A receiver-extension approach to robust full waveform inversion

Ludovic Métivier, Univ. Grenoble Alpes, CNRS, LJK and Romain Brossier, Univ. Grenoble Alpes, ISTerre*

SUMMARY

We propose a receiver-extension approach to full waveform inversion, which, contrary to source-extension methods, can be used directly with a time-domain formulation. We introduce the receiver position as the extension parameter. We use a nested loop strategy to minimize the resulting misfit function: the inner loop corrects for the receiver position using global optimization while the outer loop corrects for the subsurface velocity using local optimization. First results illustrate the interest of the approach.

INTRODUCTION

Mitigating cycle skipping in full waveform inversion (FWI) to enhance the robustness of this technology has been an important and challenging topic since it was introduced as a high resolution seismic imaging technique by Lailly (1983) and Tarantola (1984). Beyond low-to-high frequency continuation and time/offset windowing, two major interpretations of the cycle skipping issue have led to two main approaches to mitigate it.

In the first interpretation, cycle skipping is seen as the result of the non-convexity of the L^2 misfit function usually employed to measure the data-fit. Accordingly, reformulations of FWI based on alternative misfit functions have been proposed to mitigate this non-convexity: cross-correlation (Luo and Schuster, 1991), deconvolution (Luo and Sava, 2011), and more recently optimal transport distances (Yang and Engquist, 2018; Métivier et al., 2019) are instances of these propositions.

In the second interpretation, cycle skipping is seen as a consequence of the nonlinearity of FWI, which resorts to the reduced space approach employed to make the PDE constrained optimization problem behind FWI affordable. Therefore, relaxing the reduced space approach by introducing additional "unphysical" degrees of freedom to the inversion is thought as a mean to mitigate cycle skipping. The associated methods are referred to as extended-model approaches.

In theses approaches, two main classes of strategies exist. The first are based on velocity parameter extension, and are generalization of migration velocity analysis strategies (MVA) (Symes, 2008). The second have been recently gathered as "matched source waveform inversion" (MSWI) by (Huang et al., 2018b), in an effort to unify the concepts behind different methods proposed in the past few years. In MSWI, the additional degrees of freedom are introduced at the source level. These degrees of freedom can be related to a time-receiver extension (adaptive waveform inversion of Warner and Guasch (2016)), a time-space extension (wavefield reconstruction inversion proposed by van Leeuwen and Herrmann (2013, 2016); Aghamiry et al. (2019)), or a space-only extension (volume source based extended waveform inversion introduced by Huang et al. (2018b,a)).

While the theory around MVA based methods has been intensively developed during the last decades, with either acquisition dependent extension, or subsurface offset/time-lag extension, their application to field data seems still limited today. The main reasons are related to (1) the additional cost required to build the image volumes on which these methods are based, and (2) the lack of robustness in complex media where complex wavepath between sources and receivers are observed.

On the other hand, more recent MSWI techniques appear to show interesting promises. AWI has been successfully applied to time-domain FWI on field data. However it has been recognized it can still suffer from cycle skipping in complex media (Huang et al., 2018b). Time-space extension (WRI) and space extension seem to exhibit the best properties in terms of robustness. However their main drawback is the difficulty to apply them in the frame of time-domain FWI, which is the current industry standard for FWI applications.

The two latter methods rely on the alternate reconstruction between the velocity model and the extended source. The extended source reconstruction step is performed in the L^2 sense, which requires the solution of a normal system associated with the wave equation operator. While this can be easily implemented in the frequency domain (especially when factorization of the wave equation operator is already available via direct linear solvers), this is much more challenging in the time-domain. A first extension to time-domain was proposed in (Wang et al., 2016), relying on an approximation to the solution of this normal equation. In a recent study, it is shown that this approximation is relatively crude, and an alternative technique is proposed, however to the price of higher computational cost (approximately 8 times the price of conventional FWI in the first steps of the process, see Aghamiry et al., 2020).

