Wave Equation Radon Tomography for Early Arrivals

Amr Ibrahim†, Gerard T. Schuster* Sherif M Hanafy*, Physical Science and Engineering Division, King Abdullah University of Science and Technology, Thuwal 23955, Saudia Arabia* and Department of Physics, Faculty of Science, Beni-Suef University, Beni-Suef, Egypt†*

SUMMARY

We present the theory of wave-equation Radon tomography (WRT) where the slopes and zero-intercept time of early arrivals in the $\tau - p$ domain are inverted for the subsurface velocity structure. The early arrivals are windowed in a shot gather, but they are still too wiggly to avoid local minima with a full waveform inversion (FWI) method. To reduce their complexity, a local linear Radon $\tau - p$ transform is applied to the events to focus them into few points. These points, which identify the slopes and zero-intercept time of the early arrivals, are picked to give the slowness coordinate p_i^{obs} at the zero-intercept time τ_i. The misfit function $\varepsilon = \sum_{i=1}^{P}(p_i - p_i^{obs})^2 + \sum_{i=1}^{P}(\tau_i - \tau_i^{obs})^2$ is computed and a gradient optimization method is used to find the optimal velocity model that minimizes ε. Results with synthetic data and field data show that WRT can accurately reconstruct the near-surface P-wave velocity model and converges faster than other wave-equation methods.

INTRODUCTION

There are two main methods for inverting velocity models from seismic data: ray-based tomography and wave-equation inversion. The ray-based tomography methods are computationally inexpensive but are restricted in resolution because of the high-frequency approximation. In contrast, the wave-equation inversion methods avoid the high-frequency assumption by using numerical solutions to the wave equation to explain the recorded data. In theory, the model resolution can be as high as a large fraction of a wavelength. The problem with such methods, such as the full waveform inversion (FWI) method, is that it suffers from cycle skipping problems where the inverted model gets stuck in local minima far from the actual model. Also, FWI must explain the amplitudes $d(\mathbf{g},t)$ of the observed arrivals at $\mathbf{g}$, which is almost impossible unless all of the physics of attenuation, anisotropy, and elasticity are included in the forward modeling simulations.

To reduce the complexity of the data and avoid fitting amplitudes, wave equation traveltime inversion methods (Luo and Schuster, 1991b,a; Woodward, 1989, 1992) were developed inverts for traveltimes of specified arrivals using numerical solutions to the wave equation. No high-frequency approximation is required, and for a few selected events, this method provided moderate resolution of the velocity model. The wave-equation tomography (WT) method is denoted as skeletonized wave equation inversion because the complicated input data of wiggly traces are skeletonized as simple traveltimes of selected events. This simplification resulted in an objective function with low-order complexity and robust convergence.

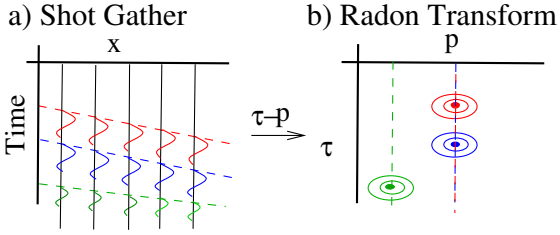

Figure 1: Shot gather $d(x_g,t|x_s,0)_s$ on the left and its Radon transform $D(\tau,p)_s$ on the right.

