Normalized nonzero-lag crosscorrelation elastic full-waveform inversion

Zhendong Zhang¹, Tariq Alkhalifah¹, Zedong Wu¹, Yike Liu², Bin He², and Juwon Oh³

ABSTRACT

Full-waveform inversion (FWI) is an attractive technique due to its ability to build high-resolution velocity models. Conventional amplitude-matching FWI approaches remain challenging because the simplified computational physics used does not fully represent all wave phenomena in the earth. Because the earth is attenuating, a sample-by-sample fitting of the amplitude may not be feasible in practice. We have developed a normalized nonzero-lag crosscorrelation-based elastic FWI algorithm to maximize the similarity of the calculated and observed data. We use the first-order elastic-wave equation to simulate the propagation of seismic waves in the earth. Our proposed objective function emphasizes the matching of the

phases of the events in the calculated and observed data, and thus, it is more immune to inaccuracies in the initial model and the difference between the true and modeled physics. The normalization term can compensate the energy loss in the far offsets because of geometric spreading and avoid a bias in estimation toward extreme values in the observed data. We develop a polynomial-type weighting function and evaluate an approach to determine the optimal time lag. We use a synthetic elastic Marmousi model and the BigSky field data set to verify the effectiveness of the proposed method. To suppress the short-wavelength artifacts in the estimated S-wave velocity and noise in the field data, we apply a Laplacian regularization and a total variation constraint on the synthetic and field data examples, respectively.

INTRODUCTION

Full-waveform inversion (FWI) is instrumental for building a high-resolution model for the subsurface (Virieux and Operto, 2009). Conventional FWI attempts to match the waveforms of the predicted data and those observed in the field in amplitude and phase. However, it has inherent weaknesses in its application to field data. The L_2 -norm objective function requires that the initial model should be kinematically accurate; otherwise, the inverse problem yields models corresponding to local minima when using a gradient-based optimization. In practical situations, the initial model cannot be guaranteed to generate kinematically accurate data. There are several strategies to build a kinematically accurate initial model: (1) Match the kinematic information in data or model domains in a moveout analysis approach (Alkhalifah, 1997; Sun and Alkhalifah, 2017), migration velocity analysis using differential semblance optimization (Symes and Carazzone, 1991; Shen et al.,

2005), reflection-based waveform inversion (Xu et al., 2012; Wu and Alkhalifah, 2015; Guo et al., 2017), or full traveltime inversion (Luo et al., 2016). (2) Compare arrivals in a global sense (tomography). The L_2 -norm objective function requires that the maximum mismatch of the predicted and observed data should not exceed a half-cycle for all events; otherwise, the adjoint source is cycle skipped. The correlation-based objective function (Van Leeuwen and Mulder, 2008, 2010; Routh et al., 2011; Choi and Alkhalifah, 2012; Chi et al., 2015; Wu and Alkhalifah, 2018; Zhang et al., 2018a) or adaptive FWI (Warner and Guasch, 2014) can compare the arrivals globally within a time window and in some cases are free of cycle skipping. (3) Find a simplified representation of the complex data. The philosophy is to find equivalent data that preserve the key features of the original one but have fewer oscillations in time. The optimal transport approach (Métivier et al., 2016; Yang et al., 2016), the Fourier domain method (Solano et al., 2013), and the skeletonized dispersion curve inversion (Zhang et al., 2015,

Manuscript received by the Editor 2 February 2018; revised manuscript received 25 August 2018; published ahead of production 03 October 2018; published online 23 November 2018.

¹King Abdullah University of Science and Technology, Department of Physical Science and Engineering, Thuwal 23955-6900, Saudi Arabia. E-mail: zhendong.zhang@kaust.edu.sa; tariq.alkhalifah@kaust.edu.sa; zedong.wu@kaust.edu.sa.

²Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China. E-mail: ykliu@mail.iggcas.ac.cn; hebin@mail.iggcas.ac.cn.

³Chonbuk National University, Department of Resources and Energy Engineering, 567, Baekje-daero, Deokjin-gu, Jeonju-si, Jeollabuk-do 54896, Republic of Korea. E-mail: juwon24@jbnu.ac.kr.

© 2019 Society of Exploration Geophysicists. All rights reserved.

2016) fall into this category. (4) Joint inversion. Using different physical models as regularization can reduce the nonlinearity of the inverse problem (Kang et al., 2016). Early-arrival traveltime and waveform joint inversion are widely used in practical applications of FWI (Zhang and Chen, 2014; Zhou et al., 2015). Besides, incorporating gravity, electromagnetic, and well information into regular FWI is drawing more and more attention. As an amplitude-matching method, the L_2 -norm objective function also requires that the estimated source wavelet should have accurate magnitudes and the observed data are free of noise. Besides, the real earth is at least attenuative; neither acoustic nor elastic modeling method can provide accurate simulations of the waves in the earth. As a result, phase- or traveltime-matching objective functions are more feasible in the application.

