Accurate implementation of two-way wave-equation operators

Ali Almomin

ABSTRACT

I present a complete derivation of wave-equation operators for nonlinear modeling, linearized modeling and migration, tomographic forward and adjoint operators and wave-equation migration velocity analysis (WEMVA) operators. The derivation is done in time domain using a chain of simple linear operators. The results show that all linearizations and adjoints are correct for any media.

INTRODUCTION

Seismic modeling, imaging, analysis and inversion utilizes a form of the wave-equation. Commonly, the different operators are derived starting from some assumption about the imaging condition and/or the medium. This often results in an inaccurate linearization which has the wrong scale and units. One common issue is that most effort is put into getting the pair-operators to be correctly adjoints but not necessarily consistent with the nonlinear operator that they are derived from. Furthermore, the implementations in time domain are not accurate adjoints of the operators which results in sub-optimal convergence rates in inversion (Ji, 2009). Finally, the source injection and boundary conditions are normally ignored in all these operators resulting in further hindrance for the inversion.

I show the details of deriving wave-equation operators, which are not always obvious from the theory. The derivation is performed in time domain using a chain of simple linear operators. The adjoint of each operator can be tested independently which simplifies the problem. Moreover, the boundary conditions, data interpolation and source injection are taken care of in this derivation.

THEORY

Nonlinear Modeling

The first and most important operator to get correctly is the nonlinear modeling operator. Although I call it nonlinear modeling operator (referring to its relationship with medium parameters), this operator is in fact linear with the source function, i.e. the right-hand side of the wave equation. Once I get the operator and its adjoint correctly, I can easily implement simple modifications to the source function to create the other operators. First, I write the forward modeling in compact notation as:

$$\mathbf{d} = \mathbf{F}\mathbf{f}, \tag{1}$$

where $\mathbf{d}$ is the modeled data, $\mathbf{F}$ is the modeling operator, and $\mathbf{f}$ is the source function. For simplicity, I will take care of spatial and temporal difference between the source, data and propagation grid outside the operator $\mathbf{F}$. So, I can rewrite the forward modeling of one experiment as:

$$\mathbf{d} = \mathbf{K_r'}\mathbf{L_r'}\mathbf{F}\mathbf{L_s}\mathbf{K_s}\mathbf{f}, \tag{2}$$

where $\mathbf{K}$ is a zero-padding operator from the source/receiver spatial grid to the propagation spatial grid and $\mathbf{L}$ is an interpolation operator from the source/receiver temporal grid to the propagation temporal grid. The subscripts $_s$ and $_r$ denotes the source and receiver sides, respectively, and the $'$ denotes the adjoint. It is important to keep the adjoint process in mind when designing the forward operator. When performing the adjoint, we want to interpolate the data before injecting it, hence the $\mathbf{L_r'}$. One of the common mistakes is to replace the $\mathbf{L_r'}$ with a subsampling operator. This replacement either causes the adjoint to be noisy (by injecting zeros instead of interpolating the data) or an incorrect adjoint (by interpolating the data). We can see now that the input and output of the operator $\mathbf{F}$ have the size of the propagation grid, both in space and time. Therefor, the operator $\mathbf{F}$ can be derived by solving the two-way wave-equation:

$$\left(\nabla^2 - \mathbf{v}^{-2}\frac{\partial^2}{\partial t^2}\right)\mathbf{p}(t) = \mathbf{f}(t), \tag{3}$$

where t is time, $\mathbf{v}$ is the velocity model, $\mathbf{p}$ is the propagated wavefield, and $\mathbf{f}$ is the source wavefield. By using finite-difference for the time derivative, the propagation can be rewritten as:

$$\nabla^2\mathbf{p}(it) - \frac{\mathbf{v}^{-2}}{\Delta t^2}\left(\mathbf{p}(it+1) - 2\mathbf{p}(it) + \mathbf{p}(it-1)\right) = \mathbf{f}(it). \tag{4}$$

I then rearrange the previous equation to have a recursive time-stepping scheme as follows:

$$\mathbf{p}(it+1) = \mathbf{v}^2\Delta t^2\left(\nabla^2\mathbf{p}(it) - \mathbf{f}(it)\right) + 2\mathbf{p}(it) - \mathbf{p}(it-1). \tag{5}$$

One common mistake is to inject the source before applying the Laplacian, which results in solving the equation for another "effective" source. Another common mistake is to not scale the source by the proper coefficient. Both this mistakes affect the linearity between this operator and the following operators.

