

A guided tour of multiparameter full waveform inversion with multicomponent data: from theory to practice

S. OPERTO, Y. GHOLAMI, V. PRIeux, and A. RIBODETTI, *Université de Nice Sophia Antipolis*
R. BROSSIER, L. METIVIER, and J. VIRIEUX, *Université de Grenoble*

Building high-resolution models of several physical properties of the subsurface by multiparameter full waveform inversion (FWI) of multicomponent data will be a challenge for seismic imaging for the next decade. The physical properties, which govern propagation of seismic waves in visco-elastic media, are the velocities of the P- and S-waves, density, attenuation, and anisotropic parameters. Updating each property is challenging because several parameters of a different nature can have a coupled effect on the seismic response for a particular propagation regime (from transmission to reflection). This is generally referred to as trade-off or crosstalk between parameters. Moreover, different parameter classes can have different orders of magnitude or physical units and footprints of different strength in the wavefield, which can make the inversion poorly conditioned if it is not properly scaled. These difficulties raise the issue of a suitable parameterization for multiparameter FWI, where the term parameterization must be understood as a set of independent parameter classes that fully describe the subsurface properties. Many combinations of parameters can be viewed and this choice is not neutral as the parameterization controls the trade-off between parameters and the local resolution with which they can be reconstructed. Once this parameterization is selected, the subset of parameter classes in the parameterization, that can be reliably updated during the inversion, must be identified to avoid overparameterization of the optimization problem. The purpose of this tutorial is to provide a comprehensive overview of the promise, pitfalls, and open questions underlying multiparameter FWI. We first review the main FWI ingredient that controls the trade-off between parameters, namely the radiation pattern of the so-called virtual sources, and some tools for analyzing these trade-offs. Then, we present some illustrative examples of multiparameter FWI, which should provide some guidelines to choose a suitable parameterization for FWI in visco-acoustic anisotropic media. We conclude by proposing a data-driven and model-driven workflow for visco-elastic anisotropic FWI of multicomponent marine data, which has been inspired by a real data case study from Valhall.

Introduction

In its conventional form, full waveform inversion (FWI) is a nonlinear data-fitting procedure, which aims to minimize the misfit between recorded and modeled seismic data to iteratively update an initial model of the subsurface (see Virieux and Opertom for a review and relevant references). The aim is to exploit the full information content of the data to derive high-resolution quantitative models of the subsurface. Seismic modeling consists of solving the full two-way wave equation in order to involve in the optimization

process as many wave phenomena as possible (diving waves, pre- and postcritical reflections, multiscattering, free-surface and internal multiples, P-to-S converted waves). Taking into account all these wave phenomena contributes to improve the subsurface illumination provided by incomplete surface acquisition and, hence, the resolution of the imaging. This improved illumination is also conducive to the reconstruction of several physical properties of the subsurface because the imprint of these properties on the seismic response is often better decoupled when several propagation regimes are recorded. Assuming a visco-elastic subsurface, the most general representation of the subsurface involves the density, the attenuation of P- and S- waves and the 21 elastic coefficients of the stiffness tensor. Some simplification in the physics of seismic wave propagation are commonly considered. The most aggressive one considers the Earth as a fluid parameterized by the density (ρ) and the compressional-wave velocity (V_p), or any combination of these two parameters such as impedance ($I_p = \rho V_p$) and V_p . Often, ρ is not involved in the inversion, which is a reasonable assumption when only diving waves and supercritical reflections are involved in the inversion (the so-called early arrival FWI), while it is a more debatable approximation when FWI is applied to seismic reflection data, as density has a significant effect on the reflection amplitudes. A more realistic description of the subsurface introduces elastic effects. In the isotropic approximation, the medium is parameterized by ρ and the Lamé parameters λ and μ . In most geologic environments, anisotropic effects should be taken into account: transverse isotropy and orthorhombic anisotropy introduce, in addition to ρ , five and nine independent parameter classes in the subsurface parameterization, respectively.

Since the re-emergence of FWI technology in exploration geophysics around one decade ago, most developments have crystallized around the building of a V_p model from low frequencies in the acoustic approximation. This model is generally used as an improved background model for wave-equation depth migration. This simplification in the subsurface description results from two main reasons. The first one is of computational nature. Elastic modeling is generally 2-3 orders of magnitude more computationally intensive than the acoustic counterpart because the elastic wave equation requires more wavefield components to be computed and potential low values of the shear velocity (V_s) in the near surface require a refinement of the computational mesh during seismic modeling. The second, more fundamental, reason is related to the ill-posedness of the inverse problem. It is well acknowledged that the misfit function contains many local minima, which result from the incomplete illumination of the subsurface provided by surface acquisitions, the lack of

low frequencies and/or the inaccuracy of the initial models, the incomplete modeling of the wave physics, and noise. This nonlinearity increases when multiple classes of parameters are involved in the inversion because more degrees of freedom are introduced in the model space and parameters of a different nature can have coupled effects of different strength on the kinematic and dynamic attributes of the data (referred to as trade-off or crosstalk between parameters).

Despite these difficulties, multiparameter FWI of multicomponent and streamer data is back at the forefront in an attempt to narrow the gap between quantitative seismic imaging and reservoir characterization and to improve our abilities in terms of fluid characterization. One key issue is to define a suitable parameterization of the subsurface that fully describes the subsurface properties under some judicious physical approximation such that crosstalk between parameters of different nature can be managed. The second key issue is to choose within this parameterization a subset of parameters, which can be reliably reconstructed by FWI, in order to avoid overparameterization of the inversion. Once this choice has been performed, a third issue is to scale these different parameters, such that the numerical optimization is as well conditioned as possible. Some hierarchical data-driven and model-driven strategies can be chosen to reduce as much as possible the nonlinearity of multiparameter FWI and steer the inversion toward the global minimum. Finally, an efficient and versatile multiparameter FWI code must be implemented with adequate optimization and regularization algorithms. In particular, a robust estimation of the Hessian must be performed to scale the gradients and correct them for crosstalk between parameters.

The purpose of this tutorial is to provide a short overview of the fundamental principles of multiparameter FWI to give a flavor of the pitfalls and promises of this technology. We shall show several applications of multiparameter FWI on a multiparameter synthetic model inspired by the geology of Valhall Field to illustrate the discussion. We shall then review the data-driven and model-driven FWI workflow that was applied on real 4C ocean-bottom-cable (OBC) data from Valhall to show the feasibility of visco-elastic anisotropic FWI of multicomponent wide-aperture data.

Multiparameter FWI: Principles

In FWI, the perturbation model Δm , which is added to the starting model at each nonlinear iteration, has the following expression

$$\Delta m = -\gamma H^{-1} \nabla C, \tag{1}$$

where ∇C is the gradient of the misfit function, H is the Hessian and γ is the step length. In multiparameter FWI, Δm and ∇C are vectors of dimension $N_x \times N_m$, where N_m denotes the number of parameter classes estimated during the inversion (for example, wave velocity, density, attenuation) and N_x is the number of nodes in the inversion mesh. In the following, a model parameter will be denoted by $m_{i,j}$ where i and j denote the position of the model parameter in the subsurface and the parameter class, respectively.

Multiparameter gradient and radiation pattern of virtual sources. When the functional to be minimized is the least-squares norm of the data residuals, namely, the difference between the calculated and the recorded data, $\Delta d = d_c - d_o$, the gradient at the position of the model parameter $m_{i,j}$ is simply the sum over receivers of the zero-lag correlation between the data residuals and the partial derivatives of the modeled wavefield extracted at the receiver positions:

$$\nabla C_{m_{i,j}} = \left(\frac{\partial d_c}{\partial m_{i,j}} \right)^T \Delta d^* = \sum_r \sum_t \frac{\partial d_{c,r,t}}{\partial m_{i,j}} \Delta d_{r,t}, \tag{2}$$

