Two-way wave-equation operators for non-constant density acoustic isotropic media

Ettore Biondi and Ossian O'Reilly

ABSTRACT

We derive two-way wave-equation operators in the time domain for isotropic non-constant density media with a finite-difference scheme. We present the chains of linear operators necessary for non-linear modeling, linearized modeling, and non-constant density migration. We also show that radiation patterns obtained from linearized modeling agree with theoretical results. With these correct forward-adjoint operators, multi-parameter full waveform inversion for non-constant density media can be pursued.

INTRODUCTION

Seismic data inversion is and has been one of the most challenging problems encountered by geophysicists. Since its first envision by Tarantola (1984), full waveform inversion (FWI) was derived in the isotropic non-constant density acoustic approximation. Moreover, it is known that multi-parameter FWI is now essential to correctly match real data amplitudes (Operto et al., 2013). Therefore, having reliable non-linear modeling operators for non-constant density is very important in case of acoustic approximation. In addition, obtaining accurate adjoint operators is fundamental to achieve optimal convergence rates during seismic data inversion (Ji, 2009).

We implement a chain of linear operators for two-way wave-equation non-linear and linearized modeling for non-constant density media. The derivation is carried out in the time domain with a staggered-grid finite-difference scheme. For the linearized modeling operator the radiation patterns are compared with the theoretical ones (Aki and Richards, 2002). These accurate operators are extremely important in case we perform modeling and FWI in acoustic approximation.

ACOUSTIC WAVE EQUATION AND BORN APPROXIMATION

As shown in the Appendix, the non-constant density wave equation for acoustic media can be derived assuming that shear stresses are null and can be written as:

$$\left[\frac{1}{K(\mathbf{r})}\frac{\partial^2}{\partial t^2} - \nabla\cdot\frac{1}{\rho(\mathbf{r})}\nabla\right] p(\mathbf{r},t) = s(\mathbf{r},t), \tag{1}$$

where $K(\mathbf{r})$ and $\rho(\mathbf{r})$ are the medium's bulk modulus and density, respectively; $p(\mathbf{r},t)$ is the propagating pressure, $s(\mathbf{r},t)$ is the source term and $\mathbf{r}$ is the position vector defining each point in the medium. By perturbing the medium's properties and neglecting the higher-order terms in the equation we can derive the linearized wave equation, also known as the Born approximation. This wave equation is linear with respect to property perturbations $\delta K(\mathbf{r})$ and $\delta p(\mathbf{r})$, and is defined as follows:

$$\left[\frac{1}{K_0(\mathbf{r})}\frac{\partial^2}{\partial t^2} - \nabla\cdot\frac{1}{\rho_0(\mathbf{r})}\nabla\right]\delta p(\mathbf{r},t) = \frac{\delta K(\mathbf{r})}{K_0(\mathbf{r})^2}\frac{\partial^2 p_0(\mathbf{r},t)}{\partial t^2} - \nabla\cdot\frac{\delta\rho(\mathbf{r})}{\rho_0(\mathbf{r})^2}\nabla p_0(\mathbf{r},t), \tag{2}$$

where $K_0(\mathbf{r})$ and $p_0(\mathbf{r})$ are the background medium's properties, $p_0(\mathbf{r},t)$ is the pressure field propagated through the background medium, and $\delta p(\mathbf{r},t)$ is the scattered pressure field. In Equation 2 the two right-hand side terms are also commonly called secondary sources. These two terms represent the pressure scattering off perturbations in the bulk modulus and density respectively. Because our model is composed of K and ρ, we have two different imaging conditions for the two parameters, and are given by:

$$\delta\tilde{K}(\mathbf{r}) = \frac{1}{K_0(\mathbf{r})^2}\left[\frac{\partial^2 p_0(\mathbf{r})}{\partial t^2} \star \delta p'(\mathbf{r})\right](0), \tag{3}$$

$$\delta\tilde{\rho}(\mathbf{r}) = \frac{1}{\rho_0(\mathbf{r})^2}\left[\nabla p_0(\mathbf{r}) \mathbin{\dot{\star}} \nabla\delta p'(\mathbf{r})\right](0), \tag{4}$$

where $[f\star g](0)$ and $[f\mathbin{\dot{\star}} g](0)$ denote the zero-lag crosscorrelation and the zero-lag crosscorrelation of the scalar product of two functions respectively, and:

$$\delta p'(\mathbf{r},t) = \sum_{i=1}^{N_d}\tilde{g}(\mathbf{r},t,\mathbf{r}_i;K,\rho) * \delta d(\mathbf{r}_i,t), \tag{5}$$

in which $\delta p'(\mathbf{r},t)$ represents the back-propagated data perturbations or residual during an inversion, obtained by convolving the $\delta d(\mathbf{r}_i,t)$ with the anticausal Green's function $\tilde{g}(\mathbf{r},t,\mathbf{r}_i;K,\rho)$ for all the N_g receivers and a given source. The reader interested in the derivation of these equations can find them in the Appendix.

FORWARD AND ADJOINT PROPAGATION

The non-constant density wave equation in acoustic approximation (1) is linear with respect to the source term on the right-hand side. In fact, as shown by Almomin (2013) for the constant-density case, we can write a linear operator as:

$$\mathbf{d} = \tilde{\mathbf{F}}\mathbf{w}, \tag{6}$$

where $\mathbf{d}$ is model data, $\tilde{\mathbf{F}}$ that is the propagator which is non-linear with respect to bulk modulus and density, and $\mathbf{w}$ is the source wavelet. The propagator can be split into a chain of linear operators as:

$$\mathbf{d} = \mathbf{R}^*\mathbf{L}_g^*\mathbf{F}\mathbf{L}_s\mathbf{R}\mathbf{w}, \tag{7}$$

where $\mathbf{R}$ is a time interpolator that resamples the source wavelet from seismic sampling to finer sampling to have stable propagation; $\mathbf{L}_s$ is a space interpolation to inject the wavelet at the source position into the model, $\mathbf{F}$ is the operator that solve equation 1; $\mathbf{L}_g^*$ is again a space interpolator that extracts the pressure field at the receiver location; and $\mathbf{R}^*$ is the adjoint of the time interpolator that transforms the extracted data from fine sampling to seismic sampling.

To derive the actual form of $\mathbf{F}$, we need to discretize both time and space in equation 1. Using a simple second-order approximation for the time derivative (i.e., using a leapfrog time integrator), we can write the following recursive equation for each point in the model grid:

$$\mathbf{p}(it) = \mathbf{q}(it-1) + \mathbf{C}\mathbf{p}(it-1) - \mathbf{p}(it-2), \tag{8}$$

where $\mathbf{p}(it)$ denotes the pressure field at the time step it, $\mathbf{q}(it-1)$ is the scaled source wavelet obtained as follows:

$$\mathbf{q}(it-1) = \mathbf{S}\mathbf{w}(it-1) = \mathbf{K}\Delta t^2 \mathbf{w}(it-1), \tag{9}$$

and the operator $\mathbf{C}$ is given by:

$$\mathbf{C} = \mathbf{K}\Delta t^2 \nabla \cdot \frac{1}{\rho} \nabla, \tag{10}$$

in which $\mathbf{K}$ is a diagonal operator containing the bulk modulus of the medium; Δt is the fine propagation time sampling; and $\nabla \cdot \frac{1}{\rho} \nabla$ denotes an operator that computes the divergence of a density-scaled gradient of $\mathbf{p}(it-1)$. It is known that central different operators produce inaccurate derivative computation that introduces numerical error during the propagation (Mattsson, 2012); therefore, we implement the $\nabla \cdot \frac{1}{\rho} \nabla$ operator on a staggered grid (Figure 1). This operator in 2D can be written as follows:

$$\nabla \cdot \frac{1}{\rho} \nabla = \mathbf{D}_x^- \mathbf{B}_x \mathbf{D}_x^+ + \mathbf{D}_z^- \mathbf{B}_z \mathbf{D}_z^+, \tag{11}$$

where $\mathbf{D}^+$ and $\mathbf{D}^-$ are the forward and backward first-order derivative operators respectively, and $\mathbf{B}$ operators are diagonal matrices that contain the staggered inverse of the density values along the main diagonal. To compute these values we use a simple average on the inverse of the density along each direction.

