

# Velocity model building from seismic reflection data by full-waveform inversion

Romain Brossier<sup>1\*</sup>, Stéphane Operto<sup>2</sup> and Jean Virieux<sup>1</sup>

<sup>1</sup>ISTerre, Univ. de Grenoble Alpes - CNRS, BP 53, 38041 Grenoble Cedex 09, France, and <sup>2</sup>Géoazur, Univ. de Nice Sophia Antipolis - CNRS, 250 rue Albert Einstein, 06560 Valbonne, France

Received April 2013, revision accepted February 2014

## ABSTRACT

Full-waveform inversion is re-emerging as a powerful data-fitting procedure for quantitative seismic imaging of the subsurface from wide-azimuth seismic data. This method is suitable to build high-resolution velocity models provided that the targeted area is sampled by both diving waves and reflected waves. However, the conventional formulation of full-waveform inversion prevents the reconstruction of the small wavenumber components of the velocity model when the subsurface is sampled by reflected waves only. This typically occurs as the depth becomes significant with respect to the length of the receiver array. This study first aims to highlight the limits of the conventional form of full-waveform inversion when applied to seismic reflection data, through a simple canonical example of seismic imaging and to propose a new inversion workflow that overcomes these limitations. The governing idea is to decompose the subsurface model as a background part, which we seek to update and a singular part that corresponds to some prior knowledge of the reflectivity. Forcing this scale uncoupling in the full-waveform inversion formalism brings out the transmitted wavepaths that connect the sources and receivers to the reflectors in the sensitivity kernel of the full-waveform inversion, which is otherwise dominated by the migration impulse responses formed by the correlation of the downgoing direct wavefields coming from the shot and receiver positions. This transmission regime makes full-waveform inversion amenable to the update of the long-to-intermediate wavelengths of the background model from the wide scattering-angle information. However, we show that this prior knowledge of the reflectivity does not prevent the use of a suitable misfit measurement based on cross-correlation, to avoid cycle-skipping issues as well as a suitable inversion domain as the pseudo-depth domain that allows us to preserve the invariant property of the zero-offset time. This latter feature is useful to avoid updating the reflectivity information at each non-linear iteration of the full-waveform inversion, hence considerably reducing the computational cost of the entire workflow. Prior information of the reflectivity in the full-waveform inversion formalism, a robust misfit function that prevents cycle-skipping issues and a suitable inversion domain that preserves the seismic invariant are the three key ingredients that should ensure well-posedness and computational efficiency of full-waveform inversion algorithms for seismic reflection data.

**Key words:** Data processing, Full-waveform inversion, Velocity analysis.

---

\*E-mail: romain.brossier@ujf-grenoble.fr

## INTRODUCTION

Full-waveform inversion (FWI) is a data-fitting procedure that aims to reconstruct high-resolution velocity models (as well as other parameters that govern wave propagation) of the subsurface from the full information contained in seismic data (Lailly 1983; Tarantola 1984; Pratt and Worthington 1990; Virieux and Operto 2009). FWI relies on the diffraction tomography principle (Devaney 1982; Wu and Toksöz 1987), which states that in the framework of the single-scattering approximation, the spatial frequency vector  $\mathbf{k}$  sampled in the image at a diffractor point is related to the local wavelength and the scattering angle  $\theta$  according to:

$$\mathbf{k} = \frac{2\omega}{c_0} \cos\left(\frac{\theta}{2}\right) \mathbf{n}, \quad (1)$$

where  $\omega$  is the angular frequency,  $c_0$  is the local wave speed and  $\mathbf{n}$  is the direction formed by the sum of the two slowness vectors, which can be associated with the rays that connect the source and the receiver to the diffractor point.

When a limited offset range with respect to the depth of investigation is spanned by narrow-azimuth reflection acquisition, only small values of  $\theta$  are sampled by the acquisition if the background model is smooth; namely, if it does not generate back-scattering. In this case, only the high-(spatial) frequency components of the subsurface model can be retrieved according to equation (1) (the spatial frequencies will henceforth be referred to as wavenumbers, allowing a slight abuse of language). This limited bandwidth reconstruction of the wavenumber spectrum gives rise to the well-known scale separation between the macro-velocity model built by traveltimes tomography or migration-velocity analysis and reflectivity built by migration (Jannane *et al.*, 1989). FWI aims to fill this gap, by merging the velocity-model building and the migration tasks into an integrated workflow when broadband sources and wide-aperture acquisition systems are available. Indeed, when these two conditions are satisfied, the redundant influence of the frequency and the scattering angle on the wavenumber coverage allows for continuous sampling of a broadband of wavenumbers from the primary downgoing wavefields, as shown by equation (1). When narrow-azimuth reflection acquisition geometries are considered, only the high wavenumbers are reconstructed from the primary wavefields by conventional FWI, which behaves as a non-linear least-squares migration. This dual behaviour of FWI was illustrated, for example, by an application to real wide-aperture ocean bottom cable data from Valhall by Prieux *et al.* (2011): a broadband velocity model was successfully imaged in the

first 1.5 km in depth above low-velocity gas layers, where both diving waves and reflections sample the medium, while a deep reflector below the reservoir level sampled by pre-critical reflections shows a migration-like reconstruction.

To establish better-posed problems to retrieve the long wavelength of the velocity model from reflected waves, several approaches have been proposed based on image-domain optimization, such as differential semblance optimization (Symes and Carazzone 1991) and wave-equation migration velocity analysis (Sava and Biondi 2004). These approaches attempt to maximize the focusing or coherency of the energy in the image domain, instead of minimizing the data residuals in the data domain. The coherency or focusing can be measured in common-image gathers in angle, offset or in extended-domains that are built during the migration. The extra dimensions in the extended domain are either a time-shift or subsurface offset-shift (Faye and Jeannot 1986; Sava and Fomel 2006). The main drawback of these approaches is the computational burden that is associated with the migration task and the building of the extended common-image gathers at each iteration of the velocity update. This cost currently prevents the direct extension to three dimensions of these approaches when seismic modelling is performed by solving the two-way wave equation. In addition, the generally poor cross-line sampling of conventionally acquired seismic data often precludes the construction of the extended imaging condition in the cross-line direction, due to aliasing effects (Zhou, Howard and Mifflin 2011). Recent applications of such image-domain and mixed image/data-domain approaches have shown promising results (Yang and Sava 2011; Almomin and Biondi 2012; Biondi and Almomin 2012; Sun and Symes 2012), although they remain limited to two-dimensional geometries and rely on some approximations to keep the computer cost affordable. Data-domain approaches have also been proposed to mitigate non-linearities of classical least-squares FWI formulation, such as migration-based traveltimes tomography (MBTT) (Chavent, Clément and Gómez 1994; Xu *et al.*, 2012). This method relies on a difference-based data-fitting approach, although it still requires a migration step at each iteration. The governing idea relies on the explicit decomposition of the subsurface model into a background part, which we seek to update and a singular part, which is assumed to be known. Forcing this decomposition explicitly in the inversion formalism allows the singular part of the subsurface to be processed as primary sources in depth, which brings out the transmission wavepaths from the reflectors to the shots and receivers at the surface in the sensitivity

kernel of the inversion. These transmission wavepaths feed the inversion with wide-scattering angles, which are amenable to the update of the long wavelengths of the subsurface that are located between the buried sources and the surface, just as in cross-hole waveform tomography.

Indeed, in this approach, a new reflectivity must be generated by migration according to the velocity update performed at each non-linear FWI iteration and the two-step workflow is iterated until the migrated image satisfies some focusing criteria.