We now propose a skeletonized inversion procedure that provides moderate resolution in the velocity model $c(\mathbf{x})$, accounts for multiples, and the computational expense is comparable to that of standard WT. The skeletonized data are the coordinates (τ_i, p_i) of the maximal values of the Radon transform $D(\tau_i, p_i)_s$ of the shot gathers, where the shot gather is indicated by the s index. As an example, Figure 1 shows that the shot gather in a) has a Radon transform in b) that consists of just three large amplitudes. See the filled red, green, and blue circles in Figure 1b. Picking the p_i values of the three filled circles gives the simplified objective function $\varepsilon = \sum_{i=1}^{3}(p_i - p_i^{obs})^2$, and a gradient optimization method can be used to invert for the velocity model $c(\mathbf{x})$. The Fréchet derivative $\partial p_i/\partial c(\mathbf{x})$ in the gradient $\partial\varepsilon/\partial c(\mathbf{x})$ is derived using the implicit function theorem (Luo and Schuster, 1991b,a). We will denote this procedure as wave equation Radon tomography (WRT) because it uses the wave equation inverts the data in the Radon transform domain. WRT is developed for the linear Radon transform, but it can easily be adapted to the parabolic, and hyperbolic Radon transforms (Bickel, 2000; Ibrahim and Sacchi, 2014a,b, 2015; Ibrahim, 2015). Unlike the previous methods of $\tau - p$ inversion which assumes a layered (Stoffa et al., 1981; Diebold and Stoffa, 1981; Stoffa, 1989) or dipping-layer model (Milkereit et al., 1985), and WRT is valid for any velocity model.

This paper is arranged into four sections. After the introduction, the theory section derives the WRT method for both transmitted arrivals and reflection arrivals. The third section presents numerical and field data examples, followed by the final section that provides the conclusions.

THEORY

Assume a point source at the end of a linear survey line that overlies an acoustic medium, where transmitted waves propagate to the geophones located at $\mathbf{g} = (x_g, 0) \in B_o$ on the horizontal free surface. The actual trace $d(x_g,t|x_s,0)$ is excited at $t = 0$ by a vertical-displacement point source at $\mathbf{s}$ and recorded at $\mathbf{g}$ by a vertical particle-velocity geophone. For convenience,

10.1190/segam2018-2998360.1

the scalar coordinates in $d(x_g,t|x_s,0) \rightarrow d(\mathbf{g},t|\mathbf{s},0)$ and $g(x,t|x_s,0) \rightarrow d(\mathbf{x},t|\mathbf{s},0)$ will sometimes be replaced with vector notation with the understanding that the data are recorded on the horizontal plane defined by $z=0$. The trial image point is denoted as $\mathbf{x}=(x,z)$. Applying a Radon transform to the data $d(x_g,t|x_s,0)$ in Figure 1a gives the skeletal $\tau-p$ data $D(\tau,p)_s$ in Figure 1b:

$$D(\tau,p)_s = \int_{-\infty}^{\infty} d(x_g, t=\tau+x_g p|x_s,0)dx_g, \quad (1)$$

where $D(\tau,p)_s$ is an implicit function of the assumed velocity model $c(\mathbf{x})$. If the Radon data are sparse in the $\tau-p$ domain, then the strategy is to pick the skeletal coordinates (τ_i,p_i) of the maximal values of $|D(\tau,p)_s|$ and invert them for the velocity model. In this case we use a gradient optimization method to find the velocity $c(\mathbf{x})$ that minimizes the simple misfit :

$$\varepsilon = \frac{1}{2}\sum_{i=1}^{P}(\overbrace{p_i-p_i^{obs}}^{residual=\Delta p_i})^2, \quad (2)$$

where p_i^{obs} is the observed p coordinate of a maximum in $|D(\tau,p)_s|$, p_i is the predicted one computed by a numerical solution to the wave equation, and there are P extremums of $D(\tau,p)$ (see Figure 1b). Here, we assume that $\tau_i=\tau_i^{obs}$ but this assumption can be relaxed by using the general misfit function

$$\varepsilon = \frac{(1-\beta)}{2}\sum_{i=1}^{P}(p_i-p_i^{obs})^2+\frac{\beta}{2}\sum_{i=1}^{P}(\tau_i-\tau_i^{obs})^2, \quad (3)$$

where $1 \geq \beta \geq 0$ is a scalar that weights the different misfit functions. For simplicity of exposition, we will choose $\beta=0$ and apply the steepest descent method to find the velocity model that minimizes the simpler objective function in equation 2:

$$\begin{aligned} c(\mathbf{x})^{(k+1)} &= c(\mathbf{x})^{(k)} - \alpha\frac{\partial\varepsilon}{\partial c(\mathbf{x})}, \\ &= c(\mathbf{x})^{(k)} - \alpha\overbrace{\sum_{i=1}^{P}\Delta p_i\frac{\partial p_i}{\partial c(\mathbf{x})}}^{\gamma(\mathbf{x})=gradient}, \end{aligned} \quad (4)$$

