Wave-equation migration velocity analysis with extended common-image-point gathers

Tongning Yang & Paul Sava
Center for Wave Phenomena, Colorado School of Mines

ABSTRACT

Wave-equation migration velocity analysis (WEMVA) is an image-domain velocity model building technique based on band-limited wave propagation and designed especially for complex subsurface environments. It exploits the coherency of reflection events measured in extended images produced by a cross-correlation imaging condition with non-zero lags. Conventional approaches use either space-lags or time-lag common image gathers, in which only partial information of the extended images is used for velocity updates.
We propose an WEMVA approach using the complete information from both space-lags and time-lags of extended images. With this approach, the velocity model building benefits both from the robustness of using the time-lag information and from the high resolution of using the space-lags information. Such an implementation is facilitated by using extended common-image-point gathers (CIPs) constructed sparsely along reflections and defined jointly for space- and time-lags. These CIPs avoid the bias towards nearly-horizontal reflectors so that steeply dipping events are well preserved in the gathers and the corresponding information related to velocity can be used. Also, the computation of the extended images can be avoided in areas where the velocity is known, e.g., inside salt bodies, or areas where the signal-to-noise ratio is too low, e.g., in shadow zones. This significantly reduces the cost of constructing extended images.
A velocity estimation process based on these images requires an objective function based on an operator penalizing the distortion of the images caused by velocity errors. Such an objective function can be designed using the differential semblance principle. The objective function built in this way is uni-modal with respect to the model, thus preventing the inversion from being trapped in local minima. The smoothness of the function around the global minima facilitates the use of gradient-type solvers for achieving convergence towards the true model. The velocity estimation process requires computing the gradient of the objective function which links image errors to velocity model updates. One key component for the construction of the gradient is the adjoint scattering operator which we construct in the framework of frequency-domain downward continuation. Such an operator is formulated by applying the Born linearization to the single square-root equation, and it serves as the foundation for image-domain wavefield tomography algorithms.

Key words: wavefield-extrapolation, velocity analysis, extended images, objective function

1 INTRODUCTION

The focus of seismic exploration is in regions characterized by complex subsurface structure. In such regions, wavefield-based migration is a powerful tool for accurately imaging the earth's interior that has become a commercial routine. However, since wavefield-based migration is sensitive to velocity error, the quality of the final image greatly depends on the accuracy of the velocity model. Thus, a key challenge for imag-

ing in complex geology is an accurate determination of the velocity model in the area under investigation.

In practice, all velocity model building techniques require an information carrier to connect the input (the quantity we can measure from the data), to the output (the unknown quantity we want to resolve). Ray-based methods are techniques which use high-frequency asymptotic rays as the information carrier (Bishop et al., 1985; Zhou, 2003; Stork, 1992; Liu, 1997); while wavefield-based methods are techniques which use band-limited wavefields as the information carrier (Tarantola, 1984; Pratt, 1999; Sava & Biondi, 2004a; Operto et al., 2007; Shen & Symes, 2008; Xie & Yang, 2008).

In complex geology, ray-based methods often fail to produce a good velocity model for several reasons. First, ray-based methods cannot handle complicated wave propagation phenomena such as multi-pathing caused by the complexity of the subsurface. Second, sharp velocity contrasts cause instability for ray-based methods. In addition, the high-frequency assumption embedded in ray-based methods is inconsistent with the finite-frequency characteristic of wavefield-based migration. A wavefield-based velocity analysis technique is expected to provide a more accurate, robust and consistent solution because of its accurate description of wave propagation in a complex subsurface, capability of handling strong velocity variation, and a frequency band that is consistent with wavefield-based migration. Although ray-based methods have the advantage of computational efficiency, wavefield-based methods are more appropriate than ray-based methods for building a velocity model in complex geologic areas. However, the expensive computational cost, reliance on the starting model, and convergence of the results at local minimum are all issues that need to be solved for practical application of wavefield-based methods (Etgen et al., 2009; Virieux & Operto, 2009).

Wavefield-based velocity analysis techniques can be classified by the domain in which the optimization problem is formulated. In the data domain, wavefield-based velocity estimation techniques are described as full- waveform inversion (FWI) (Tarantola, 1987; Mora, 1988; Song et al., 1995), while they are referred to as wavefield-based MVA when implemented in the image domain (Biondi & Sava, 1999; Shen et al., 2003; Albertin et al., 2006; Shen & Symes, 2008; Symes, 2008). The essential difference between these two categories is that FWI defines residuals by directly measuring the misfit between observed and predicted data, while WEMVA defines residuals by measuring the coherency and focusing of extended images. Then, the energy of the residual in either domain is backprojected via band-limited wave propagation into velocity updates.

WEMVA requires residuals defined in the image domain as the input. Thus, one needs first to measure the deficiency of migrated images using focusing or semblance properties. Biondi & Sava (1999) and Sava & Biondi (2004a,b) develop the framework of WEMVA and illustrate the construction of a linearized migration operator and its adjoint. Shen et al. (2003) illustrate the WEMVA method using the differential semblance optimization (DSO). Shen & Calandra (2005) and Shen & Symes (2008) implement a similar DSO scheme for survey-sinking and for shot-profile migration, respectively, and analyze properties of the differential semblance for offset and angle gathers.

Earlier research on image-domain wavefield tomography uses space-lag common-image-gathers (CIGs), which include only partial information of migrated images (Sava & Biondi, 2004a; Shen & Symes, 2008). The additional information embedded in time-lag CIGs, however, has yet to be fully exploited (Yang & Sava, 2009). An approach that simultaneously uses all lags information from the images can benefit from the high resolution of space-lag gathers and the robustness of time-lag gathers, thus having the potential to render an accurate and high-resolution velocity model in areas of complex geology.

Here, we first present the algorithms for migration with conventional and extended images, as well as their adjoint algorithms. We then explain the key steps of the procedures and analyze the required computational cost. Also, we present algorithms of linearized migration with conventional and extended images, as well as their adjoint algorithms. Linearized migration and its adjoint are critical components for image-domain wavefield tomography. We also formulate objective functions of image-domain tomography designed particularly for CIPs. Using the semblance principle, we can construct different objective functions to penalize the residual moveout of events in CIPs. Finally, we demonstrate the characteristics of the objective function by synthetic examples.

2 MIGRATION WITH EXTENDED IMAGES

Under the assumption of single scattering at discontinuities in the subsurface, we can describe seismic migration as a two-steps imaging process: wavefield reconstruction followed by the application of an imaging condition.

Seismic waves generated from the source, represented by the source wavefield, propagate in the medium and interact with discontinuities. The reflected seismic waves, represented by the receiver wavefield, propagate back to the surface and are recorded by geophones. These two wavefields kinematically coincide at discontinuities. Therefore, after we reconstruct the wavefields by solving a wave equation numerically, we can extract the reflectivity information from the reconstructed wavefields using a conventional imaging condition (Claerbout, 1985),

$$R\left(\mathbf{x}\right) = \sum_{\omega} \overline{W_s\left(\mathbf{x},\omega\right)} W_r\left(\mathbf{x},\omega\right) \ , \tag{1}$$

or an extended imaging condition,

$$R\left(\mathbf{x},\boldsymbol{\lambda},\tau\right) = \sum_{\omega} \overline{W_s\left(\mathbf{x}-\boldsymbol{\lambda},\omega\right)} W_r\left(\mathbf{x}+\boldsymbol{\lambda},\omega\right) e^{2i\omega\tau} \ . \tag{2}$$

The image R is a function of space coordinates $\mathbf{x}$, of the space-lag $\boldsymbol{\lambda}$ and time-lag extensions τ (Rickett & Sava, 2002; Sava & Fomel, 2006; Sava & Vasconcelos, 2009). The complete set of extended images is a multi-dimensional hypercube, which

allows us to simultaneously access the semblance and focusing information about the images.

Seismic modeling process which maps the reflectivity of the subsurface into data **D** recorded on the surface can be formulated using a linear operator **L** applied to the reflectivity model represented by the image R:

$$\mathbf{D} = \mathbf{L}R \ . \tag{3}$$

Since the modeling is a linear process, the adjoint process can be directly obtained. Such an adjoint process is referred to as migration, and can be conceptually formulated as applying the operator $\mathbf{L}^*$ to data **D** and to obtain the image R. If we use matrix notation, we have

$$R = \mathbf{L}^*\mathbf{D} \ . \tag{4}$$

The pseudo-codes of the algorithm for migration and modeling with both conventional and extended images are detailed in boxes 1 and 2, respectively. The algorithms of modeling and migration are formulated in this work as adjoint pairs. Such algorithm pairs ensure stability if we use the operators in solving a linear inverse problem by conjugate-gradient methods. The particular implementation we illustrate is based on one-way wavefield-extrapolation shot-record migration, but one can generalize the algorithm to other imaging setups such as two-way wave-equation shot-record migration. The algorithms presented here perform computations for each frequency independently, which leads to a straightforward parallelization.

Box 1

MIGRATION WITH CONVENTIONAL IMAGES

```
set     R(x) = 0
        ω = ω_min ... ω_max {
read      W_s(x)
set       W_r(x) = 0
read      T(x,y)
          z = z_min ... z_max {
store       W_r(x,y,z) = T(x,y)
W.E.        T(x,y) = E^-[T(x,y)]
          }
I.C.      R(x) += conj(W_s(x)) W_r(x)
        }
write   R(x)
```

MODELING WITH CONVENTIONAL IMAGES

```
read    R(x)
        ω = ω_min ... ω_max {
read      W_s(x)
set       W_r(x) = 0
I.C.      W_r(x) += W_s(x) R(x)
set       T(x,y) = 0
          z = z_max ... z_min {
W.E.        T(x,y) = E^+[T(x,y)]
inject      T(x,y) += W_r(x,y,z)
          }
write     T(x,y)
        }
```

In the algorithms detailed in boxes 1 and 2, W.E. denotes wavefield extrapolation, I.C. denotes imaging condition, $\mathcal{E}^+$ represents a casual wavefield extrapolator, while $\mathcal{E}^-$ represents an anti-casual wavefield extrapolator. The source wavefield W_s is precomputed before migration and stored in the disk, but it can also be computed inside the algorithms. In the migration algorithm, we first recursively downward continue the receiver wavefield by applying the extrapolator $\mathcal{E}^-$ and store the extrapolated wavefield at every depth level. We only need to store the wavefield for one frequency, which makes the calculation manageable. After the wavefield reconstruction is accomplished, we extract the conventional image by equation 1, or the extended images by equation 2. The final image is the stack of images over all frequencies.

In the modeling algorithm, the reflectivity is injected into the receiver wavefield using the adjoint of the imaging condition in equations 1 or 2. Next, we recursively upward continue the full receiver wavefield from bottom to top by applying the adjoint extrapolator $\mathcal{E}^+$ and inject the wavefield constructed from the model at every depth level. After the wavefield reconstruction is finished, we output the wavefield at the surface.

Box 2

MIGRATION WITH EXTENDED IMAGES

```
set    R(c, λ, τ) = 0
       ω = ω_min ... ω_max {
read     W_s(x)
set      W_r(x) = 0
read     T(x, y)
         z = z_min ... z_max {
store      W_r(x, y, z) = T(x, y)
W.E.       T(x, y) = E^-[T(x, y)]
         }
         loop λ, τ {
I.C.       R(c, λ, τ) += e^{-2iωτ} conj(W_s(c−λ)) W_r(c+λ)
         }
       }
write  R(c, λ, τ)
```

MODELING WITH EXTENDED IMAGES

```
read   R(c, λ, τ)
       ω = ω_min ... ω_max {
read     W_s(x)
set      W_r(x) = 0
         loop λ, τ {
I.C.       W_r(c+λ) += W_s(c−λ) e^{+2iωτ} R(c, λ, τ)
         }
set      T(x, y) = 0
         z = z_max ... z_min {
W.E.       T(x, y) = E^+[T(x, y)]
inject     T(x, y) += W_r(x, y, z)
         }
write    T(x, y)
       }
```

The computational cost for such implementations can also be split between the wavefield reconstruction and the imaging condition. Since we use a frequency-domain recursive extrapolator to reconstruct the wavefields, the cost is proportional to the size of the model and the number of frequencies. Thus, it can be estimated as:

$$N_{\text{W.R.}} \sim N_\omega N_{\mathbf{x}} \ , \tag{5}$$

where N_ω represents the number of frequencies and $N_{\mathbf{x}}$ represents the number of samples along the space axes.