I will now do two steps to simplify the recursive scheme into a lower triangular matrix and to simplify its recursive adjoint. First, I will shift the time sample of the propagated wavefield by one to use the computed one as reference. Second, I will take the different times of p as common factors to clarify the coefficients of the recursive operator. These changes result in the following equation:

$$\mathbf{p}(it) = -\mathbf{v}^2\Delta t^2\mathbf{f}(it) + \left(\mathbf{v}^2\Delta t^2\nabla^2 + 2\right)\mathbf{p}(it-1) - \mathbf{p}(it-2). \tag{6}$$

I can now define the coefficients as the following operators:

$$\begin{aligned}\mathbf{C}_1 &= -\mathbf{v}^2\Delta t^2\\ \mathbf{C}_2 &= \mathbf{v}^2\Delta t^2\nabla^2 + 2\mathbf{I}\\ \mathbf{C}_3 &= -\mathbf{I},\end{aligned} \tag{7}$$

where $\mathbf{I}$ is identity operator. Equation 6 can be rewritten as:

$$\mathbf{p}(it) = \mathbf{C}_1\mathbf{f}(it) + \mathbf{C}_2\mathbf{p}(it-1) + \mathbf{C}_3\mathbf{p}(it-2). \tag{8}$$

Although the previous equation describes the forward time-stepping correctly, its adjoint is not a recursive operator. The reason for this problem is the $\mathbf{C}_1$ operator does not commute with the other coefficient operators. To avoid this problem, I need to pull the coefficient $\mathbf{C}_1$ outside the recursion operator and into a separate operator, which I am going to call $\mathbf{S}$. In other words, I am going to break the operator $\mathbf{F}$ into two operators: a scaling operator and a recursive operator. The scaling operator $\mathbf{S}$ is simply multiple copies of $\mathbf{C}_1$ to multiply all time steps. This changes equation 2 into:

$$\mathbf{d} = \mathbf{K_r'L_r'GSL_sK_sf}, \tag{9}$$

whereas the operator $\mathbf{G}$ is the recursive operator described by:

$$\mathbf{p}(it) = \mathbf{q}(it) + \mathbf{C}_2\mathbf{p}(it-1) + \mathbf{C}_3\mathbf{p}(it-2), \tag{10}$$

where $\mathbf{q}$ is the source wavefield after applying the operator $\mathbf{S}$. The operators $\mathbf{S}$ and $\mathbf{L}_s$ commute because the interpolation is on the time axis only. Therefore, I can rewrite equation 9 as:

$$\mathbf{d} = \mathbf{K_r'L_r'GL_sSK_sf}, \tag{11}$$

which is more computationally efficient. I can now write the adjoint modeling as:

$$\mathbf{f} = \mathbf{K_s'S'L_s'G'L_rK_rd}. \tag{12}$$

The operator $\mathbf{G}'$ can be described by:

$$\mathbf{q}(it) = \mathbf{p}(it) + \mathbf{C}_2'\mathbf{q}(it+1) + \mathbf{C}_3'\mathbf{q}(it+2), \tag{13}$$

and the operator $\mathbf{C}_2'$ is:

$$\mathbf{C}_2' = (\mathbf{v}^2\Delta t^2\nabla^2 + 2\mathbf{I})' = (\nabla^2)'\mathbf{v}^2\Delta t^2 + 2\mathbf{I}. \tag{14}$$

equations 12 to 14 show why doing the forward propagation with the time axis reversed does not give the proper adjoint due to three steps. First, applying operator $\mathbf{S}'$ should take place after $\mathbf{G}'$. Second, the Laplacian operator in $\mathbf{C}_2'$ comes after scaling by $\mathbf{v}^2\Delta t^2$. Finally, the Laplacian operator is not self-adjoint.

