A 2D nonlinear inversion of well-seismic data

2011 Inverse Problems 27 055005

(http://iopscience.iop.org/0266-5611/27/5/055005)

Inverse Problems **27** (2011) 055005 (26pp) doi:10.1088/0266-5611/27/5/055005

A 2D nonlinear inversion of well-seismic data

Ludovic Métivier[1,2], Patrick Lailly[1], Florence Delprat-Jannaud[1] and Laurence Halpern[2]

[1] IFP, 1 et 4 avenue de Bois-Préau, 92852 Rueil-Malmaison Cedex, France
[2] LAGA, UMR 7539 CNRS, Université Paris 13, 93430 Villetaneuse, France

E-mail: florence.delprat-jannaud@ifpen.fr

Received 19 May 2010, in final form 18 November 2010
Published 29 March 2011
Online at stacks.iop.org/IP/27/055005

Abstract

Well-seismic data such as vertical seismic profiles are supposed to provide detailed information about the elastic properties of the subsurface at the vicinity of the well. Heterogeneity of sedimentary terrains can lead to far from negligible multiple scattering, one of the manifestations of the nonlinearity involved in the mapping between elastic parameters and seismic data. We present a 2D extension of an existing 1D nonlinear inversion technique in the context of acoustic wave propagation. In the case of a subsurface with gentle lateral variations, we propose a regularization technique which aims at ensuring the stability of the inversion in a context where the recorded seismic waves provide a very poor illumination of the subsurface. We deal with a huge size inverse problem. Special care has been taken for its numerical solution, regarding both the choice of the algorithms and the implementation on a cluster-based supercomputer. Our tests on synthetic data show the effectiveness of our regularization. They also show that our efforts in accounting for the nonlinearities are rewarded by an exceptional seismic resolution at distances of about 100 m from the well. They also show that the result is not very sensitive to errors in the estimation of the velocity distribution, as far as these errors remain realistic in the context of a medium with gentle lateral variations.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Seismic prospecting, i.e. the idea of using seismic waves to get information about the subsurface, is understood as a high resolution technique compared with potential methods such as gravimetry. Geophysicists, as many other physicists, always are in quest of a higher resolution: they want to get as much detail as possible of the object they inspect. However, resolution is not defined just by the physical means used for scanning the object. The way the

data are processed, i.e. algorithms, for the object reconstruction, can have a strong influence on the resolution. This is the field on which our work is to be placed. Another important point is the ability to get quantitative information about the composition of the object. This is the second motivation for our paper.

Data acquisition for seismic prospecting mainly concerns the acquisition of surface data. However, well-seismic data often are interesting for detailed studies of reservoirs. If the data quality is usually better compared with that of surface data, the part of the subsurface that is illuminated, that is the part that we can image, is of very limited extension. Classically, we distinguish the following.

- The acquisition of vertical seismic profiles (VSP) in which, the receivers covering some depth range within the well, the source is located at the surface at the vicinity of the borehole; for such acquisitions, waves usually propagate along directions close to the vertical so that most of the imaging techniques [19] rely on a 1D propagation assumption.
- The acquisition of walkaway data in which, the receivers still covering some depth range within the well, the source is moved, shot after shot, at different offsets from the borehole; by using such acquisitions, we expect to enlarge the part of the subsurface that can be imaged.

In the following, we will make a distinction between different zones associated with different depth ranges: the 'overburden' and the 'target' separated by the horizontal plane at the depth of the shallowest receiver. The target itself will be subdivided into the 'upper target' corresponding to the depth interval covered by the receivers and the 'lower target' below. Well-seismic data (figure 1) show a first arrival which is most often a downgoing wave. This arrival is usually followed by many downgoing waves: these waves have undergone (at least) a 'reflection'[3] at some interface in the overburden before propagating down again. The presence of many downgoing waves reflects the vertical heterogeneity of the overburden. Well-seismic data also show upgoing waves which are the result of 'reflections' that have taken place within either the upper or lower target (or in both). However, the tremendous heterogeneity of sedimentary layers leads to a very complex wave propagation and this complexity is not revealed by the visual inspection of the data; what we see is in fact the result of interferences between a multitude of wavefronts.

The procedure which is commonly used for the imaging of surface seismic data is called migration [5, 8, 9, 25]. Its application requires an estimation of the velocity distribution in the subsurface. This procedure is also often applied to image well-seismic data. But such an imaging technique turns out to be inadequate for our data that have shown many downgoing waves illuminating the target whereas migration assumes that a single wavefront illuminates the target. This is the reason why the standard procedure for imaging VSP data prefers to account for the complexity of the illuminating wavefield at the expense of a tremendous simplification of the geometrical aspects of wave propagation (precisely at the expense of a 1D propagation assumption). However, it is often attempted to preprocess the data so as to get closer to the 1D propagation assumption. Should we want to account for the multiple scattering effects ([33]) in wave propagation, we would have to apply a nonlinear imaging technique such as the nonlinear inversion of VSP data proposed in [22] in a 1D framework. Here, the goal is to reconstruct the acoustic impedance profile while accounting for the nonlinear character of the operator that maps the profile and the wavefield. This nonlinear character is far from negligible as soon as the impedance profile shows important heterogeneity thus giving rise to important multiple scattering [16].

[3] We use quotes to highlight that such a description relies on a geometrical optics view of wave propagation. In our paper, we want to promote waveform inversion and therefore we question such a view.

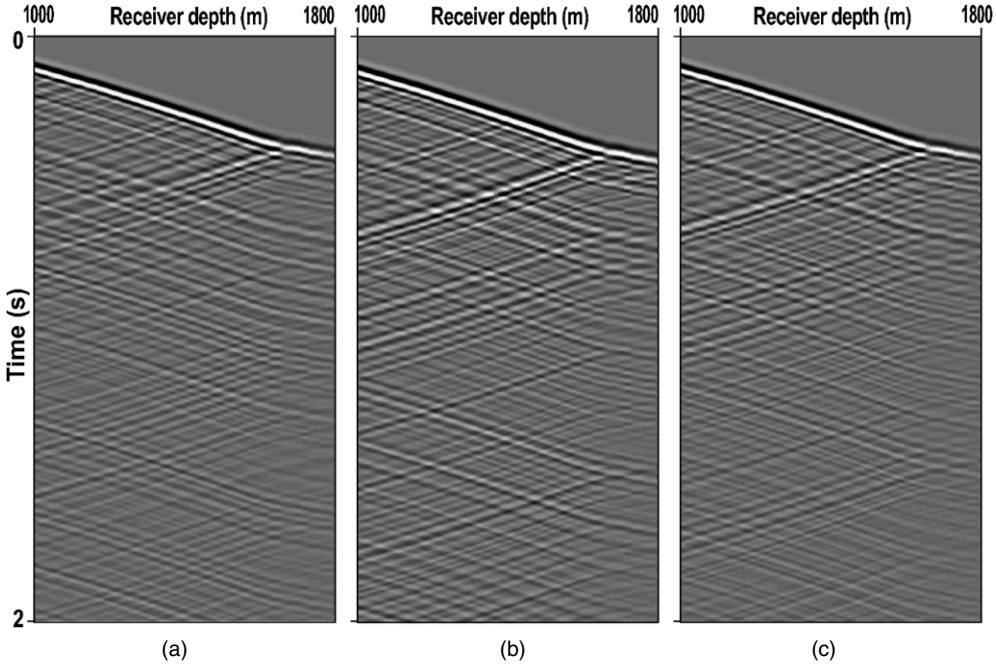

Figure 1. Typical walkaway data for receiver ranging between 1000 and 1800 m. The considered offsets are −360 m (left), 0 m (center) and 360 m (right).

The nonlinear inverse problem in which we reconstruct the impedance profile for known velocity distribution should not be confused with the inverse problem in which we want to reconstruct the propagation velocity distribution from seismic data, a problem extensively studied by many authors (see for instance [32] for an overview). For the 1D nonlinear inversion of VSP data, measuring the depth in terms of vertical traveltime allows us to get rid of the estimation of the velocity distribution (except that we have to express the depths of the different receivers in terms of vertical traveltime, but the conversion is simple since the VSP data allow us to evaluate precisely the difference in traveltimes between successive receivers).

In our paper, we want to set up a 2D extension of the waveform inversion developed in [22]. We want not only to get rid of the unrealistic 1D propagation assumption but also want to estimate the impedance distribution (as opposed to a single profile) in some vicinity of the well and this with a vertical resolution (to be understood as separation power) much better than the one obtained for surface seismic data. As in [22], we will have to reconstruct the illuminating wavefield in the same time as we reconstruct the impedance distribution. Indeed the estimated impedance distribution depends on the illuminating wavefields and, in some situations, in a very sensitive way [15]. The reconstruction of this illuminating wavefield is in fact one of the major difficulties for our extension since, contrary to the 1D propagation framework and the migration context, the geometries of the illuminating wavefronts are unknown (figure 2 (left)). This difficulty will constrain us to only consider media that vary gently along the horizontal direction. In addition, we will restrict ourselves to acoustic wave propagation and assume the velocity distribution $c(x, z)$ to be known.

In section 2, we introduce, in a rather general context, the formalism for the inversion of walkaway data and study how to struggle against the indetermination inherent in this

inverse problem. In section 3, we present the numerical method used for the solution of the inverse problem. Compared with the 1D inverse problem we are faced to two important differences: the necessity to introduce lateral boundaries in the propagation domain and the tremendous inflation of unknowns. This inflation will lead us to be very careful for the numerical implementation. In section 4, we evaluate, by means of tests on synthetic data, the performance of our method with special emphasis on the accessible vertical resolution.

2. The 2D extension of the 1D problem

We thus consider walkaway data. Different datasets are acquired by moving the source at different offsets from the well, supposed here to be vertical with horizontal location $x = 0$. These distances will be called offsets throughout the paper. One dataset is thus associated with a specific offset. We denote by S the number of such datasets. The receivers, located at depths denoted by $z_1, \dots, z_Q$, record, as a function of time t, the particle displacement velocity vector. In a 2D context, this vector is characterized by its horizontal and vertical components. Our data are thus composed of the measurements $D^{s,q}(t)$ associated with receiver q and dataset (or offset) number s, for $t \in [0, T]$ where T is the recording duration. We denote by D the vector in $(L^2(]0, T[))^{2\times S\times Q}$ composed of such measurements. As in [22], we restrict ourselves to the imaging of the target and thus want to get rid of the overburden whose base is at depth z_1. This leads us to introduce a boundary condition at this depth. The target can only be imaged down to a depth $z_{\max}$. This depth can be computed given the recording duration T and the propagation velocity distribution.

