Image-space wave-equation tomography in the generalized source domain

Yaxun Tang, Claudio Guerra and Biondo Biondi

ABSTRACT

We extend the theory of image-space wave-equation tomography to the generalized source domain, where a smaller number of synthesized shot gathers are either generated by data-space phase encoding or image-space phase encoding. We demonstrate how to evaluate the wave-equation forward tomographic operator and its adjoint in these new domains. We compare the gradients of the image-space misfit functional obtained using both data-space and image-space encoded gathers with that obtained using the original shot gathers. We show that with those encoded shot gathers, we can obtain a similar gradient as that computed in the original shot-profile domain, but at lower computational cost. The saving in cost is important for putting this theory into practical applications. We illustrate our examples on a simple model with Gaussian anormalies in the subsurface.

INTRODUCTION

Wave-equation tomography has the potential to accurately estimate the velocity model in complex geological senarios where ray-based traveltime tomography often fails. Wave-equation based tomography uses band-limited wavefields instead of wide-band rays as carriers of information, thus it is robust even in the presence of strong velocity contrast and immune from multi-pathing issues. Generally speaking, wave-equation tomography can be classified into two different categaries based on the domian where it minimizes the residual. The domain can either be the data space or the image space. For the data-space based approach, it directly compares the modeled waveform with the recorded waveform, and is widely known as the waveform inversion or data-space wave-equation tomography (Tarantola, 1987; Mora, 1989; Woodward, 1992; Pratt, 1999). The main disadvantage of the data-space approach is that in complex areas, the recorded waveforms can be very complicated and are usually of low signal to noise ratio (S/N), so matching the full waveform might be extremely difficult. On the other hand, the image-space based approach, also known as image-space wave-equation tomography, minimizes the residual in the image domain obtained after migration. The image domain is often much simpler than the data domain, because even with a relatively inaccurate velocity, migration is able to (partially) collapse diffractions and enhance the S/N, thus the image-space wave-equation tompgraphy

has the potential to mitigate some of the difficulties that we encounter in the data-space based approach. Another advantage of the image-space based approach is that the more efficient one-way wave-equation extrapolator can be utilized, whereas in waveform inversion, the one-way propagator is difficult (if not imposible) to find its usage because of its inability to model the multiple arrivales, though some tweaks can be employed so that the one-way propagater can be applied in the case of turning wave tomography (Shragge, 2007).

However, despite of its theoretical advantages, image-space wave-equation tomography is still computationally challenging. Each iteration of tomographic velocity updating is computationally expensive and often converges slowly. Practial applications are still rare and small in scale (Biondi and Sava, 1999; Shen et al., 2005; Albertin et al., 2006). The goal of this paper is to extend the theory of image-space wave-equation tomography from the conventional shot-profile domain (Shen, 2004) to the generalized source domain, where a smaller number of synthesized shot gathers make the tomographic velocity update substentially faster. The generalized source domain can be obtained either by data-space phase encoding or image-space phase encoding. By data-space phase encoding, we mean that the synthesized shot gathers are obtained by linear combination of the original shot gather after some kind of phase encoding, in particular, here we mainly consider plane-wave phase encoding (Whitmore, 1995; Zhang et al., 2005; Duquet and Lailly, 2006; Liu et al., 2006) and random phase encoding (Romero et al., 2000). As the encoding process is done in the data space, we call it data-space phase encoding. By image-space phase encoding, we mean that the synthesized gathers are obtained by prestack exploding reflector modeling (Biondi, 2006, 2007; Guerra and Biondi, 2008b). where several subsurface-offset-domain common-image gathers (SODCIGs) are simultaneously demigrated to generate area source and area receiver gathers. To attenuate the cross-talk, the SODCIGs have to be encoded, e.g., by random phase encoding. This encoding process is done in the image space, so we call it image-space phase encoding. We show that in these generalized source domains, we can obtain gradients, which are used for updating the velocity model, similar to that obtained in the original shot-profile domain, but with less computational cost.

This paper is organized as follows: We first briefly review the theory of image-space wave-equation tomography, we review the formulation of the objective function and its gradient; then we discuss how to evaluate the forward tomographic operator and its adjoint in the orignal shot-profile domain, the latter is an important component in computing the gradient; then we extend the theory to the generalized source domain; finally, we show examples on a simple synthetic model.

IMAGE-SPACE WAVE-EQUATION TOMOGRAPHY

Image-space wave-equation tomography is a non-linear inverse problem, it is based on an optimization that tries to find an optimal background slowness that minimizes the residual field, $\mathbf{\Delta I}$, defined in the image space. The residual field is derived from

the background image, $\mathbf{I}$, which is computed with a background slowness (or the current estimate of the slowness). The residual field measures the correctness of the background slowness, its minimum (under some norm, e.g. ℓ_2) is achieved when a correct background slowness has been used for migration. There are many choices of the residual field, such as residual moveout in the Angle-Domain Common-Image Gathers (ADCIGs), differential semblence in the ADCIGs, reflection angle stacking power (in this case we have to maximize the residual field, or minimize the minus stacking power), and etc.. Here we follow a similar definition as in Biondi (2008), and define a general form of the residual field as follows:

$$\mathbf{\Delta I} = \mathbf{I} - \mathbf{F}(\mathbf{I}), \tag{1}$$

where $\mathbf{F}$ is a focusing operator, which measures the focusing of the migrated image. For example, in the wave-equation migration velocity analysis (WEMVA) method (Sava, 2004), the focusing operator is the linearized residual migration operator defined as follows:

$$\mathbf{F}(\mathbf{I}) = \mathbf{R}[\rho]\mathbf{I} \approx \mathbf{I} + \mathbf{K}[\Delta\rho]\mathbf{I}, \tag{2}$$

where ρ is the ratio between the background slowness $\widehat{\mathbf{s}}$ and the true slowness $\mathbf{s}$, and $\Delta\rho = 1 - \rho = 1 - \frac{\widehat{\mathbf{s}}}{\mathbf{s}}$; $\mathbf{R}[\rho]$ is the residual migration operator; $\mathbf{K}[\Delta\rho]$ is the differential residual migration operator defined as follws:

$$\mathbf{K}[\Delta\rho] = \Delta\rho \left.\frac{\partial \mathbf{R}[\rho]}{\partial \rho}\right|_{\rho=1}. \tag{3}$$

The linear operator $\mathbf{K}[\Delta\rho]$ applies different phase rotations to the image for different reflection angle and geological dips (Biondi, 2008). In the Differential Semblance Optimization (DSO) method (Shen, 2004), the focusing operator takes the following form:

$$\mathbf{F}(\mathbf{I}) = (\mathbf{1} - \mathbf{O})\,\mathbf{I}, \tag{4}$$

where $\mathbf{O}$ is the DSO operator either in the subsurface offset domain or in the angle domain. Subsurface offset domain DSO focuses the energy at zero offset while angle domain DSO flattens the ADCIGs.

In general, if we choose ℓ_2 norm, the objective function to minimize can be written as follows:

$$J = \frac{1}{2}||\mathbf{\Delta I}||_2 = \frac{1}{2}||\mathbf{I} - \mathbf{F}(\mathbf{I})||_2, \tag{5}$$

where $||\cdot||_2$ stands for the ℓ_2 norm. Gradient based optimization techniques such as Quasi-Newton method and Conjugate Gradient method can be used to minimize the objective function J, which is non-linear because the image $\mathbf{I}$ is a non-linear function of the slowness $\mathbf{s}$. To use the gradient based methods, we have to provide the gradient of J with respect to the slowness $\mathbf{s}$. The gradient reads

$$\nabla J = \Re\left(\left(\frac{\partial \mathbf{I}}{\partial \mathbf{s}} - \frac{\partial \mathbf{F}(\mathbf{I})}{\partial \mathbf{s}}\right)'(\mathbf{I} - \mathbf{F}(\mathbf{I}))\right), \tag{6}$$

where $'$ denotes the adjoint. For the DSO method, the linear operator $\mathbf{O}$ is independent of the slowness, so we have

$$\frac{\partial \mathbf{F}(\mathbf{I})}{\partial \mathbf{s}} = (\mathbf{1} - \mathbf{O})\frac{\partial \mathbf{I}}{\partial \mathbf{s}}. \tag{7}$$

Substituting Equations 4 and 7 into Equation 6 yields

$$\nabla J_{\mathrm{DSO}} = \Re\left(\left(\frac{\partial \mathbf{I}}{\partial \mathbf{s}}\right)' \mathbf{O}'\mathbf{O}\mathbf{I}\right) \tag{8}$$

For the WEMVA method, the gradient is slightly more complicated, because in this case, the focusing operator is also dependent on the slowness $\mathbf{s}$. However, one can simplify it by assuming that the focusing operator is performed on the background image $\widehat{\mathbf{I}}$ instead of $\mathbf{I}$, i.e.,

$$\mathbf{F}(\widehat{\mathbf{I}}) = \widehat{\mathbf{I}} + \mathbf{K}[\widehat{\Delta\rho}]\widehat{\mathbf{I}}, \tag{9}$$

With this assumption, we get the "classic" WEMVA gradient as follows:

$$\nabla J_{\mathrm{WEMVA}} = \Re\left(-\left(\frac{\partial \mathbf{I}}{\partial \mathbf{s}}\right)' \mathbf{K}[\widehat{\Delta\rho}]\widehat{\mathbf{I}}\right) \tag{10}$$

The complete WEMVA gradient can also be derived following the method described by Biondi (2008).