The cost for the imaging condition is controlled by several factors: the frequency band, the number of locations we choose to construct the extended images, and the number of lags involved. Thus, it can be estimated as:

$$N_{\text{I.C.}} \sim N_\omega N_{\mathbf{c}} N_{\boldsymbol{\lambda}} N_\tau \ , \tag{6}$$

where $N_{\mathbf{c}}$ represents the number of CIPs and $N_{\boldsymbol{\lambda}}$, N_τ represent the number of space- and time-lags. If we take the conventional space-lags CIGs and extended CIPs as examples, the computational cost for the imaging condition is

$$N_{\text{I.C.}} \sim N_\omega N_z N_{h_x} N_{h_y} N_{\lambda_x} N_{\lambda_y} \ , \tag{7}$$

where N_{h_x} and N_{h_y} represent the number of horizontal locations at which we construct the gathers along in-line and cross-line directions. If we denote the cost of constructing space-lags common-image gathers as $C_{\boldsymbol{\lambda}}$ and the cost of constructing space- and time-lags common-image-point gathers as $C_{\boldsymbol{\lambda}\tau}$, from the equation 6 and 7, we obtain the ratio:

$$\frac{C_{\boldsymbol{\lambda}}}{C_{\boldsymbol{\lambda}\tau}} \sim \frac{N_z N_{h_x} N_{h_y}}{N_{\mathbf{c}} N_{\lambda_z} N_\tau} \ . \tag{8}$$

Suppose $N_z = N_x = N_y = 500$, $N_{h_x} = N_{h_y} = 50$, $N_{\mathbf{c}} = 500$, $N_{\lambda_z} = 20$, $N_\tau = 100$, then $\frac{C_{\boldsymbol{\lambda}}}{C_{\boldsymbol{\lambda}\tau}} \sim 1.2$, which means the cost of computing CIPs is about the same as the cost of computing CIGs. However, we can drop the vertical space-lag axis if we are imaging nearly-horizontal reflectors, or we can eliminate one of the lag axes if we have information about the reflector dip. Then, $\frac{C_{\boldsymbol{\lambda}}}{C_{\boldsymbol{\lambda}\tau}} \sim 25$, thus we achieve a significant reduction in the computational cost of the imaging condition. Therefore, using CIPs for velocity analysis is particularly attractive, as they provide information about the velocity model at smaller cost. Furthermore, as discussed by Sava & Vasconcelos (2009), CIPs have other advantages over CIGs, mainly related to image sampling that is more consistent with the underlying geologic structure.

We migrate the Sigsbee 2A dataset (Paffenholz et al., 2002) as an example. Figure 1(a) shows the background velocity model, constructed by subtracting the perturbation shown in Figure 1(b) from the true velocity model. The image obtained is shown in Figure 1(c), overlain with the dots indicating the locations where CIPs are constructed. The $*$ indicate the positions of CIPs shown in Figure 2(a)-2(d). Since we use the background model for migration, we see that events in all CIPs are not well focused and exhibit residual moveout in λ_x-τ panels. In addition, in Figures 2(c) and 2(d), the CIPs are constructed in subsalt area with an uneven illumination, we see that the CIPs contain significant artifacts. In general, the salt creates uneven illumination and shadow zones, which generate artifacts and distort the images. The distortion is similar to the imperfections caused by the velocity error. If we use such gathers for velocity analysis, we may obtain an incorrect result. Thus, in complex area, illumination compensation is necessary for velocity model building. This important research direction falls outside the scope of this paper.

3 LINEARIZED WAVEFIELD-BASED MIGRATION VELOCITY ANALYSIS WITH EXTENDED IMAGES

To formulate the velocity analysis in the image-domain, we first need to construct an objective function for the optimization problem. In general, there are two approaches we can use to seek the solution. The first approach uses a linearized image perturbation. The second approach, based on differential semblance optimization, is discussed in the next section. To obtain

an image perturbation, one starts by applying the Taylor expansion to an image as follows:

$$R(\mathbf{x}, s_c) = R(\mathbf{x}, s_b) + \Delta R(\mathbf{x}) \ , \tag{9}$$

where

$$s_c = s_b + \Delta s \ . \tag{10}$$

Here, s_c denotes the correct slowness and s_b and Δs denote background slowness and slowness perturbation, respectively. The image perturbation is related to the slowness perturbation by a linearized scattering operator $\mathbf{A}$:

$$\Delta R(\mathbf{x}) = \mathbf{A}\Delta s \equiv \frac{\partial R(\mathbf{x}, s)}{\partial s}|_{s=s_b} \Delta s \ . \tag{11}$$

Since the image perturbation is proportional to the slowness perturbation, the velocity analysis can be performed by minimizing the image perturbation. Thus, the corresponding objective function, which defines a linearized optimization problem, is formulated as:

$$J(\Delta s) = \frac{1}{2}\|\Delta R - \mathbf{A}\Delta s\|_2^2 \ . \tag{12}$$

The operator $\mathbf{A}$ in equations 11 and 12 is used to obtain an image perturbation from a slowness perturbation:

$$\Delta R = \mathbf{A}\Delta s \ . \tag{13}$$

We can directly construct the adjoint MVA process similarly to migration. In this way, we obtain the adjoint pair of operators. The adjoint process describes the backprojection from image perturbation to slowness perturbation:

$$\Delta s = \mathbf{A}^* \Delta R \ , \tag{14}$$

where $\mathbf{A}^*$ represents the adjoint operator.

The pseudo-codes of the algorithm for forward and adjoint linearized MVA with both conventional and extended images are detailed in boxes 3 and 4, respectively. The algorithms are constructed based on the one-way wave-equation shot-record migration, and implemented in the frequency domain.

Box 3

FORWARD SCATTERING WITH CONVENTIONAL IMAGES

set $\Delta R(\mathbf{x}) = 0$
read $\Delta s(\mathbf{x})$
$\omega = \omega_{min} \dots \omega_{max}\{$
read $W_s(\mathbf{x})$
read $W_r(\mathbf{x})$
set $\overline{T_s(x,y)} = 0$
set $T_r(x,y) = 0$
$z = z_{min} \dots z_{max}\{$
W.S. $\overline{T_s(x,y)} \mathrel{+}= \mathcal{S}^-\left[\overline{W_s(x,y,z)}, \Delta s(x,y,z)\right]$
W.S. $T_r(x,y) \mathrel{+}= \mathcal{S}^-\left[W_r(x,y,z), \Delta s(x,y,z)\right]$
W.E. $\overline{T_s(x,y)} = \mathcal{E}^+\left[\overline{T_s(x,y)}\right]$
W.E. $T_r(x,y) = \mathcal{E}^-\left[T_r(x,y)\right]$
store $\overline{\Delta W_s(x,y,z)} = \overline{T_s(x,y)}$
store $\Delta W_r(x,y,z) = T_r(x,y)$
}
I.C. $\Delta R(\mathbf{x}) \mathrel{+}= W_r(\mathbf{x})\,\overline{\Delta W_s(\mathbf{x})}$
I.C. $\Delta R(\mathbf{x}) \mathrel{+}= \overline{W_s(\mathbf{x})}\Delta W_r(\mathbf{x})$
}
write $\Delta R(\mathbf{x})$

ADJOINT SCATTERING WITH CONVENTIONAL IMAGES

read $\Delta R(\mathbf{x})$
set $\Delta s(\mathbf{x}) = 0$
$\omega = \omega_{min} \dots \omega_{max}\{$
read $W_s(\mathbf{x})$
read $W_r(\mathbf{x})$
I.C. $\overline{\Delta W_s(\mathbf{x})} = \overline{W_r(\mathbf{x})}\Delta R(\mathbf{x})$
I.C. $\Delta W_r(\mathbf{x}) = W_s(\mathbf{x})\,\Delta R(\mathbf{x})$
set $\overline{T_s(x,y)} = 0$
set $T_r(x,y) = 0$
$z = z_{max} \dots z_{min}\{$
inject $\overline{T_s(x,y)} \mathrel{+}= \overline{\Delta W_s(x,y,z)}$
inject $T_r(x,y) \mathrel{+}= \Delta W_r(x,y,z)$
W.E. $\overline{T_s(x,y)} = \mathcal{E}^-\left[\overline{T_s(x,y)}\right]$
W.E. $T_r(x,y) = \mathcal{E}^+\left[T_r(x,y)\right]$
W.S. $\Delta s(x,y,z) \mathrel{+}= \mathcal{S}^+\left[W_s(x,y,z), \overline{T_s(x,y)}\right]$
W.S. $\Delta s(x,y,z) \mathrel{+}= \mathcal{S}^+\left[\overline{W_r(x,y,z)}, T_r(x,y)\right]$
}
}
write $\Delta s(\mathbf{x})$

Box 4

FORWARD SCATTERING WITH EXTENDED IMAGES

set $\Delta R(\mathbf{c},\boldsymbol{\lambda},\tau)=0$
read $\Delta s(\mathbf{x})$
$\omega=\omega_{min}\dots\omega_{max}\{$
read $W_s(\mathbf{x})$
read $W_r(\mathbf{x})$
set $\overline{T_s(x,y)}=0$
set $T_r(x,y)=0$
$z=z_{min}\dots z_{max}\{$
W.S. $\overline{T_s(x,y)}\mathrel{+}=\mathcal{S}^-\left[\overline{W_s(x,y,z)},\Delta s(x,y,z)\right]$
W.S. $T_r(x,y)\mathrel{+}=\mathcal{S}^-\left[W_r(x,y,z),\Delta s(x,y,z)\right]$
W.E. $\overline{T_s(x,y)}=\mathcal{E}^+\left[\overline{T_s(x,y)}\right]$
W.E. $T_r(x,y)=\mathcal{E}^-\left[T_r(x,y)\right]$
store $\overline{\Delta W_s(x,y,z)}=\overline{T_s(x,y)}$
store $\Delta W_r(x,y,z)=T_r(x,y)$
}
loop $\boldsymbol{\lambda},\tau\{$
I.C. $\Delta R(\mathbf{c},\boldsymbol{\lambda},\tau)\mathrel{+}=e^{+2i\omega\tau}W_r(\mathbf{c}+\boldsymbol{\lambda})\,\overline{\Delta W_s(\mathbf{c}-\boldsymbol{\lambda})}$
I.C. $\Delta R(\mathbf{c},\boldsymbol{\lambda},\tau)\mathrel{+}=e^{-2i\omega\tau}\overline{W_s(\mathbf{c}-\boldsymbol{\lambda})}\Delta W_r(\mathbf{c}+\boldsymbol{\lambda})$
}
}
write $\Delta R(\mathbf{c},\boldsymbol{\lambda},\tau)$

