F027

Building Blocky Models With Full Waveform Inversion

A. Guitton* (GeoImaging Solutions)

SUMMARY

Deriving blocky models from Full Waveform Inversion (FWI) might be required in some geological settings. These models might be also needed for imaging with wave-equation based techniques such as Reverse Time Migration (RTM).
Enforcing blockiness in the model space amounts to enforcing sparseness in the derivative of the model. Sparseness can be obtained by using long tail distributions such as the l1 norm or Cauchy function. By using different directions for the derivatives, the model can be made blocky in more than one direction: horizontally, vertically, or both. Whereas the Cauchy function produces better blocky models, the l1 norm has the advantage of being convex and easy to parameterize.

Introduction

Full Waveform Inversion (FWI) delivers high resolution images of physical properties such as velocity and impedance. One of the main challenge of FWI is the presence of local minima which, when gradient-based techniques are used, forces us to use multi-scale approaches to mitigate their effects (Bunks et al., 1995; Sirgue and Pratt, 2004). Having to run the inversion in a multi-scale fashion increases the cost of FWI, since many scales are needed to converge towards an acceptable solution. Because of its cost, FWI is usually carried out at lower frequencies than a standard migration: if a Reverse Time Migration (RTM) can be ran at 40Hz, a FWI will be limited to frequencies below 20Hz.

If a frequency-based FWI is chosen, schemes to limit the number of frequency slices exist to save computation time. Running FWI at lower frequencies has an impact on the level of details that can be seen in our images. It can also introduce some artifacts , related to the Gibbs phenomenon, into the final results in the form of low-frequency ripples. In some geological settings and for certain applications, having sharp boundaries might be required. If velocity fields derived from FWI are used for RTM, for instance, and strong velocity contrasts exist (such as between salt and sediments), then sharp boundaries are needed. In this paper, we propose adding regularization terms to the objective function of FWI to enforce blockiness in the final velocity model. We show how this model can be made blocky in any direction, vertically, horizontally, or both. We also show how different functions in the regularization terms (ℓ^1 norm and Cauchy) affect the level of details in the image. In the next section, we introduce our expended objective function that includes the blocky regularization terms. Then we illustrate our approach with a 2-D example inspired by a typical Gulf of Mexico setting where salt bodies are present.

Theory

In this section, we first present our FWI objective function and gradient without regularization and introduce the different terms that enforce blockiness in the final model. We use a time domain approach for solving the scalar acoustic wave equation (parameterized in terms of P-wave velocity v_p):

$$\frac{\partial^2 u(\mathbf{x},t)}{\partial t^2} - v_p(\mathbf{x})^2 \nabla^2 u(\mathbf{x},t) = v_p(\mathbf{x})^2 s(\mathbf{x},t). \tag{1}$$

with the source term $s(\mathbf{x},t) = \delta(\mathbf{x}-\mathbf{x}_s)S(t)$ where $S(t)$ is the source function at $\mathbf{x}_s$ and u the pressure field. The goal of waveform inversion is to derive physical properties of the Earth, such as P-wave velocity, S-wave velocity, or density. These properties can be related to the presence of hydrocarbons in the subsurface and their estimation is one of the most important goal in seismic processing. In practice, we try to minimize the function

$$f(\mathbf{m}) = \|\mathbf{u}^{obs} - \mathbf{u}^{mod}\|^2 \tag{2}$$

where $\mathbf{m}$ is a vector of physical properties (what we are looking for), $\mathbf{u}^{obs}$ the observed and $\mathbf{u}^{mod}$ the modeled data. Tarantola (1984) gives the definition of the gradient for the acoustic Equation (2) for each component of $\mathbf{m}$ (equal to v_p only in this case).

$$\nabla f(\mathbf{m}_n) = \frac{-2}{\mathbf{v_p}_n^3} \sum_{shots} \sum_{t} \frac{\partial^2 \overrightarrow{\mathbf{u}_n}}{\partial t^2} \cdot \overleftarrow{\delta \mathbf{u}_n} \tag{3}$$

where $\overleftarrow{\delta \mathbf{u}_n}$ is the *backward* propagated residual at iteration n such that $\delta \mathbf{u_n} = \mathbf{u}^{obs} - \mathbf{u}_n^{mod}$ and $\overrightarrow{\mathbf{u}_n}$ is the *forward* propagated synthetic source.