No matter which gradient we choose to back-project the slowness perturbeation, we have to evaluate the adjoint of the linear operator $\frac{\partial \mathbf{I}}{\partial \mathbf{s}}$, which defines a linear mapping from the slowness perturbation $\mathbf{\Delta s}$ to the image perturbation $\mathbf{\Delta I}$. This is easy to see by expanding the image $\mathbf{I}$ around a background slowness $\widehat{\mathbf{s}}$ as follows:

$$\mathbf{I} = \widehat{\mathbf{I}} + \frac{\partial \mathbf{I}}{\partial \mathbf{s}}(\mathbf{s} - \widehat{\mathbf{s}}) + \cdots, \tag{11}$$

where $\widehat{\mathbf{I}}$ is the background image computed with the background slowness $\widehat{\mathbf{s}}$. Keeping only the zero and first order terms, we get the linearized operator $\frac{\partial \mathbf{I}}{\partial \mathbf{s}}$ defined as follows:

$$\mathbf{\Delta I} = \frac{\partial \mathbf{I}}{\partial \mathbf{s}}\mathbf{\Delta s} = \mathbf{T}\mathbf{\Delta s}, \tag{12}$$

where $\mathbf{\Delta I} = \mathbf{I} - \widehat{\mathbf{I}}$ and $\mathbf{\Delta s} = \mathbf{s} - \widehat{\mathbf{s}}$. $\mathbf{T} = \frac{\partial \mathbf{I}}{\partial \mathbf{s}}$ is the wave-equation tomographic operator. The tomographic operator can either be defined in the source and receiver domain (Sava, 2004) or in the shot-profile domain (Shen, 2004). In next section we follow a similar approach disscussed by Shen (2004) and review the forward and adjoint tomographic operator in the shot profile domain, in the subsequent sections, we generalize the expression of the tomographic operator to generalized source domains.

THE TOMOGRAPHIC OPERATOR IN THE SHOT-PROFILE DOMAIN

For the conventional shot-profile migration, both source and receiver wavefields are downward continued with the following one-way wave equations (Claerbout, 1971):

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) D(\mathbf{x}, \mathbf{x}_s, \omega) = 0 \\ D(x, y, z = 0, \mathbf{x}_s, \omega) = \overline{f_s(\omega)\delta(\mathbf{x} - \mathbf{x}_s)} \end{cases}, \tag{13}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) U(\mathbf{x}, \mathbf{x}_s, \omega) = 0 \\ U(x, y, z = 0, \mathbf{x}_s, \omega) = Q(x, y, z = 0, \mathbf{x}_s, \omega) \end{cases}, \tag{14}$$

where the overline stands for complex conjugate; $D(\mathbf{x}, \mathbf{x}_s, \omega)$ is the source wavefield for a single frequency ω at image point $\mathbf{x} = (x, y, z)$ with the source located at $\mathbf{x}_s = (x_s, y_s, 0)$, $U(\mathbf{x}, \mathbf{x}_s, \omega)$ is the receiver wavefield for a single frequency ω at image point $\mathbf{x}$ for the source located at $\mathbf{x}_s$. $s(\mathbf{x})$ is the slowness at $\mathbf{x}$, $\mathbf{k} = (k_x, k_y)$ is the spatial wavenumber vector, $f_s(\omega)$ is the freqnency dependent source signature and $\overline{f_s(\omega)\delta(\mathbf{x} - \mathbf{x}_s)}$ defines the point source function at $\mathbf{x}_s$, which serves as the boundary condition of equation 13. $Q(x, y, z = 0, \mathbf{x}_s, \omega)$ is the shot gather for the shot located at $\mathbf{x}_s$, which serves as the boundary condition of equation 14. To produce the image, the following cross-correlation imaging condition is used:

$$I(\mathbf{x}, \mathbf{h}) = \sum_{\mathbf{x}_s} \sum_{\omega} D(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega) U(\mathbf{x} + \mathbf{h}, \mathbf{x}_s, \omega), \tag{15}$$

where $\mathbf{h} = (h_x, h_y, h_z)$ is the subsurface half offset.

The perturbed image at some image point can be derived by a simple application of the chain rule to equation 15:

$$\begin{aligned} \Delta I(\mathbf{x}, \mathbf{h}) \;=\; & \sum_{\mathbf{x}_s} \sum_{\omega} \Big(\Delta D(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega) \widehat{U}(\mathbf{x} + \mathbf{h}, \mathbf{x}_s, \omega) + \\ & \widehat{D}(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega) \Delta U(\mathbf{x} + \mathbf{h}, \mathbf{x}_s, \omega) \Big), \end{aligned} \tag{16}$$

where $\widehat{D}(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega)$ and $\widehat{U}(\mathbf{x} + \mathbf{h}, \mathbf{x}_s, \omega)$ are the background source and receiver wavefields computed with the background slowness $\widehat{s}(\mathbf{x})$; $\Delta D(\mathbf{x} - \mathbf{h}, \mathbf{x}_s, \omega)$ and $\Delta U(\mathbf{x} + \mathbf{h}, \mathbf{x}_s, \omega)$ are the perturbed source wavefield and perturbed receiver wavefield, which are the results of the slowness perturbation $\Delta s(\mathbf{x})$. The perturbed source and receiver wavefields satisfy the following linearized (with respect to slowness) one-way wave equations, respectively(see Appendix A for derivations)

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 \widehat{s}^2(\mathbf{x}) - |\mathbf{k}|^2}\right) \Delta D(\mathbf{x}, \mathbf{x}_s, \omega) = \frac{-i\omega \Delta s(\mathbf{x})}{\sqrt{1 - \frac{|\mathbf{k}|^2}{\omega^2 \widehat{s}^2(\mathbf{x})}}} \widehat{D}(\mathbf{x}, \mathbf{x}_s, \omega) \\ \Delta D(x, y, z = 0, \mathbf{x}_s, \omega) = 0 \end{cases}, \tag{17}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2\widehat{s}^2(\mathbf{x}) - |\mathbf{k}|^2}\right)\Delta U(\mathbf{x},\mathbf{x}_s,\omega) = \frac{-i\omega\Delta s(\mathbf{x})}{\sqrt{1-\frac{|\mathbf{k}|^2}{\omega^2\widehat{s}^2(\mathbf{x})}}}\widehat{U}(\mathbf{x},\mathbf{x}_s,\omega) \\ \Delta U(x,y,z=0,\mathbf{x}_s,\omega) = 0 \end{cases} . \tag{18}$$

Equations 17 and 18 are linear with respect to slowness perturbation, thus they define linear mappings from the slowness perturbation $\Delta s(\mathbf{x})$ to the perturbed source wavefield $\Delta D(\mathbf{x},\mathbf{x}_s,\omega)$ and the perturbed receiver wavefield $\Delta U(\mathbf{x},\mathbf{x}_s,\omega)$, respectively. Therefore, recursively solving equations 17 and 18 give us the perturbed source and receiver wavefields. The perturbed source and receiver wavefields are then used in equation 16 to generate the perturbed image $I(\mathbf{x},\mathbf{h})$, where the background source and receiver wavefields are precomputed by solving the recursive equations 13 and 14 with a background slowness $\widehat{s}(\mathbf{x})$. Appendix B gives a more detailed matrix representation of how to evaluate the forward tomographic operator $\mathbf{T}$.

To compute the adjoint tomographic operator $\mathbf{T}'$, we first apply the adjoint of the imaging condition in equation 16 to get the perturbed source and receiver wavefields $\Delta D(\mathbf{x},\mathbf{x}_s,\omega)$ and $\Delta U(\mathbf{x},\mathbf{x}_s,\omega)$ as follows

$$\Delta D(\mathbf{x},\mathbf{x}_s,\omega) = \sum_{\mathbf{h}} \Delta I(\mathbf{x},\mathbf{h})\overline{\widehat{U}(\mathbf{x}+\mathbf{h},\mathbf{x}_s,\omega)} \tag{19}$$

$$\Delta U(\mathbf{x},\mathbf{x}_s,\omega) = \sum_{\mathbf{h}} \Delta I(\mathbf{x},\mathbf{h})\overline{\widehat{D}(\mathbf{x}-\mathbf{h},\mathbf{x}_s,\omega)} \tag{20}$$

Then we solve the adjoint equations of 17 and 18 to get the slowness perturbation $\Delta s(\mathbf{x})$. Again, in order to solve the adjoint equations of 17 and 18, the background source wavefield $\widehat{D}(\mathbf{x},\mathbf{x}_s,\omega)$ and the background receiver wavefield $\widehat{U}(\mathbf{x},\mathbf{x}_s,\omega)$ have to be computed in advance. Appendix B gives a more detailed matrix representation of how to evaluate the adjoint tomographic operator: $\mathbf{T}^{'}$.

