Wave Equation-Based Local Traveltime Inversion

Y. Hu1, L. G. Han2, Y. S. Liu3, and Z. Y. Jin4

1School of Resources and Geosciences, China University of Mining and Technology, Xuzhou, China, 2College of Geo-Exploration Science and Technology, Jilin University, Changchun, Jilin, China, 3Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China, 4SinoProbe Center, Chinese Academy of Geological Sciences, Beijing, China

Abstract Full waveform inversion (FWI) is a strongly nonlinear optimization problem, which suffers from cycle skipping when the initial velocity model is not good enough or the seismic data lack low frequencies. Traveltime tomography is often used to invert the low-wave number components of a velocity model, but traveltime inversion results show a low-resolution model. To bridge the inversion wave number gap between the traveltime tomography and the FWI, a wave equation-based local traveltime inversion is proposed. In this case, the cross-correlation algorithm is applied to calculate the traveltime differences between the localized synthetic and observed data. To mitigate traveltime crosstalk noise caused by multiple seismic events, a sliding Gaussian window is applied to transform the seismic data into local domain. In this way, the traveltime information of different seismic events can be simultaneously used to improve the inversion results. The numerical examples show that the combination of Local Traveltime Inversion (LTI) and FWI successfully mitigate the cycle skipping and produce satisfactory inversion results, even if the seismic data lack low frequencies and the initial velocity model is far from the true one.

Plain Language Summary The high resolution of the velocity inversion is important for oil and gas exploration. However, the full waveform inversion results may be trapped into the local minima. This paper introduces a wave equation-based local traveltime inversion to provide a good initial model for the FWI method. The numerical results verify the correctness of our idea.

1. Introduction

Full waveform inversion (FWI) aims to minimize the waveform difference between the synthetic and observed data, which can make full use of the amplitude and phase information. During the inversion process, it utilizes the adjoint-state method to update the velocity model that are obtained by a zero-lag cross correlation between the forward propagated source and backward propagated receiver wavefields (Lailly, 1983; Plessix, 2006; Tarantola, 1984). However, FWI is a strongly nonlinear optimization problem, which heavily depends on the low frequencies of seismic data and initial models (Alkhalifah, 2016; Virieux & Operto, 2009; Wu et al., 2014). Unfortunately, the frequencies below 4 Hz in real seismic data are usually contaminated with noise or even not available (Alkhalifah & Choi, 2014; Wang & Rao, 2006). Under such circumstances, FWI can suffer from severe cycle skipping when the phase difference between the synthetic and observed data is more than half a period.

To mitigate the FWI cycle skipping, a lot of approaches have been proposed, which includes multiscale strategy, weak nonlinear misfit, and low-frequency demodulation. The common used multiscale strategy is conducting the FWI from low-frequency to high-frequency components (Bunks et al., 1995; Pratt et al., 1998). In this way, it can successfully bridge the inversion wave number gap between the initial model and the high-resolution results. Besides, wavefield separation can mitigate the nonlinearity of the FWI misfit and avoid falling into the local minima, such as refraction data-based early arrival tomography (Sheng et al., 2006; Yang et al., 2018); damping function-based layer stripping inversion (Choi & Alkhalifah, 2015; Choi & Alkhalifah, 2018; Luo et al., 2018); and Laplace domain-based multiscale FWI (Ha et al., 2015; Shin & Cha, 2008; Shin & Cha, 2009). However, the low frequencies are very expensive and are usually contaminated by noise (Chen et al., 2020; Wu et al., 2014). If the seismic data lack low frequencies, the multiscale FWI still suffers from cycle skipping. To retrieve the low frequency from seismic data, a lot of signal demodulation methods have been incorporated into the FWI process. The Hilbert envelope of the seismic signal contains abundant low frequencies; under such circumstances, the envelope inversion is a promising method to reconstruct the low-wave number components of the subsurface velocity models.
Traveltime inversion is insensitive to the cycle skipping. It can estimate the low-wave number components of the velocity model that can then be used as an initial model for FWI (Hale, 2013; Liu & Zhang, 2017; Ma & Hale, 2013; Sun & Alkhalifah, 2019). However, traveltime inversion is associated with low resolution and limited detectable depth, which cannot meet the requirements for a conventional FWI initial model (Feng & Schuster, 2018; Feng & Wang, 2015; Sheng et al., 2006). To bridge the inversion wave number gap between the traveltime inversion and conventional FWI, a wave equation traveltime inversion was proposed by Luo and Schuster (1991), which measures the time lag at the maximum of the cross correlation between the synthetic and observed data. However, the calculation of the cross correlation-based traveltime difference strongly depends on the selection of seismic events (Zhang & Wang, 2009; Zhou et al., 1995, 1997). Van Leeuwen and Mulder (2010) proposed a weighted norm of the cross-correlation misfit. It is more robust to measure the phase difference between the synthetic and observed data. To improve the low-wave number inversion of the velocity models, a correlation-based reflection FWI was proposed by Chi et al. (2015), which behaves in a more linear way than the traditional waveform misfit. In addition, phase unwrapping-based FWI misfit is similar to the traveltime tomography, which can successfully retrieve the macrostructures of the subsurface velocity model (Choi & Alkhalifah, 2015; Djebbi & Alkhalifah, 2014; Vesevaker, 2017; Vesevaker & Poggiagliolmi, 2015). To avoid the influence of the amplitude information on the velocity inversion, a full traveltime misfit was proposed by Luo et al. (2016). However, these wave equation-based travetime inversion methods require manual picking of the relevant arrivals from the synthetic and observed data. Also, the wave equation-based traveltime inversion achieves a lower-resolution model.

