A deconvolution-based objective function for wave-equation inversion

Simon Luo and Paul Sava, Center for Wave Phenomena, Colorado School of Mines

SUMMARY

We propose a new objective function for wave-equation inversion that seeks to minimize the norm of the weighted deconvolution between synthetic and observed data. Compared to more the conventional difference-based objective function which minimizes the norm of the residual between synthetic and observed data, the deconvolution-based objective function is less susceptible to cycle skipping and local minima. Compared to a crosscorrelation-based objective function, the deconvolution-based objective function is less sensitive to a bandlimited or non-impulsive source function, which may result in a nonzero gradient of the objective function even when the constructed velocity model matches the true model.

INTRODUCTION

The construction of a subsurface velocity model is a central problem in exploration seismology. Recently there has been much effort devoted toward wave-equation inversion methods, which can provide high resolution velocity models.

Wave-equation inversion methods can be divided into two categories: data domain methods (Lailly, 1984; Tarantola, 1984, 1986; Mora, 1987; Pratt et al., 1998; Shin and Ha, 2008), and image domain methods (Shen et al., 2003; Sava and Biondi, 2004). Data domain methods minimize the residual between synthetic and observed seismograms, while image domain methods optimize some measure of image quality (e.g., flatness of common-image gathers). See Symes (2008) for an overview and discussion of these methods.

Data domain methods can be further categorized as either traveltime-based inversion (Luo and Schuster, 1991; Zhang and Wang, 2009), which matches the phase information contained in observed seismograms; or full-waveform inversion (Lailly, 1984; Tarantola, 1984; Mora, 1987), which matches both the amplitude and phase information in observed data. Because of this strict data matching, full-waveform inversion has high resolution, but is susceptible to problems such as cycle skipping and local minima, especially when the data lack low frequencies. In addition, if we synthesize data using the acoustic wave equation, any inconsistencies in the amplitudes between observed and synthetic data due to non-acoustic effects, such as density variations and converted waves, negatively impact the inversion result. In comparison, traveltime-based inversion is less susceptible to cycle skipping, and in general the objective function tends to have fewer local minima compared to that of full-waveform inversion. Also, because it emphasizes the phase information contained in data, traveltime inversion is less affected by inability to correctly model seismic amplitudes. However, traveltime inversion generally has lower resolution compared to full-waveform inversion.

Nonetheless, traveltime-based inversion may be useful when an initial velocity model is not available, or may be used to build an initial model. One example of a traveltime-based objective function is the norm of the weighted crosscorrelation of synthetic and observed data (van Leeuwan and Mulder, 2010). While this crosscorrelation-based objective function is less susceptible to cycle skipping, it may be sensitive to a bandlimited or non-impulsive source function. For a bandlimited or non-impulsive source, when the constructed velocity model matches the true model, the crosscorrelation is centered at zero time lag, but it is not confined to zero lag. Consequently, this may result in artifacts in the gradient, even with the correct velocity model.

To resolve this issue, we propose a new objective function for wave-equation inversion that minimizes the norm of the weighted deconvolution, rather than the crosscorrelation, of synthetic and observed data. As for the crosscorrelation-based objective function, the deconvolution-based objective function is less susceptible to cycle skipping compared to the conventional difference-based objective function used in full-waveform inversion. In addition, it can also avoid potential problems with the crosscorrelation-based objective function when using a non-impulsive source function. The method is fully automatic as it does not require traveltime picking, and it is simple to implement as a variation of an existing full-waveform inversion implementation.

THEORY

To compute the gradient of a functional $\mathcal{J}(m) = \mathcal{H}(u(m))$ with respect to the model parameter m, we use the adjoint-state method (Tarantola, 2005; Plessix, 2006). Here, $\mathcal{H}(u(m))$ is the objective function we seek to minimize.