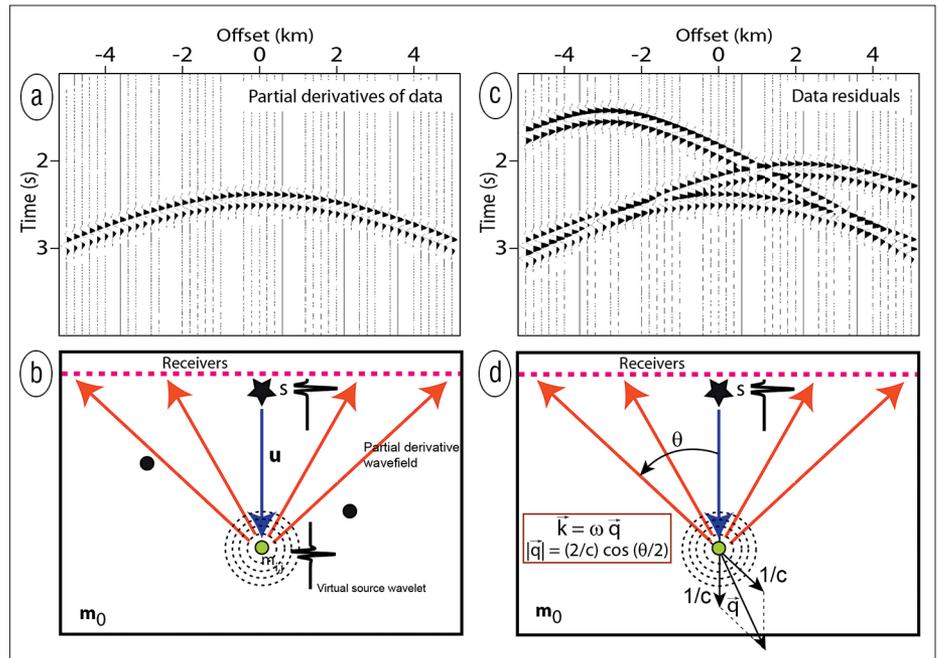


Figure 1. Physical interpretation of the gradient in the framework of the single-scattering Born approximation. The true medium contains three diffractor points—filled circles in (b)—and the background model m_0 is homogeneous. The data residuals recorded at receivers—thick dashed line in (b)—correspond to the wave velocity scattered by the three diffractor points (c). The partial derivative of the wavefield with respect to the model parameter located at the position of the middle point diffractor—green circle in (b)—corresponds to the wavefield scattered by this point diffractor normalized by the value of the wavespeed perturbation. The radiation pattern of the virtual source located at the point diffractor is illustrated by the thin dashed lines (here an isotropic radiation pattern) in (b). The zero-lag correlation between the data residuals (c) and the partial derivative wavefield taken at the receiver positions (a) provides the unscaled contribution of the source s at the position $m_{i,j}$. (d) The wavenumber vector \mathbf{k} injected in the gradient is labeled. Its orientation is given by the sum of the source and receiver slowness vectors and its modulus is controlled by the scattering angle θ , the local waves velocity c and the angular frequency ω .

where T denotes the transpose of a matrix, r is the receiver index and t is the time index (time can be indifferently replaced by frequency depending whether FWI is implemented in the time domain or in the frequency domain). In the framework of diffraction tomography, the data residuals can be interpreted as the wavefield scattered by all the missing heterogeneities in the initial model (Figure 1c), while the partial derivatives of the modeled wavefield represent the wavefield scattered by a small perturbation of the parameter m_{ij} , all the other parameters, $m_{k,l}, k \neq i, l \neq j$, being kept fixed (Figure 1b). The cross-correlation in Equation 2 aims to pick, in the data residuals, the piece of information that must be transformed into a model perturbation at the position i . However, the gradient does not provide the correct amplitude of the model perturbation, first because the gradient does not have the units of the model perturbation and second because the gradient associated with one parameter class can be affected by crosstalk from the other parameter classes. The Hessian and γ in Equation 1 aim to assign the correct units to the model perturbations and to remove these cross-talk effects.

It is worth recalling that the scattering angle θ at the diffractor point controls, together with the local wavelength λ_0 , the modulus of the local wavenumber vector k that is constrained by the imaging procedure (Equation 27 and Figure 1d in Fergues and Lambare, 1997). The wavenumber vector describes the spatial-frequency component that is locally injected at the position of the diffractor point along the spatial direction defined by the sum of the source and receiver slowness vectors. The wide θ associated with the forward-scattering regime (transmission paths associated with diving waves and supercritical reflections) contribute to imaging the long wavelengths of the subsurface, while the small θ associated with the backward-scattering regime (short-spread reflected waves) contribute to imaging the short wavelengths of the subsurface.

The point diffractor at the position of the model parameter m_{ij} behaves as a secondary virtual source (Figure 1b), whose radiation (or scattering) pattern depends on the subsurface parameterization and on the parameter class in this parameterization. The radiation pattern of the virtual source governs the amplitude variations of the partial derivative wavefield as a function of θ (it

can also affect the phase as in attenuation imaging performed with complex-valued wave velocity), and hence apply some filtering on the wavenumber spectrum of the subsurface, in addition to the filtering resulting from the limited bandwidth of the source and the limited θ illumination provided by the source-receiver geometry. The differences between the gradients associated with different parameter classes result only from their different radiation patterns. Therefore, if virtual sources associated with two different parameter classes have radiation patterns which significantly overlap over a significant range of θ , some trade-off (or crosstalk) between these parameters impacts their gradients.

The radiation patterns of virtual sources are illustrated numerically in Figure 2. We consider a homogeneous acoustic background medium parameterized by two classes of parameter: (V_p, ρ) and (I_p, V_p) , respectively. For the two parameterizations, we compute the wavefield scattered by one parameter perturbation, keeping the other parameter fixed at

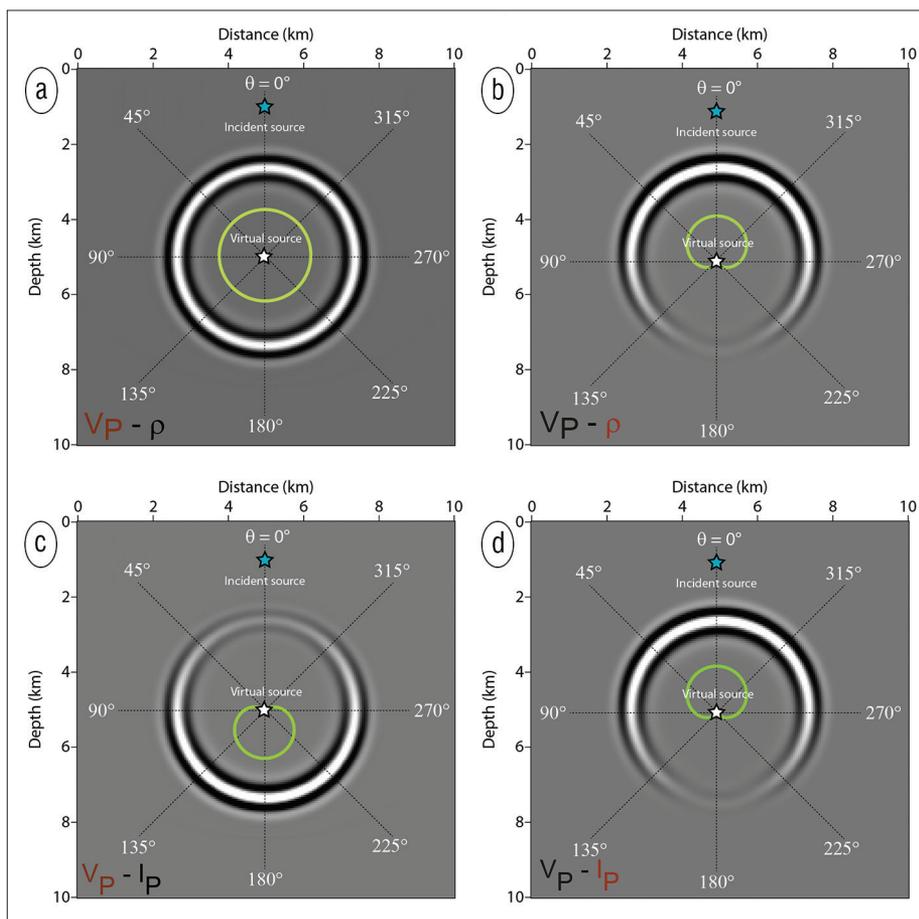


Figure 2. Radiation patterns of the virtual sources of the partial derivative wavefields for an acoustic medium parameterized by (V_p, ρ) in (a) and (b) and (V_p, I_p) in (c) and (d). The figure shows a snapshot of a wavefield that propagates in a homogeneous background model and which is scattered by a point perturbation located in the middle of the grid. Only one model parameter of the subsurface parameterization is perturbed, the other one being kept fixed. (a) V_p perturbation, with ρ fixed. (b) ρ perturbation, with V_p fixed. (c) V_p perturbation, with I_p fixed. (d) I_p perturbation, with V_p fixed. The scattering angle is labeled by θ . The virtual source at the point diffractor is denoted by the white-filled star. The incident source, at the vertex of the model perturbation, is denoted by the blue-filled star. The amplitude variations of the wavefront around the point diffractor result from the radiation pattern of the virtual sources. The analytical radiation patterns derived in the framework of the ray + Born approximation are superimposed (green curves).

its background value. The amplitude variations around the diffractor point of the scattered wavefield are representative of the radiation pattern of the virtual source.

The wavefield scattered by a V_p perturbation in the (V_p, ρ) parameterization has an isotropic pattern; the corresponding virtual source behaves as a pressure source (Figure 2a). In contrast, the amplitudes of the wavefield scattered by the ρ perturbation decrease progressively from the small θ to the wider ones (Figure 2b); the corresponding virtual source behaves as a vertical force pointing upward in the direction of the excitation source. These two radiation patterns are consistent with the fact that V_p governs both the traveltimes and amplitudes of waves at all θ , while ρ has a dominant influence on the amplitude of the precritical reflections at small θ . A broad spectrum of V_p can be reconstructed from broadband sources and wide-azimuth/wide-aperture acquisition geometries because the virtual pressure source does not apply any filtering on the V_p gradient, while only the short-to-intermediate wavelengths of ρ can be retrieved because of the high-pass filtering applied on the ρ gradient by the vertical-force virtual source.