The adjoint operator of $\tilde{\mathbf{F}}$ can be easily obtained by taking the adjoint of the chain of operators of Equation 7 that is:

$$\tilde{\mathbf{w}} = \mathbf{R}^* \mathbf{L}_s^* \mathbf{S} \mathbf{P}^* \mathbf{L}_g \mathbf{R} \mathbf{d}, \tag{12}$$

where we have split the propagation operator as:

$$\mathbf{F} = \mathbf{P}\mathbf{S}, \tag{13}$$

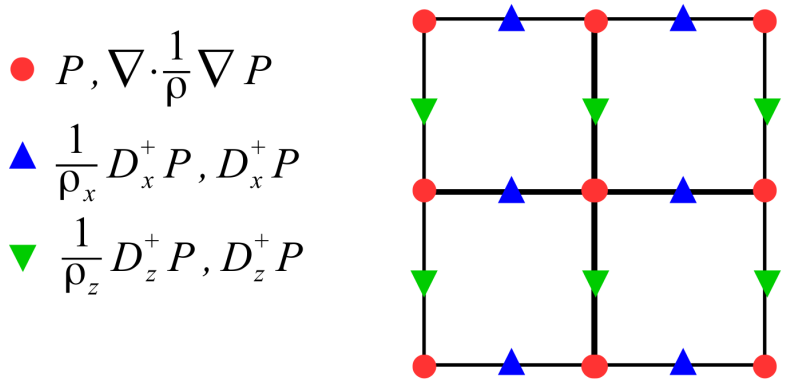


Figure 1: 2D staggered-grid scheme used for computing the $\nabla \cdot \frac{1}{\rho} \nabla$. The pressure and the output of this operator lay on the regular grid (red circles). Instead, the first-order derivative in the two spatial dimension and their density-scaled versions are computed on two different staggered grids (blue and green triangles). [**NR**]

which is a chain of a scaling operator $\mathbf{S}$ and a recursive operator $\mathbf{P}$. The adjoint of the recursive operator is given by:

$$\mathbf{q}(it) = \mathbf{p}(it+1) + \mathbf{C}^* \mathbf{q}(it+1) - \mathbf{q}(it+2), \tag{14}$$

where we are back-solving a system of equations for $\mathbf{q}$, which corresponds to back propagate the data from the receivers; and the adjoint of $\mathbf{C}$ is written as follows:

$$\mathbf{C}^* = \left(\nabla \cdot \frac{1}{\rho} \nabla \right)^* \mathbf{K} \Delta t^2. \tag{15}$$

The adjoint divergence of the scaled gradient is defined as:

$$\left(\nabla \cdot \frac{1}{\rho} \nabla \right)^* = \mathbf{D}_x^{+*} \mathbf{B}_x \mathbf{D}_x^{-*} + \mathbf{D}_z^{+*} \mathbf{B}_z \mathbf{D}_z^{-*}. \tag{16}$$

As shown in the Appendix, if we assume the pressure field is null outside the computational domain, we have:

$$\mathbf{D}^{+*} = -\mathbf{D}^-, \tag{17}$$
$$\mathbf{D}^{-*} = -\mathbf{D}^+, \tag{18}$$

that cause the operator $\nabla \cdot \frac{1}{\rho} \nabla$ to be self-adjoint.

Forward and adjoint propagation examples

We run two numerical experiments in which we created a model with two velocity constrasts, one case these variations are caused by density changes only; whereas, in

the second case we vary the bulk modulus only. Figure 2 shows the two models used for the numerical tests. In these tests we use a single source placed at the origin, and evenly spaced receivers at the surface with receiver interval of 50 m. In both examples we use a Ricker wavelet with central frequency of $10Hz$. We use 8th-order accuracy staggered-grid operators and 5 points per wavelength for the minimum one which is $30m$. Moreover, as explained in Almomin (2013), to damp spurious reflections from the model boundary, we add the absorbing boundaries described by Israeli and Orszag (1981) in the recursive equations used for propagating the pressure field. Figure 3 displays the data obtained from the two numerical simulations. In the top panel, we observe the data generated by propagating the source in the varying density model; in the bottom panel, we show the data for the varying bulk modulus. In both data, we clearly distinguish the direct arrival and two primary reflections. It is notable that for the varying density case the impedance contrasts are negative; and also that phase rotations are present as we increase receiver offset. When we vary the compressibility only, using the same velocity contrasts of the varying density case, we obtain positive impedance contrasts and do not observe any phase rotations for increasing offset.

Since we also have at our disposal the adjoint operator of the propagator (Equation 12), we run an adjoint experiment with the varying density model and the data of Figure 3a. As expected, for a simple geometry, where a single source and multiple receivers are deployed, the output of the adjoint operator is a scaled version of the source wavelet injected for recording the data of Figure 3a (Figure 4). We also plot the forward and adjoint pressure fields for three different times in Figure 5. The forward pressure field is propagating in the positive time direction; whereas, the adjoint field is collapsing at the receiver locations as we increase the propagation time. In both cases we have reflections coming from the two contrasts. In the adjoint wavefield, we notice the propagating receivers have a strong directed arrival because this wave is in phase for all the receivers.

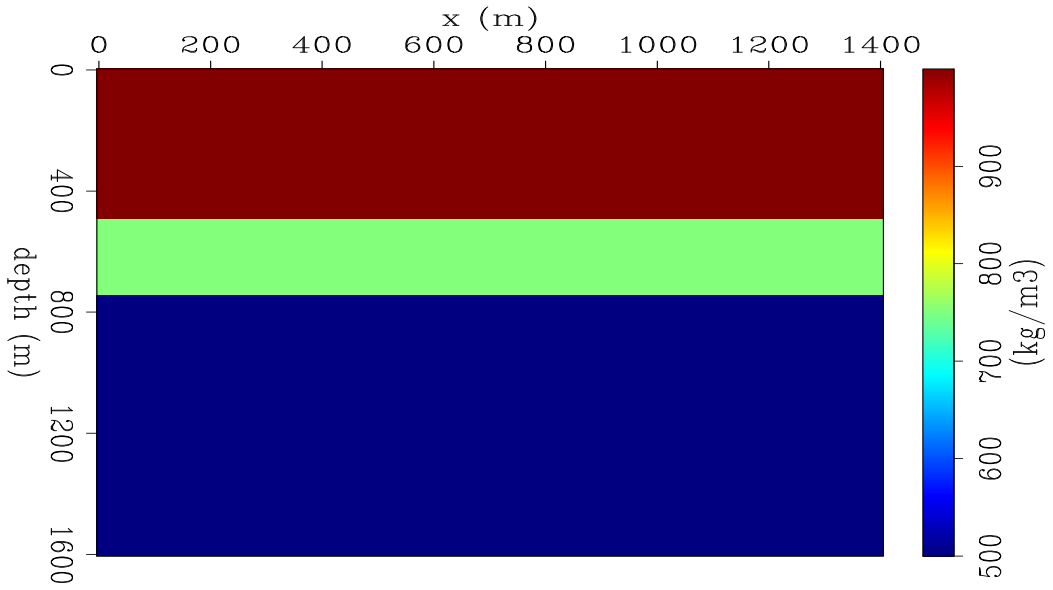
BORN AND RTM OPERATORS

To write the chain of linear operators for the Born approximation and its adjoint operator (also known as RTM operator), we can use most of the operators employed in forward and adjoint propagation. However, we need to take care of the scattering condition and its adjoint shown in Equations 2, 3, and 4. The Born operator can be written as follows:

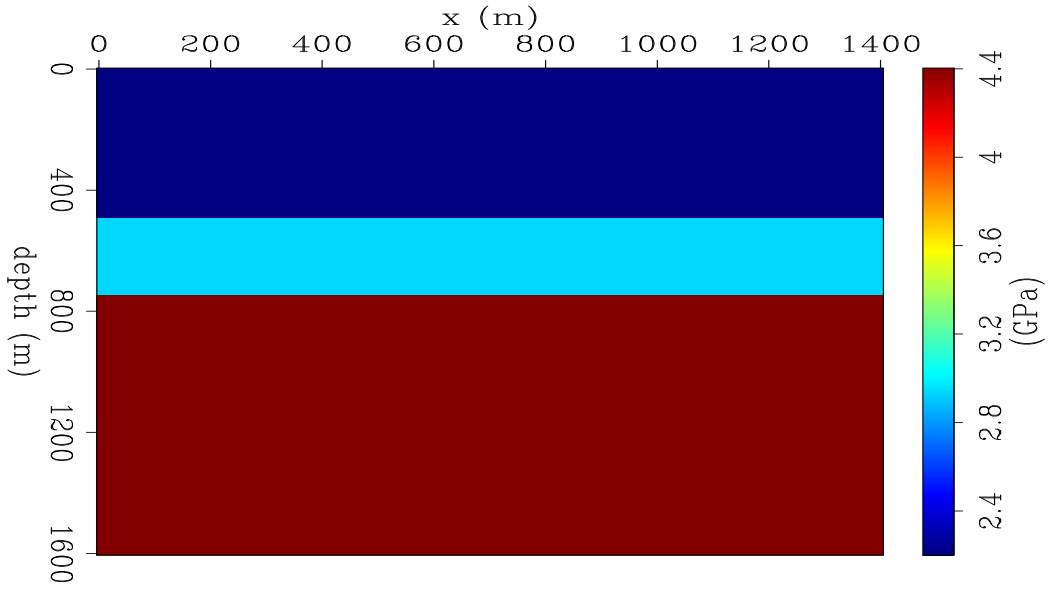
$$\Delta\mathbf{d} = \mathbf{B}\Delta\mathbf{m} = \mathbf{B}\begin{bmatrix} \Delta\mathbf{K} \\ \Delta\boldsymbol{\rho} \end{bmatrix}, \tag{19}$$

where we obtain perturbed data $\Delta\mathbf{d}$ from model perturbations $\Delta\mathbf{m}$ in the background compressibility and density. In this equation $\Delta\mathbf{K}$ and $\Delta\boldsymbol{\rho}$ represent two vectors containing the bulk modulus and density perturbations for all the points in the model, respectively. The operator $\mathbf{B}$ is expanded into the following series of linear operators:

$$\Delta\mathbf{d} = \mathbf{R}^*\mathbf{L}_g^*\mathbf{F}\mathbf{R}\begin{bmatrix}\ddot{\mathbf{P}}_\mathbf{0} & \tilde{\mathbf{P}}_\mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{M_K} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\boldsymbol{\rho}}\end{bmatrix}\begin{bmatrix}\Delta\mathbf{K} \\ \Delta\boldsymbol{\rho}\end{bmatrix}, \tag{20}$$

