Interpretation®

Tutorial for Wave Equation Inversion of Skeletonized Data

Journal:	<i>Interpretation</i>
Manuscript ID	INT-2016-0241.R1
Manuscript Type:	2016-05 Skeletonized/sparse/multiscale geophysical inversion for the interpreter
Date Submitted by the Author:	23-Feb-2017
Complete List of Authors:	Lu, Kai; King Abdullah University of Science and Technology, Physical Science and Engineering li, jing; Jilin University, ; Earth Science Guo, Bowen; King Abdullah University of Science and Technology, Earth science and engineering Fu, Lei; King Abdullah University of Science and Technology, Earth Science and Engineering Program Schuster, Jerry; King Abdullah University for Science and Technology, Earth Science Department;
Keywords:	tomography, inversion, wave equation
Subject Areas:	Tutorials/pitfalls and lessons learned, Interpretation concepts, algorithms, methods, and tools

SCHOLARONE™
Manuscripts

TUTORIAL FOR WAVE EQUATION INVERSION OF SKELETONIZED DATA

Kai Lu¹, Jing Li¹, Bowen Guo¹, Lei Fu¹, and Gerard Schuster¹

¹King Abdullah University of Science and Technology (KAUST),
Physical Science and Engineering Division (PSE),
Thuwal, 23955-6900, Saudi Arabia

Original paper date of submission: 16-Dec-2016

Revised paper date of submission: 23-Feb-2017

ABSTRACT

Full waveform inversion of seismic data is often plagued by cycle skipping problems so that an iterative optimization method often gets stuck in a local minimum. To avoid this problem we simplify the objective function so that the iterative solution can quickly converge to a solution in the vicinity of the global minimum. The objective function is simplified by only using parsimonious and important portions of the data, which are defined as skeletonized data. We now present a mostly non-mathematical tutorial that explains the theory of wave-equation inversion of skeletonized data. We also show its effectiveness with examples.

INTRODUCTION

Conventional FWI inverts for a velocity model by using an objective function that minimizes the L^2 norm of the residuals between the predicted and the observed traces (Tarantola, 2005; Virieux and Operto, 2009). However, such a misfit function is highly non-linear and the iterations often get stuck in local minima. In order to mitigate this problem, a skeletonized representation of the data such as first-arrival traveltimes (Luo and Schuster, 1991a,b; Woodward, 1992; Zhou et al., 1995) can be inverted to obtain the low-to-intermediate wavenumber details of the background velocity model. Using the wave equation to invert traveltimes is known as the wave equation traveltime inversion (WT) method where the misfit function is quasi-linear and enjoys better convergence properties than conventional full waveform inversion (FWI).

For the traveltimes and waveform inversion strategy (Zhou et al, 1995), the early iterations use WT to estimate the low-wavenumber parts of the model, and then a waveform inversion method is used to estimate the detailed parts of the velocity model. An example is shown in Figure 1, where the waveform misfit function ζ is defined as the sum of the squared residuals

$$\zeta = \frac{1}{2} \sum_{s,g,t} [\Delta P(\mathbf{g}, \mathbf{s}, t)]^2, \quad (1)$$

where $\Delta P(\mathbf{g}, \mathbf{s}, t) = P_{\text{obs}}(\mathbf{g}, \mathbf{s}, t) - P(\mathbf{g}, \mathbf{s}, t)$ is the data residual. Here, $P_{\text{obs}}(\mathbf{g}, \mathbf{s}, t)$ and $P(\mathbf{g}, \mathbf{s}, t)$ represent the observed and the predicted traces for a point source at $\mathbf{s}$ and a geophone at $\mathbf{g}$ respectively, and the model in this example is a two-layer medium so that the scattered data contain both primaries and multiples from the reflector and free surface. The velocity of the top layer is $V = 1.0$ km/s. A gradient optimization method will seek to move to the global minimum of ζ at $V = 1.0$ km/s for the top layer by using a starting velocity model (in this case the starting model is $V = 1.5$ km/s on the RHS of the Figure 1b) to compute the calculated trace $P(\mathbf{g}, \mathbf{s}, t)$. The starting velocity model $\mathbf{m}(\mathbf{x})$ is updated by a gradient optimization method such as the steepest descent method

$$\mathbf{m}(\mathbf{x})^{(k+1)} = \mathbf{m}(\mathbf{x})^{(k)} - \alpha \sum_{s,g,t} \Delta P(\mathbf{g}, \mathbf{s}, t)^{(k)} \frac{\partial \Delta P(\mathbf{g}, \mathbf{s}, t)^{(k)}}{\partial \mathbf{m}(\mathbf{x})}, \quad (2)$$

where α is the step length, $\mathbf{m}(\mathbf{x})^k$ is the model parameter at the k^{th} iteration.

Unfortunately, the gradient method stops at the local minimum nearest to the starting model (rightmost red circle in Figure 1b) and cannot proceed further unless special adjustments are made.

