Reflection full waveform inversion

YAO Gang¹ & WU Di^{2*}¹ Department of Earth Science and Engineering, Imperial College London, London SW7 2BP, UK;² The Unconventional Natural Gas Institute, China University of Petroleum, Beijing 102249, China

Received November 9, 2016; accepted July 31, 2017; published online September 18, 2017

Abstract Because of the combination of optimization algorithms and full wave equations, full-waveform inversion (FWI) has become the frontier of the study of seismic exploration and is gradually becoming one of the essential tools for obtaining the Earth interior information. However, the application of conventional FWI to pure reflection data in the absence of a highly accurate starting velocity model is difficult. Compared to other types of seismic waves, reflections carry the information of the deep part of the subsurface. Reflection FWI, therefore, is able to improve the accuracy of imaging the Earth interior further. Here, we demonstrate a means of achieving this successfully by interleaving least-squares RTM with a version of reflection FWI in which the tomographic gradient that is required to update the background macro-model is separated from the reflectivity gradient using the Born approximation during forward modeling. This provides a good update to the macro-model. This approach is then followed by conventional FWI to obtain a final high-fidelity high-resolution result from a poor starting model using only reflection data. Further analysis reveals the high-resolution result is achieved due to a deconvolution imaging condition implicitly used by FWI.

Keywords Full-waveform inversion, Reflection full-waveform inversion, Least-squares reverse-time migration, Tomographic gradient, Reflectivity gradient, Deconvolution imaging condition

Citation: Yao G, Wu D. 2017. Reflection full waveform inversion. *Science China Earth Sciences*, 60: 1783–1794, doi: [10.1007/s11430-016-9091-9](https://doi.org/10.1007/s11430-016-9091-9)

1. Introduction

As the advance of computer hardware and three-dimensional wave equation simulation technique (Zhang and Yao, 2013; Liu, 2013; Wang et al., 2014; Yao et al., 2016), the three-dimensional full-waveform inversion (FWI) has been successfully applied to field data sets to update the velocity macro-model (Warner et al., 2013; Kapoor et al., 2013). Because of the combination of optimization algorithms and full wave equations, FWI can, in turn, offer more accurate velocity model with higher resolution than ray-based tomography and therefore dramatically improve pre-stack depth migration (PSDM) results (Ratcliffe et al., 2011; Silva et al., 2016). Since FWI recovers the physical parameters of the subsurface, for example, the P-wave velocity, the inversion result can be used to interpret the physical property of the

subsurface directly. One of the successful applications is to identify the shallow gas hazards (Bright et al., 2015). Besides the successes in the field of oil and gas exploration, FWI has been successfully applied to other branches of geophysics, for example, crust structure geophysics (Operto et al., 2006; Kamei et al., 2012) and volcanology (Arnoux et al., 2017).

However, a successful application of conventional FWI on field data must meet some tough conditions: firstly, long-offset data are normally required with turning energy reaching to the velocity target; secondly, the initial model must be close to the true model; and thirdly, low frequencies, typically reaching down to around 3 to 4 Hz are required in the field data. These conditions apply because conventional FWI works to update the macro-model at long wavelengths principally by using transmitted arrivals, and these then are especially vulnerable to the detrimental effects of cycle skipping. In practice, conventional FWI has been most-productively applied in this way to OBC and similar surveys that are dominated

* Corresponding author (email: wudi@cup.edu.cn)

by refracted waves. However, conventional FWI does not so easily update the macro-model when it is applied to reflection-dominated data sets, for example, moderate-offset towed-streamer surveys targeting deeper reservoirs. Nevertheless, the reflection wave in this type of surveys is the only information source to explore the deeper Earth structure.

There are two main reasons. Firstly, the initial model must be extremely accurate in order to migrate reflectors correctly into depth, at which point conventional FWI comes gradually to resemble PSDM plus post-stack inversion albeit in a formulation that properly honors the wave equation. Secondly, the gradient generated by conventional FWI typically favors updating reflector amplitudes more than it favors updating the background macro-velocity model.

To deal with this problem, Xu et al. (2012) proposed the concept of reflection FWI, which is one version of FWI aiming to correct the background macro-velocity model with pure reflection data. In their method, the high-frequency component of the velocity perturbation in the model is given by the true-amplitude reverse-time migration (Zhang et al., 2007) in their implementation. Although this true-amplitude migration provides relatively accurate AVO, in theory, it does not guarantee the modeled data during de-migration can match the observed data in terms of amplitude. Zhou et al. (2015) and Wu and Alkhalifah (2015) built a joint objective function by updating the long-wavelength background velocity and the short-wavelength velocity interface alternately or simultaneously. This is achieved by separating the gradient of FWI into the background velocity component and the velocity interface component. Their examples show the effectiveness of the background velocity updates. All the objective functions of the methods above are built in the data domain. The objective function for recovering the background velocity can be built in the image domain as well (Symes and Carazzone, 1991; Shen et al., 2005; Symes, 2008). The superiority of the schemes in the image domain is they are immune to the detrimental effect of the subsurface structure. However, this type of methods requires multiple elimination.