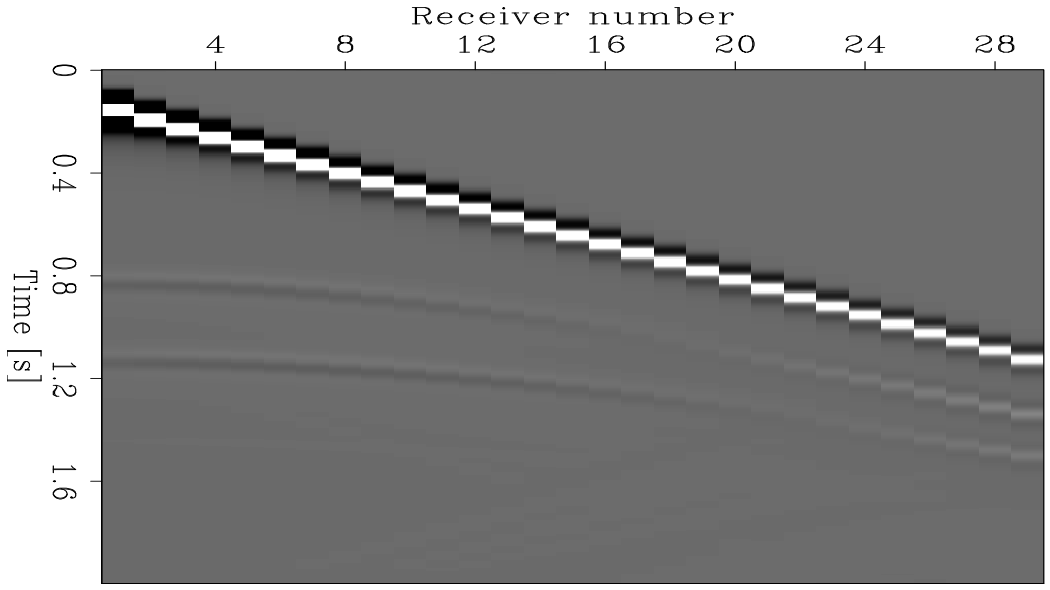


(a)

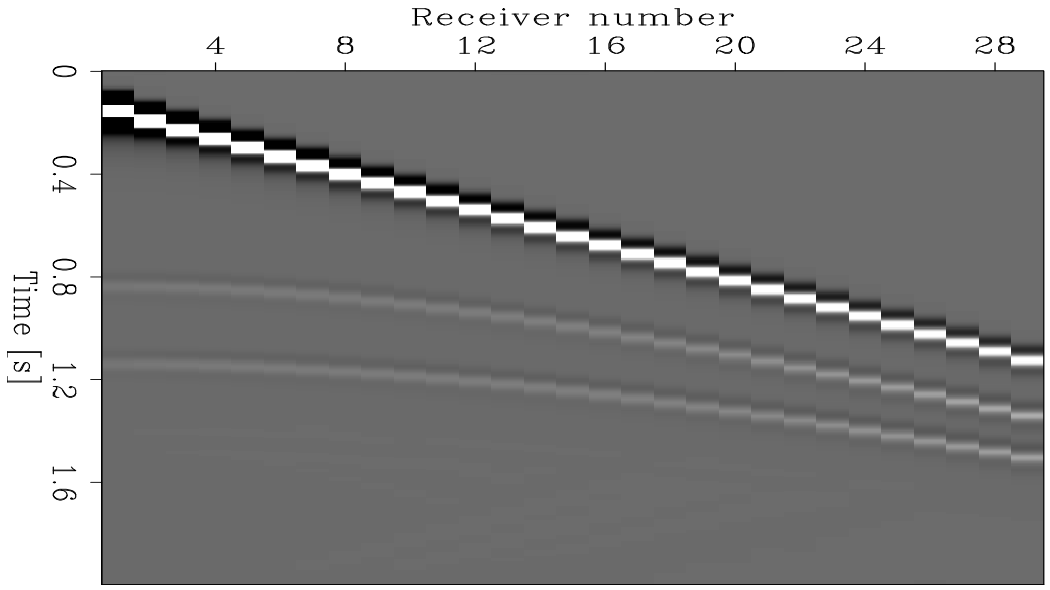


(b)

Figure 2: Density and bulk modulus models used in two different numerical experiments having the same velocity contrasts. (a) Density model used in the first experiment where the bulk modulus is kept constant and equal to $2.2GPa$. (b) Bulk modulus model used in the second experiment where the density is $1000kg/m^3$ [**ER**]



(a)



(b)

Figure 3: Data resulting from the numerical tests. (a) Data from the varying density model of Figure2a. In this case the impedance contrasts are negative and we also observe phase rotations for increasing offset. (b) Data from the varying density model of Figure 2b. In this model the impedance contrasts are positive. Moreover, we do not observe phase rotations. [**ER**]

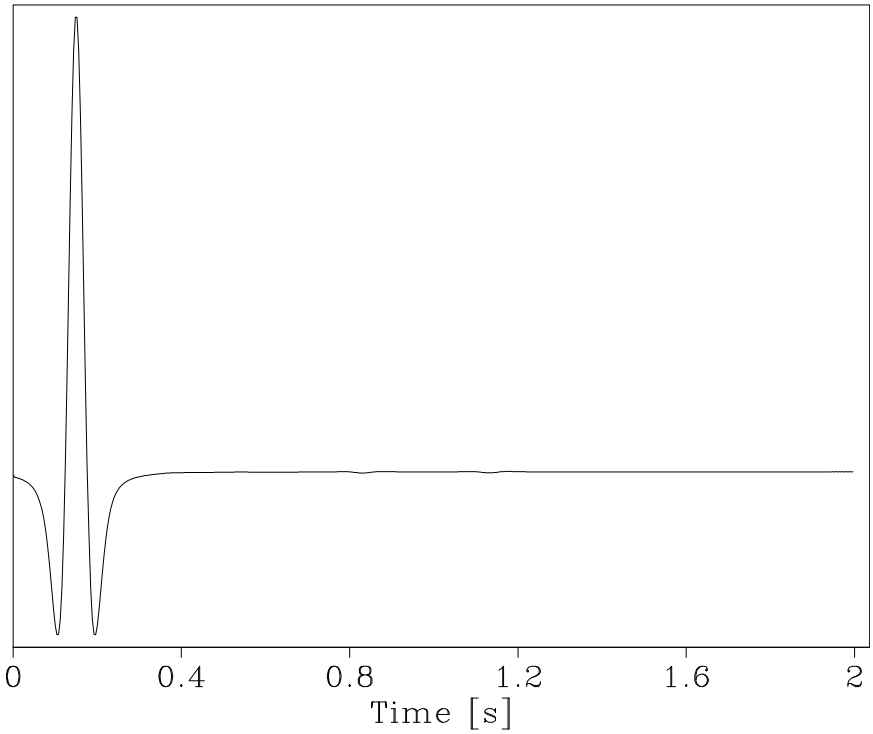


Figure 4: Adjoint source wavelet obtained using the adjoint propagator on the modeled data of Figure 3a with the varying density model of Figure 2a. [**ER**]

where we notice part of the forward chain of operators used in Equation 7, which takes care of propagating the scattered pressure coming from the other operator on the right of it. The matrix of operators containing $\mathbf{M_K}$ and $\mathbf{M_\rho}$ is a scaling-spreading operator that scales the two model perturbations with the inverse of the square of the background property (i.e., $1/K_0(\mathbf{r})^2$ and $-1/\rho_0(\mathbf{r})^2$ in Equation 2) and spreads these scaled perturbations to all time steps. The two stacked operators $\ddot{\mathbf{P}}_\mathbf{0}$ and $\tilde{\mathbf{P}}_\mathbf{0}$ contain both the forward propagated pressure field in the background properties $\mathbf{K}_0$ and $\boldsymbol{\rho}_0$ obtained as follows:

$$\mathbf{p}_0 = \mathbf{R}^*\mathbf{F}\mathbf{L}_s\mathbf{R}\mathbf{w} \tag{21}$$

downsampling the propagated source wavefield by applying $\mathbf{R}^*$ to save computational time when calculating the scattered pressure. Note that after scattering, we need to resample it with $\mathbf{R}$. To maintain the same amplitudes, we need to have $\mathbf{R}^*\mathbf{R} \approx \mathbf{I}$. The scattering operator $\ddot{\mathbf{P}}_\mathbf{0}$, which accounts for perturbations in the compressibility, is obtained as:

$$diag(\ddot{\mathbf{P}}_\mathbf{0}) = \mathbf{D_2}\mathbf{p}_0, \tag{22}$$

where $\mathbf{D_2}$ outputs the second-order time derivative of the forward pressure field. The other scattering operator $\tilde{\mathbf{P}}_\mathbf{0}$, which accounts for density perturbations, is more complex than the $\ddot{\mathbf{P}}_\mathbf{0}$ operator, and compose a chain of linear operators written as