2.1. The forward problem

We thus consider waves propagating within a domain corresponding to depths larger than z_1 and introduce a boundary condition at this depth. We also have to specify conditions at the other boundaries. Use of absorbing layers at the vicinity of these other boundaries is essential for a realistic modeling. We have chosen the perfectly matched layer (PML) technique in the framework proposed in [6, 7] for the 2D and 3D Maxwell equations. Let

$$\mathcal{O} =]-a, a[\times]z_1, z_{\max}[\tag{1}$$

be the domain of interest. This is the domain where the velocity $c(x, z)$ and impedance $I(x, z)$ distributions are defined. The PML domain with width l is defined as

$$\Omega =]-X, X[\times]z_1, Z[\,, \tag{2}$$

where we have set $X = a + l$ and $Z = z_{\max} + l$. We extend the original velocity and impedance distributions to the PML domain by setting (similar expressions also hold for the velocity distribution):

$$\begin{cases} I(x, z) = I(a, z) & \text{for } x > a, \\ I(x, z) = I(-a, z) & \text{for } x < -a, \\ I(x, z) = I(x, z_1) & \text{for } z > z_1. \end{cases} \tag{3}$$

Within domain Ω, the Euler equation with absorption is written as

$$\begin{cases} \dfrac{I}{c}\partial_t u_x - \partial_x(p_x + p_z) + \eta_x \dfrac{I}{c} u_x = 0, \\ \dfrac{I}{c}\partial_t u_z - \partial_z(p_x + p_z) + \eta_z \dfrac{I}{c} u_z = 0, \\ \dfrac{1}{Ic}\partial_t p_x - \partial_x u_x + \dfrac{1}{Ic}\eta_x p_x = 0, \\ \dfrac{1}{Ic}\partial_t p_z - \partial_z u_z + \dfrac{1}{Ic}\eta_z p_z = 0, \end{cases} \tag{4}$$

with zero initial conditions and the following boundary conditions:

$$\begin{cases} (p_x + p_z)(x, z_1, t) = h(x, t), \\ (p_x + p_z)(x, Z, t) = 0, \\ (p_x + p_z)(-X, z, t) = 0, \\ (p_x + p_z)(X, z, t) = 0. \end{cases} \tag{5}$$

The principle of the PML technique consists in splitting the pressure function $p(x, z, t)$ into two subfunctions $p_x(x, z, t)$ and $p_z(x, z, t)$, so as to obtain four equations involving spatial derivatives along a single direction. Then, an absorption term associated with this direction is introduced ($\eta_x(x)$ or $\eta_z(z)$). These absorption terms vanish within the domain of interest $\mathcal{O}$ and grow according to a polynomial function within the PML layer. They are expressed as

$$\eta_x(x) = \begin{cases} 0 & \text{if } x \in [-a, a] \\ -\dfrac{\eta_{0x}(x+a)^3}{l^3} & \text{if } x \in [-a-l, -a] \\ \dfrac{\eta_{0x}(x-a)^3}{l^3} & \text{if } x \in [a, a+l] \end{cases} \tag{6}$$

$$\eta_z(z) = \begin{cases} 0 & \text{if } z \in [z_1, z_{\max}] \\ \dfrac{\eta_{0z}(z-z_{\max})^3}{l^3} & \text{if } z \in [z_{\max}, z_{\max}+l] \end{cases} \tag{7}$$

where η_{0x} and η_{0z} are constants that have been set up in an empiric way following the numerical experiments given in [20]. Thus, by adding the two equations in p_x and p_z, we retrieve the standard acoustic wave equation (written in the form of Euler equations), within the domain of interest $\mathcal{O}$. In addition, Dirichlet homogeneous boundary conditions for the pressure are specified at the boundaries of domain Ω (except for the part corresponding to z_1).

The above equations are written for a generic boundary condition $h(x, t)$ at $z = z_1$. However, we will consider different displacement velocity wavefields $u^s(x, z, t)$ solutions of the above equations with the boundary conditions $h^s(x, t)$. In what follows, these boundary conditions will be referred to as the *pressure conditions*. We call *parameter* a pair composed of an impedance distribution $I(x, z)$ defined on $\mathcal{O}$ and of a sequence of S pressure conditions $h^s(x, t)$, $s = 1, \ldots, S$. We denote by H the vector whose components are the S pressure conditions. The forward problem thus consists in computing the *seismic response* to such a parameter, given the velocity distribution. Precisely, the seismic response is described by the synthetic data $U^{s,q}(t)$ for $s = 1, \ldots, S$ and $q = 1, \ldots, Q$ that are the displacement velocity vectors (with the components $u_x^{s,q}(t)$ and $u_z^{s,q}(t)$) observed by the receiver at z_q in the dataset number s.

We introduce the forward modeling operator

$$\mathcal{M} : (I, H) \longmapsto U = \left\{U^{s,q}(t), s = 1, \ldots, S; q = 1, \ldots, Q, t \in [0, T]\right\}.$$

Let us point out again that the solution of the forward problem requires the propagation velocity distribution to be known within domain $\mathcal{O}$ (and only within this domain).

2.2. The 2D inverse problem: an ill-posed problem

We aim at retrieving the parameters whose seismic response best matches the recorded data. The inverse problem can thus be formulated as

$$\min_{I,H} J(I, H), \tag{8}$$

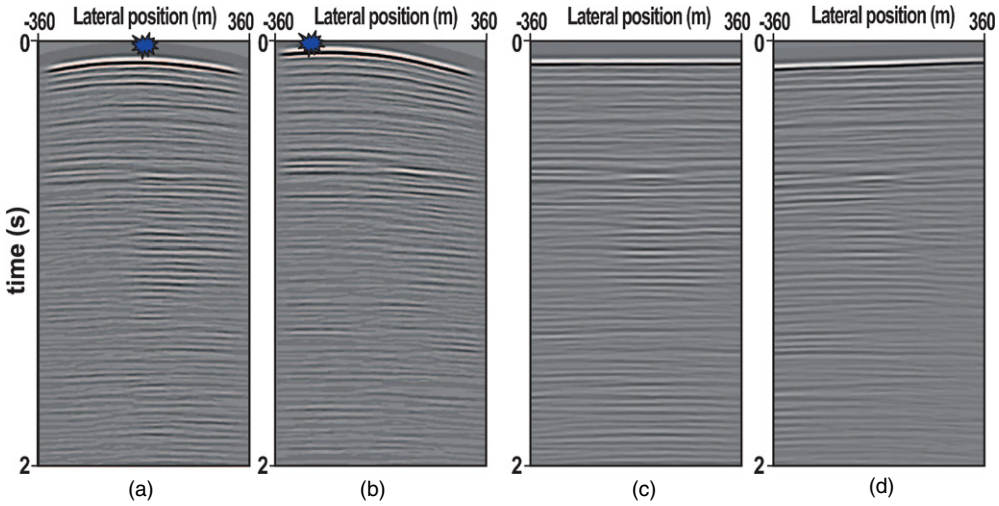

Figure 2. The pressure conditions for a source located at $x = 0$ m (a) and a source located at $x = 250$ m (b). The pressure conditions for slant stacked data (right) for an illumination angle of 0° (c) and 6° (d). Note that in figure (d) the downgoing wavefronts, whose direction is roughly parallel to the direct wave, overwhelm the small amplitude upgoing wavefronts thus making them hardly distinguishable.

where $J(I, H)$ is the so-called *seismic misfit function* defined by

$$J(I, H) = \|D - \mathcal{M}(I, H)\|^2, \tag{9}$$

where $\|\ \|$ is the norm in $(L^2(]0, T[))^{2\times S\times Q}$.
Thus formulated, the inverse problem appears indetermined: obviously we cannot retrieve the pressure conditions associated with each dataset and therefore we cannot image the target. Indeed, we lack information about the geometries of the illuminating wavefronts. To overcome indetermination, we need prior information. This prior information will be accounted for by means of an appropriate regularization.

2.3. The regularized 2D inverse problem

If we consider media that are slowly varying along the x direction, we may integrate in the original objective function (9) a regularization term $R_1(I)$ accounting for the lateral variations of the impedance distribution: $R_1(I) = \|\partial_x I\|^2_{L^2(]-a,a[\times]z_1,z_{\max}[)}$. But this is not enough: we also need prior information about the pressure conditions, an information required, let us point it out again, for the estimation of the impedance distribution.

As for the 1D inverse problem, we want, through the pressure conditions, to account for the complexity of the wavefield that illuminates the target. The big difference, which is in fact a major difficulty for our extension, stems from the fact that the wavefield is only known at $x = 0$ (location of the well) and we can hardly conceive a reliable method to extrapolate the wavefield when we move away from $x = 0$; basically, we lack information about the wavefront geometries. Such wavefronts are shown in figures 2(a) and (b). However, we underline that such information is in practice unavailable: the information displayed in figure 2 is the result of the modeling experiment presented in section 4.

To overcome this difficulty, we apply the walkaway data a Radon transform known as slant stacking in the geophysical literature [13]. We thus obtain well data associated with a

close to plane wave excitation (which would be similar to a pressure condition at $z = 0$). The excitation would be a genuine plane wave if the point sources were infinitely close to each other while covering the whole of the surface and if the propagation velocity at $z = 0$ were constant. Under the latter assumption, the angle θ_0 made by the propagation direction with the vertical is given by $\sin\theta_0 = \frac{c_0\Delta t}{\Delta s}$, where

- c_0 is the velocity at $z = 0$;
- Δs is the distance between two consecutive sources (we thus assume here equidistant sources);
- Δt is the time delay imposed from one source to the next: this delay tunes, as seen in the above formula, the propagation direction of the plane wave.

Thus, we can transform our S original recordings into $\hat{S}$ slant stacked datasets. By doing so, we fix Δt^s, $s = 1, \ldots, \widehat{S}$, to be used. The pressure condition at $z = z_1$ associated with slant stacked data will be denoted by $\hat{h}^s(s,t)$, $s = 1, \ldots, \widehat{S}$. We introduce the vector $\mathcal{H}$ gathering the pressure conditions $\hat{h}^s(s,t)$, $s = 1, \ldots, \widehat{S}$. If we choose $\widehat{S} = S$, we may expect no information loss. The angles θ_0 (which in fact depend on index s) will be called *illumination angles*; they will be chosen depending on the interval $[z_1, z_Q]$ where the receivers are located and on the part of the subsurface where we want to enhance the resolution of the seismic imaging. Here, we just mention this point which would deserve a whole study.

If we deal with a 1D medium, we can have precise information about the propagation angle θ_1 of the plane wave at depth z_1 using the classical formula: $\frac{\sin\theta_1}{c_1} = \frac{\sin\theta_0}{c_0}$ (c_1 is the propagation velocity at z_1).

However, the media we will consider are only close to 1D so that, at depth z_1, the wavefronts are only approximately plane. We thus assume the pressure condition to vary slowly along lines that are orthogonal to vector γ^s (superscript s refers to the dataset obtained by slant stacking with the time delay Δt^s) defined by

$$\gamma^s(x) = \begin{bmatrix} \sin\theta_1^s(x) \\ -\cos\theta_1^s(x) \end{bmatrix} \text{ with } \sin\theta_1^s(x) = \frac{c(x,z_1)\Delta t^s}{\Delta s}.$$

The pressure conditions at $z = z_1$ associated with the slant stacked data can thus be assumed to vary gently along directions that are orthogonal to $\gamma^s(x)$. This leads us to introduce a second regularization term $R_2(\mathcal{H})$ defined by

$$R_2(\mathcal{H}) = \sum_{s=1}^{\widehat{S}} \left\| \gamma^s \cdot \nabla \hat{h}^s \right\|^2,$$

where $\|\cdot\|$ denotes the norm in $L^2(]-a,a[\times]0,T[)$.