To mitigate the local minima problem, we need to begin our iterations at a starting velocity model that is close to the true model at $V = 1.0$ km/s, which is often not possible. An alternative is to simplify the misfit function to eliminate most of its bumpiness and avoid getting stuck far from the global minimum. One way to do this is to find the velocity model that minimizes the sum of the squared traveltimes as depicted in Figure 1c. Here there is only one simple traveltimes for each trace, rather than thousands of amplitudes that need to be explained by the velocity model. In this case, the velocity model will be much simpler than the actual velocity model because it only has to explain traveltimes. Therefore, the complicated waveform misfit function in equation 1 is replaced by the traveltimes misfit function

$$\zeta = \frac{1}{2} \sum_{s,g} [\Delta\tau(\mathbf{g}, \mathbf{s})]^2, \quad (3)$$

which will be much simpler, i.e. less wiggly, than the waveform misfit function, where $\Delta\tau(\mathbf{g}, \mathbf{s}) = t_{\text{obs}}(\mathbf{g}, \mathbf{s}) - t_{\text{cal}}(\mathbf{g}, \mathbf{s})$ is the traveltimes residual. Figure 1d plots the much simpler traveltimes misfit function associated with the Figure 1c traveltimes, and the associated steepest descent formula is

$$\mathbf{m}(\mathbf{x})^{(k+1)} = \mathbf{m}(\mathbf{x})^{(k)} - \alpha \sum_{s,g} \Delta\tau(\mathbf{g}, \mathbf{s})^{(k)} \frac{\partial \Delta\tau(\mathbf{g}, \mathbf{s})^{(k)}}{\partial \mathbf{m}(\mathbf{x})}. \quad (4)$$

There are no local minima in this case so a gradient optimization method will quickly converge to the global minima at $V = 1.0$ km/s. In fact, this inverted velocity model, also known as a velocity tomogram, can be used as a starting velocity model for FWI to mitigate the local minima problem. Another simplification, i.e. skeletonization, of the traces, is to low-pass filter the data as shown in Figure 1e. In this case the waveform misfit function in Figure 1f is much simpler than that for the original data in Figure 1b.

We can also mute all but the early arrivals and get an even simpler misfit function. This strategy of simplifying the data (traveltime and /or early arrivals or lower frequencies) is defined as skeletonization or multiscaling of the data (Luo and Schuster, 1991a,b; Bunks et al., 1995).

Previous to Luo and Schuster (1991), the Fréchet derivative $\frac{\partial \Delta \tau}{\partial m}$ in equation 4 could not be computed unless the high-frequency approximation of ray tracing was used. This restriction was lifted by the implicit function theorem (Luo and Schuster, 1991a; Luo and Schuster, 1991b) which says

$$\frac{\partial \Delta \tau}{\partial m} = - \frac{\frac{\partial(\Phi_{\Delta \tau})}{\partial m}}{\frac{\partial(\Phi_{\Delta \tau})}{\partial \Delta \tau}}, \quad (5)$$

where Φ is the time derivative of the cross-correlation between the observed traces $P_{\text{obs}}(\mathbf{g}, \mathbf{s}, t)$ and the predicted traces $P(\mathbf{g}, \mathbf{s}, t)$ at the geophone position $\mathbf{g}$:

$$\begin{aligned} \Phi(m, \tau)|_{\tau=\Delta \tau} &= \int dt P(\mathbf{g}, \mathbf{s}, t) \dot{P}_{\text{obs}}(\mathbf{g}, \mathbf{s}, t + \Delta \tau) \\ &= P(\mathbf{g}, \mathbf{s}, t) \otimes \dot{P}_{\text{obs}}(\mathbf{g}, \mathbf{s}, t)|_{\tau=\Delta \tau} = 0, \end{aligned} \quad (6)$$

where $\otimes$ denotes cross-correlation in time and τ is the time lag; here, the $\dot{P}_{\text{obs}}$ denotes differentiation of P with respect to the time variable. The cross-correlation function, also known as the connective function, achieves its maximum and its time derivative becomes zero at the time lag value $\tau = \Delta \tau$, which temporally aligns the predicted trace with the recorded trace. If the predicted velocity is wrong, then this lag value is non-zero because the arrival time of the predicted trace does not match the observed arrival time on the

recorded trace. Here, the predicted pressure field value $P(\mathbf{g}, \mathbf{s}, t)$ honors the acoustic wave equation for a specified velocity and density model.

In equation 5, $\frac{\partial(\Phi_{\Delta\tau})}{\partial m}$ is given by

$$\frac{\partial(\Phi_{\Delta\tau})}{\partial m} = \frac{\partial P(\mathbf{g}, \mathbf{s}, t)}{\partial m} \otimes \dot{P}_{\text{obs}}(\mathbf{g}, \mathbf{s}, t)|_{\tau=\Delta\tau}, \quad (7)$$

and can be computed using the Fréchet derivative $\frac{\partial P(\mathbf{g}, t')}{\partial m}$ of the pressure field (Luo and Schuster, 1991a), which is nothing more than a weighted reverse time migration operation. The other term in equation 5

$$\frac{\partial(\Phi_{\Delta\tau})}{\partial \Delta\tau} = P(\mathbf{g}, \mathbf{s}, t) \otimes \ddot{P}_{\text{obs}}(\mathbf{g}, \mathbf{s}, t)|_{\tau=\Delta\tau}, \quad (8)$$

