On the role of the deconvolution imaging condition in full waveform inversion

Wenyong Pan, Gary Margrave, Kris Innanen

ABSTRACT

Full Waveform Inversion (FWI) employs full waveform information to inverse the subsurface properties through a iterative process. This method has been widely studied in recent years but still cannot be practiced effectively in industry. FWI with a steepest descent method assumes the Hessian matrix as an identity matrix and suffers from slow convergence rate. The Hessian matrix can compensate the geometrical spreading effects and then improve the convergence rate of FWI. While calculating the inverse Hessian matrix directly is thought to be unfeasible because of its extensively computational cost. Even though the researchers have developed various methods to approximate the full Hessian matrix, this problem remains to be addressed. It is known that FWI and Reverse Time Migration (RTM) share the same algorithmic structure and the gradient calculation in FWI is formally identical to a RTM image with a cross-correlation imaging condition. In this research, we found that auto-correlation of the forward modeling wavefields, namely, the source illumination is actually equivalent to the diagonal part of the pseudo-Hessian. And the gradient scaled by the auto-correlation of the forward modeling wavefields is equivalent to a RTM image based on the deconvolution imaging condition. Furthermore, deconvolution imaging condition based gradient is much more close to the reflectivity. Hence, it is possible for us to estimate the model perturbation through the traditional impedance inversion method. Combing FWI and traditional impedance inversion forms the Iterative Modelling Migration and Inversion (IMMI) method by Margrave et al.(2012). Finally, we practiced this strategy on a portion of Marmousi model and the phase encoding method was introduced to construct the gradient and diagonal pseudo-Hessian. And the iteration-dependent ray parameter setting strategy in the iterative process has also been involved to reduce the computational burden and balance the update.

INTRODUCTION

Full waveform Inversion (FWI) is a very important method to build the velocity model for high resolution seismic imaging (Tarantola, 1984; Virieux and Operto, 2009; Margrave et al., 2011) by minimizing the difference between the synthetic data and observed data. It has been widely observed that the calculation of the gradient of a least-squares FWI objective function corresponds to migrating the data residuals(Gao et al., 2012). Reverse time migration and full waveform inversion share the same algorithmic structure (Shin et al., 2001) and the gradient calculation in FWI is formally identical to constructing a Reverse Time Migration (RTM) image with a cross-correlation imaging condition. Margrave et al. (2011) states that any other wave-equation migration technology is also able to calculate an approximate gradient.

The classical imaging principle given by Claerbout (1971) states that: the reflector exists where the downgoing wavefields and upgoing wavefields coincide in time and space, which is equivalent to a zero-lag cross-correlation of the two wavefields (Claerbout, 1971;

Kaelin and Guitton, 2006; Chattopadhyay and McMechan, 2008). While this cross-correlation imaging principle actually provides a blurred image for that it employs the adjoint of the forward modeling operator (Lailly, 1983). The imaging problem can be posed as an inverse problem based on the minimization of a least-squares function (Tang, 2009) to deblur the migrated image using the Hessian operator. Thus, the imaging problem is related to what are discussed in waveform inversion problem(Ben-Hadj-Ali et al., 2011).

The gradient based method, which simplifies the Hessian matrix to an identity matrix, has been proved to be a crude strategy in scaling. The un-scaled or blurred image can be enhanced considerably by multiplying the inverse Hessian. Pratt et al. (1998) showed that the Hessian can sharpen the blurred images obtained by the less expensive gradient method and the diagonal terms of the approximate Hessian is a zero-lag correlation of the scattered waves, which represent the geometrical spreading effects as the scattering point moves away from the sources and receivers. While the explicit calculation of the Hessian matrix is extremely expensive.

In this research, we found that the cross-correlation based gradient can be enhanced obviously by the source-side illumination, which forms the standard deconvolution imaging condition in RTM. And the source illumination is actually equivalent to the diagonal part of the pseudo-Hessian, which was proposed by (Shin et al., 2001a). Furthermore, the deconvolution imaging condition actually transforms the unit of the gradient from the square of the amplitude to the reflectivity, which makes it possible for us to combine FWI with traditional impedance inversion method. We applied the Iterative Modelling Migration and Inversion (IMMI) method by Margrave et al. (2012a) on a modified Marmousi model. The phase encoding method was also introduced to construct the gradient and diagonal pseudo-Hessian and preconditioning the phase encoded gradient using the diagonal pseudo-Hessian is equivalent to the deconvolution imaging condition (Pan et al., 2013a,b,c, 2014). In the iterative process, we varied the ray parameter regularly during the iterative process, namely, iteration-dependent ray parameter setting strategy (Pan et al., 2014). The effectiveness for combing these strategies was presented by the inverted result.

THEORY AND METHOD

The deconvolution imaging condition and cross-correlation imaging condition are reviewed firstly in this part. And the deconvolution imaging condition effects are presented based on Marmousi model. And we compared and discussed the cross-correlation based gradient and deconvolution based gradient. And we proposed to estimate the model perturbation using the traditional impedance inversion method. The combination of deconvolution based gradient and impedance inversion method forms the Iterative Modelling Migration and Inversion (IMMI) method by Margrave et al. (2012a).

Deconvolution and cross-correlation imaging conditions

The approximation to reflectivity I is always given as the ratio of upgoing wavefields $U(\mathbf{r},\omega)$ to downgoing wavefields $D(\mathbf{r},\omega)$ (Claerbout, 1971):

$$I(\mathbf{r}) = \int_{\omega}^{\omega_N} d\omega \frac{U(\mathbf{r},\omega)}{D(\mathbf{r},\omega)}, \tag{1}$$

where $\mathbf{r} = (x, y, z)$ is the Euclidean coordinate of a subsurface position, ω is the temporal frequency, and ω_N is the Nyquist frequency. Equation (1) is unstable when $D(\mathbf{r}, \omega)$ is 0 (or very close to 0). To avoid this instability, the complex conjugate of downgoing wavefields D^* ("$*$" denotes the complex conjugate) is always multiplied in denominator and numerator of equation (1) (Guitton et al., 2007; Scheleicher et al., 2008), which gives:

$$I(\mathbf{r}) = \int_{\omega}^{\omega_N} d\omega \frac{D^*(\mathbf{r}, \omega)U(\mathbf{r}, \omega)}{D^*(\mathbf{r}, \omega)D(\mathbf{r}, \omega) + \lambda A_{max}}, \tag{2}$$

where λA_{max} is further stable term, λ is the damping factor, a small constant value and A_{max} means the maximum value of $D^*(\mathbf{r}, \omega)D(\mathbf{r}, \omega)$. Equation (2) is equivalent to equation (1) multiplied by an optimal Wiener filter (Guitton et al., 2007). The term $D^*(\mathbf{r}, \omega)D(\mathbf{r}, \omega)$ in the denominator is the auto-correlation of the downgoing wavefields, which is also called the source illumination. And the unit of source illumination is the square of amplitude, which means that it is independent of the frequency ω. Thus, the source illumination can be approximated as a constant A_D^2 and we can take it out from the integration. Then equation (2) becomes:

$$I(\mathbf{r}) \simeq \frac{1}{A_D^2} \int_{\omega}^{\omega_N} d\omega D^*(\mathbf{r}, \omega)U(\mathbf{r}, \omega) \simeq \int_{\omega}^{\omega_N} d\omega D^*(\mathbf{r}, \omega)U(\mathbf{r}, \omega), \tag{3}$$