To push FWI to work better with pure reflection data, we

here explore alternating inversion schemes that include two steps. The first step builds temporary reflectors, which are rebuilt after the second step. The second step is then to update the background velocity model using the reflections generated by the temporary reflectors. By alternating the two steps, the background velocity model can be successively improved, and the reflections become more focused. A key additional ingredient is to formulate the methodology so that it does not require an accurate starting model, and to modify the raw gradient so that it favors macro-velocity updates. Then conventional FWI can lead to the final high-accuracy and high-resolution velocity model using the reflection data. Thus, deeper Earth's information can be captured more accurately with this method.

2. Theory

Conventional FWI is implemented by minimizing the objective function of the L2-norm of the data residual, which is the difference between the predicted and observed data (Tarantola, 1984). This minimization is generally achieved with an iterative inversion algorithm that updates the model guided by the direction of the local gradient of the objective function. If such a scheme is applied to pure reflection data with a smooth initial velocity model, then conventional FWI progresses as illustrated in Figure 1.

In the first iteration (Figure 1a), forward wavefield is generated by forward modeling of the source wavelet in the first step. Because the initial velocity model is smooth, the forward wavefield, therefore, consists only of transmitted arrivals. The back-propagated wavefield is generated by back-propagating the residual wavefield in the second step. Reflection wave does not exist in the first step; therefore, the residual consists only of reflection energy of the observed data in the second step. Relative to the amplitudes in the forward wavefield, the back-propagated wavefield will have amplitudes that are of the order of the reflectivity R . The cross-correlation of these wavefields will build a reflector into the model which will also have a reflectivity of order R . Although

Figure 1 Behavior of conventional FWI for pure reflection data using a smooth starting model. (a) At the first iteration, only transmitted energy exists in both the forward and back-propagated wavefields. Cross-correlating these builds interface and does not update velocity. (b) At subsequent iterations, reflected energy exists in both wavefields. Cross-correlation can now update the velocity model, but the magnitude of this tomographic effect varies as R^2 , whereas updates to the reflectivity vary as R .

FWI is being used, the dominant effect will be to perform a version of reverse-time migration (RTM) on the observed data using the starting model as the migration velocity.

During the second iteration, the forward wavefield will consist of a transmitted wavefield with a similar amplitude to that in the first iteration, plus a reflected wavefield of amplitude R . The back-propagated wavefield will contain its own transmitted wavefield which will be of order R , plus a reflection of this wavefield which will be of order R^2 . Cross-correlating the forward and backward wavefields will then continue to build the reflector with an update that is proportional to R and will begin to update the macro-model everywhere above the reflector with an update that is proportional to R^2 . The former acts as reflectivity migration; the latter acts as velocity tomography. Now, since $|R| \ll 1$, the update of the reflectivity will be much stronger than the update of the background velocity. Successive iteration will attempt to match the observed data by adjusting the reflectivity, and will minimally improve the match by adjusting the macro-model.

From this heuristic analysis, it can be seen that there are two key issues that need to be solved in this system. One of these is to build reflectors in their correct locations in the absence of the correct macro-model, and the other is to find an effective means to update the macro-model usefully when the fundamental algorithm being used focuses the majority of its effort upon short-wavelength reflectivity rather than longer-wavelength background velocity. To solve the latter problem, we must enhance the tomographic aspects of FWI with respect to its migration aspects, and to solve the former problem, we must rebuild the reflectivity as the macro-model evolves and improves in order to move the reflectivity to the more accurate location during the iterations.

To do this, we here use an alternating inversion scheme. In the first step, we use a least-squares reverse-time migration (LSRTM) (Yao and Jakubowicz, 2012; Yao and Wu, 2015) to build reflectors with the near offset data set. The predicted data generated from the reflectivity model can match partial recorded data with the same offset in the least-squares sense. However, the far offset data cannot be well matched in terms of the traveltime or phase since a time difference exists between the predicted and recorded data which can be seen from Figure 2. If the time difference is greater than zero, the background velocity above the reflector is slower than its true velocity. Otherwise, the background velocity is over high. This offers reliable information guiding the direction of background velocity updates. In the second step, we use a modified wavefield for FWI to generate the gradient of the macro-model without updating the reflectivity. Having updated the macro-model, the location of the reflectors in the reflectivity model built in the first iteration cannot match the updated macro-model. Thus we set the reflectivity to zero. We then re-migrate in the first step, and iterate this two-stage

Figure 2 The schematic illustration of data selection and data fit in reflection FWI. In its first step, the short offset data is selected for LSRTM. The predicted data of the reflector generated by LSRTM fits the observed data in a least-square sense. However, the predicted data does not fit the far offset data. If the time difference between the predicted data and observed data is greater than zero, then the background velocity above the reflector is lower than its true velocity. Otherwise, the velocity is higher than its true velocity.

process, until an optimal solution is achieved. In the first step, an empirical criterion for the short offset selection is: starting from the point at zero-offset and after the direct arrival, and then gradually increasing the offset with the increase of time. As a result, the mute shape for the LSRTM data looks like a cone or half an ellipse. This selection can guarantee that the mute width is small for early arrives and becomes larger for the later arrivals. Generally speaking, the range of data for LSRTM is smaller than that for background update.