TOMOGRAPHY WITH THE ENCODED WAVEFIELDS

It is clear from previous sections that the cost for computing the gradient of the objective function J in the original shot-profile domain is at least twice the cost as that of a shot-profile migration, because to compute the perturbed wavefields, the background wavefields are required. As minimizing the objective function J requires a considerable number of gradient and function evaluations, image-space wave-equation tomography in the conventional shot-profile domain seems to be infeasible for large scale 3-D applications, even with nowadays computer resources. To save the cost and to make this powerful method more practical, we extend the wave-eqaution tomography theory to generalized source domains, where a smaller number of synthesized shot gathers are used for computing the gradient. We discuss two different strategies generate the generalized shot gathers, i.e., the data-space phase encoding and the image-space phase encoding. Both of which can achieve a considerable amount of

data reductions while still keeping the necessary kinematic information for velocity analysis.

Data-space encoded wavefields

The data-space encoded shot gathers are obtained by linear combination of the original shot gathers after phase encoding. For simplicity, we mainly consider the plane-wave phase-encoding and the random phase-encoding. Because of the linearity of the one-way wave equation respect to the wavefield, the encoded source and receiver wavefields also satisfy the same one-way wave equations defined by Equations 13 and 14, but with different boundary conditions:

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) \widetilde{D}(\mathbf{x}, \mathbf{p}_s, \omega) = 0 \\ \widetilde{D}(x, y, z = 0, \mathbf{p}_s, \omega) = \sum_{\mathbf{x}_s} \overline{f_s(\omega)\delta(\mathbf{x} - \mathbf{x}_s)\alpha(\mathbf{x}_s, \mathbf{p}_s, \omega)} \end{cases}, \tag{21}$$

and

$$\begin{cases} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) \widetilde{U}(\mathbf{x}, \mathbf{p}_s, \omega) = 0 \\ \widetilde{U}(x, y, z = 0, \mathbf{x}_s, \omega) = \sum_{\mathbf{x}_s} Q(x, y, z = 0, \mathbf{x}_s, \omega)\alpha(\mathbf{x}_s, \mathbf{p}_s, \omega) \end{cases}, \tag{22}$$

where $\widetilde{D}(\mathbf{x}, \mathbf{p}_s, \omega)$ and $\widetilde{U}(\mathbf{x}, \mathbf{p}_s, \omega)$ are the encoded source and receiver wavefields respectively; $\alpha(\mathbf{x}_s, \mathbf{p}_s, \omega)$ is the phase-encoding function. In the case of plane-wave phase encoding, $\alpha(\mathbf{x}_s, \mathbf{p}_s, \omega)$ is defined as

$$\alpha(\mathbf{x}_s, \mathbf{p}_s, \omega) = e^{i\omega \mathbf{p}_s \mathbf{x}_s}, \tag{23}$$

where $\mathbf{p}_s$ is the ray parameter for the source plane waves on the surface. In the case of random phase encoding, the phase function is

$$\alpha(\mathbf{x}_s, \mathbf{p}_s, \omega) = e^{i\gamma(\mathbf{x}_s, \mathbf{p}_s, \omega)}, \tag{24}$$

where $\gamma(\mathbf{x}_s, \mathbf{p}_s, \omega)$ is a random sequence in $\mathbf{x}_s$ and ω. $\mathbf{p}_s$ defines the index of different realizations of the random sequence (Tang, 2008). The final image is obtained by applying the cross-correlation imaging condition and summing all $\mathbf{p}_s$'s:

$$I_{\text{de}}(\mathbf{x}, \mathbf{h}) = \sum_{\mathbf{p}_s} \sum_{\omega} \widetilde{D}(\mathbf{x} - \mathbf{h}, \mathbf{p}_s, \omega)\widetilde{U}(\mathbf{x} + \mathbf{h}, \mathbf{p}_s, \omega). \tag{25}$$

It has been shown by Etgen (2005) and Liu et al. (2006) that, in the case of plane-wave phase-encoding migration, by stacking a considerable number of $\mathbf{p}_s$, the plane-wave migrated image is almost identical to the shot-profile migrated image. If the original shots are well sampled, the number of plane waves required for migration is generally much smaller than the number of the original shot gathers (Etgen, 2005). On the other hand, random-phase encoding function is not very effective in attenuating the cross-talk, especially when many sources are simultaneously encoded (Romero et al.,

2000; Tang, 2008). However, if many realizations of the random sequences are used, the final stacked image would also be approximately the same as the shot-profile migrated image. Therefore, the following relation approximately holds, i.e., with the data-space encoded gathers, we obtain a similar image as that computed by the more expensive shot-profile migration:

$$I(\mathbf{x},\mathbf{h}) \approx I_{\rm de}(\mathbf{x},\mathbf{h}). \tag{26}$$

From Equation 25, the perturbed image can be easily obtained as follows:

$$\begin{aligned}\Delta I_{\rm de}(\mathbf{x},\mathbf{h}) &= \sum_{\mathbf{p}_s}\sum_{\omega}\Big(\Delta\widetilde{D}(\mathbf{x}-\mathbf{h},\mathbf{p}_s,\omega)\widehat{\widetilde{U}}(\mathbf{x}+\mathbf{h},\mathbf{p}_s,\omega)+ \\ &\quad \widehat{\widetilde{D}}(\mathbf{x}-\mathbf{h},\mathbf{p}_s,\omega)\Delta\widetilde{U}(\mathbf{x}+\mathbf{h},\mathbf{p}_s,\omega)\Big),\end{aligned} \tag{27}$$

where $\widehat{\widetilde{D}}(\mathbf{x},\mathbf{p}_s,\omega)$ and $\widehat{\widetilde{U}}(\mathbf{x},\mathbf{p}_s,\omega)$ are the data-space encoded background source and receiver wavefields; $\Delta\widetilde{D}(\mathbf{x},\mathbf{p}_s,\omega)$ and $\Delta\widetilde{U}(\mathbf{x},\mathbf{p}_s,\omega)$ are the perturbed source and receover wavefields in the data-space phase-encoding domain, which satisfy the perturbed one-way wave equations defined by Equations 17 and 18. The tomographic operator $\mathbf{T}$ and its adjoint $\mathbf{T}'$ can be implemented in a similar mannar as discussed in Appendix B and C by replacing the original wavefileds with the data-space phase encoded wavefields.

Image-space encoded wavefields

The image-space phase encoded gathers are obtained using the prestack exploding reflector modeling method introduced by Biondi (2006) and Biondi (2007). The general idea of this method is to model the data and the corresponding source function that are related to only one event in the subsurface, where a single unfocused SODCIG (obtained with an inaccurate velocity model) is used as the initial condition for the recursive upward continuation by the following one-way wave equations:

$$\begin{cases}\left(\frac{\partial}{\partial z}-i\sqrt{\omega^2\hat{s}^2(\mathbf{x})-|\mathbf{k}|^2}\right)Q_D(\mathbf{x},\omega;x_m,y_m)=I_D(\mathbf{x},\mathbf{h};x_m,y_m)\\ Q_D(x,y,z=z_{\max},\omega;x_m,y_m)=0\end{cases}, \tag{28}$$

and

$$\begin{cases}\left(\frac{\partial}{\partial z}-i\sqrt{\omega^2\hat{s}^2(\mathbf{x})-|\mathbf{k}|^2}\right)Q_U(\mathbf{x},\omega;x_m,y_m)=I_U(\mathbf{x},\mathbf{h};x_m,y_m)\\ Q_U(x,y,z=z_{\max},\omega;x_m,y_m)=0\end{cases}, \tag{29}$$

where $I_D(\mathbf{x},\mathbf{h};x_m,y_m)$ and $I_U(\mathbf{x},\mathbf{h};x_m,y_m)$ are the isolated SODCIGs at horizontal location (x_m,y_m) for a single reflector, which are suitable for the initial conditions for the source and receiver wavefields, respectively. They are obtained by rotating the original unfocused SODCIGs according to the apparent geological dip of the reflector.

This rotation maintains the correct velocity information needed for migration velocity analysis, especially for dipping reflectors (Biondi, 2007). By collecting the wavefields at the surface, we obtain the areal source data $Q_D(x, y, z = 0, \omega; x_m, y_m)$ and the areal receiver data $Q_U(x, y, z = 0, \omega; x_m, y_m)$ for a single reflecor and a single SODCIG located at (x_m, y_m).

