A SLIM APPROXIMATE HESSIAN FOR PERTURBATION VELOCITY INVERSION WITH INCOMPLETE REFLECTION DATA

BINGHONG HE and GUOCHEN WU

School of GeoSciences, China University of Petroleum (Huadong), Qingdao, Shandong 266555, P.R. China. hbh9517@163.com

(Received June 9, 2014; revised version accepted May 8, 2015)

ABSTRACT

He, B. and Wu, G., 2015. A slim approximate Hessian for perturbation velocity inversion with incomplete reflection data. *Journal of Seismic Exploration*, 24: 281-304.

A recent velocity model building is the full waveform inversions (FWI) that allow to recover the long-scale structures through the refraction waves and diving waves, and the short-scale structures which provide the high-resolution component through the reflection waves. However, incomplete seismic data include non-geological artifacts in the gradient for velocity update. The strong off-diagonal elements of approximate Hessians are important to reflection FWI with incomplete data; however, it is difficult to implement an approximate Hessian using the forward modeling method because of the cost of the computation efficiency. In this study, we investigate the ability of an approximate Hessian to remove artifacts that are caused by incomplete reflection data. In order to reduce the costs associated with calculating the Hessian, the large model is separated into sparse sub-models, and an alternative slim approximate Hessian is implemented sequentially on these sub-models. Afterwards, The complete model is obtained from sub-model using the radial point interpolation method (RPIM).

A two-dimensional flat-layers synthetic example provides a reasonable test case for our method. We find that the slim approximate Hessian removes non-geophysical artifacts as effectively as the approximate Hessian, but has the advantages of greater cost-efficiency and lower memory requirements.

KEY WORDS: full waveform inversion, slim approximate Hessian, reflection, incomplete data.

0963-0651/15/$5.00

INTRODUCTION

Seismic inversions obtain the geophysical features of lithological layers and boundaries, describing the geological structure of reservoirs and their interfaces. Conventional seismic inversions, such as the acoustic impedance inversion, elastic impedance inversion, and the AVO (amplitude versus offset) inversion are implemented using convolutions based on one-dimensional ray theory (Hampson et al., 2012; Connolly, 1999, 2007). These methods are well developed, as they are cost-efficient compared to wave equation methods. However, model waves based on one-dimensional ray theory are a gross simplification of waves propagated in the real three-dimensional complex earth system. As full waveform inversions are implemented based on the wave equation, varied waves are incorporated into the inversion process, including direct waves, diving waves, diffraction waves, refraction waves, reflection waves, and scattering (Virieux and Operto, 2009). The classic full waveform inversion (FWI) is the reflection FWI proposed in the time domain (Tarantola, 1984; Lailly, 1983). With the development of seismic modeling in the frequency domain (Pratt and Worthington, 1990), full waveform inversion in the frequency domain can take advantage of multi-scale inversion and parallel computation techniques (Pratt et al., 1998). An alternative technique (Shin and Cha, 2008) proposed a full waveform inversion in the Laplace and Fourier-Laplace domains in order to avoid cycle-skipped portions of the data.

Full waveform inversion is a procedure that has a wide range of applications from long-scale to short-scale structures, and from shallow to deep inversions (Fichtner, 2010). When the data range is observed over a wide aperture from near offset to far offset, the long-scale structures for shallow sections will be recovered from the refraction and diving waves (Shipp and Singh, 2002), and reflection data provide the inversion for short-scale structures, which correspond to the higher resolution components of velocity. To account for the absence of low frequencies in models and data, the velocity is always split to provide a smooth background velocity and oscillatory component (reflectivity). Symes and Kern (1994) proposed a reflection waveform inversion using differential semblance analysis to provide a data fit, followed by updates to smooth the background velocity. Xu et al. (2012) recommended a similar reflection inversion method using Green's function, and applied this method to Gulf of Mexico data. Currently, most research is focused on smooth background velocities for migration images. Further study on higher resolution models with high-wavenumbers would therefore be beneficial. In this study, we assume that the observed data include both the refraction and reflection waves, and implement the refraction FWI and reflection FWI accordingly. This involves inversion of the long-scale structures using refraction data, and of the short-scale structures using reflection data. Perturbations in velocity reflect short-scale structures.

The classical objective function for full waveform inversion is built based on the match between observed and simulated data. The comprehensive solutions based on Newton methods are prohibitively expensive due to the computational costs of forward modeling. Fortunately, gradient-based methods are very efficient, owing to the ease of computing gradients using the adjoint-state method, especially for large datasets. Since gradient calculation performances improve with reverse time migration, for reflection inversion the high-wavenumber structures are always mixed with artifacts that are not geologically reasonable (Ma, 2012). Furthermore, incomplete observed seismic data, due to a limited acquisition system or inappropriate seismic processing, would add further artifacts. Sirgue (2004) had investigated that the wavenumber illumination depended on the half offset-to-depth ratio of the source-receiver pair to the target. The larger the offset-to-depth ratio, the smaller is the wavenumber. Therefore, the refraction FWI which focus on the shallow and large-offset seismic waves always excludes the high-wavenumber non-geophysical artifacts. And then it is more reliable to apply the refraction FWI to marine and land seismic data (Plessix et al., 2010, 2012; Ravaut et al., 2004; Shipp and Singh, 2002; Vigh, 2012) than the reflection FWI.

Nemeth et al. (1999) discussed the suppression of non-geological artifacts in a diffraction model using least squares migration. Least squares migration is analogous with the first iteration of a perturbation update by the Gauss-Newton method using reflection data. Although the background velocity and perturbation velocity are updated alternately according to reflection FWI (Xu et al., 2012), only the perturbation velocity is updated if a reliable background velocity has been obtained. we focus on the perturbation velocity update, and we do not discuss the background velocity update in this paper. Additionally, a slim approximate Hessian method is proposed to conserve computational memory and storage for field seismic data. Compared with the approximate Hessian method and the filtering method, this slim method works well for both a test dataset and field data applications.

HESSIAN RESPONSE

The classical objective function for full waveform inversion is defined as the least squares of the data misfit:

$$J[m] = \tfrac{1}{2} \| u(m) - d \|^2 \ , \tag{1}$$

where u(m) and d are the synthetic and observed datasets, respectively.

For reflection FWI, in order to avoiding local solution the inversion processes were separated into two stages based on the supposition that the velocity model can be divided into a smooth background section and a

perturbation (van Leeuwen and Mulder, 2009, 2010). A variable projection (Golub and Pereyra, 1973) provides a way forward for this least squares solution. The full seismic wavefield can be described as:

$$u = F[m] \quad .$$

Using the Born approximation, the linear modeling operator can be expressed as

$$F[m] = \partial u/\partial m$$

$$\partial u[m] = (\partial u/\partial m)\big|_m m_p = F[m]m_p \quad , \tag{2}$$

where m_p is the perturbation model. For the least squares migration, it behaves as reflectivity. In this study, m_p is the perturbation velocity, and $r = m_p/m$ stands for reflectivity, where m is the smooth background model.

Neglecting multi-scattering, the input observed data should include only born data without direct wave. Observed data d_{obs} include direct wave and the simulation data $F[m]m_p$ by Born equation do not include direct wave. In order to match the observed data with the simulation data, we get the new observed data without the direct wave using equation as:

$$\bar{d}[m_0] = d_{obs} - F[m_0] \quad . \tag{3}$$

According to the Appendix, the perturbation component is expressed as,

$$m_{p,k}[m] = (F^*[m]F[m])^{-1}F^*[m]\bar{d}[m] = (H_a[m])^{-1}F^*[m]\bar{d}[m] \quad , \tag{4}$$

where $H_a[m] = F^*[m]F[m]$.

The term $F^*[m]F[m]$ in eq. (4) is the normal operator, which is also called the approximate Hessian, and is the linear part of the Hessian. Eq. (4) indicates that the perturbation model for a given background is similar to the least squares migration, while the term $F^*[m]\bar{d}[m]$ corresponds to reverse time migration. The update using refraction is performed in a similar way to the perturbation component calculation using reflection. The difference is that the data for refraction FWI, in data session $\delta d = F[m] - d$ is residual, and the observed data, d, include refraction, reflection, diffraction, diving waves, and scattering waves. In contrast, for reflection FWI, the input data, $\bar{d}[m]$, are reflected waves.