Crosscorrelation-type objective functions emphasize matching the phases of the predicted data and the observed data, and thus they are insensitive to amplitude issues arising from the simplified physics we use to represent the medium (Choi and Alkhalifah, 2012). Zero-lag crosscorrelation still has the potential to have local minima. Van Leeuwen and Mulder (2008, 2010) introduce a non-zero-lag crosscorrelation to make the objective function even more convex and verify its effectiveness on transmission data. Wu and Alkhalifah (2018) improve the nonzero-lag crosscorrelation approach by adding a normalization term and a selective operation to make the objective function even more convex. Instead of selecting the positive values in crosscorrelation, we propose a squared and an envelope crosscorrelation objective function (Chi et al., 2014; Wu et al., 2014; Oh and Alkhalifah, 2017), which preserves the advantage of convexity. The crosscorrelation-based objective functions with a properly selected time lag have only one local minimum for single-event data. However, reflection data often include far more than one event, in which case the convexity of the objective function is also influenced by the time lag and the penalty function. A large time lag can generate crosstalks between consecutive events, but a small time lag can lead the inversion to converge to a local minimum. The choice of an optimal time lag was often related to the wavelength (Van Leeuwen and Mulder, 2008, 2010). The weighting function used to penalize the data with fewer similarities is usually a linear or a Gaussian function. A novel polynomial-type weighting

function is introduced that requires fewer input parameters and is guaranteed to be smoothly varying within the time lag.

In current practice, due to an inherent trade-off between parameters, for example, P-wave velocity and density, the density model is usually not updated. Besides, due to the varying wavefield wavelengths, the estimated P- and S-wave velocities can have different spatial resolutions. Common ways to reduce the null space in multi-parameter inversion include choosing the proper parameterization (Alkhalifah and Plessix, 2014; Oh and Alkhalifah, 2016) and adding additional constraints to the inversion (Asnaashari et al., 2013; Zhang et al., 2017, 2018b). The first- or second-order spatial derivatives (i.e., Laplacian) can be used to generate a smooth estimation of the subsurface and also can suppress some short-wavelength artifacts in the estimate of S-wave velocity. The total variation (TV) regularization can generate blocky models, and it also can suppress the noise in the observed data (Guitton, 2012; Lin and Huang, 2014).

This paper is divided into four sections. After the “Introduction,” the theory of the normalized nonzero lag crosscorrelation objective function is described in which we analyze the convexity of the proposed objective function and its dependency on the selected time lag. We use a synthetic elastic model and the BigSky field data to verify the effectiveness of the proposed method and draw the conclusions from the numerical tests.

THEORY

Objective functions intend to measure the mismatch between the predicted and the observed data. One of the most intuitive measurements is the L_2 -norm distance, which is given by

$$\phi(m) = \|d^p - d^o\|^2, \quad (1)$$

where ϕ measures the differences and d^p and d^o are the predicted and observed data, respectively.

The inverse problem is constrained by the first-order elastic-wave equation, which is given by (Virieux, 1986; Vigh et al., 2014)

$$\begin{pmatrix} \rho \mathbf{I}_3 & 0 \\ 0 & \mathbf{C}^{-1} \end{pmatrix} \frac{\partial \Psi}{\partial t} - \begin{pmatrix} 0 & E^T \\ E & 0 \end{pmatrix} \Psi - \mathbf{f} = 0, \quad (2)$$

where $\Psi = (v_1, v_2, v_3, \sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6)$ is a vector containing three particle velocities and six stresses, $\mathbf{I}_3$ indicates a 3×3 identity matrix, $\mathbf{C}$ is the stiffness matrix, E denotes the space differentiation, and $\mathbf{f}$ is the source.

Due to the oscillatory nature of seismic waves, the L_2 -norm objective inversion suffers from cycle skipping when the mismatches between the predicted and the observed data are larger than the half-cycle as illustrated in Figure 1. The predicted data (the green curves) are shifted by one cycle ($\pm\lambda$), a half-cycle ($\pm 0.5\lambda$), and zero (as indicated by τ on the label) compared with the observed data (the red curves). The residual wavefield (the blue curves) that is calculated by a wiggle-to-wiggle subtraction of these two data sets can have a quite different waveform with the actual wavelet for large shifts. The corresponding normalized data misfit (indicated by ϕ) has three local minima, which are at $\tau = \pm\lambda, 0$. As

Figure 1. Cycle skipping in L_2 -norm objective functions. The predicted data (the green curves) are shifted by one wavelength (a and e) compared with the observed data (the red curves), and the subtractions have more than one residual wavelet (the blue curves). The critical points of cycle skipping are a half-wavelength (b and d). One straightforward remedy is comparing two data with shifts.

expected, the half-wavelength ($\tau = \pm 0.5\lambda$) is the critical value for cycle skipping. The initial model is not guaranteed to generate predicted data, which should satisfy the half-wavelength criterion in practice. A natural remedy to this problem is to compare two events with shifts in time. Theoretically, the critical value for cycle skipping in this approach is enlarged by the time window used.