To solve the problem with the current wave equation-based traveltime inversion, a wave equation-based local traveltime inversion (LTI) is proposed to build a good initial velocity model for the conventional FWI. The LTI misfit can take full advantage of multiple seismic events traveltime information, such as turning wave, reflection wave, refraction wave, and scattering wave. To calculate simultaneously the traveltime difference with multiple seismic events, a sliding Gaussian window is used to localize the seismic events and avoid traveltime crosstalk noise. In this case, the local traveltime shifts between the localized synthetic and observed can be calculated automatically by selecting the maximum of the cross correlation (Díaz & Sava, 2015).

The outline of this paper is as follows: First, we show how to transform the seismic data into local domains and calculate the local traveltime difference between the synthetic and observed data. Then, the local traveltime differences are used as a LTI misfit. After that, the LTI gradients and adjoint sources are derived. Numerical examples are then used to demonstrate the effectiveness of the LTI results on mitigating the cycle skipping. Finally, the conclusions of the wave equation-based local traveltime inversion are derived.

2. Local Traveltime Differences

The cross correlation is often used to measure the traveltime differences or the waveform similarity between the two corresponding seismic events (Maggi et al., 2009; Yi et al., 2019; Zhang et al., 2018). However, it is highly sensitive to multiple seismic events. If the seismic signal contains more than one events, the cross-correlation method cannot obtain a physically meaningful traveltime difference between the two seismic traces. Therefore, traditional wave equation-based traveltime inversion methods have to separate the seismic events into different wave arrivals. Considering this problem, a sliding Gaussian window is proposed to localize the seismic data, and thus, we can measure traveltime differences with multiple seismic events, simultaneously. The forward local transformation and inverse local transformations are defined as

\[ U(k, t) = u(t)h(t - k), \]

\[ u(t) = \int u(k, t)h(t - k)dk, \]

where \( U(k, t) \) is the localized seismic data; \( u(t) \) is the seismic data in the time-space domain; \( k \) is the time displacement of the Gaussian window; and \( h(t) \) is a Gaussian window. To make sure that the localized seismic data can transform back to the time-space domain correctly, the Gaussian window function
must satisfy \( \| h(t) \|_2 = 1 \). If the length of Gaussian window is too long, in this case, it can cover one whole seismic event, but it will also produce severe cross-talk noise. If we choose a short Gaussian window to localize the seismic data, it can effectively mitigate the cross-talk noise, but it may separate one seismic event into different parts. Therefore, the length of the Gaussian window function is an important parameter for local traveltime calculation, and it will be discussed in the next section.

In Figure 1a, there is a 100 ms time delay between the synthetic and observed data. According to the Equation 1, the synthetic and observed seismic waveform can be transformed into the local domain with a 0.4 s Gaussian window, which is shown in Figures 1b and 1c, respectively. In this case, a cross-correlation function can be applied to measure the local traveltime differences between the localized synthetic and observed data. The cross-correlation function with the localized synthetic and observed data is defined as

\[
C(k, \tau) = \frac{\int [U(k, t) D(k, t + \tau)] dt}{\int \left[ \frac{1}{2} |U(k, t)|^2 dt \right] \int \left[ \frac{1}{2} |D(k, t)|^2 dt \right]}
\]  

(3)

where \( U(k, t) \) and \( D(k, t) \) are the localized synthetic and observed data. According to Equation 3, the local traveltime difference (\( \Delta \tau \)) can be obtained by seeking the maximum value of \( C(k, \tau) \).

\[
C(k, \Delta \tau) = \max_{\tau} \{ C(k, \tau) \mid \tau \in [-T_0, T_0] \}
\]

(4)

where \( T_0 \) is the estimated maximum traveltime difference between the localized synthetic and observed data and \( \Delta \tau \) corresponds to the traveltime differences. If the \( \Delta \tau = 0 \), it means that the localized synthetic data have been matched with the localized observed data. It also indicates the correct velocity has been found.

Now, we show the local traveltime difference in Figure 2b. The true and calculated traveltime differences are denoted by blue and red lines, respectively. It shows that the red line almost overlaps with the blue one, which means that the calculated traveltime differences are correct. From Figure 2, it proves that the cross correlation is suitable for calculating traveltime differences in the local domain.