In this study, we want to propose yet another extension approach, amenable to time-domain FWI, without a drastic increase of the computational cost. To avoid the solution of a normal system associated with the wave propagation operator, we investigate a receiver extension approach rather than a source extension approach. Our additional degree of freedom is taken as the receivers position. This is motivated by the idea that time shifts due to incorrect velocities in the data might be compensated by relocating the receivers. The misfit function we consider depends thus on the velocity model and the receivers position. It is built as the sum between a L^2 data matching term, and a focusing term for the receiver position.

Similarly as in MSWI methods, this misfit function is minimized in an alternate way: the inner loop reconstructs the receivers position, while the outer loop updates the velocity model. However, for MSWI, the extended source is found as the solution of a linear L^2 problem (convergence in a single Newton step). In our approach, the receiver position is the solution of a nonlinear L^2 problem, which is the minimizer of a non-convex misfit function. To compute it, we adopt a global optimization approach: for each receiver, the optimal position is determined using a simple grid-search algorithm. For the

© 2020 Society of Exploration Geophysicists

10.1190/segam2020-3425103.1

outer loop, a standard local optimization approach is used. We show on a simple example how this can lead to the minimization of a convex misfit function with respect to the velocity parameter. Then, we illustrate the efficiency of this approach on the Marmousi II model.

METHODOLOGY

Formalism

Conventional L^2 FWI can be formulated as

$$\min_m f[m] = \frac{1}{2}\sum_{r,s}\int_0^T |R(x_r)u_s[m](t) - d_{obs,r,s}(t)|^2 dt, \quad (1)$$

where T is the recording time, $d_{obs,s,r}(t)$ denotes the observed seismic trace associated with the source/receiver couple s,r, $u_s[m](x,t)$ is the forward wavefield, solution of the general wave equation

$$A(m)u_s = b_s, \quad (2)$$

with b_s a given source term, and $R(x_r)$ is a restriction operator extracting the values of the wavefield at the receiver location x_r:

$$R(x_r)u_s = \int_\Omega \delta(x - x_r)u(x,t)dx, \quad (3)$$

where $\delta(x)$ denotes the delta Dirac function.

Our idea is to consider the receiver-extended FWI problem

$$\min_{m,\Delta x} f[m,\Delta x] = \frac{1}{2}\sum_{r,s}\int_0^T |R(x_r + \Delta x_{r,s})u_s[m](t) - d_{obs,s,r}(t)|^2 dt + \frac{1}{2}\|\Delta x\|_\alpha^2 \quad (4)$$

where Δx is a vector of spatial corrections for receiver positions (one correction per source/receiver couple), and $\|.\|.$ is a weighted norm

$$\|\Delta x\|_\alpha^2 = \sum_{r,s}\alpha_{r,s}^2\Delta x_{r,s}^2. \quad (5)$$

with $\alpha_{r,s}$ a weighting vector. We solve this extended problem with the usual nested loop approach. The inner loop is

$$\min_{\Delta x} f[m,\Delta x]. \quad (6)$$

We denote $\overline{\Delta x}[m]$ the solution of this inner loop. Then the outer loop is given by

$$\min g[m] = f\left[m, \overline{\Delta x}[m]\right]. \quad (7)$$

We use a l-BFGS local optimization solver to minimize $g[m]$. We thus need the gradient of $g[m]$. By chain rule, we have

$$\nabla g[m] = \frac{\partial f}{\partial m}[m,\overline{\Delta x}[m]] + \frac{\partial f}{\partial \Delta x}[m,\overline{\Delta x}[m]]\frac{\partial \overline{\Delta x}[m]}{\partial m}. \quad (8)$$

However, the second term of the right hand side cancels by the definition of $\overline{\Delta x}[m]$, and we have

$$\nabla g[m] = \frac{\partial f}{\partial m}[m,\overline{\Delta x}[m]]. \quad (9)$$

Therefore the gradient of $g[m]$ is nothing else than the conventional L^2 FWI problem considering receiver positions corrected by $\overline{\Delta x}[m]$. This gradient is built by the correlation of incident and adjoint wavefields following the adjoint-state strategy (Plessix, 2006). The adjoint fields are computed as the backpropagation in time of the residuals from the corrected receiver positions. The residuals are computed as the difference between observed data (at the true receivers position) and calculated data extracted at the corrected receivers position.

