversion by coherency optimisation: 58th Ann. Internat. Mtg., Soc. Expl. Geophys.

Symes, W. W., and Carrazone, J. J., 1991, Velocity inversion by differential semblance optimization: Geophysics, 56(5).

Versteeg, R., and Geoltrain, S., 1990, Sensibility of prestack depth migration to the velocity model: PSI 1990 Annual Report, IFP, Rueil Malmaison, France.

Versteeg, R., Ehinger, A., and Geoltrain, S., 1991, Sensitivity of migration coherency panels to the velocity model: PSI 1991 Annual Report, IFP, Rueil Malmaison, France.

Unité de Recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 LE CHESNAY Cedex (France)

Unité de Recherche INRIA Lorraine Technopôle de Nancy-Brabois - Campus Scientifique
615, rue du Jardin Botanique - B.P. 101 - 54602 VILLERS LES NANCY Cedex (France)
Unité de Recherche INRIA Rennes IRISA, Campus Universitaire de Beaulieu 35042 RENNES Cedex (France)
Unité de Recherche INRIA Rhône-Alpes 46, avenue Félix Viallet - 38031 GRENOBLE Cedex (France)
Unité de Recherche INRIA Sophia Antipolis 2004, route des Lucioles - B.P. 93 - 06902 SOPHIA ANTIPOLIS Cedex (France)

EDITEUR
INRIA - Domaine de Voluceau - Rocquencourt - B.P. 105 - 78153 LE CHESNAY Cedex (France)

ISSN 0249 - 6399