Accepted Manuscript

A new full waveform inversion method based on shifted correlation of the envelope and its implementation based on OPENCL

Zedong Wu, Tariq Alkhalifah, Zhendong Zhang, Faisal Alonaizi, Majed Almalki

PII: S0098-3004(18)30375-3

DOI: <https://doi.org/10.1016/j.cageo.2019.04.007>

Reference: CAGEO 4262

To appear in: *Computers and Geosciences*

Received Date: 13 April 2018

Revised Date: 28 March 2019

Accepted Date: 15 April 2019

Please cite this article as: Wu, Z., Alkhalifah, T., Zhang, Z., Alonaizi, F., Almalki, M., A new full waveform inversion method based on shifted correlation of the envelope and its implementation based on OPENCL, *Computers and Geosciences* (2019), doi: <https://doi.org/10.1016/j.cageo.2019.04.007>.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

A new full waveform inversion method based on shifted correlation of the envelope and its implementation based on OPENCL

Zedong Wu^{a,*}, Tariq Alkhalifah^a, Zhendong Zhang^a, Faisal Alonaizi^b, Majed Almalki^b

^a *Seismic wave analysis group, King Abdullah university of science and technology, Thuwal, Saudi arabia.*

^b *Oil and Gas research center, King Abdulaziz city for science and technology, Riyadh, Saudi Arabia.*

Abstract

Standard full waveform inversion (FWI) attempts to minimize the difference between observed and modeled data. When the initial velocity is kinematically accurate, FWI often converges 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 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 compare the envelope of the modeled and observed data instead of comparing the modeled and observed data directly. However, if the initial velocity is not sufficient, the correlation of the envelope of the modeled and observed data might still be zero. To mitigate this issue, we propose to maximize both the zero-lag correlation of the envelope and the non-zero-lag correlations of the envelope. A weighting function with maximum value at zero lag and decays away from zero lag is introduced to balance the role of the lags. The resulting

[☆] Authorship statement: Tariq Alkhalifah, Faisal Alonaizi, and Majed Almalki proposed the idea to obtain an efficient implementation of Full waveform inversion based on OPENCL and formulated the project; All the authors were involved in the discussion of the proposed method; Zedong Wu implemented the method based on OPENCL.

*Corresponding author

Email address: zedong.wu@kaust.edu.sa (Zedong Wu)

objective function is less sensitive to the choice of the maximum lag allowed and has a wider region of convergence with respect 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 a modified adjoint source. The implementation of this algorithm was performed on an AMD GPU platform based on OPENCL and provided a 14 times speedup over a CPU implementation based on OPENMP. 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.

Keywords: Acoustic wave equation, Waveform inversion, Envelope, Shifted correlation, OPENCL and MPI.

2010 MSC: 00-01, 99-00

1. Introduction

Full waveform inversion (FWI) has recently gained a lot of attention as a tool that can provide a high-resolution velocity estimation of the subsurface. However, it strongly depends on the initial velocity model that has to accurately explain the kinematics of the wavefield within a half cycle of the observed data. Otherwise, it will converge to a local minimum instead of a global one, which reflects the cycle-skip between the real and predicted data.

There are many solutions that have been proposed recently to address the issue of cycle-skipping. One family of solutions suggests extending the model space [1, 2, 3] with additional degrees of freedom that will allow 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. However, an extension of the model of any type usually forces the inversion to work in an extended domain, which is relatively expensive. In addition, the penalty is applied in many forms while

allowing the basic measure of difference between the observed and modeled data to impose a classic FWI in the inversion. Another family of methods is based on the global measurement of the distance between the modeled and observed data [4, 5, 6], and yet 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 [7, 8]. More generally, some researchers propose different methods to measure the phase difference between the modeled and observed data [9, 10, 11, 12]. Others, such as [13] developed an adaptive data selective method, that constrains the inversion to data that is not cycle-skipped by a measure of the traveltime lag through crosscorrelation. Another group of methods is based on measuring the quality or the difference of an extended image [14, 15, 16, 17, 18]. For data without sufficient low frequency information, some researchers have proposed generating artificial low frequencies by approximating the data in a transformed domain [19, 20, 21, 22]. Envelope inversion [23, 24, 25, 26, 27, 28, 29, 30, 31, 32] can reduce the number of local minima by comparing the envelope of data, which induces low frequency information. For reflection dominated data, [33] and [34] developed a method based mainly on the work of [35] to invert for smooth velocity models using modeled reflected energy from an image and referred to it as reflected waveform inversion (RWI). The idea is based on migration followed by de-migration to predict the reflections. Since the de-migration is obtained from the image, for an imperfect velocity, the modeled data from the image will have residuals, ideally at far offsets. As a result, RWI mainly inverts the propagator (smooth) part of the model, like migration velocity analysis (MVA) but without the need for extended images or angle gathers. [36] 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. [37] proposed similar ideas that invert for both velocity and impedance. In previous work, we proposed inverting for the background and perturbation simultaneously [38, 39], which can utilize diving waves, first-order reflection and even multi-scattering energy together [40] and later on, [41] extended the idea to the elastic case.

Reflected waveform inversion can update along the wave-path of both diving and reflected waves which can reduce the cycle-skipping problem. However, reflected waveform inversion might still suffer from cycle-skipping problems due to limitations in the objective function.

