

## Robust Full Waveform Inversion of surface waves

*I. Masoni<sup>1,2</sup>, R. Brossier<sup>1</sup>, J.-L. Boelle<sup>2</sup>, M. Macquet<sup>1</sup> and J. Virieux<sup>1</sup>*

<sup>1</sup>*Université Grenoble Alpes, ISTerre, Grenoble, France*

<sup>2</sup>*E&P, TOTAL, France*

### ABSTRACT

We have investigated a local optimization procedure for seismic imaging by full waveform inversion of surface waves. These waves are highly energetic when considering near-surface imaging and present dispersive effects as well as significant forward scattering. These propagation behaviours have led us to consider alternative data domains where the data fitting should be performed. The frequency-slowness domain as well as the frequency-wavenumber domain have been considered. We have shown how to build the gradient of the misfit function in these two domains based on the adjoint approach. The adjoint source term has been estimated with a systematic procedure based on Lagrangian multipliers, making the cost of the procedure identical to the one formulated in the standard time-offset domain.

A simple 2D synthetic example shows that the reconstruction of the shear velocity is possible, although real applications could require a more extensive investigation including windowing and filtering, as a hierarchical strategy appears to be necessary when considering surface waves with high variations in amplitude.

A linearised approach is therefore possible and quite efficient. This avoids the picking of dispersion curves and attempts to fill in the velocity spectrum, extending our search domain to models with lateral velocity variations as done previously using global model search.

**Keywords:** Surface waves, Inverse theory, Dispersion curve analysis, Velocity spectrum, Full waveform inversion

### INTRODUCTION

The recorded seismic signals at the free surface provide a complex signature we may need to analyze and decipher. Of the different recorded phases in seismograms or traces, surface waves are the most energetic ones and can provide interesting information on targeted zones from near surface to deep structures. In exploration seismics, these phases are considered as noise, although they have been intensively used in seismology for lithospheric reconstruction (van der Lee, 2002), mantle imaging (Nolet, 1977; Romanowicz, 2003) and for global Earth imaging (Lerner-Lam and Jordan, 1983). Shear wave velocity (and bulk velocity and density to a lesser extent) influences the propagation of surface waves through the dispersion relation. Seismic imaging with these surface waves proceeds in a rather different way compared to using body waves as they are dispersive in heterogeneous media (Aki and Richards, 2002) and provide forward scattering interaction with the free surface when propagating almost horizontally.

Through time-frequency analysis based on moving Gaussian filter (Cara, 1973), disper-

sion curve analysis is built at each station: the fundamental mode is picked manually or automatically for different frequencies. Phase differences are obtained between two stations when considering earthquakes. From these 2D mapping of apparent Rayleigh wave phase velocities between two stations at different frequencies, a 3D mapping of the S-wave phase velocities is reconstructed, using the different vertical variations of normal modes with frequency. Although difficult, the detection and the interpretation of the fundamental mode is easier at large scale than at short scale as for subsurface imaging. This has led to consider surface waves as ground rolls and, therefore, noise, although they carry the most important part of the energy in subsurface imaging.

When considering active sources for the emission of seismic waves at shallow depths from tens of meters to few hundreds of meters, the interpretation of surface waves has been promoted as the spectral analysis of surface waves (SASW) using two stations and many active sources as hammers. The picking of dispersion curves done through frequency-wavenumber or slant analysis provides phase velocity diagrams enabling the reconstruction of a local 1D shear-wave model (McMechan and Yedlin, 1981; Heisey et al., 1982; Nazarian and Stokoe, 1984; Stokoe et al., 1994). This method has been efficiently used in many geotechnical engineering investigations leading to 1D S-wave profiles. Later, an extension to many stations as for refraction seismic investigations, through the multichannel analysis of surface waves (MASW) (Park et al., 1999), has allowed improved reconstructions of the 1D S-wave velocity profile. Phase analysis on multiple stations enables a better detection of unwanted energy bursts and a better detection/delineation of the dispersion curve which is used for the profile reconstruction. This approach, intrinsically valid for laterally invariant media, has been extended to smoothly varying medium but is still strongly limited by the requirement to pick and interpret dispersion curves as well as the local constant-velocity layer-based representation of the subsurface (Socco and Strobbia, 2004; Boiero and Socco, 2010; Bergamo et al., 2012). Real applications have been performed in different subsurface environments, but detection of low velocity zones is still difficult for the dispersion-curves interpretation, especially when considering higher modes (Gabriels et al., 1987; Louie, 2001; O’Neil, 2004; Park et al., 2005).

Therefore, mainly the fundamental mode of Rayleigh waves has been used in most applications and efficient inversion schemes have been designed for recovering the S-wave velocity profile. Linearised approaches (Herrmann, 1987) are used and complemented by pure Monte-Carlo perturbation (Shapiro et al., 1997). Semi-global search such as the neighbourhood algorithm has been promoted by Wathelet (2008); Wathelet et al. (2008) for more efficient exploration of the model space with constraints. Very few studies include waveform inversion of surface waves using non-local investigations (Ryden and Park, 2006; Maraschini and Foti, 2010) where one escapes the problem of detecting and interpreting dispersion curves. In these investigations, the extensive exploration of the model space is limited to a rather simple model description. Other studies based on linearised approaches do not focus on surface waves but on the entire seismogram: the so-called full waveform inversion includes both body and surface waves and turns out to be difficult even on synthetic data (Gélis et al., 2007; Romdhane et al., 2011; Schäfer et al., 2013) for near-surface targets. Brossier et al. (2009) have shown that preconditioning the data by time damping allows successful reconstruction of acceptable models while considering the entire waveform on a section of the SEG/EAGE overthrust crustal model using elastic wave propagation with a free surface. The contribution of the surface waves is unclear as one wonders if the body waves are mainly driving the imaging process.

Specific developments of local optimization of surface waveforms deserve an investigation starting from the definition of the misfit function between observed seismograms and synthetic ones at the free surface. In order to avoid local minima, an analysis of different misfit functions has been performed by Masoni et al. (2013) in the framework of a linearised formulation which can be considered as an extension from the dispersion curve analysis towards an interpretation of the full velocity spectrum. Other investigations have also shown potential interests (Pérez Solano et al., 2014; Yuan et al., 2014).

In this paper, we shall detail the different misfit functions to be used when considering the waveforms of surface waves. After reviewing the elastic wave equation solved with a finite difference scheme to allow lateral variations of material properties, we formulate the optimization procedure using the adjoint formulation for the minimisation of the three misfit functions we have considered. We show how to solve the related adjoint wave equation with a specific source term. We finally proceed with a simple synthetic example to show how we can reconstruct the velocity spectrum starting from an initial model we construct. A complete workflow starting from a MASW analysis to a FWI procedure is beyond the scope of this paper devoted to the design of the optimization scheme in the particular case of surface waves.

