REVIEW

A review on reflection-waveform inversion

Gang Yao[1] · Di Wu[2] · Shang-Xu Wang[2]

Abstract

Full-waveform inversion (FWI) utilizes optimization methods to recover an optimal Earth model to best fit the observed seismic record in a sense of a predefined norm. Since FWI combines mathematic inversion and full-wave equations, it has been recognized as one of the key methods for seismic data imaging and Earth model building in the fields of global/regional and exploration seismology. Unfortunately, conventional FWI fixes background velocity mainly relying on refraction and turning waves that are commonly rich in large offsets. By contrast, reflections in the short offsets mainly contribute to the reconstruction of the high-resolution interfaces. Restricted by acquisition geometries, refractions and turning waves in the record usually have limited penetration depth, which may not reach oil/gas reservoirs. Thus, reflections in the record are the only source that carries the information of these reservoirs. Consequently, it is meaningful to develop reflection-waveform inversion (RWI) that utilizes reflections to recover background velocity including the deep part of the model. This review paper includes: analyzing the weaknesses of FWI when inverting reflections; overviewing the principles of RWI, including separation of the tomography and migration components, the objective functions, constraints; summarizing the current status of the technique of RWI; outlooking the future of RWI.

Keywords Full-waveform inversion · Reflection-waveform inversion · Tomographic component · Migration component · Travel time-based objective function · Waveform-based objective function · Constraint

1 Introduction

Seismic data imaging is a key task of seismic data processing. It began with post-stack time migration (Bednar 2005), then pre-stack time and depth migration (e.g., Claerbout 1971; Claerbout and Doherty 1972; Gazdag 1978; Schneider 1978; Stolt 1978; Gazdag and Sguazzero 1984; Ma and Ji 1988; Hill 1990; Wu and Huang 1992; Ristow and Ruhl 1994; Popovici 1996; Stoffa et al. 1990), and reached reverse-time migration recently (e.g., Lailly 1983; Whitmore 1983; Baysal et al. 1983; Zhang et al. 2011). Currently, full-waveform inversion (FWI) represents the latest development

of seismic data imaging and becomes a hotspot in this field. One reason is that it fills the resolution gap between tomography and migration, which is illustrated in Fig. 1. FWI is an optimization process that seeks an Earth model, e.g., P-wave velocity, to best fit the waveform of seismic record in a sense of the norm defined in the objective function. The concept and framework of FWI were proposed by Tarantola (1984), who is a French professor in geophysics. However, limited by the capacity of the computers in that time, Tarantola (1984) did not demonstrate FWI with numerical examples but derived the gradient of bulk modulus, density and source wavelets using functional analysis in the acoustic assumption. This derivation employed the adjoint-state method (Plessix 2006), and thus the gradient can be computed by backward propagating the adjoint source; it avoids formulating, storing and inverting the gigantic Jacobian matrix. Consequently, FWI became a feasible technique later.

Although the framework of conventional FWI was defined in the paper of Tarantola (1984), its title is "Inversion of seismic reflection data"; therefore, the author's main objective is to invert reflection data. However, the

---

Edited by Jie Hao

✉ Shang-Xu Wang
   wangsx@cup.edu.cn

[1] Unconventional Petroleum Research Institute, State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum (Beijing), Beijing 102249, China

[2] College of Geophysics, State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum (Beijing), Beijing 102249, China
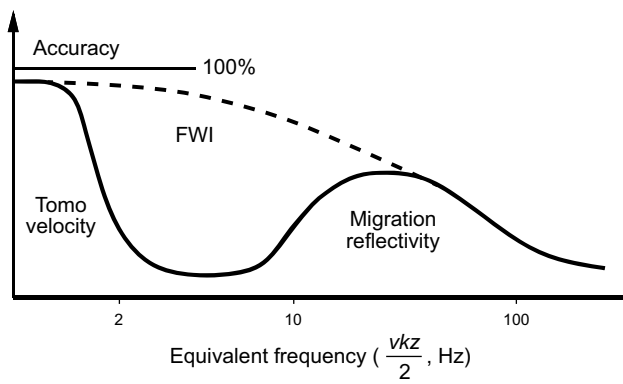
Fig. 1 Schematic illustration of the accuracy and resolution range of tomography, migration and full-waveform inversion. The solid black curve indicates the accuracy and resolution range of the results of tomography and migration (Claerbout 1985). There is a gap between 2 and 10 Hz, which can be filled by the results of full-waveform inversion with large-offset refractions and low-frequency signals (Biondi and Almomin 2013)

conventional FWI proposed by Tarantola is ineffective for inverting reflection data. This will be analyzed fully in the subsequent section.

The second milestone of FWI was made by geophysicist Pratt (1999), who successfully applied FWI on an ultrasonic data set from a cross-hole geometry. Pratt used matrix algebra to reformulate the theory of FWI, which was derived with functional analysis by Tarantola (1984). It makes the theory of FWI more understandable. Pratt (1999) also pointed out the maximum resolution of FWI is the resolution of half wavelength, which is much higher than that of tomography, i.e., the width of the first Fresnel zone. Unfortunately, restricted by the speed and RAM size of computers at that time, Pratt solved the 2D acoustic wave equation in the frequency domain (Štekl and Pratt 1998). Compared to wavefield modeling in the time domain, there are two advantages in the frequency domain: firstly, factorization of the Helmholtz equation needs only once for multi-shot modeling; secondly, it is convenient to achieve attenuation by introducing velocity in complex numbers (Mueller 1983).

The third milestone of FWI was achieved by geophysicist Sirgue in BP (Sirgue et al. 2009), who successfully applied 3D frequency domain FWI on the oil field Valhall using 3D OBC data. It was the first time to demonstrate that FWI is capable of fixing background velocity and recovering the high-frequency structures with industry seismic data (see Fig. 14 of Virieux and Operto 2009). Since then, FWI has attracted wide attention and heavy research investments from both academia and industry. It is difficult to directly solve a 3D wave equation in the frequency domain. Sirgue achieved the frequency domain modeling indirectly by solving time-domain wave equations with finite-difference methods and applying discretized Fourier transform to extract a

few interesting frequencies used for inversion (Sirgue et al. 2008). Afterward, Warner et al. (2013a) applied 3D time-domain anisotropic acoustic wave equation to FWI and successfully inverted the OBC data from the Tommeliten field in the North Sea. The inverted model reflects the gas-cloud boundary clearly and improves the accuracy and continuity of the anticline in the depth migration image significantly.

The open mark of the FWI workshop in SEG of that year (2013) is "Full-waveform inversion has emerged as the final and ultimate solution to the Earth resolution and imaging objective." It emphasized that FWI is the final and ultimate solution of seismic data inversion and imaging. This point of view was questioned by some peers working on other seismic imaging methods, for instance, professor Weglein. He (2013) argued that acoustic FWI uses "wrong mathematic model" (acoustic wave equation), "wrong data" (pure pressure data) and "wrong method" (iterative methods) to invert a more complex Earth model. To some extent, Weglein's argument about the early stage of FWI, i.e., the theory of Tarantola (1984) makes sense. The theory of Tarantola (1984) is based on the acoustic wave equation to invert P-wave velocity and density from pure pressure data using iterative methods. However, the Earth is not pure acoustic. Except for P-wave velocity and density, it includes S-wave velocity and other seismicity parameters. Consequently, acoustic wave equation cannot simulate wave propagation accurately. In addition, pure pressure data are insufficient for recovering P-wave, S-wave velocities and density simultaneously. Furthermore, gradient-based iterative methods could lead FWI to converge a local minimum due to the nonlinearity of FWI.

The main reasons for the successes on FWI achieved by Sirgue and Warner et al. include: firstly, fine-tuned inversion codes; secondly, good inversion strategies; thirdly, high-quality seismic data. Their inversion strategies are elaborated in their papers. Except for high signal-to-noise ratio, OBC/OBS data have some favorable features by FWI: firstly, although the Earth has elastic property, thanks to sea-water filtering, the record by the hydrophones of OBC/OBS is similar to seismic data simulated by using acoustic wave equation; secondly, since the separation between OBC/OBS and air-guns, the offset of records can be very large, which helps to record transmitted (refractions and turning waves) seismic waves from a large depth; finally, the OBC/OBS data include rich low-frequency signals (< 3 Hz). These features help FWI to recover background velocity and mitigate cycle skipping.

Conventional FWI proposed by Tarantola (1984) fixes background velocity using refractions and turning waves mainly; therefore, conventional FWI needs these data in large offsets to repair the deep part of the model. A rule of thumb is that the maximum depth of background that FWI can fix is about a 1/5–1/3 of the maximum offset, which is the maximum penetration depth of turning waves (e.g.,

Zhou et al. (2015) pointed out that the maximum penetration depth is about 1.5 km for the turning waves in a record of an offset of 6 km from the Valhall field). FWI is a nonlinear problem, the reasons of which include: firstly, the wavefields in the wave equation have a nonlinear relationship with the model parameters; secondly, seismic signals oscillate around the zero value, the faster the oscillation (higher frequency), the stronger the nonlinearity. FWI is a large-scale inverse problem, which may contain millions of unknowns. In addition, the computational cost is high for solving wave equations numerically. Thereby, FWI currently is based on local-gradient methods, e.g., steepest-descent and conjugate-gradient methods. In order to avoid trapping in a local minimum due to nonlinearity, FWI usually begins from the lowest available frequency and then gradually increases the frequency. This strategy is called multi-scale proposed by Bunks et al. (1995). This also is the reason why FWI needs low frequencies.

In practice, due to large errors in the initial model, the time lag of the events in predicted data and observed data may be larger than half a period; even the lowest frequency still suffers cycle skipping. Consequently, solving cycle skipping turns into an important topic in FWI studies. Currently, the first category of proposed methods for cycle skipping is to introduce objective functions that have wider convex region; for example, the objective function based on penalizing the nonzero lags of cross-correlation or deconvolution [e.g., Adaptive Waveform Inversion—AWI (Warner and Guasch 2016; Guasch et al. 2019; van Leeuwen and Mulder 2010; Luo and Sava 2011; Zhu and Fomel 2016)], optimal transport distance functions (Métivier et al. 2016; Yang and Engquist 2018; Yang et al. 2018), Wavefield Reconstruction Inversion—RWI (van Leeuwen and Herrmann 2013; da Silva and Yao 2017), etc.

The second category of these methods aims to create lower-frequency signals through signal processing means, for instance, the envelope of seismic data (Wu et al. 2014; Chen et al. 2018a, b), the modulation of two close frequencies (Beat-tone, Hu 2014), downward frequency extrapolation (Warner et al. 2013b; Chen et al. 2019). The methods belonging to the third category utilize the travel time of seismic events to formulate the objective function, e.g., Jiao et al. (2015). Currently, the most widely used method for retraveling the travel-time information is dynamic time/image warping—DTW/DIW from the field of image processing (Hale 2013). Besides, there are some other methods, for example, intermediate data by Wang et al. (2016) and Yao et al. (2019b).

The most direct and efficient method for solving cycle skipping is to improve the acquisition systems for generating and recording lower-frequency signals. One successful example is the ultra-low-frequency marine seismic source, Wolfspar, designed by BP. It can produce the seismic signal as low as 1.6 Hz (Dellinger et al. 2016), which leads to their huge success of applying FWI on the Atlantis oil field in the Gulf of Mexico (Shen et al. 2017). BGP created an onshore controlled vibrator source, which can produce low frequencies as low as 1.5 Hz (Baeten et al. 2013). Unfortunately, so far not many successful FWI examples on land seismic data have been reported due to some complex factors, including low SNR, surface waves, near-surface heterogeneity and elastic effects (Mei and Tong 2015; Cheng et al. 2017; Sedova et al. 2019).

