

1D scalar full waveform inversion inferring convergence properties with analytic and numerical examples

Wenyong Pan, Kris Innanen

ABSTRACT

Formulated as a least-squares form, Full waveform inversion (FWI) seeks to minimize the difference between the modeling data and the observed data and estimate the subsurface parameters. It has been widely studied in recent years, but some problems still remain to be addressed. In this research, we performed the analytic analysis of 1D scalar FWI. The analysis to this simplest condition can help us achieve some new ideas and discoveries in FWI. A simple two-interface model and a homogeneous background model are used as the true velocity model and initial velocity model respectively. And two iterations are performed for analysis based on some optimal assumptions. We found that: (1) after the first iteration, the placement error at the second interface is influenced by the velocity contrast and interfaces distance; (2) after the second iteration, the placement error at the second interface become smaller for small velocity contrast, but may become larger for large velocity contrast; (3) and the noises produced in the cross-correlation have a negative influence to the amplitude recovery of the second interface, which will decrease the convergence rate of FWI; (4) but the noises have no significant influence to the placement error of the second interface.

GENERAL PRINCIPLE OF FULL WAVEFORM INVERSION

As a least-squares local optimization, full waveform inversion seeks to minimize the difference between the synthetic data and observed data (Lailly, 1983; Tarantola, 1984) and update the model iteratively. The misfit function ϕ is given in a least-squares norm:

$$\phi \left(s_0^{(n)}(\mathbf{r}) \right) = \frac{1}{2} \int d\omega \left(\sum_{\mathbf{r}_s} \sum_{\mathbf{r}_g} \|\delta P \left(\mathbf{r}_g, \mathbf{r}_s, \omega | s_0^{(n)}(\mathbf{r}) \right)\|_2 \right), \quad (1)$$

where $s_0^{(n)}(\mathbf{r}) = \frac{1}{(c_0^{(n)}(\mathbf{r}))}$ are the model parameters, the square of the slowness in the n th iteration, and $c_0^{(n)}(\mathbf{r})$ is the velocity. δP mean the data residuals, the difference between the observed data $P \left(\mathbf{r}_g, \mathbf{r}_s, \omega | s_0^{(n)} \right)$ and synthetic data $G \left(\mathbf{r}_g, \mathbf{r}_s, \omega | s_0^{(n)} \right)$, $\|\cdot\|_2$ indicates the $\ell - 2$ norm.

The minimum value of the misfit function is sought in the vicinity of the starting model $s_0(\mathbf{r})$ and the updated model can be written as the sum of the starting model and a model perturbation $\delta s_0^{(n)}(\mathbf{r})$ (Virieux and Operto, 2009).

$$s^{(n)}(\mathbf{r}) = s_0^{(n)}(\mathbf{r}) + \mu^{(n)} \delta s_0^{(n)}(\mathbf{r}), \quad (2)$$

where $\mu^{(n)}$ is the step length in n th iteration, which is a scalar constant used to scale the model perturbation and can be obtained through a line search method (Gauthier et al., 1986; Pica et al., 1990).

Applying a second order Taylor-Lagrange development of the misfit function and then taking partial derivative with respect to the model parameters give the model perturbation as:

$$\delta s_0^{(n)} = - \int d\mathbf{r}' H^{(n)-}(\mathbf{r}, \mathbf{r}') g^{(n)}(\mathbf{r}), \quad (3)$$

where $g^{(n)}(\mathbf{r}')$ is the gradient and $H^{(n)-}(\mathbf{r}', \mathbf{r})$ is the Hessian matrix.

For gradient, the first order derivative of the misfit function ϕ with respect to the model parameters can be obtained by a zero-lag correlation between the data residuals and the first order partial derivative wavefields. Then apply a perturbation derivation based on the Born approximation, the sensitive matrix can be written as:

$$\frac{\delta G(\mathbf{r}_g, \mathbf{r}_s, \omega | s_0^{(n)})}{\delta s_0^{(n)}(\mathbf{r})} = -\omega^2 G(\mathbf{r}_g, \mathbf{r}, \omega | s_0^{(n)}) G(\mathbf{r}, \mathbf{r}_s, \omega | s_0^{(n)}), \quad (4)$$

Then we can get the gradient:

$$g^{(n)}(\mathbf{r}) = \sum_{\mathbf{r}_s, \mathbf{r}_g} \int d\omega \Re(\omega^2 \mathcal{F}_s(\omega) G(\mathbf{r}, \mathbf{r}_s, \omega) G(\mathbf{r}_g, \mathbf{r}, \omega) \delta P^*), \quad (5)$$

where $G(\mathbf{r}, \mathbf{r}_s, \omega)$ and $G(\mathbf{r}_g, \mathbf{r}, \omega)$ are the source-side and receiver-side Green's functions, respectively. And $\mathcal{F}_s(\omega)$ is the source signature. Then 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, which avoids the direct computation of the partial derivative wavefields.

ANALYTIC ANALYSIS OF FWI—1ST ITERATION

Having introduced the basic theory and framework of FWI, we can analyze how FWI works beginning with the simplest condition: analytic example of 1D scalar FWI. In this simplest possible problem, the initial or reference model is a homogeneous model with the scalar velocity of C_0 , as shown in Figure 1. And the gradient based method is used to reconstruct a 1D true velocity model with two interfaces located at the depths of Z_1 and Z_2 and the velocity varies from C_0 to C_1 and C_2 at interfaces of Z_1 and Z_2 respectively, as shown in Figure 1. Through the derivation and analysis, it is possible for us to understand how well and how quickly FWI does its job and address the current problems for FWI. So, let's examine FWI in the first iteration.

For a gradient based method, we use the form of $s^{(n)}(\mathbf{r}) = s_0^{(n)}(\mathbf{r}) + \mu^{(n)} g^{(n)}(\mathbf{r})$ to update the velocity model, where $\mu^{(n)}$ is the step length, a constant value, which traditionally determined through a line search method. The gradient for this 1D scalar problem can be written as:

$$g(z) = \sum d\omega \Re(G(z, 0, \omega) G(z, 0, \omega) \delta P^*(0, 0, \omega)), \quad (6)$$

As indicated by equation (6), to obtain the gradient, we need to calculate two Green's functions of $G(z, 0, \omega)$ and $G(0, z, \omega)$ and the complex conjugate of data residual δP^* . We