3. The Choice of the Window Length

It is known that the correlation-based traveltime difference is sensitive to the neighboring arrivals. If a short Gaussian window is applied, it cannot completely cover one whole seismic event waveform. On the other hand, if a long Gaussian window is applied, it may result in cross-talk noise with the neighboring arrivals. Therefore, it is very important to analyze the influence of the Gaussian window length on the traveltime...
differences. In Figure 3, we show the local traveltime differences with different lengths of the Gaussian window. From Figure 3b, the calculated traveltime difference with a 200 ms Gaussian window is not correct. This is mainly because a short sliding Gaussian window separated the one whole wavelet into many parts, which cannot correctly measure the traveltime difference between the two seismic events. In Figures 3b–3e, as we gradually increase the length of the Gaussian window, the red line gradually matches with the blue one. However, if the length of the Gaussian window is long, it results in lower time resolution (Figure 3e).

Next, we discuss the influence of dominant frequency on the accuracy of the local traveltime differences. In Figure 4, we show the traveltime misfit between the calculated and the true traveltime differences. When the dominant frequency of the seismic wavelet is 10 Hz, the best length of the Gaussian window function should be set to 400 ms (Figure 4, blue line). Additionally, if we gradually reduce the dominant frequency of the seismic wavelet, the suitable window length for LTI misfit should gradually increase. When the dominant frequency of the Ricker wavelet is 4 Hz, the best length of the Gaussian window function should be close to
700 ms. Often, observed data lack low frequencies below 4 Hz; therefore, the frequency band of 4–10 Hz is very important for the conventional FWI. In this case, the length of the Gaussian window can be selected in the range of 400–700 ms. In addition, a multiwindow strategy can be applied in LTI, where the length of the Gaussian window is related to the offset.

In this part, we calculate the traveltime differences with multiple seismic events; the localized synthetic and observed wavelets are shown in the Figure 5. In the local domain, the original seismic waveforms were separated into a lot of independent time series (Figures 5b and 5c). In this way, it can effectively mitigate the cross-talk noise. In Figure 5d, the true and calculated traveltime delay almost matched with each other, which means the cross correlation-based traveltime difference is correct. The numerical results show that the local traveltime difference can effectively mitigate the cross-talk noise between the neighboring arrivals, even if one trace contains multiple seismic events.

From Figure 6b, we show that a short window with 200 ms cannot cover one whole seismic event; in this case, the local traveltime differences are not correct, which is similar to Figure 3a. Gradually increasing the length of the Gaussian window, the calculated traveltime differences slowly match the true traveltime differences (Figures 6c and 6d). However, if the length of the Gaussian window is long, the traveltime difference may produce cross-talk noise, which is caused by the neighboring seismic arrivals (Figure 6e). Therefore, the length of the Gaussian window function is very important.

4. The LTI Misfit and Its Gradient

In the local traveltime inversion (LTI) misfit, we assume that the velocity perturbations cause only traveltime shifts and amplitude differences as long as each arrival in the observed data set has a corresponding arrival in the synthetic data (Luo et al., 2016). Under such assumption, the LTI misfit is to search subsurface parameters by minimizing the local traveltime differences between the localized synthetic and observed data:

$$\sigma = \frac{1}{2} \sum_{ns} \sum_{nr} W |\Delta \tau|^2 dk,$$

where $W$ is a weighting factor; $\Delta \tau \in [-T_0, T_0]$ is the estimated local traveltime differences between the localized synthetic and observed data; $ns$ is the number of seismic sources; and $nr$ is the number of

![Figure 4](image1.png)

**Figure 4.** Traveltime misfit with different lengths of the Gaussian window and different types of the dominant frequency.

![Figure 5](image2.png)

**Figure 5.** (a) Seismic data in the time-space domain; (b) synthetic data in the local domain; (c) observed data in the local domain; and (d) traveltime difference between the synthetic and observed data.
receivers, which are equally distributed on the velocity model surface. The LTI misfit measures the travel-time shifts with multiple seismic events, which can obtain better inversion results than single events traveltime inversion. The derivative of the misfit function with respect to the velocity parameters is

$$\frac{\partial \sigma}{\partial v} = \sum_{n} \sum_{r} \int W|\Delta \tau| \frac{\partial \Delta \tau}{\partial v} dk.$$  (6)

To obtain the traveltime Frechét derivative in Equation 6, we need to find a connective function between the local traveltime differences and the localized pressure wavefield. In this way, the traveltime Frechét derivative can be changed to a waveform Frechét derivative by using the chain rule. The adjoint-state theory then can be used to calculate the LTI gradient. The local traveltime differences were calculated by selecting the maximum value of the cross correlation between the localized synthetic and observed waveforms; therefore, the derivative of the cross-correlation function \(C(k, \tau)\) with respect to \(\tau\) could be used to link the local traveltime differences to the localized wavefield pressure. In this case, the derivative of the correlated time series is zero at the true traveltime difference \(\Delta \tau\) (Luo & Schuster, 1991):

$$\frac{\partial C(k, \Delta \tau)}{\partial \tau} = \frac{\int \frac{\partial D(k, t + \Delta \tau)U(k, t)}{\partial \tau} dt}{\int [U(k, t)]^2 dt [D(k, t + \tau)]^2 dt} = 0$$  (7)