is the second-time derivative of the cross-correlation function. Plugging equations 5-8 into equation 4 gives the final form of the update formula:

$$m(\mathbf{x})^{(k+1)} = m(\mathbf{x})^{(k)} - \alpha \sum_{s,g} \frac{\Delta\tau(\mathbf{g}, \mathbf{s})^{(k)}}{E} \frac{\partial P(\mathbf{g}, \mathbf{s}, t)^{(k)}}{\partial m} \otimes \dot{P}_{\text{obs}}(\mathbf{g}, \mathbf{s}, t)^{(k)}|_{\tau=\Delta\tau}, \quad (9)$$

where the normalization term is $E = \dot{P}(\mathbf{g}, \mathbf{s}, t) \otimes \dot{P}_{\text{obs}}(\mathbf{g}, \mathbf{s}, t)|_{\tau=\Delta\tau}$ (Luo and Schuster, 1991b), and $\frac{\partial P(\mathbf{g}, \mathbf{s}, t)}{\partial m}$ is the Fréchet derivative associated with the model parameter m . The gradient on the right side of equation 9 is interpreted as the weighted reverse time migration of the recorded data, where each trace is weighted by its associated traveltime residual.

In summary, until 1991 the skeletal Fréchet derivative $\frac{\partial \Delta\tau}{\partial m}$ in equation 4 could not be computed using the wave equation because the skeletal data $\Delta\tau$ did not explicitly appear in the wave equation. Defining a connection between the recorded data

$P_{\text{obs}}(\mathbf{g}, \mathbf{s}, t)$ and the skeletal residual $\Delta\tau$ by the correlation in equation 6 leads to the skeletal Fréchet derivative in equation 5. Thus, $\frac{\partial\Delta\tau}{\partial\mathbf{m}}$ can now be computed by calculating the Fréchet derivative $\frac{\partial P}{\partial\mathbf{m}}$ with finite-difference solutions to the wave equation. Each update in equation 9 is computed by reverse time migration (RTM) of the weighted data, where the weight is proportional to the traveltimes residual $\Delta\tau$.

We now present examples where skeletonized parts of the data are inverted with the wave equation. They all follow the unified skeletonized inversion procedure. The first step is to identify the skeletonized portion of the data, which should be much simpler than the recorded traces, i.e. fewer oscillations in offset and time, and should be strongly sensitive to variations of the model parameter of interest. The second step is to design a connective function, which connects the skeletonized data with the waveform. Usually the connective function is a cross-correlation between the predicted and the observed data in a specified domain. The skeletal-data residual is equal to the lag value where the cross-correlation achieves its extreme value. The final step is to calculate the skeletal Fréchet derivative associated with the parameter to be inverted, and this is used in equation 9 to update the model.

EXAMPLES OF SKELETONIZED INVERSION

Four examples of skeletonized inversion are now presented, where we identify skeletonized data that are most sensitive to the parameter of interest. The examples are

for migration velocity analysis, wave-equation dispersion inversion, wave-equation traveltimes inversion, and wave-equation Q inversion.

A) Migration velocity analysis where $\Delta\tau \rightarrow \Delta z(x)$ and $m \rightarrow v(x')$. Here, $v(x')$

is the migration velocity and Δz is the depth shift in the migration image that is needed to align the events in the common image gather (CIG) with one another.

Wave-equation migration velocity analysis (WEMVA) (Biondi et al., 1999; Sava and Biondi, 2004; Shen and Symes, 2008) inverts for the migration velocity that maximizes the similarities of CIGs. If the migration velocity is incorrect, then the specified events will not be at the same depth in the CIG.

In the following two numerical examples, the CIGs are computed in the plane-wave domain using plane-wave migration (Whitemore, 1995; Duquet et al., 2001; Zhang et al., 2005; Liu et al., 2006). The WEMVA method assumes that the correct migration velocity will lead to flat events in the plane-wave CIGs because the migrated events from the same reflection point are at the same depth for different angles of plane-wave incidence. On the other hand, for an inaccurate migration velocity, the events in the plane-wave CIGs will have depth shifts or residual moveouts Δz . The objective function of the velocity inversion method is defined as

$$\zeta = \frac{1}{2} \sum_j [\Delta z_j(x)]^2, \quad (10)$$

where j denotes the plane-wave index, and the update formula is

$$s(x')^{(k+1)} = s(x')^{(k)} - \alpha \sum_j \Delta z_j(x)^{(k)} \frac{\partial \Delta z_j(x)^{(k)}}{\partial s(x')}, \quad (11)$$

where $s(\mathbf{x}')$ represents the slowness model. The Fréchet derivative $\frac{\partial \Delta z(\mathbf{x})}{\partial s(\mathbf{x}')}$ is derived by defining a connective function (Luo and Schuster, 1991b). In this case, the connective function is the z-directional derivative of the local cross-correlation between a plane-wave migration image $m_j(\mathbf{x})$ and a pilot image $m_0(\mathbf{x})$:

$$\Phi(s(\mathbf{x}'), \tilde{z})|_{\tilde{z}=\Delta z} = \int dz m_j(x, z) \dot{m}_0(x, z + \Delta z) = 0, \quad (12)$$

where j denotes the plane-wave index, x and z are the horizontal and vertical coordinates of the image position $\mathbf{x}$, and $\tilde{z}$ is the space lag in the correlation. When $\tilde{z} = \Delta z$, which means the shift value aligns the migration traces to be at the same depth, the cross-correlation achieves the maximum value, and its z-directional derivative becomes zero.