where α is the step length. To derive the formula for the skeletal Fréchet derivative $\frac{\partial p_i}{\partial c(\mathbf{x})}$ we use the connective function $\Phi(\tau_i,\Delta p_i,c(\mathbf{x}))$, which correlates the observed data $D(\tau,p)_s^{obs}$ with the predicted data $D(\tau,p)_s$:

$$\Phi(\tau_i,\Delta p_i,c(\mathbf{x})) = \int dp' D(\tau_i,p')_s D(\tau_i,p'-\Delta p_i)_s^{obs}, \quad (5)$$

where the extremum value of $\bar{\Phi}(\tau,p,c(\mathbf{x}))=\partial\Phi/\partial p=0$ is satisfied at the extremal coordinates $(\tau_i^*,\Delta p_i^*)$ for $i\in[1,2,...P]$:

$$\begin{aligned} \bar{\Phi}(\tau_i^*,\Delta p_i^*,c(\mathbf{x})) &= \int dp' D(\tau_i^*,p')_s \overbrace{\frac{\partial D(\tau_i^*,p'-p)_s^{obs}}{\partial p}}^{\bar{D}(\tau_i^*,p'-\Delta p_i^*)_s^{obs}}|_{p=\Delta p_i^*}, \\ &= 0, \end{aligned} \quad (6)$$

and the bar symbol denotes the derivative with respect to the p variable. In this case the implicit function theorem says the *skeletal* Fréchet derivative (Luo and Schuster, 1991a) is

$$\begin{aligned} \frac{\partial p}{\partial c(\mathbf{x})} &= -\frac{\partial\bar{\Phi}(\tau_i^*,\Delta p_i^*,c(\mathbf{x}))}{\partial c(\mathbf{x})}\Big/\frac{\partial\bar{\Phi}(\tau_i^*,\Delta p_i^*,c(\mathbf{x}))}{\partial p}, \\ &= -\frac{1}{A}\int dp'\frac{\partial D(\tau_i^*,p')_s}{\partial c(\mathbf{x})}\bar{D}(\tau_i^*,p'-\Delta p_i^*)_s^{obs}, \end{aligned} \quad (7)$$

where the integration in p' acts as a dot product between the predicted Fréchet derivative $\frac{\partial D(\tau_i^*,p')_s}{\partial c(\mathbf{x})}$ and the observed data $\bar{D}(\tau_i^*,p'-\Delta p_i^*)_s^{obs}$. Here,

$$\begin{aligned} A &= \frac{\partial\bar{\Phi}(\tau_i^*,\Delta p_i^*,c(\mathbf{x}))}{\partial p}, \\ &= \int dp' D(\tau_i^*,p')_s\frac{\partial^2 D(\tau^*,p'-p)_s^{obs}}{\partial p^2}|_{p=\Delta p_i^*}, \end{aligned} \quad (8)$$

is a normalization factor and from equation 1 we have the *data* Fréchet derivative in the $\tau-p$ domain:

$$\frac{\partial D(\tau_i^*,p)_s}{\partial c(\mathbf{x})} = \int_{B_o}\frac{\partial d(x_g,\tau_i^*+x_g p|x_s,0)}{\partial c(\mathbf{x})}dx_g. \quad (9)$$

The Fréchet derivative $\frac{\partial d(x_g,t|x_s,0)}{\partial c(\mathbf{x})}$ in the x_g-t domain on the rightside of equation 9 is (Luo and Schuster, 1991a)

$$\frac{\partial d(x_g,t|x_s,0)}{\partial c(\mathbf{x})} = \frac{2}{c(\mathbf{x})^3}\dot{g}(\mathbf{x},t|x_g,0)\star\dot{d}(\mathbf{x},t|x_s,0), \quad (10)$$

where the time derivative is indicated by the dot symbol, $g(\mathbf{x},t|x_g,0)$ is the acoustic Green's function for a point source at $\mathbf{g}=(x_g,0)$ excited at time 0, and the response is recorded by the geophone at $\mathbf{x}$ with listening time t. Plugging equation 10 into equation 9 gives

