A robust full waveform inversion based on a shifted correlation of the envelope of wavefields

Zedong Wu, Faisal Alonaizi**, Tariq Alkhalifah*, Zhendong Zhang*, and and Majed Almalki***
** King Abdullah University of Science and Technology, ** King Abdulaziz City for Science and Technology*

SUMMARY

The standard full waveform inversion (FWI) attempts to minimize the difference between the observed and modeled data. When the initial velocity is kinematically accurate, FWI often converges to to the best velocity model, usually of a high-resolution nature. However, when the modeled data using an initial velocity is far from the observed data, conventional local gradient based methods converge to a solution near the initial velocity instead of the global minimum. This is also known as the cycle skipping problem, which results in a zero correlation when observed and modeled data are not correlated. To reduce the cycle-skipping problem, we can compare the envelope of the modeled and observed data instead of the original data. However, when the initial velocity is not good enough, the correlation of the envelope of the modeled and observed data do not contribute accurately to the gradient. To mitigate this issue, we suggest to maximize not only the zero-lag correlation of the envelope but also the non-zero-lag correlation of the envelope. A weighting function, which has its maximum value at zero lag and decays away from zero lag, is proposed to balance the role of the lags. The resulting objective function is less sensitive to the choice of the maximum lag allowed and has a wider radius of convergence compared to standard FWI and envelope inversions. The implementation has the same computational complexity as conventional FWI as the only difference in the calculation is related to the modified adjoint source. We implement this algorithm on the AMD GPU based on OPENCL and obtained about a 14 fold speed up compared to a CPU implementation based on OPENMP. At last, several numerical examples are shown to demonstrate the proper convergence of the proposed method. Application to the Marmousi model shows that this method converges starting with a linearly increasing velocity model, even with data free of frequencies below 3 Hz.

INTRODUCTION

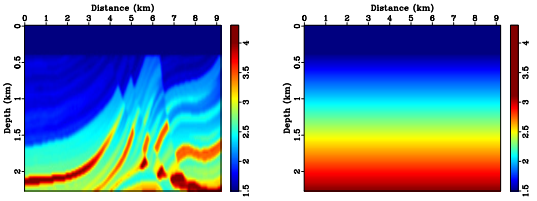
Recently, full waveform inversion (FWI), as a potential velocity model building tool, has gained a lot of momentum. FWI is iteratively capable of admitting high-resolution velocity models, provided that we start with a kinematically accurate (with respect to the minimum frequency available) initial guess of the model. However, when the initial velocity can not accurately explain the kinematics of the wavefield within a half cycle of the observed data, FWI usually converges to a local minimum instead of a global one, induced by the cycle skip between the real and predicted data. This cycle-skipped predicted data results in an inaccurate velocity model, and usually this velocity model includes artifacts needed to produce reflections that fit the observed data.

