Algorithmic and practical aspects of TV regularization for joint time-lapse FWI

Théo Rolin, Didier Auroux, Laure Combe, S. Operto

2025 EAGE ANNUAL 86TH CONFERENCE & EXHIBITION TOULOUSE | FRANCE

Algorithmic and practical aspects of TV regularization for joint time-lapse FWI

T. Rolin[1], D. Auroux[2], L. Combe[1], S. Operto[1]

[1] University Côte d'Azur (UCA) - CNRS - Geoazur; [2] University Côte d'Azur (UCA) - LJAD

Summary

Joint time-lapse FWI manages non repeatability artifacts induced by compaction and acquisition changes by jointly inverting the baseline and monitor data for the simultaneous update of the baseline and monitor models. In this framework, model-difference regularization links together the baseline and monitor subproblems. The edge-preserving total-variation (TV) regularization should be suitable to reconstruct localized model changes in the monitor model. TV regularization minimizes the 1-norm of the gradient of the subsurface model and hence is non differentiable. The alternating-direction method of multipliers (ADMM) and proximal methods provide a suitable framework to implement TV regularization in the 1-norm regularized 2-norm time-lapse FWI. The principle of ADMM is to separate the 2-norm and the 1-norm subproblems in the objective with alternating directions and recast the 1-norm subproblem as a denoising problem. Within one ADMM iteration, the denoising reduces to an element-wise soft thresholding while model parameters are updated with one or several inner iterations of 2-norm regularized FWI. ADMM however requires the careful tuning of the penalty and thresholding parameters, and raises the issue of the convergence rate. This study reviews the algorithm of ADMM-based TV-regularized joint time-lapse FWI and illustrates the tuning of the hyper parameters with two synthetic examples.

Algorithmic and practical aspects of TV regularization for joint time-lapse FWI

Introduction

Joint time-lapse (TL) FWI has been proposed to mitigate non repeatability artifacts induced by production-induced compaction and/or acquisition changes (Maharramov et al., 2016). Joint TL-FWI jointly inverts the baseline and monitor data for the simultaneous update of the baseline and monitor models. In this framework, a key ingredient is the regularization term implementing prior assumption on the model changes since it links together the baseline and the monitor FWI subproblems. Among the possible choice of regularization, the edge-preserving total-variation (TV) regularization promotes blockiness (i.e., piecewise constant) velocity perturbations (e.g., Aghamiry et al., 2019). Since the model changes are expected to be localized and generated by a given geological process (e.g., fluid migration), the blockiness assumption seems reasonable to represent time-lapse effects as opposed to Tikhonov regularization whose smoothing effects would make the detection of mild perturbations challenging. TV regularization involves the minimization of the non-differentiable $\ell 1$ norm of the velocity model gradient. A classical approach to bypass non differentiability is to smooth the TV operator near the zero gradient (Maharramov et al., 2016, their equation 13). Alternatively, we implement TV regularization in (TL-)FWI with the alternating-direction method of multipliers (ADMM) (Boyd et al., 2010) and proximal algorithms (Combettes and Pesquet, 2011; Aghamiry et al., 2021). The governing idea is to separate the least-squares ($\ell 2$) subproblem and the non-smooth ($\ell 1$) subproblem in $\ell 1$-regularized TL-FWI by introducing auxiliary variables such that the $\ell 1$ subproblem can be recast as a denoising problem whose closed-form solution is computed efficiently with proximal algorithms. Then, the different classes of variables (subsurface parameters, auxiliary variables, and Lagrange multipliers) are updated in an alternated mode to decompose the global problem into a series of easy-to-solve subproblems. The objective of this study is to provide a first assessment of ADMM-based TV-regularization in a TL-FWI context. More precisely, ADMM involves two main hyper parameters, a penalty parameter and a thresholding parameter. We illustrate heuristically how to tune these hyper parameters with 2D synthetic examples. We implement FWI in the frequency domain for computational efficiency when considering 3D targets of moderate size, flexibility provided by compact volume of data, and easy implementation of attenuation (Operto and Miniussi, 2018). We perform the synthetic experiments with a basic parallel approach where TV regularization and FWI are applied to the baseline and monitor subproblems separately. However, implementation and tuning of TV regularization in the parallel approach and in the joint approach rely on the same recipes, the key difference being the coupling between the baseline and the monitor subproblems in the joint approach induced by the TV regularization applied on the model difference. We first review the algorithm of joint ADMM-based TV-regularized TL-FWI. Then, we present the results of the two synthetic examples before concluding with perspectives.