We propose a normalized nonzero-lag crosscorrelation-based objective function that aims to maximize the similarity of the predicted and observed data. The proposed objective function compares the predicted and observed data globally within the range of predefined time lags ($\pm\tau$). Due to the oscillatory nature of seismic waves (positive and negative), a simple crosscorrelation has negative contributions to the objective function (Wu and Alkhalifah, 2018). To eliminate the conflicts of negative and positive contributions, we proposed two choices of the measurements, which are envelope (equation 3) and squared (equation 4) crosscorrelation. The envelope version of the nonzero lag crosscorrelation is given by

$$\phi_e(m) = -\iiint W(\tau) \frac{C[E(d^p(t)), E(d^o(t)), \tau]}{\sqrt{\int E^2(d^p(t))dt} \sqrt{\int E^2(d^p(t+\tau))dt}} \times d\tau dx_r dx_s dt, \quad (3)$$

where τ is the time lag used for the crosscorrelation, $E(d) = \sqrt{d^2 + H(d)^2}$ denotes the envelope operator, $H(d)$ is the Hilbert transform of d , $C[d^p, d^o, \tau] = \int d^p(t) d^o(t+\tau) dt$ denotes the crosscorrelation in time, and $W(\tau)$ is the weighting function, which will be explored later. The normalization terms can compensate the energy loss at far offsets.

The squared version of the proposed objective function is given by

$$\phi_s(m) = -\frac{1}{2} \iiint \left(W(\tau) \frac{C[d^p(t), d^o(t), \tau]}{\sqrt{\int d^p(t)^2 dt} \sqrt{\int d^o(t+\tau)^2 dt}} \right)^2 \times d\tau dx_r dx_s dt. \quad (4)$$

We introduce a polynomial-type weighting function that satisfies the following boundary conditions: $W|_{\pm\tau} = 0$; $W|_0 = 1$;

Figure 2. Proposed weighting function. It is determined by the predefined maximum time lag and reduces to zero at the shifting boundaries.

$W'|_{\pm\tau} = 0$; $W'|_0 = 0$. Figure 2 shows one example of the weighting function. The Gaussian-type weighting function needs another parameter, the standard derivation of the Gaussian function, to decide the shape of the penalty function. A narrow Gaussian function will remove most of the coherent events, and a wide Gaussian function cannot decay fast enough to reach zero at the predefined maximum time lag. The proposed weighting function does not need an additional parameter, and it is guaranteed to be smooth within the selected maximum time lag. In addition, the computational cost is bounded by the predefined maximum time lag τ .

The proposed objective functions (equations 3 and 4) have only one local minimum for a single arrival (i.e., first arrivals) in theory. To verify its effectiveness, we plot the normalized data misfit versus velocity errors for the L_2 norm (equation 1), envelope (equation 3), and squared crosscorrelation (equation 4) in Figure 3. The model used here is a homogeneous model with a maximum of 50% velocity error. The sources are located in the top of the model, and the receivers are deployed in the bottom of the model, in which case only transmission data are recorded as illustrated in Figure 3a. Figure 3b indicates that for the same parameter set, the L_2 -norm objective function has more than one local minimum; the proposed objective functions have only one local minimum. The envelope-crosscorrelation objective function varies smoothly as a function of velocity error as expected. The squared one has the same sharp curvature as the L_2 -norm objective function. To preserve the advantage of high resolution, we use the squared one in the following examples.

Figure 3. (a) Schematic diagram of the geometry and (b) normalized misfit versus velocity errors for transmission data. The L_2 -norm objective function has more than one local minimum, whereas the crosscorrelation-based objective functions have only one local minimum in such a transmission acquisition system.

In practice, the observed data include more than one arrival, and the maximum time lag τ should be delicately selected to avoid crosstalks between nearby arrivals and allow the proper intersection of certain arrivals in the predicted and observed data. There is no explicit formula for an optimal choice of τ , but it can be roughly decided by the autocorrelation of the observed data. We use a revised Marmousi model (more details can be found in the numerical example) to illustrate this property. The actual model here is the revised Marmousi model, and the initial model is generated by $v_{\text{input}} = \alpha v_{\text{true}} + 1.5(1 - \alpha)$; the initial model has the maximum and minimum errors when $\alpha = 0.1$ and $\alpha = 1$, respectively. Input velocities are linear combinations of the actual Marmousi model and a constant value that can generate multiple reflections in the data. Figure 4 shows that smaller time lags ($\tau = 0.0$ s and $\tau = 0.05$ s) cannot guarantee sufficient intersections of the predicted and observed data, in which case the objective function performs similar to wiggle-to-wiggle measurements; large time lags ($\tau = 1.3$ s) introduce more crosstalk between nearby arrivals. It is easy to understand that an optimal time lag should allow a maximum shift before neighboring arrivals intersect with each other. From the autocorrelation of the observed data, we can roughly decide the optimal maximum time lag, which should be just shorter than the second peak of the autocorrelation.

