Modified Tomographic Full Waveform Inversion using the variable projection method

Guillaume Barnier, Ettore Biondi, and Biondo Biondi

ABSTRACT

We propose a modified tomographic full waveform inversion (TFWI) optimization scheme that allows us to avoid the original nested-loop approach and reduce the number of inversion parameters. We use the variable projection method to solve for the linear component of the inverse problem. We show the convergence to the correct velocity model on a synthetic dataset lacking frequencies below 10 Hz.

INTRODUCTION

Almomin (2016a) showed the potential of TFWI in simultaneously recovering all model scale components while being immune to cycle-skipping. However, this technique is hampered by three major challenges. Global convergence has not yet been mathematically proved even though it has been heuristically observed. Moreover, it heavily relies on the user to tune the many inversion parameters present in the original nested-scheme approach. Finally, it is highly computational intensive. Therefore, this current formulation is hardly applicable on production field datasets.

To overcome the first and second issues, we propose a new formulation in which the extended modeling term simply ensures phase matching between observed and predicted data. During our optimization process, we reduce the contribution of the additional extended term over iterations while ensuring the convergence of the total objective function. Since our cost function is quadratic with respect to the extended component, we use the variable projection method to minimize it (Rickett, 2013; Huang and Symes, 2015). We also demonstrate the potential of our new formulation on a 2D synthetic example in which the low frequency content is removed.

We first describe the mathematical difference between our proposed algorithm and the original nested-loop formulation in Biondi and Almomin (2014). We then apply our technique on a 2D synthetic model similar to the one shown in Mora (1989) and compare the results to the ones obtained using FWI.

TFWI THEORY

We review the original TFWI formulation proposed by Biondi and Almomin (2014), and analyze the potential reason for its slow convergence rate. Then we describe our

proposed method to potentially overcome this issue.

Original formulation

In the TFWI optimization algorithm devised by Biondi and Almomin (2014), the minimized objective function is effectively the standard FWI data misfit,

$$\Phi_{FWI}(\mathbf{m}) = \frac{1}{2}\|\mathbf{f}(\mathbf{m}) - \mathbf{d^{obs}}\|_2^2, \tag{1}$$

where $\mathbf{f}$ is the wave-equation operator, $\mathbf{m}$ is the velocity model, and $\mathbf{d^{obs}}$ represents the observed data. The optimization is performed in a nested scheme, using an extended Born modeling operator, an auxillary model composed of a background model $\mathbf{b}$, and an extended perturbation (i.e., reflectivity) $\tilde{\mathbf{p}}$. This nested scheme is designed to protect the algorithm from cycle-skipping encountered in standard FWI workflows, especially when low frequencies are missing in the data.

To understand why the original scheme may suffer from slow convergence rate, we propose to follow step by step one full iteration of the "outer" loop, say from i to $i+1$ (equation 1). That is, to compute the model update from $\mathbf{m}^i \mapsto \mathbf{m}^{i+1} = \mathbf{m}^i + \Delta\mathbf{m}^i$. In our notation, the upper script indices correspond to the outer loop iteration number, while the subscript indices correspond to the inner loop iteration number. In order to find the model update $\Delta\mathbf{m}^i$ needed at iteration i, the following objective function is minimized,

$$\begin{aligned}\Phi^i_{TFWI}(\mathbf{b}, \tilde{\mathbf{p}}) &= \frac{1}{2}\|\mathbf{f}(\mathbf{m}^i) + \tilde{\mathbf{B}}(\mathbf{b})\tilde{\mathbf{p}} - \mathbf{d^{obs}}\|_2^2 + \frac{\epsilon^2}{2}\|\mathbf{g}(\tilde{\mathbf{p}})\|_2^2 \\ &= \frac{1}{2}\|\tilde{\mathbf{B}}(\mathbf{b})\tilde{\mathbf{p}} - \Delta\mathbf{d}_i\|_2^2 + \frac{\epsilon^2}{2}\|\mathbf{g}(\tilde{\mathbf{p}})\|_2^2, \end{aligned} \tag{2}$$

where $\tilde{\mathbf{B}}$ is the extended Born modeling operator, and $\mathbf{g}$ is an operator enhancing the non-physical energy of the the extended reflectivity. This last operator can be linear or non-linear with respect to the extended reflectivity. $\Delta\mathbf{d}_i = \mathbf{d^{obs}} - \mathbf{f}(\mathbf{m}^i)$ is kept constant during the minimization of Φ^i. This assumption might be the cause of the slow convergence rate of the original scheme. Note that Φ^i is non-quadratic with respect to the background $\mathbf{b}$, but quadratic with respect to the extended perturbation $\tilde{\mathbf{p}}$. Minimizing equation 2 is thus a "nonlinear" optimization problem. The inner inversion starts by setting $\mathbf{b}_0 = \mathbf{m}^i$ and $\tilde{\mathbf{p}}_0 = \mathbf{0}$. Minimizing equation 2 is performed using a gradient-based descent method. At each inner step, a scale mixing algorithm (wavenumber filtering) is applied to both gradients $\nabla_{\mathbf{b}}\Phi^i$ and $\nabla_{\mathbf{p}}\Phi^i$ (the non-extended component of $\nabla_{\tilde{\mathbf{p}}}\Phi^i$) to ensure that $\mathbf{b}$ is only updated with low-wavenumber components, and that $\tilde{\mathbf{p}}$ is only updated with higher-wavenumber components. The wavenumber cut-off is based on the dominant frequency in the data as

well as the average velocity of the initial model (Almomin, 2016). Let $\mathbf{b}^i_{opt}$ and $\tilde{\mathbf{p}}^i_{opt}$ be the solutions found at the last inner iteration of outer loop i (i.e., after minimizing equation 2). A final low-wavenumber bandpass filter is applied to both $\mathbf{b}^i_{opt}$ and $\mathbf{p}^i_{opt}$, where $\mathbf{p}^i_{opt}$ is the non-extended component of $\tilde{\mathbf{p}}^i_{opt}$. The new model $\mathbf{m}^{i+1}$ is given by