Wave propagation in an arbitrary medium characterized by the slowness $s(\mathbf{x})$ is governed by the Helmholtz equation:

$$\left[-\omega^2 s^2(\mathbf{x}) - \nabla^2\right] u(e,\mathbf{x},\omega,m) = f_s(e,\mathbf{x},\omega), \qquad (1)$$

where e is the shot number, $\mathbf{x}$ is the position in space, ω is the wavefield frequency, $f_s(e,\mathbf{x},\omega)$ is the source function, and $u(e,\mathbf{x},\omega,m)$ is the wavefield. The dependency of the functional $\mathcal{J}(m)$ on the model parameter m is through the state variable $u(e,\mathbf{x},\omega,m)$. We write the functional $\mathcal{F}$ linking the model parameter space and the state variable space as

$$\mathcal{F}(\mathbf{x},\omega,u,m,f_s) = \mathcal{L}(\mathbf{x},\omega,m)u(e,\mathbf{x},\omega,m) - f_s(e,\mathbf{x},\omega), \quad (2)$$

where

$$\mathcal{L}(\mathbf{x},\omega,m) = \left[-\omega^2 s^2(\mathbf{x}) - \nabla^2\right], \qquad (3)$$

is the Helmholtz operator, and

$$m(\mathbf{x}) = s^2(\mathbf{x}). \qquad (4)$$

Deconvolution-based objective function

To solve for the state variable u, we impose the condition

$$\mathcal{F}(\mathbf{x}, \omega, u, m, f_s) = 0, \tag{5}$$

to obtain

$$\mathcal{L}(\mathbf{x}, \omega, m)u(e, \mathbf{x}, \omega, m) = f_s(e, \mathbf{x}, \omega). \tag{6}$$

Thus, the state variable u represents the wavefield simulated using the forward modeling operator $\mathcal{L}$, with the source function f_s. To obtain the adjoint-state variable $a(e, \mathbf{x}, \omega, m)$, we solve the adjoint-state equation:

$$\left[\frac{\partial \mathcal{F}(\mathbf{x}, \omega, u, m, f_s)}{\partial u}\right]^* a(e, \mathbf{x}, \omega, m) = \frac{\partial \mathcal{H}(u)}{\partial u}, \tag{7}$$

where * denotes the adjoint. For a derivation of this equation, see Plessix (2006). From equation 2, we have

$$\frac{\partial \mathcal{F}(\mathbf{x}, \omega, u, m, f_s)}{\partial u} = \mathcal{L}(\mathbf{x}, \omega, m), \tag{8}$$

and from the adjoint-state equation,

$$\mathcal{L}^*(\mathbf{x}, \omega, m)a(e, \mathbf{x}, \omega, m) = \frac{\partial \mathcal{H}(u)}{\partial u}. \tag{9}$$

Thus, the adjoint-state variable a represents the wavefield simulated using the adjoint modeling operator $\mathcal{L}^*$, with the adjoint source ${}^{\partial \mathcal{H}}/_{\partial u}$ given by the derivative of the objective function $\mathcal{H}$ with respect to the state variable u.

The gradient of the functional $\mathcal{J}$ with respect to the model parameter m is given by

$$\frac{\partial \mathcal{J}(m)}{\partial m} = \sum_{\mathbf{e}} \sum_{\omega} \omega^2 \overline{u(e, \mathbf{x}, \omega, m)} a(e, \mathbf{x}, \omega, m), \tag{10}$$

which is simply the scaled zero-lag correlation of the wavefield simulated using the forward modeling operator $\mathcal{L}$ and source function f_s, with the wavefield simulated using the adjoint modeling operator $\mathcal{L}^*$ and adjoint source function ${}^{\partial \mathcal{H}}/_{\partial u}$.

Because the state variable u is independent of the choice of objective function $\mathcal{H}$, the gradient of the functional $\mathcal{J}$ ultimately depends on the adjoint-state variable a, which is computed using the adjoint source.