Since the size of the migrated image volume can be very big in practice and there are usually many reflectors in the subsurface, modeling each reflector and each SODCIG once a time may generate a data set even bigger than the original data set. One strategy to reduce the cost is to model several reflectors and several SODCIGs simultaneously (Biondi, 2006), however, this process generates unwanted crosstalk. As discussed by Guerra and Biondi (2008b,a), random phase encoding could be used to attenuate the crosstalk. The randomly encoded areal source and areal receiver wavefields can be computed as follows:

$$\left\{ \begin{array}{l} \left(\frac{\partial}{\partial z} - i\sqrt{\omega^2 \hat{s}^2(\mathbf{x}) - |\mathbf{k}|^2}\right) Q_D(\mathbf{x}, \mathbf{p}_m, \omega) = \widetilde{I}_D(\mathbf{x}, \mathbf{h}) \\ Q_D(x, y, z = z_{\max}, \mathbf{p}_m, \omega) = 0 \end{array} \right. , \tag{30}$$

and

$$\left\{ \begin{array}{l} \left(\frac{\partial}{\partial z} - i\sqrt{\omega^2 \hat{s}^2(\mathbf{x}) - |\mathbf{k}|^2}\right) Q_U(\mathbf{x}, \mathbf{p}_m, \omega) = \widetilde{I}_U(\mathbf{x}, \mathbf{h}) \\ Q_U(x, y, z = z_{\max}, \mathbf{p}_m, \omega) = 0 \end{array} \right. , \tag{31}$$

where $\widetilde{I}_D(\mathbf{x}, \mathbf{h})$ and $\widetilde{I}_U(\mathbf{x}, \mathbf{h})$ are the encoded SODCIGs after rotations, they are defined as follows:

$$\widetilde{I}_D(\mathbf{x}, \mathbf{h}) = \sum_{x_m} \sum_{y_m} I_D(\mathbf{x}, \mathbf{h}, x_m, y_m) \beta(\mathbf{x}, x_m, y_m, \mathbf{p}_m, \omega), \tag{32}$$

$$\widetilde{I}_U(\mathbf{x}, \mathbf{h}) = \sum_{x_m} \sum_{y_m} I_U(\mathbf{x}, \mathbf{h}, x_m, y_m) \beta(\mathbf{x}, x_m, y_m, \mathbf{p}_m, \omega), \tag{33}$$

where $\beta(\mathbf{x}, x_m, y_m, \mathbf{p}_m, \omega) = e^{i\gamma(\mathbf{x}, x_m, y_m, \mathbf{p}_m, \omega)}$ is chosen to be the random phase-encoding function with $\gamma(\mathbf{x}, x_m, y_m, \mathbf{p}_m, \omega)$ being a uniformly distributed random sequence in $\mathbf{x}$, x_m, y_m and ω; the variable $\mathbf{p}_m$ is the index of different realizations of the random sequence. Recursively solving Equations 30 and 31 give us the encoded areal source data $Q_D(x, y, z = 0, \mathbf{p}_m, \omega)$ and areal receiver data $Q_D(x, y, z = 0, \mathbf{p}_m, \omega)$ collected on the surface.

The synthesized new data sets are downward continued using the same one-way wave equation defined by Equations 13 and 14 (with different boundry conditions) as follows:

$$\left\{ \begin{array}{l} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) \widetilde{D}(\mathbf{x}, \mathbf{p}_m, \omega) = 0 \\ \widetilde{D}(x, y, z = 0, \mathbf{p}_m, \omega) = \overline{Q_D(x, y, z = 0, \mathbf{p}_m, \omega)} \end{array} \right. , \tag{34}$$

and

$$\left\{ \begin{array}{l} \left(\frac{\partial}{\partial z} + i\sqrt{\omega^2 s^2(\mathbf{x}) - |\mathbf{k}|^2}\right) \widetilde{U}(\mathbf{x}, \mathbf{p}_m, \omega) = 0 \\ \widetilde{U}(x, y, z = 0, \mathbf{x}_s, \omega) = Q_U(x, y, z = 0, \mathbf{p}_m, \omega) \end{array} \right. , \tag{35}$$

where $\widetilde{D}(\mathbf{x},\mathbf{p}_m,\omega)$ and $\widetilde{U}(\mathbf{x},\mathbf{p}_m,\omega)$ are the downward continuted areal source and areal receiver wavefields for realization $\mathbf{p}_m$. The image is produced by cross-correlating the two wavefields and summing all $\mathbf{p}_m$'s as follows:

$$I_{\rm me}(\mathbf{x},\mathbf{h}) = \sum_{\mathbf{p}_m}\sum_{\omega} \widetilde{D}(\mathbf{x},\mathbf{p}_m,\omega)\widetilde{U}(\mathbf{x},\mathbf{p}_m,\omega). \tag{36}$$

The crosstalk artifacts can be further attenuated if the number of $\mathbf{p}_m$ is large, therefore, approximately, the image obtained by migrating the image-space encoded gathers is kinematically equivalent to the image obtained in the shot-profile domain.

From Equation 36, the perturbed image can thus be easily obtained as follows:

$$\begin{aligned}\Delta I_{\rm me}(\mathbf{x},\mathbf{h}) &= \sum_{\mathbf{p}_m}\sum_{\omega}\Big(\Delta\widetilde{D}(\mathbf{x}-\mathbf{h},\mathbf{p}_m,\omega)\widehat{\widetilde{U}}(\mathbf{x}+\mathbf{h},\mathbf{p}_m,\omega)+ \\ &\quad \widehat{\widetilde{D}}(\mathbf{x}-\mathbf{h},\mathbf{p}_m,\omega)\Delta\widetilde{U}(\mathbf{x}+\mathbf{h},\mathbf{p}_m,\omega)\Big),\end{aligned} \tag{37}$$

where $\widehat{\widetilde{D}}(\mathbf{x},\mathbf{p}_m,\omega)$ and $\widehat{\widetilde{U}}(\mathbf{x},\mathbf{p}_m,\omega)$ are the image-space encoded background source and receiver wavefields; $\Delta\widetilde{D}(\mathbf{x},\mathbf{p}_m,\omega)$ and $\Delta\widetilde{U}(\mathbf{x},\mathbf{p}_m,\omega)$ are the perturbed source and receover wavefields in the image-space phase-encoding domain, which satisfy the perturbed one-way wave equations defined by Equations 17 and 18. The tomographic operator $\mathbf{T}$ and its adjoint $\mathbf{T}'$ can be implemented in a similar mannar as discussed in Appendix B and C by replacing the original wavefileds with the image-space phase encoded wavefields.

NUMERICAL EXAMPLES

We test the image-space wave-equation tomography in the generalized source domain on a simple model which contains only one reflector located at $z = 1500$ m. Figure 1 shows the correct slowness model. The background slowness is constant which is equal to 1/2000 s/m; there are two Gaussian anormalies located at $(x = -800, z = 800)$ and $(x = 800, z = 800)$ respectively: the left one is with 5% higher slowness, while the right one is with 5% lower slowness. We model 401 shots, the shots are ranging from -4000 m to 4000 m with a shot interval equals to 20 m. The receiver locations are also ranging from -4000 m to 4000 m, but with a 10 m inverval. The receivers are fixed for all shots to mimic a land acquisition geometry.

Figure 2 shows the migrated images in different domains computed with a background slowness $\widehat{s} = 1/2000$ s/m: Figure 2(a) is obtained by migrating the original 401 shot gathers. Because of the inaccuracy of the slowness model, we can identify the mispositioning of the reflectors, especially beneath the Gaussian anormalies. Figure 2(b) is obtained by migrating the data-space plane-wave-source encoded gathers, where 61 plane waves are migrated; The result is almost identical to that in Figure 2(a); Figure 2(c) is obtained by migrating the image-space encoded gathers. The

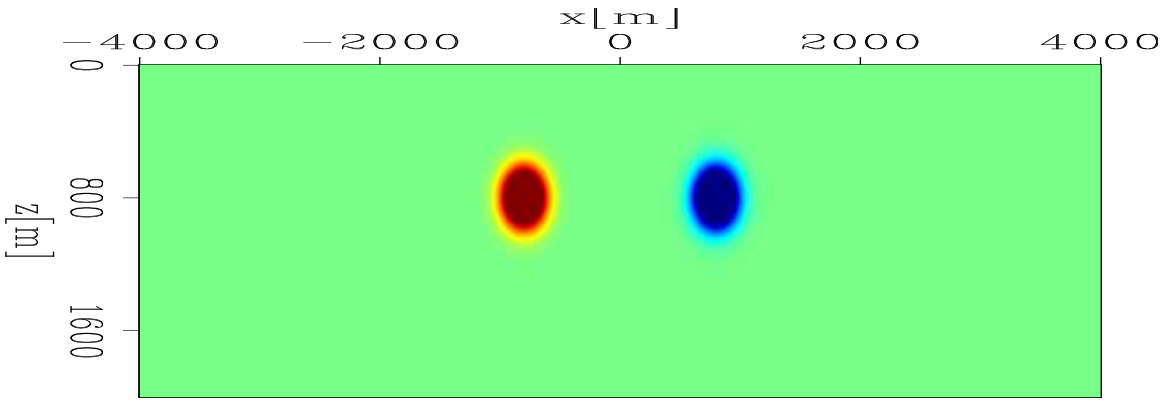

Figure 1: The correct slowness model. The slowness model consists of a constant background slowness (1/2000 s/m) and two 5% Gaussian anormalies.

image-space encoded areal source and receiver data are generated by simultaneously modeling 100 randomlly encoded unfocused SODCIGs and 2 realizations are used for the random sequence, therefore we end up with 20 image-space encoded areal gathers (there are 1001 points in x). The kinematics of the result looks almost the same as that in Figure 2(a). However, also notice the wavelet squeezing effect, and the random noise in the background caused by the random phase encoding.