ADJOINT SCATTERING WITH EXTENDED IMAGES

read $\Delta R(\mathbf{x})$
set $\Delta s(\mathbf{x})=0$
$\omega=\omega_{min}\dots\omega_{max}\{$
read $W_s(\mathbf{x})$
read $W_r(\mathbf{x})$
loop $\boldsymbol{\lambda},\tau\{$
I.C. $\overline{\Delta W_s(\mathbf{c}-\boldsymbol{\lambda})}\mathrel{+}=e^{-2i\omega\tau}\overline{W_r(\mathbf{c}+\boldsymbol{\lambda})}\Delta R(\mathbf{c},\boldsymbol{\lambda},\tau)$
I.C. $\Delta W_r(\mathbf{c}+\boldsymbol{\lambda})\mathrel{+}=e^{+2i\omega\tau}W_s(\mathbf{c}-\boldsymbol{\lambda})\,\Delta R(\mathbf{c},\boldsymbol{\lambda},\tau)$
}
set $\overline{T_s(x,y)}=0$
set $T_r(x,y)=0$
$z=z_{max}\dots z_{min}\{$
inject $\overline{T_s(x,y)}\mathrel{+}=\overline{\Delta W_s(x,y,z)}$
inject $T_r(x,y)\mathrel{+}=\Delta W_r(x,y,z)$
W.E. $\overline{T_s(x,y)}=\mathcal{E}^-\left[\overline{T_s(x,y)}\right]$
W.E. $T_r(x,y)=\mathcal{E}^+\left[T_r(x,y)\right]$
W.S. $\Delta s(x,y,z)\mathrel{+}=\mathcal{S}^+\left[W_s(x,y,z),\overline{T_s(x,y)}\right]$
W.S. $\Delta s(x,y,z)\mathrel{+}=\mathcal{S}^+\left[\overline{W_r(x,y,z)},T_r(x,y)\right]$
}
}
write $\Delta s(\mathbf{x})$

In the algorithms detailed in boxes 3 and 4, W.S. denotes wavefield scattering, $\mathcal{S}^+$ represents a casual wavefield scattering operator, and $\mathcal{S}^-$ represent an anti-casual wavefield scattering operator. These scattering operators characterize the wavefield scattered from the slowness perturbation via the background wavefield (Sava & Vlad, 2008). The background wavefields must be precomputed beforehand.

In the forward MVA algorithm, the total scattered wavefield consists of the scattered wavefield extrapolated from the previous depth level above added to the scattered wavefield generated from the slowness perturbation at the current depth level. At a certain depth level, we first apply the scattering operator $\mathcal{S}$ to obtain the scattered wavefields for the source and receiver, respectively. The imaging conditions cross-correlate the complex conjugate of the source wavefield and receiver wavefield. Taking the complex conjugate is not a linear operation in the complex domain. We can directly compute the complex conjugate of the scattered source wavefield and then downward continue the scattered wavefields by the wavefield extrapolator. After the computation for scattered wavefields is done, we use the conventional or extended imaging conditions to extract the image perturbation. In both cases, we apply the imaging conditions twice, to different combinations of the scattered and background wavefields, respectively, then sum the results from these two computations.

In the adjoint MVA algorithm, we first construct the perturbed wavefields from the image perturbations by the adjoint of the imaging conditions. Notice that we also directly compute the complex conjugate of the perturbed source wavefield to avoid taking the complex conjugate of the wavefield. Next, we inject the perturbed wavefield into the total scattered wavefield extrapolated from the previous depth level below, and upward continue the new scattered wavefield. The slowness perturbation at the current depth level is obtained from the total scattered wavefield by the adjoint of the scattering operator.

The computational cost for the MVA process consists of three parts: wavefield reconstruction (W.R.), wavefield scattering (W.S.), and imaging condition (I.C.). The cost for both wavefield reconstruction and scattering can be estimated by equation 5, scaled by 2 since we need to compute them twice, for the source and receiver sides. The cost for imaging condition can be estimated by equation 6, scaled by 2 for the same reason.

We use the Sigsbee 2A dataset as an example again with the background model in Figure 1(a) and with the slowness perturbation in Figure 1(b) to create the image perturbation. We construct the image perturbation for CIPs located at the same positions denoted by the $*$ in Figure 1(c), as shown in Figures 3(a)- 3(d). Comparing the image perturbation of CIPs to those from background velocity model in Figure2(a)-(d), we notice that the events in Figure 3(a)-3(d) are more oscillatory because we use less frequencies in constructing the image perturbation to save cost. Also, the image perturbation has a phase difference because of the linearization of the wavenumber with respect to slowness in the scattering operator. Next, we apply the adjoint operator to the image perturbation to obtain the backprojection, which is equivalent to the gradient of the corresponding objective function given by equation 12. The backprojection provides a straightforward way to study the forward and adjoint scattering operators. The main fac-

tors that control the shape of the backprojection are (1) the frequency content of the background wavefields, (2) the type of source (e.g., point or plane-wave) from which we generate the background wavefields, and (3) the type of perturbation constructed in the image space. Figure 4(a)-4(c) show the backprojection from a single shot at $x = 12.3$ km using different frequency bands. We can see that with increasing frequency band used, the magnitude of the backprojection become stronger. Figure 5(a)-5(c) show the backprojection from all shots using different frequency bands. Similarly, the magnitude of the backprojection is increasing with more frequencies used. In Figure 5(c), we see that backprojection has a complicated shape due to the complexity of the background model.

In practice, the true slowness perturbation is never known, so the image perturbation is also unknown. However, one can construct a linearized image perturbation to approximate the true one. Thus, the key element for this approach is the construction of the linearized image perturbation. Sava (2003) proposes using Stolt residual migration as the solution, while Yang & Sava (2009) use focusing of the time-lag extended images to address the problem. Generally, the linearized-image perturbation approach can effectively avoid the common cycle-skipping problem for wavefield tomography. The construction of the linearized image perturbation, however, requires a quantity that can be measured directly from the image. Such a quantity is difficult to measure for all types of extended images. An alternative to this approach is to construct penalties of the migrated images itself, as discussed in the following section.

4 DIFFERENTIAL SEMBLANCE OPTIMIZATION WITH EXTENDED IMAGES

An alternative approach to formulate image-domain wavefield tomography relies on applying a penalty function to extended images. The penalty function measures the coherency of images and highlights and penalizes the image imperfections caused by velocity errors. In general, we can formulate an objective function for optimization by first identifying characteristics of the images corresponding to correct velocity. Such features define an ideal shape for the images relative to which we can measure image imperfections. Since the velocity error distorts the wavefields reconstructed in the subsurface, the images constructed with an incorrect velocity does not conform with the ideal shape. This part of the energy in the images is the residual. Therefore, we require the penalty function to annihilate the energy of images corresponding to the ideal shape, and to preserve and highlight the energy of images departing from the ideal shape. As a result, a penalized image is equivalent to the data misfit defined in the image domain. Therefore, the velocity optimization problem is formulated by minimizing the image residual, and the corresponding objective function

$$J(s) \equiv \frac{1}{2}\|\mathbf{P}[R(\mathbf{x},\boldsymbol{\lambda},\tau)]\|_2^2 \,. \tag{15}$$

In general, $R(\mathbf{x},\boldsymbol{\lambda},\tau)$ are extended images with different extensions and $\mathbf{P}$ is the penalty function.

According to the semblance principle, if data are migrated with the correct velocity model, the subsurface reflectors are imaged at a fixed location for different seismic experiments regardless of the geological structure (Al-Yahya, 1989). Therefore, we can use the image semblance to quantify these imperfections and to update the velocity model. This process is referred to as semblance analysis. Symes & Carazzone (1991) introduce differential semblance as a particular implementation of semblance analysis, and the corresponding velocity analysis process, often referred to as differential semblance optimization (DSO). The underlying idea is to analyze the difference of semblance between neighboring traces within a CIGs or a CIPs. The main advantage of DSO is the convexity of the corresponding objective function, which greatly reduces the difficulty of handling the local minima in common wavefield-tomography methods in the context of wave-equation migration velocity analysis. The DSO approach has been applied to conventional CIGs (Shen et al., 2003; Shen & Calandra, 2005; Shen & Symes, 2008).

4. 1 Objective function for extended images

For CIGs constructed in the space-lag domain, Shen et al. (2003) and Shen & Symes (2008) derive the formula to describe the kinematics of a reflector at depth z_0 using extended images by

$$R(z,\boldsymbol{\lambda}) = \delta(\boldsymbol{\lambda})\,\delta(z - z_0) \,, \tag{16}$$

where z_0 is the depth of the reflector. The reflections should be focused at zero space-lag because the reflectors are imaged at the same subsurface location for different shots when the velocity is correct. Hence, we measure the focusing of events in CIGs, and penalize the residual moveout, which is the energy not concentrated at zero space-lag. The penalty function is formulated as

$$\mathbf{P}_{\boldsymbol{\lambda}}[R] = |\boldsymbol{\lambda}|R \,. \tag{17}$$

For CIPs constructed in the space- and time-lags domain, Sava & Vasconcelos (2009) describe the kinematics of the extended images by

$$R(\boldsymbol{\lambda},\tau) = \delta(\mathbf{q}\cdot\boldsymbol{\lambda})\,\delta(v\tau) \,, \tag{18}$$

where $\mathbf{q}$ denotes a unit vector parallel to the reflector and at the intersection of the reflection and reflector planes. When correct velocity v is used for imaging, an event in a CIPs is illuminated by multiple shots and is represented by a line at $\tau = 0$ oriented at an angle parallel to the reflector normal. Thus, we also measure the focusing along the trajectory oriented at an angle parallel to the reflector normal. Similar to the conventional DSO approach, we penalize the residual moveout spread from the tilted line at $\tau = 0$. The penalty function formulated in this way is

$$\mathbf{P}_{(\boldsymbol{\lambda},\tau)}[R] = |\tau|R \,. \tag{19}$$

Such a penalty function has the advantages that it is straightforward and requires no information about the slope of the reflection.