Application of the approximate Hessian, H_a, follows the well-known Gauss-Newton method. In this paper, $H_a = F^*F$, and the explicit format is

$$H_a(x,y) = \sum_{\omega} \sum_{x_s} \sum_{x_r} [\partial u(\omega,x_s,x_r,x)/\partial m(x)][\partial u(\omega,x_s,x_r,y)/\partial m(y)] \quad . \tag{5}$$

Rewriting using Green's function, this becomes

$$H_a(x,y) = \sum_{\omega} \omega^4 \sum_{x_s} |s(\omega)|^2 G(\omega,x_s,x) G^*(\omega,x_s,y) \sum_{x_r} G(\omega,x,x_r) G^*(\omega,y,x_r) \quad . \tag{6}$$

Owing to the cost of forward modeling and the huge storage required for the approximated Hessian, gradient-based methods provide simple solution methods, such as the steepest descent and conjugate gradient methods. For the steepest descent method, a scalar is used instead of the approximate Hessian. In detail, the perturbation component is propagation with $F^*[m]\bar{d}[m]$ shortened to $F^*\bar{d}$. Setting $y = x$, we obtain the diagonal approximate Hessian:

$$H_a(x,x) = \sum_{\omega} \omega^4 \sum_{x_s} |s(\omega)|^2 |G(\omega,x_s,x)|^2 \sum_{x_r} |G(\omega,x,x_r)|^2 \quad . \tag{7}$$

Fig. 1 shows a model with three layers and a total of 51 × 51 nodes. The approximate Hessian is shown in Fig. 2. Using a multi-shot system, the diagonal elements in the approximate Hessian are overwhelmingly dominant (Figs. 2c and 2e) and the response of one point is focused on a small zone surrounding this point (Fig. 3). Compared with a 51-shot system, the off-diagonal elements play important role for the one-shot system, although these elements are far smaller than the diagonal elements (Figs. 2d and 2f). Nevertheless, the response of one point can affect the full model space (Fig. 4).

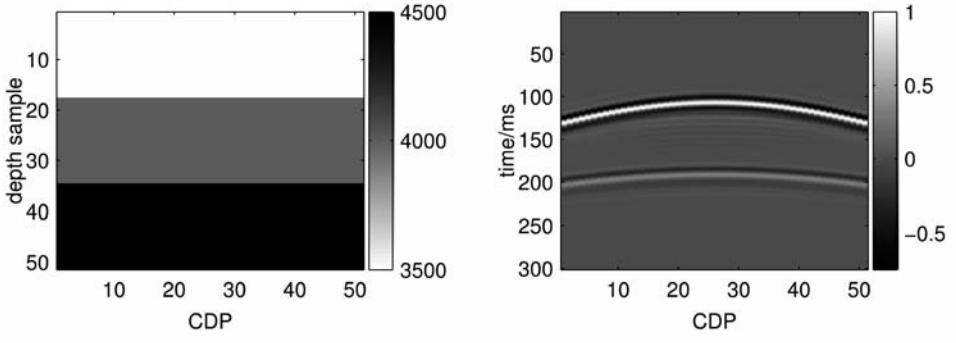

Fig. 1. Model and records: (a) velocity, the model contains three layers with 51 × 51 nodes and (b) records (one-shot) from 51 receivers.

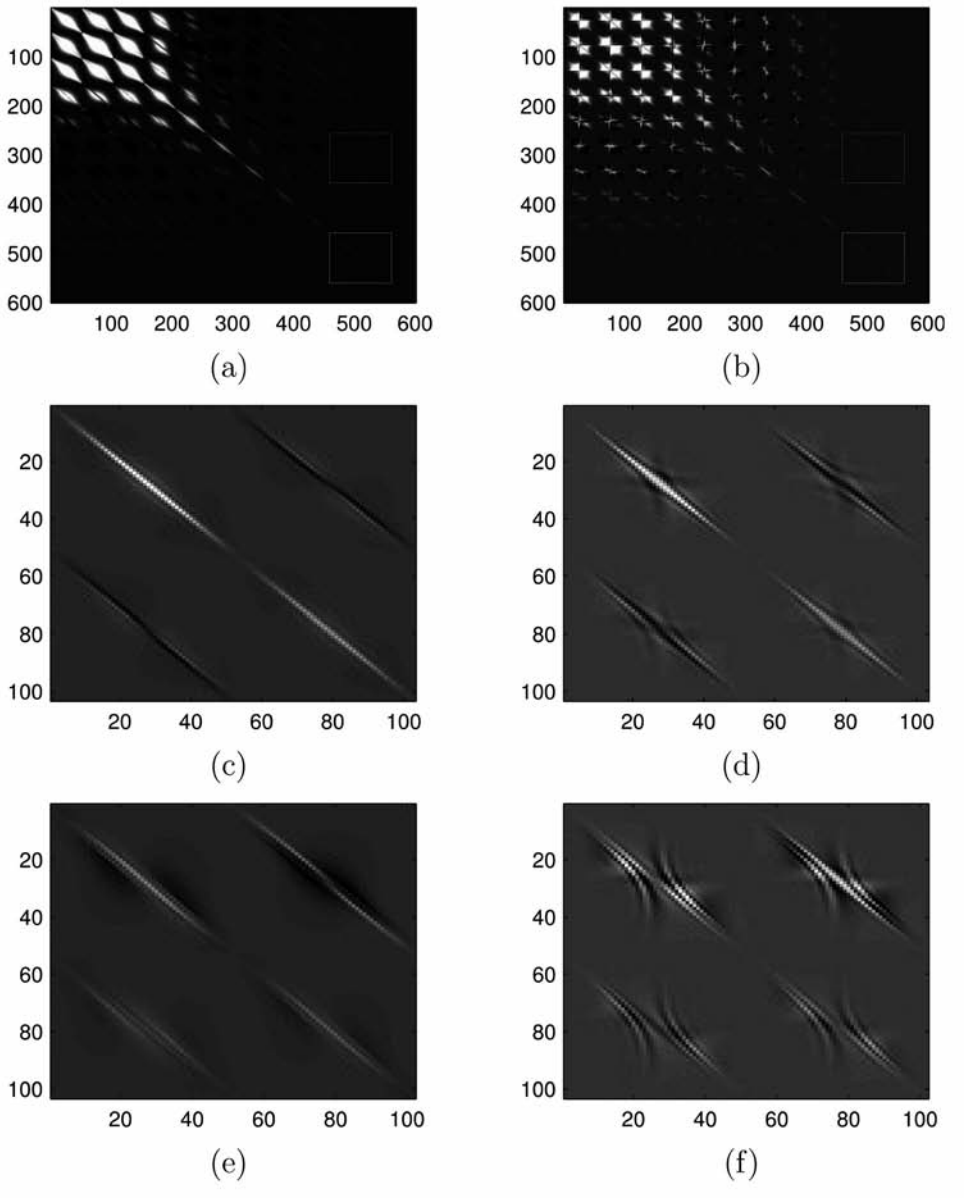

Fig. 2. Hessian responses (one-shot system) of regular receivers: (a) part of the approximate Hessian (51-shot system); (b) part of the approximate Hessian (one-shot system); and (c) and (d) are the diagonal elements of (a) and (b), and (e) and (f) are the off-diagonal elements of (a) and (b), respectively.

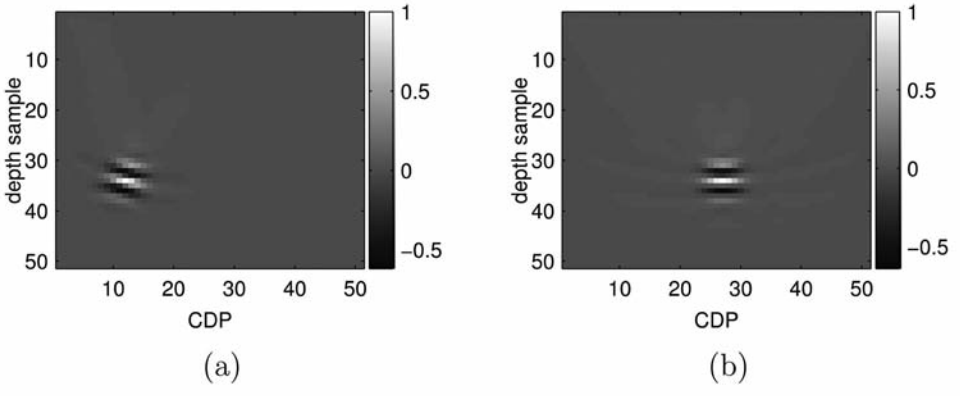

Fig. 3. The response at one point (51-shot and 51 receivers): (a) sample depth (z) = 33, CDP (x) = 11 and (b) z = 33, x = 26.