Implementation

The key point to a fast implementation to the method is related to the computation of the receiver position correction vector $\overline{\Delta x}[m]$. We first recognize that the inner problem 6 is separable in each component $\Delta x_{r,s}$. We have

$$f[m,\Delta x] = \sum_{r,s} f_{r,s}[m,\Delta x_{r,s}], \quad (10)$$

$$f_{r,s}[m,\Delta x_{r,s}] = \int_0^T |R(x_r + \Delta x_{r,s})u_s[m](t) - d_{obs,s,r}(t)|^2 dt + \frac{\alpha_{r,s}^2}{2}|\Delta x_{r,s}|^2. \quad (11)$$

Thus we minimize each misfit function $f_{r,s}[m,\Delta x_{r,s}]$ separately with respect to $\Delta x_{r,s}$. As $f_{r,s}[m,\Delta x_{r,s}]$ is not convex in $\Delta x_{r,s}$, we use a global optimization approach. We implement a grid search brute force algorithm: for a given range of variation of $\Delta x_{r,s}$ the misfit function $f_{r,s}[m,\Delta x_{r,s}]$ is evaluated and we select the receiver correction which provides the smallest value.

The regularization weights $\alpha_{r,s}$ are computed as

$$\alpha_{r,s} = \frac{\|d_{obs,r,s}\|_\infty}{\|L\|}, \quad (12)$$

where L is the maximum distance along with the receivers are allowed to be relocated.

To make this strategy affordable in terms of computational cost, we restrain the receiver correction to lateral variations only (horizontal for surface acquisition). As we consider for now 2D FWI problems, this makes each global optimization problem depending on a single degree of freedom. In addition, we store the incident wavefields $u_s(x,t)$ in the vicinity of the acquisition that for each receiver correction we can extract the calculated data at the corrected receiver position without solving a new wave equation. We will see that the resulting computational overhead is negligible compared to a standard L^2 implementation in the 2D examples we investigate here.

ANALYSIS

We first analyze the map of the extended misfit function $f[m,\Delta x]$ in the simplest case of a single couple source/receiver in a transmission configuration. The source and receivers are in two wells 50 m apart from each other, at 50 m depth. The source emits a Ricker pulse with 250 Hz central frequency. We consider a 2D constant density acoustic wave propagation model. The reference medium is homogeneous with a velocity at 2000 m.s^{-1}. We build a reference trace $d_{obs}(t)$ accordingly. We use it to compute the map of $f[m,\Delta x]$ presented in Figure 1 for velocities m and receiver position correction Δx ranging from 1000 m.s^{-1} to 3000 m.s^{-1} and -40 m to +40 m respectively. We see in Figure 1 that the function $f[m,\Delta x]$ is not convex. The minimum at $m = 2000$ m.s^{-1} and $\Delta x = 0$ m is located within a narrow valley of attraction surrounded by two important barriers. However, the profile of the misfit function $g[m]$ is almost convex in m (Fig.2). According to equation 7 this profile is extracted for the minimum value reached in Δx