When we move from the (V_p, ρ) parameterization to the (I_p, V_p) parameterization, the radiation pattern of V_p changes because the influence of V_p and ρ on the seismic response at small-to-intermediate θ is now affected by I_p ; in the (I_p, V_p) parameterization, the V_p parameter generates scattering at wide-to-intermediate θ only (Figure 2c), while the impedance parameter generates scattering at small-to-intermediate θ (Figure 2d). This parameterization honors the scale uncoupling between the velocity macromodel building task and the

reflectivity imaging by migration, which are the two main ingredients of the conventional seismic reflection imaging workflow.

The analytical expression of the radiation patterns can be derived in the framework of the asymptotic ray+Born waveform inversion (Equations B-7 and B-8 in Forgues and Labare for the previous example). These analytical radiation patterns are superimposed on the snapshot of the partial derivative wavefields in Figure 2 to check the consistency between the amplitude variations of the partial derivative wavefields and the analytical radiation patterns of their virtual sources.

The fact that the radiation pattern of V_p changes, when the subsurface parameterization is modified, reflects how the trade-off between parameters changes as these parameters are combined differently in the subsurface parameterization. This is shown, from the mathematical viewpoint, by the chain rule of derivatives.

Role of the multiparameter Hessian. In Equation 1, the Hessian matrix or normal operator H is a square matrix of order $N_x \times N_m$. If the model parameters are first sorted as a function of their position in the mesh and second as a function of the parameter class, H is an $N_m \times N_m$ block band-diagonal matrix. An example of multiparameter Hessian is shown in Figure 3a for a VTI medium in the acoustic approximation, which is parameterized by the vertical wave velocity (V_{p0}) and the Thomsen parameters δ and ϵ . According to the normal-equation system, $H\Delta m = -\nabla C$, and the block band-diagonal pattern of H , it is worth noting that the gradient of the misfit function with respect to one parameter class is a weighted sum of the model perturbations associated with

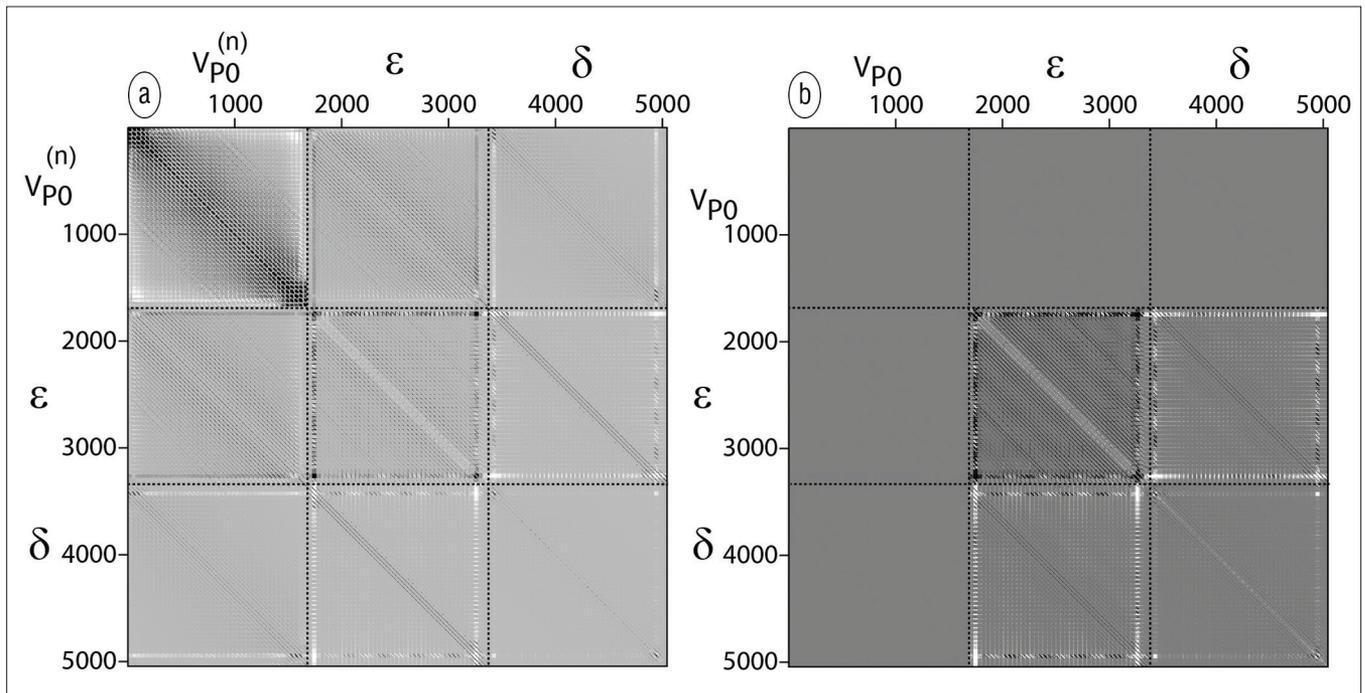


Figure 3. Hessian for a VTI medium in the acoustic approximation. The subsurface medium is parameterized by V_{p0} , δ and ϵ . The Hessian is a block matrix. The diagonal blocks describe the correlation between partial derivative wavefields with respect to model parameters belonging to the same class. The off-diagonal blocks describe the trade-off between parameters belonging to different classes averaged over the full data set. In (a), the Hessian has been computed when the vertical wave velocity is adimensionalized, unlike in (b). The Hessian in (a) fairly reflects the influence of the model parameters in the seismic response, unlike in (b).

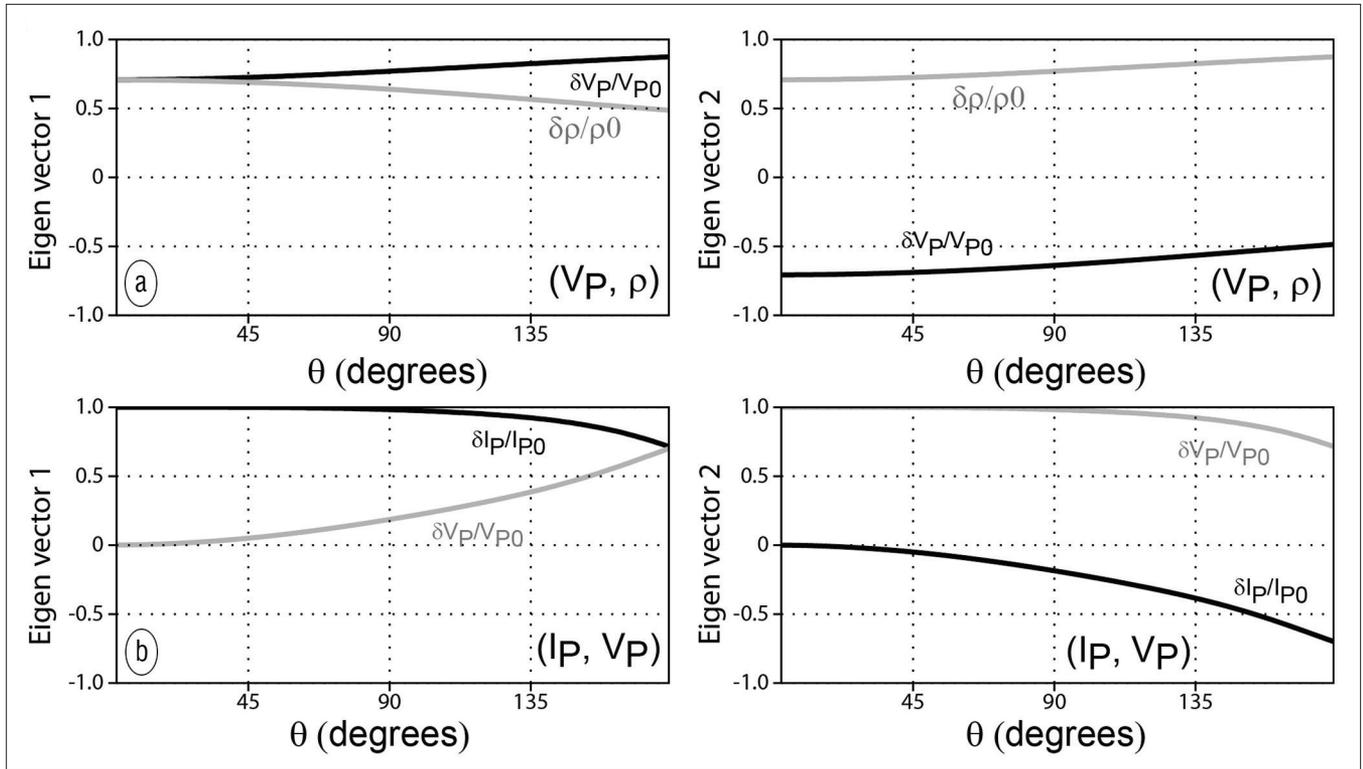


Figure 4. Eigenvalue analysis of the asymptotic Hessian for acoustic media. (a) Normalized eigenvector versus scattering angle for the (V_p, ρ) parameterization (left = first eigenvector; right = second eigenvector). The two components of the eigenvector associated with parameters V_p and ρ are plotted. (b) Same as (a) for the (I_p, V_p) parameterization (from Fougues and Lambare).

each parameter class. This is the mathematical translation of the trade-off between model parameters, which has been analyzed in the previous section based on physical grounds. No parameter trade-off would exist if the off-diagonal blocks of the Hessian were zero.