Reflection FWI with LSRTM is an optimization process, whereby the objective function can be defined as

$$\phi(v, r) = \frac{1}{2} \sum_{\mathbf{x}_r, \omega} |Pu_r(\mathbf{x}, \omega) - d(\mathbf{x}_r, \omega)|_2^2, \quad (1)$$

subject to

$$Au_s(\mathbf{x}, \omega) = s(\mathbf{x}_s), \quad (2)$$

and

$$Au_r(\mathbf{x}, \omega) = j\omega r(\mathbf{x}) u_s(\mathbf{x}, \omega), \quad (3)$$

where v denotes the background velocity, r is reflectivity, A is the wave equation operator, $u_s(\mathbf{x}, \omega)$ and $u_r(\mathbf{x}, \omega)$ represent the predicted source and reflected wave fields respectively, and are functions of v , P is the picking matrix used to extract the record at the receivers, $d(\mathbf{x}_r, \omega)$ is the observed data at the frequency of ω at the receiver $\mathbf{x}_r$, j is the imaginary unit. In eq. (3), $j\omega$ is corresponding to the first derivative in the time domain which guarantees that the reflectivity r has zero phase. The derivation of eqs. (2) and (3) can be found in Appendix 1 (available at <http://link.springer.com>) or Wu et al. (2016). The virtual source wavefield on the right side of eq. (3) is the background wavefield u_s which is free of multiples. Thus the equation system is a version of Born approximation. As a result, the procedure of de-multiple is required for this method. However, the whole system becomes more linear and consequently, the convergence of the inversion is boosted.

If the local gradient methods, for example, the steepest-descent and conjugate-gradient methods, is used to minimize the objective function $\phi(v, r)$, then the key step is to find the gradient of the objective function for the model parameters. The reflectivity and velocity gradients of the objective function can be easily derived using adjoint state methods (Plessix, 2006). u_s and u_r are state variables. We define the adjoint state variables λ_{u_s} and λ_{u_r} which are corresponding to the two state variables respectively. The constrained objective function $\phi(v, r)$ can be transferred to an unconstrained objective function through the Lagrange multiplier method,

$$\begin{aligned} \ell(v, r, u_s, u_r, \lambda_{u_s}, \lambda_{u_r}) = & \frac{1}{2} \sum_{\mathbf{x}_r, \omega} |Pu_r(\mathbf{x}, \omega) - d(\mathbf{x}_r, \omega)|_2^2 \\ & - \langle \lambda_{u_s}, Au_s(\mathbf{x}, \omega) - s(\mathbf{x}_s) \rangle \\ & - \langle \lambda_{u_r}, Au_r(\mathbf{x}, \omega) - j\omega r(\mathbf{x})u_s(\mathbf{x}, \omega) \rangle, \end{aligned} \quad (4)$$

where $\langle \cdot, \cdot \rangle$ represents the inner product. The sign of subtraction, which connects the inner product term and the objective function, in eq. (4) can be substituted by the sign of addition. The substitution does not alter the final outcome of the gradient. Differentiating the objective function with respect to the state variables, u_r , and setting the result to zero,

$$\frac{\partial \ell}{\partial u_r} = 0, \quad (5)$$

gives the adjoint state equation,

$$A^\dagger \lambda_{u_r}(\mathbf{x}, \omega) = P^T(Pu_r(\mathbf{x}, \omega) - d(\mathbf{x}_r, \omega)), \quad (6)$$

where $\dagger$ represents conjugate transpose which is equal to reverse-time propagation in the time domain. Since the picking matrix, P , normally is real-valued, which means only the transpose exists. For a single receiver acquisition, the picking matrix is a sparse matrix with non-zero values only on the diagonal. Its transpose, therefore, is equal to itself. Similarly,

$$\frac{\partial \ell}{\partial u_s} = 0, \quad (7)$$

gives

$$A^\dagger \lambda_{u_s}(\mathbf{x}, \omega) = -j\omega r(\mathbf{x})\lambda_{u_r}(\mathbf{x}, \omega). \quad (8)$$

From eqs. (6) and (8), we can see that the adjoint-state variable λ_{u_r} , which is denoted as the long-dashed line in Figure 1b, is obtained by back propagating the residual. So that λ_{u_r} is the incident wavefield during the back propagation. The adjoint-state variable λ_{u_s} , which is represented by the short dashed line in Figure 1b, is obtained by back propagating the virtual source, which is the negative of the first-order derivative of λ_{u_r} with respect to time further weighted by reflectivity r . Thus, λ_{u_s} is the reflection wavefield during the back propagation.

The gradient of the objective function with respect to the

reflectivity r and the velocity v can be found as:

$$\frac{\partial \ell}{\partial r} = \sum_{\omega} -j\omega u_s^\dagger(\mathbf{x}, \omega)\lambda_{u_r}(\mathbf{x}, \omega), \quad (9)$$

and

$$\begin{aligned} \frac{\partial \ell}{\partial v} = & \sum_{\omega} \left(-\left(\frac{\partial A}{\partial v} \right)^\dagger u_s^\dagger(\mathbf{x}, \omega)\lambda_{u_s}(\mathbf{x}, \omega) \right) \\ & + \sum_{\omega} \left(-\left(\frac{\partial A}{\partial v} \right)^\dagger u_r^\dagger(\mathbf{x}, \omega)\lambda_{u_r}(\mathbf{x}, \omega) \right). \end{aligned} \quad (10)$$