Equation 7 successfully connects the localized seismic waveform to the local traveltime differences. In this case, the local traveltime derivative with respect to velocity is

$$\frac{\partial \Delta \tau}{\partial v} = -\frac{\frac{\partial \bar{C}_{\Delta \tau}}{\partial \Delta \tau}}{\frac{\partial \bar{C}_{\Delta \tau}}{\partial v}} = \frac{1}{\int [D(k, t + \Delta \tau)]^2 dt} \int \frac{\partial D(k, t + \tau)}{\partial \tau} \left[ \frac{\partial U(k, t)}{\partial v} \right] dt,$$  (8)

where \(\bar{C}_{\Delta \tau} = \frac{\partial (C_{\Delta \tau})}{\partial \Delta \tau}\); \(\bar{D}(k, t + \Delta \tau) = \frac{\partial [D(k, t + \Delta \tau)]}{\partial \tau}\); and \(\bar{U}(k, t) = \frac{\partial [U(k, t)]}{\partial \tau}\). Substituting Equations 1 and 8 into Equation 6, we obtain

$$\frac{\partial \sigma}{\partial v} = \int \left( \frac{W \Delta \tau(k) \bar{D}(k, t + \Delta \tau)}{\int [D(k, t + \Delta \tau)]^2 dt} \right) h(t - k) dk \frac{\partial u(t)}{\partial v} dt.$$  (9)

Therefore, the LTI adjoint source is

$$f_s = \int \left( \frac{W \Delta \tau(k) \bar{D}(k, t + \Delta \tau)}{\int [D(k, t + \Delta \tau)]^2 dt} \right) h(t - k) dk.$$  (10)
According to the Equation 2, the LTI adjoint source can be transform back to the time-space domain. Similar to conventional FWI, the LTI gradient can also be computed by the zero-lag cross correlation between the forward wavefield and the backward wavefield. Therefore, the gradient of the LTI misfit can be denoted as

$$\frac{\partial \sigma(v)}{\partial v} = \frac{2}{\nu^3} \int \frac{\partial^2 P_{f} t}{\partial t^2} P_{b} dt,$$

where $P_f$ is the forward wavefield and $P_b$ is the backward wavefield, which is computed by the reverse time propagation of the adjoint source at the receiver location.

### 5. Analysis of the Weighting Factor

This section demonstrates the importance of the weighting factor in the LTI misfit function. To give a theoretical proof, we assume that the velocity perturbations cause only traveltime shift and amplitude difference between the two corresponding seismic events. Under such an assumption,

$$D(k, t + \Delta t) = \epsilon(k) U(k, t)$$

where $\epsilon(k) = \max(D(k, t))/\max(U(k, t))$. Under the assumption of small velocity perturbations, the partial derivative of the localized seismic data can be denoted as

$$\frac{\partial[D(k, t + \Delta t)]}{\partial t} = \frac{D(k, t + \Delta t) - D(k, t)}{\Delta t} = \frac{[\epsilon(k)U(k, t) - D(k, t)]}{\Delta t}.$$

As a result, the adjoint source for LTI is

$$f_s = \int \left( \frac{W[\epsilon(k)U(k, t) - D(k, t)]}{\epsilon(k)[U(k, t)]^2} dt \right) h(t-k) dk.$$

To eliminate the denominator and stabilize the inversion updates, we define a weighting factor as $W = \int \epsilon(k) [\hat{U}(k, t)]^2 dt$. In this case, the weighting factor corresponds to the energy of the localized synthetic data, which helps to attenuate the traveltime noise caused by the neighboring seismic arrivals and the weak amplitude seismic events. The weighting factor also reduces the LTI misfit return back to the conventional FWI misfit, only if the inverted model is close to the true model, such as Equation 14. Therefore, the LTI misfit not only can retrieve the low wave number components with large traveltime differences but also construct the high-resolution inversion results with seismic waveforms.

### 6. Numerical Test

In the numerical test, a Marmousi velocity model (Figure 7a) was used to test the wave equation-based local traveltime inversion (LTI) and the conventional FWI. The initial velocity model is estimated by linear gradient model (Figure 7b) with the velocity range from 1,500 to 4,000 m/s. The size of the Marmousi model is $3 \times 10$ km, with the grid interval of 25 m. There are 60 shots equally distributed on the model surface with 400 receivers distributed also along the model surface to record the seismic wavefield. The recording time of the synthetic and the observed data is 4 s with a time interval of 2 ms. In this article, we use a Ricker wavelet with the dominant frequency of 10 Hz to conduct the seismic forward modeling. To demonstrate the low-frequency independence of the LTI method, a low-cut filter with cutoff frequency of 0–4 Hz was used in the numerical tests.

