

# Wave-equation dispersion inversion

Jing Li, Zongcai Feng and Gerard Schuster

*King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia. E-mail: [inter.lijing@gmail.com](mailto:inter.lijing@gmail.com)*

Accepted 2016 December 8. Received 2016 December 7; in original form 2016 January 3

## SUMMARY

We present the theory for wave-equation inversion of dispersion curves, where the misfit function is the sum of the squared differences between the wavenumbers along the predicted and observed dispersion curves. The dispersion curves are obtained from Rayleigh waves recorded by vertical-component geophones. Similar to wave-equation traveltime tomography, the complicated surface wave arrivals in traces are skeletonized as simpler data, namely the picked dispersion curves in the phase-velocity and frequency domains. Solutions to the elastic wave equation and an iterative optimization method are then used to invert these curves for 2-D or 3-D *S*-wave velocity models. This procedure, denoted as wave-equation dispersion inversion (WD), does not require the assumption of a layered model and is significantly less prone to the cycle-skipping problems of full waveform inversion. The synthetic and field data examples demonstrate that WD can approximately reconstruct the *S*-wave velocity distributions in laterally heterogeneous media if the dispersion curves can be identified and picked. The WD method is easily extended to anisotropic data and the inversion of dispersion curves associated with Love waves.

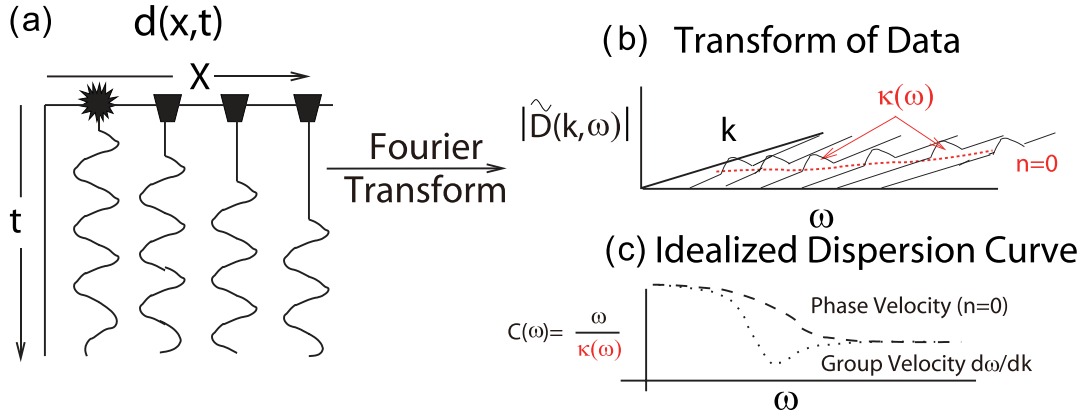
**Key words:** Inverse theory; Surface waves and free oscillations; Seismic tomography.

## 1 INTRODUCTION

Inverting surface waves for the *S*-wave velocity model falls into two categories: (1) the classical method of inverting dispersion curves (Evison *et al.* 1959; Park *et al.* 1998; Xia 2014) for a 1-D layered medium, and (2) full waveform inversion (Groos *et al.* 2014; Solano *et al.* 2014; Dou & Ajo-Franklin 2014; Bohlen *et al.* 2015; Yuan *et al.* 2015) for 2-D and 3-D media. The classical method accurately inverts for a 1-D *S*-wave velocity model, but becomes less accurate with increasing lateral heterogeneity in the subsurface. The 1-D assumption is not satisfied for some practical applications, so partial remedies are spatial interpolation of 1-D velocity models (Yamanaka & Ishida 1996; Xia *et al.* 1999; Beaty *et al.* 2002; Tian *et al.* 2003; Dal Moro 2015; Pan *et al.* 2016b) and laterally constrained inversion (Socco *et al.* 2010; Bergamo & Socco 2012). In comparison, full waveform inversion (FWI) can theoretically account for any lateral heterogeneity, but it is computationally expensive and can easily get stuck in local minima associated with the objective function (Tarantola 1984). To avoid falling into a local minimum, the initial model should be smooth and time-damping strategies can be used at the early iterations (Sheng *et al.* 2006; Sears *et al.* 2008; Brossier *et al.* 2009; Romdhane *et al.* 2011). However, there are no fail-safe strategies for always avoiding local minima in the context of FWI with surface waves. A partial FWI method is that of Solano *et al.* (2014) who used the magnitude spectra of surface waves as the input data. Results with some synthetic data showed this to sometimes be a robust and efficient method for reconstructing the *S*-wave velocity model at the near surface. Another surface wave inversion strategy is proposed by Yuan *et al.*

(2015), who developed a wavelet multi-scale adjoint method for the joint inversion of both surface and body waves. Synthetic tests showed that this approach can avoid cycle skipping for some models. The role of attenuation in FWI with surface waves was studied by Groos *et al.* (2014). They concluded that the estimation of a priori quality factors is critical for inverting seismic waves in the near-surface zone. Instead of inverting Rayleigh waves, Pan *et al.* (2016a) proposed to invert Love waves in the time domain in order to reconstruct the *S*-wave velocity model.

To avoid the assumption of a layered medium and also mitigate FWI's sensitivity to local minima, we present a skeletonized inversion method that inverts the dispersion curves of surface waves for 2-D or 3-D velocity models. The picked dispersion curves are skeletonized data (Schuster 2015; Zhang *et al.* 2016) that tend to make the objective function simpler, and hence this new method, denoted as wave-equation dispersion inversion (WD), enjoys better convergence properties than FWI. This is similar to wave-equation traveltime inversion (WT; Luo & Schuster 1991b), except that picked dispersion curves rather than picked traveltimes are the input data. The WD procedure is more robust than FWI because it replaces complicated surface wave arrivals with simple dispersion curves in the wavenumber-frequency  $k - \omega$  or phase-velocity-frequency  $C(\omega) - \omega$  domains (Fig. 1). In addition, the surface-wave inversion method presented in this paper is the adjoint-state implementation of WD, which is different from Zhang *et al.* (2015, 2016), who used a difference approximation to calculate the misfit gradient model. Hence, our WD method is more general and more than an order-of-magnitude faster for complicated models.



**Figure 1.** (a) Common shot gather  $d(x, t)$ , (b) the fundamental ( $n = 0$ ) dispersion curves for Rayleigh waves in the  $k - \omega$  domain and (c)  $C(\omega) - \omega$  domains. Here the phase velocity is  $C(\omega) = \omega/\kappa(\omega)$  and  $\kappa(\omega)$  is the skeletonized data (Zhang *et al.* 2015; Schuster 2017).

In the following sections, we first present the theory of WD and a workflow for its practical implementation. The WD method is then validated with tests on both synthetic data and field data. The field data are from 2-D surveys conducted in Saudi Arabia and Tanzania, East Africa. The final section presents a summary and discussion.

## 2 THEORY OF SKELETONIZED INVERSION OF SURFACE WAVES

The key steps in WD are similar to the WT algorithm proposed in Luo & Schuster (1991a), except the WD algorithm inverts for the  $S$ -wave velocity model  $v_s$  from the dispersion curves by executing the following steps. (1) Pick the dispersion curve and compute the dispersion-curve misfit function, (2) define a connective function that connects the observed dispersion curve residuals with the predicted data; this allows for the derivation of the Fréchet derivative of the picked wavenumbers with respect to the shear-slowness model, (3) compute the gradient of the misfit function with respect to the shear-slowness model using the elastic wave equation, and (4) update the  $S$ -wave velocity  $v_s$  model with the iterative steepest-descent method or conjugate-gradient method.

The mathematical details for the WD method are now presented, where  $d(\mathbf{g}, t)$  denotes the input shot gathers of vertical particle-velocity traces excited by a vertical-component force at  $\mathbf{s} = (x_s, 0)$  on the surface and recorded at  $\mathbf{g} = (x_g, 0)$ . To reduce notational clutter, the dependency of  $d(\mathbf{g}, t)$  on the source location  $\mathbf{s}$  is silent and we assume a 2D medium and a 2D seismic survey. We utilize the high-resolution linear Radon transform (LRT) method to get  $\tilde{D}(k, \omega)$  in the wavenumber-frequency domain (Luo *et al.* 2008). Here, the skeletonized data consist of the picked dispersion curve  $\kappa(\omega)$  of the fundamental Rayleigh mode, which is obtained by taking the maximum of the magnitude spectrum of  $\tilde{D}(k, \omega)$  shown as the red dashed line in Fig. 1b. This method can be easily adapted to inverting higher-order modes and different components of the recorded data. It is also straightforward to adapt the methodology to the inversion of dispersion curves associated with Love waves.

(i) *Misfit function.* The WD method inverts for the  $S$ -wave velocity model which predicts observed data that minimize the dispersion misfit function  $\epsilon$ :

$$\epsilon = \frac{1}{2} \sum_{\omega} \overbrace{(\kappa(\omega) - \kappa(\omega)^{\text{obs}})^2}^{\text{residual} = \Delta\kappa(\omega)}. \quad (1)$$

Here,  $\kappa(\omega)$  represents the predicted dispersion picked from the simulated spectrum  $\tilde{D}(k, \omega)$  and  $\kappa(\omega)^{\text{obs}}$  denotes the dispersion curve obtained from the recorded spectrum  $\tilde{D}(k, \omega)^{\text{obs}}$ .

The gradient  $\gamma(\mathbf{x})$  is given by

$$\gamma(\mathbf{x}) = \frac{\partial \epsilon}{\partial s(\mathbf{x})} = \sum_{\omega} \Delta\kappa(\omega) \frac{\partial \kappa(\omega)}{\partial s(\mathbf{x})}, \quad (2)$$

so that the optimal shear-slowness model  $s(\mathbf{x})$  is obtained from the steepest-descent formula:

$$s(\mathbf{x})^{(k+1)} = s(\mathbf{x})^{(k)} - \alpha \sum_{\omega} \Delta\kappa(\omega) \frac{\partial \kappa(\omega)}{\partial s(\mathbf{x})}, \quad (3)$$

where  $\alpha$  is the step length by any backtracking line-search method (Nocedal & Wright 1999) and the superscript  $(k)$  denotes the  $k$ th iteration. For pedagogical simplicity, we assume a single shot gather, but the misfit function includes an additional summation over different shot gathers if more than one shot gather is used. In practice, a preconditioned conjugate gradient method can be used (Luo & Schuster 1991b).

