2011 CSPG CSEG CWLS Convention

# Full Waveform Inversion with Total Variation Regularization

Amsalu Y. Anagaw*
Department of Physics, University of Alberta, Edmonton, Alberta, Canada
aanagaw@ualberta.ca
and
Mauricio D. Sacchi
Department of Physics, University of Alberta, Edmonton, Alberta, Canada

## Summary

Waveform inverse problems are mathematically ill-posed and, therefore, regularization methods are required to obtain stable and unique solutions. The Total Variation (TV) regularization method is used to resolve sharp interfaces obtaining solutions where edges and discontinuities are preserved. TV regularization accomplishes these goals by imposing sparsity on the gradient of the model parameters. Full waveform inversion is carried out using the L-BFGS method in the frequency domain by selecting a limited number of frequencies from low to higher frequency. Tests with the Marmousi data set are utilized to highlight our numerical results.

## Introduction

Full waveform inversion (FWI) based on the least-squares principle is in general mathematically ill-posed and, therefore, regularization methods are required to obtain unique and stable solutions. Unlike quadratic regularization methods that tend to produce models where discontinuities are blurred or smoothed, non-quadratic regularization methods can provide high-resolution images. In this paper a Total variation (TV) regularization method (Rudin et al., 1992) is considered to impose desired features on the estimated seismic image. The TV regularization method is a good candidate if the physical quantity to estimate has sharp edge boundaries and is blocky. This regularization method tries to reconstruct continuous profiles of the model parameter by enforcing a sparseness constraint on its gradient, thereby, preserving its edges and discontinuities while suppressing artifacts due to noise.

FWI is carried out in a limited set of frequencies. The inversion of the limited set of frequencies can be carried out in a sequential approach starting from low to higher frequency data or simultaneously (Pratt et al., 1998; Sirgue and Pratt, 2004; Hu et al., 2009; Virieux and Operto, 2009). The forward problem for FWI is computed by solving the full wave-equation using the finite difference method. FWI is a local gradient-based optimization problem requiring the gradient of the objective function, which can be minimized using the adjoint-state method (Plessix, 2006). To solve the actual inversion process, the Limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) optimization method was employed. Better results and an accelerated convergence rate can be obtained when the gradient of the cost function is scaled by the diagonal of Hessian matrix, or pseudo-Hessian matrix (Shin et al, 2001a).

## Theory and/or Method

Full waveform inversion requires the minimization of the objective function defined by $l_2$ norm of the residual between the observed data $d^{obs}$ and the model data $d^{cal}$

$$J(m) = \frac{1}{2}\sum_{\omega}\sum_{s,r}^{ns,nr}\left\| d^{cal}(m,\ \omega) - d^{obs}(\omega)\right\|^2 \qquad \mathbf{1}$$

where $m$ is the slowness (inverse of velocity) model parameter we seek to retrieve, ω is the angular frequency, and $ns$ and $nr$ represent the number of sources and receivers respectively. The described inverse problem is ill-posed and requires regularization to stabilize the solution. The regularized cost function is formulated as

$$J(m) = \frac{1}{2}\sum_{\omega}\sum_{s,r}^{ns,nr}\left\| d^{cal}_{s,r}(m,\ \omega) - d^{obs}_{s,r}(\omega)\right\|^2 + \mu\frac{\left\|\vec{\nabla}[m]\right\|_1}{\left\|\vec{\nabla}[m_{n-1}]\right\|_1} \qquad \mathbf{2}$$

The first term is the $l_2$ misfit norm, which represents the error between the observations and modeled data. The second term represents the regularization term, which, in this case, is the $l_1$ norm of the gradient of slowness model. The positive parameter μ is the regularization or trade-off parameter that determines the relative balance of the two terms in expression (2). Note that $m_{n-1}$ is the model parameter of the previous iteration. The minimization of equation (2) leads to solutions where $\vec{\nabla}[m]$ is sparse. By promoting sparsity on $\vec{\nabla}[m]$, the slowness model becomes blocky, therefore, preserving edges.

However, the $l_1$ norm is non-differentiable at 0. To avoid this singularity at 0, our numerical implementation uses the following expression to approximate the $l_1$ norm of the gradient via the following differentiable functional

$$J(m) = \frac{1}{2}\sum_{\omega}\sum_{s,r}^{ns,nr}\left\| d^{cal}_{s,r}(m,\ \omega) - d^{obs}_{s,r}(\omega)\right\|^2 + \mu\frac{\left\|\sqrt{\left|\vec{\nabla}(m)\right|^2 + \alpha^2}\right\|_1}{\left\|\sqrt{\left|\vec{\nabla}(m_{n-1})\right|^2 + \alpha^2}\right\|_1} \qquad \mathbf{3}$$

The parameter α ensures the stability of the solution while also controlling its smoothness. The gradient of equation (3) with respect to $m$ is given by

$$\nabla_m J(m) = \frac{1}{2}\Re e\left(\sum_{\omega}\sum_{s,r}^{ns,nr}\left[\frac{\partial d^{cal}_{s,r}(m,\ \omega)}{\partial m}\right]^*(d^{cal}_{s,r}(m,\ \omega) - d^{obs}_{s,r}(\omega))\right) + \mu\nabla_m R(m) \qquad \mathbf{4}$$

where, $\nabla_m R(m)$ is the gradient of the total variation regularization, second term of Eqn [3].