To mitigate the cycle skipping problem, many solutions have been proposed in recent years. One family of solutions suggests extending the model space (Symes, 2008; Biondi and Almomin, 2014; Huang and Symes, 2015), with additional degrees of freedom that allows us to fit the data beyond the physical model limitations. A penalty on the unphysical nature of the model extension provides a path to correct the kinematics of the wavefield to make it suitable for FWI to converge. Another family of methods are based on the global measurement of the distance between the modeled data and observed data (Warner and Guasch, 2014; Métivier et al., 2016; Yang et al., 2018). Another family of solutions to the cycle skipping problem involves minimizing the lag of the maximum of the correlation between the observed and modeled data (Luo and Schuster, 1991; Chi et al., 2015). More generally, some proposed various methods to measure the phase difference between the modeled data and observed data (Ma and Hale, 2013; Yang et al., 2015; Jiao et al., 2015). On the other hand, Bi and Lin (2014) developed an adaptive data selective method, which constrains the inversion to data that are not cycle-skipped by a measure of the traveltime lag through crosscorrelation. Another group of methods are based on measuring the quality or the difference of an extended image (Shen et al., 2003; Sava and Biondi, 2004; Biondi and Symes, 2004; Zhang and Schuster, 2013; Alkhalifah and Wu, 2017). For data without low enough frequency information, some have proposed to generate artificial low frequencies by approximating the data in a transformed domain (Shin and Cha, 2008; Hu, 2014; Choi and Alkhalifah, 2015; Li and Demanet, 2016). For reflection dominated data, Xu et al. (2012) and Zhou et al. (2012) developed a method based mainly on the work of Plessix et al. (1995) to invert for smooth velocity models using modeled reflected energy from an image. They referred to the method as reflected waveform inversion (RWI). Wang et al. (2013) implemented the same approach in the frequency domain to utilize a sequential frequency implementation, which they thought was necessary, to avoid the nonlinearity caused by an incorrect image. Zhou et al. (2015) proposed similar ideas, which invert for both velocity and impedance. In previous work, Alkhalifah and Wu (2016b); Wu and Alkhalifah (2015) proposed to invert for the background and perturbation simultaneously, which can utilize diving waves, first order reflection and even multi-scattering energy (Alkhalifah and Wu, 2016a) together. Guo and Alkhalifah (2017) extended the idea to the elastic case. One modification in the objective function recently came as an extension to the fitting space in the data domain (Van Leeuwen and Mulder, 2010). Wu and Alkhalifah (2018) proposed an extension correlation of the observed and modeled data.

In this paper, we propose to maximize the correlation with a time lag between the envelope of the modeled data and the envelope of the observed data. In this case, there is no need to introduce the selective function because the envelope of the signal is always positive. The resulting algorithm has similar gradient calculation complexity and cost to that of standard FWI, which makes the proposed method easy to implement in

10.1190/segam2018-2991067.1

a high performance computing platform, like the one we use here. Thus, we implement this new inversion on a unique GPU computing platform referred to as SANAM, which is a GPU cluster hosted by King Abdulaziz City for Science and Technology (KACST), Saudi Arabia. There are two AMD FirePro S10000 graphic cards in each node. As the GPUs represent the main floating computing power, it is important to fully utilize the two GPUs. Here we will also discuss our attempt in a high-performance implementation of our FWI method on the AMD GPU based on OPENCL programming.



(a) (b)

Figure 1: (a) The exact velocity for Marmousi model. (b) The linearly increasing initial velocity.

THE OBJECTIVE FUNCTION BASED ON AN EXTENDED CORRELATION OF THE ENVELOPE

The standard waveform inversion can be formulated in the following optimization problem (Lailly, 1983; Tarantola, 1984; Virieux and Operto, 2009):

$$\min_v J_0 = \sum_{\vec{x}} \int_t \frac{1}{2} \left[p(\vec{x},t) - g(\vec{x},t)\right]^2 dt, s.t. \frac{1}{v^2} p_{tt} - \Delta p = f, \quad (1)$$

in which, v is the P-wave velocity, f is the source wavelet, g is the observed data and p is the modeled wavefield corresponding to velocity v. This classic objective function (1) is highly sensitive to the amplitude of the wavefield, which is hard to accurately simulate considering our typical acoustic assumption of the medium and our ignorance of attenuation, among other shortcomings in the modeling process. An amplitude independent objective function (Choi and Alkhalifah, 2012) is given by maximizing the cross correlation between the modeled and observed wavefields, which can be expressed as

$$\min_v J_1 = -\sum_{\vec{x}} \frac{\int_t p(\vec{x},t) g(\vec{x},t) dt}{\sqrt{\int_t p(\vec{x},t)^2 dt \int_t g(\vec{x},t)^2 dt}}. \quad (2)$$