As can be seen from eq. (9), the gradient of the objective function with respect to the reflectivity is obtained by zero-lag cross-correlation between the derivative of predicted source wavefield and the back-propagated wavefield. There are two terms for the gradient of the objective function with respect to the background velocity in eq. (10). The first term is the zero-lag cross-correlation between the incident wavefield during forward modeling and the reflected wavefield during back propagation. The second term is the zero-lag cross-correlation between the reflected wavefield during forward modeling and the incident wavefield during back propagation. For the acoustic wave equation with constant density,

$$A = \frac{1}{v^2} \frac{\partial^2}{\partial t^2} - \nabla^2 = -\frac{\omega^2}{v^3} - \nabla^2. \quad (11)$$

Thus,

$$\left(\frac{\partial A}{\partial v} \right)^\dagger = 2\frac{\omega^2}{v^3}. \quad (12)$$

For simplicity, the derivation above is performed in the frequency domain. However, the numerical implementation is carried out in the time domain to take the advantage of the flexibility and full frequency band of the time domain data.

Eqs. (9) and (10) give out the gradient update of the background velocity for LSRTM and FWI, respectively. The gradient-based inversion algorithm such as the steepest-descent and conjugate-gradient methods can be used to fulfill the first step, i.e. LSRTM. For the implementation details of LSRTM, readers are referred to Yao and Wu (2015). The second step which is updating the background velocity by using FWI can be implemented from eq. (10).

A two-layer model is employed to demonstrate the sensitivity of the gradient to the reflector and background velocity between conventional FWI and reflection FWI. The velocity of the upper layer and the bottom layer is 2000 and 2400 m s⁻¹ respectively. The interface is at the depth of 800 m. 151 shots are excited from 1500 to 7500 m by every 40 m on the surface. The source wavelet is a 10 Hz Ricker wavelet. The split-spread acquisition geometry is used with a maximum offset of 1500 m. We carried out two group tests. The first group test is to use an initial velocity model of 1950 m s⁻¹. Consequently, FWI should increase the initial velocity model. Figure 3a and b show the gradient for the first and second iteration of conventional FWI. If we take a

Figure 3 Comparison of the gradients of conventional FWI and reflection FWI from a two-layer model. The width and depth of the model are 9 and 1.3 km, respectively. The velocity of the top layer is 2 km s^{-1} while the velocity of the bottom layer is 2.4 km s^{-1} . The interface is at the depth of 0.8 km. (a)–(d) show the gradients from the starting model of a constant velocity of 1.95 km s^{-1} ; (e)–(h) are the counterparts from the starting model of a constant velocity of 2.053 km s^{-1} . (a) and (e) show the gradient of the first iteration of conventional FWI; (b) and (f) are the counterpart of the second iteration of the conventional FWI. (c) and (g) show the gradient of the first iteration of reflection FWI in the step of building the reflector. (d) and (h) show the gradient of the first iteration of reflection FWI in the step of updating the background velocity.

close look at the results, we can find that the gradient for the first iteration only reflects the loss information of the reflector while the gradient for the second iteration still mainly is dominated by the same information. The gradient in the first iteration is free of the component of the background velocity while in the second iteration it is still much weaker than the reflector component. Figure 3a and b demonstrate that conventional FWI with pure reflection data is much more sensitive to the reflector than the background velocity. This observation is consistent with the diagram in Figure 1 and its relevant statements.

Figure 3c shows the gradient of the reflectivity for the first iteration with near offset data of 400 m from reflection FWI using eq. (9) while Figure 3d shows the gradient of the background velocity for the first iteration with all offset data from reflection FWI using eq. (10). In Figure 3d, the gradient of the velocity is negative, which means the inversion will increase the velocity. This matches the fact that the initial velocity model is slower than the true model. The gradi-

ent of the velocity is quite smooth, which reveals its inherent long-wavelength characteristic. This is because the two wavefields for cross-correlation in eq. (10) propagate with an opposite direction and thus the opening angle is close to 180° . The wavenumber of the gradient can be obtained from

$$|k| = \frac{2\omega}{v} \cos \frac{\theta}{2}, \quad (13)$$

where θ is the opening angle of the cross-correlated wavefields (Virieux and Operto, 2009). Thus, as this angle close to 180° , the wavenumber is nearly zero, which means long wavelength. As can be seen from Figure 3c and d, reflection FWI proposed in this paper overcomes the high sensitivity of conventional FWI to the reflector and is capable of extracting a correct gradient of the background velocity. Figure 3e–h demonstrate the second group test with an initial model of 2053 m s^{-1} . Again we can observe the fact that the gradient of conventional FWI with pure reflection data mainly reflects the loss information of the reflector whilst reflection FWI can

reveal the loss information of the background velocity.

3. Application to synthetic examples

3.1 Three layer model

A layered model (Figure 4a) is chosen as the first example. A smooth low-velocity anomaly is embedded in a high-velocity layer with a speed of 2400 m s^{-1} . The background velocity is 2000 m s^{-1} . 80 shots simulated by the acoustic wave equation are excited from 3000 to 7000 m every 50 m on the surface. The maximum offset is 2400 m with split spread acquisition geometry. One shot record is shown in Figure 4b. The direct wave has been muted and only reflection wavefield is used for FWI. There is no obvious reflection energy because the low-velocity anomaly is quite smooth. Since the absorbing boundary is adopted, the synthetic record is multiple free. For real data application, multiples can be attenuated by pre-processing. One importance consideration of using multiple free data is: multiples will increase the nonlinearity of the objec-

tive function; however, by considering the computational cost for optimizing the objective function, the local gradient algorithms are widely used for solving the linear equations, more precisely speaking the convex problem (Wang, 2010; Yuan and Sun, 2010).