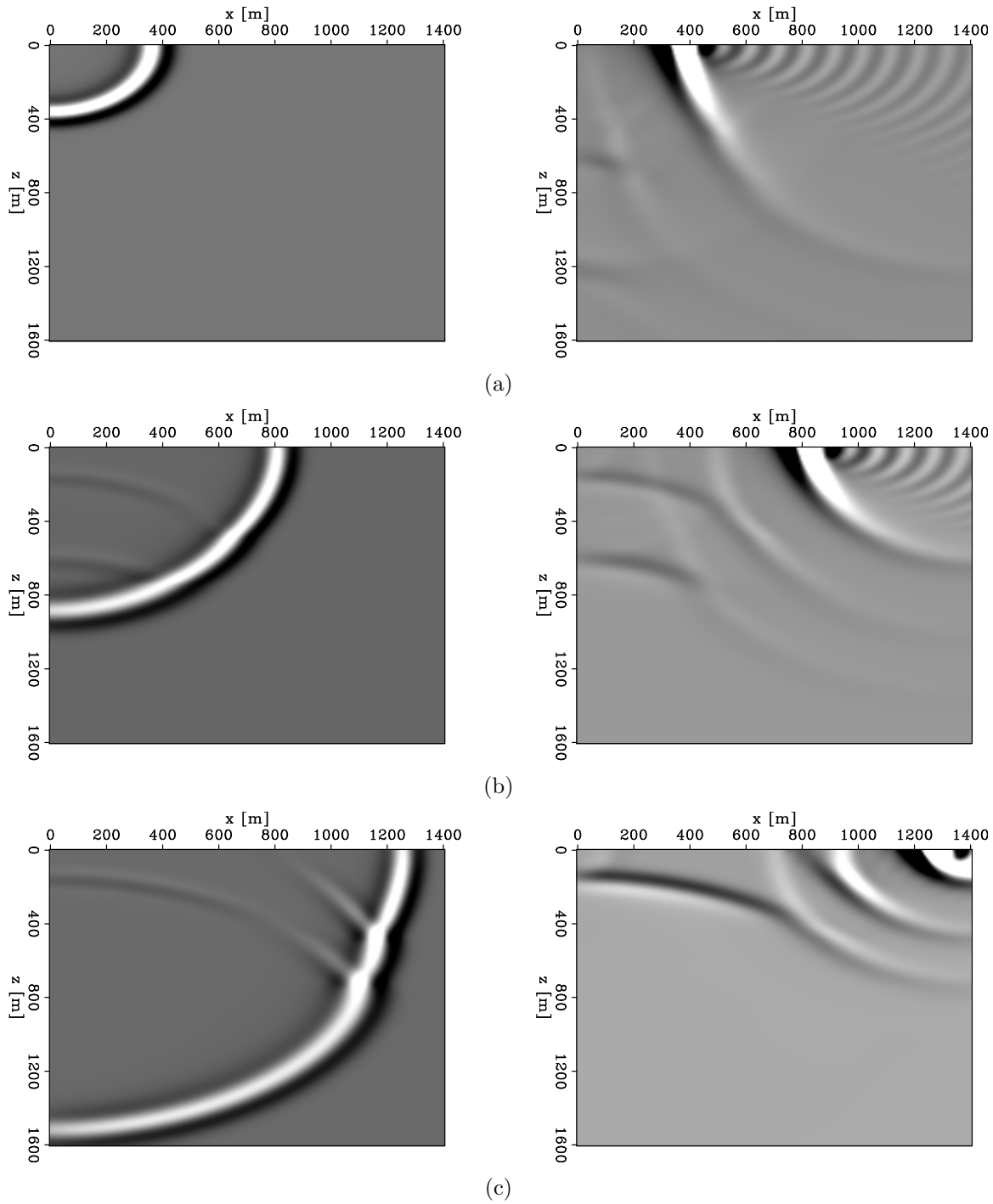


Figure 5: Forward (left panels) and adjoint (right panels) pressure fields at (a) $0.4s$, (b) $0.7s$, and (c) $1.0s$. Since we back-solve the system of equations for the propagation, the adjoint wavefield is running backward in time. [**ER**]

follows:

$$\tilde{\mathbf{P}}_{\mathbf{0}} = \begin{bmatrix} \mathbf{D}_{\mathbf{z}}^{-} & \mathbf{D}_{\mathbf{x}}^{-} \end{bmatrix} \begin{bmatrix} diag(\mathbf{D}_{\mathbf{z}}^{+} p_0) & \mathbf{0} \\ \mathbf{0} & diag(\mathbf{D}_{\mathbf{x}}^{+} p_0) \end{bmatrix} \begin{bmatrix} \mathbf{G}_{\mathbf{z}}^{+} \\ \mathbf{G}_{\mathbf{x}}^{+} \end{bmatrix}, \tag{23}$$

where the first stacked operators $\mathbf{G}$ stagger the density perturbations into the two staggered grids shown in Figure 1. Then, we multiply the staggered density perturbations with the gradient of the pressure field and afterward compute the divergence with the two $\mathbf{D}^{-}$ operators in the respective direction. The RTM operator is simply the adjoint of the chain of operator of Equation 20 that is:

$$\begin{bmatrix} \Delta\tilde{\mathbf{K}} \\ \Delta\tilde{\boldsymbol{\rho}} \end{bmatrix} = \begin{bmatrix} \mathbf{M}_{\mathbf{K}}^{*} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{\boldsymbol{\rho}}^{*} \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{P}}_{\mathbf{0}} \\ \tilde{\mathbf{P}}_{\mathbf{0}}^{*} \end{bmatrix} \mathbf{R}^{*}\mathbf{F}^{*}\mathbf{L}_{g}\mathbf{R}\Delta\mathbf{d}, \tag{24}$$

where $\ddot{\mathbf{P}}_{\mathbf{0}}$ is self-adjoint, because it is a real diagonal operator, and the chain $\mathbf{R}^{*}\mathbf{F}^{*}\mathbf{L}_{g}\mathbf{R}$ is back-propagating data perturbations in the background model. The matrix of operators at the end of the chain is scaling the result of the output from the chain on the right of it, and taking the sum of it for each model point, which corresponds to zero-lag cross-correlation. We then clearly see that the chain $\mathbf{M}_{\mathbf{K}}^{*}\ddot{\mathbf{P}}_{\mathbf{0}}\mathbf{R}^{*}\mathbf{F}^{*}\mathbf{L}_{g}\mathbf{R}$ is computing Equation 3.

We have to verify that the chain $\mathbf{M}_{\boldsymbol{\rho}}^{*}\tilde{\mathbf{P}}_{\mathbf{0}}^{*}\mathbf{R}^{*}\mathbf{F}^{*}\mathbf{L}_{g}\mathbf{R}$ is actually equal to Equation 4. To this end we take the adjoint of equation 23 that is equal to:

$$\tilde{\mathbf{P}}_{\mathbf{0}}^{*} = \begin{bmatrix} \mathbf{G}_{\mathbf{z}}^{+*} & \mathbf{G}_{\mathbf{x}}^{+*} \end{bmatrix} \begin{bmatrix} diag(\mathbf{D}_{\mathbf{z}}^{+} p_0)^{*} & \mathbf{0} \\ \mathbf{0} & diag(\mathbf{D}_{\mathbf{x}}^{+} p_0)^{*} \end{bmatrix} \begin{bmatrix} \mathbf{D}_{\mathbf{z}}^{-*} \\ \mathbf{D}_{\mathbf{x}}^{-*} \end{bmatrix}, \tag{25}$$

and using the property of staggered-grid operators (17 and 18), it can be rewritten as:

$$\tilde{\mathbf{P}}_{\mathbf{0}}^{*} = -\begin{bmatrix} \mathbf{G}_{\mathbf{z}}^{+*} & \mathbf{G}_{\mathbf{x}}^{+*} \end{bmatrix} \begin{bmatrix} diag(\mathbf{D}_{\mathbf{z}}^{+} p_0)^{*} & \mathbf{0} \\ \mathbf{0} & diag(\mathbf{D}_{\mathbf{x}}^{+} p_0)^{*} \end{bmatrix} \begin{bmatrix} \mathbf{D}_{\mathbf{z}}^{+} \\ \mathbf{D}_{\mathbf{x}}^{+} \end{bmatrix}, \tag{26}$$

where we see that we are calculating the gradient of the back-propagated data perturbations; and then multiplying it with the gradient of forward pressure field with both gradients computed on the staggered-grids. We apply the adjoint of the stacked staggering operators $\mathbf{G}$ that shift the inputs from the staggered grids to the regular grid and compute the sum of the two outputs. All together we now see that the chain $\mathbf{M}_{\boldsymbol{\rho}}^{*}\tilde{\mathbf{P}}_{\mathbf{0}}^{*}\mathbf{R}^{*}\mathbf{F}^{*}\mathbf{L}_{g}\mathbf{R}$ is computing Equation 4.