The objective functions J_0 and J_1 are highly nonlinear with respect to the velocity model such that conventional local gradient based methods often converge to local minima if the initial velocity is far from the exact one; there are not low enough frequencies in the data or the maximum offset is not large enough. This is often attributed to the cycle skipping between the predicted and observed data. Instead of correlating the modeled data with observed data directly, we can correlate the envelope of the modelled data and the shift of the envelope of the observed data. In this case, the correlation will always be positive due to the positive nature of the envelope function. Thus, we can avoid the introduction of the selective function. More precisely, we suggest to extend the objective function J_1 by correlating the envelope of modeled data with the shifted envelope of the observed data as follows:

$$\min_v J_2 = -\sum_{\vec{x}} \sum_{\tau} W(\tau) C(p,g,\vec{x},\tau), \quad (3)$$

in which $W(\tau)$ is a weighting function and the shifted cross-correlation between the envelope of modeled data and the envelope of the observed data is defined as:

$$C(p,g,\vec{x},\tau) = \frac{\int_t E(p)(\vec{x},t) E(g)(\vec{x},t+\tau) dt}{\sqrt{\int_t E(p)(\vec{x},t)^2 dt \int_t E(g)(\vec{x},t+\tau)^2 dt}}, \quad (4)$$

where $E(p)$ denotes the square envelope of the function of p, which is defined as:

$$E(p)(\vec{x},t) = p(\vec{x},t)^2 + Hp(\vec{x},t)^2. \quad (5)$$

Here, Hp is the Hilbert transform of p. Taking into consideration the smoothness and the compact support requirement, we suggest the following weighting function:

$$W(\tau) = \begin{cases} 2(\frac{|\tau|}{\max|\tau|})^3 - 3(\frac{|\tau|}{\max|\tau|})^2 + 1, & \tau \le \max|\tau|, \\ 0, & \tau > \max|\tau|, \end{cases} \quad (6)$$

which enhances the wavefield comparison. In this case, the objective function maximizes the correlation, not only at zero lag, but also at non-zero lag. We can control the maximum shift by making the $\max|\tau|$ depend on the quality of the initial velocity. In the case where the modeled data are far from the observed data, we choose large $\max|\tau|$. Due to the positive feature of the envelope function, the shifted correlation between the envelope of the modeled data and the envelope of the observed data will always be positive to guarantee the positive contribution to the objective function. Let us denote

$$\begin{aligned} R_1(\vec{x},t,\tau) &= \frac{E(g)(\vec{x},t+\tau)}{\sqrt{\int_t E(p)(\vec{x},t)^2 dt \int_t E(g)(\vec{x},t+\tau)^2 dt}} \\ &- \frac{\int_t E(p)(\vec{x},t) E(g)(\vec{x},t+\tau) dt E(p)(\vec{x},t)}{\sqrt{\int_t E(p)(\vec{x},t)^2 dt}^3 \sqrt{\int_t E(g)(\vec{x},t+\tau)^2 dt}} \end{aligned} \quad (7)$$

Then, the adjoint source $R(\vec{x},t)$ is given by

$$\begin{aligned} R(\vec{x},t) &= -2 \sum_{\tau} W(\tau) p(\vec{x},t) R_1(\vec{x},t,\tau) \\ &+ 2H\left(\sum_{\tau} W(\tau) H(p)(\vec{x},t) R_1(\vec{x},t,\tau) \right). \end{aligned} \quad (8)$$

Thus, only three Hilbert transforms are required even though the objective function is the summation of time shifted correlation on the envelope. As a result, the gradient is given by the adjoint state method (Plessix, 2006) as follows:

$$\nabla_v J_2 = \frac{2}{v^3} \int_t p_{tt}(\vec{x},t) \lambda(\vec{x},t) dt, \quad (9)$$

where λ is the back propagated wavefield with the adjoint source $R(\vec{x},t)$. As we can see from the above gradient, the proposed method has the same cost as standard FWI for each

10.1190/segam2018-2991067.1

iteration. We only need to execute one forward modeling and one backward modeling operations per iteration. The only difference is in the computation of the adjoint source. Due to the simple form of this objective function, it can be easily combined with other methods such as reflected waveform inversion(Wu and Alkhalifah, 2015; Alkhalifah and Wu, 2016b,a) to help mitigate the cycle skipping issue for these methods. vspace-0.45cm