$$\begin{aligned} \mathbf{m}^{i+1} &= \mathbf{m}^i + \Delta\mathbf{m}^i \\ &= \mathbf{m}^i + \mathbf{F}\left(\mathbf{b}^i_{opt} + \mathbf{p}^i_{opt} - \mathbf{m}^i\right) \\ &= \mathbf{F}\left(\mathbf{b}^i_{opt} + \mathbf{p}^i_{opt}\right) + (\mathbf{I} - \mathbf{F})\,\mathbf{m}^i, \end{aligned} \tag{3}$$

where $\mathbf{F}$ is a low-pass filter in the wavenumber domain, and $\mathbf{I}$ is the identity operator. So far, at the end of outer loop i that we just described above, we found a pair of variables $\mathbf{b}^i_{opt}$ and $\tilde{\mathbf{p}}^i_{opt}$ such that

$$\mathbf{f}(\mathbf{m}^i) + \tilde{\mathbf{B}}(\mathbf{b}^i_{opt})\tilde{\mathbf{p}}^i_{opt} \approx \mathbf{d}^{obs}, \tag{4}$$

which corresponds to minimizing the data fitting term of equation 2 with the constraint that most of the non-physical energy of $\tilde{\mathbf{p}}^i_{opt}$ has been reduced to zero, which in turn means that $\tilde{\mathbf{p}}^i_{opt}$ is not extended (i.e., $\tilde{\mathbf{p}}^i_{opt} \approx \mathbf{p}^i_{opt}$). If we also assume that $\mathbf{d}^{obs} = \mathbf{f}(\mathbf{m}_{true})$, equation 4 becomes

$$\mathbf{f}(\mathbf{m}^i) + \mathbf{B}(\mathbf{b}^i_{opt})\mathbf{p}^i_{opt} \approx \mathbf{f}(\mathbf{m}_{true}). \tag{5}$$

Even though this updating scheme has been shown to work heuristically, there is no mathematical jusification to update $\mathbf{m}^i$ using equation 3. Equation 5 can be seen as a first-order Taylor expansion neither about $\mathbf{m}^i$ nor $\mathbf{b}^i_{opt}$ since they are different. However, if we had allowed the first term of equation 5 to vary as $\mathbf{b}^i$, the Taylor expansion could be justified (and therefore using equation 3 to update $\mathbf{m}^i$ would seem reasonable) as long as $\mathbf{p}^i_{opt}$ is "small" enough. This stepping method does not ensure that the new model decreases the FWI objective function (equation 1). This phenomenon has been observed in Almomin (2016b).

Modified formulation

In our formulation we propose to minimize the following objective function

$$\Phi(\mathbf{m}, \tilde{\mathbf{p}}) = \frac{1}{2}\|\mathbf{f}(\mathbf{m}) + \tilde{\mathbf{B}}(\mathbf{m})\tilde{\mathbf{p}} - \mathbf{d}^{\mathbf{obs}}\|_2^2 + \frac{\epsilon^2}{2}\|\mathbf{D}\tilde{\mathbf{p}}\|_2^2, \tag{6}$$

where $\mathbf{D}$ is a linear operator with respect to $\tilde{\mathbf{p}}$ that enhances the non-physical extended energy of $\tilde{\mathbf{p}}$. This equation differs from equation 2 because $\mathbf{f}(\mathbf{m})$ is not kept constant and $\mathbf{g} = \mathbf{D}$ is linear with respect to $\tilde{\mathbf{p}}$. Since Φ is quadratic with respect to $\tilde{\mathbf{p}}$, we use the variable projection method to solve equation 6 (Golub and Pereyra, 1973; Rickett, 2013; Huang and Symes, 2015), which corresponds to minimizing the following objective function

$$\Phi(\mathbf{m}) = \frac{1}{2}\|\mathbf{f}(\mathbf{m}) + \tilde{\mathbf{B}}(\mathbf{m})\tilde{\mathbf{p}}_{opt}(\mathbf{m}) - \mathbf{d}^{\mathbf{obs}}\|_2^2 + \frac{\epsilon^2}{2}\|\mathbf{D}\tilde{\mathbf{p}}_{opt}(\mathbf{m})\|_2^2, \tag{7}$$

where $\tilde{\mathbf{p}}_{opt}$ is an extended perturbation model, defined as the minimizer of the following objective function $\Phi_{\mathbf{m}}$,

$$\Phi_{\mathbf{m}}(\tilde{\mathbf{p}}) = \frac{1}{2}\|\tilde{\mathbf{B}}(\mathbf{m})\tilde{\mathbf{p}} - \left(\mathbf{d}^{\mathbf{obs}} - \mathbf{f}(\mathbf{m})\right)\|_2^2 + \frac{\epsilon^2}{2}\|\mathbf{D}\tilde{\mathbf{p}}\|_2^2. \tag{8}$$

For a fixed $\mathbf{m}$, $\Phi_{\mathbf{m}}$ reaches its minimum for

$$\tilde{\mathbf{p}}_{opt}(\mathbf{m}) = \left[\tilde{\mathbf{B}}^*(\mathbf{m})\tilde{\mathbf{B}}(\mathbf{m}) + \epsilon^2\mathbf{D}^*\mathbf{D}\right]^{-1}\tilde{\mathbf{B}}^*(\mathbf{m})\left(\mathbf{d}^{\mathbf{obs}} - \mathbf{f}(\mathbf{m})\right), \tag{9}$$

where $*$ denotes adjoint operators. Note that $\tilde{\mathbf{p}}_{opt}$ also depends nonlinearly on $\mathbf{m}$. The data residual component on the right side of equation 7 is a modified FWI objective function where an additional term is used to ensure the phase alignment between modeled and observed data. During the optimization process we slowly reduce the contribution of this additional term by adding a regularization term on the right side of equation 7. Therefore, finding the minimum of this equation is equivalent to minimizing Φ_{FWI} (equation 1).