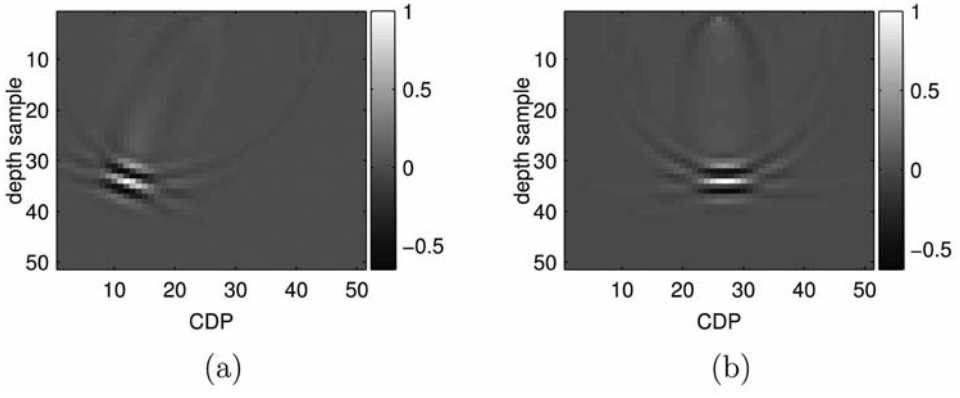

Fig. 4: The response at one point (one-shot and 51 receivers): (a) sample depth (z) = 33, CDP (x) = 11 and (b) z = 33, x = 26.

THE PERTURBATION COMPONENT INVERSION USING INCOMPLETE REFLECTION DATA

Within the scope of seismic exploration, seismic signals are picked up by a considerable number of receivers arranged on the ground or sea floor above a geological formation. Since it is not always possible to place receivers or explosive sources in the theoretically best positions, especially for mountainous areas, rivers, and city buildings, seismic data are usually irregular and incomplete. In this section, we discuss the perturbation component inversion with incomplete reflection data.

Slim approximate Hessian for perturbation component

The diagonal elements represent the illumination of the source-receiver system. Since calculations are simpler for the diagonal elements, the approximate Hessian is usually replaced with its diagonal elements. However, for incomplete data, the approximate Hessian typically has strong off-diagonal components (Nemeth et al., 1999) that must be considered when removing artifacts. Fig. 5a shows incomplete data for one-shot with only six receivers arranged on the surface. Figs. 5b and 5d show parts of the approximate Hessian according to the source and six-receiver system. It can be clearly seen that the off-diagonal elements are greater than the diagonal elements. Indeed the response of one of the off-diagonal elements is too strong to justify considering the diagonal elements only (Fig. 6).

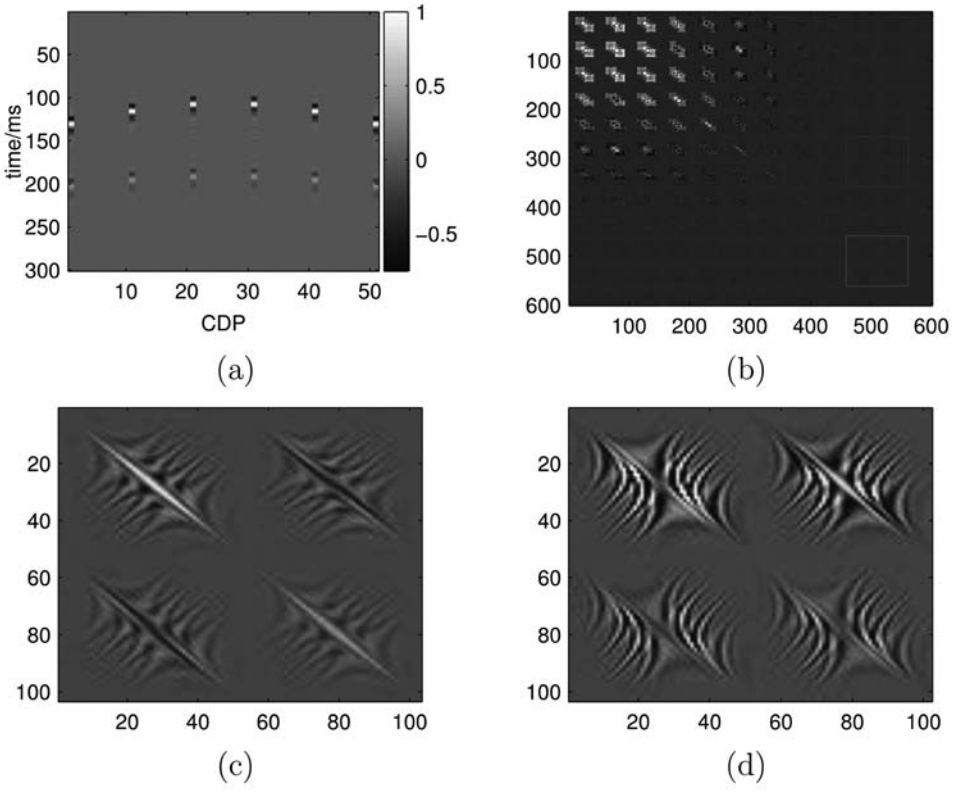

Fig. 5. Incomplete reflection and approximate Hessian: (a) incomplete reflection data; (b) part of the approximate Hessian (one-shot); (c) diagonal elements of (b); (d) off-diagonal elements of (b).

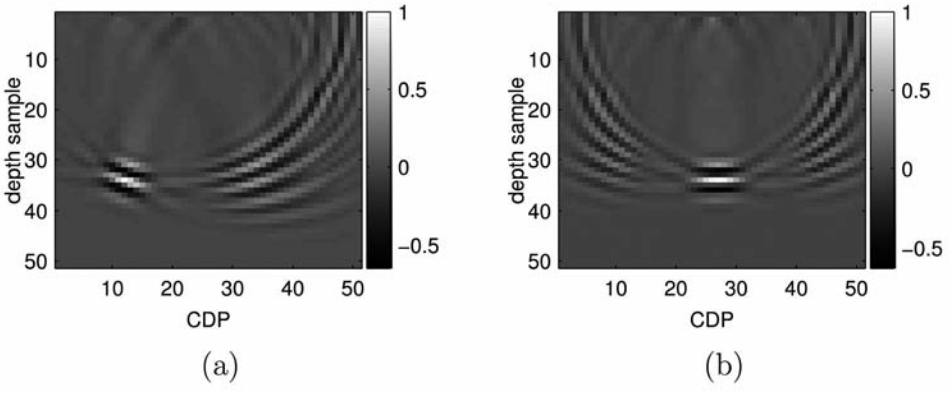

Fig. 6. The response of one point (one-shot and only six receivers): (a) sample depth (z) = 33, CDP (x) = 11 and (b) z = 33, x = 26.

In order to reduce the costs associated with calculating the Hessian, image-guild FWI (IGFWI) (Hale, 2009) uses subsurface structures to constrain the inversion of a sparse model. With respect to the sparse model, a projected Hessian matrix saves both computational time and memory, and Ren et al. (2012) decomposed the explicit Hessian into the local angle domain to avoid the significant inversion problems. In this study, we introduce an alternative slim approximate Hessian to reduce the cost of approximate Hessian calculations and keep the strong off-diagonal components for incomplete data.

We rewrite the term $m_p[m]$ as $m_{p,k}[m]$, standing for the perturbation velocity after k iterations, based on the background velocity m. And we give the new term $m^{1/2}_{p,k}[m]$ in model domain,

$$m^{1/2}_{p,k}[m] = F^*[m]\bar{d}[m] \quad . \tag{8}$$

In order to reduce the cost of $H_a[m]$, we introduce a split operator P_i to split the model-domain term $m^{1/2}_{p,k}[m]$ to sub-models, marked as $m^{1/2}_{pi,k}[m]$, which are also in model domain. The relation between $m^{1/2}_{p,k}[m]$ and $m^{1/2}_{pi,k}[m]$ is

$$m^{1/2}_{p,k}[m_k] = \sum_{i=1}^{N} m^{1/2}_{pi,k}[m_k] \quad , \tag{9}$$

and

$$m^{1/2}_{pi,k}[m] = m^{1/2}_{p,k}[m_k]P_i = F^*[m]\bar{d}[m]P_i \quad . \tag{10}$$

The split operator P_i can takes different forms, including wavelet transform, cubic splines, or image-guided interpolation. In this paper, we implement the split operator P_i as a linear interpolation, satisfying the relation

$$\sum_{i=1}^{N} P_i(\mathbf{x}) = 1 \quad .$$

Therefore the perturbation velocity in sub-model domain is written as:

$$m_{pi,k}[m] = (H_a[m])^{-1}(m^{1/2}_{p\,i,k}[m]) \quad . \tag{11}$$

In fact, we need a complete model for velocity update and fit the observed data. We introduce a shape function ϕ to interpolate the sub-model to the complete model.

$$m_{p,k}[m](x) = \sum_{i=1}^{N} \phi(x_j) m_{pi,k}[m](x_j) \quad , \tag{12}$$

where n is the number of the grid points in the support domain of point (x). The radial point interpolation method (RPIM) (Liu and Gu, 2005; Liu et al., 2005) is implemented for the interpolation, which is developed using the radial basis function (RBF). And the trial function for RPIM is written as

$$m_{p,k}[m](x) = \sum_{i=1}^{N} R_j(x)a(j) = \mathbf{R}^T(\mathbf{x})\mathbf{a} \quad , \tag{13}$$

where $R_i(x)$ is the RBF and n is the number of radial basis functions (RBF). There are many types of RBF, such as the multi-quadric (MQ) function, the Gaussian function, and the thin plate spline (TPS) function. The MQ function is applied in this study:

$$R_i(x,y) = [r_i^2 + (a_c d_c)^2]^q \quad , \tag{14}$$

where a_c and q are shape parameters. A standard MQ-RBF ($q = 0.5$) is used for the RPIM, and distance is given by $r_i^2 = (x - x_0)^2 + (y - y_0)^2$.