One modification in the objective function recently came as an extension to the fitting space in the data domain [42]. [43] proposed an extension correlation of the observed and modeled data. However, this requires a selective function to remove the negative contribution in the objective function. In this work, we propose maximizing the correlation with a time lag between the envelope of the modeled data and the envelope of the observed data. By combining this extension with a proper weighting function, there will be no need to introduce the selective function because the envelope of the signal is always positive. Numerical examples illustrate that the proposed method can have a much larger region of convergence (basin of attraction) than standard FWI and envelope inversions. Implementing the extension in the data domain makes the resulting algorithm has a similar gradient calculation and cost, compared to standard FWI. The only difference here is in the adjoint source, which makes the proposed method easy to implement in 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. Here we will also discuss our attempt at a high-performance implementation of our FWI methods on the AMD GPU based on OPENCL programming.

2. The objective function based on an extended correlation of the envelope

Standard waveform inversion can be formulated in the following optimization problem [44, 45, 46]:

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

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 . Here we consider only a single source case for simplicity. We only need to sum the contribution of different sources in the case of multiple sources. The summation of $\vec{x}$ is over all receiver locations. 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. A globally amplitude independent objective function [47] 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 no sufficiently low frequencies in the data or the maximum offset is not large enough. This is often attributed to the cycle skipping between predicted and observed data. An example is shown in Figure 1(a), in which the red and blue curves represent the observed and modeled data respectively. The correlation between the red and blue curves equals zero, as there is no intersection between the energy corresponding to the two events. In this case, gradient methods cannot improve the velocity as the objective function J_1 (Figure 1(b)) is flat

Figure 1: An illustration of the cycle skipping problem in the time domain. (a) Modeled and observed data and their envelope. (b) The global correlation objective function with different constant velocity. (c) The shift of the observed data to the modeled data. (d) The shift of the envelope of the observed data to the envelope of the modeled data.

95 around that velocity. In order to solve this problem and allow for an interaction
between modeled and observed data, it might be possible to shift the observed
data to the modeled data and correlate them, as shown in Figure 1(c). However,
shifting the two data and correlating them directly could produce a correlation
that contributes negatively to the objective function. To remove the negative
100 contribution, [43] proposed an extra selective function to the shifted correlation.
Instead of correlating of the modeled wavefield with shifted observed data di-
rectly, we can correlate the envelope of the modelled data and the shift of the
envelope of the observed data as shown in Figure 1(d). In this case, the correla-
tion will always be positive due to the positive nature of the envelope function.
105 Thus, we can avoid the introduction of the selective function. More precisely,
we suggest extending the objective function J_1 by correlating the envelope of
the 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 the modeled data and the envelope of the observed data is
110 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 [23, 24, 25, 27, 32]:

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

Here, Hp is the Hilbert transform of p . Taking into consideration of the smooth-
ness and the compact support requirement, we suggest the following weighting
115 function (see Figure 2):

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

Figure 2: An illustration of $w(\tau)$.

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 in which the modeled data is far from the
 120 observed data as shown in Figure 1(a), 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.

3. The gradient calculation

125 In order to solve the optimization problem (3) with a local-based optimization method, we derive the gradient of the objective function. We perturb the velocity v by δv , so that

$$\delta J_2 = - \sum_{\vec{x}} \sum_{\tau} W(\tau) \delta C(p, g, \vec{x}, \tau), \quad (7)$$

where

$$\begin{aligned} \delta C(p, g, \vec{x}, \tau, \vec{h}) = & \frac{\int_t \delta E(p)(\vec{x}, t) E(g)(\vec{x}, t + \tau) dt}{\sqrt{\int_t E(p)(\vec{x}, t)^2 dt \int_t E(g)(\vec{x}, t + \tau)^2 dt}} \\ & - \frac{\int_t E(p)(\vec{x}, t) E(g)(\vec{x}, t + \tau) dt \int_t E(p)(\vec{x}, t) \delta E(p)(\vec{x}, t) dt}{\sqrt{\int_t E(p)(\vec{x}, t)^2 dt}^3 \sqrt{\int_t E(g)(\vec{x}, t + \tau)^2 dt}}. \end{aligned} \quad (8)$$

Let us denote

$$R_1(\vec{x}, t, \tau) = \frac{E(g)(\vec{x}, t + \tau)}{\sqrt{\int_t E(p)(\vec{x}, t)^2 dt} \sqrt{\int_t E(g)(\vec{x}, t + \tau)^2 dt}} - \frac{\int_t E(p)(\vec{x}, t) E(g)(\vec{x}, t + \tau) dt E(p)(\vec{x}, t)}{\sqrt{\int_t E(p)(\vec{x}, t)^2 dt}^3 \sqrt{\int_t E(g)(\vec{x}, t + \tau)^2 dt}}. \quad (9)$$

Then, we will have

$$\begin{aligned} \delta C(p, g, \vec{x}, \tau, \vec{h}) &= \int_t \delta E(p)(\vec{x}, t) R_1(\vec{x}, t, \tau) dt, \\ &= \int_t (2p(\vec{x}, t) \delta p(\vec{x}, t) + 2H(p)(\vec{x}, t) \delta H(p)(\vec{x}, t)) R_1(\vec{x}, t, \tau) dt, \\ &= \int_t [2p(\vec{x}, t) - H(2H(p)(\vec{x}, t) R_1(\vec{x}, t, \tau))] \delta p(\vec{x}, t) dt. \end{aligned} \quad (10)$$

We extract from the above the terms that form the adjoint source $R(\vec{x}, t)$ given by

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

In the above formulation, we utilized the linearity of the Hilbert transform.

