A comparative review of wavefield tomography methods

Paul Sava
Center for Wave Phenomena, Colorado School of Mines

ABSTRACT
Waveform inversion and wave-equation migration velocity analysis share one common feature: both techniques exploit information available in the same source and receiver wavefields reconstructed at all times and at all locations in space. These two wavefields are reconstructed from the same source function and observed data as solutions to the same wave-equation. Both techniques invert for the same model parameters and are formulated in the same inverse problem framework coupled with gradient calculations using the adjoint state method.
The main distinction between these techniques is in the way in which they handle the many seismic experiments representing a seismic survey. For the techniques in the waveform inversion family, an objective function is formulated by comparing wavefields for individual experiments; then, a compound objective function is formed by summing the objective functions of all the experiments. Many wavefield comparison measures are possible, including differencing and simple or oriented cross-correlations. For the techniques in the wave-equation migration velocity analysis family, an objective function is formulated after we accumulate for all experiments, e.g. through stacking, different wavefield comparison measures. Here, too, we can campare the wavefields using different measures, including simple and oriented cross-correlations. This summation over experiments produces images that simply function as proxies for the underlying seismic wavefields.

Key words: wavefield, tomography, imaging, adjoint-state method

1 INTRODUCTION

Seismic wavefield tomography methods applied to subsurface imaging are becoming increasingly popular. This trend is driven by the need to exploit every bit of information contained in the acquired data in order to constrain complex geologic structures and to extract detailed information about the geometry and physical properties of the bodies under investigation. Although much cheaper and more popular for the time being, alternative methods based on ray tracing suffer from numerical instability in complex geology and from difficulties in handling multi-pathing. Here, I restrict my attention to wavefield tomography methods and seek to analyze the various techniques in this category using a common theoretical framework.

Wavefield tomography methods have been extensively studied and can be separated into two main families:

(i) **Waveform inversion** – a family of techniques designed to characterize the subsurface by directly mapping its physical properties (e.g. the velocity). This methodology has a long history and remains an active subject of research to-date (Tarantola, 1984; Gauthier et al., 1986; Mora, 1987, 1988, 1989; Pratt and Worthington, 1990; Crase et al., 1990; Woodward, 1992; Bunks et al., 1995; Pratt, 1999; Plessix et al., 2000; Sirgue and Pratt, 2004; Operto et al., 2004; Tarantola, 2005; Plessix, 2006; Min and Shin, 2006; Shin and Min, 2006; Vigh and Starr, 2008; Shin and Ha, 2008; Plessix, 2009; Brossier et al., 2009; Virieux and Operto, 2009; van Leeuwen and Mulder, 2010; Symes, 2010; Alkhalifah and Choi, 2014).

(ii) **Wave-equation migration velocity analysis** – a family of techniques designed to alternate between identifying interfaces in the subsurface through **wave-equation migration** and constraining the physical properties of the medium (e.g. velocity) that allow placement of reflectors at their correct position in the subsurface. Both migration and migration velocity analysis also have a long history reported in publications covering fundamental imaging concepts (Berkhout, 1982; Clærbout, 1985; Yilmaz, 2001; Biondi, 2006), various implementation strategies, e.g. using the acoustic wave-equation in reverse-time migration (Kosloff and Baysal, 1983; Baysal et al., 1983; McMechan, 1983), various imaging conditions (de Bruin et al., 1990; Rickett and Sava, 2002; Sava and Fomel, 2003; Biondi and Symes, 2004; Fomel, 2004; Sava and Fomel, 2005; Stolk and Symes, 2004; Biondi and Symes, 2004; Sava and Fomel, 2006; Yang and Sava, 2010; Sava and Vasconcelos, 2011; Sava and Vlad, 2011), and differ-

ent approaches for velocity analysis (Mulder and ten Kroode, 2002; Shen et al., 2003; Sava and Biondi, 2004a,b; Shen et al., 2005; Albertin et al., 2006; Maharramov and Albertin, 2007; Soubaras and Gratacos, 2007; Shen and Symes, 2008; Symes, 2009; Zhang and Wang, 2009; Yang and Sava, 2008, 2009; Biondi, 2010; Yang and Sava, 2011; Perrone et al., 2013; Biondi and Almomin, 2012, 2013; Shan and Wang, 2013; Tang and Biondi, 2013; Weibull and Arntsen, 2013; **?**).

Both wavefield tomography families of method have are effective for characterizing the subsurface properties, which justifies their existing and growing popularity. However, a lot has been argued around their different implementation details, but mostly about their ability to exploit the information contained in the observed data.

My goal in this paper is to explore comparatively the wavefield tomography methods and address some of the following questions. What is really the difference between these two classes of techniques? In both cases wave propagation is governed by the same physics (i.e. the same wave-equation), the boundary conditions are the same (i.e. the same data), the model parameters are the same, etc. If we operate with the same wave-equation with the same boundary and initial conditions, i.e. we use the same wavefields, why is one class of methods considered to be "full wavefield" and the other one not so? What makes the difference, if there is one?

In the following, I proceed to describe these wavefield tomography methods from the perspective of an inverse problem. This point of view leads to a common framework for which wavefield tomography implemented in different domains become just special cases, and wavefield imaging just a proxy for analyzing seismic wavefields. This common point of view allows us to compare and contrast without ambiguity the different techniques based on their shared elements. I do not intend in this paper to provide a historic perspective on the development of the various techniques, which is why in the following I restrict references to a minimum.

The paper consists of several sections. I begin by reviewing the adjoint state method applied to the wavefields tomography problem, which is a special case of the general methodology summarized in Appendix A. Then, I discuss the wavefield tomography methods based on the separation introduced earlier:

- First, I discuss waveform inversion, i.e. a family of wavefield tomography methods that assume a-priori knowledge about locations in space where different wavefields match one-another. I include in this discussion a review of wavefield comparison methods, some of which are detailed in Appendix B.
- Second, I discuss wave-equation migration velocity analysis, i.e. the family of wavefield tomography methods that do not assume knowledge about where to compare the wavefields, but extract such information from the wavefields themselves. The basic building blocks in this category are images, either conventional or extended, derived from different wavefield comparison mechanisms.

These two families of wavefield tomography methods can also be described as **data-domain** for the former and **image-domain** for the later. Although convenient, this nomenclature obscures the fact that both families of techniques derive model information from the same wavefields, and also use the same seismic wavefields to propagate information from the measurement location to all locations in space. However, this terminology highlights the fact that waveform inversion and wave-equation migration velocity analysis belong to the same general class of methods.

2 THE ADJOINT-STATE METHOD FOR SEISMIC WAVEFIELDS

The adjoint-state method summarized in Appendix A (Plessix, 2006), can be applied directly to the seismic wavefield tomography problem. In this case, we seek to find the extreme relative to model parameters of an objective function $\mathcal{J}$ defined using wavefields obtained as solutions of a wave-equation, i.e. subject to physical constraints $\mathcal{F}$. The wave operator $\mathbf{L}$ could be, for example, the acoustic wave-equation

$$\mathbf{L} = \frac{1}{v^2}\frac{\partial^2}{\partial t^2} - \nabla^2 \; . \tag{1}$$

The model parameters m could be defined as slowness squared:

$$m\left(\mathbf{x}\right) = \frac{1}{v^2\left(\mathbf{x}\right)} \; . \tag{2}$$

The objective function depends on the source and receiver wavefields, $u_s\left(e,\mathbf{x},t\right)$ and $u_r\left(e,\mathbf{x},t\right)$, obtained by solving the wave-equation 1 from a known source function and observed data, $d_s\left(e,\mathbf{x},t\right)$ and $d_r\left(e,\mathbf{x},t\right)$, respectively. The wavefields are a function of the experiment index e, space $\mathbf{x}$ and time t. The seismic experiment could be based on a point source, a line or plane source, or other shot encodings etc.

In the wavefield tomography problem we are operating with two physical constraints, $\mathcal{F}_s = 0$ and $\mathcal{F}_r = 0$, accounting for the fact that both wavefields u_s and u_r are solutions of the wave-equation 1. The physical constraints are defined using the wave operator $\mathbf{L}$ by

$$\begin{bmatrix} \mathcal{F}_s \\ \\ \mathcal{F}_r \end{bmatrix} = \begin{bmatrix} \mathbf{L} & 0 \\ \\ 0 & \mathbf{L}^\top \end{bmatrix} \begin{bmatrix} u_s \\ \\ u_r \end{bmatrix} - \begin{bmatrix} d_s \\ \\ d_r \end{bmatrix} = \begin{bmatrix} 0 \\ \\ 0 \end{bmatrix} \; , \tag{3}$$

thus capturing the idea that the variables u_s and u_r are wavefields obtained as solutions to a wave-equation. The wavefields u_s and u_r are the only ones that can be computed given the known source function d_s and observed data d_r. In the following, I assume that the source function is known, although this may not always be the case in practice. The symbol $^\top$ indicates that we are using the adjoint operator to construct the receiver wavefield, i.e. we solve the equation backward in time.

For the two mentioned physical constraints, we need the variables $a_s\left(e,\mathbf{x},t\right)$ and $a_r\left(e,\mathbf{x},t\right)$ to construct the **augmented functional**

$$\mathcal{H} = \mathcal{J} - \begin{bmatrix} \mathcal{F}_s^\top & \mathcal{F}_r^\top \end{bmatrix} \begin{bmatrix} a_s \\ \\ a_r \end{bmatrix} \; . \tag{4}$$

which we intend to minimize using a gradient-based optimization method. Using the definitions from Appendix A, the quantities u_s and u_r are **state variables**, and the quantities a_s and a_r are **adjoint variable**.

2.1 State equations

As explained in Appendix A, we find the extreme of the augmented functional $\mathcal{H}$ by searching in the direction of its gradient relative to the model parameters m using

$$\frac{\partial \mathcal{H}}{\partial m} = \frac{\partial \mathcal{J}}{\partial m} - \begin{bmatrix} \left(\frac{\partial \mathcal{F}_s}{\partial m}\right)^\top & \left(\frac{\partial \mathcal{F}_r}{\partial m}\right)^\top \end{bmatrix} \begin{bmatrix} a_s \\ \\ a_r \end{bmatrix} . \quad (5)$$

The state variables u_s and u_r are computed by solving the **state equations** obtained by setting the partial derivatives of the augmented functional relative to the adjoint variables a_s and a_r to zero

$$\begin{bmatrix} \frac{\partial \mathcal{H}}{\partial a_s} \\ \\ \frac{\partial \mathcal{H}}{\partial a_r} \end{bmatrix} = \begin{bmatrix} 0 \\ \\ 0 \end{bmatrix} . \quad (6)$$

Using equation 4, the condition 6 is equivalent to

$$\begin{bmatrix} \frac{\partial \mathcal{J}}{\partial a_s} \\ \\ \frac{\partial \mathcal{J}}{\partial a_r} \end{bmatrix} = \begin{bmatrix} \mathcal{F}_s \\ \\ \mathcal{F}_r \end{bmatrix} . \quad (7)$$

Since $\mathcal{J}$ is not a function of the adjoint variables, a_s and a_r, we can write

$$\begin{bmatrix} \mathcal{F}_s \\ \\ \mathcal{F}_r \end{bmatrix} = \begin{bmatrix} 0 \\ \\ 0 \end{bmatrix} , \quad (8)$$

to obtain the state equations

$$\begin{bmatrix} \mathbf{L} & 0 \\ \\ 0 & \mathbf{L}^\top \end{bmatrix} \begin{bmatrix} u_s \\ \\ u_r \end{bmatrix} = \begin{bmatrix} d_s \\ \\ d_r \end{bmatrix} . \quad (9)$$

As stated earlier, the state variables u_s and u_r are wavefields obtained as solutions of the wave-equation solved forward and backward in time, respectively.

2.2 Adjoint equations

Likewise, the adjoint variables a_s and a_r are computed by solving the **adjoint equations** obtained by setting the partial derivatives of the augmented functional relative to the state variables u_s and u_r to zero;

$$\begin{bmatrix} \frac{\partial \mathcal{H}}{\partial u_s} \\ \\ \frac{\partial \mathcal{H}}{\partial u_r} \end{bmatrix} = \begin{bmatrix} 0 \\ \\ 0 \end{bmatrix} . \quad (10)$$

Using equation 4, the condition 10 is equivalent to

$$\begin{bmatrix} \left(\frac{\partial \mathcal{F}_s}{\partial u_s}\right)^\top & \left(\frac{\partial \mathcal{F}_r}{\partial u_s}\right)^\top \\ \\ \left(\frac{\partial \mathcal{F}_s}{\partial u_r}\right)^\top & \left(\frac{\partial \mathcal{F}_r}{\partial u_r}\right)^\top \end{bmatrix} \begin{bmatrix} a_s \\ \\ a_r \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathcal{J}}{\partial u_s} \\ \\ \frac{\partial \mathcal{J}}{\partial u_r} \end{bmatrix} . \quad (11)$$

Since the physical constraints $\mathcal{F}_s$ and $\mathcal{F}_r$ do not depend on the state variables u_r and u_s, respectively, and the physical constraints are defined in terms of a wave-equation, we can write

$$\begin{bmatrix} \left(\frac{\partial \mathcal{F}_s}{\partial u_s}\right)^\top & \left(\frac{\partial \mathcal{F}_r}{\partial u_s}\right)^\top \\ \\ \left(\frac{\partial \mathcal{F}_s}{\partial u_r}\right)^\top & \left(\frac{\partial \mathcal{F}_r}{\partial u_r}\right)^\top \end{bmatrix} = \begin{bmatrix} \mathbf{L}^\top & 0 \\ \\ 0 & \mathbf{L} \end{bmatrix} . \quad (12)$$

Defining **adjoint sources** based on the chosen objective function as

$$\begin{bmatrix} g_s \\ \\ g_r \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathcal{J}}{\partial u_s} \\ \\ \frac{\partial \mathcal{J}}{\partial u_r} \end{bmatrix} , \quad (13)$$

we can write the adjoint equations as

$$\begin{bmatrix} \mathbf{L}^\top & 0 \\ \\ 0 & \mathbf{L} \end{bmatrix} \begin{bmatrix} a_s \\ \\ a_r \end{bmatrix} = \begin{bmatrix} g_s \\ \\ g_r \end{bmatrix} . \quad (14)$$

The adjoint variables a_s and a_r are wavefields obtained as solutions of the wave-equation solved backward and forward in time, respectively.

2.3 Gradient

Finally, since we can write

$$\begin{bmatrix} \frac{\partial \mathcal{F}_s}{\partial m} \\ \\ \frac{\partial \mathcal{F}_r}{\partial m} \end{bmatrix} = \begin{bmatrix} \frac{\partial^2 u_s}{\partial t^2} \\ \\ \frac{\partial^2 u_r}{\partial t^2} \end{bmatrix} , \quad (15)$$

the constrained **gradient** of the augmented functional $\mathcal{H}$ is

$$\frac{\partial \mathcal{H}}{\partial m} = \frac{\partial \mathcal{J}}{\partial m} + \sum_e \begin{bmatrix} \left(\frac{\partial^2 u_s}{\partial t^2}\right)^\top & \left(\frac{\partial^2 u_r}{\partial t^2}\right)^\top \end{bmatrix} \begin{bmatrix} a_s \\ \\ a_r \end{bmatrix} . \quad (16)$$

This expression implies summation over time in addition to the summation over experiments, and thus the gradient can be interpreted as the zero time-lag of the cross-correlation between state and adjoint variables. Thus, we can write the gradient as

$$\frac{\partial \mathcal{H}}{\partial m} = \frac{\partial \mathcal{J}}{\partial m} + \sum_e \sum_\tau \delta(\tau) \left(\ddot{u}_s \star a_s + \ddot{u}_r \star a_r \right) , \quad (17)$$

where the double dot indicates second time derivative.