Numerical tests of linearized operator and its adjoint

To test the Born operator, we run two simple experiments with same background medium's properties that we used in the previous example, where we perturb only the compressibility or the density, and we use the same acquisition geometry used in the previous propagations. We use constant background values equal to $2.2GPa$ and

$1000kg/m^3$ and two single perturbations of $0.2GPa$ and $100kg/m^3$ for bulk modulus and density respectively. With these experiments we want to verify the following relations:

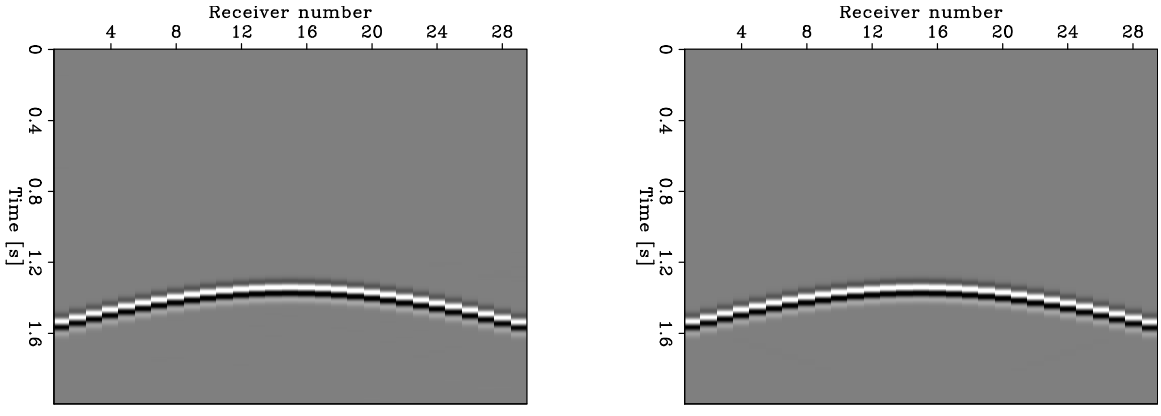
$$f(\mathbf{K}_0 + \Delta\mathbf{K}, \boldsymbol{\rho}_0) - f(\mathbf{K}_0, \boldsymbol{\rho}_0) \approx \mathbf{B}\Delta\mathbf{K}, \tag{27}$$

$$f(\mathbf{K}_0, \boldsymbol{\rho}_0 + \Delta\boldsymbol{\rho}) - f(\mathbf{K}_0, \boldsymbol{\rho}_0) \approx \mathbf{B}\Delta\boldsymbol{\rho}, \tag{28}$$

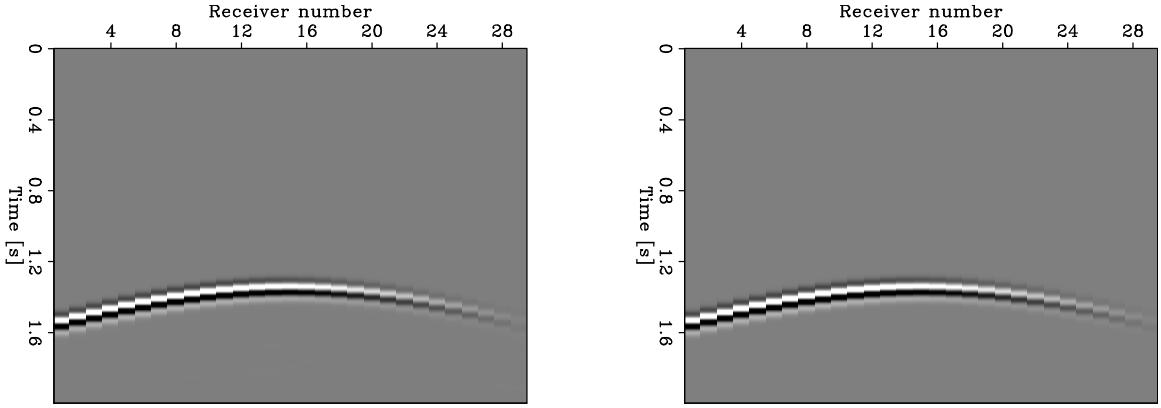
where $f(\mathbf{K}, \boldsymbol{\rho})$ represents the data generated by the propagation through the medium's properties $\mathbf{K}$ and $\boldsymbol{\rho}$. Here, we are verifying that the linearized data perturbations are actually modeled by the implemented Born operator. To verify this observation we need to have small model perturbations. Figure 6a shows the comparison of the data perturbation when we add a compressibility change in the model; while Figure 6b displays the comparison; but when we perturbe only the density model. The same clipping value is used for all the panels, and we can clearly see that for a single model perturbation the linearized scattered data matches the non-linear data perturbations.

To understand the reason why we observe different amplitude variations for increasing offsets only when perturb the density, we need to analyze the scattering pressure from the single diffraction point. Figure 7 shows the comparison of the scattered pressure and the theoretical radiation pattern (Aki and Richards, 2002) in case we perturb the bulk modulus. We observe that we have an isotropic energy scattering, which results in an isotropic back-scattered energy at the receivers (Figure 6a). Otherwise, if we look at the scattered energy from a density perturbation (Figure 8), we see an anisotropic scattered pressure. This observation is also confirmed by comparing this scattered pressure with the theoretical radiation pattern (Figure 8b). In fact, in the experiment of Figure 6b, because we are hitting the scatter not vertically, the scattered energy recorded at the receivers is affected by this anisotropic radiation pattern.

The last performed test is to apply the RTM operator to data perturbations. We apply this operator to the perturbed data obtained from a density perturbation (Figure 6b). Figure 9 shows the result of this application. We see that even if we have data generated only by a density perturbation some of the energy is leaking into the bulk modulus image. This effect is well known as parameter cross-talk (Operto et al., 2013). It is also important to notice that the energy of the density image is higher than in the compressibility image. One possible explanation could be because data contains both amplitude and phase information entangle together, and during FWI we are trying to simultaneously invert the two information that depend on a combination of physical parameters. In fact, it is known FWI gradient does not provide the correct amplitude of the model perturbations, first because the gradient does not have the correct units of the model perturbation and second because the gradient associated with one parameter can be affected by crosstalk from the other parameters. Hessian-matrix based inverse methods could be possible solutions to solve this problem (Innanen, 2014).

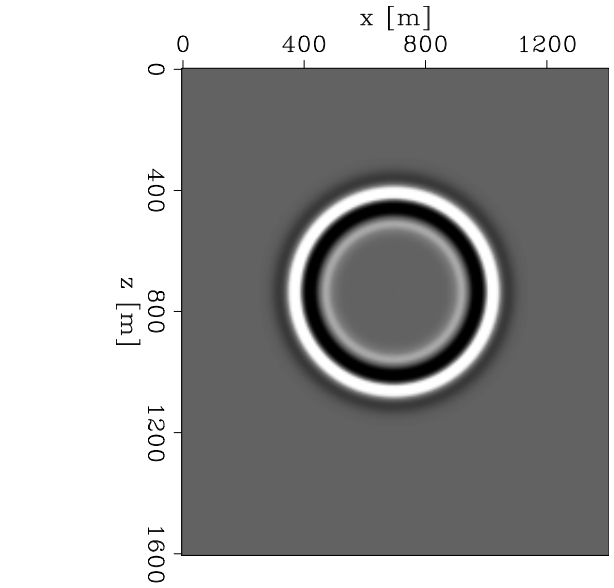


(a)

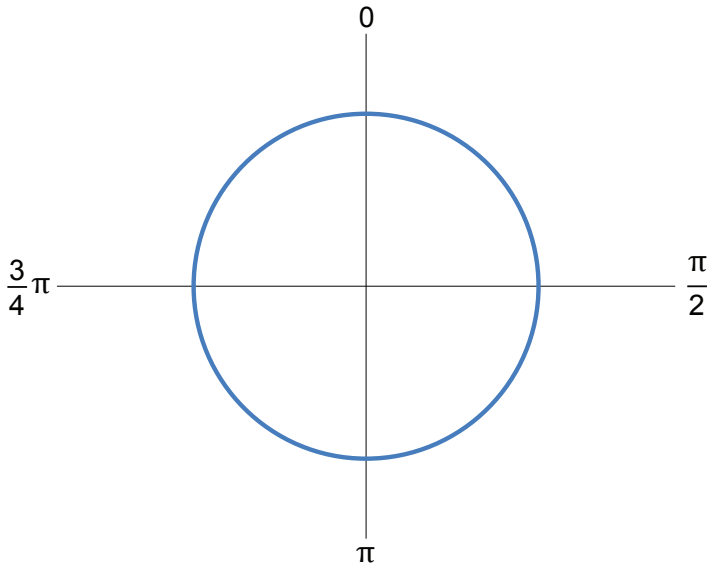


(b)

Figure 6: Comparison of the correct data perturbation given by the difference of the data obtained from a perturbed model and an unperturbed model (left panels), and the linearized data perturbation generated by the Born model operator (right panels). (a) Data comparison when we perturbe only the compressibility of the model (equation 27). (b) Same comparison but perturbing only the density model (equation 28). [**ER**]