Equation 8 can be solved by preconditioning the variable $\tilde{\mathbf{p}}$,

$$\tilde{\mathbf{q}} = \mathbf{D}\tilde{\mathbf{p}}, \tag{10}$$

which becomes

$$\Phi_{\mathbf{m}}^{prec}(\tilde{\mathbf{q}}) = \frac{1}{2}\|\tilde{\mathbf{B}}(\mathbf{m})\mathbf{E}\tilde{\mathbf{q}} - (\mathbf{d}^{\mathbf{obs}} - \mathbf{f}(\mathbf{m}))\|_2^2 + \frac{\epsilon^2}{2}\|\tilde{\mathbf{q}}\|_2^2, \tag{11}$$

where $\mathbf{E} = \mathbf{D}^{-1}$. Solution of equation 11 is given by

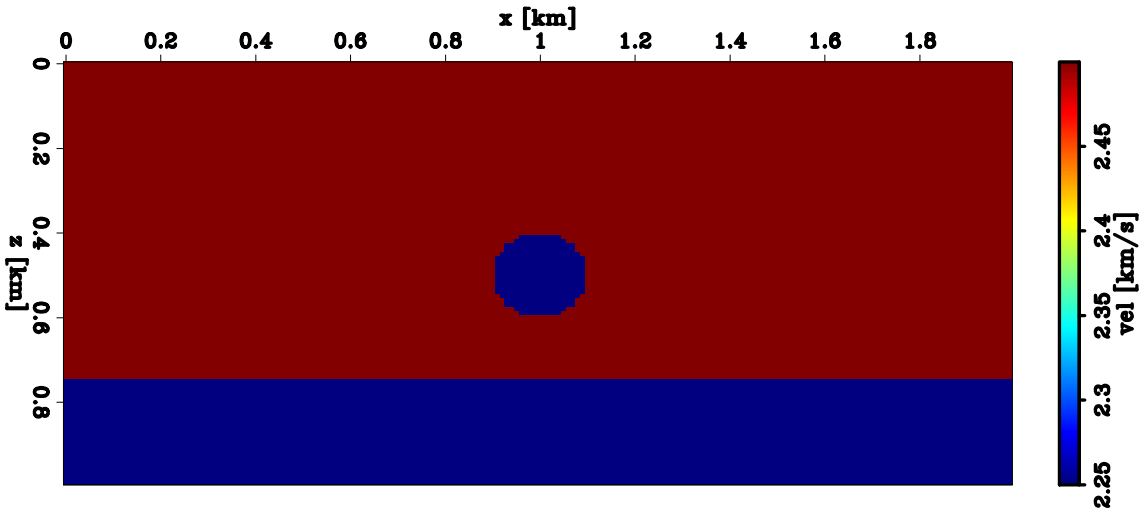


Figure 1: True velocity model similar to the one shown in Mora (1989) where an anomaly is embedded into a two layer subsurface. The velocity of the anomaly and of the second layer is 10% lower than the upper one. [**ER**]

$$\tilde{\mathbf{q}}_{opt}(\mathbf{m}) = \left[\mathbf{E}^*\tilde{\mathbf{B}}^*(\mathbf{m})\tilde{\mathbf{B}}(\mathbf{m})\mathbf{E} + \epsilon^2\mathbf{I}\right]^{-1}\mathbf{E}^*\tilde{\mathbf{B}}^*(\mathbf{m})(\mathbf{d}^{\mathbf{obs}} - \mathbf{f}(\mathbf{m})), \tag{12}$$

where $\mathbf{I}$ is the identity operator in the extended model space. We choose to precondition the problem since the operator $\mathbf{D}$ is badly conditioned as shown by Clapp (2005)

RESULTS

We apply our proposed technique on a 2D synthetic model similar to the one described in Mora (1989). Figure 1 displays the true velocity model used in this test. We generate synthetic data with a Ricker wavelet containing energy from 10 Hz up to 25 Hz (Figure 2). We use 20 shots with 100 m spacing and we place receivers every 10 m. We apply our technique and compare the results to the ones obtained with FWI. For both inversions we inject the full bandwidth without using a multiscale approach (Bunks et al., 1995). The initial model is a constant background with a velocity of the top layer. For both inversions we use a Broyden-Fletcher-Goldfarb-Shanno (BFGS) optimization scheme (Liu and Nocedal, 1989). The linear problem solved during the TFWI optimization has been preconditioned as described in the previous section.

Figure 3 displays the FWI inverted model after 25, 100, 200 iterations. We notice that for the first 25 iterations mostly the reflectivity component of the model is retrieved by the inversion algorithm (Figure 3a). As we increase iteration number the algorithm is able to obtain the short-wavenumber component of the anomaly.