Multi-parameter inversion is another important research direction of FWI. The parameters include P-wave velocity, S-wave velocity, density, anisotropic parameters (ε, δ, γ) and attenuation factor Q. The biggest challenge in multi-parameter inversion is the cross-talk between different types of parameters; it means that the seismic data response generated by the perturbation of one type of parameter could be mapped back into the model perturbation of another type of parameter, for instance, the cross-talk between P-wave velocity and S-wave velocity. In mathematics, the cross-talk is a manifestation of ill-posedness of the inversion problem. The fundamental reason is that the number of independent data is less than the number of unknowns in the model. Except better inversion strategies, e.g., optimized parameterization (Tarantola 1986; Mora 1987; Zhou et al. 2015; Kazei and Alkhalifah 2019; Pan et al. 2019), the inversion system needs extra information to constrain inversion.

As described previously, the record in large offsets drives FWI to update the background velocity model. However, the data in large offsets exist two problems: firstly, low SNR, which is caused by energy attenuation through long wave paths; secondly, strong AVO effects, AVO in the real data follows the rule of elastic (not acoustic) waves on the interfaces (Aki and Richards 1981; Shuey 1985; Castagna et al. 1998; Wang 1999), which causes that predicted data generated by acoustic wave equation cannot match the waveform of recorded data accurately (da Silva et al. 2019). Besides, with current acquisition techniques, only reflections can reach the depth of targets. The three factors lead to the birth of an important branch of FWI—RWI (also called reflection full-waveform inversion—RFWI), which utilizes reflections to recover the background velocity under the framework of FWI. In the rest of this paper, we will review the principle, algorithm and examples of RWI.

2 The weaknesses of conventional FWI for inverting reflection data

Conventional FWI is an optimization problem. In the case of the simplest wave equation, i.e., isotropic acoustic wave equation, the objective function of conventional FWI is the square of L2-norm of data difference expressed as:

$$\phi = \frac{1}{2}\|Du(v) - d\|_2^2, \tag{1}$$

which is constrained by the wave equation,

$$\left(\frac{1}{v^2}\frac{\partial^2}{\partial t^2} - \nabla^2\right)u = s, \tag{2}$$

where d denotes the observed seismic data, including refractions and reflections, u represents the simulated seismic wavefield from the velocity model v, s is the source wavelet, D denotes the receiver operator for extracting out the predicted data from u at the receivers' location, D is a pulse function or identity matrix for a single-channel recording system. If a receiver is formed by a group of detectors, D is a stack of multi-pulse functions, and thus the non-diagonal elements may not be zero. The expression of the gradient of (1) can be found in the papers of Tarantola (1984), Pratt (1999) and Virieux and Operto (2009).

If conventional FWI applies to pure reflections starting with a smooth velocity model, which does not contain interfaces (this is common in real applications), then as illustrated in Fig. 2, the processes of forward and backward propagation generate only transmitted waves in the first iteration. Consequently, the gradient generated with zero-lag cross-correlation builds interfaces but does not contain components for background velocity update. This is because the waves in the forward and backward propagations only meet at the interfaces. In the later iterations, the wave paths of the source and residual wavefields are overlapped on and above the interfaces; therefore, the gradient contains the components for both interface and background update. Based on this observation, Mora (1989) proposed "Inversion = migration + tomography," which means RWI possesses two functionalities, i.e., migration and tomography. Note that Mora did not mean that migration plus tomography is equivalent to inversion. Migration aims for building interfaces, while tomography is for background velocity update.

As depicted in Fig. 2b, in general, the reflectivity of rock interfaces is much smaller than 1, usually less than 0.05. As a result, the gradient at interfaces is one order of magnitude higher than that above the interfaces, which causes FWI based on gradient methods mainly updates interfaces rather than the background velocity model. This is one of the weaknesses of conventional FWI for inverting pure reflection data.

On the contrary, the transmission coefficient is close to 1, for instance, the transmission coefficient is 0.95 if the reflection coefficient is 0.05. As a result, the magnitude of the gradient formed by transmitted waves is relatively uniform along the wave paths; conventional FWI with transmitted waves is good at fixing background velocity as well as rebuilds the interfaces with reflections.

FWI produces a model that contains interfaces after the first iteration. The predicted reflection data from the model fit the observed reflection data in a least-squares sense. Conventional FWI is based on a trace-by-trace comparison. Consequently, if the events in predicted data arrive later than the corresponding events in observed data, then inversion increases velocity; otherwise, inversion decreases velocity. In the case that the velocity of the current model is greater than the actual velocity, the events in predicted data are intercepted by the events in observed data (Fig. 3), and the travel time of the events between the intercept points (indicating small offsets) in predicted data is larger than that in observed data. The travel time outside the intercept points (indicating large offsets) in predicted data is smaller than that in observed data. In the rest iterations, the data between the intercept points drive inversion to increase background velocity for reducing travel time. However, the data outside the intercept points push inversion to decrease background velocity for raising travel time. Consequently, the two parts of data contradict each other, which leads to inversion stuck into a local minimum. The whole inversion then mainly updates the interfaces, which is equivalent to impedance inversion rather than fixing background velocity. This is
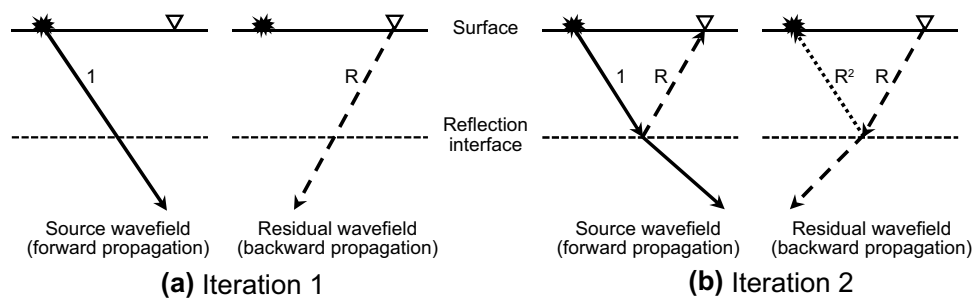

Fig. 2 Behavior of conventional FWI for pure reflection data using a smooth starting model. **a** At the first iteration, only transmitted energy exists in both the source and residual wavefields. Cross-correlating these builds interface and does not update velocity. **b** At subsequent iterations, reflected 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. Usually, R is much smaller than 1 (Yao and Wu 2017)
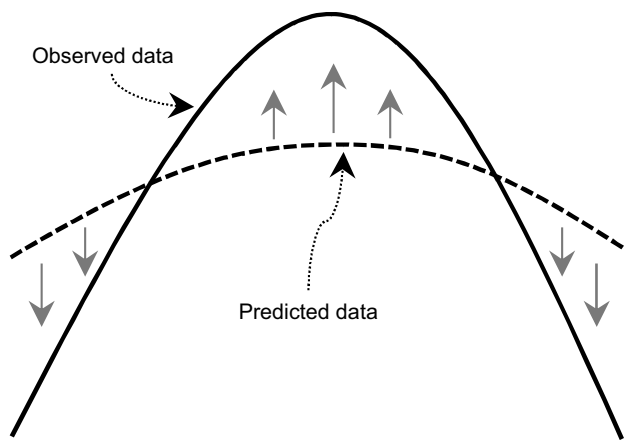
Fig. 3 The move direction of predicted data in conventional FWI. After the first iteration, the reflector that fits the observed data in a least-squares sense is built. In the subsequent iterations, the section of the event between the two intercept points indicates the update direction of the background velocity is opposite to the rest section of the event. Consequently, the inversion is stuck in a local minimum

the significance of FWI inverting reflection-dominated data (Wang and Rao 2009; Lazaratos et al. 2011; Plessix and Li 2013; Routh et al. 2016; Yao et al. 2018b). The contradiction and trapping in a local minimum are the second weakness of conventional FWI for inverting reflection data.

3 Reflection-waveform inversion

Conventional FWI inverting pure reflection data have two weaknesses: firstly, the migration component in the gradient is much stronger than the tomography component; secondly, inversion traps in a local minimum. Thus, the inversion mainly rebuilds the interfaces rather than fixes background velocity. Modifying the algorithm of conventional FWI for fixing background velocity using reflection data is the aim of reflection-waveform inversion (RWI). In order to solve the two weaknesses, studies of RWI mainly focus on the three topics: (1) separation of tomography and migration components; (2) optimizing the objective functions; (3) improving the stability of inversion by incorporating constraints.

3.1 Separation of tomography and migration components

The tomography and migration components in the gradient of conventional FWI are mixed. If the source wavefield shares the same wave path with the residual wavefield, then zero-lag cross-correlation produces the tomography component. If the tomography component is formed from transmitted waves, then its sensitivity kernel has a shape of banana (e.g., Fig. 4a, Fig. 1a of Xu et al. 2012, or Fig. 5b of Virieux

and Operto 2009). If the tomography component is formed from reflection waves, then the kernel is like "rabbit ears" (e.g., Fig. 4b, Fig. 1d of Xu et al. 2012). As the tomography component updates the model along the whole wave path, the tomography component is key for updating background velocity. As illustrated in Fig. 5b of Virieux and Opetro (2009), the resolution of the tomography component to the model is the width of the first Fresnel zone. This is because that other order Fresnel zones will be canceled out and then only the first Fresnel zones will remain if many sensitivity kernels are stacked together.

If the source wavefield travels in a different direction to the residual wavefield, then zero-lag cross-correlation generates the migration component. This situation only happens to reflections at the interfaces. The sensitivity kernel of the migration component has a shape of an ellipse, e.g., Fig. 4b, Fig. 1b of Xu et al. (2012). Multiple ellipse-shaped sensitivity kernels stack together to form the interfaces of the model.

As can be seen from Fig. 2, the migration component is much stronger than the tomography component in the gradient formed from reflections. In order to emphasize the tomographic function of FWI, it is necessary to separate the tomography component in the gradient from its migration component. Currently, the separation methods can be divided into three categories: (1) up/down wavefield separation; (2) scattering angle filtering; (3) born-based modeling.

3.1.1 Up/down wavefield separation

The gradient of the FWI system formed by Eqs. 1 and 2 can be expressed as:

$$g(\mathbf{x}) = -\frac{2}{v^3} \int_0^{T_{max}} \vec{U}(\mathbf{x}, t)\overleftarrow{U}(\mathbf{x}, t)\mathrm{d}t, \tag{3}$$

where $\vec{U}(\mathbf{x}, t)$ and $\overline{U}(\mathbf{x}, t)$ represent the first-order time derivative of the source and residual wavefields, respectively, at the location of $\mathbf{x}$ and the time of t.

