Tu BC 01

Multi-Parameter FWI - An Illustration of the Hessian Operator Role for Mitigating Trade-off between Parameter Classes

L. Métivier* (LJK), R. Brossier (ISTerre), S. Operto (Geoazur) & J. Virieux (ISTerre)

SUMMARY

Full Waveform Inversion is a powerful tool for quantitative estimation of subsurface parameters (P-wave, S-wave velocities, density, attenuation, anisotropy parameters). This methods has been applied successfully to 2D acoustic and elastic reconstructions, as well as to 3D acoustic reconstructions. Most of the applications of Full Waveform Inversion have been devoted to mono-parameter reconstructions of wave velocities. Multi-parameter Full Waveform Inversion aims at reconstructing simultaneously different class of subsurface parameters. This is a very challenging task: the similarity of the sensitivity of the data to different classes of parameters is the source of trade-off (or cross-talk) which renders the Full Waveform Inversion problem even more undetermined than in the mono-parameter context. This can related to the similarity of the diffraction patterns of different classes parameters for a given propagation regime. In order to overcome this difficulty, the role of the Hessian operator should be crucial. The off-diagonal blocks of this operator accounts for the trade-off between parameters. Incorporating the inverse Hessian operator within the Full Waveform Inversion scheme may help to alleviate this difficulty. On this basis, we provide in this study a very simple example for which we can compute exactly the Hessian operator we use to illustrate these issues.

Introduction

Full Waveform Inversion (FWI) is a powerful tool for quantitative estimation of subsurface parameters such as P-wave or S-wave velocities, density, attenuation, or anisotropy parameters. Thanks to the simultaneous development of high performance computing capabilities and wide-azimuth multi-components acquisition systems, the method has been successfully applied in 2D configurations, assuming acoustic or elastic wave propagation, as well as in 3D configurations, under an acoustic framework. See Virieux and Operto (2009) for an overview of the method and its applications.

The FWI method is based on the minimization of the misfit between recorded and simulated data. The simulated data is computed through the resolution of a Partial Differential Equation (PDE) which describes the wave propagation. This formalism is adapted to the multiple reconstruction of several classes of parameters. Such a multi-parameter inversion is of major interest. Besides the intrinsic interest of recovering different types of information on the subsurface, being able to reconstruct simultaneously the P-wave velocity and secondary attributes such as density and attenuation is a key issue to recover micro-scale parameters, for instance porosity or saturation (De Barros et al., 2012).

Multi-parameter FWI remains however a very challenging task. For mono-parameter estimation, the incompleteness of illumination of the subsurface and the inherent noise which contaminates the seismic data already renders the inverse problem strongly undetermined. Moving from mono-parameter to multi-parameter inversion increases the indetermination. The similarity of the diffraction patterns of different classes of parameters for certain illumination angles is partially responsible for trade-off (or cross-talk) between parameters. The difference in amplitude of the footprint of the parameter classes on the seismic data (primary and secondary seismic attributes) is an additional difficulty. A review of these issues is proposed in Operto et al. (2013), where it is shown that the choice of the parameterization is crucial, as it impacts directly the diffraction patterns of the different parameters. Hierarchical approaches designed in terms of illumination angles (from large offsets to narrow offsets) and a parameterization adapted to the considered propagation regime could help to alleviate these difficulties.

Operto et al. (2013) also mention the important role of the inverse Hessian operator. In a mono-parameter context, this operator helps to refocus the information on less illuminated subsurface parameters (Pratt et al., 1998). This has justified the progressive drift from gradient-based optimization methods to quasi-Newton methods (l-BFGS) (Brossier et al., 2009) and truncated Newton methods (Métivier et al., 2013a,b) for the solution of the FWI problem with extra computations. In a multi-parameter context, the off-diagonal blocks of the Hessian operator reflects the coupling between parameter classes. Therefore, incorporating its inverse within the inversion scheme should help to mitigate the trade-off between parameters during the reconstruction, aside partial illumination issues.