2.4 Model constraints

The inverse problem solution depends on our choice of objective function, $\mathcal{J}(u_s, u_r, m)$. In the following, I consider an objective function consisting of two parts: a term defined based on the state variables u_s and u_r, and a term defined based on the model parameters m:

$$\mathcal{J}(u_s, u_r, m) = \mathcal{J}_u(u_s, u_r) + \mathcal{J}_m(m) \ . \tag{18}$$

The term $\mathcal{J}_m$ expresses the constraint that we seek model parameters from a multidimensional Gaussian distribution with mean $\overline{m}$, i.e. the reference model, and covariance $\mathbf{C}_m$ represented by the weighting operator $\mathbf{W}_m = \mathbf{C}_m^{-\frac{1}{2}}$:

$$\mathcal{J}_m = \frac{1}{2} \left\| r_m(\mathbf{x}) \right\|_{\mathbf{x}}^2 \ , \tag{19}$$

where the model residual r_m is

$$r_m(\mathbf{x}) = \mathbf{W}_m(\mathbf{x}) \left(m(\mathbf{x}) - \overline{m}(\mathbf{x}) \right) \ . \tag{20}$$

With this notation, the model component of the constrained gradient is

$$\frac{\partial \mathcal{J}_m}{\partial m} = \mathbf{W}_m^\top(\mathbf{x}) \, r_m(\mathbf{x}) \ . \tag{21}$$

In the following, I assume that the model constraints do not change, but I consider different objective functions $\mathcal{J}_u$ corresponding to different acquisition geometry and different measures of similarity between the source and receiver wavefields. The goal of the next sections is to evaluate the gradient of the data component of the objective function with respect to model parameters, i.e.

$$\frac{\partial \mathcal{J}_u}{\partial m} = \sum_e \sum_\tau \delta(\tau) \left(\ddot{u}_s \star a_s + \ddot{u}_r \star a_r \right) \ . \tag{22}$$

This requires that we construct appropriate adjoint source and receiver wavefields, a_s and a_r, for various choices of the data objective function $\mathcal{J}_u$.

2.5 Summary and discussion

Computing the gradient of an objective function $\mathcal{J}$ for the wavefield tomography problem subject to physical constraints $\mathcal{F}$ provided by a wave equation $\mathbf{L}$, consists in the following steps:

- Compute the state variables, i.e. the seismic wavefields u_s and u_r, by solving the state equations, i.e. wave-equations with known source functions d_s and d_r, equation 9.
- Compute the adjoint sources g_s and/or g_r, which depend on the specific choice of objective function $\mathcal{J}_u$, equation 13.
- Compute the adjoint variables, i.e. the seismic wavefields a_s and a_r, by solving the adjoint equations which are wave-equations with source functions g_s and g_r, equation 14.
- Compute the gradient of the objective function, constrained by the specific choice of wave-equation, equation 17.

This framework remains the same for various choices of objective functions $\mathcal{J}_u$ and wave equations $\mathcal{F}$. In the following, I concentrate on the cases when

(i) we define the objective function $\mathcal{J}_u$ at a location where we know a-priori that the wavefields u_s and u_r should be identical with one-another, i.e. data-domain wavefield tomography, or

(ii) we define the objective function $\mathcal{J}_u$ at locations identified based on the coherence between the wavefields as a function of experiment, e.g. similarity as a function of shot position, i.e. image-domain wavefield tomography.

I discuss these two cases separately in the following sections.

3 DATA-DOMAIN WAVEFIELD TOMOGRAPHY

Let us consider the wavefield tomography problem under the assumption that we know positions in space where the source and receiver wavefields should be identical with one-another. For example, we can do the analysis at the source or receiver coordinates, or at another point located somewhere else in the medium where this property holds. This later case could correspond, for example, to the case of imaging inside an object of known exterior shape, e.g. in non-destructive testing, but it also provides a link to the methods described in the following section.

I illustrate the methods in this section with a simple model depicted in Figure 1(a). The source is at $\{0, 0\}$ km, and the receiver is at $\{4, 0\}$ km. The state variables are the wavefields u_s and u_r, depicted in Figures 2(a) and 2(b), respectively. Both wavefields are represented in successive frames ordered by increasing time, although the receiver wavefield (u_r) is actually generated backward in time.

As discussed earlier, the key element for the wavefield tomography problem is the objective function which controls both the adjoint sources g_s and g_r, in addition to the gradient itself. In all cases, I formulate the objective function using the fact that the source and receiver wavefields match each-other, if the model is accurate.

3.1 Difference

We can define an objective function using the difference between the source and receiver wavefields at the receiver position. This is equivalent to saying that we are comparing the simulated and observed data, and therefore this objective function corresponds to the conventional "full waveform inversion".

Using the difference between the source and receiver wavefields, we can define the data residual

$$r_d(e, \mathbf{x}, t) = \mathbf{W}_u(e, \mathbf{x}, t) \left(u_s(e, \mathbf{x}, t) - u_r(e, \mathbf{x}, t) \right) \ , \tag{23}$$

and then the objective function

$$\mathcal{J}_u = \sum_e \frac{1}{2} \left\| r_d(e, \mathbf{x}, t) \right\|_{\mathbf{x},t}^2 \ , \tag{24}$$

where the weight $\mathbf{W}_u$ is a combination of an operator $\mathbf{K_r}$ which simply restricts the source and receiver wavefields to the

known receiver position, and another operator $\mathbf{D}(t)$ weighing the data as a function of time:

$$\mathbf{W}_u(e,\mathbf{x},t) = \mathbf{K}_\mathbf{r}(e,\mathbf{x})\,\mathbf{D}(t) \ . \tag{25}$$

Using this objective function 24, we can write the adjoint sources

$$\begin{bmatrix} g_s(e,\mathbf{x},t) \\ \\ g_r(e,\mathbf{x},t) \end{bmatrix} = \begin{bmatrix} +\mathbf{W}_u^\top(e,\mathbf{x},t) \\ \\ -\mathbf{W}_u^\top(e,\mathbf{x},t) \end{bmatrix} \left[r_d(e,\mathbf{x},t)\right] \ , \tag{26}$$

which are just weighted versions of the difference between the source and receiver wavefields, restricted to the receiver position, Figure 1(c). As indicated earlier, the adjoint variables (a_s and a_r), are wavefields extrapolated using the given wave-equation backward and forward in time, respectively. In this problem, only the adjoint source wavefield is relevant because it is reconstructed backward in time, thus interacting with the forward simulated source wavefield, Figure 3(a). In contrast, the adjoint receiver wavefield is reconstructed forward in time, and thus it does not interact in principle with the backward simulated receiver wavefields except at and around the receiver. Then, the gradient is computed by correlating the corresponding source and receiver state and adjoint variables, Figure 3(c).

Since the state variables are simulated with a fairly inaccurate model (15% error relative to the correct velocity), the source and receiver wavefields at the receiver are out of phase; this situation corresponds to the frequently-discussed cycle-skipping problem of full waveform inversion. The wavefield difference at the known analysis point, which coincides with the receiver position in this example, Figure 1(c), represents the superposition of pulses simulated in the true velocity, Figure 1(b), with another pulse simulated in the background (inaccurate) velocity.

All these pulses are back-propagated in the adjoint source wavefield (a_s), and they interfere with the forward simulated wavefield (u_s) along several paths. First, wavefield coincidence occurs along the path connecting directly the source with the receiver due to the fact that one of the pulses in the back-propagated difference has actually been generated by forward propagating in the same background model. Second, wavefield coincidence also occurs along two elliptical paths corresponding to the time of the pulses sen in the observed data. The ellipse corresponding to the reflected data is the familiar migration impulse response, which indicates a straightforward link between migration and waveform inversion.

It is apparent that at least some of the elliptical shapes do not correspond to any interface in the model, but are merely an expression of crosstalk between different waves in the extrapolated wavefields that should not correlate with one-another. For wavefield tomography, most of these elliptical shapes are artifacts that interfere with the desirable update of the background model. This can be seen, in the second example, Figure 5(a). In this case, the background model contains the reflector at $z = 2$ km, which generates additional pulses in the background data, and therefore in the data difference, Figure 5(c). The adjoint source wavefield carries all these pulses back into the model, Figure 7(a), and their interference with the source wavefield leads to the gradient seen in Figure 7(c). This gradient contains both the direct and reflected wave paths, but it is also contaminated with many artifacts that do not correspond to anything physical in the model.

This observation remains true if we analyze the synchronization of the source and receiver wavefields at another location where the two wavefields are expected to be similar. For example, in the setup from Figure 9(a), the data difference from Figure 9(c) enables us to generate the adjoint source and receiver wavefields depicted in Figures 11(a) and 13(a), thus leading to the gradients seen in Figures 11(c) and 13(c), respectively. The cross-talk artifacts seen here are related to the way in which we compute the data residual, i.e. as data difference. The question is whether we can compute a different data residual that is less prone to interference between unrelated wavefield paths. One possibility is to use data cross-correlation, as discussed in the following section.

3.2 Time correlation

Using the same source and receiver wavefields (u_s and u_r), Figures 2(a) and 2(b), we can define a data residual

$$r_d(e,\mathbf{x},\tau) = \mathbf{W}_u(e,\mathbf{x},\tau)\,c(e,\mathbf{x},\tau) \tag{27}$$

based on their cross-correlation, Figure 1(d):

$$c(e,\mathbf{x},\tau) = u_s(e,\mathbf{x},t) \star u_r(e,\mathbf{x},t) \ , \tag{28}$$

and form the objective function

$$\mathcal{J}_u = \sum_e \frac{1}{2}\left\|r_d(e,\mathbf{x},\tau)\right\|^2_{\mathbf{x},\tau} \ , \tag{29}$$

where the weight $\mathbf{W}_u$ is as a combination of an operator $\mathbf{K}_\mathbf{r}$ restricting the source and receiver wavefields to the known receiver position, and another operator $\mathbf{P}(\tau)$ weighing the cross-correlation as a function of time-lag τ:

$$\mathbf{W}_u(e,\mathbf{x},\tau) = \mathbf{K}_\mathbf{r}(e,\mathbf{x})\,\mathbf{P}(\tau) \ . \tag{30}$$

The weight $\mathbf{P}$ is essential in the definition of the data residual. Its role is to nullify the cross-correlation when the wavefields are identical, i.e. when the correlation function is maximum at zero time-lag.

Using the objective function 29, the adjoint sources are

$$\begin{bmatrix} g_s(e,\mathbf{x},t) \\ \\ g_r(e,\mathbf{x},t) \end{bmatrix} = \begin{bmatrix} u_r(e,\mathbf{x},t) \\ \\ u_s(e,\mathbf{x},t) \end{bmatrix} * \left[\mathbf{W}_u^\top(e,\mathbf{x},\tau)\,r_d(e,\mathbf{x},\tau)\right] \ , \tag{31}$$

where the symbol $*$ indicates convolution (i.e. the adjoint of correlation) between the receiver and source wavefields, and the weighted residual. As indicated earlier, the adjoint variables (a_s and a_r) are wavefields extrapolated using the given wave-equation backward and forward in time, respectively. As before, only the adjoint source wavefield is relevant because it is reconstructed backward in time, thus interacting with the forward simulated source wavefield, Figure 3(b). The adjoint receiver wavefield is reconstructed forward in time, and thus

it does not interact in principle with the backward simulated receiver wavefields except at or around the receiver.

Finally, the gradient is computed by correlating the corresponding source and receiver state and adjoint variables, Figure 3(d). Comparing with the gradient of the difference objective function, Figure 3(c), the correlation gradient is corrupted by fewer artifacts. The artifacts are due to the fact that the pulse contained in the data simulated in the background model correlates with individual pulses in the observed data, Figure 1(d). The adjoint source convolves the penalized correlation function with the observed data, thus leading to additional pulses in the adjoint source wavefield and consequently to artifacts resulting from cross-talk between paths that should not correlate with one-another. A better example is in Figure 7(d) where the gradient contains visible cross-talk artifacts, albeit fewer than the ones seen for the difference objective function in Figure 7(c). This observation also remains true for the case when we analyze wavefield synchronization away from either the source and receiver positions. For example, in the setup from Figure 9(a), the wavefield correlation from Figure 9(d) enables us to generate the adjoint source and receiver wavefields depicted in Figures 11(b) and 13(b), thus leading to the gradients seen in Figures 11(d) and 13(d), respectively.

We can conclude that gradient artifacts due to cross-talk between different paths present in the reconstructed wavefields are significantly reduced if we measure wavefield similarity by cross-correlation instead of differences. This reduction is due to the fewer independent branches present in the wavefields. However, when wavefields are characterized by significant multi-pathing, some non-physical artifacts still remain in the gradient even if we reduce the total number of branches to a minimum. This is because each path in the source wavefield interacts with each branch in the receiver wavefield. Further artifact reduction requires that we exploit other wavefield features, for example their directionality, as discussed in the following section.

3.3 Oriented correlation

Using the same source and receiver wavefields (u_s and u_r), Figures 2(a) and 2(b), we can define another data residual

$$r_d\left(e,\mathbf{x},\boldsymbol{\lambda},\tau\right)=\mathbf{W}_u\left(e,\mathbf{x},\boldsymbol{\lambda},\tau\right)c\left(e,\mathbf{x},\boldsymbol{\lambda},\tau\right) \tag{32}$$

based on their oriented cross-correlation (Appendix B)

$$c\left(e,\mathbf{x},\boldsymbol{\lambda},\tau\right)=u_s\left(e,\mathbf{x}-\frac{\boldsymbol{\lambda}}{2},t\right)\star u_r\left(e,\mathbf{x}+\frac{\boldsymbol{\lambda}}{2},t\right)\ , \tag{33}$$

and form an objective function

$$\mathcal{J}_u=\sum_e\frac{1}{2}\left\|r_d\left(e,\mathbf{x},\boldsymbol{\lambda},\tau\right)\right\|^2_{\mathbf{x},\boldsymbol{\lambda},\tau}\ , \tag{34}$$

where the weight $\mathbf{W}_u$ is a combination of an operator $\mathbf{K_r}$ restricting the source and receiver wavefields to the known receiver position, and another operator $\mathbf{P}\left(\boldsymbol{\lambda},\tau\right)$ weighing the data as a function of space-lag $\boldsymbol{\lambda}$ and cross-correlation time-lag τ:

$$\mathbf{W}_u\left(e,\mathbf{x},\boldsymbol{\lambda},\tau\right)=\mathbf{K_r}\left(e,\mathbf{x}\right)\mathbf{P}\left(\boldsymbol{\lambda},\tau\right)\ . \tag{35}$$

As in the preceding case, the weight $\mathbf{P}$ is essential in the definition of the data residual. Its role is to nullify the oriented cross-correlation when the wavefields are identical, i.e. when the correlation function is maximum at zero time-lag.