A_D^2 can be considered as a further stable term (Scheleicher et al., 2008). Thus, equation (3) becomes the so called cross-correlation imaging condition. The cross-correlation imaging condition corresponds to Claerbout's (1971) imaging principle: the reflector exists where the downgoing wavefields $D(\mathbf{r}, \omega)$ and upgoing wavefields $U(\mathbf{r}, \omega)$ coincide in time and space, which is equivalent to a zero-lag cross-correlation of the two wavefields (Claerbout, 1971; Kaelin and Guitton, 2006; Chattopadhyay and McMechan, 2008). The cross-correlation imaging condition is similar to the deconvolution imaging condition but lacks the intrinsic gain correction of the former. And the equations for reverse time migration with cross-correlation and deconvolution imaging conditions are shown in equations (4) and (5) respectively:

$$I_{cross}(\mathbf{r}) = \sum_{\mathbf{r}_s} \int d\omega \Re\{\omega^2 f_s(\omega) G(\mathbf{r}, \mathbf{r}_s, \omega) G(\mathbf{r}_g, \mathbf{r}, \omega) \psi^*(\omega)\}, \tag{4}$$

$$I_{dec}(\mathbf{r}) = \frac{\sum_{\mathbf{r}_s} \int d\omega \Re\{\omega^2 f_s(\omega) G(\mathbf{r}, \mathbf{r}_s, \omega) G(\mathbf{r}_g, \mathbf{r}, \omega) \psi^*(\omega)\}}{\sum_{\mathbf{r}_s} \int d\omega \Re\{\omega^4 |f_s(\omega)|^2 |G(\mathbf{r}, \mathbf{r}_s, \omega)|^2 \psi^*(\omega)\} + \lambda A_{max}}, \tag{5}$$

where I_{cross} and I_{dec} mean the images based on cross-correlation and deconvolution imaging conditions respectively, $f_s(\omega)$ is the source function, $\mathbf{r}_s = (x_s, y_s, z_s)$ and $\mathbf{r}_g = (x_g, y_g, z_g)$ are the positions of sources and receivers, $G(\mathbf{r}_s, \omega)$ indicate the source wavefields, $G(\mathbf{r}_g, \mathbf{r}, \omega)\psi^*(\omega)$ mean the backpropagated wavefields from all receivers by adjoint state technique, $\psi^*(\omega)$ is the complex conjugate of the observed data, $\Re\{\cdot\}$ means the real part of the value and A_{max} indicates the maximum value of the source illumination.

Analytic Solution of the Imaging Conditions

Figure 1 shows an analytic analysis for the two imaging conditions. Assuming that the reflector is embedded in a homogeneous background with a constant velocity of c_0. The wavenumber becomes $k_0 = \omega/c_0$. The analytic expression of the downgoing wavefields ca be written as:

$$D = A\frac{e^{ik_0\tilde{\mathbf{r}}_s}}{4\pi\tilde{\mathbf{r}}_s}, \tag{6}$$

where $\tilde{\mathbf{r}}_s = \sqrt{(x-x_s)^2 + (y-y_s)^2 + (z-z_s)^2}$ indicate the distance from the source to reflector. When the incident wavefields illuminate the reflector, the scattered wavefields or upgoing wavefields can be recorded at the receiver position:

$$U = AR\frac{e^{ik_0(\tilde{\mathbf{r}}_s+\tilde{\mathbf{r}}_g)}}{4\pi\left(\tilde{\mathbf{r}}_s+\tilde{\mathbf{r}}_g\right)}, \tag{7}$$

where R is the true reflection coefficient and $\tilde{\mathbf{r}}_g$ denotes the distance from the receiver to the reflector. The wavefields backpropagated from the receiver can be expressed as:

$$\tilde{U} = AR\frac{e^{ik_0\tilde{\mathbf{r}}_s}}{4\pi\tilde{\mathbf{r}}_s}, \tag{8}$$

So, the analytic solution of the cross-correlation imaging condition can be obtained by cross-correlating the downgoing wavefields and the complex conjugate of the backpropagated wavefields:

$$I_{cross} = D\tilde{U}^* = \frac{1}{\tilde{\mathbf{r}}_s^2}\frac{A^2}{(4\pi)^2}R, \tag{9}$$

And the analytic solution of the deconvolution imaging condition can be written as:

$$I_{dec} = \frac{\tilde{U}D^*}{DD^*} = R, \tag{10}$$

It can be seen that the unit of the cross-correlation imaging condition is the square of the amplitude A^2. However, the unit of the deconvolution imaging condition is reflectivity R. Further, we can notice that the image by the cross-correlation imaging condition is poorly scaled and the reflectivity estimation is off by a spatially variant factor $\tilde{\mathbf{r}}_s^{-2}$, which can be compensated in deconvolution imaging condition. Another advantage of the deconvolution imaging condition is that it can improve resolution and suppress crosstalk noise, which cannot be observed obviously (Poole et al., 2010).

Here, we presented a numerical example based on the Marmousi model with $2km$ in depth and $4.5km$ in wide. Firstly, we use one shot record to image and examine the imaging result. Figure 2 shows the imaging results based on cross-correlation imaging condition when the source is located at $0.75km$ (a), $2.25km$ (b) and $3.75km$ (c) respectively. And Figure (d), (e) and (f) are the corresponding source illuminations for different source locations.

Fig. 3a, b and c show the single shot imaging results based on deconvolution imaging condition without damping term λA_{max} in equation (5). We can see that some extremely large values appear, as indicated by the blue areas in Fig. 3a and c. Fig. 3d, e and f show the imaging results based on stabilized deconvolution imaging condition with $\lambda = 0.01$. The damping term suppressed the noise at far offset and stabilized the deconvolution imaging

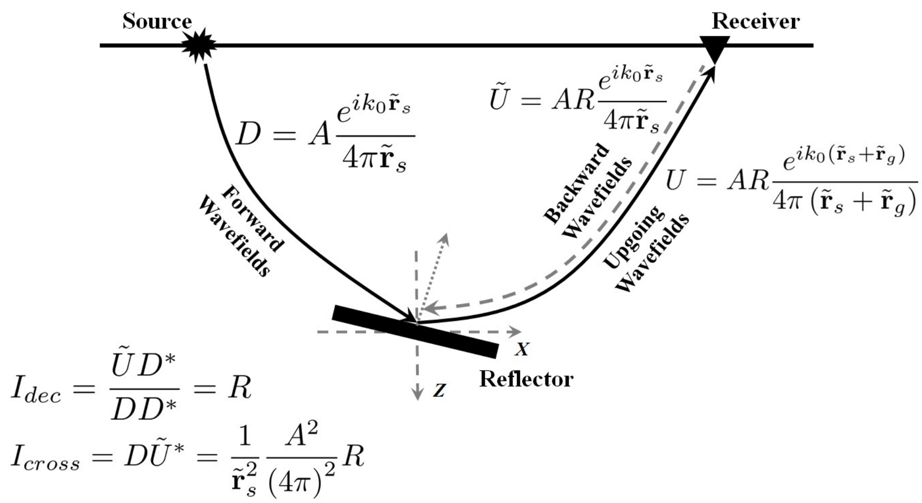

FIG. 1. Analytic solution of the cross-correlation and deconvolution imaging conditions. k_0 is the wavenumber, $\tilde{\mathbf{r}}_s$ and $\tilde{\mathbf{r}}_g$ indicate the distances from the source and receiver to the reflector, A and R indicate the amplitude and the reflection coefficient, I_{cross} and I_{dec} are the images based on cross-correlation and deconvolution imaging conditions respectively.

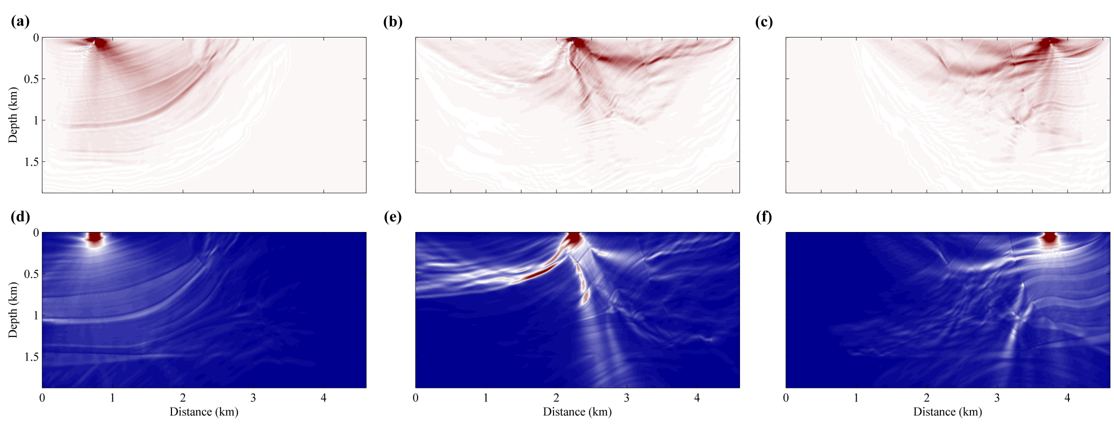

FIG. 2. Imaging results based on cross-correlation imaging condition. (a), (b) and (c) are the Imaging results based on cross-correlation imaging condition when the source is located at $0.75km$, $2.25km$ and $3.75km$ respectively. And (d), (e) and (f) are the source illuminations corresponding to different source locations.