Although the model is very simple, it is difficult to obtain a satisfactory FWI result by using pure reflection without knowing the long-wavelength background velocity. An ideal inversion should recover the velocity above the first reflector and between the two reflectors, the low-velocity anomaly and the accurate depth of the reflector. We do not expect to recover the velocity beneath the second reflector because no reflected energy is recorded beneath the second reflector. Besides, the amplitude of reflection is not only affected by the variation of velocity and density but also by other factors, such as elastic properties, attenuation, and anisotropy. It is not reasonable to assume that the absolute value of reflection amplitude includes any useful information of P-wave velocity in FWI. As the absolute value of reflection amplitude is mapping to the reflectivity in our method, the background velocity

Figure 4 A simple but difficult-to-invert velocity model. (a) True model. (b) A shot record showing only reflections. (c) True model after smoothing. (d) Model recovered by conventional FWI starting from smoothed model-this result is not possible for field data. (e) 1950 m s^{-1} constant-velocity starting model-incorrect everywhere. (f) Model recovered by conventional FWI starting from constant velocity-the long-wavelength model is not recovered. (g) Initial model recovered by reflection FWI. (h) Final model recovered by reflection FWI-close to the ideal result in (d).

update mainly depends on the traveltimes of reflection. That is also because during the background velocity updating the gradient mainly represents the loss information of background velocity (Figure 3c and f), which means the update mostly depends on the traveltimes of reflections.

Figure 4c shows the smooth true velocity model which contains the correct long-wavelength background velocity. The conventional FWI will lead to the inversion result as shown in Figure 4d taking the true model as the initial model. Comparing Figure 4d and a, we know that if the correct background velocity information is included in the initial model, conventional FWI can migrate the pure reflection data to the correct position and generate the accurate reflection boundary. Further, the reflected energy from the accurate reflection boundary can penetrate the low-velocity anomaly and be recorded by the receivers. By using the recorded data containing the low-velocity anomaly information conventional FWI can recover the anomaly feasibly.

Figure 4e shows another initial model with a constant velocity of 1950 m s^{-1} . It varies from the true model with a velocity of 2000 m s^{-1} for the first and third layers. The conventional FWI result is given in Figure 4f. The numerical test illustrates that more iterations or slight alteration of the inversion strategy such as multi-scale have little benefit for the improvement of the inversion result. FWI built the reflector at the incorrect depth in the initial iterations; therefore, the tomography part cannot be fully performed. The conventional FWI is equivalent to LSRTM which maps the reflectivity to the velocity model.

Figure 4g shows the reflection FWI result proposed in this paper with a constant velocity model (Figure 4e). The maximum offset is 2400 m. In each alternating iteration near offset data set of 800 m is used to update the reflectivity using eq. 9 with three iterations. Then the full offset data set is applied to update the background velocity using eq. 10 with three iterations. A few iterations can balance the computation cost and the convergence of the objective function during the reflectivity and background velocity updating. More iterations lead to a better convergence, however, the computation cost increases dramatically with the limited improvement of the inversion result. On the contrary, fewer iterations can reduce

the computational cost but convergence is compromised. The reflectivity is reset to zero after one alternating iteration and it will be rebuilt in next alternating iteration. That is because the reflectivity model is not compatible with the updated velocity model anymore. Figure 4g is the velocity model after 30 alternating iterations. Comparing with the initial model (Figure 4e), reflection FWI retrieves the velocity in the first and second layers quite well and the low-velocity anomaly is recovered as well. Taking the reflection FWI result (Figure 4g) as the initial model, conventional FWI improves the inversion result further (Figure 4h). From Figure 4g and h, we can see that conventional FWI recovers the high-frequency component in the model, i.e., the interface, and modifies the low-velocity anomaly based on reflection FWI after 50 iterations.

Migration with the velocity model obtained from FWI can be used to check the quality of the inversion result. Figure 5 shows the reverse-time migration image. Figure 5a is the migration image with true velocity model while Figure 5b is the migration image with the reflection FWI velocity model (Figure 3g). We limit the model display within 7.5 km, where it is not influenced by edge effect.

By measuring the two migration images, the error of the depth of the upper reflector is almost zero. The maximum error in the absolute depth of the lower reflector is 12 m; this is produced principally by the incomplete resolution of the shallow velocity anomaly. The maximum error in the absolute amplitude of the migrated image is 3% which is also related to the shallow anomaly. The source wavelet used here has a maximum frequency of about 10 Hz, which has a wavelength of 250 m so that the spatial resolution of the velocity model should be around 125 m. Note however that resolution concerns the ability to distinguish two discrete anomalies rather than the ability to place a single anomaly in its correct position which can be done much more accurately.