Using objective function 34, the adjoint sources are

$$\begin{bmatrix}g_s\left(e,\mathbf{x}-\frac{\boldsymbol{\lambda}}{2},t\right)\\ g_r\left(e,\mathbf{x}+\frac{\boldsymbol{\lambda}}{2},t\right)\end{bmatrix}=\begin{bmatrix}u_r\left(e,\mathbf{x}+\frac{\boldsymbol{\lambda}}{2},t\right)\\ u_s\left(e,\mathbf{x}-\frac{\boldsymbol{\lambda}}{2},t\right)\end{bmatrix}*\left[\mathbf{W}_u^\top\left(e,\mathbf{x},\boldsymbol{\lambda},\tau\right)r_d\left(e,\mathbf{x},\boldsymbol{\lambda},\tau\right)\right] \tag{36}$$

where the symbol $*$ indicates oriented convolution (i.e. the adjoint of oriented correlation) between the receiver and source wavefields, and the weighted residual. As indicated earlier, the adjoint variables (a_s and a_r) are wavefields extrapolated using the given wave-equation backward and forward in time, respectively. The major difference between the adjoint sources in equations 31 and 36 is that in the later case the sources are oriented according to the direction of the incoming waves, while in the former they are not. Therefore, the oriented cross-correlation enables us to beam energy in the direction of the incoming wavefield, as opposed to sending waves in all possible directions, thus further reducing the possibility of cross-talk between different wave paths present in the wavefield.

For example, the adjoint source wavefields for the example in Figure 1(a) are in Figure 3(b) for the simple correlation and in Figure 4(a) for the oriented correlation. The energy from the source obtained by the simple correlation propagates in all directions, while the energy for the source obtained by oriented correlation propagates in a specific direction, matching the direction of the incoming waves. The gradient of the oriented correlation, Figure 4(c), is correspondingly cleaner of artifacts compared with the gradient of the simple correlation, Figure 3(d).

The example in Figure 5(a) demonstrates this property even better. In this case, we are seeing waves beamed back from the adjoint source with different directions, corresponding to the different paths captured by the source wavefields, Figure 7(b) and Figure 8(a). Not surprisingly, the gradient for the oriented cross-correlation, Figure 8(c) is cleaner than the gradient of the simple cross-correlation, Figure 7(d).

These conclusions remain true if the analysis point is located away from either the source or the receiver, Figure 9(a). The oriented cross-correlation adjoint source and receiver wavefields, Figures 12(a) and 14(a), beam energy in specific directions toward the source and receivers, while the simple cross-correlation adjoint source and receiver wavefields, Figures 11(b) and 13(b), send energy in all directions.

A natural question to ask is whether we could more simply capture the directionality of the wavefield using just the zero time-lag of the correlation. This implies that we could attempt to formulate the wavefield tomography problem using the slice at $\tau=0$ from the oriented cross-correlation de-

picted in Figure 1(e). This is a special case of the oriented cross-correlation

$$c\left(e,\boldsymbol{\lambda}\right)=c\left(e,\boldsymbol{\lambda},\tau=0\right)\ ,$$

as is also the case of the simple cross-correlation:

$$c\left(e,\tau\right)=c\left(e,\boldsymbol{\lambda}=\mathbf{0},\tau\right)\ .$$

For the zero time-lag oriented cross-correlation, we can define the residual

$$r_d\left(e,\mathbf{x},\boldsymbol{\lambda}\right)=\mathbf{W}_u\left(e,\mathbf{x},\boldsymbol{\lambda}\right)\sum_{\tau}\delta\left(\tau\right)c\left(e,\mathbf{x},\boldsymbol{\lambda},\tau\right)\ , \quad (37)$$

from which we can formulate an objective function using the expression

$$\mathcal{J}_u=\sum_e\frac{1}{2}\left\|r_d\left(e,\mathbf{x},\boldsymbol{\lambda}\right)\right\|^2_{\mathbf{x},\boldsymbol{\lambda}}\ . \quad (38)$$

The summation over τ after multiplication with $\delta\left(\tau\right)$ indicates that we take the zero-time lag of the cross-correlation. The weight is

$$\mathbf{W}_u\left(e,\mathbf{x},\boldsymbol{\lambda}\right)=\mathbf{K}_{\mathbf{r}}\left(e,\mathbf{x}\right)\mathbf{P}\left(\boldsymbol{\lambda}\right)\ , \quad (39)$$

indicating that, as before, we restrict the cross-correlation at the receiver position and we also penalize it as a function of space-lag using the operator $\mathbf{P}$. Then, the source and receiver adjoint sources are

$$\begin{bmatrix} g_s\left(e,\mathbf{x}-\frac{\boldsymbol{\lambda}}{2},t\right)\\ g_r\left(e,\mathbf{x}+\frac{\boldsymbol{\lambda}}{2},t\right)\end{bmatrix}=\begin{bmatrix} u_r\left(e,\mathbf{x}+\frac{\boldsymbol{\lambda}}{2},t\right)\\ u_s\left(e,\mathbf{x}-\frac{\boldsymbol{\lambda}}{2},t\right)\end{bmatrix}\left[\mathbf{W}_u^{\top}\left(e,\mathbf{x},\boldsymbol{\lambda},\tau\right)r_d\left(e,\mathbf{x},\boldsymbol{\lambda},\tau\right)\right]\ . \quad (40)$$

These expressions resemble the ones for the full oriented cross-correlation, except for the fact that the wavefields are multiplied with the residual, instead of convolved, as is the case in equation 36.

The adjoint source wavefield for the zero time-lag oriented cross-correlation is capable to capture partially the directionality of the wavefields, as seen by comparing Figure 4(b) with Figure 4(a). However, this is not always the case. The limited extent of the space-lag domain may not be sufficient to capture the correlation for events that require a longer space-lag axis. This is visible with the oriented cross-correlation from Figure 5(e). The zero time-lag captures (partially) one of the arrivals at the receiver, but not others. As a consequence, the adjoint source wavefield does not represent correctly the reflected path, Figure 8(b), thus leading to a distorted gradient, Figure 8(d). This drawback is even more striking in the example from Figure 9(a). The adjoint source and receiver wavefields, Figures 12(b) and 14(b), are essentially amplified noise and do not represent correct propagation paths, thus producing distorted gradients, Figures 12(d) and 14(d).

3.4 Summary and discussion

The derivations in this section lead to the following conclusions:

- The wavefield cross-correlation reduces the number of gradient artifacts in comparison with the wavefield difference. This is due to the reduced number of non-physical events present in the adjoint sources, and therefore in the adjoint wavefields.
- The oriented cross-correlation further reduces the number of gradient artifacts by restricting wavefield propagation in the direction from where waves are observed at the analysis point.
- Zero time-lag oriented cross-correlation can also capture wavefield directionality. However, this kind of analysis requires a larger space-lag domain, compared with the full oriented cross-correlation. In this case, we are trading time-lags (more) for space-lags (none), thus increasing the potential to interfere with events propagating in different directions.

It appears that the oriented cross-correlation provides significant advantages over the other measures of wavefield similarity considered here. This does not mean, however, that other wavefield comparison methods cannot provide other, perhaps more significant, advantages.

In all situations considered so far, I have assumed that we know where we should compare the source and receiver wavefields. Common candidates for this analysis are the source and receiver positions, as well as some interfaces with known geometry. This analysis also allows for arbitrary source functions, e.g. point sources, plane sources, randomly-encoded sources, etc.

More generally, we can ask the following question: where can we conduct a similar analysis when we do not know a-priori any point where the wavefields should match? Such analysis implies that we use additional information, as discussed next.

4 IMAGE-DOMAIN WAVEFIELD TOMOGRAPHY

The methods described in the preceding section assume that we know one or more locations where the source and receiver wavefields should match. Under this assumption, one can formulate objective functions exploiting different forms of comparison between multi-dimensional wavefields. We can construct such objective functions for each experiment separately, and then sum them all together, as is the case in Equations 24, 29, 34 and 38.

When the analysis position is unknown, we need to define the objective function by relying on the so-called **semblance principle**. A convenient formulation of this principle in the context of this paper is the following: *the model used for wavefield reconstruction is accurate if the source and receiver wavefields match each-other for all experiments, at the same locations*.

What this means in practice is that we neither define an objective function based on the wavefields themselves, nor based on the cross-correlation between the source and receiver

wavefields of individual experiments, as is the case with all formulations described in the preceding section. Using the semblance principle, we define an objective function based on a combination of the cross-correlations between source and receiver wavefields for all experiments representing a seismic survey. In other words, we stack all wavefield cross-correlations before we define an objective function,

$$\mathcal{J} = \mathbf{F}\left[\sum_e c\left(u_s\left(e\right), u_r\left(e\right)\right)\right] \ , \tag{41}$$

instead of stacking the objective functions defined on the cross-correlation of source and receiver wavefields of individual experiments:

$$\mathcal{J} = \sum_e \mathbf{F}\left[c\left(u_s\left(e\right), u_r\left(e\right)\right)\right] \ . \tag{42}$$

Here, $\mathbf{F}$ is just a notation indicating that some operator acts on individual or combinations of cross-correlations c of individual wavefields, $u_s\left(e\right)$ and $u_r\left(e\right)$. Stacking is, of course, just one possibility; one could use other strategies to combine various images, e.g. image warping. We can exploit this general idea in several ways, as discussed next.

4.1 Stack power

Starting with the correlation between the source and receiver wavefields (u_s and u_r) at all locations in space

$$c\left(e, \mathbf{x}, \tau\right) = u_s\left(e, \mathbf{x}, t\right) \star u_r\left(e, \mathbf{x}, t\right) \ , \tag{43}$$

we can define **the image** as the stack over experiments of the zero time-lag cross-correlation:

$$i\left(\mathbf{x}\right) = \sum_e \sum_\tau \delta\left(\tau\right) c\left(e, \mathbf{x}, \tau\right) \ . \tag{44}$$

The multiplication by $\delta\left(\tau\right)$ followed by summation over τ indicates that we window the zero cross-correlation time-lag. Using this image obtained through a process referred-to as a conventional imaging condition, we can define a residual

$$r_d\left(\mathbf{x}\right) = \mathbf{W}_u\left(\mathbf{x}\right) i\left(\mathbf{x}\right) \ , \tag{45}$$

and then an objective function

$$\mathcal{J}_u = \frac{1}{2}\left\|r_d\left(\mathbf{x}\right)\right\|^2_{\mathbf{x}} \tag{46}$$

maximizing the image strength, also known as the stack power. The weight $\mathbf{W}_u$ has the purpose of rebalancing the strength of the image, for example in areas of poor illumination, and is only a function of position in space:

$$\mathbf{W}_u\left(\mathbf{x}\right) = \mathbf{K}\left(\mathbf{x}\right) \ . \tag{47}$$

The objective function 46 is based on exactly the same wavefields as the ones exploited in the earlier sections. The fundamental difference is that the summation over experiments occurs inside the norm, as opposed to the preceding cases when the summation occurs outside the norm.

The source and receiver adjoint sources are

$$\begin{bmatrix} g_s\left(e, \mathbf{x}, t\right) \\ \\ g_r\left(e, \mathbf{x}, t\right) \end{bmatrix} = \begin{bmatrix} u_r\left(e, \mathbf{x}, t\right) \\ \\ u_s\left(e, \mathbf{x}, t\right) \end{bmatrix} \left[\mathbf{W}_u^\top\left(\mathbf{x}\right) r_d\left(\mathbf{x}\right)\right] \ , \tag{48}$$

and the adjoint source and receiver wavefields are obtained as usually by extrapolation backward and forward in time, respectively, Figures 18(a) and 20(a). Using the state and adjoint variables, we can obtain the gradient of the objective function for a specific source/receiver pair, Figures 18(c) and 20(c).

Since this objective function 46 exploits just a portion of the cross-correlation function comparing the source and receiver wavefields, we can ask whether we could not extract different or additional information by using more of the cross-correlation function?

4.2 Time-lag images

Using the temporal cross-correlation between the source and receiver wavefields,

$$c\left(e, \mathbf{x}, \tau\right) = u_s\left(e, \mathbf{x}, t\right) \star u_r\left(e, \mathbf{x}, t\right) \ , \tag{49}$$

we can define an alternative type of image if we preserve the entire correlation function in the output. In this case, we construct so-called **time-lag images** by summing this cross-correlation over experiments

$$i\left(\mathbf{x}, \tau\right) = \sum_e c\left(e, \mathbf{x}, \tau\right) \ , \tag{50}$$

defined at all positions in space. Using such time-lag images, we can construct the residual

$$r_d\left(\mathbf{x}, \tau\right) = \mathbf{W}_u\left(\mathbf{x}, \tau\right) i\left(\mathbf{x}, \tau\right) \ , \tag{51}$$

and the objective function

$$\mathcal{J}_u = \frac{1}{2}\left\|r_d\left(\mathbf{x}, \tau\right)\right\|^2_{\mathbf{x},\tau} \ . \tag{52}$$

The weight

$$\mathbf{W}_u\left(\mathbf{x}, \tau\right) = \mathbf{K}\left(\mathbf{x}\right) \mathbf{P}\left(\tau\right) \tag{53}$$

has the purpose of restricting the analysis to some image points through the mask operator $\mathbf{K}$, and also to apply a penalty to the correlation function through the operator $\mathbf{P}$. We expect that there will be points in space where the cross-correlations for all experiments maximize at zero time-lag, and therefore so will their stack over experiments. Figure 16(a) shows an example of time-lag gather for the example described in Figure 15(a). Since the velocity is not correct, the maximum of the cross-correlation stack does not occur at zero time-lag or at the correct depth of the interface. We can use this information to formulate an objective function.

Using objective function 52, we can define the adjoint sources

$$\begin{bmatrix} g_s\left(e, \mathbf{x}, t\right) \\ \\ g_r\left(e, \mathbf{x}, t\right) \end{bmatrix} = \begin{bmatrix} u_r\left(e, \mathbf{x}, t\right) \\ \\ u_s\left(e, \mathbf{x}, t\right) \end{bmatrix} * \left[\mathbf{W}_u^\top\left(\mathbf{x}, \tau\right) r_d\left(\mathbf{x}, \tau\right)\right] \ , \tag{54}$$

which enable us to construct the adjoint source and receiver wavefields, Figures 18(b) and 20(b). The major difference with the stack power case is that in this case the source and receiver wavefields are convolved with the time-lag image stack. Figures 18(d) and 20(d) show the gradient computed by correlating the source and receiver state and adjoint variables, respectively. Not surprisingly, the gradient update originates at the depth of the strongest correlation stack, which is different from the true depth of the interface.

As in the preceding section, we can also ask the question whether we can capture directional information in the images, and then beam this energy back in the appropriate direction after the appropriate penalty has been applied. I discuss this case next.

4.3 Space-lag images

An approach resembling the oriented correlation discussed in the preceding section leads to so-called **space-lag images**. In this case, we exploit the oriented correlation between the source and receiver wavefields

$$c\left(e,\mathbf{x},\boldsymbol{\lambda},\tau\right)=u_s\left(e,\mathbf{x}-\frac{\boldsymbol{\lambda}}{2},t\right)\star u_r\left(e,\mathbf{x}+\frac{\boldsymbol{\lambda}}{2},t\right)\ , \tag{55}$$

and form an image by stacking its zero time-lag over experiments:

$$i\left(\mathbf{x},\boldsymbol{\lambda}\right)=\sum_e\sum_\tau\delta\left(\tau\right)c\left(e,\mathbf{x},\boldsymbol{\lambda},\tau\right)\ . \tag{56}$$

As suggested by the semblance principle, the zero space-lag of each individual cross-correlation must be located at the same point, if the model used for wavefield reconstruction is accurate. This is exactly the same idea used for the stack power or for the time-lag images. What is different in this case is that we construct oriented correlations capturing the propagation direction of the source and receiver wavefields, so we can use the formed image for other purposes, for example to decompose them as a function of reflection angles. Since in the ideal case the space-lag images are focused at zero space-lag, we can write the residual

$$r_d\left(\mathbf{x},\boldsymbol{\lambda}\right)=\mathbf{W}_u\left(\mathbf{x},\boldsymbol{\lambda}\right)i\left(\mathbf{x},\boldsymbol{\lambda}\right)\ , \tag{57}$$

based on which we can write the objective function

$$\mathcal{J}_u=\frac{1}{2}\left\|r_d\left(\mathbf{x},\boldsymbol{\lambda}\right)\right\|^2_{\mathbf{x},\boldsymbol{\lambda}}\ . \tag{58}$$

As before, the weight

$$\mathbf{W}_u\left(\mathbf{x},\boldsymbol{\lambda}\right)=\mathbf{K}\left(\mathbf{x}\right)\mathbf{P}\left(\boldsymbol{\lambda}\right) \tag{59}$$

has the dual purpose of restricting the analysis to some image points through the mask operator $\mathbf{K}$, and to penalize the image by emphasizing its components that are not concentrated at zero space-lag through the penalty operator $\mathbf{P}$.