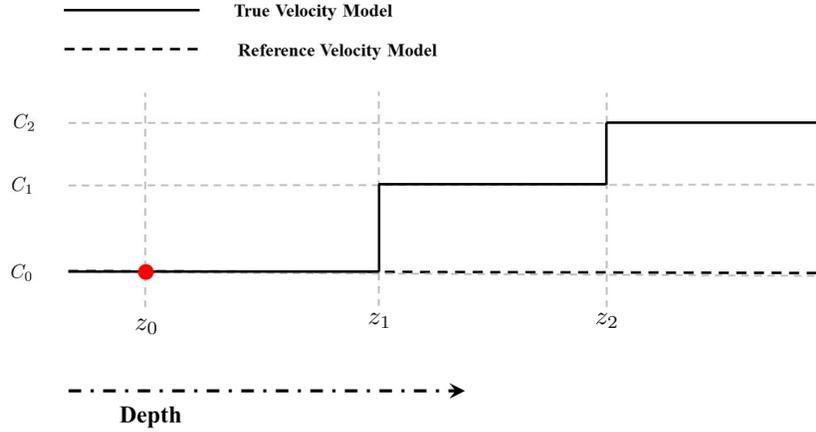


FIG. 1. The velocity models used in the first iteration of FWI. The dash line is the homogeneous reference model and the solid line is the true velocity model.

know that the Green's functions in the homogeneous medium can be written as:

$$G(z, 0, \omega) = G(0, z, \omega) = \frac{e^{ik_0 z}}{i2k_0}, \quad (7)$$

The synthetic data in the reference model only includes the direct wave. So, it can be

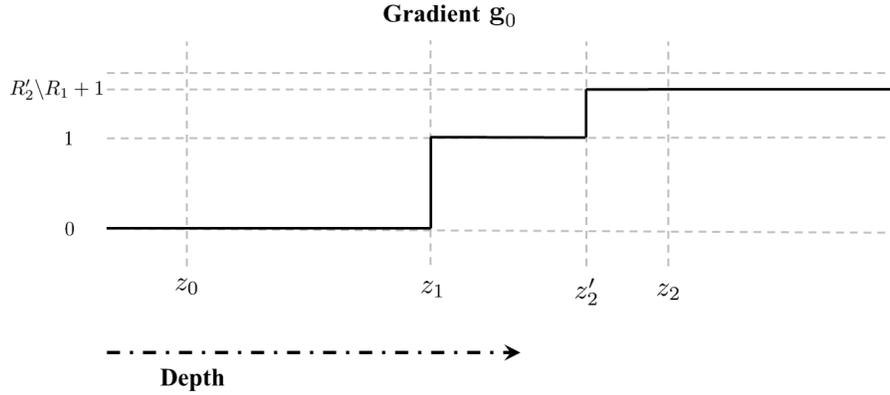


FIG. 2. The gradient \mathbf{g}_0 for the first iteration.

written as $G(0, 0, \omega) = \frac{1}{i2k_0}$, where $k_0 = \frac{\omega}{C_0}$ is the wavenumber and ω is the angular frequency. While for the data $D(\omega)$ observed in the real velocity model, it includes the directive wave and the reflections from the two interfaces. Then, we have:

$$D(\omega) = \frac{1}{i2k_0} + R_1 \frac{e^{i2k_0 z_1}}{i2k_0} + R_2' \frac{e^{i2k_0 z_1}}{i2k_0} e^{i2k_1(z_2 - z_1)}, \quad (8)$$

where $R_1 = \frac{C_1 - C_0}{C_1 + C_0}$ is the reflection coefficient at the first interface and $R_2' = (1 - R_1^2) R_2$ is the reflection coefficient at the second interface and $R_2 = \frac{C_2 - C_1}{C_2 + C_1}$. So, we can get the complex conjugate of the data residual as:

$$\delta P^*(0, 0, \omega) = [D(\omega) - G(0, 0, \omega)]^* = -R_1 \frac{e^{-2k_0 z_1}}{i2k_0} - R_2' \frac{e^{-i2k_0 z_1}}{i2k_0} e^{-i2k_1(z_2 - z_1)}, \quad (9)$$

Then the full gradient can be given as:

$$g(z) = - \int d\omega \omega^2 \frac{e^{i2k_0 z}}{i2k_0^2} \left\{ R_1 \frac{e^{-i2k_0 z_1}}{i2k_0} + R_2' \frac{e^{i2k_0 z_1}}{i2k_0} e^{-i2k_1(z_2-z_1)} \right\}, \quad (10)$$

$$= - \frac{R_1 C_0^2}{4} \int d\omega \frac{e^{i2k_0(z-z_1)}}{i2k_0} + \frac{R_2' C_0^2}{4} \int d\omega \frac{e^{i2k_0(z-Z_2')}}{i2k_0}.$$

where $Z_2' = z_1 + \frac{k_1}{k_0}(z_2 - z_1)$. Then we replace $d\omega$ with $\frac{C_0}{2} \times d(2k_0)$ and we have:

$$g(z) = \frac{R_1 C_0^3}{8} \int d(2k_0) \frac{e^{i2k_0(z-z_1)}}{i2k_0} + \frac{R_2' C_0^3}{8} \int d(2k_0) \frac{e^{i2k_0(z-Z_2')}}{i2k_0}, \quad (11)$$

$$= \frac{\pi R_1 C_0^3}{4} H(z - z_1) + \frac{\pi R_2' C_0^3}{4} H(z - Z_2').$$

where H means the Heaviside function. We notice that the gradient we get consists of two Heaviside functions (step functions), which are like the two interfaces missed in the reference model. If we add the gradient with appropriate scaling, it will be a great improvement for the velocity reconstruction. We can normalize gradient by dividing $\frac{\pi R_1 C_0^3}{4}$, then we get:

$$g_0(z) = H(z - z_1) + \frac{R_2'}{R_1} H(z - Z_2'), \quad (12)$$

The gradient g_0 is shown in Figure 2. We also notice that the placement of the first step provided by the gradient is consistent with the placement of the first interface in the true velocity model, which is z_1 . While the placement of the second step Z_2' is not right. And the error E between the right placement z_2 and the wrong placement is:

$$E = z_2 - Z_2' = \left(1 - \frac{C_0}{C_1}\right) (z_2 - z_1), \quad (13)$$

We can see that the value of E is negative and the larger the difference between C_0 and C_1 or z_2 and z_1 , the larger this error becomes.

For scaling the gradient, here we can set the step length to make the first interface of the updated velocity model match the real velocity model. In this condition, the step length becomes $\mu = \left(\frac{1}{C_1^2} - \frac{1}{C_0^2}\right)$ and model perturbation becomes:

$$\delta s_0^{(n)} = \left(\frac{1}{C_1^2} - \frac{1}{C_0^2}\right) H(z - z_1) + \left(\frac{1}{C_1^2} - \frac{1}{C_0^2}\right) \frac{R_2'}{R_1} H(z - Z_2'), \quad (14)$$

Then we can add this model perturbation to reference model directly for getting the updated model in the first iteration. In the new updated model, the placement of the first interface is $Z_2' = z_1 + \frac{k_1}{k_0}(z_2 - z_1)$ and the velocity of third layer becomes $C_2' = C_1 + (C_1 - C_0) \frac{R_2'}{R_1} = C_1 + (C_1 + C_0) R_2'$. Because $C_2' - C_2 < 0$, the updated velocity model after the first iteration is as shown in Figure 3. From Figure 2, it can be seen that the updated velocity model after first iteration match the true velocity model well at the first interface. While the velocity and placements at the second interface are both not right. So, how does FWI works for the next iteration? We can examine the second iteration of FWI using a similar method.