Now we introduce regularization terms to enforce blockiness in the model. Enforcing blockiness in the model can be translated into the need to have the derivative of the model to be sparse. This derivative can be in the horizontal or vertical direction (Almonin, 2010). A regularization term with these two properties, enforcing sparseness in the derivative of the model space, is as follows:

$$g(\mathbf{m}) = \sum_{i=1}^{2} \alpha_i |\mathbf{D}_i \mathbf{m}|^s \tag{4}$$

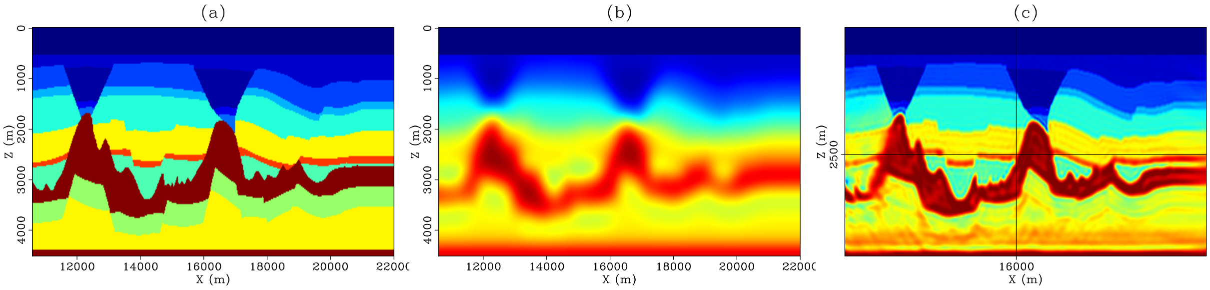

Figure 1 *(a) True model. (b) Starting model for FWI. (c) Recovered model without regularization.*

where $\mathbf{D}_1$ is the operator representation of the vertical derivative $\partial/\partial z$, $\mathbf{D}_2$ is the operator representation of the horizontal derivative $\partial/\partial x$, $|.|^s$ a function enforcing sparseness in the derivatives, and α_i a constant which can take the value 0 or 1, depending on our choice of direction for blockiness. For the functions enforcing sparseness, we are using both the ℓ^1 norm and Cauchy function (Sacchi and Ulrych, 1995):

$$|\mathbf{x}|^1 = \sum_{j=1}^{n} |x_i| \text{ and } |\mathbf{x}|^c = \frac{1}{2}\sum_{j=1}^{n} \log\left(1+\left(\frac{x_i}{\gamma}\right)^2\right) \quad (5)$$

where γ is a parameter to be selected. Adding the regularization terms to the objective function in Equation (2) we obtain:

$$f_{reg}(\mathbf{m}) = \|\mathbf{u}^{obs} - \mathbf{u}^{mod}\|^2 + \varepsilon \sum_{i=1}^{2} \alpha_i |\mathbf{D}_i \mathbf{m}|^s \quad (6)$$

where ε is a constant to be chosen a-priori. Now, for the computation of the gradient with the new regularization term, we have

$$\nabla f_{reg}(\mathbf{m}_n) = \nabla f(\mathbf{m}_n) + \varepsilon \sum_{i=1}^{2} \alpha_i \mathbf{D}_i^* \mathbf{r}_i^s \quad (7)$$

where $\mathbf{D}^*$ is the adjoint of $\mathbf{D}$, $\mathbf{r}_i = \mathbf{D}_i \mathbf{m}_n$ and each element of the residual vector $\mathbf{r}_i$ is given by

$$r^1 = \begin{cases} 1 & if \quad r > 0 \\ 0 & if \quad r = 0 \\ -1 & if \quad r < 0 \end{cases} \text{ and } r^c = \frac{r}{\gamma^2 + r^2} \quad (8)$$

for the ℓ^1 norm and Cauchy function, respectively. The addition of the regularization terms to the objective function and gradient has a negligible effect on the overall cost of FWI since the regularization operates on the model space only. For the non-linear inversion, we use the L-BFGS method of Nocedal (1980).