## FORWARD MODELLING

In a 2D medium, seismic waves follow the elastodynamics equation which can be recast into first-order hyperbolic partial differential equations with particle velocity vector  $(v_x, v_z)$  and symmetrical stress tensor, recasted as the vector  $(\sigma_{xx}, \sigma_{zz}, \sigma_{xz})$ , as spatially varying unknowns to be constructed. We consider a time formulation where these equations can be written explicitly as

$$\begin{aligned}
 \rho \frac{\partial v_x}{\partial t} &= \frac{1}{\rho} \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xz}}{\partial z} + f_{v_x} \\
 \rho \frac{\partial v_z}{\partial t} &= \frac{1}{\rho} \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{zz}}{\partial z} + f_{v_z} \\
 \frac{\partial \sigma_{xx}}{\partial t} &= (\lambda + 2\mu) \frac{\partial v_x}{\partial x} + \lambda \frac{\partial v_z}{\partial z} + g_{\sigma_{xx}} \\
 \frac{\partial \sigma_{zz}}{\partial t} &= (\lambda + 2\mu) \frac{\partial v_z}{\partial z} + \lambda \frac{\partial v_x}{\partial x} + g_{\sigma_{zz}} \\
 \frac{\partial \sigma_{xz}}{\partial t} &= \mu \left( \frac{\partial v_x}{\partial z} + \frac{\partial v_z}{\partial x} \right) + g_{\sigma_{xz}},
 \end{aligned} \tag{1}$$

where Lamé coefficients  $\lambda$  and  $\mu$  as well as the density  $\rho$  characterise the medium properties and can spatially change. We have made the spatial dependence implicit in the writing of these equations. External sources could be applied as forces  $(f_{v_x}, f_{v_z})$  or as an internal time-derivative of stress tensor  $(g_{\sigma_{xx}}, g_{\sigma_{zz}}, g_{\sigma_{xz}})$ . We have discretised these equations at nodes of a staggered finite-difference grid. The spatial derivatives are approximated by a fourth-order central finite-difference stencil (Levander, 1988) and the explicit time integration relies on a second-order leap-frog scheme.

This system could be given to the vector  $u^t = (v_x, v_z, \sigma_{xx}, \sigma_{zz}, \sigma_{xz})^t$  with a compact

expression

$$\partial_t u(t, \mathbf{x}) = A_x(\mathbf{x}, m) \partial_x u(t, \mathbf{x}) + A_z(\mathbf{x}, m) \partial_z u(t, \mathbf{x}) + s(t, \mathbf{x}), \quad (3)$$

where differential operators  $A_x$  and  $A_z$  depend on the model parameters  $m$  at the each position  $\mathbf{x}$  in the model. The transpose of a vector is denoted by the symbol  $^t$ . A conservative form could be designed by considering symmetrical compliance operator  $S(m, \mathbf{x})$  leading to an implicit differential system

$$S(\mathbf{x}, m) \partial_t u(t, \mathbf{x}) = H_x(\mathbf{x}) \partial_x u(t, \mathbf{x}) + H_z(\mathbf{x}) \partial_z u(t, \mathbf{x}) + S(m, \mathbf{x}) s(t, \mathbf{x}), \quad (4)$$

where symmetrical differential operators  $H_x$  and  $H_z$  are composed of values one or zero and do not depend on the model properties (Burrige, 1996). Please note that this transformation modifies the source term in this conservative expression. Expression (4) is not very useful for modelling as it turns out to be implicit in time, but is quite interesting for the formulation of the inversion because it is self-adjoint and terms depending on both model properties and spatial derivatives can be split.

Boundary conditions are crucial for proper surface wave simulations as well as absorbing boundary conditions for mimicking the extended infinite medium. For flat free surfaces, we have considered the stress imaging condition (Levander, 1988; Robertsson, 1996) which presents a good compromise between computational cost and accuracy, although a limited number of propagated wavelengths can be considered to avoid accumulation of numerical dispersion. The free surface interface goes through nodes sampling stress components  $(\sigma_{xx}, \sigma_{zz})$  and velocity component  $v_x$ . Finally, the three other edges of the simulation box are implemented with Perfectly Matched Layers (Bérenger, 1994) to mimic a non-reflecting infinite medium.

This finite difference modelling tool will be the forward modelling kernel in the optimization scheme we shall consider, whichever data domain is used for comparing observed and synthetic data.

## MISFIT FUNCTION ANALYSIS

FWI is a data-fitting imaging technique used to derive quantitative models of the subsurface considering the entire set of waves, including surface waves (Virieux and Operto, 2009). It aims to solve the ill-posed inverse problem of deriving the model parameters that describe the subsurface from observed seismic data. At the oil & gas exploration scale, FWI is up to now mainly used to image at depth considering an acoustic approximation to delineate high resolution velocity models for migration, only using body waves.

The method aims to match the full observed seismogram with a corresponding synthetic seismogram computed from an initial velocity model, solving a local optimization problem. To measure the residual between the observed and the calculated seismograms, the L2 norm of the difference is conventionally used, fitting both amplitude and phase information of the waveforms (Tarantola, 1984; Pratt et al., 1998). This L2 norm generally focuses on the sample-to-sample difference between the observed and the computed data in a given domain, generally the time-offset domain  $(t, h)$ , or frequency-offset  $(\omega, h)$  for frequency-domain FWI, leading to the following misfit function

$$C_{t,h} = \frac{1}{2} (d_{obs}(t, h) - d_{cal}(t, h))^2, \quad (5)$$

where we assume an implicit summation over sources and receivers. By exploiting the full information content of the data and employing a strict data-matching approach, this method is able to produce very high resolution results, but may therefore not be very robust or stable. FWI is indeed susceptible to non-linearity problems such as cycle skipping and local minima, which reduce the convexity of the misfit function (Mulder and Plessix, 2008). As a result, the initial predicted velocity model needs to be accurate enough so that it lies within the small basin of attraction of the global minimum, meaning that it has to be kinematically correct. Otherwise the minimisation easily ends up in the nearest minimum without ever reaching the global minimum, and converges to a wrong velocity model. It is believed that the initial model has to explain the data to within half a wavelength to avoid cycle skipping and converge.

In order to relax this constraint, we may either design a more robust comparison criterion between observed and synthetic data or we may consider comparing observed and synthetic data on a sample-by-sample strategy in different domains. In the following, we shall initially analyze different misfit functions before going more in detail on alternative domains for sample-to-sample comparisons of data with a main focus on surface waves.

## Robust misfit design

Designing a more robust misfit comparison between observed data and synthetic data could prevent cycle skipping problems and local minima issues related to the essentially oscillating nature of seismic waves. Extracting time-shifts to be minimised by the cross-correlation of windowed waveforms has been promoted by Luo and Schuster (1991); Tromp et al. (2005). van Leeuwen and Mulder (2010); Brossier et al. (2014) use a maximisation of the cross-correlation function between observed and calculated data to focus energy close to zero-time lag. These comparisons based on cross-correlation between two signals are robust with respect to cycle-skipping but face ambiguity when considering several events. Zero-lag cross-correlation to minimise the effect of amplitude in the misfit (Routh et al., 2011) or a phase-only misfit function (Bednar et al., 2007) could mitigate amplitude effects but are sensitive to cycle-skipping unless unwrapping is performed adequately. Integration of the signal preventing the oscillation appears to be sensitive to noise and amplitude prediction (Donno et al., 2013). The use of the envelope of the signal also appears to relax the initial model accuracy but cycle-skipping is still present (Bozdag et al., 2011). Deconvolution is also an alternative and has been proposed by Luo and Sava (2011).