© 2020 Society of Exploration Geophysicists

10.1190/segam2020-3425103.1

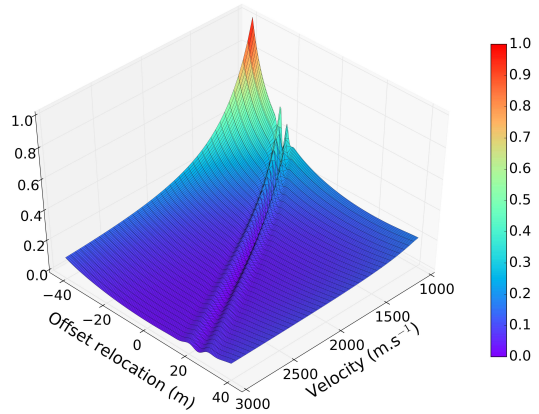

Figure 1: Misfit map of $f[m,\Delta x]$ in the single couple source/receiver transmission case. The function is not convex. The minimum is located in a narrow valley of attraction at position $\Delta x = 0$ and $m = 2000$ m.s^{-1}.

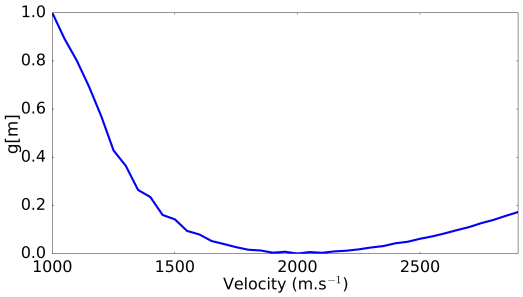

Figure 2: Misfit profile of $g[m]$ in the single couple source/receiver transmission case, depending on the range set for the relocation The function is almost convex in m and reaches its minimum at $m = 2000$ m.s^{-1}.

for each value of m. This is the motivation for the receiver-extension method studied here: if this global minimum in Δx can be computed efficiently, then we can define a function $g[m]$ which might present less local minima.

Note that if the velocity is too fast compared to the true one, the correction requires to move the receivers away from the sources, potentially outside the computational box. One can extend the computational box in this case. Another preferable option is to work with initial models slower than expected true velocities, such that receiver corrections bring receivers closer to the sources, and gradually correct the medium until the receivers have reached their true position.

Next, we illustrate the gradient building in a similar cross-hole configuration with an array of receivers in the right well instead of a single receiver (Fig.3). The gradient of $g[m]$ is compared with the conventional L^2 gradient. The computation is done in a medium slower than the true one, and the receiver position correction bring the receivers closer to the source, forming a semi circle around the source. The residuals are backpropagated from the receivers at their corrected position, which explains the gradient shape. Although interesting, these results are obtained for transmission acquisition. Surface acquisition configuration is of more interest for seismic exploration. In this case, we expect that relocating horizontally the receivers can account for kinematic mismatch of diving waves, hence help also to yield a more convex misfit function in terms of velocity. To test this, we consider a linearly increasing velocity model $v_P(z) = v_{P,0} + \gamma z$, similarly as in Mulder and Plessix (2008). We consider a reference model for $v_{P,0}$=2000 m.s^{-1} and $\gamma = 0.7$s^{-1}. We build reference data in this model considering 2D constant density acoustic wave propagation, a domain size of 17 km by 3.5 km and a fixed-spread surface acquisition at 50 m depth with a single source at $x = 8.5$ km and 170 receivers each 100 m. We use this reference data to compute misfit function maps depending on $v_{P,0}$ and γ as shown in Figure 4. As can be seen, the misfit function $g[m]$, although not fully convex, exhibits a much smoother landscape with a large valley of attraction towards the global minimum, compared with the L^2 misfit function.

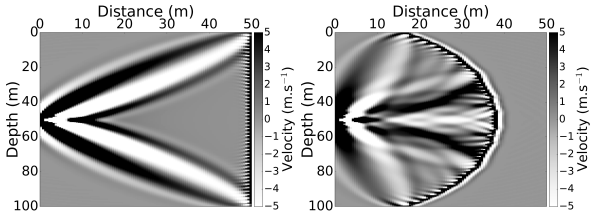

Figure 3: Velocity gradients in a transmission case with one source and one array of receivers. Left: standard L^2 gradient. Right: gradient using the receiver-extension method. To compensate for the kinematic mismatch, the receivers have been relocalized closer from the source, forming a semi-circular shape. The residuals are backpropagated from the corrected positions to assemble the gradient.