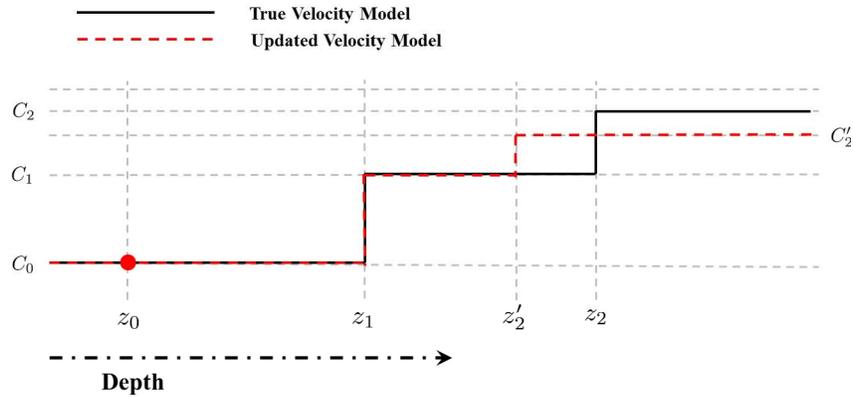


FIG. 3. Updated velocity model after first iteration as indicated by the red-dash line.

ANALYTIC ANALYSIS OF FWI—2ND ITERATION

Then we can go head to see what will happen for the second iteration. To calculate the gradient, we also need to calculate the Green’s functions and data residuals. While comparing with the Green’s functions in the homogeneous medium, the Green’s functions in different parts of the new updated velocity model are different, as shown in Figure 3. We can analyze the gradient case by case. In order to make it easier to understand, we transform the 1D velocity to a 2D velocity model for analyzing the Green’s functions and data residuals.

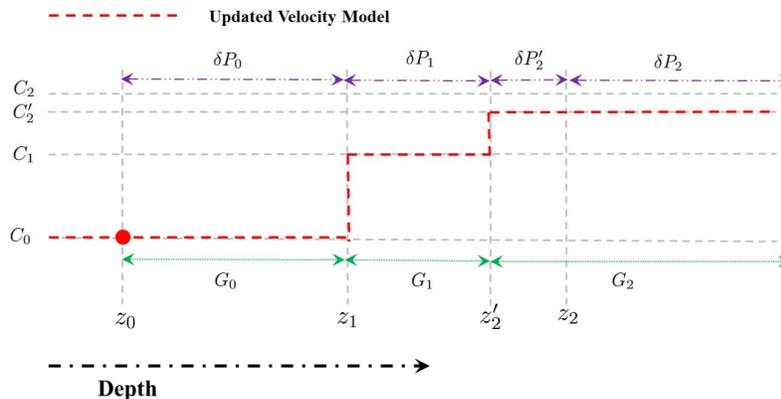
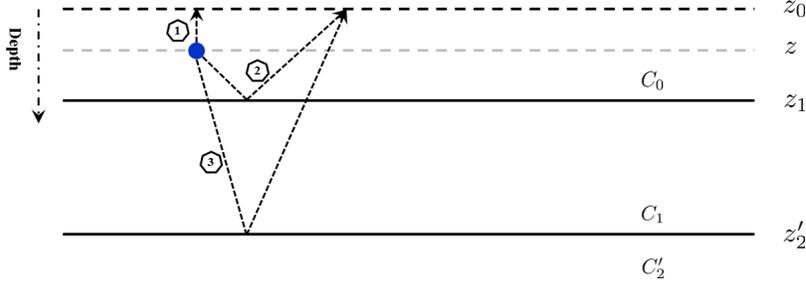


FIG. 4. Green’s functions and data residuals in the second iteration.

Case 1 when $z_1 > z > z_0$

Firstly, for G_0 , when $z_1 > z > z_0$, as shown in Figure 4, we have:

$$\begin{aligned}
 G_0(z, 0, \omega) &= G(0, z, \omega) = G_0^1(z, 0, \omega) + G_0^2(z, 0, \omega) + G_0^3(z, 0, \omega), \\
 &= \frac{e^{ik_0 z}}{i2k_0} + R_1 \frac{e^{ik_0(2z_1-z)}}{i2k_0} + R_2' \frac{e^{ik_0(2z_1-z)}}{i2k_0} e^{ik_1(Z_2'-z_1)}.
 \end{aligned} \tag{15}$$

FIG. 5. Green's functions when $z_1 > z > z_0$.

where R_1 is the reflection coefficient at Z_1 and $R_2'' = (1 - R_1^2)R_2'''$, where $R_2''' = \frac{C_2' - C_1}{C_2' + C_1}$ is the reflection coefficient at Z_2' . The observed data and synthetic data can be written as:

$$D_{obs}(\omega) = \frac{1}{i2k_0} + R_1 \frac{e^{i2k_0 z_1}}{i2k_0} + R_2' \frac{e^{i2k_0 z_1}}{i2k_0} e^{i2k_1(z_2 - z_1)}, \quad (16)$$

$$D_{syn}(\omega) = \frac{1}{i2k_0} + R_1 \frac{e^{i2k_0 z_1}}{i2k_0} + R_2'' \frac{e^{i2k_0 z_1}}{i2k_0} e^{i2k_1(Z_2' - z_1)}, \quad (17)$$

And the complex conjugate of the data residual is:

$$\delta P_2^*(0, 0, \omega) = [D_{obs}(\omega) - D_{syn}(\omega)]^* = -R_2' \frac{e^{-i2k_0 z_1}}{i2k_0} e^{-i2k_1(z_2 - z_1)} + R_2'' \frac{e^{-i2k_0 z_1}}{i2k_0} e^{-i2k_1(Z_2' - z_1)}, \quad (18)$$

Then, we can set the gradient as:

$$g(z) = \int d\omega \omega^2 A \delta P^*(0, 0, \omega), \quad (19)$$

where

$$A = \frac{e^{i2k_0 z}}{4k_0^2} \left\{ 1 + R_1^2 e^{i2k_0(2z_1 - 2z)} + (R_2''')^2 e^{i2k_0(2z_1 - 2z)} e^{i2k_1(2Z_2' - 2z_1)} \right\} \\ - \frac{e^{i2k_0 z}}{4k_0^2} \left\{ 2R_1 R_2'' e^{i2k_0(2z_1 - 2z)} e^{i2k_1(2Z_2' - 2z_1)} + 2R_1 e^{i2k_0(2z_1 - 2z)} + 2R_2'' e^{-i2k_0(z_1 - z)} e^{i2k_1(Z_2' - z_1)} \right\},$$

and the complex conjugate of the data residual is:

$$\delta P^*(0, 0, \omega) = -R_2' \frac{e^{-i2k_0 z_1}}{i2k_0} e^{-i2k_1(z_2 - z_1)} + R_2'' \frac{e^{-i2k_0 z_1}}{i2k_0} e^{-i2k_1(Z_2' - z_1)},$$

