Do we really need low frequencies in waveform inversion?

Paul Sava, Esteban Díaz and Tongning Yang, Center for Wave Phenomena, Colorado School of Mines*

Wavefield tomography (WT) constructs velocity models using seismic wavefields (Tarantola, 1984; Pratt, 1999; Sirgue and Pratt, 2004; Plessix, 2006; Symes, 2009). Models are typically updated by matching simulated and recorded data. This match imposes the strong assumption that the chosen wave-equation is consistent with the physics of wave propagation in the earth. Thus, a significant part of WT is dedicated to removing the components of the observed data that are inconsistent with the assumptions made about the wave-equation.

A key component of WT is the objective function (OF) measuring the match between simulated and recorded data. We assume that the OF is convex, thus enabling its minimization using gradient-based techniques. The gradient calculation uses the adjoint state method (Plessix, 2006; Symes, 2009). This method consists of 4 steps: (1) compute the *state variables*, i.e. seismic wavefields obtained from the source by forward modeling; (2) compute the *adjoint source* based on the OF and the state variables; (3) compute the *adjoint state variables*, i.e. seismic wavefields obtained from the adjoint source by backward modeling; (4) compute the *gradient* using the state and adjoint state variables.

The main sources of information in WT are the source wavelet $f_s(e,\mathbf{x},\omega)$ (hereby assumed to be known) and the observed data $f_r(e,\mathbf{x},\omega)$. The WT state variables, i.e. the source and receiver wavefields, $u_s(e,\mathbf{x},\omega)$ and $u_s(e,\mathbf{x},\omega)$, are obtained by solving the wave-equation:

$$\begin{bmatrix} \mathcal{L}(\mathbf{x},\omega,m) & 0 \\ 0 & \mathcal{L}^*(\mathbf{x},\omega,m) \end{bmatrix} \begin{bmatrix} u_s \\ u_r \end{bmatrix} = \begin{bmatrix} f_s \\ f_r \end{bmatrix} . \quad (1)$$

$\mathcal{L}$ denotes the wave operator, m are the model parameters (e.g slowness squared), e is the experiment index, ω is the frequency, and $\mathbf{x}$ are space coordinates $\{x,y,z\}$.

Data domain: We formulate data-domain WT (dWT) as an inverse problem based on an OF defined using the difference between the source and receiver wavefields (Tarantola, 1984; Pratt, 1999):

$$\mathcal{J}_D = \sum_e \frac{1}{2} \| K_D \left(u_s(e,\mathbf{x},\omega) - u_r(e,\mathbf{x},\omega) \right) \|^2_{\mathbf{x},\omega} . \quad (2)$$

Here, $K_D(e,\mathbf{x})$ is a mask operator restricting the wavefields to the receiver positions, i.e. to the observed and simulated data. This common OF suffers, among other things, from cycle skipping due to the oscillatory nature of the subtracted wavefields, Figure 2(b). This problem is usually addressed by bootstrapping the frequency from low to high, thus requiring that low frequency data ($\sim$ 1Hz) are acquired in the field. If this condition is satisfied, we can define the adjoint source, $g_s(e,\mathbf{x},\omega)$, based on the source and receiver wavefields

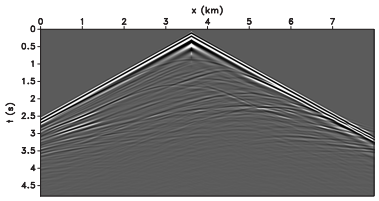
$$g = K_D \overline{K_D} \left(u_s - u_r \right) , \quad (3)$$

compute the adjoint state variable, $a_s(e,\mathbf{x},\omega)$, using backward modeling,

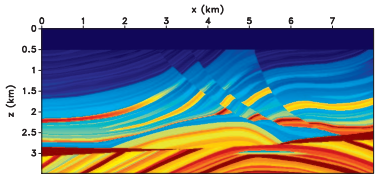
$$\mathcal{L}^*(\mathbf{x},\omega,m)\, a_s = g_s , \quad (4)$$

and evaluate the OF gradient by correlating the state (u_s) and adjoint state (a_s) variables (Plessix, 2006):

$$\nabla_m \mathcal{J}_D(\mathbf{x}) = \sum_{e,\omega} \omega^2 \left(\overline{u_s} a_s \right) . \quad (5)$$

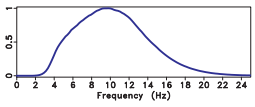


(a)

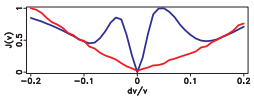


(b)

Figure 1: (a) A shot with short acquisition aperture (no diving waves), and (b) the Marmousi model.



(a)



(b)

Figure 2: (a) The data spectrum, and (b) the OF for dWT (blue) and iWT (red).