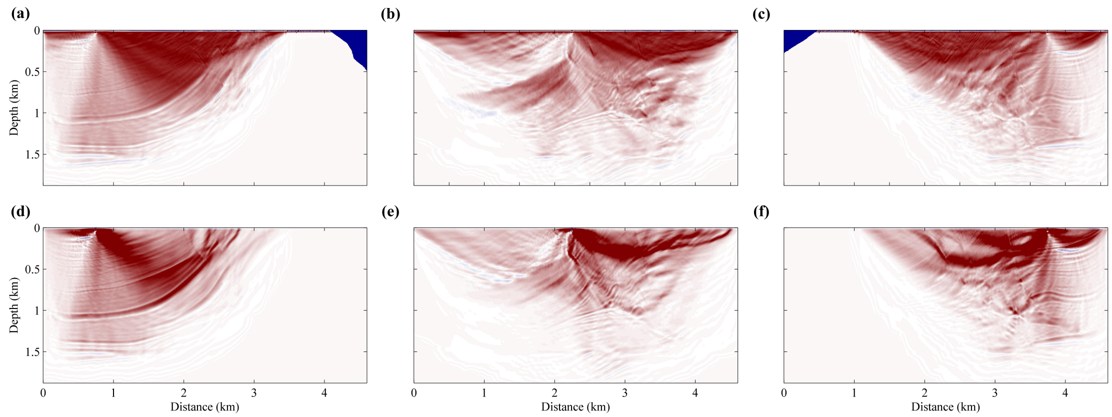

FIG. 3. Imaging results based on deconvolution imaging condition. (a), (b) and (c) are the Imaging results based on deconvolution imaging condition when the source is located at $0.75km$, $2.25km$ and $3.75km$ respectively. And (d), (e) and (f) are the imaging results based on stabilized deconvolution imaging condition.

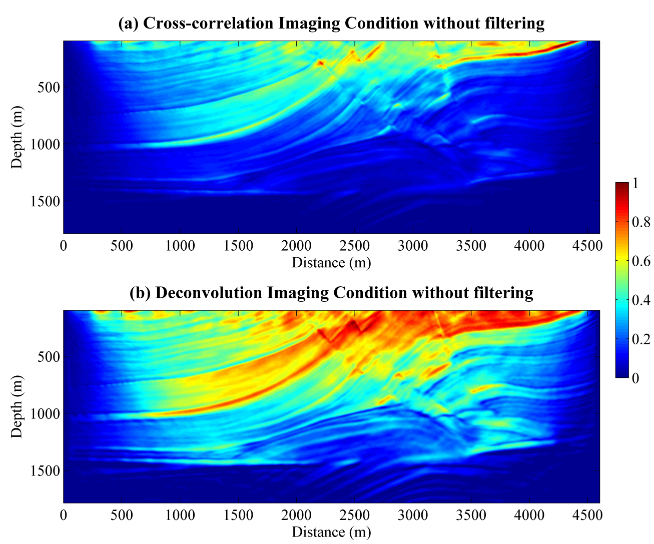

FIG. 4. RTM imaging results before filtering based on cross-correlation imaging condition (a) and deconvolution imaging condition (b).

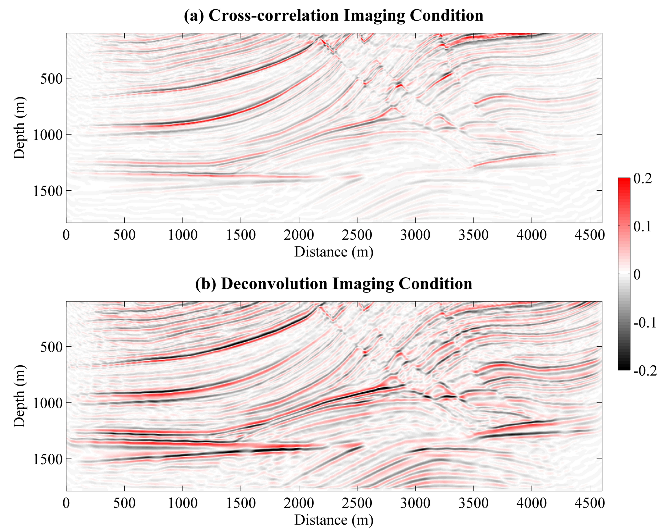

FIG. 5. RTM imaging results after filtering based on cross-correlation imaging condition (a) and deconvolution imaging condition (b).

condition. Compared to the imaging results in Fig.2, it can be seen that the deconvolution imaging condition can enhance the deep reflectors amplitudes of the imaging results.

Fig. 4a and b show the imaging results by cross-correlation imaging condition and deconvolution imaging condition respectively before filtering by 18 sources distributed from $0.25km$ to $4.5km$ with a spacing of $0.15km$. And Fig. 5a and b show the imaging results after filtering based on different imaging conditions. We can see that amplitudes of the deep reflectors in Fig. 5b have been obviously enhanced, compared to Fig. 5a.

FWI gradient based on deconvolution imaging condition

Full waveform inversion estimates the subsurface parameters through an iterative process by minimizing the difference between the synthetic data **Bu** and observed data **d** (Lailly, 1983; Tarantola, 1984) .The misfit function ϕ is formulated in a least-squares form (Virieux and Operto, 2009):

$$\phi = \frac{1}{2}\|\mathbf{d} - \mathbf{B}\mathbf{u}\|_2, \tag{11}$$

where **B** is the forward modeling operator, **u** mean the wavefields and $\|\cdot\|_2$ means the $\ell - 2$ norm.

The minimum value of the misfit function is sought in the vicinity of the starting model $m_0(\mathbf{r})$ and the updated model can be written as the summation of the starting model and a model perturbation $\delta m_k(\mathbf{r})$ (Virieux and Operto, 2009).

$$m_{k+1}(\mathbf{r}) = m_k(\mathbf{r}) + \mu_k \delta m_k(\mathbf{r}), \tag{12}$$

where μ_k is the step length in kth iteration, which is a scalar constant used to scale the model perturbation $\delta m_k(\mathbf{r})$ and can be obtained through a line search method (Gauthier et al., 1986; Pica et al., 1990).

Then we can apply a second order Taylor-Lagrange development of the misfit function and take partial derivative with respect to the model parameter. When the equation towards to zero, the misfit function approaches its minimum value and the model perturbation can be expressed as:

$$\delta m = -H^{-1} g, \tag{13}$$

where $g = \frac{\partial \phi}{\partial m}$ is the gradient and $H = \frac{\partial^2 \phi}{(\partial m)^2}$ is the Hessian matrix. For gradient, the first order partial derivative of the misfit function ϕ with respect to the model parameters, can be obtained by a zero-lag correlation between the complex conjugate of the data residuals and the first order partial derivative wavefields:

$$g = \frac{\partial \phi}{\partial m} = \Re \left(\mathbf{J}^T \Delta d^* \right), \tag{14}$$

where $\mathbf{J} = \frac{\partial \mathbf{u}}{\partial m}$ is the Jacobian matrix (or sensitive matrix), Δd^* denote the complex conjugate of the data residuals. T and $*$ indicate the transpose and complex conjugate respectively. While it is extremely expensive to calculate the Jacobian matrix or the first order partial derivative wavefields directly. For the Jacobian matrix, we can reexamine the wave equation firstly:

$$\mathbf{B}(\mathbf{r}, \omega)\mathbf{u}(\mathbf{r}, \omega) = f(\mathbf{r}, \omega), \tag{15}$$

where $\mathbf{r} = (x, y, z)$ means the subsurface position, ω is the angular frequency and $f(\mathbf{r}, \omega)$ is the source term. And according to the virtual source theory (Pratt et al., 1998; Virieux and Operto, 2009), we can take partial derivative on both sides of equation (15):

$$\mathbf{B} \frac{\partial \mathbf{u}}{\partial m} = -\frac{\partial \mathbf{B}}{\partial m} \mathbf{u}, \tag{16}$$