Difference-based objective function

We first consider the conventional difference-based objective function, which is defined as the squared l_2-norm of the difference between synthetic and observed data:

$$\mathcal{H}_{DIF}(u_s) = \frac{1}{2} \sum_{e, \mathbf{x}, \omega} ||K(e, \mathbf{x})(u_s(e, \mathbf{x}, \omega, \mathbf{m}) - u_o(e, \mathbf{x}, w)||_2^2, \tag{11}$$

where u_o is the observed (i.e., recorded) wavefield, u_s is the synthetic wavefield, and $K(e, \mathbf{x})$ is a masking operator that limits the wavefields to the receiver locations. Applying the mask to a wavefield gives the corresponding data. From the adjoint-state equation, the adjoint source is obtained by taking the partial derivative of the objective function $\mathcal{H}_{DIF}$ with respect to the state variable u_s. For the difference-based objective function,

$$\frac{\partial \mathcal{H}_{DIF}(u_s)}{\partial u_s} = K(e, \mathbf{x})\overline{K(e, \mathbf{x})}\Re[u_s(e, \mathbf{x}, \omega, \mathbf{m}) - u_o(e, \mathbf{x}, \omega)]. \tag{12}$$

The adjoint source for the difference-based objective function is simply the residual between the synthetic and observed data.

Crosscorrelation-based objective function

Next we consider the crosscorrelation-based objective function proposed by van Leeuwan and Mulder (2010). The objective function is defined as the squared l_2-norm of the weighted crosscorrelation of synthetic and observed data:

$$\mathcal{H}_{COR}(u_s) = \frac{1}{2} \sum_{e, \mathbf{x}, \tau} ||K(e, \mathbf{x})P(\tau)c(e, \mathbf{x}, \tau)||_2^2, \tag{13}$$

where $c(e, \mathbf{x}, \tau)$ is the crosscorrelation:

$$c(e, \mathbf{x}, \tau) = \sum_{\omega} \overline{u_s(e, \mathbf{x}, \omega, \mathbf{m})} u_o(e, \mathbf{x}, \omega) e^{2i\omega\tau}, \tag{14}$$

and $P(\tau)$ is a penalty function. Generally, $P(\tau)$ should be chosen to penalize energy at nonzero time lags τ in the crosscorrelation. One simple penalty function, as suggested by van Leeuwan and Mulder (2010), is the time lag within a window:

$$P(\tau) = \begin{cases} \tau & \text{if } |\tau| \leq \tau_0 \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

The parameter τ_0 is the maximum allowable time lag, whose purpose is to prevent energy in the crosscorrelation at unreasonably large time lags to influence the gradient. Its value should be chosen based on the expected maximum traveltime error.

The adjoint source is given by the partial derivative of the objective function $\mathcal{H}_{COR}$ with respect to the state variable u_s. For the crosscorrelation-based objective function,

$$\frac{\partial \mathcal{H}_{COR}(u_s)}{\partial u_s} = K(e, \mathbf{x})\overline{K(e, \mathbf{x})} \times \tag{16}$$

$$\sum_{\tau} P(\tau)\overline{P(\tau)}\Re\left[u_o(e, \mathbf{x}, \omega)c(e, \mathbf{x}, \tau)e^{-2i\omega\tau}\right]. \tag{17}$$

This adjoint source can be interpreted as the observed data shifted by the traveltime difference between the synthetic and observed data, which is given by the crosscorrelation $c(e, \mathbf{x}, \tau)$. If the observed and synthetic data match, the energy in the crosscorrelation is maximized at zero time lag, and the energy at zero time lag is subsequently annihilated by the penalty function.

The goal is for the adjoint source, and as a result the gradient, to be zero when the constructed velocity model matches the true model. However, with the penalty function given in equation 15, this is true only for an impulsive, i.e., infinite bandwidth, source function. For a bandlimited or non-impulsive source, the crosscorrelation peak for the correct velocity model is centered at zero time lag, but is not confined to zero lag. Consequently, the adjoint source and the gradient are nonzero even when the model is correct.

To resolve this issue, one possibility is to choose a different penalty function that is zero over a window centered about zero time lag. This approach, however, effectively reduces the resolution of the inversion, as any traveltime differences between

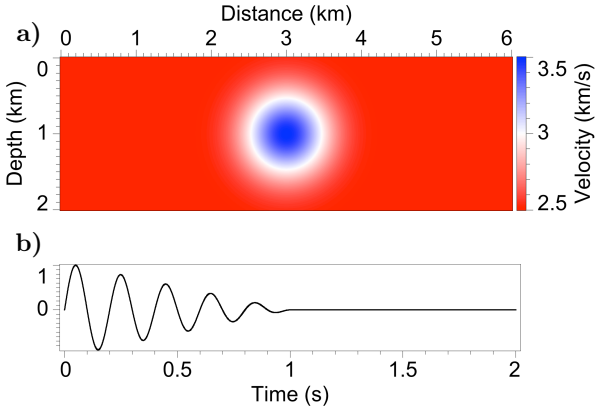

Figure 1: The true velocity model (a) and the source function (b). The trial velocity model is a constant 2.5 km/s.

the observed and synthetic data that fall within the zero-valued window cannot be resolved.