$$\frac{\partial D(\tau_i^*,p)_s}{\partial c(\mathbf{x})} = \frac{2}{c(\mathbf{x})^3}\times \overbrace{\int_{B_o}dx_g[\dot{g}(\mathbf{x},t|x_g,0)\star\dot{d}(\mathbf{x},t|x_s,0)]|_{t=\tau_i^*+x_g p}}^{Radon\ transform\ of\ predicted\ shot\ gather}, \quad (11)$$

which says that the data Fréchet derivative in the $\tau-p$ domain is the Radon transform of the predicted shot gather for a point scatterer at the trial image point $\mathbf{x}=(x,z)$. For $i=1$ $\gamma(\mathbf{x})=\Delta p_1\frac{\partial p_1}{\partial c(\mathbf{x})}$, which is the $\tau-p$ dot product between the shifted $\tau-p$ data $\bar{D}(\tau_1^*,p-\Delta p_1^*)_s^{obs}$ and the predicted shot gather $\frac{\partial D(\tau_1^*,p)_s}{\partial c(\mathbf{x})}$ for a point scatterer at $\mathbf{x}$. This interpretation is similar to the interpretation of generalized diffraction stack migration (GDM), where the migration image at $\mathbf{x}$ is the x_g-t dot product between the predicted and the actual shot gathers. Here, the predicted shot gather is that for a point scatterer at $\mathbf{x}$ (see Chapter 13 in Schuster (2017)). However, the problem with the GDM implementation is that the point-scatterer response must be computed and stored for each subsurface point. This can be prohibitively expensive so the reverse time migration approach should be preferred for implementation. In summary, the WRT method for early arrival P-wave is implemented in the following steps.

10.1190/segam2018-2998360.1

1. Assume a starting velocity model $c(\mathbf{x})^{(1)}$ and set the iteration index $k = 1$.
2. Apply the Radon transform to each recorded shot gather $d(x_g,t)_s^{(obs)}$ to get $D(\tau,p)_s^{obs}$. To focus the first arrival P-wave in $|D(\tau,p)_s^{obs}|$, we window the input data along the linear events in the shot gather.
3. Repeat step 2, except apply it to the predicted shot gathers $d(x_g,t)_s$ generated by a numerical solution to the acoustic wave equation, where the background velocity model is $c(\mathbf{x})^{(k)}$.
4. Cross-correlate the observed and predicted data in the Radon domain to estimate the skeletal data residuals $\Delta p = p - p^{obs}$ and $\Delta\tau = \tau - \tau_{obs}$ and shift the observed data in Radon domain using Δp and $\Delta\tau$ values.
5. Backpropagate the shifted data weighted by the Radon misfit $\frac{\beta}{2}\Delta\tau + \frac{1-\beta}{2}\Delta p$ to compute the gradient $\gamma(x)$.
6. Compute the velocity update $c(\mathbf{x})^{(k+1)}$ using equation 4. Increment the value of the iteration index by $k \rightarrow k+1$, and repeat steps 3-5 until the RMS residual $\frac{\beta}{2}\|\Delta\tau\|_2^2 + \frac{1-\beta}{2}\|\Delta p\|_2^2$ falls below a specified threshold value.

EXAMPLES

Synthetic Example

The WRT method is tested on acoustic data generated for the two-layer model shown in Figure 2a . The model consists of a near-surface layer with an irregular interface, which is discretized into 26×240 grid points with a gridpoint spacing of 2 m. There are 30 shots gathers with 8 m spacing and 120 geophones spaced every 1 m on the free surface. The source wavelet is a Ricker wavelet with a peak frequency of 30 Hz. In this example, the dominant wavelength is about 33 m. The initial velocity model is shown in Figure 2b. The velocity tomograms obtained by the WT and WRT methods are shown in Figures 2c and 2d, respectively. Figure 3 shows the comparison between the rates of residual reduction for the WT and WRT methods.