In order to illustrate differences between different misfit functions, we consider the cross-correlation  $Xcorr(\tau)$  between the observed and synthetic data

$$Xcorr(\tau) = \sum_{t=1}^{nt} d_{obs}(t + \tau) d_{cal}(t) \quad (6)$$

and the normalised cross-correlation  $Xcorr_N(\tau)$

$$Xcorr_N(\tau) = \frac{Xcorr(\tau)}{\|d_{obs}\| \|d_{cal}\|}. \quad (7)$$

This cross-correlation function (7) is insensitive to the amplitudes of the wavelet and allows to access the phase of the data. van Leeuwen and Mulder (2010); Brossier et al. (2014)

used this function to define a weighted norm of the normalised cross-correlation through the expression

$$C_{W_i} = \frac{1}{2} \sum_{\tau=1}^{n\tau} (W_i(\tau) X_{corr_N}(\tau))^2, \quad (8)$$

where the weight  $W_i(\tau)$  is applied at each time sample. The width and the convexity of these misfit functions are modified depending on the applied weighting. Two types of weighting are considered through the following expressions

$$\begin{aligned} W_1(\tau) &= (\tau/\tau_{max})^2 \\ W_2(\tau) &= -e^{-\alpha\tau^2}. \end{aligned} \quad (9)$$

The first weighting  $W_1$  linearly penalises the values further away from zero-lag value while the second one concentrates on values near by the zero-lag (van Leeuwen and Mulder, 2010). If the observed and the synthetic data match, the energy in the cross-correlation is maximised at zero-lag, and the energy at zero-lag is subsequently annihilated by the penalty function. The weighted norm can only be zero if the wavelet is a delta pulse and all cross-talk is excluded by choosing an appropriate maximal shift. However, the cross-correlation function suffers from cross-talk between different events in traces which leads to significant energy away from zero-lag even for perfectly fitting data (this is also true for the auto-correlation of the signal). This effect limits the application of such a misfit to windowed data or when one arrival dominates, which could be the case for surface waves.

In order to illustrate the impact of the three different misfit functions  $C_{(t,h)}$ ,  $C_{W_1}$ ,  $C_{W_2}$ , we consider a simple model with two parameters which are perturbed in order to extensively sample the misfit function: the velocity of the first layer and the thickness of this layer (Figure 1). Using finite-difference modelling the true data values are computed, as well as the synthetic data values for each model sampling the two-parameter space. The range of the perturbed layer depth is of 15 m to 25 m, while the true depth is 20 m, and the perturbed layer velocity ranges from to 480 m/s, while the true value is 450 m/s. Figure 2 illustrates the difficulties of the standard sample-by-sample misfit function and the advantages of considering normalised cross-correlations at the expense of an additional parameter coming from the weight.

These cross-correlation based misfit functions can be quite interesting alternatives when starting from crude initial models although the expected resolution characterised by the

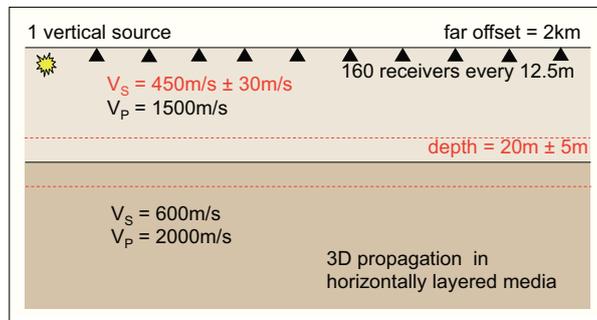


Figure 1: Design of a simple model with two parameters defined as the depth of the interface and the S-wave velocity in the upper layer

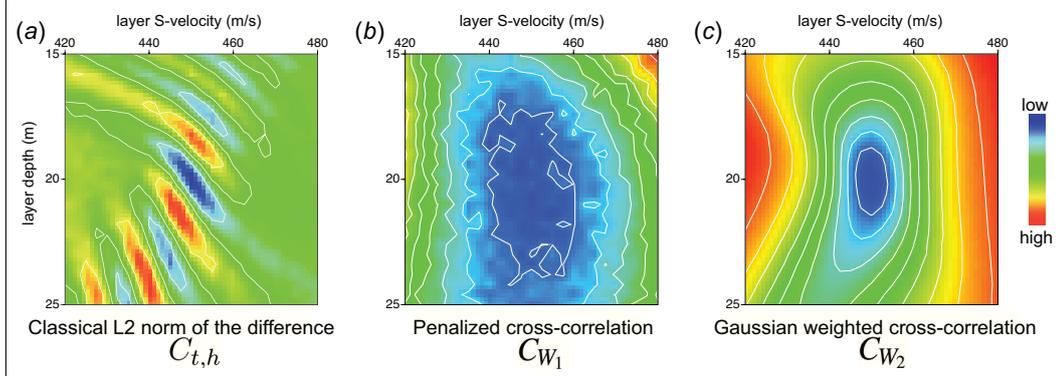


Figure 2: Misfit values for the three different misfit functions we have investigated. On the left panel (a), one can see that the classical misfit function has many local minima while the middle panel (b), penalising values far from zero-lag, is more noisy. The right panel (c) presents a nice smooth valley, given by the Gaussian windowing which needs to be carefully designed.

shape of the attraction basin may be quite poor. Furthermore, cross-talk effects may require to window the data and the inaccuracy of the source wavelet could mitigate this robustness as the cross-correlation is quite sensitive to the spectrum of the data. Further research would be required to tackle these problems. In the following, we focus on the sample-to-sample strategy, considering alternative domains.

### Sample-by-sample strategy

We may benefit from dense acquisition sampling for an improved analysis of the sample-by-sample comparison. Alternative misfit functions can be used following similar strategies as those used in dispersion curve analysis. McMechan and Yedlin (1981) have transformed data in the frequency-slowness domain while Gabriels et al. (1987) have considered data transformed to the frequency-wavenumber domain. This leads us to propose the following misfit functions for FWI, first in the  $\omega - p$  domain

$$C_{\omega,p} = \frac{1}{2} (|d_{obs_N}(\omega, p)| - |d_{cal_N}(\omega, p)|)^2, \quad (10)$$

and then in the  $\omega - k$  domain after two Fourier transforms of the data

$$C_{\omega,k} = \frac{1}{2} (|d_{obs_N}(\omega, k)| - |d_{cal_N}(\omega, k)|)^2. \quad (11)$$

The data is normalized before being transformed to the alternative domain. The phase has also been removed from these two new misfit functions, and only the amplitude of the velocity spectrum is considered. This strategy is related to dispersion curve analysis where maximum-energy picks allow the extraction of the apparent dispersion velocity at each frequency (Pérez Solano et al., 2014). A related misfit design has been promoted by Maraschini et al. (2010) for the particular case of layered media in order to avoid detection and picking of dispersion curves.

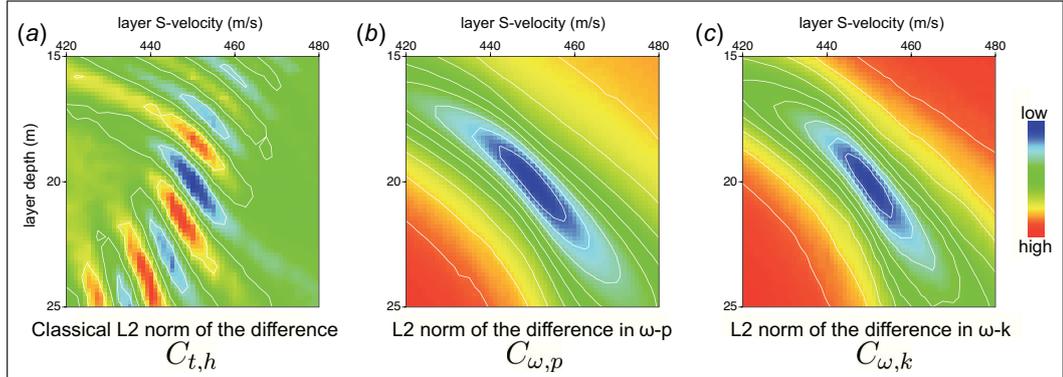


Figure 3: Sample-by-sample differential misfit values for the three different domains we consider. The left panel (a) is the same as the one in Figure 2, while the middle panel (b) is for the  $(\omega, k)$  domain and the right panel for the  $(\omega, p)$  domain. Relation  $k = \omega p$  between slowness and wavenumber leads to a similar structure of the misfit functions.