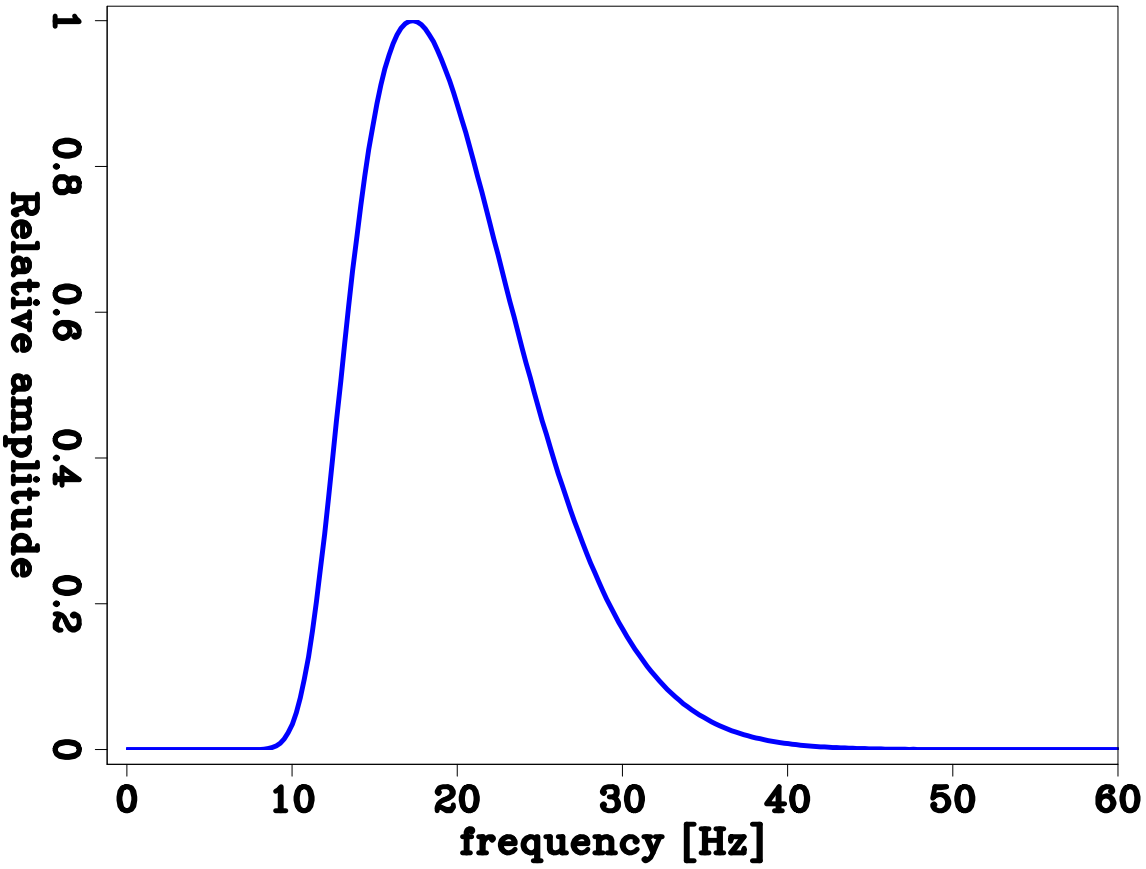


Figure 2: Spectrum of the used wavelet. [**ER**]

Although most of the data are matched after 100 iterations (Figure 4), it is necessary additional 100 iterations to further improve the tomographic part of the model.

Figure 5 shows the TFWI inverted model after 5, 20, and 40 iterations. Because most of the reflection component of the data is matched by the optimal extended reflectivity, the modified TFWI algorithm inverts directly for the tomographic component of the model. Additionally, the reflection part of the model is correctly inverted. In fact, the reflector below the anomaly is present in the TFWI inverted model. As the algorithm progresses, the contribution of the extended reflectivity decreases (Figure 6). We observe the same inversion behavior as for the FWI result. In fact, despite that most of the data are matched after 20 iterations, the model is still changing significantly with additional 20 iterations.

To compare the FWI and TFWI inverted model we plot a vertical and a horizontal velocity profiles passing through the center of the anomaly along with the true model (Figure 7). Both inversion algorithms achieve similar results that in good agreement with the true velocity model. The ringing effects are caused by the limited bandwidth nature of the recorded data. In addition, the velocity contrast of the reflector below the anomaly is incorrectly placed due to the depth uncertainty present in the data.

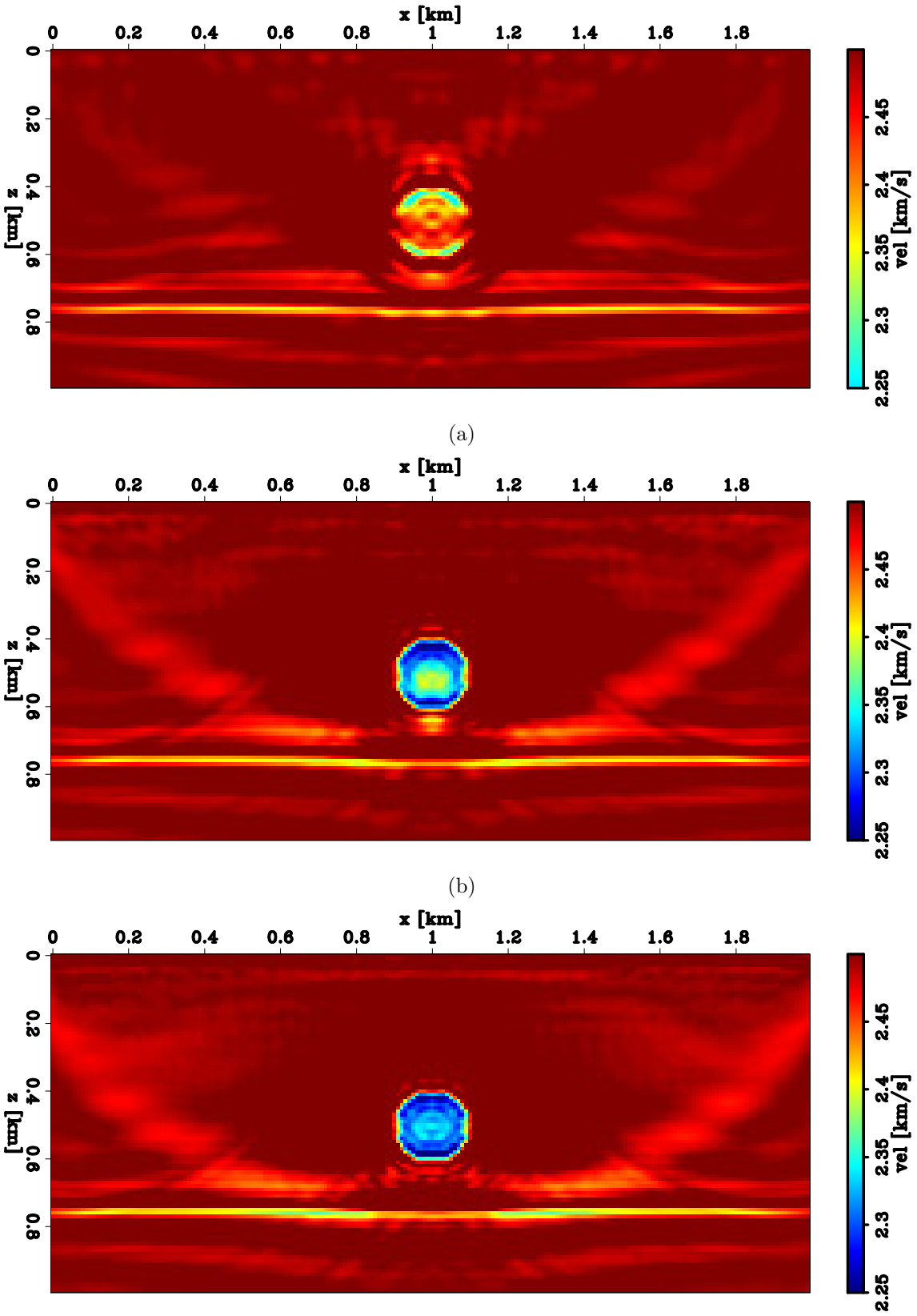
To understand the advantage of preconditioning the linear inversion, we compare the convergence curves at the first non-linear iteration of the TFWI problem for preconditioned and un-preconditioned linear inversion (Figure 8). These curves show that the rate of convergence of the linear problem is greatly improved thanks to the preconditioning approach used.