Deconvolution-based objective function

An alternative approach is to use a deconvolution-based objective function. We define the deconvolution-based objective function as the squared l_2-norm of the weighted deconvolution of synthetic and observed data:

$$\mathcal{H}_{DEC}(u_s) = \frac{1}{2}\sum_{e,\mathbf{x},\tau} ||K(e,\mathbf{x})P(\tau)d(e,\mathbf{x},\tau)||_2^2, \qquad (18)$$

where $d(e,\mathbf{x},\tau)$ is the deconvolution:

$$d(e,\mathbf{x},\tau) = \sum_{\omega} \frac{\overline{u_o(e,\mathbf{x},\omega)}u_s(e,\mathbf{x},\omega,\mathbf{m})e^{2i\omega\tau}}{\overline{u_o(e,\mathbf{x},\omega)}u_o(e,\mathbf{x},\omega)+\varepsilon^2}, \qquad (19)$$

where ε is a constant for stabilization. The adjoint source is given by the partial derivative of the objective function $\mathcal{H}_{DEC}$ with respect to u_s:

$$\frac{\partial \mathcal{H}_{DEC}(u_s)}{\partial u_s} = K(e,\mathbf{x})\overline{K(e,\mathbf{x})} \times \qquad (20)$$

$$\sum_{\tau} P(\tau)\overline{P(\tau)}\Re\left[\frac{u_o(e,\mathbf{x},\omega)d(e,\mathbf{x},\tau)e^{-2i\omega\tau}}{\overline{u_o(e,\mathbf{x},\omega)}u_o(e,\mathbf{x},\omega)+\varepsilon^2}\right]. \qquad (21)$$

The interpretation of this adjoint source is similar to that of the crosscorrelation-based objective function, with the important distinction being that the traveltime shift between the synthetic and observed data is now represented by the deconvolution $d(e,\mathbf{x},\tau)$ instead of the crosscorrelation $c(e,\mathbf{x},\tau)$.

Given synthetic and observed data that differ by only a traveltime shift, in the limit of infinite bandwidth, the deconvolution of these data is a shifted delta function. If the constructed velocity model matches the true model, the energy in the deconvolution is both centered and confined to zero time lag, and is completely annihilated by the penalty function $P(\tau)$ given in equation 15. This is not the case for the crosscorrelation-based objective function. Thus, we expect the deconvolution-based objective function to provide a more reasonable gradient and

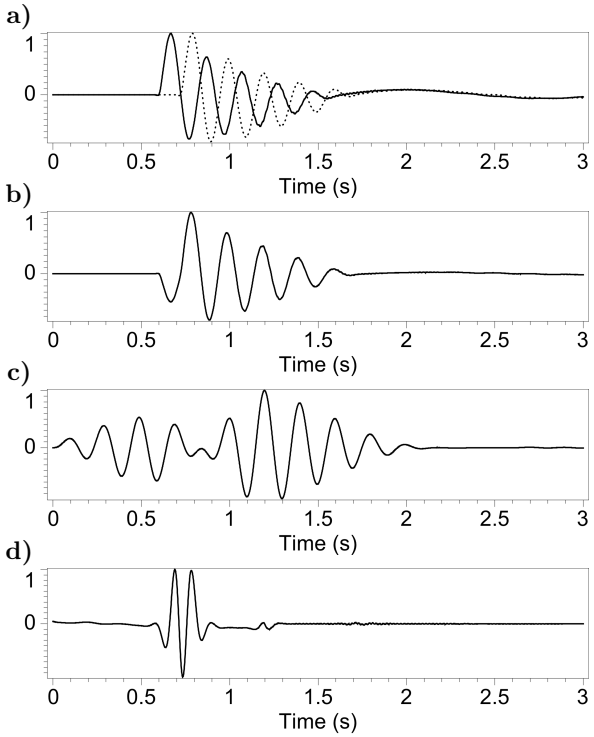

Figure 2: Observed (solid line) and synthetic data (a) for the velocity model shown in Figure 1, and the adjoint source for the difference-based (b), convolution-based (c), and deconvolution-based (d) objective functions.

faster convergence compared to the crosscorrelation-based objective function when using a bandlimited or non-impulsive source function.