If the source and residual wavefields are decomposed into upward and downward wavefields, then the migration and tomography components can be expressed as:

$$g_{\mathrm{mig}}(\mathbf{x}) = -\frac{2}{v^3} \int_0^{T_{max}} \left(\vec{U}_{\mathrm{U}}(\mathbf{x}, t)\overline{U}_{\mathrm{D}}(\mathbf{x}, t) + \vec{U}_{\mathrm{D}}(\mathbf{x}, t)\overline{U}_{\mathrm{U}}(\mathbf{x}, t) \right)\mathrm{d}t, \tag{4}$$

and

$$g_{\mathrm{tomo}}(\mathbf{x}) = -\frac{2}{v^3} \int_0^{T_{max}} \left(\vec{U}_{\mathrm{U}}(\mathbf{x}, t)\overline{U}_{\mathrm{U}}(\mathbf{x}, t) + \vec{U}_{\mathrm{D}}(\mathbf{x}, t)\overline{U}_{\mathrm{D}}(\mathbf{x}, t) \right)\mathrm{d}t, \tag{5}$$
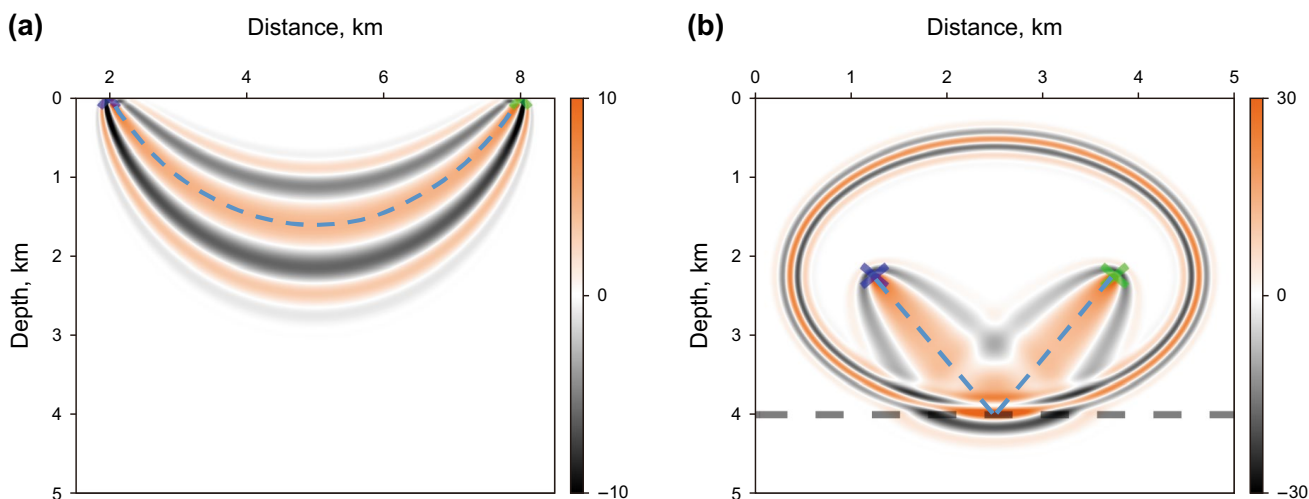
(a)

(b)

Fig. 4 The sensitivity kernels of **a** transmitted waves and **b** reflected waves in conventional FWI. The sensitivity kernel of transmitted waves has a shape of banana because the velocity usually increases with the increase in depth. On the contrary, the sensitivity kernel of reflected waves includes the oval-shaped migration component and the rabbit-ear-like tomographic component. The dashed gray line represents the interface. The blue curves indicate the ray paths of the transmitted and reflected waves. The blue and green crosses "×" represent the location of the source and the receiver, respectively
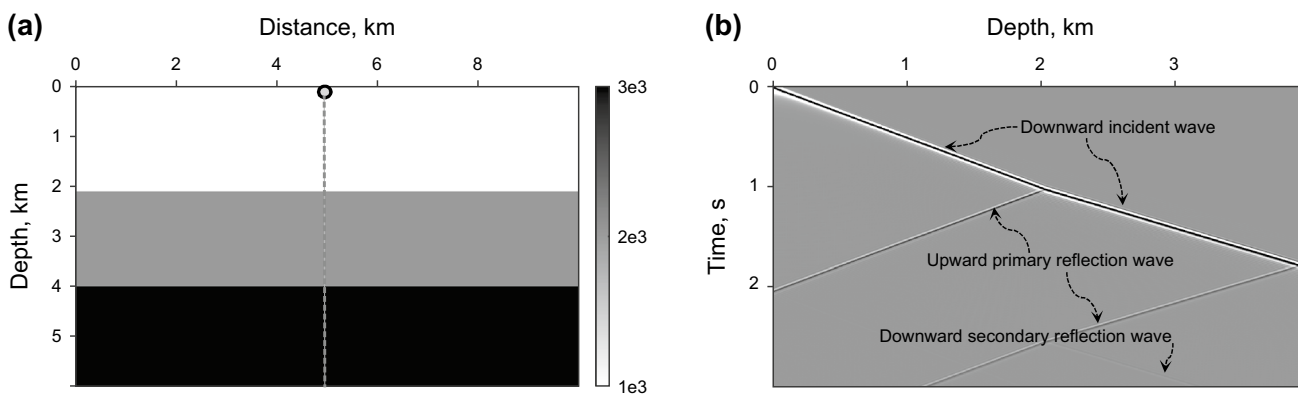
(a)

(b)

Fig. 5 A vertical seismic record profile. **a** A three-layer velocity model. A 25 Hz Ricker wavelet is used as the source signature. The black circle on the surface indicates the source location; the dashed vertical line represents the distribution of the receiver array. **b** shows the vertical profile, in which all the down-going events tilt toward right while all the up-going events tilt toward left

where the subscripts U and D represent the upward and downward wavefields, respectively. As can be seen from Fig. 5, the separation of the upward and downward wavefields can be achieved using Fourier transform along the time and depth directions, in which the part of $\omega k_z \geq 0$ corresponds the downward wavefield while the part of $\omega k_z < 0$ is the upward wavefield (Hu and McMechan 1987). However, Fourier transform along the time and depth axes is prohibited in the time-domain FWI. Liu et al. (2011) derived an image condition that removes the background noise in the RTM image effectively and efficiently. This imaging condition only carries out 1D Fourier transform along the depth axis. The background noise in the RTM image is the tomography component in the FWI gradient, while the rest part is

the migration component. Wang et al. (2013) utilized this imaging condition to separate the two components, which can be expressed as

$$g_{\text{mig}}(\mathbf{x}) = -\frac{2}{v^3} \int_0^{T_{max}} \left(\vec{U}_+^*(\mathbf{x}, t) \overleftarrow{U}_-(\mathbf{x}, t) + \vec{U}_-^*(\mathbf{x}, t) \overleftarrow{U}_+(\mathbf{x}, t) \right) dt$$

(6)

and

$$g_{\text{tomo}}(\mathbf{x}) = -\frac{2}{v^3} \int_0^{T_{max}} \left(\vec{U}_+^*(\mathbf{x}, t) \overleftarrow{U}_+(\mathbf{x}, t) + \vec{U}_-^*(\mathbf{x}, t) \overleftarrow{U}_-(\mathbf{x}, t) \right) dt,$$

(7)

where the superscript * represents complex conjugate, $U_+(\mathbf{x}, t)$ and $U_-(\mathbf{x}, t)$ can be computed by taking inverse Fourier transformation of

$$U_+\left(k_z, t\right) = \begin{cases} U\left(k_z, t\right), & k_z \geq 0 \\ 0, & k_z < 0 \end{cases}, \tag{8}$$

and

$$U_-\left(k_z, t\right) = \begin{cases} 0, & k_z \geq 0 \\ U\left(k_z, t\right), & k_z < 0 \end{cases}. \tag{9}$$

Equations 8 and 9 imply that $U_+(\mathbf{x}, t)$ and $U_-(\mathbf{x}, t)$ can be computed using Hilbert transform alternatively (Liu et al. 2011; Fei et al. 2015; Chi et al. 2017; Lian et al. 2018). The mechanism of Eqs. 6 and 7 for the separation of the incident and reflection waves is based on the assumption that the incident wave travels downwards, while the reflection wave propagates upwards (Fig. 5). However, if the strata have a large dipping angle (e.g., Figure 12 of Liu et al. 2011), this method fails to separate the wavefields. Consequently, the tomography component contains the migration component, vice versa. Since the migration component usually is much stronger than the tomography component, this leakage does not affect the result of RTM very much, but the leakage in the tomography component could lead to RWI updating background velocity along the wrong direction significantly. Wang et al. (2018) from CGG demonstrated that the leakage can cause inversion to update the model toward the wrong direction at the steep flanks of salt bodies. In order to mitigate this issue, Irabor and Warner (2016) proposed to split the wavefield into two parts, one of which travels along the vertical axis and the other of which propagates along the horizontal axis, and then applied Eqs. 6 and 7 to separate the two parts into the upward and downward waves. In the isotropic acoustic media, the vertical and horizontal waves can be simulated by using the equation system,

$$\begin{cases} \frac{1}{v^2} \frac{\partial^2 u_x}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = \frac{s}{2} \\ \frac{1}{v^2} \frac{\partial^2 u_z}{\partial t^2} - \frac{\partial^2 u}{\partial z^2} = \frac{s}{2} \end{cases}, \tag{10}$$

where u_z and u_x represent the vertical and horizontal wavefields.

3.1.2 Scattering angle filtering

The migration component is generated when the source wavefield meets the residual wavefield around the interfaces; the tomography component is formed when the waves meet on the wave path. As illustrated in Fig. 6, the two types of components can be separated according to the scattering angle
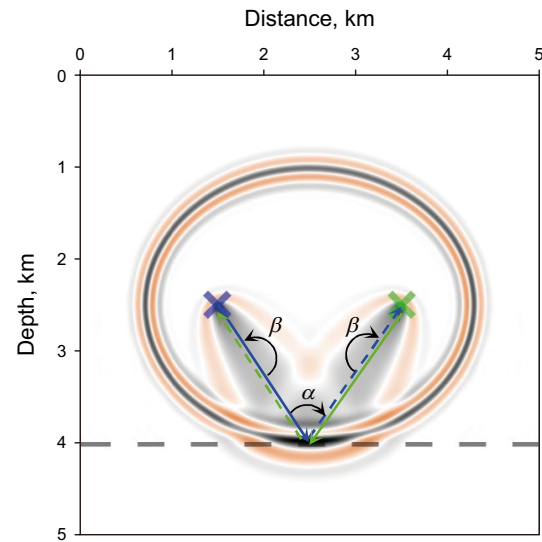

Fig. 6 The gradient formed by the reflection from a pair of "source-receiver" configuration. The interface of the two-layer model is at a depth of 4 km. The blue and green "×" indicate the location of the source and the receiver, respectively. The solid blue line represents the path of the incident wavefield in the source wavefield, while the dashed blue line is for the reflection. The green lines are the counterpart in the residual wavefield. α represents the angle between the source and residual wavefields at the interface; β is for their angle at the rest of the wave path

between the source and residual waves: if the scattering angle is close to 180°, the cross-correlation of the two wavefields forms the tomography component; otherwise, the migration component is generated. Thus, the scattering angle can be used as an indicator to separate the two components. Three methods have been proposed for scattering angle filtering. The first filtering method is based on the time-lag gather (Khalil et al. 2013; Alkhalifah 2014). The gradient in Eq. 3 with time lag can be expressed as:

$$g(\mathbf{x}, \tau) = -\frac{2}{v^3} \int_0^{T_{max}} U_S(\mathbf{x}, t - \tau) U_R(\mathbf{x}, t + \tau)\mathrm{d}t, \tag{11}$$

where U_S and U_R represent the first-order time derivative of the source and residual wavefields, respectively, and τ denotes the time lag. In isotropic media, the wavenumber, frequency and velocity have a relationship of

$$\left|\mathbf{k}_S\right| = \left|\mathbf{k}_R\right| = \frac{\omega}{v(\mathbf{x})}, \tag{12}$$

where $\mathbf{k}_s$ and $\mathbf{k}_R$ represent the wavenumbers of the source and residual wavefields, ω denotes frequency and $v(\mathbf{x})$ is the velocity at the location of $\mathbf{x}$. As depicted in Fig. 7, the relationship of the wavenumber of the gradient, the scattering angle, frequency and velocity can be expressed as
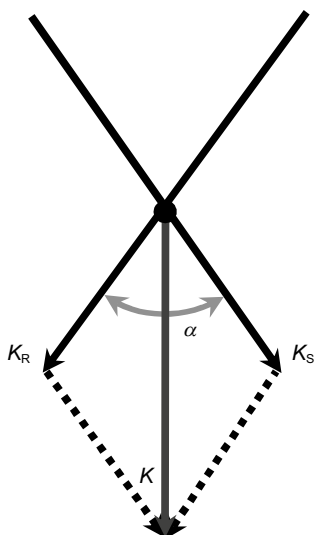
Fig. 8 The transform between data domains in the scattering angle filtering process based on time-shift gathers