We illustrate on a simple numerical example the role of the inverse Hessian operator in a 2D frequency-domain acoustic framework. The experiment is designed as the reconstruction of a Gaussian P-wave velocity (v_p) perturbation from homogeneous v_p and density (ρ) models. In a (v_p, ρ) parameterization, v_p has an isotropic radiation pattern, while the radiation pattern of ρ decreases rapidly for increasing illumination angles (Fig. 1). The similarity of the radiation patterns of these two parameters in the reflection regime explains partially the trade-off. In our experiment, we thus investigate two different configurations. The first one is designed with a perfect acquisition system (receivers and sources located all around the domain). The second one is designed with a surface acquisition system. In the first case, the seismic data contain both transmitted and reflected waves. The presence of transmitted waves reduces the trade-off between v_p and ρ as the transmitted waves are mainly sensitive to v_p. In the second case, the absence of the transmitted waves strengthens the trade-off between v_p and ρ. In both cases, we show how the inverse Hessian reduce this trade-off.

Methodology

We consider a 2D acoustic frequency-domain framework. We give a formulation for a single frequency ω (the extension to the multi-frequency context is straightforward by summation). Let m gathers the two subsurface parameters: $m = (v_p, \rho)$. The FWI problem is defined as

$$\min_m f(m) = \frac{1}{2} \sum_{s=1}^{S} \|Ru_s(m) - d_s\|, \tag{1}$$

where S is the number of data-sets, $u_s(m)$ and d_s are respectively the simulated wavefield and the data-set associated with the sth source, and R is a restriction operator which maps the wavefields $u_s(m)$ to the receiver locations. We define $u_s(m)$ as the solution of the forward problem

$$-\omega^2 u - \rho c \ div \left(\frac{1}{\rho} \ grad \ u \right) = \varphi_s, \tag{2}$$

where φ_s denotes the source s. Newton methods for the resolution of the inverse problem (1) are based on a local quadratic approximation of $f(m)$ such that

$$q(\Delta m) = f(m) + \nabla f(m)^T \Delta m + \Delta m^T H(m) \Delta m, \tag{3}$$

where $\nabla f(m)$ and $H(m)$ are respectively the gradient and the Hessian operator of the misfit function $f(m)$. In this case, the solution of this quadratic optimization is simply the solution of the linear system

$$H(m)\Delta_m = -\nabla f(m), \tag{4}$$

where we see the inverse Hessian operator being applied to the gradient. In practice, the exact computation of $H(m)^{-1}$ is intractable, and a rough approximations is used (quasi-Newton method, see for instance the l-BFGS approximation proposed by Byrd et al. (1995)), or an incomplete resolution of the linear system is performed (truncated Newton method, see Métivier et al. (2013a,b)). In this study, we want to handle a more precise approximation of the inverse Hessian operator. To this purpose, the size of the two test cases is defined so that $H(m)$ can be computed exactly. This computation is performed row by row, using second-order adjoint formulas (Métivier et al., 2013b). We focus on the Gauss-Newton part $H_{GN}(m)$ of the Hessian operator and neglect second-order terms. We perform a SVD decomposition of this operator, and we solve the linear system (4) through a truncated SVD method.

Numerical experiments

We consider a 600×600 m domain discretized over a 31×31 grid. A mixed-staggered grid approach is implemented for the discretization of the forward problem (2) (Hustedt et al., 2004). Perfectly Matched Layers (Berenger (1994)) are added on the four boundaries of the domain to simulate an infinite extension. The initial velocity model $v_{p,0}$ is homogeneous, as well as the initial density model ρ_0. The $v_{p,0}$ value is set to 1500 m.s^{-1}. The ρ_0 value is to 1431 kg.m^{-3} following the Gardner Law. The exact velocity model v_p contains a Gaussian perturbation of maximum amplitude 100 m.s^{-1}, located at the center of the domain (Fig. 1). The exact density model ρ is identical to the initial density model ρ_0. We perform a broadband inversion: 17 data-sets from 4 Hz to 12 Hz, each 0.5 Hz, are inverted simultaneously. We consider two different acquisition systems. The first is designed with 112 sources and 112 receivers are located on the four sides of the domain (perfect acquisition). For this configuration, both reflected and transmitted waves are recorded. The second one is a surface acquisition system: only 28 sources and 28 receivers are used, located on top, at the depth $z = 20$ m. For this data-set, the reflected waves are dominant. The corresponding Gauss-Newton operators are presented in Figure 2. The operators are composed of 4 main blocks. The top left block (respectively the bottom right block) is associated with second-order derivatives of the misfit function with respect to v_p (respectively with respect to ρ). The off-diagonal blocks correspond to cross-derivatives between v_p and ρ, and express the trade-off