Method

The functional of joint time-lapse FWI with TV regularization can be formulated as

$$\min_{\mathbf{m}} \mathscr{F}(\mathbf{m}) + \lambda \left(\|\nabla_x \Delta\mathbf{m}\|_1 + \|\nabla_z \Delta\mathbf{m}\|_1 \right). \tag{1}$$

$\mathscr{F}(\mathbf{m}) = \mathscr{F}^b(\mathbf{m}_b) + \mathscr{F}^m(\mathbf{m}_m) = \frac{1}{2}\sum_\omega \sum_s \left\{ \|\mathbf{S}(\omega,\mathbf{m}_b)\mathbf{f}_s - \mathbf{d}_s^b(\omega)\|_2^2 + \|\mathbf{S}(\omega,\mathbf{m}_m)\mathbf{f}_s - \mathbf{d}_s^m(\omega)\|_2^2 \right\}$ is the sum of the baseline and monitor data misfit functions; Parameters $\mathbf{m} = (\mathbf{m}_b, \mathbf{m}_m)$ gather baseline and monitor models and represent squares slowness; $\Delta\mathbf{m} = \mathbf{m}_m - \mathbf{m}_b$ represents time-lapse model changes; $\mathbf{S} = \mathbf{P}\mathbf{A}^{-1}(\mathbf{m})$ is the modeling operator with $\mathbf{P}$ the observation operator, $\mathbf{A} = \omega^2 \mathrm{diag}(\mathbf{m}) + \nabla^2$ the Helmholtz operator, ω the angular frequency; $\mathbf{f}_s$ is the source; $\mathbf{d}_s$ is the recorded data; $\lambda \in \mathbb{R}_+$ is the penalty parameter controlling the weight of the regularization term; ∇ and ∇^2 are the gradient and Laplace operators, respectively. To separate the $\ell 2$ and the $\ell 1$ subproblems and formulate the later as a denoising problem, we recast (1) as a constrained problem:

$$\min_{\substack{\mathbf{m},\mathbf{p} \\ \mathbf{p}=(\mathbf{p}_x,\mathbf{p}_z)}} \mathscr{F}(\mathbf{m}) + \lambda \left(\|\mathbf{p}_x\|_1 + \|\mathbf{p}_z\|_1 \right) \quad s.t \quad \mathbf{p} = \nabla\Delta\mathbf{m}, \tag{2}$$

where $\mathbf{p} = (\mathbf{p}_x, \mathbf{p}_z)$ are auxiliary variables. We solve (2) with the augmented Lagrangian (AL) method, which combines a penalty function and a Lagrangian term.

This leads to the saddle-point problem:

$$\min_{\substack{\mathbf{m},\mathbf{p}\\ \mathbf{p}=(\mathbf{p}_x,\mathbf{p}_z)}} \max_{\mathbf{q}=(\mathbf{q}_x,\mathbf{q}_z)} \mathscr{F}(\mathbf{m}) + \lambda\left(\|\mathbf{p}_x\|_1 + \|\mathbf{p}_z\|_1\right) + \frac{\rho_x}{2}\|\mathbf{p}_x - \nabla_x\Delta\mathbf{m}\|_2^2 + \frac{\rho_z}{2}\|\mathbf{p}_z - \nabla_z\Delta\mathbf{m}\|_2^2 + \langle\mathbf{q}_x, \mathbf{p}_x - \nabla_x\Delta\mathbf{m}\rangle + \langle\mathbf{q}_z, \mathbf{p}_z - \nabla_z\Delta\mathbf{m}\rangle,$$