In this study, we use a simple numerical experiment of seismic imaging to state the problem of velocity model building from short-spread reflection data by FWI. We first highlight the intrinsic limits of the conventional form of FWI in this context. Then, we show how prior knowledge of the reflectivity allows for a better-posed optimization problem. The Appendix complements the numerical experiment with a more formal consideration of the problem statement through an analysis of the sensitivity kernels of frequency-domain FWI and the derivation of the misfit function of the FWI when the uncoupling between the macro-model and the reflectivity is forced in the formalism. In spite of the improvement provided by prior knowledge of the reflectivity, we also show how FWI fails to converge to the correct solution if the reflectivity is not updated at each iteration of the FWI. This prompts us to reformulate the FWI in a suitable domain, in which seismic invariants are preserved, such as zero-offset traveltimes (Landa, Beydoun and Tarantola 1989; Snieder *et al.*, 1989; Plessix *et al.*, 2012). In this case, the migration step at each non-linear FWI iteration can be avoided, hence considerably reducing the computational cost of the entire workflow. We also show that all of these strategies do not prevent the use of suitable misfit functions based on cross-correlations, to avoid cycle-skipping issues (Luo and Schuster 1991; van Leeuwen and Mulder 2010). A suitable functional, an adequate inversion domain and relevant prior knowledge of the reflectivity should be combined and these are the three key features that constitute the proposed workflow of FWI of reflection data.

## THEORY

The classical FWI problem attempts to minimize the differences between observed and computed data, in the least-squares sense. This leads to the minimization of the following misfit function:

$$C_{diff} = \sum_b \sum_t \frac{1}{2} \left( d_{cal}(t, h) - d_{obs}(t, h) \right)^2, \quad (2)$$

where  $d_{cal}(t, h)$  and  $d_{obs}(t, h)$  are the computed and observed data, respectively, at time  $t$  and source-receiver offset  $h$ . An implicit sum over the sources is also contained in equation (2). The derivative of the misfit function with respect to the velocity, i.e., the gradient of the misfit function, gives:

$$G_{diff} = \sum_b \sum_t \frac{\partial d_{cal}(t, h)}{\partial m} \Delta d(t, h), \quad (3)$$

where  $\partial d_{cal}(t, h)/\partial m$  is the Fréchet derivative of the data at time  $t$  and at offset  $h$  and  $\Delta d(t, h)$  is the data misfit at time  $t$  and offset  $h$ . In practice, these derivatives are never explicitly computed and the gradient is directly computed with the adjoint-state method (Plessix 2006).

van Leeuwen and Mulder (2008) introduced a weighted-time cross-correlation misfit function that was defined as:

$$C_{XcorN} = \sum_b \sum_\tau \frac{1}{2} \left( P(\tau) XcorN(\tau, h) \right)^2, \quad (4)$$

where the function  $P(\tau)$  is a weighting/penalization function and the variable  $\tau$  is the time-lag of the normalized cross-correlation function  $XcorN(\tau, h)$ , defined as:

$$XcorN(\tau, h) = \frac{\sum_t d_{cal}(t, h) d_{obs}(t + \tau, h)}{\|d_{cal}(h)\| \|d_{obs}(h)\|}. \quad (5)$$

We note that the proposed misfit function extends the zero-lag-only cross-correlation of Routh *et al.* (2011). Therefore, misfit function (4) can handle time (phase) delay that is larger than half a period without any ambiguity.

Deriving the misfit function with respect to the velocity gives the gradient of the time-normalized cross-correlation misfit function:

$$\begin{aligned} \frac{\partial C_{XcorN}}{\partial m_i} &= \sum_b \sum_t \frac{\partial d_{cal}(t, h)}{\partial m_i} \sum_\tau P(\tau)^2 XcorN(\tau, h) \\ &\quad \times \left( \frac{d_{obs}(\tau + t, h)}{\|d_{cal}(h)\| \|d_{obs}(h)\|} - \frac{XcorN(\tau, h) d_{cal}(t, h)}{\|d_{cal}(h)\|^3 \|d_{obs}(h)\|} \right), \end{aligned} \quad (6)$$

where we define the expression:

$$Xcor(\tau, h) = \sum_t d_{cal}(t, h) d_{obs}(t + \tau, h). \quad (7)$$

Note that this gradient can be computed at the same computational cost as the classical gradient (equation (3)). Only the source of the adjoint-state equation (i.e., the residual term in equation (6)) needs to be modified (Tromp, Tape and Liu, 2005).

In the following experiments, these misfit functions and gradient formulations are used in a two-dimensional time-domain acoustic inversion algorithm, in which the bound

constrained quasi-Newton optimization relies on the L-BFGS-B method for both Hessian estimation and the line-search satisfying Wolfe's conditions (Byrd, Lu and Nocedal, 1995).

## RESULTS

We consider a surface reflection acquisition geometry in a square box of dimension  $500\text{ m} \times 500\text{ m}$ . The true subsurface model is composed of a homogeneous background model with a velocity of  $1500\text{ m/s}$  and a density of  $1500\text{ kg/m}^3$ , to which we have added a density reflector at  $340\text{ m}$  in depth. The density along the reflector is  $2500\text{ kg/m}^3$  (Fig. 1a,b). Twenty-one explosive sources and 101 receivers are evenly distributed all along the surface, with  $25\text{-m}$  and  $5\text{-m}$  spacings, respectively. The twenty-one shot gathers are involved in the inversion during all of the following tests, although we illustrate the data fit with the first shot gather only (the shot position at  $(x = 0, z = 0)$ ). A Ricker wavelet of central frequency  $26\text{ Hz}$  is used to generate a data set for inversion. The first shot gather shows the direct wave and the reflection from the density-contrast reflector (Fig. 1c). The direct wave is muted and only the reflected wave is considered for the inversion (Fig. 1d). In all the following tests, the starting velocity model is homogeneous with an overestimated velocity of  $1800\text{ m/s}$ . During inversion, we only update the velocity either in the depth domain or in the pseudo-depth domain, keeping the density model fixed. We will use two different background density models for FWI, with and without the reflector, to highlight the benefit that is provided by prior knowledge of the reflectivity in the FWI of the reflection data. All of the results are displayed at the 15<sup>th</sup> iteration of the model update or at convergence, if this is reached beforehand.

### 3.1 Inversion with no prior knowledge of the reflectivity

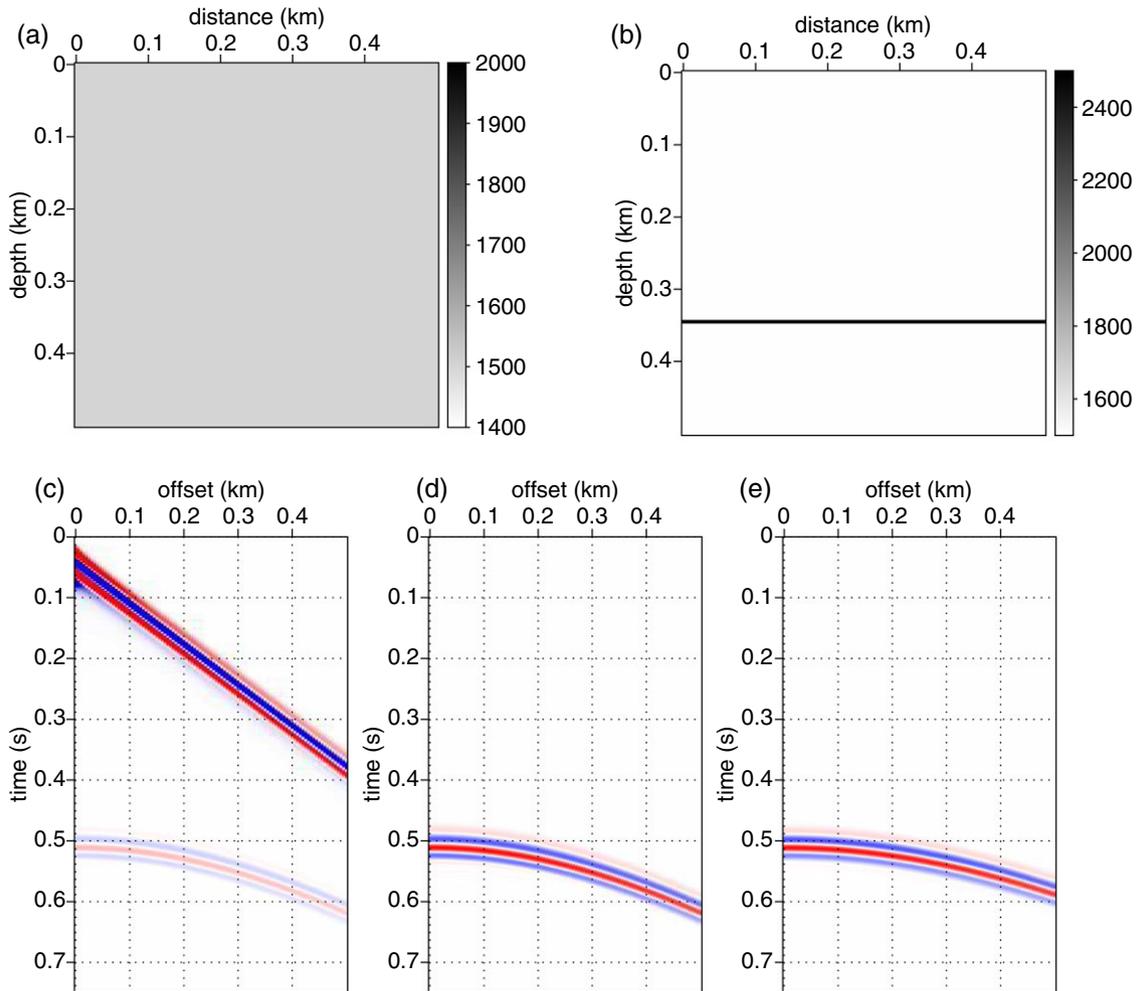
In the first test, the background density model that is used for inversion is homogeneous, which means that no reflection is generated by the starting subsurface model. In such a case, only the difference-based inversion can be used, as the cross-correlation misfit is zero in the starting model when the direct arrival is muted. During the first iteration, the inversion behaves as a least-squares migration and generates a short-scale velocity perturbation that mimics the impedance contrast of the true subsurface model, in an attempt to match the reflected wave (Fig. 2a). However, this velocity perturbation is mispositioned in depth, because the velocity in the starting model is overestimated and the inversion of the short spread