We point out that, even in the case of a 1D medium, this regularization term is not exact; indeed, in this case, the pressure wavefield at $z = z_1$ is composed of wavefronts propagating according to angle θ_1 but also of wavefronts propagating according to angle $-\theta_1$, the latter wavefronts corresponding to 'reflections' that reach z_1 under the form of upgoing waves. Use of inhomogeneous transparent boundary conditions[4] such as

$$\frac{\sqrt{I}}{c}(x,z_1,t)\partial_t P(x,z_1,t) + \partial_z\sqrt{I}P(x,z_1,t) = h(x,t) \tag{10}$$

in replacement of the Dirichlet boundary conditions (5) would have been better suited for a correct handling of prior information about the pressure conditions at z_1. However, use of Dirichlet boundary conditions leads to simpler formulas. In addition, our prior information

[4] More sophisticated inhomogeneous transparent boundary conditions can be found in [3]. For the sake of simplicity, we just give the simplest version.

associated with the Dirichlet boundary condition is not unreasonable (figure 2(d)): indeed the amplitude of the downgoing part of the wavefield is much larger than the amplitude of the upgoing contribution.

With these elements, we end up with the regularized formulation of the inverse problem written as

$$\min_{I,\mathcal{H}} J^{\sigma}(I,\mathcal{H}) \tag{11}$$

with

$$J^{\sigma}(I,\mathcal{H}) = J(I,\mathcal{H}) + \frac{\sigma_D^2}{\sigma_I^2} R_1(I) + \frac{\sigma_D^2}{\sigma_{\mathcal{H}}^2} R_2(\mathcal{H}) \tag{12}$$

In equation (12), σ_D, σ_I and $\sigma_{\mathcal{H}}$ represent the uncertainties associated with the seismic data and the pieces of prior information associated with the impedance distribution and pressure conditions, respectively. In fact, these uncertainties can hardly be estimated beforehand. We have thus decided to set the values for $\frac{\sigma_D}{\sigma_I}$ and $\frac{\sigma_D}{\sigma_{\mathcal{H}}}$ using a trial and error approach with the aim of getting a balance between the three components of the regularized objective function at the optimum. Our strategy would be justified if the assumption of an uncorrelated Gaussian additive noise corrupting the different pieces of information would be realistic .... Although more appropriate strategies ([31], p 211) could have been used, ours has the advantage of being straightforward.

We thus have replaced the ill-posed problem (8) by a regularized version that will hopefully be well posed. Precisely, we hope that if the uncertainties σ_D, σ_I and $\sigma_{\mathcal{H}}$ are small enough, then the computed impedance distribution will be 'close' to the distribution we look for. We explain in the next section the reasons on which our expectations rely.

2.4. A justification of the regularization

The understanding of the mechanisms that allows the reconstruction of the impedance distribution from well-seismic data is essential if we want to know to what extent the solution of problem (11) provides an approximation of the impedance distribution we are looking for.

2.4.1. The reconstruction of a 1D medium using plane wave excitations. In the first step, we consider the case of noise free data (the case of noise corrupted data will be skimmed over at the end of this subsection).

We thus consider a 1D medium excited by a plane wave propagating according to the illumination angle θ_0. We denote by $\widetilde{p}(k_x, z, \omega)$, $\widetilde{u}_x(k_x, z, \omega)$ and $\widetilde{u}_z(k_x, z, \omega)$ the 2D Fourier transforms of $p(x, z, t)$, $u_x(x, z, t)$ and $u_z(x, z, t)$, respectively. The Fourier transform $\widetilde{p}(k_x, z, \omega)$ is the solution of the Helmholtz equation:

$$-\frac{c\,\omega^2}{I}\left(\frac{1}{c^2} - \left(\frac{k_x}{\omega}\right)^2\right)\tilde{p} - \partial_z\left(\frac{c}{I}\partial_z\tilde{p}\right) = 0, \tag{13}$$

while $\widetilde{u}_x(k_x, z, \omega)$ and $\widetilde{u}_z(k_x, z, \omega)$ are linked to $\widetilde{p}(k_x, z, \omega)$ by

$$\begin{cases} \omega\tilde{u}_x = -\dfrac{c}{I}k_x\tilde{p} \\ -\mathrm{i}\omega\tilde{u}_z = -\dfrac{c}{I}\partial_z\tilde{p}. \end{cases} \tag{14}$$

The functions $\tilde{u}_x(k_x, z, \omega)$, $\tilde{u}_z(k_x, z, \omega)$ and $\tilde{p}(k_x, z, \omega)$ do not vanish, this is a specificity of the regime associated with a plane wave excitation, only for values of k_x/ω such that $\frac{k_x}{\omega} = \frac{\cos\theta_0}{c_0}$, so that equations (14) can be rewritten as

$$\begin{cases} \tilde{u}_x = -\dfrac{c}{I}\dfrac{\cos\theta_0}{c_0}\tilde{p} \\ -\mathrm{i}\omega\tilde{u}_z = -\dfrac{c}{I}\partial_z\tilde{p}. \end{cases}$$

We introduce the functions $\hat{u}_x(x, z, \omega)$, $\hat{u}_z(x, z, \omega)$ and $\hat{p}(x, z, \omega)$ corresponding to the inverse Fourier transforms of the functions $\tilde{u}_x(k_x, z, \omega)$, $\tilde{u}_z(k_x, z, \omega)$ and $\tilde{p}(k_x, z, \omega)$, respectively. By integrating in the variable x the above equations, we obtain the equations

$$\begin{cases} \hat{u}_x(0, z, \omega) = -\dfrac{c}{I}\dfrac{\cos\theta_0}{c_0}\hat{p}(0, z, \omega) \\ -\mathrm{i}\omega\hat{u}_z(0, z, \omega) = -\dfrac{c}{I}\partial_z\hat{p}(0, z, \omega) \end{cases}$$

and by dividing these two equations we obtain

$$\partial_z\left(\ln\hat{p}(0, z, \omega)\right) = -\mathrm{i}\omega\frac{\cos\theta_0}{c_0}\frac{\hat{u}_z(0, z, \omega)}{\hat{u}_x(0, z, \omega)}.$$

In other words, our measurements yield the function $\hat{p}(0,\ z,\ \omega)$ over the measurement interval, for whatever value of ω and of the illumination angle, but up to a multiplicative constant. Integrating equation (13) over k_x yields

$$-\frac{c\,\omega^2}{I}\left(\frac{1}{c^2} - \left(\frac{\cos\theta_0}{c_0}\right)^2\right)\hat{p}(0,\ z,\ \omega) - \partial_z\left(\frac{c}{I}\partial_z\hat{p}\right)(0,\ z,\ \omega) = 0. \tag{15}$$

This equation allows us to understand the mechanism for the reconstruction of the impedance profile from well-seismic data. Two cases are to be considered: the cases $z \in [z_1, z_Q]$ and $z > z_Q$.

In the former case, for whatever pair (ω, θ_0), I turns out to be the solution of a differential equation (derived from (15)) that can be written on the whole interval $\left[z_1, z_Q\right]$ since we know $\hat{p}(0,\ z,\ \omega)$ on this interval. As a result, the logarithmic derivative of I (we thus assume here I to be differentiable) is determined up to an additive constant or, in other words, I is determined up to a multiplicative constant. Having the information for a single ω and a single illumination angle is enough to reconstruct the impedance profile on this interval (up to a multiplicative constant). In this reasoning, we have assumed a continuum of receivers which is not realistic. However, we easily realize that, if the distance between successive receivers is small, the reconstruction of the impedance profile will be accurate: the resolution (in the sense of separation power) is governed by the vertical sampling between receivers; we can obtain a high vertical resolution even for a small value of ω. In practice, we record $\widehat{p}$ on the interval $[z_1, z_Q]$ for a multitude of values of ω and for several illumination angles: the great amount of redundancy in the information will be very helpful when dealing with noise corrupted data. However, this redundancy is of no help to remove the indetermination associated with the unknown multiplicative constant since there is one such constant for each (ω, θ_0) pair.

The reconstruction of the impedance profile for $z > z_Q$ relies on a fully different mechanism. Here, the understanding of the reconstruction mechanism is to be found in the wave equation formulated as an evolution equation in time (that is the equation satisfied by the inverse Fourier transform of $\widehat{p}$), starting from zero initial conditions. The important point consists in realizing that our recordings provide us with the information associated with both the Dirichlet and Neumann boundary conditions at $z = z_Q$, the latter up to a multiplicative

constant since the value of the impedance at $z = z_Q$ is unknown. The inverse problem which consists in retrieving the impedance profile for $z > z_Q$ from the two boundary conditions has been the object of extensive studies especially the one in [1] and [2]. We underline two important results.

- In the case of noise free data, the uniqueness of the solution (again here up to a multiplicative constant), a result obtained by different authors using different assumptions. We refer to [10] for a review of these results.
- The stability of the solution of the inverse problem when the data are noise corrupted [2]. In fact, this stability strongly depends on the complexity of the profile to be retrieved (this complexity is quantified by the total variation of the profile) and on the wavelet smoothness. In particular, standard seismic wavelets, whose frequency content is negligible below 8 Hz, are not appropriate to recover the low frequency trend of the impedance profile.

The results given above stand for a single illumination angle. In practice, the information for different angles is available; again there is some redundancy in the data. Again this redundancy will be very helpful when dealing with noise corrupted data but will be of no help for removing the indetermination associated with the unknown multiplicative constant and for the recovery of the low frequency trend for depths larger than z_Q.

The results concerning the uniqueness are obtained by exhibiting an analytic expression for the solution of the inverse problem. This analytic expression is of limited interest as soon as we deal with noise corrupted data. In practice, least-squares-based methods such as the one proposed in 1D in [22] (which can be easily generalized to the case of 1D media excited by an oblique plane wave) are better suited, in particular for the separation of signal and noise using the information redundancy. Our motivation for reviewing the theoretical results above was to give elements for apprehending the behavior of inversion software packages whether they are based on an analytic or a least-squares-based approach.

2.4.2. The 2D inversion of walkaway data. This section aims at justifying the efficiency of the regularization we have introduced to produce a reliable estimation of impedance distributions with gentle lateral velocity variations by the inversion of the slant stacked walkaway data.