(ii) *Connective function.* The analytic expression for the Fréchet derivative  $\frac{\partial \kappa(\omega)}{\partial s(\mathbf{x})}$  in eq. (2) is obtained by forming a connective function (see Appendix A) that relates the residual  $\Delta\kappa(\omega)$  to the shear-slowness model  $s(\mathbf{x})$ . This connective function is defined to be the cross-correlation between the predicted  $\tilde{D}(k, \omega)$  and conjugated observed  $\tilde{D}(k, \omega)_{\text{obs}}^*$  spectra in the  $k - \omega$  domain (Luo & Schuster 1991a,b):

$$\Phi(\Delta\kappa, s(\mathbf{x})) = \text{Real} \left\{ \int \tilde{D}(k + \Delta\kappa, \omega)_{\text{obs}}^* \tilde{D}(k, \omega) dk \right\}, \quad (4)$$

where  $\Delta\kappa$  is the lag variable along the wavenumber axis. Let  $\Delta\kappa(\omega)$  denote the dispersion curve residual for the actual background  $v_s$  model, in which case, the derivative of  $\Phi$  with respect to  $\Delta\kappa(\omega)$  should be zero. That is,

$$\dot{\Phi}(\Delta\kappa, s(\mathbf{x})) = \text{Real} \left\{ \int \dot{\tilde{D}}(k + \Delta\kappa, \omega)_{\text{obs}}^* \tilde{D}(k, \omega) dk \right\} = 0, \quad (5)$$

where  $\dot{\tilde{D}}(k, \omega)_{\text{obs}} = \frac{\partial \tilde{D}(k, \omega)_{\text{obs}}}{\partial k}$  and  $\Delta\kappa$  is now the wavenumber lag that aligns the predicted and observed spectra with one another for a specified  $\omega$ . This wavenumber lag is also equal to the dispersion curve residual in eq. (1), so eq. (5) connects the slowness model with the dispersion curve which will be used to derive the Fréchet derivative  $\frac{\partial \kappa(\omega)}{\partial s(\mathbf{x})}$ .

(iii) *Fréchet derivative.* For eq. (5), the implicit function theorem implies that  $\kappa$  is an implicit function of  $s(\mathbf{x})$  so that

$$d\dot{\Phi} = \frac{\partial \dot{\Phi}}{\partial s} ds + \frac{\partial \dot{\Phi}}{\partial \kappa} d\kappa = 0. \quad (6)$$

Rearranging this equation gives the Fréchet derivative

$$\frac{\partial \kappa}{\partial s(\mathbf{x})} = -\frac{\partial \dot{\Phi}/\partial s}{\partial \dot{\Phi}/\partial \kappa}, \quad (7)$$

where the denominator is the normalization term

$$A = \frac{\partial \dot{\Phi}(\Delta\kappa, s(\mathbf{x}))}{\partial \kappa} = \text{Real} \left\{ \int \ddot{D}(k + \Delta\kappa, \omega)_{\text{obs}}^* \ddot{D}(k, \omega) dk \right\}, \quad (8)$$

and the numerator is

$$\frac{\partial \dot{\Phi}(\Delta\kappa, s(\mathbf{x}))}{\partial s(\mathbf{x})} = \text{Real} \left\{ \int \ddot{D}(k + \Delta\kappa, \omega)_{\text{obs}}^* \frac{\partial \ddot{D}(k, \omega)}{\partial s(\mathbf{x})} dk \right\}. \quad (9)$$

Plugging eqs (8) and (9) into eq. (7) gives the Fréchet derivative:

$$\begin{aligned} \frac{\partial \kappa}{\partial s(\mathbf{x})} &= -\frac{\partial \dot{\Phi}/\partial s}{\partial \dot{\Phi}/\partial \kappa} = \frac{-1}{A} \text{Real} \left\{ \int \ddot{D}(k + \Delta\kappa, \omega)_{\text{obs}}^* \frac{\partial \ddot{D}(k, \omega)}{\partial s(\mathbf{x})} dk \right\} \\ &= \frac{-1}{A} \text{Real} \left\{ \int \hat{D}(\mathbf{g}, \omega)_{\text{obs}}^* \frac{\partial D(\mathbf{g}, \omega)}{\partial s(\mathbf{x})} d\mathbf{x}_g \right\}. \end{aligned} \quad (10)$$

Eq. (10) follows by eq. (A2) in Appendix A where  $\hat{D}(\mathbf{g}, \omega)_{\text{obs}}^*$  is the weighted conjugated data defined in Appendix A.

$$\hat{D}(x_g, \omega)_{\text{obs}}^* = \frac{-ix_g e^{-i\Delta\kappa x_g}}{2\pi} D(x_g, \omega)_{\text{obs}}^*, \quad (11)$$

where the geophone locations  $\mathbf{g} = (x_g, 0)$  are on the horizontal free surface, and  $D(\mathbf{g}, \omega)$  is the inverse Fourier transform of  $\ddot{D}(k, \omega)$ . The Fréchet derivative  $\frac{\partial D(\mathbf{g}, \omega)}{\partial s(\mathbf{x})}$  is derived in Appendix B.

For Rayleigh waves, we assume that only the vertical-component of particle velocity  $D(\mathbf{g}, \omega)$  is recorded so its Fréchet derivative with respect to the shear slowness  $s(\mathbf{x})$  is given as eq. (B6):

$$\begin{aligned} \frac{\partial D(\mathbf{g}, \omega)}{\partial s(\mathbf{x})} &= -4s(\mathbf{x})\rho(\mathbf{x}) \left\{ \frac{\partial G(\mathbf{g}|\mathbf{x})_{zz}}{\partial z} \frac{\partial D(\mathbf{x}, \omega)_x}{\partial x} \right. \\ &\quad + \frac{\partial G(\mathbf{g}|\mathbf{x})_{xz}}{\partial x} \frac{\partial D(\mathbf{x}, \omega)_z}{\partial z} \\ &\quad \left. - \frac{1}{2} \left( \frac{\partial G(\mathbf{g}|\mathbf{x})_{xz}}{\partial z} + \frac{\partial G(\mathbf{g}|\mathbf{x})_{zz}}{\partial x} \right) \right. \\ &\quad \left. \times \left( \frac{\partial D(\mathbf{x}, \omega)_z}{\partial x} + \frac{\partial D(\mathbf{x}, \omega)_x}{\partial z} \right) \right\}, \end{aligned} \quad (12)$$

where  $G(\mathbf{g}, \omega|_{\mathbf{x}})_{xz}$ , and  $G(\mathbf{g}, \omega|_{\mathbf{x}})_{zz}$  are two components of the harmonic Green's tensor (Aki & Richards 1980) for a vertical point-source displacement. Plugging eq. (12) into eq. (10) gives the final form for the Fréchet derivative:

$$\begin{aligned} \frac{\partial \kappa}{\partial s(\mathbf{x})} &= \frac{4s(\mathbf{x})\rho(\mathbf{x})}{A} \text{Real} \left\{ \overbrace{\frac{\partial D(\mathbf{x}, \omega)_x}{\partial x}}^{\text{source}=f(\mathbf{x}, s, \omega)_1} \right. \\ &\quad \times \overbrace{\int \hat{D}(\mathbf{g}, \omega)_{\text{obs}}^* \left( \frac{\partial G(\mathbf{g}|\mathbf{x})_{zz}}{\partial z} \right) d\mathbf{x}_g}^{\text{backprojected data}=B(\mathbf{x}, s, \omega)_1^*} \\ &\quad + \overbrace{\frac{\partial D(\mathbf{x}, \omega)_z}{\partial z}}^{\text{source}=f(\mathbf{x}, s, \omega)_2} \overbrace{\int \hat{D}(\mathbf{g}, \omega)_{\text{obs}}^* \frac{\partial G(\mathbf{g}|\mathbf{x})_{xz}}{\partial x} d\mathbf{x}_g}^{\text{backprojected data}=B(\mathbf{x}, s, \omega)_2^*} \end{aligned}$$

$$\begin{aligned} &\times \overbrace{\int \hat{D}(\mathbf{g}, \omega)_{\text{obs}}^* \left( \frac{\partial G(\mathbf{g}|\mathbf{x})_{xz}}{\partial z} \frac{\partial G(\mathbf{g}|\mathbf{x})_{zz}}{\partial x} \right) d\mathbf{x}_g}^{\text{backprojected data}=B(\mathbf{x}, s, \omega)_3^*} \left\{ \right. \\ &\quad \left. - \frac{1}{2} \left( \frac{\partial D(\mathbf{x}, \omega)_z}{\partial x} + \frac{\partial D(\mathbf{x}, \omega)_x}{\partial z} \right) \right\}, \end{aligned} \quad (13)$$

where  $G(\mathbf{g}|\mathbf{x})_{ij}$  is the Green's tensor for the  $i$ th component of the displacement field at  $\mathbf{x} = (x, z)$ , due to a point source at  $\mathbf{g} = (x_g, 0)$  in the  $j$ th direction,  $s(\mathbf{x})$  is the shear-slowness model for the background medium, and  $A$  is defined in eq. (8). Here,  $f(\mathbf{x}, s, \omega)_i$  for  $i \in (1, 2, 3)$  is the downgoing source field at  $\mathbf{x}$  that originates at  $\mathbf{s}$ , and  $B(\mathbf{x}, s, \omega)_i$  for  $i \in (1, 2, 3)$  represents the backprojected field at  $\mathbf{x}$ . If the membrane model is assumed (Tanimoto 1990; Tromp & Dahlen 1993), then the Fréchet derivative assumes a simplified scalar form similar to that of the acoustic wave equation.

(iv) *Gradient Update.* Plugging eq. (13) into eq. (3) gives the steepest-descent formula for updating the  $S$ -wave velocity model from the dispersion-curve residuals in a shot gather:

$$s(\mathbf{x})^{(k+1)} = s(\mathbf{x})^{(k)} - \alpha \text{Real} \left\{ \overbrace{\sum_{\omega} \sum_{i=1}^3 \frac{4s(\mathbf{x})\rho(\mathbf{x})\Delta\kappa(\omega)}{A} f(\mathbf{x}, s, \omega)_i B(\mathbf{x}, s, \omega)_i^*}^{\text{misfit gradient for CSG}} \right\}. \quad (14)$$

The above equation says that the slowness update is computed by a weighted zero-lag correlation between the backprojected data and the source fields. In the space-time domain, the modified source wavelet is defined as  $\Delta\kappa(\omega)W(\omega)$ . If there is more than one source, then there will be an additional summation over source coordinates.

In summary, the slowness model is updated by first calculating the magnitude spectra of the dispersion curves for both the observed and predicted data, then  $\Delta\kappa(\omega)$  is computed for each shot gather. The weighted observed data  $\hat{D}(\mathbf{g}, \omega)_{\text{obs}}$  for each source are migrated, where the forward propagated source has the weighted source spectrum  $W(\omega)\Delta\kappa(\omega)$ . The gradients for each migrated shot gather are added together to get the slowness update.

