Frequency-domain full waveform inversion with updates based on nonlinear sensitivities

Yu Geng, Wenyong Pan and Kristopher A. Innanen, CREWES project, Dept. of Geoscience, University of Calgary*

SUMMARY

Although full waveform inversion (FWI) is a highly nonlinear inverse problem, it is usually solved as a local optimization problem under a linear approximation, where small angle backscattered data (via the residuals) are treated linearly. Adding nonlinearity within each update may have important consequences for convergence rates and parameter accuracy. One approach is to include higher-order scattering terms into the sensitivities during the construction of the gradient, by varying not the current but the updated model at each iteration. By applying inverse scattering theory, this additional sensitivity term can be computed from the data residuals at the current iteration. A nonlinear frequency-domain FWI inversion scheme, with an inner and an outer loop, implementing this idea is presented here. A perturbation is inverted from the data residuals within the inner loop, and the descent direction based on the nonlinear sensitivity to update the model is computed involving this perturbation in the outer loop. We test this nonlinear FWI on acoustic single-parameter Marmousi synthetics. The inverted results vary depending on data frequency ranges and initial models, but we conclude that the nonlinear FWI has the capability to generate high resolution model estimates in both shallow and deep regions, and to converge rapidly, relative to a benchmark FWI approach involving the standard gradient.

INTRODUCTION

Although FWI (Lailly, 1983, Tarantola, 1984, Virieux and Operto, 2009) is a highly nonlinear inverse problem, it is usually solved as a local optimization problem in the framework of the Born approximation. The fundamentally nonlinear data-model relationship is accounted mainly through iteration and updating of linearized relationships. Linearization error in seismic inversion can be significant in many cases (Wu and Zheng, 2014), especially when large-contrast and spatially sustained perturbations are involved. The former situation prevails when large-angle reflection information is being considered, which, in reflection configurations, is the regime in which many high-priority elastic and/or anisotropic parameters are distinguished (Innanen, 2015). The Fréchet derivative, or sensitivity, restricts the resolution of FWI and could cause a strong deficit of low wavenumber components in the updated model due to the lack of large-aperture illumination and low frequency information. Furthermore, cycle-skipping artifacts occur when the Born approximation is no longer valid, leading the optimization to a local minimum (Virieux and Operto, 2009).

To establish a better-posed reconstruction of low wavenumber model components, and help FWI to converge to a global minimum, several approaches can be employed, e.g., by retrieving the low frequency information in the data as in Laplace-domain and Laplace-Fourier-domain inversion (Shin and Cha, 2008, 2009), envelop inversion (Wu, Luo, and Wu, 2014); by building the low wavenumber background model through reflection information alone (Xu et al., 2012, Brossier, Operto, and Virieux, 2015), or, together with refractions (Wang et al., 2015, Zhou et al., 2015); and by building the background model and perturbation simultaneously, in the data domain or mixed data/image domain (Sun and Symes, 2012, Albertin, Shan, and Washbourne, 2013, Biondi and Almomin, 2014, Wu and Alkhalifah, 2015, Alkhalifah and Wu, 2016).

Mitigation of linearization errors within each iteration could have a significant impact on convergence rates and inversion results. Using nonlinear sensitivities, e.g., in resistivity inversion (Mcgillivray and Oldenburg, 1990), optical imaging (Kwon and Yazici, 2010) and seismic inversion (Wu and Zheng, 2014, Innanen, 2015) has been considered. Here, we consider an extension of frequency domain FWI (e.g., Sirgue and Pratt, 2004) to incorporate nonlinear sensitivities, building on the procedure introduced by Innanen (2015). A two-loop inversion scheme is employed, in which, in the inner iterations, a perturbation is determined from the data residuals using linear inversion, and this is used, in the outer iterations, to determine a descent direction which in principle anticipates some of the curvature of the objective function caused by data-model nonlinearity. In wave physics terms, the introduction of higher order sensitivities involves transmission wave paths from each scattering point to both sources and receivers at the surface. The application on the Marmousi model shows that this nonlinear FWI converges more rapidly than does FWI with a conventional gradient, and appears to be adept at reconstructing low wavenumber components even absent rich low frequencies.