We start by pointing out that our problem is hampered by the same indetermination ('up to some multiplicative constant') as the problem studied in subsection 2.4.1. Indeed, if we consider the seismic response to parameters $(I, \mathcal{H})$, we do not change this seismic response by multiplying the impedance distribution I by some constant and each of the $\widehat{S}$ pressure conditions by the same constant. This leads us to introduce equivalence classes and to define a representative for each equivalence class by imposing the condition $I(0,\ z_1) = I^{\text{ref}}$, where I^{ref} is a value to be chosen at our convenience.

We now change the definition of pressure conditions: instead of considering the Dirichlet boundary condition (5) at $z = z_1$, we impose the downgoing wavefield at that depth by means of an outgoing boundary condition, that is a condition that makes the reflection coefficient small for a wide interval of incidence angles (see (10) for a simple version).

For fixed positive real λ and μ, we introduce the set $M_{\lambda,\,\mu}$ of equivalence class representatives that satisfy the inequalities

$$\begin{cases} \|\partial_x I\| \leqslant \lambda, \\ \displaystyle\sum_{s=1}^{\widehat{S}} \|\nabla \hat{h}^s \cdot \gamma^s\|^2 \leqslant \mu, \end{cases}$$

where $\|\|$ denotes various L^2 norms (they have been previously specified). We consider the set $E_{\lambda,\mu}$ of parameters in $M_{\lambda,\mu}$ that match our well-seismic data (assumed here to be noise free) and note that this set has a single element if $\lambda = \mu = 0$. Our regularization technique should thus be effective if we admit that the 'size' of the set $E_{\lambda,\mu}$ does not rapidly increase when λ and μ increase. Note that the larger the number of illuminating angles, the slower the increase of the 'size' of this set. From a formal standpoint, we can define the 'size' $S(E_{\lambda,\mu})$ of this set by considering the quantity defined by $\sup_{I,I'} \left\| I - I' \right\|_{L^2(]-a,a[\times]z_1,z_{\max}[}$. With this definition, we focus our attention on the reliability of the reconstructed impedance distributions considering pressure conditions just as intermediate unknowns. Also with this definition we implicitly consider the L^2 norm as an appropriate measure of the accuracy for the reconstruction of the impedance distribution. This is a subjective but not an unreasonable choice. Thus, our expectation concerning the stability of our inverse problem relies on the assumption of a continuous evolution of $S(E_{\lambda,\mu})$ with the (λ, μ) pair. Such an assumption can only make sense if the uniqueness in the recovery of the impedance distribution for known pressure conditions is ensured. As far as we know, there is no proof of such a result (except for the 1D problem) and the nonlinearity involved in the forward map seems to be a major obstacle for the obtention of such a result.

However, in practice, and even though the pressure condition is specified as a Dirichlet boundary condition (and not as the transparent boundary condition (10)), provided that the regularization weights are strong enough, our regularization turns out to be effective for a stable reconstruction of the impedance as we will show in section 4.

3. The numerical method for the solution of the inverse problem

3.1. The numerical solution of the forward problem

The PML technique we have used consists in an extension to acoustic wave propagation in heterogeneous media of the conceptual framework proposed in [6, 7], where the author dealt with electromagnetic wave propagation in a homogeneous medium. Studying the stability of the forward problem formalized by equations (4) and (5) is beyond the scope of our paper. However, the interested reader can refer to the work of [23], this work following the one of [28].

The numerical solution of the forward problem relies on a discretization of equations (4) and (5). The aim of our technique is the improvement of the seismic resolution: a fine discretization of domain Ω is then required. Using a high-order finite-difference scheme in this context would make the computation heavier without improving the accuracy.

We have used the explicit centered second-order scheme proposed in [34].

3.2. The discrete inverse problem

The discrete inverse problem consists in the minimization of the objective function $J^\sigma(I, \mathcal{H})$ in which the functions have been replaced by discrete functions, namely vectors, the forward map by its discrete version based on Yee's finite-difference scheme and the regularization terms by their discrete version involving a finite-difference approximation of the gradients. In order to alleviate the formulas that will follow, we explain the solution of the inverse problem using the formalism of the original (not discretized) problem.

Let us underline the size of the discrete inverse problem: the example presented in section 4 involves typically a million of unknowns. In addition, the evaluation of the value of the objective function for some parameter requires the solution of $\widehat{S}$ wave equations

discretized using a very fine grid. If we want the computation to be feasible, the solution of the discrete inverse problem thus calls for an efficient numerical implementation on a cluster-based supercomputer.

3.3. An original optimization method

The minimization of a non-quadratic objective function that involves about a million of unknowns is not a simple job!

Great care has to be taken to choose the minimization method. Classically, for huge-size non-quadratic problems, we preferably attempt at using quasi-Newton techniques. Indeed, information about the Hessian of the objective function, which would be very complicated, is not required. Such methods only require the evaluation of the objective function and of its gradient for some parameter.

The adjoint state method [11, 21] allows the computation of the gradient of the objective function $\nabla J^\sigma(I, \mathcal{H})$ at, practically, the same cost as the evaluation of this objective function. Note that the presence of unknowns that are of different physical nature (an impedance distribution and pressure conditions) and that act on the seismic response in quite different ways leads to an ill-conditioned optimization problem. Practically, our attempts [24] for a straightforward minimization of the objective function using a quasi-Newton algorithm were led to failure (negligible decrease of the objective function).

3.3.1. An interlocked optimization technique. To overcome the above-mentioned difficulty, we have set up a technique exploiting the quadratic dependence of the objective function on the pressure conditions. Precisely, we are going to use a quasi-Newton algorithm for the minimization of a new objective function that only involves the impedance distribution as unknown. We explain below how we have proceeded.

We introduce a new objective function $\widetilde{J}^\sigma(I)$ defined by

$$\widetilde{J}^\sigma(I) = J^\sigma(I, \overline{\mathcal{H}}(I)), \tag{16}$$

with

$$\begin{cases} \overline{\mathcal{H}}(I) = \arg\min_{\mathcal{H}} J^\sigma(I, \mathcal{H}), \\ \mathcal{H} = (\hat{h}^1, \ldots, \hat{h}^{\hat{S}}). \end{cases}$$

We can easily realize that problem (8) is equivalent to problem

$$\min_I \widetilde{J}^\sigma(I). \tag{17}$$

Our idea is to use the quasi-Newton procedure (specifically a l-BFGS algorithm [26, 27]) to minimize $\widetilde{J}^\sigma(I)$). At each iteration, this algorithm asks for the computations of the objective function and its gradient. These computations require the knowledge of the pressure conditions that are the components of $\overline{\mathcal{H}}(I)$. The computation of these pressure conditions is carried out using a conjugate gradient algorithm (specifically the algorithm described in [18]); indeed these pressure conditions are the solution of a quadratic problem.

In addition, the gradient of the new objective function is straightforwardly obtained from the evaluation of $\nabla J^\sigma(I, \mathcal{H})$, that is with no additional complexity compared with a straightforward minimization of the original objective function using a quasi-Newton algorithm.

Indeed, a disturbance δI brought to I gives rise to a disturbance $\delta\overline{\mathcal{H}}$ of $\overline{\mathcal{H}}(I)$ and to a disturbance $\delta\widetilde{J}^\sigma$ of $\widetilde{J}^\sigma(I)$ given by

$$\delta\widetilde{J}^\sigma = \frac{\partial J^\sigma}{\partial I}(I, \overline{\mathcal{H}}(I)) \cdot \delta I + \frac{\partial J^\sigma}{\partial \mathcal{H}}(I, \overline{\mathcal{H}}(I)) \cdot \delta\overline{\mathcal{H}}.$$

Since $\frac{\partial J^\sigma}{\partial \mathcal{H}}(I, \overline{\mathcal{H}}(I)) = 0$, we can avoid the complicated computation of $\delta\overline{\mathcal{H}}$ and the computation of $\nabla\widetilde{J}_\sigma$ just requires the computation of $\frac{\partial J^\sigma}{\partial I}(I, \overline{\mathcal{H}}(I))$. However, the computation of $\overline{\mathcal{H}}(I)$ requires the computation of $\frac{\partial J^\sigma}{\partial \mathcal{H}}(I, \mathcal{H})$. In summary, the kernel is the computation of $\nabla J^\sigma(I, \mathcal{H})$.

3.4. Parallel computation of $\nabla J^\sigma(I, \mathcal{H})$ by the adjoint state technique

The gradients of the regularization terms R_1 and R_2 are straightforwardly obtained. But the gradient of $J(I, \mathcal{H})$ is not trivial. The adjoint state technique yields the value of this gradient at the expense of the solutions of two wave equations: one associated with the forward problem and one associated with the adjoint problem. We thus avoid the computation of the Jacobian matrix which would require tremendous computer resources. This technique has been developed by Lions [21] for the optimal control of systems governed by PDEs and by Chavent [11] for the identification of distributed systems. A review of its applications to inverse problems in geophysics is given in [29].

We give below a sketch of the adjoint state method. In order to alleviate the notations, we have removed the superscript s (we have done so previously when explicating the forward problem) but the calculations have to be applied for each of the $\widehat{S}$ values of s. We thus consider a vector $\mathcal{H}$ involving a single pressure condition h: $\mathcal{H} = \{h\}$. Also, we only describe the computation of the gradient for an impedance distribution defined on the PML domain Ω and a pressure condition defined on $[-X, X]$; from this computation, we easily obtain the gradient with respect to the functions associated with the original domain $\mathcal{O}$ by the application of the chain rule.

3.4.1. The adjoint state. There are eight dual variables corresponding to the system (4)–(5) for the pressure condition $h(x, t)$. The first four are the functions of (x, z, t) while the others are traces on hyperplanes that can be expressed as the functions of the first four. Let λ_i, $i = 1, 4$, be the first four functions. They are the solutions of the wave equations:

$$\begin{cases} \dfrac{I}{c}\partial_t\lambda_1 - \partial_x\lambda_3 - \eta_x\dfrac{I}{c}\lambda_1 = f_x(x, z, t) \\ \dfrac{I}{c}\partial_t\lambda_2 - \partial_z\lambda_4 - \eta_z\dfrac{I}{c}\lambda_2 = f_z(x, z, t) \\ \dfrac{1}{Ic}\partial_t\lambda_3 - \partial_x\lambda_1 - \eta_x\dfrac{1}{Ic}\lambda_3 = 0 \\ \dfrac{1}{Ic}\partial_t\lambda_4 - \partial_z\lambda_2 - \eta_z\dfrac{1}{Ic}\lambda_4 = 0, \end{cases} \tag{18}$$

to be integrated backward in time from zero final conditions $t = T$, with the boundary conditions

$$\begin{cases} \lambda_3(-X, z, t) = 0, \quad \lambda_3(X, z, t) = 0 \\ \lambda_4(x, z_1, t) = 0, \quad \lambda_4(x, Z, t) = 0 \end{cases} \tag{19}$$

and with

$$\begin{cases} f_x(x, z, t) = 2\displaystyle\sum_{q=1}^{Q}\left(u_x(0, z_q, t) - D_x^q(t)\right)\delta_{M_q}(x, z) \\ f_z(x, z, t) = 2\displaystyle\sum_{q=1}^{Q}\left(u_z(0, z_q, t) - D_z^q(t)\right)\delta_{M_q}(x, z). \end{cases} \tag{20}$$

In the above equations, $\delta_{M_q}(x, z)$ is the Diract delta function centered at the coordinates $(0, z_q)$ of the receiver q.