Coefficients matrix **a** can be determined to be satisfied at these n nodes surrounding the point of interest **x**. This leads to n linear equations, one for each node. The matrix form of these equations can be expressed as

$$\mathbf{m}_{pi,k}[m] = \mathbf{R}_0\mathbf{a} \quad , \tag{15}$$

where the moment matrix of RBFs is:

$$\mathbf{R}_0 = \begin{bmatrix} R_1(r_1) & R_2(r_1) & \cdots & R_n(r_1) \\ R_1(r_2) & R_2(r_2) & \cdots & R_n(r_2) \\ \cdots & \cdots & \cdots & \cdots \\ R_1(r_n) & R_2(r_n) & \cdots & R_n(r_n) \end{bmatrix} . \tag{16}$$

The RPIM shape function Φ can be expressed:

$$\Phi^T(\mathbf{x}) = \mathbf{R}^T(\mathbf{x})\mathbf{R}_0^{-1} . \tag{17}$$

Eq. (12) can be rewritten as:

$$m_{p,k}[m](x) = \Phi \mathbf{m}_{pi,k}[m] . \tag{18}$$

The choice of the split operator

The conventional splitting method is to split the full model into smaller models, gradually moving through full model. In this paper, we define N = 2, and therefore there are two sub models (Figs. 7a and 7b). For a 2D velocity model, the split operator P_i is expressed as:

$$P_1 = \begin{bmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & \cdots & 1 & 0 & \cdots & 0 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 0 & \cdots & 0 & 1 & \cdots & 1 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & \cdots & 1 \end{bmatrix} . \tag{19}$$

This study provides two additional methods for splitting the model. Figs. 7c and 7d show sub-models covering the entire model range, (the "rectangle-model"). In this situation, the split operator P_i is expressed as:

$$P_1 = \begin{bmatrix} 1 & 0 & 1 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & 1 & \cdots & 1 & 0 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & 0 & \cdots & 0 & 1 \end{bmatrix} . \tag{20}$$

Another splitting method is executed in both two dimensions. Figs. 7e and 7f show a staggered method for splitting the model (the "stagger-model"). The split operator P_i is:

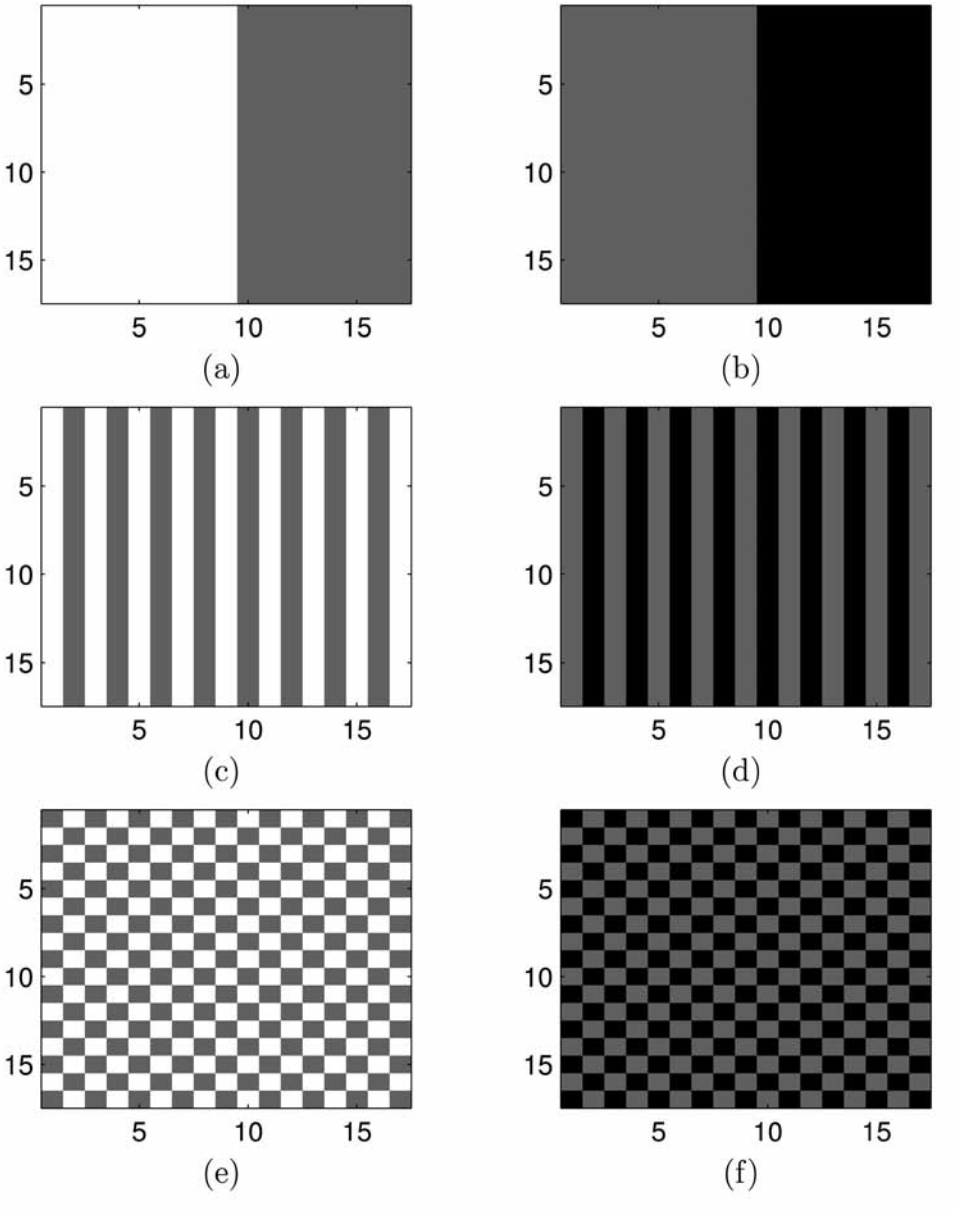

Fig. 7. Three methods of model splitting are shown. (a) and (b): conventional methods; (c) and (d): rectangle-model method; and (e) and (f): stagger-model methods. The total model is the sum of (a) and (b), (c) and (d), or (e) and (f). The white and black nodes are the efficient nodes.

$$P_1 = \begin{bmatrix} 1 & 0 & 1 & \cdots & 0 & 1 \\ 0 & 1 & 0 & \vdots & \vdots & \vdots \\ 1 & 0 & 1 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & 0 & \cdots & 1 & 0 \\ 1 & 0 & 1 & \cdots & 0 & 1 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 0 & 1 & 0 & \cdots & 1 & 0 \\ 1 & 0 & 1 & \vdots & \vdots & \vdots \\ 0 & 1 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 1 & \cdots & 0 & 1 \\ 0 & 1 & 0 & \cdots & 1 & 0 \end{bmatrix} . \quad (21)$$

As the number of nodes is approximately halved for both methods, the slim approximate Hessian needs only a quarter of the memory compared to the full model. Additionally, forward modeling is more efficient.

The strategy for the alternative slim approximation Hessian

The strategy for the alternative slim approximation Hessian was designed as follows:

Stage 1 involves executing $F^*\bar{d}$ for the tentative complete model to obtain $m^{1/2}_{p,k}[m]$:

$$F^*\bar{d} = \int_0^T u_b(T - t)(\partial^2 u_s/\partial t^2) \quad , \quad (22)$$

where $u_b(T - t)$ is the back-propagated reflection, which combines with the wave equation:

$$m[\delta^2 u_b(T - t)/\delta t^2] - \nabla^2 u_b(T - t) = \sum_{x_r} -\bar{d}(x_r,t) \quad . \quad (23)$$

In stage 2, the tentative complete model, $m^{1/2}_{p,k}[m]$ and the background velocity m , is split into sub-models, respectively. Since the nodes in the rectangle-model are regular, it is simple to apply the finite-difference operator for modeling. In contrast, it is difficult to apply the finite-difference operator to the stagger-model. Therefore, in this study, we adopt the rectangle-model.