An elastic approximation to the earth reduces the errors caused by the simplified physics used in FWI. However, estimating elastic parameters makes the inverse problem even more ill-posed. Regularization techniques can reduce the null space of the ill-posed inverse problem (Asnaashari et al., 2013; Wang et al., 2013; Duan and Sava, 2016). With some assumptions made of the subsurface (such as smoothness or sharpness), an additional penalty or a constraint can be added to the data-fitting objective or the estimated parameters, respectively. Usually the first- or second-order Tikhonov regularization can generate a smooth estimation of the subsurface (Tikhonov, 1963), and a TV regularization can preserve sharp boundaries in the estimated models. The first-order Tikhonov regularization is given by

$$R_I(m) = \|\nabla m\|^2, \quad (5)$$

Figure 4. Normalized misfit versus τ for multiple arrivals. The optimal time lag should avoid local minima.

where the vector m denotes the estimated model in each iteration. It provides a smooth model within the optimization step to fit the data.

Guitten (2012) proposes a L_1 -norm model penalty to preserve sharp edges and contrasts for FWI. From a minimization point of view, a smoothed TV prior is more appropriate because of its convexity, which is given by

$$R_{TV}(m) = \int \sqrt{\epsilon^2 + \|\nabla m\|^2} dx, \quad (6)$$

where the gradient is given by $g_m(R_{TV}) = \text{div}(\nabla m / \sqrt{\epsilon^2 + \|\nabla m\|^2})$ and ϵ is a predefined parameter that makes equation 6 differentiable. A large ϵ transforms the smoothed TV gradient to a scaled Laplacian (scaled by $1/\epsilon$).

In this study, the estimated S-wave velocity has high-wavenumber artifacts caused by the short wavelength of S-wave wavefields. To suppress these artifacts, we add a Laplacian penalty term in the objective function. A TV regularization is applied in the field data test because of its effectiveness in suppressing noise. The exact TV regularization ($\epsilon = 0$ in equation 6) is implemented by an efficient primal-dual hybrid gradient algorithm (Zhu and Chan, 2008). A more advanced facies-based regularization is also applicable if the prior information is available (Zhang et al., 2017, 2018b).

The gradient of the misfit (objective) function in equation 4 can be calculated from the adjoint-state method:

$$g_s(m) = \left(\frac{\partial d^p(t)}{\partial m} \right)^T \int_{\tau} W(\tau)^T W(\tau) \frac{C[d^p(t), d^o(t), \tau]}{\sqrt{\int d^p(t)^2 dt} \sqrt{\int d^o(t + \tau)^2 dt}} \left(\frac{d^o(t + \tau)}{\sqrt{\int d^p(t)^2 dt} \sqrt{\int d^o(t + \tau)^2 dt}} - \frac{d^p(t) \int d^p(t) d^o(t + \tau) dt}{\sqrt{\int d^p(t)^2 dt}^3 \sqrt{\int d^o(t + \tau)^2 dt}} \right) d\tau. \quad (7)$$

The model is updated iteratively using the L-Broyden–Fletcher–Goldfarb–Shanno (BFGS) method (Liu and Nocedal, 1989), which is written as

$$\mathbf{m} = \mathbf{m}_0 - \lambda \mathbf{H} \mathbf{g}, \quad (8)$$

where λ is the step length calculated using a standard line-search method and $\mathbf{H}$ is the approximated Hessian matrix.

NUMERICAL EXAMPLES

Modified elastic Marmousi model

We use a modified Marmousi model to verify the effectiveness of the proposed objective function. The actual S-wave velocity is generated by setting $V_S = V_P / \sqrt{3} + 0.1(V_P - 2.4)$ (Figure 5b). Two low-velocity zones (indicated by the dashed circles in Figure 5a) in the P-wave velocity are intended to test whether there is crosstalk between inverted V_P and V_S . There are 110 explosive sources and 660 receivers placed at constant increments. The recorded data are 2C particle velocities. The staggered finite-difference method is implemented to solve the elastic-wave equation (Virieux, 1986).

Initial models in Figure 6 are 1D linear gradient velocity models, which are far from the exact ones ($V_S = V_P/\sqrt{3}$).

The source wavelet is a Ricker wavelet (peak frequency is 5 Hz) without frequencies less than 3 Hz, in which case a conventional L_2 -norm-based FWI cannot converge to a global minimum because of the poor initial model and the relatively high minimum frequency. A mono time lag, $\tau = 0.06s$, is used to get a better initial model for conventional elastic FWI. Only one frequency band, 3–5 Hz, is used for the crosscorrelation-objective (equation 4) inversion to save computational cost. A Laplacian regularization term is added to the original objective function to remove some short-wavelength artifacts. The objective function used in the example is given by

$$\phi = \phi_d + \alpha \|\Delta \mathbf{m}\|^2, \quad (9)$$

where ϕ_d is the measurement of data misfit (i.e., equation 1 or 4), $\Delta \mathbf{m}$ is the Laplacian of the model vector (V_P and V_S), and α controls

the contribution of the regularization term. The regularization term has a 20% of the data misfit (ϕ_d) contribution to the total objective function (ϕ) in this example.