According to the geophysicists' vocabulary, the computation of the adjoint state amounts to backpropagate the seismic residuals associated with the different receivers.

3.4.2. Computation of the gradient. Once the adjoint state has been computed, it is easy to obtain the gradient of the misfit function. The formulas (see [24]) that give $\nabla J^\sigma(I, \mathcal{H})$ are, for the $\left(u_x^s, u_z^s, p_x^s, p_z^s\right)$ solution of (4), (5) for a pressure condition $h^s(x, t)$, and $\left(\lambda_1^s, \lambda_2^s, \lambda_3^s, \lambda_4^s\right)$ the corresponding adjoint state,

$$\partial_I J(I, \{\hat{h}^1, \ldots, \hat{h}^{\hat{S}}\}) = \sum_{s=1}^{\hat{S}} \int_0^T \lambda_1^s \frac{1}{I} \partial_x \left(p_x^s + p_z^s\right) + \lambda_2^s \frac{1}{I} \partial_z \left(p_x^s + p_z^s\right) \mathrm{d}t$$
$$- \sum_{s=1}^{\hat{S}} \int_0^T \lambda_3^s \frac{1}{I} \partial_x u_x^s + \lambda_4^s \frac{1}{I} \partial_z u_z^s \, \mathrm{d}t. \tag{21}$$

On the other hand,

$$\partial_{h^s} J(I, \{\hat{h}^1, \ldots, \hat{h}^{\hat{S}}\}) = -\frac{I}{c} \lambda_2^s(x, z_1, t). \tag{22}$$

We refer to [24] for the expression of the gradient of the discrete objective function.

3.4.3. Parallelization of the computations using domain decomposition. Use of parallel computing is essential if we want to avoid an unacceptable elapsed time. The most computation intensive part of the algorithm consists in the solution of the forward and adjoint equations. Here, domain decomposition comes in. Domain Ω is subdivided into P subdomains Ω_i. A dedicated processor is in charge of the computations for the solution of the forward and adjoint equations in domain Ω_i. Of course, if we leave aside the communications between processors, the smaller the size of the subdomain, the smaller the computation time is. The parallelization technique is the same for the forward and adjoint equations: we describe the technique for the forward equations only.

Processor i is in charge of the solution of equations (4) within subdomain Ω_i. The only noticeable change lies in the boundary condition that $p_x + p_z$ has to meet. Indeed, domain decomposition introduces fictitious boundaries within domain Ω. These fictitious boundaries separate adjacent subdomains. To specify the boundary condition at these fictitious boundaries, we have to make each subdomain partly cover its neighbors. If we accept the idea of exchanging information between subdomains at each time step, the domain decomposition technique becomes straightforward for Yee's scheme. The covering is composed of two rows (or two columns), the points constituting the boundary of one subdomain corresponding to the row (or column) of the nearest interior points of the neighbor subdomain.

This domain decomposition method allows an efficient parallelization of the computations for solving the forward and adjoint equations. Communication is required at each time step but the volume of information to be exchanged at a specific time step is small since only data within the coverings are to be exchanged, so that we end up with a good 'scaling'. Our numerical implementation allows an effective solution of large size 2D problems as shown in the next section.

In the case of small subdomains, the memory of the computation cores may allow the storage of the solution of the forward problem. So that the computation of the gradient (especially computation of the integral in (21)) is straightforward, there is no need of recomputing this solution by a backward integration in time (synchronous with the solution

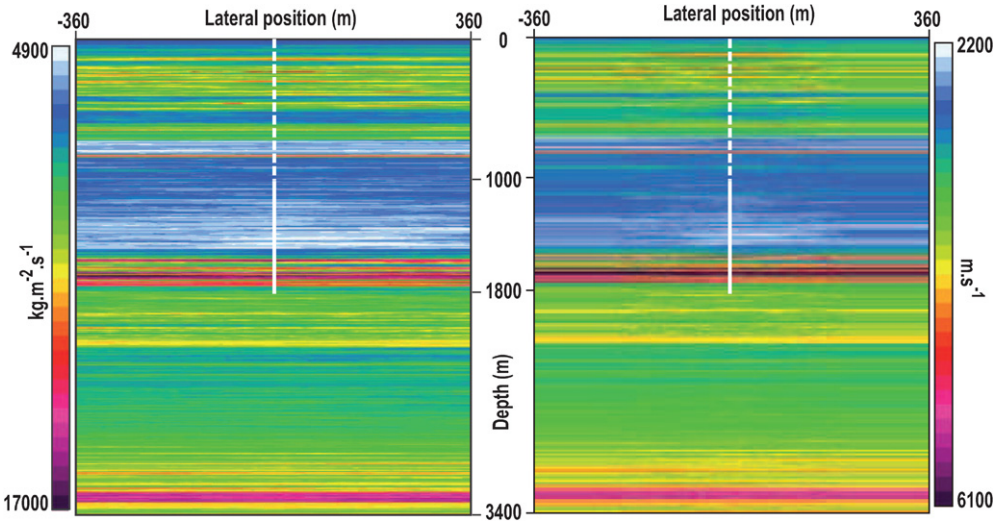

Figure 3. The exact impedance (left) and velocity (right) models. The well is displayed (dashed line) along with the interval where receivers have been placed (solid line).

of the adjoint equation) of the wave equation as it is usually done. This saves some 30% of the computing time required for the computation of the gradient.

4. Evaluation of the method

The goal of this section is to evaluate the performance of the method using synthetic data. Special attention will be paid to the resolution obtained.

4.1. The test model and data

The subsurface model we have used (impedance and velocity distribution) is displayed in figure 3. It is defined over a 720 m wide and 3400 m deep domain. The model shows gentle lateral variations. For the synthetic data generation, 100 receivers were placed in a vertical well (at $x = 0$) every 8 m from $z = 1000$ m. The seismic sources were shot at the surface from $x = -360$ m to $x = 360$ m every 24 m. The seismic wavelet is a Ricker function with 25 Hz central frequency (figure 4). We then applied the original data (figure 1) a Radon transform so as to obtain data corresponding to a pseudo-plane wave excitation for illumination angles ranging, at $z = 0$, between -6° and 6° with a 1° sampling. The transformed data are displayed in figure 5.

Our aim is to invert for the impedance distribution given the 13 transformed datasets and the velocity distribution, for $z \geqslant z_1 = 1000$ m. In order to preserve seismic resolution, we keep the fine spatial discretization, namely $\Delta x = 12$ m and $\Delta z = 8$ m, used for modeling the seismic data.

In all the experiments presented below, the optimization algorithm is initiated using a constant impedance distribution and zero pressure conditions (except for the situation where these pressure conditions are known).

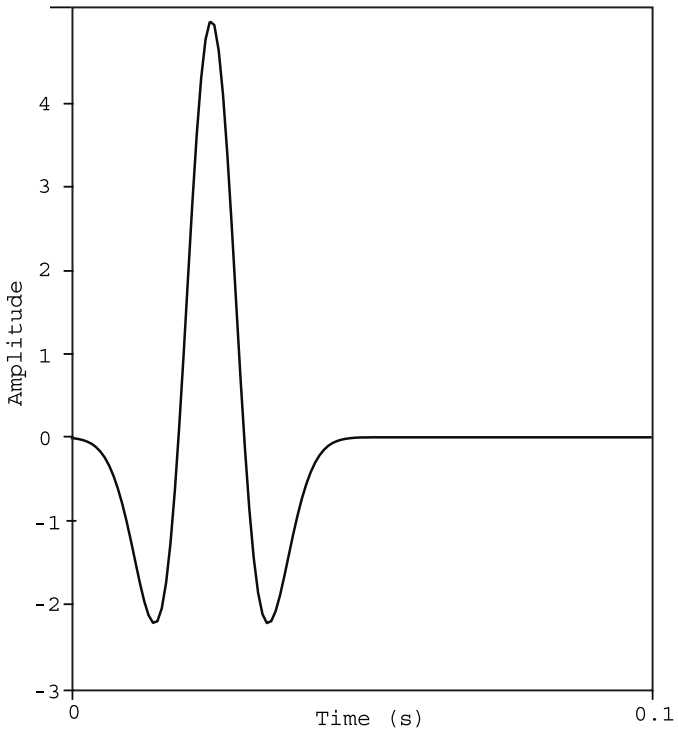

Figure 4. The seismic wavelet used for our experiments.

4.2. Inversion with known pressure conditions

We start with an inversion with known pressure conditions; this experiment will serve as a reference for the test of our method in which the pressure conditions also have to be retrieved. This reference is important since we deal here with 2D nonlinear inversion, a topic in which we lack theoretical background.

To carry out this artificial experiment, we saved the pressure conditions at depth z_1 (top of the overburden) when modeling the wavefields generated by the shots. We then added these different recordings after the application of a shot-dependent delay so as to simulate the slant stacking procedure for the 13 considered illumination angles. Some of the so-obtained pressure conditions are displayed in figures 2(c) and (d).

In this inversion experiment, we want to make sure that the optimization algorithm yields a model that matches the seismic data and, in case of success, to evaluate to what extent the inversion allows a recovery of the impedance distribution. It should be noted that

- we do not regularize here; in other words, we set $\sigma_D/\sigma_I = \sigma_D/\sigma_{\mathcal{H}} = 0$;
- we skip the inner loop of our interlocked optimization technique (indeed there is no need for computing the pressure conditions since they are known).