Applying a linear moveout (LMO) to obtain data in the  $\omega - p$  domain or applying a wavenumber Fourier transform to obtain data in the  $\omega - k$  domain separate events through slopes or wavenumbers: it might allow the extraction of kinematic information in a more robust manner. Furthermore the stacking involved in the transformation may also make the misfit function more efficient in the presence of noise. In these domains, we may have a better identification of frequency-dependent dispersion effects. Moreover, because we consider the modulus of the data, the misfit function is no more sensitive to the phase of the source wavelet, but should not limit the capacity of fitting the kinematic properties contained in the data. The simple example we have considered may help illustrate the effects of these transformations. Figure 3 displays the misfit function structure where one can observe major differences between the  $(t, h)$  domain and both the transformed  $(\omega, p)$  and  $(\omega, k)$  domains, which have a broad attraction basin. Of course, this structure may be due to the simple description of the model space.

We may now consider how to update the model parameters while considering the comparison of data in these different domains, but still doing the forward modelling in the  $(t, h)$  domain using the finite-difference engine we have selected through time integration of the differential system (3).

## OPTIMIZATION APPROACH

We shall proceed by a linearized approach for updating the model, and we may therefore estimate the gradient operator of each misfit function and eventually the effect of the Hessian operator on the model perturbation vector. In this work, we shall concentrate on the estimation of the gradient and we shall consider quasi-Newton approaches for updating the model (Byrd et al., 1995). We shall proceed following Plessix (2006) by constructing a Lagrangian function for each domain in which the seismic data are expressed.

### Misfit in the domain $(t, h)$

We shall first consider the generic scalar product

$$\langle f(t, \mathbf{x}) | g(t, \mathbf{x}) \rangle_{T, \Omega^2} = \int_T \iint_{\Omega^2} dt d\mathbf{x} f(t, \mathbf{x})^\dagger g(t, \mathbf{x}), \quad (12)$$

over the time window  $[0, T]$  and the 2D space  $\Omega^2$ . Note that symbol  $\dagger$  denotes the transpose complex conjugate of a complex number vector if  $f(t, \mathbf{x})$  is defined in  $\mathbb{C}^n$ .

The misfit function will be defined still assuming an implicit summation over sources through the expression

$$\begin{aligned} C_{t,h}(m, d_{cal}) &= \frac{1}{2} \|d_{obs}(t, h) - d_{cal}(t, h)\|_{T, \Omega_r}^2 \\ &= \frac{1}{2} \langle d_{obs}(t, h) - d_{cal}(t, h) | d_{obs}(t, h) - d_{cal}(t, h) \rangle_{T, \Omega_r}, \end{aligned} \quad (13)$$

where  $\Omega_r$  is the receiver space.

We shall consider the following Lagrangian

$$\begin{aligned} \mathcal{L}(m, d_{cal}, u, \xi, \zeta) &= C_{t,h}(m, d_{cal}) + \langle \xi(t, h) | d_{cal}(t, h) - R_h u(t, \mathbf{x}) \rangle_{T, \Omega_r} \\ &+ \langle \zeta(t, \mathbf{x}) | S \partial_t u(t, \mathbf{x}) - H_x \partial_x u(t, \mathbf{x}) - H_z \partial_z u(t, \mathbf{x}) - S s(t, \mathbf{x}) \rangle_T \\ &+ \langle \zeta(0, \mathbf{x}) | u(0, \mathbf{x}) \rangle, \end{aligned} \quad (14)$$

which must have derivatives equal to zero for the different arguments which are considered as independent (Chavent, 2009). Please note that we have considered here the differential system (4) with an initial value condition, and that spatial dependency on operators  $S$ ,  $H_x$  and  $H_z$  have been removed for compactness.

Zeroing the derivative with respect to the vector  $\zeta$  gives the state equation of wave propagation at all spatial positions and all times, while zeroing the derivative with respect to the vector  $\xi$  will define the state equation which projects the wavefield  $u(t, \mathbf{x})$  at receiver positions through the operators  $R_h$ .

Zeroing the derivative with respect to the state variable  $u(t, \mathbf{x})$  and  $d_{cal}(t, h)$  leads to the definition of two adjoint variables

$$\xi(t, h) = d_{obs}(t, h) - d_{cal}(t, h) = \Delta d(t, h) \quad (15)$$

and

$$\begin{aligned} S \partial_t \zeta(t, \mathbf{x}) - H_x \partial_x \zeta(t, \mathbf{x}) - H_z \partial_z \zeta(t, \mathbf{x}) &= \sum_{\Omega_r} R_h^t \xi(t, h) \\ \text{subject to } \zeta(T, \mathbf{x}) &= 0 \end{aligned} \quad (16)$$

Note that the prolongation operator  $R^t$  is the transpose of the operator  $R$ . Note also that because of the self-adjoint expression of the wave-equation (4), the  $\zeta$  adjoint equation is a wave-equation problem similar to the state equation, but subjected to a final condition. This adjoint-equation is classically solved using the change of variable  $\tau = T - t$ , giving

$$\begin{aligned} -S \partial_\tau \zeta(T - \tau, \mathbf{x}) - H_x \partial_x \zeta(T - \tau, \mathbf{x}) - H_z \partial_z \zeta(T - \tau, \mathbf{x}) &= \sum_{\Omega_r} R_h^t (\Delta d(T - \tau, h)) \\ \text{subject to } \zeta(0, \mathbf{x}) &= 0. \end{aligned} \quad (17)$$

Finally, one can observe that the partial differential equation of the adjoint system is very close to the one of the forward wavefield with a specific source term, gathering the residues at receivers positions.

One can deduce the gradient by taking the derivative of the Lagrangian with respect to each model parameter  $m_i$  through the expression

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial m_i} &= \frac{\partial C}{\partial m_i} = \langle \zeta(t, \mathbf{x}) | \frac{\partial S(m)}{\partial m_i} \partial_t u(t, \mathbf{x}) \rangle_T \\ &= \int_T dt \zeta(t, \mathbf{x})^t \frac{\partial S(m)}{\partial m_i} \partial_t u(t, \mathbf{x})\end{aligned}\quad (18)$$

away from active sources. The expression shows that the gradient is the zero-lag cross-correlation between the adjoint wavefield and the derivative of the incident wavefield multiplied by a local scattering operator expressed with parameter derivatives of the compliance matrix. The scattering operator expresses both the illumination of the scattering point by the incident wavefield and the re-emission capacity towards the receiver through the adjoint wavefield.

### Misfit in the domain $(\omega, p)$

The misfit function is expressed in the  $(\omega, p)$  transformed domain through the expression

$$\begin{aligned}C_{\omega,p}(m, \bar{d}_{cal_N}) &= \| |\bar{d}_{obs_N}(\omega, p)| - |\bar{d}_{cal_N}(\omega, p)| \|_{F,P}^2 \\ &= \langle |\bar{d}_{obs_N}(\omega, p)| - |\bar{d}_{cal_N}(\omega, p)| | |\bar{d}_{obs_N}(\omega, p)| - |\bar{d}_{cal_N}(\omega, p)| \rangle_{F,P}\end{aligned}\quad (19)$$

where  $F$  is the space of considered frequencies and  $P$  the space of considered slownesses.