Figure 2 shows a synthetic result for WEMVA using plane-wave CIGs. Using the initial velocity model shown in Figure 2b, the plane-wave CIGs have strong residual moveouts as depicted in Figure 2d. After 10 iterations, the inverted velocity model in Figure 2c recovers most of the low-wavenumber components of the true velocity in Figure 2a and produces a more flattened CIGs as depicted in Figure 2e.

Figure 3 describes the result of applying the WEMVA method to marine data recorded in the Gulf of Mexico. The initial velocity model shown in Figure 3a is obtained by inverting the traveltimes of the first arrivals. After 10 iterations, the inverted velocity model is shown in Figure 3b. The plane-wave CIGs are more

flattened using the inverted velocity in Figure 3d than those using the initial velocity (Guo and Schuster, 2017).

B) *Wave-equation dispersion inversion where $\Delta\tau \rightarrow \Delta\kappa$ and $m \rightarrow v_s(x)$.*

Here, $v_s(x)$ is the shear velocity and $\Delta\kappa$ is the residual between the predicted and observed dispersion curves of a shot gather. The dispersion curves are those for the fundamental mode of the Rayleigh wave and are computed by windowing the predicted (or observed) Rayleigh wave arrivals in the shot gather, and applying Fourier transforms along the x and t axes to get the spectrum (Figure 4). The dispersion curve for the fundamental mode is picked from the spectrum by identifying the largest amplitudes that are closest to the κ axis. Analogous to wave-equation traveltime inversion, the complicated surface-wave arrivals in traces are skeletonized as simpler data, namely the picked dispersion curves in the $\kappa - \omega$ domain. The elastic wave equation is then used to invert these curves for 2D or 3D velocity models where the objective function is the sum of the squared residuals $\Delta\kappa$ as

$$\zeta = \frac{1}{2} \sum_{\omega} [\Delta\kappa(\omega)]^2, \quad (13)$$

and the update formula is

$$s(x)^{k+1} = s(x)^k - \alpha \sum_{\omega} \Delta\kappa(\omega)^k \frac{\partial \Delta\kappa(\omega)^k}{\partial s(x)}, \quad (14)$$

where $s(x)$ represents the shear-slowness. In this case, the connective function is the derivative of the cross-correlation between the predicted spectrum $D(\kappa, \omega)$ and the conjugated observed spectrum $D(\kappa, \omega)_{\text{obs}}^*$ in the $\kappa - \omega$ domain:

$$\Phi(s(\mathbf{x}), \tilde{\kappa})|_{\tilde{\kappa}=\Delta\kappa} = \text{Real}[\int d\kappa D(\kappa, \omega) \dot{D}(\kappa + \Delta\kappa, \omega)^*_{\text{obs}}] = 0, \quad (15)$$

where $\tilde{\kappa}$ is the wavenumber lag in the correlation, and $\dot{D}(\kappa, \omega)^*_{\text{obs}} = \frac{\partial D(\kappa, \omega)^*_{\text{obs}}}{\partial \kappa}$; for notation simplicity, the shot variable is silent. Setting $\tilde{\kappa} = \Delta\kappa$, says that the wavenumber shift aligns the predicted and observed spectra with one another for a specified ω ; thus, the cross-correlation achieves the maximum value and its wavenumber derivative becomes zero. This method is denoted as skeletonized wave-equation dispersion inversion (WD) (Li and Schuster, 2016; Li et al., 2017a), and is less prone to the cycle skipping problems of full waveform inversion (FWI).

A 2D blind¹ model test is conducted to test the proposed idea. The synthetic data are generated by using a Ricker source wavelet with a 10-Hz peak frequency. There are 120 shot gathers, each with 120 receivers at a 5 m interval. Figure 5a is the common offset gather² (COG) with a source-receiver offset of 50 m, indicating that the blind model includes two layers and two faults at offsets of 200 m and 500 m. Figure 5c is the P-wave tomogram calculated by inverting the first-arrival traveltimes. The S-wave tomogram calculated by the WD method is shown in Figure 5b, which is consistent with the COG (Figure 5a) and the P-wave tomogram (Figure 5c). The actual model is a three-layer model with the actual interface indicated by the dashed line in Figures 5b and 5c. The actual shear velocities in the top, middle and lower layers are 300 m/s, 500 m/s and 700 m/s, respectively.

¹ The author was given the data to invert but did not know the velocity model at the time.

² A common offset gather refers such a gather that all traces within it have the same source-receiver offset.