The l-BFGS algorithm makes the seismic misfit function decrease by six orders of magnitude after 5000 iterations and by more than four orders of magnitude after 1000 iterations. The computed impedance distribution matches the exact impedance distribution pretty well in the upper target (figure 6 (right)). However, our inversion result has no quantitative value in the lower target. This is not surprising in view of section 2.4: the low frequency trend of the impedance distribution cannot be retrieved in the lower target. Although the reconstructed

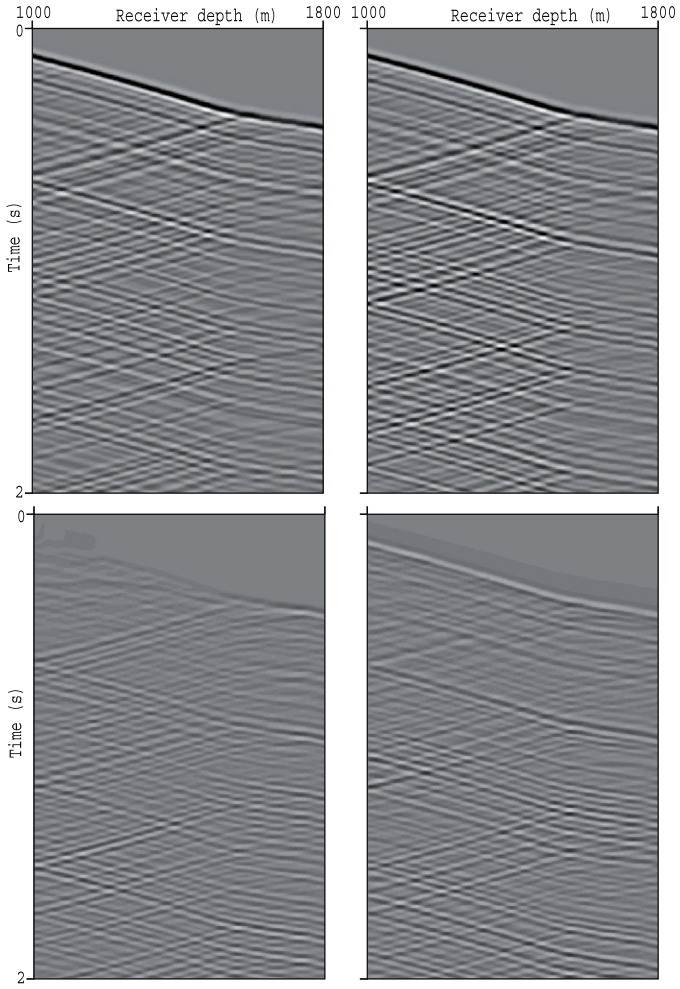

Figure 5. The slant-stacked walkaway data (the vertical component at the top and the horizontal component at the bottom) for illumination angles 0° (left) and +6° (right). The gain has been multiplied by a factor 10 to display the horizontal component.

impedance distribution has no quantitative value (figure 6(left)), we can however recognize some features present in the exact impedance distribution. We will analyze in greater detail the reconstruction of the impedance in the lower target in the next subsection. But let us highlight here the difference between our nonlinear inversion (figure 6(left)) and a migration-like imaging as displayed in figure 7. Leaving aside the quantitative aspects of the reconstruction, the vertical resolution is obviously much higher in our nonlinear inversion result than in the migration-like result. The vertical resolution in migration is governed by the frequency bandwidth of the source [4] whereas, in our nonlinear inversion, it is governed by the sampling between receivers in the upper target. In the lower target, the resolution accessible by our nonlinear inversion cannot be quantified beforehand. However, the fact that it is much enhanced compared with migration can be understood as a kind of super-resolution phenomenon [17] caused by multiple scattering [14]. We will pay considerable attention to the vertical resolution in the next subsection.

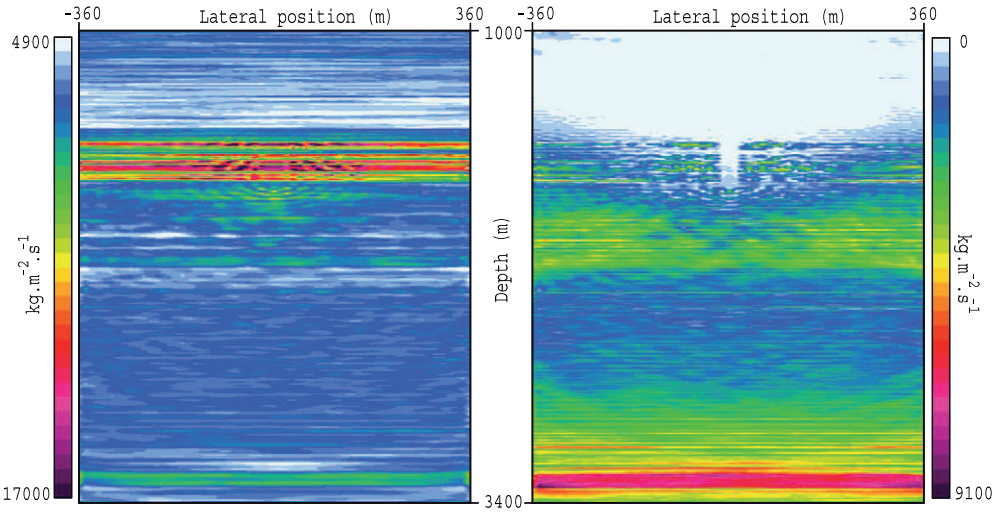

Figure 6. The computed impedance model (left) and the difference with the exact model (right) for known pressure conditions. Note when comparing this result to the exact model in figure 3 that, here, only the estimated part of the subsurface (i.e. below the first receiver at 1000 m) is displayed whereas in figure 3, the subsurface is displayed from depth 0.

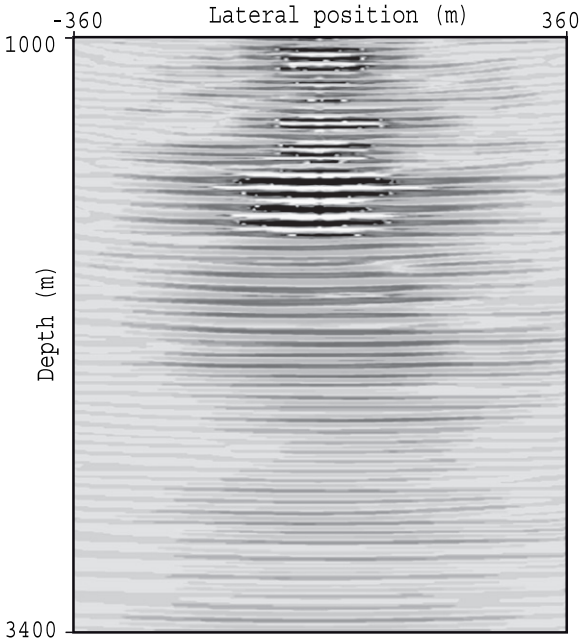

Figure 7. A migration-like imaging. More precisely we display the gradient of the seismic misfit function for a homogeneous reference impedance model. Positive amplitudes are displayed in black whereas negative amplitudes are displayed in white.

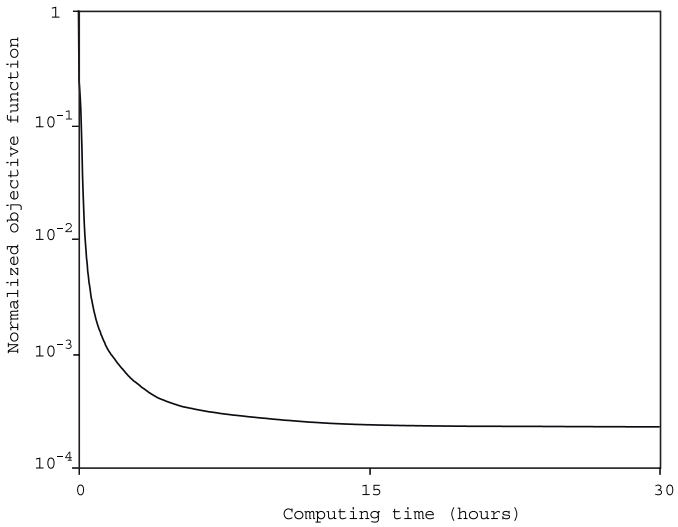

Figure 8. Decrease of the normalized objective function (logarithmic scale) as a function of the computing time.

4.3. Test of our method (unknown pressure conditions)

Encouraged by the results presented in the previous section, we aim here at recovering simultaneously the impedance distribution and the 13 pressure conditions from the well data associated with the 13 illumination angles. Let us point out that, although we deal with synthetic data, this experiment is far from an inversion crime, e.g. [30], contrary to the experiment shown in the previous subsection. Of course, the question is to what extent the lack of information about the pressure conditions degrades the reconstruction of the impedance distribution. Special attention will be paid to the resolution of the imaging technique. But, before addressing these important questions, we need to trust the optimization scheme which constructs the result. Figure 8 displays the behavior of the normalized objective function (we normalize using the value of the objective function, which is also here the seismic misfit function, at iteration 0) during the first 1000 l-BFGS iterations (which correspond to some 30 h CPU time using 128 cores of our cluster[5]). We observe a decrease of this normalized objective function by almost four orders of magnitude, which is comparable to the experiment with known pressure conditions although we deal here with conflicting pieces of information. This illustrates the effectiveness of our interlocked optimization scheme.

4.3.1. An analysis of the resolution. The comparison between the computed and exact models (figures 9 and 3, left) shows that the prior information about the impedance distribution has been accounted for while allowing some lateral variations. Figure 9 (right) also displays the differences between these two models: we observe that a reasonable estimation of the impedance distribution is obtained only in the upper target. Again, this is no surprise according to what has been explained in subsection 2.4. Also, by comparison with the experiment with the known pressure condition, we observe negligible degradation even though the reliability of the quantitative estimation degrades a little when we move away from the well.

[5] More specifically, we have used 8 nodes composed of 16 computation cores.

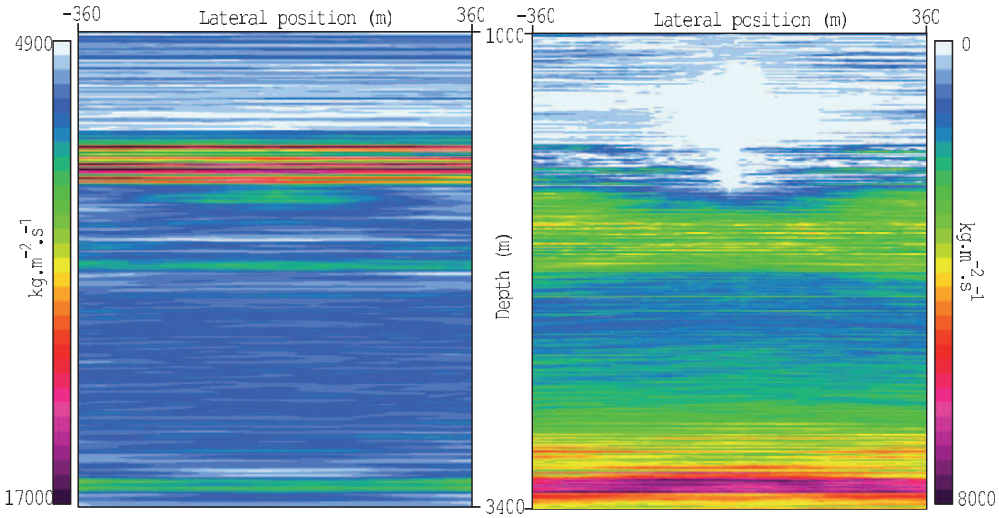

Figure 9. The computed impedance model (left) and the difference with the exact model (right) for unknown pressure conditions. As in figure 6, only the estimated part of the subsurface (i.e. below the first receiver at 1000 m) is displayed.