Stage 3 involves applying the operator $F^*F = H_a[m_1]$ to the first sub-model $m^{1/2}_{p1,k}[m]$ to obtain the model $m_{p1,k}[m]$. And then it is able to get the complete velocity model $m_{p,k}[m]$ using the radial point interpolation method (RPIM).

In Stage 4, turning to the next iteration, $m^{1/2}_{p,k+1}[m]$ is split into sub-models

and the operator F*F = $H_a[m_2]$ is applied to the second sub-model to obtain $m_{p2,k+1}[m]$. RPIM is then used to obtain the total model, $m_{p,k+1}$.

Fig. 8 shows part of the approximate Hessian for the total model and the sub-model. The responses of one point on the sub-model, shown in Fig. 9, are consistent with responses for the total model.

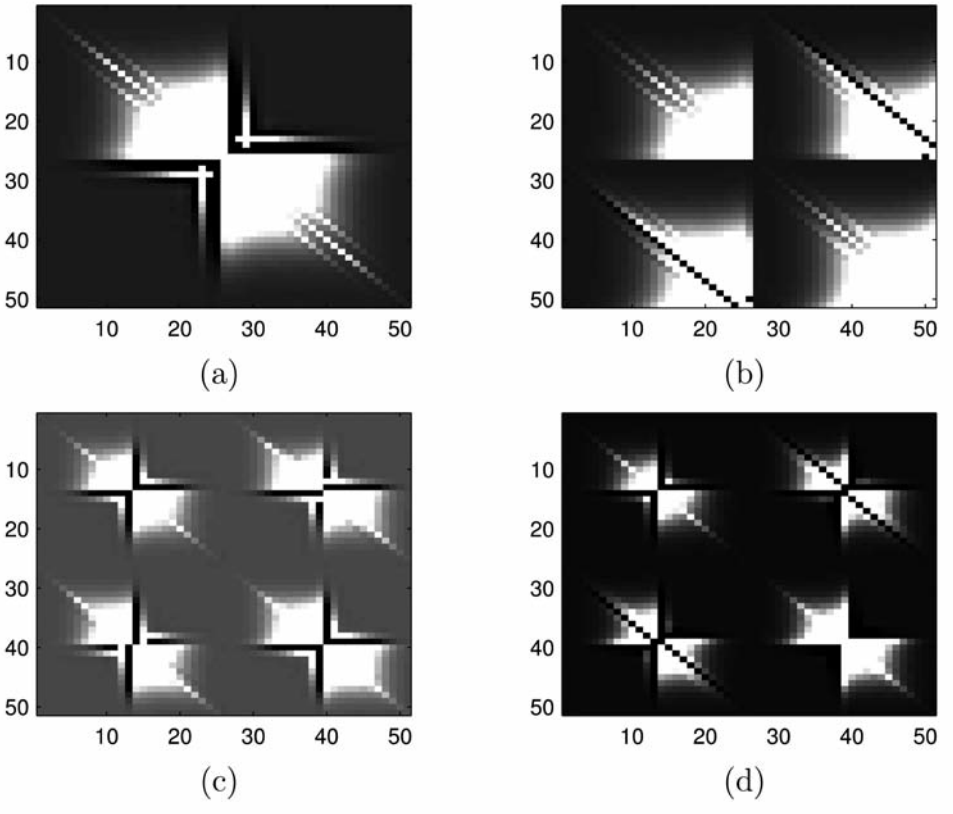

Fig. 8. Details of the approximate Hessian for (a) the total model, (b) the sub-model in Fig. 7a, (c) the sub-model in Fig. 7c, and (d) the sub-model in Fig. 7e.

Filter for perturbation component calculation in the wavenumber domain

As artifacts behave very differently for geological information in the wavenumber domain, a filter was designed to remove artifacts, and the filtering method was compared to the slim approximate Hessian method. The application of this filter involves three steps. First, we transfer the gradient from the space domain to the wavenumber domain by a two-dimensional Fourier transform:

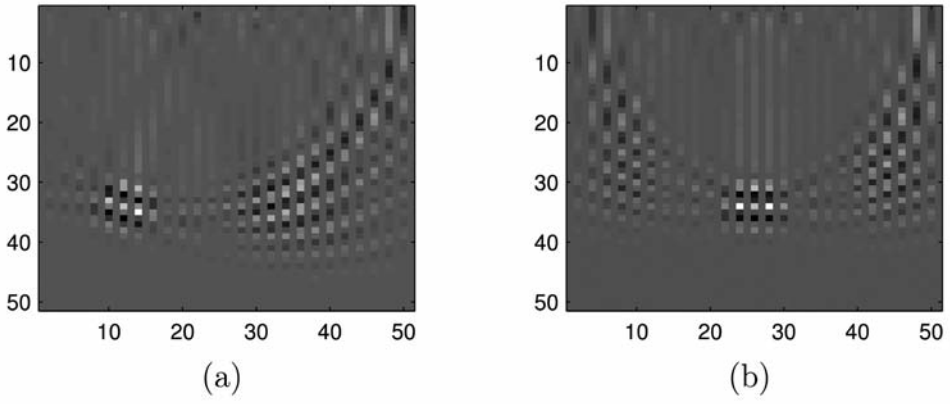

Fig. 9. The response of one point in a sub-model (one-shot and only six receivers). (a) Distraction position: sample depth (z) = 33, CDP (x) = 11 and (b) distraction position: z = 33, x = 26.

$$m_p(k_x,k_z) = \int\int m_p(x,z)e^{-jk_x x}\, e^{-jk_z z}\, dxdz \quad , \tag{24}$$

where $m_p(k_x,k_z)$ is the perturbation component in the wavenumber domain, k_x is the horizontal wavenumber, and k_z is the vertical wavenumber. Secondly, a filter, $h(k_x,k_z)$, defined in the wavenumber domain, is applied to the perturbation component:

$$m_p'(k_x,k_z) = h(k_x,k_z) \int\int m_p(x,z)e^{-jk_x x}\, e^{-jk_z z}\, dxdz \quad . \tag{25}$$

Thirdly, we obtain the gradient in the space domain by a two-dimensional inverse Fourier transform:

$$\begin{aligned} m_p'(x,z) &= (1/4\pi) \int\int h(k_x,k_z) \\ &\quad \times \int\int m_p(x,z)e^{-jk_x x}\, e^{-jk_z z}\, dxdz\, e^{\,jk_x x} e^{\,jk_z z}\, dk_x dk_z \quad . \end{aligned} \tag{26}$$

The filter is designed in the wavenumber domain:

$$h(k_x,k_z) = \begin{cases} 1 & k_x = k_z\tan(\theta-\phi) \\ [\arctan(k_x/k_z)+\theta+\phi]/(2\phi+\varepsilon) & [k_z\tan(\theta+\phi) \le k_x \le k_z\tan(\theta+\phi) \\ 0 & k_x = k_z\tan(\theta+\phi) \end{cases} , \tag{27}$$

where θ is determined relative to the geological structure. And ϕ is a small transitional angle between the remain and mute angle region.

Fig. 10a shows the reverse time migration, the result of F*$\bar{d}$. With the slim approximate Hessian, valid structure is retained and most artifacts are removed (Fig. 10b). For the filter, when $\theta = 45°$, many artifacts are retained (Fig. 10c). However, when $\theta = 10°$, the alias of the migrated section is prohibitive (Fig. 10d). Therefore, the filter could not replace the slim approximate Hessian for incomplete reflection data.