Fig. 7 Schematic illustration of the relationship between the wavenumbers of source and residual wavefields and the gradient. $\mathbf{k}_S$, $\mathbf{k}_R$ and $\mathbf{k}$ represent the wavenumber of the source and residual wavefields and the gradient. α is the angle between the source and residual wavefields

$$\frac{|\mathbf{k}|^2}{\omega^2} = \frac{4}{v^2(\mathbf{x})} \cos^2\left(\frac{\alpha}{2}\right). \tag{13}$$

The counterpart of Eq. 13 for the gradient with time lag shown in Eq. 11 can be expressed as

$$\frac{|\mathbf{k}|^2}{\omega_\tau^2} = \frac{4}{v^2(\mathbf{x})} \cos^2\left(\frac{\alpha}{2}\right), \tag{14}$$

where ω_τ represents the frequency of the time-lag gather. Using the variable replacement

$$\xi = \frac{\tau}{2} v(\mathbf{x}), \tag{15}$$

Equation 14 can be expressed as

$$\frac{|\mathbf{k}|^2}{\psi^2} = \cos^2\left(\frac{\alpha}{2}\right), \tag{16}$$

where ψ represents the counterpart of ξ in the Fourier domain. As can be seen from Eq. 16, the scattering angle is independent of velocity; thus, we can achieve the scattering angle filtering using Eq. 16 to any velocity model (Khalil et al. 2013; Alkhalifah 2014).

As can be seen from Fig. 8, the scattering angle filtering based on the time-lag gather includes five steps: (1) forming the gradient with time lag using Eq. 11, and stacking the gradient of all shots; (2) applying coordinate transform on the stacked gather using Eq. 15; (3) taking Fourier transform along all axes; (4) calculating the scattering angle using Eq. 16 and performing scattering angle
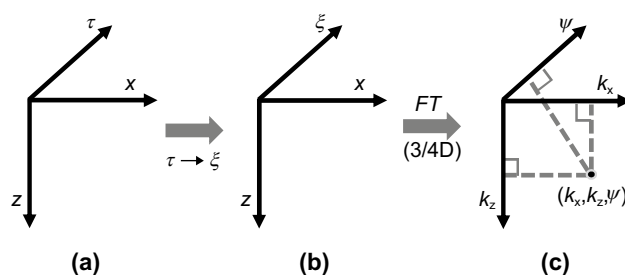
filtering; (5) taking inverse Fourier transform along all wavenumber axes, and then stacking all frequencies to produce the migration and tomography components. One simple angle filtering function can be found in Fig. 2d of Yao et al. (2018a).

The analysis above shows that the scattering angle filtering based on the time-lag gather has limitation: firstly, assuming the velocity is isotropic; secondly, relatively large memory and computation cost is caused due to generating the time-lag gather and high-dimension Fourier transform; Thirdly, the coordinate transform needs resampling, and thus simple interpolation, e.g., spline interpolation, could introduce alias. To improve the efficiency, Wu and Alkhalifah (2017) demonstrated that the separation can be achieved using spatial Fourier transform and simple extra computation if the inversion is carried out in the Frequency domain and the scattering angle is assumed independent on velocity.

The second method of scattering angle filtering is achieved in the plane-wave domain. The key point of scattering angle filtering is to compute the angle between the source and residual wavefields. A natural thought for the angle measurement is in the plane-wave domain. Thus, Yao et al. (2018a) proposed to Fourier transform the source and residual wavefields into the wavenumber domain and then achieve the scattering angle filtering by measuring the angle of the elements in the wavefields. The gradient expressed in Eq. 3 in the wavenumber domain becomes a convolution form,

$$\tilde{g}(\mathbf{k}) = -\frac{2}{v^3} \sum_t \tilde{U}_S(\mathbf{k}_S, t) * \tilde{U}_R(\mathbf{k}_R, t)$$
$$= -\frac{2}{v^3} \sum_t \sum_i \sum_j \tilde{U}_S(\mathbf{k}_{Si}, t) \tilde{U}_R(\mathbf{k}_{Rj}, t), \tag{17}$$

where i and j are the indexes of the first-order time derivative of the source and residual wavefields in the wavenumber domain, respectively, and * represents convolution. The process of convolution is illustrated in Fig. 9. Each element of wavefields in the wavenumber domain represents
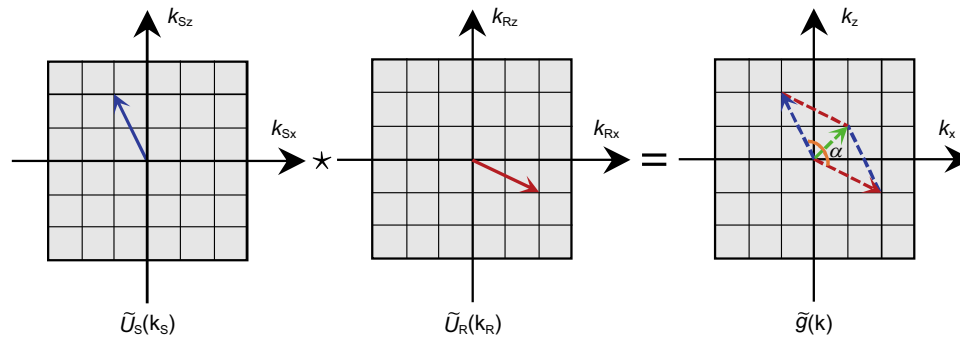
Fig. 9 Sketch of the wavenumber mapping in the convolution process to form the gradient in the plane-wave domain

a plane-wave with a particular wavelength and a dipping angle; therefore, the angle between a pair of plane waves can be computed easily using the equation,

$$\cos \alpha = \frac{\mathbf{k}_S \cdot \mathbf{k}_R}{|\mathbf{k}_S||\mathbf{k}_R|}. \tag{18}$$

The filtered gradient then can be expressed as

$$\hat{g}(\mathbf{k}) = -\frac{2}{v^3} \sum_t \sum_i \sum_j f(\alpha) \tilde{U}_S(\mathbf{k}_{Si}, t) \tilde{U}_R(\mathbf{k}_{Rj}, t), \tag{19}$$

where $f(\alpha)$ is the filtering function, the choice of which can be found in Yao et al. (2019a).

As can be seen from the derivation above, scattering angle filtering in the plane-wave domain does not assume an isotropic or homogenous media; therefore, it is an effective algorithm. However, the convolution process has large computation cost.

The third method of scattering angle filtering is non-stationary smoothing proposed by Yao et al. (2019a). Equation 13 can be rewritten as

$$|\mathbf{k}| = \frac{2\omega}{v} \cos \frac{\alpha}{2}. \tag{20}$$

Equation 20 implies that for a fixed velocity and frequency, the wavenumber of the gradient, i.e., the inverse of wavelength, equals the cosine of half the scattering angle; in the other words, the wavenumber depends on the scattering angle monotonically. As illustrated in Fig. 10, after determined the maximum scattering angle of the migration component and the minimum scattering angle of the tomography component, they are distributed in different rings. Thus, scattering angle filtering can be achieved using non-stationary smoothing. If the minimum scattering angle α_{min} is fixed, the maximum wavenumber of the tomography component can be computed using the equation,

$$k_{\text{tomo\_max}} = \frac{2\omega}{v} \cos \frac{\alpha_{min}}{2}. \tag{21}$$

Thereby, we can design a smoother to filter out the wavenumber larger than $k_{\text{tomo\_max}}$. A simple choice of the smoother is to design a Gaussian function, the standard deviation of which is

$$\sigma = \sqrt{\frac{-\log(\beta)}{2\pi^2 k_{\text{tomo\_max}}^2}}, \tag{22}$$

where β is the preset value of the Gaussian function at $k_{\text{tomo\_max}}$ in the Fourier domain, e.g., 0.3. Then, the tomography component at one element of the gradient for this frequency can be computed by multiplying the Gaussian function with the gradient and then stacking the results. In the multiplication, the center of the Gaussian function is coincident with this element. Yao et al. (2019a) pointed out that this method is very suitable for frequency domain FWI. For time-domain FWI, we can divide the seismic data into several frequency bands, each of which uses one reference frequency to compute $k_{\text{tomo\_max}}$; then, this band is smoothed with this reference frequency.

Compared to the scattering angle filtering in the plane-wave domain, this method has much smaller computation cost. However, similar to the method based on the time-lag gather, this method assumes an isotropic model. In addition, as Gaussian functions have a relative wide transition zone, partial energy of the tomography component can be removed unnecessarily.

3.1.3 Born modeling

Another widely used method for separating the migration and tomography components is Born modeling (Xu et al. 2012; Wu and Alkhalifah 2015; Sun et al. 2017; Yao and Wu 2017). As can be seen from Fig. 6, the zero-lag cross-correlation of the incident waves of the source wavefield and the reflected waves of the residual wavefield forms one
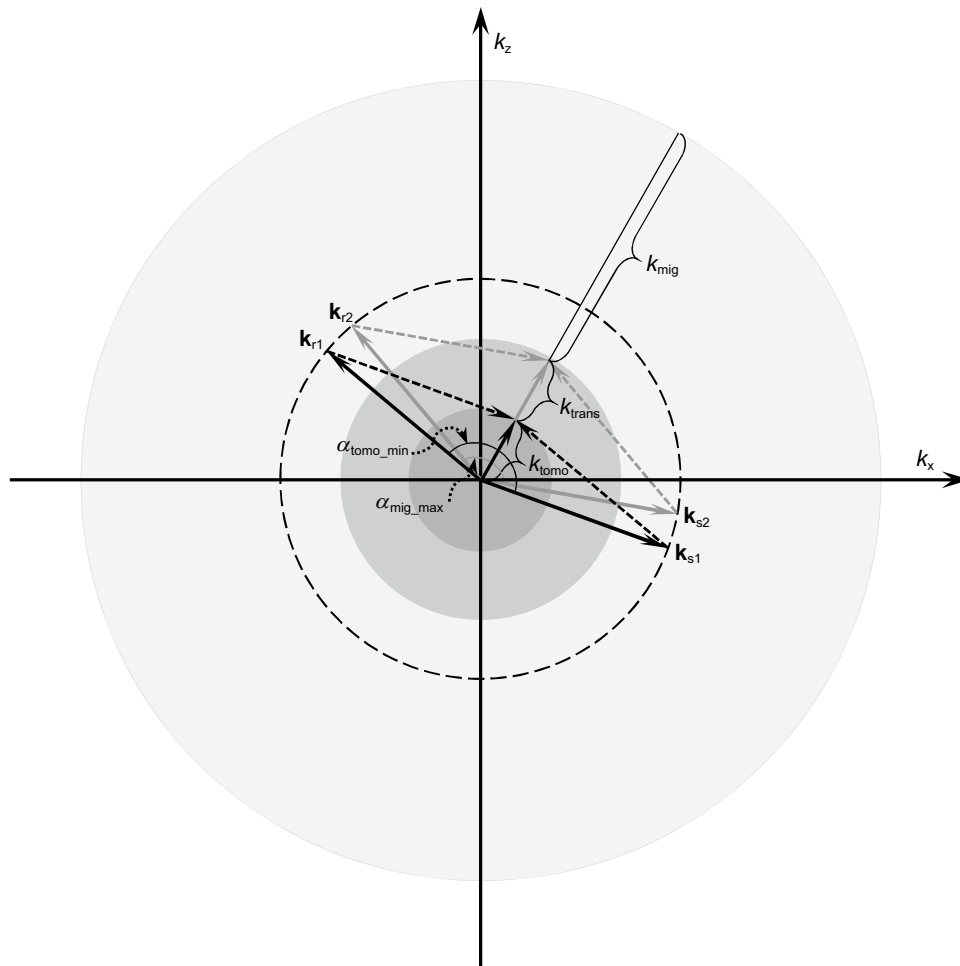
Fig. 10 Schematic diagram for the relationship between opening angles and wavenumber distribution. α_{tomo_min} indicates the minimum opening angle for the tomographic component, while α_{mig_max} indicates the maximum opening angle for the migration component. $\mathbf{k}_S$ and $\mathbf{k}_R$ are the wavenumber of the source and residual wavefields, respectively. They sit on the dashed circle. The tomographic component sits inside the inner dark-gray-filled circle while the migration component locates inside the outer light-gray ring. The transition zone is the middle-gray ring between them (Yao et al. 2019a)