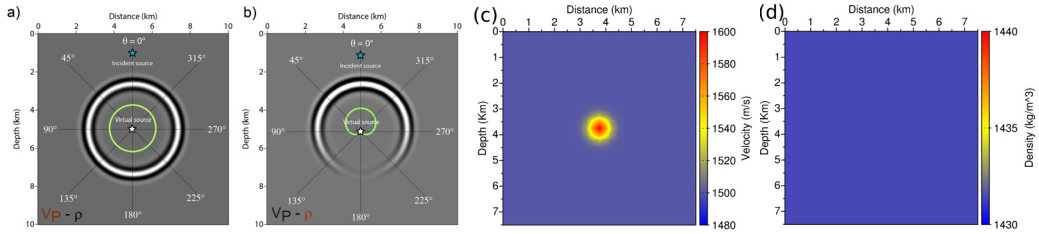

Figure 1 *Radiation pattern for $v_p(a)$ and $\rho(b)$ for a (v_p, ρ) parameterization. Homogeneous v_p model with a Gaussian anomaly (c), constant density model (d).*

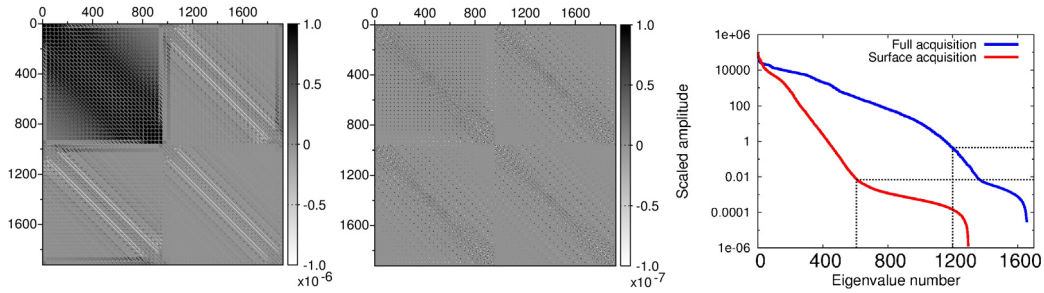

Figure 2 *Hessian operator for perfect acquisition (a), surface acquisition (b). Eigenvalues associated with the Hessian operators for perfect and surface acquisition, with and without damping (c).*

between these two parameters. In the perfect acquisition configuration, we see that the sensitivity to v_p is dominant. This is mainly related to the fact that the density has a very limited effect on the transmitted waves. This part of the signal constrains mainly v_p. For the surface acquisition configuration, the average amplitude of the two diagonal blocks is similar. In this case, reflected waves are dominant in the data, and the sensitivity of the seismic data to these two parameters is comparable. We therefore expect the trade-off to be stronger for the surface acquisition than for the perfect acquisition system.

The eigenvalues of the operators are presented in Figure 2 in logarithmic scale. We detect the presence of small negative eigenvalues (probably due to round-off errors) which explains the shape of the spectrum tending toward $-\infty$ because of the log scale. The dot lines added on the spectrum in Figure 2 represent the level at which the truncation is performed for the truncated SVD resolution of the linear system (4). This level is determined through a trial and error approach. Below a certain level, the inversion becomes unstable. Because the Gauss-Newton operator associated with the surface acquisition configuration has much more eigenvalues close to 0 the level of truncation is more severe in this case: only 600 eigenvectors are taken into account, while 1300 can be used in the perfect acquisition case.

We compare the model update corresponding to the gradient, and the one obtained after truncated SVD. Because we are mainly interested in the trade-off between parameters, we focus on the relative amplitude of v_p and ρ updates. In each case, we divide these updates by the maximum value of the v_p model update. The results are presented in Figure 3. The narrow aperture illumination provided by the surface acquisition leads to a significant blurring of the v_p update caused by a deficit of small wave-number content. Moreover, the increased trade-off between v_p and ρ in the case of the surface acquisition generates as expected a more significant leakage of the v_p perturbation in the ρ update. In both cases, the inverse Hessian operator refocuses the computed v_p perturbation and reduces efficiently the trade-off.