Figure 3 shows the image perturbations obtained by applying the forward tomographic operator $\mathbf{T}$ in different domains. For this example, we assume that we know the correct slowness perturbation $\Delta\mathbf{s}$, which is obtained by subtracting the background slowness $\widehat{\mathbf{s}}$ from the correct slowness $\mathbf{s}$. Figure 3(a) shows the image perturbation computed with the original 401 shot gathers, notice the relative 90 degree phase rotation compared to the background image shown in Figure 2(a); Figure 3(b) is the result obtained by using 61 data-space plane-wave encoded gathers, it is almost identical to Figure 3(a); Figure 3(c) shows the result computed with 20 image-space encoded gathers, the kinematics are also similar to those in Figure 3(a).

Figure 4 illustrates the predicted slowness perturbations by applying the adjoint tomographic operator $\mathbf{T}'$ to the image perturbations obtained in Figure 3. For comparison, Figure 4(a) shows the correct slowness perturbation, i.e., $\Delta\mathbf{s} = \mathbf{s} - \widehat{\mathbf{s}}$; Figure 4(b) is the predicted slowness perturbation obtained in the original shot-profile domain using all 401 shot gathers; Figure 4(c) is the result obtained using all 61 data-space plane-wave encoded gathers, it is almost identical to Figure 4(b); Figure 4(d) shows the result obtained using all 20 image-space encoded gathers, which is also similar to Figure 4(b). However, note that Figure 4(d) shows slightly less focused result than Figure 4(b), the reason for this is effect of the unwanted crosstalk

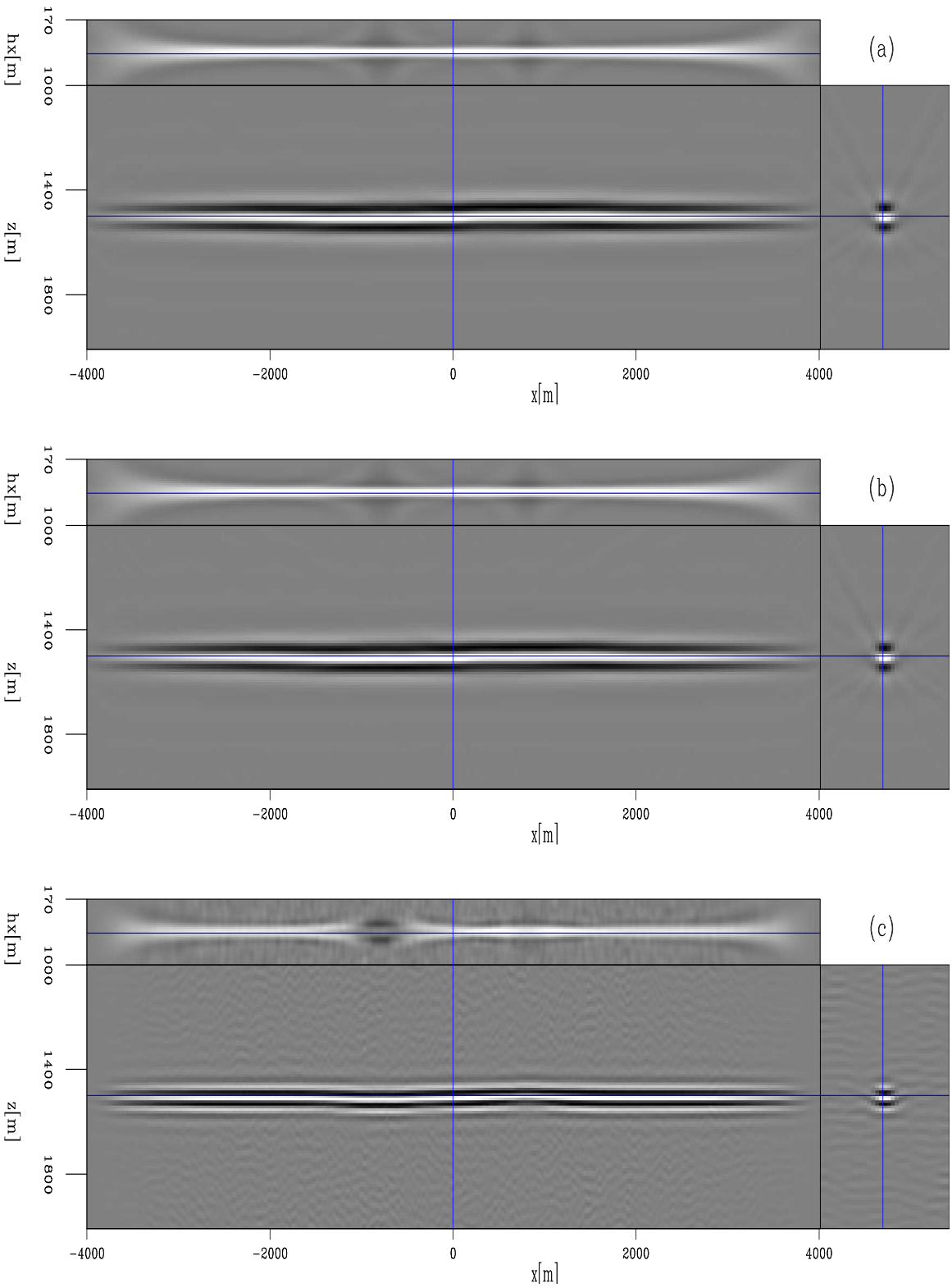

Figure 2: Migrated image cubes with a constant background slowness ($\widehat{s} = 2000$ s/m). Panel (a) is the result obtained in the original shot-profile domain; Panel (b) is the result obtained by migrating 61 plane waves, while panel (c) is obtained by migrating 40 image-space encoded areal gathers.