And the gradient becomes:

$$\begin{aligned}
 g(z) = & \frac{1}{4k_0^2} \int d\omega \omega^2 \frac{1}{i2k_0} \left(R_2' e^{i2k_0 \left(z - \left(z_1 + \frac{k_1}{k_0} (z_2 - z_1) \right) \right)} + R_2' R_1^2 e^{-i2k_0 \left(z - \left(z_1 - \frac{k_1}{k_0} (z_2 - z_1) \right) \right)} \right) \\
 & + \frac{1}{4k_0^2} \int d\omega \omega^2 \frac{1}{i2k_0} \left(R_2' (R_2'')^2 e^{-i2k_0 \left(2 - \left(z_1 + \frac{k_1}{k_0} (2Z_2' - z_2 - z_1) \right) \right)} \right) \\
 & + \frac{1}{4k_0^2} \int d\omega \omega^2 \frac{1}{i2k_0} \left(2R_1 R_1' R_2'' e^{-i2k_0 \left(z - \left(z_1 + \frac{k_1}{k_0} (Z_2' - z_2) \right) \right)} \right) \\
 & + \frac{1}{4k_0^2} \int d\omega \omega^2 \frac{1}{i2k_0} \left(2R_2' R_1 e^{-i2k_1 (z_2 - z_1)} + R_2' R_2'' e^{-i2k_1 (Z_2' - z_2)} \right) \\
 & - \frac{1}{4k_0^2} \int d\omega \omega^2 \frac{1}{i2k_0} \left(R_2'' e^{i2k_0 \left(z - \left(z_1 + \frac{k_1}{k_0} (Z_2' - z_1) \right) \right)} \right) \\
 & + \frac{1}{4k_0^2} \int d\omega \omega^2 \frac{1}{i2k_0} \left(-R_2'' R_1^2 e^{-i2k_0 \left(z - \left(z_1 + \frac{k_1}{k_0} (Z_2' - z_1) \right) \right)} - (R_3'')^3 e^{-i2k_0 \left(z - \left(z_1 + \frac{k_1}{k_0} (Z_2' - z_1) \right) \right)} \right) \\
 & + \frac{1}{4k_0^2} \int d\omega \omega^2 \frac{1}{i2k_0} \left(2R_1 (R_2'')^2 e^{-i2k_0 (z - z_1)} - 2R_1 R_2'' e^{-i2k_1 (Z_2' - z_1)} - 2(R_2'')^2 \right), \tag{20}
 \end{aligned}$$

Because the reference model and the true velocity model match well when $z_1 > z > z_0$, the best condition is that the gradient should be 0. While something actually exists. These things are the low frequencies produced in the process of cross-correlation. When cross-correlating the forward modeling wave-fields and back-propagated wave-fields, the backscattered waves can also meet the cross-correlation imaging conditions. They are not the real gradient what we want. They are the low frequencies noise just like those in reverse time migration. Most of them are weak along the wave path but also produce influence to the gradient. So, we can neglect this part temporarily for analysis convenience.

Case 2 when $Z_2' > z > z_1$

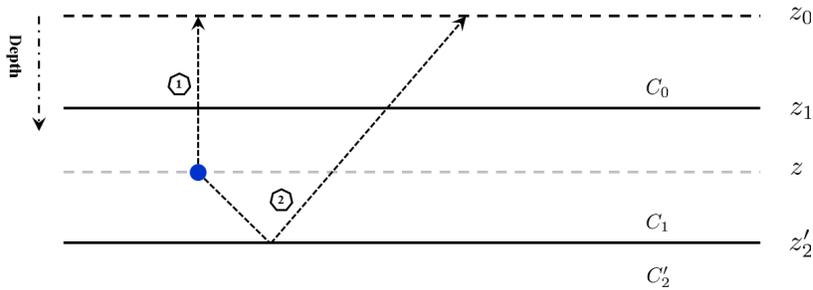


FIG. 6. Green's functions when $Z_2' > z > z_1$.

The second condition is that when $Z_2' > z > z_1$. We can get the Green's functions as:

$$G(z, 0, \omega) = T_{10} \frac{e^{i6k_1(z - z_1)}}{i2k_1} e^{ik_0 z_1} + R_2''' T_{10} \frac{ik_1(2Z_2' - z_1 - z)}{i2k_1} e^{ik_0 z_1}, \tag{21}$$

$$G(0, z, \omega) = T_{01} \frac{e^{ik_1(z - z_1)}}{i2k_1} e^{ik_0 z_1} + T_{01} R_2''' \frac{e^{ik_1(2Z_2' - z_1 - z)}}{i2k_1} e^{ik_0 z_1}, \tag{22}$$

The data residual is as listed in equation (21). Here, we can also neglect the terms with high order reflection coefficients. The gradient is:

$$\begin{aligned}
 g(z) &= \int d\omega \omega^2 G(z, 0, \omega) G(0, z, \omega) \delta P^*(0, 0, \omega) \\
 &= T_{01} T_{10} \frac{1}{4k_1^2} \int d\omega \omega^2 \left(R_2' \frac{e^{i2k_1(z-z_2)}}{i2k_0} - R_2'' \frac{e^{i2k_1(z-Z_2')}}{i2k_0} + R_2'' R_2' \frac{e^{i2k_1(Z_2'-z_2)}}{i2k_0} \right) \\
 &+ T_{01} T_{10} \frac{1}{4k_1^2} \int d\omega \omega^2 \left(-(R_2'')^2 \frac{1}{i2k_0} + R_2'' \frac{i2k_1(Z_2'-z_2)}{i2k_0} - R_2^2 R_2''' \frac{e^{i2k_1(Z_2'-z)}}{i2k_0} \right).
 \end{aligned} \tag{23}$$

Case 3 when $z > Z_2$

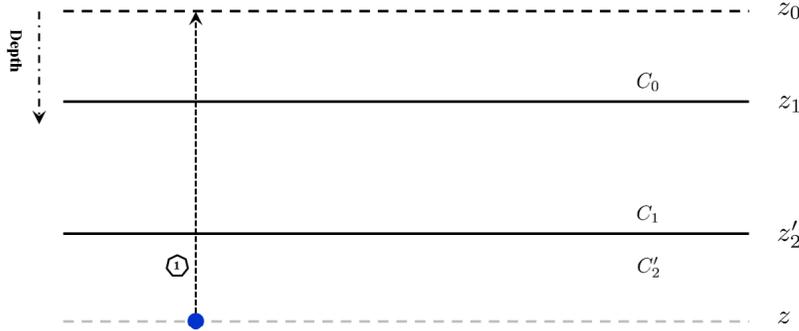


FIG. 7. Green's functions when $z > Z_2$.