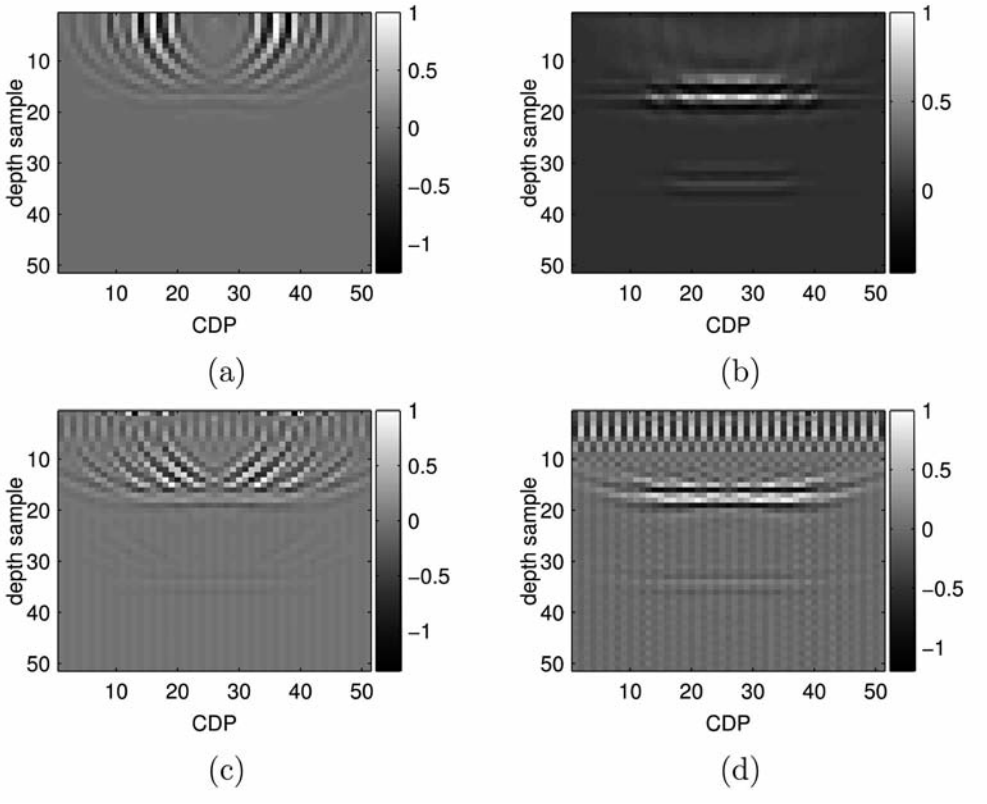

Fig. 10. The perturbation component inversion with the one-shot and six receivers system: (a) the result of F*d; (b) the perturbation component inversion with the slim approximate Hessian; and (c) and (d) are the perturbation component inversions using $\theta = 45°$ and $\theta = 10°$ for the filter method, respectively.

FIELD DATA APPLICATION

For field data acquisition, there are many reasons for incomplete data, including scarcity of sources. As high-density acquisition will increase the cost, sources are often sparsely arranged on the ground or sea floor. The second factor is limitations in the receiver array, and the third factor is the discontinuation of the seismic records, such as breakpoint, due to irregularity

of receiver arrays caused by buildings, rivers and other complex topography. In this study, we used a land-surface dataset to test the slim Hessian method. The two-dimensional seismic data are extracted from a three-dimensional survey, including 48 sources arranged in two shot arrays (Fig. 11). For each shot array, there are 24 surface sources (the triangles in Fig. 11). The sparse and irregular sources are one cause of artifacts. Fig. 12 shows the observed reflected data for the full waveform inversion, which does not include Rayleigh surface waves, direct waves, and refraction waves. The reflection is obviously discontinuous, leading to strong artifacts in the inversion.

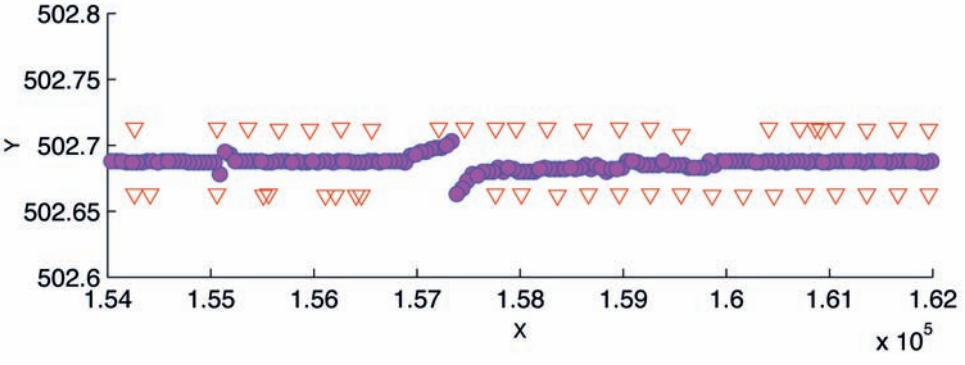

Fig. 11. Acquisition system: the triangles are sources, and the filled circles are receivers.

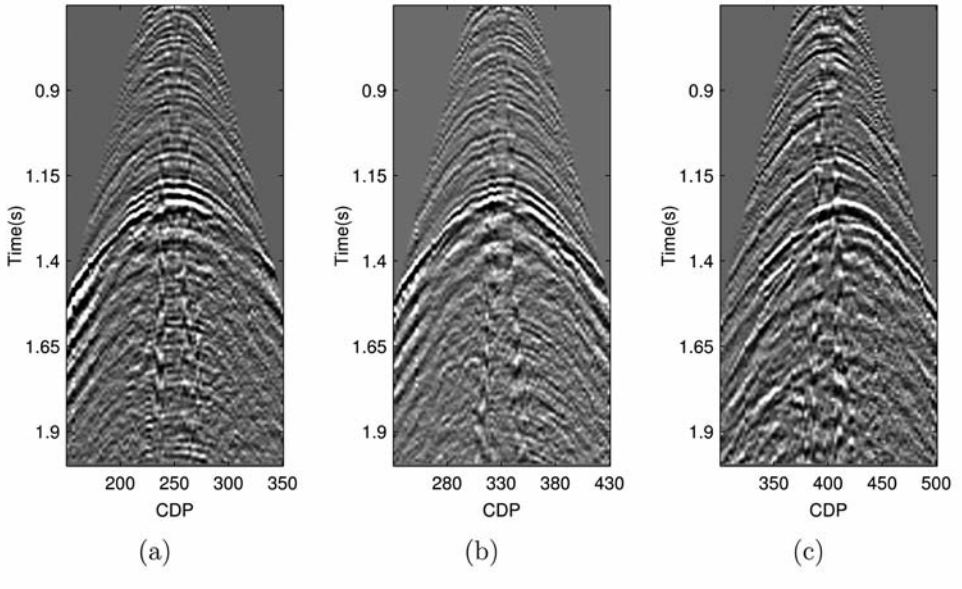

Fig. 12. Common-shot gathers (observed seismic data): (a) shot position CDP252, (b) shot position CDP332, and (c) shot position CDP403.

As energy decreases with depth, making deep velocity inversions difficult, a cubic weight factor (dependent on depth) is applied to the gradient. Although a series of researchers had paid attention to build the initial velocity (Lambaré, 2005, 2008), for this field data, the initial model is obtained by NMO analysis, which includes 161 x 1201 nodes, and is shown in Fig. 13. Fig. 14 shows the reflectivity obtained by reverse time migration (i.e., the result of $F^*\bar{d}$). Following the method described above, for the second stage, the total reflectivity model is divided to two sub-models (Fig. 15). For Fig. 15a, there are 161 x 601 nodes, and for Fig. 15b, there are 161 x 600 nodes, keeping the same structure for the total model and sub-models. For comparison, a filter with $\theta = 15°$ is applied to Fig. 14 to remove the artifact. However, the valid steeply dipping structures are also removed (Fig. 16). When applying the slim approximate Hessian to Fig. 14, we retain these valid structures but remove the artifacts (Fig. 17).

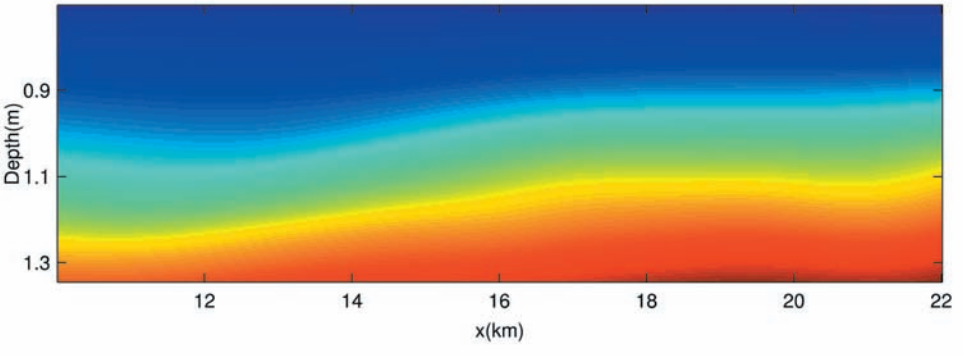

Fig. 13. Initial velocity profile.

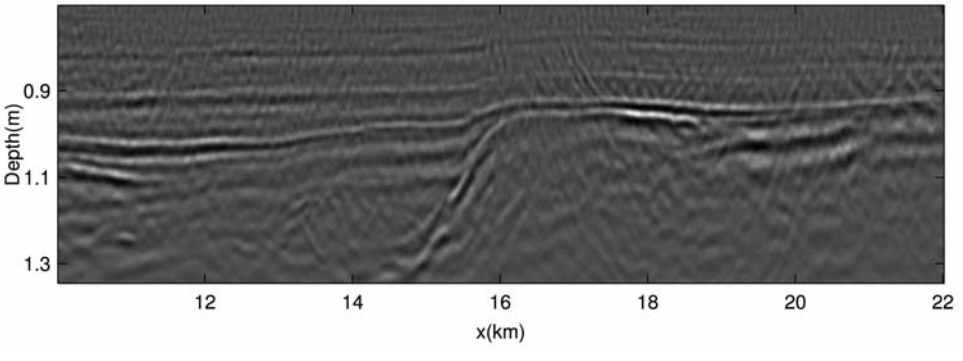

Fig. 14. Reflectivity profile using operator $F^*\bar{d}$. There are 161 x 1201 nodes.