Thus, only three Hilbert transforms are required even though the objective

function is the summation of the time-shifted correlation on the envelope. The perturbed wavefield satisfies the wave equation:

$$\frac{1}{v^2} \delta p_{tt} - \Delta \delta p = \frac{2}{v^3} p_{tt} \delta v. \quad (12)$$

As a result, the gradient is given by the adjoint state method [48] as follows:

$$\nabla_v J_2 = \frac{2}{v^3} \int_t p_{tt}(\vec{x}, t) \lambda(\vec{x}, t) dt, \quad (13)$$

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 iteration. We only need to execute one forward modeling and one backward modeling operation per iteration. The only difference is in the computation of the adjoint source. Due to the simple form of this objective function, it can easily be combined with other methods such as reflected

145 waveform inversion[39, 38, 40] to help mitigate the cycle-skipping issue for these methods.

4. High-performance implementation

As mentioned earlier, the proposed technique has been implemented on a high-performance computing platform referred to as SANAM, which is a multi-GPU cluster hosted by King Abdulaziz City for Science and Technology (KACST) in Riyadh, Saudi Arabia. SANAM has been ranked second in the worldwide list of the most energy-efficient computers when it was launched. it consists of 210 nodes, each of which has two Intel(R) Xeon(R) CPU E5-2650@2.00GHz CPUs with 128GB of memory. Each node uses two AMD FirePro S10000 graphic cards with four graphic processors for acceleration. There are only 6 GB of memory for each AMD FirePro S10000 graphic card, which is not practical storing the entire wavefield. One option is to move the wavefield and store it on the CPU. However, this will result in a heavy burden on the communication between the CPU and GPU. To avoid frequent communication issues between the CPU and GPU, we need to solve the storage problem. The random boundary condition [49] provides an efficient solution to the storage problem for gradient calculation. To improve the efficiency and accuracy of the objective function calculation, we propose using regular modeling in calculating the objective function and adjoint source. The random boundary condition is utilized to mitigate the storage problem for the gradient calculation. The detailed algorithm is summarized below as Algorithm 1.

In the forward modeling process, to minimize the usage of the velocity interpolation, we suggest using hierarchical modelling, where the same grid is used as much as possible. For relatively low-frequency modelling, we utilize the relatively low order finite difference. If the frequency is slightly higher, we utilize relatively high order finite difference method. For an even higher frequency, we use the pseudo-spectral method. To implement the free surface boundary condition for the pseudo-spectral method, we use the technique proposed in [50].

Algorithm 1 GPU implementation for FWI**Require:** The source f , and seismic velocity v

- Forward modeling with an absorbing boundary condition; the modelled data is saved at the receiver locations
- Calculate the objective function and adjoint source
- Forward modeling with the random boundary condition; the snapshot at the last two time steps p is stored
- Backward modeling with the last time step of p with the random boundary condition to recover p . Meanwhile, backward modelling with an absorbing boundary condition using the adjoint source to obtain λ . Apply the image condition at each time step

For the absorbing boundary condition, we use the hybrid absorbing boundary condition[51], which works well for both the spectral and the finite-difference
 175 methods. In the three dimensional finite difference implementation, we use share memory for the first two axes to reduce access to the global memory because the memory access to the global memory is really slow for GPU. For the same reason, we use register memory to store the neighbor value in the third dimension and roll the register memory on the third dimension. For the implementation of
 180 Fast Fourier Transform(FFT), we utilize a highly optimized FFT library called clFFT (<https://github.com/clMathLibraries/clFFT>).

5. Numerical Examples

To demonstrate that the objective function (3) can adequately address the cycle-skipping issue clearly, we start with a simple example. The exact velocity
 185 shown in Figure 3(a) is a constant background of 1.5 km/s with a Gaussian anomaly in the middle.

$$v_e = 1.5 + 1.5 \exp \left(-\frac{(x_1 - 3.75)^2 + (x_2 - 3.75)^2}{1.2^2} \right). \quad (14)$$

The initial velocity shown in Figure 3(b) is a constant velocity of 1.5 km/s. The space sampling is 0.03 km. The source is located in the middle of the left boundary and the receivers are located on the right boundary with a maximum offset of 1.5 km. The modeled data with the exact and initial velocities are shown in Figure 4(a) and Figure 4(b), respectively. We can see that the modeled and observed data are cycle skipped. The envelope of the observed and modeled data with the initial velocity is shown in Figures 4(c) and 4(d), respectively. We can see that even the envelope of the observed and modeled data with the initial velocity is cycle skipped. To show this clearly, we extract the traces in the middle and plot them in Figure 5. To show the effect of the cycle-skipping problem on the various objective functions, we consider the linear combination velocity $v(t) = v_e t + v_i(1 - t)$. The objective function with different t values for different forms of the objective function is shown in Figure 6. It can be realized from Figure 6 that the objective function, which computes the zero-lag correlation of the envelope function has a larger convergence region than the standard zero-lag correlation. The new objective function has an even larger convergence region than the objective function which computes the zero-lag correlation of the envelope function.