The third condition is when $z > Z_2$. As shown in Figure 6, the Green's functions only consist of one component. So, we can write the Green's functions as:

$$G_2(z, 0, \omega) = T_{21} T_{10} \frac{e^{ik_2(z-Z_2')}}{i2k_2} e^{ik_1(Z_2'-z_1)} e^{ik_0 z_1}, \tag{24}$$

$$G_2(0, z, \omega) = T_{01} T_{12} \frac{e^{ik_2(z-Z_2')}}{i2k_2} e^{ik_1(Z_2'-z_1)} e^{ik_0 z_1}, \tag{25}$$

where $T_{21} = \frac{2C_1}{C_1+C_2}$, $T_{10} = \frac{2C_0}{C_1+C_0}$, $T_{01} = \frac{2C_1}{C_1+C_0}$ and $T_{12} = \frac{2C_2}{C_1+C_2}$ are the transmission coefficients. The observed data and synthetic data can be written as:

$$D_{obs}(\omega) = \frac{1}{i2k_0} + R_1 \frac{e^{i2k_0 z_1}}{i2k_0} + R_2' \frac{e^{i2k_0 z_1}}{i2k_0} e^{i2k_1(z_2-z_1)}, \tag{26}$$

$$D_{syn}(\omega) = \frac{1}{i2k_0} + R_1 \frac{e^{i2k_0 z_1}}{i2k_0} + R_2'' \frac{e^{i2k_0 z_1}}{i2k_0} e^{i2k_1(Z_2'-z_1)}, \tag{27}$$

And the complex conjugate of the data residual is:

$$\delta P_2^*(0, 0, \omega) = [D_{obs}(\omega) - D_{syn}(\omega)]^* = -R_2' \frac{e^{-i2k_0 z_1}}{i2k_0} e^{-i2k_1(z_2-z_1)} + R_2'' \frac{e^{-i2k_0 z_1}}{i2k_0} e^{-i2k_1(Z_2'-z_1)}, \tag{28}$$

Then, we can get the gradient:

$$\begin{aligned}
 g(z) &= \int d\omega \omega^2 G(z, 0, \omega) G(0, z, \omega) \delta P^*(0, 0, \omega) \\
 &= -T_{21} T_{10} T_{01} T_{12} \frac{1}{4k_2^2} \int d\omega \omega^2 \left(e^{i2k_2(z-Z'_2)} e^{i2k_1(Z'_2-z_1)} e^{i2k_0 z_1} \right) \\
 &\quad \times \left(-R'_2 \frac{e^{-i2k_0 z_1}}{i2k_0} e^{-i2k_1(z_2-z_1)} + R''_2 \frac{e^{-i2k_0 z_1}}{i2k_0} e^{-i2k_1(Z'_2-z_1)} \right).
 \end{aligned} \tag{29}$$

Then after a series of derivation, the gradient becomes:

$$\begin{aligned}
 g(z) &= -T_{21} T_{10} T_{01} T_{12} \frac{1}{4k_2^2} \int d\omega \omega^2 \left(-R'_2 \frac{e^{i2k_2(z-Z'_2)}}{i2k_0} e^{i2k_1(Z'_2-z_2)} + R''_2 \frac{e^{i2k_2(z-Z'_2)}}{i2k_0} \right) \\
 &= T_{21} T_{10} T_{01} T_{12} R'_2 \frac{\omega^2 C'_2}{8k_2 k_0} \int (2k_2) \left(\frac{e^{i2k_2(z-(Z'_2-\frac{k-1}{k_2}(Z'_2-z_2)))}}{i2k_2} \right) \\
 &\quad - T_{21} T_{10} T_{01} T_{12} R''_2 \frac{\omega^2 C'_2}{8k_2 k_0} \int (2k_2) \frac{e^{i2k_2(z-Z'_2)}}{i2k_2}.
 \end{aligned} \tag{30}$$

We can notice that gradient can be written as a summation of two Heaviside functions:

$$g(z) = T_{21} T_{10} T_{01} T_{12} \frac{\pi R'_2 C_0 (C'_2)^2}{4} H\left(z - \left(Z'_2 - \frac{C'_2}{C_1} (Z'_2 - z_2)\right)\right) - T_{21} T_{10} T_{01} T_{12} \frac{\pi R'_2 C_0 (C'_2)^2}{4} H(z - Z'_2). \tag{31}$$

We can conclude that the gradient is proportional to:

$$g(z) \propto -H(z - Z'_2) + \frac{R'_2}{R''_2} = -H(z - Z'_2) + \frac{R_2}{R'_2}, \tag{32}$$

where $Z''_2 = \left(Z'_2 - \frac{C'_2}{C_1} (Z'_2 - z_2)\right)$, $R''_2 = (1 - R_1^2) R_2'''$ and $R_2''' = \frac{C_2 - 2C_1 - C_0}{C_2 + C_1}$. We can see that the first step in gradient is negative and while the second step is positive. And the placement of the first step is right, while the placement of the second step is wrong. And the placement error E' is:

$$E' = z_2 - Z''_2 = \left(1 - \frac{C'_2}{C_1}\right) \left(1 - \frac{C_0}{C_1}\right) (z_2 - z_1), \tag{33}$$

we can see that this equation is similar to equation (16). And the placement error of the second step is influenced by the differences of C'_2 and C_1 , C_0 and C_1 , z_1 and z_2 . We also note that for equation (16) $E > 0$, which means that the placement of inverted second interface is between z_1 and z_2 ($z_1 < Z''_2 < z_2$). While for E' , it is negative ($E' < 0$), so the $Z''_2 > z_2$. How about the difference of E' and E ? We can compare $\left(|1 - \frac{C'_2}{C_1}| - 1\right)$ with 0. Then,

$$\left(|1 - \frac{C'_2}{C_1}| - 1\right) = \left(\frac{4C_0}{C_1 + C_0} \times \frac{C_2 - C_1}{C_2 + C_1} - 1\right) = \left(\frac{C_2(3C_0 - C_1 - 1) - 5C_1 C_0 - C_1^2}{(C_0 + C_1)(C_2 + C_1)}\right), \tag{34}$$

It is easy to get that if $C_1 \geq 3C_0$, which means that $|1 - \frac{C'_2}{C_1}| < 1$, the error of the second interface placement in the second iteration is smaller than that in the first iteration,

which is $|E'| < |E|$. And if $0 < C_1 < 3C_0$ and $C_2 < \frac{5C_1C_0 + C_1^2}{(3C_0 - C_1 - 1)}$, the error of the second interface placement also becomes smaller in the second iteration. We can conclude that for small velocities' contrasts among C_0 , C_1 and C_2 , the placement error becomes smaller when increasing iterations. While for large velocities contrast among C_0 , C_1 and C_2 , the placement error may increase. We know that $C_2 > C'_2$, so $R_2 > R''_2$ and $\frac{R-2}{R''_2} > 1$. So the