The WD method is also applied to a near-surface field data recorded over the Qademah fault, which is shown in Figure 6. We denote this as a possible fault because it is characterized by a low-velocity zone that linearly trends for many kilometers. A common offset gather (COG) is shown in Figure 7a with the source-receiver offset of 50 m. The dashed lines in Figure 7a indicates the location of the Qademah fault, which is consistent with the lateral velocity decrease in the P-wave velocity tomogram of Figure 7b. The P-wave velocity tomogram is computed by inverting the P-wave first-arrival traveltimes with a ray-based tomography method. Then, all shot gathers are transformed into the $\kappa - \omega$ domain by a linear Radon transform (LRT) method and the maximum energy values of the dispersion curves are picked. Figure 7c shows the S-wave velocity tomogram obtained from 1D WD inversion (Li and Hanafy, 2016), where the horizontal layered medium is assumed, which can be used to roughly estimate the position of the Qademah fault. The WD method does not require the horizontal layered medium assumption, and is used to invert the picked dispersion curves to give the 2D S-wave velocity tomogram shown in Figure 7d. It is obvious that there is a low-velocity zone on the downthrown side of the fault. This is consistent with the P-wave velocity tomogram in Figure 7b and the COG profile in Figure 7a for $150 \text{ m} < x < 300 \text{ m}$ (Li and Schuster, 2016).

C) **Wave equation traveltime inversion where $\Delta\tau \rightarrow \Delta\tau$ and $m \rightarrow v(x, z)$.**

Here, $v(x, z)$ is the P-wave velocity and $\Delta\tau$ is the residual between the predicted and observed first-arrival traveltimes of a shot gather. The first-arrival traveltimes are

computed by windowing the computed (or observed) early arrivals in the shot gather as shown in Figure 8, picking the times (or correlating the predicted and observed traces to get the residual), and using equation 9 to update the velocity model.

A synthetic example is illustrated to demonstrate the method. The true P-velocity model is shown in Figure 9a, which is a complex model with an irregular interface indicated by the white dashed line. The initial velocity model is the gradient velocity model. The traveltimes of the first-arrivals are picked and inverted, and the resulting WT tomogram is shown in Figure 9b. The horizontal velocity profiles at the depth of 25 m are shown in Figure 10, where the black, blue, and red curves represent the velocity of the true model, the initial model, and the WT tomogram respectively. The tomogram accurately recovers low-to-intermediate wavenumber information, while the high wavenumber details are missing.

The WT method is applied to a field dataset from a Gulf of Aqaba experiment. A 2D resistivity profile was also acquired at the same location parallel to the seismic profile. Figure 11 shows COGs with different offsets, where a discontinuity appears at $x = 145$ m (the area indicated by the black dashed ellipse). The time delays of the surface waves in the distorted area (ellipse in Figure 11b) implies that the shear velocity is slower, and suggests the location of a possible fault. Figure 12a shows the WT tomogram after 10 iterations with a gradient starting model. The pink ellipse area represents a laterally variable velocity, which might be interpreted as a colluvial wedge associated with the fault. The recorded resistivity data were inverted using

Res2DInv software to generate the resistivity tomogram in Figure 12b. The location of the resistivity anomaly between offsets 130m and 145 m indicates a potential fault, which is consistent with the WT result.

D) Wave equation Q inversion where $\Delta\tau \rightarrow \Delta f$ and $m \rightarrow Q(x)$. Here, $Q(x)$ is the Rayleigh-wave attenuation parameter, and Δf is the residual between the predicted and observed frequencies of the spectral peaks (Dutta and Schuster, 2016a; Dutta, 2016b) of a fundamental Rayleigh wave (see Figure 13c). The spectral peaks are computed by windowing the computed (or observed) surface waves in the shot gather (see Figure 13a), applying Fourier transforms along the x - and t -axes, and picking the peak frequencies of the spectra. Now, the complicated surface-wave arrivals are skeletonized as simpler data, namely the amplitude spectrum of the windowed fundamental-mode Rayleigh-wave arrivals. The centroid or peak frequency shifts between the observed and the predicted Rayleigh-wave arrivals are the skeletal residuals, and the attenuation model is sought that minimizes the sum of the squared frequency residuals

$$\zeta = \frac{1}{2} \sum_{s,\omega} [\Delta f(\mathbf{g}, s)]^2. \quad (16)$$

The optimal attenuation model is found by the steepest descent method:

$$\eta(\mathbf{x})^{(k+1)} = \eta(\mathbf{x})^{(k)} - \alpha \sum_{\mathbf{g},s} \Delta f(\mathbf{g}, s)^{(k)} \frac{\partial \Delta f(\mathbf{g}, s)^{(k)}}{\partial \eta(\mathbf{x})}, \quad (17)$$

where $\eta(\mathbf{x})$ is the relaxation rate, and approximately equals to $1/Q(\mathbf{x})$ for realistic geologic models (Li et al, 2017b). The connective function is defined as the

derivative of the cross-correlation between the predicted and the observed spectral peak frequencies of the fundamental Rayleigh-wave arrivals

$$\Phi(Q(\mathbf{x}), \tilde{f})|_{\tilde{f}=\Delta f} = \int df D(\mathbf{g}, \mathbf{s}, f) \dot{D}(\mathbf{g}, \mathbf{s}, f + \Delta f)^*_{\text{obs}} = 0, \quad (18)$$