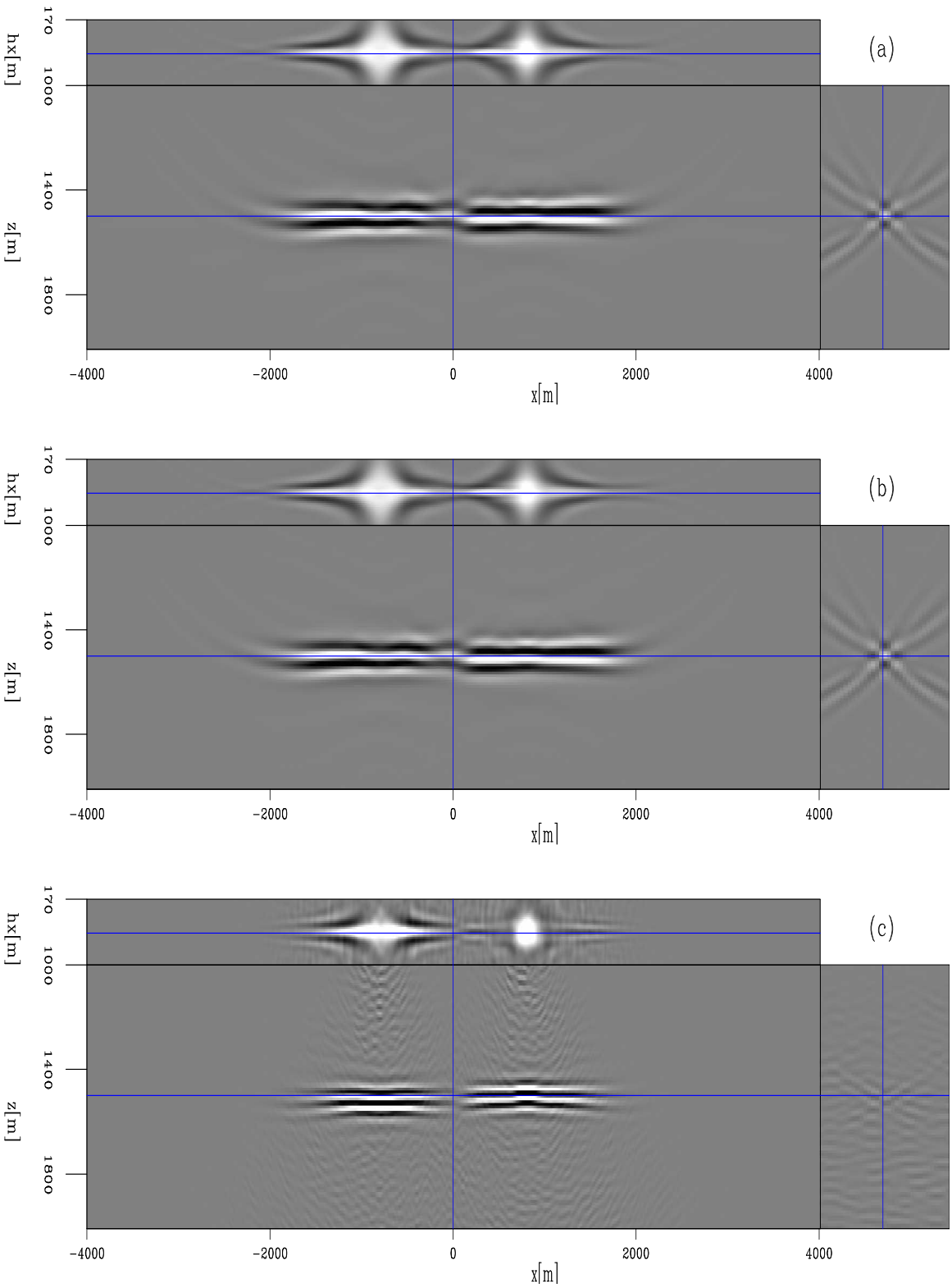

Figure 3: The image perturbations obtained by applying of the forward tomographic operator **T** on the correct slowness perturbations in different domains. Panel (a) shows the image perturbation obtained using the original shot gathers, while panels (b) and (c) are obtained using the data-space encoded gathers and image-space encoded gathers, respectively.

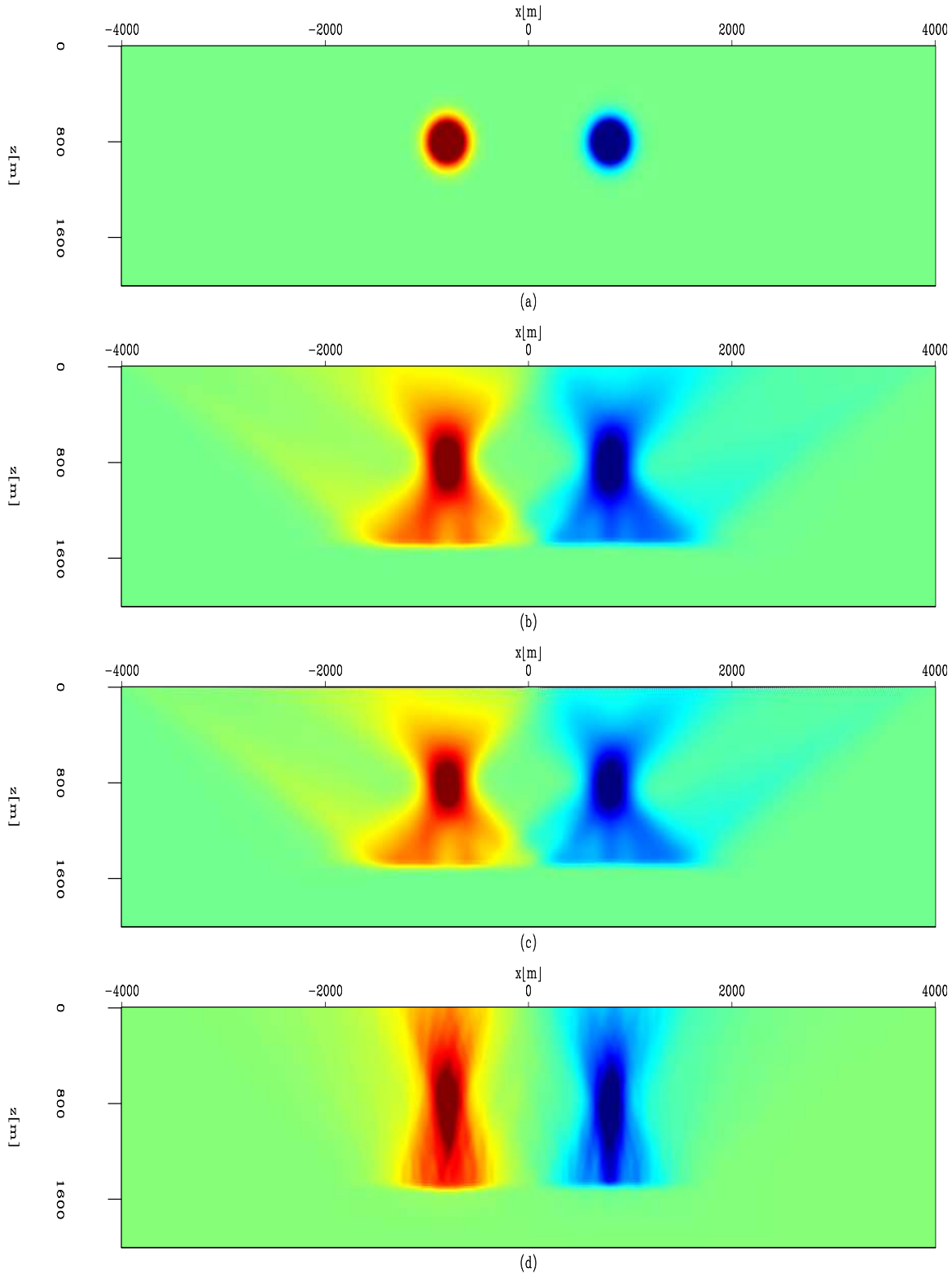

Figure 4: The slowness perturbation obtained by applying of the adjoint tomographic operator $\mathbf{T}'$ on the image perturbations in Figure 3. Panel (a) shows the correct slowness perturbation which is obtained by taking the difference between the correct slowess s and the background slowness $\widehat{s}$; Panel (b) shows the estimated slowness perturbation by back-projecting the image perturbation shown in Figure 3

CONCLUSIONS

REFERENCES

Albertin, U., P. Sava, J. Etgen, and M. Maharramov, 2006, Adjoint wave-equation velocity analysis: 76th Ann. Internat. Mtg., Expanded Abstracts, 3345–3349, Soc. of Expl. Geophys.

Biondi, B., 2006, Prestack exploding-reflectors modeling for migration velocity analysis: 76th Ann. Internat. Mtg., Expanded Abstracts, 3056–3060, Soc. of Expl. Geophys.

———, 2007, Prestack modeling of image events for migration velocity analysis: **SEP-131**, 101–118.

———, 2008, Automatic wave-equation migration velocity analysis: **SEP-134**, 65–78.

Biondi, B. and P. Sava, 1999, Wave-equation migration velocity analysis: 69th Ann. Internat. Mtg., Expanded Abstracts, 1723–1726, Soc. of Expl. Geophys.

Claerbout, J. F., 1971, Towards a unified theory of reflector mapping: Geophysics, **36**, 467–481.

Duquet, B. and P. Lailly, 2006, Efficient 3D wave-equation migration using virtual planar sources: Geophysics, **71**, S185–S197.

Etgen, J. T., 2005, How many angles do we really need for delayed-shot migration: 75th Ann. Internat. Mtg., Expanded Abstracts, 1985–1988, Soc. of Expl. Geophys.

Guerra, C. and B. Biondi, 2008a, New phase-encoding schemes for wave-equation migration: **SEP-136**.

———, 2008b, Prestack exploding reflector modeling: The crosstalk problem: **SEP-134**, 79–92.

Liu, F., D. W. Hanson, N. D. Whitmore, R. S. Day, and R. H. Stolt, 2006, Toward a unified analysis for source plane-wave migration: Geophysics, **71**, S129–S139.

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

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.

Romero, L. A., D. C. Ghiglia, C. C. Ober, and S. A. Morton, 2000, Phase encoding of shot records in prestack migration: Geophysics, **65**, 426–436.

Sava, P., 2004, Migration and Velocity Analysis by Wavefield Extrapolation: PhD thesis, Stanford University.

Shen, P., 2004, Wave-equation Migration Velocity Analysis by Differential Semblance Optimization: PhD thesis, Rice University.

Shen, P., W. Symes, S. Morton, and H. Calandra, 2005, Differential semblance velocity analysis via shot-profile migration: 75th Ann. Internat. Mtg., Expanded Abstracts, 2249–2252, Soc. of Expl. Geophys.

Shragge, J., 2007, Waveform inversion by one-way wavefield extrapolation: Geophysics, **72**, A47–A50.

Tang, Y., 2008, Modeling, migration and inversion in the generalized source and receiver domain: **SEP-136**.

Tarantola, A., 1987, Inverse problem theory: Methods for data fitting and model parameter estimation: Elsevier.