The last step is to add the boundary condition. In my derivation, I will use an absorption boundary that removes waves coming back from the boundaries as described by Israeli and Orszag (1981). For this boundary condition, I need to define the quantity $\mathbf{w}$ which will have a value of 0 where there is no absorbing and linearly increase across the absorbing layer to a maximum value of $0.15\frac{\Delta t}{\Delta s}\mathbf{v}$ at the outer edges where Δs is the spatial sampling in the direction of absorbing. Now, I can apply the absorbing condition after propagating each time step as:

$$\mathbf{p}(it) = \mathbf{p}'(it) + \mathbf{w}\left(\mathbf{p}(it-1) - \mathbf{p}(it) - \mathbf{v}\Delta t\frac{\partial}{\partial s}\mathbf{p}(it-1)\right), \tag{15}$$

where $\frac{\partial}{\partial s}$ is the spatial first derivative in the direction of absorbing and $\mathbf{p}'(it)$ is the time slice before applying the boundary condition. I can arrange the previous equation by taking the time slices of $\mathbf{p}$ as common factors:

$$\mathbf{p}(it) = (1-\mathbf{w})\,\mathbf{p}'(it) + \mathbf{w}\left(1 - \mathbf{v}\Delta t\frac{\partial}{\partial s}\right)\mathbf{p}(it-1). \tag{16}$$

Next, I substitute the value of $\mathbf{p}'(it)$ from equation 8 into the previous equation:

$$\mathbf{p}(it) = (1-\mathbf{w})\left(\mathbf{C}_1\mathbf{f}(it) + \mathbf{C}_2\mathbf{p}(it-1) + \mathbf{C}_3\mathbf{p}(it-2)\right) + \mathbf{w}\left(1 - \mathbf{v}\Delta t\frac{\partial}{\partial s}\right)\mathbf{p}(it-1). \tag{17}$$

Again, I rearrange the equation and take the time slices of $\mathbf{p}$ and $\mathbf{f}$ as common factors to get:

$$\mathbf{p}(it) = \mathbf{C}_4\mathbf{f}(it) + \mathbf{C}_5\mathbf{p}(it-1) + \mathbf{C}_6\mathbf{p}(it-2), \tag{18}$$

where the new coefficients are:

$$\begin{aligned}
\mathbf{C}_4 &= (\mathbf{w}-\mathbf{I})\mathbf{v}^2\Delta t^2 \\
\mathbf{C}_5 &= (\mathbf{I}-\mathbf{w})\left(\mathbf{v}^2\Delta t^2\nabla^2 + 2\mathbf{I}\right) + \mathbf{w}\left(\mathbf{I} - \mathbf{v}\Delta t\frac{\partial}{\partial s}\right) \\
\mathbf{C}_6 &= \mathbf{w}-\mathbf{I}.
\end{aligned} \tag{19}$$

All I have to do now is to redefine the scaling operator $\mathbf{S}$ to be multiple copies of $\mathbf{C}_4$ instead of $\mathbf{C}_1$ and the operator $\mathbf{G}$ to have the coefficients $\mathbf{C}_5$ and $\mathbf{C}_6$ instead of $\mathbf{C}_2$ and $\mathbf{C}_3$, respectively. Also, operator $\mathbf{G}'$ requires the proper adjoint of $\mathbf{C}_5$ which can be written as:

$$\mathbf{C}_5' = \left((\nabla^2)'\mathbf{v}^2\Delta t^2 + 2\mathbf{I}\right)(\mathbf{I}-\mathbf{w}) + \left(\mathbf{I} - \left(\frac{\partial}{\partial s}\right)'\mathbf{v}\Delta t\right)\mathbf{w}. \tag{20}$$

One possible adjustment to the forward operator is to use a normalized interpolation operators. The reason for the normalization is that we do not want the amplitudes to change for the same model and wavelet if we model with finer propagation time steps. The normalization is easy in this case since the ratio between the wavelet time sampling and the propagation time sampling is a constant.

Born Operator

The Born operator relates perturbation in the velocity (i.e. the image), $\mathbf{r}$, to changes in the data. In the theory, we see that the Born operator is simply the solution to the two-way wave-equation but with a virtual source. The virtual source is the background wavefield $\mathbf{u_{s0}}$ after scattering by the perturbation $\mathbf{r}$ and a second derivative in time. Therefore, I first need to compute the background wavefield similarly to how I computed the data in equation 11 but without the last truncation step:

$$\mathbf{u_{s0}} = \mathbf{L_s'GL_sSK_sf}, \tag{21}$$

Notice that the operator $\mathbf{G}$ is now followed by the operator $\mathbf{L_s'}$ instead of $\mathbf{L_r'}$ since I am using this wavefield as a secondary source instead of recording it as data. It is important to keep the operator $\mathbf{L_s'}$ for two reasons. First, to have proper adjoint (similar explanation to the nonlinear modeling operator). Second, to reduce the number of times we do the expensive scattering/imaging step. Once $\mathbf{u_{s0}}$ is computed, I construct the operator $\mathbf{U_{s0}}$ which has a model space of the image and a data space of scattered wavefield. In other words, this operator is the adjoint of imaging condition. With a second time derivative operator $\mathbf{T}$, I can write the forward Born modeling operator as:

$$\mathbf{\Delta d} = \mathbf{K_r'L_r'GL_sST\widehat{U_{s0}}r}, \tag{22}$$

where the hat $\widehat{\ }$ indicates an extended image space with either offset or time lags. The adjoint of Born operator, i.e. the RTM operator, can be written as:

$$\mathbf{r} = \mathbf{\widehat{U_{s0}'}T'S'L_s'G'L_rK_r\Delta d}. \tag{23}$$

Tomographic Operator

The tomographic operator relates changes in the propagation velocity, $\mathbf{b}$, to changes in the data using the perturbation (or the image) $\mathbf{r}$. Similar to the Born operator, the tomographic operator is also a solution to the two-way wave-equation but with two virtual sources: one source is the background wavefield $\mathbf{u_{s0}}$ after scattering by the propagation-velocity perturbation, $\mathbf{\Delta b}$, and a second derivative in time and the second source is the perturbed wavefield $\mathbf{u_{\Delta s}}$ after scattering by the propagation-velocity perturbation $\mathbf{\Delta b}$ and a second derivative in time as well. The perturbed wavefield $\mathbf{u_{\Delta s}}$ can be computed as:

$$\mathbf{u_{\Delta s}} = \mathbf{L_s'GL_sST\widehat{U_{s0}}r}. \tag{24}$$

Once $\mathbf{u_{\Delta s}}$ is computed, I construct the operator $\mathbf{U_{\Delta s}}$ the same way I constructed $\mathbf{U_{s0}}$. Moreover, I need to define an operator $\mathbf{R}$ which similar to the $\widehat{\mathbf{U_{s0}}}$ but with the model space being the background (unscattered) wavefield and the perturbation $\mathbf{r}$ being in the operator itself instead of being in the model space. Now, I can write the forward tomographic operator as:

$$\mathbf{\Delta d} = \mathbf{K_r'L_r'GL_sSTU_{\Delta s}\Delta b} + \mathbf{K_r'L_r'GL_sSTRL_s'GL_sSTU_{s0}\Delta b}. \tag{25}$$

The adjoint of the tomographic operator can be written as:

$$\mathbf{\Delta b} = \mathbf{U}'_{\mathbf{\Delta s}}\mathbf{T}'\mathbf{S}'\mathbf{L}'_{\mathbf{s}}\mathbf{G}'\mathbf{L}_{\mathbf{r}}\mathbf{K}_{\mathbf{r}}\mathbf{\Delta d} + \mathbf{U}'_{\mathbf{s0}}\mathbf{T}'\mathbf{S}'\mathbf{L}'_{\mathbf{s}}\mathbf{G}'\mathbf{L}_{\mathbf{s}}\mathbf{R}'\mathbf{T}'\mathbf{S}'\mathbf{L}'_{\mathbf{s}}\mathbf{G}'\mathbf{L}_{\mathbf{r}}\mathbf{K}_{\mathbf{r}}\mathbf{\Delta d}. \tag{26}$$

WEMVA Operator

The WEMVA operator relates changes in the propagation velocity, $\mathbf{b}$, to changes in the perturbation, $\mathbf{r}$, using the data. The WEMVA operator is slightly different than the tomographic operators because both its model space and data space are in the "image" domain (either $\mathbf{b}$ or $\mathbf{r}$) and both the source function and the data are part of the operator itself. However, as the theory shows, the adjoint WEMVA operator has the same expression as the adjoint of the tomographic operator after switching the model space with some components of the operator. Hence, I need to compute the receiver-side background wavefield as:

$$\mathbf{u}_{\mathbf{r0}} = \mathbf{S}'\mathbf{L}'_{\mathbf{s}}\mathbf{G}'\mathbf{L}_{\mathbf{r}}\mathbf{K}_{\mathbf{r}}\mathbf{\Delta d}. \tag{27}$$