We note the linear move-out (LMO) transform of the wavefield at receivers  $u(t, h)$ , known also as the slant stack, as

$$\tilde{u}(t, p) = \int_{\Omega_r} dh u(t + ph, h),\quad (20)$$

and the Fourier transform of  $\tilde{u}(t, p)$  with the compact notation

$$\bar{u}(\omega, p) = \int_T dt \tilde{u}(t, p) e^{-i\omega t}.\quad (21)$$

We construct the related Lagrangian  $\mathcal{L}(m, u, d_{cal}, d_{cal_N}, \tilde{d}_{cal_N}, \bar{d}_{cal_N}, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5)$  with five adjoint wavefields related to applied transformations, normalisation and differentiation as the following expression

$$\begin{aligned}\mathcal{L} &= C_{\omega,p}(m, \bar{d}_{cal_N}) + \langle \lambda_1(\omega, p) | \bar{d}_{cal_N}(\omega, p) - \int_T dt \tilde{d}_{cal_N}(t, p) e^{-i\omega t} \rangle_{F,P} \\ &+ \langle \lambda_2(t, p) | \tilde{d}_{cal_N}(t, p) - \int_{\Omega_r} dh d_{cal_N}(t + ph, h) \rangle_{T,P} \\ &+ \langle \lambda_3(t, h) | d_{cal_N}(t, h) - \frac{d_{cal}(t, h)}{\|d_{obs}(h)\|} \rangle_{T, \Omega_r} \\ &+ \langle \lambda_4(t, h) | d_{cal}(t, h) - R_h u(t, \mathbf{x}) \rangle_{T, \Omega_r} \\ &+ \langle \lambda_5(t, \mathbf{x}) | S \partial_t u(t, \mathbf{x}) - H_x \partial_x u(t, \mathbf{x}) - H_z \partial_z u(t, \mathbf{x}) - Ss(t, \mathbf{x}) \rangle_T \\ &+ \langle \lambda_5(0, \mathbf{x}) | u(0, \mathbf{x}) \rangle,\end{aligned}\quad (22)$$

We may proceed in the same way as for the  $(t, h)$  domain. The derivative with respect to each adjoint variable  $\lambda_i$  will provide states equations of computation and transformations requested on data. The derivatives with respect to the variable  $\bar{d}_{cal_N}$  provides the definition of the adjoint variable  $\lambda_1$  through the expression

$$\lambda_1(\omega, p) = (\bar{d}_{obs_N}(\omega, p) - \bar{d}_{cal_N}(\omega, p)) \frac{\bar{d}_{cal_N}(\omega, p)^\dagger}{\|\bar{d}_{cal_N}(\omega, p)\|}. \quad (23)$$

Note that  $\bar{d}_{cal_N}(\omega, p)^\dagger / \|\bar{d}_{cal_N}(\omega, p)\|$  represents the opposite of the phase of  $\bar{d}_{cal_N}(\omega, p)$ , meaning that if  $\bar{d}_{cal_N}(\omega, p) = Re^{i\phi}$ , then  $\bar{d}_{cal_N}(\omega, p)^\dagger / \|\bar{d}_{cal_N}(\omega, p)\| = e^{-i\phi}$ .

The derivative with respect to the variable  $\tilde{d}_{cal_N}$  provides the adjoint variables  $\lambda_2$  through the expression

$$\lambda_2(t, p) = \int_F d\omega \lambda_1(\omega, p) e^{i\omega t}, \quad (24)$$

giving the adjoint of the Fourier transform of the  $\lambda_1$  variable, an inverse Fourier transform if  $F$  is considering the whole frequency range. The derivative with respect to the variable  $d_{cal_N}$  requires a change of variable  $t' = t + ph$  and we assume that boundaries have no impact on the interval defined for variable  $t$  and for  $t'$  and we can obtain the following expression

$$\lambda_3(t, h) = \int_P dp \lambda_2(t - ph, p), \quad (25)$$

giving the adjoint of the LMO transform of  $\lambda_2$ , an inverse LMO transform if  $P$  is considering the whole range of slownesses. The derivative with respect to the variable  $d_{cal}$  gives us

$$\lambda_4(t, h) = \frac{\lambda_3(t, h)}{\|d_{obs}(h)\|}. \quad (26)$$

Finally the derivative with respect to the variable  $u(t, \mathbf{x})$  gives the partial differential equation of the adjoint variable  $\lambda_5$

$$\begin{aligned} S\partial_t \lambda_5(t, \mathbf{x}) - H_x \partial_x \lambda_5(t, \mathbf{x}) - H_z \partial_z \lambda_5(t, \mathbf{x}) &= \sum_{\Omega_r} R_h^t \lambda_4(t, h) \\ \text{subject to } \lambda_5(T, \mathbf{x}) &= 0. \end{aligned} \quad (27)$$

One can see that the source of the adjoint wavefield can be recursively deduced from the different transformations/normalisation (23,24,25,26,27), making the gradient computation (18) feasible in the  $(\omega, p)$  domain.

### Misfit in the domain $(\omega, k)$

We can proceed in exactly the same way for the  $(\omega, k)$  domain. The misfit function is expressed in the  $(\omega, k)$  transformed domain through the expression

$$\begin{aligned} C_{\omega, k}(m, \bar{d}_{cal_N}) &= \|\bar{d}_{obs_N}(\omega, k) - \bar{d}_{cal_N}(\omega, k)\|_{F, K}^2 \\ &= \langle |\bar{d}_{obs_N}(\omega, k)| - |\bar{d}_{cal_N}(\omega, k)| \rangle_{F, K} \langle |\bar{d}_{obs_N}(\omega, k)| - |\bar{d}_{cal_N}(\omega, k)| \rangle_{F, K} \end{aligned} \quad (28)$$

where  $F$  is the space of considered frequencies and  $K$  the space of considered wavenumbers. We note the wavenumber Fourier transform of the wavefield at receivers  $u(t, h)$

$$\tilde{u}(t, k) = \int_{\Omega_r} dh u(t, h) e^{-ikh} \quad (29)$$

and the Fourier transform of  $\tilde{u}(t, k)$  with the compact notation

$$\bar{u}(\omega, p) = \int_T dt \tilde{u}(t, k) e^{-i\omega t}. \quad (30)$$

We construct the related Lagrangian  $\mathcal{L}(m, u, d_{cal}, d_{cal_N}, \bar{d}_{cal_N}, \bar{d}_{cal_N}, \mu_1, \mu_2, \mu_3, \mu_4, \mu_5)$  with five adjoint wavefields related to applied transformations, normalisation and differentiation as the following expression

$$\begin{aligned} \mathcal{L} = & C_{\omega, k}(m, \bar{d}_{cal_N}) + \langle \mu_1(\omega, k) | \bar{d}_{cal_N}(\omega, k) - \int_T dt \bar{d}_{cal_N}(t, k) e^{-i\omega t} \rangle_{F, K} \\ & + \langle \mu_2(t, k) | \bar{d}_{cal_N}(t, k) - \int_{\Omega_r} dh d_{cal_N}(t, h) e^{-ikh} \rangle_{T, K} \\ & + \langle \mu_3(t, h) | d_{cal_N}(t, h) - \frac{d_{cal}(t, h)}{\|d_{obs}(h)\|} \rangle_{T, \Omega_r} \\ & + \langle \mu_4(t, h) | d_{cal}(t, h) - R_h u(t, \mathbf{x}) \rangle_{T, \Omega_r} \\ & + \langle \mu_5(t, \mathbf{x}) | S \partial_t u(t, \mathbf{x}) - H_x \partial_x u(t, \mathbf{x}) - H_z \partial_z u(t, \mathbf{x}) - Ss(t, \mathbf{x}) \rangle_T \\ & + \langle \mu_5(0, \mathbf{x}) | u(0, \mathbf{x}) \rangle, \end{aligned} \quad (31)$$

Following the same strategy, the adjoint expressions are successively obtained

$$\mu_1(\omega, k) = \bar{d}_{obs_N}(\omega, k) - \bar{d}_{cal_N}(\omega, k) \frac{\bar{d}_{cal_N}(\omega, k)^\dagger}{\|\bar{d}_{cal_N}(\omega, k)\|}, \quad (32)$$

giving the residues,

$$\mu_2(t, k) = \int_F d\omega \mu_1(\omega, k) e^{i\omega t}, \quad (33)$$

giving an adjoint temporal Fourier transform,

$$\mu_3(t, h) = \int_K dk \mu_2(t, k) e^{ikh}, \quad (34)$$

giving an adjoint spatial Fourier transform,

$$\mu_4(t, h) = \frac{\mu_3(t, h)}{\|d_{obs}(h)\|}, \quad (35)$$

giving the normalisation

$$\begin{aligned} S \partial_t \mu_5(t, \mathbf{x}) - H_x \partial_x \mu_5(t, \mathbf{x}) - H_z \partial_z \mu_5(t, \mathbf{x}) &= \sum_{\Omega_r} R_h^t \mu_4(t, h) \\ \text{subject to } \mu_5(T, \mathbf{x}) &= 0 \end{aligned} \quad (36)$$

giving the partial differential equation of the  $\mu_5$  adjoint variable.