Whitmore, N. D., 1995, An Imaging Hierarchy for Common Angle Plane Wave Seismogram: PhD thesis, University of Tulsa.
Woodward, M. J., 1992, Wave-equation tomography: Geophysics, **57**, 15–26.
Zhang, Y., J. Sun, C. Notfors, S. Grey, L. Chemis, and J. Young, 2005, Delayed-shot 3D depth migration: Geophysics, **70**, E21–E28.

APPENDIX A

This appendix derives the perturbed one-way wave equation with respect to the slowness.

APPENDIX B

This appendix demonstrates a matrix representation of the forward tomographic operator $\mathbf{T}$

$$\mathbf{D}_{z+\Delta z} = \mathbf{E}_z(\mathbf{s}_z)\mathbf{\Delta D}_z, \tag{B-1}$$

where

$$\mathbf{E}_z(\mathbf{s}_z) = e^{-ik_z(\mathbf{s}_z)\Delta z} = e^{-i\sqrt{\omega^2\mathbf{s}_z^2 - |\mathbf{k}|^2}} \tag{B-2}$$

The perturbed source wavefield at some depth level can be derived from the background wavefield by a simple application of the chain rule to equation B-1:

$$\mathbf{\Delta D}_{z+\Delta z} = \mathbf{E}_z(\widehat{\mathbf{s}}_z)\mathbf{\Delta D}_z + \mathbf{\Delta E}_z(\widehat{\mathbf{s}}_z)\widehat{\mathbf{D}}_z, \tag{B-3}$$

where $\widehat{\mathbf{D}}_z$ is the background source wavefield and $\mathbf{\Delta E}_z$ represents the perturbed extrapolator, it can be obtained by a formal linearization with respect to slowness of the extrapolator defined in equation B-2

$$\begin{aligned} \mathbf{E}_z(\mathbf{s}_z) = e^{-ik_z(\mathbf{s}_z)\Delta z} &\approx e^{-i\Delta z\widehat{k}_z} + e^{-i\Delta z\widehat{k}_z}\left(-i\Delta z\,\frac{dk_z}{d\mathbf{s}_z}\bigg|_{\mathbf{s}_z=\widehat{\mathbf{s}}_z}\right)\mathbf{\Delta s}_z \\ &= \mathbf{E}_z(\widehat{\mathbf{s}}_z) + \mathbf{E}_z(\widehat{\mathbf{s}}_z)\left(-i\Delta z\,\frac{dk_z}{d\mathbf{s}_z}\bigg|_{\mathbf{s}_z=\widehat{\mathbf{s}}_z}\right)\mathbf{\Delta s}_z \end{aligned} \tag{B-4}$$

where $\widehat{k}_z = k_z(\widehat{\mathbf{s}}_z)$. So the perturbed extrapolator reads

$$\mathbf{\Delta E}_z(\widehat{\mathbf{s}}_z) = \mathbf{E}_z(\widehat{\mathbf{s}}_z)\left(-i\Delta z\,\frac{dk_z}{d\mathbf{s}_z}\bigg|_{\mathbf{s}_z=\widehat{\mathbf{s}}_z}\right)\mathbf{\Delta s}_z \tag{B-5}$$

Substitute equation B-5 into B-3, we get

$$\mathbf{\Delta D}_{z+\Delta z} = \mathbf{E}_z(\widehat{\mathbf{s}}_z)\mathbf{\Delta D}_z + \mathbf{E}_z(\widehat{\mathbf{s}}_z)\left(-i\Delta z \left.\frac{dk_z}{d\mathbf{s}_z}\right|_{\mathbf{s}_z=\widehat{\mathbf{s}}_z}\right)\widehat{\mathbf{D}}_z\mathbf{\Delta s}_z, \tag{B-6}$$

Let us define a scattering operator $\mathbf{G}_z$ that interacts with the background wavefield as follows:

$$\mathbf{G}_z(\widehat{\mathbf{D}}_z,\widehat{\mathbf{s}}_z) = \left(-i\Delta z \left.\frac{dk_z}{d\mathbf{s}_z}\right|_{\mathbf{s}_z=\widehat{\mathbf{s}}_z}\right)\widehat{\mathbf{D}}_z = \frac{-i\omega\Delta z}{\sqrt{1-\frac{|\mathbf{k}|^2}{\omega^2\widehat{\mathbf{s}}_z^2}}}\widehat{\mathbf{D}}_z \tag{B-7}$$

Then the perturbed source wavefield for depth level $z+\Delta z$ can be rewritten as follows:

$$\mathbf{\Delta D}_{z+\Delta z} = \mathbf{E}_z(\widehat{\mathbf{s}}_z)\mathbf{\Delta D}_z + \mathbf{E}_z(\widehat{\mathbf{s}}_z)\mathbf{G}_z(\widehat{\mathbf{D}}_z,\widehat{\mathbf{s}}_z)\mathbf{\Delta s}_z. \tag{B-8}$$

We can further write out the recursive equation B-8 for all depth levels in the following matrix form:

$$\begin{pmatrix} \mathbf{\Delta D}_0 \\ \mathbf{\Delta D}_1 \\ \mathbf{\Delta D}_2 \\ \vdots \\ \mathbf{\Delta D}_n \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{E}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{E}_{n-1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{\Delta D}_0 \\ \mathbf{\Delta D}_1 \\ \mathbf{\Delta D}_2 \\ \vdots \\ \mathbf{\Delta D}_n \end{pmatrix} + \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{E}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{E}_{n-1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{G}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{G}_n \end{pmatrix} \begin{pmatrix} \mathbf{\Delta s}_0 \\ \mathbf{\Delta s}_1 \\ \mathbf{\Delta s}_2 \\ \vdots \\ \mathbf{\Delta s}_n \end{pmatrix}$$

Or in a more compact notation

$$\mathbf{\Delta D} = \mathbf{E}(\widehat{\mathbf{s}})\mathbf{\Delta D} + \mathbf{E}(\widehat{\mathbf{s}})\mathbf{G}(\widehat{\mathbf{D}},\widehat{\mathbf{s}})\mathbf{\Delta s}. \tag{B-9}$$

The solution of equation B-9 can be formally written as follows:

$$\mathbf{\Delta D} = (\mathbf{1}-\mathbf{E}(\widehat{\mathbf{s}}))^{-1}\,\mathbf{E}(\widehat{\mathbf{s}})\mathbf{G}(\widehat{\mathbf{D}},\widehat{\mathbf{s}})\mathbf{\Delta s}. \tag{B-10}$$

Similarly, the perturbed receiver wavefield satisfy the following recursive relation:

$$\mathbf{\Delta U}_{z+\Delta z} = \mathbf{E}_z(\widehat{\mathbf{s}}_z)\mathbf{\Delta U}_z - \mathbf{E}_z(\widehat{\mathbf{s}}_z)\mathbf{G}_z(\widehat{\mathbf{U}}_z,\widehat{\mathbf{s}}_z)\mathbf{\Delta s}_z, \tag{B-11}$$

where $\mathbf{G}_z(\widehat{\mathbf{U}}_z,\widehat{\mathbf{s}}_z)$ is the scattering operator interacts with the background receiver wavefield as follows:

$$\mathbf{G}_z(\widehat{\mathbf{U}}_z,\widehat{\mathbf{s}}_z) = \left(-i\Delta z \left.\frac{dk_z}{d\mathbf{s}_z}\right|_{\mathbf{s}_z=\widehat{\mathbf{s}}_z}\right)\widehat{\mathbf{U}}_z = \frac{-i\omega\Delta z}{\sqrt{1-\frac{|\mathbf{k}|^2}{\omega^2\widehat{\mathbf{s}}_z^2}}}\widehat{\mathbf{U}}_z. \tag{B-12}$$

We can further write out the recursive equation B-12 for all depth levels in the following matrix form:

$$\begin{pmatrix} \mathbf{\Delta U}_0 \\ \mathbf{\Delta U}_1 \\ \mathbf{\Delta U}_2 \\ \vdots \\ \mathbf{\Delta U}_n \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{E}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{E}_{n-1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{\Delta U}_0 \\ \mathbf{\Delta U}_1 \\ \mathbf{\Delta U}_2 \\ \vdots \\ \mathbf{\Delta U}_n \end{pmatrix} - \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{E}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{E}_{n-1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{G}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{G}_n \end{pmatrix} \begin{pmatrix} \mathbf{\Delta s}_0 \\ \mathbf{\Delta s}_1 \\ \mathbf{\Delta s}_2 \\ \vdots \\ \mathbf{\Delta s}_n \end{pmatrix}$$

Or in a more compact notation

$$\mathbf{\Delta U} = \mathbf{E}(\widehat{\mathbf{s}})\mathbf{\Delta U} + \mathbf{E}(\widehat{\mathbf{s}})\mathbf{G}(\widehat{\mathbf{U}}, \widehat{\mathbf{s}})\mathbf{\Delta s}. \tag{B-13}$$