THEORY

Nonlinear sensitivity

In this paper, the space/frequency-domain isotropic acoustic wave equation with constant density is used to describe wave motion:

$$\left(\omega^2 s(\mathbf{r}) + \nabla^2\right) G(\mathbf{r}, \mathbf{r}_s, \omega) = -\delta(\mathbf{r} - \mathbf{r}_s), \qquad (1)$$

where ω is the frequency, $s(\mathbf{r})$ is the squared slowness, $\mathbf{r}_s$ is the source location, and $G(\mathbf{r},\mathbf{r}_s,\omega)$ is the Green's function. For simplicity, we hereafter omit ω in all wave expressions, e.g., we use $G(\mathbf{r},\mathbf{r}_s)$ rather than $G(\mathbf{r},\mathbf{r}_s,\omega)$.

At the nth iteration, FWI seeks to minimize the misfit function

$$\phi(s)=\frac{1}{2}\sum_{\mathbf{r}_s}\sum_{\mathbf{r}_g}\sum_{\omega}\left\|\delta d(s_n)\right\|^2, \qquad (2)$$

which is the L^2 norm of the data residual $\delta d(s_n)$, which is the difference between the recorded data $d(\mathbf{r}_g,\mathbf{r}_s)$ at receiver $\mathbf{r}_g$ and modeled data $P(\mathbf{r}_g,\mathbf{r}_s\,|\,s_n)=FG(\mathbf{r}_g,\mathbf{r}_s\,|\,s_n)$, where F is the source spectrum. It does so by iteratively updating the current model, $s_n(\mathbf{r})$ with a perturbation $\Delta s_n(\mathbf{r})$ along some appropriate descent direction with a certain step length α_n. This is assumed to converge to the true velocity model $s(\mathbf{r})$.

Although the descent direction is different for different optimization methods, the gradient of the misfit function is always involved in the calculation. At a given frequency, the gradient is

$$g_n(\mathbf{r})=-\sum_{\mathbf{r}_s,\mathbf{r}_g}\mathrm{Re}\left(F\frac{\partial G(\mathbf{r}_g,\mathbf{r}_s\,|\,s_n)}{\partial s(\mathbf{r})}\delta d^*(\mathbf{r}_g,\mathbf{r}_s\,|\,s_n)\right), \qquad (3)$$

where * stands for the complex conjugate, and $\partial G/\partial s$ is the sensitivity determined by varying the field $G(\mathbf{r}_g,\mathbf{r}_s\,|\,s_n)$ in the current medium iterate $s_n(\mathbf{r})$:

$$\frac{\partial G(\mathbf{r}_g,\mathbf{r}_s\,|\,s_n)}{\partial s(\mathbf{r})}=-\omega^2G(\mathbf{r}_g,\mathbf{r}\,|\,s_n)G(\mathbf{r},\mathbf{r}_s\,|\,s_n). \qquad (4)$$

Equation (4) reflects the "true" sensitivity of the medium only to the extent that $s_n(\mathbf{r})$ reflects the true medium properties. Prior to convergence, a different sensitivity – in fact, one which implicitly or explicitly depends on the residuals, is expected if we instead vary the field in the medium iterate under construction, $s_{n+1}(\mathbf{r})$. At the n+1th iteration, this updated model is $s_{n+1}(\mathbf{r})=s_n(\mathbf{r})+\Delta s_n(\mathbf{r})$. Varying the medium by a δs localized at $\mathbf{r}$ in $s_{n+1}(\mathbf{r})$, the perturbation caused by this added variation δs and $\Delta s_n(\mathbf{r})$ can be written (Innanen, 2015)

$$\begin{aligned}\delta G(\mathbf{r}_g,\mathbf{r}_s\,|\,s_{n+1},\delta s)=&-\delta s\left(\omega^2G(\mathbf{r}_g,\mathbf{r}\,|\,s_n)G(\mathbf{r},\mathbf{r}_s\,|\,s_n)\right.\\&+\delta G(\mathbf{r}_g,\mathbf{r}\,|\,s_n,\Delta s_n)G(\mathbf{r},\mathbf{r}_s\,|\,s_n)\\&\left.+G(\mathbf{r}_g,\mathbf{r}\,|\,s_n)\delta G(\mathbf{r},\mathbf{r}_s\,|\,s_n,\Delta s_n)\right)+\ldots\end{aligned}, \qquad (5)$$