which can be reformulated in a more convenient scaled form as

$$\min_{\substack{\mathbf{m},\mathbf{p}\\ \mathbf{p}=(\mathbf{p}_x,\mathbf{p}_z)}} \max_{\bar{\mathbf{q}}=(\bar{\mathbf{q}}_x,\bar{\mathbf{q}}_z)} \mathscr{F}(\mathbf{m}) + \lambda\left(\|\mathbf{p}_x\|_1 + \|\mathbf{p}_z\|_1\right) + \frac{\rho_x}{2}\|\mathbf{p}_x - \nabla_x\Delta\mathbf{m} + \bar{\mathbf{q}}_x\|_2^2 + \frac{\rho_z}{2}\|\mathbf{p}_z - \nabla_z\Delta\mathbf{m} + \bar{\mathbf{q}}_z\|_2^2 - \frac{\rho_x}{2}\|\bar{\mathbf{q}}_x\|_2^2 - \frac{\rho_z}{2}\|\bar{\mathbf{q}}_z\|_2^2,$$

where $\bar{\mathbf{q}}_x = \frac{1}{\rho_x}\mathbf{q}_x, \bar{\mathbf{q}}_z = \frac{1}{\rho_z}\mathbf{q}_z$ are the scaled version of the Lagrange multipliers $\mathbf{q}_x$ and $\mathbf{q}_z$ by the penalty parameters ρ_x and ρ_z. In this study, we use $\rho_x = \rho_z = \rho$. Finally, the primal and dual variables, $\mathbf{m}$, $\mathbf{p}$, $\mathbf{q}$ are updated with alternating directions in the frame of ADMM (Boyd et al., 2010):

$$\mathbf{m}^{k+1} = \arg\min_{\mathbf{m}} \mathscr{F}(\mathbf{m}) + \frac{\rho}{2}\|\mathbf{p}^k - \nabla\Delta\mathbf{m} + \bar{\mathbf{q}}^k\|_2^2, \tag{3}$$

$$\mathbf{p}^{k+1} = \arg\min_{\mathbf{p}} \frac{\rho}{2}\|\mathbf{p} - \nabla\Delta\mathbf{m}^{k+1} + \bar{\mathbf{q}}^k\|_2^2 + \lambda\|\mathbf{p}\|_1, \tag{4}$$

$$\bar{\mathbf{q}}^{k+1} = \bar{\mathbf{q}}^k + \mathbf{p}^{k+1} - \nabla\Delta\mathbf{m}^{k+1}. \tag{5}$$

We refer the recurrence, eqs. 3-5, to as the ADMM iteration, which represents the outer loop of the algorithm. Within an ADMM iteration, the dual variables are updated with one steepest-ascent iteration and represent the running sum of constraint violations in iteration, eq. 5. The auxiliary primal variables $\mathbf{p}$ are the unknowns of a denoising problem, eq. 4. Proximal algorithms give their closed-form solutions as the element-wise soft thresholding of the gradient of $\mathbf{m}$ minus the dual variables (Combettes and Pesquet, 2011; Aghamiry et al., 2021):

$$\mathbf{p}^{k+1} = \text{shrink}(\nabla\Delta\mathbf{m}^{k+1} - \mathbf{q}^k, \gamma) = \text{sign}(\nabla\Delta\mathbf{m}^{k+1} - \mathbf{q}^k)\max\left(|\nabla\Delta\mathbf{m}^{k+1} - \mathbf{q}^k| - \gamma, 0\right), \tag{6}$$