branch of the tomography component; the other branch is generated by the reflected waves of the source wavefield and the incident waves of the residual wavefield. Thus, if the reflected waves are separated from the incident waves, then the tomography and migration components can be isolated. Born modeling is a simple and effective method to separate reflected waves from incident waves. Born modeling for the isotropic wave equation can be expressed as

$$\begin{cases} \left[\frac{1}{v_b^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right] u_b = s \\ \left[\frac{1}{v_b^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right] u_\delta = -m_\delta \frac{\partial^2}{\partial t^2} u_b \end{cases}, \tag{23}$$

where u_b and u_δ represent the incident and scattering wavefields, v_b and m_δ denote background velocity and model

perturbation, respectively. As illustrated in Fig. 6, the forward scattering waves of the source and residual wavefields travel in different directions; therefore, they have no contribution to the gradient. However, the backward scattering waves, i.e., mainly the reflected waves, contribute to the tomography component by cross-correlating with the incident waves of the other wavefield.

Compared to other methods, the Born modeling method can be applied to anisotropic media and arbitrarily complex media; therefore, it becomes the most widely used method for the separation. However, it cannot simulate higher-order scattering waves.

3.2 Objective functions of reflection-waveform inversion

The objective functions of reflection-waveform inversion can be divided into two main categories: the objective functions based on waveform and the objective functions based on travel time. The objective functions based on waveform can be expressed as

$$\phi = \frac{1}{2}\left\|Du_\delta\left(v_b, m_\delta\right) - d\right\|_2^2, \tag{24}$$

where d represents the observed reflection data, u_δ denotes the predicted reflection wavefields, D is the receiver operator that is used for extracting the predicted data at the receiver position from u_δ, v_b represents background velocity, m_δ is the reflector that generates reflections. Once we have an objective function, e.g., Eq. 24, combining it with a wave equation, e.g., Eqs. 2 or 23, can form an inversion system. Currently, restricted by the huge computation cost, FWI is based on local-gradient methods, e.g., steepest-descent methods and conjugate-gradient methods. As a result, computing the gradient is a key step. The formula of the gradient can be derived using the adjoint-state method (Plessix 2006) or the matrix algebra (Pratt 1999).

Currently, RWI is achieved mainly by alternatively updating v_b and m_δ. Its flowchart is shown in Fig. 11: firstly, setting reflectors m_δ as zero; secondly, building temporary reflectors m_δ with 3–5 iterations; thirdly, updating background velocity v_b with a few iterations; the three steps form a big loop. In order to avoid the weakness of RWI illustrated in Fig. 3, one commonly used method is to build reflectors m_δ with short-offset data. As can be seen from Fig. 12, the predicted data by the reflectors m_δ can fit the observed data in these short offsets while the
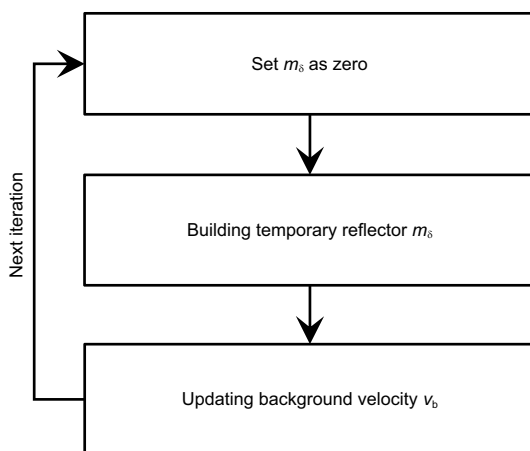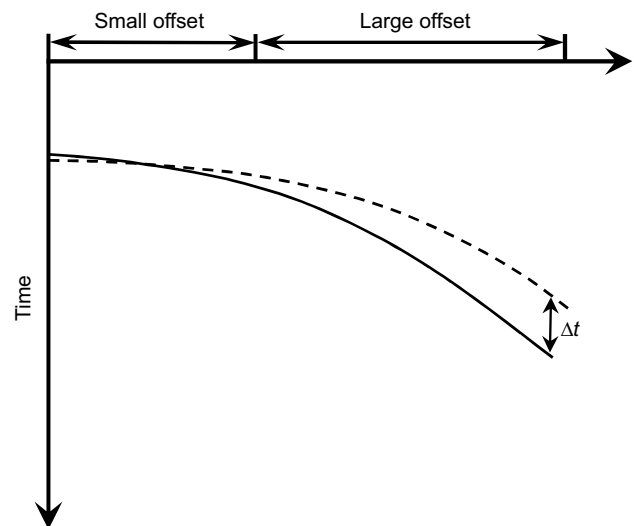

Fig. 12 Data offset selection for reflection-waveform inversion, the small offset data are used for building the temporal reflector m_δ, while the large-offset data are applied for updating the background velocity v_b

large-offset residuals contain the information for background velocity update.

The main objective of RWI is for fixing background velocity. Since the background velocity affects the travel time of seismic events much more than amplitudes, the travel-time information in the seismic data is more sensitive to the error of background velocity than amplitudes. Consequently, RWI can be achieved by using the objective functions based on travel time:

$$\phi = \frac{1}{2}\|\Delta\tau\|_2^2, \tag{25}$$

where $\Delta\tau$ represents the travel-time differences between the reflection events in the observed and predicted data. For a shot gather containing multiple reflection events, cross-correlation with a sliding window can be used to compute the local travel-time difference, $\Delta\tau$ (Hale 2009). Hale (2013) pointed out that dynamic time/image warping (DTW/DIW) is more stable than cross-correlation for multiple events, especially in case of fast travel-time variation. When computing the gradient of the objective function in Eq. 25 with respect to background velocity, there are two choices: the first is based on connective function that gives the relationship between the travel-time difference $\Delta\tau$ and predicted data (Luo and Schuster 1991; Chi et al. 2015); the other is to combine instantaneous phase and phase-only FWI, in which DTW/DIW can be used to compute travel-time difference $\Delta\tau$, $\Delta\tau$ multiplied by frequency is a stable means to calculate the instantaneous phase (Jiao et al. 2015; Sun et al. 2017).

Compared to the objective functions based on waveform, the objective functions based on travel time can

Fig. 11 The flowchart of reflection-waveform inversion in an alternating scheme

mitigate the influence of amplitudes on inversion; therefore, the influence of the physics that cannot be predicted by acoustic wave equations, e.g., s-waves and attenuation, can be mitigated. However, if the initial model is relatively accurate, the objective functions based on waveform can give a more accurate result than the objective functions based on travel time, as the latter gives up the dynamic information in the data (Chi et al. 2015).

In addition, since the objective functions based on travel time are convex, they are immune to cycle skipping, for instance, Fig. 7d of Chi et al. (2016). However, this good feature is true only if the travel time is computed correctly. In case of large errors in the initial model, wrong pairs of events can be selected from the predicted and observed data to compute the travel-time difference.

The second step of RWI builds temporary reflectors m_δ at a wrong depth to fit the data in short offsets; therefore, the likelihood of cycle skipping in large offsets using waveform-based objective functions is reduced significantly compared to conventional FWI. An example is illustrated in Fig. 13, which shows the width of the convex region in the waveform-based objective function is several times wider than that of conventional FWI. Wang et al. (2018) pointed out the likelihood can be further reduced by splitting the long-offset seismic data gather into two parts and then inverting them separately.

Except the two types of objective functions of RWI based on pure reflection data, Zhou et al. (2015) proposed a joint inversion of reflection and transmission data:

$$\phi = \frac{1}{2}\left\|W^d\left(d_{obs}^{div} - d_{pred}^{div}\right)\right\|_2^2 + \frac{1}{2}\left\|W^r\left(d_{obs}^{refl} - d_{pred}^{refl}\right)\right\|_2^2, \quad (26)$$

where the superscripts div and refl represent the transmission and reflection data, the subscripts obs and pred denote observed and predicted data, W^d and W^r are the weighting of transmission and reflection data. The objection function is a function of reflectors, i.e., the impedance perturbation in Zhou et al. (2015), and background velocity. Zhou et al. (2015) updated reflectors and background velocity alternatively, where reflectors are built of reflections while the background velocity is updated with both reflection and transmission data. Different from Yao and Wu (2017), Zhou et al. (2015) did not reset the reflectors to zero in each inversion loop.