where the first term can be interpreted as a scattering process involving one interaction with the variation δs only, and $\delta G(\mathbf{r},\mathbf{r}_s\,|\,s_n,\Delta s_n)$ and $\delta G(\mathbf{r}_g,\mathbf{r}\,|\,s_n,\Delta s_n)$ are the scattered wavefields related to $\Delta s_n(\mathbf{r})$,

$$\delta G(\mathbf{r},\mathbf{r}_s\,|\,s_n,\Delta s_n)=-\omega^2\int d\mathbf{r}'G(\mathbf{r},\mathbf{r}'\,|\,s_n)\Delta s_n(\mathbf{r}')G(\mathbf{r}',\mathbf{r}_s\,|\,s_{n+1})$$

$$\delta G(\mathbf{r}_g,\mathbf{r}\,|\,s_n,\Delta s_n)=-\omega^2\int d\mathbf{r}'G(\mathbf{r}_g,\mathbf{r}'\,|\,s_n)\Delta s_n(\mathbf{r}')G(\mathbf{r}',\mathbf{r}\,|\,s_{n+1})$$

which includes multiple interactions with $\Delta s_n(\mathbf{r})$ and one interaction with δs in the remaining terms in equation (5). The resulting sensitivity is

$$\frac{\partial G(\mathbf{r}_g,\mathbf{r}_s\,|\,s_{n+1})}{\partial s(\mathbf{r})}=\left(\frac{\partial G(\mathbf{r}_g,\mathbf{r}_s\,|\,s_{n+1})}{\partial s(\mathbf{r})}\right)_0+\left(\frac{\partial G(\mathbf{r}_g,\mathbf{r}_s\,|\,s_{n+1})}{\partial s(\mathbf{r})}\right)_1+\ldots, \qquad (6)$$

where the index refers to the order of the term in the perturbation $\Delta s_n(\mathbf{r})$,

$$\left(\frac{\partial G(\mathbf{r}_g,\mathbf{r}_s\,|\,s_{n+1})}{\partial s(\mathbf{r})}\right)_0=-\omega^2G(\mathbf{r}_g,\mathbf{r}\,|\,s_n)G(\mathbf{r},\mathbf{r}_s\,|\,s_n), \qquad (7)$$

$$\begin{aligned}\left(\frac{\partial G(\mathbf{r}_g,\mathbf{r}_s\,|\,s_{n+1})}{\partial s(\mathbf{r})}\right)_1+\left(\frac{\partial G(\mathbf{r}_g,\mathbf{r}_s\,|\,s_{n+1})}{\partial s(\mathbf{r})}\right)_2+\ldots&\\=-\omega^2\left(\delta G(\mathbf{r}_g,\mathbf{r}\,|\,s_n,\Delta s_n)G(\mathbf{r},\mathbf{r}_s\,|\,s_n)\right.&\\\left.+G(\mathbf{r}_g,\mathbf{r}\,|\,s_n)\delta G(\mathbf{r},\mathbf{r}_s\,|\,s_n,\Delta s_n)\right)&\end{aligned}. \qquad (8)$$

When $\Delta s_n(\mathbf{r})$ is zero, the higher order term (8) vanish, and the sensitivity reduces to the conventional form in (4).

Calculation of $\Delta s_n(\mathbf{r})$ before n+1th iteration

All of the Green's functions used in the new sensitivity in (6) are still calculated in the nth model $s_n(\mathbf{r})$. However, to calculate the higher order terms (8) in the new sensitivities, $\Delta s_n(\mathbf{r})$ is required. This quantity is unknown, but it has a relationship with the nth data residual. Innanen (2015) showed the perturbation $\Delta s_n(\mathbf{r})$ can be expressed with a series related to the nth data residual $\delta d(s_n)$, according to inverse scattering theory,

$$\delta d(s_n)=-F\omega^2\int d\mathbf{r}'G(\mathbf{r}_g,\mathbf{r}'\,|\,s_n)\Delta s_n(\mathbf{r}')G(\mathbf{r}',\mathbf{r}_s\,|\,s_n)+\ldots. \qquad (9)$$