CONCLUSIONS AND FUTURE WORK

We highlighted issues associated with the original implementation of the TFWI algorithm and proposed a new optimization scheme to potentially overcome these problems. In this new method we let the non-linear modeling operator vary during optimization. We also make use of the variable projection method to optimize the linear component independently of the non-linear one. Moreover, we showed the advantage of preconditioning the linear inversion to improve the convergence rate of the linear problem. On a model composed of two layers and a velocity anomaly we demonstrated the consistency of the proposed TFWI algorithm with FWI. Despite the lack of energy below 10 Hz in the generated data, FWI did not cycle skip in this case and both methods retrieved similar velocity models. Future work will involve the verification of global convergence on a more complicated model in which FWI fails to attain the global minimum. In addition, a complete comparison of different TFWI algorithms will be performed.

REFERENCES

Almomin, A., 2016a, Tomographic full waveform inversion: Stanford University.



(a)

(b)

(c)

Figure 3: FWI results after (a) 25, (b) 100, and (c) 200 iterations, respectively. [**CR**]

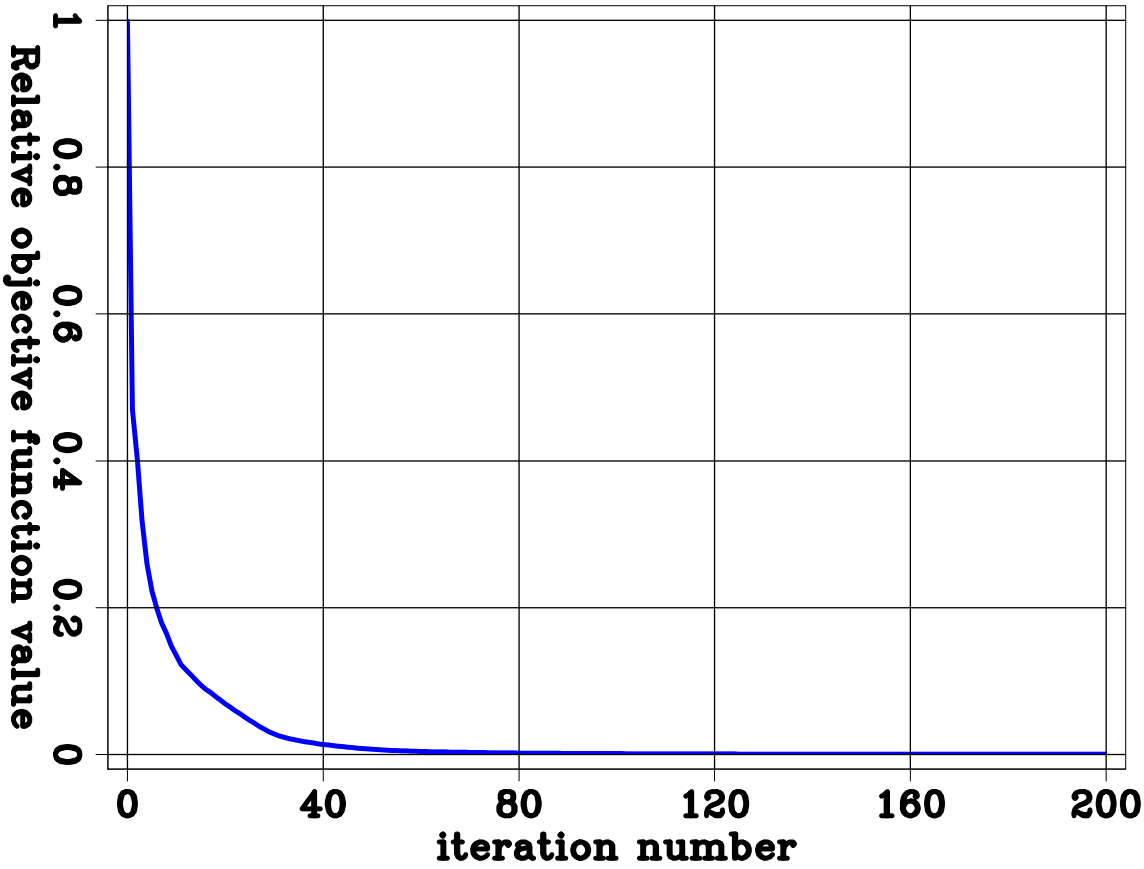
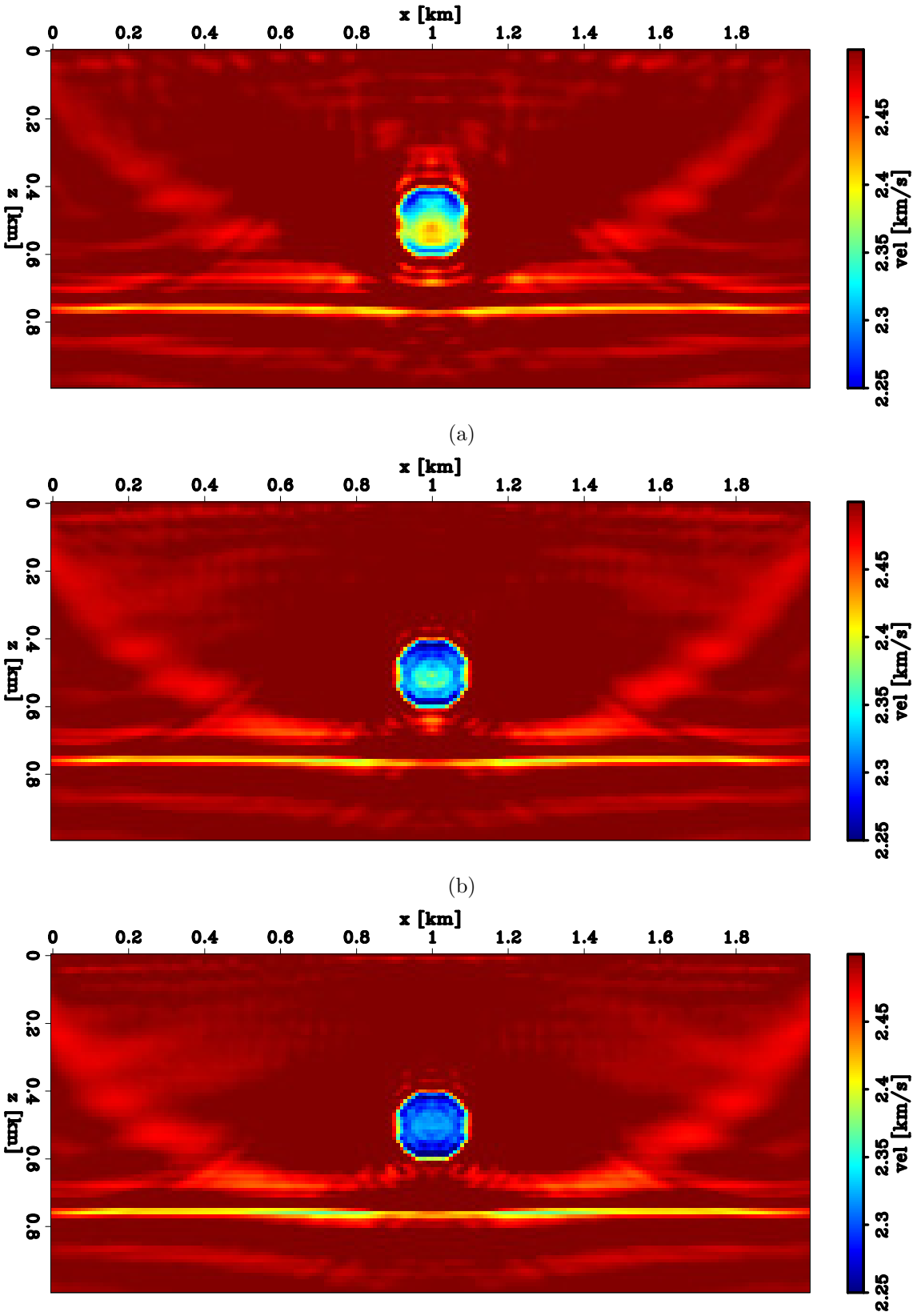