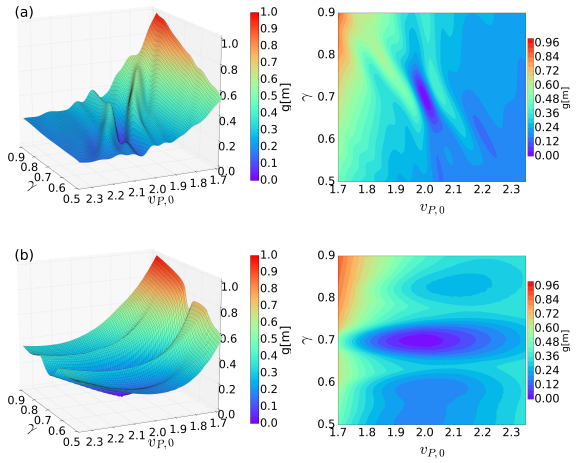

Figure 4: Misfit maps and contours for a linearly increasing in depth velocity model depending on the background velocity $v_{P,0}$ and the velocity gradient γ. The L^2 misfit function (a) presents numerous local minima. The landscape of $g[m]$ (b) appears relatively smoother with a large valley of attraction toward the global minimum and two local minima on both sides.

A FIRST VALIDATION: MARMOUSI II MODEL

Encouraged by these results, we present a first validation of our approach on a modified version of the Marmousi II model

© 2020 Society of Exploration Geophysicists

10.1190/segam2020-3425103.1

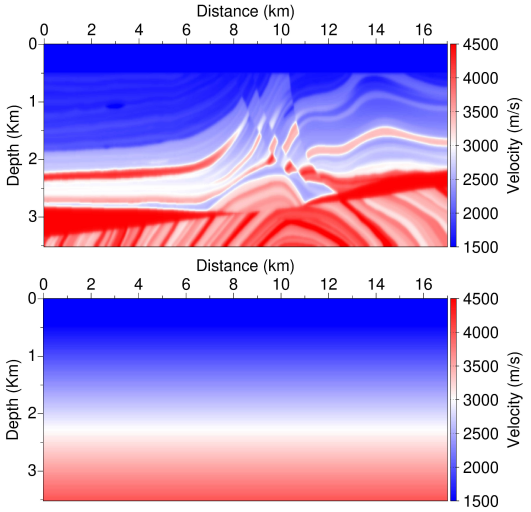

Figure 5: Exact Marmousi II model (top) and 1D initial model (bottom).

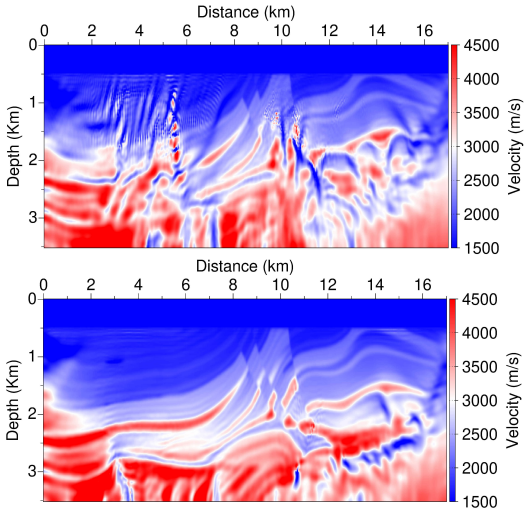

Figure 6: Reconstructed models using L^2 approach (top) and receiver-extension approach (bottom). The L^2 reconstruction is severely cycle skipped while the reconstruction by the receiver-extension method is satisfactory in the mainly illuminated zone.