Examples

Our example is derived from a velocity model used by Etgen and Regone (1998) (Figure 1a). This model is blocky by design and fits our intended model after inversion very well. We modeled 38 shots, 300 m. apart, with a fix spread of receiver recording at every surface grid point. We use five frequency bands for FWI: 0-4, 0-8, 0-10, 0-14, 0-18Hz with a maximum of 30 function evaluations for each scale (the number of iterations will vary).

The starting model for FWI is shown in Figure 1b and is obtained by smoothing the true model. If no regularization is used, we obtain the model in Figure 1c: FWI is successful at recovering the main layers, but we can notice some velocity variations within the constant velocity layers due to the our frequency limitation. Now, we show the results of blocky regularization with the ℓ^1 norm when blockiness is enforced in the x direction (Figure 2a), z direction (Figure 2b), and both x and z directions (Figure 2c). The black horizontal and vertical lines indicate the locations of the slices extracted from these models

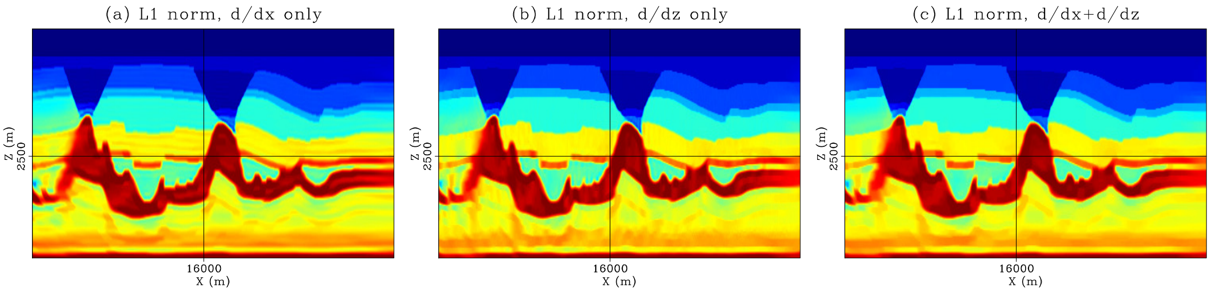

Figure 2 *Blocky inversion using ℓ^1 norm for, (a) $\partial/\partial x$ only, (b) $\partial/\partial z$ only and (c), $\partial/\partial x + \partial/\partial z$.*

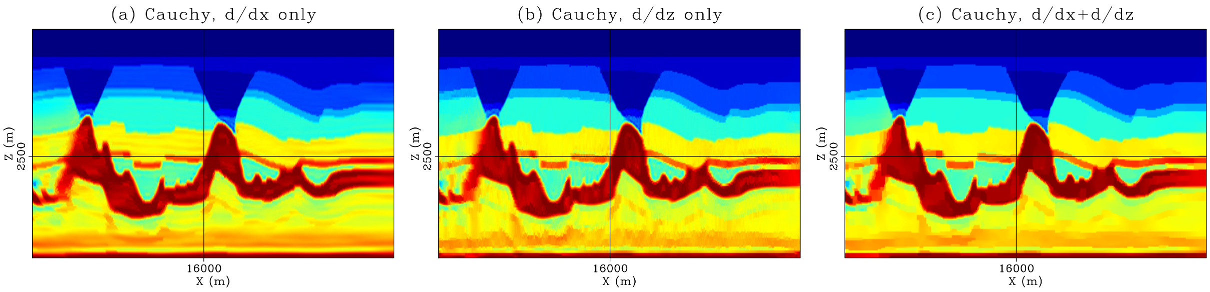

Figure 3 *Blocky inversion using Cauchy function for, (a) $\partial/\partial x$ only, (b) $\partial/\partial z$ only and (c), $\partial/\partial x + \partial/\partial z$.*