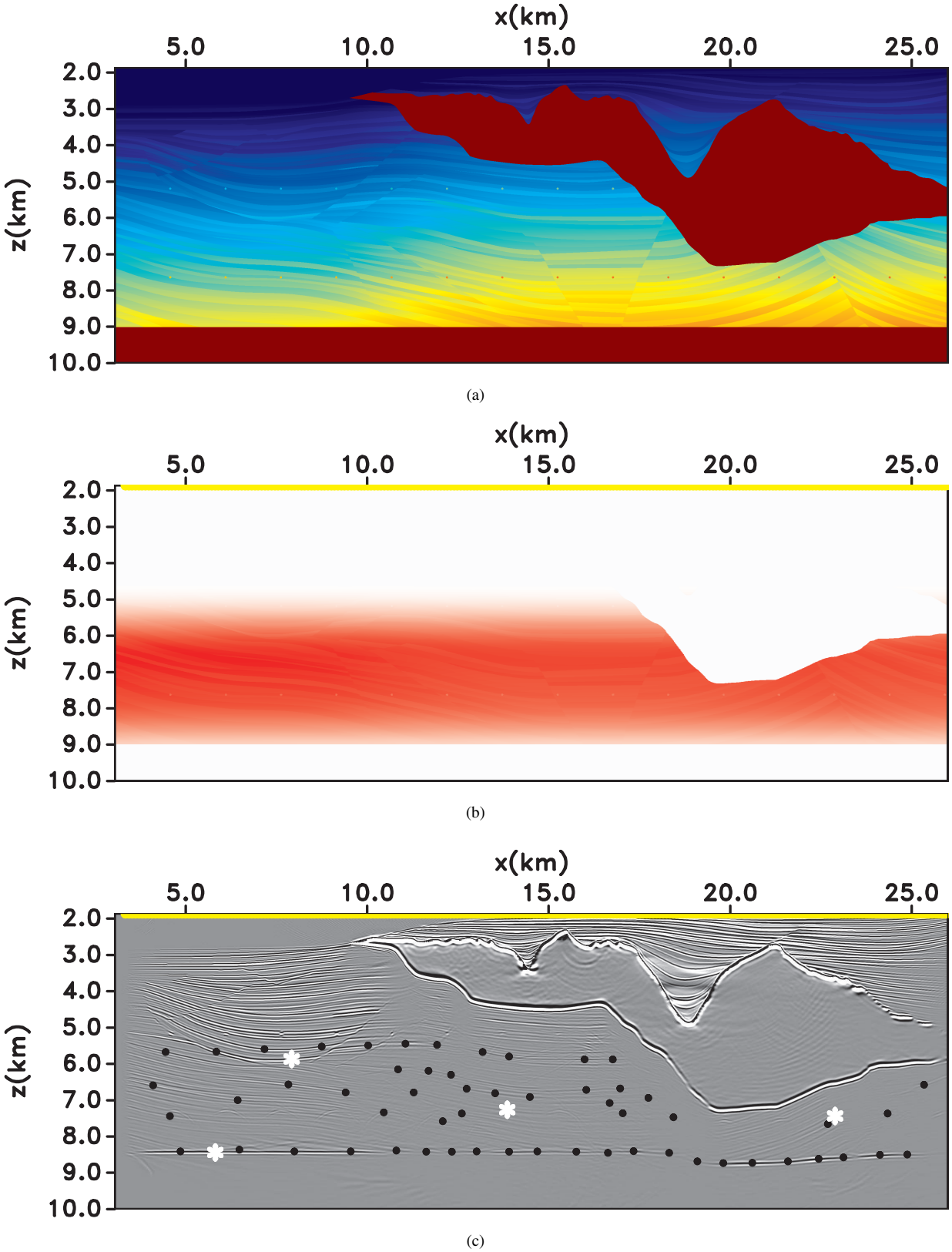
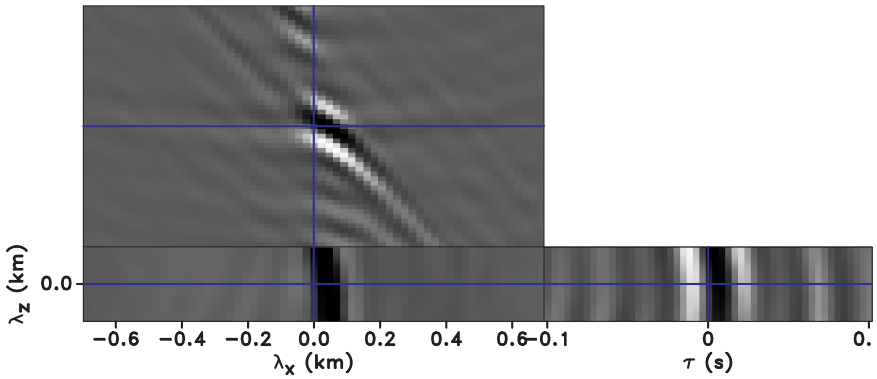
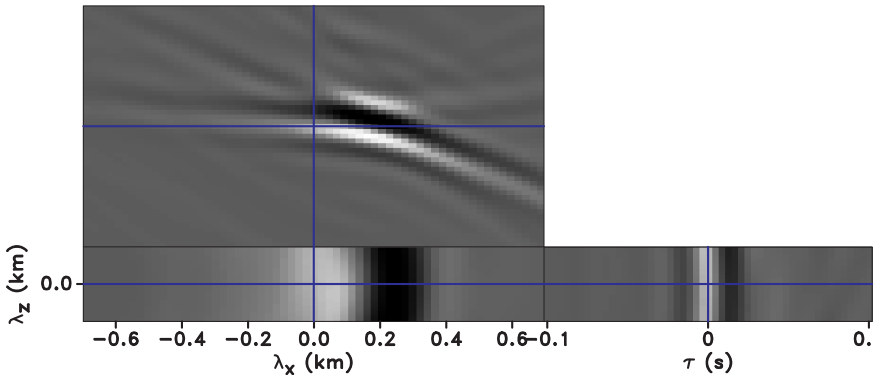


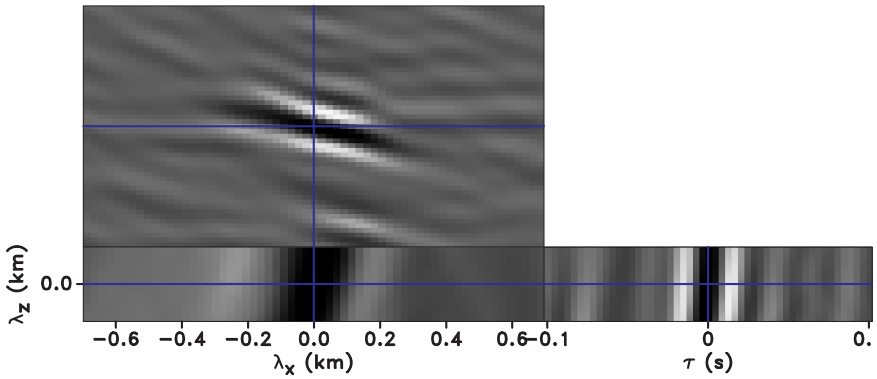
Figure 1. (a) Sigsbee 2A synthetic model, (b) slowness perturbation and (c) migrated image with background velocity model. The dots indicate the positions where CIPs are constructed. The $*$ indicate the positions of the CIPs shown in Figures 2 and 3.



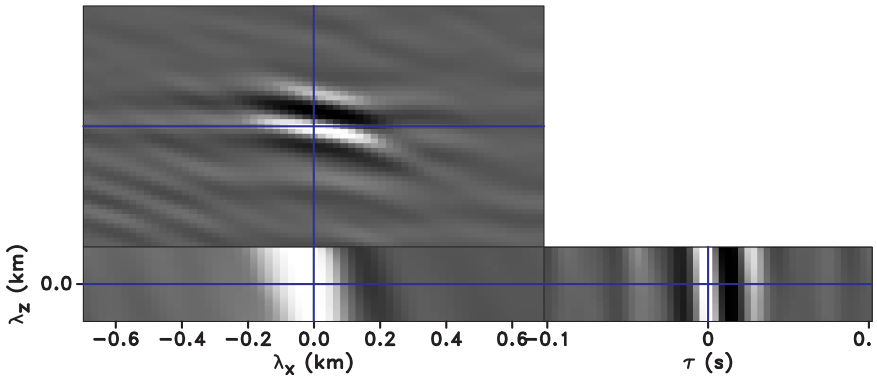
(a)



(b)

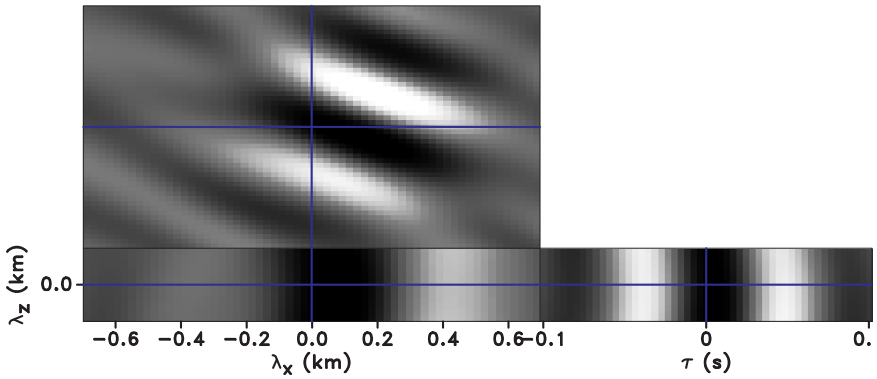


(c)

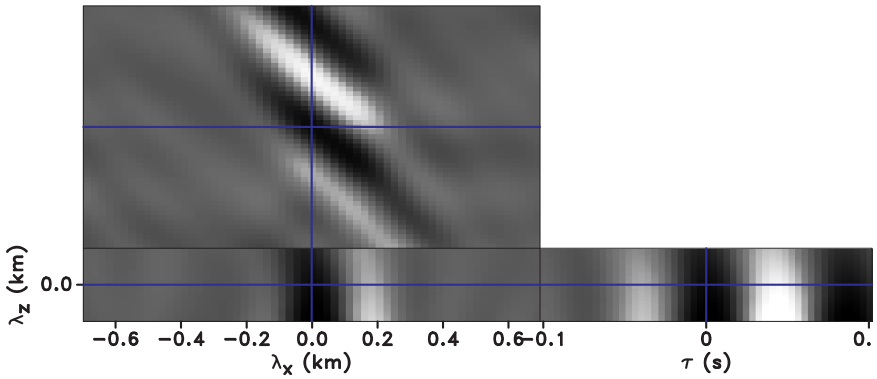


(d)

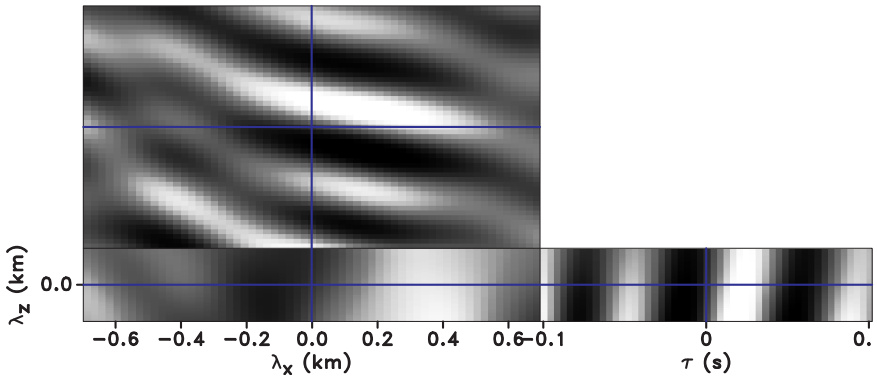
Figure 2. CIPs chosen from subsurface at (a) $x = 5.81$ km, $z = 8.44$ km (b) $x = 7.92$ km, $z = 5.87$ km, (c) $x = 22.89$ km, $z = 7.44$ km, and (d) $x = 14$ km, $z = 6.62$ km.



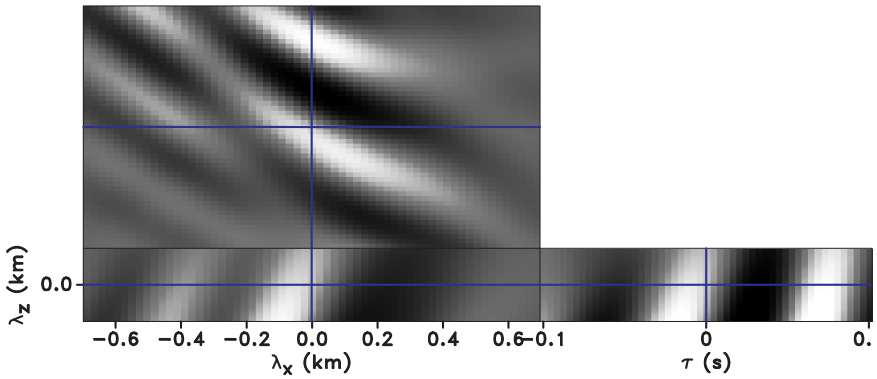
(a)



(b)