Finally, we can construct the adjoint sources

$$\begin{bmatrix}g_s\left(e,\mathbf{x}-\frac{\boldsymbol{\lambda}}{2},t\right)\\ g_r\left(e,\mathbf{x}+\frac{\boldsymbol{\lambda}}{2},t\right)\end{bmatrix}=\begin{bmatrix}u_r\left(e,\mathbf{x}+\frac{\boldsymbol{\lambda}}{2},t\right)\\ u_s\left(e,\mathbf{x}-\frac{\boldsymbol{\lambda}}{2},t\right)\end{bmatrix}\left[\mathbf{W}_u^\top\left(\mathbf{x},\boldsymbol{\lambda}\right)r_d\left(\mathbf{x},\boldsymbol{\lambda}\right)\right]\ , \tag{60}$$

which enable us to reconstruct the adjoint source and receiver wavefields, Figures 19(a) and 21(a). In this case, the wavefields are multiplied with the residual, and not convolved, as is the case in equation 54. Using the state and adjoint wavefields, we can generate the source and receiver gradients depicted in Figures 19(c) and 21(c).

The space-lag images build on the zero time-lag of the oriented cross-correlation between source and receiver wavefields. As discussed earlier, this captures the directionality of the wavefields and beam energy back from the image toward the corresponding source and receivers, but it requires that we span a fairly large space-lag domain. A more economical approach can be formulated based on complete oriented cross-correlation functions.

4.4 Extended images

We can consider the most general form of oriented wavefield cross-correlation

$$c\left(e,\mathbf{x},\boldsymbol{\lambda},\tau\right)=u_s\left(e,\mathbf{x}-\frac{\boldsymbol{\lambda}}{2},t\right)\star u_r\left(e,\mathbf{x}+\frac{\boldsymbol{\lambda}}{2},t\right)\ , \tag{61}$$

and define an extended image

$$i\left(\mathbf{x},\boldsymbol{\lambda},\tau\right)=\sum_e c\left(e,\mathbf{x},\boldsymbol{\lambda},\tau\right) \tag{62}$$

that depends both on space-lags $\boldsymbol{\lambda}$ and on time-lag τ. This type of image generalizes both the previously defined time-lag and space-lag images, Equations 50 and 56, as well as the conventional image, equation 44. Using these extended images, we can define the residual

$$r_d\left(\mathbf{x},\boldsymbol{\lambda},\tau\right)=\mathbf{W}_u\left(\mathbf{x},\boldsymbol{\lambda},\tau\right)i\left(\mathbf{x},\boldsymbol{\lambda},\tau\right)\ , \tag{63}$$

and the objective function

$$\mathcal{J}_u=\frac{1}{2}\left\|r_d\left(\mathbf{x},\boldsymbol{\lambda},\tau\right)\right\|^2_{\mathbf{x},\boldsymbol{\lambda},\tau}\ , \tag{64}$$

where the weight represents both a mask operator $\mathbf{K}$, restricting the analysis to some points in the image, as well as a penalty operator $\mathbf{P}$, exploiting the parts of the extended image based on correlations that do not align at the same position in the subsurface:

$$\mathbf{W}_u\left(\mathbf{x},\boldsymbol{\lambda},\tau\right)=\mathbf{K}_\mathbf{c}\left(\mathbf{x}\right)\mathbf{P}\left(\boldsymbol{\lambda},\tau\right)\ . \tag{65}$$

Unlike for the special case of space-lag images, here we can exploit the extended image at specific points identified in the subsurface, as indicated by subscript $\mathbf{c}$ in the mask operator, and still capture the directionality of the propagating wavefields. This approach also leads to computational savings,

since we only need to compute oriented cross-correlations at specific points in the subsurface, as opposed to everywhere in space. Using the extended images, we can formulate adjoint sources

$$\begin{bmatrix} g_s\left(e,\mathbf{x}-\dfrac{\boldsymbol{\lambda}}{2},t\right) \\ g_r\left(e,\mathbf{x}+\dfrac{\boldsymbol{\lambda}}{2},t\right) \end{bmatrix} = \begin{bmatrix} u_r\left(e,\mathbf{x}+\dfrac{\boldsymbol{\lambda}}{2},t\right) \\ u_s\left(e,\mathbf{x}-\dfrac{\boldsymbol{\lambda}}{2},t\right) \end{bmatrix} * \left[\mathbf{W}_u^\top\left(\mathbf{x},\boldsymbol{\lambda},\tau\right) r_d\left(\mathbf{x},\boldsymbol{\lambda},\tau\right)\right] , \tag{66}$$

where the symbol $*$ indicates convolution in time. Using the source and receiver adjoint sources, we can construct the adjoint source and receiver wavefields, Figures 19(b) and 21(b), and then the source and receiver gradients, Figures 19(d) and 21(d).

4.5 Summary and discussion

This section allows us to draw the following conclusions:

- We can exploit the semblance principle to define objective functions usable for wavefield tomography. The residual underlying the objective function is defined after combining, e.g. by stacking, the cross-correlations for all the experiments in a seismic survey.
- As was the case for the analysis at known positions in the subsurface, we can compare the wavefields in many different ways, e.g. using simple or oriented time cross-correlations. Oriented correlations potentially lead to cleaner gradients, and thus to more robust inversion.
- Oriented cross-correlations can be constructed at any location in the image, but a computationally efficient process restricts such correlations to some points in space identified by the strength of the zero time-lag correlation.
- Wavefield imaging and wavefield tomography are closely related to one-another, since both use the same wavefields, and the wavefield images generated in the context of the semblance principle are just pre-requisites to wavefield tomography.

Wavefield tomography formulated by assuming known positions where the source and receiver wavefields match, is often restricted to transmitted (diving) waves with analysis conducted at either the source or at the receiver positions. This often limits the depth of penetration and does not allow for model constraints deep in the subsurface. In contrast, wavefield analysis under the semblance principle enables us to probe the model at all depths, thus giving us more flexibility and better control on the model parameters deep in the model.

Key components of the analysis at arbitrary points in the subsurface are the penalty functions. These are designed to transform the images obtained after stacking over independent experiments into residuals that can be exploited in inversion. There are several possible definitions of the penalty functions:

- We can define penalties using analytic functions. These functions are designed to wipe-out ideal images, and transform everything else into a residual for optimization. Examples include the time-lag penalty $\mathbf{P}\left(\tau\right) = |\tau|$ and the space-lag penalty $\mathbf{P}\left(\boldsymbol{\lambda}\right) = |\boldsymbol{\lambda}|$. These common penalty functions assume idealized images (infinite frequency band, perfect illumination, etc), assumptions that are unlikely to be valid in practice.
- Alternatively, we can define penalties using the images themselves. We could simply evaluate through modeling how images constructed in a realistic frequency band behave under the influence of a velocity model and subject to partial illumination due to overburden geology or due to the acquisition geometry.

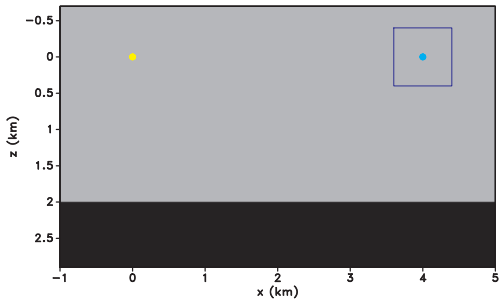
5 CONCLUSIONS

The words "wavefield tomography" describe a family of techniques that share one main feature: all use the same wavefields reconstructed using a given wave-equation from observed data. The data-domain methods (waveform inversion) and the image-domain methods (wave-equation migration velocity analysis) are thus exploiting the same wavefields, just in different ways and under different assumptions. The main difference between these methods is that in the first category we compare wavefields for each seismic experiment independently, while in the second category we compare the wavefields after summation over experiments, i.e. after we form an image. The seismic images are, therefore, just proxies for the underlying seismic wavefields and they provide a convenient mechanism for defining an objective function.

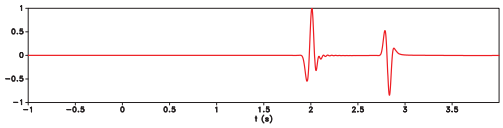
In all situations, we take advantage of similar wavefield comparison procedures, which include among others difference and (oriented) cross-correlations, although other comparison mechanisms are possible. By wavefield comparison, we formulate objective functions whose gradients reduce to the cross-correlation of state and adjoint wavefields reconstructed everywhere in the subsurface as solutions to a given wave-equation. In complex geology characterized by multipathing, one particular wavefield may contain multiple branches that should only correlate with specific branches from the matching wavefield. If this requirement is violated, the gradients are corrupted with non-physical events that can distort the model updates and hamper the optimization. Wavefield comparison by differencing is particularly prone to such artifacts; oriented cross-correlations limit artifact occurrence by directing waves in specific directions.

6 ACKNOWLEDGMENTS

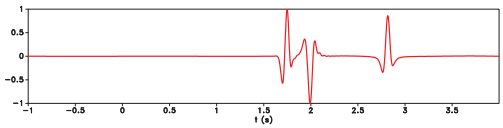
Partial funding for this research is provided by KACST. The reproducible numeric examples in this paper use the Madagascar open-source software package (Fomel et al., 2013) freely available from http://www.ahay.org. This research was supported in part by the Golden Energy Computing Organization at the Colorado School of Mines using resources acquired with financial assistance from the National Science Foundation and the National Renewable Energy Laboratory.



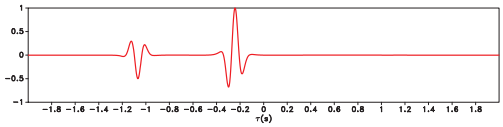
(a)



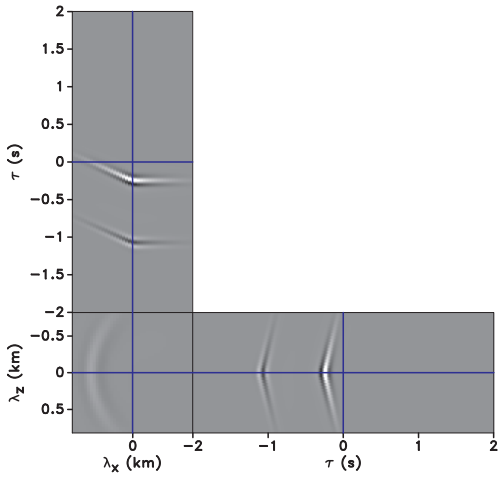
(b)



(c)

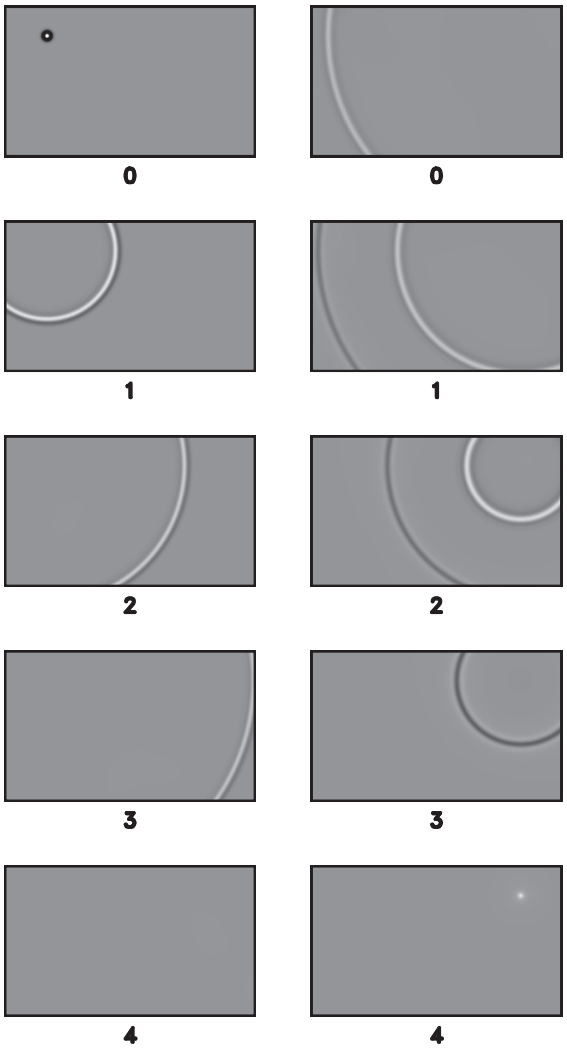


(d)



(e)

Figure 1. Example 1: (a) Velocity model, (b) observed data, (c) data difference, (d) data cross-correlation, and (e) oriented wavefield correlation at the analysis point indicated by the box in panel (a).



(a) (b)

Figure 2. Example 1: (a) The source wavefield u_s and (b) the receiver wavefield u_r, ordered by time as indicated by the panel index.