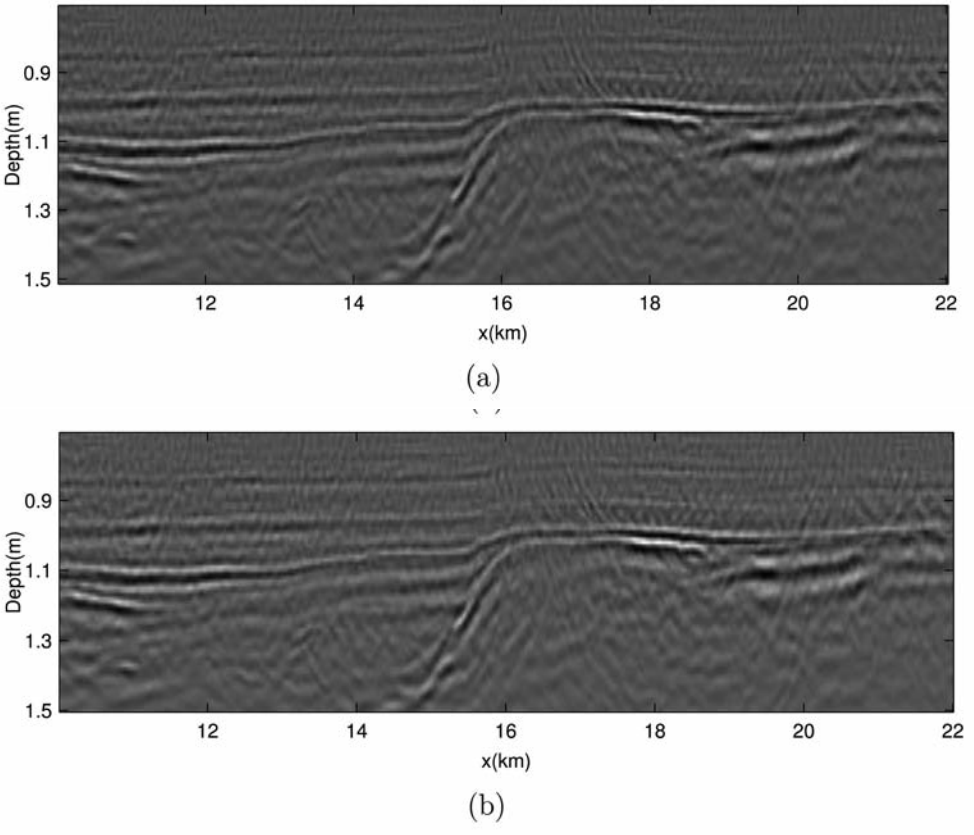

Fig. 15. Sub-models: (a) sub-model 1, nodes 161 x 601 and (b) sub-model 2, nodes 161 x 600.

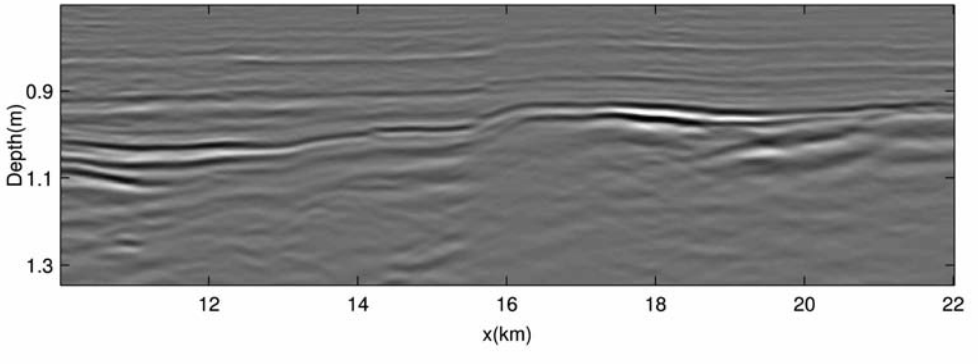

Fig. 16. The gradient for perturbation inversion with filter.

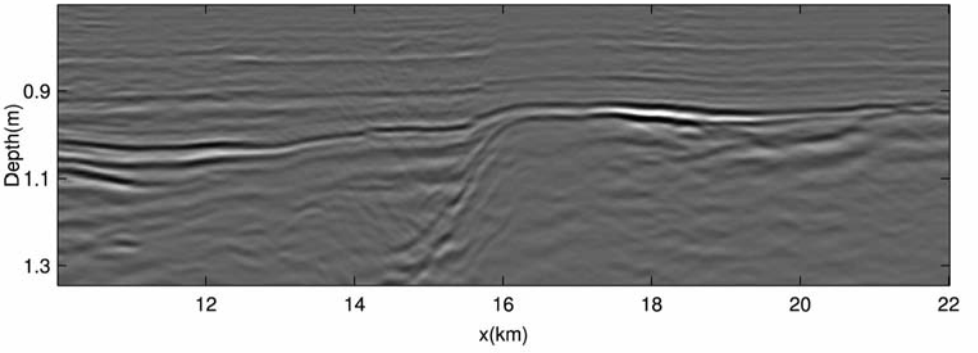

Fig. 17. The gradient for perturbation inversion with the slim approximate Hessian.

Without any preconditioning the inversion velocity is mix with the artifacts (Fig. 18). The inversion velocity obtained by steepest decent method, filtering method and the slim approximate method are shown in Fig. 19. The inversion velocity by the steepest descent method keeps the steeply dipping structure (around x = 16 km), but introduces too many non-geological artifacts. With the filtering method, the high horizontal wavenumber components are reduced; however, this method is not workable for the lateral geological structure. The slim approximate Hessian method retains the information in off-diagonal elements, but removes the artifacts caused by incomplete data. Furthermore, the steeply dipping geological structures are also recovered.

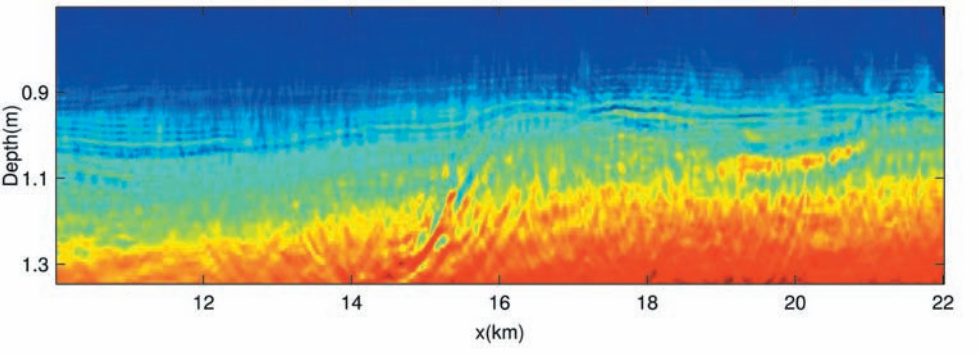

Fig. 18. Inversed velocity by full waveform inversion with no preconditioning.

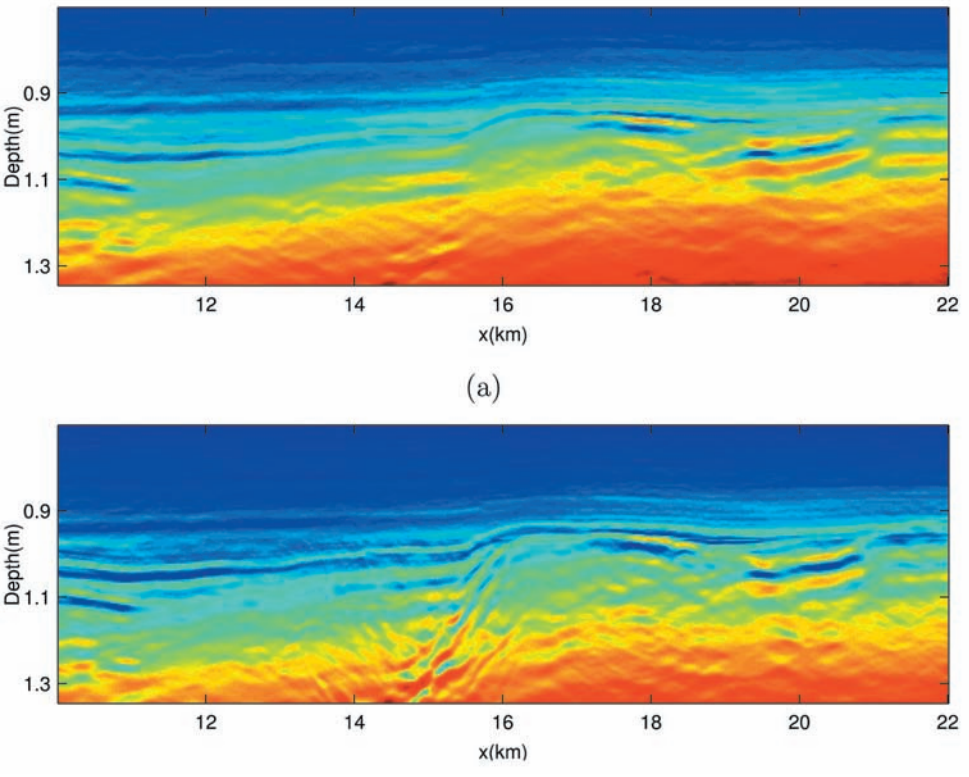

Fig. 19. Inversed velocity by full waveform inversion with (a) a filter, and (b) a slim approximate Hessian.