Instead of using standard inverse scattering series techniques (Weglein et al., 2003), taking only the first order term, the perturbation $\Delta s_n(\mathbf{r})$ can be approximated by the solution $\Delta s(\mathbf{r})$ of a further, second data fitting scheme in a least squares sense, by minimizing

$$\phi(\Delta s)=\frac{1}{2}\sum_{\mathbf{r}_s}\sum_{\mathbf{r}_g}\left\|\delta d(s_n)-\delta P(\mathbf{r}_g,\mathbf{r}_s\,|\,s_n)\right\|^2, \qquad (10)$$

where $\delta P(\mathbf{r}_g,\mathbf{r}_s\,|\,s_n)$ is the scattered data obtained by sampling the scattered wavefield $\delta P(\mathbf{r},\mathbf{r}_s\,|\,s_n)$ at receiver $\mathbf{r}_g$, which is found by solving the wave equation with a virtual secondary source $-\omega^2\Delta s(\mathbf{r}')P(\mathbf{r}',\mathbf{r}_s\,|\,s_n)$:

$$\left(\omega^2s_n(\mathbf{r})+\nabla^2\right)\delta P(\mathbf{r},\mathbf{r}_s\,|\,s_n)=-\omega^2\Delta s(\mathbf{r}')P(\mathbf{r}',\mathbf{r}_s\,|\,s_n). \qquad (11)$$

Gradient with higher-order sensitivities

Substituting the perturbation $\Delta s(\mathbf{r})$ into the nonlinear sensitivity (6), the nth model update for $s_n(\mathbf{r})$ is

$$g_n(\mathbf{r}) = \sum_{\mathbf{r}_s,\mathbf{r}_g} \mathrm{Re}\Big(\omega^2 \delta d^*(s_n)\big(G(\mathbf{r}_g,\mathbf{r}\,|\,s_n)P(\mathbf{r},\mathbf{r}_s\,|\,s_n) + \delta P(\mathbf{r}_g,\mathbf{r}\,|\,s_n)P(\mathbf{r},\mathbf{r}_s\,|\,s_n) + G(\mathbf{r}_g,\mathbf{r}\,|\,s_n)\delta P(\mathbf{r},\mathbf{r}_s\,|\,s_n)\big)\Big), \quad (12)$$

where $\delta P(\mathbf{r},\mathbf{r}_s\,|\,s_n)$ is as in (11) and $\delta P(\mathbf{r}_g,\mathbf{r}\,|\,s_n)$ is the scattered wavefield obtained by solving the wave equation with virtual secondary source $-\omega^2 \Delta s(\mathbf{r}')G(\mathbf{r}_g,\mathbf{r}'\,|\,s_n)$. In this approximation, only up to the 1st order term in (6) is included. The perturbation $\Delta s(\mathbf{r})$ can, lead to a direct update using the perturbation without line searching as

$$\tilde{s}_{n+1}(\mathbf{r}) = s_n(\mathbf{r}) + \Delta s(\mathbf{r}). \quad (13)$$

By exchanging $\delta d(s_n)$ with $\delta d(\tilde{s}_{n+1})$ in (12), we obtain the gradient for updating $\tilde{s}_{n+1}(\mathbf{r})$ with the higher order sensitivities in the model $\tilde{s}_{n+1}(\mathbf{r})$,

$$\begin{aligned} g_{n+1}(\mathbf{r}) &= \sum_{\mathbf{r}_s,\mathbf{r}_g} \mathrm{Re}\Big(\omega^2 \delta d^*(\tilde{s}_{n+1})\big(G(\mathbf{r}_g,\mathbf{r}\,|\,s_n)P(\mathbf{r},\mathbf{r}_s\,|\,s_n) \\ &+ \delta P(\mathbf{r}_g,\mathbf{r}\,|\,\tilde{s}_{n+1})P(\mathbf{r},\mathbf{r}_s\,|\,s_n) + G(\mathbf{r}_g,\mathbf{r}\,|\,s_n)\delta P(\mathbf{r},\mathbf{r}_s\,|\,\tilde{s}_{n+1})\big)\Big) \\ &\approx \sum_{\mathbf{r}_s,\mathbf{r}_g} \mathrm{Re}\Big(\omega^2 \delta d^*(\tilde{s}_{n+1})\big(G(\mathbf{r}_g,\mathbf{r}\,|\,\tilde{s}_{n+1})P(\mathbf{r},\mathbf{r}_s\,|\,\tilde{s}_{n+1}) \\ &- \delta P(\mathbf{r}_g,\mathbf{r}\,|\,\tilde{s}_{n+1})\delta P(\mathbf{r},\mathbf{r}_s\,|\,\tilde{s}_{n+1})\big)\Big) \end{aligned} \quad (14)$$