HIGH PERFORMANCE IMPLEMENTATION

We aim to implement the proposed approach on a high performance computing platform referred to as SANAM, which is a GPU cluster hosted by King Abdulaziz City for Science and Technology (KACST), Saudi Arabia. SANAM consists of 210 nodes, each of which has two Intel(R) Xeon(R) CPU E5-2650@ 2.00GHz CPUs with 128GB memory. Each nodes use two AMD FirePro S10000 graphic cards with four graphic processors for acceleration. There are only 6 gigabytes memory for each AMD FirePro S10000 graphic card, which is not practical to store the whole wavefield. One option is to move the wavefield and store it on the CPU. However, that will result in a heavy burden imposed on the communication between CPU and GPU. To avoid the frequent communication issues between CPU and GPU, we need to solve the storage problem. Random boundary condition (Shen and Clapp, 2015) provides an efficient solution of the storage problem for the gradient calculation. However, it can not provide accurate enough wavefields for the objective function and the adjoint source calculation. So we propose to do a regular modeling and calculate the objective function and adjoint source. The random boundary condition is utilized to mitigate the storage problem for the gradient calculation. The detail algorithm is summarized in Algorithm 1.

Algorithm 1 GPU implementation for FWI

Require: The source f, and seismic velocity v

- Forward modelling with an absorbing boundary condition and save the modeled data at the receiver locations.
- Calculate the objective function and adjoint source.
- Forward modelling with a random boundary condition and store a snapshot of the wavefield, p, at the last two time steps.
- Backward modelling with the last time step of p with a random boundary condition to recover p. Meanwhile, backward modelling with absorbing boundary condition using the adjoint source to obtain λ. Apply the imaging condition at each time step

For the forward modeling, we suggest to use a hierarchical modelling scheme to minimize the usage of a velocity interpolation step. That means we use the same grid as much as possible. For a relative low-frequency modelling, we will utilize the relative low order finite difference. If the frequency is a little bit higher, we will utilize a higher order finite difference method. For even higher frequency, we use the pseudo-spectral approach. For the implementation of the free surface boundary condition for pseudo-spectral, we use the method proposed in Wu and Alkhalifah (2016). For the absorbing boundary condition, we use the hybrid absorbing boundary condition (Liu and Sen, 2010), which works well for both spectral and finite difference methods. For the three dimensional finite difference implementation, we use shared memory for the first two axises to reduce the access to the global memory because the access to the global memory for GPU is really slow. For the same reason, we use register memory to store the neighboring value in the third dimension and roll the register memory on the third dimension. For the implementation of Fast Fourier transform (FFT), we utilize the highly optimized FFT library called clFFT (`https://github.com/clMathLibraries/clFFT`).

NUMERICAL EXAMPLES

We will consider the benchmark Marmousi model. The exact velocity is shown in Figure 1a. To remove the source singularity, we add a water layer and assume the initial velocity is accurate in the water. The time sampling is 0.0025 s while the space sampling is 0.024 km. The source is located on the surface with 0.216 km space sampling. All the points on the earth surface are serving as receivers. The starting model is a linearly increasing model shown in Figure 1b. We first model the data with a source wavelet that has no frequencies lower than 3 Hz. Starting from the initial velocity, we apply the proposed method with $\max|\tau| = 0.25$ s. The inverted model is shown in Figure 2a. After that, we start from the inverted velocity in Figure 2a and apply the standard FWI. The inverted velocity is shown in Figure 2b. Compared to the exact velocity in 1a, the inverted velocity shown in Figure 2b is reasonably good. For comparison, the velocity shown in Figure 2c is the inverted velocity using standard Full waveform inversion starting from the linearly increasing initial velocity in Figure 1b. As shown in Figure 2c, the standard Full waveform inversion does not produce good inverted results due to the cycle skipping issue. To further compare, we apply the standard envelope inversion followed by standard Full waveform inversion starting from the initial velocity shown in Figure 1b. The inverted velocity is shown in Figure 2d. Compared to Figure 2c, the inverted velocity has improved a lot. However, the inverted velocity (Figure 2b) produced by the new method still shows better accuracy. To show the approximate accuracy of the modeled data to the observed data, we plot the shot gather for a source located in the middle of the model. Figure 4a shows the observed data. Figure 4b shows the modeled data with the initial velocity. Figure 4c shows the residual of the modeled data using the initial velocity. The residual of the new method followed by standard FWI is shown in Figure 4d. The residual of the standard envelope inversion followed by standard FWI is shown in Figure 4e. The residual of the standard standard FWI is shown in Figure 4f. All these figures are plotted at the same scale. From these figures, we can see that the new method followed by standard FWI produces the smallest residual. The standard FWI does not improve the residual a lot due to the severe cycle-skipping issue. The standard envelope inversion followed by