CONCLUSIONS

The reflection FWI makes use of high-resolution seismic waves and achieves the perturbation model inversion. The widely used gradient-based methods are very efficient due to the ease of computing gradients using the adjoint-state method. However, for incomplete reflection data, the inversion velocities are always mixed with non-geological artifacts. The strong off-diagonal elements of approximate Hessians are important to remove the artifacts using reflection FWI with incomplete data. In order to reduce the cost involved in approximate Hessian calculations, a slim approximate Hessian method is proposed. By this method, the total model is divided into small sub-models, and a slim approximate Hessian is implemented sequentially. A filtering method was introduced for comparison, but the filtering method could not match the slim approximate Hessian for a complex geological structure with steep dip. When applying the slim approximate Hessian method to field seismic

data, we obtain a structurally continuous and higher-resolution velocity inversion as effectively as the approximate Hessian, but more cost-effectively and with lower memory requirements. This method represents a practical solution for velocity inversion that offers an effective alternative to prohibitively expensive calculations.

REFERENCES

Connolly, P., 1999. Elastic impedance. The Leading Edge, 18: 438-452.

Connolly, P., 2007. Reflections on elastic impedance. The Leading Edge, 26: 1360-1368.

Fichtner, A., 2010. Full Seismic Waveform Modelling and Inversion. Advances in Geophysical and Environmental Mechanics and Mathematics. Springer Verlag, Berlin, Heidelberg.

Golub, G.H. and Pereyra, V., 1973. The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. SIAM J. Numer. Anal.,10: 413-432.

Hale, D., 2009. Image-guided blended neighbor interpolation. CWP Report.

Hampson, D.P., Russell, B.H. and Bankhead, B., 2012. Simultaneous inversion of pre-stack seismic data. Expanded Abstr., 82nd Ann. Internat. SEG Mtg., Las Vegas: 1633-1637.

Lailly, P., 1983. The seismic inverse problem as a sequence of before stack migrations. Conf. on Inverse Scattering: Theory and Scattering, Tulsa, OK. SIAM, Philadelphia.

Lambaré, G., 2008. Stereotomography. Geophysics, 73(5): VE25-VE34.

Lambaré, G. and Alerini, M., 2005. Semi automatic PP-PS Stereotomography: application to the synthetic Valhall dataset. Expanded Abstr., 75th Ann. Internat. SEG mtg., Houston: 943-946.

Liu, G.-R. and Gu, Y.T., 2005. An Introduction to Meshfree Methods and Their Programming. Springer Verlag, Heidelberg.

Liu, G.-R., Zhang, G.Y., Gu, Y.T. and Wang, Y.Y., 2005. A meshfree radial point interpolation method (RPIM) for three-dimensional solids. Computat. Mechan., 36: 421-430.

Ma, Y., 2012. Waveform-based velocity estimation from reflection seismic data. Ph.D. thesis, Colorado School of Mimes, Boulder.

Nemeth, T., Wu, C. and Schuster, G.T., 1999. Least-squares migration of incomplete reflection data. Geophysics, 64: 208-221.

Plessix, R.E., Baeten, G. and de Maag, J.W., 2010. Application of acoustic full waveform inversion to a low-frequency large-offset land data set. Expanded Abstr., 80th Ann. Internat. SEG Mtg., Denver.

Plessix, R.E., Baeten, G., de Maag, J.W., ten Kroode, F. and Rujie, Z., 2012. Full waveform inversion and distance separated simultaneous sweeping: a study with a land seismic data set. Geophys. Prosp., 60: 733-747.

Pratt, G., Shin, C. and Hicks, G.J., 1998. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. Geophys. J. Internat., 133: 341-362.

Pratt, R.G. and Worthington, M.H., 1990. Inverse theory applied to multi-source crosshole tomography. Part 1: acoustic wave-equation method. Geophys. Prosp., 38: 287-310.

Ravaut, C., Operto, S., Improta, L., Virieux, J., Herrero, A. and Dell'Aversana, P., 2004. Multiscale imaging of complex structures from multifold wide-aperture seismic data by frequency-domain full-waveform tomography: application to a thrust belt. Geophys. J. Internat., 159: 1032-1056.

Ren, H., Wu, R.S. and Wang, H., 2012. Least square migration with Hessian in the local angle domain. Expanded Abstr., 82nd Ann. Internat. SEG Mtg., Las Vegas: 3010-3014.

Shin, C. and Cha, Y.H., 2008. Waveform inversion in the Laplace domain. Geophys. J. Internat., 173: 922-931.

Shipp, R.M. and Singh, S.C., 2002. Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data. Geophys. J. Internat., 151: 325-344.

Sirgue, L. and Pratt, R.G., 2004. Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies. Geophysics, 69: 231-248.

Symes, W.W. and Kern, M., 1994. Inversion of reflection seismograms by differential semblance analysis: algorithm structure and synthetic examples. Geophys. Prosp., 42: 565-614.

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

van Leeuwen, T. and Mulder, W.A., 2009. A variable projection method for waveform inversion. Extended Abstr., 71st EAGE Conf., Amsterdam.

van Leeuwen, T. and Mulder, W.A., 2010. A comparison of seismic velocity inversion methods for layered acoustics. Inverse Probl., 26 (1), 015008.

Vigh, D., 2012. Velocity determination for land data by full waveform inversion. Extended Abstr., 74th EAGE Conf., Copenhagen: 1-5.

Virieux, J. and Operto, S., 2009. An overview of full-waveform inversion in exploration geophysics. Geophysics, 74(6): WCC1-WCC26.

Xu, S., Wang, D., Chen, F. and Zhang, Y., 2012. Full waveform inversion for reflected seismic data. Extended Abstr., 74th EAGE Conf., Copenhagen.

APPENDIX

The objective function can be written as

$$J[m] = \tfrac{1}{2} \left\| F[m]m_p[m] - \bar{d}[m] \right\|^2 , \tag{A-1}$$

where $m_p[m]$ and $\bar{d}[m]$ are dependent on the background, m. For a given perturbation m_p, we obtain the misfit (Symes and Kern, 1994):

$$\begin{aligned} \delta J[m]\big|_{\overline{m_p}} &= \langle D_m F[m]\delta m, F[m]m_p[m] - \bar{d}[m] \rangle \\ &= \langle \delta m, D_m F^*[m, m_p[m]](F[m]m_p[m] - \bar{d}[m]) \rangle . \end{aligned} \tag{A-2}$$

The gradient for the background component is

$$\delta J[m]/\delta m = D_m F^*[m, m_p[m]](F[m]m_p[m] - \bar{d}[m]) , \tag{A-3}$$

which performs as the second order in the Hessian Matrix.

According the Sobolev scale of Hilbert norms, the s-th Sobolev norm is given by Symes and Kern (1994) as

$$\left\| \delta m \right\| = \left\| \Lambda^{-2s} \delta m \right\| , \tag{A-4}$$

where Λ^{-2s} is a smooth operator, such as the inverse of the Laplace operator.

In order to calculate J[m], we need to estimate the perturbation velocity $m_p[m]$:

$$(\delta J[m]/\delta m)\big|_m = F^*[m](F[m]m_p[m] - \bar{d}[m]) \quad . \tag{A-5}$$

Setting the gradient to zero gives

$$F^*[m]F[m]m_p[m] = F^*[m]\bar{d}[m]) \quad , \tag{A-6}$$

and we then obtain the perturbation component as

$$m_p[m] = (F^*[m]F[m])^{-1}F^*[m]\bar{d}[m] = (H_a)^{-1}F^*[m]\bar{d}[m] \quad . \tag{A-7}$$