### 3 WORK FLOW FOR WAVE-EQUATION DISPERSION INVERSION

Fig. 2 depicts the workflow for the WD method.

(i) Use window muting to remove the early-arrival body waves, backscattered data and higher modes of the Rayleigh waves in the shot gather and then apply a 1-D Fourier transform along the time axis of the shot gather to get its frequency-domain spectrum. The same muting is applied to both the observed data and the predicted shot gather computed by a finite-difference solution to the elastic wave equation.

(ii) Apply the LRT to the spectra of predicted and observed data to get the phase-velocity image in the  $\omega - C$  domain (Luo *et al.* 2008), where  $C$  is the phase velocity of the surface waves. The fundamental dispersion-curves are automatically picked according to the maximum amplitudes of the magnitude spectrum. However, in some field data, the dispersion curves still contain higher-order modes, so there will be discontinuities in the dispersion curves. We utilize the adjacent point-smoothing method to remove the residual higher-order modes as follows. Firstly, we compare the phase velocities at adjoining points in the dispersion curves, and if they differ

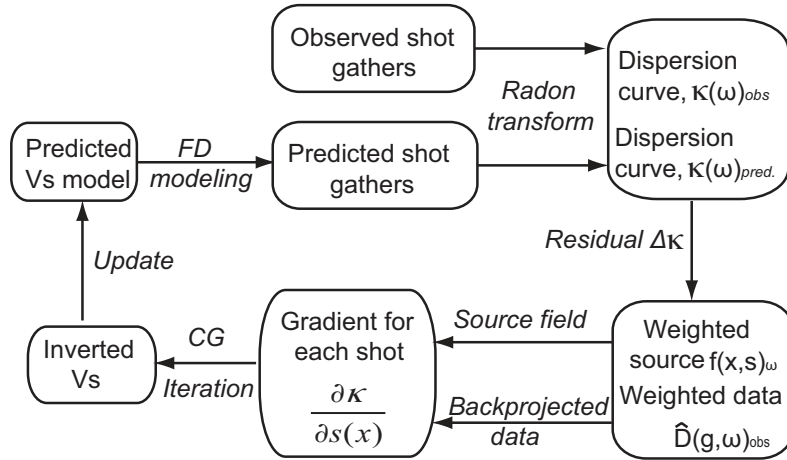


Figure 2. The work flow for implementing the WD method.

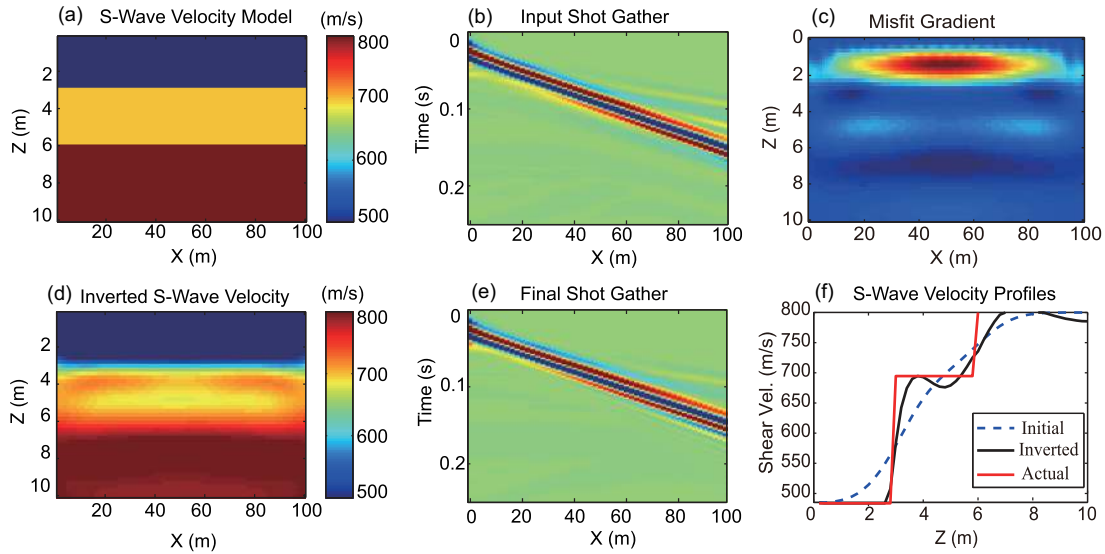


Figure 3. (a) True  $S$ -wave velocity model, (b) common shot gather (CSG), (c) misfit gradient for all of the shot gathers, (d) inverted  $S$ -wave velocity model, (e) CSG associated with the final inverted model and (f)  $S$ -wave velocity profile at  $X = 50$  m.

by a specified threshold, we set the value of the larger one equal to the lower one. Then, the sum of the squared residuals in eq. (1) can be computed from the predicted and observed skeletonized data  $\kappa(\omega)$  and  $\kappa(\omega)^{\text{obs}}$  ( $\kappa(\omega) = \omega/C(\omega)$ ). A limitation to this method is that some of the low-frequency information in the fundamental mode might be lost due to the overlap between fundamental and higher-order modes.

(iii) Use eq. (A3) to calculate the weighted data  $\hat{D}(\mathbf{g}, \omega)_{\text{obs}}$ , which can be used to compute the backprojected data in eq. (13). The forward propagated source field is computed by a finite-difference solution to the elastic wave equation.

(iv) Estimate the step-length  $\alpha$  by any backtracking line-search method (Nocedal & Wright 1999).

(v) The gradients for each migrated shot gather are added together to get the  $S$ -wave slowness update (eq. 14). The background  $S$ -wave slowness model is updated and the above steps are repeated until the residual falls below a specified value.

## 4 NUMERICAL TESTS

The WD method only inverts for the  $S$ -wave velocity model and the initial  $P$ -wave velocity model is defined as  $v_p = \sqrt{3}v_s$ . Three

synthetic data sets and field data from seismic surveys in Saudi Arabia and Tanzania are now inverted by the WD method to demonstrate both its benefits and limitations.

### 4.1 Simple three-layer model

A three-layer model is shown in Fig. 3(a) where the  $S$ -wave velocity increases with depth. For input data, 20 vertical-component shot gathers with a shot spacing of 4 m (see Fig. 3b) are computed by solving the 2-D elastic wave equation with 50 geophones located every 2 m on the surface. The dominant frequency of the source wavelet is 30 Hz with useful frequencies between 10 and 80 Hz. The initial velocity model for WD is the linear gradient model described by the blue dashed line in Fig. 3(f). The dispersion curves are estimated using the procedure described in the previous section. The fundamental-mode dispersion curve is picked for each shot gather in the  $\omega - C$  domain and the associated wavenumber residuals are inverted by the WD method using an iterative gradient solver. The misfit gradient for all of the shot gathers in Fig. 3(c) suggests that the shallow velocity distribution to the depth range of 3–6 m is accurately reconstructed. After 20 iterations, the reconstructed model is shown in Fig. 3(d) where the predicted shot gather in

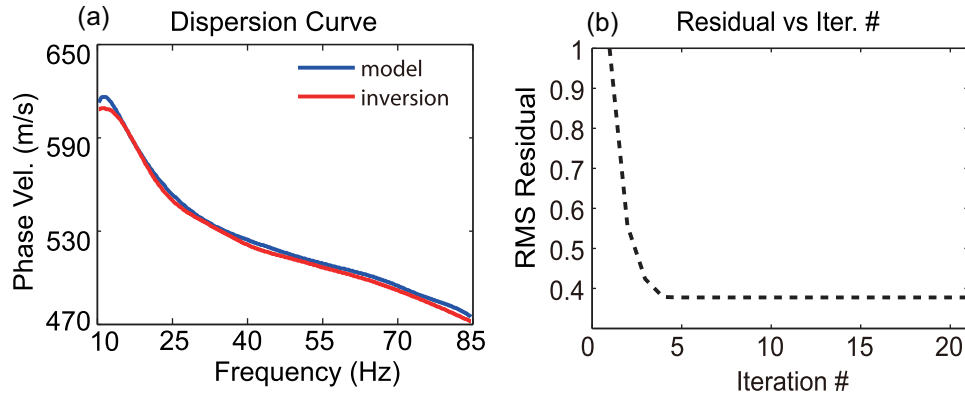


Figure 4. Plots of (a) model and inverted dispersion curves, and (b) the data misfit values after 21 iterations of the WD method.

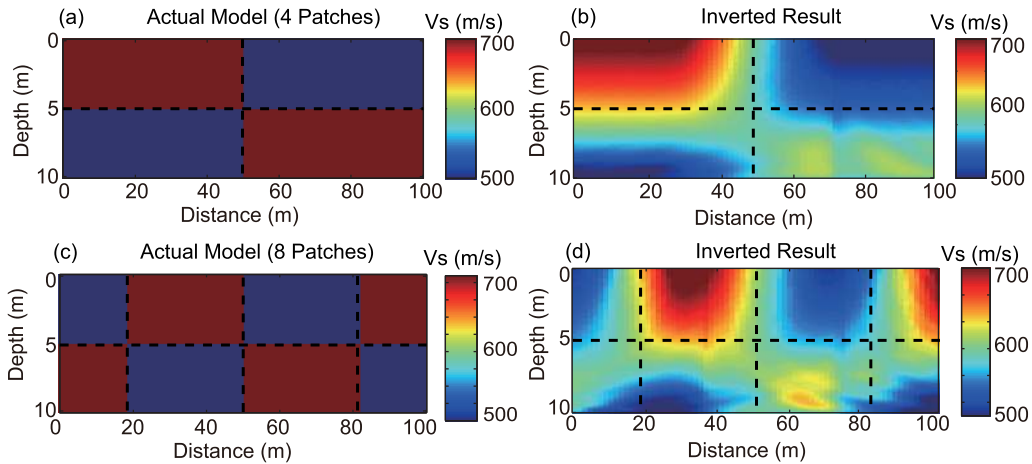


Figure 5. Checkerboard (a,c)  $S$ -wave velocity models. (b,d) The  $S$ -wave velocity tomograms are inverted from the dispersion curves for the (a) and (c) models.

Fig. 3(e) closely resembles the observed shot gather illustrated in Fig. 3(b). This is consistent with Fig. 3(f), where the 1-D inverted  $S$ -wave velocity profile (blue curve) over the centre of the model ( $x = 50$  m) is a good approximation to the true  $S$ -wave velocity profile (red curve). The inverted velocity profile mostly agrees with that of the actual model (Fig. 4a) and Fig. 4(b) shows rapid convergence of the RMS residual with an increase in iteration number.