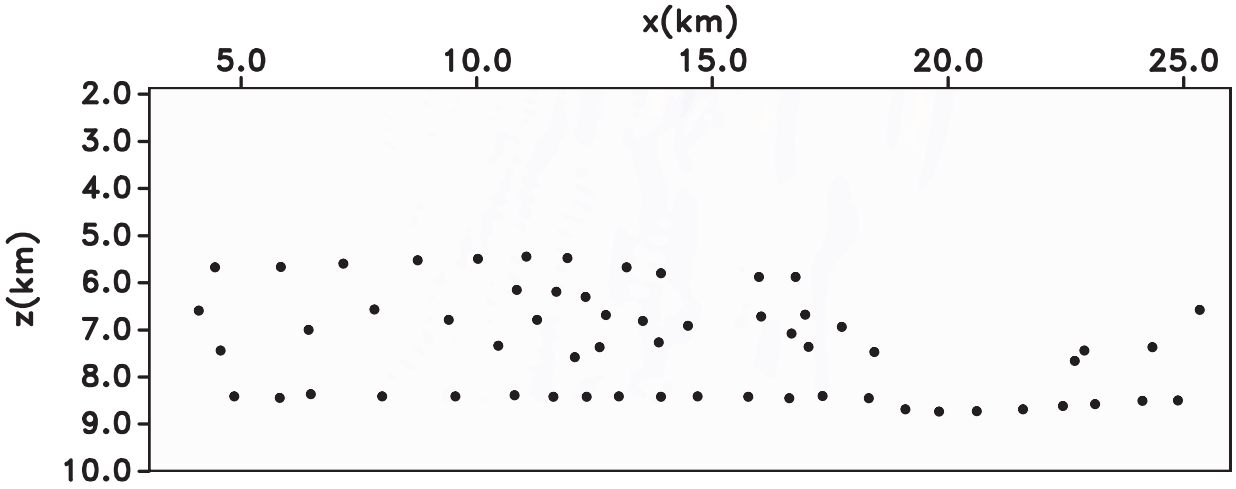


(c)

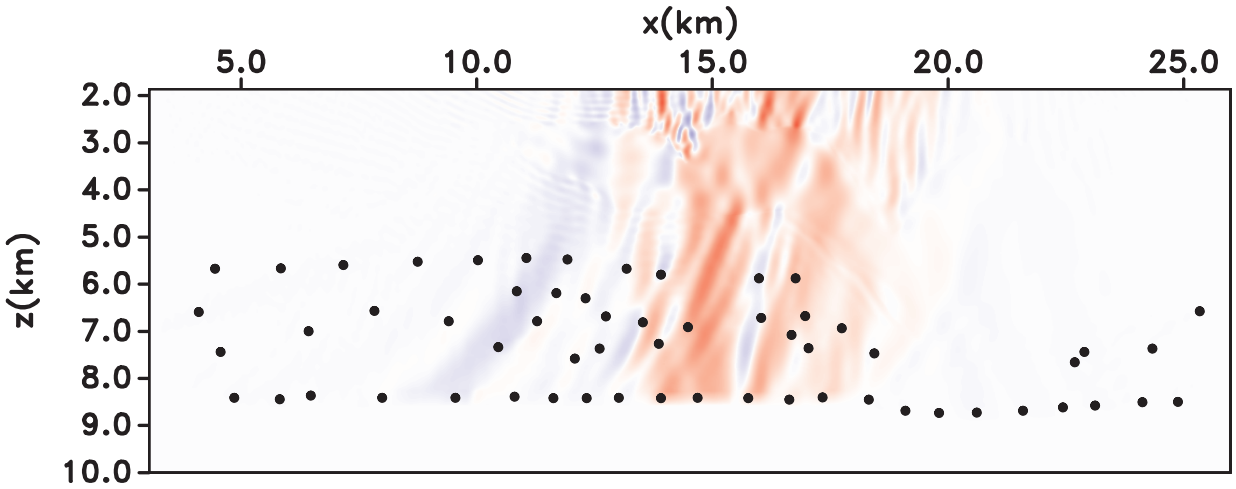


(d)

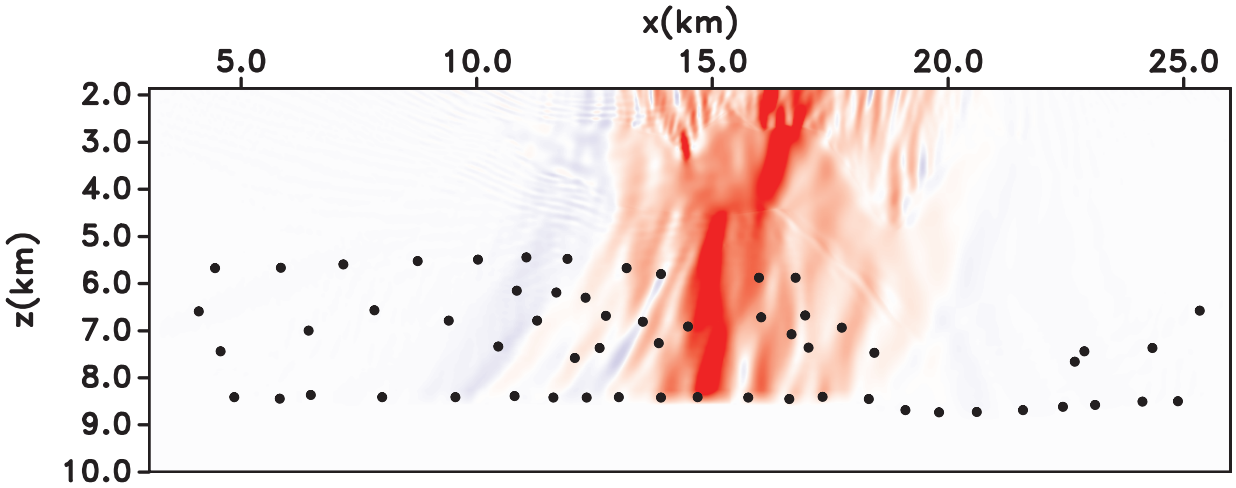
Figure 3. Image perturbation of CIPs chosen from subsurface at (a) $x = 5.81$ km, $z = 8.44$ km (b) $x = 7.92$ km, $z = 5.87$ km, (c) $x = 22.89$ km, $z = 7.44$ km, and (d) $x = 14$ km, $z = 6.62$ km.



(a)



(b)



(c)

Figure 4. Backprojection from the image perturbation of all CIPs (a) using a single frequency of 3.5 Hz and a single shot at $x = 12.3$ km, (b) using a frequency band from 3-4.5 Hz and a single shot at $x = 12.3$ km, (c) using a frequency band from 3-8.5 Hz and a single shot at $x = 12.3$ km.

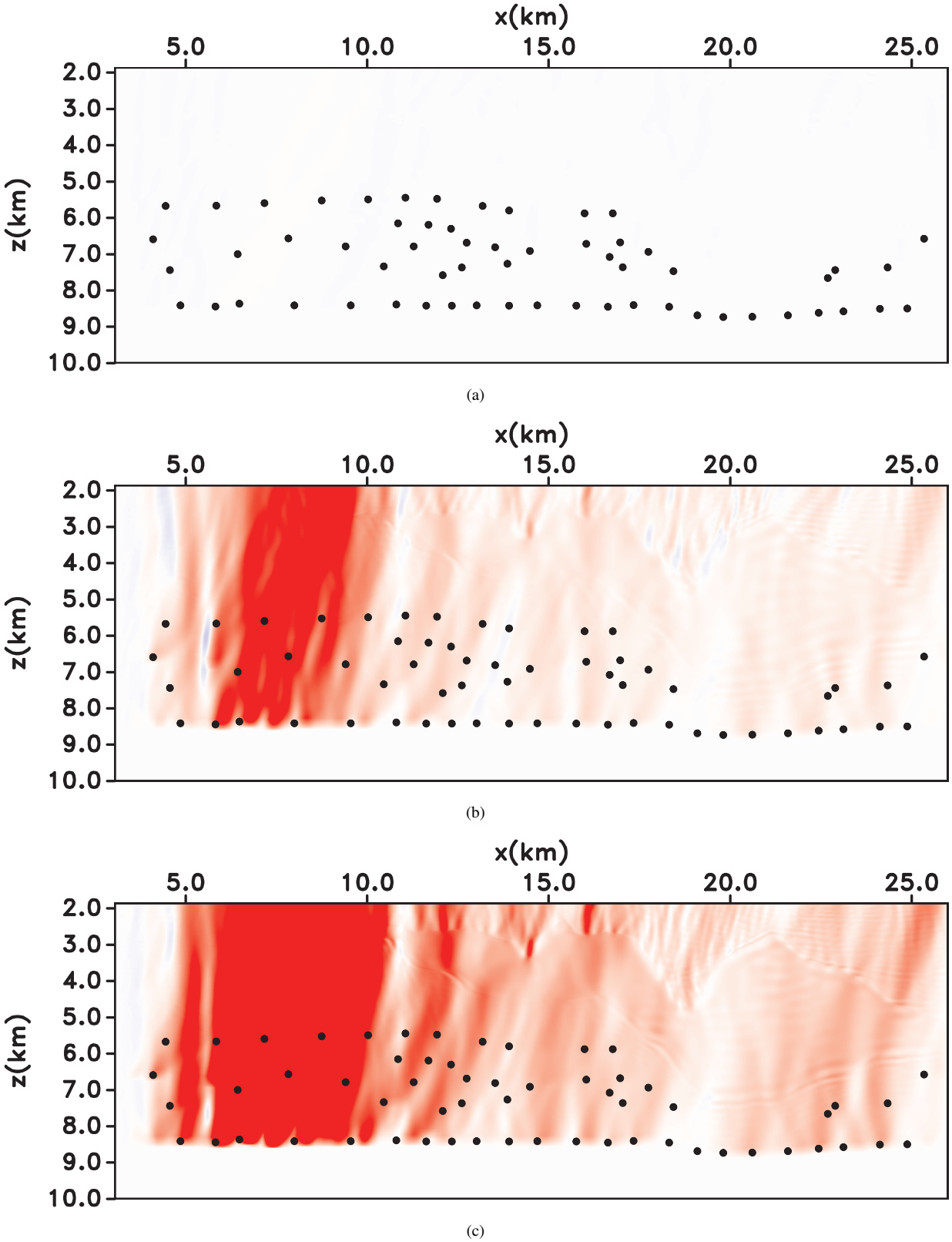


Figure 5. Backprojection from the image perturbation of all CIPs (a) using a single frequency of 3.5 Hz and all shots, (b) using a frequency band from 3-4.5 Hz and all shots, (c) using a frequency band from 3-8.5 Hz and all shots.

Since the residual moveout extends in both the space- and time-lags direction, we can also penalize the energy not concentrated at $\tau = 0$ and spread in the space-lag direction. Thus, we can formulate another penalty function as:

$$\mathbf{P}_{(\boldsymbol{\lambda},\tau)}[R] = |\mathbf{q}\cdot\boldsymbol{\lambda}|R\ , \tag{20}$$

The information about the unit vector $\mathbf{q}$ can be extracted from CIPs by measuring the slope of the event in the $\lambda_x - \lambda_z$ panel, or just by measuring the reflector slope in conventional images.

Furthermore, we can combine the penalty functions above so that we simultaneously penalize the residual moveout in both the space- and time-lags direction. The penalty function is thus

$$\mathbf{P}_{(\boldsymbol{\lambda},\tau)}[R] = R\sqrt{(v\tau)^2 + (\mathbf{q}\cdot\boldsymbol{\lambda})^2}\ . \tag{21}$$

4. 2 Event isolation for DSO

When we design the various penalty functions, we implicitly assume that there is only one event in the CIPs. As the τ axis of CIPs represents a time-shift applied to the reconstructed wavefields, however, it is possible that nearby reflections can also be imaged in a given CIPs, although these nearby events do not go through the origin of the lag space. Consider an example shown in Figure 6, which consists of three reflectors embedded in a constant medium. Figures 7(a) and 7(b) shows the CIPs constructed in the middle reflector for correct and incorrect velocity models, respectively. For the CIPs corresponding to correct velocity, Figure 7(a), the event crossing $\tau = 0$ is focused at $\boldsymbol{\lambda} = 0$. For the CIPs corresponding to incorrect velocity, Figure 7(b), the event crossing $\tau = 0$ is characterized residual moveout. Ideally, we expect a CIPs constructed with correct velocity to be focused at $\tau = 0$ and $\mathbf{q}\cdot\boldsymbol{\lambda} = 0$, which is how we have defined the penalty functions shown in Figure 8(a), 9(a) and 10(a). However, when multiple reflectors are presented in the image, multiple reflections will also be part of the CIPs. Of all these events, the only event we need to penalize is the one on which we have picked the CIPs, i.e. the event that goes through $\tau = 0$ and $\boldsymbol{\lambda} = 0$. All other events should not be penalized. Therefore, it is necessary to isolate just the event on which we have constructed the CIP for the application of the penalty function.