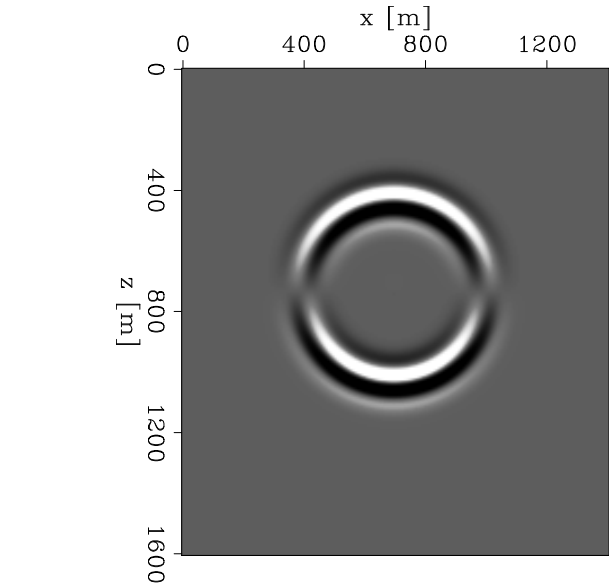


(a)

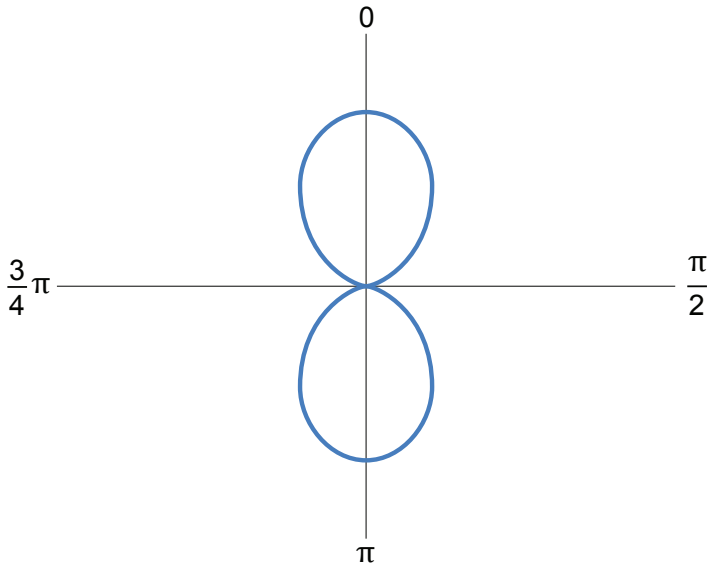


(b)

Figure 7: (a) Radiated energy from a single bulk modulus perturbation point. [**ER**](b) Theoretical radiation pattern of a vertical incident wave on a bulk modulus perturbation (Aki and Richards, 2002). [**NR**]



(a)



(b)

Figure 8: (a) Radiated energy from a single density perturbation point. [**ER**](b) Theoretical radiation pattern of a vertical incident wave on a density perturbation (Aki and Richards, 2002). [**NR**]

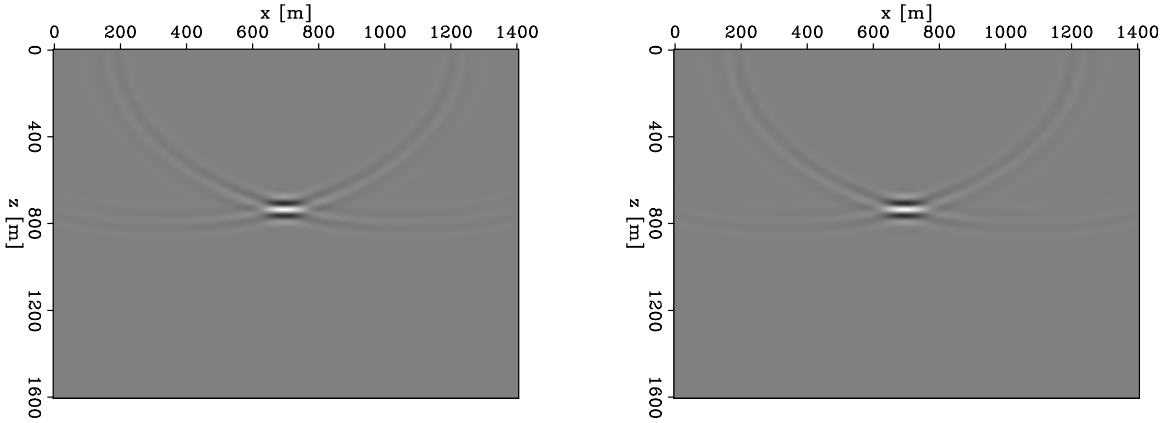


Figure 9: Images obtained applying the non-constant density RTM operator. The images look the same but the scales are different. The density image (right panel) is stronger than the bulk modulus image (left panel) by six orders of magnitude. [**ER**]

CONCLUSIONS

We implemented and discussed the two-way wave-equation operators for non-constant density isotropic media. We explain that non-linear and linearized modeling can be obtained as a chain of simple linear operators. Synthetic tests illustrate how these operators work and also that scattered pressure field is consistent with the theoretical radiation patterns. These consistent linearized wave-equation operators are important when applying a gradient-based inverse scheme in the context of a multi-parameter FWI.

ACKNOWLEDGEMENT

We thank Joseph Stefani and Mark Meadows for the useful discussions and comments on the non-constant density propagation and inversion.

APPENDIX

In this Appendix we derive most of the results used in the paper.

Derive the acoustic wave equation from elastodynamic equations

First, it is important to derive the acoustic wave equation in non-constant density media such that we fully understand the approximations made when considering

acoustic propagation. From Newton's second law of motion we can write:

$$\begin{aligned}
\rho(\mathbf{r})\frac{\partial^2 u_x(\mathbf{r},t)}{\partial t^2} &= \frac{\partial \tau_{xx}(\mathbf{r},t)}{\partial x} + \frac{\partial \tau_{xy}(\mathbf{r},t)}{\partial y} + \frac{\partial \tau_{xz}(\mathbf{r},t)}{\partial z} + f_x(\mathbf{r},t),\\
\rho(\mathbf{r})\frac{\partial^2 u_y(\mathbf{r},t)}{\partial t^2} &= \frac{\partial \tau_{yx}(\mathbf{r},t)}{\partial x} + \frac{\partial \tau_{yy}(\mathbf{r},t)}{\partial y} + \frac{\partial \tau_{yz}(\mathbf{r},t)}{\partial z} + f_y(\mathbf{r},t),\\
\rho(\mathbf{r})\frac{\partial^2 u_z(\mathbf{r},t)}{\partial t^2} &= \frac{\partial \tau_{zx}(\mathbf{r},t)}{\partial x} + \frac{\partial \tau_{zy}(\mathbf{r},t)}{\partial y} + \frac{\partial \tau_{zz}(\mathbf{r},t)}{\partial z} + f_z(\mathbf{r},t),
\end{aligned} \tag{29}$$

where $\rho(\mathbf{r})$ is the density of the medium, $u_i(\mathbf{r},t)$ is the particle displacement in the i direction, $\tau_{ii}(\mathbf{r},t)$ and τ_{ij} are the normal and shear stresses applied to the orthogonal plane to i direction, respectively; and $f_i(\mathbf{r},t)$ represents the external forces along the i direction. Because we are considering only acoustic isotropic media the shear stresses can be neglected, and the normal stresses are given by the following constitutive relations:

$$\begin{aligned}
\tau_{xx}(\mathbf{r},t) &= \lambda(\mathbf{r})\left(\frac{\partial u_x(\mathbf{r},t)}{\partial x} + \frac{\partial u_y(\mathbf{r},t)}{\partial y} + \frac{\partial u_z(\mathbf{r},t)}{\partial z}\right) + 2\mu(\mathbf{r})\frac{\partial u_x(\mathbf{r},t)}{\partial x},\\
\tau_{yy}(\mathbf{r},t) &= \lambda(\mathbf{r})\left(\frac{\partial u_x(\mathbf{r},t)}{\partial x} + \frac{\partial u_y(\mathbf{r},t)}{\partial y} + \frac{\partial u_z(\mathbf{r},t)}{\partial z}\right) + 2\mu(\mathbf{r})\frac{\partial u_y(\mathbf{r},t)}{\partial y},\\
\tau_{zz}(\mathbf{r},t) &= \lambda(\mathbf{r})\left(\frac{\partial u_x(\mathbf{r},t)}{\partial x} + \frac{\partial u_y(\mathbf{r},t)}{\partial y} + \frac{\partial u_z(\mathbf{r},t)}{\partial z}\right) + 2\mu(\mathbf{r})\frac{\partial u_z(\mathbf{r},t)}{\partial z},
\end{aligned} \tag{30}$$

where $\lambda(\mathbf{r})$ is the Lamé's first parameter, and $\mu(\mathbf{r})$ is the Lamé's second parameter also known as shear modulus of the medium.