Similar to Zhou et al. (2015), Wu and Alkhalifah (2015) formed the objective function with both transmission and reflection data:

$$\phi = \frac{1}{2}\left\|d_{pred}^{div} + d_{pred}^{refl} - d_{obs}\right\|_2^2. \quad (27)$$

The inversion system is constrained by a Born wave equation, such as Eq. 23. Wu and Alkhalifah (2015) did not update reflectors and background velocity alternatively but simultaneously, i.e., the update direction is formed by weighted velocity and reflector gradients. In the early stage of inversion, the weighting to the velocity gradient is small; therefore, the velocity update is smooth, i.e., background velocity update. The weighting to the velocity gradient is
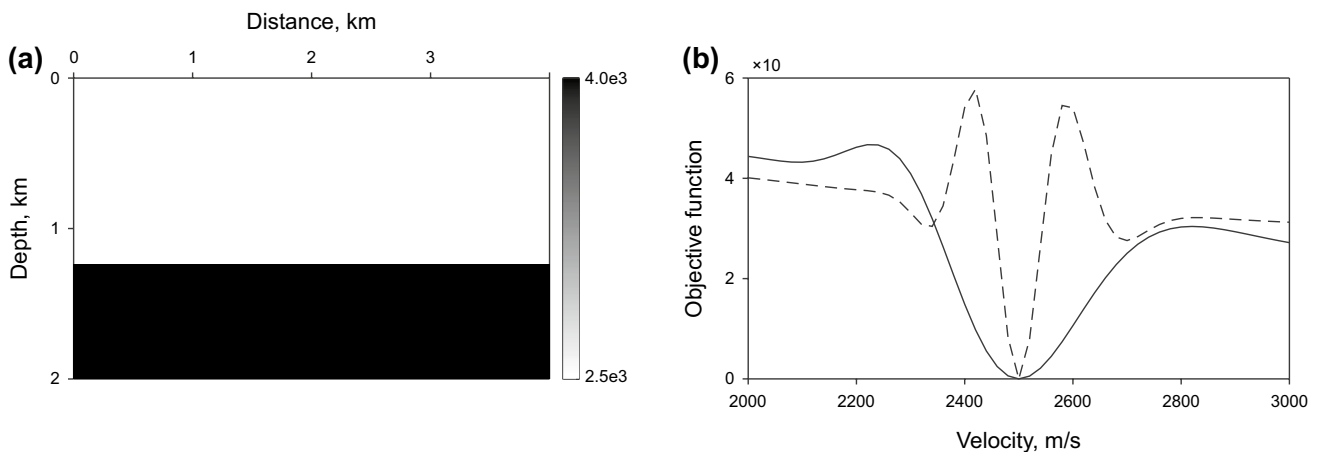

Fig. 13 The waveform-based objective function of RWI. **a** The two-layer velocity model used for the test. The 10 Hz Ricker wavelet used as the source signature is located in the middle of the surface of the model; the receiver array is evenly distributed on the surface of the model. **b** The waveform-based objective function calculated by using Eq. 22. The dashed curve represents the functional value calculated by changing the velocity of the first layer but fixing the depth of the interface. The solid curve indicates the functional value by changing both the velocity of the first layer and the depth of the interface. The depth is determined by keeping the normal incident two-way travel time from the source to the interface as that of the true model. The dashed curve approximately represents the situation of conventional FWI, while the solid curve illustrates the situation of RWI, the flowchart of which is shown in Fig. 10. The predicted data are scaled by a factor so that they have the same L2 norm as the observed data for the central trace. In this example, the convex zone around the global minimum for the solid curve is 3.75 times as wide as that of the dashed curve

then increased gradually as the inversion progress, which means inversion adds more and more high-frequency components into the velocity model.

3.3 Constraints in reflection-waveform inversion

Generally, RWI decomposes the Earth model into a background velocity model v_b and a reflector model m_δ; therefore, compared to conventional single-parameter FWI, the number of unknown parameters in the system of RWI is doubled. In addition, conventional FWI utilized all seismic records including refractions, but RWI uses much fewer data. The two factors lead to that RWI is more ill-posed than conventional single-parameter FWI. One manifestation of the ill-posedness is that an inverted background velocity model is full of artifacts but still fits the observed data, for example, Fig. 7a of Yao et al. (2019c).

In order to mitigate the ill-posedness, the inversion system has to incorporate extra information about the velocity model. Since the inversion uses all information from seismic data, the extra information only comes from nonseismic sources, for example, geology, well-logs or other geophysical methods. So far, the research on adding extra information has been studied inadequately. Recently, Yao et al. (2019c) introduced structural-oriented smoothing to constrain RWI. The constraint is a piece of geologic information: the rock properties vary slowly in one stratum but can be fast in different strata. This geological information can be realized by smoothing along strata. This smoothing can be achieved by using anisotropic smoothing. There are two commonly used methods for the anisotropic smoothing: the first is solving an anisotropic diffusion equation (Fehmers and Höcker 2003); the second is solving an anisotropic smoothing system (Hale 2011). In inversion, the anisotropic smoothing can be applied to the gradient of background velocity acting as preconditioning (Lewis et al. 2014; Lee and Pyun 2018); or the smoothing can be used on the background velocity. The latter is called shaping, which directly manipulates the model to possess a particular feature (Fomel 2007). The shaping process can be added as the last step in a loop of RWI (Yao et al. 2019c).

Figure 14 shows an example using the constraint of structural-oriented smoothing. The orientation of strata is picked up from the reflector model m_δ automatically, e.g., Fig. 14c. Solving the anisotropic diffusion equation achieves the smoothness in a stratum but keeps the variation between strata. As can be seen from Fig. 14e, the ill-posedness of the inversion system leads to many artificial anomalies in the inversion result. These anomalies break the geological empirical law, i.e., slow variation of rock properties in a stratum. As a result, structural smoothing can suppress the artificial anomalies and force the velocity of the inversion result varies slowly along the structure.

The constrained result follows the background trend of the true model (Fig. 14f).

4 Conclusions and future perspective

So far, the concept and theory of RWI have become mature. There are many successful applications on marine seismic data (Sun et al. 2017; Wang et al. 2018). The preprocesses of seismic data include de-multiples and deghost. These ensure the input data satisfy Born approximation and mitigate the nonlinearity of the inversion system caused by multiples. The separation of tomography and migration components is achieved using Born modeling. Both waveform-based (Wang et al. 2018) and travel time-based (Sun et al. 2017) objective functions succeeded to recover background velocity. Similar to conventional FWI, the industry is trying to apply RWI on land seismic data. However, there exist many difficulties including: firstly, low signal-to-noise ratio in land seismic data, it is often hard to observe hyperbolic-shape events in raw data; secondly, near-surface complexities, including severe heterogeneity, undulate topography and weather layer, which cause modeling kernels to simulate wavefields incorrectly; thirdly, current FWI commonly uses acoustic wave equations as the modeling kernel, and thus there exists large waveform discrepancy between the predicted data and the observed land data that contain strong elastic effects.

Based on the analysis above, except separation between the migration and tomography components, objective functions and inversion constraints described in previous sections, RWI includes other study topics: firstly, seismic data preprocesses, which aims to remove all energy unable to be modeled by the modeling kernel in the data, as well as the residuals not generated by the parameters of inversion; secondly, multi-parameter inversion, including P-wave velocity, S-wave velocity, anisotropic parameters and attenuation factor Q. RWI based on elastic wave equations has been explored preliminarily (e.g., Guo and Alkhalifah 2017; Li et al. 2019; Ren et al. 2019).

In a more general scope, for the past thirty years, especially the last decade, the research on FWI has made significant progress on the theory of FWI and many successful applications have been achieved in the industry. However, FWI has not reached the expectation of the industry. The reasons may include:

Gigantic computation cost FWI is achieved using local or global search methods to find the optimal solution so that the inversion solves wave equations numerically in hundreds of times. Mitigating this problem, except the advance of computer hardware, optimization on inversion algorithms, numerical wavefield modeling and computer codes are important research directions. Among them,
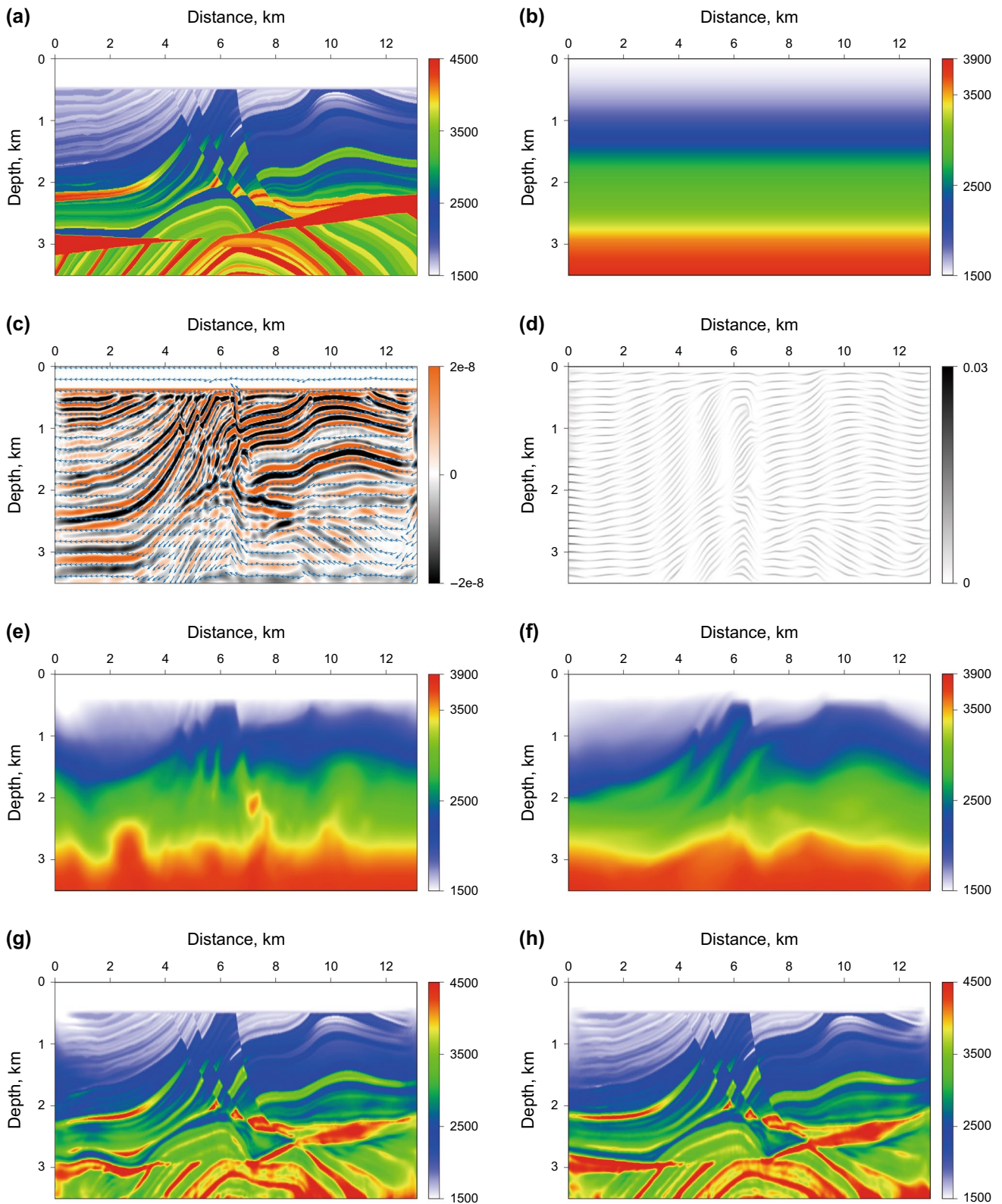
Fig. 14 Application of reflection-waveform inversion to the Marmousi Model. **a** The true velocity of Marmousi model. **b** The initial model of background velocity for inversion. The recorded data set is comprised of 128 shots in a uniform spacing of 100 m generated by using a 10 Hz Ricker wavelet. **c** The reflector model generated in the first loop of inversion. The blue arrows indicate the orientation of strata. **d** The response of spikes from structural smoothing by solving an anisotropic diffusion equation in a time of 400 s. **e** The inversion result without structural smoothing. **f** The inversion result with structural smoothing. **g** and **h** shows the results of conventional FWI starting from **e** and **f**, respectively (Yao et al. 2019c)

improving the inversion and modeling algorithms are the key directions because computer code optimization, for example, load balance between processes, CPU vectorization, cache reuse and GPU acceleration (e.g., Wang et al. 2019), can only speed up the inversion in tens of times or less generally.