where $\tilde{f}$ represents the frequency lag in the correlation, $D(\mathbf{g}, \mathbf{s}, f)$ and $D(\mathbf{g}, \mathbf{s}, f + \Delta f)^*_{\text{obs}}$ are the predicted and the observed spectral peak frequencies of the fundamental Rayleigh wave at the specified receiver $\mathbf{g}$ and source $\mathbf{s}$ locations, and $\dot{D}(\mathbf{g}, \mathbf{s}, f + \Delta f)^*_{\text{obs}} = \frac{\partial D(\mathbf{g}, \mathbf{s}, f + \Delta f)^*_{\text{obs}}}{\partial f}$. When $\tilde{f} = \Delta f$, which means the frequency shift Δf aligns the predicted and observed spectra with one another, the cross-correlation achieves the maximum value, and its frequency derivative becomes zero. Solutions to the viscoelastic wave-equation are used to compute the predicted Rayleigh-wave arrivals and the gradient at every iteration (Li et al, 2017b).

A numerical example with a complex near-surface model is illustrated in Figure 14. The true V_s and Q models are shown in Figures 14a and 14b, respectively. The smoothed true S-wave velocity model shown in Figure 14c is used as the background velocity. The observed data are generated by 40 evenly distributed shots and 70 receivers at intervals of 2 m on the surface. The Q tomogram after 21 iterations is shown in Figure 14d, which roughly agrees with the true Q model in Figure 14b.

The proposed method is now tested on the same field data used in the WD section. The starting Q model is homogeneous with $Q = 1000$ and the inverted Q tomogram is shown in Figure 15a. A geological relation between the S-wave velocity and the Q

tomograms says that the high attenuation regions in the Q tomogram (low Q values) correspond to the low S-wave velocity regions (De Meersman 2013; Zhu and Harris 2015). Previous work by Zhang et al. (2015) demonstrated that areas with high V_p/V_s ratios tend to have low Q values (or high attenuation) while the variations of Q_p and Q are the same, which mostly agrees with what can be observed in Figure 15a.

To check the feasibility of the inverted Q tomogram, it is compared to a common offset gather (COG) profile (Hanafy et al., 2015). Figure 15c shows a COG profile using the processed data for an offset of 50 m. The black dashed line in this figure shows the location of the fault which is between 250-300 m. The locations of the Q anomalies in the Q tomogram and the low-velocity area in the S-wave tomogram are consistent with the location of the fault, as delineated in the COG profile.

SUMMARY

We presented an overview of wave equation inversion of skeletonized data. The skeletal data form a simplified objective function that is much reduced in the number of local minima and complexity. This results in an iterative gradient optimization method that typically enjoys robust convergence to the vicinity of the actual model without getting stuck in a local minimum. The liability is that the resulting tomogram is only of moderate resolution because of the reduced complexity of the skeletal data. This model with moderate resolution can be used as a good starting model for full waveform inversion to give a more detailed estimate of the actual model.

Examples of four types of skeletonization were presented, using both synthetic data and field data. For the synthetic cases, the inverted models closely resembled smoothed representations of the actual models. The field data results could not be compared to the actual models, but they provided models that were reasonably consistent with the data, e.g. COGs, or the known geology of the area. This suggests that the interpreter can identify portions $\tilde{d}$ of the data that are most sensitive to the model parameter m of interest, and then use skeletonized inversion to invert $\tilde{d}$ for m . Some examples include AVO parameters and geometry attributes in the migration image.

ACKNOWLEDGMENTS

The research reported in this publication was supported by funding from King Abdullah University of Science and Technology (KAUST). We thank the financial support from the sponsors of the Consortium of Subsurface Imaging and Fluid Modeling (CSIM). 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.

REFERENCE

Biondi, B., P. Sava, et al., 1999, Wave-equation migration velocity analysis: SEG Technical Program Expanded Abstracts 1999, 1723-1726.

Bunks, C., F. M. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: *Geophysics*, 60, 1457–1473.

De Meersman, K., 2013, S-waves and the near surface: A time-lapse study of S-wave velocity and attenuation in the weathering layer of an Alberta heavy oil field: *The Leading Edge*, 32, 40–47.

Duquet, B., P. Lailly, and A. Ehinger, 2001, 3D plane wave migration of streamer data: *SEG Technical Program Expanded Abstracts 2001*, 1033-1036.

Dutta, G., and G. T. Schuster, 2016a, Wave-equation Q tomography: *Geophysics*, 81, 471–484.

Dutta, G., 2016b, Skeletonized wave-equation inversion for Q: *SEG Technical Program Expanded Abstracts 2016*, 3618-3623.

Guo, B., and G. T. Schuster, 2017, Wave-equation migration velocity analysis using plane-wave common image gathers: *Geophysics*, accepted.

Hanafy, M. S., A. Altheyab, G. T. Schuster, 2015, Controlled noise seismology: *SEG Technical Program Expanded Abstracts 2015*, 5102–5107

Li, J. and M. S. Hanafy, 2016. Skeletonized inversion of surface wave: active source versus controlled noise comparison: *Interpretation*, 3, 11–19.