#### 4.2 Checkerboard model

Elastic shot gathers are computed for two different checkerboard models (see Figs 5a and c) to assess the resolution capabilities of the WD method. The average background  $S$ -wave velocity is  $600 \text{ m s}^{-1}$  in both models. To construct the checkerboard perturbations, the  $S$ -wave velocity is changed by  $\pm 15$  per cent, a 30-Hz Ricker source wavelet is used for the synthetics, and 2 per cent random noise is added to the elastic shot gathers. Here, the wavelength ranges between 16 m and 23 m, with an average value of 20 m. The first checkerboard model has two rectangular velocity anomalies where each one has the area of  $50 \times 5 \text{ m}^2$  (Fig. 5a), while the second one has two layers with a patch size of  $25 \times 5 \text{ m}^2$  (Fig. 5c). There are 25 vertical-component shot gathers computed by solving the 2-D elastic wave equation with 50 geophones located every 2 m and shots every 4 m on the surface.

According to the WD workflow in Fig. 2, the reconstructed  $S$ -wave velocity tomograms are shown in Figs 5(b) and (d) for

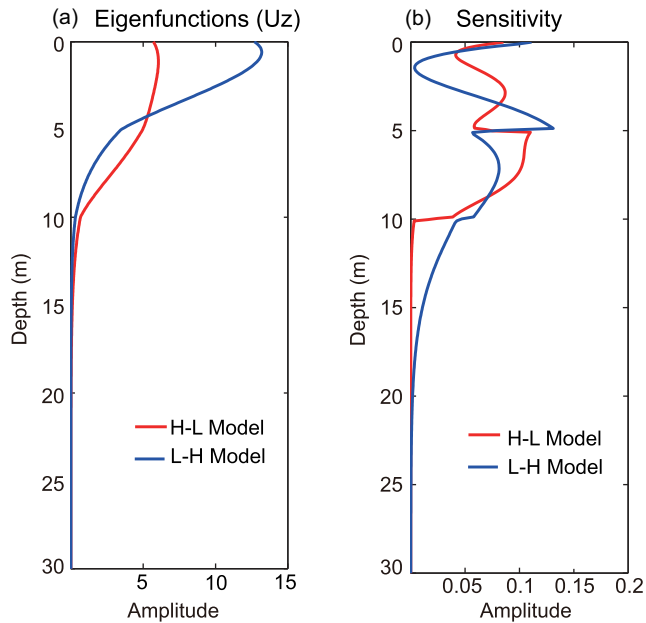
the checkerboard models, where the white dashed lines depict the velocity interfaces. These tomograms suggest that the WD method is able to accurately estimate the  $S$ -wave velocity variations in the horizontal direction, but resolution decreases with depth. This is because surface waves propagate horizontally along the surface and attenuate in depth.

We now employ the vertical displacement eigenfunctions of Rayleigh waves (Denolle *et al.* 2012) and the Jacobian sensitivity matrix (Xia *et al.* 1999) to explain the sensitivity of Rayleigh waves with depth (see Appendix C).

(i) Fig. 6(a) plots the amplitude of the Rayleigh-wave eigenfunctions ( $r_2$ ) against depth for the vertical displacement component at the peak frequency. For the low-to-high-velocity (LH) patches in Figs 5(a) and (c), the eigenfunction amplitudes in the second layer (high) are less than half those in the first layer (low). This suggests that the WD inversion method can best estimate the width and depth of the shallowest velocity anomaly. However, the low-velocity anomalies in the high-to-low-velocity (HL) patch tomogram are more accurately reconstructed than the high-velocity anomalies in the LH patches. Comparison of the sensitivity plots, in the depth range of 5–10 m (second layer), says that the HL model (red line) is characterized by eigenvectors with a larger magnitude than in the LH model.

(ii) Similar to Fig. 6(a), Fig. 6(b) plots a row vector of the Jacobian matrix at the centre frequency. This vector defines the sensitivity of phase velocity values with respect to variations in the  $S$ -wave





**Figure 6.** Sensitivity analysis for checkerboard models at the peak frequency with (a) eigenfunctions associated with the Rayleigh waves and (b) components of the Jacobian matrix, respectively.

velocity at different depths. It confirms that the data associated with the HL checkerboard model are more sensitive to *S*-wave velocity variations than in the LH model in the second layer.

#### 4.3 Low-velocity mineral model

The objective of this synthetic test is to determine how well WD can reconstruct the blue low-velocity anomalies in Fig. 7(a). These

anomalies are based on realistic mineral deposits seen in an open-pit mine. The input data are computed by solving the 2-D elastic wave equation for 100 shot gathers, with the shots evenly distributed on the surface at 2 m intervals; the source wavelet for each shot is a 10-Hz Ricker wavelet. The seismic waves from each shot are recorded by 100 receivers spaced at the same interval as the shots. The ensemble of 1-D WD profiles is displayed as the *S*-wave velocity tomogram in Fig. 7(d). In this case, the wavenumber residuals from each shot gather were inverted assuming a 1-D velocity model beneath the location of each shot. There is a rough but imprecise correspondence between the tomogram and the actual model in Fig. 7(a). To generate a more accurate tomogram, the 2-D WD method is used to invert the data. In this case, only 25 shot gathers are employed with an 8 m shot interval. Fig. 7(b) depicts the initial gradient model, and the fundamental dispersion curve is picked for each shot gather and inverted by the WD method to give the reconstructed model shown in Fig. 7(c). This *S*-wave tomogram shows much better correspondence to the actual model than does the 1-D tomogram in Fig. 7(d).

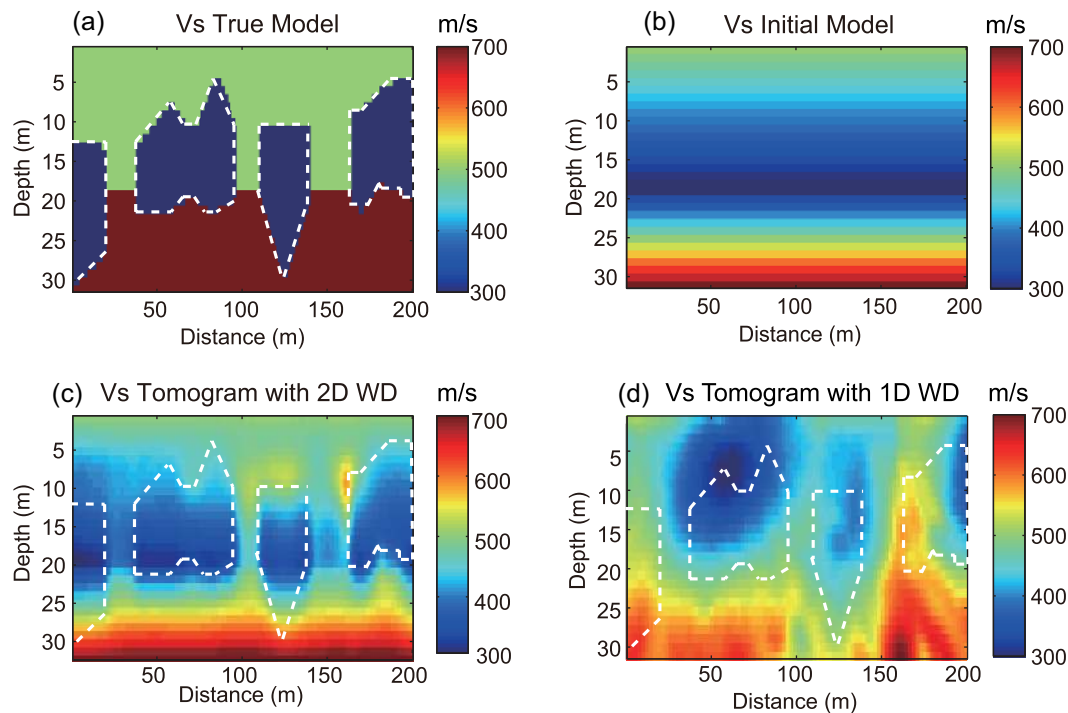
The predicted and observed dispersion curves are plotted against iteration number in Fig. 8. After 25 iterations, the normalized misfit residual decreased to 0.3 and shows an acceptable fit to the data.

## 5 FIELD DATA TESTS

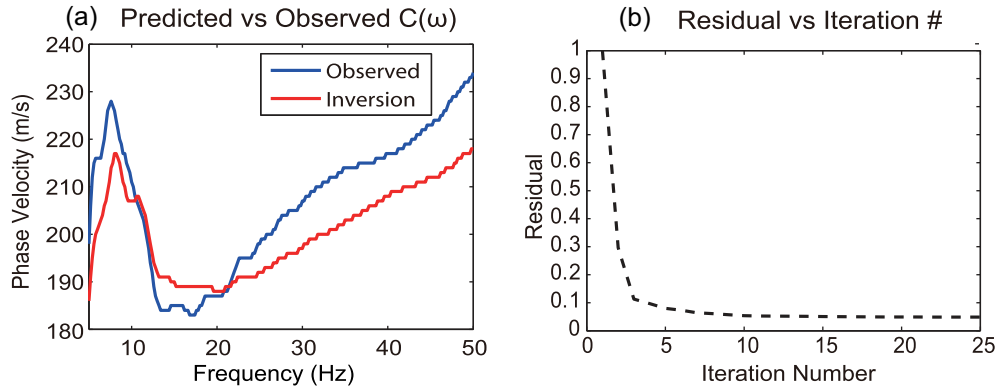
The field data include controlled noise source (CNS) seismic data recorded over the Qademah fault in Saudi Arabia and active source data recorded near Olduvai Gorge in Tanzania.