In order to assess the vertical resolution of the method, we display, in figure 10, the impedance profiles corresponding to different horizontal locations. We observe the very good reconstruction of the impedance in the upper target, even for locations as remote as 200 m. In this upper target, the very fine vertical variations of the impedance are retrieved: the vertical resolution is limited by the sampling interval we have chosen (we should have chosen an even finer interval . . .). In the lower target, even though the low frequency trend is lost, the vertical variations of the impedance are recovered in great detail. In this part, our inversion seems to provide a vertical resolution much better than the one provided by standard seismic imaging techniques such as migration. Again, the enhancement of the vertical resolution can be explained by the multiple illumination of points in the subsurface caused by multiple scattering [14]. This explanation is confirmed by the relative degradation of vertical resolution for large depths. Thus, our efforts for solving a nonlinear inverse problem are rewarded by a significant enhancement of the vertical resolution. However, if the vertical heterogeneity of the medium had been less severe, the enhancement of the vertical resolution would have been less significant.

In order to analyze the lateral resolution, we first display (figure 11) horizontal impedance profiles corresponding to different depths. For $z = 1400$ m or $z = 1600$ m, the profiles show how good the estimation is. However, the accuracy of the quantitative estimation degrades as we move away from the well. At $z = 2000$ m, in spite of the loss of the low frequency trend, we still recover, although with poor resolution, the horizontal variations. However, the horizontal resolution degrades very fast for larger depths. In the transition zone, that is in the depth range at the vicinity of the boundary between the upper and lower targets ($z = 1800$ m), the quality of the quantitative estimation degrades very quickly as we move away from the well. This is not that surprising since this zone corresponds to the transition between two different reconstruction mechanisms as explained in subsection 2.4.1 for 1D models.

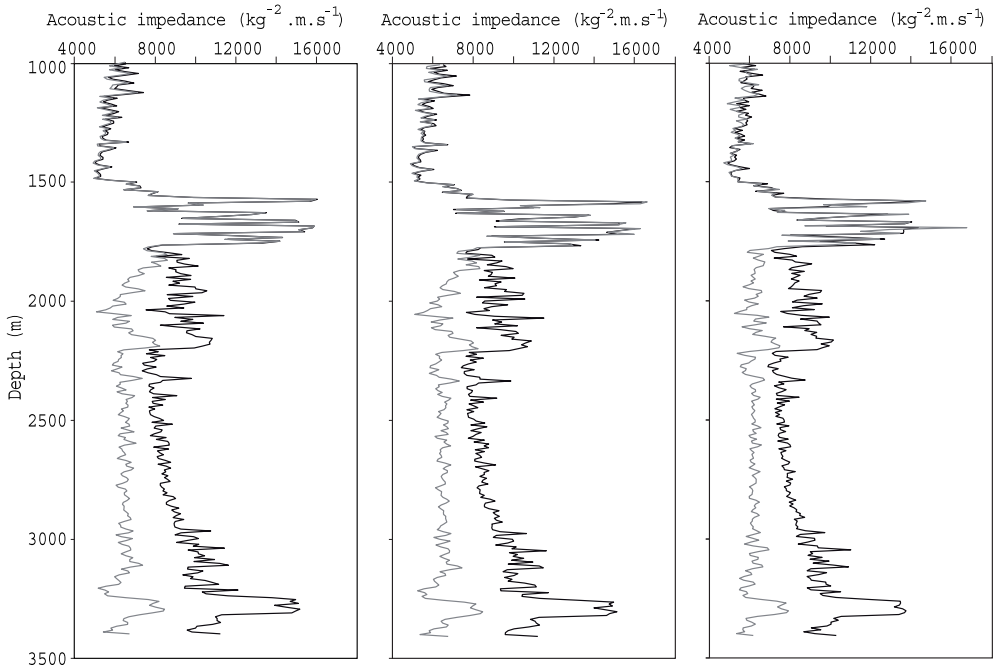

Figure 10. Vertical impedance profiles for unknown pressure conditions: $x = 0$ m (left), $x = 100$ m (center), $x = 200$ m (right). Exact and reconstructed profiles are displayed in black and gray, respectively.

4.4. Sensitivity to the velocity model

Practical use of our method requires the estimation of the distribution velocity. A traveltime analysis of the seismic events seen on the recorded data allow an accurate estimation of the velocity profile at the well location within the depth interval covered by receivers. However, even though the traveltime inversion of well-seismic data ([12]) can be of some help, we cannot expect a very accurate estimation of the velocity at some hundred meters of the well. In this section, we address the problem of the sensitivity of the inversion result to an inaccurate velocity model.

To this aim, we use a 1D velocity model such that the velocity profile is the same as the one of the exact model at the well within the interval covered by receivers and is set to a constant below (the constant is the velocity at the deeper receiver). The objective function is decreased by almost 4 magnitudes: we match the data almost as well as when using the exact velocity model. The recovered impedance distribution is displayed in figure 12 and some vertical profiles are shown in figure 13. The reconstructed impedance is still quite good. However, we observe some degradation in the resolution along with, in the lower target, some depth-varying shifts between the exact and reconstructed model. Those shifts are the consequences of an erroneous time to depth conversion caused by the very erroneous constant velocity model used in the lower target and the fit between the profiles would have appeared much better if they had been displayed in terms of a depth measured in traveltime.

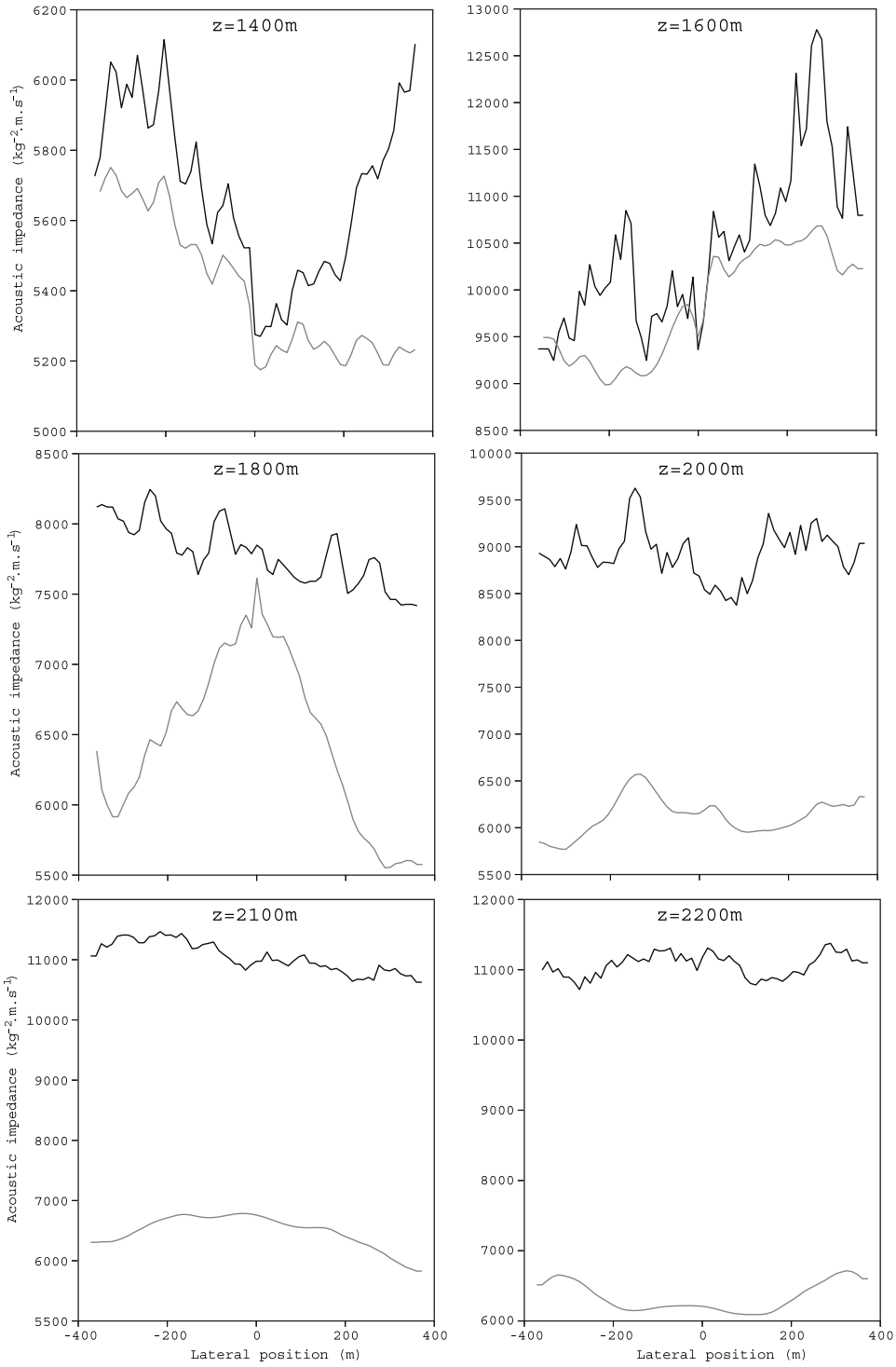

Figure 11. Horizontal impedance profiles with unknown pressure conditions for different depths. Exact and reconstructed profiles are displayed in black and gray, respectively.

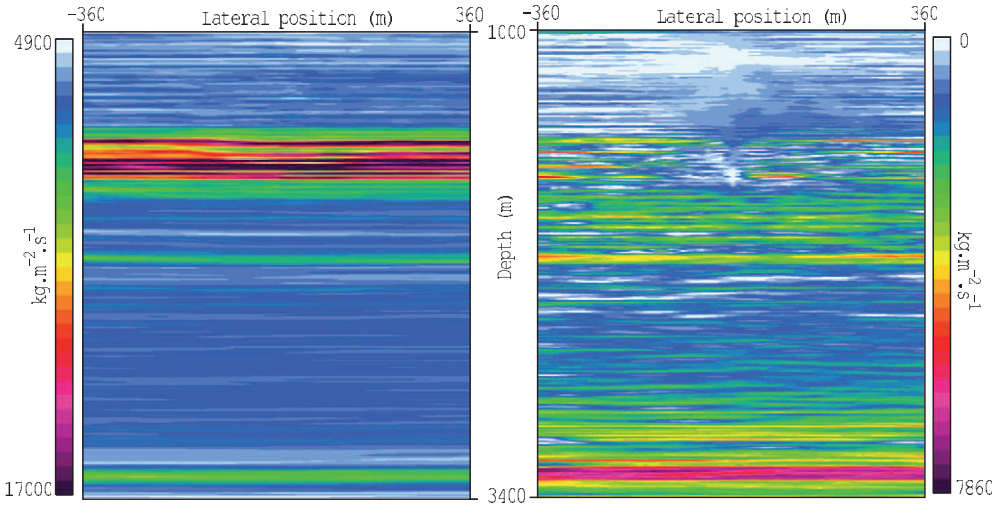

Figure 12. The computed impedance model (left) and the difference with the exact model (right) for an erroneous velocity model.