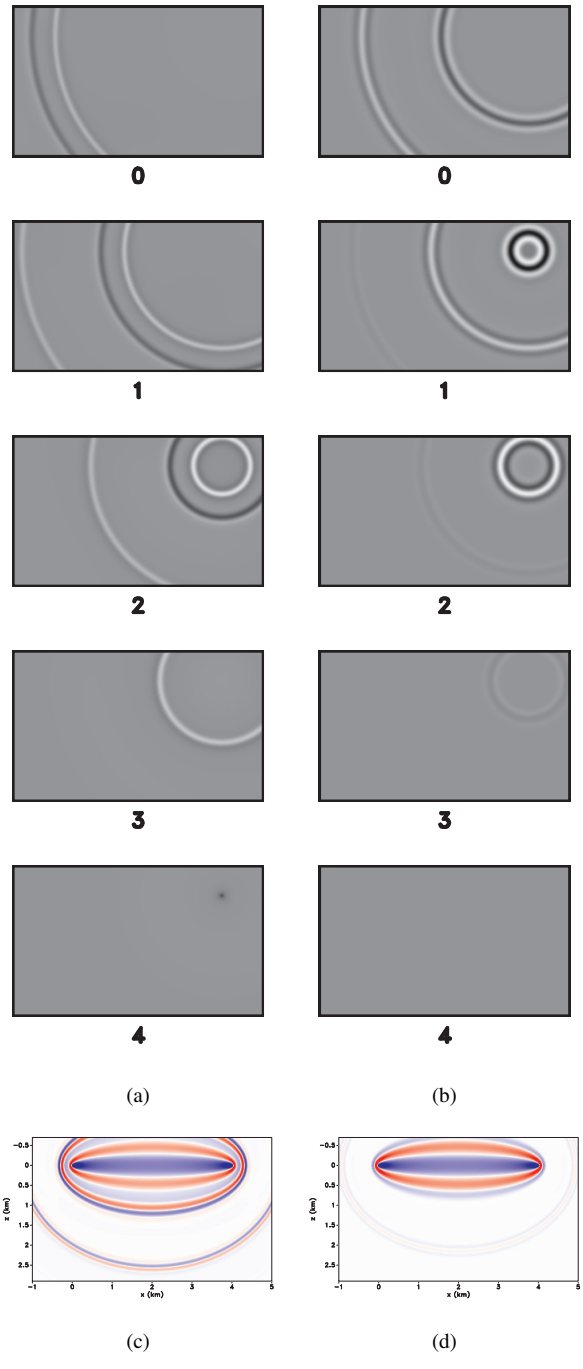


Figure 3. Example 1: (a)-(b) adjoint source wavefields and (c)-(d) and gradients for objective functions defined using difference and simple time correlation, respectively

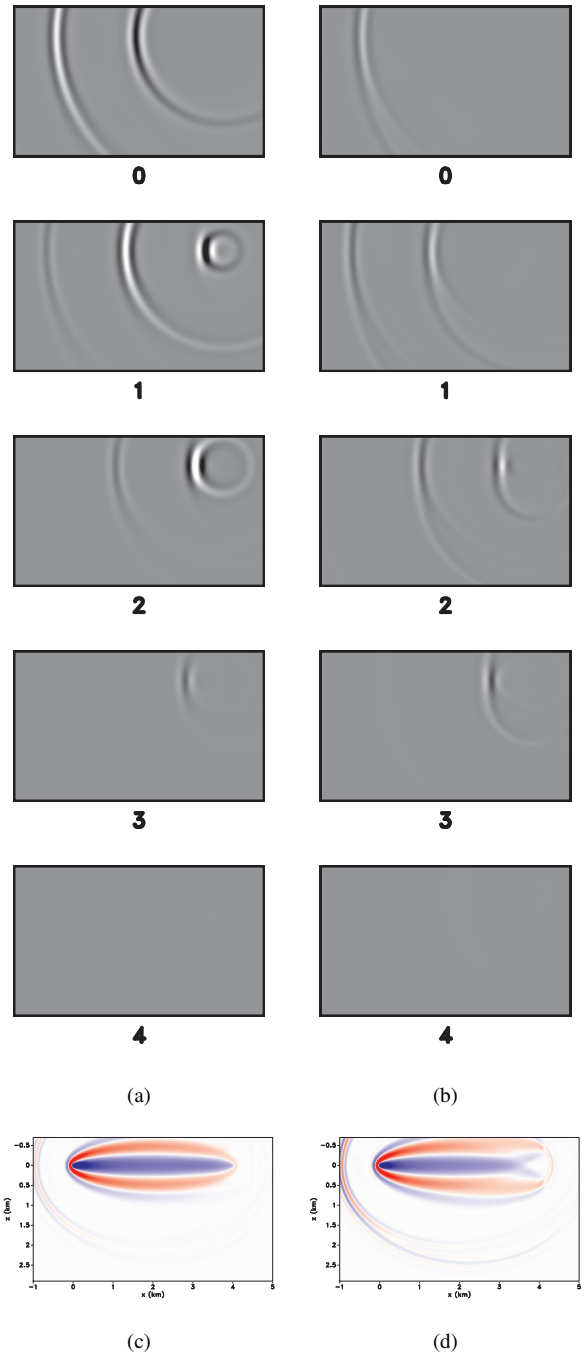
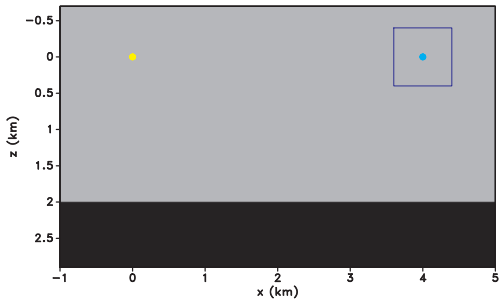
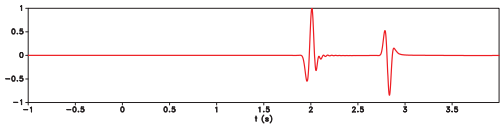


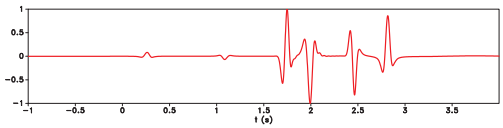
Figure 4. Example 1: (a)-(b) adjoint source wavefields and (c)-(d) and gradients for objective functions defined using oriented correlations and zero time-lag oriented correlations, respectively



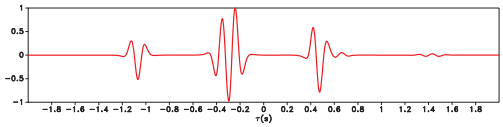
(a)



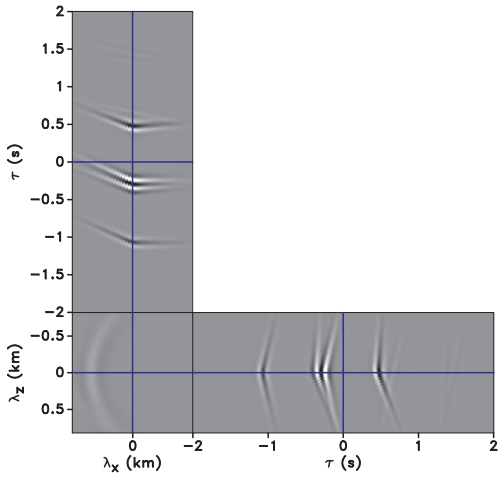
(b)



(c)

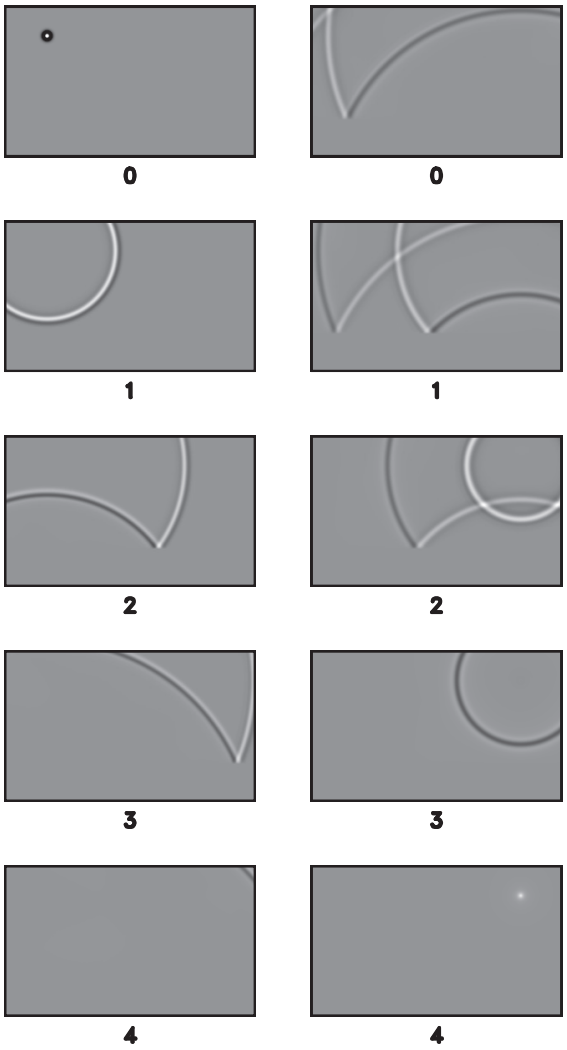


(d)



(e)

Figure 5. Example 2: (a) Velocity model, (b) observed data, (c) data difference, (d) data cross-correlation, and (e) oriented wavefield correlation at the analysis point indicated by the box in panel (a).



(a) (b)

Figure 6. Example 2: (a) The source wavefield u_s and (b) the receiver wavefield u_r, ordered by time as indicated by the panel index.

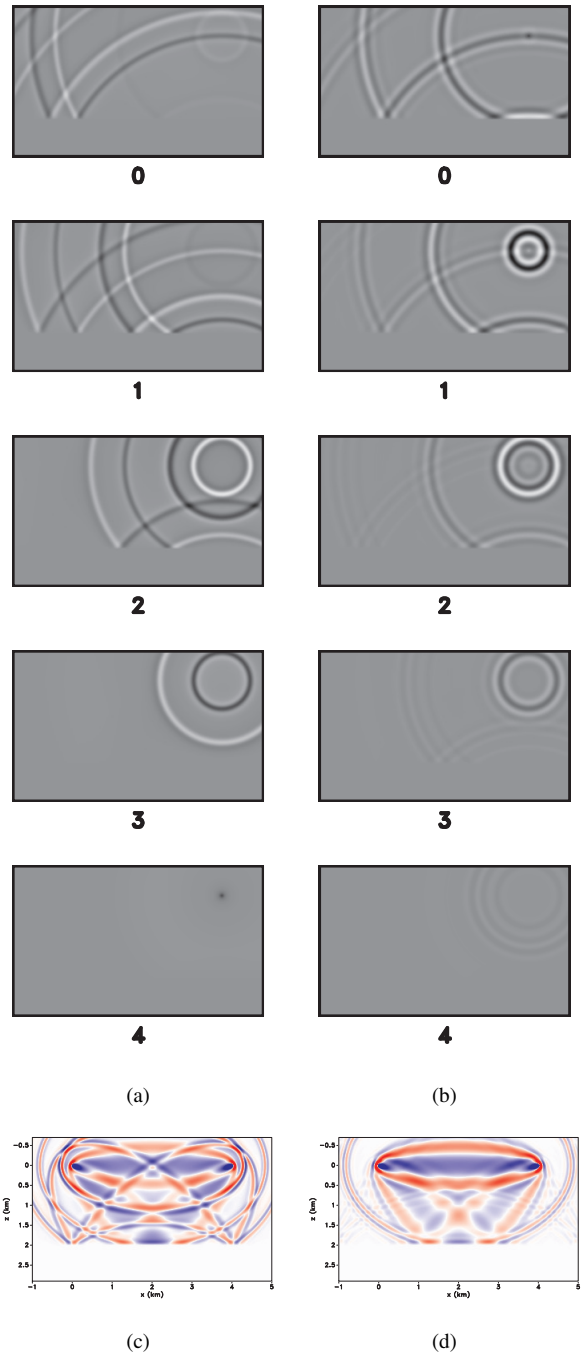


Figure 7. Example 2: (a)-(b) adjoint source wavefields and (c)-(d) and gradients for objective functions defined using difference and simple time correlation, respectively

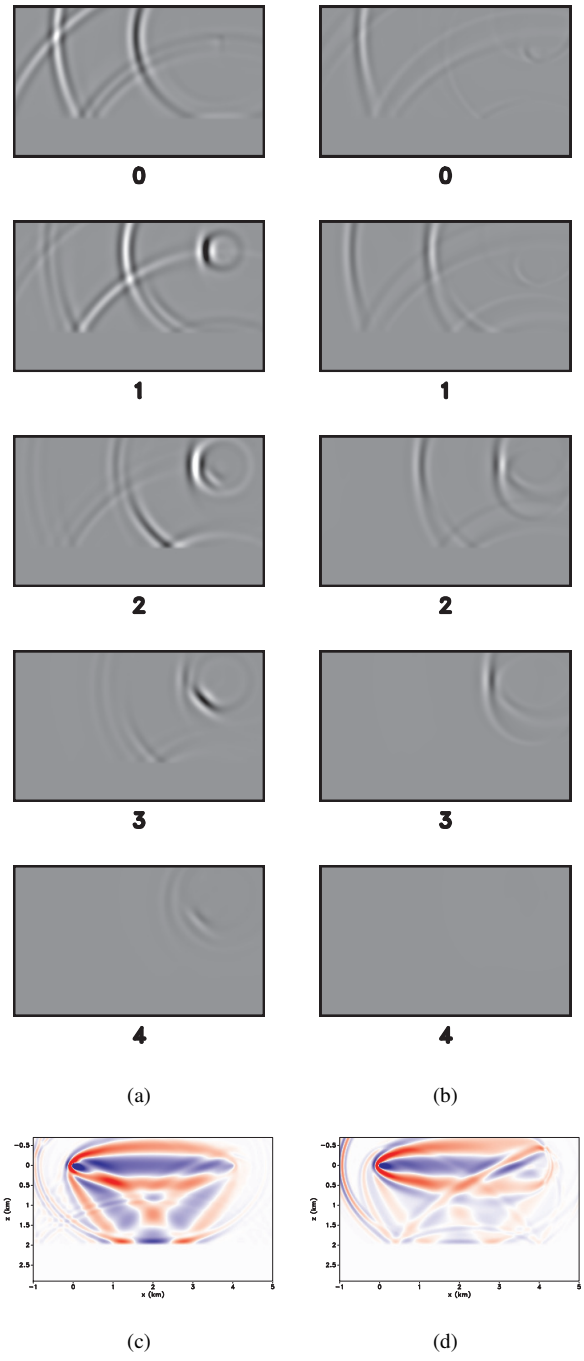
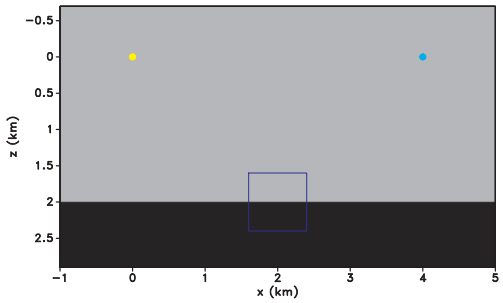
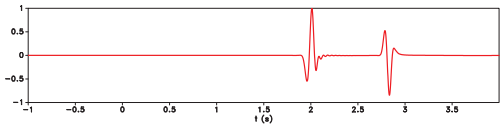


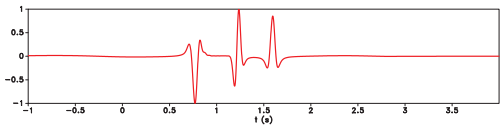
Figure 8. Example 2: (a)-(b) adjoint source wavefields and (c)-(d) and gradients for objective functions defined using oriented correlations and zero time-lag oriented correlations, respectively



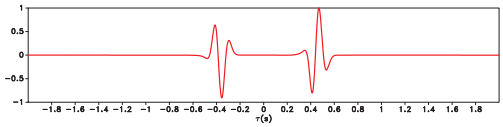
(a)



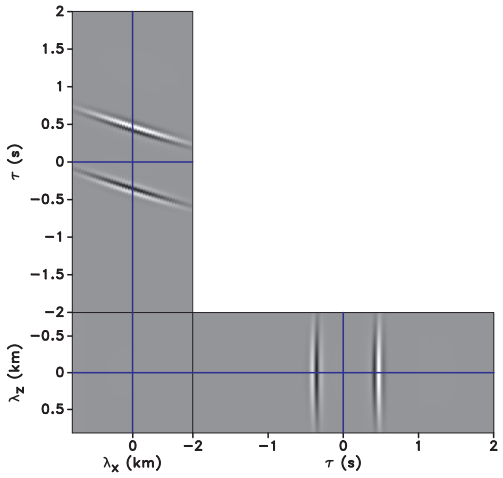
(b)



(c)

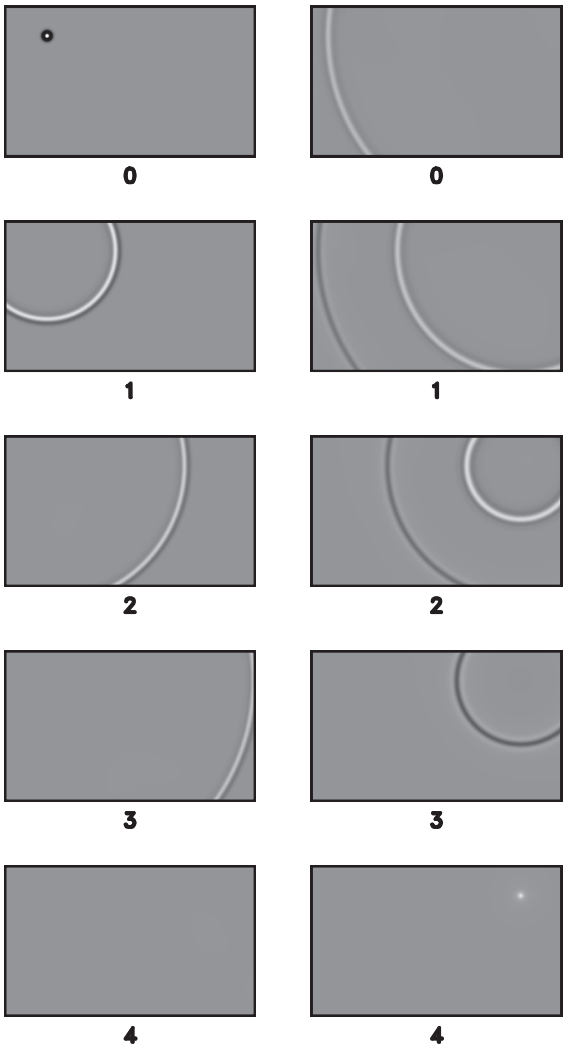


(d)



(e)

Figure 9. Example 3: (a) Velocity model, (b) observed data, (c) data difference, (d) data cross-correlation, and (e) oriented wavefield correlation at the analysis point indicated by the box in panel (a).