reflection fails to update the long wavelengths of the velocity model, because of the short scattering angles that are spanned by the acquisition (equation (1)). During the next iterations, the FWI probably attempts to update the long wavelengths of the subsurface from the upgoing wavefields back-scattered by the short-scale perturbation generated during the first iteration (the role of the back-scattered waves generated by the reflectivity during FWI of reflection data will be made clear later in the study and in the Appendix). However, the velocity reconstruction becomes trapped into a local minimum after the first iteration, because of cycle-skipping issues on the one hand and velocity-depth ambiguity on the other. The statement that the FWI attempts to update the long-to-intermediate wavelengths during the late iterations is supported by the fact that the FWI clearly attempts to match the moveout of the reflection hyperbola over the iterations, at the expense of the short-offset branch of the reflection (compare the short-offset and long-offset traveltimes of the reflection arrival computed in the true model (Fig. 1d), in the initial model with the correct zero-offset time and the wrong moveout reflection (Fig. 1e) and in the final FWI model (Fig. 2b)).

### 3.2 Inversion with prior knowledge of the reflectivity

In the following tests, the density model contains the reflector information. The correct value of the homogeneous density background is known, as well as the density contrast that defines the reflector. Our aim is to build a starting subsurface model that generates a reflection in the modelled wavefield, such that the reflection honours the zero-offset traveltimes of the observed data. The depth of the reflector in the density model is therefore positioned at an erroneous depth of  $410\text{ m}$ , with the true depth being  $340\text{ m}$  according to the overestimated wave speed in the starting velocity model ( $1800\text{ m/s}$ ). This density model could have been inferred from an amplitude-versus-offset analysis or from a stack section computed in the starting velocity model. The first shot-gather computed in this starting model (Fig. 1e) shows the accurate zero-offset traveltimes of the reflection but the wrong moveout, due to the erroneous velocity above the reflector.

A key change relative to the previous test (Fig. 2a,b) is the presence of the reflector in the starting model, which completely modifies the sensitivity kernel of the waveform inversion at the first iteration (Fig. 3a,b). Without prior knowledge of the reflectivity, the kernel shows a high sensitivity at all of the spatial positions, corresponding to a constructive correlation of the downgoing wavefield emitted by the source and the downgoing back-propagated wavefields emitted by



**Figure 1** (a, b) True velocity (a) and density (b) models. (c) Shot gather computed in the true model for the first shot located at  $(x = 0, z = 0)$ . (d) Same as (c), after the mute of the direct arrival. (e) Same as (d), with the shot gather computed in the subsurface model of the overestimated velocity (1800 m/s) with the density reflector positioned at the correct zero-offset traveltime (hence, at the erroneous depth, according to the velocity error).

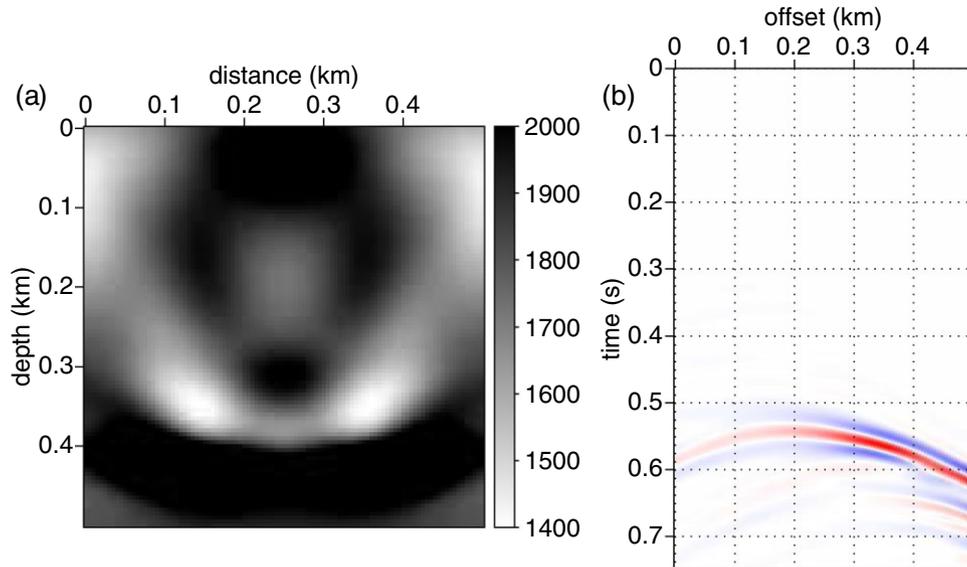
the residuals of the reflected waves at the receiver positions (Fig. 2a). Indeed, these spatial positions are distributed along an isochrone surface in the framework of migration, along which the arrival times of the incident wavefield and of the back-propagated wavefields coincide.

When prior knowledge of the reflectivity is used, the sensitivity of the waveform inversion along the isochrone surface shown in Fig. 3(a) is augmented with two wavepaths that connect the source and the receiver to the reflector (Fig. 2(b)).

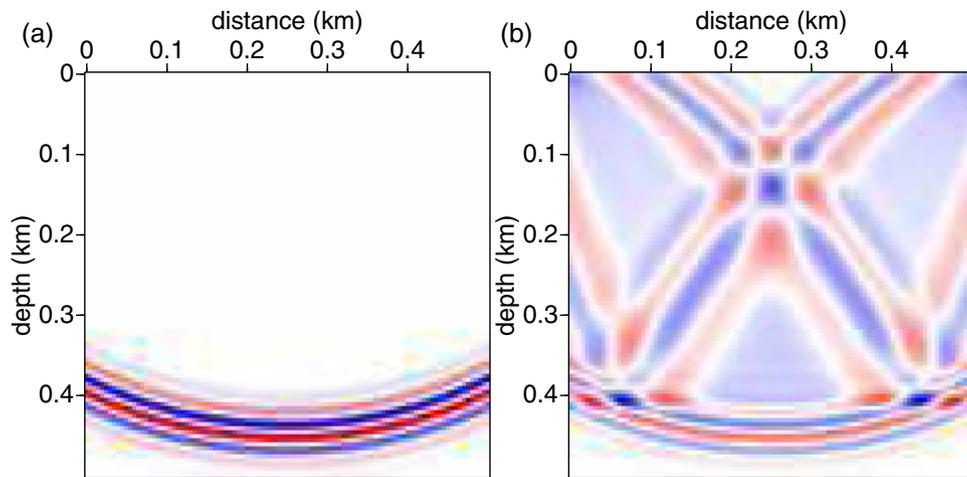
In this setting, the reflectivity can be understood as a back-scattering secondary source in depth that emits upgoing waves towards the surface, where they are recorded by the seismic sources and receivers. The first Fresnel zone of these wavepaths (the central isochrone) describes the for-

ward scattering or transmission regime, which is amenable to the update of the long wavelengths of the subsurface, with a resolution given by the width of these first Fresnel zones. In other words, the prior knowledge of the reflectivity allows for combining migration-like inversion of the reflection data and tomography-like inversion of the transmitted waves using virtual (secondary) sources at the reflector positions.