The linear term of the Hessian matrix is formed by the sum over the data of the zero-lag correlation of the partial derivative wavefields taken at the receiver positions:

$$H_{(i,j),(k,l)} = \sum_i \sum_r \left(\frac{\partial d_{c_{t,r}}}{\partial m_{i,j}} \right) \left(\frac{\partial d_{c_{t,r}}}{\partial m_{k,l}} \right), \quad (3)$$

where (i,j) and (k,l) denote the model parameters of kind j located at position i and the model parameter of kind l located at position k , respectively. When $i = k$ and $j = l$, the (diagonal) Hessian coefficients correspond to the autocorrelation of the partial derivative wavefield and, hence, correct the gradient for wave-amplitude effects. When $i \neq k$ and $j = l$, the Hessian coefficients correspond to the correlations between partial derivative wavefields associated with nearby parameters of the same nature. These correlations describe the limited bandwidth effects, which hamper the spatial resolution of FWI and which result from the limited bandwidth of the source, the limited source-receiver geometry, and the directivity of the virtual sources. When $j \neq l$, the Hessian coefficients represent the correlations between partial derivative wavefields associated with parameters of different nature and, hence, represent the trade-off between these

parameters. These trade-off effects vary with θ according to the radiation pattern of the virtual sources. As the Hessian coefficients are computed by summation over the source-receiver pairs and over time, the above-mentioned effects (geometrical spreading, limited bandwidth effects, crosstalk between parameters) are corrected in an average sense if the full data set is processed in one go. Some weighting operator (defined as a function of offset, time or θ if this information is accessible) can be introduced in the data misfit function to balance the contribution of each arrival in the Hessian. Alternatively, multiscale approaches of FWI can be designed to hierarchically process different subsets of the data, which can be sorted as a function of the propagation regime. For example, one may start inverting the early arriving phases to update the long wavelengths of the parameters that have an influence at wide θ , and invert later arrivals in a second step to update the short wavelengths of those parameters, which have a dominant influence at short θ . A dedicated parameterization might be chosen during each hierarchical step to minimize trade-off effects and the number of parameter classes involved in the inversion.

Eigenvalue analysis of the asymptotic Hessian. Eigenvalue analysis of the Hessian in the high-frequency regime provides invaluable insights on the trade-off between parameters during waveform inversion. The asymptotic ray + Born waveform inversion relies on similar principles as FWI in the sense that the gradient of the misfit function relies on the single-scattering Born approximation (Figure 1). However, the Green's functions are computed with ray theory rather than

with full-wave methods. This allows one to explicitly introduce some key attributes such as traveltimes, amplitudes, and angles in the kernels of the gradient and of the Hessian, and to derive the analytical expression of the radiation patterns as a function of θ . Moreover, the high-frequency approximation allows diagonalization of the Hessian through a judicious weighting of the misfit function (e.g., Forgues and Lambare).

In acoustic and isotropic elastic media, the scattering kernel in the asymptotic Hessian reduces to a 2×2 and 3×3 square symmetric matrix, on which an eigenvalue decomposition can be applied. For a two-parameter problem, the conditioning number of this matrix (the ratio between the two eigenvalues) identifies whether the information carried out by the second eigenvector is significant for a given θ . If this information is not significant (below the noise level), only one parameter can be reconstructed. For each eigenvalue, the eigenvector components (each component is associated with a parameter class) plotted as a function of θ provide insight into the relative weight of each parameter class in the subsurface reconstruction. If the two components of the first eigenvector have significant values over the same range of θ , trade-off between the two parameters is expected. Diagonalization of the scattering kernel leads to a parameterization that cancels out the contribution of the off-diagonal blocks of the multiparameter Hessian and, hence, remove parameter trade-off. The drawback is that this parameterization depends on θ ; that implies that the parameterization should be ideally adapted to each seismic arrival and each subsurface position during inversion.

As an illustrative example, the eigenvectors of the scattering kernels are shown in Figure 4 for the two parameterizations (V_p, ρ) and (I_p, V_p) . For the (V_p, ρ) parameterization, the two components of the first eigenvector have the same value for zero-offset reflection ($\theta=0^\circ$), and the relative weight of the component associated with V_p increases with θ as the one of ρ decreases (Figure 4a). This trend reflects that the trade-off between V_p and ρ decreases as the influence of ρ on the data decreases when θ increases. A new parameterization, based on a weighted sum of the two components of the first eigenvector and a weighted difference of the two components of the second eigenvector, would be necessary to uncouple these

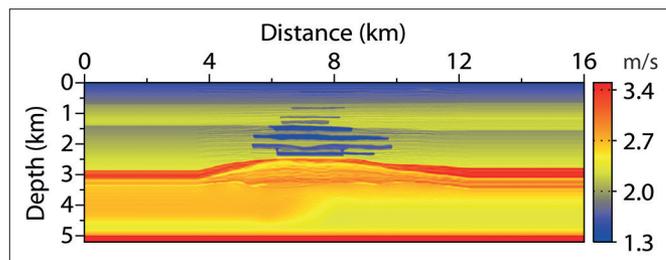


Figure 5. Synthetic Valhall velocity model. The velocity structure is characterized by a pile of low-velocity gas layers above the reservoir level at 2.5 km in depth. The water depth is 50 m. Anisotropy reaches 15 in Valhall. The acquisition geometry represents an ocean-bottom cable survey, with a maximum offset of 16 km. For such maximum offset, only the first 1.5 km in depth of the subsurface, above the gas layers, are sampled by both diving waves and short-spread reflections. Short-spread reflections mainly sample the subsurface at the reservoir level.

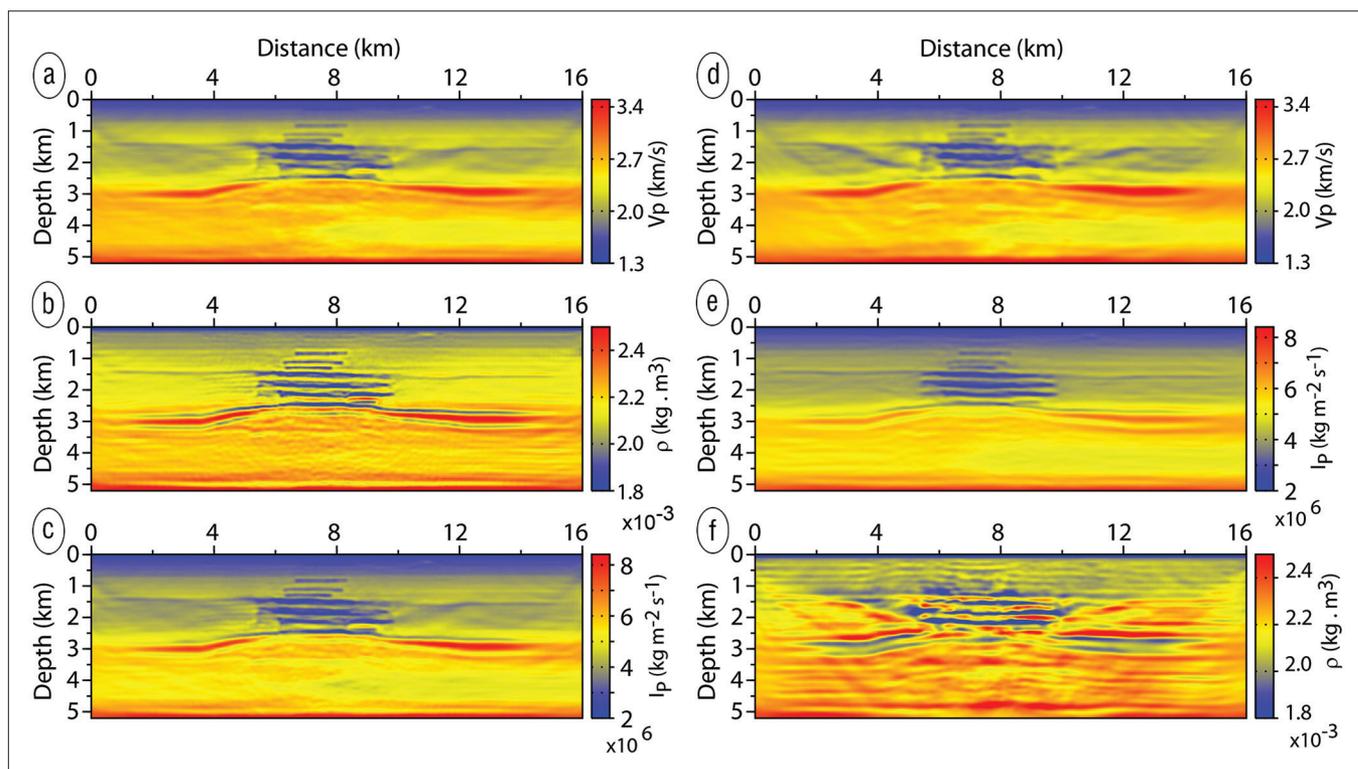


Figure 6. Synthetic Valhall model. Acoustic FWI. Joint update of V_p (a) and ρ (b) for the (V_p, ρ) parameterization. (c) Impedance model inferred from the multiplication of the models shown in (a) and (b). Joint update of V_p (d) and I_p (e) for the (I_p, V_p) parameterization. (f) Density model inferred from the division of the models shown in (e) and (d). The quality of the impedance model in (c) is higher than the one in (e) (from Prieux et al.).