(a) (b)

Figure 10. Example 3: (a) The source wavefield u_s and (b) the receiver wavefield u_r, ordered by time as indicated by the panel index.

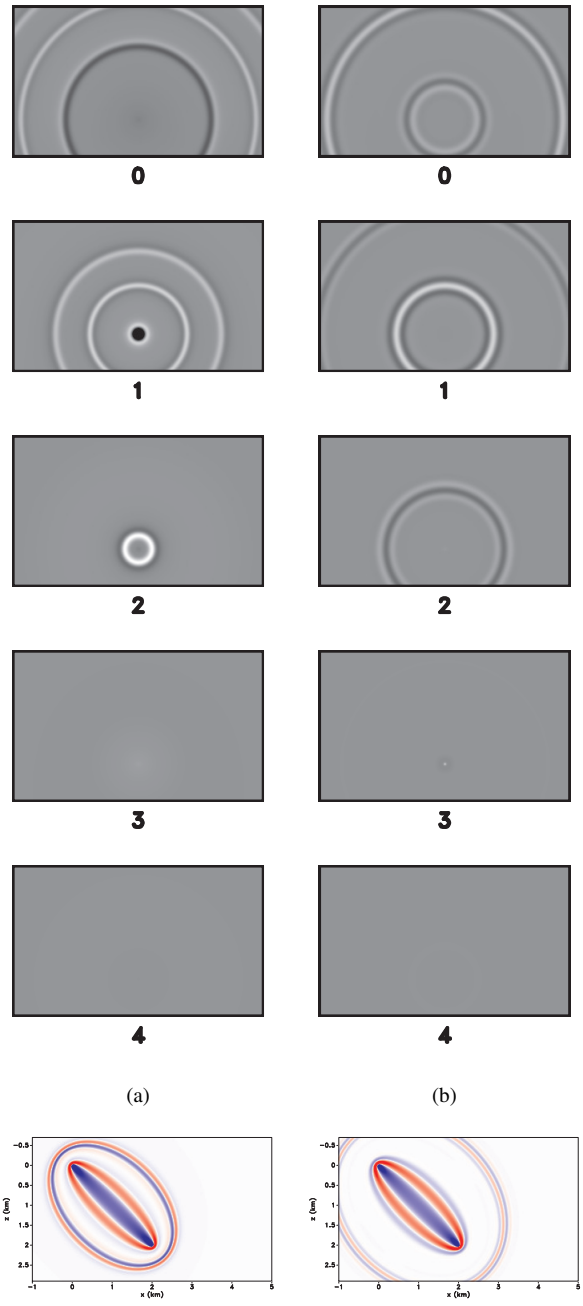


Figure 11. Example 3: (a)-(b) adjoint source wavefields and (c)-(d) and gradients for objective functions defined using difference and simple time correlation, respectively

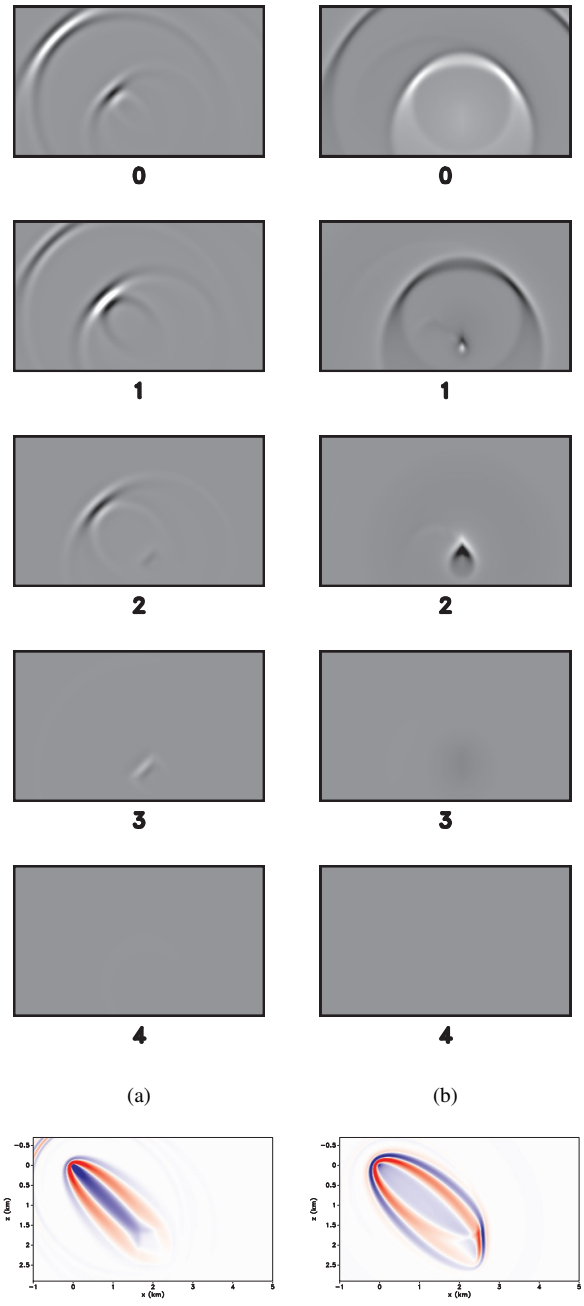


Figure 12. Example 3: (a)-(b) adjoint source wavefields and (c)-(d) and gradients for objective functions defined using oriented correlations and zero time-lag oriented correlations, respectively

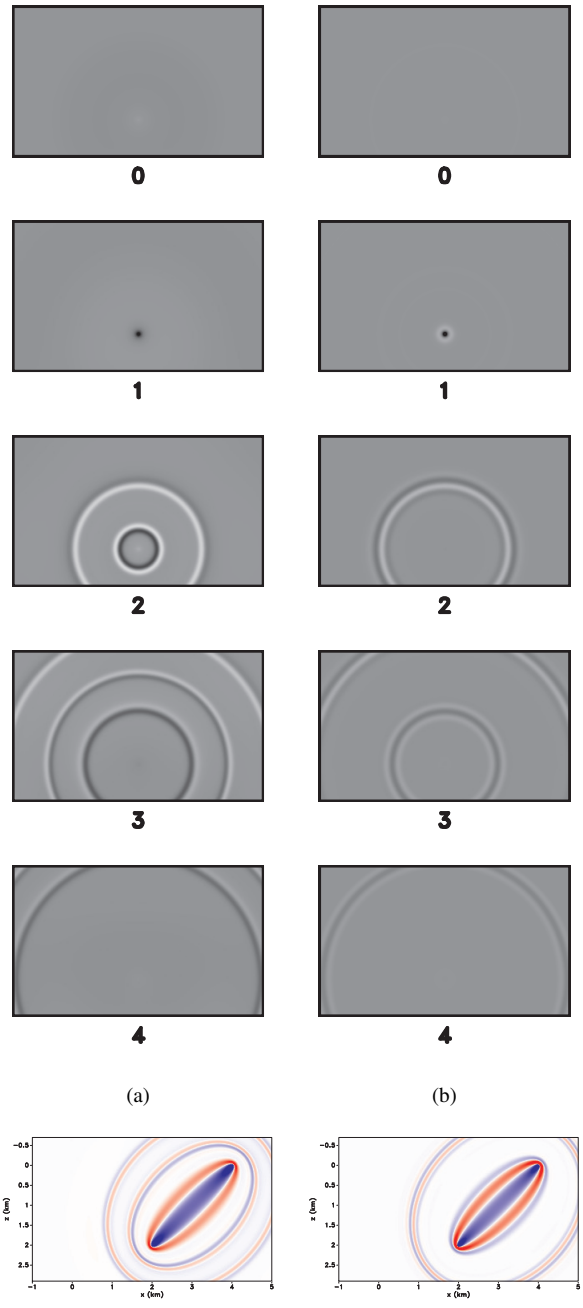


Figure 13. Example 3: (a)-(b) adjoint receiver wavefields and (c)-(d) and gradients for objective functions defined using difference and simple time correlation, respectively

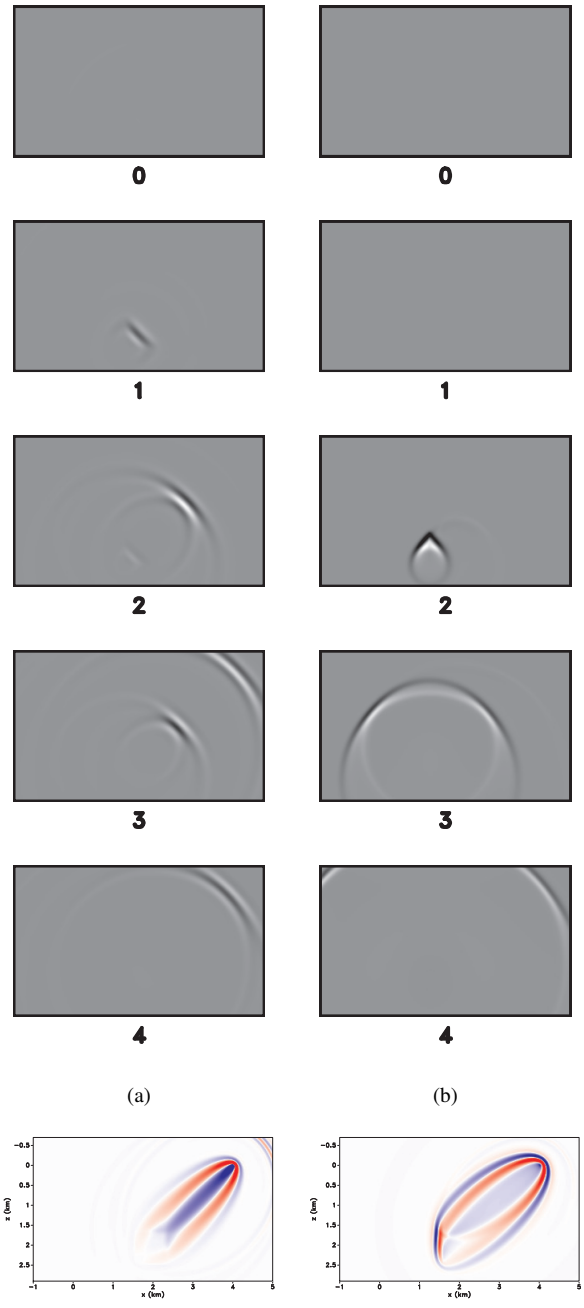
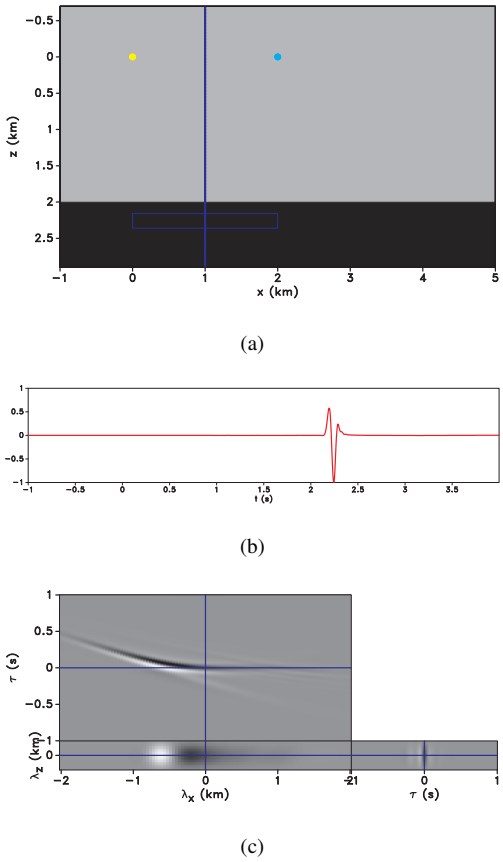


Figure 14. Example 3: (a)-(b) adjoint receiver wavefields and (c)-(d) and gradients for objective functions defined using oriented correlations and zero time-lag oriented correlations, respectively

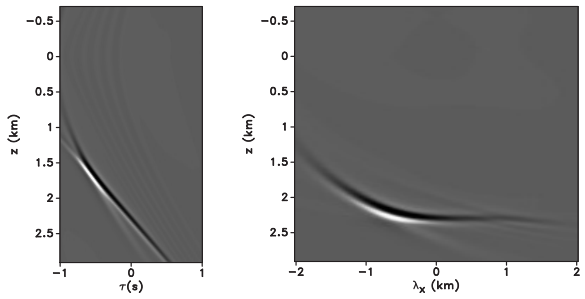


(a)

(b)

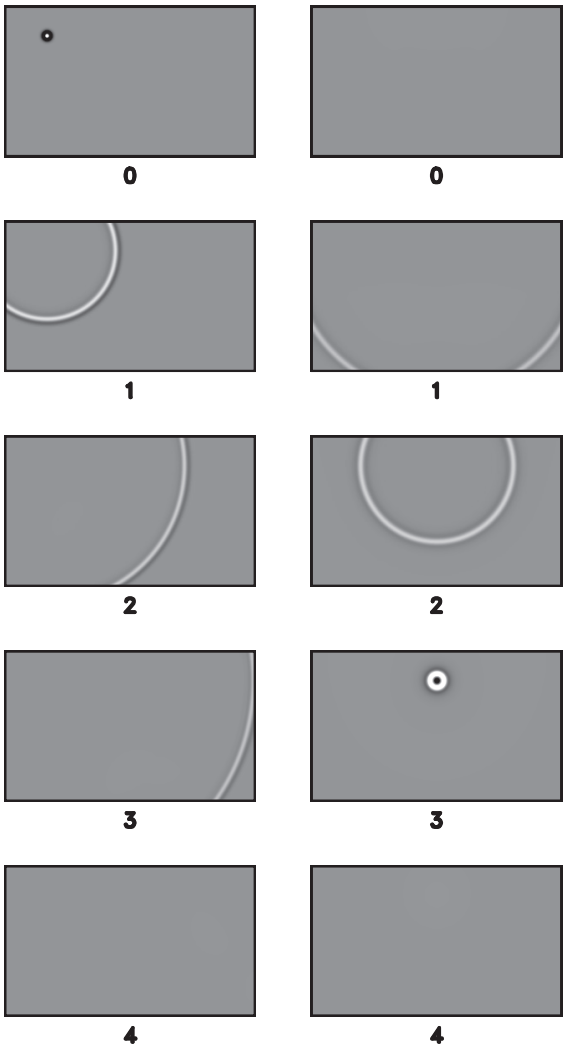
(c)

Figure 15. Example 4: (a) Velocity model, (b) observed data, and (c) oriented wavefield correlation at the analysis point indicated by the box in panel (a).