These concepts are described from a more formal viewpoint in the Appendix, where the gradient of the misfit function for the update of the background model is developed with the adjoint-state method when a reflectivity model is assumed to be known. We show, in particular, how the explicit uncoupling between the background model and the



**Figure 2** (a) Final FWI velocity model inferred from the 1800 m/s starting velocity model, the homogeneous background density model and the data-difference misfit. The 21 shot gathers are involved in the inversion. (b) First shot gather computed in the model in (a) after the muting of the direct arrival. This modelled shot gather can be compared with the true shot gather shown in Fig. 1(d): the inversion remained locked in a local minimum. See text for interpretation.

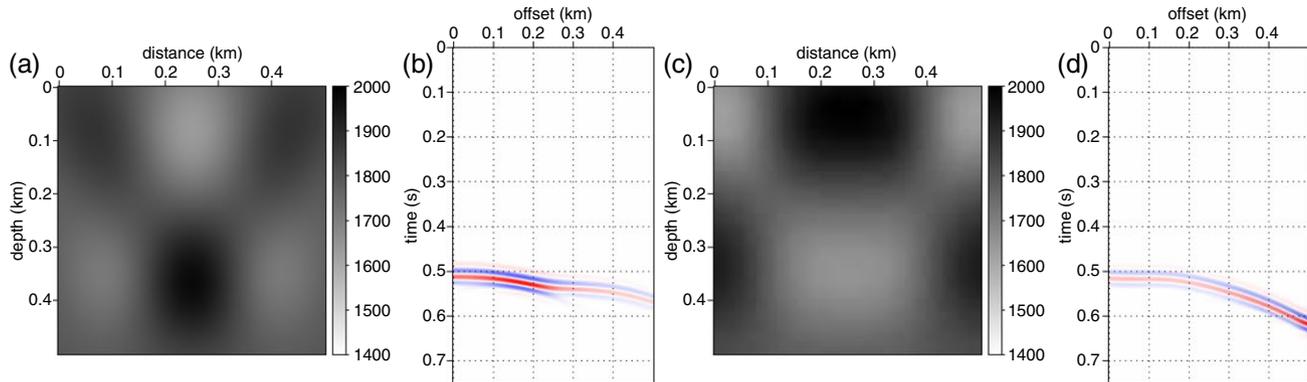


**Figure 3** Sensitivity kernels for the first source ( $x = 0$  m) and the last receiver ( $x = 500$  m) for the homogeneous density (a) and the density model that contains the reflector at 410 m in depth (b).

reflectivity in the FWI formalism allows for the removal of the migration wavepaths from the sensitivity kernel of the FWI, to the benefit of the transmission wavepaths. This is a key issue for realistic applications (i.e., a complex subsurface model with several reflectors), because the footprint of the low-amplitude transmission wavepaths, which are generated by secondary sources at the reflector positions, is often dominated by high-amplitude migration wavepaths, which are

generated by primary sources located at the shot and receiver positions.

Figure 4 shows the results of the inversion using difference-based and correlation-based misfit functions when some prior knowledge of the reflectivity is set in a fixed-density model, such that the zero-offset traveltime is honoured. In this case, FWI failed to converge and stopped after a few iterations. This convergence failure arises because as



**Figure 4** Final FWI velocity models and the corresponding first shot gather after the muting of the direct wave. The density model contains the reflector at 410 m in depth. FWI was performed using the difference misfit (a,b) and time cross-correlation misfit (c,d) in the classical depth domain. Note that in both cases the inversion stops converging after a few iterations.

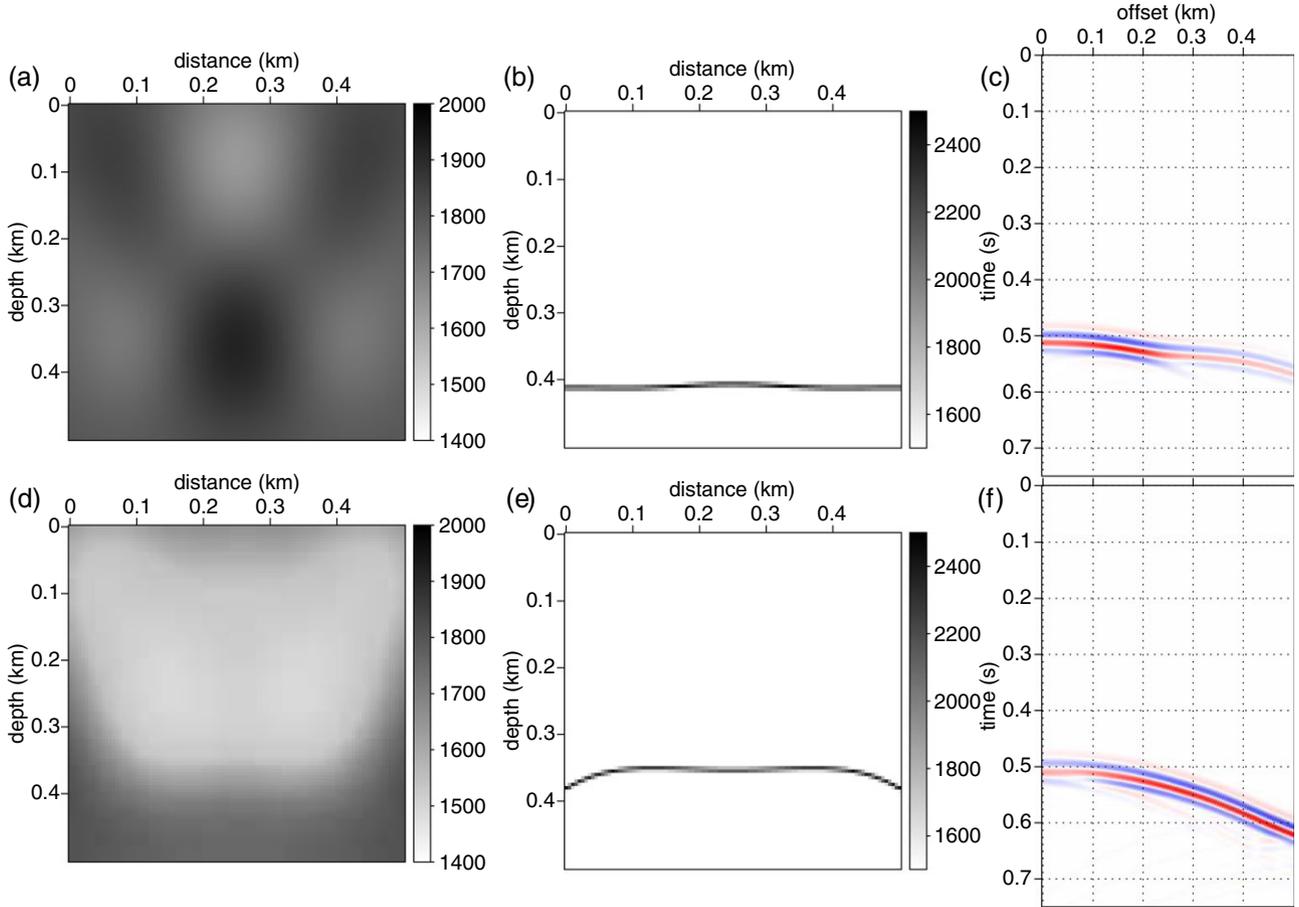
soon as the inversion updates the velocities over the iteration (Fig. 4a,c), the FWI would have needed the necessary degrees of freedom to update the depth of the reflector in the density model accordingly, such that the invariance of the zero-offset traveltimes is preserved. Here, we kept the density model fixed because we want to avoid performing migration at each FWI iteration. This is the main reason for the failure of the convergence in conjunction with the cycle skipping associated with inaccurate wave speeds. It is worth stressing the differences between the modelled data that were generated from the final FWI models obtained when the density model does or does not contain the reflector (Fig. 4b and Fig. 2b, respectively). In the first case, the short-offset data are matched because the inversion is ‘frozen’ by the prior information of the reflectivity in the density model, which honours zero-offset reflection. The match of the short-offset data is preserved at the expense of the match of the long-offset data, which is relatively poor (Fig. 4b). In the second case, the reflectivity was generated close to the first iteration, to match as well as possible the reflection in an average sense. During subsequent iterations, it appears that the inversion was mostly driven by the transmission wavepaths that were generated by the wave speed reflectivity during the first iteration, as suggested by the better match of the long-offset reflection relative to that of the short offset (Fig. 2b).