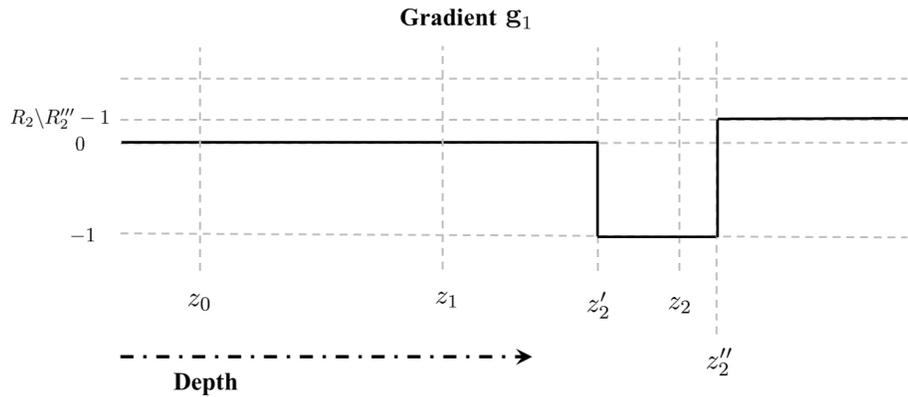


FIG. 8. Gradient g_1 in the second iteration.

gradient g_1 is like the form shown in Figure 7. If we select the optimal step length to make the second layer of the new updated velocity model match the true velocity model well, the updated velocity after the second iteration is like the form shown by the green-dash line in Figure 8.

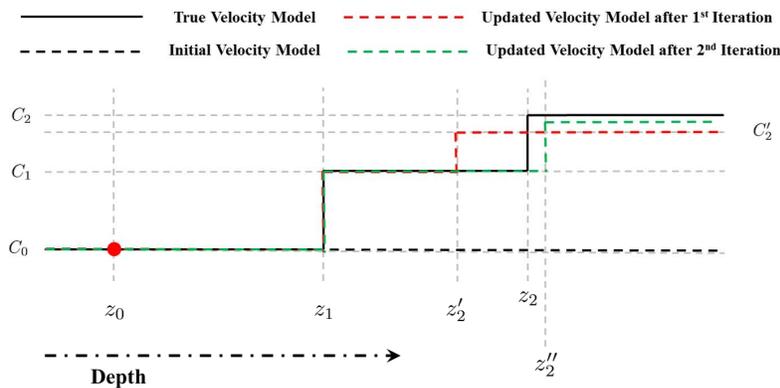


FIG. 9. The inverted velocity model after the second iteration.

NUMERICAL VERIFICATION

To verify our prediction, we performed a series of numerical tests. The 1D numerical example is shown in Figure 9 (a). The solid black line in Figure 9 (a) is the true velocity model with 4000m in horizontal. It consists of two interfaces with the locations of $Z_1 = 977.5m$ and $Z_2 = 1564m$. The velocities of the three layers are $C_0 = 1500m/s$, $C_1 = 3000m/s$ and $C_2 = 4500m/s$ respectively. And the reference model is homogeneous model with the velocity of $C_0 = 1500m/s$, as shown by the red-dash line in Figure 9 (a). Then we performed the first iteration of FWI. The data residual is shown in Figure 9 (b).

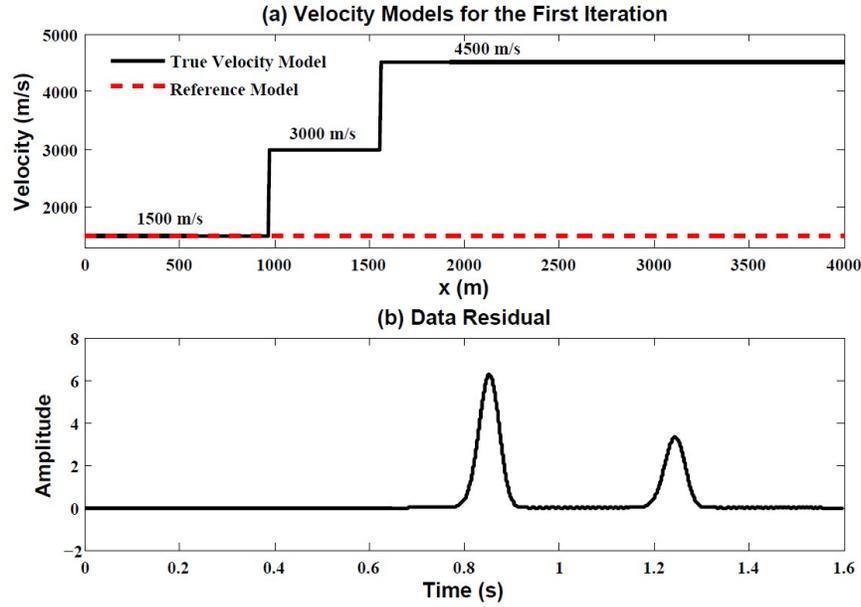


FIG. 10. (a) True velocity model (the solid-black line) and the reference model (the red-dash line); (b) Data residual.

Two impulses caused by the two interfaces are present. Then we calculate the gradient g_0 in the first iteration, as shown in Figure 10. The red line and green line indicate the placements of interface 1 (z_1) and interface 2 (z_2) of the true velocity model, respectively. And the black line is the predicted placement of the second interface in the inverted velocity model using equation $Z'_2 = z_1 + \frac{k_1}{k_0}(z_2 - z_1) = 1270.75m$. We can see that the predicted placement Z'_2 matches the second spike very well, which means that our prediction in the first iteration is right. And we also need to transform the spike to "step" through integration over time. We can also get similar result in the "step" like gradient. To update the velocity model, we need the step length the scale the gradient. Here, for analysis convenience, we can set the step length to make the updated velocity match the true velocity model well at the first interface z_1 . Then, the step length can be set as $\mu = (C_1 - C_0) = 1500$. Then the velocity perturbation becomes:

$$(C_1 - C_0)H(z - z_1) + (C_1 - C_0) \frac{R - 2'}{R_1} H(z - Z'_2) = 1500H(z - 977.5m) + 266.7H(z - 1270.75m), \quad (35)$$

And then, we add this velocity perturbation to the reference model directly and get the updated velocity model indicated by the red-dash line in Figure 11 (a). And Figure 11(b) shows the data residual and two spikes are present. The first negative spike is caused by the second interface Z'_2 in the updated velocity model and the second positive spike is caused by the second interface Z_2 in the true velocity model.

Then, we calculate the gradient g_1 in the second iteration. We can see that the predicted placement of the second interface Z''_2 match the second spike of the gradient very well. What's more, Z''_2 and the error $E' < E$. These are all in consistent with our predictions.