and displayed in Figures 4 and 5. As expected, the ℓ^1 norm regularization is able to constrain the model in any chosen direction: the velocity variations within constant velocity layers seen in Figure 1c are attenuated in Figures 2b and 2c, thanks to the regularization in the z direction. These improvements are particularly visible in Figure 5 (top row, two right panels). The effects of the regularization in the x direction are mostly visible in Figure 4 where the horizontal blockiness is improved (red curve, top panel in Figure 4).

If the Cauchy function is used instead of the ℓ^1 norm, we obtain Figure 3. Similar to what we observed with the ℓ^1 norm, blockiness is improved in whatever direction the regularization is applied. Now, comparing the two regularization schemes (Cauchy or ℓ^1), we can notice that the Cauchy function provides blockier models if the two directions are regularized: this is clearly visible by comparing the top and bottom rows of Figures 4 and 5. If blockiness in z is sought after, then the ℓ^1 norm is preferred since reflectors look more continuous in Figure 2b than in Figure 3b.

From the computational side, two elements play in favor of the ℓ^1 norm regularization. First, it is a convex function, whereas Cauchy is not (second derivative is not positive everywhere). Therefore, from a minimization point of view where gradient based methods are used, the ℓ^1 norm is more indicated. Second, the Cauchy function requires an extra parameter evaluation (γ in Equation 5) that can be tricky to estimate. In our case, we chose it by trial and error. One may argue that the ℓ^1 norm can't be minimized with the L-BFGS method either, since it is not twice continuously differentiable. However, considering that we never find the global minimum and that the estimated Hessian in L-BFGS is an approximate one, then using the ℓ^1 norm seems valid as exemplified in our results. To alleviate some of these approximations, the Huber (Guitton and Symes, 2003) or hybrid norm could be used instead.

Conclusion

Having blocky models might be a requirement of FWI in geological settings where strong velocity contrasts are present. We present regularization schemes that enforce blockiness by making the derivative of the velocity field sparse in one or two directions. Sparseness is achieved by using the ℓ^1 norm or Cauchy function. Our results indicate that both measures yield blocky models, but that the Cauchy function

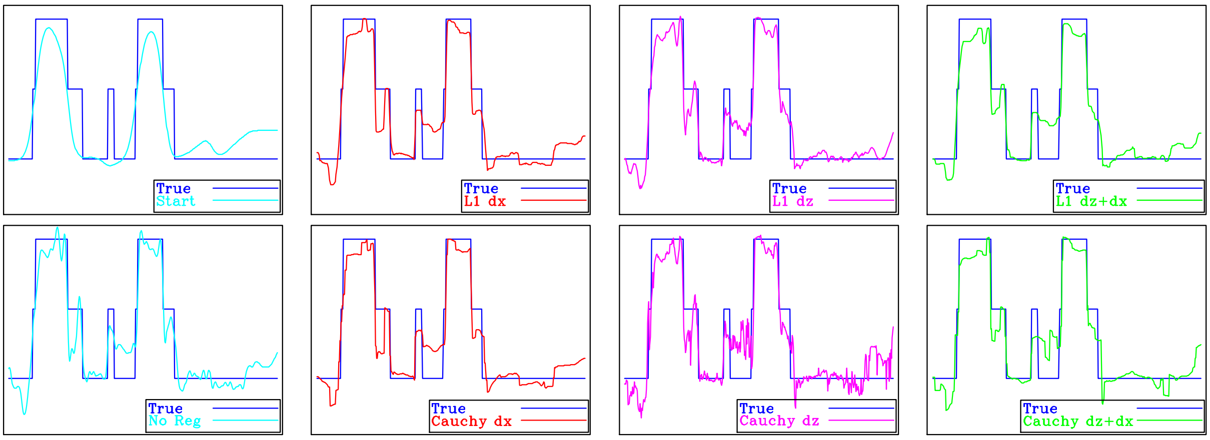

Figure 4 *Constant depth slices (Z=2500m) for the true model (dark blue) and FWI results without regularization (bottom-left), with ℓ^1 norm regularization (3 top right panels) and with Cauchy function regularization (3 bottom right panels). Red is for $\partial/\partial x$ only, magenta for $\partial/\partial z$ only, and green for $\partial/\partial x + \partial/\partial z$.*