In this paper, the full waveform inversion is carried out using the L-BFGS method. The forward problem is solved using a direct solver based on an LU decomposition of the finite-difference Helmholtz operator into a lower and upper LU triangular decomposition (PARDISO) (Schenk and Gartner, 2004; Schenk et al., 2007, 2008). This operator is quite sparse and, therefore, storable in memory. The main advantage of this method is that once the decomposition is performed and available for a given angular frequency ω and background velocity, the forward problem is efficiently solved for multiple sources using the forward and backward substitutions. The same procedure is applied for back-propagated wavefield. The gradient of the cost function is computed by zero-lag correlation between the forward-propagated wavefield and the back-propagated residual wavefield. Computationally, both wavefields are computed by solving two forward problems.

## Examples

For full waveform inversion, the Marmousi velocity model in Figure 1 [a] is used to generate the data. The subsequently smoothed velocity mode is shown in Figure 1 [b]. The latter is also the starting model of the process. To ensure convergence of the optimization, the starting model has been chosen close enough to the original model by smoothing the original velocity model with Hanning window. The inversion is then carried out by employing the Total variation regularization method and without the regularization method. The data set consist of 96 shots and 192 receivers. A small amount of noise, signal to noise ratio (SNR)=10, was added to synthetic data in the time domain data. In the frequency domain, three discrete frequencies were selected (3.5, 7.6, 12.5 Hz) for the inversion.

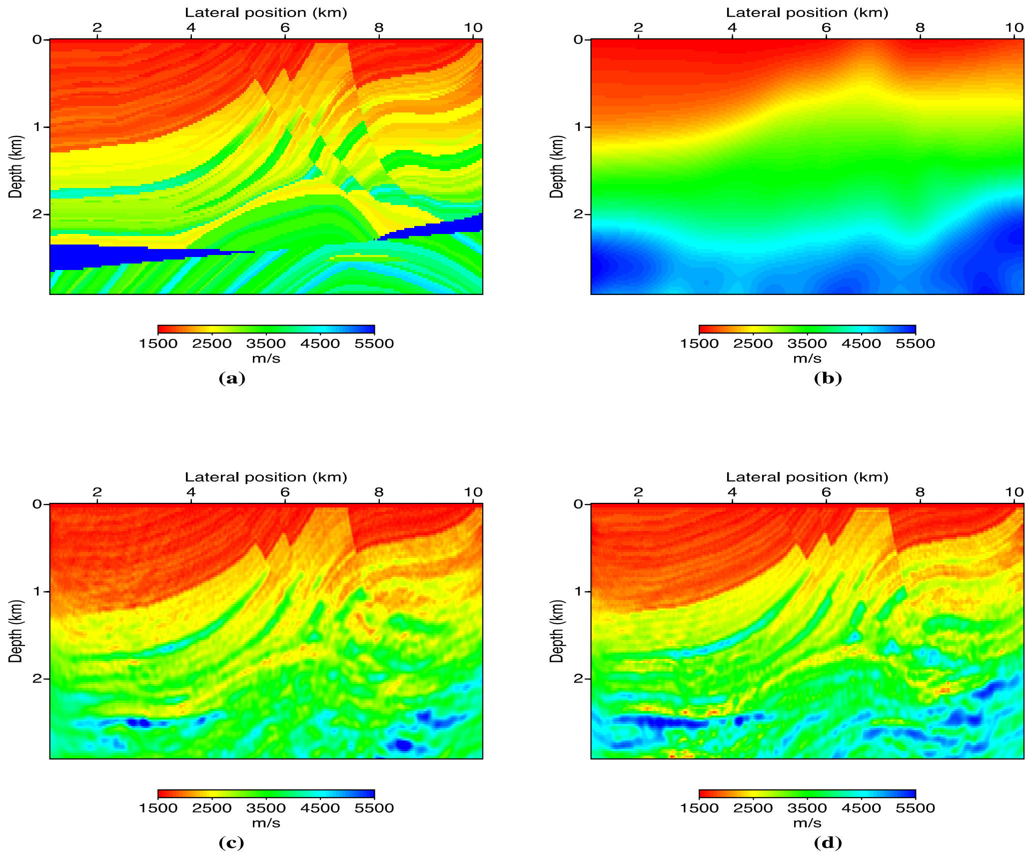


Figure 1 Full waveform inversion with and without Total Variation regularization. (a) True Marmousi velocity model. (b) Smoothed velocity model used as a starting model. FWI without and with regularization method (c) and (d) for frequency 12.5 Hz.