### 5.1 Qademah fault controlled noise seismic data

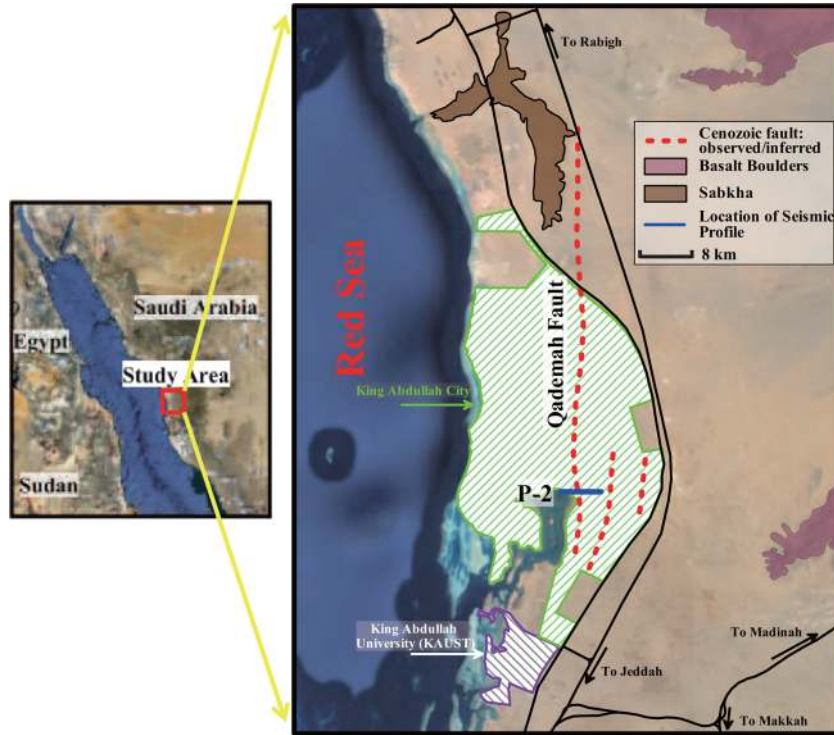
A controlled noise survey is conducted across the Qademah fault. The location of the field experiment is shown in Fig. 9 and the long red dashed line in this figure is the Qademah fault which is aligned



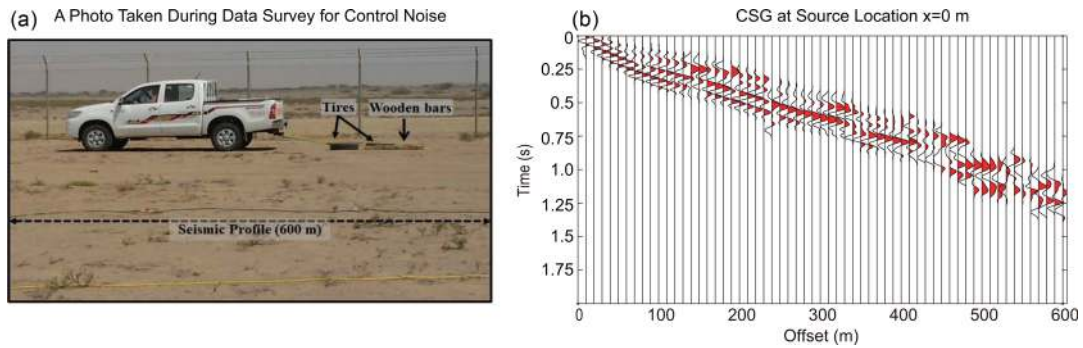
**Figure 7.** (a) *S*-wave velocity true model; (b) initial model; (c) 2-D WD *S*-wave velocity tomogram after 15 iterations; (d) 1-D *S*-wave velocity tomogram.



**Figure 8.** (a)  $C(\omega)$  for predicted (red) and observed (blue) dispersion curves for a CSG, and (b) dispersion misfit residual plotted against iteration number.



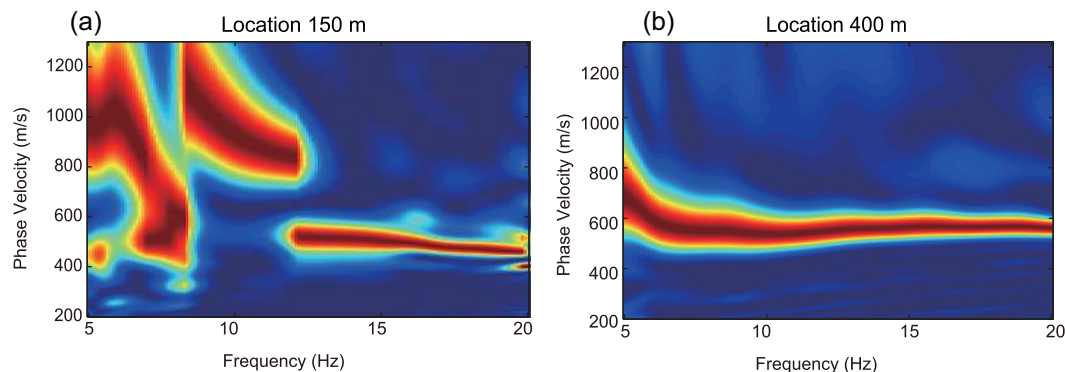
**Figure 9.** The seismic survey line (blue line marked P-2) across the Qadimah fault (Hanafy *et al.* 2015).



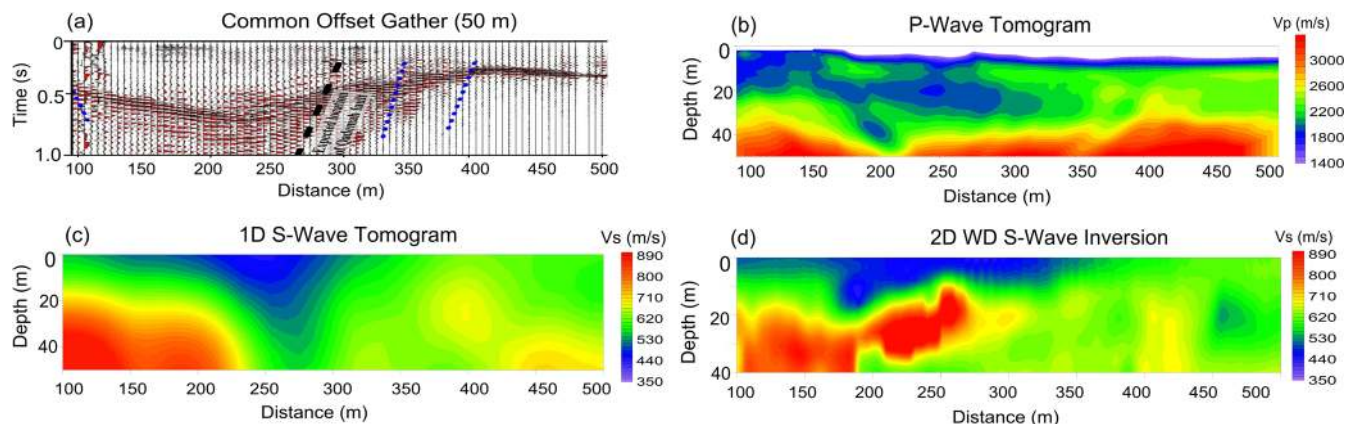
**Figure 10.** (a) Photo of the CNS truck, and (b) common shot gather after applying all the processing steps.

along the north-south direction. The blue solid line is the survey line which consists of 60 receivers at a 10 m spacing with 60 shots recorded at 10 m intervals. The controlled noise is created by a noise-making truck driven around the survey line for 2 hours; the driving speed ranged between 20 and 25 km  $\text{hr}^{-1}$  (Hanafy *et al.* 2015). The

wooden bars and tires are tied behind the truck to create additional noise (Fig. 10a), and the resulting seismic noise is recorded at each of the traces. Then, the traces are broken up into small windows, and each window of arrivals is correlated with the corresponding window of arrivals in other traces to give a virtual CSG (Hanafy



**Figure 11.** Qadimah fault CNS data dispersion curves at different source offsets of (a) 150 m and (b) 400 m.

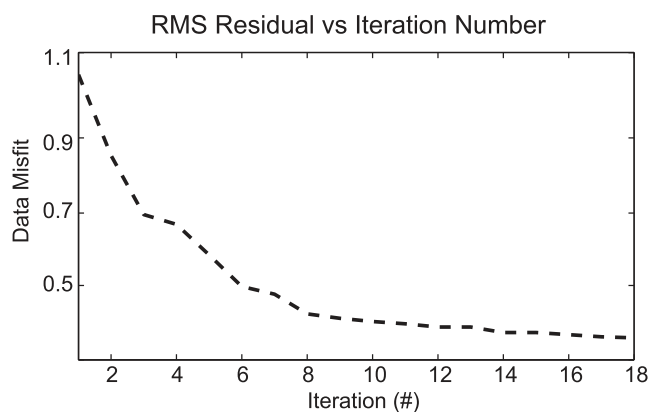


**Figure 12.** Qadimah COG and associated tomograms (Li & Hanafy 2016). (a) common offset gather (COG, offset = 50 m), (b) *P*-wave velocity tomogram inverted from the first-arrival traveltimes, (c) 1-D *S*-wave velocity tomogram inverted from the dispersion-curve data and (d) *S*-wave velocity tomogram inverted by 2-D WD inversion of the wavenumber residuals.

*et al.* 2015). Stacking the virtual CSGs for the same source position gives the virtual shot gather after band-pass filtering and a window mute (Fig. 10b).

A common offset gather (COG) is shown in Fig. 12(a) with the source–receiver offset of 50 m. The dashed lines in Fig. 12(a) indicate the location of the Qadimah fault, which is consistent with the lateral velocity decrease in the *P*-wave velocity tomogram in Fig. 12(b). The *P*-wave velocity tomogram is computed by inverting the *P*-wave first-arrival traveltimes with a ray-based tomography method. Then, all virtual shot gathers are transformed into the  $\omega - C$  domain by a Fourier transform in time and the LRT method, and the maximum energy values of the dispersion curves are picked. The typical dispersion curves at different shot locations (150 and 450 m) show that the fundamental mode is mostly continuous from 5 to 20 Hz (Fig. 11). In addition, the higher-mode dispersion curve in Fig. 11(a) can be removed with the adjacent point smoothing method. According to the detection depth of about  $1/3$  wavelength ( $z = \frac{v}{3f}$ ) (Rix & Leipski 1991), the maximum depth for a reliable estimate of  $v_s$  is about 40 m.

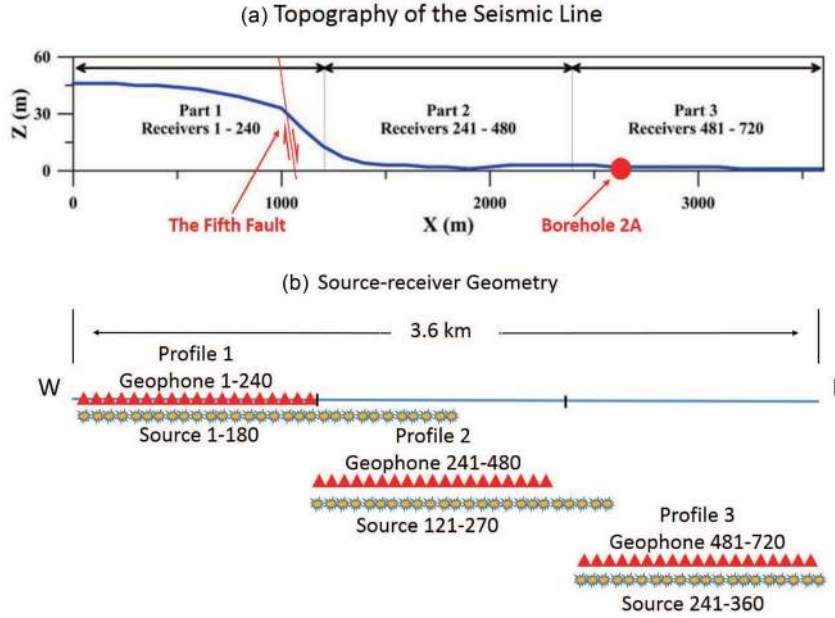
Fig. 12(c) shows the *S*-wave velocity tomogram obtained from 1-D WD inversion (Li & Hanafy 2016). This tomogram roughly estimates the position of the Qadimah fault and the velocity structure is smoothed with low resolution. Then, the 2-D WD method is applied to the picked dispersion curves to give the *S*-wave velocity tomogram in Fig. 12(d) after 18 iterations in Fig. 13, where there is a low-velocity zone on the downthrown side of the fault. This is consistent with the *P*-wave velocity tomogram in Fig. 12(b) and the COG profile in Fig. 12(a) for  $150 \text{ m} < x < 300 \text{ m}$ .