EXAMPLES

The Marmousi model is used to test the use of the nonlinear gradients (12) and (14), hereafter FOFWI and NFWI respectively. FWI solved with conventional gradients, by steepest descent (SD) and truncated Gaussian-Newton methods (GN) (Metivier et al., 2013, Pan, Innanen, and Liao, 2017) will be used for comparison. Two initial models and different frequency ranges are used to test convergence and resolution characteristics of the method. Synthetic data are generated in the frequency domain with 461 receivers and 46 sources along the surface. The true model is shown in Figure 1a, which has a 500*m* added water layer on top of the classical Marmousi model.

The first initial model is obtained from the true velocity using a Gaussian smoother (Figure 1b). With 3 frequencies (4Hz, 6.6Hz, 14.9Hz) starting at 4Hz, 10 iterations per frequency are used in both conventional and nonlinear FWI (updating only the velocity under the added water layer). The SD FWI result is shown in Figure 1c. 5 inner iterations are used for the GN FWI method (Figure 1d) and in FOFWI (Figure 1e) and NFWI (Figure 1f). For the first frequency, Figure 2a shows the residual norm vs. # of iterations, and Figure 2b shows the model error norm vs. # of iterations, with red, cyan blue line and blue plus line stands for FWI, GN FWI, NFWI and FOFWI result, respectively. FWI with the nonlinear gradient converges more rapidly than conventional SD FWI. Because of the lack of low frequencies, GN FWI performs similarly to SD FWI, but the convergence of both FOFWI and NFWI outperforms SD FWI; the NFWI result is particularly close to the true model.

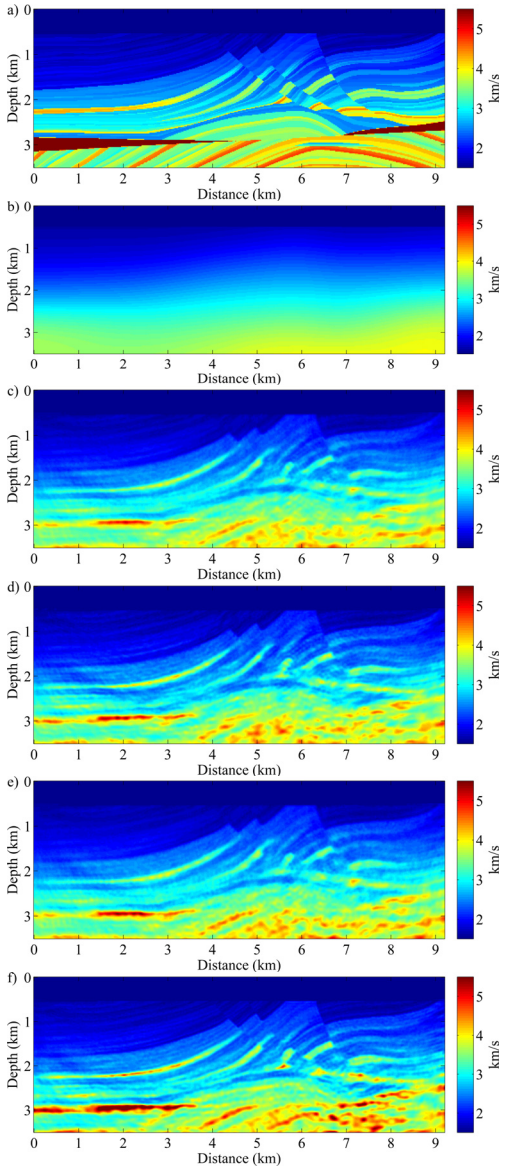

Figure 1: a) True Marmousi velocity model; b) initial velocity model. Inversion results with conventional gradient c) using SD and d) GN method, e) FOFWI and f) NFWI with 10 iterations for 3 frequencies starting from 4Hz to 15Hz.