Field Data Example

In the western coast of Saudia Arabia, an earthquake rupture located near the Gulf of Aqaba can be seen at the surface (Klinger et al., 1999). In order to investigate the geometry of the fault and the alluvial sediments across the fault, a seismic survey was carried out near the Gulf of Aqaba in November 2013. The tomogram estimated by travel time tomography is shown in Figure 4a for reference. The seismic profile consists of traces recorded by 120 vertical-component geophones deployed at 2.5 m intervals along the survey line. A 90-kg accelerated-weight drop was used as a source at every geophone position to record 120 common shot gathers. The initial velocity model for both WT and WRT is a gradient velocity model shown in Figure 4b, which is discretized into 75x300 grid-points with the grid size of 1 m. In order to reduce the complexity of the input data and improve convergence, we use a multi-scale inversion approach with both the WT and WRT

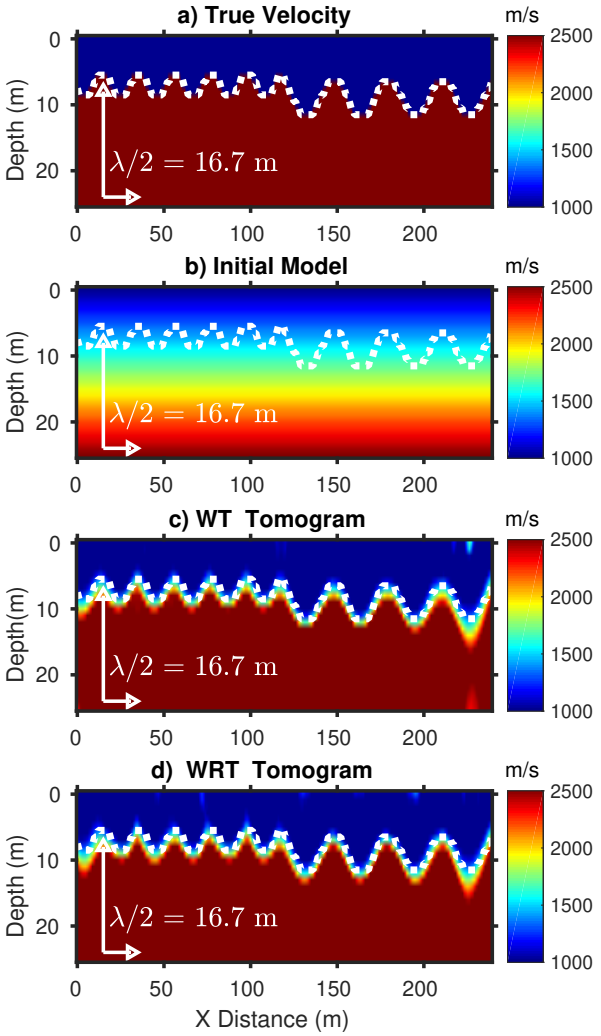

Figure 2: Synthetic data results: a) true velocity model, b) starting velocity model, c) P-velocity tomogram using the WT method, and d) P-velocity tomogram using the WRT method.

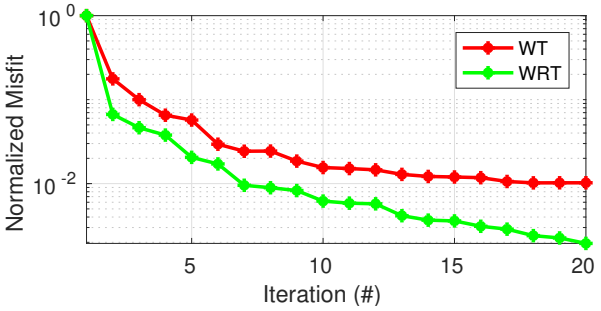

Figure 3: Normalized misfit versus iteration number for the synthetic model.