With this analysis and with a one-dimensional model in mind, we perform inversion starting from a constant background velocity. In this example, there are 112 shots regularly placed on all the boundaries of the region and all the points on all boundaries act as receivers. We first analyze the gradient of the different methods with the initial constant background. We mute out the gradient near the source locations. The gradient of the proposed method with a maximum shift of 1.8 s is shown in Figure 7(a). To compare with Figure 7(a), we display the gradient with a maximum shift of 0 s (standard envelope inversion[24, 25]) and the gradient of standard full waveform inversion are shown in Figures 7(b) and 7(c), respectively. Figure 7(a) shows a smooth gradient with the correct sign, while Figures 7(b) and 7(c) show gradients marred with cycle skipping. For the new method, starting from a constant background velocity, we only execute five iterations with a maximum shift of 1.8 s. This is sufficient because there

Figure 3: (a) The exact velocity. (b) The initial velocity.

is no cycle skipping with the new objective function. The inverted velocity is shown in Figure 8(a). Next, we plot the envelope of the modeled and observed data. The maximum distance between the modeled and observed data is about 0.3 s. Starting from Figure 8(a), another five iterations with a maximum shift of 0.3 s are utilized to obtain Figure 8(b). Then we update another five iterations starting from Figure 8(b) with the maximum shift of 0s, which is the standard envelope inversion, to obtain Figure 8(c). After that, we apply 10 iterations of standard full waveform inversion. The final inverted velocity is shown in Figure 8(d). For comparison, the inverted velocity with 30 iterations using the standard envelope objective function is shown in Figure 8(e). The inverted velocity with 30 iterations using standard Full waveform inversion is shown in Figure 8(f). By comparing Figures 8(d), 8(e), and 8(f), we can see that the newly proposed method results in a better converged result even with fewer number of iterations. To show the accuracy clearly, we take out the velocity profile at the middle and plot it in Figure 9.

Next, we will consider the more complicated Marmousi model. The exact velocity is shown in Figure 10(a). To avoid the issue that the gradient will have large amplitudes at the source location, we add a water layer and assume that 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

Figure 4: (a) The observed data. (b) The modeled data with the initial velocity. (c) The envelope of the observed data . (d) The envelope of the modeled data with the initial velocity.

Figure 5: The data comparison (red curve: the modeled data with the initial velocity; blue curve: the observed data; green curve: the envelope of the modeled data with the initial velocity; purple curve: the envelope of the observed data.).

Figure 6: The comparison of different objective function with different linearly combined velocity.

Figure 7: The gradient comparison with the constant background velocity. (a) The weighted envelope inversion with maximum shift ($\max |\tau|$) of 1.8 s. (b) The standard envelope inversion. (c) The standard Full waveform inversion.

Figure 8: (a) The inverted velocity of weighted envelope inversion with maximum shift ($\max |\tau|$) of 1.8 s starting from constant background velocity. (b) The inverted velocity with maximum shift of 0.3 s starting from (a). (c) The inverted velocity using standard envelope inversion starting from (b). (d) The inverted velocity of standard Full waveform inversion starting from (c). (e) The inverted velocity with standard envelope inversion starting from constant background velocity. (f) The inverted velocity with standard Full waveform inversion starting from constant background velocity.

Figure 9: The comparison of the velocity profile in the middle (black-dashed line: exact velocity; black-solid line: maximum time shift of 1.8s; blue-solid line: maximum time shift of 0.3s; red-solid line: maximum time shift of 0s; green-solid line: followed standard Full waveform inversion; blue-dashed line: standard envelope inversion; red-dashed line: standard Full waveform inversion).

km space sampling. All the points on the Earth's surface act as receivers. The starting model is a linearly increasing velocity model as shown in Figure 10(b).

240 We first model the data with a source wavelet (shown in Figure 11(a)) that has no frequencies lower than 3 Hz. The frequency spectrum of the source wavelet is shown in Figure 11(b). Starting from the initial velocity, we apply the proposed method with $\max|\tau| = 0.25$ s. The inverted model is shown in Figure 12(a). We then start from the inverted velocity in Figure 12(a) and apply the standard
245 FWI. The inverted velocity is shown in Figure 12(b). Compared to the exact velocity in 10(a), the inverted velocity shown in Figure 12(b) is reasonably good. For comparison, the velocity shown in Figure 12(c) is the inverted velocity using standard full waveform inversion starting from the linearly increasing initial velocity in Figure 10(b). As shown in Figure 12(c), the standard full waveform
250 inversion does not produce good inverted results due to the cycle-skipping issue. Furthermore, we apply the standard envelope inversion followed by standard full waveform inversion starting from the initial velocity shown in Figure 10(b). The inverted velocity is shown in Figure 12(d). Compared to Figure 12(c), the inverted velocity has significantly improved. However, the inverted velocity
255 (Figure 12(b)) produced by the weighted envelope objective function still shows

Figure 10: (a) The exact velocity for Marmousi model. (b) The linearly increasing initial velocity.

better accuracy. Figure 13 shows the comparison of the velocity profile at 3km, 5km, and 7km. To show the approximate accuracy of the modeled data to the observed data, we plot the shot gather when the source is located in the middle of the model. Figure 14(a) shows the observed data. Figure 14(b) shows the modeled data with the initial velocity. Figure 14(c) shows the residual (difference from the observed data) of the modeled data using the initial velocity. The residual of weighted envelope inversion followed by standard FWI is shown in Figure 14(d). The residual of the standard envelope inversion followed by standard FWI is shown in Figure 14(e). The residual of the standard FWI is shown in Figure 14(f). All these figures are plotted at the same color scale. From these figures, we can see that the weighted envelope inversion followed by standard FWI produces the smallest difference. The standard FWI does not significantly improve the residual due to the severe cycle-skipping issue. The standard envelope inversion followed by standard FWI shows a much smaller residual than standard FWI. However, it is still slightly larger than the weighted envelope inversion followed by standard FWI.