Figure 7 shows inverted models using the proposed objective function (equation 4). One of the predefined low-velocity zones in V_P can be captured. The inverted V_S has higher spatial resolution. These two models are used as initial models for an L_2 -norm objective elastic FWI to improve their resolutions. The final inversion results are shown in Figure 8. Two low-velocity zones in V_P have been captured. The estimated parameters without regularization are shown in Figure 9 for comparison. The Laplacian regularization in equation 9 can suppress the wiggles caused by the short-wavelength nature of S-waves and also can improve the quality of inverted V_P and V_S . As an intuitive quality control (QC) factor, vertical profiles across the low-velocity zones (indicated by the red arrows) are plotted in Figures 10 and 11. It is evident that the inverted velocities match the model well.

Figure 5. True velocity models: (a) V_P and (b) V_S . The dashed circles indicate low-velocity zones in V_P .

Figure 6. Initial velocity models: (a) V_P and (b) V_S . They are 1D linearly increasing models.

Figure 7. Inverted velocity models of the proposed squared crosscorrelation (equation 4): (a) V_P and (b) V_S . The low-velocity zones are showing in V_P .

BigSky field data

The field data used to verify the effectiveness of our proposed method come from the BigSky Carbon Sequestration Partnership, which is a U.S. CO₂ storage project (Clochard et al., 2017). The data were collected by a 3D multicomponent seismic survey (3D-9C) with a minimum frequency of 15 Hz, which is a challenge for conventional FWI. The 3D survey geometry is plotted in Fig-

ure 12, and a 2D inline across the central area is selected to allow us to use high frequencies. There is a well at the edge of the survey area that provides the initial velocities and references to the inverted velocities. A raw data set of the vertical component is plotted in Figure 13. It is preprocessed by applying random noise attenuation and surface wave suppression. Figure 14 shows the preprocessed vertical component of the observed data with a frequency band of 15–30 Hz. The initial velocity models shown in Figure 15 are

Figure 8. Inverted velocity models of L_2 -norm elastic FWI: (a) V_P and (b) V_S . The low-velocity zones have been captured in V_P . V_P and V_S have a similar spatial resolution thanks to the regularization term.

Figure 9. Inverted velocity models of L_2 -norm elastic FWI without the regularization: (a) V_P and (b) V_S . Some wiggles due to the short wavelength of S-waves can be seen in V_S .

Figure 10. Vertical profile across the low-velocity zone at a depth of 0.4 km: (a) V_P and (b) V_S . The velocity zone is well-reconstructed, and there is less crosstalk between V_P and V_S .

Figure 11. Vertical profile across the low-velocity zone at a depth of 0.7 km: (a) V_P and (b) V_S . Most of the structures are captured correctly in the inverted results.

calculated from the provided well log using the Backus theory (Lindsay and Van Koughnet, 2001). We applied a multiscale inversion from low to high frequencies (15–20, 15–30, 15–40, and 15–45 Hz) with a fixed time lag, which is 0.01 s. The TV regularization is implemented as a separate optimization process (Zhu and Chan, 2008). In each frequency band, we first update the velocities until the gradient is small and then we apply TV regularization to the updated velocities (Alkhalifah et al., 2018). These two inversions are repeated for four times in each inversion stage. A strong lateral smooth filter is applied to the gradient because of the weak lateral variation in this area. Figure 16 shows the final estimated velocities using the proposed approach. Two low P-wave velocities zones at depths of 0.7 and 1 km have been captured as shown in the vertical profiles in Figure 17. Notice that the well is far from the survey line. The estimated S-wave velocity is not well-constrained because there are not many converted waves found in the observed data as shown in Figure 14. As a QC factor, we plot the normalized data misfit versus iterations of the first inversion stage in Figure 18. After updating the velocities with several iterations, we apply the TV constraint (the green lines indicate when the TV constraints are applied). The data misfit is overall decreasing, and the TV constraint helps to reduce the misfit even more. We also plot the predicted data from the initial model and the estimated model in Figures 19 and 20 to highlight the improvements. Predicted data from the inverted velocities have similar moveouts with the observed data for the first arrivals. Reflections of the predicted and observed data match. In the field data example, we did not apply a conventional L_2 -norm-based elastic FWI to further improve the estimated velocities because of the weaknesses of L_2 -norm measurements discussed in the paper. To further verify the improvements of the initial velocities, we apply reverse time migration (RTM) to the observed data (Etgen et al., 2009) and obtain the PP- and PS-images. Figure 21 shows the RTM images from the initial velocities (from the well logs), and Figure 22 shows the images using the estimated velocities as migration velocities. These two figures are plotted using the same parameters, and their differences should come from the differences in velocities used for migration. The yellow arrows indicate where the coherency in the images are improved. The PP- and PS-images from the estimated velocities have more continuous reflectors than the one from the initial velocities, which verify the effectiveness of the proposed inversion approach.

Figure 13. One example raw data set, vertical component. There exist strong near-source noise.

Figure 14. Preprocessed observed data, vertical component. Strong surface waves are muted.

Figure 12. The 3D survey geometry. The blue circles and red stars are receivers and sources, respectively. We choose one 2D line for inversion as indicated by the arrow. There is a well at the edge of the survey area.

