Acoustic multiparameter Full-waveform Inversion of diving and reflected waves through a hierachical scheme

W. Zhou, ISTerre, Université Grenoble Alpes, R. Brossier, ISTerre, Université Grenoble Alpes, S. Operto, Géoazur-CNRS, Université Nice Sophia Antipolis, J. Virieux, ISTerre, Université Grenoble Alpes*

SUMMARY

Most of the successful applications of FWI have been mainly driven by the information carried out by diving waves and subcritical reflections, as the classical formalism appears to be unable to exploit reflected waves to build the long-to-intermediate wavelengths of the velocity structure. Alternative approaches have been recently revisited to focus on short-spread reflections only using some prior knowledge of the reflectivity. The present study presents a unified formalism, which aims to update the low wavenumbers of the velocity model by joint inversion of the diving waves and the full reflected wavefield. A hierarchical scheme assuming a scale seperation is implemented to alternatively update the high-wavenumber impedance model through a non-linear migration-like approach and low wavenumber velocity model through our unified formulation.

INTRODUCTION

Full Waveform Inversion (FWI) is an appealing method for quantitative high-resolution subsurface imaging (Virieux and Operto, 2009). In most applications, classical FWI approaches only invert for the velocity parameter (V_P) to avoid updating too many unknowns, assuming that a density (ρ) background model is constant or known. If the initial V_P model allows to predict the data without cycle skipping and very long offsets make transmission data available, classical FWI is theoretically amenable to update both the kinematics (smooth variations) and kinetics (rapid variations) properties of the reconstructed V_P model. However, other parameters such as the density and/or impedance can have a significant impact on the amplitudes of the reflected wavefield in presence of sharp contrasts. Limiting the inversion to the velocity update keeping the other parameters fixed might introduce artifacts in the final V_P model. Therefore, it appears that multi-parameter inversion is required for reliable high-frequency quantitative imaging from both diving/transmitted and reflected waves.

According to the diffraction pattern associated with parameters of different nature (see Operto et al. (2013) for a recent analysis), the sensitivity of FWI to parameters varies with the subsurface parameterization that is used for inversion. In particular, the velocity-impedance parameterization ($V_P - I_P$ with $I_P = V_P\,\rho$) makes FWI sensitive mainly to transmitted energy for V_P update and reflected energy for I_P update. With such parameterization, cross-talks between the two classes of parameters are reduced, at the expense of the resolution with which the V_P model is updated. In contrast, the $V_P - \rho$ parametrization enables FWI to exploit the full range of diffracting angles to update V_P, while FWI is only sensitive to the short-spread reflections for the ρ update. Although high-resolution V_P models can be built, this parameterization leads to significant cross-talks between V_P and ρ during the inversion of the short-spread reflections. In order to reduce such ambiguity, the (V_P, I_P) parameterization might be more suitable for FWI of reflected waves. The fact that V_P and I_P are updated within two distinct wavenumber bands (low and high, respectively) during FWI prompts us to recast the inversion formulation under the assumption of an explicit scale separation between a smooth velocity background model and a short-wavelength impedance model.

Gradients of the FWI misfit function can be analyzed under the diffraction tomography principle (Devaney, 1982), which states that the imaged wavenumber $\vec{\mathbf{k}}$ at a point diffractor is linked to the angular frequency ω and local diffraction angle θ through the relation : $\vec{\mathbf{k}} = 2\omega/c_0 \cos(\theta/2)\vec{\mathbf{n}}$, where c_0 denotes the local wavespeed. The recovered wavenumber spectrum is therefore determined by the sampling of ω and θ. As both diving and reflected waves generally illuminate the shallow part of the subsurface, leading to a broad diffraction-angle coverage, the low wavenumber spectrum (i.e., long-wavelength variations) is well reconstructed in this area, whereas deeper targets are mainly illuminated by short-spread reflections, which favor the reconstruction of short wavelengths at the expense of longer ones. Although low frequencies may balance the deficit of wide-angle coverage to update the low wavenumbers of the subsurface, in practice the lowest available frequency is not sufficiently small to overcome this issue.