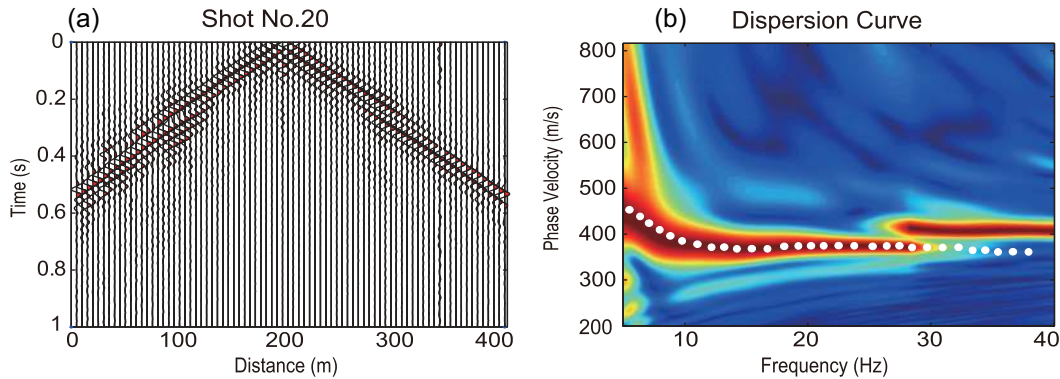
**Figure 13.** Dispersion misfit residual plotted against iteration number for the Qadimah fault seismic data.

Compared with the *P*-wave velocity tomogram, the 2-D WD tomogram has higher horizontal resolution. For example, below the depth of 30 m, the *P*-wave velocity tomogram shows a smoothly varying velocity along the horizontal direction, but the *S*-wave velocity tomogram indicates a strong lateral heterogeneity at the same depth. The dispersion curves at different shot locations (150 and 450 m) in Fig. 11 also show that the phase-velocity strongly varies over the same frequency range. As the surface waves enter the fault zone there is strong dispersion in the surface wave arrivals. It is difficult to assess the accuracy of the 2-D WD tomogram, but it





**Figure 14.** (a) Topography along Lines 1–3, and (b) the source–receiver geometry for profiles 1–3, where the red triangles and yellow stars represent the locations of receivers and sources, respectively.



**Figure 15.** (a) CSG of shot No. 20, and (b) phase velocity image obtained from the linear Radon transform (white dots depict the picked fundamental dispersion-curve).

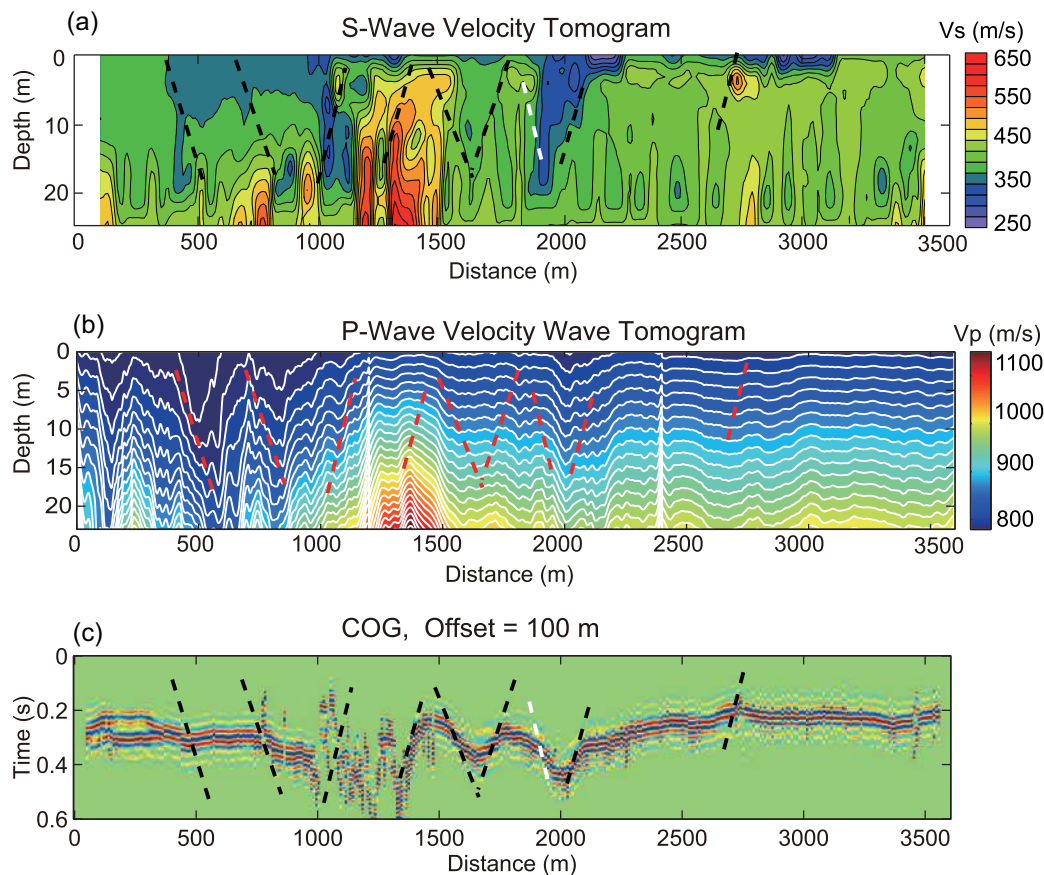
appears to have much more complexity than the simpler 1-D tomogram in Fig. 12(c). In fact, the 1-D tomogram appears to be too simple to fully explain the complexity of events in the COG profile (Fig. 12a).

## 5.2 East Africa data

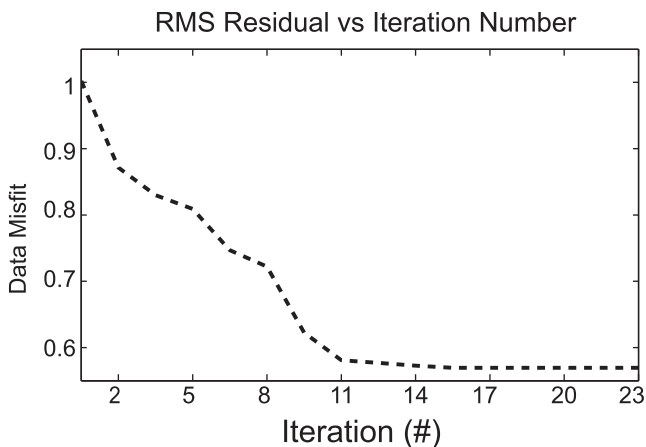
Another field data test is carried out on seismic data recorded over a basin near Olduvai Gorge, Tanzania, which is one of the earliest hominid sites where fossil and stone tool evidence dates back to more than 2.2 million years. A 2-D seismic survey with a total length of 3.6 km is recorded and the survey line straddles the major Fifth Fault. Fig. 14 shows the topography and the source–receiver geometry, respectively, where the Fifth Fault is located at the sharp change in elevation at around  $X = 1000$  m. The survey line consists of 3 profiles, each of which has a length of 1.2 km, with 240 channels at 5 m receiver intervals. In-line shooting is carried out for all 3 profiles, with 120 shots at 10 m intervals for each profile. We use a 200 lb accelerated weight drop source, and stack 30 shot records at each shot location. However, the data suffer severely from a low signal-to-noise ratio caused by the weathering layer and wind

noise. To reduce this noise, we bury geophones at the depth 0.5 m, after which, the data quality is improved. Fig. 15 depicts a typical shot gather and the associated phase velocity image. We can pick the surface-wave fundamental-mode dispersion curve (white dots in Fig. 15b) with the strongest amplitudes and also use the adjacent point-smoothing method mentioned in the workflow section. The fundamental dispersion curves of all shot gathers are picked along the maximum of the phase-velocity spectra. The higher modes also can be removed according to the adjacent point smooth method.

Simultaneously inverting all of the dispersion curves by the 2-D WD method gives the  $S$ -wave velocity tomogram shown in Fig. 16(a), where the dashed lines are interpreted as fault structures. The Fifth Fault and other faults are indicated in the  $S$ -wave velocity tomogram. This diagnosis is consistent with the interpretation of faults by the  $P$ -wave velocity tomogram in Fig. 16(b) and the COG profile in Fig. 16(c). In addition, the  $S$ -wave tomogram computed by the WD method provides a more detailed velocity structure in the shallow regions. Here, a fault is suggested by a sudden delay in the onset of surface waves at the dashed lines. The combination of both the  $S$ -wave and  $P$ -wave velocity tomograms and the COG strongly suggest the presence of fault-like structures which gives



**Figure 16.** (a) *S*-wave velocity tomogram obtained by inverting the dispersion curves in the East Africa data shown in the previous figure, (b) *P*-wave velocity tomogram by wave-equation traveltimes inversion of first arrivals, and (c) COGs at source–receiver offset of 100 m for data recorded by an East Africa survey. The dashed white faults are more reliably interpreted than the dashed red faults.



**Figure 17.** Dispersion misfit residual plotted against iteration number for the East Africa field data.

a new interpretation to the geology in this basin. Fig. 17 plots the normalized data misfit values for 23 iterations.

## 6 SUMMARY

We present the theory for wave-equation inversion of dispersion curves (WD), where the misfit function is the sum of the squared differences between the wavenumbers along the predicted and observed dispersion curves. The *S*-wave velocity model is updated by migrating the weighted data, where the weight is proportional to the

wavenumber residual. It largely overcomes the expense of finding the Fréchet derivative by a finite-difference approximation. Numerical simulations suggest that WD inversion is effective for selected 2-D velocity models where the dispersion curves can be readily identified. These 2-D tomograms are more accurate than the ones inverted by assuming a local 1-D velocity model over each common shot gather. Tests on both the Qadimah fault and East Africa data suggest that velocity models can be inverted to reveal the presence of faults.