Next, we use a linear initial model (Figure 3a) to test NFWI. 5 frequencies (2Hz, 3.3Hz, 5.5Hz, 9Hz and 14.9Hz) are used, with 10 iterations per frequency. Figure 3b-c show the SD and GN FWI results respectively, and Figure 3d shows the NFWI result (5 inner iterations are used for both

GN FWI and NFWI). Profiles along x = 4km, 6km are shown in Figure 4, with red, cyan and blue line stands for FWI, GN FWI and NFWI result, respectively. For the first frequency, Figure 5a shows the residual norm vs. # of iterations, and Figure 5b shows the model error norm vs. # of iterations. We observe that, for the given number of iterations, with frequency as low as 2Hz, SD FWI can only reconstruct rough model structures, while GN FWI can help converge much faster towards the true model. But, NFWI outperforms both, especially in the deeper regions.

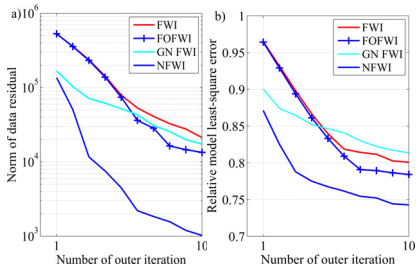

Figure 2: a) Norm of data residual vector vs. number of iterations; b) relative model least-squares error vs. number of iterations.

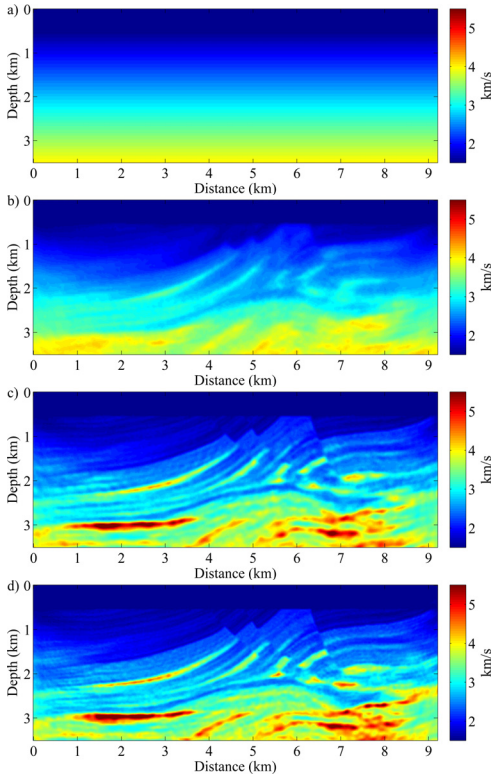

Figure 3: a) Initial velocity model; Inversion result with conventional gradient by b) SD and c) GN, d) NFWI with 10 iterations for 5 frequencies starting from 2Hz to 15Hz.

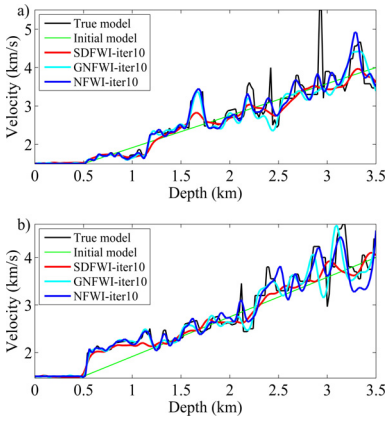

Figure 4: Velocity profiles along a) x = 4km and b) x = 6km.

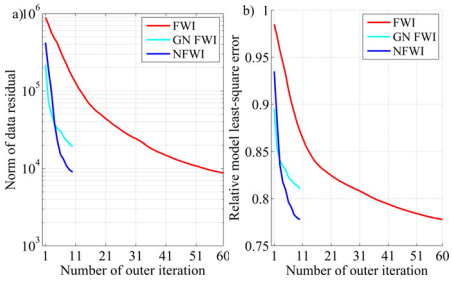

Figure 5: a) Norm of data residual vector vs. number of iterations; b) relative model least-squares error vs. number of iterations.

CONCLUSIONS

In this study, by computing both zeroth and higher order sensitivity terms, we have constructed a two-loop nonlinear frequency domain FWI. In this approach, at each frequency, a linear inversion is used to obtain a rough perturbation model from the current data residual before calculating the nonlinear FWI gradient. The resulting nonlinear FWI appears to provide accurate and well-resolved inversion results. Tests on the Marmousi model with different initial models and frequency bands illustrate the convergence characteristics of the nonlinear FWI.