The right hand side of the above equation is always referred to as "scattered sources" or "secondary Born sources". It underlines the fact that the scattered wavefields due to the perturbations in the model parameters can be interpreted as the wavefields generated by a set of secondary body sources. Isolating the Jacobian matrix on the left hand side of gives:

$$\mathbf{J} = \frac{\partial \mathbf{u}}{\partial m} = -\mathbf{B}^{-1} \frac{\partial \mathbf{B}}{\partial m} \mathbf{u}, \tag{17}$$

Then substituting equation (17) into equation (14) and the gradient can be expressed as:

$$g = -\Re \left\{ \left(\mathbf{B}^{-1} \frac{\partial \mathbf{B}}{\partial m} \mathbf{u} \right)^T \Delta d^* \right\}, \tag{18}$$

Because $\frac{\partial \mathbf{B}}{\partial m} = \omega^2$, expansion of the transpose in the above equation results in the final expression for the gradient:

$$g = -\sum_{\omega, \mathbf{r}_s} \omega^2 \Re \left\{ \mathbf{u}^T \otimes \left(\mathbf{B}^{-1} \right)^T \Delta d^* \right\}, \tag{19}$$

where $\otimes$ means cross-correlation. Hence, the gradient can be calculated using the adjoint state method by applying a zero-lag convolution between the forward modeling wavefields and back-propagated data residuals (Lailly, 1983; Tarantola, 1984), which avoids the direct computation of the first order partial derivative wavefields. The analytic expression of the gradient can be obtained as:

$$g = -\sum_{\omega,\mathbf{r}_s} \omega^2 \Re \left\{ A \frac{e^{ik_0 \tilde{\mathbf{r}}_s}}{4\pi \tilde{\mathbf{r}}_s} \otimes AR \frac{e^{-ik_0 \tilde{\mathbf{r}}_s}}{4\pi \tilde{\mathbf{r}}_s} \right\} = -\sum_{\omega,\mathbf{r}_s} \omega^2 \Re \left\{ \frac{1}{\tilde{\mathbf{r}}_s^2} \frac{A^2}{16\pi^2} R \right\}, \tag{20}$$

And furthermore, we can notice that the gradient is indeed a poorly scaled image and it decays as $\tilde{\mathbf{r}}_s^{-2}$ for this analytic example.

The Hessian matrix is the second order partial derivative of the misfit function with respect to the model parameters. Under the assumption of small model perturbation, the nonlinear term in the Hessian matrix can be ignored and the linear term forms the approximate Hessian, which can be written as a scalar product between two Jacobian matrices:

$$H_a = \mathbf{J}^* \mathbf{J}, \tag{21}$$

Substituting equation (17) into the above equation forms:

$$H_a = \left(\mathbf{B}^{-1} \frac{\partial \mathbf{B}}{\partial m} \mathbf{u} \right)^* \left(\mathbf{B}^{-1} \frac{\partial \mathbf{B}}{\partial m} \mathbf{u} \right) \tag{22}$$

The above equation can be expanded as:

$$H_a = \omega^4 \Re \left(\mathbf{u}^* \mathbf{u} \left(\mathbf{B}^{-1} \right)^* \left(\mathbf{B}^{-1} \right) \right), \tag{23}$$

The pseudo-Hessian proposed by (Shin et al., 2001a) is constructed by two virtual sources:

$$H_{pseudo} = f_{virtual}^* f_{virtual} = \left(\frac{\partial \mathbf{B}}{\partial m} \mathbf{u} \right)^* \left(\frac{\partial \mathbf{B}}{\partial m} \mathbf{u} \right) = \omega^4 \Re \left(\mathbf{u}^* \mathbf{u} \right), \tag{24}$$

Compared to the approximation Hessian H_a, we can see that the pseudo-Hessian H_{pseudo} actually ignores the back propagation operator $\left(\mathbf{B}^{-1} \right)^* \left(\mathbf{B}^{-1} \right)$. The model perturbation can be obtained by plugging equation (14) and equation (21) into equation (13):

$$\delta m = -(\mathbf{J}^* \mathbf{J} + \lambda I)^{-1} (\mathbf{J}^T \Delta d^*), \tag{25}$$

where I is an identity matrix and λ is an constant value. And λI is stable term which makes Hessian matrix invertible. When considering high frequency asymptotics, the Hessian matrix is highly diagonal dominant and the inverse of the approximate Hessian can be replaced by its reciprocal:

$$\delta m = -\frac{\mathbf{J}^T \Delta d^*}{\mathbf{J}^* \mathbf{J} + \lambda I}, \tag{26}$$

Plugging the gradient and approximate Hessian into the above equation gives:

$$\delta m = -\Re \left\{ \frac{\left(\mathbf{B}^{-1} \frac{\partial \mathbf{B}}{\partial m} \mathbf{u} \right)^T \Delta d^*}{\left(\mathbf{B}^{-1} \frac{\partial \mathbf{B}}{\partial m} \mathbf{u} \right)^* \left(\mathbf{B}^{-1} \frac{\partial \mathbf{B}}{\partial m} \mathbf{u} \right)} \right\}, \tag{27}$$

After a further derivation, the model perturbation becomes:

$$\delta m = -\Re \left\{ \frac{1}{\left(\mathbf{B}^{-1} \right)^* \left(\mathbf{B}^{-1} \right)} \frac{\sum_{\omega,\mathbf{r}_s} \omega^2 \Re \left\{ \mathbf{u}^T \otimes \left(\mathbf{B}^{-1} \right)^T \Delta d^* \right\}}{\omega^4 \mathbf{u}^* \mathbf{u}} \right\}, \tag{28}$$

The convolution of the forward modeling wavefields $\mathbf{u}^*\mathbf{u}$ compensate the geometrical spreading effects on the source-side and the backward propagation operator $\left(\mathbf{B}^{-1}\right)^*\left(\mathbf{B}^{-1}\right)$ compensates the geometrical spreading effects on the receiver-side. Under the assumption of infinite receiver coverage, this part can be approximated as an constant. And the diagonal part of $\mathbf{u}^*\mathbf{u}$ is the auto-correlation of the forward modeling wavefields and the gradient preconditioned by $\mathbf{Diag}\left(\mathbf{u}^*\mathbf{u}\right)$ is equivalent to the deconvolution imaging condition in reverse time migration:

$$\delta m \simeq -\Re\left\{\frac{\sum_{\omega,\mathbf{r}_s}\omega^2\Re\left\{\mathbf{u}^T\otimes\left(\mathbf{B}^{-1}\right)^T\Delta d^*\right\}}{\mathbf{Diag}\left(\omega^4\mathbf{u}^*\mathbf{u}\right)}\right\}, \tag{29}$$

The analytic expression for the pseudo-Hessian can be written as:

$$H_{pseudo} = \Re\left\{\omega^4\mathbf{u}^*(\mathbf{r}_s',\omega)\mathbf{u}(\mathbf{r}_s'',\omega)\right\} = \Re\left\{\omega^4\left(A\frac{e^{-ik_0\tilde{\mathbf{r}}_s'}}{4\pi\tilde{\mathbf{r}}_s'}\right)\left(A\frac{e^{ik_0\tilde{\mathbf{r}}_s''}}{4\pi\tilde{\mathbf{r}}_s''}\right)\right\}, \tag{30}$$

The diagonal part of the pseudo-Hessian can be obtained when $\mathbf{r}_s' = \mathbf{r}_s''$:

$$\begin{aligned}\mathbf{Diag}\left(H_{pseudo}\right) &= \Re\left\{\omega^4\mathbf{u}^*(\mathbf{r}_s,\omega)\mathbf{u}(\mathbf{r}_s,\omega)\right\}\\ &= \Re\left\{\omega^4\left(A\frac{e^{-ik_0\tilde{\mathbf{r}}_s}}{4\pi\tilde{\mathbf{r}}_s}\right)\left(A\frac{e^{ik_0\tilde{\mathbf{r}}_s}}{4\pi\tilde{\mathbf{r}}_s}\right)\right\}\\ &= \Re\left\{\omega^4\frac{A^2}{16\pi^2\tilde{\mathbf{r}}_s^2}\right\}\end{aligned} \tag{31}$$

Hence, we can obtain the analytic expression for the model perturbation by inserting equation (20) and (31) into equation (29):

$$\delta m \simeq \Re\left\{\frac{-\sum_{\omega,\mathbf{r}_s}\omega^2\frac{1}{\tilde{\mathbf{r}}_s^2}\frac{A^2}{16\pi^2}R}{\omega^4\frac{A^2}{16\pi^2\tilde{\mathbf{r}}_s^2}}\right\} = -\Re\left\{\sum_{\omega}\omega^2 R\right\}, \tag{32}$$

It is obvious that the model perturbation based on the deconvolution imaging condition compensate the energy loss in the cross-correlation based gradient and estimate the reflectivity directly.