To update the velocity macromodel from reflection data, image-domain strategies (e.g., Symes and Carazzone, 1991) aim to maximize a semblance criterion in the migrated domain. Alternatively, recent data-domain strategies (e.g., Xu et al., 2012; Brossier et al., 2013; Ma and Hale, 2013), inspired by Chavent et al. (1994), rely on a scale separation between the velocity macromodel and prior knowledge of the reflectivity to emphasize the transmission regime in the sensitivity kernel of the inversion. However, all these strategies focus on reflected waves, discarding the low-wavenumber information carried out by diving waves. As diving waves and subcritical reflections provide a significant part of the recorded energy, high-resolution tomographic methods should take advantage of all types of waves.

In this study, we propose a multi-parameter FWI based on the $V_P - I_P$ parameterization, which iteratively alternates the updating of the I_P model (keeping V_P fixed) and that of the V_P model (keeping I_P fixed). The misfit function associated with the I_P inversion involves the reflected waves only, leading to a least-square migration-like algorithm to build the high-wavenumber I_P model. Due to the specific chosen parameterization, this inversion does not inject low wavenumbers in the I_P model, even in the shallow part. On the other hand, for the V_P update, we design a new misfit function that allows us

DOI http://dx.doi.org/10.1190/segam2014-1634.1

to use both diving and reflected energy during the updating of the small-to-intermediate wavenumbers of the velocity model (Zhou et al., 2014). Assuming that the reflectivity model that is derived from the previous I_P inversion is known and separating explicitly the contributions of the reflected waves and the diving waves in the misfit function allow us to remove the first-order migration isochrones associated with reflected waves in the sensitivity kernel of FWI. Without this explicit separation, the energetic high wavenumbers from those migration isochrones would be introduced into the V_P model, although the $V_P - I_P$ parameterization is used: this is not desirable for velocity macromodel building and would violate the scale-separation assumption.

Both forward modeling and inversion are performed in the time domain, although the formulation is developed here in the frequency domain for sake of compactness. The synthetic Valhall model with surface long-offset survey will illustrate the concepts of our approach.

THEORY

In classical FWI, the misfit function is defined through the L2 norm of the data residual

$$C_{clasFWI} = 0.5\|d - Ru\|^2, \tag{1}$$

where the observed data are denoted by d and the modeled wavefield is denoted by u sampled at receiver positions (the real-valued sampling operator R). All types of waves are involved in this definition, including diving, direct, reflected waves and so on. As a criterion to assess the quality of the proposed model, a lower value of the misfit function is searched by using the local descent direction w.r.t. model parameters m, generally termed as the gradient :

$$G_{clasFWI} = \lambda^\dagger \frac{\partial B}{\partial m} u, \tag{2}$$

where $\dagger$ denotes the adjoint operation, B the forward problem operator and λ the adjoint wavefield. In the time domain, this formalism transfers to a zero-lag cross-correlation of fields u and λ with coefficients $\partial B/\partial m$. To understand the behavior of the gradient, let us consider a two-reflector model prospected with surface survey ($z = 0$km). The initial velocity slightly differs from the true velocity. The adjoint field associated with the direct wave residual λ^d generates a first Fresnel zone within $z = \pm 0.5$km, while the one associated with the reflected wave residual λ^r generates two secondary Fresnel zones (also known as migration isochrones), through the correlations with u (Figure 1-a). The first Fresnel zone contributes to long-wavelength update, while the isochrones delineate the reflector positions. Due to the trade-off between the depth of reflector and background velocity, the imaged depths of reflectors can be hardly corrected ($z = 0.84, 1.7$km vs true depths$= 0.80, 1.6$km) and the velocity remains erroneous ($= 1600$m/s vs true value=1500m/s). From this simple test, we show that the classical FWI fails to retrieve the long-wavelength of velocity contained in the reflected waves, and falls into local minima due to the trade-off effect. We therefore rely on the scale separation between the velocity macromodel and high-wavenumber reflectivity model, and aim to develop a formalism that enables us to eliminate the high-wavenumber first-order isochrones in the gradient (2).