However, there is an issue on the receiver-side because the order of operators between the scattering/image and second time derivative will not be the same as the tomographic operator. To resolve this issue, I can simply include the time derivative in the computation of the receiver-side background wavefield and redefine it as:

$$\mathbf{u}_{\mathbf{r0}} = \mathbf{T}'\mathbf{S}'\mathbf{L}'_{\mathbf{s}}\mathbf{G}'\mathbf{L}_{\mathbf{r}}\mathbf{K}_{\mathbf{r}}\mathbf{\Delta d}. \tag{28}$$

Then, I construct the operator $\mathbf{U}_{\mathbf{r0}}$ the same way I constructed the operator $\mathbf{U}_{\mathbf{s0}}$ but using the receiver-side background wavefield. I can now define the adjoint WEMVA operator as:

$$\mathbf{\Delta b} = \mathbf{U}'_{\mathbf{r0}}\mathbf{L}'_{\mathbf{s}}\mathbf{G}\mathbf{L}_{\mathbf{s}}\mathbf{S}\mathbf{T}\widehat{\mathbf{U}_{\mathbf{s0}}}\mathbf{\Delta r} + \mathbf{U}'_{\mathbf{s0}}\mathbf{T}'\mathbf{S}'\mathbf{L}'_{\mathbf{s}}\mathbf{G}'\mathbf{L}_{\mathbf{s}}\widehat{\mathbf{U}_{\mathbf{r0}}}\mathbf{\Delta r}, \tag{29}$$

and the forward WEMVA operator as:

$$\mathbf{\Delta r} = \widehat{\mathbf{U}'_{\mathbf{r0}}}\mathbf{L}'_{\mathbf{s}}\mathbf{G}\mathbf{L}_{\mathbf{s}}\mathbf{S}\mathbf{T}\mathbf{U}_{\mathbf{s0}}\mathbf{\Delta b} + \widehat{\mathbf{U}'_{\mathbf{s0}}}\mathbf{T}'\mathbf{S}'\mathbf{L}'_{\mathbf{s}}\mathbf{G}'\mathbf{L}_{\mathbf{s}}\mathbf{U}_{\mathbf{r0}}\mathbf{\Delta b}. \tag{30}$$

SYNTHETIC EXAMPLE

A simple synthetic model with a Gaussian anomaly (representing $\Delta\mathbf{b}$) and a spike (representing $\mathbf{r}$) is used for the synthetic examples. The true velocity is shown in Figure 1. The background velocity is 3 km/s and the Gaussian anomaly is located at 1 km in z-axis and x-axis. The perturbation spike is located at 2 km in z-axis and x-axis. Both anomalies have an amplitude of 300 m/s. The sampling for both spatial axes is 10 m. A Ricker wavelet with a fundamental frequency of 15 Hz is used to model the data. The source spacing is 50 m and the receiver spacing is 10 m.

To test the RTM operator, I start from a propagation velocity model that is the same as the true model except for the perturbation spike. The result of applying