3.2 Marmousi model

A simple geometrical model with a layered structure is recovered successfully by using reflection FWI in the previous example. As a general and more complicated model, Mar-

Figure 5 Reverse-time depth migration: (a) using the true model, (b) using the inverted mode from reflection FWI.

mousi model (Figure 6) will be used for verifying the effectiveness of the reflection FWI scheme proposed here. The Marmousi model is easy to invert using long-offsets, turning energy, low frequencies, and an accurate starting model. Here we attempt to invert this model using only short offsets, muting the data to remove refracted energy, using a 10 Hz Ricker source wavelet, and an inaccurate starting model. This is a much more challenging proposition, and it much more closely represents the reality of most field data sets than does the approach typically used to invert Marmousi.

509 shots generated by an acoustic wave equation are excited with the source and receiver spacing of 25 m and 12.5 m respectively. Figure 7a shows one shot record in the middle of the model. Figure 7b shows the data that we invert in which all early-arrival refracted energy has been muted. The maximum offset that we include in the inversion is 3500 m which are of a similar size to the depth to the target.

The wavelet is expected to be correct in order to obtain accurate migration image although the inaccurate wavelet can be corrected by inversion (Pratt, 1999). As the most important goal of inversion is updating the background velocity, we only expect the reflector generated in the first step at a reasonable depth and the amplitude (waveform) of predicted data matches that of the recorded data. It is not necessary that the wiggle shape of the reflector is accurate. Thus a certain deviation of the wavelet does little harm to background velocity update, which is quite different from conventional FWI (Delprat-Jannaud and Lailly, 2005).

Figure 8a is the one-dimensional inaccurate initial model which excludes the correct background velocity. Taking this model as the starting model conventional FWI hardly recovers the true velocity without the aid of long-offset refraction and low-frequency data (as shown in Figure 7b). Figure 8b is the retrieval velocity model by conventional FWI with 30 multi-scale iterations. Comparing Figures 6 and 8b, we can find that the inversion recovers the basic geometry of the faults and the anticlines. However, there is still a significant

discrepancy of the depth and geometry of the interfaces between the inverted model and the true model. This is because the major effort of the inversion is devoted to migrating the reflection data to generate the high-frequency stratigraphic interfaces and the background velocity has not been recovered correctly.

Figure 8c is the background velocity obtained from reflection FWI proposed in this paper. Data with offset smaller than 1250 m are used for the reflectivity model inversion whilst the full offset data are used for background velocity model updating. Comparing the recovered background velocity (Figure 8c) with the true model (Figure 6) and the initial model (Figure 8a), we can find that reflection FWI produces a good background velocity, where the background trend has been well recovered. The reflectivity model and the smooth background velocity model are stored and updated independently in this inversion. The forward wavefield and the backward wavefield are generated in two steps by solving two wave equations sequentially: in the first step, generating the incident wavefield with the background velocity, i.e. solving eq. (2) or (6); in the second step, generating the first-order reflected wavefield by using the first-order time derivative of

Figure 6 The true P-wave velocity of Marmousi model.

Figure 7 Synthetic data from Marmousi. (a) Raw record; (b) muted record, in which direct and refracted arrivals are removed. In the inversion, the maximum offset is limited to 3500 m.

Figure 8 Invert the Marmousi model. (a) 1D inaccurate starting model. (b) Conventional FWI started from (a). (c) Macro-model recovered by using reflection FWI with the starting model in (a). (d) Conventional FWI following method (c).

the scaled incident wavefield as the virtual source, i.e. solving eq. (3) or (8). In fact, the simulation of the reflection wavefield is based on Born approximation and thus it cannot correctly deal with multiple reflections. The reflectivity and background velocity are updated iteratively during the inversion. In this alternating iteration scheme, three iterations are applied for building the reflectivity model and another three iterations for building the background velocity. The reflectivity is reset to zero after one alternating iteration and will be rebuilt during next alternating iteration. After many alternating iterations, the smooth background velocity model has been updated successively and the reflectivity is abandoned after each alternating iteration. In this way, reflection FWI concentrates on the macro-background velocity update. Although it does not produce the high-frequency velocity interface information, the result includes the traveltime information which meets the requirement of depth migration. At least it is accurate enough within the bandwidth of reflection waves. Once again we have to point out that this result is obtained from pure reflection data whose offset is short and equivalent to the depth of the target and none refraction energy is introduced.

Taking the background velocity model from reflection FWI (Figure 8c) as the initial model, we get the updated velocity from conventional FWI by using the pure reflection data. The same inversion strategy is applied as Figure 8b with 30 multi-scale iterations. Comparing with the true model (Figure 6) and the inversion result from the 1-D initial model (Figure

8(b)), Figure 8d gives the most satisfying result except the boundary areas and the bottom right portion. The boundary effect results from incomplete data and the problem in the bottom right portion is because reflection FWI over updates this part (i.e. two blue anomalies in Figure 8c). This becomes another challenge of reflection FWI: reflection data only record the information within a limited aperture. Reflection inversion is more underdetermined than refraction inversion and thus even if there is an over update, the objective function will not have a noticeable response.