Figure 11: (a) The exact velocity for Marmousi model. (b) The linearly increasing initial velocity.

Figure 12: (a) The inverted velocity using the proposed method with maximum shift of 0.25s. (b) The inverted velocity using standard FWI starting from Figure 12(a). (c) The inverted velocity using standard FWI starting from Figure 10(b). (d) The inverted velocity using standard envelope inversion followed by standard FWI.

Figure 13: 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.)

6. Performance discussion

As previously mentioned, we are testing the implementation on SANAM. To simplify the program, we assume one shot can be fitted into one single GPU, which is reasonable because the frequency of the data used in full waveform inversion is often reasonably low and, as previously discussed, 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 compare the performance for one single GPU and one single CPU of eight cores. Since forward modeling from a random boundary condition dominates the computation for a single iteration, we will focus the comparison on forward modeling with a random boundary condition. The relative hardware and software information is summarized in Table 1. The space mesh samples are $256 * 256 * 256$ and we have extrapolated for 1000 time steps. The relative computing time for the CPU with different OPENMP threads and the GPU is shown in Table 2. It can be seen that we achieved a speed up of 80 compared to one single CPU core and the GPU code is 14 times faster than one whole CPU.

Figure 14: 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: Hardware and software information

Hardware	Facts
Software	
CPU	Intel(R) Xeon(R) CPU E5-2650@ 2.00GHz CPU; 8 core, 128GB memory
GPU	AMD FirePro S10000; 2×3GB GDDR5 Memory; ECC: enabled; 5.91 TFLOPS single precision float calculation
OPENCL	AMD-APP-SDK-v2.8-lnx64
Compiler	gcc 4.3

Table 2: Computing time comparison

Setting	CPU (1 thread)	CPU (2 threads)	CPU (4 threads)	CPU (8 threads)	GPU
Time(s)	655	367	209	116	8

7. 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 the 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, which produces 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.

8. Acknowledgements

We thank KAUST and KACST for their support. We thank the SWAG group for collaborative environment. The numerical experiments were performed on the super computer SANAM, which is hosted by KACST, Saudi Arabia. We also thank the support team of SANAM for their useful help. Our special thanks are extended to the associate editor Sjoerd de Ridder and two anonymous reviewers for their fruitful suggestions and comments.

References

- [1] W. W. Symes, Migration velocity analysis and waveform inversion, *Geophysical Prospecting* 56 (6) (2008) 765–790. doi:10.1111/j.1365-2478.2008.00698.x.
URL <http://dx.doi.org/10.1111/j.1365-2478.2008.00698.x>
- [2] B. Biondi, A. Almomin, Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion, *Geophysics* 79 (3) (2014) WA129–WA140. arXiv:<http://dx.doi.org/10.1190/geo2013-0340.1>, doi:10.1190/geo2013-0340.1.
URL <http://dx.doi.org/10.1190/geo2013-0340.1>
- [3] G. Huang, W. W. Symes, Full waveform inversion via matched source extension, in: 85th Ann. Internat. Mtg, 2015, pp. 1320–1325. arXiv:<http://library.seg.org/doi/pdf/10.1190/segam2015-5872566.1>, doi:10.1190/segam2015-5872566.1.
URL <http://library.seg.org/doi/abs/10.1190/segam2015-5872566.1>
- [4] M. Warner, L. Guasch, Adaptive waveform inversion: Theory, in: SEG Technical Program Expanded Abstracts 2014, 2014, pp. 1089–1093. arXiv:<http://library.seg.org/doi/pdf/10.1190/segam2014-0371.1>, doi:10.1190/segam2014-0371.1.
URL <http://library.seg.org/doi/abs/10.1190/segam2014-0371.1>
- [5] L. Métivier, R. Brossier, Q. Mrigot, E. Oudet, J. Virieux, Measuring the misfit between seismograms using an optimal transport distance: application to full waveform inversion, *Geophysical Journal International* 205 (1) (2016) 345–377. arXiv:<http://gji.oxfordjournals.org/content/205/1/345.full.pdf+html>, doi:10.1093/gji/ggw014.
URL <http://gji.oxfordjournals.org/content/205/1/345.abstract>
- [6] Y. Yang, B. Engquist, J. Sun, B. F. Hamfeldt, Application of optimal

transport and the quadratic wasserstein metric to full-waveform inversion, *GEOPHYSICS* 83 (1) (2018) R43–R62. [arXiv:https://doi.org/10.1190/geo2016-0663.1](https://doi.org/10.1190/geo2016-0663.1), doi:10.1190/geo2016-0663.1.

URL <https://doi.org/10.1190/geo2016-0663.1>

- [7] Y. Luo, G. T. Schuster, Wave equation traveltime inversion, *Geophysics* 56 (5) (1991) 645–653. [arXiv:http://dx.doi.org/10.1190/1.1443081](http://dx.doi.org/10.1190/1.1443081), doi:10.1190/1.1443081.

URL <http://dx.doi.org/10.1190/1.1443081>

- [8] B. Chi, L. Dong, Y. Liu, Correlation-based reflection full-waveform inversion, *Geophysics* 80 (4) (2015) R189–R202. [arXiv:http://dx.doi.org/10.1190/geo2014-0345.1](http://dx.doi.org/10.1190/geo2014-0345.1), doi:10.1190/geo2014-0345.1.