Again, one can see that the source of the adjoint wavefield  $\mu_5$  can be recursively deduced from the different transformations and normalisation (32,33,34,35,36), making the gradient computation feasible (18) in the  $(\omega, k)$  domain.

## ILLUSTRATIVE SYNTHETIC EXAMPLE

We shall consider a simple synthetic example where an oblique velocity gradient is assumed in the upper layer as illustrated in Figure 4a. The shear velocity increases laterally by  $0.5 \text{ s}^{-1}$ , and the vertically by  $5 \text{ s}^{-1}$ . The P-wave velocity is determined by a constant poisson ratio, and the density is homogeneous. The acquisition simulated is a fixed-spread consisting of 24 vertical sources and 48 multi-component receivers evenly spaced on the surface of the 600 m long model. The resulting data in the  $t - x$  domain is shown in Figure 5. We can see quite dispersive surface waves coming from the gradient in the true model. We consider an initial model with two layers of constant velocities: we have eliminated the gradient in the upper layer as can be seen in Figure 4b. Differences with the initial data are quite strong even for the phase evolution of surface waves and one may wonder how the FWI will perform on these datasets. One may also look at the data in the phase velocity spectrum as shown in Figure 6 which is the standard domain when analysing the dispersion of surface waves.

We perform the FWI reconstruction of the shear velocity both in the  $\omega - p$  domain and in the  $\omega - k$  domain. Both approaches extract information from the data in the frequency range we consider. Figures 4c-4d show the reconstructed models, while the fit of the data is shown in Figure 5 viewed in the domain where the data is collected, and in Figure 6 which highlights the dispersion of the surface waves. The depth of penetration of the surface waves controls the depth at which we can reconstruct the velocity which reaches the of depth 20 m of the first layer.

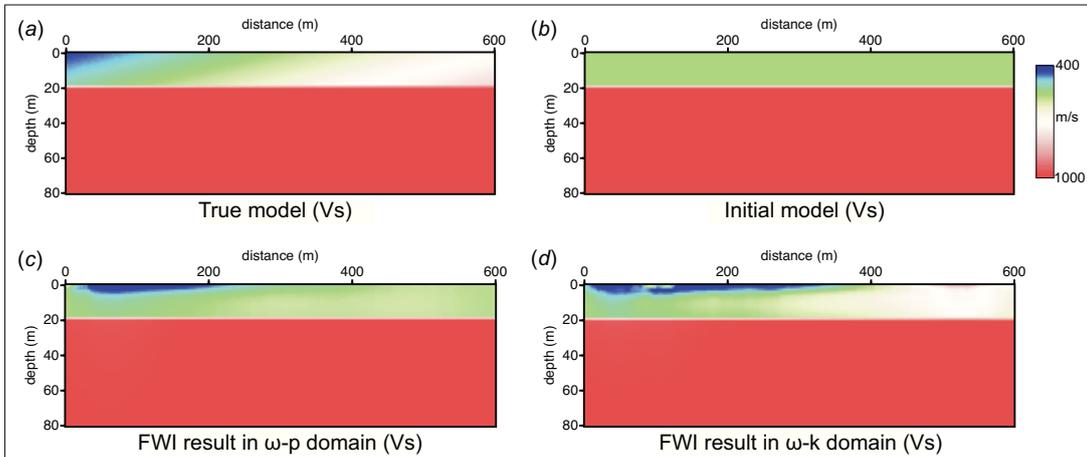


Figure 4: Model description and reconstruction: the top panel represents the true model with an oblique gradient (a) and the initial two-layered model (b). The bottom panel represents the reconstructed model when performing FWI in the  $\omega - p$  domain (c) and the reconstructed model when performing FWI in the  $\omega - k$  domain (d). Please note that in this example the horizontal gradient reconstruction is more successful for FWI in the  $\omega - k$  domain.

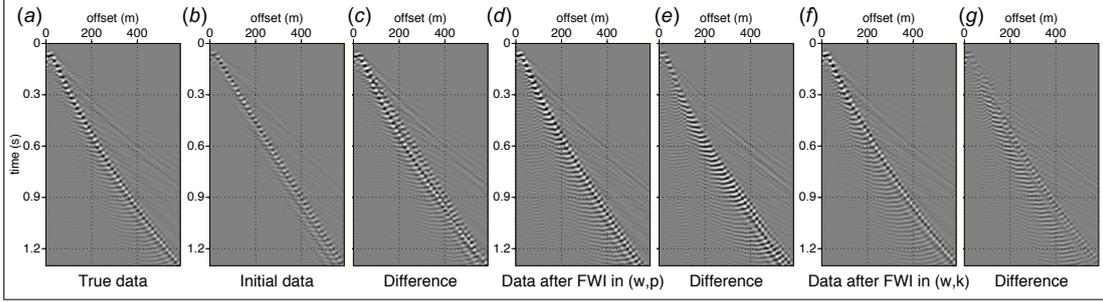


Figure 5: The true data with strong energetic dispersive surface waves (a); the initial data for the two-layered model (b) and the difference to the true data (c); the final data after FWI in the  $\omega - p$  domain (d) and the difference to the true data (e); and the final data after FWI in the  $\omega - k$  domain (f) and the difference to the true data (g). In this example, the FWI performs better in the  $\omega - k$  domain than in the  $\omega - p$  domain.