The cross-correlation misfit function (Fig. 4c,d) leads to a much better kinematic fit and in particular at long offsets, probably because the inversion was not hampered by cycle-skipping issues. However, the FWI failed to generate a satisfying velocity model, as the inversion was steered towards the update of the long velocity wavelengths without the possibility of updating the depth of the reflectivity accordingly.

### 3.3 Depth versus pseudo-time domain for Full-waveform inversion of reflection data

The previous experiment failed to converge to the true velocity model because of the inconsistency between the velocity model and the fixed reflectivity for the matching of the recorded reflected wavefield over the FWI iterations. Recall that the reflector in the background density model was positioned at a depth such that the zero-offset traveltime of the reflection computed in the starting velocity model and in the background density model matched that of the observed reflection computed in the true subsurface model. As long as the velocities of the FWI model are updated, this zero-offset traveltime match is not satisfied anymore if the reflector in the density model is not updated accordingly.

As already mentioned, one remedy would be to estimate the reflectivity from impedance analysis or migration after each iteration of the velocity update (Chavent, Clément and Gómez 1994; Xu *et al.*, 2012). Another solution, which is chosen here to avoid the computational burden that results from migration, is the performing of the inversion in an alternative domain. The vertical traveltime domain, which was also called the pseudo-time by Plessix *et al.* (2012), was proposed by Alkhalifah, Fomel and Biondi (2001) for seismic processing in anisotropic media and Ma and Alkhalifah (2011) formulated a modelling technique directly in this alternative domain. The idea of exploiting the invariant properties of seismic data for velocity model building was first proposed by Landa, Beydoun and Tarantola (1989) and was exploited later by Snieder *et al.* (1989). This pseudo-time domain allows the velocity/depth ambiguity of the classical seismic processing to be better dealt with, by exploiting seismic invariants, such as the zero-offset traveltime. In the pseudo-time domain,



**Figure 5** FWI in the vertical time (pseudo-depth) domain. (a,b) Final FWI velocity (a) and density (b) models obtained with the difference misfit. The FWI was trapped in a local minimum because of the initial inaccurate wave speed, which leads to a mispositioning in depth of the reflector in the density model. (c) First shot gather computed in the final FWI model of (a,b). The FWI failed to match the moveout of the reflection. (d,e) As for (a,b) for the time cross-correlation misfit. The time cross-correlation misfit function allows us to overcome the cycle-skipping issue: the wavespeed was successfully updated, which leads to the correct positioning in depth of the density reflector. (f) First shot gather computed in the final FWI model of (d,e). The reflection is matched over the full range of offsets. The FWI models are shown in the depth domain but the inversion was performed in the vertical time domain.

a change of variable transforms the local depth ( $z$ ) coordinate along the vertical axis into a local zero-offset vertical traveltimes ( $\tau$ ). Denoting the velocity in the vertical direction by  $v(z)$  at a given horizontal position, the vertical traveltimes is defined as:

$$\tau = \int_0^z \frac{dz'}{v(z')}. \quad (8)$$

Conversely, the expression of the depth  $z$  as a function of the pseudo-depth-domain velocity  $\tilde{v}(\tau)$  is given by:

$$z = \int_0^\tau \tilde{v}(\tau') d\tau'. \quad (9)$$

Plessix *et al.* (2012) showed that it is possible to avoid formally deriving the gradient of the misfit function in the  $\tau$  domain.

Instead, he proposed to use the chain rule of the derivatives to infer the gradient of the misfit function in the  $\tau$  domain from the gradient of the misfit function in the depth domain, through the expression:

$$\frac{\partial C}{\partial \tilde{v}(\tau_i)} = \frac{\partial C}{\partial v(z_i)} - \int_{z_i}^{z_{max}} \frac{1}{v(z')} \frac{\partial v(z)}{\partial z} \frac{\partial C}{\partial v(z)} dz. \quad (10)$$

This chain rule allows the gradient of the misfit function in the  $\tau$  domain to be implemented in a relatively simple way, from the existing modelling and inversion tools in the depth domain: the gradient in the depth domain  $\partial C / \partial v(z_i)$  is first computed with usual depth-domain tools and then implemented in equation (10) to retrieve the pseudo-depth gradient  $\partial C / \partial \tilde{v}(\tau_i)$ .

Plessix *et al.* (2012) showed the interest in the use of FWI in the pseudo-time domain to deal with the velocity/depth ambiguity, when initial models are derived from reflection tomography and already contain large velocity contrasts.

Figure 5 shows the inversion results in the vertical travel-time domain  $\tau$ . Note that the results are plotted in the depth domain, but the inversion is performed in the  $\tau$  domain. In this case, the difference-based misfit function fails to converge to the true velocity model, as the inversion is hampered by cycle skipping at intermediate and long offsets, due to the wrong moveout of the reflection computed in the starting model (Fig. 5a–c). In contrast, the inversion performed with the cross-correlation-based misfit function in the  $\tau$  domain succeeds in decreasing the velocity to the true value of 1500 m/s above the reflector (Fig. 5d–f). As the density model was kept fixed during the inversion in the  $\tau$  domain, the depth of the reflector was modified in the density model after time-to-depth conversion according to the velocity update found by the inversion. This update of the reflector position preserves the zero-offset traveltimes, hence moving the reflector to its correct position as the velocities become more accurate. Please note that even if we force the zero-offset traveltime to be matched in the pseudo-time domain, a robust measure of the error to avoid cycle-skipping artefacts remains necessary. Indeed, we show that cycle-skipping artefacts occur when the difference-based misfit function is used and when the starting model is not accurate enough to predict the moveout of the reflection with an error that does not exceed half of the dominant period. The same issue remains when the inversion is performed in the depth domain and the reflectivity is updated at each iteration, as proposed by Xu *et al.* (2012).

## DISCUSSION AND CONCLUSIONS

This study proposes a robust and computationally efficient wave-equation-based seismic inversion workflow for velocity-model building from reflection data. The seismic workflow combines the following three critical key ingredients: (1) prior knowledge of the reflectivity, which allows to combine a migration-like imaging problem with a transmission-like imaging problem. The transmission regime is necessary, to make the inversion sensitive to the large wavelengths of the subsurface; (2) a robust correlation-based misfit function, which avoids cycle-skipping issues; and (3) a suitable inversion domain, which preserves seismic invariants, such as zero-offset traveltimes and hence performing reflectivity updates over inversion iterations is not required as the velocity model

is updated. This domain is obtained by transforming the depth coordinate into the vertical traveltime.

The recent developments in acquisition design and processing to obtain broadband seismic data should relax the requirement of such robust misfit measurements, if cycle-skipping can be avoided. However, the prior knowledge of the reflectivity and the pseudo-time domain should remain necessary to exploit the reflected waves.

Even if beyond the scope of this study, the construction of the short-wavelength reflectivity model is a critical step. Amplitude-versus-offset analysis can be used for some applications, although it can also be problematic for shallow targets on which the direct waves strongly interfere with the reflected waves. Depth or time migration can also be used to delineate the reflectivity, although the focusing can be limited for strongly erroneous velocity models or complex targets. We could also rely on a classical FWI formalism to retrieve the impedance parameter in the depth domain, assuming a known velocity background. The choice of the method should be application-dependent and deserves a dedicated study.