URL <http://dx.doi.org/10.1190/geo2014-0345.1>

- [9] E. Bozda, J. Trampert, J. Tromp, Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements, *Geophysical Journal International* 185 (2) (2011) 845–870. doi:10.1111/j.1365-246X.2011.04970.x.

URL <http://dx.doi.org/10.1111/j.1365-246X.2011.04970.x>

- [10] Y. Ma, D. Hale, Wave-equation reflection traveltime inversion with dynamic warping and full-waveform inversion, *Geophysics* 78 (6) (2013) R223–R233. doi:10.1190/geo2013-0004.1.

- [11] H. Yang, Y. Wu, X. Chen, Full waveform inversion overcoming cycle skipping by phase correlation shifting, in: 77th Annual EAGE Meeting, EAGE, Expanded Abstracts, 2015, p. TuP214.

- [12] K. Jiao, D. Sun, X. Cheng, D. Vigh, Adjustive full waveform inversion, in: 85th Ann. Internat. Mtg, 2015, pp. 1091–1095. [arXiv:http://library.seg.org/doi/pdf/10.1190/segam2015-5901541.1](http://library.seg.org/doi/pdf/10.1190/segam2015-5901541.1), doi:10.1190/segam2015-5901541.1.

URL <http://library.seg.org/doi/abs/10.1190/segam2015-5901541.1>

- [13] H. Bi, T. Lin, Impact of adaptive data selection on full waveform inversion, in: 84th Ann. Internat. Mtg, 2014, pp. 1094–1098. [arXiv:http://library.seg.org/doi/pdf/10.1190/segam2014-0310.1](http://arxiv.org/abs/10.1190/segam2014-0310.1), doi:10.1190/segam2014-0310.1.
- 370 URL <http://library.seg.org/doi/abs/10.1190/segam2014-0310.1>
- [14] P. Shen, W. W. Symes, C. C. Stolk, Differential semblance velocity analysis by waveequation migration, in: 73th Ann. Internat. Mtg, 2003, pp. 2132–2135. [arXiv:http://library.seg.org/doi/pdf/10.1190/1.1817759](http://arxiv.org/abs/10.1190/1.1817759), doi:10.1190/1.1817759.
- 375 URL <http://library.seg.org/doi/abs/10.1190/1.1817759>
- [15] P. Sava, B. Biondi, Wave-equation migration velocity analysis. i. theory, Geophysical Prospecting 52 (6) (2004) 593–606. doi:10.1111/j.1365-2478.2004.00447.x.
- URL <http://dx.doi.org/10.1111/j.1365-2478.2004.00447.x>
- 380 [16] B. Biondi, W. Symes, Angle domain common image gathers for migration velocity analysis by wavefield continuation imaging, Geophysics 69 (5) (2004) 1283–1298. [arXiv:http://dx.doi.org/10.1190/1.1801945](http://arxiv.org/abs/10.1190/1.1801945), doi:10.1190/1.1801945.
- URL <http://dx.doi.org/10.1190/1.1801945>
- 385 [17] S. Zhang, G. Schuster, Image-domain full waveform inversion, in: 83rd Ann. Internat. Mtg, 2013, pp. 861–865. [arXiv:http://library.seg.org/doi/pdf/10.1190/segam2013-1238.1](http://arxiv.org/abs/10.1190/segam2013-1238.1), doi:10.1190/segam2013-1238.1.
- URL <http://library.seg.org/doi/abs/10.1190/segam2013-1238.1>
- [18] T. Alkhalifah, Z. Wu, Migration velocity analysis using pre-stack wave fields, Geophysical Prospecting 65 (3) (2017) 639–649. doi:10.1111/1365-2478.12439.
- 390 URL <http://dx.doi.org/10.1111/1365-2478.12439>
- [19] C. Shin, Y. H. Cha, Waveform inversion in the laplace domain, Geophysical Journal International 173 (3) (2008) 922–931. [arXiv:http://](http://arxiv.org/abs/10.1111/1365-2478.12439)