Insufficient information restricted by the technology and space of acquisition, the acquired seismic data are incomplete in spatial coverage and frequencies. Petroleum exploration and exploitation do not only require P-wave velocity but also other seismic properties, such as density and S-wave velocity and anisotropic parameters. Currently, in successful applications, FWI usually produces only P-wave velocity, in which the density is constrained by using Gardener's relation, and the anisotropic parameters are preset. If FWI is used for multi-parameter inversion, then it turns into a severe under-determined problem. Consequently, the inversion results are still far different from the true models but fit the recorded data nicely. This is caused by insufficient information in the seismic data. Mathematic theory pointed out that the solution for an under-determined problem is to introduce more constraint information, which includes mathematic regularization and extra information. Common mathematic regularization includes Tikhonov, total variation and so on. The extra information includes seismic and non-seismic information. "two wide and one high" (i.e., broad frequency band, wide azimuth and high shot and receiver density) acquisition and multi-component acquisition can provide more seismic information for inversion. Theoretic tests show that the ill-posedness of multi-parameter inversion cannot be fully removed with a better acquisition. Extra non-seismic information, including well-logs, geologic laws, is an important information supplement. Thus, merging the multi-source information into FWI is an important research direction.

Incomplete theory Currently, the theory of FWI is still under the framework of Tarantola's theory proposed in 1984. It includes localized gradient methods and Born approximation. FWI is a nonlinear problem; therefore, localized gradient methods have limitations. The Born approximation is used to compute the gradient. Many geologic bodies missed in the starting model cannot be represented using Born approximation, because Born approximation is valid only for weak scattering caused by small volume and magnitude. Global inversion is a way to solve this problem because it does not need to compute a gradient. But its gigantic computation cost forbids its use. Except the incomplete inversion theory, the forward modeling theory is also incomplete. The existing wave equations cannot describe the propagation of wavefields in the real world accurately. Thus, exploring the inversion and modeling theories is also an important research direction.

Unsophisticated acquisition technology Professor Teng Jiwen, who is a fellow of Chinese Academy of Sciences, pointed out that geophysics is a science of observation essentially; therefore, reliable information, or incompleteness and erroneous of information cannot be remedied by using any mathematic tricks or image processes. In the past several years, the significant advances of FWI in real data applications are mainly because of the progress in acquisition technology. In the applications on the salt models of the Gulf of Mexico, OBN acquisition and marine controlled source, i.e., Wolfspar, ensure the maximum offset exceeds 25 km and the valid signal lower than 2 Hz. However, there is a lack of breakthroughs in land acquisition, for instance low signal-to-noise ratio in the data and inconsistency of different shots. The quality of land data may be the main reason for restricting successful applications of FWI on land data. Consequently, developing acquisition technology is a key research direction for FWI as well as for other seismic methods.

References

Aki K, Richards PG. Quantitative seismology: theory and methods. San Francisco: W. H. Freeman and Co; 1981. https://doi.org/10.1002/gj.3350160110.

Alkhalifah T. Scattering-angle based filtering of the waveform inversion gradients. Geophys J Int. 2014;200(1):363–73. https://doi.org/10.1093/gji/ggu379.

Baeten G, de Maag JW, Plessix RE, Klaassen R, Qureshi T, Kleemeyer M, et al. The use of low frequencies in a full-waveform inversion and impedance inversion land seismic case study. Geophys Prospect. 2013;61(4):701–11. https://doi.org/10.1111/1365-2478.12010.

Baysal E, Kosloff DD, Sherwood JWC. Reverse time migration. Geophysics. 1983;48(11):1514–24. https://doi.org/10.1190/1.1441434.

Bednar JB. A brief history of seismic migration. Geophysics. 2005;70(3):3–20. https://doi.org/10.1190/1.1926579.

Biondi B, Almomin A. Tomographic full-waveform inversion (TFWI) by combining FWI and wave-equation migration velocity analysis. Lead Edge. 2013;32(9):1074–80. https://doi.org/10.1190/tle32091074.1.

Bunks C, Saleck F, Zaleski S, Chavent G. Multiscale seismic waveform inversion. Geophysics. 1995;60(5):1457–73. https://doi.org/10.1190/1.1443880.

Castagna JP, Swan HW, Foster DJ. Framework for AVO gradient and intercept interpretation. Geophysics. 1998;63(3):948–56. https://doi.org/10.1190/1.1444406.

Chen GX, Wu RS, Chen SC. Reflection multi-scale envelope inversion. Geophys Prospect. 2018a;66(7):1258–71. https://doi.org/10.1111/1365-2478.12624.

Chen GX, Wu RS, Wang YQ, Chen SC. Multi-scale signed envelope inversion. J Appl Geophys. 2018b;153:113–26. https://doi.org/10.1016/j.jappgeo.2018.04.008.

Chen WY, Jial XF, Hu CH. Time-domain multiscale FWI based on a time-shift method and low-frequency data reconstruction. In: SEG Technical Program Expanded Abstracts; 2019. p. 1570–4. https://doi.org/10.1190/segam2019-3215657.1.

Cheng X, Jiao K, Sun D, Xu Z, Vigh D, El-Emam A. High-resolution Radon preconditioning for full-waveform inversion of land seismic data. Interpretation. 2017;5(4):SR23–33. https://doi.org/10.1190/int-2017-0020.1.

Chi B, Dong L, Liu Y. Correlation-based reflection full-waveform inversion. Geophysics. 2015;80(4):R189–202. https://doi.org/10.1190/geo2014-0345.1.

Chi B, Gao K, Huang L. Least-squares reverse time migration guided full-waveform inversion. In: SEG Technical Program Expanded Abstracts; 2017. p. 1471–5. https://doi.org/10.1190/segam2017-17775494.1.

Claerbout JF. Toward a unified theory of reflector mapping. Geophysics. 1971;36(3):467–81. https://doi.org/10.1190/1.1440185.

Claerbout JF, Doherty SM. Downward continuation of moveout-corrected seismograms. Geophysics. 1972;37(5):741–68. https://doi.org/10.1190/1.1440298.

Claerbout J. Imaging the Earth's interior. Oxford: Blackwell; 1985.

da Silva NV, Yao G. Wavefield reconstruction inversion with a multiplicative cost function. Inverse Probl. 2017;34(1):015004. https://doi.org/10.1088/1361-6420/aa9830.

da Silva NV, Yao G, Calderón Agudo Ò, Stronge G, Warne M. 3D elastic semi-global waveform inversion—estimation of VP to VS ratio. In: SEG Technical Program Expanded Abstracts; 2019. p. 1440–4. https://doi.org/10.1190/segam2019-3216359.1.

Dellinger J, Ross A, Meaux D, Brenders A, Gesoff G, Etgen J, et al. An FWI-friendly ultralow-frequency marine seismic source. In: SEG Technical Program Expanded Abstracts; 2016. p. 4891–5. https://doi.org/10.1190/segam2016-13762702.1.

Fehmers G, Höcker C. Fast structural interpretation with structure-oriented filtering. Geophysics. 2003;68(4):1286–93. https://doi.org/10.1190/1.1598121.

Fei TW, Luo Y, Yang J, Liu H, Qin F. Removing false images in reverse time migration: The concept of de-primary. Geophysics. 2015;80(6):S237–44. https://doi.org/10.1190/geo2015-0289.1.

Fomel S. Shaping regularization in geophysical-estimation problems. Geophysics. 2007;72(2):R29–36. https://doi.org/10.1190/1.2433716.

Gazdag J. Wave equation migration with the phase-shift method. Geophysics. 1978;43(7):1342–51. https://doi.org/10.1190/1.1440899.

Gazdag J, Sguazzero P. Migration of seismic data by phase shift plus interpolation. Geophysics. 1984;49(2):124–31. https://doi.org/10.1190/1.1441643.

Guasch L, Warner M, Ravaut C. Adaptive waveform inversion. Practice. 2019;84(3):R447–61. https://doi.org/10.1190/geo2018-0377.1.

Guo Q, Alkhalifah T. Elastic reflection-based waveform inversion with a nonlinear approach. Geophysics. 2017;82(6):R309–21. https://doi.org/10.1190/geo2016-0407.1.

Hale D. A method for estimating apparent displacement vectors from time-lapse seismic images. Geophysics. 2009;74(5):V99–107. https://doi.org/10.1190/1.3184015.

Hale D. Structure-oriented bilateral filtering of seismic images. In: SEG Technical Program Expanded Abstracts; 2011. p. 3596–600. https://doi.org/10.1190/1.3627947.

Hale D. Dynamic warping of seismic images. Geophysics. 2013;78(2):S105–15. https://doi.org/10.1190/geo2012-0327.1.

Hill NR. Gaussian beam migration. Geophysics. 1990;55(11):1416–28. https://doi.org/10.1190/1.1442788.

Hu LZ, McMechan GA. Wave-field transformations of vertical seismic profiles. Geophysics. 1987;52(3):307–21. https://doi.org/10.1190/1.1442305.

Hu W. FWI without low frequency data—beat tone inversion. In: SEG Technical Program Expanded Abstracts; 2014. p. 1116–20. https://doi.org/10.1190/segam2014-0978.1.

Irabor K, Warner M. Reflection FWI. In: SEG Technical Program Expanded Abstracts; 2016. p. 1136–40. https://doi.org/10.1190/segam2016-13944219.1.

Jiao K, Sun D, Cheng X, Vigh D. Adjustive full waveform inversion. In: SEG Technical Program Expanded Abstracts; 2015. p. 1091–5. https://doi.org/10.1190/segam2015-5901541.1.

Kazei V, Alkhalifah T. Scattering radiation pattern atlas: what anisotropic elastic properties can body waves resolve. J Geophys Res Sol Earth. 2019;124(3):2781–811. https://doi.org/10.1029/2018jb016687.

Khalil A, Sun J, Zhang Y, Poole G. RTM noise attenuation and image enhancement using time-shift gathers. In: SEG Technical Program Expanded Abstracts; 2013. p. 3789–93. https://doi.org/10.1190/segam2013-0600.1.

Lailly P. The seismic inverse problem as a sequence of before stack migration. In: Conference on inverse scattering, theory and applications. SIAM; 1983.

Lazaratos S, Chikichev I, Wang K. Improving the convergence rate of full wavefield inversion using spectral shaping. In: SEG Technical Program Expanded Abstracts; 2011. p. 2428–32. https://doi.org/10.1190/1.3627696.

Lee D, Pyun S. Adaptive preconditioning of full-waveform inversion based on structure-oriented smoothing filter. In: SEG Technical Program Expanded Abstracts; 2018. p. 1048–52. https://doi.org/10.1190/segam2018-2996620.1.

Lewis W, Amazonas D, Vigh D, Coates R. Geologically constrained full-waveform inversion using an anisotropic diffusion based regularization scheme: application to a 3D offshore Brazil dataset. In: SEG Technical Program Expanded Abstracts; 2014. p. 1083–8. https://doi.org/10.1190/segam2014-1174.1.

Li Y, Guo Q, Li Z, Alkhalifah T. Elastic reflection waveform inversion with variable density. Geophysics. 2019;84(4):R553–67. https://doi.org/10.1190/geo2017-0722.1.

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

Liu F, Zhang G, Morton SA, Leveille JP. An effective imaging condition for reverse-time migration using wavefield decomposition. Geophysics. 2011;76(1):S29–39. https://doi.org/10.1190/1.3533914.

Luo S, Sava P. A deconvolution-based objective function for wave-equation inversion. In: SEG Technical Program Expanded Abstracts; 2011. p. 2788–92. https://doi.org/10.1190/1.3627773.

Luo Y, Schuster GT. Wave-equation traveltime inversion. Geophysics. 1991;56(5):645–53. https://doi.org/10.1190/1.1443081.

Ma Z, Ji S. All dip finite-difference migration with scalar wave equation. J Chin Geophys. 1988;31(4):678–86 **(in Chinese)**.