(Martin et al., 2006). The model is 17 km long and 3.5 km deep. A fixed spread acquisition is used with 128 sources located each 125 m and 170 receivers located each 100 m from 50 m to 16.95 km. We build reference data using a high pass filter Ricker function with no signal below 2.5 Hz. We consider a 1D initial model linearly increasing in depth from 1500 m.s^{-1} to 4000 m.s^{-1} (Fig 5). We implement a simple vertical scaling linear in depth as a preconditioner. The results obtained using a L^2 approach and the extended-receiver approach are presented in Figure 6. We see the severe cycle skipping due to the 1D initial model and the absence of low-frequencies in the data when the L^2 approach is used. In comparison, the extended-receiver approach mitigates efficiently the cycle skipping: the reconstruction is quite close from the true model in the zone of main illumination (artifacts are still present on the lateral and bottom edges).

To complement, we present in Figure 7 the evolution of the average receiver position correction $\sum_r |\Delta x_{r,s}|/N_r$ along the iterations for three shot positions: 0.05 km, 8.5 km, 16.95 km. As expected, the corrections are larger at the beginning of the iterations and reach almost 0 at the end of the inversion. In terms of computation cost, we note that the gradient building step represents an approximate 1.5 s increase, on a total of 36 s for the conventional L^2 approach, hence an approximate 4.1 % increase. The computational overhead is thus negligible in this 2D configuration.

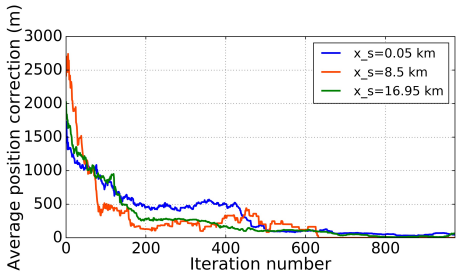

Figure 7: Evolution of the average receiver position correction $\sum_r |\Delta x_{r,s}|/N_r$ along the inversion iterations (outer loop) for three shots locations: 0.05 km, 8.5 km and 16.95 km.

CONCLUSION AND PERSPECTIVES

Our receiver extension approach to mitigate cycle skipping in FWI is adapted to time-domain FWI, contrary to WRI and volume-based MSWI. The receiver position is added as a free parameter, in a nested loop approach. The outer loop updates for the velocity using local optimization while the inner loop updates the receiver positions using grid search. Preliminary results show how this approach can mitigate cycle skipping: the misfit function minimized in the outer loop exhibits a more convex profile with respect to wave velocity. Many questions remain opened, the first being the extension to 3D imaging. Assuming the receivers can move only along the surface, a rough computational complexity analysis of the global optimization loop gives, for this 2D study $O(N_{t,Nyquist} \times N^2)$, while the modeling loop is in $O(N_{t,CFL} \times N^2)$. Here N is the average number of discretization points in one spatial direction, while $N_{t,Nyquist}$ (resp. $N_{t,CFL}$) is the number of discretization points in time according to Nyquist sampling (resp. according to the CFL condition). A straightforward implementation in 3D would give $O(N_{t,Nyquist} \times N^4)$ against $O(N_{t,CFL} \times N^3)$ for the modeling, which is less favorable. However, the receiver position to be investigated during the global optimization loop could be restricted in a vicinity of the true position, at least after a certain number of iterations of the outer loop, as suggested by Figure 7

ACKNOWLEDGMENTS

This study was partially funded by the SEISCOPE consortium (`https://seiscope2.osug.fr`), sponsored by AKERBP, CGG, CHEVRON, EQUINOR, EXXON-MOBIL, JGI, PETROBRAS, SCHLUMBERGER, SHELL, SINOPEC, SISPROBE and TOTAL. This study was granted access to the HPC resources of CIMENT infrastructure (`https://ciment.ujf-grenoble.fr`) and CINES/IDRIS/TGCC under the allocation 046091 made by GENCI.