The solution of equation B-13 can be formally written as follows:

$$\mathbf{\Delta U} = (\mathbf{1} - \mathbf{E}(\widehat{\mathbf{s}}))^{-1}\, \mathbf{E}(\widehat{\mathbf{s}})\mathbf{G}(\widehat{\mathbf{U}}, \widehat{\mathbf{s}})\mathbf{\Delta s}. \tag{B-14}$$

Therefore, the perturbed image can be obtained as follows:

$$\begin{pmatrix} \mathbf{\Delta I}_0 \\ \mathbf{\Delta I}_1 \\ \mathbf{\Delta I}_2 \\ \vdots \\ \mathbf{\Delta I}_n \end{pmatrix} = \begin{pmatrix} \widehat{\mathbf{U}}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{U}}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \widehat{\mathbf{U}}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \widehat{\mathbf{U}}_n \end{pmatrix} \begin{pmatrix} \mathbf{\Delta D}_0 \\ \mathbf{\Delta D}_1 \\ \mathbf{\Delta D}_2 \\ \vdots \\ \mathbf{\Delta D}_n \end{pmatrix} + \begin{pmatrix} \widehat{\mathbf{D}}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{D}}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \widehat{\mathbf{D}}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \widehat{\mathbf{D}}_n \end{pmatrix} \begin{pmatrix} \mathbf{\Delta U}_0 \\ \mathbf{\Delta U}_1 \\ \mathbf{\Delta U}_2 \\ \vdots \\ \mathbf{\Delta U}_n \end{pmatrix}$$

Or in a more compact notation

$$\mathbf{\Delta I} = \mathrm{diag}\left(\widehat{\mathbf{U}}\right)\mathbf{\Delta D} + \mathrm{diag}\left(\widehat{\mathbf{D}}\right)\mathbf{\Delta U}. \tag{B-15}$$

After substituting equations B-10 and B-14 into equation B-15

$$\begin{aligned} \mathbf{\Delta I} \;=\; & \Big(\mathrm{diag}\left(\widehat{\mathbf{U}}\right)(\mathbf{1} - \mathbf{E}(\widehat{\mathbf{s}}))^{-1}\,\mathbf{E}(\widehat{\mathbf{s}})\mathbf{G}(\widehat{\mathbf{D}}, \widehat{\mathbf{s}}) + \\ & \mathrm{diag}\left(\widehat{\mathbf{D}}\right)(\mathbf{1} - \mathbf{E}(\widehat{\mathbf{s}}))^{-1}\,\mathbf{E}(\widehat{\mathbf{s}})\mathbf{G}(\widehat{\mathbf{U}}, \widehat{\mathbf{s}})\Big)\mathbf{\Delta s}, \end{aligned} \tag{B-16}$$

we finally get the following matrix representation of the forward tomographic operator $\mathbf{T}$:

$$\begin{aligned} \mathbf{T} \;=\; & \mathrm{diag}\left(\widehat{\mathbf{U}}\right)(\mathbf{1} - \mathbf{E}(\widehat{\mathbf{s}}))^{-1}\,\mathbf{E}(\widehat{\mathbf{s}})\mathbf{G}(\widehat{\mathbf{D}}, \widehat{\mathbf{s}}) + \\ & \mathrm{diag}\left(\widehat{\mathbf{D}}\right)(\mathbf{1} - \mathbf{E}(\widehat{\mathbf{s}}))^{-1}\,\mathbf{E}(\widehat{\mathbf{s}})\mathbf{G}(\widehat{\mathbf{U}}, \widehat{\mathbf{s}}) \end{aligned} \tag{B-17}$$

APPENDIX C

This appendix demonstrates a matrix representation of the adjoint tomographic operator $\mathbf{T}'$. Since the slowness perturbation $\mathbf{\Delta s}$ is linearly related to the perturbed wavefields, $\mathbf{\Delta D}$ and $\mathbf{\Delta U}$, to obtain the back-projected slowness perturbation, we first need to get the back-projected perturbed wavefields from the perturbed image $\mathbf{\Delta I}$. From equation B-15, the back-projected perturbed source and receiver wavefields are obtained as follows:

$$\mathbf{\Delta D} = \overline{\mathrm{diag}\left(\widehat{\mathbf{U}}\right)}\mathbf{\Delta I} \tag{C-1}$$

and

$$\mathbf{\Delta U} = \overline{\mathrm{diag}\left(\widehat{\mathbf{D}}\right)}\mathbf{\Delta I} \tag{C-2}$$

Then the adjoint equations of equations B-10 and B-14 are used to get the back-projected slowness perturbation $\mathbf{\Delta s}$. Let us first look at the adjoint equation of equation B-10

$$\mathbf{\Delta s}_D = \mathbf{G}'(\widehat{\mathbf{D}}, \widehat{\mathbf{s}})\mathbf{E}'(\widehat{\mathbf{s}})\left(\mathbf{1} - \mathbf{E}'(\widehat{\mathbf{s}})\right)^{-1}\mathbf{\Delta D} \tag{C-3}$$

We define a temporary wavefield $\mathbf{\Delta P}_D$ that satisfies the following equation:

$$\mathbf{\Delta P}_D = \mathbf{E}'(\widehat{\mathbf{s}})\left(\mathbf{1} - \mathbf{E}'(\widehat{\mathbf{s}})\right)^{-1}\mathbf{\Delta D} \tag{C-4}$$

After some simple algebra, the above equation can be rewritten as follows

$$\mathbf{\Delta P}_D = \mathbf{E}'(\widehat{\mathbf{s}})\mathbf{\Delta P}_D + \mathbf{E}'(\widehat{\mathbf{s}})\mathbf{\Delta D} \tag{C-5}$$

Substitute equation C-1 into equation C-5, we get:

$$\mathbf{\Delta P}_D = \mathbf{E}'(\widehat{\mathbf{s}})\mathbf{\Delta P}_D + \mathbf{E}'(\widehat{\mathbf{s}})\overline{\text{diag}\left(\widehat{\mathbf{U}}\right)}\mathbf{\Delta I} \tag{C-6}$$

Therefore, $\mathbf{\Delta P}_D$ can be obtained by recursively upward continuation, where $\mathbf{\Delta D} = \overline{\text{diag}\left(\widehat{\mathbf{U}}\right)}\mathbf{\Delta I}$ serves as the boundary condition. The back-projected slowness perturbation from the perturbed source wavefield is then obtained by applying the ajoint of the scattering operator $\mathbf{G}(\widehat{\mathbf{D}}, \widehat{\mathbf{s}})$ to the wavefield $\mathbf{\Delta P}_D$

$$\mathbf{\Delta s}_D = \mathbf{G}'(\widehat{\mathbf{D}}, \widehat{\mathbf{s}})\mathbf{\Delta P}_D \tag{C-7}$$

Similarly, the adjoint equation of equation B-14 reads

$$\mathbf{\Delta s}_U = \mathbf{G}'(\widehat{\mathbf{U}}, \widehat{\mathbf{s}})\mathbf{E}(\widehat{\mathbf{s}})'\left(\mathbf{1} - \mathbf{E}(\widehat{\mathbf{s}})'\right)^{-1}\mathbf{\Delta U} \tag{C-8}$$

We can also define a temporary wavefield $\mathbf{\Delta P}_U$ that satisfies the following equation

$$\mathbf{\Delta P}_U = \mathbf{E}(\widehat{\mathbf{s}})'\left(\mathbf{1} - \mathbf{E}(\widehat{\mathbf{s}})'\right)^{-1}\mathbf{\Delta U} \tag{C-9}$$

After rewriting it, we get the following recursive form

$$\begin{aligned}\mathbf{\Delta P}_U &= \mathbf{E}(\widehat{\mathbf{s}})'\mathbf{\Delta P}_U + \mathbf{E}(\widehat{\mathbf{s}})'\mathbf{\Delta U} \\ &= \mathbf{E}(\widehat{\mathbf{s}})'\mathbf{\Delta P}_U + \mathbf{E}(\widehat{\mathbf{s}})'\overline{\text{diag}\left(\widehat{\mathbf{D}}\right)}\mathbf{\Delta I}.\end{aligned} \tag{C-10}$$

And the back-projected slowness perturbation from the perturbed receiver wavefield is then obtained by applying the adjoint of the scattering operator $\mathbf{G}(\widehat{\mathbf{U}}, \widehat{\mathbf{s}})$ to the wavefield $\mathbf{\Delta P}_U$

$$\mathbf{\Delta s}_U = \mathbf{G}'(\widehat{\mathbf{U}}, \widehat{\mathbf{s}})\mathbf{\Delta P}_U \tag{C-11}$$

The total back-projected slowness perturbation is obtained by adding $\mathbf{\Delta s}_D$ and $\mathbf{\Delta s}_U$ together

$$\mathbf{\Delta s} = \mathbf{\Delta s}_D + \mathbf{\Delta s}_U. \tag{C-12}$$