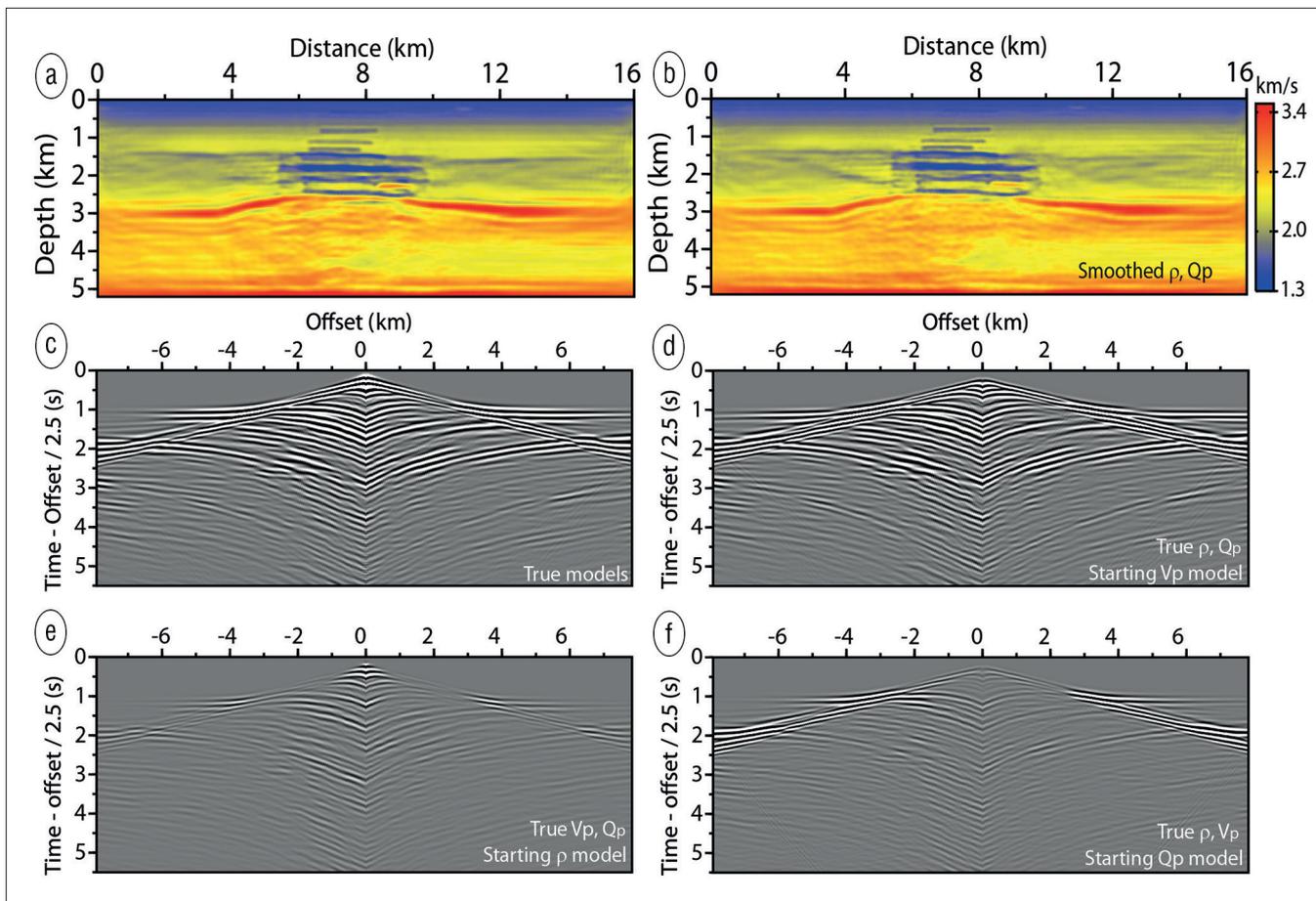


Figure 7. Synthetic Valhall model showing influence of ρ and Q_p on the seismic response. Monoparameter FWI for V_p keeping fixed the ρ and Q_p models. (a) The ρ and Q_p background models are the true models (the residuals result only from missing V_p perturbations). (b) The ρ and Q_p background models are smooth models (the residuals result from missing V_p , ρ and Q_p perturbations). The two V_p FWI models are of similar quality. (c) Synthetic seismograms computed in the true models. (d) Differences between the seismograms shown in (a) and seismograms computed in the true ρ model, in the true Q_p model and in the smooth V_p model. (e) Same as (d) but seismograms are computed in the true V_p model, in the true Q_p models and in the smooth ρ model. (f) Same as (e) but seismograms are computed in the true V_p model, in the true ρ model and in the smooth Q_p model (from Prioux et al.).

two parameters. In the (I_p, V_p) parameterization, the relative weight of V_p is zero for zero-offset reflection, while the one for I_p is maximum, and the relative weight of V_p increases with θ as the one of I_p decreases (Figure 4b). This opposite behavior highlights the natural uncoupling between these two parameters, a distinct advantage which is balanced by the more limited resolution with which V_p can be reconstructed when the (I_p, V_p) parameterization is chosen. The inferences drawn from this eigenanalysis are indeed consistent with those revealed by the analysis of the radiation patterns (Figure 2). Note, however, that a reliable sensitivity analysis based on the eigenvalue analysis of the Hessian requires a suitable scaling (or, adimensionalization) of the model parameters to remove from the partial derivative wavefields the bias introduced by the unit of the parameters.

Which parameterization for multiparameter FWI?

Now that we have stated the crux of multiparameter FWI, we illustrate the influence of the subsurface parameterization on multiparameter FWI (Figure 5). Frequency-domain FWI is performed in the frequency band 3.5–6.7 Hz. More

details on the multiparameter models and the FWI setups are provided in Prioux et al. (2011), Gholami et al. (2013), and Prioux et al. (2013).

Can density or impedance attenuation be retrieved? We first perform multiparameter acoustic FWI with two parameterizations, (V_p, ρ) and (I_p, V_p) , whose pros and cons are discussed. The starting models are obtained by smoothing the true models and the two inversions were initiated from the same data residuals. Results of the multiparameter FWI are shown in Figure 6 for the two parameterizations. Our conclusion is that the (V_p, ρ) parameterization is the most suitable because it favors a broadband (and, hence, more accurate) reconstruction of the dominant parameter V_p (compare Figures 6a and 6d), although we notice crosstalk artifacts in the low-velocity gas layers: underestimated values of V_p correlate with overestimated values of ρ (Figures 4a–4c in Prioux et al.). However, multiplication of the V_p and ρ models after inversion provides a reliable I_p model, suggesting that the trade-off artifacts were a posteriori removed (Figure 6c). The a posteriori building of the I_p model is sound because both V_p and ρ models have a similar wavenumber content; FWI

mainly updated the small-to-intermediate wavelengths of V_p and ρ because the initial models already contains the long wavelengths, which cannot be updated in the ρ model. The (I_p, V_p) parameterization leads to satisfying I_p and V_p models (Figure 6(d-e)), which are, however, of more limited accuracy than those inferred from the (V_p, ρ) parameterization (Prioux et al.). Moreover, building a ρ model from the V_p and I_p models didn't provide a reliable model (Figure 6f) because the V_p and I_p models do not have the same wavenumber content (the high wavenumbers are updated only in the I_p model) and no correlation is expected between the errors of the two models.

If narrow-azimuth reflection data are considered, the two parameterizations might be combined in a two-step workflow. New FWI technologies are currently being developed for reflection data, where updating the velocity macromodel is alternated with updating the reflectivity in the iterations of the FWI workflow (Xu et al., 2012). The governing idea is to exploit prior constraints on the reflectivity in the velocity model building task to emphasize the transmitted wavepaths between reflectors and sources and between reflectors and receivers in the sensitivity kernel of the FWI and, hence, steer the inversion toward the update of the long wavelengths. Once the macromodel is updated, a new reflectivity can be generated by migration or by one iteration of conventional FWI (i.e., the FWI formulation that emphasizes the migration isochrones in its sensitivity kernel). As the seismic reflection FWI workflow honors the scale uncoupling between the macromodel building and the reflectivity imaging, one might view to update the velocity macro model using the (V_p, ρ) parameterization to take advantage of the broad radiation pattern of V_p , while the I_p parameter can be updated by conventional monoparameter FWI with the (I_p, V_p) parameterization for a more reliable accounting for the reflection amplitudes without trade-off issues.

Attenuation is another parameter that can have a significant influence on the data, although it is marginally considered during FWI. By influence of a parameter on the data we mean the wavefield perturbation that would be generated by a representative perturbation of the parameter (i.e., a perturbation of amplitude and size that is expected to be lacking in the initial model). If the amplitudes of the wavefield perturbations are

lower than the noise level, FWI will fail to update this parameter.