10.1190/segam2018-2998360.1

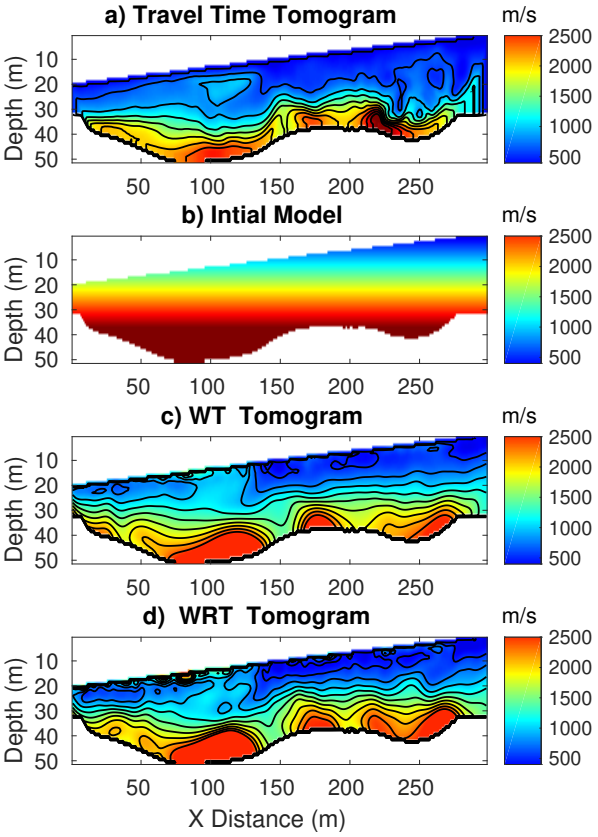

Figure 4: Aqaba field data results: a) P-velocity tomogram using ray-based travel time inversion, b) starting velocity model, c) P-velocity tomogram using the WT method, and d) P-velocity tomogram using the WRT method.

methods. The windowed early arrivals are band-pass filtered using 10, 30 and 50 Hz center frequencies. Starting from the lowest frequency, each tomogram estimated is used as an initial model for the inversion using the higher frequency band. Figure 5 shows that WRT reduces the residuals faster than the WT method. Comparison of the traces predicted by the WRT tomogram to observed common offset gathers shows that the WRT tomogram more closely predicts the recorded data than predicted by the Figure 4a ray-based tomogram.

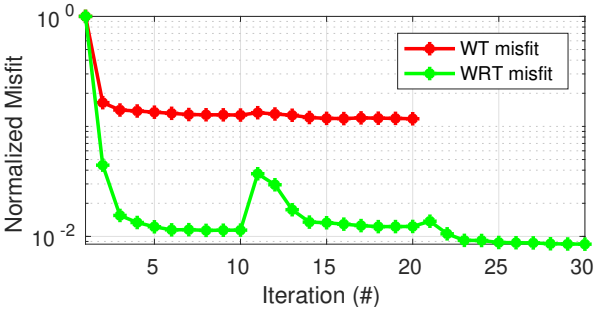

Figure 5: Normalized misfit versus iteration number for the Aqaba field data.

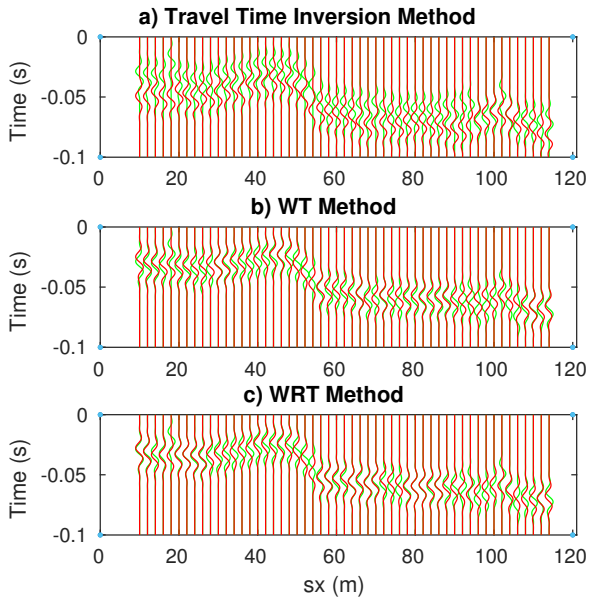

Figure 6: Comparison between predicted (red) and observed (green) common offset gathers for Aqaba data, a) travel time inversion method, b) WT method, and c) WRT method.