Building long-wavelength velocity model

Due to the usual frequency band and offset-range that are considered in seismic exploration, the recorded direct/diving waves, and hence the first Fresnel zone in the FWI gradient (2), sample only the shallow part of the subsurface, while the migration isochrones generated by the reflection residuals sample the subsurface at all depths. In this setting, the long wavelengths can be only updated in the shallow part of the subsurface. Some approaches have been proposed to update the long wavelengths from reflected waves (Xu et al., 2012; Brossier et al., 2013). However, these approaches do not make use of the long-wavelength information carried out by diving waves. In this study, we propose a new integrated formulation of FWI that jointly invert diving and reflected waves for macro-velocity model building.

$$C_{newFWI} = 0.5\|W^d(d^d - Ru_0)\|^2 + 0.5\|W^r(d^r - R\delta u)\|^2, \tag{3}$$

where d^d and d^r denote the observed direct and reflected arrivals, u_0 and δu denote the background and scattered part of the modeled wavefield, respectively. The scattered field δu is computed from a reflectivity model, which is assumed to be known. In this definition, the misfit function has been split into two L2 norms: this splitting honors the explicit separation between the contributions of the direct/diving waves and reflections part of the data in the FWI formulation. Without this separation, the high-wavenumber first-order isochrones would not removed from the FWI gradient. Weighting operators W^d and W^r are also introduced to balance the respective contributions of the diving and reflected waves.

The gradient of this misfit function (7) can be derived through the Lagrangian formulation of the adjoint-state method:

$$\begin{aligned}\mathscr{L}(m_0, u_0, \delta u, a_1, a_2) &= C_{newFWI} + \langle a_1, B(m_0)u_0 - s\rangle \\ &+ \langle a_2, B(m_0, \delta m)\delta u + [B(m_0, \delta m) - B(m_0)]u_0\rangle,\end{aligned} \tag{4}$$

where m_0 denotes the current velocity model, δm the given reflectivity model, s the source term and a_1, a_2 the adjoint fields, which are found by zeroing the partial derivatives of (4) w.r.t. the state variables u_0 and δu. The field a_2, which satisfies $B(m_0, \delta m)^\dagger a_2 = -R^{\mathrm{T}} W^{r\mathrm{T}} W^r \Delta d^{r*}$ with the conjugate reflected wave residual Δd^{r*} as the source term, can be decomposed into a background part λ_0^r computed in m_0 and a perturbation part $\delta\lambda^r$ scattered by δm. The field a_1, which satisfies $B(m_0)^\dagger a_1 = -[B(m_0, \delta m) - B(m_0)]^\dagger a_2 - R^{\mathrm{T}} W^{d\mathrm{T}} W^d \Delta d^{d*}$, can also be decomposed into a background part λ_0^d generated by the direct/diving wave residuals Δd^{d*} and a scattered part $\delta\lambda^r$ associated with the scattering source term $[B(m_0) - B(m_0, \delta m)]a_2$ from the perturbation model. The superscript of fields λ denotes the type of adjoint source (direct/diving versus reflection residuals). The gradient of (7) for m_0 update is given by

$$\begin{aligned}G_{newFWI} &= \partial_{m_0}\mathscr{L} \approx a_1^\dagger\, \partial B\, u_0 + a_2^\dagger\, \partial B\, \delta u \\ &= \lambda_0^{d\dagger}\, \partial B\, u_0 + \delta\lambda^{r\dagger}\, \partial B\, u_0 + \lambda_0^{r\dagger}\, \partial B\, \delta u + \delta\lambda^{r\dagger}\, \partial B\, \delta u\end{aligned} \tag{5}$$

DOI http://dx.doi.org/10.1190/segam2014-1634.1

where ∂B is a shorthand for the diffraction pattern $\partial_{m_0} B(m_0, \delta m)$. Some terms are neglected as they only appear at $\delta m \neq 0$ positions. This formalism (5) indicates a combination of first Fresnel zones associated with diving waves (1st. term) and reflected waves (from 2nd. and 3rd. terms), as well as isochrones resulting from higher-order scattering (from 2nd. to 4th. terms) (Figure 1-b,c).