If we consider isotropic stress fields the pressure field is equal to each normal stress, and:

$$p(\mathbf{r},t) = \tau_{xx}(\mathbf{r},t) = \tau_{yy}(\mathbf{r},t) = \tau_{zz}(\mathbf{r},t). \tag{31}$$

Therefore, the scalar pressure field can be expressed as:

$$p(\mathbf{r},t) = \frac{1}{3}(\tau_{xx}(\mathbf{r},t) + \tau_{yy}(\mathbf{r},t) + \tau_{zz}(\mathbf{r},t)) + p_s(\mathbf{r},t), \tag{32}$$

where $p_s(\mathbf{r},t)$ is an external pressure source; therefore, using Equation 30, we have:

$$p(\mathbf{r},t) = \left(\lambda(\mathbf{r}) + \frac{2}{3}\mu(\mathbf{r})\right)\nabla\cdot\mathbf{u}(\mathbf{r},t) + p_s(\mathbf{r},t) = K(\mathbf{r})\nabla\cdot\mathbf{u}(\mathbf{r},t) + p_s(\mathbf{r},t), \tag{33}$$

where we have introduced a new elastic property of the medium $K(\mathbf{r})$ called bulk modulus or compressibility, which measures the medium's resistance to uniform compression. Because $K(\mathbf{r})$ is time independent, taking the second-order time derivative of equation 33, and substituting Newton's second law (i.e., Equation 29) we can write:

$$\frac{\partial^2 p(\mathbf{r},t)}{\partial t^2} = K(\mathbf{r})\nabla\cdot\frac{1}{\rho(\mathbf{r})}\nabla p(\mathbf{r},t) + \frac{\partial^2 p_s(\mathbf{r},t)}{\partial t^2} + \nabla\cdot\frac{\mathbf{f}(\mathbf{r},t)}{\rho(\mathbf{r})}. \tag{34}$$

Reordering Equation 34 we can express it as the common wave equation:

$$\left[\frac{1}{K(\mathbf{r})}\frac{\partial^2}{\partial t^2} - \nabla\cdot\frac{1}{\rho(\mathbf{r})}\nabla\right] p(\mathbf{r},t) = s(\mathbf{r},t), \tag{35}$$

where $s(\mathbf{r},t)$ represents both the volume injection and external force per unit volume (i.e., last two right-side terms in Equation 34, respectively). From this last relation, we can easily derive the well-known acoustic isotropic constant-density wave equation.

Born approximation and adjoint conditions for non-constant density media

Before deriving the Born approximation and the adjoint conditions for non-constant density, we first introduce the Green's function of Equation 35:

$$\left[\frac{1}{K(\mathbf{r})}\frac{\partial^2}{\partial t^2} - \nabla\cdot\frac{1}{\rho(\mathbf{r})}\nabla\right] g(\mathbf{r},t,\mathbf{r}',t';K,\rho) = \delta(\mathbf{r}-\mathbf{r}')\delta(t-t'). \tag{36}$$

Considering time invariant medium's properties the solution of the acoustic non-constant density wave equation can be expressed as:

$$p(\mathbf{r},t) = \int g(\mathbf{r},t,\mathbf{r}';K,\rho) * s(\mathbf{r}',t)d\mathbf{r}', \tag{37}$$

where $*$ denotes time convolution. During active seismic experiments sources can be considered as points in space with a given source signature $w(t)$ (i.e., $s(\mathbf{r},t) = w(t)\delta(\mathbf{r}-\mathbf{r}_s)$), therefore, the previous equation becomes:

$$p(\mathbf{r},t) = g(\mathbf{r},t,\mathbf{r}_s;K,\rho) * w(t). \tag{38}$$

Finally, because seismic receivers can be considered as recording points, the data are given by:

$$d(\mathbf{r}_g,t) = p(\mathbf{r},t)\delta(\mathbf{r}-\mathbf{r}_g). \tag{39}$$

To derive the Born approximation for non-constant density isotropic acoustic media we perturbe the background medium's properties $K_0(\mathbf{r})$ and $\rho_0(\mathbf{r})$:

$$K(\mathbf{r}) = K_0(\mathbf{r}) + \delta K(\mathbf{r}), \tag{40}$$

$$\rho(\mathbf{r}) = \rho_0(\mathbf{r}) + \delta\rho(\mathbf{r}), \tag{41}$$

where $\delta K(\mathbf{r})$ and $\delta\rho(\mathbf{r})$ are the bulk modulus and density perturbations respectively. Considering small perturbations respect to the background properties, we can assume the pressure field propagating in the perturbed medium is given by:

$$\left[\frac{1}{K_0(\mathbf{r})+\delta K(\mathbf{r})}\frac{\partial^2}{\partial t^2} - \nabla\cdot\frac{1}{\rho_0(\mathbf{r})+\delta\rho(\mathbf{r})}\nabla\right] (p_0(\mathbf{r},t)+\delta p(\mathbf{r},t)) = s(\mathbf{r},t). \tag{42}$$

After expanding all terms in the left-side of the previous equation and dropping high-order terms, the perturbed pressure field $\delta p(\mathbf{r},t)$ is given by:

$$\left[\frac{1}{K_0(\mathbf{r})}\frac{\partial^2}{\partial t^2} - \nabla\cdot\frac{1}{\rho_0(\mathbf{r})}\nabla\right]\delta p(\mathbf{r},t) = \frac{\delta K(\mathbf{r})}{K_0(\mathbf{r})^2}\frac{\partial^2 p_0(\mathbf{r},t)}{\partial t^2} - \nabla\cdot\frac{\delta\rho(\mathbf{r})}{\rho_0(\mathbf{r})^2}\nabla p_0(\mathbf{r},t), \quad (43)$$

where we have used:

$$\frac{1}{m_0(\mathbf{r})+\delta m(\mathbf{r})} = \frac{1}{m_0(\mathbf{r})} - \frac{\delta m(\mathbf{r})}{m_0(\mathbf{r})^2}, \quad (44)$$

and $p_0(\mathbf{r},t)$ corresponds to the pressure field propagated in the background medium:

$$\left[\frac{1}{K_0(\mathbf{r})}\frac{\partial^2}{\partial t^2} - \nabla\cdot\frac{1}{\rho_0(\mathbf{r})}\nabla\right]p_0(\mathbf{r},t) = s(\mathbf{r},t). \quad (45)$$

In Equation 43 the right-side two terms describe the scattered pressure caused by a compressibility and density perturbations, respectively; and they can be considered as secondary sources generated by the energy coming from the primary source $s(\mathbf{r},t)$. The Born modeled data are then obtained as:

$$\delta d(\mathbf{r}_g,t) = \delta p(\mathbf{r},t)\delta(\mathbf{r}-\mathbf{r}_g). \quad (46)$$

To simplify the discussion we continue it in the frequency domain and discretize both in time and space all the variables so far introduced. Given these two observations we can write Equation 38 as:

$$P(\mathbf{x}_i,\omega_k) = G(\mathbf{x}_i,\omega_k,\mathbf{x}_s;K,\rho)W(\omega_k) \quad i=1,\dots,M \quad k=1,\dots,N_t, \quad (47)$$

which provides the pressure field for all M points in space and N_t frequencies. Assuming Equation 39, seismic data are found as:

$$D(\mathbf{x}_g,\omega_k) = P(\mathbf{x}_i=\mathbf{x}_g,\omega_k) \quad g=1,\dots,N_g, \quad (48)$$

where N_g denotes the number of active receivers for a given source.

The solution of Equation 43 can be found by multiplying the unperturbed Green's function with the two secondary source terms for all the points in space:

$$\begin{gathered}\Delta P(\mathbf{x}_i,\omega_k) = \textstyle\sum_{j=1}^{M} G(\mathbf{x}_i,\omega_k,\mathbf{x}_j;K_0,\rho_0)\\ \left(\frac{\Delta K(\mathbf{x}_j)}{-K_0(\mathbf{x}_j)^2}\omega^2 - \nabla\cdot\frac{\Delta\rho(\mathbf{x}_j)}{\rho_0(\mathbf{x}_j)^2}\nabla\right)G(\mathbf{x}_j,\omega_k,\mathbf{x}_s;K_0,\rho_0)W(\omega_k)\end{gathered} \quad (49)$$

Because we want to find two separate linear kernels for compressibility and density perturbations we write the total perturbed pressure field as:

$$\Delta P(\mathbf{x}_i,\omega_k) = \Delta P_K(\mathbf{x}_i,\omega_k) + \Delta P_\rho(\mathbf{x}_i,\omega_k), \quad (50)$$

where

$$\Delta P_K(\mathbf{x}_i,\omega_k) = -\sum_{j=1}^{M} G(\mathbf{x}_i,\omega_k,\mathbf{x}_j;K_0,\rho_0)\omega^2 G(\mathbf{x}_j,\omega_k,\mathbf{x}_s;K_0,\rho_0)W(\omega_k)\frac{\Delta K(\mathbf{x}_j)}{K_0(\mathbf{x}_j)^2}, \quad (51)$$

$$\Delta P_\rho(\mathbf{x}_i,\omega_k) = -\sum_{j=1}^{M} G(\mathbf{x}_i,\omega_k,\mathbf{x}_j;K_0,\rho_0)\nabla\cdot\frac{\Delta\rho(\mathbf{x}_j)}{\rho_0(\mathbf{x}_j)^2}\nabla G(\mathbf{x}_j,\omega_k,\mathbf{x}_s;K_0,\rho_0)W(\omega_k). \quad (52)$$