10.1190/segam2018-2991067.1

standard FWI shows much smaller residual than standard FWI. However, it is still a little bit bigger than the newly proposed method followed by standard FWI.

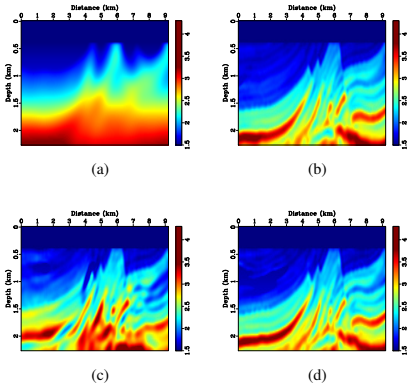


Figure 2: (a) The inverted velocity using the proposed method with maximum shift of 0.25s. (b) The inverted velocity using standard FWI starting from Figure 2a. (c) The inverted velocity using standard FWI starting from Figure 1b. (d) The inverted velocity using standard envelope inversion followed by standard FWI.

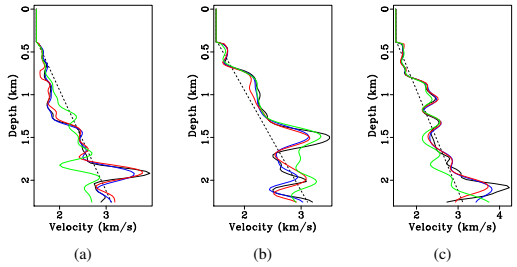


Figure 3: The velocity profile. (a) 3 km. (b) 5 km.(c) 7 km.(black-dashed line: initial velocity; black-solid line:exact velocity; green-solid line: standard FWI. red-solid line: standard envelope inversion followed by standard FWI. blue-solid line: the new method followed by standard FWI.)

PERFORMANCE DISCUSSION

As mentioned before, we are testing the implementation on SANAM. For simplicity of the programming, we assume that one shot calculation can be fitted into one single GPU, which is reasonable because the frequency of data used in Full-waveform inversion is often reasonably low and, as discussed previously, we utilize the random boundary condition to store the wavefield. In this case, the scalability with respect to the number of shots is trivial. So we will compare the performance for one single GPU and one single CPU of 8 cores. Since the forward modeling from a random boundary condition dominates the computation for a single iteration, we will focus the comparison to the forward modeling with random boundary condition. The space mesh samples are $256 * 256 * 256$ and we have extrapolated for 1000 time steps. The relative computing time for

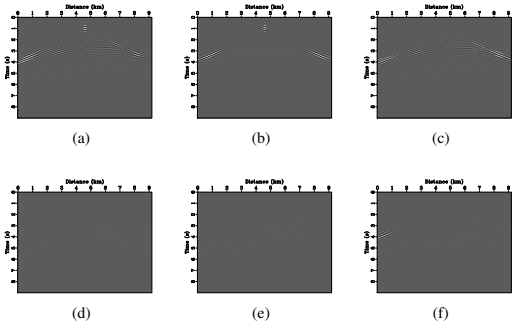


Figure 4: Comparison of shot gather. (a) The observed data. (b) The modelled data with the initial velocity. (c) The error of the modelled data with the initial velocity. (d) The error of the modelled data with the new method followed by standard FWI. (e) The error of the modelled data with the inverted velocity using standard envelope inversion followed by standard FWI. (f) The error of the modelled data with the inverted velocity using standard FWI.