Conclusions

Multi-parameter FWI is a challenging task. The possibility of reconstructing different classes of parameter from seismic data strongly depends on the propagation regime related with the acquisition geometry, and on the parameterization of the subsurface which is used. The trade-off resulting from the

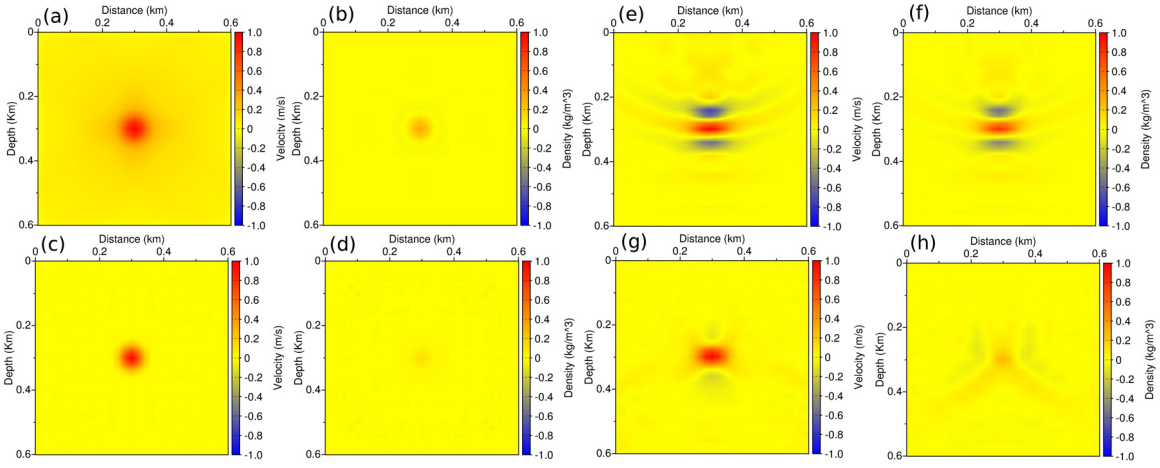

Figure 3 *Perfect acquisition case: model updates corresponding to the gradient for v_p (a), ρ (b); model updates after truncated SVD for v_p (c), ρ (d). Surface acquisition case: model updates corresponding to the gradient for v_p (e), ρ (f); model updates after truncated SVD for v_p (g), ρ (h)*

overlapping of the radiation patterns for specific propagation regimes can however be limited through a proper approximation of the inverse Hessian operator. In this study, we show for a simple case in a 2D frequency-domain framework, how the acquisition geometry plays a role in the expression of these trade-off. We show how the inverse Hessian operator can help to reduce these trade-off by an appropriate scaling of the model updates.

Acknowledgments

This study was funded by the SEISCOPE II consortium (*http://seiscope2.osug.fr*), sponsored by BP, CGG, CHEVRON, EXXON-MOBIL, JGI, PETROBRAS, SAUDI ARAMCO, SHELL, STATOIL, TOTAL and WESTERNGECO. This study was granted access to the HPC facilities of CIMENT (Université Joseph Fourier Grenoble), and of GENCI-CINES under Grant 2011-046091 of GENCI (Grand Equipement National de Calcul Intensif).

References

Berenger, J.P. [1994] A perfectly matched layer for absorption of electromagnetic waves. *Journal of Computational Physics*, **114**, 185–200.

Brossier, R., Operto, S. and Virieux, J. [2009] Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics*, **74**(6), WCC105–WCC118, doi:10.1190/1.3215771.

Byrd, R., Lu, P. and Nocedal, J. [1995] A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific and Statistical Computing*, **16**, 1190–1208.

De Barros, L., Dupuy, B., O'Brien, G., Virieux, J. and Garambois, S. [2012] *Using a poroelastic theory to reconstruct subsurface properties: numerical investigation*, Seismic Waves, Research and Analysis, Intech, Masaki Kanao (Ed.). ISBN 978-953-307-944-8, Intech, doi:DOI: 10.5772/28106.

Hustedt, B., Operto, S. and Virieux, J. [2004] Mixed-grid and staggered-grid finite difference methods for frequency domain acoustic wave modelling. *Geophysical Journal International*, **157**, 1269–1296.

Métivier, L., Bretaudeau, F., Brossier, R., Operto, S. and Virieux, J. [2013a] Full waveform inversion and the truncated newton method: quantitative imaging of complex subsurface structures. *Geophysical Prospecting*, **submitted**.

Métivier, L., Brossier, R., Virieux, J. and Operto, S. [2013b] Full waveform inversion and the truncated newton method. *SIAM Journal On Scientific Computing*, **35(2)**, B401–B437.

Operto, S. et al. [2013] A guided tour of multiparameter full waveform inversion for multicomponent data: from theory to practice. *The Leading Edge*, **September, Special section Full Waveform Inversion**, 1040–1054.

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

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