In Equation 52, the perturbed pressure field depends on the divergence of the product of the density perturbation and gradient of the background pressure field. This operation is linear; however, it contains multiple terms that can be eliminated under certain assumptions. We can rewrite this equation using the following relation:

$$\psi\nabla\cdot(\mathbf{v}) = -(\nabla\psi)\cdot\mathbf{v} + \nabla\cdot(\psi\mathbf{v}), \quad (53)$$

as:

$$\begin{aligned}&\Delta P_\rho(\mathbf{x}_i,\omega_k,\mathbf{x}_s;K_0,\rho_0) = \\ &\textstyle\sum_{j=1}^{M}\frac{\Delta\rho(\mathbf{x}_j)}{\rho_0(\mathbf{x}_j)^2}\nabla G(\mathbf{x}_i,\omega_k,\mathbf{x}_j;K_0,\rho_0)\cdot\nabla G(\mathbf{x}_j,\omega_k,\mathbf{x}_s;K_0,\rho_0)W(\omega_k) \\ &-\textstyle\sum_{j=1}^{M}\nabla\cdot\frac{\Delta\rho(\mathbf{x}_j)}{\rho_0(\mathbf{x}_j)^2}G(\mathbf{x}_i,\omega_k,\mathbf{x}_j;K_0,\rho_0)\nabla G(\mathbf{x}_j,\omega_k,\mathbf{x}_s;K_0,\rho_0)W(\omega_k).\end{aligned} \quad (54)$$

Using the divergence theorem, the second right-side term of Equation 54 can be written as a surface integral along the medium's boundaries; assuming homogeneous conditions or only internal density perturbations this term vanishes. Therefore, knowing that we are recording data only at the surface (Equation 46), we can write two linear kernels respect to the medium perturbations as:

$$\Delta D_K(\mathbf{x}_g,\omega_k) = \sum_{j=1}^{M}\left[\frac{\partial\Delta P_K(\mathbf{x}_g,\omega_k)}{\partial\Delta K(\mathbf{x}_j)}\right]\Delta K(\mathbf{x}_j), \quad (55)$$

$$\Delta D_\rho(\mathbf{x}_g,\omega_k) = \sum_{j=1}^{M}\left[\frac{\partial\Delta P_\rho(\mathbf{x}_g,\omega_k)}{\partial\Delta\rho(\mathbf{x}_j)}\right]\Delta\rho(\mathbf{x}_j), \quad (56)$$

where the total data perturbation of Equation 46 is given by:

$$\Delta D(\mathbf{x}_g,\omega_k) = \Delta D_K(\mathbf{x}_g,\omega_k) + \Delta D_\rho(\mathbf{x}_g,\omega_k). \quad (57)$$

The Born operator is applied to model perturbations and is returning data perturbation $\Delta D(\mathbf{x}_g,\omega_k)$ for all frequencies N_t and receivers N_g. Therefore, the adjoint of it is applied to data perturbation and provides back model perturbations, that mathematically corresponds to:

$$\Delta\tilde{K}(\mathbf{x}_j) = \sum_{k=1}^{N_t}\sum_{i=1}^{N_d}\left[\frac{\partial\Delta P_K(\mathbf{x}_g,\omega_k)}{\partial\Delta K(\mathbf{x}_j)}\right]^*\Delta D(\mathbf{x}_i,\omega_k), \quad (58)$$

$$\Delta\tilde{\rho}(\mathbf{x}_j) = \sum_{k=1}^{N_t}\sum_{i=1}^{N_d}\left[\frac{\partial\Delta P_\rho(\mathbf{x}_g,\omega_k)}{\partial\Delta\rho(\mathbf{x}_j)}\right]^*\Delta D(\mathbf{x}_i,\omega_k). \quad (59)$$

Substituting Equations 51 and 54 into the two previous relations we have:

$$\Delta\tilde{K}(\mathbf{x}_j) = -\frac{1}{K_0(\mathbf{x}_j)^2}\sum_{k=1}^{N_t}\sum_{i=1}^{N_d}\omega^2 W(\omega_k)^* G(\mathbf{x}_j,\omega_k,\mathbf{x}_s;K_0,\rho_0)^* \\ G(\mathbf{x}_j,\omega_k,\mathbf{x}_i;K_0,\rho_0)^*\Delta D(\mathbf{x}_i,\omega_k) \tag{60}$$

$$\Delta\tilde{\rho}(\mathbf{x}_j) = \frac{1}{\rho_0(\mathbf{x}_j)^2}\sum_{k=1}^{N_t}\sum_{i=1}^{N_d}\left[\nabla G(\mathbf{x}_j,\omega_k,\mathbf{x}_s;K_0,\rho_0)W(\omega_k)\right]^* \cdot \\ \nabla G(\mathbf{x}_j,\omega_k,\mathbf{x}_i;K_0,\rho_0)^*\Delta D(\mathbf{x}_i,\omega_k), \tag{61}$$

where we have used the reciprocity property of the Green's function:

$$G(\mathbf{x}_i,\omega_k,\mathbf{x}_j;K_0,\rho_0) = G(\mathbf{x}_j,\omega_k,\mathbf{x}_i;K_0,\rho_0). \tag{62}$$

Equations 60 and 61 in the time and continuous domain are:

$$\delta\tilde{K}(\mathbf{r}) = \frac{1}{K_0(\mathbf{r})^2}\sum_{i=1}^{N_d}\left[\frac{\partial^2 p_0(\mathbf{r})}{\partial t^2}\star\delta p_i'(\mathbf{r})\right](0), \tag{63}$$

$$\delta\tilde{\rho}(\mathbf{r}) = \frac{1}{\rho_0(\mathbf{r})^2}\sum_{i=1}^{N_d}\left[\nabla p_0(\mathbf{r})\mathbin{\dot{\star}}\nabla\delta p_i'(\mathbf{r})\right](0), \tag{64}$$

where $[f\star g](0)$ and $[f\mathbin{\dot{\star}} g](0)$ denote the zero-lag crosscorrelation and the zero-lag crosscorrelation of the scalar product of two functions, respectively; $p_0(\mathbf{r},t)$ is given by Equation 38, and $\delta p_i'(\mathbf{r},t)$ is obtained as:

$$\delta p_i'(\mathbf{r},t) = \tilde{g}(\mathbf{r},t,\mathbf{r}_i;K,\rho) * \delta d(\mathbf{r}_i,t), \tag{65}$$

in which we are convolving the anticasual Green's function with the data perturbation. Because we are considering a single source, we can rewrite Equations 63 and 64 as:

$$\delta\tilde{K}(\mathbf{r}) = \frac{1}{K_0(\mathbf{r})^2}\left[\frac{\partial^2 p_0(\mathbf{r})}{\partial t^2}\star\delta p'(\mathbf{r})\right](0), \tag{66}$$

$$\delta\tilde{\rho}(\mathbf{r}) = \frac{1}{\rho_0(\mathbf{r})^2}\left[\nabla p_0(\mathbf{r})\mathbin{\dot{\star}}\nabla\delta p'(\mathbf{r})\right](0), \tag{67}$$

where:

$$\delta p'(\mathbf{r},t) = \sum_{i=1}^{N_d}\tilde{g}(\mathbf{r},t,\mathbf{r}_i;K,\rho) * \delta d(\mathbf{r}_i,t). \tag{68}$$

Property of staggered-grid operators

Staggered-grid derivative operators have a useful property that can be used when applying their adjoints. We define the forward and backward first-order derivative with second-order accuracy as:

$$\left.\frac{df(x)}{dx}\right|_{\mathbf{x}_{i+\frac{1}{2}}} = \frac{f(\mathbf{x}_{i+1})-f(\mathbf{x}_i)}{\Delta x}, \tag{69}$$

$$\left.\frac{df(x)}{dx}\right|_{\mathbf{x}_{i-\frac{1}{2}}} = \frac{f(\mathbf{x}_i)-f(\mathbf{x}_{i-1})}{\Delta x}, \tag{70}$$

respectively. If we assume the function $f(x)$ vanishes outside the computational domain, we can write the operator matrix $\mathbf{D}^+$ of Equation 69 using transient convolution (Claerbout, 2014) as:

$$\mathbf{D}^+ = \begin{bmatrix} -1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 \end{bmatrix}. \tag{71}$$

Taking the adjoint of this operator we have:

$$\mathbf{D}^{+*} = \begin{bmatrix} -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & -1 \end{bmatrix} =$$

$$- \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 1 \end{bmatrix} = -\mathbf{D}^-, \tag{72}$$

where $\mathbf{D}^-$ denotes the operator matrix of Equation 70. From relation 72 follows that:

$$\mathbf{D}^{-*} = -\mathbf{D}^+. \tag{73}$$

These equalities hold also in case higher-accuracy finite-difference operators are used.

REFERENCES

Aki, K. and P. G. Richards, 2002, Quantitative seismology, volume **1**.

Almomin, A., 2013, Accurate implementation of two-way wave-equation operators: SEP-Report, **149**, 281–288.

Claerbout, J. F., 2014, Geophysical image estimation by examples.

Innanen, K. A., 2014, Reconciling seismic avo and precritical reflection fwi–issues in multiparameter gradient-based updating: Presented at the 76th EAGE Conference and Exhibition, European Association of Geoscientists and Engineers.

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.

Mattsson, K., 2012, Summation by parts operators for finite difference approximations of second-derivatives with variable coefficients: Journal of Scientific Computing, **51**, 650–682.

Operto, S., Y. Gholami, V. Prieux, A. Ribodetti, R. Brossier, L. Metivier, and J. Virieux, 2013, A guided tour of multiparameter full-waveform inversion with multicomponent data: From theory to practice: The Leading Edge, **32**, 1040–1054.

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