© 2020 Society of Exploration Geophysicists

10.1190/segam2020-3425103.1

REFERENCES

Aghamiry, H., A. Gholami, and S. Operto, 2019, Improving full-waveform inversion by wavefield reconstruction with alternating direction method of multipliers: Geophysics, **84**, no. 1, R139–R162, doi: https://doi.org/10.1190/geo2018-0093.1.

Aghamiry, H., A. Gholami, and S. Operto, 2020, Accurate and efficient wavefield reconstruction in the time domain: Geophysics, **85**, no. 2, A7–A12, doi: https://doi.org/10.1190/geo2019-0535.1.

Huang, G., R. Nammour, and W. W. Symes, 2018a, Source-independent extended waveform inversion based on space-time source extension: Frequency-domain implementation: Geophysics, **83**, no. 5, R449–R461, doi: https://doi.org/10.1190/geo2017-0333.1.

Huang, G., R. Nammour, and W. W. Symes, 2018b, Volume source-based extended waveform inversion: Geophysics, **83**, no. 5, R369–R387, doi: https://doi.org/10.1190/geo2017-0330.1.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Conference on Inverse Scattering: Theory and Application, 206–220.

Luo, S., and P. Sava, 2011, A deconvolution-based objective function for wave-equation inversion: 81st Annual International Meeting, SEG, Expanded Abstracts, 2788–2792, doi: https://doi.org/10.1190/1.3627773.

Luo, Y., and G. T. Schuster, 1991, Wave-equation traveltime inversion: Geophysics, **56**, 645–653, doi: https://doi.org/10.1190/1.1443081.

Martin, G. S., R. Wiley, and K. J. Marfurt, 2006, Marmousi2: An elastic upgrade for Marmousi: The Leading Edge, **25**, 156–166, doi: https://doi.org/10.1190/1.2172306.

Métivier, L., R. Brossier, Q. Mérigot, and E. Oudet, 2019, A graph space optimal transport distance as a generalization of Lp distances: Application to a seismic imaging inverse problem: Inverse Problems, **35**, 085001, doi: https://doi.org/10.1088/1361-6420/ab206f.

Mulder, W., and R. E. Plessix, 2008, Exploring some issues in acoustic full waveform inversion: Geophysical Prospecting, **56**, 827–841, doi: https://doi.org/10.1111/j.1365-2478.2008.00708.x.

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**, 495–503, doi: https://doi.org/10.1111/j.1365-246X.2006.02978.x.

Symes, W. W., 2008, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, 765–790, doi: https://doi.org/10.1111/j.1365-2478.2008.00698.x.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266, doi: https://doi.org/10.1190/1.1441754.

van Leeuwen, T., and F. Herrmann, 2016, A penalty method for PDE-constrained optimization in inverse problems: Inverse Problems, **32**, 1–26.

van Leeuwen, T., and F. J. Herrmann, 2013, Mitigating local minima in full-waveform inversion by expanding the search space: Geophysical Journal International, **195**, 661–667, doi: https://doi.org/10.1093/gjihttps://doi.org//ggt258.

Wang, C., D. Yingst, P. Farmer, and J. Leveille, 2016, Full-waveform inversion with the reconstructed wavefield method: 86th Annual International Meeting, SEG, Expanded Abstracts, 1237–1241, doi: https://doi.org/10.1190/segam2016-13870317.1.

Warner, M., and L. Guasch, 2016, Adaptive waveform inversion: Theory: Geophysics, **81**, no. 6, R429–R445, doi: https://doi.org/10.1190/geo2015-0387.1.

Yang, Y., and B. Engquist, 2018, Analysis of optimal transport and related misfit functions in full-waveform inversion: Geophysics, **83**, no. 1, A7–A12, doi: https://doi.org/10.1190/geo2017-0264.1.

© 2020 Society of Exploration Geophysicists