A direct implementation of eq. (5) would require a scattered-field formulation of the forward problem, for which the source terms of the scattered fields δu and a_1 are expensive to compute in the time domain. Therefore, we propose a cheaper implementation of G_{newFWI} (which should be a reasonable approximation):

$$G_{newFWI} \approx G_{clasFWI \text{ in } m_0, \delta m} - \lambda_0^{r\dagger}\, \partial_{m_0} B(m_0)\, u_0 \qquad (6)$$

Three computations are necessary: [1] Compute the gradient of the classical FWI misfit function (2) with model $m_0, \delta m$, using weighted data residuals $W^d \Delta d^d + W^r \Delta d^r$ instead of Δd in eq. (1). This term contains all terms in eq. (5) plus undesired first-order isochrones $\lambda_0^{r\dagger}\, \partial B\, u_0$. [2] Compute $\lambda_0^{r\dagger}\, \partial_{m_0} B(m_0)\, u_0$ using $W^r \Delta d^r$ as an approximation to $\lambda_0^{r\dagger}\, \partial B\, u_0$ obtained from [1]. [3] Subtract the two terms to discard the first-order isochrones. Note that [1] and [2] could be performed in parallel, demanding two-times more computational resources than classical FWI, for the same real computational time.

High-order scattering effect and inversion parametrization

The above workflow is unable to suppress high-order isochrones generated by the constructive correlation of multi-scattered wavefields (Figure 1-b). Fortunately, these isochrones have smaller amplitudes than those of first-order isochrones removed by our workflow (2nd. term on the RHS of (6)), and most part of them destructively interfere each other when multiple sources and receivers are considered. However, not all these high-order isochrones can be suppressed in the same way, as some of them show the same behavior as the first-order isochrones. Efficient filtering of these high-order isochrones can be performed through a judicious subsurface parametrization for gradient computation. If the V_P gradient is built with the $V_P - \rho$ parameterization, a broad range of wavelengths are imaged according to the isotropic diffraction pattern of V_P for this parameterization (Operto et al., 2013). In this case, the isochrones cannot be fully filtered out (Figure 1-b). In contrast, if the $V_P - I_P$ parameterization is used, only the long-to-intermediate wavelengths of V_P are reconstructed from the wide diffraction angles (Figure 1-c), while the short-to-intermediate wavelengths of the subsurface are mapped into the I_P model, according to the diffraction pattern of these two parameter classes. Therefore, the $V_P - I_P$ parameterization is chosen in the following to reduce the imprint of isochrones in the V_P models.

Inversion of $V_P - I_P$ through a hierarchical scheme

We propose to implement the $V_P - I_P$ inversion through a hierarchical scheme. Starting from a smooth V_P model, we perform the I_P inversion using only weighted reflected wave residual. The misfit function for this I_P inversion is

$$C_{I_P FWI} = 0.5 \| W^r (d^r - R \delta u) \|^2 \qquad (7)$$

This inversion builds a high-wavenumber initial I_P model which is used as the known reflectivity model for the subsequent V_P update. For V_P reconstruction we only invert for a smooth model through the workflow (6), regardless any high-wavenumber perturbations. In this way we decouple the contribution of the two parameters and make the inversion more robust to recover the true models. We then perform a new inversion of the reflected wavefield to update the reflectivity model according to the updated V_P model, as the new V_P would introduce cycle-skipping without updating the reflectivity model (Brossier et al., 2013). Therefore, the overall multi-parameter inversion is implemented by alternating inversions of I_P and V_P in an iterative manner.