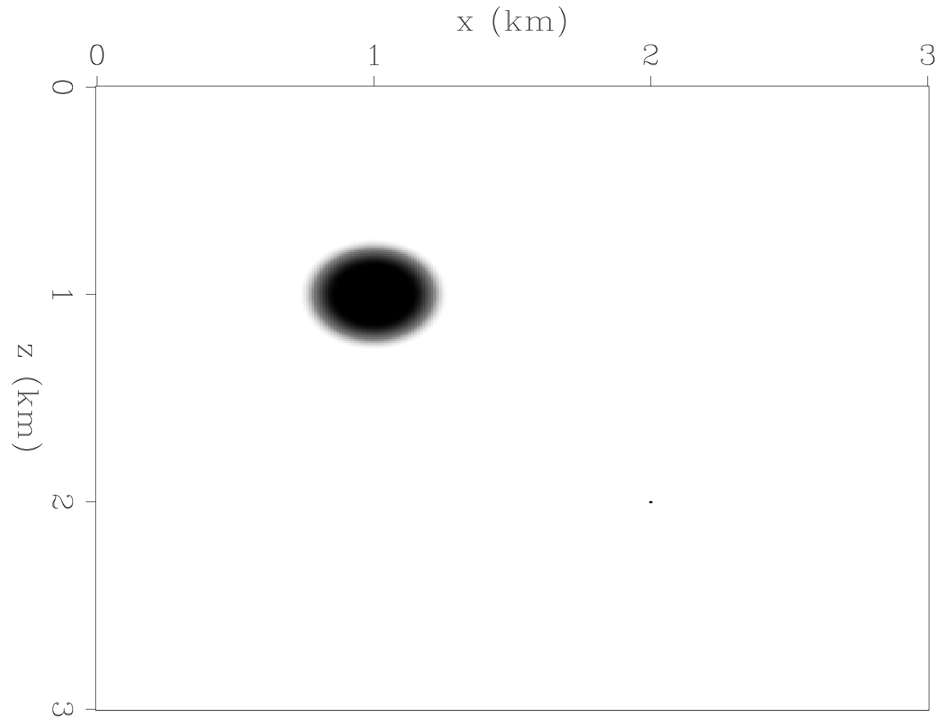

Figure 1: The true velocity model. [**ER**]

the RTM operator is shown in Figure 2. Next, I test the tomographic operator by computing the data residual due to removing the Gaussian anomaly. I use the RTM results as perturbation to estimate the update of the propagation velocity. The results of using the adjoint tomographic operator is shown in Figure 3. Finally, I test the WEMVA operator by computing the image residual due to removing the Gaussian anomaly. The results of using the adjoint WEMVA operator is shown in Figure 4. Both operators have the correct direction for the anomaly.

CONCLUSIONS

I provide the details of deriving two-way wave-equation operators while taking care of practical issues such as boundary conditions and source injection. The results of all the operators are consistent in terms unit and scale including the original wave-equation. Getting consistent operators is important when trying to use joint inversion without manually or empirically guessing a ratio between the outputs of different operators. Also, getting correct linearizations and adjoint can help the convergence rate of the inversion.

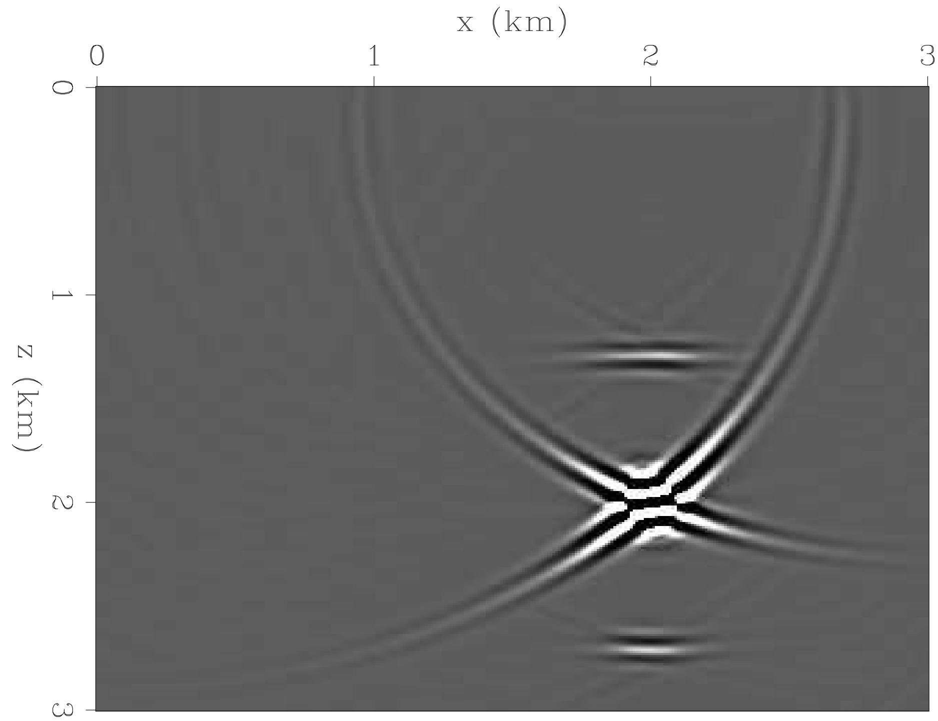

Figure 2: The RTM image. [**ER**]