In Figure 8, we show the LTI misfit between synthetic and observed data, in which the correlation-based method finds it difficult to calculate the local traveltime differences. Therefore, an energy weighting factor is needed to attenuate the cross-talk noise caused by the neighboring seismic arrivals (Equation 5). In this way, the traveltime difference of weak signals will be attenuated. In Figure 8a, the first arrival of observed and synthetic data almost has a 0.1 s traveltime difference. If a short Gaussian window was applied, it cannot cover one whole seismic event. In this case, a short Gaussian window only measures the traveltime difference between the two incomplete seismic waveforms (Figure 8b), which is similar to Figure 6b. As we gradually increase the length of the Gaussian window, the LTI misfit successfully measures that the true
traveltime difference (Figures 8c and 8d). However, if the length of the Gaussian window is too long, it produces cross-talk noise between the neighboring seismic arrivals and also shows the low time resolution (Figure 8e).

In Figure 9, we show the conventional FWI results. When the seismic data lack low frequencies below 4 Hz, the conventional FWI result suffers from cycle skipping, especially for the left side of the Marmousi model (Figure 9a). And then, Figure 9a was used as an initial velocity model for the high-frequency band inversion. The final conventional FWI result is shown in Figure 9b. In this test, we show that the conventional FWI method cannot successfully retrieve the detail structures, if the seismic data lack low frequencies and the initial model is far from the true model.

The LTI results with different lengths of the Gaussian window are shown in Figure 10. Comparison of Figures 9a and 10 shows that the LTI results look different, especially when the length of the Gaussian window is less than 1,100 ms. Thus, the LTI results seem closer to the true Marmousi model than the conventional FWI result. However, when the length of Gaussian window is too long, the LTI result will be severely contaminated by the cross-talk noise (Figure 10d). The LTI results were applied as initial velocity models for the high-frequency band FWI. Comparison of Figures 9b and 11b, we demonstrate that the LTI + FWI result is better than the conventional FWI result. However, as we increase the length of the Gaussian window to 1,500 ms, the LTI + FWI converges to a poor result (Figure 11d), but it is still closer to the true Marmousi model than the conventional FWI result (Figures 9b).
The length of the Gaussian window is an important parameter for the LTI method. If a short Gaussian window is applied, it cannot completely cover one whole seismic event waveform. Whereas if a long Gaussian window is applied, it may have low time resolution and produce cross-talk noise. To solve these problem, a multiwindow inversion strategy is proposed. In this case, the length of Gaussian window is related to the offset, which can be defined as

$$M = \begin{cases} 
0.1 \times \frac{(N)(T)}{(N)}, & n \leq 0.1 \times (N) \\
\frac{(T)(n)}{(N)}, & 0.1 \times (N) < n < (N) 
\end{cases} \quad (15)$$

Figure 9. Conventional FWI result: (a) low-frequency band inversion result (4–6 Hz) and (b) high-frequency band inversion result (4–10 Hz), using (a) as an initial model.

Figure 10. The LTI results with different lengths of the Gaussian window function (4–6 Hz), (a) 400 ms, (b) 700 ms, (c) 1,100 ms, and (d) 1,500 ms.
where \( n \) is the offset, \( N \) is the maximum offset, and \( T \) is the maximum length of the Gaussian window; in this case, it is defined by the maximum traveltime difference between the first arrival and direct wave of the observed data. In this numerical test the maximum offset \( N = 10 \) km and \( T = 1,500 \) ms.

The multiwindow LTI + FWI result is shown in Figure 12, which can successfully retrieve the structures of the Marmousi model. Comparison of Figures 10 and 12a shows that the cycle skipping problem of the conventional FWI result has been mitigated, especially on the left side of the Marmousi model. This is mainly because the multiwindow LTI can successfully invert both low-wave number and high-wave number components at the same misfit: in this case, the far offset seismic data with a long Gaussian window, which retrieves the low-wave number components of the velocity model by using the diving wave with large traveltime differences, and the near offset seismic data with a short Gaussian window, which updates the high wave number components by using the reflection wave with small traveltime difference. Then, the multiwindow LTI result was used as an initial velocity model for the high-frequency band FWI. The final inversion results show in Figure 12b, which verifies that the multiwindow LTI + FWI method is an effective approach for mitigating the cycle skipping.

To test the effect of an inaccurate amplitude seismic wavelet on the multiwindow LTI method, we used two seismic wavelets with different amplitudes (Figure 13a) to generate synthetic and observed data, respectively. Comparing Figures 12a and 13b, we show that the multiwindow LTI with an inaccurate amplitude wavelet still can successfully retrieve the macrostructures of the Marmousi model. Based on this test, we demonstrate that the multiwindow LTI has the ability to work even with the amplitude errors in the source wavelet. This is mainly because we normalized the amplitude between synthetic and observed data in the misfit function; in this way, it is more suitable for the realistic seismic data inversion. Furthermore, we also test the effect of Gaussian noise on the multiwindow LTI method. Comparing of Figures 12a and 14b, the multiwindow LTI also obtains accurate macrostructures of the Marmousi model; the noisy contaminated inversion result is similar to the noise-free one. The noisy test result demonstrates that the multiwindow LTI method is reasonably robust to the Gaussian noise.