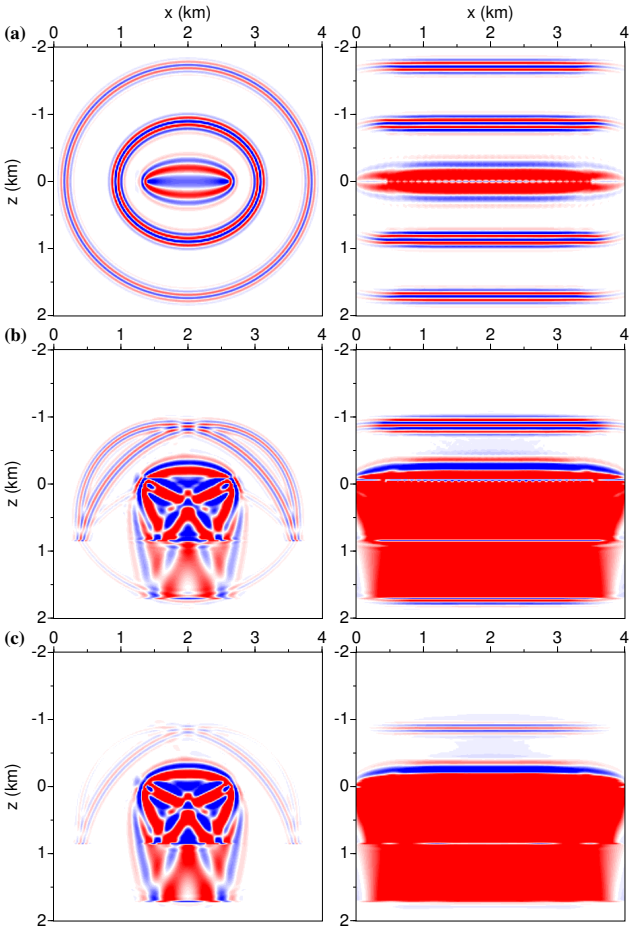

Figure 1: Preconditioned gradients of V_P. Left: One source-receiver pair. Right: Full acquisition. (a) Classical FWI gradient computed without prior reflectivity. (b-c) New FWI gradient computed with prior reflectivity. (b) $V_P - \rho$ parameterization (c) $V_P - I_P$ parameterization. Note that more long-wavelength content is retrieved in (b), (c) than in (a), and that the unwanted high-order isochrones are damped in (c).

NUMERICAL EXPERIMENTS

We illustrate this workflow on the synthetic 2D Valhall model (Figure 2-a,b). The initial models (Figure 2-c,d) are built from 1D models. This models is extracted from a highly smoothed version of the true model, but deliberately chosen such that

DOI http://dx.doi.org/10.1190/segam2014-1634.1

no cycle-skipping occurs on direct and diving waves. Sources and receivers are put in the water layer to mimic an offshore acquisition with the maximum offset equal to 6km.

The first high-wavenumber I_P model (Figure 2-e) is built by using the reflected waves data, and iterating with l-BFGS optimization. The misfit value of the reflected waves (i.e. eq. (7)) is decreased to 50% within only 5 iterations as the I_P inversion is a quite linear problem. However, the number of iteration has to remain small to avoid overfitting the data with the wrong velocity model, which introduce significant artifacts to the I_P model and degenerate the quality of the reflectivity information.

This first impedance model is then used as fixed prior perturbation model to compute the gradient of V_P, derived from both diving and reflected waves with our proposed unified formulation (Figure 2-f). Because the high-wavenumber components have been suppressed or canceled out, the gradient is dominated by long-wavelength updates that should be able to reconstruct a smooth V_P model. As the gradient is computed in the first I_P model, which is yet unable to focalize energy of the deep reflectors, the long wavelength update is still limited above the anticlinal structure, and focalized to the low velocity zones of the gas layers. The alternate updates of velocity and impedance should be able to focalize energy to deeper reflectors with iterations, making the update of velocity at greater depth with iterations.

As a comparison, we also show the classical FWI gradients for ρ (Figure 2-g) and V_P (Figure 2-h), respectively, computed in the same initial model. As expected, the classical approach is mainly sensitive to small diffraction angles of reflected waves for ρ, leading to the ρ gradient dominated by high-wavenumber components. On the contrary, due to the sensitivity to the full diffraction angle range for V_P, the gradient of V_P contains a full spectrum of wavenumbers. However, the low-wavenumber components are limited within the depth range sample by diving waves at about $z < 1.5$km, meaning that the low-wavenumber components contained in reflected waves are not well used.

This example shows how our method is able to exploit the low-wavenumber components from both diving and reflected waves.