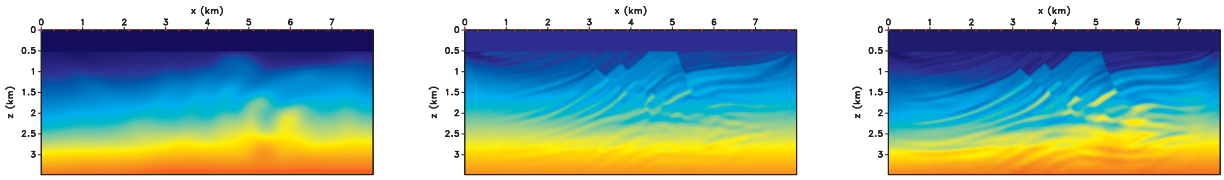
Image domain: We formulate image-domain WT (iWT) using extended wave-equation imaging (Sava and Vasconcelos, 2011). The OF is

$$\mathcal{J}_I = \frac{1}{2} \| K_I P(\boldsymbol{\lambda},\tau)\, r(\mathbf{x},\boldsymbol{\lambda},\tau) \|^2_{\mathbf{x},\boldsymbol{\lambda},\tau} , \quad (6)$$

where $r(\mathbf{x},\boldsymbol{\lambda},\tau)$ are extended images

$$r = \sum_{e,\omega} \overline{T(\boldsymbol{\lambda})}\ \overline{u_s(e,\mathbf{x},\omega)} T(\boldsymbol{\lambda})\ u_r(e,\mathbf{x},\omega)\, e^{2i\omega\tau} , \quad (7)$$

$T(\boldsymbol{\lambda})$ indicates space shift and $\boldsymbol{\lambda}$ and τ are cross-correlation lags. The mask $K_I(\mathbf{x})$ restricts the evaluation of the OF



(a) (b) (c)

Figure 3: (a) iWT model starting from a $v(z)$ model, (b) dWT model starting from the same $v(z)$ model and (c) dWT model starting from the iWT model.

to some image locations, and $P(\boldsymbol{\lambda}, \tau)$ is a penalty operator applied in the extended space (Symes, 2009). The extend image $r(\mathbf{x}, \boldsymbol{\lambda}, \tau)$ is just a proxy for the source and receiver wavefields. This OF does not suffer from the cycle-skipping problem, Figure 2(b). The source and receiver adjoint sources, $g_s(e, \mathbf{x}, \omega)$ and $g_r(e, \mathbf{x}, \omega)$, can be written as (Yang and Sava, 2011)

$$\begin{bmatrix} g_s \\ g_r \end{bmatrix} = \begin{bmatrix} K_I \overline{K_I} \sum\limits_{\tau, \boldsymbol{\lambda}} \overline{T} \left(P \overline{P}\, r \right) T \overline{u_r} e^{-2i\omega\tau} \\ K_I \overline{K_I} \sum\limits_{\tau, \boldsymbol{\lambda}} T \left(P \overline{P}\, r \right) \overline{T} u_s e^{-2i\omega\tau} \end{bmatrix} , \qquad (8)$$

and are used to simulate the adjoint state variables, $a_s(e, \mathbf{x}, \omega)$ (backward) and $a_r(e, \mathbf{x}, \omega)$ (forward),

$$\begin{bmatrix} \mathcal{L}^*(\mathbf{x}, \omega, m) & 0 \\ 0 & \mathcal{L}(\mathbf{x}, \omega, m) \end{bmatrix} \begin{bmatrix} a_s \\ a_r \end{bmatrix} = \begin{bmatrix} g_s \\ g_r \end{bmatrix} . \qquad (9)$$

The OF gradient is the correlation of the state (u_s, u_r) and adjoint state (a_s, a_r) variables: (Plessix, 2006):

$$\nabla_m \mathcal{J}_I(\mathbf{x}) = \sum_{e, \omega} \omega^2 \left(\overline{u_s} a_s + \overline{u_r} a_r \right) . \qquad (10)$$

Discussion: Both dWT and iWT use the same wavefields and wave operators, thus describing two forms of WT which differ essentially just in the definition of the OF. The image-domain OF is smooth and allows convergence from a poor starting model, even if low frequencies (i.e. $<\sim$ 1Hz) are not present in the data – we avoid cycle skipping. The data-domain OF is more abrupt near the correct model, which enables convergence to a good quality model – we achieve high resolution. Therefore, dWT and iWT complement each-other and can be cascaded to obtain model updates even when the starting model is wrong and when low frequency data are not recorded. We conclude that the requirement that very low frequency data are necessary for WT is simply an artifact of the definition of the data-domain OF. Using alternative OFs in the image-domain, we can construct models that are close enough to enable convergence of data-domain WT in the more conventional seismic band.

Example: We illustrate the method with data, Figure 1(a), simulated in the Marmousi model, Figure 1(b), using relatively short offsets and in a frequency band above 3Hz, Figure 2(a), thus eliminating offsets and frequencies that enable a successful start of conventional dWT. Our dWT is initialized either with a $v(z)$ model or with a model obtained by iWT, Figure 3(a). Figures 3(b) and 3(c) show the dWT results using the two starting models, respectively. Both starting models generate reasonable results in the upper part of the model, but only the iWT model is sufficiently close to allow a good reconstruction in the deeper part.

Conclusion: WT can be formulated in the data-domain or in the image-domain based on the same wavefields and wave-equation. iWT can construct good starting models for dWT, since its OF is not sensitive to cycle-skipping. Conversely, dWT can construct high-resolution models, since its OF is steep in the vicinity of the global minimum. Cascading iWT and dWT eliminates the need for extremely low-frequency in the acquired data.

Acknowledgments: We acknowledge the sponsors of the Center for Wave Phenomena at Colorado School of Mines. The reproducible numeric examples use the Madagascar open-source package (http://www.ahay.org).

REFERENCES

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

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

Sava, P., and I. Vasconcelos, 2011, Extended imaging condition for wave-equation migration: Geophysical Prospecting, **59**, 35–55.

Sirgue, L., and R. Pratt, 2004, Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies: Geophysics, **69**, 231–248.

Symes, W., 2009, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, 765–790.

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

Yang, T., and P. Sava, 2011, Waveform inversion in the image domain: Presented at the 73rd Mtg., Abstracts, Eur. Assoc. Expl. Geophys.