The relative influence of velocity, density and attenuation on the data is illustrated by a series of wave simulations in the visco-acoustic Valhall model parameterized by V_p , ρ and Q_p , the quality factor (Figure 7). The smallest value of Q_p is 60 in the shallow sediments and in the gas layers. Smoothed versions of the true V_p and ρ models were generated to mimic initial FWI models, while the background Q_p model is homogeneous with a value of 150 (Prioux et al.). The influence of each parameter class on the data residuals is shown by the difference between the seismograms computed in the true model (Figure 7c) and the seismograms computed in the smooth model of the assessed parameter and in the true models of the remaining parameters. As expected, V_p shows the strongest influence over the full range of offsets (Figure 7d). The density and Q_p have an influence on the data of similar strength, except that the influence of ρ is dominant at short offsets (Figure 7e), where short-spread reflections

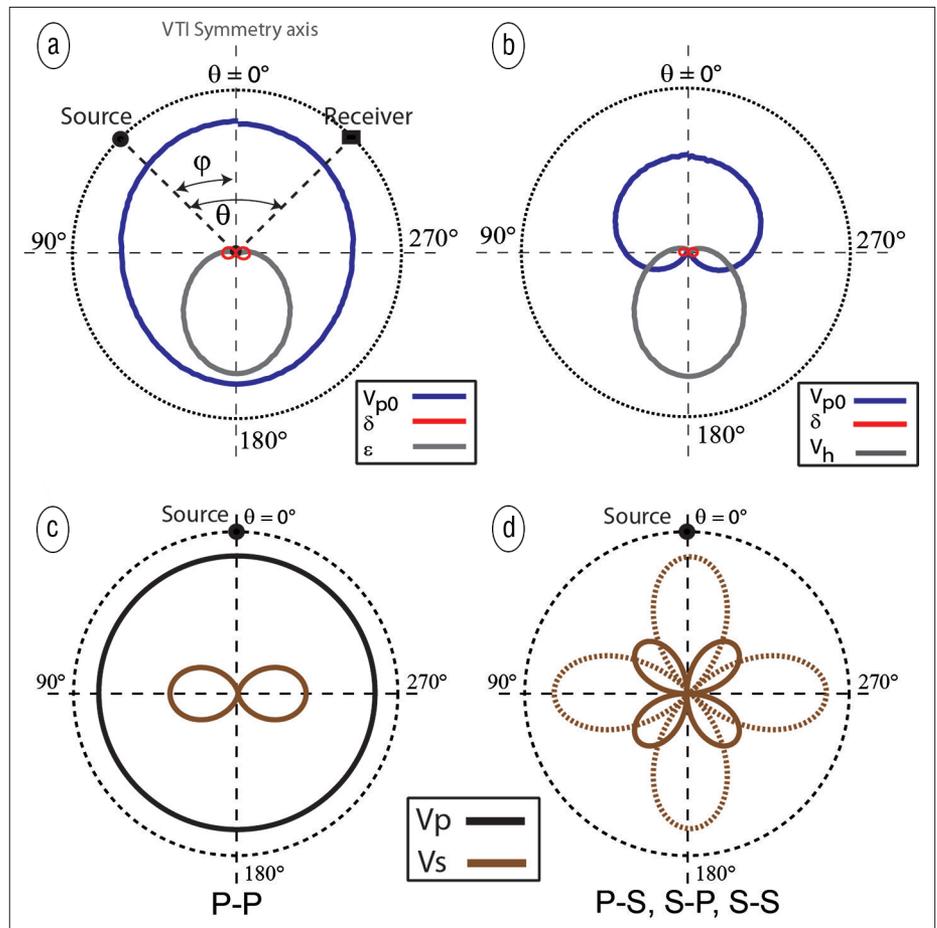


Figure 8. Radiation patterns in VTI acoustic media parameterized by (a) $(V_{p0}, \delta, \epsilon)$ and (b) (V_{p0}, δ, V_h) . The radiation patterns were computed for the source-receiver configuration representative of a surface acquisition: the source and receiver are moving in opposite direction on the dot circle from the vertical symmetry axis (Gholami et al.). Radiation patterns of V_p and V_s in elastic isotropic media for a (V_p, V_s, ρ) parameterization (from Forgues and Lambare). (c) PP scattering mode. The radiation patterns of V_p and V_s are plotted with black and brown lines, respectively. (d) PS, SP and SS scattering modes. Only V_s generates scattering for these modes. The solid line corresponds to the PS and SP modes, while the dashed line corresponds to the SS mode.

are recorded, while that of Q_p is dominant at long offsets, where diving waves and supercritical reflections are recorded after many propagated wavelengths (Figure 7f). We perform monoparameter FWI for V_p using the true and the smooth ρ and Q_p models as background models, respectively. Although the influence of ρ and Q_p is not negligible, the two FWI V_p models are of similar quality (Figure 7a and 7b), that shows that FWI is mainly driven by phase effects at least at low frequencies and that monoparameter FWI for velocity can provide reliable results by keeping secondary ρ and Q_p parameters fixed. However, updating these secondary parameters is possible because their influence in the data is higher than the noise level both in marine and land environments. Recent examples of attenuation reconstruction by FWI in marine and land environments are shown in Malinowski et al. (2011) and Prioux et al. The results of Malinowski et al. and Prioux et al. confirm that the reconstruction of Q_p is more reliable in the shallow part of the subsurface, which is illuminated by diving waves.

Anisotropic FWI. Although anisotropy is often taken into account in seismic modeling during FWI, anisotropic parameters other than one wave velocity are generally kept fixed during inversion. We now discuss which parameterization is suitable for imaging vertical transverse isotropic (VTI) media in the acoustic approximation, and which parameters can be reliably updated (Plessix and Cao, 2011; Gholami et al., 2013.). We investigate two kinds of parameterization: in the first one, the subsurface is parameterized by one wavespeed (here, V_{p0}), δ and ε , while in the second one, ε is replaced by the horizontal velocity: $V_h = V_{p0}\sqrt{1+2\varepsilon}$.

The radiation patterns of the three parameter classes are shown in Figure 8a and 8b for the two parameterizations. In the $(V_{p0}, \delta, \varepsilon)$ parameterization, V_{p0} has an isotropic radiation pattern, while ε generates scattering at large θ ; only the long-to-intermediate wavelengths of ε can be updated and crosstalk between V_{p0} and ε is expected at large θ (Figure 8a).

The parameter δ generates a small amount of scattering at intermediate θ that makes it challenging to update during FWI challenging (Plessix and Cao, 2011). In the (V_{p0}, δ, V_h) parameterization, the influence of V_{p0} is now limited to small-to-intermediate θ because V_h gathers the influence of V_{p0} and ε on the seismic response at large θ ; crosstalk between V_{p0} and V_h is expected at intermediate θ only (Figure 8b).

We first compare the results of monoparameter FWI for V_{p0} , when the $(V_{p0}, \delta, \varepsilon)$ and the (V_{p0}, δ, V_h) parameterizations are used (Figure 9). The two inversions started from the same data residuals and the smooth background models of δ and ε or V_h are kept fixed during iterations. The V_{p0} model obtained with the $(V_{p0}, \delta, \varepsilon)$ parameterization is reliable (Figure 9a), while the one obtained with the (V_{p0}, δ, V_h) parameterization shows overestimated perturbations and mispositioning of reflectors in depth (Figure 9b). The first inversion succeeds because the radiation pattern of V_{p0} spans the full range of θ and, hence, a broad band of wavenumbers can be updated. Moreover, the background model of ε describes the long wavelengths of the subsurface sufficiently accurately to make the footprint of this parameter in the data residuals negligible. In contrast, the second inversion fails because V_h is kept fixed and V_{p0} has now a radiation pattern that is limited to the small-to-intermediate θ . This prevents the update of the long wavelengths, causing the mispositioning in depth of the reflectors. Second, the footprint of V_h in the data residuals at intermediate θ , where the radiation patterns of V_{p0} and V_h overlap, is not negligible. This footprint has been interpreted by the inversion as missing V_{p0} perturbations due to crosstalk between the two wavespeeds, leading to strongly overestimated V_{p0} perturbations. This interpretation is supported by the success of multiparameter FWI with the (V_{p0}, δ, V_h) parameterization for the joint update of V_{p0} and V_h (Figure 9c and 9d). The V_h model is smoother than the V_{p0} model, which is consistent with their respective radiation patterns (Figure 8b). The V_{p0} model inferred from the (V_{p0}, δ, V_h) parameterization