In order to verify the quality of the background velocity, we compare the inverted reflectivity models (Figure 9) started from the 1-D velocity model, the recovered model of reflection FWI and the true velocity model. We can confirm that the location of reflectors from the recovered model of reflection FWI matches that from the true model except for the bottom right portion. On the contrary, there is large displacement between the location of reflectors from the 1-D starting model and the true model even in the shallow part. This comparison further confirms that our reflection FWI is capable of restoring the macro-background velocity with high quality.

Furthermore, we compare the recorded data with the predicted data generated from the initial model and the recovered model of reflection FWI. Figure 10a is one shot gather from the initial model (Figure 8a) and the true model (Figure 6). The gather is displayed by alternating the shot from the two models by every ten traces. We can find that the reflection events are not aligned up especially for far offset, where the

Figure 9 Inverted reflectivity models started from (a) the 1-D velocity model, (b) the recovered model of reflection FWI, and (c) the true velocity model.

misplacement is more obvious. Figure 10b shows the comparison of the shot generated from the recovered velocity model of reflection FWI (Figure 8c) and the true model. It is obviously shown that the reflection events match much better than that in Figure 10a. This suggests that our reflection FWI can recover the background velocity accurately.

Finally, we compare the common image gather (CIG) of the depth migration from the initial velocity model (Figure 8a), the recovered velocity model of reflection FWI (Figure 8c) and the true velocity model (Figure 6) at the distance of 1.975 and 8.75 km. It can be obviously seen from Figure 11a that the CIG of the depth migration from the initial velocity model contains many migrated events curved downward because of the over high velocity of the left part in the initial model. On the contrary, there are many migrated events curved upward because of the over low velocity of the right part. These issues are nicely addressed in the CIGs generated from velocity model of reflection FWI. The CIGs become flat and are quite close to the CIGs generated from the true model. The background velocity model built with reflection FWI proposed here has been confirmed again to be accurate.

4. Resolution analysis

As can be seen from Figure 8d, conventional FWI with reflection data and an accurate background velocity can produce an extremely high resolution. For example, the fine layer, approximately 35 m thick, near the faults is clearly recovered; however, the half wavelength of the 10 Hz Ricker source wavelet, in the faulted region with velocities of 2300 m s^{-1} , is about 89 m. This implies that FWI can produce even higher resolution images than half the wavelength, which is the resolution of normal migration.

The fundamental reason is that a cross-correlation imaging condition is usually used in normal migration while FWI im-

Figure 10 Comparison of the predicted data of (a) the initial model and (b) the final inverted model of reflection FWI with the recorded data. In this figure, the predicted data and the recorded data are displayed alternatively in order to emphasize the misalignment of the events. From left to right, the recorded data are plotted first and followed by the predicted data. The position of the source is located at 3.1875 km. The dashed lines indicate the mute position of the direct arrivals and refractions.

Figure 11 Comparison of the common image gathers generated by the initial velocity model ((a), (d)), the recovered velocity model with reflection FWI ((b), (e)), and the true velocity model ((c), (f)). (a)–(c) show the gather at the distance of 1.875 km; (d)–(f) show the gather at the distance of 8.75 km.

plicitly applies a deconvolution imaging condition, which can be mathematically shown as

$$\delta v(\mathbf{x}) = -\frac{v^3(\mathbf{x})}{2} \sum_{\omega} \frac{u_r(\mathbf{x}, \omega)}{\omega^2 u_s(\mathbf{x}, \omega)}, \quad (14)$$

where $\delta v(\mathbf{x})$ is velocity update, $v(\mathbf{x})$ is the background velocity at imaging point $\mathbf{x}$, $u_s(\mathbf{x}, \omega)$ and $u_r(\mathbf{x}, \omega)$ are the source and reflected wavefields at imaging point $\mathbf{x}$ respectively. The derivation is based on the linearization of the velocity update (perturbation) δv and reflection data (perturbation) u_r from the Born approximation of the wave equation. For detailed derivation, readers are referred to the Appendix B in Yao and Jakubowicz (2016), where the deconvolution image condition of LSRTM is derived.

5. Conclusions

A scheme of reflection FWI, which is achieved by alternating LSRTM and modified FWI, is proposed for macro-velocity recovery from reflection-dominated datasets. It is less likely to become trapped in a local minimum due to cycle-skipping than is refraction FWI. This is because the first step, LSRTM, can always find a reflector model to fit an average position of recorded data, which brings the predicted data very close to the recorded data; in addition, to the same target, the travel-time of reflections is less than that of refractions.

Since reflection FWI aims to mainly recover the background velocity, it does not provide a complete solution for reflection-dominated data. However, this method can be used successfully to precondition the starting model such that conventional FWI with the reflection data can then proceed towards an accurate final model. This approach appears to provide a robust and general solution for reflection FWI. In addition, the resolution analysis reveals that conventional FWI implicitly applies a deconvolution imaging condition,

which contributes to the high resolution produced by FWI.

Acknowledgements The authors greatly appreciate the two reviewers' and editors' constructive comments and suggestions from the fundamental theory to the implementation details, which helped to clarify and improve this manuscript significantly. This work is partly supported by the National Natural Science Foundation of China (Grant No. 41504106 & 41274099), the Science Foundation of China University of Petroleum (Beijing) (Grant No. 2462015YJRC012) and State Laboratory of Petroleum Resource and Prospecting (Grant No. PRP/indep-3-1508).