A limitation of this method is that it requires picking of the fundamental dispersion-curves in the shot gathers. Accurate picking of the fundamental dispersion-curve can be prone to errors if there is a strong overlap with higher-order modes and other coherent events such as backscattered surface waves. Therefore we recommend separation of the different modes and elimination of noise prior to the application of WD to the wavenumber residuals. Another limitation is the *S*-velocity tomogram will have less resolution than that of a successful FWI tomogram. Therefore, the WD tomogram might be used as a good starting model for FWI.

## ACKNOWLEDGEMENTS

We thank the financial support from the sponsors of the Consortium of Subsurface Imaging and Fluid Modeling (CSIM). We also thank KAUST for providing funding by the CRG grant OCRF-2014-CRG3-2300. For computer time, this research used the resources of the IT Research Computing Group and the Supercomputing Laboratory at KAUST. We thank them for providing the computational resources required for carrying out this work.

## REFERENCES

- Aki, K. & Richards, P.G., 1980. *Quantitative Seismology: Theory and Methods*, vol. 1, W. H. Freeman.
- Beatty, K.S., Schmitt, D.R. & Sacchi, M., 2002. Simulated annealing inversion of multimode Rayleigh wave dispersion curves for geological structure, *Geophys. J. Int.*, **151**, 622–631.
- Bergamo, P., Boiero, D. & Socco, L.V., 2012. Retrieving 2D structures from surface-wave data by means of space-varying spatial windowing, *Geophysics*, **77**, EN39–EN51.
- Bohlen, T., Forbriger, T., Groos, L., Schäfer, M. & Metz, T., 2015. Applications of elastic full waveform inversion to shallow seismic surface waves, in *EGU General Assembly Conference Abstracts*, **17**, 3512.
- Brossier, R., Operto, S. & Virieux, J., 2009. Seismic imaging of complex on-shore structures by 2D elastic frequency-domain full-waveform inversion, *Geophysics*, **74**, 105–118.
- Dal Moro, G., 2015. Joint analysis of Rayleigh-wave dispersion and HVSR of lunar seismic data from the Apollo 14 and 16 sites, *Icarus*, **254**, 338–349.
- Denolle, M.A., Dunham, E.M. & Beroza, G.C., 2012. Solving the surface-wave eigenproblem with Chebyshev spectral collocation, *Bull. seism. Soc. Am.*, **102**, 1214–1223.
- Dou, S. & Ajo-Franklin, J.B., 2014. Full-wavefield inversion of surface waves for mapping embedded low-velocity zones in permafrost, *Geophysics*, **79**, 107–124.
- Evison, F.F., Orr, R. & Ingham, C., 1959. Thickness of the Earth's crust in Antarctica, *Nature*, **183**, 306–308.
- Groos, L., Schäfer, M., Forbriger, T. & Bohlen, T., 2014. The role of attenuation in 2D full-waveform inversion of shallow-seismic body and Rayleigh waves, *Geophysics*, **79**, R24–R261.
- Hanafy, M.S., Altheyab, A. & Schuster, G.T., 2015. Controlled noise seismology, in *2015 SEG Annual Meeting*, Society of Exploration Geophysics, pp. 5102–5107.
- Haskell, N.A., 1953. The dispersion of surface waves on multilayered media, *Bull. seism. Soc. Am.*, **43**, 17–34.
- Lai, C.G. & Rix, G.J., 1998. Simultaneous inversion of Rayleigh phase velocity and attenuation for near-surface site characterization, *PhD thesis*, School of Civil and Environmental Engineering, Georgia Institute of Technology, pp. 114–125.
- Li, J. & Hanafy, M.S., 2016. Skeletonized inversion of surface wave: active source versus controlled noise comparison, *Interpretation*, **3**, 11–19.
- Luo, Y. & Schuster, G.T., 1991a. Wave equation inversion of skeletonized geophysical data, *Geophys. J. Int.*, **105**, 289–294.
- Luo, Y. & Schuster, G.T., 1991b. Wave-equation traveltime inversion, *Geophysics*, **56**, 645–653.
- Luo, Y., Xia, J., Miller, R.D., Xu, Y., Liu, J. & Liu, Q., 2008. Rayleigh-wave dispersive energy imaging using a high-resolution linear Radon transform, *Pure appl. Geophys.*, **165**, 903–922.
- Mora, P., 1988. Nonlinear two-dimensional elastic inversion of multioffset seismic data, *Geophysics*, **52**, 1211–1228.
- Nocedal, J. & Wright, S., 1999. *Numerical Optimization*, Springer Verlag.
- Pan, Y., Xia, J., Xu, Y., Gao, L. & Xu, Z., 2016a. Love-wave waveform inversion in time domain for shallow shear-wave velocity, *Geophysics*, **81**, R1–R14.
- Pan, Y., Xia, J., Xu, Y., Xu, Z., Cheng, F., Xu, H. & Gao, L., 2016b. Delineating shallow S-wave velocity structure using multiple ambient-noise surface-wave methods: an example from western Junggar, China, *Bull. seism. Soc. Am.*, **106**(2), 327–336.
- Park, C. B., Miller, R.D. & Xia, J., 1998. Imaging dispersion curves of surface waves on multi-channel records, *SEG Expanded Abstracts*, 1377–1380.
- Rix, G.J. & Leipski, A.E., 1991. Accuracy and resolution of surface wave inversion: recent advances in instrumentation, data acquisition and testing in soil dynamics, *Geotech.*, **29**, 17–32.
- Romdhane, A., Grandjean, G., Brossier, R., Rejiba, F., Operto, S. & Virieux, J., 2011. Shallow-structure characterization by 2D elastic full-waveform inversion, *Geophysics*, **76**, R81–R93.
- Schuster, G.T., 2015. Skeletonized wave-equation dispersion inversion: dispersion misfit function, 2015 Annual CSIM Report, 62–65.
- Schuster, G.T., 2017. *Seismic Inversion*, SEG Publishing.
- Sears, T.J., Singh, S. & Barton, P., 2008. Elastic full waveform inversion of multi-component OBC seismic data, *Geophys. Prospect.*, **56**, 84–862.
- Sheng, J., Leeds, A., Buddensiek, M. & Schuster, G.T., 2006. Early arrival waveform tomography on near-surface refraction data, *Geophysics*, **71**, U47–U57.
- Socco, L.V., Foti, S. & Boiero, D., 2010. Surface-wave analysis for building near-surface velocity model-established approaches and new perspectives, *Geophysics*, **75**, 83–102.
- Solano, C.P., Donno, D. & Chauris, H., 2014. Alternative waveform inversion for surface wave analysis in 2D media, *Geophys. J. Int.*, **198**, 1359–1372.
- Tanimoto, T., 1990. Modelling curved surface wave paths: membrane surface wave synthetics, *Geophys. J. Int.*, **102**, 89–100.
- Tarantola, A., 1984. Inversion of seismic reflection data in the acoustic approximation, *Geophysics*, **49**, 1259–1266.
- Thomson, W.T., 1950. Transmission of elastic waves through a stratified solid medium, *J. Appl. Phys.*, **21**, 89–93.
- Tian, G., Steeples, D.W., Xia, J. & Spikes, K.T., 2003. Useful resorting in surface-wave method with the autogugge, *Geophysics*, **68**, 1900–1908.
- Tromp, J. & Dahlen, F.A., 1993. Variational principles for surface wave propagation on a laterally heterogeneous earth-III. potential representation, *Geophys. J. Int.*, **112**, 195–209.
- Xia, J., 2014. Estimation of near-surface shear-wave velocities and quality factors using multichannel analysis of surface-wave methods, *J. Appl. Geophys.*, **103**, 140–151.
- Xia, J., Miller, R.D. & Park, C.B., 1999. Estimation of near-surface shear-wave velocity by inversion of Rayleigh waves, *Geophysics*, **64**, 691–700.
- Yamanaka, H. & Ishida, H., 1996. Application of genetic algorithms to an inversion of surface-wave dispersion data, *Bull. seism. Soc. Am.*, **86**, (2), 434–444.
- Yuan, Y.O., Simons, F.J. & Bozda, E., 2015. Multiscale adjoint waveform tomography for surface and body waves, *Geophysics*, **80**, R281–R302.
- Zhang, Z.D., Liu, Y., Schuster, G.T. & Hanafy, S.M., 2015. Wave equation inversion of skeletonized surface waves, in *2015 SEG Annual Meeting*, Society of Exploration Geophysics, pp. 2391–2395.
- Zhang, Z.D., Schuster, G.T., Liu, Y., Hanafy, S.M. & Li, J., 2016. Wave equation dispersion inversion using a difference approximation to the dispersion-curve misfit gradient, *J. Appl. Geophys.*, **133**, 9–15.

## APPENDIX A: CORRELATION IDENTITY

In eq. (9), the functions in the integrand can be replaced by their Fourier transforms:

$$\begin{aligned}\dot{D}(k + \Delta k, \omega)_{\text{obs}} &= \frac{1}{2\pi} \int \left[ i x'_g D(x'_g, \omega)_{\text{obs}} e^{i(k + \Delta k)x'_g} \right] dx'_g, \\ \tilde{D}(k, \omega) &= \frac{1}{2\pi} \int D(x_g, \omega) e^{ikx_g} dx_g, \end{aligned} \quad (\text{A1})$$

to give

$$\begin{aligned}\frac{\partial \Phi(\Delta k, s(\mathbf{x}))}{\partial s(\mathbf{x})} &= \frac{1}{4\pi^2} \text{Real} \left\{ \int dx_g \left\{ \int dx'_g \left[ \int e^{ik(x_g - x'_g)} dk \right] \right. \right. \\ &\quad \times (-ix'_g) D(x'_g, \omega)_{\text{obs}}^* e^{-i\Delta k x'_g} \left. \left. \frac{\partial D(x_g, \omega)}{\partial s(\mathbf{x})} \right\} \right\}, \\ &= \frac{1}{4\pi^2} \text{Real} \left\{ \int dx_g \left\{ \int dx'_g [2\pi \delta(x_g - x'_g)] \right. \right. \\ &\quad \times (-ix'_g) D(x'_g, \omega)_{\text{obs}}^* e^{-i\Delta k x'_g} \left. \left. \frac{\partial D(x_g, \omega)}{\partial s(\mathbf{x})} \right\} \right\}, \\ &= \frac{1}{2\pi} \text{Real} \left\{ \int dx_g \left[ -ix_g D(x_g, \omega)_{\text{obs}}^* e^{-i\Delta k x_g} \right] \frac{\partial D(x_g, \omega)}{\partial s(\mathbf{x})} \right\}, \\ &= \text{Real} \left\{ \int dx_g \left[ \frac{\tilde{D}(x_g, \omega)_{\text{obs}}^*}{2\pi} D(x_g, \omega)_{\text{obs}} \right] \frac{\partial D(x_g, \omega)}{\partial s(\mathbf{x})} \right\}, \end{aligned} \quad (\text{A2})$$

where the weighted conjugated data function is

$$\hat{D}(x_g, \omega)_{\text{obs}}^* = \frac{-ix_g e^{-i\Delta x_g}}{2\pi} D(x_g, \omega)_{\text{obs}}^*. \quad (\text{A3})$$

The offset coordinate  $x_g$  in the argument of  $\hat{D}(x_g, \omega)_{\text{obs}}^*$  in eq. (A3) will be replaced by the vector notation  $\mathbf{g} = (x_g, 0)$ .