Figure 15. Initial velocities from the well logs: (a) V_P and (b) V_S .

Figure 16. Inverted velocities from the proposed approach: (a) V_P and (b) V_S .

Figure 17. Vertical profiles of the velocities: (a) V_P and (b) V_S . Notice that the well is far from the survey line.

Figure 18. Reduced data misfit (%) versus iterations at the first stage. The velocities are updated first and then followed by applying the TV regularization. The green line indicates when TV constraint is applied.

Figure 19. Predicted data from the initial velocities, vertical component. It is far from the observed data.

Figure 20. Predicted data from the inverted velocities, vertical component. Some key reflections are matching with the ones in the observed data.

Figure 21. The RTM images using the initial velocities as migration velocities: (a) PP-image and (b) PS-image.

Figure 22. The RTM images using the estimated velocities as migration velocities. The reflectors are more continuous as indicated by the yellow arrows than the ones got from the initial velocities: (a) PP-image and (b) PS-image.

CONCLUSION

The proposed normalized nonzero-lag crosscorrelation objective function is more convex than the conventional L_2 -norm objective function. Nonzero lags allow us to compare arrivals within a pre-defined time window, which extends the limitation of half-cycle difference limitation in conventional L_2 -norm FWI. We proposed a polynomial-type weighting function to penalize the uncorrelated data, which is determined by the time lag only. The proposed objective function has just one local minimum for single arrivals (e.g., the first arrivals). However, it might also have more than one local minimum for reflections. An optimal time lag that guarantees the intersection of particular arrivals and avoids crosstalk between two neighboring arrivals can be determined by the autocorrelation of the observed data. The synthetic and field examples verify the effectiveness of the proposed objective function. The field example also shows that the proposed objective function is relatively immune to noise recorded in the near offsets of the observed data thanks to the normalization term. The estimated velocities from the proposed method can improve the coherency of reflectors in RTM images.

ACKNOWLEDGMENTS

We thank J. Etgen and four anonymous reviewers for improving the quality of the paper. We thank B. Sun and Y. Choi for their helpful discussions. We thank KAUST for its support and specifically the seismic wave analysis group members for their valuable insights. For computer time, this research used the resources of the Supercomputing Laboratory at King Abdullah University of Science and Technology (KAUST) in Thuwal, Saudi Arabia. We thank Vecta Oil and Gas and especially B. Devault for the BigSky data and the helpful discussions.

DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.

REFERENCES

- Alkhalifah, T., 1997, Velocity analysis using nonhyperbolic moveout in transversely isotropic media: *Geophysics*, **62**, 1839–1854, doi: [10.1190/1.1444285](https://doi.org/10.1190/1.1444285).
- Alkhalifah, T., and R.-É. Plessix, 2014, A recipe for practical full-waveform inversion in anisotropic media: An analytical parameter resolution study: *Geophysics*, **79**, no. 3, R91–R101, doi: [10.1190/geo2013-0366.1](https://doi.org/10.1190/geo2013-0366.1).
- Alkhalifah, T., B. Sun, and Z. Wu, 2018, Full model wavenumber inversion: Identifying sources of information for the elusive middle model wavenumbers: *Geophysics*, **83**, no. 6, R597–R610, doi: [10.1190/geo2017-0775.1](https://doi.org/10.1190/geo2017-0775.1).
- 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, doi: [10.1190/geo2012-0104.1](https://doi.org/10.1190/geo2012-0104.1).
- Chi, B., L. Dong, and Y. Liu, 2014, Full waveform inversion method using envelope objective function