References

- Arnoux G M, Toomey D R, Hooft E E E, Wilcock W S D, Morgan J, Warner M, VanderBeek B P. 2017. Seismic evidence that black smoker heat flux is influenced by localized magma replenishment and associated increases in crustal permeability. *Geophys Res Lett*, 5: 1687–1695
- Bright D, Jones C E, Belhassen Y, Brasil R, Macintyre H, França C. 2015. Full waveform inversion for shallow hazard identification on a narrow azimuth dataset. Mardid: 77th EAGE Conference and Exhibition. WS10-C03
- Delprat-Jannaud F, Lailly P. 2005. A fundamental limitation for the reconstruction of impedance profiles from seismic data. *Geophysics*, 70: R1–R14
- Kamei R, Pratt R G, Tsuji T. 2012. Waveform tomography imaging of a megasplay fault system in the seismogenic Nankai subduction zone. *Earth Planet Sci Lett*, 317–318: 343–353
- Kapoor S, Vigh D, Wiarda E, Alwon S. 2013. Full waveform inversion around the world. London: 75th EAGE Conference and Exhibition. We-11-03
- Liu Y. 2013. Globally optimal finite-difference schemes based on least squares. *Geophysics*, 78: T113–T132
- Operto S, Virieux J, Dessa J X, Pascal G. 2006. Crustal seismic imaging from multifold ocean bottom seismometer data by frequency domain full waveform tomography: Application to the eastern Nankai trough. *J Geophys Res*, 111: B09306
- Plessix R E. 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophys J Int*, 167: 495–503
- Pratt R G. 1999. Seismic waveform inversion in the frequency domain, Part

- 1: Theory and verification in a physical scale model. *Geophysics*, 64: 888–901
- Ratcliffe A, Win C, Vinje V, Conroy G. 2011. Full waveform inversion: A North Sea OBC case study. San Antonio: 81st Annual International Meeting. SEG. 2384–2388
- Shen P, Symes W W, Morton S, Calandra H. 2005. Differential semblance velocity analysis via shot profile migration. Huston: 75th Annual International Meeting. SEG. 2249–2252
- Symes W W. 2008. Migration velocity analysis and waveform inversion. *Geophys Prospecting*, 56: 765–790
- Symes W W, Carazzone J J. 1991. Velocity inversion by differential semblance optimization. *Geophysics*, 56: 654–663
- Silva N V, Ratcliffe A, Vinje V, Conroy G. 2016. A new parameter set for anisotropic multiparameter full-waveform inversion and application to a North Sea data set. *Geophysics*, 81: U25–U38
- Tarantola A. 1984. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, 49: 1259–1266
- Virieux J, Operto S. 2009. An overview of full-waveform inversion in exploration geophysics. *Geophysics*, 74: WCC1–WCC26
- Wang Y. 2010. Computational Methods for Inverse Problems and Their Applications. Beijing: Higher Education Press
- Wang Y, Liang W, Nashed Z, Li X, Liang G, Yang C. 2014. Seismic modeling by optimizing regularized staggered-grid finite-difference operators using a time-space-domain dispersion-relationship-preserving method. *Geophysics*, 79: T277–T285
- Warner M, Ratcliffe A, Nangoo T, Morgan J, Umpleby A, Shah N, Vinje V, Štekli I, Guasch L, Win C, Conroy G, Bertrand A. 2013. Anisotropic 3D full-waveform inversion. *Geophysics*, 78: R59–R80
- Wu D, Yao G, Cao J, Wang Y. 2016. Least-squares RTM with L1 norm regularisation. *J Geophys Eng*, 13: 666–673
- Wu Z, Alkhalifah T. 2015. Simultaneous inversion of the background velocity and the perturbation in full-waveform inversion. *Geophysics*, 80: R317–R329
- Xu S, Wang D, Chen F, Zhang Y, Lambare G. 2012. Full Waveform Inversion for Reflected Seismic Data. Copenhagen: 74th EAGE Conference and Exhibition. EAGE. W024
- Yao G, Jakubowicz H. 2012. Non-linear least-squares reverse-time migration. Las Vegas: 82nd SEG Annual Meeting. 1–5
- Yao G, Jakubowicz H. 2016. Least-squares reverse-time migration in a matrix-based formulation. *Geophys Prospecting*, 64: 611–621
- Yao G, Wu D. 2015. Least-squares reverse-time migration for reflectivity imaging. *Sci China Earth Sci*, 58: 1982–1992
- Yao G, Wu D, Debens H A. 2016. Adaptive finite difference for seismic wavefield modelling in acoustic media. *Sci Rep*, 6: 30302
- Yuan Y X, Sun W Y. 2010. Optimization Theories and Methods (in Chinese). Beijing: Science Press
- Zhang J H, Yao Z X. 2013. Optimized finite-difference operator for broadband seismic wave modeling. *Geophysics*, 78: A13–A18
- Zhang Y, Xu S, Bleistein N, Zhang G. 2007. True-amplitude, angle-domain, common-image gathers from one-way wave-equation migrations. *Geophysics*, 72: S49–S58
- Zhou W, Brossier R, Operto S, Virieux J. 2015. Full waveform inversion of diving & reflected waves for velocity model building with impedance inversion based on scale separation. *Geophys J Int*, 202: 1535–1554

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.