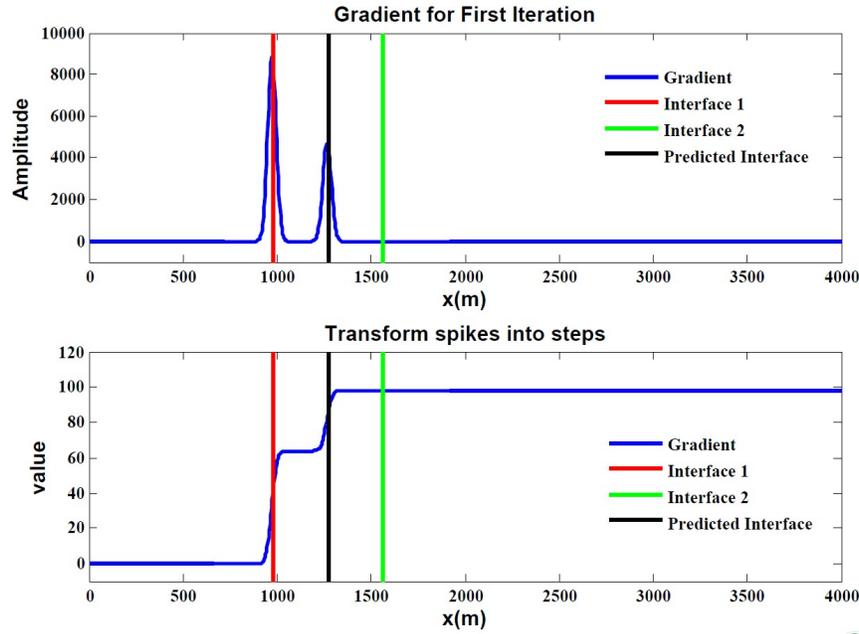


FIG. 11. Gradient g_0 in the first iteration.

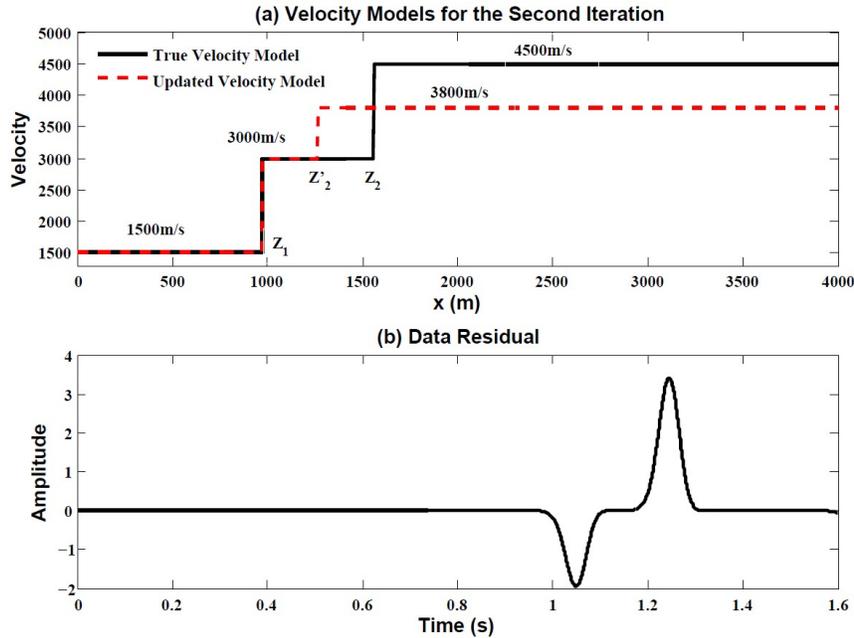
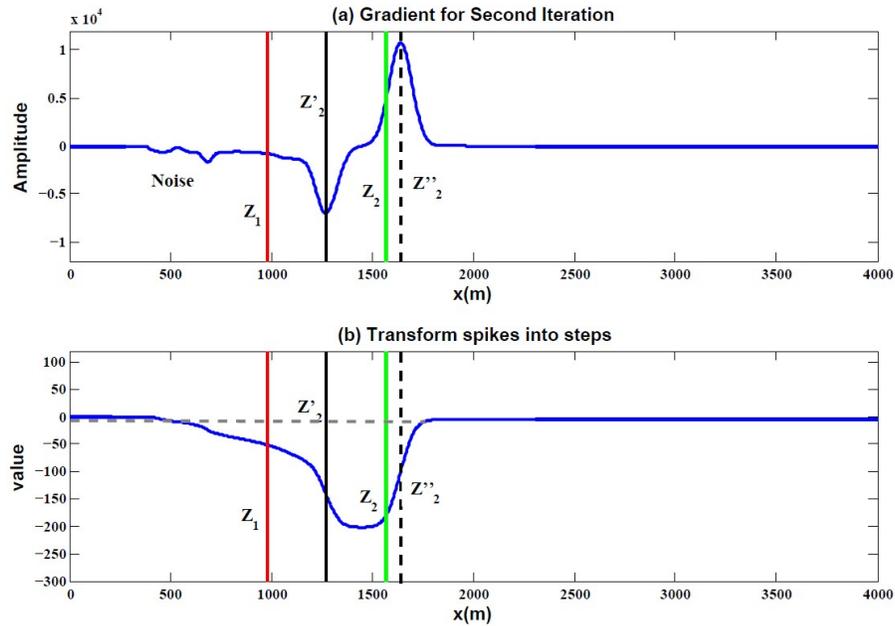


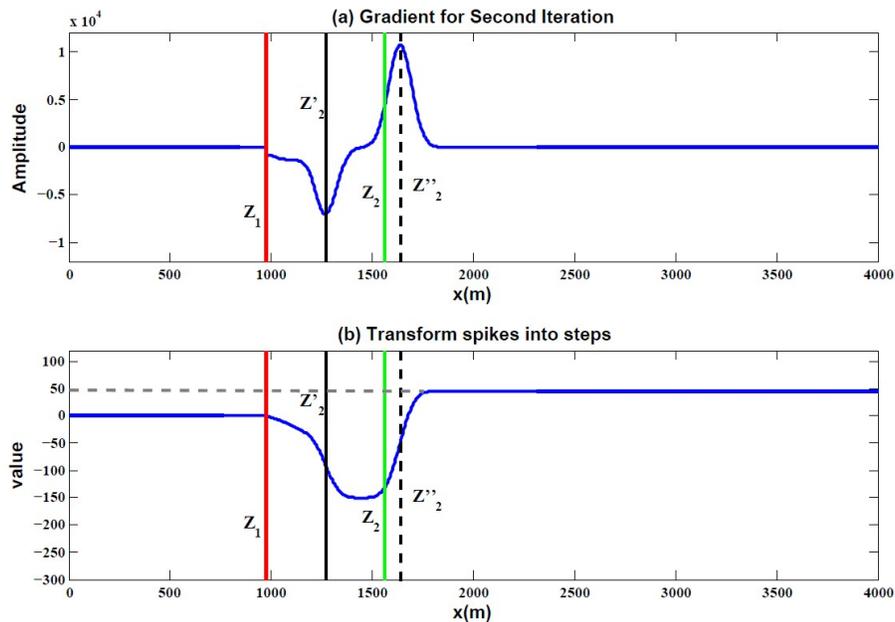
FIG. 12. (a) Velocity models for the second iteration; (b) data residual.