It is well known that cross-correlation functions suffer from cross-talk when applied to multi-arrival data. The cross-correlation should therefore be applied to appropriate domains where the arrivals can be separated. To keep computer costs low, the data-domain separation can involve hyperbolic and/or nonhyperbolic moveout corrections and slant-stack and/or double slant-stack on the source and receiver sides (double beamforming). Alternative error measurements such as dynamic warping can also be investigated, as this has been shown to provide robust results with multi-arrival seismic data (Hale 2012, 2013). Of course, the price of migration can be paid to separate the data directly in the image domain, and to link this study with image-oriented approaches.

## ACKNOWLEDGEMENTS

The authors are grateful to the anonymous reviewers and to the associate editor for their comments that have helped us to improve the manuscript. This study was partially performed using the computing facilities of the SCCI department of CIMENT (Univ. Joseph Fourier, Grenoble). This study is partially supported by the SEISCOPE consortium (<http://seiscope.oca.eu>), which was sponsored by BP, CGGVeritas, ENI, Exxon-Mobil, Petrobras, Saudi Aramco, Shell, Statoil, and Total and the

SEISCOPE II consortium (<http://seiscope2.osug.fr>), which is sponsored by BP, CGG, Chevron, Exxon-Mobil, JGI, Petrobras, Saudi Aramco, Schlumberger, Sinopec, Shell, Statoil, Total and Woodside. The authors would like to thank L. Métivier (LJK, Univ. Grenoble Alpes - CNRS) and R.E. Plessix (Shell) for fruitful discussions.

## REFERENCES

- Alkhalifah T., Fomel S., and Biondi B. 2001. The space-time domain: theory and modelling for anisotropic media. *Geophysical Journal International* **144**, 105–113.
- Almomin A. and Biondi B. 2012. Tomographic full waveform inversion: Practical and computationally feasible approach. *SEG Technical Program Expanded Abstracts* **500**, 1–5.
- Biondi B. and Almomin A. 2012. Tomographic full waveform inversion (TFWI) by combining full waveform inversion with wave-equation migration velocity analysis. *SEG Technical Program Expanded Abstracts*, **547**, 1–5.
- Byrd R. H., Lu P. and Nocedal J. 1995. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific and Statistical Computing* **16**, 1190–1208.
- Chavent G., Clément F. and Gómez S. 1994. Automatic determination of velocities via migration-based traveltimes waveform inversion: A synthetic data example. *SEG Technical Program Expanded Abstracts* **328**, 1179–1182.
- Devaney A. J. 1982. A filtered backprojection algorithm for diffraction tomography. *Ultrasonic Imaging* **4**, 336–350.
- Faye J. P. and Jeannot J. P. 1986. Prestack migration velocities from focusing depth analysis. *SEG Expanded Abstracts*, 438–440.
- Hale D. 2012. Dynamic warping of seismic images. *CWP Report*, **723**.
- Hale D. 2013. Dynamic warping of seismic images. *Geophysics* **78**(2), S105–S115.
- Jannane M., Beydoun W., Crase E., Cao D., Koren Z., Landa, E. et al., 1989. Wavelengths of Earth structures that can be resolved from seismic reflection data. *Geophysics* **54**(7), 906–910.
- Lailly P. 1983. The seismic inverse problem as a sequence of before stack migrations. In: *Conference on Inverse Scattering, Theory and application, Society for Industrial and Applied Mathematics, (Philadelphia)*, eds by R. Bednar and Weglein), pp. 206–220.
- Landa E., Beydoun W. and Tarantola A. 1989. Reference velocity estimation from prestack waveforms: coherency optimization by simulated annealing. *Geophysics* **54**, 984–990.
- Luo Y. and Schuster G. T. 1991. Wave-equation travel time tomography. *Geophysics* **56**, 645–653.
- Ma X. and Alkhalifah T. 2011. Wavefield extrapolation in the pseudo-depth domain. *Expanded Abstracts*, **73<sup>rd</sup> EAGE Annual meeting**, p. A014 .
- Pageot D., Operto S., Vallée M., Brossier R. and Virieux J. 2013. A parametric analysis of two-dimensional elastic full waveform inversion of teleseismic data for lithospheric imaging. *Geophysical Journal International* **193**(3), 1479–1505.
- Plessix R.E., Milcik P., Corcoran C., Kuehl H. and Matson K. 2012. Full waveform inversion with a pseudotime approach. *Expanded Abstracts*, **74<sup>th</sup> EAGE Annual meeting**, p. W012.
- Plessix R. E., 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International* **167**(2), 495–503.
- Pratt R. G. and Worthington M. H. 1990. Inverse theory applied to multi-source cross-hole tomography. Part I: acoustic wave-equation method. *Geophysical Prospecting* **38**, 287–310.
- Prieux V., Brossier R., Gholami Y., Operto S., Virieux J., Barkved O.I. and Kommedal J.H. 2011. On the footprint of anisotropy on isotropic full waveform inversion: the Valhall case study. *Geophysical Journal International* **187**, 1495–1515.
- Routh P., Krebs J., Lazaratos S., Baumstein A., Lee S., Cha Y. H. et al., 2011. Encoded simultaneous source full-wavefield inversion for spectrally shaped marine streamer data. *SEG Technical Program Expanded Abstracts* **30**(1), 2433–2438.
- Sava P. and Biondi B. 2004. Wave-equation migration velocity analysis. I Theory. *Geophysical Prospecting* **52**(6), 593–606.
- Sava P. and Fomel S. 2006. Time-shift imaging condition in seismic migration. *Geophysics* **71**(6), S209–S217.
- Snieder R., Xie M. Y., Pica A. and Tarantola A. 1989. Retrieving both the impedance contrast and background velocity: a global strategy for the seismic reflection problem. *Geophysics* **54**(8), 991–1000.
- Sun D. and Symes W. W. 2012. Waveform inversion via non-linear differential semblance optimization. *SEG Technical Program Expanded Abstracts*, pp. 1–7.
- Symes W. W. and Carazzone J. J. 1991. Velocity inversion by differential semblance optimization. *Geophysics*, **56**, 654–663.
- Tarantola A. 1984. Inversion of seismic reflection data in the acoustic approximation. *Geophysics* **49**(8), 1259–1266.
- Tromp J., Tape C. and Liu Q. 2005. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International* **160**, 195–216.
- van Leeuwen T. and Mulder W.A. 2008. Velocity analysis based on data correlation. *Geophysical Prospecting* **56**(6), 791–803.
- van Leeuwen T. and Mulder W. A. 2010. A correlation-based misfit criterion for wave-equation traveltimes tomography. *Geophysical Journal International* **182**(3), 1383–1394.
- Virieux J. and Operto S. 2009. An overview of full waveform inversion in exploration geophysics. *Geophysics* **74**(6), WCC1–WCC26.
- Woodward M. J. 1992. Wave-equation tomography. *Geophysics* **57**, 15–26.
- Wu R. S. and Toksöz M. N. 1987. Diffraction tomography and multisource holography applied to seismic imaging. *Geophysics* **52**, 11–25.
- Xu S., Wang D., Chen F., Lambaré G. and Zhang Y. 2012. Inversion on reflected seismic wave. *SEG Technical Program Expanded Abstracts* **509**, 1–7.
- Yang T. and Sava P. 2011. Wave-equation migration velocity analysis with time-shift imaging. *Geophysical Prospecting* **59**(4), 635–650.
- Zhou Z., Howard M. and Mifflin C. 2011. Use of RTM full 3D subsurface angle gathers for subsalt velocity update and image optimization: Case study at Shenzi field. *Geophysics* **76**(5), WB27–WB39.