- 395 //gji.oxfordjournals.org/content/173/3/922.full.pdf+html, doi:
10.1111/j.1365-246X.2008.03768.x.
URL <http://gji.oxfordjournals.org/content/173/3/922.abstract>
- [20] W. Hu, Fwi without low frequency data - beat tone inversion, in: 84th Ann.
Internat. Mtg, 2014, pp. 1116–1120. arXiv:[http://library.seg.org/
400 doi/pdf/10.1190/segam2014-0978.1](http://library.seg.org/doi/pdf/10.1190/segam2014-0978.1), doi:10.1190/segam2014-0978.1.
URL <http://library.seg.org/doi/abs/10.1190/segam2014-0978.1>
- [21] Y. Choi, T. Alkhalifah, Unwrapped phase inversion with an exponential
damping, *Geophysics* 80 (5) (2015) R251–R264. arXiv:[http://dx.doi.
org/10.1190/geo2014-0498.1](http://dx.doi.org/10.1190/geo2014-0498.1), doi:10.1190/geo2014-0498.1.
405 URL <http://dx.doi.org/10.1190/geo2014-0498.1>
- [22] Y. E. Li, L. Demanet, Full-waveform inversion with extrapolated low-
frequency data, *Geophysics* 81 (6) (2016) R339–R348. arXiv:[http://dx.
doi.org/10.1190/geo2016-0038.1](http://dx.doi.org/10.1190/geo2016-0038.1), doi:10.1190/geo2016-0038.1.
URL <http://dx.doi.org/10.1190/geo2016-0038.1>
- 410 [23] E. Bozdag, J. Tromp, J. Trampert, Misfit functions for full wave-
form inversion based on instantaneous phase and envelope mea-
surements, *Geophysical Journal International* 185 (2) (2011) 845–
870. arXiv:[http://oup.prod.sis.lan/gji/article-pdf/185/2/845/
1969004/185-2-845.pdf](http://oup.prod.sis.lan/gji/article-pdf/185/2/845/1969004/185-2-845.pdf), doi:10.1111/j.1365-246X.2011.04970.x.
415 URL <https://doi.org/10.1111/j.1365-246X.2011.04970.x>
- [24] B. Chi, L. Dong, Y. Liu, Full waveform inversion method using
envelope objective function without low frequency data, *Journal of Applied Geophysics* 109 (2014) 36 – 46. doi:[http:
//dx.doi.org/10.1016/j.jappgeo.2014.07.010](http://dx.doi.org/10.1016/j.jappgeo.2014.07.010).
420 URL [http://www.sciencedirect.com/science/article/pii/
S0926985114002031](http://www.sciencedirect.com/science/article/pii/S0926985114002031)
- [25] R.-S. Wu, J. Luo, B. Wu, Seismic envelope inversion and modulation signal
model, *Geophysics* 79 (3) (2014) WA13–WA24. arXiv:<http://dx.doi.org/10.1190/geo2014-0498.1>

org/10.1190/geo2013-0294.1, doi:10.1190/geo2013-0294.1.

425 URL <http://dx.doi.org/10.1190/geo2013-0294.1>

- [26] C. Huang, L.-G. Dong, B.-X. Chi, Elastic envelope inversion using multicomponent seismic data with filtered-out low frequencies, *Applied Geophysics* 12 (3) (2015) 362–377. doi:10.1007/s11770-015-0499-8.

URL <https://doi.org/10.1007/s11770-015-0499-8>

- 430 [27] J. Song, Y. Shi, J. Xu, J. Li, Y. Li, X. Cui, Elastic full waveform inversion with envelope based misfit function, in: 85th Ann. Internat. Mtg, Soc. of Expl. Geophys., 2015, pp. 1479–1484.

- [28] J. Luo, R.-S. Wu, Seismic envelope inversion: reduction of local minima and noise resistance, *Geophysical Prospecting* 63 (3) (2015) 597–614. arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/1365-2478.12208>, doi:10.1111/1365-2478.12208.

435 URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/1365-2478.12208>

- [29] J.-W. Oh, T. Alkhalifah, Full waveform inversion using envelope-based global correlation norm, *Geophysical Journal International* 213 (2) (2018) 815–823. arXiv:oup/backfile/content_public/journal/gji/213/2/10.1093/gji/ggy031/1/ggy031.pdf, doi:10.1093/gji/ggy031.

440 URL <http://dx.doi.org/10.1093/gji/ggy031>

- [30] D. Borisov, R. Modrak, F. Gao, J. Tromp, 3d elastic full-waveform inversion of surface waves in the presence of irregular topography using an envelope-based misfit function, *GEOPHYSICS* 83 (1) (2018) R1–R11. arXiv:<https://doi.org/10.1190/geo2017-0081.1>, doi:10.1190/geo2017-0081.1.

445 URL <https://doi.org/10.1190/geo2017-0081.1>

- [31] G.-X. Chen, R.-S. Wu, Y.-Q. Wang, S.-C. Chen, Multi-scale signed envelope inversion, *Journal of Applied Geophysics* 153 (2018) 113 – 126. doi:<https://doi.org/10.1016/j.jappgeo.2018.04.008>.

450

URL <http://www.sciencedirect.com/science/article/pii/S0926985117307772>

- [32] D. Borisov, R. Modrak, F. Gao, J. Tromp, 3d elastic full-waveform inversion
 455 of surface waves in the presence of irregular topography using an envelope-
 based misfit function, *Geophysics* 83 (3) (2018) R1–R11. doi:doi.org/
 10.1190/geo2017-0081.1.
- [33] S. Xu, D. Wang, F. Chen, G. Lambare, Y. Zhang, Inversion on reflected
 seismic wave, in: 82th Ann. Internat. Mtg, Soc. of Expl. Geophys., 2012,
 460 pp. 1–7.
- [34] H. Zhou, L. Amundsen, G. Zhang, Fundamental issues in full waveform
 inversion, in: 82th Ann. Internat. Mtg, Soc. of Expl. Geophys., 2012, pp.
 1–5.
- [35] R. Plessix, Y. De Roeck, G. Chavent, Automatic and simultaneous mi-
 465 gration velocity analysis and waveform inversion of real data using a
 MBTT/WKBJ formulation, in: 75th Ann. Internat. Mtg, Soc. of Expl.
 Geophys., 1995, pp. 1099–1102.
- [36] S. Wang, F. Chen, H. Zhang, Y. Shen, Reflection-based full waveform in-
 version (RFWI) in the frequency domain, in: 83th Ann. Internat. Mtg, Soc.
 470 of Expl. Geophys., 2013, pp. 877–881.
- [37] W. Zhou, R. Brossier, S. Operto, J. Virieux, Full waveform inversion of
 diving and reflected waves for velocity model building with impedance in-
 version based on scale separation, *Geophysical Journal International* 202 (3)
 (2015) 1535–1554. arXiv:[http://gji.oxfordjournals.org/content/](http://gji.oxfordjournals.org/content/202/3/1535.full.pdf+html)
 475 [202/3/1535.full.pdf+html](http://gji.oxfordjournals.org/content/202/3/1535.full.pdf+html), doi:10.1093/gji/ggv228.
 URL <http://gji.oxfordjournals.org/content/202/3/1535.abstract>
- [38] T. Alkhalifah, Z. Wu, The natural combination of full and image-based
 waveform inversion, *Geophysical Prospecting* 64 (1) (2016) 19–30.