We also notice that in Figure 12 (a), before the first interface Z_1 , some noise is present. These are the noise that we neglect in equation (25). If we integrate the gradient over time, we notice that the value of the second step is negative. While according our prediction, because $\frac{R-2}{R_2''} > 1$, the value of the second step should be positive, as shown in Figure 7.

Then, we set the gradient to be 0 when $z < Z_1$ and eliminate these noise. And then we integrate over time again and we found that the value of the second step changes to positive,

FIG. 13. Gradient g_1 in the second iteration.

which is consistent with our prediction (as shown in Figure 13). So, we can conclude that the noise present in the gradient produces a negative influence on the amplitude recovery, but has no obvious influence on the placement recovery. These noises will decrease the convergence rate of FWI.

FIG. 14. Gradient g_1 after elimination of the noise.

The second numerical example is used to verify that the placement error E' in the second iteration will become larger compared with the placement E in the first iteration. The velocity models are shown in Figure 14 (a). In Figure 14 (a), the blue-dash line is the

initial velocity model. The black-solid line is the true velocity model. And the red-dash line is the velocity model after the first iteration. It can be seen that the velocity contrast between C_2 and C_1 is very large.

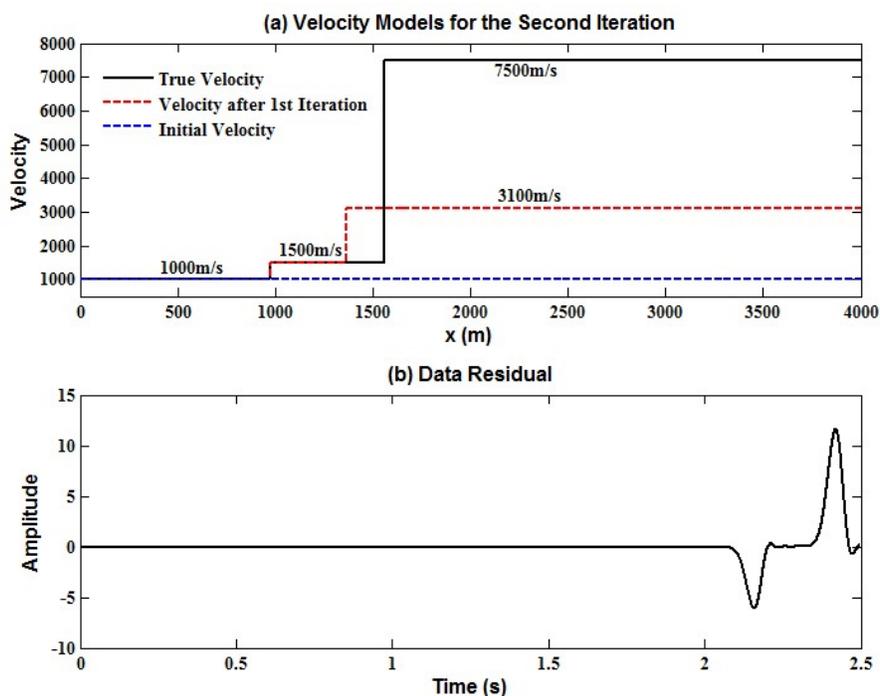


FIG. 15. (a) Velocity models. The blue-dash line is the initial velocity model. The black-solid line is the true velocity model. And the red-dash line is the velocity model after 1st iteration. (b) The data residual in the second iteration.

The gradient g_1 obtained in the second iteration for this numerical example is shown in Figure 15. We can see that the placement error E' in the second iteration is close to the placement error E in the first iteration or even larger. Through numerical calculation, we get that $E' = 208.5370m$ is actually larger than $E = 195.5034m$, which testifies our discussions and predictions in equation (40) further.

CONCLUSION

In this research, we analyzed the 1D scalar FWI using the analytic method. And then we use the numerical examples to testify our predictions. Combining the two methods helps understand how FWI does its job clearly. And several interesting conclusions have been achieved. The placement errors in the first and second iterations are all influenced by the velocity contrast and interfaces distance. And what's more, the placement error in the second iteration becomes smaller than that of the first iteration for small velocity contrast. While for large velocity contrast, this placement error may become larger. Another thing found in this research is that the noise produced in the cross-correlation produce a negative influence to the amplitude recovery of the second interface, which will decrease the convergence rate of FWI.

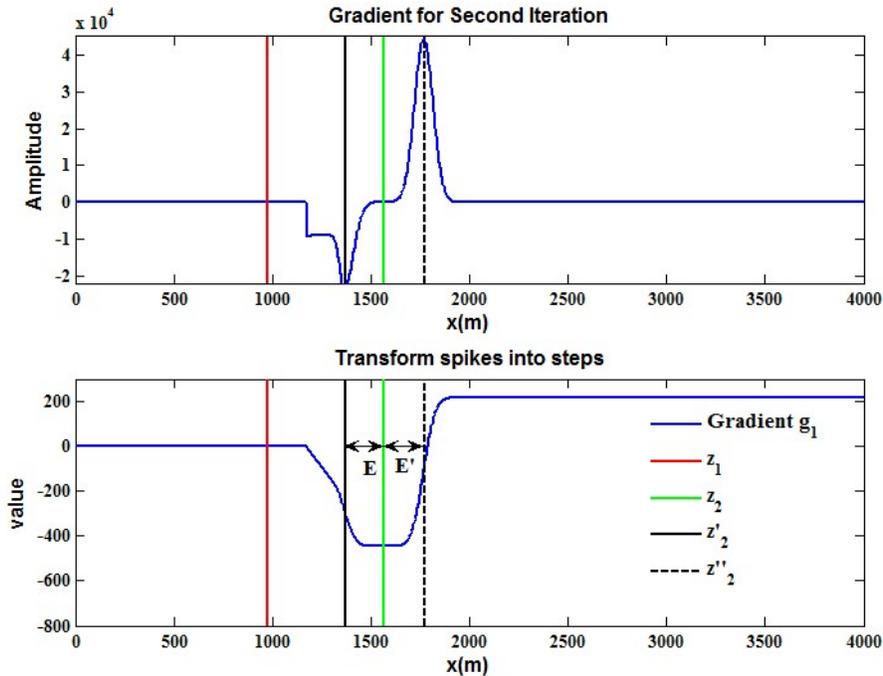


FIG. 16. Gradient g_1 in the second iteration. The red and green lines are the interface locations in the true velocity model. And the black-solid and black-dash lines indicate the locations of the second interface in the first and second iterations respectively.

ACKNOWLEDGEMENTS

This research was supported by the Consortium for Research in Elastic Wave Exploration Seismology (CREWES).

REFERENCES

- Gauthier, O., Virieux, J., and Tarantola, A., 1986, Two-dimensional nonlinear inversion of seismic waveforms: numerical results: *Geophysics*, **51**, 1387–1403.
- Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migration, 206–220.
- Pica, A., Diet, J. P., and Tarantola, A., 1990, Nonlinear inversion of seismic reflection data in a laterally invariant medium: *Geophysics*, **55**, 284–292.
- 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.