Li, J. and G. T. Schuster, 2016, Skeletonized wave equation of surface wave dispersion inversion: *SEG Technical Program Expanded Abstracts 2016*, 3630–3635.

Li, J., Z. Feng and G. T. Schuster, 2017a, Wave-equation dispersion inversion: Geophys. J. Inter., 208, 1567-1578.

Li, J., G. Dutta and G. T. Schuster, 2017b, Wave-equation Qs inversion of skeletonized surface waves: Geophys. J. Inter., doi: 10.1093/gji/ggx051.

Liu, F., D. W. Hanson, N. D. Whitmore, R. S. Day, and R. H. Stolt, 2006, Toward a unified analysis for source plane-wave migration: Geophysics, 71, S129-S139.

Luo, Y. and G. T. Schuster, 1991a, Wave equation inversion of skeletonized geophysical data: Geophys. J. Inter., 105, 289–294.

Luo, Y. and G. T. Schuster, 1991b, Wave equation traveltime inversion: Geophysics, 56, 645–653.

Sava, P. and B. Biondi, 2004, Wave-equation migration velocity analysis. I. Theory: Geophysical Prospecting, 52, 593-606.

Schuster, G. T., 2017. Seismic inversion, SEG Publishing: Tulsa, Oklahoma.

Shen, P. and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, 73, VE49-VE59.

Tarantola, A., 2005, Inverse problem theory and methods for model parameter estimation: Society for Industrial and Applied Mathematics, Philadelphia, PA.

Virieux, J. and S. Operto, 2009, An overview of full-waveform inversion in exploration: *Geophysics*, 74, WCC1–WCC26.

Whitmore, N. D., 1995, An imaging hierarchy for common-angle plane wave seismograms: PhD thesis, University of Tulsa.

Woodward, M. J., 1992, Wave-equation tomography: *Geophysics*, 57, 15-26.

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

Zhang, L., D. Zhu, and X. Zhang, 2015, Seismic attributes method for prediction of unconsolidated sand reservoirs of heavy oil: *The Open Fuels & Energy Science Journal*, 8, 1–10.

Zhang, Z. D., G. T. Schuster, Y. Liu, S. M. Hanafy, and J. Li, 2016. Wave equation dispersion inversion using a difference approximation to the dispersion-curve misfit gradient: *Journal of Applied Geophysics*, 133, 9–15.

Zhou, C., W. Cai, Y. Luo, G. T. Schuster, and S. Hassanzadeh, 1995, Acoustic wave equation traveltimes and waveform inversion of crosshole seismic data: *Geophysics*, 60, 765–773.

Zhu, T., and J. M. Harris, 2015, Improved estimation of P-wave velocity, S-wave velocity, and attenuation factor by iterative structural joint inversion of crosswell seismic data: *Journal of Applied Geophysics*, 123, 71–80.

LIST OF FIGURES

Figure 1. (a), (c) and (e) are synthetic 10 Hz data, traveltimes data and 2 Hz data, respectively, and (b), (d) and (f) are the associated misfit functions plotted against the hypothetical velocity value of the first layer in the two-layer model. The correct value of V is 1.0 km/s, where the misfit function is zero. Notice the fewer local minima (red circles) for the 2-Hz data than for the 10-Hz data.

Figure 2: (a) True velocity model, (b) initial velocity of the migration velocity analysis (MVA), (c) inverted velocity by the plane-wave MVA method, (d) plane-wave CIGs migrated using the initial velocity model and (e) plane-wave CIGs migrated with the inverted velocity model (Guo and Schuster, 2017).

Figure 3: (a) Initial velocity model obtained by inverting the traveltimes of the first arrivals, (b) inverted velocity model using the plane-wave MVA method, (c) plane-wave CIGs migrated using the initial velocity model and (d) plane-wave CIGs migrated using the inverted velocity model (Guo and Schuster, 2017).

Figure 4. (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 $\kappa(\omega)$ denotes the skeletonized data (Zhang et al., 2015; Li et al., 2017a; Schuster, 2017).

Figure 5: Blind model test: (a) COG profile; (b) S-wave tomogram; (c) P-wave tomogram.

Figure 6. The location of the field experiment. The red dashed line in this figure is the Qademah fault which is aligned along the north-south direction. The blue solid line is the survey line which consists of 60 geophones at a 10 m spacing, where shots are fired at each geophone locations to generate 60 shot gathers.

Figure 7. Qademah COG and associated tomograms (Li and Schuster 2016): (a) common offset gather (COG, offset=50 m), (b) P-wave velocity tomogram inverted from the first-arrival traveltimes, (c) 1D S-wave velocity tomogram inverted from the dispersion-curve data, and d) S-wave velocity tomogram inverted from the 2D WD inversion of the dispersion-curves data.

Figure 8. (a) An original shot gather, where the direct, refraction, air, Rayleigh, and reflection waves are indicated, and (b) the skeletonized data, where only the first arrivals are extracted.

Figure 9. (a) The true velocity model for synthetic example, where the white dashed line represents the interface interpreted in (b) the WT tomogram.