CONCLUSION

We presented the theory of wave equation Radon tomography that inverts for the velocity model from the skeletal data in the $\tau - p$ domain. Applying the linear Radon transform to a shot gather simplifies the complicated multiple arrivals to a few maximal amplitudes. This assumes that the complicated early arrivals have a quasi-linear moveout in local space-time windows. Picking the peak values (τ_i, p_i) and inverting them by a gradient optimization method can lead to robust convergence, where the cycle-skipping problems in FWI are largely avoided. Moreover, picking the peak values of the tau-p transform can be less time consuming than picking traveltimes for the WT method. This methodology can be adapted to reflections by applying the hyperbolic Radon transform to the reflection data. One tradeoff with the WRT method is that only the slopes and reflection times, not amplitudes, are explained so that WRT is unlikely to achieve the high-resolution capability of an ideal FWI method. Another trade-off is related to the quasi-linear moveout assumption. This assumption necessitates processing complicated non-linear arrivals in narrow spatial windows which reduces the focusing abilities of the Radon transform.

ACKNOWLEDGMENTS

We thank the sponsors for supporting 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 Supercomputing Laboratory at KAUST.

10.1190/segam2018-2998360.1

REFERENCES

Bickel, S. H., 2000, Focusing aspects of the hyperbolic Radon transform: Geophysics, **65**, 652–655, https://doi.org/10.1190/1.1444762.

Diebold, J. B., and P. L. Stoffa, 1981, The traveltime equation, tau-p mapping, and inversion of common midpoint data: Geophysics, **46**, 238–254, https://doi.org/10.1190/1.1441196.

Ibrahim, A., 2015, Separating simultaneous seismic sources using robust inversion of Radon and migration operators: Ph.D. thesis, University of Alberta.

Ibrahim, A., and M. D. Sacchi, 2014a, Eliminating blending noise using fast apex shifted hyperbolic Radon transform: 76th Annual International Conference and Exhibition, EAGE, Expanded Abstracts, ELI2–01, https://doi.org/10.3997/2214-4609.20141455.

Ibrahim, A., and M. D. Sacchi, 2014b, Simultaneous source separation using a robust Radon transform: Geophysics, **79**, no. 1, V1–V11, https://doi.org/10.1190/geo2013-0168.1.

Ibrahim, A., and M. D. Sacchi, 2015, Fast simultaneous seismic source separation using Stolt migration and demigration operators: Geophysics, **80**, no. 6, WD27–WD36, https://doi.org/10.1190/geo2015-0044.1.

Klinger, Y., L. Rivera, H. Haessler, and J.-C. Maurin, 1999, Active faulting in the Gulf of Aqaba: New knowledge from the Mw7.3 earthquake of 22 November 1995: Bulletin of the Seismological Society of America, **89**, 1025–1036.

Luo, Y., and G. T. Schuster, 1991a, Wave equation inversion of skeletalized geophysical data: Geophysical Journal International, **105**, 289–294, https://doi.org/10.1111/j.1365-246X.1991.tb06713.x.

Luo, Y., and G. T. Schuster, 1991b, Wave-equation traveltime inversion: Geophysics, **56**, 645–653,. https://doi.org/10.1190/1.1443081.

Milkereit, B., W. D. Mooney, and W. M. Kohler, 1985, Inversion of seismic refraction data in planar dipping structure: Geophysical Journal International, **82**, 81–103, https://doi.org/10.1111/j.1365-246X.1985.tb05129.x.

Schuster, G. T., 2017, Seismic Inversion: SEG Books.

Stoffa, P., 1989, Tau-p: A plane wave approach to the analysis of seismic data: Kluwer Academic Publishers.

Stoffa, P. L., J. B. Diebold, and P. Buhl, 1981, Inversion of seismic data in the τ–p plane: Geophysical Research Letters, **8**, 869–872, https://doi.org/10.1029/GL008i008p00869.

Woodward, M., 1989, Wave-equation tomography: Ph.D. thesis, Stanford University.

Woodward, M. J., 1992, Wave-equation tomography: Geophysics, **57**, 15–26, https://doi.org/10.1190/1.1443179.