Mei J, Tong Q. A practical acoustic full waveform inversion workflow applied to a 3D land dynamite survey. In: SEG Technical Program Expanded Abstracts; 2015. p. 1220–4; https://doi.org/10.1190/segam2015-5850377.1.

Métivier L, Brossier R, Mérigot Q, Oudet E, Virieux J. Measuring the misfit between seismograms using an optimal transport distance: application to full waveform inversion. Geophys J Int. 2016;205(1):345–77. https://doi.org/10.1093/gji/ggw014.

Mora P. Nonlinear two-dimensional elastic inversion of multioffset seismic data. Geophysics. 1987;52(9):1211–28. https://doi.org/10.1190/1.1442384.

Mora P. Inversion = migration + tomography. Geophysics. 1989;54(12):1575–86. https://doi.org/10.1190/1.1442625.

Mueller G. Rheological properties and velocity dispersion of a medium with power-law dependence of Q on frequency. J Geophys Zeitschrift für Geophysik. 1983;54(1):20–9.

Pan W, Innanen KA, Geng Y, Li J. Interparameter trade-off quantification for isotropic-elastic full-waveform inversion with various model parameterizations. Geophysics. 2019;84(2):R185–206. https://doi.org/10.1190/geo2017-0832.1.

Plessix RE. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. Geophys J Int. 2006;167(2):495–503. https://doi.org/10.1111/j.1365-246X.2006.02978.x.

Plessix RÉ, Li Y. Waveform acoustic impedance inversion with spectral shaping. Geophys J Int. 2013;195(1):301–14. https://doi.org/10.1093/gji/ggt233.

Popovici AM. Prestack migration by split-step DSR. Geophysics. 1996;61(5):1412–6. https://doi.org/10.1190/1.1444065.

Pratt R. Seismic waveform inversion in the frequency domain, Part 1: theory and verification in a physical scale model. Geophysics. 1999;64(3):888–901. https://doi.org/10.1190/1.1444597.

Ren Z, Li Z, Gu B. Elastic reflection waveform inversion based on the decomposition of sensitivity kernels. Geophysics. 2019;84(2):R235–50. https://doi.org/10.1190/geo2018-0220.1.

Ristow D, Ruhl T. Fourier finite-difference migration. Geophysics. 1994;59(12):1882–93. https://doi.org/10.1190/1.1443575.

Routh P, Neelamani R, Lu R, Lazaratos S, Braaksma H, Hughes S, et al. Impact of high-resolution FWI in the Western Black Sea: revealing overburden and reservoir complexity. Lead Edge. 2016;36(1):60–6. https://doi.org/10.1190/tle36010060.1.

Schneider WA. Integral formulation for migration in two and three dimensions. Geophysics. 1978;43(1):49–76. https://doi.org/10.1190/1.1440828.

Sedova A, Royle G, Allemand T, Lambaré G, Hermant O. High-frequency acoustic land full-waveform inversion: a case study from the Sultanate of Oman. First Break. 2019;37(1):75–81.

Shen X, Ahmed I, Brenders A, Dellinger J, Etgen J, Michell S. Salt model building at Atlantis with full-waveform inversion. In: SEG Technical Program Expanded Abstracts; 2017. p. 1507–11. https://doi.org/10.1190/segam2017-17738630.1.

Shuey R. A simplification of the Zoeppritz equations. Geophysics. 1985;50(4):609–14. https://doi.org/10.1190/1.1441936.

Sirgue L, Barkved OI, van Gestel JP, Askim OJ, Kommedal JH. 3D waveform inversion on Valhall wide-azimuth OBC. In: 71st EAGE conference; 2009. p. 1–4. https://doi.org/10.3997/2214-4609.201400395.

Sirgue L, Etgen JT, Albertin U. 3D frequency domain waveform inversion using time domain finite difference methods. In: 70th

EAGE conference; 2008. p. 1–4. https://doi.org/10.3997/2214-4609.20147683.

Štekl I, Pratt RG. Accurate viscoelastic modeling by frequency-domain finite differences using rotated operators. Geophysics. 1998;63(5):1779–94. https://doi.org/10.1190/1.1444472.

Stoffa PL, Fokkema JT, Freire RMdL, Kessinger WP. Split-step Fourier migration. Geophysics. 1990;55(4):410–21. https://doi.org/10.1190/1.1442850.

Stolt RH. Migration by Fourier transform. Geophysics. 1978;43(1):23–48. https://doi.org/10.1190/1.1440826.

Sun D, Jiao K, Cheng X, Xu Z, Zhang L, Vigh D. Born modeling based adjustive reflection full-waveform inversion. In: SEG Technical Program Expanded Abstracts; 2017. p. 1460–5. https://doi.org/10.1190/segam2017-17664306.1.

Tarantola A. Inversion of seismic reflection data in the acoustic approximation. Geophysics. 1984;49(8):1259–66. https://doi.org/10.1190/1.1441754.

Tarantola A. A strategy for nonlinear elastic inversion of seismic reflection data. Geophysics. 1986;51(10):1893–903. https://doi.org/10.1190/1.1442046.

van Leeuwen T, Herrmann FJ. Mitigating local minima in full-waveform inversion by expanding the search space. Geophys J Int. 2013;195(1):661–667. https://doi.org/10.1093/gji/ggt258.

Van Leeuwen T, Mulder WA. A correlation-based misfit criterion for wave-equation traveltime tomography. Geophys J Int. 2010;182(3):1383–94. https://doi.org/10.1111/j.1365-246X.2010.04681.x.

Virieux J, Operto S. An overview of full-waveform inversion in exploration geophysics. Geophysics. 2009;74(6):WCC1–26. https://doi.org/10.1190/1.3238367.

Wang F, Chauris H, Donno D, Calandra H. Taking advantage of wave field decomposition in full waveform inversion. In: 75th EAGE conference; 2013. p. 1–4. https://doi.org/10.3997/2214-4609.20130415.

Wang M, Xie Y, Xu WQ, Loh FC, Xin K, Chuah BL, et al. Dynamic-warping full-waveform inversion to overcome cycle skipping. In: SEG Technical Program Expanded Abstracts; 2016. p. 1273–7. https://doi.org/10.1190/segam2016-13855951.1.

Wang P, Zhang Z, Wei Z, Huang R. A demigration-based reflection full-waveform inversion workflow. In: SEG Technical Program Expanded Abstracts; 2018. p. 1138–42. https://doi.org/10.1190/segam2018-2997404.1.

Wang Y. Approximations to the Zoeppritz equations and their use in AVO analysis. Geophysics. 1999;64(6):1920–7. https://doi.org/10.1190/1.1444698.

Wang YH, Rao Y. Reflection seismic waveform tomography. J Geophys Res-Sol Earth. 2009;114:B03304. https://doi.org/10.1029/2008JB005916.

Wang ZY, Huang JP, Liu DJ, Li ZC, Yong P, Yang ZJ. 3D variable-grid full-waveform inversion on GPU. Pet Sci. 2019;2019:1–14. https://doi.org/10.1007/s12182-019-00368-2.

Warner M, Guasch L. Adaptive waveform inversion: theory. Geophysics. 2016;81(6):R429–45. https://doi.org/10.1190/geo2015-0387.1.

Warner M, Nangoo T, Shah N, Umpleby A, Morgan J. Full-waveform inversion of cycle-skipped seismic data by frequency down-shifting. In: SEG Technical Program Expanded Abstracts; 2013a. p. 903–7. https://doi.org/10.1190/segam2013-1067.1.

Warner M, Ratcliffe A, Nangoo T, Morgan J, Umpleby A, Shah N, et al. Anisotropic 3D full-waveform inversion. Geophysics. 2013b;78(2):R59–80. https://doi.org/10.1190/geo2012-0338.1.

Weglein AB. A timely and necessary antidote to indirect methods and so-called P-wave FWI. Lead Edge. 2013;32(10):1192–204. https://doi.org/10.1190/tle32101192.1.

Whitmore ND. Iterative depth migration by backward time propagation. In: SEG Technical Program Expanded Abstracts; 1983. p. 382–5. https://doi.org/10.1190/1.1893867.

Wu RS, Huang LJ. Scattered field calculation in heterogeneous media using a phase-screen propagator. In: SEG Technical Program Expanded Abstracts; 1992, p. 1289–92. https://doi.org/10.1190/1.1821974.

Wu R, Luo J, Wu B. Seismic envelope inversion and modulation signal model. Geophysics. 2014;79(3):WA13–24. https://doi.org/10.1190/geo2013-0294.1.

Wu Z, Alkhalifah T. Simultaneous inversion of the background velocity and the perturbation in full-waveform inversion. Geophysics. 2015;80(6):R317–29. https://doi.org/10.1190/geo2014-0365.1.

Wu Z, Alkhalifah T. Efficient scattering-angle enrichment for a nonlinear inversion of the background and perturbations components of a velocity model. Geophys J Int. 2017;210(3):1981–92. https://doi.org/10.1093/gji/ggx283.

Xu S, Wang D, Chen F, Lambaré G, Zhang Y. Inversion on reflected seismic wave. In: SEG Technical Program Expanded Abstracts; 2012. p. 1–7. https://doi.org/10.1190/segam2012-1473.1.

Yang Y, Engquist B. Analysis of optimal transport and related misfit functions in full-waveform inversion. Geophysics. 2018;83(1):A7–12. https://doi.org/10.1190/geo2017-0264.1.

Yang Y, Engquist B, Sun J, Hamfeldt BF. Application of optimal transport and the quadratic Wasserstein metric to full-waveform inversion. Geophysics. 2018;83(1):R43–62. https://doi.org/10.1190/geo2016-0663.1.

Yao G, da Silva NV, Warner M, Kalinicheva T. Separation of migration and tomography modes of full-waveform inversion in the plane wave domain. J Geophys Res-Sol Earth. 2018a;123(2):1486–501. https://doi.org/10.1002/2017JB015207.

Yao G, da Silva NV, Wu D. Sensitivity analyses of acoustic impedance inversion with full-waveform inversion. J Geophys Eng. 2018b;15(2):461–77. https://doi.org/10.1088/1742-2140/aaa980.

Yao G, da Silva NV, Kazei V, Wu D, Yang C. Extraction of the tomography mode with nonstationary smoothing for full-waveform inversion. Geophysics. 2019a;84(4):R527–37. https://doi.org/10.1190/geo2018-0586.1.

Yao G, da Silva NV, Warner M, Wu D, Yang C. Tackling cycle skipping in full-waveform inversion with intermediate data. Geophysics. 2019b;84(3):R411–27. https://doi.org/10.1190/geo2018-0096.1.

Yao G, da Silva NV, Wu D. Reflection-waveform inversion regularized with structure-oriented smoothing shaping. Pure appl Geophys. 2019c. https://doi.org/10.1007/s00024-019-02265-6.

Yao G, Wu D. Reflection full waveform inversion. Sci China Earth Sci. 2017;60(10):1783–94. https://doi.org/10.1007/s11430-016-9091-9.

Zhang Y, Zhang H, Zhang G. A stable TTI reverse time migration and its implementation. Geophysics. 2011;76(3):WA3–11. https://doi.org/10.1190/1.3554411.

Zhou W, Brossier R, Operto S, Virieux J. Full waveform inversion of diving & reflected waves for velocity model building with impedance inversion based on scale separation. Geophys J Int. 2015;202(3):1535–54. https://doi.org/10.1093/gji/ggv228.

Zhu H, Fomel S. Building good starting models for full-waveform inversion using adaptive matching filtering misfit. Geophysics. 2016;81(5):U61–U72. https://doi.org/10.1190/geo2015-0596.1