Table 1: Computing time comparison

Setting	CPU (1 thread)	CPU (4 threads)	CPU (8 threads)	GPU
Time(s)	655	209	116	8

CPU with different OPENMP threads and the GPU is shown in Table 1. Basically, we can achieve a speed up of 80 compared to one single CPU core. The GPU code is 14 times faster than one whole CPU.

CONCLUSIONS

We proposed a new approach to mitigate the cycle-skipping problem. Instead of maximizing only the zero lag correlation in the objective function, we maximize the correlation over time lag of the envelope of the data. We apply a weighting function over the lag axis that allows for a correlation of the observed and predicted data up to a user defined maximum shift in time, giving larger weights near zero lag. Since we are maximizing the correlation of the shifted envelope instead of the original data, the correlation will always contribute to the objective function positively. With the help of the weighting function applied to the lag, we obtain better convergence behavior. Numerical examples confirm these features, in which the inversion is less dependent on the initial velocity model and can tackle the cycle-skipping problem. This implementation is optimized on a unique GPU cluster in which we obtain reasonable scalability of our parallel implementation.

ACKNOWLEDGEMENTS

We thank KAUST for its support and we thank the SWAG group for collaborative environment. The numerical experiments were performed on the super computer SANAM, which is hosted by King Abdulaziz City for Science and Technology, Saudi Arabia. We also thank the support team of SANAM for their useful help.

10.1190/segam2018-2991067.1

REFERENCES

Alkhalifah, T., and Z. Wu, 2016a, Multiscattering inversion for low-model wavenumbers: Geophysics, **81**, no. 6, R417–R428, https://doi.org/10.1190/geo2015-0650.1.

Alkhalifah, T., and Z. Wu, 2016b, The natural combination of full and image-based waveform inversion: Geophysical Prospecting, **64**, 19–30.

Alkhalifah, T., and Z. Wu, 2017, Migration velocity analysis using pre-stack wave fields: Geophysical Prospecting, **65**, 639–649, https://doi.org/10.1111/1365-2478.12439.

Bi, H., and T. Lin, 2014, Impact of adaptive data selection on full waveform inversion: 84th SEG Annual International Meeting, SEG, Expanded Abstracts, 1094–1098.

Biondi, B., and A. Almomin, 2014, Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion: Geophysics, **79**, no. 3, WA129–WA140, https://doi.org/10.1190/geo2013-0340.1.

Biondi, B., and W. Symes, 2004, Angle domain common image gathers for migration velocity analysis by wavefield continuation imaging: Geophysics, **69**, 1283–1298, https://doi.org/10.1190/1.1801945.

Chi, B., L. Dong, and Y. Liu, 2015, Correlation-based reflection full-waveform inversion: Geophysics, **80**, no. 4, R189–R202, 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, https://doi.org/10.1111/j.1365-2478.2012.01079.x.

Choi, Y., and T. Alkhalifah, 2015, Unwrapped phase inversion with an exponential damping: Geophysics, **80**, no. 5, R251–R264, https://doi.org/10.1190/geo2014-0498.1.

Guo, Q., and T. Alkhalifah, 2017, Elastic reflection-based waveform inversion with a nonlinear approach: Geophysics, **82**, no. 6, R309–R321, https://doi.org/10.1190/geo2016-0407.1.

Hu, W., 2014, Fwi without low frequency data—Beat tone inversion: 84th SEG Annual International Meeting, SEG, Expanded Abstracts, 1116–1120.

Huang, G., and W. W. Symes, 2015, Full waveform inversion via matched source extension: 85th SEG Annual International Meeting, SEG, Expanded Abstracts, 1320–1325.

Jiao, K., D. Sun, X. Cheng, and D. Vigh, 2015, Adjustive full waveform inversion: 85th SEG Annual International Meeting, SEG, Expanded Abstracts, 1091–1095.