where $\gamma = \lambda/\rho$. Soft thresholding of a linear function is illustrated in Fig. 1a. Note the bias generated by this thresholding. Parameters $\mathbf{m}$ are updated at ADMM iteration k by performing one or several inner iterations of FWI with a $\ell 2$ regularization term, eq. 3. In this study, we perform these inner FWI iterations with the quasi-Newton l-BFGS method using the SEISCOPE optimization toolbox (Métivier and Brossier, 2016). In summary, the algorithm involves two main hyper parameters: the penalty parameter ρ controlling the relative weight between the data misfit term and the regularization term in the $\mathbf{m}$ subproblem, eq. 3, and the soft thresholding parameter γ, eq. 6. Another key point is indeed the number of outer ADMM iterations and the number of inner l-BFGS iterations per ADMM iteration. In the next section, we provide some practical insights on the tuning of these parameters when we perform TL-FWI with a basic parallel approach. In the parallel approach, we apply the above ADMM algorithm for $\mathbf{m}_b$ and $\mathbf{m}_m$ separately rather than simultaneously and we apply the TV regularization on $\mathbf{m}_b$ and $\mathbf{m}_m$ rather than on their differences. Assessment of TV regularization in the frame of the joint approach is ongoing.

Numerical tests

We first validate our algorithm in a favorable setting by considering a simple 2D 3 $km \times 3$ km inclusion model discretized with a grid interval of 50 m (Fig. 2). The baseline model contains a large circular inclusion in a homogeneous background model. The monitor model contains two additional small elliptical inclusions (Fig. 2a,e). The starting model of both baseline and monitor inversions is the homogeneous background model. Full-aperture acquisition involves 29 sources and receivers along the four edges of the model. Source signature is a delta function assigning the same weight to each spectral component. Frequency-domain FWI of 23 frequencies between 3 Hz and 40 Hz is performed in one go for noise-free data. The FWI results for the baseline and monitor models are shown when TV regularization is used or not in Fig. 2(b-d),(f-h). The role of the TV regularization is to filter out the

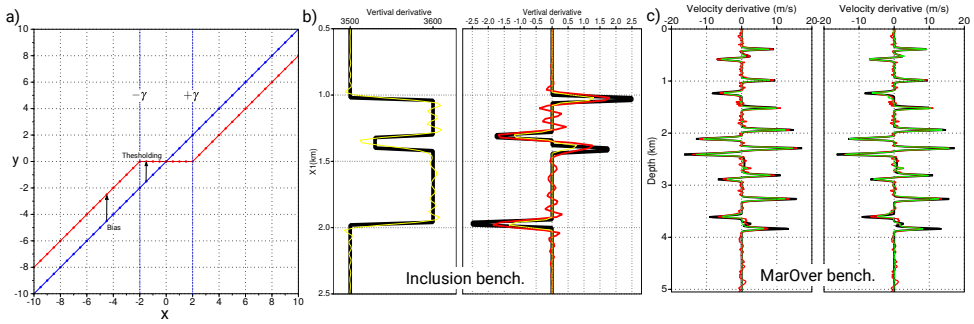

Figure 1: (a) Soft thresholding of linear function ($\gamma = 2$). (b) Inclusion benchmark. (Left) Vertical profile across negative time-lapse perturbation of true (black) and unregularized FWI (yellow) models. (Right) Velocity derivative of true (black) and unregularized FWI (red) models. The effect of soft thresholding (γ=0.5, eq. 6) on the derivative of the unregularized FWI profile is illustrate for representative subsurface contrasts (yellow curve). Noise is filtered out at the expense of contrasts. (c) MarOverthrust benchmark. (Left) Velocity derivative of true model (black), unregularized FWI model (red) and soft thresholding of unregularized FWI model profile (γ=1.2) (yellow). (Right) Yellow curve is now the profile of the final TV-regularized FWI model. The goal is to preserve noise removal across iterations while removing amplitude bias induced by soft thresholding.