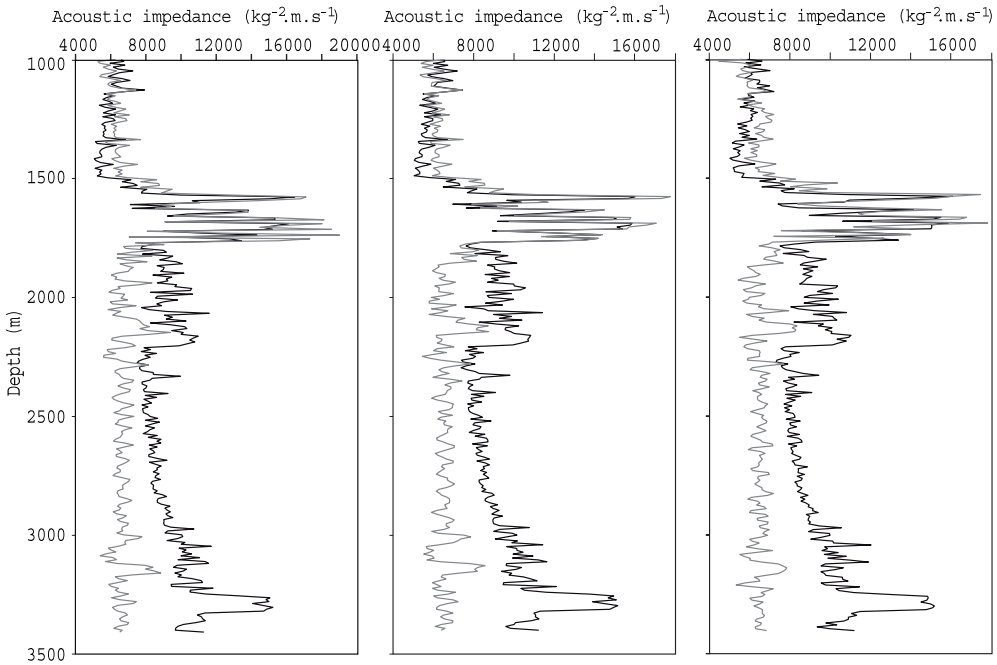

Figure 13. Vertical impedance profiles for an erroneous velocity model. Exact and reconstructed profiles are displayed in black and gray, respectively.

5. Conclusion and perspectives

In the framework of 2D acoustic wave propagation, we have proposed a method that attempts at recovering the impedance distribution from walkaway well-seismic data given the velocity distribution. This method is based on a nonlinear waveform inversion. From a physical standpoint, this amounts to account for the multiple scattering caused by the heterogeneity of sedimentary terrains.

Basically, our method is an extension of a classical least-squares-based 1D inversion of vertical seismic profiles. Our 2D extension calls for an appropriate regularization of the seismic misfit function. This regularization only makes sense in the case of media with gentle lateral variations and is based on a preprocessing that transforms the recorded data into the seismic responses to quasi-plane waves excitations associated with different illumination angles. Our method gets rid of the overburden, that is the part of the subsurface above the depth of the shallower receiver, by reconstructing also the boundary conditions at that depth for the different considered illumination angles.

The so-formulated nonlinear inverse problem involves a tremendous number of unknowns (typically one million), an ill-conditioned objective function, and a forward map based on a numerical solution of the wave equation. Our quest for a high resolution imaging technique led us to deal with finely sampled impedance distributions. In this situation, a second-order finite-difference scheme is appropriate for the numerical solution of the forward problem. The minimization of the objective function calls for a dedicated optimization method to overcome the difficulties arising from ill-conditioning. This method is an extension of the l-BFGS optimization technique and makes use of a gradient computed by means of the adjoint state. On a problem of this complexity, a cluster-based supercomputer is essential as well as an implementation based on a domain decomposition. Our implementation, although straightforward, turned out to be very effective.

Our method has been designed as a 2D extension of a well-understood 1D nonlinear inversion. Based on this connection, an analysis of the stability of the impedance reconstruction would be possible if there were some theorem stating, for 2D acoustic wave propagation, the uniqueness of the impedance reconstruction for known boundary condition. As far as we know, such a theorem is not available today. However, if uniqueness cannot be expected, our problem then loses its sense. Thus, we have assumed this uniqueness and, by analogy with the 1D problem, foreseen the impedance reconstruction mechanism. It turns out that, in the upper target (the part corresponding to the depth interval covered by receivers) we obtain a quantitative estimation of the impedance distribution (up to a multiplicative constant). The vertical resolution is governed by the sampling between receivers. In the lower target, we lose the low frequency trend and this, of course, partly hampers the quantitative estimation. Our numerical experiments on synthetic data confirm and complete the above-mentioned results. In particular, some lateral resolution can be expected only in the upper target. Although the reliability of the quantitative estimation and the lateral resolution degrade as we move away from the well, our method yields a valuable result at about 100 m away from the well. In the lower target, the reconstructed impedance, although with poor lateral resolution, is still valuable. The vertical resolution is much better than the one obtained by standard (linear) seismic imaging. This is a consequence of the multiple illumination caused by multiple scattering and of our efforts in accounting for these nonlinear effects. These results were obtained using the exact velocity distribution.

An important question is the sensitivity of the result in the chosen velocity distribution. In practice, the velocity profile can be identified with very good accuracy along the well in the depth interval covered by receivers. The difficulty is the obtention of the velocity

distribution around and below the well. We have carried out an experiment to mimic this situation. The conclusion is that the result is not very sensitive to the unknown features of the velocity distribution as long as we deal with a close to 1D medium and with close to vertical illuminations; we have mainly observed a slight degradation of the resolution.

Of course important work remains to be done. We have already mentioned the problem concerning the uniqueness of the solution of the nonlinear inverse problem in 2D. From a more practical standpoint, a 3D extension, use of the inhomogeneous transparent boundary condition in replacement of our 'pressure conditions' and update of the estimated velocity distribution during the inversion would be very interesting extensions. Also, how to design a seismic acquisition so as to optimize the reliability of the quantitative estimation and the spatial resolution at a given location in the subsurface is a point on which we have very limited answers and which, basically, remains an open question.

References

[1] Bamberger A 1978 Analyse, contrôle et identification de certains systèmes *PhD Thesis* Université de Paris VI, France

[2] Bamberger A, Chavent G and Lailly P 1979 About the stability of the inverse problem in the 1D wave equation: application to the interpretation of seismic profiles *J. Appl. Math. Optim.* **5** 1–47

[3] Bamberger A, Engquist B, Halpern L and Joly P 1988 Higher order paraxial wave equation approximations in heterogeneous media *SIAM J. Appl. Math.* **48** 129–54

[4] Berkhout A J 1985 Seismic resolution *Handbook of Geophysical Exploration* (London: Geophysical Press)

[5] Berkhout A J 1985 *Seismic Migration* (Amsterdam: Elsevier)

[6] Bérenger J P 1994 A perfectly matched layer for the absorption of electromagnetic waves *J. Comput. Phys.* **114** 185–200

[7] Bérenger J P 1996 Three-dimensional perfectly matched layer for the absorption of electromagnetic waves *J. Comput. Phys.* **127** 363–79

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

[9] Bleistein N 1987 On the imaging of reflectors in the earth *Geophysics* **52** 931–42

[10] Burridge R 1980 The Gelfand–Levitan, the Marchenko and the Gopinath–Sondhi integral equations of inverse scattering theory, regarded in the context of inverse impulse problem *Wave Motion* **2** 305–23

[11] Chavent G 1971 Analyse fonctionnelle et identification de coefficients répartis dans les équations aux dérivées partielles *PhD Thesis* Université de Paris, France

[12] Chiu S K L and Stewart R R 1986 Tomographic determination of three-dimensional seismic velocity structure using well logs, vertical seismic profiles, and surface seismic data *Geophysics* **52** 1085–98

[13] Schultz P S and Claerbout J F 1978 Velocity estimation and downward continuation by wavefront synthesis *Geophysics* **43** 691–714

[14] Delprat-Jannaud F and Lailly P 2004 The insidious effects of fine-scale heterogeneity in reflection seismology *J. Seismic Explor.* **13** 39–84

[15] Delprat-Jannaud F and Lailly P 2005 A fundamental limitation for the reconstruction of impedance profiles from seismic data *Geophysics* **70** R1–14

[16] Delprat-Jannaud F and Lailly P 2008 Wave propagation in heterogeneous media: the effects of fine-scale heterogeneity *Geophysics* **73** T37–49

[17] Fink M 1997 Time-reversed acoustics *Phys. Today* **50** 34–40

[18] Gilbert J Ch 1999 Module N1CG1, http://www-rocq.inria.fr/estime/modulopt/modulopt.html

[19] Hardage A B 1983 Vertical seismic profiling *Handbook of Geophysical Exploration* (London: Geophysical Press)

[20] Hu F Q 1996 On absorbing boundary conditions for linearized Euler equations by a perfectly matched layer *J. Comput. Phys.* **129** 201–19

[21] Lions J L 1971 Optimal control of systems governed by partial differential equations (New York: Springer-Verlag)

[22] Macé D and Lailly P 1986 Solution of the VSP one-dimensional inverse problem *Geophys. Prospect.* **34** 1002–21

[23] Métivier L 2009 Utilisation des équations Euler-PML en milieu hétérogène borné pour la résolution d'un problème inverse en géophysique *ESAIM: Proc.* **27** 156–70

[24] Métivier L 2009 Une méthode d'inversion non linéaire pour l'imagerie sismique haute resolution *PhD Thesis* Université de Paris XIII, France
[25] Miller D, Oristaglio M and Beylkin G 1987 A new slant on seismic imaging: migration and integral geometry *Geophysics* **52** 943–64
[26] Byrd R H, Lu P and Nocedal J 1995 A limited memory algorithm for bound constrained optimization *SIAM J. Sci. Stat. Comput.* **16** 1190–208
[27] Zhu C, Byrd R H and Nocedal J 1997 L-BFGS-B: Algorithm 778: FORTRAN routines for large scale bound constrained optimization *ACM Trans. Math. Softw.* **23** 550–60
[28] Petit-Bergez S 2006 Problèmes faiblement bien posés: discrétisation et applications *PhD Thesis* Université de Paris XIII, France
[29] Plessix R E 2006 A review of the adjoint-state method for computing of a functional with geophysical applications *Geophys. J. Int.* **167** 495–503
[30] Shen P and Symes W W 2008 Automatic velocity analysis via shot profile migration *Geophysics* **73** VE49–60
[31] Tarantola A 1987 *Inverse Problem Theory* (Amsterdam: Elsevier)
[32] Virieux J and Operto S 2009 An overview of full-waveform inversion in exploration geophysics *Geophysics* **74** WCC1–26
[33] Weglein A B, Gasparotto F A, Carvalho P M and Stolt R H 1997 An inverse-scattering series method for attenuating multiples in seismic reflection data *Geophysics* **62** 1875–989
[34] Yee K S 1966 Numerical solution of initial boundary value problems involving Maxwell's equations in istropic media *IEEE Trans. Antennas Propag.* **14** 302–7