The full waveform inversion in frequency domain is carried out sequentially from low to high frequency. For each frequency, the FWI runs for 30 iterations. Figure 1 [c] is the reconstructed velocity model for 12.5 Hz without using regularization and Figure 1 [d] is the reconstructed velocity model for 12.5 Hz using TV regularization. These two numerical inversions show that the L-BFGS algorithm reproduces results that are comparable to the original velocity model. In both cases, the upper most part of the velocity model, the low velocity region, was reconstructed with high degree of accuracy. As expected, the lower or deeper part of the Marmousi model lacks resolution and it is improperly scaled by unregularized inversion. Edges in the deeper part of the model are, however, improved by making use of the TV regularization. Figure 2 shows the vertical velocity profiles extracted at lateral position 5.2km. This plot confirms that the edges of the velocity model are better resolved with the TV regularization.

The parameter $\alpha$ introduced in the regularization function controls the behavior of the nonlinearity of the model, thereby controlling the sharp edges, discontinuities and the smoothness of the model. One of the difficulties in the numerical simulation is finding the best combination of $\mu$ and $\alpha$. There is no successful heuristic way to determine these values. After several trails these values were adjusted to vary through the iterations in such a way that an appropriate weights are given for each frequencies and iteration. Finding optimal trade-off parameters is a major effort of our research.

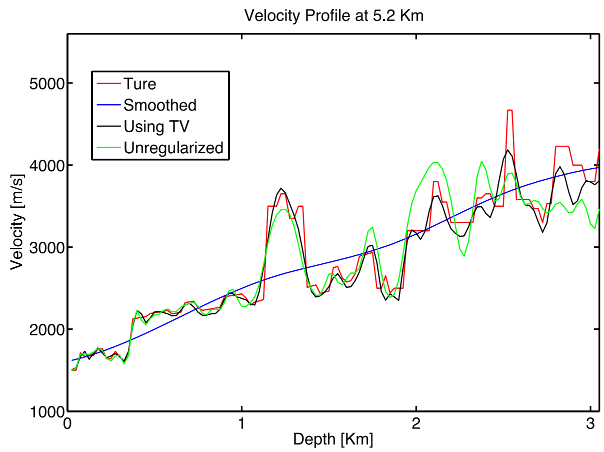


Figure 2. Vertical velocity profile extracted at 5.2km. Red is the true velocity profile, blue is the smooth velocity profile used as a staring model, green is the recovered velocity profile using the unregularized inversion and black is the reconstructed velocity profile using TV.

## Conclusions

We presented a non-linear sequential frequency domain 2D full waveform inversion method in a constant density medium that uses Total variation regularization to constrain seismic velocities. The proposed method uses TV in conjunction with the L-BFGS optimization method to retrieve a model of the subsurface with preserved edges and discontinuities. Numerical results with synthetic data, using only a few frequencies, appear to confirm that the regularization via TV produces images of high resolution, particularly in the deeper part of the model. This is because the regularization method was chosen to have the ability to resolve sharp edges and interfaces of the medium. The method expands the work proposed in the areas of imaging using TV regularization via single-scattering Born inversion to retrieve a model of the subsurface with preserved edges and discontinuities (Anagaw et al., 2010).

## Acknowledgements

We thank the authors of PARDISO LU solver (http://www.pardiso-project.org/) for making their package available to us. Thanks to the sponsors of SAIG for funding this research and David Bonar for discussing this work with us.

**References**

Anagaw, A. Y. and M. D. Sacchi, 2010, Edge-preserving seismic imaging using the total variation method. *GeoCanada* 2010, 440

Hu, W., A. Abubakar, and T. M. Habashy, 2009, Simultaneous multifrequency inversion of full-waveform seismic data. *Geophysics* 74 R1–14.

Nocedal, J [1980] Updating Quasi-Newton Matrices with Limited Storage . *Mathematics of Computation* 35, 773-782.

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 (2), 495–503.

Pratt, R., C. Shin, and G. Hicks, 1998, Gauss-Newton and full Newton methods in frequency space seismic waveform inversion. *Geophysical Journal International* 133 (2), 341–362.

Rudin, L., S. Osher, and E. Fatemi, 1992, Nonlinear total variation based noise removal algorithms. *Physica D* 259–268.

Schenk, O. and K. Gartner, 2004, Solving unsymmetric sparse systems of linear equations with PARDISO. *Future Generation Computer Systems* 20 (3), 475–487. *International Conference on Computational Science* (ICCS 2002), Amsterdam, Netherlands, APR 21-24, 2002.

Shin, C., Jang, S. H., and Min, D. J, 2001a, Improved amplitude preservation for prestack depth migration by inverse scattering theory. *Geophysical Prospecting*, 49, 592–606.

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

Virieux, J. and S. Operto,2009, An overview of full-waveform inversion in exploration geophysics. *Geophysics* 74 (6), WCC1–WCC26.