Phase Encoded Gradient and Phase Encoded Pseudo-Hessian

While it is still extremely expensive to construct the gradient using the traditional shot-profile method. Simultaneous source technique was introduced in seismic imaging for reducing the extensively computational buren firstly, namely, delayed-shot migration (Zhang et al., 2005). As we mentioned above, the construction of the gradient is equivalent to a migration process. Hence, it is also possible for us to create the gradient using multisource technique for improving the efficient of FWI.

While when the sources are densely distributed, the multisource method can introduce seriously crosstalk noise arising from undesired interactions between unrelated source and receiver wavefields. The phase encoding method can be employed to disperse or shift these unwanted crosstalk terms. The linear phase encoding technique is performed by applying linear phase shifts (or time delays in time domain) to the shot records. The phase shift function $\gamma(x_s, p, w) = \omega p(x_s - x_0)$ is controlled by the ray parameter (or slant parameter) p and source's position x_s. Generally, sufficient ray parameters can disperse or reduce these crosstalk terms effectively. And a common-receiver gather can be transformed into a single trace from a linear source wavefields by $\tau - p$ transform (Zhang et al., 2005):

$$\tilde{\mathbf{d}}\left(\mathbf{r}_g, \mathbf{r}_s, p, \omega\right) = \int \mathbf{d}\left(\mathbf{r}_g, \mathbf{r}_s, \omega\right) e^{i\omega p(x_s - x_0)} d\mathbf{r}_s \tag{33}$$

where $\mathbf{r}_g = (x_g, y_g = 0, z_g = 0)$ and $\mathbf{r}_s = (x_s, y_s = 0, z_s = 0)$ mean the locations of the receivers and sources. Actually, equation (12) can be considered as a Fourier transform, replacing the wavenumber k with ωp. This theory can be used to determine the ray parameter spacing $\triangle p$ for slant stacking (Zhang et al., 2005).

The slant gradient with ray parameter p_i can be expressed as:

$$\tilde{g}\left(\mathbf{r}, p_i, \omega\right) = \sum_{\omega} \sum_{\mathbf{r}_s} \sum_{\mathbf{r}'_s} \Re \left\{ \omega^2 \mid A(\omega) \mid^2 \tilde{G}\left(\mathbf{r}, \mathbf{r}_s, \omega\right) \bar{G}^*\left(\mathbf{r}, \mathbf{r}'_s, \omega\right) e^{i\omega p_i (\mathbf{r}_s - \mathbf{r}'_s)} \right\}, \tag{34}$$

In the above equation, when $\mathbf{r}_s = \mathbf{r}'_s$, the linear phase encoded gradient $\tilde{g}\left(\mathbf{r}, p_i, \omega\right)$ is equal to the conventional shot-profile gradient $g_{sp}\left(\mathbf{r}, \omega\right)$. And when $\mathbf{r}_s \neq \mathbf{r}'_s$, the linear phase encoded gradient $\tilde{g}\left(\mathbf{r}, p_i, \omega\right)$ becomes the cross terms g_{cross}. So, the linear phase encoded image can be written as a summation of the conventional shot-profile gradient and the crosstalk term:

$$\tilde{g}\left(\mathbf{r}, p_i, \omega\right) = g_{sp}\left(\mathbf{r}, \omega\right) + g_{cross}, \tag{35}$$

To disperse the crosstalk term in the above equation, we can construct the gradient by slant stacking a set of ray parameters:

$$\tilde{g}\left(\mathbf{r}, \mathbf{p}^g, \omega\right) = \sum_{\omega} \sum_{\mathbf{r}_s} \sum_{\mathbf{r}'_s} \sum_{i=1}^{N_p^g} \Re \left\{ \omega^2 \mid A(\omega) \mid^2 \tilde{G}\left(\mathbf{r}, \mathbf{r}_s, \omega\right) \bar{G}^*\left(\mathbf{r}, \mathbf{r}'_s, \omega\right) e^{i\omega p_i^g (\mathbf{r}_s - \mathbf{r}'_s)} \right\}, \tag{36}$$

where i and N_p^g indicate the ray parameter index and the maximum number of the ray parameters used to construct the phase encoded gradient. And $\mathbf{p}^g$ means the ray parameters vector:

$$\mathbf{p}^g = [\ p_1^g \quad p_2^g \quad p_3^g \quad \cdots \quad p_{N_p^g}^g\ ], \tag{37}$$

We can also use the phase encoding method to construct the pseudo-Hessian, which can be expressed as:

$$\tilde{H}_{pseudo}\left(\mathbf{p}^H, \omega\right) = \sum_{\omega} \sum_{\mathbf{r}_s} \sum_{\mathbf{r}'_s} \sum_{i=1}^{N_p^H} \Re \left\{ \omega^4 \mid A(\omega) \mid^2 \tilde{G}\left(\mathbf{r}, \mathbf{r}_s, \omega\right) \tilde{G}^*\left(\mathbf{r}', \mathbf{r}'_s, \omega\right) e^{i\omega p_i^H (\mathbf{r}_s - \mathbf{r}'_s)} \right\}, \tag{38}$$

where N_p^H is the maximum number of ray parameters for constructing the phase encoded pseudo-Hessian and $\mathbf{p}^H$ is the ray parameter vector:

$$\mathbf{p}^H = [\ p_1^H \quad p_2^H \quad p_3^H \quad \cdots \quad p_{N_p^H}^H\ ], \tag{39}$$

Similarly, the phase encoded pseudo-Hessian can be written as a summation of the exact pseudo-Hessian and the crosstalk term:

$$\tilde{H}_{pseudo}\left(\mathbf{p}^H,\omega\right) = H_{pseudo}\left(\omega\right) + H_{cross}, \tag{40}$$

And the diagonal part of the phase encoded pseudo-Hessian can be obtained by setting $\mathbf{r} = \mathbf{r}'$:

$$\textbf{Diag}\left(\tilde{H}_{pseudo}\left(\mathbf{p}^H,\omega\right)\right) = \sum_{\omega}\sum_{\mathbf{r}_s}\sum_{\mathbf{r}'_s}\sum_{i=1}^{N_p^H}\Re\left\{\omega^4 \mid A(\omega) \mid^2 \tilde{G}\left(\mathbf{r},\mathbf{r}_s,\omega\right)\tilde{G}^*\left(\mathbf{r},\mathbf{r}'_s,\omega\right)e^{i\omega p_i^H\left(\mathbf{r}_s-\mathbf{r}'_s\right)}\right\}, \tag{41}$$

So, preconditioning the phase encoded gradient using the diagonal part of the phase encoded pseudo-Hessian can give a crude approximation of the reflectivity:

$$\tilde{R}\left(\mathbf{p}^g,\mathbf{p}^H,\omega\right) = \frac{\tilde{g}\left(\mathbf{r},\mathbf{p}^g,\omega\right)}{\textbf{Diag}\left(\tilde{H}_{pseudo}\left(\mathbf{p}^H,\omega\right)\right) + \lambda\tilde{A}_{max}}, \tag{42}$$

The above equation can be named as phase encoded deconvolution imaging condition, which is actually equivalent to the shot-profile deconvolution imaging condition with sufficient ray parameters in $\mathbf{p}^g$ and $\mathbf{p}^H$.

Model Perturbation Estimation Using Impedance Inversion Method

As discussed above, deconvolution imaging condition actually transforms the unit of the gradient from square of amplitude to reflectivity. Now, it is able for us to convert the reflectivity to impedance using the theory in impedance inversion. Recall that at normal incidence, the reflection coefficients and the impedance have the relationship of:

$$R_n = \frac{I_{n+1} - I_n}{I_{n+1} + I_n} = \frac{\Delta I}{I_{n+1} + I_n}, \tag{43}$$

where R_n means the reflection coefficients at interface n, I_n and I_{n+1} are the impedance at n and $n+1$ layers respectively, and ΔI is the impedance difference. If I_n is close to I_{n+1}, we can assume that:

$$I_{n+1} + I_n \simeq 2I_n, \tag{44}$$

And then the impedance perturbation can be approximated as:

$$\Delta I \simeq 2I_nR_n, \tag{45}$$

So, during iterative cycles of FWI, the impedance update can be written as:

$$I_{k+1} = I_k + \Delta I_k = I_k + 2I_kR_k, \tag{46}$$

where I_k and I_{k+1} are the impedances in the kth and $k+1$th iterations and R_k is the reflectivity in the kth iteration.