Figure 11. The LTI + FWI results, using Figure 11 as initial velocity models, (a) 400 ms, (b) 700 ms, (c) 1,100 ms, and (d) 1,500 ms.
7. Discussion

Under the assumption of small velocity perturbations $\Delta v$, which causes a small traveltime shifts and amplitude differences. In this case, we measure a local traveltime shift $\Delta \tau$ for each time window, which is able to handle traveltime inversion for multiple seismic events using the same misfit function and with weak cross-talk noise. Therefore, the application of multiple seismic events at the same time, which has the potential to reconstruct higher-resolution velocity models. The traveltime information is more linearly related to long-wavelength velocity perturbations; however, the amplitude components in the misfit are also significant to the LTI method. This is mainly because seismic waves propagate in often complex media. As we calculated the local traveltime difference, it may produce traveltime noise. Therefore, the amplitude in the LTI misfit is regarded as a weighting factor to attenuate the traveltime noise, which guarantees that the LTI misfit value decreases steadily. In addition, the amplitude weighting factor in the misfit also can reduce the effect of the Gaussian noise and help obtain low-wave number components of the velocity model, even if the seismic data are heavily contaminated with noise.

The LTI misfit was established by using the local transform; therefore, it needs to transform the seismic data into the local domain, trace by trace. After we finish the computation in the local domain, we use the inverse local transformation to transform the adjoint source back to the time-space domain. In this case, we only need to transform the synthetic and observed data into the local domain and do not need to process the forward wavefield, which saves a lot of computational time. The forward and inverse local transformation process requires additional computation, and the computational cost is similar with the wave equation finite difference forward modeling in this article. Fortunately, the forward and inverse local transformation is a redundant transform; in this way, we can increase the time sampling interval to improve the computation efficiency. In addition, the theory of sparsely resampling and compress sensing can be studied to further improve the computational efficiency. Therefore, with the application of the sparsely resampling strategy,
the cost of the local transformation in comparison with the wave equation finite difference forward modeling is small.

The length of the Gaussian window is an important parameter for the LTI method. If a long Gaussian window is applied, it can measure a large traveltime difference between the synthetic and observed data; in this case, the LTI gradient will be rich with low-wave number components. Whereas if a shorter Gaussian window is applied, the LTI gradient will be in rich in higher-wave number components. Therefore, to make full use of advantages of the far offset seismic data, we propose a multiwindow strategy, which allows for the length of the Gaussian window to increase with offset. Based on the multiwindow inversion strategy, we only need to extract the maximum traveltime difference between the first arrival and direct wave in the observed data to automatically generate different lengths of the Gaussian window for different offset seismic data. However, for different situations, the length of the Gaussian window should be changed a little. It is still difficult to choose an optimal length Gaussian window for the LTI method, which requires further study.

8. Conclusions

To improve the resolution of the traveltime inversion and mitigate the cycle skipping of the FWI, a wave equation-based local traveltime inversion (LTI) is proposed. In this case, a local transformation based Gaussian window function is used to localize the seismic waveforms. In this way, it can effectively mitigate the influence of cross-talk noise between the neighboring seismic arrivals. In the LTI misfit, a weighting factor is applied to attenuate the local traveltime noise with weak amplitude seismic events. In this way, it can successfully guarantee the stability of the LTI misfit to converge to the global minimum. In addition, a multi-window strategy is applied to reduce the influence of the unsuitable length of the Gaussian window on the LTI result, which allows for updates of low- and high-wave number components of the velocity model using the same misfit. It also significant for improving the resolution of the LTI result. From the numerical examples, the LTI misfit only measures the local traveltime differences, which is able to recover a kinematically accurate velocity model, even starting with a quite-inaccurate initial guess and seismic data lacking low frequencies.

Conflict of Interest

The authors report no conflict of interest.

Data Availability Statement

Data are available through a repository (https://zenodo.org/record/3742440#.XovKXdNipZs).

References

Alkhalifah, T. (2016). Full model wavenumber inversion (FMWI): An emphasize on the appropriate wavenumber continuation. Geophysics, 81(3), R89–R98. https://doi.org/10.1190/geo2015-0537.1
Alkhalifah, T., & Choi, Y. (2014). From tomography to full-waveform inversion with a single objective function. Geophysics, 81(5), R261–R274. https://doi.org/10.1190/geo2015-0353.1

Bhardwaj, P., Mulder, W., & Drijkoningen, G. (2016). Full waveform inversion with an auxiliary bump functional. Geophysical Journal International, 206(2), 1076–1092. https://doi.org/10.1093/gji/ggw129

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

Bunks, C., Saleeck, F., Zaleski, S., & Chavent, G. (1995). Multiscale seismic waveform inversion. Geophysics, 60(5), 1457–1473. https://doi.org/10.1190/1.1443880

Chen, G., Chen, S., & Yang, W. (2020). Reflection waveform inversion based on full-band seismic data reconstruction for salt structure inversion. Geophysical Journal International, 220(1), 235–247. https://doi.org/10.1093/gji/ggg442