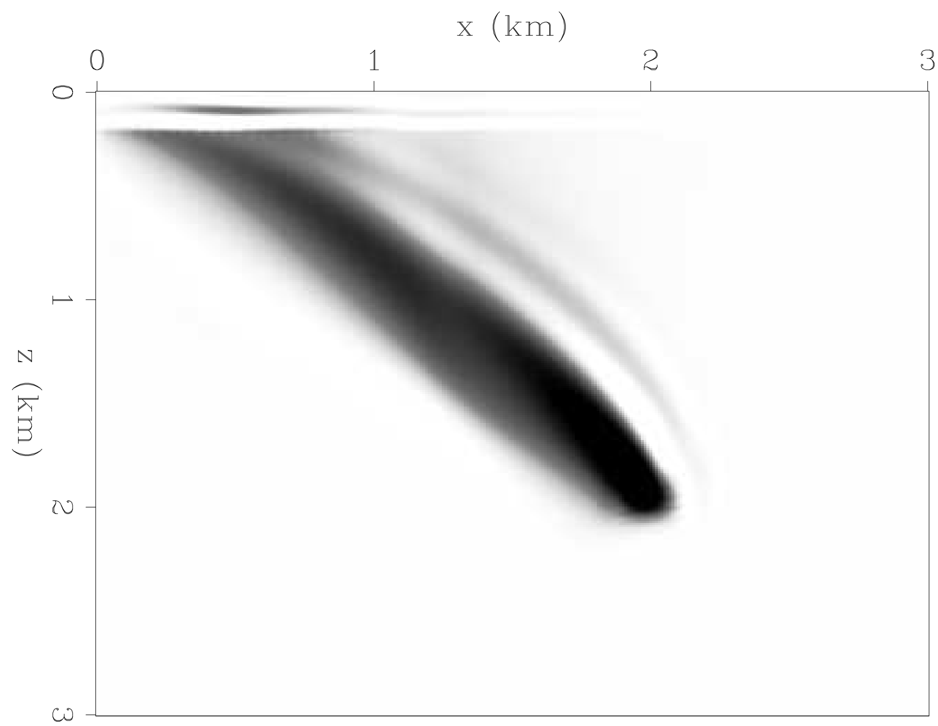

Figure 3: The tomographic adjoint on the data residual. [**ER**]

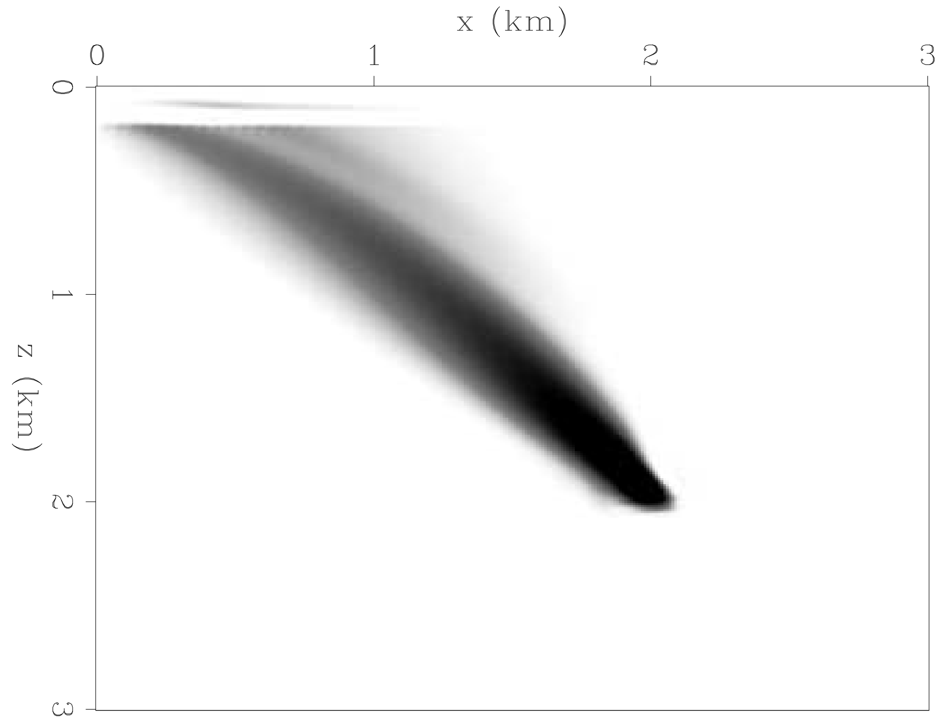

Figure 4: The WEMVA adjoint on the image residual. [**ER**]

REFERENCES

Israeli, M. and S. A. Orszag, 1981, Approximation of Radiation Boundary Conditions: Journal of Computational Physics, **41**, 115–135.

Ji, J., 2009, An exact adjoint operation pair in time extrapolation and its application in least-squares reverse-time migration: Geophysics, **74**, H27–H33.