Figure 4: Relative FWI objective function (equation 7). [**CR**]



(a)

(b)

(c)

Figure 5: Modified TFWI results after (a) 5, (b) 20, and (c) 40 iterations, respectively. [**CR**]

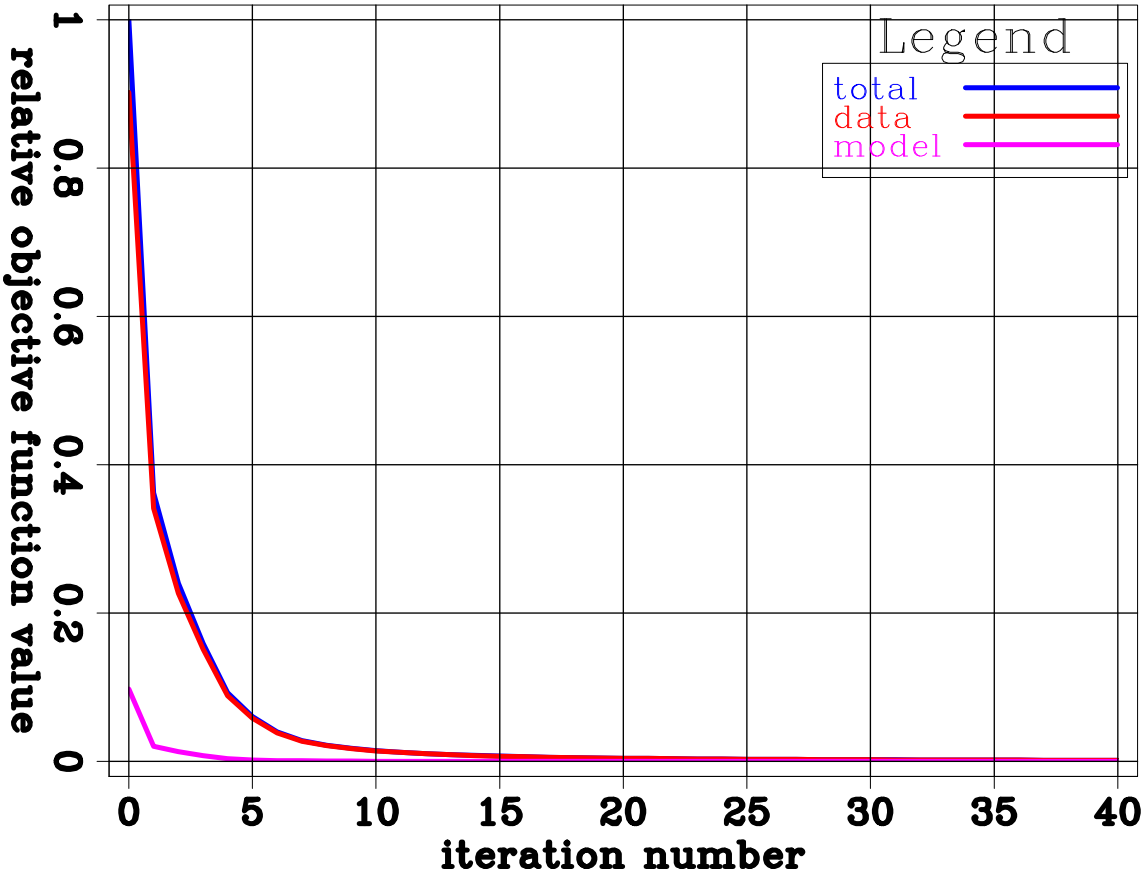
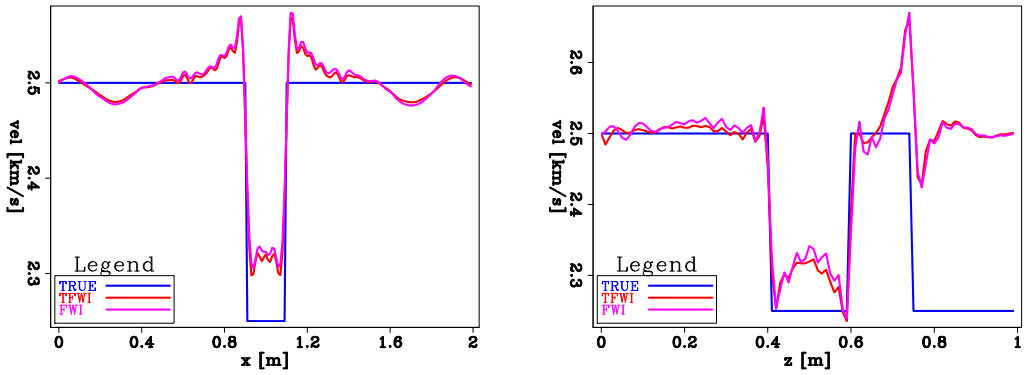