To isolate the irrelevant events, we simply apply a mask to CIPs and mute the events that do not contain the origin of the lag space. The key step is to construct the mask whose shape matches the trajectory of the event we want to preserve. we determine the trajectory by finding the similarity of the selected between neighboring traces using cross-correlation method. The maximum cross-correlation value corresponds to the shift between the signals at which they reach the maximum similarity. An example of the application of this procedure is shown in Figure 7(c).

For the illustration of penalty functions, we use the example shown in Figure 6. The CIPs corresponding to correct and erroneous velocities are shown in Figures 7(a) and 7(b), respectively. Figures 8(a) and 8(c) show the penalty function given by equation 19 and the CIPs in Figure 7(c) after the application of the penalty. The energy focused at $\tau = 0$ is annihilated, while the energy at nonzero τ is enhanced, thus defining the image residual. Figures 9(a) and 9(c) show the penalty function given by equation 20 and the CIPs in Figure 7(c) after the application of the penalty. In this case, we remove the energy focused at $\mathbf{q}\cdot\boldsymbol{\lambda} = 0$, and enhance the energy outside the trajectory of $\mathbf{q}\cdot\boldsymbol{\lambda} = 0$. Figures 8(a) and 8(c) show the penalty function given by equation 21 and the CIPs in Figure 7(c) after the application of the penalty. The penalized image is similar to the one shown in 9(c) in that the residual moveout spread in $\boldsymbol{\lambda}$ direction is emphasized. However, the residual moveout spread in τ direction is also enhanced.

Next, to demonstrate the characteristics of the objective functions corresponding to different penalties, we migrate data with models obtained by scaling the correct model with a constant factor ranging from 0.75 to 1.25. For each situation, we construct the CIPs then apply the penalty functions and compute the values of the corresponding objective functions. To emphasize the importance of event isolation,we apply the penalty functions to the un-masked CIPs and obtain the result as shown in Figure 11(a). The thick solid, thin solid and dashed curves correspond to penalty functions given by equation 19, 20 and 21, respectively. We see that without the event isolation, none of the objective functions reach the minimum at the correct velocity. This observation demonstrates that the objective function incorrectly penalize the image imperfections and thus fails to measure the coherency of the events in CIPs. As a result, the objective function implemented in this way cannot produce correct velocity updates. Next, we apply the penalty functions to CIPs with event isolation and obtain the result shown in Figure 11(b). Similarly, the thick solid, thin solid and dashed lines correspond to penalty functions given by equation 19, 20 and 21, respectively. We see that the objective functions for different penalties share similar characteristics. First, for all objective functions, there is only one minimum and it occurs at the correct velocity model. The uni-modal character of the curves ensures the existence of a unique global minimum for the inversion. Second, the objective function is also monotonically increasing with the velocity error except for the objective function with penalty given by equation 21. One possible reason might be that the uneven illumination of the dipping reflectors causes uneven energy in the CIPs. The monotonically increasing objective functions facilitate the use of gradient-type methods in the inversion.

5 CONCLUSIONS

We discuss image-domain wavefield-based migration velocity analysis using extended images analyzed in common-image-point gathers constructed at sparse locations in the image. We present the algorithms of migration and modeling using extended images, as well as the forward and adjoint linearized scattering using extended images. These algorithms are necessary components for the velocity updating procedure. To formulate the velocity optimization problem, a penalty function is required to measure the coherency of the images and to penalize imperfections caused by velocity errors. The penalty func-

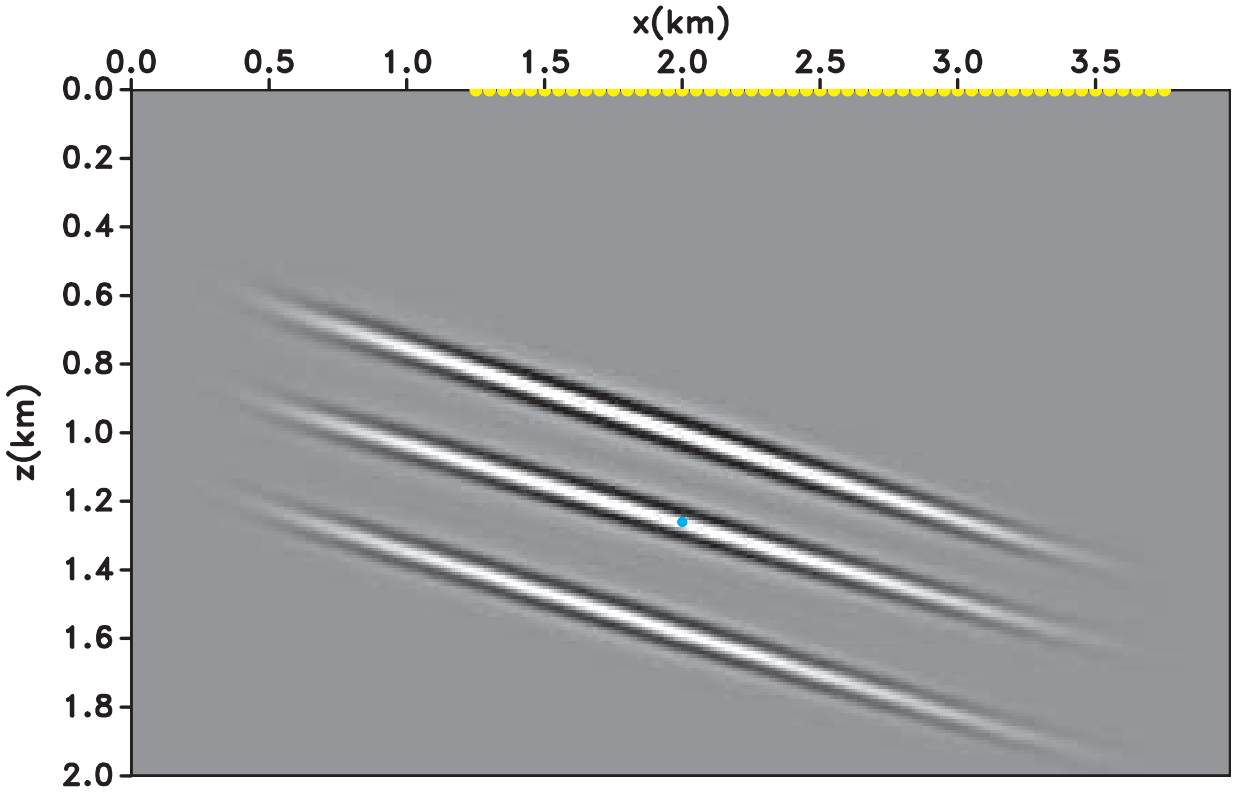
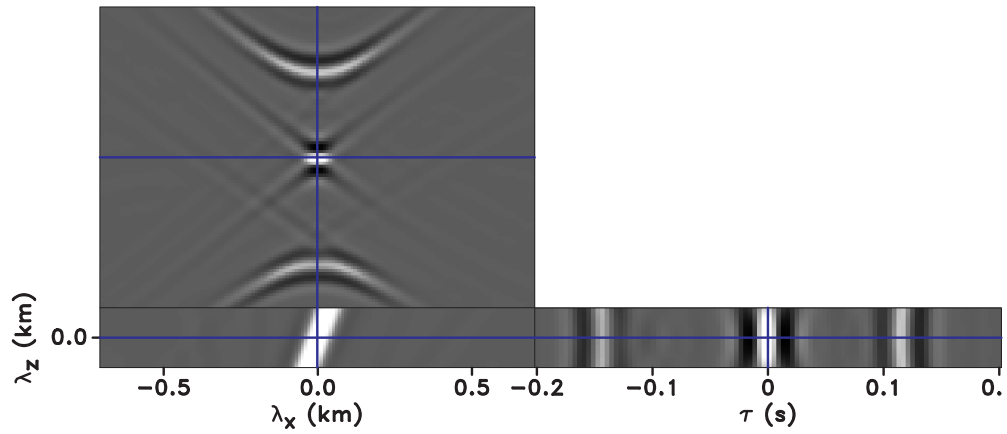
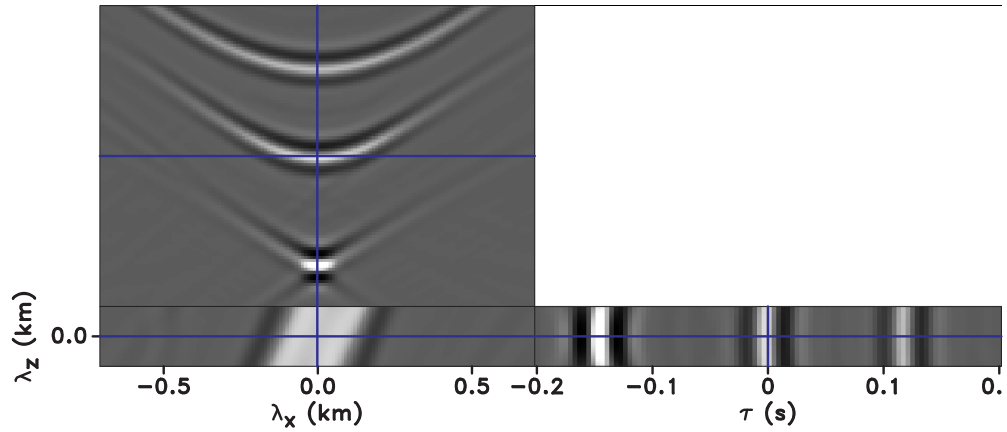


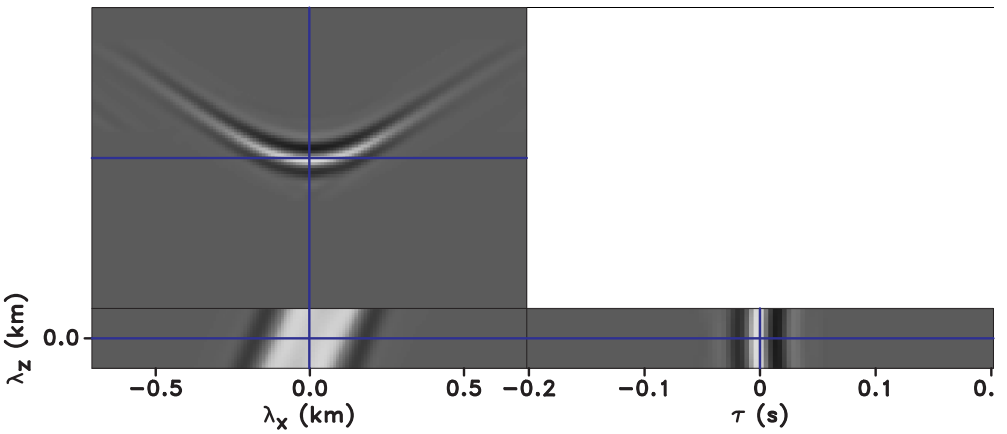
Figure 6. Synthetic example of a dipping reflector in a constant velocity model. The dot indicates the location where we construct CIPs.