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

Li, Y. E., and L. Demanet, 2016, Full-waveform inversion with extrapolated low-frequency data: Geophysics, **81**, no. 6, R339–R348, https://doi.org/10.1190/geo2016-0038.1.

Liu, Y., and M. K. Sen, 2010, A hybrid scheme for absorbing edge reflections in numerical modeling of wave propagation: Geophysics, **75**, no. 2, A1–A6, https://doi.org/10.1190/1.3295447.

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

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

Metivier, L., R. Brossier, Q. Mrigot, E. Oudet, and J. Virieux, 2016, Measuring the misfit between seismograms using an optimal transport distance: Application to full waveform inversion: Geophysical Journal International, **205**, 345–377.

Plessix, R., Y. De Roeck, and G. Chavent, 1995, Automatic and simultaneous migration velocity analysis and waveform inversion of real data using a MBTT/WKBJ formulation: 75th SEG Annual International Meeting, SEG, Expanded Abstracts, 1099–1102, https://doi.org/10.1190/1.1887624.

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, https://doi.org/10.1111/j.1365-246X.2006.02978.x.

Sava, P., and B. Biondi, 2004, Wave-equation migration velocity analysis—I. Theory: Geophysical Prospecting, **52**, 593–606, https://doi.org/10.1111/j.1365-2478.2004.00447.x.

Shen, P., W. W. Symes, and C. C. Stolk, 2003, Differential semblance velocity analysis by waveequation migration: 73th SEG Annual International Meeting, SEG, Expanded Abstracts, 2132–2135, https://doi.org/10.1190/1.1817759.

Shen, X., and R. G. Clapp, 2015, Random boundary condition for memory-efficient waveform inversion gradient computation: Geophysics, **80**, no. 6, R351–R359, https://doi.org/10.1190/geo2014-0542.1.

Shin, C., and Y. H. Cha, 2008, Waveform inversion in the laplace domain: Geophysical Journal International, **173**, 922–931, https://doi.org/10.1111/j.1365-246X.2008.03768.x.

Symes, W. W., 2008, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, 765–790, 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, https://doi.org/10.1190/1.1441754.

Van Leeuwen, T., and W. A. Mulder, 2010, A correlation-based misfit criterion for wave-equation traveltime tomography: Geophysical Journal International, **182**, 1383–1394, https://doi.org/10.1111/j.1365-246X.2010.04681.x.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration Geophysics: Geophysics, **74**, no. 6, WCC1–WCC26, https://doi.org/10.1190/1.3238367.

Wang, S., F. Chen, H. Zhang, and Y. Shen, 2013, Reflection-based full waveform inversion (RFWI) in the frequency domain: 83rd SEG Annual International Meeting, SEG, Expanded Abstracts, 877–881, https://doi.org/10.1190/segam2013-0671.1.

Warner, M., and L. Guasch, 2014, Adaptive waveform inversion: Theory: 84th SEG Annual International Meeting, SEG, Expanded Abstracts, 1089–1093, https://doi.org/10.1190/segam2014-0371.1.

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, https://doi.org/10.1190/geo2014-0365.1.

Wu, Z., and T. Alkhalifah, 2016, The optimized gradient method for full waveform inversion and its spectral implementation: Geophysical Journal International, **205**, 1823–1831, https://doi.org/10.1093/gji/ggw112.

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, https://doi.org/10.1190/geo2016-0649.1.

Xu, S., D. Wang, F. Chen, G. Lambare, and Y. Zhang, 2012, Inversion on reflected seismic wave: 84th SEG Annual International Meeting, SEG, Expanded Abstracts, 1–7, https://doi.org/10.1190/segam2012-1473.1.

Yang, H., Y. Wu, and X. Chen, 2015, Full waveform inversion overcoming cycle skipping by phase correlation shifting: 77th Annual International Conference and Exhibition, EAGE, Extended Abstracts, TuP214, https://doi.org/10.3997/2214-4609.201412768.