Figure 10. The comparison of the WT tomogram and true velocity model profile at the depth of 25 m, where the black, blue, and red curves represent the true, initial, and WT velocities.

Figure 11. Common offset gathers for different offsets: (a) offset = 0 m, (b) offset = 20m, and (c) offset = 40 m.

Figure 12. a) The inverted WT tomograms computed from the Aqaba data, and b) resistivity tomogram inverted by Res2DInv. A low-resistivity anomaly is shown around $x = 145$ m.

Figure 13. (a) A common shot gather (CSG) comparing the fundamental-mode Rayleigh-wave arrivals with and without Q. The blue dashed line indicates the window used to extract these arrivals for WQ, (b) comparison between a predicted and an observed surface-wave arrival, and (c) comparison between their amplitude spectra.

Figure 14. (a) True velocity, (b) Q models used for generating the observed data, (c) velocity model used for WQ, and (d) inverted Q tomogram from WQ.

Figure 15. (a) Q tomogram computed by WQ , (b) the ratio (V_p / V_s) of the seismic field data, and (c) common offset gather (offset=50 m) profile after data processing.

Figure 1. (a), (c) and (e) are synthetic 10 Hz data, traveltime data and 2 Hz data, respectively, and (b), (d) and (f) are the associated misfit functions plotted against the hypothetical velocity value of the first layer in the two-layer model. The correct value of V is 1.0 km/s, where the misfit function is zero. Notice the fewer local minima (red circles) for the 2-Hz data than for the 10-Hz data.

177x181mm (300 x 300 DPI)

Figure 2: (a) True velocity model, (b) initial velocity of the migration velocity analysis (MVA), (c) inverted velocity by the plane-wave MVA method, (d) plane-wave CIGs migrated using the initial velocity model and (e) plane-wave CIGs migrated with the inverted velocity model (Guo and Schuster, 2017).

202x153mm (300 x 300 DPI)

Figure 3: (a) Initial velocity model obtained by inverting the traveltimes of the first arrivals, (b) inverted velocity model using the plane-wave MVA method, (c) plane-wave CIGs migrated using the initial velocity model and (d) plane-wave CIGs migrated using the inverted velocity model (Guo and Schuster, 2017).

196x180mm (300 x 300 DPI)

Figure 4. (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 $\kappa(\omega)$ denotes the skeletonized data (Zhang et al., 2015; Li et al., 2017a; Schuster, 2017).

128x62mm (300 x 300 DPI)

Figure 5: Blind model test: (a) COG profile; (b) S-wave tomogram; (c) P-wave tomogram.

178x219mm (300 x 300 DPI)

Figure 6. The location of the field experiment. The red dashed line in this figure is the Qadimah fault which is aligned along the north-south direction. The blue solid line is the survey line which consists of 60 geophones at a 10 m spacing, where shots are fired at each geophone locations to generate 60 shot gathers.

182x146mm (300 x 300 DPI)

Figure 7. Qadema COG and associated tomograms (Li and Schuster 2016): (a) common offset gather (COG, offset=50 m), (b) P-wave velocity tomogram inverted from the first-arrival traveltimes, (c) 1D S-wave velocity tomogram inverted from the dispersion-curve data, and d) S-wave velocity tomogram inverted from the 2D WD inversion of the dispersion-curves data.

200x116mm (300 x 300 DPI)

Figure 8. (a) An original shot gather, where the direct, refraction, air, Rayleigh, and reflection waves are indicated, and (b) the skeletonized data, where only the first arrivals are extracted.

190x136mm (300 x 300 DPI)

Figure 9. (a) The true velocity model for synthetic example, where the white dashed line represents the interface interpreted in (b) the WT tomogram.

138x76mm (300 x 300 DPI)

Figure 10. The comparison of the WT tomogram and true velocity model profile at the depth of 25 m, where the black, blue, and red curves represent the true, initial, and WT velocities.

191x139mm (300 x 300 DPI)

Figure 11. Common offset gathers for different offsets: (a) offset = 0 m, (b) offset = 20m, and (c) offset = 40 m.

196x171mm (300 x 300 DPI)

Figure 12. a) The inverted WT tomograms computed from the Aqaba data, and b) resistivity tomogram inverted by Res2DInv. A low-resistivity anomaly is shown around $x = 145$ m.

165x124mm (300 x 300 DPI)

Figure 13. (a) A common shot gather (CSG) comparing the fundamental-mode Rayleigh-wave arrivals with and without Q . The blue dashed line indicates the window used to extract these arrivals for WQ, (b) comparison between a predicted and an observed surface-wave arrival, and (c) comparison between their amplitude spectra.

149x101mm (300 x 300 DPI)

Figure 14. (a) True velocity, (b) Q models used for generating the observed data, (c) velocity model used for WQ, and (d) inverted Q tomogram from WQ.

140x97mm (300 x 300 DPI)

Figure 15. (a) Q tomogram computed by WQ , (b) the ratio (V_p/V_s) of the seismic field data, and (c) common offset gather (offset=50 m) profile after data processing.

155x135mm (300 x 300 DPI)