(a) (b)

Figure 16. Example 4: Common image gathers as a function of (a) experiment, (b) time-lag and (c) space-lag.



(a) (b)

Figure 17. Example 4: (a) The source wavefield u_s and (b) the receiver wavefield u_r, ordered by time as indicated by the panel index.

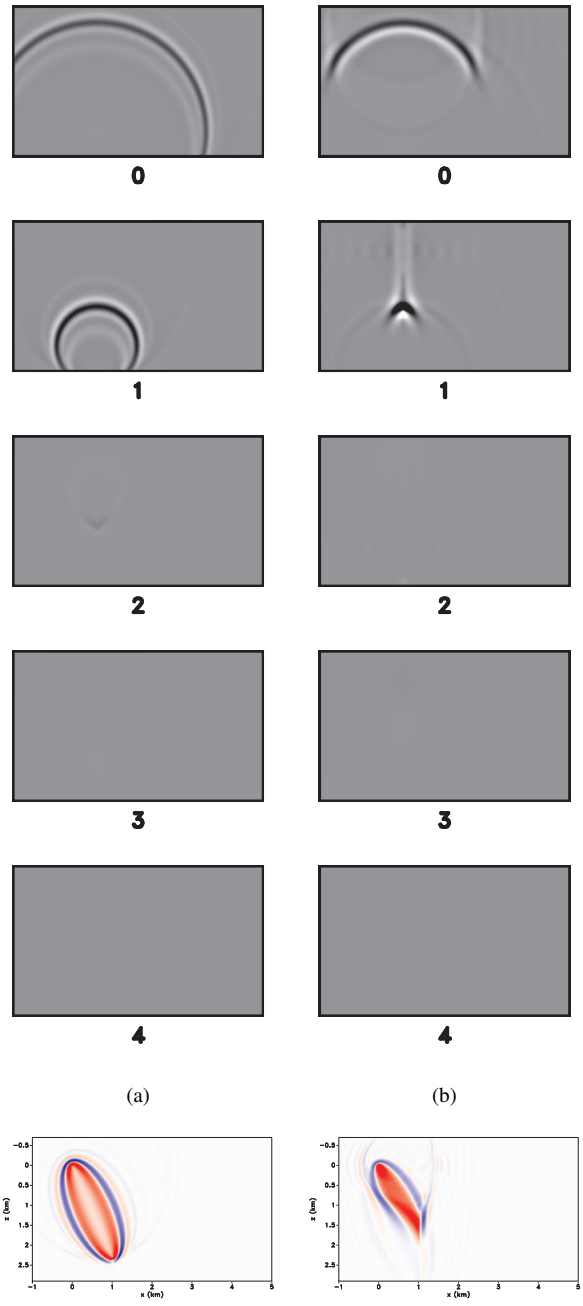


Figure 18. Example 4: (a)-(b) adjoint source wavefields and (c)-(d) and gradients for objective functions defined using shot-gathers and time-lag gathers, respectively

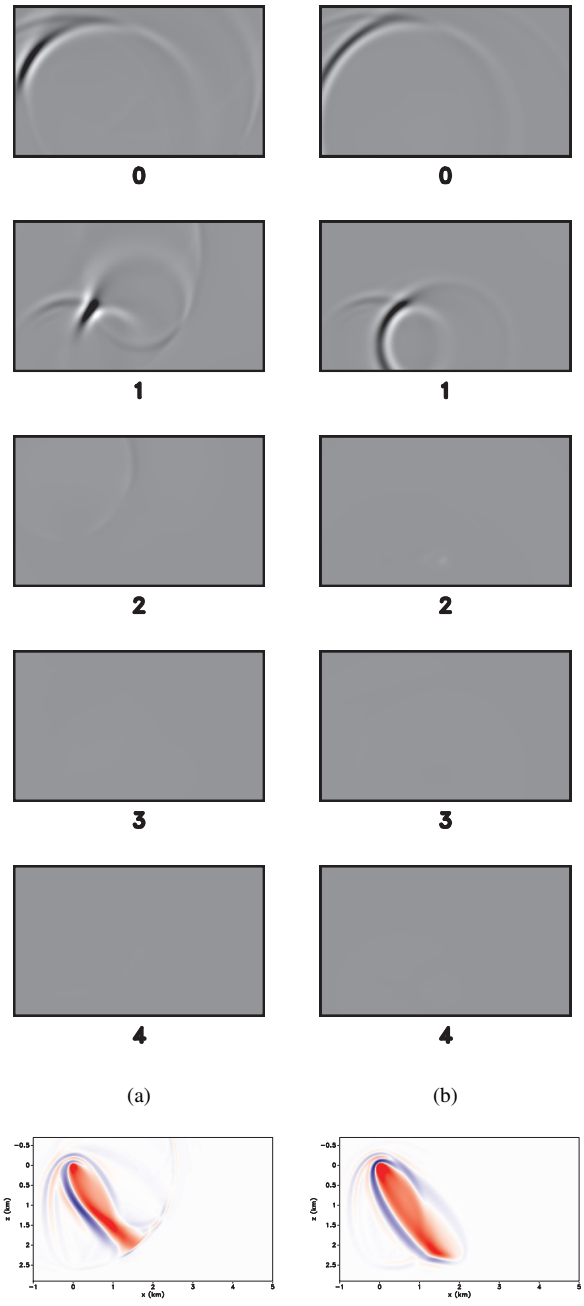


Figure 19. Example 4: (a)-(b) adjoint source wavefields and (c)-(d) and gradients for objective functions defined using zero time-lag oriented correlations and oriented correlations, respectively

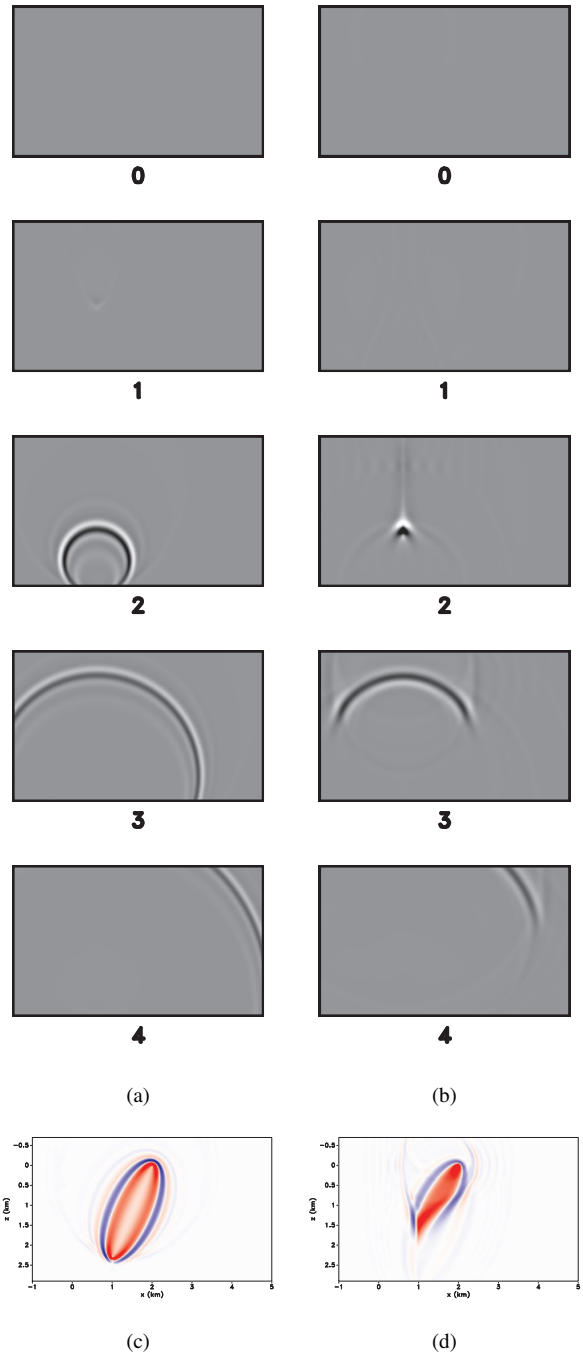


Figure 20. Example 4: (a)-(b) adjoint receiver wavefields and (c)-(d) and gradients for objective functions defined using shot-gathers and time-lag gathers, respectively

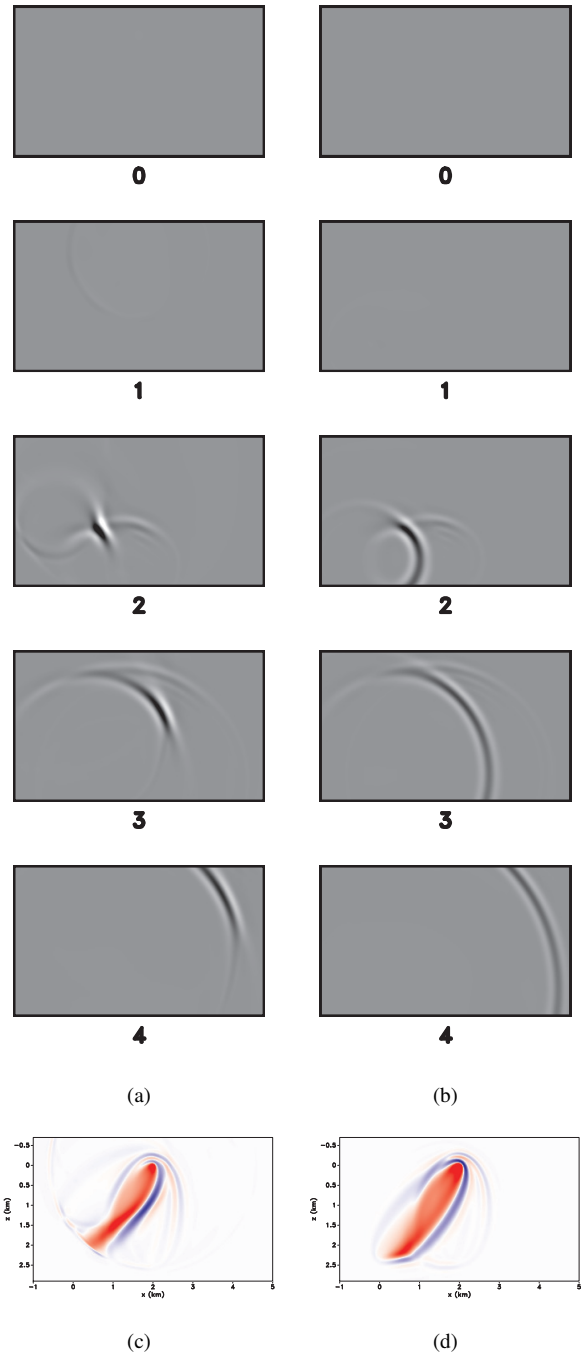


Figure 21. Example 4: (a)-(b) adjoint receiver wavefields and (c)-(d) and gradients for objective functions defined using zero time-lag oriented correlations and oriented correlations, respectively

REFERENCES

Albertin, U., P. Sava, J. Etgen, and M. Maharramov, 2006, Image differencing and focusing in wave-equation velocity analysis: Presented at the 68th Mtg., Abstracts, Eur. Assoc. Expl. Geophys.

Alkhalifah, T., and Y. Choi, 2014, From tomography to full-waveform inversion with a single objective function: Geophysics, **79**, R55–R61.

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 1514–1524.

Berkhout, A. J., 1982, Imaging of acoustic energy by wave field extrapolation: Elsevier.

Biondi, B., 2006, 3-D Seismic Imaging: Society of Exploration Geophysicists.

———, 2010, Velocity estimation by image-focusing analysis: Geophysics, **75**, U49–U60.

Biondi, B., and A. Almomin, 2012, Tomographic full waveform inversion (TFWI) by combining full waveform inversion with wave-equation migration velocity anaylisis: SEG Technical Program Expanded Abstracts, 1–5.

———, 2013, Tomographic full waveform inversion (TFWI) by extending the velocity model along the time-lag axis: SEG Technical Program Expanded Abstracts, 1031–1036.

Biondi, B., and W. Symes, 2004, Angle-domain common-image gathers for migration velocity analysis by wavefield-continuation imaging: Geophysics, **69**, 1283–1298.

Brossier, R., S. Operto, and J. Virieux, 2009, Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion: Geophysics, **74**, WCC105–WCC118.

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

Clærbout, J. F., 1985, Imaging the Earth's Interior: Blackwell Scientific Publications.

Crase, E., A. Pica, M. Noble, J. McDonald, and A. Tarantola, 1990, Robust elastic nonlinear waveform inversion: Application to real data: Geophysics, **55**, 527–538.

de Bruin, C. G. M., C. P. A. Wapenaar, and A. J. Berkhout, 1990, Angle-dependent reflectivity by means of prestack migration: Geophysics, **55**, 1223–1234.

Fomel, S., 2004, Theory of 3-D angle gathers in wave-equation imaging: 74th Annual International Meeting, SEG, Expanded Abstracts, Soc. of Expl. Geophys., 1053–1056.

Fomel, S., P. Sava, I. Vlad, Y. Liu, and V. Bashkardin, 2013, Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments: Journal of Open Research Software, **1**, e8.

Gauthier, O., J. Virieux, and A. Tarantola, 1986, Two-dimensional nonlinear inversion of seismic waveforms - Numerical results: Geophysics, **51**, 1387–1403.

Kosloff, D. D., and E. Baysal, 1983, Migration with the full acoustic wave equation: Geophysics, **48**, 677–687.

Maharramov, M., and U. Albertin, 2007, Localized image-difference wave-equation tomography: Presented at the 75th Annual International Meeting, SEG, Expanded Abstracts.

McMechan, G. A., 1983, Migration by extrapolation of time-dependent boundary values: Geophys. Prosp., **31**, 413–420.

Min, D.-J., and C. Shin, 2006, Refraction tomography using a waveform-inversion back-propagation technique: Geophysics, **71**, R21–R30.

Mora, P., 1988, Elastic wave-field inversion of reflection and transmission data: Geophysics, **53**, 750–759.

———, 1989, Inversion = migration + tomography: Geophysics, **54**, 1575–1586.

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

Mulder, W. A., and A. P. E. ten Kroode, 2002, Automatic velocity analysis by differential semblance optimization: Geophysics, **67**, 1184–1191.

Operto, S., C. Ravaut, L. Improta, J. Virieux, A. Herrero, and P. DellAversana, 2004, Quantitative imaging of complex structures from dense wide-aperture seismic data by multiscale traveltime and waveform inversions: a case study: Geophys. Prosp., **52**, 625–651.

Perrone, F., J. Pannizardi, and P. Sava, 2013, Wavefield tomography based on local image correlations: Geophysical Prospecting, submitted for publication.

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.

———, 2009, Three-dimensional frequency-domain full-waveform inversion with an iterative solver: Geophysics, **74**, WCC53–WCC61.

Plessix, R. E., F. ten Kroode, and W. Mulder, 2000, Automatic crosswell tomography by differential semblance optimization: 70th Ann. Internat. Mtg, Soc. of Expl. Geophys., 2265–2268.

Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model: Geophysics, **64**, 888–901.