Table1. Pseudo code for IMMI method

BEGIN $\leftarrow I_0$, initial model;
WHILE $\epsilon \leq \epsilon_{min}$ or $k \leq k_{max}$

1. Identify the ray parameter $p^g_{i,k}$ for constructing the phase encoded gradient
2. Identify the frequency band $f^k = f_0 \rightarrow f_{max}, f_{interval}$, every n iterations
3. Generate the data residual δP and apply low-pass filtering $\delta \tilde{P} = \textbf{low\_pass}\left(\delta P, f^k\right)$
4. Create the phase encoded gradient $\tilde{g}_k\left(p^g_{i,k}, \omega\right)$
5. **FOR** $i = 1$ to N^H_p, every 1 or m iterations

    Construct the diagonal part of the phase encoded pseudo-Hessian: $\textbf{Diag}\left(\tilde{H}^k_{pseudo}\right)$

    END FOR
6. Impedance Perturbation Estimation:

    $\Delta I = 2 I_k \Re \left\{ \frac{\tilde{g}_k\left(p^g_{i,k}, \omega\right)}{\textbf{Diag}\left(\tilde{H}^k_{pseudo}(\mathbf{p}^H, \omega)\right)/N^H_p + \lambda A_{max}} \right\}$
7. Calculate the step length μ_k using the line search method
8. Update the impedance:

    $I_{k+1} = I_k + 2\mu_k I_k \Re \left\{ \left(\textbf{Diag}\left(\tilde{H}^k_{pseudo}\left(\mathbf{p}^H, \omega\right)\right)/N^H_p + \lambda A_{max}\right)^{-1} \tilde{g}_k\left(p^g_{i,k}, \omega\right) \right\}$
9. Calculate the relative least-squares error:

    $\epsilon = \frac{\|I_{k+1} - I_{true}\|_2}{\|I_{true}\|_2}$

END WHILE

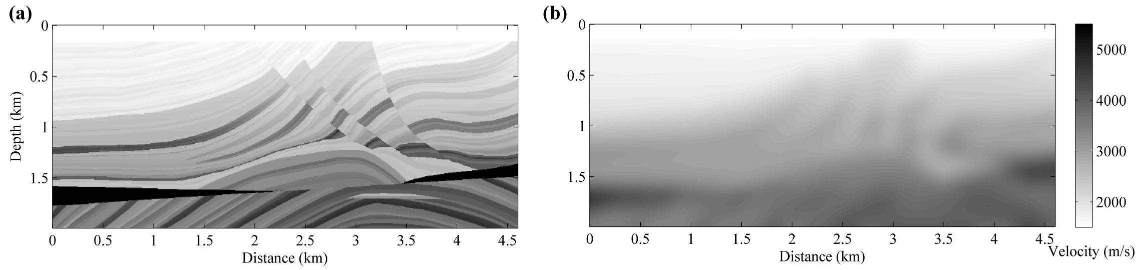

FIG. 6. (a) True Velocity Model; (b) Initial Velocity Model.

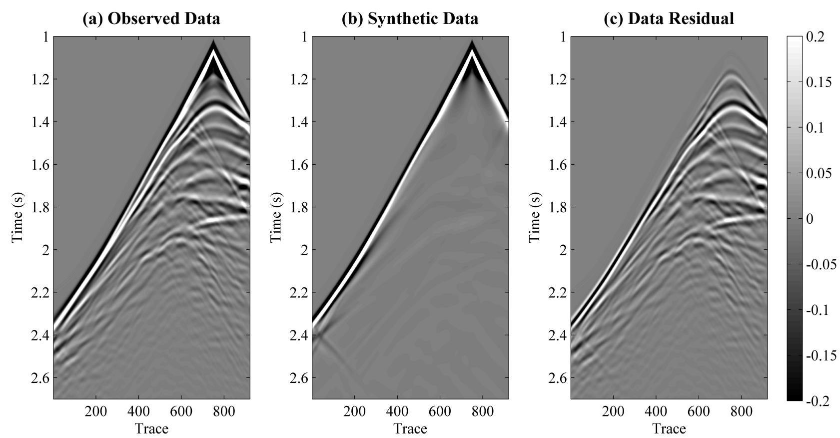

FIG. 7. Shot records. (a) Observed data; (b) Synthetic data; (c) Data residual. The source location is $[3.75km, 0.025km]$

Iterative Modeling Migration and Inversion (IMMI) Method

As proposed by Margrave et al. (2012b), the estimation of the reflectivity using deconvolution imaging condition allows us to combine FWI with Standard Methodology (SM), which forms the Iterative Modeling Migration Inversion (IMMI) method. The pseudo code of the IMMI method is presented in Table 1.

NUMERICAL EXAMPLES

To verify the advantage and effectiveness of deconvolution imaging condition for FWI, a numerical test was performed using a 2D Marmousi dataset. The true velocity model is shown in Fig.6a and Fig.6b is the initial velocity model obtained by applying a 2D Gaussian convolutional smoother to the true velocity model with half width of $30m$. The model is $4.6km$ in wide and $2km$ in depth. We generate the synthetic data with 18 shots and 920 receivers located at the top surface of the model. The source interval is $50m$ and receiver interval is $5m$. And the dominant frequency of the source wavelet is $10Hz$. Fig. 7a, b and c show the observed data, synthetic data and data residual when the source is located at $(3.75km, 0.025km)$.

Fig.8a, b and c show the source illumination, cross-correlation based gradient, deconvolution based gradient respectively, when the source is located at $(x = 2.250km, z = 0.025km)$. Fig.9a, b and c show the source illumination, cross-correlation based gradient, deconvolution based gradient respectively, when the source is located at $(x = 3.750km, z = 0.025km)$.

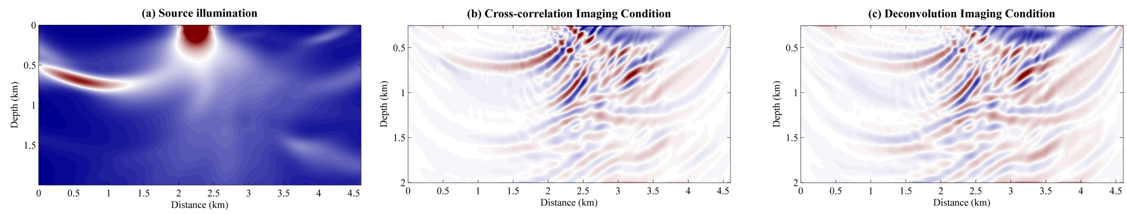

FIG. 8. (a) Source illumination; (b) Gradient by cross-correlation imaging condition; (c) Gradient by deconvolution imaging condition. The source location is $(2.25km, 0.025km)$.

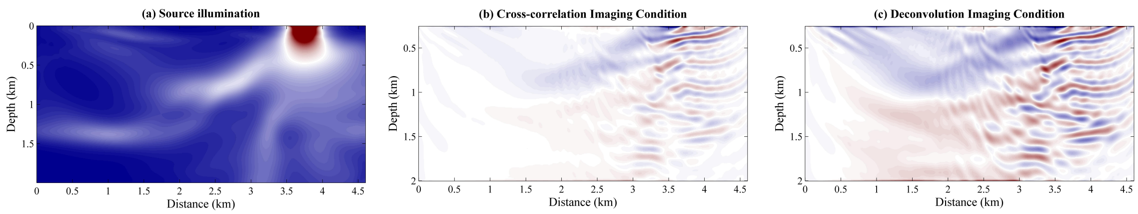

FIG. 9. (a) Source illumination; (b) Gradient by cross-correlation imaging condition; (c) Gradient by deconvolution imaging condition. The source location is $(3.75km, 0.025km)$.

The damping factor λ for the stabilized deconvolution imaging condition is 0.01. It can be seen that the amplitudes of the gradient based on deconvolution imaging condition is stronger than that of the gradient based on cross-correlation imaging condition, especially for the deep reflectors.