Gibbs effects generated by the sharp high-cut truncature of the wavenumber spectrum. In this example, we perform 70 ADMM iterations, 10 l-BFGS iterations by ADMM iteration, and we use $\gamma = 0.5$. The initial value of the penalty parameter ρ has been tuned such that the two terms in eq.3 have roughly the same weight at the beginning of the inversion. Moreover, we progressively decrease ρ in iterations (multiplication by 0.8 every two ADMM iterations) as the data misfit decreases to guarantee sufficient data fit at convergence point. These heuristic rules are consistent with those discussed in Maharramov et al. (2016). For this simple test, we perfectly reconstruct the time-lapse model changes with the basic TV-regularized parallel approach. Soft thresholding is further illustrated with this example in Fig. 1b.

We present a more realistic example with a 2D marine version of the SEG/EAGE overthrust model

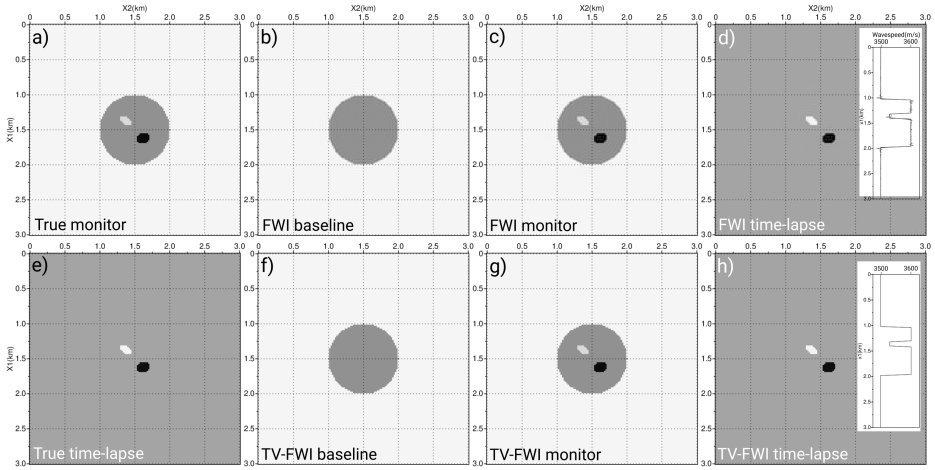

Figure 2: (a) True monitor model. (b-c) FWI baseline (b) and monitor (c) models. Note Gibbs effects in (b-c). (d) TL-FWI model changes ((b) minus (c)) to be compared with true time-lapse model changes in (e). Inset: vertical profiles across negative (-70 m/s) anomaly of (a) and (c) (black an,d gray, respectively). (f-h) Same as (b-d) when TV regularization is used.

on which we superimpose several time-lapse model perturbations (Fig. 3). Accurate starting model has been generated by Gaussian smoothing of true model with horizontal and vertical correlation lengths of 250 m. We perform multi-scale inversion by frequency continuation between 3 Hz and 37.5 Hz. We start applying TV regularization from the 7 Hz frequency. The source signature is a delta function. A fixed-spread seabed acquisition involves 97 reciprocal sources on the seabed spaced 200 m apart and 400 reciprocal receivers near the surface spaced 25 m apart (without surface multiples). Similar tun-

ing has been used as for inclusion test (γ=0.5 and progressive decrease of ρ). However, we increase the number of l-BFGS iterations from 10 to 20. Although the TV-regularization significantly filters out noise in both baseline and monitor models (Fig.3c-d versus Fig.3e-f), the velocity changes inferred from the TV-regularized FWI models still contain some noise although the velocity changes are fairly well identified (Fig.3f). This highlights the potential benefit of joint time-lapse FWI compared to the parallel approach even in this favorable setting.