## APPENDIX.

GRADIENT OF FULL-WAVEFORM  
INVERSION MISFIT FUNCTIONS WITH  
PRIOR KNOWLEDGE OF THE  
REFLECTIVITY

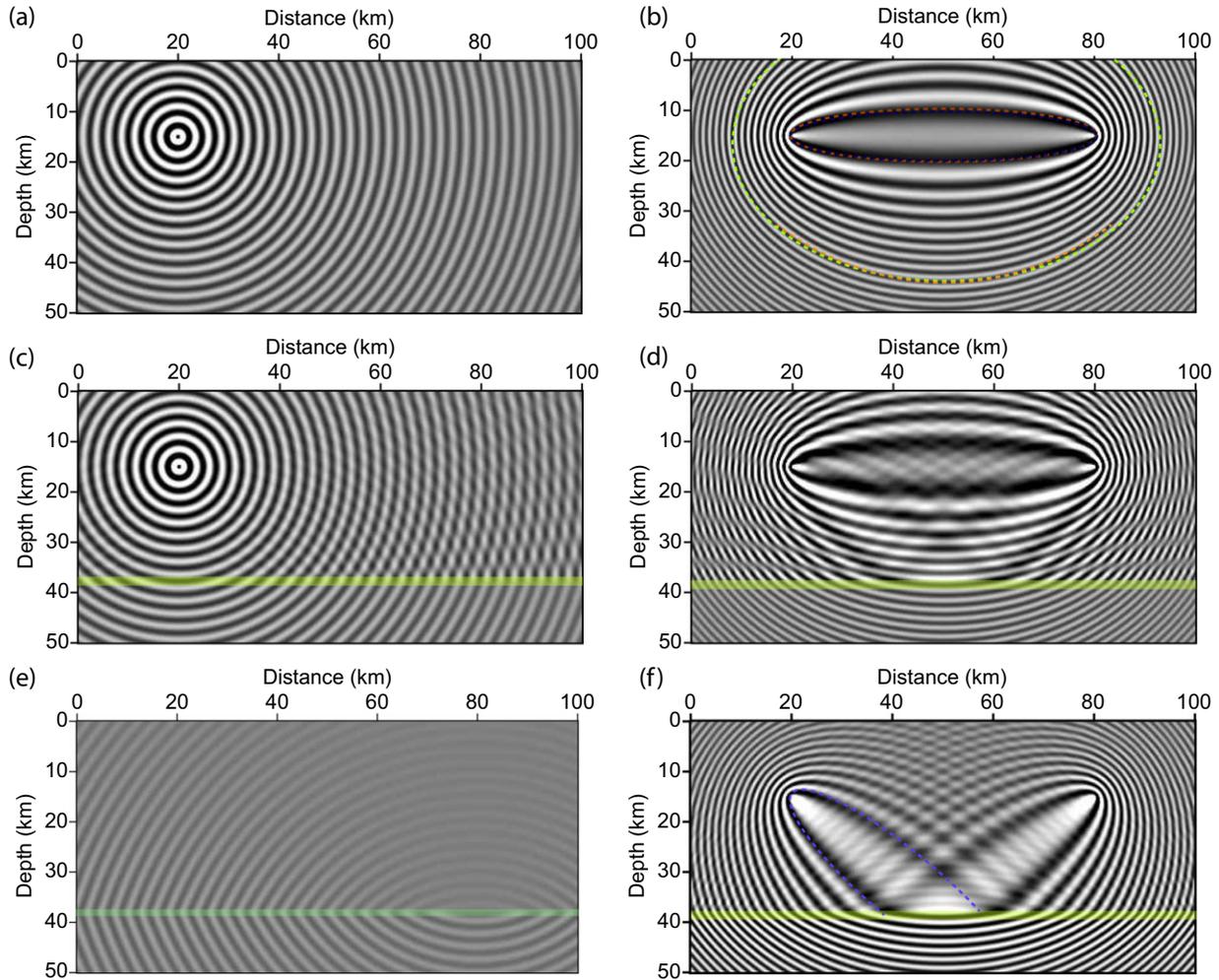
In this Appendix, we first clarify why explicitly forcing the separation between the background model and the reflectivity in the FWI formalism is important to build a suitable sensitivity kernel for the inversion of reflection data for realistic applications. We then derive the gradient of the misfit function for the background model update with the Lagrangian formulation of the adjoint-state method.

Let us first consider the sensitivity kernel of the FWI for monochromatic Green functions when the background model is homogeneous (Fig. A1b). This sensitivity kernel (or wavepath, to follow the terminology of Woodward (1992) is computed by multiplying the background wavefields coming from the shot and receiver positions and it represents the iso-phase or iso-time surfaces on which a data residual recorded at an arbitrary traveltime would be back-projected. Indeed, when the wavepath is built from monochromatic Green functions (i.e., for a sinusoidal time excitation of infinite support) (Fig. A1a), the sensitivity kernel contains the iso-phase surfaces that are associated with all of the traveltimes, hence providing the complete picture of the resolution power carried out by all of the arrivals in the seismogram (Fig. A1b). At shallow depths, the sensitivity kernel represents the first Fresnel zone associated with the forward-scattering regime: all of the diffractions generated in this Fresnel zone constructively interfere with each other to build the first arrival. Conversely, first-arrival residuals are back-projected onto this first Fresnel zone during transmission tomography. The broad width of the first Fresnel zone makes the tomography-like FWI amenable to update to the long-to-intermediate wavelengths of the subsurface from the wide-aperture components of the data. The maximum depth sampled by the first Fresnel zone is controlled by the lowest frequency of the source and the maximum source-receiver offset and it rarely reaches the depths of the deepest targeted structures. For significant depths with respect to the source-receiver offset, the outer fringes of the sensitivity kernel represent secondary Fresnel zones (referred to as migration wavepaths, or isochrone surfaces in this study) that are associated with the back-scattering regime: all of the diffractions generated along an isochrone surface constructively interfere with each other to build a reflection arrival. Conversely, the reflected arrivals are back-projected onto the corresponding isochrone surface during migration. The width

of these isochrones, equation (1), highlights the resolution of the migration-like FWI in the spatial direction  $\vec{n}$  perpendicular to the isochrone surface (the vertical direction in the case of a flat reflector). As the width of the isochrones is relatively limited and decreases with depth as the scattering angle decreases for a fixed source-receiver offset, the long wavelengths of the subsurface cannot be updated in the deep part of the subsurface.

As FWI proceeds over non-linear iterations of reflection data, short-wavelength components are superimposed on the smooth background model through the migration process underlying FWI. The sensitivity kernel of the FWI is then modified accordingly. If we assume now that the background model is composed of a homogeneous background model plus a reflector, the full sensitivity kernel (Fig. A1d) is formed by the primary sensitivity kernel (Fig. A1b) plus a secondary sensitivity kernel that results from the seismic response of the reflector. This secondary sensitivity kernel is built on the one hand by the multiplication of the background wavefield coming from the shot position with the scattered wavefield coming from the receiver position and on the other hand by the multiplication of the incident background wavefield coming from the receiver position with the scattered wavefield coming from the shot position. This secondary sensitivity kernel is amenable to the update of the long-to-intermediate wavelengths of the subsurface, because it is formed by transmitted wavepaths from the reflector to the surface, just as in cross-hole tomography. A key issue is that this secondary sensitivity kernel is generally dominated by the primary sensitivity kernel (the difference between Fig. A1b and Fig. A1d represents the footprint of this secondary sensitivity kernel), because the amplitudes of the direct wavefields (Fig. A1a) are significantly higher than those of the reflected wavefield (Fig. A1d). Extraction of the secondary sensitivity kernel from the full sensitivity kernel through the subtraction of the primary sensitivity kernel of Fig. (A1b) from the full sensitivity kernel of Fig. A1(d) highlights more clearly the resolution power of the information carried by the depth-to-surface transmitted wavepaths (Fig. A1f). The resulting wavepath is dominated by the first Fresnel zones associated with a set of virtual sources located at the reflector, and virtual point receivers located at the shot position and at the real receiver position, respectively. Note also that the outer fringes of the wavepath should also make FWI amenable to the high-resolution imaging of steep reflectors, such as the flanks of salt domes.