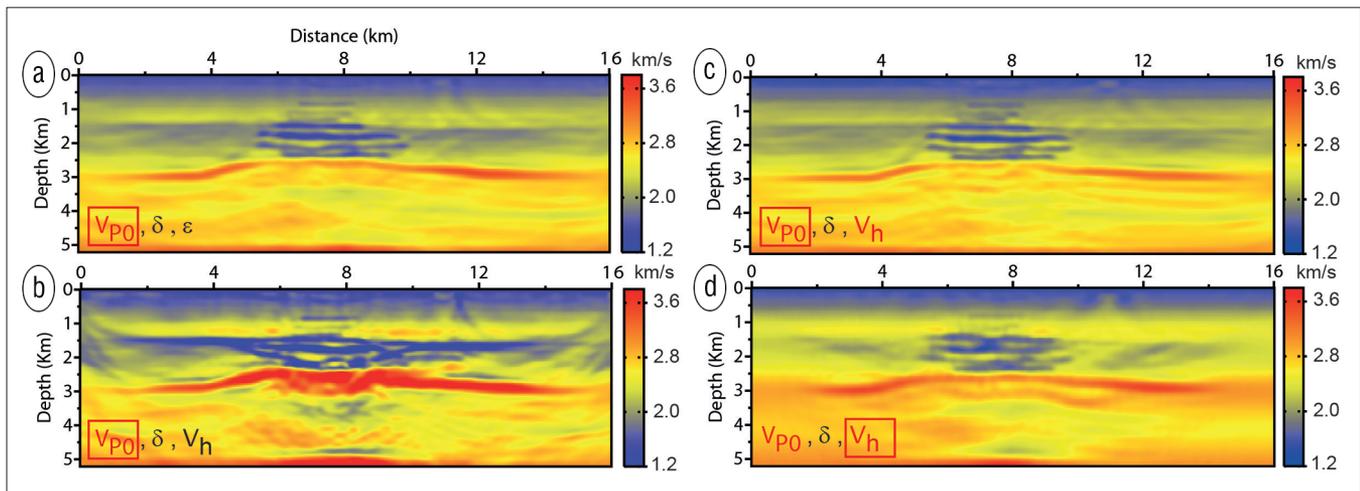


Figure 9. Synthetic Valhall model. VTI acoustic FWI. The parameters involved in the inversion are red. The box parameter is shown. (a) Final V_{p0} model with the $(V_{p0}, \delta, \varepsilon)$ parameterization. Only V_{p0} is updated, while the smooth δ and ε background models are kept fixed. A velocity model of similar quality would have been obtained by replacing V_{p0} by V_h or V_{NM0} in the parameterization. (b) Same as (a) but FWI is performed with the (V_{p0}, δ, V_h) parameterization. The inversion clearly converges toward a local minimum. Final V_{p0} (c) and V_h (d) models in the (V_{p0}, δ, V_h) parameterization. Compared to (b), V_{p0} and V_h are jointly updated during inversion (Gholami et al.).

(Figure 9c) is close to the one inferred from the $(V_{p0}, \delta, \varepsilon)$ parameterization (Figure 9a), despite the different radiation patterns of V_{p0} (Figure 8a and 8b). This results because a significant part of the long-wavelength components, which cannot be retrieved with the (V_{p0}, δ, V_h) parameterization, are already present in the initial model.

In summary, a parameterization that combines one wave velocity with Thomsen's parameters is suitable for the update of the selected velocity (vertical, horizontal, or normal moveout velocity) over a broad wavenumber spectrum. The ability of FWI to update ε depends on the accuracy of the initial ε model because this accuracy will control whether the influence of ε on the data residuals is significant and sticks out from the noise. Regardless of resolution issues, a parameterization that combines two velocities is more suitable for multiparameter FWI because the trade-off between them can be more easily managed (see also the eigenanalysis of the Hessian in Plessix and Cao, 2011).

Data-driven and model-driven strategies for visco-elastic imaging. All the discussion so far has been limited to the acoustic approximation and pressure data recorded by hydrophones. To evolve beyond the acoustic approximation, multicomponent recordings provide a significant added-value for multiparameter FWI, in particular in marine environments where the information carried out by the data on V_s can be difficult to extract.

The radiation patterns of V_p and V_s are shown in Figure 8c and Figure 8d for an isotropic elastic medium and for the four scattering modes: PP, SP, PS, and SS (see Forgues and Lambare for a more exhaustive analysis of elastic parameterizations). A V_s perturbation generates scattering only in the PP mode, with an isotropic radiation pattern (Figure 8c). A V_p perturbation generates scattering in the P- mode at intermediate θ , with a much smaller amplitude than the one of V_p . According to the limited influence of the V_s parameter on the PP scattering mode, a broadband V_p model can be estimated from the hydrophone in the acoustic approximation. This model can be used as a starting model for the elastic inversion of the hydrophone data, which allows for the updating of the long-to-intermediate wavelengths of V_s from the amplitude variations with offset of the P-P scattered waves. This intermediate step provides an improved starting model of V_s for the inversion of the more resolving geophone data as shown by the radiation patterns of V_s for the scattering modes PS, SP, and SS (Figure 8d). Potential trade-off between V_p and V_s occurs only in the PP mode at intermediate θ . However, this might not be a severe issue because only V_p has an influence of the P-wave traveltimes, that would favor the uncoupling between the two wavespeeds.

We have just described a data-driven and model-driven FWI workflow where the data components and the parameter classes are hierarchically processed. This hierarchy is driven by the strength of the parameter footprint in the data residuals, namely, dominant parameters are reconstructed before secondary parameters, and the resolution with which a parameter can be reconstructed, namely, long wavelengths of V_s are reconstructed from the hydrophone before the short

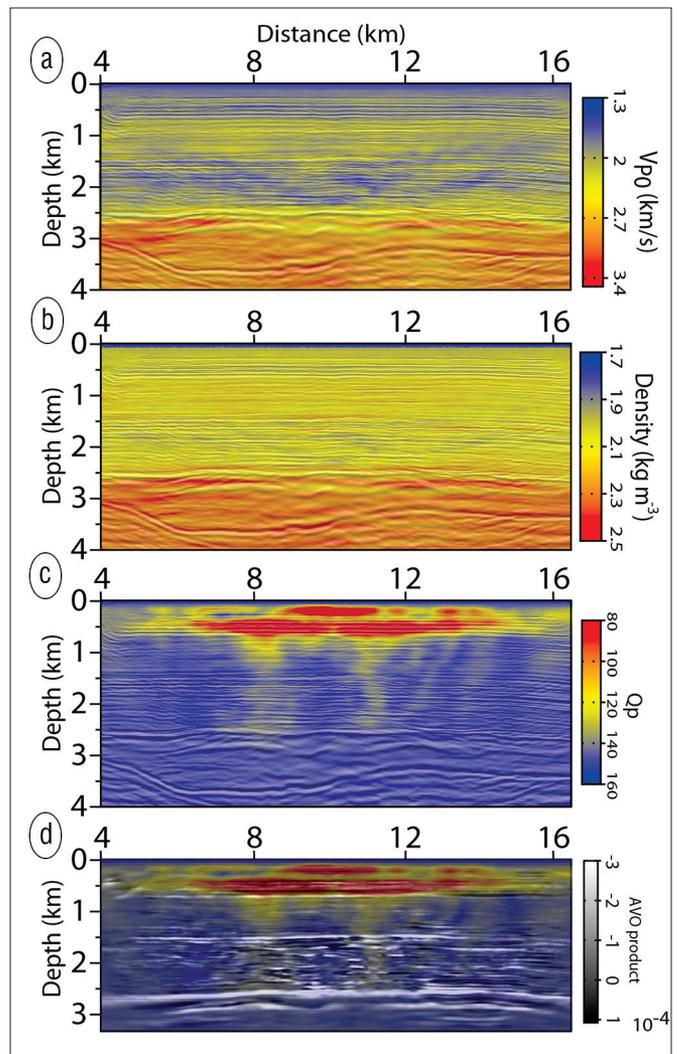


Figure 10. Multiparameter visco-elastic anisotropic FWI of multicomponent OBC data from Valhall. The updated parameters are V_{p0} , ρ , Q_p and V_s . The δ and ε background models are kept fixed. Visco-acoustic FWI of the hydrophone component: (a) V_{p0} , (b) ρ , and (c) Q_p . (d) AVO product inferred from the final V_p and V_s models of the elastic inversion with the Q_p model superimposed. A reverse time migrated image computed in the initial anisotropic model is superimposed on the FWI models shown in (a)-(c) (Prioux et al.).

wavelengths from the geophones. These levels of hierarchy are conventionally combined with other data manipulation related to frequency and θ : multiscale FWI can be designed with two nested loops over frequency and time (here, time should be ideally replaced by θ), which proceed from the low frequencies to the higher ones and from the early times to the later ones to inject progressively in the subsurface model smaller-scale heterogeneities (Brossier et al., 2009).

Our multiparameter FWI workflow is illustrated with a real data case study from Valhall in Figure 10, which shows the feasibility of the reconstruction of V_{p0} , ρ , Q_p and V_s by combined visco-acoustic and visco-elastic VTI FWI of 4C OBC data (Prioux et al., 2013). The inversion was performed in the frequency domain and was organized by successive inversions of slightly overlapping frequency groups of increasing high frequency content, with a maximum frequency of

6.7 Hz. No other data preconditioning such as time damping was applied. Thomsen's parameters δ and ε are kept fixed. The parameterizations $(V_{p0}, \rho, Q_p, \delta, \varepsilon)$ and $(V_{p0}, \rho, Q_p, \delta, \varepsilon, V_s)$ are used for the VTI FWI in the visco-acoustic and in the visco-elastic approximations, respectively. The workflow incorporates four inversion steps:

- 1) Update V_{p0} by monoparameter visco-acoustic FWI of the hydrophone data to reduce most of the data residuals associated with the dominant parameter.
- 2) Joint update V_{p0}, ρ and Q_p by visco-acoustic VTI FWI of the hydrophone data. The ρ and Q_p models are kept fixed during the subsequent steps.
- 3) Joint update V_{p0} and V_s by visco-elastic VTI FWI of the hydrophone data.
- 4) Joint update V_{p0} and V_s by visco-elastic VTI FWI of the geophone data.