- without low frequency data: *Journal of Applied Geophysics*, **109**, 36–46, doi: [10.1016/j.jappgeo.2014.07.010](https://doi.org/10.1016/j.jappgeo.2014.07.010).
- Chi, B., L. Dong, and Y. Liu, 2015, Correlation-based reflection full-waveform inversion: *Geophysics*, **80**, no. 4, R189–R202, doi: [10.1190/geo2014-0345.1](https://doi.org/10.1190/geo2014-0345.1).
- Choi, Y., and T. Alkhalifah, 2012, Application of multi-source waveform inversion to marine streamer data using the global correlation norm: *Geophysical Prospecting*, **60**, 748–758, doi: [10.1111/j.1365-2478.2012.01079.x](https://doi.org/10.1111/j.1365-2478.2012.01079.x).
- Clochard, V., N. Delépine, B. C. Devault, and K. Wangkawong, 2017, Warping of 9C seismic data and its application to 9C stratigraphic pre-stack inversion: A real case study from the big sky CO₂ project (Montana, US): 87th Annual International Meeting, SEG, Expanded Abstracts, 839–843, doi: [10.1190/segam2017-17494693.1](https://doi.org/10.1190/segam2017-17494693.1).
- Duan, Y., and P. Sava, 2016, Elastic wavefield tomography with physical model constraints: *Geophysics*, **81**, no. 6, R447–R456, doi: [10.1190/geo2015-0508.1](https://doi.org/10.1190/geo2015-0508.1).
- Etgen, J., S. H. Gray, and Y. Zhang, 2009, An overview of depth imaging in exploration geophysics: *Geophysics*, **74**, no. 6, WCA5–WCA17, doi: [10.1190/1.3223188](https://doi.org/10.1190/1.3223188).
- Guittou, A., 2012, Blocky regularization schemes for full-waveform inversion: *Geophysical Prospecting*, **60**, 870–884, doi: [10.1111/j.1365-2478.2012.01025.x](https://doi.org/10.1111/j.1365-2478.2012.01025.x).
- Guo, Q., T. Alkhalifah, and Z. Wu, 2017, Velocity building by reflection waveform inversion without cycle-skipping: 79th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.201701003](https://doi.org/10.3997/2214-4609.201701003).
- Kang, P. K., Y. Zheng, X. Fang, R. Wojcik, D. McLaughlin, S. Brown, M. C. Fehler, D. R. Burns, and R. Juanes, 2016, Sequential approach to joint flow-seismic inversion for improved characterization of fractured media: *Water Resources Research*, **52**, 903–919, doi: [10.1002/2015WR017412](https://doi.org/10.1002/2015WR017412).
- Lin, Y., and L. Huang, 2014, Acoustic-and elastic-waveform inversion using a modified total-variation regularization scheme: *Geophysical Journal International*, **200**, 489–502, doi: [10.1093/gji/ggu393](https://doi.org/10.1093/gji/ggu393).
- Lindsay, R., and R. Van Koughnet, 2001, Sequential Backus averaging: Upscaling well logs to seismic wavelengths: *The Leading Edge*, **20**, 188–191, doi: [10.1190/1.1438908](https://doi.org/10.1190/1.1438908).
- Liu, D. C., and J. Nocedal, 1989, On the limited memory BFGS method for large scale optimization: *Mathematical Programming*, **45**, 503–528, doi: [10.1007/BF01589116](https://doi.org/10.1007/BF01589116).
- Luo, Y., Y. Ma, Y. Wu, H. Liu, and L. Cao, 2016, Full-traveltime inversion: *Geophysics*, **81**, no. 5, R261–R274, doi: [10.1190/geo2015-0353.1](https://doi.org/10.1190/geo2015-0353.1).
- Métivier, L., R. Brossier, Q. Mérogot, E. Oudet, and J. Virieux, 2016, Measuring the misfit between seismograms using an optimal transport distance: Application to full waveform inversion: *Geophysical Supplements to the Monthly Notices of the Royal Astronomical Society*, **205**, 345–377, doi: [10.1093/gji/ggw014](https://doi.org/10.1093/gji/ggw014).
- Oh, J., and T. Alkhalifah, 2017, Full waveform inversion with envelope-based global correlation norm-application to real OBC data: 79th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.201701337](https://doi.org/10.3997/2214-4609.201701337).
- Oh, J.-W., and T. Alkhalifah, 2016, Elastic orthorhombic anisotropic parameter inversion: An analysis of parameterization: *Geophysics*, **81**, no. 6, C279–C293, doi: [10.1190/geo2015-0656.1](https://doi.org/10.1190/geo2015-0656.1).
- Routh, P., J. Krebs, S. Lazaratos, A. Baumstein, I. Chikichev, S. Lee, N. Downey, D. Hinkley, and J. Anderson, 2011, Full-wavefield inversion of marine streamer data with the encoded simultaneous source method: 73rd Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.20149730](https://doi.org/10.3997/2214-4609.20149730).
- Shen, P., W. H. Symes, S. Morton, A. Hess, and H. Calandra, 2005, Differential semblance velocity analysis via shot profile migration: 75th Annual International Meeting, SEG, Expanded Abstracts, 2249–2252, doi: [10.1190/1.2148164](https://doi.org/10.1190/1.2148164).
- Solano, C. P., D. Donno, and H. Chauris, 2013, 2D surface wave inversion in the F-K domain: 75th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.20130118](https://doi.org/10.3997/2214-4609.20130118).
- Sun, B., and T. Alkhalifah, 2017, Automatic WEMVA by focusing subsurface offset virtual sources: 79th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.201700837](https://doi.org/10.3997/2214-4609.201700837).