This description highlights the key role of some prior knowledge of the reflectivity in the background model of the FWI, as secondary sources in depth to update the



**Figure A1** Transmission wavepath versus migration wavepath. (a) Monochromatic wavefield in a homogeneous background model. (b) Wavepath (sensitivity kernel of the FWI) computed by multiplication of the shot monochromatic wavefield (a) and a receiver monochromatic wavefield for a source located at  $(x,z) = (80 \text{ km}, 15 \text{ km})$  (not shown in the figure). The red dashed line and the Green dashed line delineate the first Fresnel zone and an isochrone surface, on which the first-arrival residuals and the reflection residuals are back-projected during the FWI. (c) Monochromatic wavefield computed in the homogeneous background model, to which a flat reflector (thick Green solid line) at 38 km in depth was added. The monochromatic wavefield contains the direct wavefield (a) and the reflected wavefield. (d) Monochromatic wavepath computed in the homogeneous background model with the reflector. A comparison between (b) and (d) highlights the weak footprint of the wavepaths generated by a virtual secondary source that is located at the reflector position. (e) Reflected wavefield for a source located at the receiver position the amplitudes of which are significantly smaller than those of the primary wavefield (a). (f) The two secondary wavepaths generated by the virtual secondary source located at the reflector position are shown (see text for details). The dashed line delineates the first Fresnel zone of the wavepath, computed by multiplying (a) with (e).

long-to-intermediate wavelengths of the subsurface and also the difficulty associated with the extraction of this second-order information from the full sensitivity kernel of the FWI.

The trick to filter out the footprint of the high-amplitude migration wavepath from the full wavepath, to bring out the transmission wavepath in the sensitivity kernel of the FWI, is to assume that the reflectivity is known and to minimize the misfit function for the update of the background model

(e.g., Xu *et al.*, 2012). This scale separation between the background model and the reflectivity is at the heart of most seismic reflection processing.

Let us now derive the gradient of the misfit function of the reflected wavefield under the above-mentioned assumption. We use the adjoint-state method to derive the expression of the gradient of the misfit function following the Lagrangian formulation reviewed in Plessix (2006). All of the

developments are performed with a matrix formalism in the frequency domain for one source and one frequency, for the sake of compactness and simplicity. Similar expressions can be derived for multiple frequencies (or in the time domain) and multiple sources in a straightforward way, by summing the elementary contribution of each frequency and each source.

Let us consider the least-squares misfit function  $C$  given by the expression:

$$C = \frac{1}{2} \|\delta d_{cal} - \delta d_{obs}\|^2, \quad (A1)$$

where  $\delta d_{cal}$  is the modelled reflection data and  $\delta d_{obs}$  is the recorded reflection data at all receiver positions. We use the single-scattering Born approximation to derive the expression of the modelled reflection wavefield  $\delta d_{cal}$  from the known reflectivity  $\delta m$ :

$$A(m_0)\delta u_{cal} \approx -(A(m_0 + \delta m) - A(m_0))u_0, \quad (A2)$$

where the background wavefield  $u_0$  is the solution of the wave equation for the incident source  $s$  located at the shot position:

$$A(m_0)u_0 = s. \quad (A3)$$

The matrix  $A$  in equations (A2) and (A3) is the two-way wave-propagation operator: its coefficients depend on frequency (not written here) and on the background model  $m_0$ .

We have  $\delta d_{cal} = R\delta u_{cal}$ , where  $R$  is a sampling operator that extracts the values of  $\delta u_{cal}$  from the modelling space to the receiver positions in the data space. The term  $(A(m_0 + \delta m) - A(m_0))$  represents the reflectivity. For example, let us consider the scalar time-harmonic wave equation written as:

$$(\omega^2 m + \Delta)u = s, \quad (A4)$$

where the model parameter  $m$  is the squared slowness and  $u$  is the pressure wavefield. For the squared-slowness parametrization, the term  $(A(m_0 + \delta m) - A(m_0))$  reduces to  $\omega^2 \delta m$ . In the following,  $(A(m_0 + \delta m) - A(m_0))$  will be substituted by  $\omega^2 \delta m$ .

The Lagrangian, i.e., the augmented misfit function, is given by the equation:

$$\begin{aligned} L(m_0, u_0, \delta u_{cal}, \delta d_{cal}, a_1, a_2, a_3) &= \frac{1}{2} \|\delta d_{cal} - \delta d_{obs}\|^2 \\ &+ \Re \langle a_1 | A(m_0)u_0 - s \rangle \\ &+ \Re \langle a_2 | A(m_0)\delta u_{cal} + \omega^2 \delta m u_0 \rangle \\ &+ \Re \langle a_3 | \delta d_{cal} - R\delta u_{cal} \rangle, \end{aligned} \quad (A5)$$

where fields  $a_1$ ,  $a_2$  and  $a_3$  are the adjoint-state variables that are associated with three state equations. The first state equation governs the propagation of the wavefield  $u_0$  emitted by the seismic source  $s$  in the background model: equation (A3). The second state equation computes the modelled reflection wavefield  $\delta u_{cal}$  from the incident wavefield  $u_0$  with the Born approximation: equation (A2). The third state equation extracts the modelled data from the modelled wavefield at all of the receiver positions.

The adjoint-state equations are  $\partial L / \partial \delta d_{cal} = 0$ ,  $\partial L / \partial \delta u_{cal} = 0$  and  $\partial L / \partial u_0 = 0$ . After noting that the variable  $a_3$  reduces to the assemblage of the data residuals, the wave equations satisfied by the adjoint-state variables  $a_1$  and  $a_2$  are given by:

$$A(m_0)a_2^* = R^t(\delta d_{cal} - \delta d_{obs})^*, \quad (A6)$$

$$A(m_0)a_1^* = \omega^2 \delta m^t a_2^*. \quad (A7)$$

As shown by equation (A6), the variable  $a_2$  is the usual back-propagated residual wavefield from the receiver positions. The operator  $R^t$  projects the residual  $\delta d_{cal} - \delta d_{obs}$  defined in the data space to the modelling space. This back-propagated wavefield is multiplied by the reflectivity  $\delta m$  in the right-hand-side term of equation (A7), hence describing a source term where the spatial support coincides with the reflectivity and where the temporal support is the arrival time of the incident wavefield at the reflector position.

Remembering that  $\delta m$  does not depend on  $m_0$ , the gradient of the misfit function with respect to  $m_0$  is given by the expression:

$$\nabla C = \Re \left\{ u_0^t \left( \frac{\partial A(m_0)}{\partial m_0} \right)^t a_1^* \right\} + \Re \left\{ \delta u_{cal}^t \left( \frac{\partial A(m_0)}{\partial m_0} \right)^t a_2^* \right\}, \quad (A8)$$

which can be simplified as:

$$\nabla C = \omega^2 \Re \{ u_0^t a_1^* + \delta u_{cal}^t a_2^* \}, \quad (A9)$$

if we note that  $\partial A(m_0) / \partial m_0 = \omega^2$ , where  $A$  is given in equation (A4) and the model  $m_0$  is parametrized by the square of the slowness.

The first term in the gradient describes the zero-lag correlation between the background incident wavefield  $u_0$  emitted by the seismic source  $s$  from time 0 s (Fig. A1a) and the adjoint-state wavefield  $a_1$ , which is scattered upwards from the reflector position from the arrival time of  $u_0$  at the reflector position (Fig. A1e). The second term describes the zero-lag correlation between the scattered wavefield  $\delta u_{cal}$  emitted by

the reflectivity  $\delta m$  from the arrival time of  $u_0$  at the reflector position, and the background residual wavefield  $a_2$ , which is back-propagated in space from the receiver positions and in time from the recording time of the reflected wavefield.

The two terms of the gradient just described, as in equation (A9), correspond to the two sensitivity sub-kernels shown in Fig. A(1d) in which the imprint of the migration isochrones (Fig. A(1b) has been explicitly filtered out, to bring out the wavepaths that are useful for the purpose of updating the background model.

In this Appendix, for the sake of simplicity, we choose to model the reflection wavefield with the single-scattering Born approximation. An alternative is to model the full scat-

tered wavefield, to account for internal multiples with the scattered-field formulation (i.e., Pageot *et al.*, 2013). This simply amounts to replacing  $A(m_0)$  by  $A(m_0 + \delta m)$  in the left-hand-side of equation (A2).

Note that we implement the numerical experiments shown in the main body of the text with the conventional form of FWI (i.e., without forcing the explicit uncoupling between  $m_0$  and  $\delta m$  in the formalism), for the sake of simplicity. This does not question the conclusions of the numerical experiment, because the impedance contrast along the unique reflector is sufficiently sharp to make the contribution of the transmission wavepath significant in the sensitivity kernel relatively to the migration wavepath (Fig. 3).