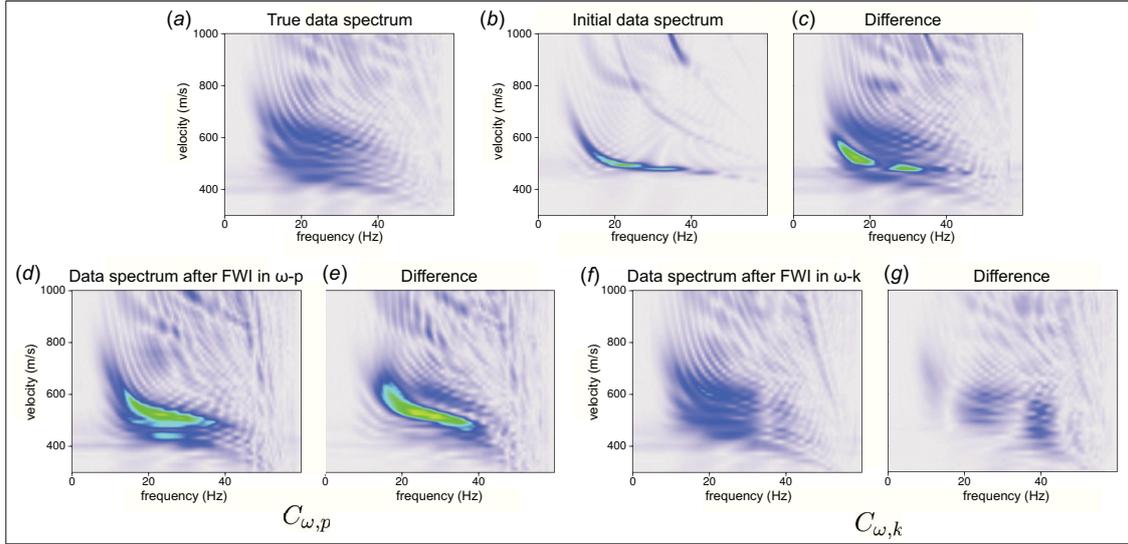


Figure 6: Represented in the top row are the true velocity spectrum (a), the initial spectrum for the simple two-layer model (b) and the velocity spectrum differences (c). The bottom row represents the final velocity spectrum for the FWI result in the  $\omega - p$  domain (d), and the difference with the true velocity spectrum (e); and the final velocity spectrum for the FWI result in the  $\omega - k$  domain (f), and the difference (g). In this example, the FWI performs better in the  $\omega - k$  domain than in the  $\omega - p$  domain.

The inversion result of the shear velocity structure shown in Figure 4c obtained with the expression of the data in the  $(\omega, p)$  domain leads to the data displayed in Figures 5d-5e and in Figures 6d-6e. One can observe that the data misfit is still quite strong. It is also interesting to examine the data in the domain related to the optimization procedure. Figure 7 shows the data in the  $(\omega, p)$  domain and how the inversion attempts to fit some particularly energetic packets leading to some strong misfits in specific parts of the domain.

The inversion result of the shear velocity structure obtained using the expression of the data in the  $(\omega, k)$  domain leads to Figures 5f-5g, where one can note that the data misfit is much better than for the  $(\omega, p)$  domain. The data is also represented in the data domain where the optimization is performed in Figure 8. The inversion performs much better than in other domains as can be seen from the difference between the true data and the recovered data (Figure 8e). The pattern one can see in the final data structure has been successfully updated from the simple initial dispersion curves, and the velocity spectrum has been filled-in.

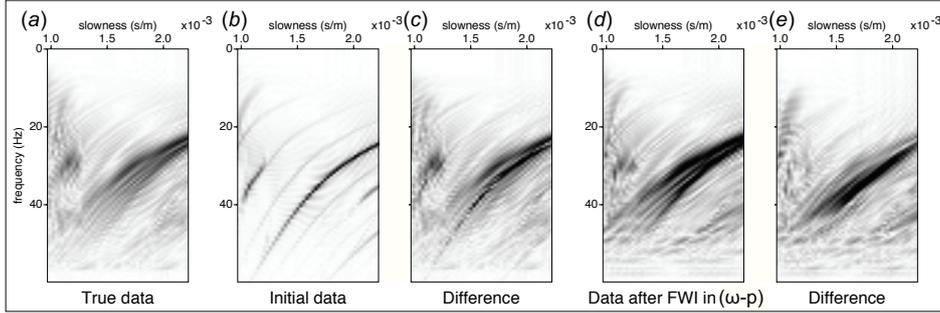


Figure 7: The true data in the  $(\omega, p)$  domain (a); the initial data and the difference to the true data (b,c); the recovered data after FWI in the  $(\omega, p)$  domain and the difference to the true data (d,e). Some particularly energetic packets can be observed.

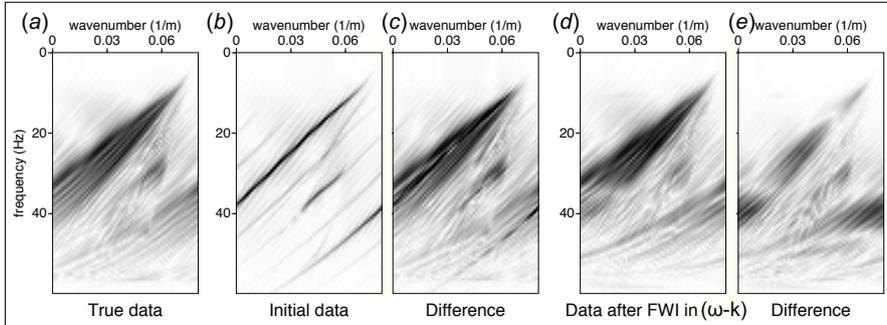


Figure 8: The true data in the  $(\omega, k)$  domain (a); the initial data and the difference to the true data (b,c); the recovered data and the difference to the true data (d,e). One can see a significantly reduced misfit between the initial data difference and the final data difference with less energetic patterns than for the  $(\omega, p)$  domain.

## CONCLUSION AND PERSPECTIVES

We have presented a systematic framework to exploit the full velocity spectrum of surface waves, while explicitly taking into account their dispersive behaviour, as part of a FWI scheme. We have formulated the inversion through a local optimization procedure, considering data transformed to alternative data domains such as the  $(\omega, p)$  domain and the  $(\omega, k)$

domain. We have shown that we are able to extract useful information from surface waves in an automatic way, and without relying on dispersion curve picking or assuming a layered medium. Furthermore, the workflow is still affordable as it relies on a local optimization procedure and avoids a systematic exploration of the model space, which would require a more intensive exploration workflow.

The synthetic example we consider has a relatively dispersive behaviour as the model contains combined vertical and lateral velocity gradients. A systematic investigation of the inversion procedure when considering more complex models will be the purpose of further work. Surface waves may have strong energetic pulses and can generate complex forward and backward scattering in complex subsurface environments. We therefore expect that specific windowing and filtering approaches may be necessary to perform inversion in more complex cases, although such operations should be relaxed in the final steps of the inversion.

## ACKNOWLEDGEMENTS

*The authors would like to thank TOTAL E&P for support and for permission to show these results. This study was partially funded by the SEISCOPE consortium (<http://seiscope2.osug.fr>), sponsored by BP, CGG, CHEVRON, EXXON-MOBIL, JGI, PETROBRAS, SAUDI ARAMCO, SCHLUMBERGER, SHELL, SINOPEC, STATOIL, TOTAL and WOODSIDE. This study was granted access to the HPC resources of the Froggy platform of the CIMENT infrastructure (<https://ciment.ujf-grenoble.fr>), which is supported by the Rhône-Alpes region (GRANT CPER07\_13 CIRA), the OSUG@2020 labex (reference ANR10 LABX56) and the Equip@Meso project (reference ANR-10-EQPX-29-01) of the programme Investissements d’Avenir supervised by the Agence Nationale pour la Recherche, and the HPC resources of CINES/IDRIS under the allocation 046091 made by GENCI. Authors appreciate fruitful discussions with Ludovic Métivier (LJK-CNRS, Université Grenoble Alpes) and Stéphane Operto (GEOAZUR, Université de Nice-Sophia Antipolis).*

## REFERENCES

- Aki, K. and Richards, P. G. (2002). *Quantitative seismology, theory and methods, second edition*. University Science Books, Sausalito, California.
- Bednar, J. B., Shin, C., and Pyun, S. (2007). Comparison of waveform inversion, part 2: phase approach. *Geophysical Prospecting*, 55(4):465–475.
- Bérenger, J.-P. (1994). A perfectly matched layer for absorption of electromagnetic waves. *Journal of Computational Physics*, 114:185–200.
- Bergamo, P., Boiero, D., and Socco, L. V. (2012). Retrieving 2d structures from surface-wave data by means of space-varying spatial windowing. *Geophysics*, 77(4):EN39–EN51.
- Boiero, D. and Socco, L. V. (2010). Retrieving lateral variations from surface wave dispersion curves. *Geophysical Prospecting*, 58:977–996.
- Bozdog, E., Trampert, J., and Tromp, J. (2011). Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements. *Geophysical Journal International*, 185(2):845–870.