Pratt, R. G., and M. H. Worthington, 1990, Inverse theory applied to multi-source cross-hole tomography. part 1: Acoustic wave-equation method: Geophys. Prosp., **38**, 287–310.

Rickett, J., and P. Sava, 2002, Offset and angle-domain common image-point gathers for shot-profile migration: Geophysics, **67**, 883–889.

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

———, 2004b, Wave-equation migration velocity analysis - II: Subsalt imaging examples: Geophysical Prospecting, **52**, 607–623.

Sava, P., and S. Fomel, 2003, Angle-domain common image gathers by wavefield continuation methods: Geophysics, **68**, 1065–1074.

———, 2005, Coordinate-independent angle-gathers for wave equation migration: 75th Annual International Meeting, SEG, Expanded Abstracts, 2052–2055.

———, 2006, Time-shift imaging condition in seismic migration: Geophysics, **71**, S209–S217.

Sava, P., and I. Vasconcelos, 2011, Extended imaging condi-

tion for wave-equation migration: Geophysical Prospecting, **59**, 35–55.
Sava, P., and I. Vlad, 2011, Wide-azimuth angle gathers for wave-equation migration: Geophysics, **76**, S131–S141.
Shan, G., and Y. Wang, 2013, RTM based wave equation migration velocity analysis: 4726–4731.
Shen, P., W. Symes, S. Morton, and H. Calandra, 2005, Differential semblance velocity analysis via shot profile migration: 74th Annual International Meeting, SEG, Expanded Abstracts, 2249–2253.
Shen, P., W. Symes, and C. C. Stolk, 2003, Differential semblance velocity analysis by wave-equation migration: 73rd Ann. Internat. Mtg., Soc. of Expl. Geophys., 2132–2135.
Shen, P., and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, VE49–VE59.
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.
Shin, C., and D.-J. Min, 2006, Waveform inversion using a logarithmic wavefield: Geophysics, **71**, R31–R42.
Sirgue, L., and R. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248.
Soubaras, R., and B. Gratacos, 2007, Velocity model building by semblance maximization of modulated-shot gathers: Geophysics, **72**, U67–U73.
Stolk, C. C., and W. W. Symes, 2004, Kinematic artifacts in prestack depth migration: Geophysics, **69**, 562–575.
Symes, W., 2009, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, 765–790.
———, 2010, Waveform inversion progress and prospects: SEG Technical Program Expanded Abstracts, 3968–3969.
Tang, Y., and B. Biondi, 2013, Subsalt imaging by target-oriented inversion-based imaging: A 3D field-data example: Geophysics, **78**, S47–S58.
Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.
———, 2005, Inverse Problem Theory and methods for model parameter estimation: SIAM.
van Leeuwen, T., and W. A. Mulder, 2010, A correlation-based misfit criterion for wave-equation traveltime tomography: Geophysical Journal International, **182**, 1383–1394.
Vigh, D., and E. W. Starr, 2008, 3D prestack plane-wave, full-waveform inversion: Geophysics, **73**, VE135–VE144.
Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, WCC1–WCC26.
Weibull, W., and B. Arntsen, 2013, Automatic velocity analysis with reverse-time migration: Geophysics, **78**, S179–S192.
Woodward, M. J., 1992, Wave-equation tomography: Geophysics, **57**, 15–26.
Yang, T., and P. Sava, 2008, Wave-equation extended images applied to semblance and depth focusing velocity analysis: 78th Annual International Meeting, SEG, Expanded Abstracts, 2351–2354.
———, 2009, Wave-equation migration velocity analysis using extended images: 79th Annual International Meeting, SEG, Expanded abstracts, 3715–3719.
———, 2010, Moveout analysis of wave-equation extended images: Geophysics, **75**, 151–161.
———, 2011, Wave-equation migration velocity analysis with time-shift imaging: Geophysical Prospecting, **59**, 635–650.
Yilmaz, O., 2001, Seismic Data Analysis, 2^{nd} edition: Society of Exploration Geophysicists.
Zhang, Y., and D. Wang, 2009, Traveltime information-based wave-equation inversion: Geophysics, **74**, WCC27–WCC36.

7 APPENDIX A: THE ADJOINT-STATE METHOD

Consider the problem of finding the extreme of a function $\mathcal{J}(d,m)$ under the constraints $\mathcal{F}(d,m)=0$:

$$min \qquad \mathcal{J}(d,m) \tag{A-1}$$

$$s.t. \qquad \mathcal{F}(d,m)=0\ . \tag{A-2}$$

The variables $d=\{d_i\}$, $1\leq i\leq N_d$, represent the data, and the variables $m=\{m_j\}$, $1\leq j\leq N_m$ represent the model. The variables N_d and N_m indicate the number of data and model parameters, respectively. The function $\mathcal{J}(d,m)$ could be an objective function of an inverse problem, and the constraints $\mathcal{F}(d,m)=\{\mathcal{F}_i(d,m)\}$ specify the fact that the data are physical realizations of the model, for example they satisfy a wave-equation.

The constrained gradient of the selected objective function can be done efficiently using the **adjoint state method** (Plessix, 2006).

7.1 Augmented functional

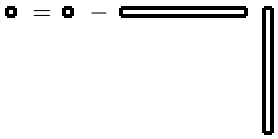
Using the method of Lagrange multipliers, the constrained minimum of $\mathcal{J}$ subject to the constraints given by $\mathcal{F}=0$ can be found by searching in the direction of the gradient of the augmented functional $\mathcal{H}$ defined by the expression

$$\mathcal{H}=\mathcal{J}-\mathcal{F}^\top a\ . \tag{A-3}$$

The parameters $a=\{a_i\}$, $1\leq i\leq N_d$ are known as Lagrange multipliers, and the symbol $^\top$ indicates the adjoint operation. Since the vectors $\mathcal{F}$ and a have equal size, the expression $\mathcal{F}^\top a$ also indicates summation over the data parameters, i.e.

$$\mathcal{F}^\top a=\sum_i \mathcal{F}_i^\top a_i\ . \tag{A-4}$$

We can represent the expression A-3 pictorially using matrices, which indicate the relative dimensions of the various variables:



$\mathcal{H}$ $\mathcal{J}$ $\mathcal{F}^\top$ a

We can find the extreme of the augmented functional $\mathcal{H}$ by searching in the direction of

- its **gradient** computed by evaluating the partial derivatives of the augmented functional relative to the model parameters m using

$$\frac{\partial\mathcal{H}}{\partial m}=\frac{\partial\mathcal{J}}{\partial m}-\left(\frac{\partial\mathcal{F}}{\partial m}\right)^\top a\ , \tag{A-5}$$

where

- the **state variables** d are computed by solving the **state equations** obtained by setting the partial derivatives of the augmented functional $\mathcal{H}$ relative to the Lagrange parameters a to zero

$$\frac{\partial\mathcal{H}}{\partial a}=0\ , \tag{A-6}$$

and

- the **adjoint variables** a are computed by solving the **adjoint equations** obtained by setting the partial derivatives of the augmented functional $\mathcal{H}$ relative to the state variables d to zero

$$\frac{\partial\mathcal{H}}{\partial d}=0\ . \tag{A-7}$$

Both $\dfrac{\partial\mathcal{H}}{\partial a}=\left\{\dfrac{\partial\mathcal{H}}{\partial a_i}\right\}$ and $\dfrac{\partial\mathcal{H}}{\partial d}=\left\{\dfrac{\partial\mathcal{H}}{\partial d_i}\right\}$ are vectors of size N_d.

7.2 State equations

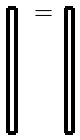
Using equation A-3, the condition A-6 is equivalent to

$$\frac{\partial\mathcal{J}}{\partial a}-\mathcal{F}=0\ . \tag{A-8}$$

Since by construction $\mathcal{J}$ is not a function of a, we have

$$\mathcal{F}=0\ , \tag{A-9}$$

which are the original physical constraints expressing the relationships between the model and data parameters:



$\mathcal{F}$ 0

which simply represent a system of N_d equations for the state variables d.

7.3 Adjoint equations

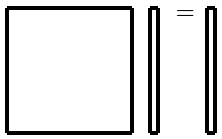
Using equation A-3, the condition A-7 is equivalent to

$$\frac{\partial\mathcal{J}}{\partial d}-\left(\frac{\partial\mathcal{F}}{\partial d}\right)^\top a=0\ , \tag{A-10}$$

therefore we can write the adjoint equations that can be used to construct the adjoint variables a given the state variables d:

$$\left(\frac{\partial\mathcal{F}}{\partial d}\right)^\top a=\frac{\partial\mathcal{J}}{\partial d}\ . \tag{A-11}$$

The quantities $\dfrac{\partial\mathcal{J}}{\partial d}$ are known as the **adjoint sources** since they are the forcing term for the adjoint equations used to compute the adjoint variables a. Pictorially, we can represent the system A-11 as



$$\left(\frac{\partial \mathcal{F}}{\partial d}\right)^{\top} \quad a \quad \frac{\partial \mathcal{J}}{\partial d}$$

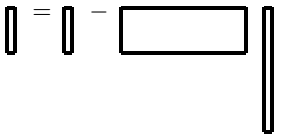
thus indicating that the adjoint variables are solved by a system of similar dimensions to the one needed to solve for the state variables.

7.4 Gradient

Finally, we can obtain the gradient of the objective function $\mathcal{J}$ relative to the model m subject to the physical constraints $\mathcal{F}$ by

$$\frac{\partial \mathcal{H}}{\partial m} = \frac{\partial \mathcal{J}}{\partial m} - \left(\frac{\partial \mathcal{F}}{\partial m}\right)^{\top} a \; . \tag{A-12}$$

Equation A-12 can be represented pictorially as



$$\frac{\partial \mathcal{H}}{\partial m} \quad \frac{\partial \mathcal{J}}{\partial m} \quad \left(\frac{\partial \mathcal{F}}{\partial m}\right)^{\top} \quad a$$

thus indicating that the gradient has the same size as the model (N_m).

8 APPENDIX B: ORIENTED CROSS-CORRELATION

One possible measure to evaluate the similarity between two multi-dimensional wavefields $a\,(\mathbf{x},t)$ and $b\,(\mathbf{x},t)$ is **cross-correlation**. A symmetric definition of this operation is

$$\begin{aligned} c\,(\mathbf{x},\tau) &= a\,(\mathbf{x},t) \star b\,(\mathbf{x},t) \\ &= \sum_t a\left(\mathbf{x},t-\frac{\tau}{2}\right) b\left(\mathbf{x},t+\frac{\tau}{2}\right) \; , \end{aligned} \tag{B-1}$$

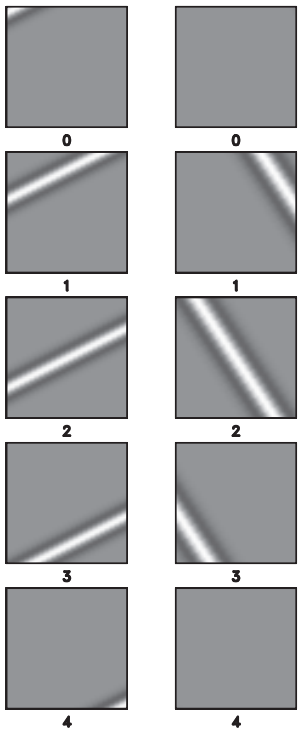
where the summation extends over the entire time domain in which the wavefields a and b are defined. The main drawback of this operation is that correlation is done point-by-point, and therefore the relation between the wavefields as a function of space is not accounted for. For example, the examples depicted in Figures B-1 and B-3 represent wavefields propagating in different directions. All synthetic wavefields are constructed such that for all 4 cases, the central trace is the same. Therefore, the central cross-correlations between the wavefields in Figure B-1 and Figures B-3 are identical, as seen in Figures 2(a) and 4(a).

Alternatively, we can formulate an **oriented cross-correlation** which accounts for the spatial distribution of the wavefield, for which the simple cross-correlation defined in equation B-1 is a special case:

$$\begin{aligned} c\,(\mathbf{x},\boldsymbol{\lambda},\tau) &= a\left(\mathbf{x}-\frac{\boldsymbol{\lambda}}{2},t\right) \star b\left(\mathbf{x}+\frac{\boldsymbol{\lambda}}{2},t\right) \\ &= \sum_t a\left(\mathbf{x}-\frac{\boldsymbol{\lambda}}{2},t-\frac{\tau}{2}\right) b\left(\mathbf{x}+\frac{\boldsymbol{\lambda}}{2},t+\frac{\tau}{2}\right) \; . \end{aligned} \tag{B-2}$$

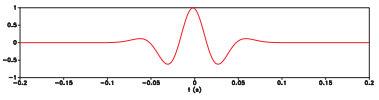
The space-lag vector $\boldsymbol{\lambda}$ is used to reach into the wavefield away from the central point where we are interested to evaluate the similarity between the two wavefields. This symmetric definition employing the vectors $+\boldsymbol{\lambda}$ and $-\boldsymbol{\lambda}$ is just one convenient possibility.

The oriented cross-correlation between the wavefields in Figure B-1 evaluated in the neighborhood of the central point of the wavefield captures the directionality of the wavefields, as seen in Figures 2(b) and 4(b). Such oriented cross-correlation can be used to filter wavefields as a function of direction, or to capture wavefield directionality during imaging, thus enabling angle-domain imaging, for example.

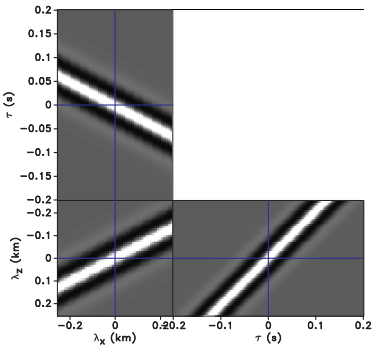


(a) (b)

Figure B-1. Example A: Synthetic plane waves propagating in different directions such that middle traces are identical.

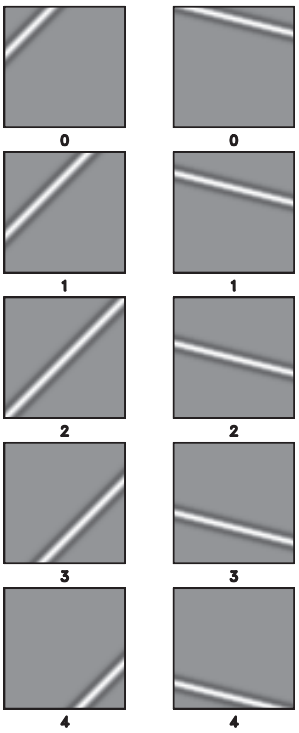


(a)



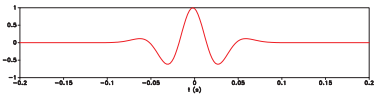
(b)

Figure B-2. Example A: (a) Simple cross-correlation and (b) oriented cross-correlation of the wavefields depicted in Figure B-1

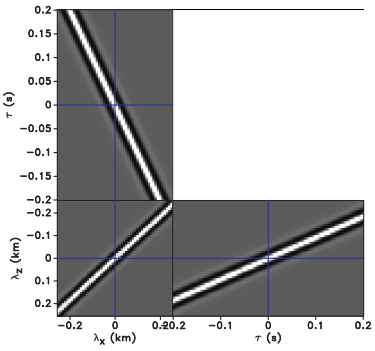


(a) (b)

Figure B-3. Example B: Synthetic plane waves propagating in different directions such that middle traces are identical.



(a)



(b)

Figure B-4. Example B: (a) Simple cross-correlation and (b) oriented cross-correlation of the wavefields depicted in Figure B-3