EXAMPLES

To compare the difference-based, crosscorrelation-based, and deconvolution-based objective functions, we compare their adjoint sources and gradients for a synthetic example. The true model shown in Figure 1a consists of a background velocity of 2.5 km/s with a Gaussian anomaly in the center, while the trial model consists of only the background velocity of 2.5 km/s. We use the tapered 5 Hz sine wave shown in Figure 1b as our source function to demonstrate the effect of a non-impulsive source. Note that for this source function, the magnitude of the velocity anomaly is large enough to produce cycle skipping for the difference-based objective function.

Figure 2a shows the observed data (solid line) modeled with the true velocity model and the synthetic data (dotted line) modeled with the trial velocity model for a single shot located at distance 3 km and depth 0.06 km, and a single receiver at distance 3 km and depth 1.94 km. Figures 2b, 2c, and 2d show the adjoint sources for the difference-, crosscorrelation-, and deconvolution-based objective functions, respectively. Note the oscillations in the adjoint sources for the difference- and crosscorrelation-based objective functions compared to that of

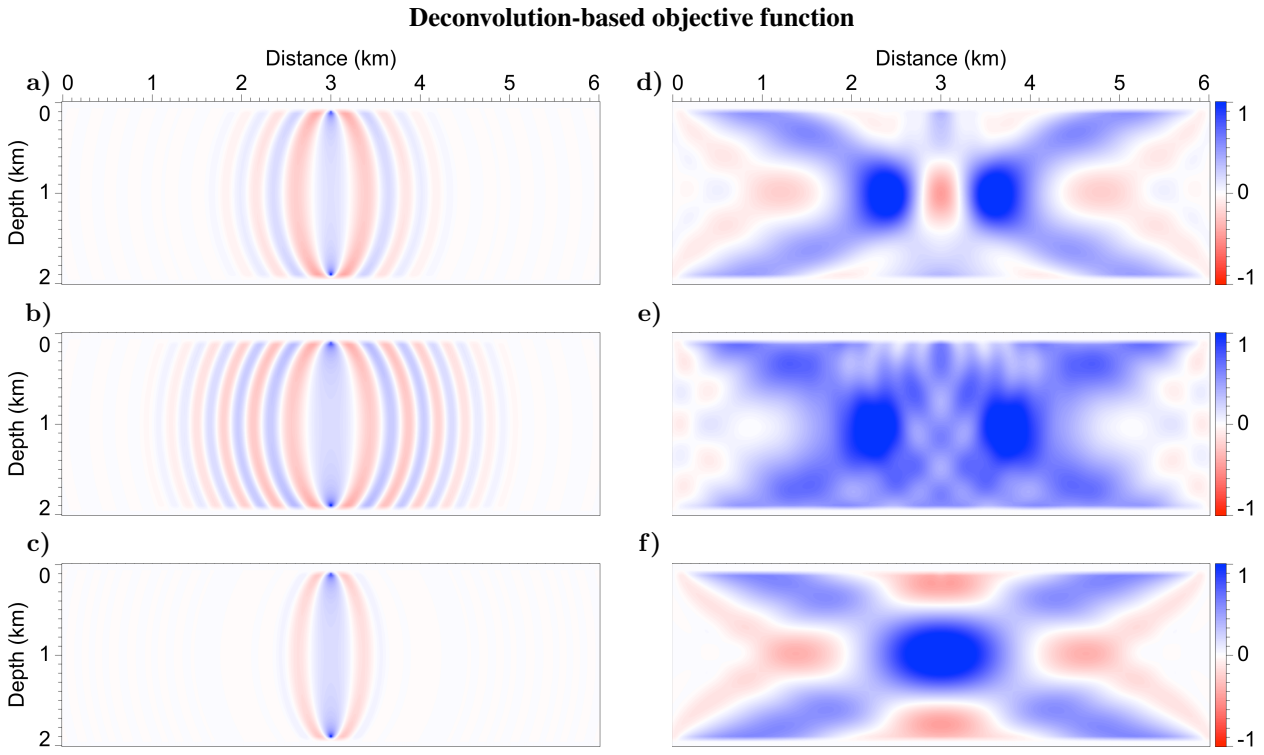

Figure 3: The sensitivity kernels for the difference-based (a), correlation-based (b), and deconvolution-based (c) objective functions contribute to the gradients of the difference-based (d), correlation-based (e), and deconvolution-based (f) objective functions.

the deconvolution-based objective function, which is more impulsive.

The sensitivity kernels and gradients for these objective functions are shown in Figure 3. Figures 3a, 3b, and 3c show sensitivity kernels computed for the difference-, crosscorrelation-, and deconvolution-based objective functions, respectively, using their corresponding adjoint sources shown in Figure 2. Figures 3d, 3e, and 3f show the gradients computed for the difference-, crosscorrelation-, and deconvolution-based objective functions, respectively, for 96 shots at depth 0.06 km and full coverage of receivers at depth 1.94 km.

Note that the oscillations in the adjoint sources for the difference- and crosscorrelation-based objective functions shown in Figure 2 are reflected in their sensitivity kernels and gradients. Also, notice that because the velocity anomaly is large enough to produce cycle skipping, the gradient of the difference-based objective function (Figure 3d) does not correctly recover the sign of the anomaly. Comparing the gradients to the correct velocity update (i.e., the difference between the true velocity model and the trial model), we observe that the gradient of the deconvolution-based objective function is closest to the correct update.

CONCLUSION