## APPENDIX B: ELASTIC GRADIENT

The gradient for the WD method is now derived. For an isotropic heterogeneous medium, the Born approximation in terms of the elastic Green's functions for a harmonic source (Mora 1987) is

$$\begin{aligned} \delta D(\mathbf{x}, \omega)_i &= - \int G(\mathbf{x}|\mathbf{x}')_{ij,j} D_{k,k}(\mathbf{x}', \omega) \delta \lambda(\mathbf{x}') d\mathbf{x}'^3 \\ &\quad - \int G(\mathbf{x}|\mathbf{x}')_{ik,j} (D(\mathbf{x}', \omega)_{k,j} + D(\mathbf{x}', \omega)_{j,k}) \delta \mu(\mathbf{x}') d\mathbf{x}'^3, \end{aligned} \quad (\text{B1})$$

where  $\delta D(\mathbf{x}, \omega)_i$  denotes the  $i$ th component of the perturbed particle velocity recorded at  $\mathbf{x}$  due to the scattering from the perturbations of Lamé parameters  $\delta \lambda$  and  $\delta \mu$ . Einstein notation is assumed where  $D_{i,j} = \frac{\partial D_i}{\partial x_j}$  for  $i, j \in (1, 2)$ .  $G_{ij}$  is the 2-D harmonic Green's tensor (Aki & Richards 1980) for the background medium with the Lamé parameters  $\lambda$  and  $\mu$ , and density  $\rho$ . We assume density  $\rho$  is a constant. Eq. (B1) yields the derivative of  $D(\mathbf{x}, \omega)_i$  with respect to  $\lambda$  and  $\mu$  at  $\mathbf{x}'$

$$\begin{aligned} \frac{\delta D(\mathbf{x}, \omega)_i}{\delta \lambda(\mathbf{x}')} &= -G(\mathbf{x}|\mathbf{x}')_{ij,j} \frac{\partial D(\mathbf{x}', \omega)_k}{\partial x_k}, \\ \frac{\delta D(\mathbf{x}, \omega)_i}{\delta \mu(\mathbf{x}')} &= -G(\mathbf{x}|\mathbf{x}')_{ik,j} \left( \frac{\partial D(\mathbf{x}', \omega)_j}{\partial x_k} + \frac{\partial D(\mathbf{x}', \omega)_k}{\partial x_j} \right). \end{aligned} \quad (\text{B2})$$

For 2-D, our interest is confined to the derivative of the vertical component of the particle velocity at  $\mathbf{g}$ , so eq. (B2) for  $i = 2$  and the derivatives of  $D(\mathbf{g}, \omega) \leftrightarrow D(\mathbf{x}, \omega)_z$  with respect to  $\lambda$  and  $\mu$  at  $\mathbf{x}$  can be written as

$$\begin{aligned} \frac{\delta D(\mathbf{g}, \omega)}{\delta \lambda(\mathbf{x})} &= - \left( \frac{\partial G(\mathbf{g}|\mathbf{x})_{xz}}{\partial x} + \frac{\partial G(\mathbf{g}|\mathbf{x})_{zz}}{\partial z} \right) \\ &\quad \times \left( \frac{\partial D(\mathbf{x}, \omega)_x}{\partial x} + \frac{\partial D(\mathbf{x}, \omega)_z}{\partial z} \right), \\ \frac{\delta D(\mathbf{g}, \omega)}{\delta \mu(\mathbf{x})} &= -2 \frac{\partial G(\mathbf{g}|\mathbf{x})_{xz}}{\partial x} \frac{\partial D(\mathbf{x}, \omega)_x}{\partial x} - 2 \frac{\partial G(\mathbf{g}|\mathbf{x})_{zz}}{\partial z} \frac{\partial D(\mathbf{x}, \omega)_z}{\partial z} \\ &\quad - \left( \frac{\partial G(\mathbf{g}|\mathbf{x})_{xz}}{\partial z} + \frac{\partial G(\mathbf{g}|\mathbf{x})_{zz}}{\partial x} \right) \\ &\quad \times \left( \frac{\partial D(\mathbf{x}, \omega)_z}{\partial x} + \frac{\partial D(\mathbf{x}, \omega)_x}{\partial z} \right), \end{aligned} \quad (\text{B3})$$

where  $D(\mathbf{x}, \omega)_x$  and  $D(\mathbf{x}, \omega)_z$  are horizontal- and vertical-component finite-difference solutions to the 2-D elastic wave equation for the background velocity model and the specified source location and source wavelet. From the definitions  $v_p = \sqrt{(\lambda + 2\mu)/\rho}$  and  $v_s = \sqrt{\mu/\rho}$  the Fréchet derivatives with respect to  $v_p$  and  $v_s$  (Mora 1987) can be obtained:

$$\begin{aligned} \frac{\delta D(\mathbf{g}, \omega)}{\delta v_p(\mathbf{x})} &= 2v_p(\mathbf{x})\rho(\mathbf{x}) \frac{\delta D(\mathbf{g}, \omega)}{\delta \lambda(\mathbf{x})}, \\ \frac{\delta D(\mathbf{g}, \omega)}{\delta v_s(\mathbf{x})} &= -4v_s(\mathbf{x})\rho(\mathbf{x}) \frac{\delta D(\mathbf{g}, \omega)}{\delta \lambda(\mathbf{x})} + 2v_s(\mathbf{x})\rho(\mathbf{x}) \frac{\delta D(\mathbf{g}, \omega)}{\delta \mu(\mathbf{x})}. \end{aligned} \quad (\text{B4})$$

Inserting eq. (B3) into eq. (B4) gives

$$\begin{aligned} \frac{\partial D(\mathbf{g}, \omega)}{\partial v_s(\mathbf{x})} &= 4v_s(\mathbf{x})\rho(\mathbf{x}) \left\{ \frac{\partial G(\mathbf{g}|\mathbf{x})_{zz}}{\partial z} \frac{\partial D(\mathbf{x}, \omega)_x}{\partial x} \right. \\ &\quad + \frac{\partial G(\mathbf{g}|\mathbf{x})_{xz}}{\partial x} \frac{\partial D(\mathbf{x}, \omega)_z}{\partial z} - \frac{1}{2} \left( \frac{\partial G(\mathbf{g}|\mathbf{x})_{xz}}{\partial z} \right. \\ &\quad \left. \left. + \frac{\partial G(\mathbf{g}|\mathbf{x})_{zz}}{\partial x} \right) \left( \frac{\partial D(\mathbf{x}, \omega)_z}{\partial x} + \frac{\partial D(\mathbf{x}, \omega)_x}{\partial z} \right) \right\}. \end{aligned} \quad (\text{B5})$$

The shear slowness is defined as  $s = 1/v_s$  and its derivative is  $\partial v_s = -s^{-2} \partial s$ , so that the Fréchet derivative of shear slowness can be expressed as:

$$\begin{aligned} \frac{\partial D(\mathbf{g}, \omega)}{\partial s(\mathbf{x})} &= -4s(\mathbf{x})\rho(\mathbf{x}) \left\{ \frac{\partial G(\mathbf{g}|\mathbf{x})_{zz}}{\partial z} \frac{\partial D(\mathbf{x}, \omega)_x}{\partial x} \right. \\ &\quad + \frac{\partial G(\mathbf{g}|\mathbf{x})_{xz}}{\partial x} \frac{\partial D(\mathbf{x}, \omega)_z}{\partial z} - \frac{1}{2} \left( \frac{\partial G(\mathbf{g}|\mathbf{x})_{xz}}{\partial z} \right. \\ &\quad \left. \left. + \frac{\partial G(\mathbf{g}|\mathbf{x})_{zz}}{\partial x} \right) \left( \frac{\partial D(\mathbf{x}, \omega)_z}{\partial x} + \frac{\partial D(\mathbf{x}, \omega)_x}{\partial z} \right) \right\}. \end{aligned} \quad (\text{B6})$$

## APPENDIX C: RAYLEIGH-WAVE SENSITIVITY ANALYSIS

The amplitudes of Rayleigh waves are related to the vertical displacement eigenfunction  $r_2(z, v, \omega)$ , which is a kernel function that is a function of velocity, frequency  $\omega$  of the fundamental Rayleigh wave, and depth  $z$  (Aki & Richards 1980). Strong amplitudes at a specified depth say that the Rayleigh waves are sensitive to  $S$ -wave velocity variations at that depth and frequency. The eigenfunction  $r_2$  can be computed from the Thomson–Haskell algorithm and the Rayleigh dispersion equation (Thomson 1950; Haskell 1953). It is constructed by a sequence of matrix multiplications involving terms that are transcendental functions of the properties (velocity and depth) of a layered medium. The roots are the wavenumbers corresponding to the modes of propagation of Rayleigh waves at each frequency. The eigenfunctions give the depth dependence of the vertical displacement which can be used to test the sensitivity of the Rayleigh-wave amplitudes to localized variations in the velocity model.

Similarly, the Jacobian matrix  $\mathbf{J}$  is also used to evaluate the sensitivity of Rayleigh waves at the depth  $z_j$  according to the partial derivative of phase velocity with respect to the  $S$ -wave velocity. The Rayleigh-wave phase velocity  $v_{R_j}$  is determined by the characteristic equation  $F$ :

$$F(f_j, v_{R_j}, v_s, v_p, \rho, h) = 0, \quad (j = 1, 2, \dots, m), \quad (\text{C1})$$

where  $f_j$  is the frequency,  $v_R$  is the phase velocity; and  $v_s$ ,  $v_p$ ,  $\rho$  and  $h$  are the  $S$ -wave velocity,  $P$ -wave velocity, density, and thickness, respectively. The  $\mathbf{J}_z$  element of the Jacobian matrix is the partial derivative of phase velocity with respect to the  $S$ -wave velocity over different depths (Lai & Rix 1998):

$$\mathbf{J}_z = \left[ \frac{\partial F / \partial v_s(z)}{\partial F / \partial v_R(z)} \right]_{z=z_i} \quad (i = 1, 2, 3 \dots n). \quad (\text{C2})$$