Chen, G. X., Wu, R. S., & Chen, S. C. (2019). Multiscale direct envelope inversion: Algorithm and methodology for application to the salt structure inversion. Earth and Space Science, 6(1), 174–190. https://doi.org/10.1002/2018EA000453

Choi, Y., & Alkhalifah, T. (2014). From tomography to full-waveform inversion. Geophysical Journal International, 199(2), 1385–1398. https://doi.org/10.1093/gji/ggu025

Choi, Y., & Alkhalifah, T. (2015). Unwrapped phase inversion with an exponential damping. Geophysics, 80(4), R189–R202. https://doi.org/10.1190/geo2014-0345.1

Choi, Y., & Alkhalifah, T. (2015). Unwrapped phase inversion with an exponential damping. Geophysics, 80(5), R251–R264. https://doi.org/10.1190/geo2014-0498.1

Djebbi, R., & Alkhalifah, T. (2014). Traveltime sensitivity kernels for wave equation tomography using the unwrapped phase. Geophysical Journal International, 197(2), 975–986. https://doi.org/10.1093/gji/ggt025

Djebbi, R., & Alkhalifah, T. (2014). Traveltime sensitivity kernels for wave equation tomography using the unwrapped phase. Geophysical Journal International, 197(2), 975–986. https://doi.org/10.1093/gji/ggt025

Feng, B., & Wang, H. (2015). Data-domain wave equation reflection traveltome tomography. Journal of Earth Science, 26(4), 487–494. https://doi.org/10.1007/s12583-015-0562-7

Feng, S., & Schuster, G. T. (2018). Transmission + reflection anisotropic wave-equation traveltome and waveform inversion. Geophysical Prospecting. https://doi.org/10.1111/1365-2478.12733

Gao, Z., Pan, Z., Gao, J., & Wu, R. (2017). A novel full waveform inversion method based on a time-shift nonlinear operator. Geophysical Journal International, 208(3), 1672–1689. https://doi.org/10.1093/gji/ggw478

Ha, W., Kang, S. G., & Shin, C. (2015). 3D Laplace-domain waveform inversion using a low-frequency time-domain modeling algorithm. Geophysics, 80(1), R1–R13. https://doi.org/10.1190/1.4893321

Hale, D. (2013). Dynamic warping of seismic images. Geophysics, 78(2), S105–S115. https://doi.org/10.1190/1.4796718

Hu, W. (2014). FWI with low frequency data-beat tone inversion. In SEG Technical Program Expanded Abstracts 2014. Denver: Society of Exploration Geophysicists, pp. 1116–1120.

Lailly, P. (1983). The seismic inverse problem as a sequence of before stack migrations. In J. B. Bednar, R. Redner, & E. Robinson (Eds.), Conference on inverse scattering: Theory and application (pp. 206–220). Philadelphia, PA: Society for Industrial and Applied Mathematics.

Li, Y. E., & Demanet, L. (2016). Full-waveform inversion with extrapolated low-frequency data. Geophysics, 81(6), R339–R348. https://doi.org/10.1190/geo2016-0626-M5

Lian, S., Yuan, S., Wang, G., Liu, T., Liu, Y., & Wang, S. (2018). Enhancing low-wavenumber components of full-waveform inversion using an improved wavefield decomposition method in the time-space domain. Journal of Applied Geophysics, 157, 10–22. https://doi.org/10.1016/j.jappgeo.2018.06.013

Liu, Z., & Zhang, J. (2017). Joint traveltome, waveform, and waveform envelope inversion for near-surface imaging. Geophysics, 82(4), R235–R244. https://doi.org/10.1190/geo2016-0356.1

Luo, J., Wu, R.-S., & Gao, F. (2018). Time-domain full waveform inversion using instantaneous phase information with damping. Journal of Geophysics and Engineering, 15(3), 1032–1041. https://doi.org/10.1088/1742-2140/aaa984

Luo, Y., Ma, Y., Wu, Y., Liu, H., & Cao, L. (2016). Full-traveltime inversion. Geophysics, 81(5), R261–R274. https://doi.org/10.1190/geo2015-0353.1

Luo, Y., & Schuster, G. T. (1991). Wave-equation traveltome inversion. Geophysics, 56(5), 645–653. https://doi.org/10.1190/1.1443081

Ma, Y., & Hale, D. (2013). Wave-equation reflection traveltome inversion with dynamic warping and full-waveform inversion. Geophysics, 78(6), R223–R233. https://doi.org/10.1190/1.40041.1

Maggi, A., Tape, C., Chen, M., Chao, D., & Tromp, J. (2009). An automated time-window selection algorithm for seismic tomography. Geophysical Journal International, 178(1), 257–281. https://doi.org/10.1111/j.1365-246X.2009.04099.x

Ovcharenko, O., Kazei, V., Mahesh, K., Peter, D., & Alkhalifah, T. (2019). Deep learning for low-frequency extrapolation from multi-offset seismic data. Geophysics, 84(4), R1001–R1013. https://doi.org/10.1190/geo2018-0884.1