Fig.10a and b are the normalized gradients with 18 sources based on cross-correlation imaging condition and deconvolution imaging condition respectively. It is obvious for us to observe that the amplitudes of the deconvolution based gradient are stronger than that of the cross-correlation based gradient. Then we can transform the gradient to impedance perturbation following equation (20) and update the initial velocity model iteratively using the model perturbation following equation (21). The estimated impedance perturbations are shown in Fig.11a and b. The inverted velocity models after the first iteration are shown in Figure 12. We can see that the updated velocity models have a big improvement comparing with the initial velocity model just after first iteration in FWI. Furthermore, the deep parts of the inverted velocity model based on deconvolution imaging condition are much better than that based on cross-correlation imaging condition. Fig. 13 shows the inverted velocities of

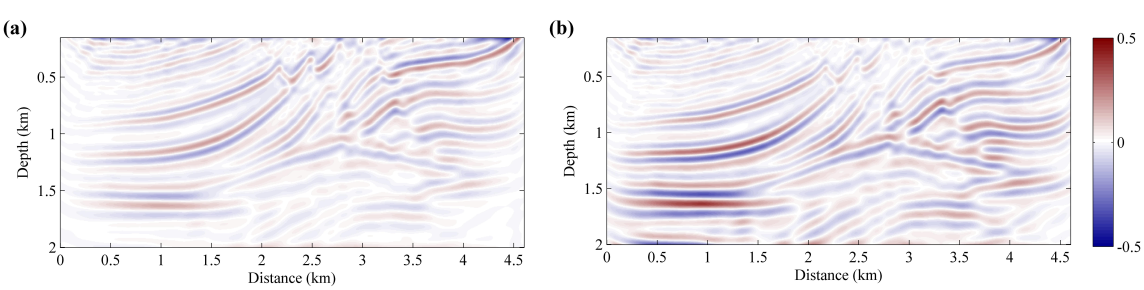

FIG. 10. (a) Gradient by cross-correlation imaging condition; (b) Gradient by deconvolution imaging condition.

4 well logs at the location of $0.5km$, $1.5km$, $2.5km$ and $3.5km$ respectively. Even though it is not very obvious, we can still recognize that the inverted velocities by deconvolution imaging condition (green lines) are more close to the true velocities than that of cross-

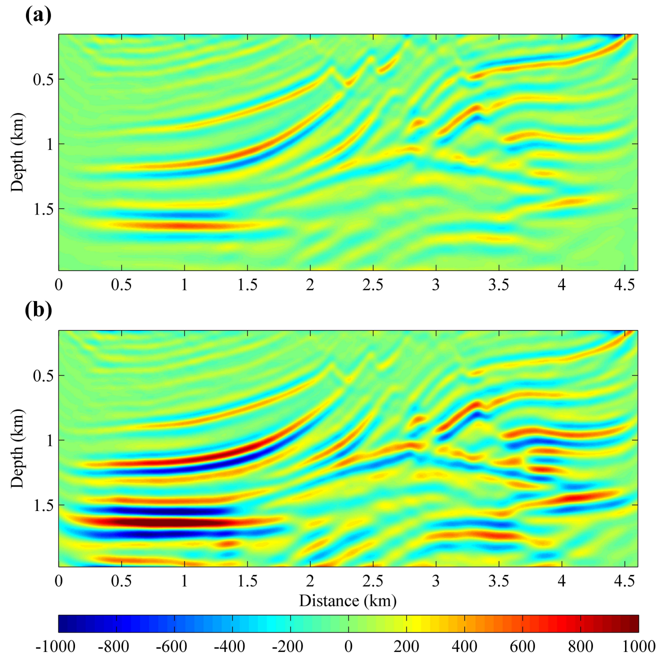

FIG. 11. (a) Impedance perturbation by cross-correlation imaging condition; (b) Impedance perturbation by deconvolution imaging condition.

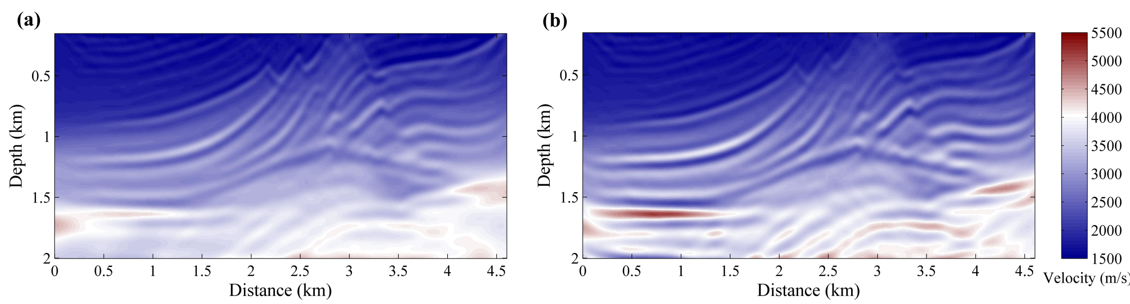

FIG. 12. (a) Inverted velocity model by cross-correlation imaging condition; (b) Inverted velocity model by deconvolution imaging condition.

correlation imaging condition (black lines).

Finally, we applied the Iterative Modelling Migration and Inversion (IMMI) method on the modified Marmousi model using phase encoded gradient and phase encoded pseudo-Hessian (Pan et al., 2013a, 2014). The true velocity model and initial velocity model are shown in Fig.14a and b. We constructed the phase encoded gradient and preconditioning the phase encoded gradient using the diagonal part of the phase encoded pseudo-Hessian is equivalent to the deconvolution imaging condition. We varied the ray parameters regularly for different iterations for reducing the computational cost and balancing the update iteratively. The impedance perturbation is estimated using the traditional impedance inversion method. We also employed the multiscale approach with increasing the frequency band from $[0Hz, 5Hz]$ to $[0Hz, 32Hz]$ with a step of $3Hz$ every 10 iterations. Fig.14c shows the inverted velocity model after 100 iterations. We can see that the structures of the velocity model are reconstructed very well. And the iterative inversion process is stable and

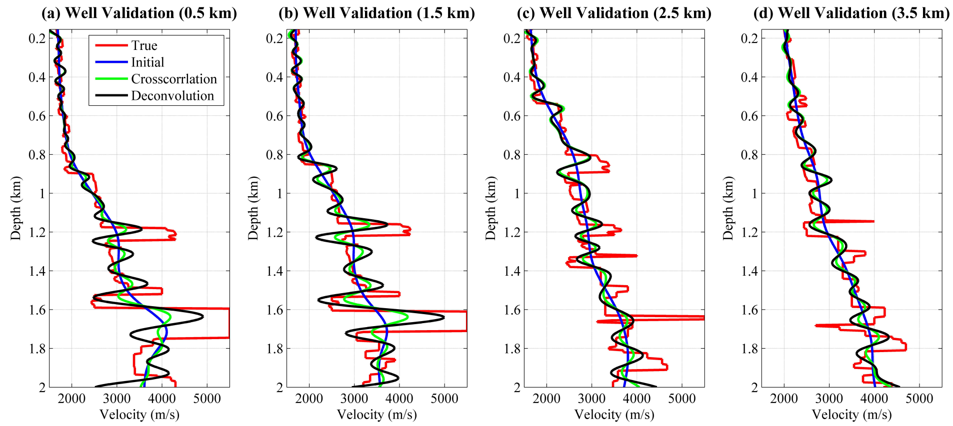

FIG. 13. Comparison of the true velocity (red), initial velocity (blue), and inverted velocity based on cross-correlation imaging condition (green) and deconvolution imaging condition (black) at the location of $0.5km$ (a), $1.5km$ (b),$2.5km$ (c) and $3.5km$ (d) respectively.

efficient, which proves the effectiveness of the IMMI method. Fig.15a, b and c compare the true velocity model (red lines), initial velocity model (black lines) and inverted velocity model (blue lines) at $0.5km$, $1km$ and $3km$ respectively. It can been that the inverted velocity model match the true velocity model very well.