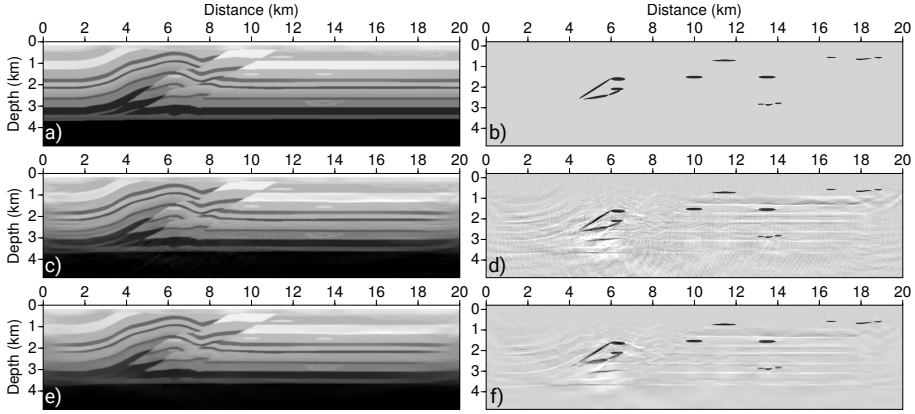

Figure 3: MarOverthrust benchmark. (a-b) True monitor model (a) and time-lapse model changes (b). (c-d) Unregularized FWI monitor model (c) and time-lapse model changes (d) (difference between monitor model and baseline model (not shown)). (e-f) Same as (c-d) for TV-regularized FWI.

Conclusion and perspectives

We have reviewed a practical algorithm for joint time-lapse FWI where a key ingredient is the non-smooth regularization linking together the baseline and the monitor subproblems. We assess heuristically the tuning of the hyper parameters, which mainly involve a penalty parameter and a threshold parameter. Future works aim at recasting this heuristic tuning in a better controlled mathematical framework with adaptive or fast ADMM. The parametrization of the joint approach can be revisited in terms of baseline model $\mathbf{m}_b$ and model changes $\Delta\mathbf{m}$, and TV regularization can be applied to both. A pragmatical approach disregarding the $\ell 2$ regularization term in the $\mathbf{m}$-subproblem and the Lagrange multipliers while performing TV denoising of the updated model at each FWI iteration may be investigated to simplify the algorithm. In a longer-term perspective, ADMM-based TV-regularized frequency-domain joint TL-FWI will be applied to 4D OBN data collected in the Ivar Aasen field (North Sea).

Acknowledgement

We thank the sponsors of the WIND consortium for their support and S. Beller (CNRS) for fruitful discussions. This work was granted access to the HPC resources of GENCI under allocation 0596.

References

Aghamiry, H., Gholami, A. and Operto, S. [2019] Implementing bound constraints and total-variation regularization in extended full waveform inversion with the alternating direction method of multiplier: application to large contrast media. *Geophysical Journal International*, **218**(2), 855–872.

Aghamiry, H., Gholami, A. and Operto, S. [2021] Full Waveform Inversion by Proximal Newton Methods using Adaptive Regularization. *Geophysical Journal International*, **224**(1), 169–180.

Boyd, S., Parikh, N., Chu, E., Peleato, B. and Eckstein, J. [2010] Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and trends in machine learning*, **3**(1), 1–122.

Combettes, P.L. and Pesquet, J.C. [2011] Proximal Splitting Methods in Signal Processing. In: Bauschke, H.H., Burachik, R.S., Combettes, P.L., Elser, V., Luke, D.R. and Wolkowicz, H. (Eds.) *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, *Springer Optimization and Its Applications*, 49, Springer New York, 185–212.

Maharramov, M., Biondi, B.L. and Meadows, M.A. [2016] Time-lapse inverse theory with applications. *Geophysics*, **81**(6), R485–R501.

Métivier, L. and Brossier, R. [2016] The SEISCOPE Optimization Toolbox: A large-scale nonlinear optimization library based on reverse communication. *Geophysics*, **81**(2), F11–F25.

Operto, S. and Miniussi, A. [2018] On the role of density and attenuation in 3D multi-parameter visco-acoustic VTI frequency-domain FWI: an OBC case study from the North Sea. *Geophysical Journal International*, **213**, 2037–2059.