- Brossier, R., Operto, S., and Virieux, J. (2009). Seismic imaging of complex on-shore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics*, 74(6):WCC105–WCC118.
- Brossier, R., Operto, S., and Virieux, J. (2014). Velocity model building from seismic reflection data by full waveform inversion. *Geophysical Prospecting*, doi = 10.1049/ip-f-1.1983.0003,.
- Burridge, R. (1996). Elastic waves in anisotropic media. Schlumberger-Doll Research.
- 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.
- Cara, M. (1973). Filtering of dispersed wave trains. *Geophysical Journal of Royal Astronomy Society*, 33:65–80.
- Chavent, G. (2009). *Nonlinear least squares for inverse problems*. Springer Dordrecht Heidelberg London New York.
- Donno, D., Chauris, H., and Calandra, H. (2013). Estimating the background velocity model with the normalized integration method. In *EAGE Technical Program Expanded Abstracts 2013*, page Tu0704.
- Gabriels, P., Snieder, R., and Nolet, G. (1987). In situ measurements of shear-wave velocity in sediments with high-mode rayleigh waves. *Geophysical Prospecting*, 35:187–196.
- Gélis, C., Virieux, J., and Grandjean, G. (2007). 2D elastic waveform inversion using Born and Rytov approximations in the frequency domain. *Geophysical Journal International*, 168:605–633.
- Heisey, J., Stokoe, K., and Meyer, A. (1982). Moduli of pavement systems from spectral analysis of surface waves. *Transportation Research Record*, 652:22–31.
- Herrmann, R. (1987). Computer programs in seismology, vol. iv: Surface waves inversion. Technical report, St. Louis University Press.
- Lerner-Lam, A. and Jordan, T. (1983). Earth structure from fundamental and higher-mode waveform analysis. *Geophysical Journal of Royal Astronomy Society*, 75:759–797.
- Levander, A. R. (1988). Fourth-order finite-difference P-SV seismograms. *Geophysics*, 53(11):1425–1436.
- Louie, J. (2001). Faster, better: Shear-wave velocity to 100 meters depth from refraction microtremor arrays. *Bulletin of Seismological Society of America*, 91:347–364.
- Luo, S. and Sava, P. (2011). A deconvolution-based objective function for wave-equation inversion. *SEG Technical Program Expanded Abstracts*, 30(1):2788–2792.
- Luo, Y. and Schuster, G. T. (1991). Wave-equation travelttime inversion. *Geophysics*, 56(5):645–653.
- Maraschini, M., Ernst, F., Fot, S., and Socco, L. V. (2010). A new misfit function for multimodal inversion of surface waves. *Geophysics*, 75(4):G31–G43.

- Maraschini, M. and Foti, S. (2010). A Monte Carlo multimodal inversion of surface waves. *Geophysical Journal International*, 182:1557–1566.
- Masoni, I., Brossier, R., Virieux, J., and Boelle, J. (2013). Alternative misfit functions for FWI applied to surface waves. In *EAGE Technical Program Expanded Abstracts 2013*, page Th P10 13. EAGE.
- McMechan, G. A. and Yedlin, M. (1981). Analysis of dispersive wave by wavefield transformation. *Geophysics*, 46:869–874.
- Mulder, W. and Plessix, R. E. (2008). Exploring some issues in acoustic full waveform inversion. *Geophysical Prospecting*, 56(6):827–841.
- Nazarian, S. and Stokoe, K. H. (1984). In situ shear wave velocities from spectral analysis of surface waves. *Proc. 8th Conf. on Earthquake Eng., San Francisco*, 3:31–38.
- Nolet, G. (1977). The upper mantle under western europe inferred from the dispersion of rayleigh modes. *Journal of Geophysics*, 43:265–286.
- O’Neil, A. (2004). Full waveform reflectivity for inversion of surface wave dispersion in shallow site investigations. In *17th EEGS Symposium on the Application of Geophysics to Engineering and Environmental Problems*.
- Park, C., Miller, R., Ryden, N., Xia, J., and Ivanov, J. (2005). Combined use of active and passive surface waves. *JEEG*, 10(3):323–334.
- Park, C., Miller, R., and Xia, J. (1999). Multichannel analysis of surface waves. *Geophysics*, 64:800–808.
- Pérez Solano, C., Donno, D., and Chauris, H. (2014). Alternative waveform inversion for surface wave analysis in 2-d media. *Geophysical Journal international*.
- 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., Shin, C., and Hicks, G. J. (1998). Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. *Geophysical Journal International*, 133:341–362.
- Robertsson, J. O. A. (1996). A numerical free-surface condition for elastic/viscoelastic finite-difference modeling in the presence of topography. *Geophysics*, 61:1921–1934.
- Romanowicz, B. (2003). Global mantle tomography: progress status in the past 10 years. *Annual Review of Earth and Planetary Sciences*, 31:303–328.
- Romdhane, A., Grandjean, G., Brossier, R., Réjiba, F., Operto, S., and Virieux, J. (2011). Shallow structures characterization by 2d elastic waveform inversion. *Geophysics*, 76(3):R81.
- Routh, P., Krebs, J., Lazaratos, S., Baumstein, A., Lee, S., Cha, Y. H., Chikichev, I., Downey, N., Hinkley, D., and Anderson, J. (2011). Encoded simultaneous source full-wavefield inversion for spectrally shaped marine streamer data. *SEG Technical Program Expanded Abstracts 2011*, 30(1):2433–2438.

- Ryden, N. and Park, C. B. (2006). Fast simulated annealing inversion of surface waves on pavement using phase-velocity spectra. *Geophysics*, 71.
- Schäfer, M., Groos, L., Forbriger, T., and Bohlen, T. (2013). 2D full waveform inversion of recorded shallow seismic rayleigh waves on a significantly 2D structure. In *in Proceedings of 19th European Meeting of Environmental and Engineering Geophysics, Expanded Abstracts, Bochum, Germany*.
- Shapiro, N., Campillo, M., Paul, A., Singh, S., Jongmans, D., and Sanchez-Sesma, F. (1997). Surface wave propagation across the mexican volcanic belt and origin of the long-period seismic-wave amplification in the valley of mexico. *Geophysical Journal International*, 128:151–166.
- Socco, L. and Strobbia, C. (2004). Surface-wave method for near-surface characterization: a tutorial. *Near Surface Geophysics*, 2:165–185.
- Stokoe, K. H., Wright, S. G., Bay, J. A., and Roesset, J. M. (1994). Characterization of geotechnical sites by sasw method. *Geophysical characterization of sites*, pages 15–25.
- 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 der Lee, S. (2002). High-resolution estimates of lithospheric thickness from missouri to massachusetts, usa. *Earth and Planetary Science Letters*, 203(1):15–23.
- van Leeuwen, T. and Mulder, W. A. (2010). A correlation-based misfit criterion for wave-equation travelttime 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.
- Wathelet, M. (2008). An improved neighborhood algorithm: parameter conditions and dynamic scaling. *Geophysical Research Letters*, 35:L09301.
- Wathelet, M., Jongmans, D., Ohrnberger, M., and Bonnefoy-Claudet, S. (2008). Array performances for ambient vibrations on a shallow structure and consequences over Vs inversion. *Journal of Seismology*, 12:1–19.
- Yuan, Y. O., Simons, F. J., and Bozdog, E. (2014). Full-waveform adjoint tomography in a multiscale perspective. In *SEG Technical Program Expanded Abstracts 2014*, pages 1194–1199.