(a)

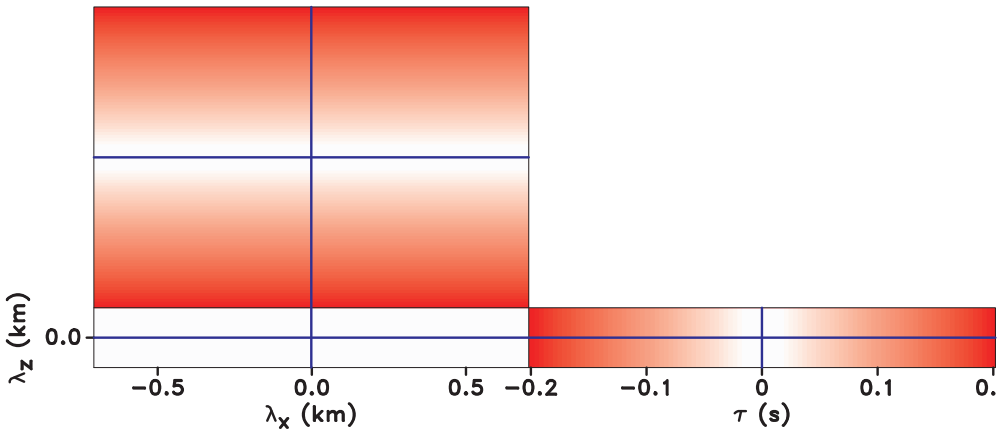


(b)

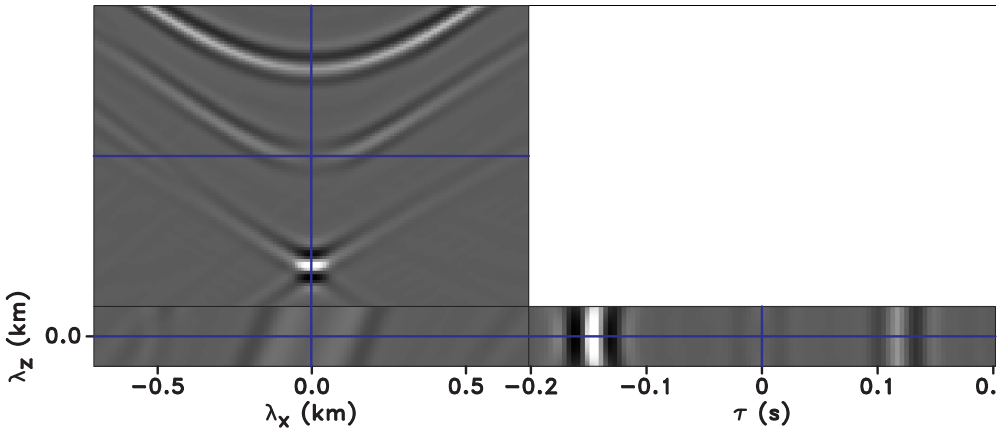


(c)

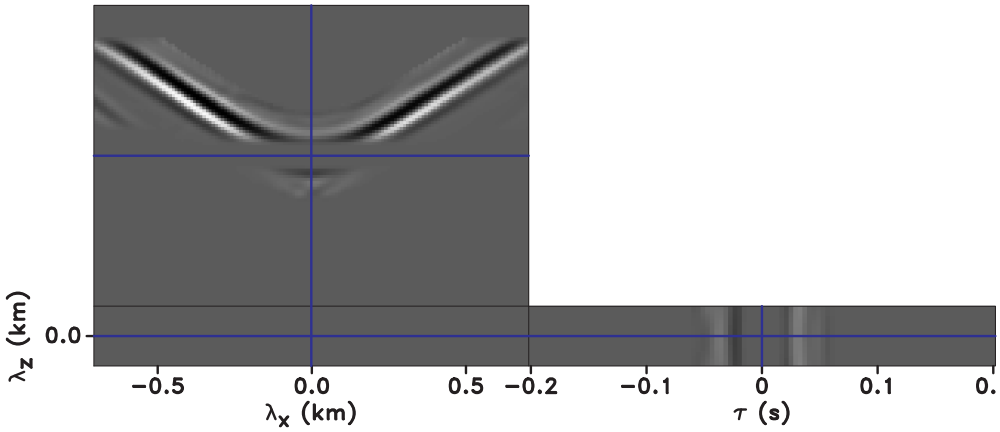
Figure 7. CIPs chosen from the middle reflector (a) for the correct velocity model. (b) for 90% of the correct velocity model. (c) after the event isolation. Notice that all the events not crossing $\tau = 0$ are eliminated.



(a)

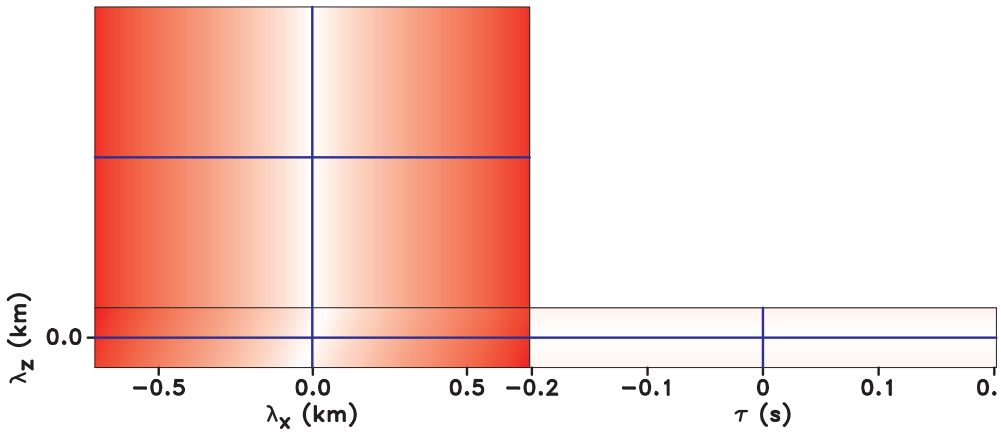


(b)

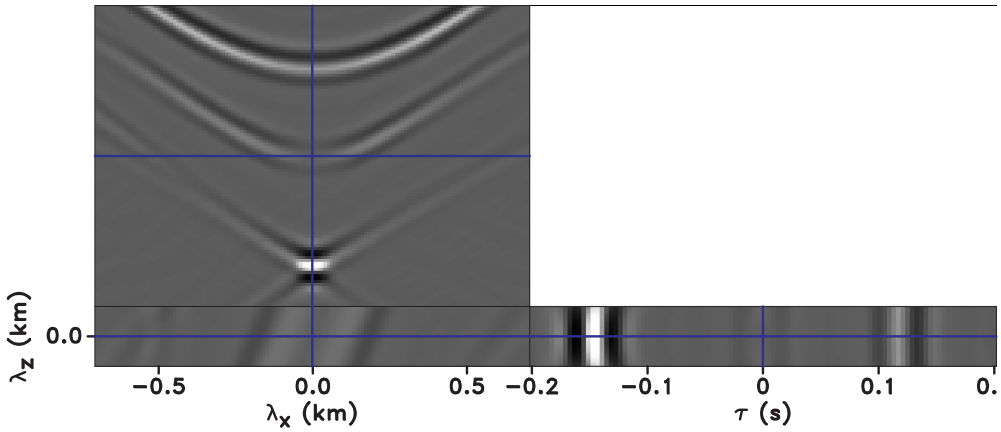


(c)

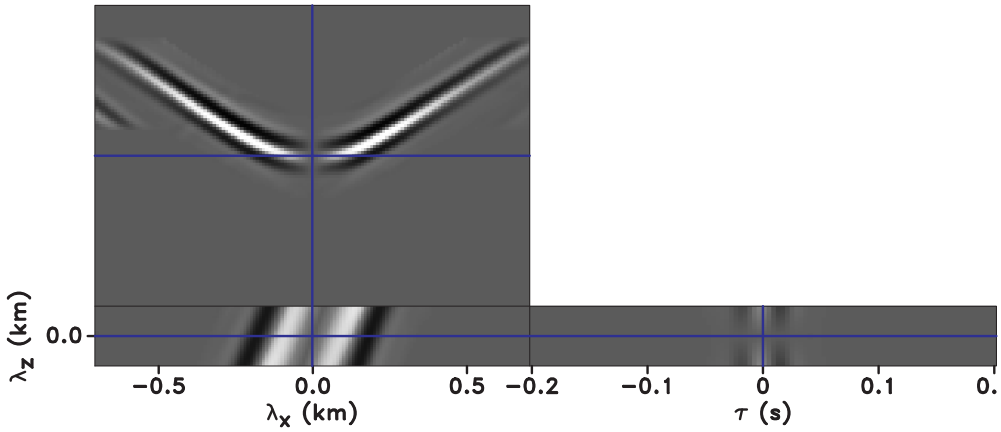
Figure 8. (a) The penalty function constructed by equation 19. (b) Penalty applied to CIPs shown in Figure 7(b) without event isolation. (c) Penalty applied to CIPs shown in Figure 7(c) with event isolation.



(a)

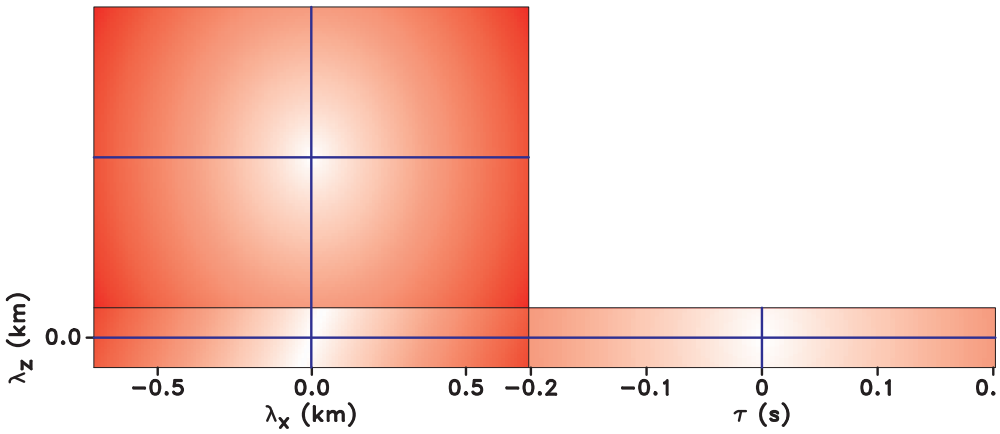


(b)

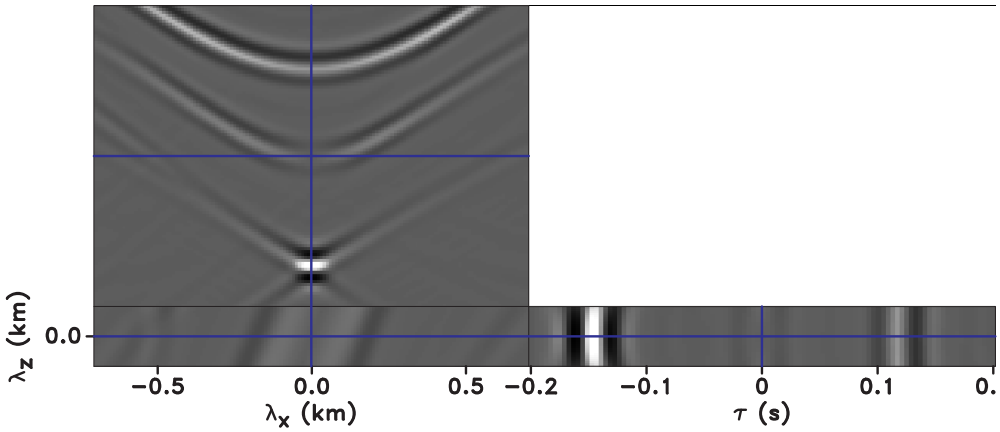


(c)

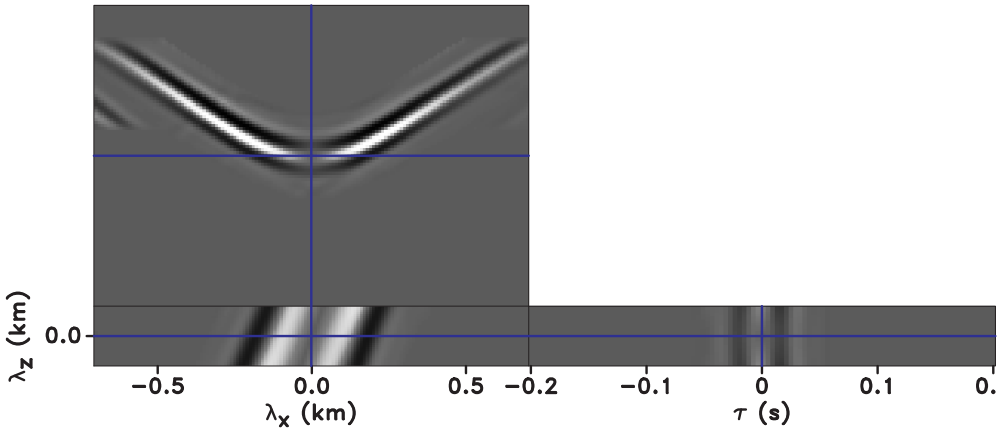
Figure 9. (a) The penalty function constructed by equation 20. (b) Penalty applied to CIPs shown in Figure 7(b) without event isolation. (c) Penalty applied to CIPs shown in Figure 7(c) with event isolation.