CONCLUSION

In this research, we analyzed the cross-correlation imaging condition and deconvolution imaging condition for RTM and we also gave an analytic solutions for the two imaging conditions. And the deconvolution imaging condition actually transforms the unit of the gradient from the square of the amplitude to the reflectivity, which enables us to use traditional impedance inversion method to estimate the impedance perturbation. Furthermore, we found that the source illumination, the auto-correlation of the forward modeling wavefields, is actually equivalent to the diagonal part of the pseudo-Hessian. The phase encoded gradient and phase encoded pseudo-Hessian are introduced to reduce the computational cost and preconditioning the phase encoded gradient using the diagonal part of the phase encoded pseudo-Hessian is also equivalent to the deconvolution imaging condition. So, we implemented the IMMI method with phase encoded gradient and phase encoded pseudo-Hessian and applied this strategy on a modified Marmousi model. The reconstructed velocity model using IMMI method matches the true velocity very well, which proves the effectiveness of the IMMI method.

ACKNOWLEDGMENTS

This research was supported by the Consortium for Research in Elastic Wave Exploration Seismology (CREWES). And I would like to thank Yu Zhang (CGGveritas, now ConocoPhillips)for his help in understanding the imaging condition.

REFERENCES

Ben-Hadj-Ali, H. S., Operto, S., and Virieux, J., 2011, An efficient frequency-domain full waveform inversion method using simultaneous encoded sources: Geophysics, **76**, WCC177–WCC188.

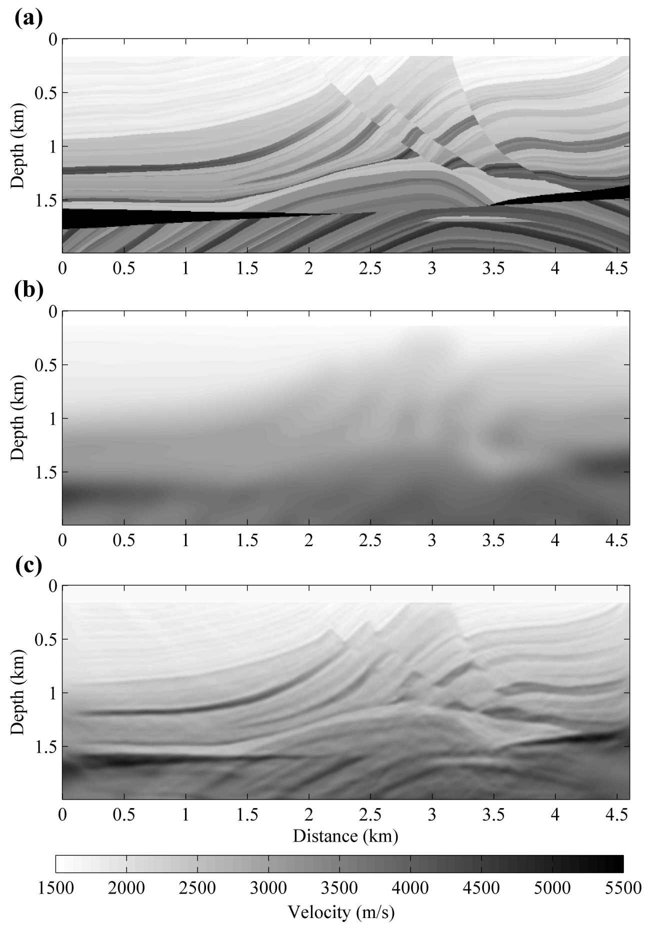

FIG. 14. (a) and (b) show the true velocity model and initial velocity model respectively. And (c) is the inverted velocity model after 100 iterations using the IMMI method.

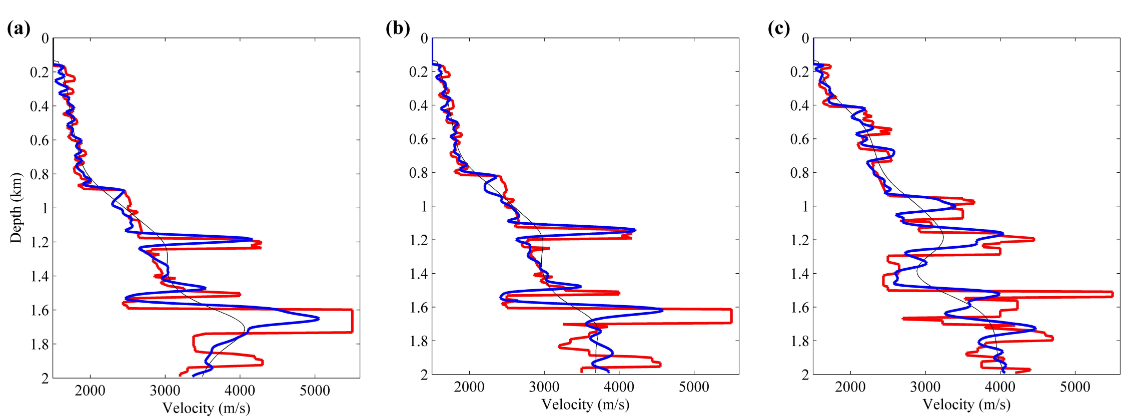

FIG. 15. Comparison of the true velocity model (bold red lines), the initial velocity model (black lines) and the inverted velocity model (bold blue lines) for the wells located at $0.5km$, $1.0km$ and $3.0km$.

Chattopadhyay, S., and McMechan, G. A., 2008, Imaging conditions for prestack reverse-time migration: Geophysics, **73**, S81–S89.

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, 467–481.

Gao, F., Williamson, R., and Houllevigue, H., 2012, Full waveform inversion by deconvolution gradient method, 1–5.

Gauthier, O., Virieux, J., and Tarantola, A., 1986, Two-dimensional nonlinear inversion of seismic waveforms: numerical results: Geophysics, **51**, 1387–1403.

Guitton, A., Valencinao, A., Bevc, D., and Claerbout, J. F., 2007, Smoothing imaging condition for shot-profile migration: Geophysics, **72**, S149–S154.

Kaelin, B., and Guitton, A., 2006, Imaging conditions for reverse time migration, 2594–2598.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migration, 206–220.

Margrave, G. F., Ferguson, R. J., and Hogan, C. M., 2011, Full waveform inversion using wave-equation depth migration with tying to wells, 2454–2458.

Margrave, G. F., Innanen, K. A., and Yedlin, M., 2012a, A prospective on full waveform inversion: CREWES Research Report.

Margrave, G. F., Innanen, K. A., and Yedlin, M., 2012b, A prospective on full-waveform inversion, 1–18.

Pan, W., Innanen, K. A., and Margrave, G. F., 2013a, Efficient full waveform inversion with phase encoded pseudo-hessian, 1–15.

Pan, W., Innanen, K. A., and Margrave, G. F., 2013b, Efficient pseudo gauss-newton full waveform inversion in the time-ray parameter domain, 1–36.

Pan, W., Innanen, K. A., and Margrave, G. F., 2013c, A comparison of different scaling methods for least-squares inversion/migration, 1–15.

Pan, W., Innanen, K. A., and Margrave, G. F., 2014, Efficient full waveform inversion in the time-ray parameter domain using iteration-dependent sets of ray parameters, 1–5.

Pica, A., Diet, J. P., and Tarantola, A., 1990, Nonlinear inversion of seismic reflection data in a laterally invariant medium: Geophysics, **55**, 284–292.

Poole, T. L., Crutis, C., Robertson, J. O. A., and Mannen, D., 2010, Deconvolution imaging conditions and cross-talk suppression: Geophysics, **75**, W1–W12.

Pratt, R. G., Shin, C., and Hicks, G. J., 1998, Gauss-newton and full newton methods in frequency-space seismic waveform inversion: Geophysical Journal International, **133**, 341–362.

Scheleicher, J., Costa, J. C., and Novais, A., 2008, A comparison of imaging conditions for wave equation shot-profile migration: Geophysics, **73**, S219–S227.

Shin, C., Jang, S., and Min, M., 2001a, Improved amplitude preservation for prestack depth migration by inverse scattering theory: Geophysical Prospecting, **49**, 592–606.

Tang, Y., 2009, Target-oriented wave-equation least-squares migration/inversion with phase-encoded hessian: Geophysics, **74**, WCA95–WCA107.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

Virieux, A., and Operto, S., 2009, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **74**, WCC1–WCC26.

Zhang, Y., Sun, J., Notfors, C., Gray, S. H., Chernis, L., and Young, J., 2005, Delayed-shot 3d depth migration: Geophysics, **70**, E21–E28.