Figure 6: Relative TFWI objective function (equation 7). The total, data, and model objective functions are plotted. [**CR**]



(a) (b)

Figure 7: Horizontal and vertical velocity profiles of FWI and TFWI results at the lastest iteration, and true model. Both profiles are passing through the center of the velocity anomaly. [**CR**]

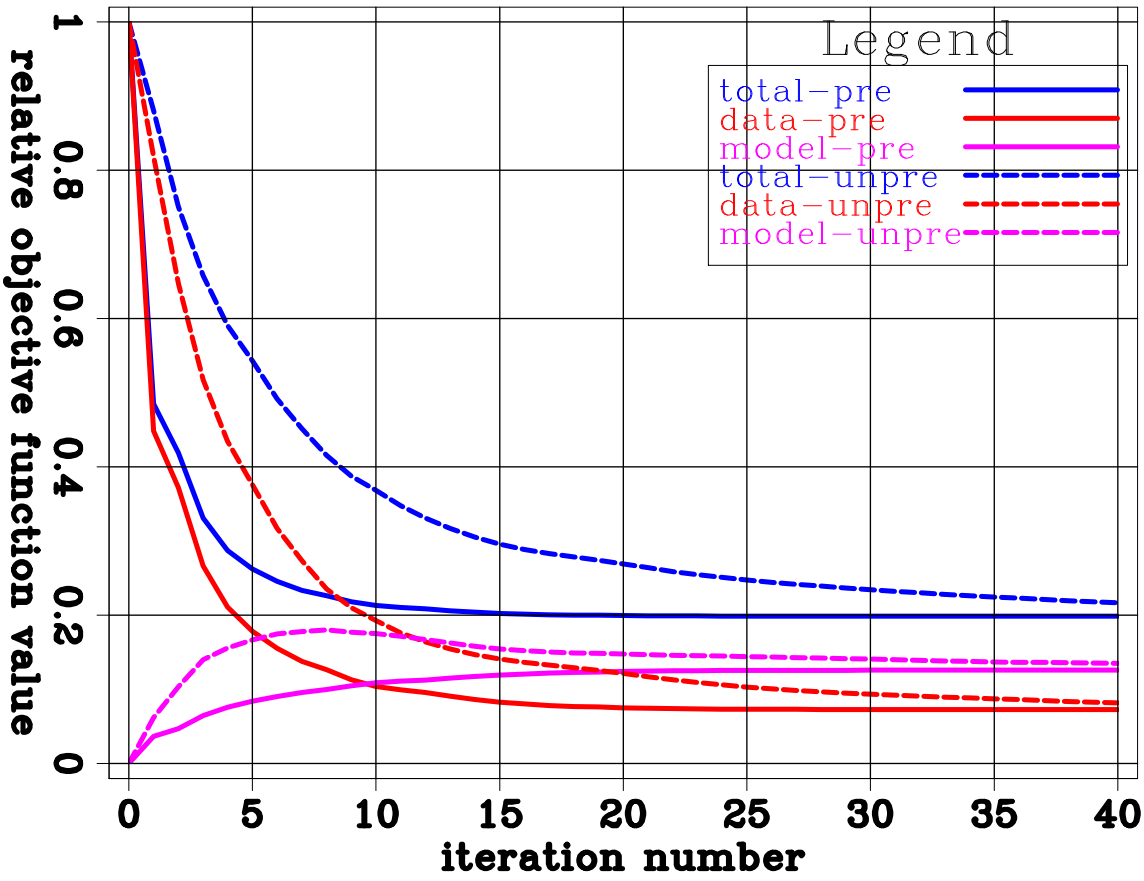


Figure 8: Convergence curves for preconditioned (solid curves) and un-preconditioned (dashed curves) linear inverse problem (equations 8 and 11) at the first non-linear iteration of the modified TFWI. [**CR**]

———, 2016b, Tomographic full waveform inversion: Presented at the 2016 SESAAI Annual Meeting.

Biondi, B. and A. Almomin, 2014, Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion: Geophysics, **79**, WA129–WA140.

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

Clapp, M., 2005, Imaging under salt: illumination compensation by regularized inversion: Stanford University.

Golub, G. H. and V. Pereyra, 1973, The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate: SIAM Journal on numerical analysis, **10**, 413–432.

Huang, Y. and W. W. Symes, 2015, Born waveform inversion via variable projection and shot record model extension: SEG Technical Program Expanded Abstracts 2015, 1326–1331, Society of Exploration Geophysicists.

Liu, D. C. and J. Nocedal, 1989, On the limited memory BFGS method for large scale optimization: Mathematical programming, **45**, 503–528.

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

Rickett, J., 2013, The variable projection method for waveform inversion with an unknown source function: Geophysical Prospecting, **61**, 874–881.