(a)

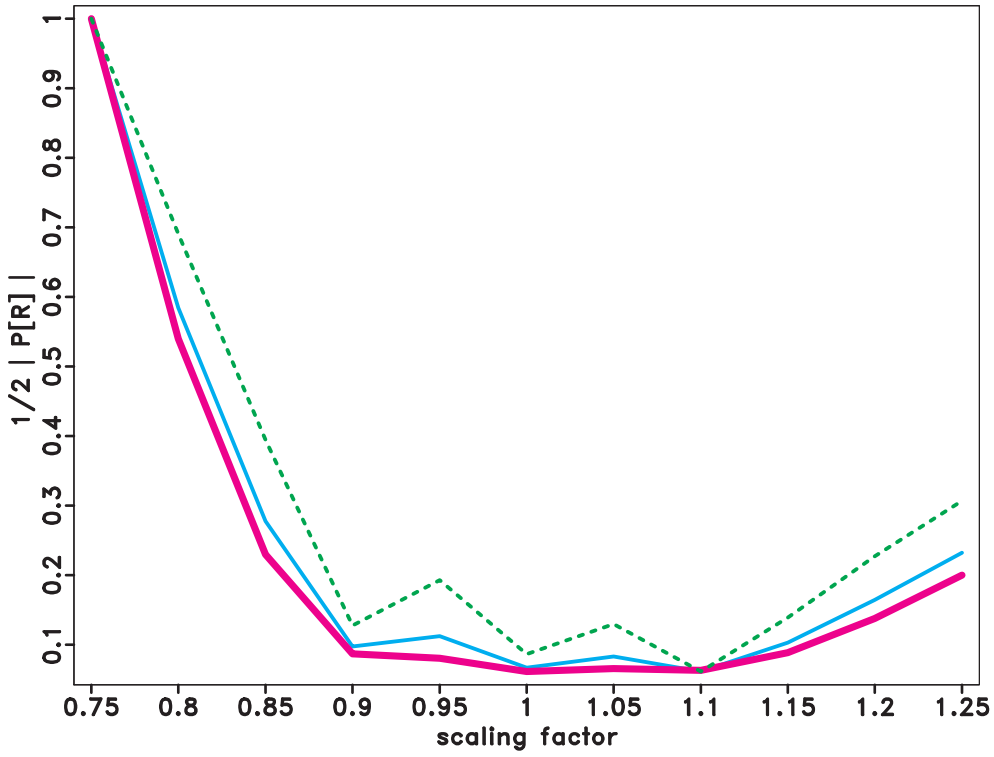


(b)

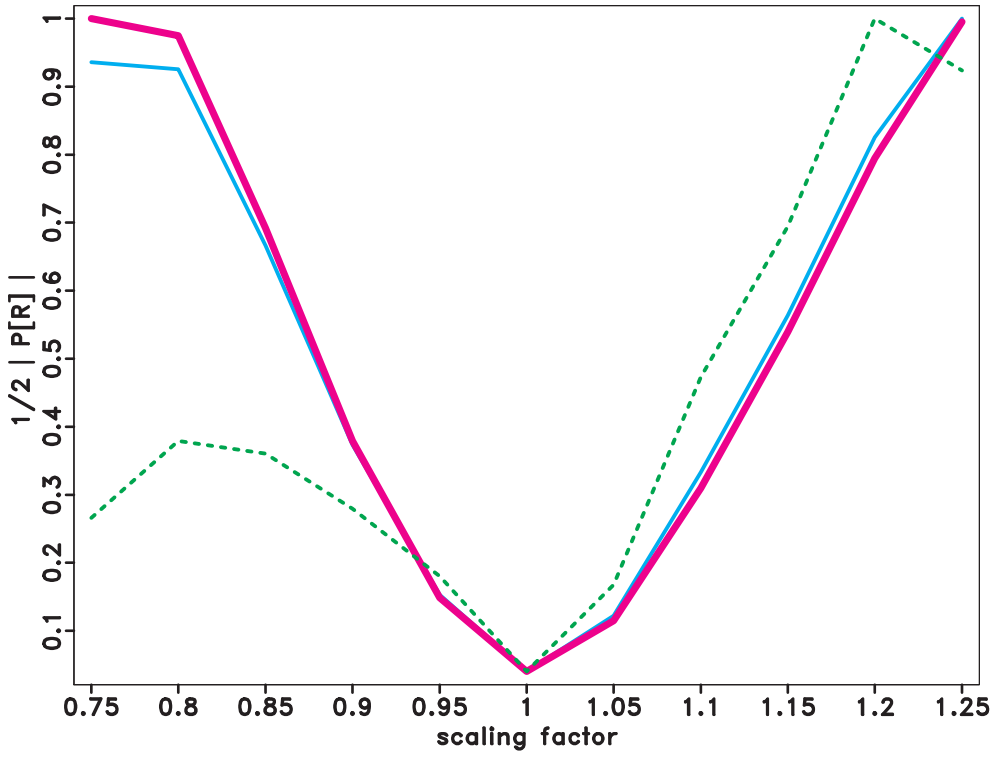


(c)

Figure 10. (a) The penalty function constructed by equation 21. (b) Penalty applied to CIPs shown in Figure 7(b) without event isolation. (c) Penalty applied to CIPs shown in Figure 7(c) with event isolation.



(a)



(b)

Figure 11. The plot of the objective functions. The dashed line corresponds to penalty given by equation 19, the thick solid line corresponds to penalty given by equation 20, the thin solid line corresponds to penalty given by equation 21. (a) The objective functions applied to CIPs without event isolation. The curves do not reach the unique minimum at the correct velocity. (b) The objective functions applied to CIPs with event isolation. All the curves reach the unique minimum at the correct velocity.

tion is designed using the semblance principle, and we present here three among the many possibilities. Synthetic examples demonstrate the uni-modal character of the objective function, which guarantees the convergence of inversion to a global minimum, and the smooth variation of the objective function around the minimum, which facilitate the use of gradient-type methods. The events from nearby reflections, however, must be isolated before the objective function is constructed since the penalty functions are designed only for the reflections at the location where the CIPs are constructed.

6 ACKNOWLEDGMENTS

The reproducible numeric examples in this paper use the Madagascar open-source software package freely available from http://www.reproducibility.org. We thank Thomas Cullison for support with CIPs location used in the examples.

REFERENCES

Al-Yahya, K. M., 1989, Velocity analysis by iterative profile migration: Geophysics, **54**, no. 06, 718–729.

Albertin, U., Sava, P., Etgen, J., and Maharramov, M., 2006, Adjoint wave-equation velocity analysis: 74th Annual International Meeting, SEG, Expanded Abstracts, 3009–3013.

Biondi, B., and Sava, P., 1999, Wave-equation migration velocity analysis: 69th Annual International Meeting, SEG, Expanded Abstracts, 1723–1726.

Bishop, T. N., Bube, K. P., Cutler, R. T., Langan, R. T., Love, P. L., Resnick, J. R., Shuey, R. T., Spindler, D. A., and Wyld, H. W., 1985, Tomographic determination of velocity and depth in laterally varying media: Geophysics, **50**, no. 06, 903–923.

Claerbout, J. F., 1985, Imaging the Earth's interior: , Blackwell Scientific Publications.

Etgen, J., Gray, S. H., and Zhang, Y., 2009, An overview of depth imaging in exploration geophysics: Geophysics, **74**, WCA5–WCA17.

Liu, Z., 1997, An analytical approach to migration velocity analysis: Geophysics, **62**, no. 04, 1238–1249.

Mora, P., 1988, Elastic wave-field inversion of reflection and transmission data: Geophysics, **53**, no. 06, 750–759.

Operto, S., Virieux, J., Amestoy, P., Giraud, J. I. L., and Ali, H., 2007, 3d finite-difference frequency-domain modeling of viscoacoustic wave propagation using a massively parallel direct solver: A feasibility study: Geophysics, **72**, no. 05, SM195–SM211.

Paffenholz, J., McLain, B., Zaske, J., and Keliher, P., 2002, Subsalt multiple attenuation and imaging: Observations from the sigsbee 2b synthetic dataset: 72nd Annual International Meeting, SEG, Expanded Abstracts, 2122–2125.

Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model: Geophysics, **64**, no. 3, 888–901.

Rickett, J., and Sava, P., 2002, Offset and angle-domain common image-point gathers for shot-profile migration: Geophysics, **67**, no. 3, 883–889.

Sava, P., and Biondi, B., 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 Fomel, S., 2006, Time-shift imaging condition in seismic migration: Geophysics, **71**, no. 6, S209–S217.

Sava, P., and Vasconcelos, I., 2009, Extended common-image-point gathers for wave-equation migration: 71st Mtg., Eur. Assoc. Expl. Geophys., Abstracts.

Sava, P., and Vlad, I., 2008, Numeric implementation of wave-equation migration velocity analysis operators: Geophysics, **73**, VE145–VE159.

Sava, P., 2003, Prestack residual migration in the frequency domain: Geophysics, **67**, no. 2, 634–640.

Shen, P., and Calandra, H., 2005, One-way waveform inversion within the framework of adjoint state differential migration: 75th Annual International Meeting, SEG, Expanded Abstracts, 1709–1712.

Shen, P., and Symes, W. W., 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, no. 05, VE49–VE59.

Shen, P., Stolk, C., and Symes, W., 2003, Automatic velocity analysis by differential semblance optimization: 73th Annual International Meeting, SEG, Expanded Abstracts, 2132–2135.

Song, Z. M., Williamson, P. R., and Pratt, R. G., 1995, Frequency-domain acoustic-wave modeling and inversion of crosshole data: Part Ii: Inversion method, synthetic experiments and real-data results: Geophysics, **60**, no. 03, 796–809.

Stork, C., 1992, Reflection tomography in the postmigrated domain: Geophysics, **57**, no. 05, 680–692.

Symes, W. W., and Carazzone, J. J., 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, no. 05, 654–663.

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**, no. 08, 1259–1266.

Tarantola, A., 1987, Inverse Problem Theory: methods for data fitting and model parameter estimation: , Elsevier.

Virieux, J., and Operto, S., 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, WCC1–WCC26.

Xie, X. B., and Yang, H., 2008, The finite-frequency sensitivity kernel for migration residual moveout and its applications in migration velocity analysis: Geophysics, **73**, no. 06, S241–S249.

Yang, T., and Sava, P., 2009, Wave-equation migration velocity analysis using extended images: 79th Annual International Meeting, SEG, Expanded abstracts.

Zhou, H., 2003, Multiscale traveltime tomography: Geophysics, **68**, no. 5, 1639–1649.