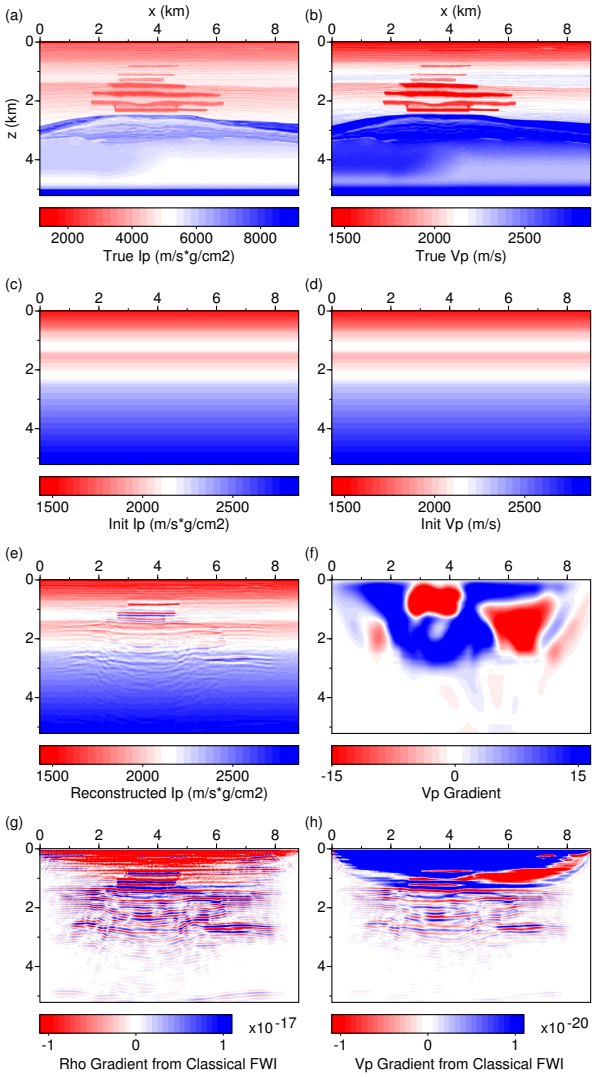

Figure 2: Example on Valhall case. (a,b) True models and (c,d) initial models. (e) The first I_P model is reconstructed by using reflected waves only. This I_P model is then fixed and used as reflectivity information to derive the V_P gradient (f), which is dominated by low-wavenumber components. The gradients of (g) ρ (g) and (h) V_P from the classical FWI exhibits high-wavenumber components and low-wavenumber components only in the shallow part of the V_P gradient, sampled by diving waves.

CONCLUSION

We propose to build the velocity macromodel by FWI by joint inversion of diving waves and reflected waves. Our approach relies on the explicit decomposition of the subsurface model into a smooth background velocity model and reflectivity, and the dataset into diving waves and short-spread reflected waves. These decompositions make the velocity gradient to be dominated by surface-to-surface and reflector-to-surface first Fresnel zones. We propose a computationally efficient implementation of this approach by subtracting two terms computed by classical FWI in two models, one with the prior reflectivity and one without. Although fictitious second-order isochrones cannot be explicitly suppressed, they are mostly removed by stacking over multiple sources and receivers. The role of the subsurface parameterization to drive the FWI toward the reconstruction of the long wavelengths is also highlighted. This unified formulation for the velocity update is integrated to a hierarchical workflow that alternates impedance and velocity updates.

ACKNOWLEDGEMENTS

This work is funded by the SEISCOPE II consortium (*http://seiscope2.osug.fr*). This study is granted access to the HPC facilities of CIMENT (Université Joseph Fourier, Grenoble) and GENCI-CINES under grand 046091. Authors appreciate fruitful discussions with L. Métivier and Y. Li.

DOI http://dx.doi.org/10.1190/segam2014-1634.1

http://dx.doi.org/10.1190/segam2014-1634.1

EDITED REFERENCES

Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2014 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES

Chavent, G., F. Clément, and S. Gómez, 1994, Automatic determination of velocities via migration-based traveltime waveform inversion: A synthetic data example: 64th Annual International Meeting, SEG, Expanded Abstracts, 1179–1182.

Devaney, A. J., 1982, A filtered backpropagation algorithm for diffraction tomography: Ultrasonic Imaging, **4**, no. 4, 336–350, doi: 10.1016/0161-7346(82)90017-7.

Ma, Y., and D. Hale, 2013, Wave-equation reflection traveltime inversion with dynamic warping and full waveform inversion: Geophysics, **78**, no. 6, R223–R233, http://dx.doi.org/10.1190/geo2013-0004.1.

Operto, S., R. Brossier, Y. Gholami, L. Métivier, V. Prieux, A. Ribodetti, and J. Virieux, 2013, A guided tour of multiparameter full waveform inversion for multicomponent data: From theory to practice: The Leading Edge, **32**, 1040–1054.

Symes, W. W., and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663, http://dx.doi.org/10.1190/1.1443082.

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, doi: 10.1190/segam2012-1473.1.

Zhou, W., R. Brossier, S. Operto, and J. Virieux, 2014, Combining diving and reflected waves for velocity model building by waveform inversion: 76th Conference & Exhibition, EAGE, Extended Abstracts, We E106 15.

DOI http://dx.doi.org/10.1190/segam2014-1634.1