Although V_{p0} has been involved in steps 2, 3 and 4, it was marginally updated relative to the one inferred from step 1. A reverse-time-migrated image, computed with VTI acoustic modeling, is superimposed on the FWI models to check the consistency between the FWI structures and the migration reflectors. The low Q_p values in the shallow part of the FWI model are consistent with the soft quaternary sediments. Some vertical structures below might be interpreted as gas chimneys. Although the density model lacks small wavenumbers, which makes its calibration against sonic logs challenging in an absolute sense, we found a good correlation between the position in depth of the main density contrasts in the FWI models and in the sonic log (Figure 19 in Prieux et al, 2013). The information carried out by V_s is shown through the AVO product, an indicator of layers with a hydrocarbon potential (Figure 10d, white layers). We found the highest values of this hydrocarbon indicator at the reservoir level at a depth of around 2.5 km. Interfaces at depths of 0.7 km and 1.5 km can also be interpreted as lithological barriers below which gas accumulates. These zones of gas accumulation might be correlated with attenuating zones between 0.7 km and 1.5 km in depth and between 7 km and 12 km in distance in Figure 10d.

Discussion and conclusion

The tutorial has reviewed the basic principles of multiparameter FWI, from which some guidelines for choosing a suitable parameterization and a suitable data-driven and model-driven workflow were inferred and some open questions were raised.

We have shown that a challenging issue in multiparameter FWI is to manage the trade-off between parameters of different nature. We have stressed the importance of the Hessian operator to properly scale the gradients associated with different parameter classes and correct them from trade-off artifacts. However, we also showed that trade-off effects vary with the local scattering angle associated with one arrival in the seismogram and one position in the subsurface. When the full information content of the data is processed

in one go during FWI, all local information is averaged in the Hessian in an indiscriminate way. Combined data-driven and model-driven approaches are often designed to separately process subsets of data and parameters, which carry information of a different nature on the subsurface. We have reviewed an application of such visco-elastic anisotropic FWI workflow on real multicomponent OBC data from Valhall. However, we believe that these heuristic approaches are currently far from being optimized and automated because a dedicated parameterization (the one which removes trade-off effects) should ideally be defined for each scattering angle spanned by a given arrival at a given position. It is unclear how to access this local information during FWI and how to design tractable algorithms which can manage it.

Acknowledging the just-mentioned limitations, we have provided some guidelines to select a suitable parameterization for multiparameter FWI. Schematically, the dilemma we often face is to choose between a parameterization which favors the resolution with which the dominant parameter is reconstructed and a parameterization that minimizes trade-off between parameters. Generally, we converge to the conclusion that the first option should be chosen when wide-aperture data are available to favor a broadband reconstruction of the velocity parameter. Secondary parameters can be involved in the inversion or kept fixed depending whether the available initial models are sufficiently-accurate to make the footprint of these parameters in the data residuals negligible. This conclusion can be, however, challenged when considering FWI strategies for seismic reflection data. These strategies honor the scale uncoupling underlying conventional seismic reflection processing by alternating the macrovelocity model building and reflectivity imaging in a two-step FWI workflow. These two tasks can be performed with two different parameterizations, one which is suitable for the reconstruction of the long-wavelengths of the wavespeed and one which is suitable for the short-scale reconstruction of the parameter that describes reflectivity (i.e., impedance).

We also highlight that parameters that have a dominant effect on amplitudes rather than on phase, such as density and attenuation, can be updated during inversion although reliable FWI results might be obtained by keeping them fixed. This results because FWI is dominated by phase effects rather than amplitude effects at least at low frequencies. The wave physics that must be considered for a reliable accounting for amplitudes is still a matter of debate, as well as our ability to push FWI at high frequencies.

Although FWI relies on a simple correlation procedure between incident and adjoint back-propagated wavefields, design of robust and efficient multiparameter FWI algorithms is deeply rooted in the field of numerical optimization. A first issue is to design inversion algorithms, which account for the Hessian without explicitly building this matrix. The quasi-Newton L-BFGS method and the truncated Newton method are two matrix-free algorithms, which account for the Hessian in multiparameter FWI with a manageable computational burden (see Métivier et al. for relevant references). Suitable multiparameter FWI algorithms should

be designed such that different subsurface parameterizations can be easily interfaced with the inversion; the subsurface parameterization, which is used to discretize the wave equation for seismic modeling, must be independent from the one that is used to solve the inverse problem. The inverse problem requires some adimensionalization to make the optimization well conditioned and independent to the physical nature of the model parameters. Examples of a well-conditioned and poorly-conditioned Hessian are shown in Figure 3. This adimensionalization is also useful to automate the tuning of the regularization hyperparameters associated with different parameter classes (i.e., the parameters which control the relative weight of the data residuals and the prior model constraints in the misfit function). During the Valhall case study, we have produced several versions of V_p , ρ and Q_p models from different parameter scaling and regularization tuning. These models allowed us to fit the data equally well, highlighting the ill-posedness of multiparameter FWI. The best models were selected based on some a priori geologic information and the calibration of the V_p model against a sonic log. We have concluded that two key methodological issues to steer the inversion toward the correct solution, aside the management of parameter trade-off, are the incorporation of reliable prior constraints in the misfit function and an automated approach of parameter scaling and FWI regularization. The *prior* information should be implemented as an error correction term in the functional rather than in the initial model to avoid introducing some kinematic bias in the subsurface model in which seismic modeling is performed (Asnaashari et al., 2013). Reliable estimation of model uncertainty remains a fundamental open question in FWI based on local optimization methods. **TLF**

References

- Asnaashari, A., R. Brossier, S. Garambois, F. Audebert, P. Thore, and J. Virieux, 2013, Regularized seismic full waveform inversion with prior model information: *Geophysics*, **78**, no. 2, R25–R36, <http://dx.doi.org/10.1190/geo2012-0104.1>.
- Brossier, R., S. Operto, and J. Virieux, 2009, Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion: *Geophysics*, **74**, no. 6, WCC105–WCC118, <http://dx.doi.org/10.1190/1.3215771>.
- Forgues, E., and G. Lambaré, 1997, Parameterization study for acoustic and elastic ray + Born inversion: *Journal of Seismic Exploration*, **6**, 253–278.
- Gholami, Y., R. Brossier, S. Operto, A. Ribodetti, and J. Virieux, 2013, Which parametrization is suitable for acoustic VTI full waveform inversion? Part 1: sensitivity and trade-off analysis: *Geophysics*, **78**, no. 2, R81–R105, <http://dx.doi.org/10.1190/geo2012-0204.1>
- Gholami, Y., R. Brossier, S. Operto, V. Prioux, A. Ribodetti, and J. Virieux, 2013, Which parametrization is suitable for acoustic VTI full waveform inversion? Part 2: application to Valhall: *Geophysics*, **78**, no. 2, R107–R124, <http://dx.doi.org/10.1190/geo2012-0203.1>
- Malinowski, M., S. Operto, and A. Ribodetti, 2011, High-resolution seismic attenuation imaging from wide-aperture onshore data by visco-acoustic frequency-domain full waveform inversion: *Geophysical Journal International*, **186**, no. 3, 1179–1204, <http://dx.doi.org/10.1111/j.1365-246X.2011.05098.x>.
- Métivier, L., R. Brossier, J. Virieux, and S. Operto, 2012, Full waveform inversion and the truncated newton method: *SIAM Journal on Scientific Computing*, **35**, no. 2, B401–B437, <http://dx.doi.org/10.1137/120877854>.
- Plessix, R. E., and Q. Cao, 2011, A parametrization study for surface seismic full waveform inversion in an acoustic vertical transversely isotropic medium: *Geophysical Journal International*, **185**, no. 1, 539–556, <http://dx.doi.org/10.1111/j.1365-246X.2011.04957.x>.
- Prioux, V., R. Brossier, Y. Gholami, S. Operto, J. Virieux, O. J. Barkved, and J. H. Kommedal, 2011, On the footprint of anisotropy on isotropic full waveform inversion: the Valhall case study: *Geophysical Journal International*, **187**, no. 3, 1495–1515, <http://dx.doi.org/10.1111/j.1365-246X.2011.05209.x>.
- Prioux, V., R. Brossier, S. Operto, and J. Virieux, 2013, Multiparameter full waveform inversion of multicomponent OBC data from Valhall. Part 1: imaging compressional wavespeed, density and attenuation: *Geophysical Journal International*, <http://dx.doi.org/10.1093/gji/ggt177>.
- Prioux, V., R. Brossier, S. Operto, and J. Virieux, 2013, Multiparameter full waveform inversion of multicomponent OBC data from valhall. Part 2: imaging compressional and shear-wave velocities: *Geophysical Journal International*, <http://dx.doi.org/10.1093/gji/ggt178>.
- Virieux, J., and S. Operto, 2009, An overview of full waveform inversion in exploration geophysics: *Geophysics*, **74**, no. 6, WCC1–WCC26, <http://dx.doi.org/10.1190/1.3238367>.
- Xu, S., D. Wang, F. Chen, G. Lambaré, and Y. Zhang, 2012, Inversion on reflected seismic wave: 82nd Annual International Meeting, SEG, Expanded Abstracts, 1–7.

Acknowledgments: This study was partly funded by the SEISCOPE consortium (<http://seiscope2.osug.fr>). We the guest editors for offering us the opportunity to participate in this special section.

Corresponding author: operto@geoazur.obs-ufv.fr