- Symes, W., and J. J. Carazzone, 1991, Velocity inversion by differential semblance optimization: *Geophysics*, **56**, 654–663, doi: [10.1190/1.1443082](https://doi.org/10.1190/1.1443082).
- Tikhonov, A. N., 1963, On the solution of ill-posed problems and the method of regularization: *Doklady Akademii Nauk, Russian Academy of Sciences*, 501–504.
- Van Leeuwen, T., and W. Mulder, 2008, Velocity analysis based on data correlation: *Geophysical Prospecting*, **56**, 791–803, doi: [10.1111/j.1365-2478.2008.00704.x](https://doi.org/10.1111/j.1365-2478.2008.00704.x).
- Van Leeuwen, T., and W. Mulder, 2010, A correlation-based misfit criterion for wave-equation traveltime tomography: *Geophysical Journal International*, **182**, 1383–1394, doi: [10.1111/j.1365-246X.2010.04681.x](https://doi.org/10.1111/j.1365-246X.2010.04681.x).
- Vigh, D., K. Jiao, D. Watts, and D. Sun, 2014, Elastic full-waveform inversion application using multicomponent measurements of seismic data collection: *Geophysics*, **79**, no. 2, R63–R77, doi: [10.1190/geo2013-0055.1](https://doi.org/10.1190/geo2013-0055.1).
- Virieux, J., 1986, P-sv wave propagation in heterogeneous media: Velocity-stress finite-difference method: *Geophysics*, **51**, 889–901, doi: [10.1190/1.1442147](https://doi.org/10.1190/1.1442147).
- Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: *Geophysics*, **74**, no. 6, WCC1–WCC26, doi: [10.1190/1.3238367](https://doi.org/10.1190/1.3238367).
- Wang, C., D. Yingst, J. Bai, J. Leveille, P. Farmer, and J. Brittan, 2013, Waveform inversion including well constraints, anisotropy, and attenuation: *The Leading Edge*, **32**, 1056–1062, doi: [10.1190/le32091056.1](https://doi.org/10.1190/le32091056.1).
- Warner, M., and L. Guasch, 2014, Adaptive waveform inversion: Theory: 84th Annual International Meeting, SEG, Expanded Abstracts, 1089–1093, doi: [10.1190/segam2014-0371.1](https://doi.org/10.1190/segam2014-0371.1).
- Wu, R.-S., J. Luo, and B. Wu, 2014, Seismic envelope inversion and modulation signal model: *Geophysics*, **79**, no. 3, WA13–WA24, doi: [10.1190/geo2013-0294.1](https://doi.org/10.1190/geo2013-0294.1).
- Wu, Z., and T. Alkhalifah, 2015, Simultaneous inversion of the background velocity and the perturbation in full-waveform inversion: *Geophysics*, **80**, no. 6, R317–R329, doi: [10.1190/geo2014-0365.1](https://doi.org/10.1190/geo2014-0365.1).
- Wu, Z., and T. Alkhalifah, 2018, Selective data extension for full-waveform inversion: An efficient solution for cycle skipping: *Geophysics*, **83**, no. 3, R201–R211, doi: [10.1190/geo2016-0649.1](https://doi.org/10.1190/geo2016-0649.1).
- Xu, S., D. Wang, F. Chen, Y. Zhang, and G. Lambare, 2012, Full waveform inversion for reflected seismic data: 74th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.20148725](https://doi.org/10.3997/2214-4609.20148725).
- Yang, Y., B. Engquist, J. Sun, and B. D. Froese, 2016, Application of optimal transport and the quadratic Wasserstein metric to full-waveform inversion: arXiv preprint arXiv:1612.05075.
- Zhang, J., and J. Chen, 2014, Joint seismic traveltime and waveform inversion for near surface imaging: 84th Annual International Meeting, SEG, Expanded Abstracts, 934–937, doi: [10.1190/segam2014-1501.1](https://doi.org/10.1190/segam2014-1501.1).
- Zhang, Z., T. Alkhalifah, and Z. Wu, 2018a, Normalized nonzero-lag cross-correlation elastic full waveform inversion: 80th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.201800889](https://doi.org/10.3997/2214-4609.201800889).
- Zhang, Z., E. Z. Naeini, and T. Alkhalifah, 2017, Facies constrained elastic full waveform inversion: 79th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: [10.3997/2214-4609.201700719](https://doi.org/10.3997/2214-4609.201700719).
- Zhang, Z., G. T. Schuster, Y. Liu, S. M. Hanafy, and J. Li, 2016, Wave equation dispersion inversion using a difference approximation to the dispersion-curve misfit gradient: *Journal of Applied Geophysics*, **133**, 9–15, doi: [10.1016/j.jappgeo.2016.07.019](https://doi.org/10.1016/j.jappgeo.2016.07.019).
- Zhang, Z.-D., T. Alkhalifah, E. Z. Naeini, and B. Sun, 2018b, Multiparameter elastic full waveform inversion with facies-based constraints: *Geophysical Journal International*, **213**, 2112–2127, doi: [10.1093/gji/ggy113](https://doi.org/10.1093/gji/ggy113).
- Zhang, Z.-D., Y. Liu, and G. Schuster, 2015, Wave equation inversion of skeletonized surface waves: 85th Annual International Meeting, SEG, Expanded Abstracts, 2391–2395, doi: [10.1190/segam2015-5805253.1](https://doi.org/10.1190/segam2015-5805253.1).
- Zhou, W., R. Brossier, S. Operto, and J. Virieux, 2015, Full waveform inversion of diving & reflected waves for velocity model building with impedance inversion based on scale separation: *Geophysical Journal International*, **202**, 1535–1554, doi: [10.1093/gji/ggv228](https://doi.org/10.1093/gji/ggv228).
- Zhu, M., and T. Chan, 2008, An efficient primal-dual hybrid gradient algorithm for total variation image restoration: UCLA CAM Report, 34.