Crosscorrelation- and deconvolution-based objective functions for wave-equation inversion are less susceptible to the cycle skipping and local minima problems that are inherent to strict data matching inversions such as full-waveform inversion using the difference-based objective function. Thus, they may be useful for building an initial velocity model, which then may be close enough to the true model for full-waveform inversion to converge to the global minimum of the difference-based objective function. However, the crosscorrelation-based objective function may be sensitive to a bandlimited or non-impulsive source function, because the crosscorrelation of synthetic and observed data produced by a non-impulsive source is not confined to zero lag even when the constructed velocity model matches the true model. In comparison, the deconvolution of these data is more impulsive, and is more confined to zero lag given the correct velocity model. For this reason, wave-equation inversion using the proposed deconvolution-based objective function may provide more reasonable gradient estimates and faster convergence compared to the crosscorrelation-based objective function.

ACKNOWLEDGEMENTS

This work was supported by the sponsors of the Center for Wave Phenomena at the Colorado School of Mines.

REFERENCES

Lailly, P., 1984, The seismic inverse problem as a sequence of before stack migration: Conference on Inverse Scattering, SIAM, 206–220.

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

Mora, P. R., 1987, Nonlinear two-dimensional elastic inversion of multioffset seismic data: Geophysics, **52**, 1211–1228.

Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503.

Pratt, G., C. Shin, and G. Hicks, 1998, Gauss-newton and full newton methods in frequency-space seismic waveform inversion: Geophysical Journal International, **113**, 341–462.

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

Shen, P., W. Symes, and C. C. Stolk, 2003, Differential semblance velocity analysis by wave-equation migration: 73rd Annual International Meeting, SEG Expanded Abstracts, 2132–2135.

Shin, C., and W. Ha, 2008, A comparison between the behavior of objective functions for waveform inversion in the frequency and laplace domains: Geophysics, **73**, VE119–VE133.

Symes, W. W., 2008, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, 765–790.

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

———, 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51**, 1893–1903.

———, 2005, Inverse Problem Theory: Society of Industrial and Applied Mathematics.

van Leeuwan, T., and W. A. Mulder, 2010, A correlation-based misfit criterion for wave-equation traveltime tomography: Geophysical Journal International, **182**, 1383–1394.

Zhang, Y., and D. Wang, 2009, Travetime information-based wave-equation inversion: Geophysics, **74**, WCC27–WCC36.