- [39] Z. Wu, T. Alkhalifah, Simultaneous inversion of the background velocity
 480 and the perturbation in full-waveform inversion, *Geophysics* 80 (6) (2015)
 R317–R329.
- [40] T. Alkhalifah, Z. Wu, Multiscattering inversion for low-model wavenum-
 bers, *Geophysics* 81 (6) (2016) R417–R428. *arXiv:*[http://dx.doi.org/](http://dx.doi.org/10.1190/geo2015-0650.1)
 10.1190/geo2015-0650.1, doi:10.1190/geo2015-0650.1.
 485 URL <http://dx.doi.org/10.1190/geo2015-0650.1>
- [41] Q. Guo, T. Alkhalifah, Elastic reflection-based waveform inversion
 with a nonlinear approach, *GEOPHYSICS* 82 (6) (2017) R309–
 R321. *arXiv:*<https://doi.org/10.1190/geo2016-0407.1>, doi:10.
 1190/geo2016-0407.1.
 490 URL <https://doi.org/10.1190/geo2016-0407.1>
- [42] T. Van Leeuwen, W. A. Mulder, A correlation-based misfit criterion
 for wave-equation traveltime tomography, *Geophysical Journal Inter-*
national 182 (3) (2010) 1383–1394. *arXiv:*[oup/backfile/content_](http://oupanbackfile/content_public/journal/gji/182/3/10.1111/j.1365-246x.2010.04681.x/2/182-3-1383.pdf)
 495 [public/journal/gji/182/3/10.1111/j.1365-246x.2010.04681.x/2/](http://oupanbackfile/content_public/journal/gji/182/3/10.1111/j.1365-246x.2010.04681.x/2/182-3-1383.pdf)
 182-3-1383.pdf, doi:10.1111/j.1365-246X.2010.04681.x.
 URL <http://dx.doi.org/10.1111/j.1365-246X.2010.04681.x>
- [43] Z. Wu, T. Alkhalifah, Selective data extension for full waveform inversion-
 an efficient solution for cycle skipping, *GEOPHYSICS* 83 (3) (2018)
 R201–R211. *arXiv:*<https://doi.org/10.1190/geo2016-0649.1>, doi:
 500 10.1190/geo2016-0649.1.
 URL <https://doi.org/10.1190/geo2016-0649.1>
- [44] P. Lailly, The seismic inverse problem as a sequence of before stack migra-
 tions, in: *Conference on inverse scattering: theory and application*, 1983,
 pp. 206–220.
- [45] A. Tarantola, Inversion of seismic reflection data in the acoustic approxi-
 505 mation, *Geophysics* 49 (8) (1984) 1259–1266. *arXiv:*<http://dx.doi.org/>

10.1190/1.1441754, doi:10.1190/1.1441754.

URL <http://dx.doi.org/10.1190/1.1441754>

[46] J. Virieux, S. Operto, An overview of full-waveform inversion in exploration
510 geophysics, *Geophysics* 74 (6) (2009) WCC1–WCC26.

[47] Y. Choi, T. Alkhalifah, Application of multi-source waveform inversion
to marine streamer data using the global correlation norm, *Geophysical
Prospecting* 60 (4) (2012) 748–758. doi:10.1111/j.1365-2478.2012.
01079.x.

515 URL <http://dx.doi.org/10.1111/j.1365-2478.2012.01079.x>

[48] R.-E. Plessix, A review of the adjoint-state method for computing the gra-
dient of a functional with geophysical applications, *Geophysical Journal
International* 167 (2) (2006) 495–503. doi:10.1111/j.1365-246X.2006.
02978.x.

520 URL <http://dx.doi.org/10.1111/j.1365-246X.2006.02978.x>

[49] X. Shen, R. G. Clapp, Random boundary condition for memory-efficient
waveform inversion gradient computation, *GEOPHYSICS* 80 (6) (2015)
R351–R359. arXiv:<https://doi.org/10.1190/geo2014-0542.1>, doi:
10.1190/geo2014-0542.1.

525 URL <https://doi.org/10.1190/geo2014-0542.1>

[50] Z. Wu, T. Alkhalifah, The optimized gradient method for full wave-
form inversion and its spectral implementation, *Geophysical Journal In-
ternational* 205 (3) (2016) 1823–1831. arXiv:/oup/backfile/content_
public/journal/gji/205/3/10.1093/gji/ggw112/3/ggw112.pdf, doi:
10.1093/gji/ggw112.

530 URL <http://dx.doi.org/10.1093/gji/ggw112>

[51] Y. Liu, M. K. Sen, A hybrid scheme for absorbing edge reflections in numer-
ical modeling of wave propagation, *GEOPHYSICS* 75 (2) (2010) A1–A6.
arXiv:<https://doi.org/10.1190/1.3295447>, doi:10.1190/1.3295447.

535 URL <https://doi.org/10.1190/1.3295447>