APPENDIX

In this appendix we derive the gradient of equation 7. To compute the gradient of equation 7, we first define

$$\mathbf{r}_d(\mathbf{m}) = \mathbf{f}(\mathbf{m}) + \tilde{\mathbf{B}}(\mathbf{m})\tilde{\mathbf{p}}_{opt}(\mathbf{m}) - \mathbf{d}^{\mathbf{obs}} \tag{13}$$

$$\mathbf{r}_m(\mathbf{m}) = \mathbf{D}\tilde{\mathbf{p}}_{opt}(\mathbf{m}). \tag{14}$$

The gradient of Φ is given by

$$\nabla\Phi(\mathbf{m}) = \left(\frac{\partial \mathbf{r}_d(\mathbf{m})}{\partial \mathbf{m}}\right)^* \mathbf{r}_d(\mathbf{m}) + \epsilon^2 \left(\frac{\partial \mathbf{r}_m(\mathbf{m})}{\partial \mathbf{m}}\right)^* \mathbf{r}_m(\mathbf{m}), \tag{15}$$

and we have

$$\begin{aligned}\left(\frac{\partial \mathbf{r}_d(\mathbf{m})}{\partial \mathbf{m}}\right)^* &= \left(\frac{\partial \mathbf{f}(\mathbf{m})}{\partial \mathbf{m}}\right)^* + \left(\frac{\partial(\tilde{\mathbf{B}}(\mathbf{m})\tilde{\mathbf{p}}_{opt}(\mathbf{m}))}{\partial \mathbf{m}}\right)^* \\ &= \mathbf{B}^*(\mathbf{m}) + \left(\frac{\partial \tilde{\mathbf{B}}(\mathbf{m})}{\partial \mathbf{m}}\tilde{\mathbf{p}}_{opt}(\mathbf{m}) + \tilde{\mathbf{B}}(\mathbf{m})\frac{\partial \tilde{\mathbf{p}}_{opt}(\mathbf{m})}{\partial \mathbf{m}}\right)^* \\ &= \mathbf{B}^*(\mathbf{m}) + \mathbf{T}^*(\mathbf{m}) + \left(\frac{\partial \tilde{\mathbf{p}}_{opt}(\mathbf{m})}{\partial \mathbf{m}}\right)^* \tilde{\mathbf{B}}(\mathbf{m})^*, \end{aligned} \tag{16}$$

where $\mathbf{T}^*(\mathbf{m}) = \left(\frac{\partial \tilde{\mathbf{B}}(\mathbf{m})}{\partial \mathbf{m}}\tilde{\mathbf{p}}_{opt}(\mathbf{m})\right)^*$. Similarly,

$$\left(\frac{\partial \mathbf{r}_m(\mathbf{m})}{\partial \mathbf{m}}\right)^* = \left(\frac{\partial \tilde{\mathbf{p}}_{opt}(\mathbf{m})}{\partial \mathbf{m}}\right)^* \mathbf{D}^*. \tag{17}$$

Equation 15 becomes

$$\begin{aligned}\nabla\Phi(\mathbf{m}) &= \left[\mathbf{B}^*(\mathbf{m}) + \mathbf{T}^*(\mathbf{m}) + \left(\frac{\partial \tilde{\mathbf{p}}_{opt}(\mathbf{m})}{\partial \mathbf{m}}\right)^* \tilde{\mathbf{B}}(\mathbf{m})^*\right]\mathbf{r}_d(\mathbf{m}) + \epsilon^2\left(\frac{\partial \tilde{\mathbf{p}}_{opt}(\mathbf{m})}{\partial \mathbf{m}}\right)^* \mathbf{D}^*\mathbf{r}_m(\mathbf{m}) \\ &= \left[\mathbf{B}^*(\mathbf{m}) + \mathbf{T}^*(\mathbf{m})\right]\mathbf{r}_d(\mathbf{m}) + \left(\frac{\partial \tilde{\mathbf{p}}_{opt}(\mathbf{m})}{\partial \mathbf{m}}\right)^* \left[\tilde{\mathbf{B}}^*(\mathbf{m})\mathbf{r}_d(\mathbf{m}) + \epsilon^2\mathbf{D}^*\mathbf{r}_m(\mathbf{m})\right]. \end{aligned} \tag{18}$$

Since $\tilde{\mathbf{p}}_{opt}$ satisfies equation 9, we have

$$\left[\tilde{\mathbf{B}}^*(\mathbf{m})\tilde{\mathbf{B}}(\mathbf{m}) + \epsilon^2\mathbf{D}^*\mathbf{D}\right]\tilde{\mathbf{p}}_{opt}(\mathbf{m}) = \tilde{\mathbf{B}}^*(\mathbf{m})(\mathbf{d}^{\mathbf{obs}} - \mathbf{f}(\mathbf{m})). \tag{19}$$

Therefore,

$$\begin{aligned}\tilde{\mathbf{B}}^*(\mathbf{m})\mathbf{r}_d(\mathbf{m}) + \epsilon^2\mathbf{D}^*\mathbf{r}_m(\mathbf{m}) &= \tilde{\mathbf{B}}^*(\mathbf{m})\left(\mathbf{f}(\mathbf{m}) + \tilde{\mathbf{B}}(\mathbf{m})\tilde{\mathbf{p}}_{opt}(\mathbf{m}) - \mathbf{d}^{\mathbf{obs}}\right) + \epsilon^2\mathbf{D}^*\mathbf{D}\tilde{\mathbf{p}}_{opt}(\mathbf{m}) \\ &= \left[\tilde{\mathbf{B}}^*(\mathbf{m})\tilde{\mathbf{B}}(\mathbf{m}) + \epsilon^2\mathbf{D}^*\mathbf{D}\right]\tilde{\mathbf{p}}_{opt}(\mathbf{m}) - \tilde{\mathbf{B}}^*(\mathbf{m})(\mathbf{d}^{\mathbf{obs}} - \mathbf{f}(\mathbf{m})) \\ &= \mathbf{0}. \end{aligned} \tag{20}$$

Finally, equation 15 reduces to

$$\nabla\Phi(\mathbf{m}) = \left[\mathbf{B}^*(\mathbf{m}) + \mathbf{T}^*(\mathbf{m})\right]\left(\mathbf{f}(\mathbf{m}) + \tilde{\mathbf{B}}(\mathbf{m})\tilde{\mathbf{p}}_{opt}(\mathbf{m}) - \mathbf{d}^{\mathbf{obs}}\right). \tag{21}$$


SINGER