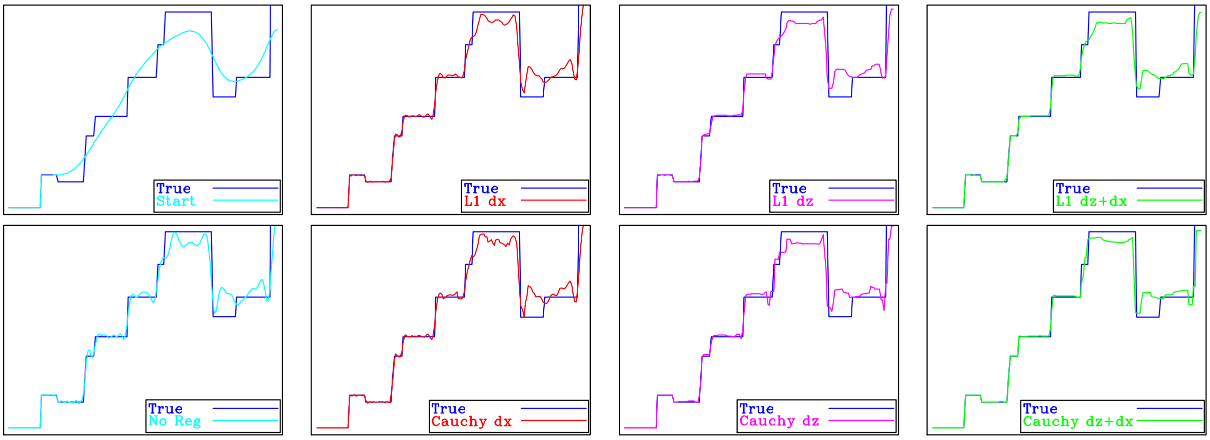

Figure 5 *Constant vertical slices (X=16000m) for the true model (dark blue) and FWI results without regularization (bottom-left), with ℓ^1 norm regularization (3 top right panels) and with Cauchy function regularization (3 bottom right panels). Red is for $\partial/\partial x$ only, magenta for $\partial/\partial z$ only, and green for $\partial/\partial x + \partial/\partial z$.*

seems to provide higher resolution models. However, from a minimization point of view, the ℓ^1 norm might be more appropriate because it is convex and easier to parameterize.

References

Almonin, A. [2010] Blocky velocity inversion by hybrid norm. *SEP Annual Report*, 143–154.

Bunks, C., Saleck, F., Zaleski, S. and Chavent, G. [1995] Multiscale seismic waveform inversion. *Geophysics*, **60**(5), 1457–1473.

Etgen, J. and Regone, C. [1998] Strike shooting, dip shooting, widepatch shooting - Does prestack depth migration care? A model study. *SEG Technical Program Expanded Abstracts*, 66–69.

Guitton, A. and Symes, W.W. [2003] Robust inversion of seismic data using the Huber norm. *Geophysics*, **68**(4), 1310–1319.

Nocedal, J. [1980] Updating Quasi-Newton matrices with limited storage. *Mathematics of Computation*, **95**, 339–353.

Sacchi, M.D. and Ulrych, T.J. [1995] High-resolution velocity gathers and offset space reconstruction. *Geophysics*, **60**(4), 1169–1177.

Sirgue, L. and Pratt, R. [2004] Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies. *Geophysics*, **69**(1), 231–248.

Tarantola, A. [1984] Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, **49**(8), 1259–1266.