Plessis, R. E. (2006). A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. Geophysical Journal International, 167(2), 495–503. https://doi.org/10.1111/j.1365-246X.2006.02979.x

Pratt, R., Shin, C., & Hicks, G. (1998). Gauss-Newton and full Newton method in frequency domain seismic waveform inversion. Geophysical Journal International, 138(2), 341–362. https://doi.org/10.1046/j.1365-246X.1998.00498.x

Sheng, J., Leedis, A., Buddensiek, M., & Schuster, G. T. (2006). Early arrival waveform tomography on near-surface refraction data. Geophysics, 71(4), U47–U57. https://doi.org/10.1190/1.2210969

Shin, C., & Cha, Y. H. (2008). Waveform inversion in the Laplace domain. Geophysical Journal International, 173(3), 922–931. https://doi.org/10.1111/j.1365-246X.2008.03786.x

Shin, C., & Cha, Y. H. (2009). Waveform inversion in the Laplace-Fourier domain. Geophysical Journal International, 177(3), 1067–1079. https://doi.org/10.1111/j.1365-246X.2009.04100.x

Sun, B., & Alkhalifah, T. (2019). Adaptive traveltome inversion. Geophysics, 84(4), U13–U29. https://doi.org/10.1190/geophysics-2018-0595.1

Tanatola, A. (1984). Inversion of seismic reflection data in the acoustic approximation. Geophysics, 49(8), 1259–1266. https://doi.org/10.1190/1.1441754

Van Leeuwen, T., & Mulder, W. A. (2010). A correlation-based misfit criterion for wave-equation traveltome tomography. Geophysical Journal International, 182(3), 1383–1394. https://doi.org/10.1111/j.1365-246X.2010.04681.x
Vesnaver, A. (2017). Instantaneous frequency and phase without unwrapping. *Geophysics, 82*(1), F1–F7. https://doi.org/10.1190/geo2016-0185.1

Vesnaver, A., & Poggiagliolmi, E. (2015). Stable phase unwrapping: A new tool for full-waveform inversion. In *Seg Workshop on Full-waveform Inversion*. Abu Dhabi: Society of Exploration Geophysicists. Retrieved from https://www.researchgate.net/publication/281409991_Stable_phase_unwrapping_A_new_tool_for_full-waveform_inversion

Virieux, J., & Operto, S. (2009). An overview of full-waveform inversion in exploration geophysics. *Geophysics, 74*(6), WCC1–WCC26. https://doi.org/10.1190/1.3238367

Wang, G., Yuan, S., & Wang, S. (2019). Retrieving low-wavenumber information in FWI: An efficient solution for cycle skipping. *IEEE Transactions on Geoscience and Remote Sensing Society, 16*(7), 1125–1129. https://doi.org/10.1109/lgrs.2019.2892998

Wang, Y., & Rao, Y. (2006). Crosshole seismic waveform tomography—I. Strategy for real data application. *Geophysical Journal International, 166*(3), 1224–1236. https://doi.org/10.1111/j.1365-246X.2006.03030.x

Wu, R. S., Luo, J. R., & Wu, B. (2014). Seismic envelope inversion and modulation signal model. *Geophysics, 79*(3), WA13–WA24. https://doi.org/10.1190/geo2013-0294.1

Yang, H., Juan, W., & Taikun, S. (2018). Jianzhong Z frequency-domain multi-scale early-arrival waveform tomography with a time-domain wavefield modeling engine. *Journal of Earth Science China, 29*(5), 679–688. https://doi.org/10.1007/s13583-018-0828-y

Yi, J., Liu, Y., Yang, Z., Lu, H., He, B., & Zhang, Z. (2019). A least-squares correlation-based full traveltime inversion for shallow subsurface velocity reconstruction. *Geophysics, 84*(4), R613–R624. https://doi.org/10.1190/geo2018-0581.1

Zhang, Y., & Wang, D. (2009). Traveltime information-based wave-equation inversion. *Geophysics, 74*(6), WCC27–WCC36. https://doi.org/10.1190/1.3243073

Zhang, Z., Alkhalifah, T., Wu, Z., Liu, Y., He, B., & Oh, J. (2018). Normalized nonzero-lag crosscorrelation elastic full-waveform inversion. *Geophysics, 84*(1), R15–R24. https://doi.org/10.1190/geo2018-0082.1

Zhou, C., Cai, W., Luo, Y., Schuster, G. T., & Hassanzadeh, S. (1995). Acoustic wave-equation traveltime and waveform inversion of crosshole seismic data. *Geophysics, 60*(3), 765–773. https://doi.org/10.1190/1.1822667

Zhou, C., Schuster, G. T., Hassanzadeh, S., & Harris, J. M. (1997). Elastic wave-equation traveltime and waveform inversion of crosshole seismic data. *Geophysics, 62*(3), 853–868. https://doi.org/10.1117/12.187494