ACKNOWLEDGEMENTS

We thank the sponsors of CREWES for support. This work was funded by CREWES and NSERC (Natural Science and Engineering Research Council of Canada) through the grant CRDPJ 379744-08.

Albertin, U., G. Shan, and J. Washbourne, 2013, Gradient orthogonalization in adjoint scattering-series inversion: SEG Technical Program Expanded Abstracts. 1058-1062.

Alkhalifah, T., and Z. D. Wu, 2016, The natural combination of full and image-based waveform inversion: Geophysical Prospecting, **64**, 19-30.

Biondi, B., and A. Almomin, 2014, Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion: Geophysics, **79**, Wa129-Wa140.

Brossier, R., S. Operto, and J. Virieux, 2015, Velocity model building from seismic reflection data by full-waveform inversion: Geophysical Prospecting, **63**, 354-367.

Innanen, K. A., 2015, Full waveform inversion updating in the presence of high angle/high contrast reflectivity: SEG Technical Program Expanded Abstracts. 1314-1319.

Kwon, K., and B. Yazici, 2010, Born expansion and Frechet derivatives in nonlinear Diffuse Optical Tomography: Computers & Mathematics with Applications, **59**, 3377-3397.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Conference on Inverse Scattering, Theory and Application, Society of Industrial and Applied Mathematics, Expanded Abstracts. 206-220.

Mcgillivray, P. R., and D. W. Oldenburg, 1990, Methods for Calculating Frechet Derivatives and Sensitivities for the Nonlinear Inverse Problem - a Comparative-Study: Geophysical Prospecting, **38**, 499-524.

Metivier, L., R. Brossier, J. Virieux, and S. Operto, 2013, Full Waveform Inversion and the Truncated Newton Method: Siam Journal on Scientific Computing, **35**, B401-B437.

Pan, W. Y., K. A. Innanen, and W. Y. Liao, 2017, Accelerating Hessian-free Gauss-Newton full-waveform inversion via l-BFGS preconditioned conjugate-gradient algorithm: GEOPHYSICS, **82**, R49-R64.

Shin, C., and Y. H. Cha, 2008, Waveform inversion in the Laplace domain: Geophysical Journal International, **173**, 922-931.

Shin, C., and Y. H. Cha, 2009, Waveform inversion in the Laplace-Fourier domain: Geophysical Journal International, **177**, 1067-1079.

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

Sun, D., and W. W. Symes, 2012, Waveform inversion via non-linear differential semblance optimization: SEG Technical Program Expanded Abstracts. 1-7.

Tarantola, A., 1984, Inversion of Seismic-Reflection Data in the Acoustic Approximation: Geophysics, **49**, 1259-1266.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, Wcc1-Wcc26.

Wang, H. Y., S. C. Singh, F. Audebert, and H. Calandra, 2015, Inversion of seismic refraction and reflection data for building long-wavelength velocity models: Geophysics, **80**, R81-R93.

Weglein, A. B., F. V. Araujo, P. M. Carvalho, R. H. Stolt, K. H. Matson, R. T. Coates, D. Corrigan, D. J. Foster, S. A. Shaw, and H. Y. Zhang, 2003, Inverse scattering series and seismic exploration: Inverse Problems, **19**, R27-R83.

Wu, R. S., J. R. Luo, and B. Y. Wu, 2014, Seismic envelope inversion and modulation signal model: Geophysics, **79**, Wa13-Wa24.

Wu, R. S., and Y. C. Zheng, 2014, Non-linear partial derivative and its De Wolf approximation for non-linear seismic inversion: Geophysical Journal International, **196**, 1827-1843.

Wu, Z. D., and T. Alkhalifah, 2015, Simultaneous inversion of the background velocity and the perturbation in full-waveform inversion: Geophysics, **80**, R317-R329.

Xu, S., D. Wang, F. Chen, G. Lambaré, and Y. Zhang, 2012, Inversion on reflected seismic wave: SEG Technical Program Expanded Abstracts. 1-7.

Zhou, W., R. Brossier, S. Operto, and J. Virieux, 2015, Full waveform inversion of diving & reflected waves for velocity model building with impedance inversion based on scale separation: Geophysical Journal International, **202**, 1535-1554.