Implementing bound constraints and Hybrid TV+Tikhonov Regularization in Wavefield Reconstruction Inversion with the Alternating Direction Method of Multiplier

MS35 Practical Aspects of Large-scale Sparsity-promoting Seismic Inversion - Part II of II

H. S. Aghamiry[1,2] A. Gholami[1] S. Operto[2],

[1] University of Tehran, Institute of Geophysics, Tehran, Iran
[2] University of Côte d'Azur, Geoazur, CNRS - IRD - OCA, Valbonne, France,

- Yesterday presentation:

  Improving the Wavefield Reconstruction Inversion (WRI) method based upon the alternating direction method of multiplier (ADMM).

  For more details,

  H. Aghamiry, A. Gholami and S. Operto, Improving full-waveform inversion by wavefield reconstruction with alternating direction method of multipliers, *Geophysics*, 84(1), R139-R162, 2019.

- Today focus:

  Which regularizer for subsurface imaging and its interfacing with ADMM-based WRI.

  For more details,

  1. Aghamiry, A. Gholami and S. Operto, Implementing bound constraints and total-variation regularization in extended full waveform inversion with the alternating direction method of multiplier: application to large contrast media, *arXiv:1902.02744*, 2019.
  2. H. Aghamiry, A. Gholami and S. Operto, Compound Regularization in Full-waveform Inversion for Imaging Piecewise Media, *arXiv:1903.04405*, 2019.

Content

Introduction

Regularized ADMM-based Wavefield Reconstruction Inversion (WRI)

- General framework of ADMM-based WRI
- Models of compound regularizers
- Implementing compound regularizers in ADMM-based WRI

Numerical example: The 2004 BP salt model

Conclusions

References

Outline

Introduction

Regularized ADMM-based Wavefield Reconstruction Inversion (WRI)

Numerical example: The 2004 BP salt model

Conclusions

References

- PDE-constrained optimization problem

$$\min_{\mathbf{m},\mathbf{u}} \quad \sum_s \|\mathbf{P}\mathbf{u}_s - \mathbf{d}_s\|_2^2, \quad \text{subject to} \quad \mathbf{A}(\mathbf{m})\mathbf{u}_s = \mathbf{b}_s, s \in [1; N_s] \quad (1)$$

where $\mathbf{A}(\mathbf{m}) = \omega^2 \text{diag}\,(\mathbf{m}) + \Delta$ is the scalar Hemmholtz operator.

- Method of Lagrange multiplier

$$\min_{\mathbf{m},\mathbf{u}} \max_{\mathbf{v}} \mathcal{L}(\mathbf{m}, \mathbf{u}_s, \mathbf{v}_s) = \min_{\mathbf{m},\mathbf{u}} \max_{\mathbf{v}} \sum_s \|\mathbf{P}\mathbf{u}_s - \mathbf{d}_s\|_2^2 + \sum_s \mathbf{v}_s^T \left[\mathbf{A}(\mathbf{m})\mathbf{u}_s - \mathbf{b}_s\right]. \quad (2)$$

- **Full-space** formulation: joint update of $\mathbf{m}$, $\mathbf{u}$, $\mathbf{v}$ (KKT system).
- **Reduced-space** unconstrained optimization (projection on the parameter space).

$$\min_{\mathbf{m}} \sum_s \|\mathbf{P}\mathbf{A}^{-1}(\mathbf{m})\mathbf{b}_s - \mathbf{d}_s\|_2^2.$$

- **Highly nonlinear** cycle skipping $\rightarrow$ Need of **extended search space**.

Wavefield Reconstruction Inversion (WRI) (van Leeuwen and Herrmann, 2013, 2016)

Extending the FWI search space

- **Wavefield Reconstruction Inversion (WRI)** extends the FWI search space with a **penalty method**.

$$\min_{\mathbf{u},\mathbf{m}} \sum_s \|\mathbf{P}\mathbf{u}_s - \mathbf{d}_s\| + \lambda \sum_s \|\mathbf{A}(\mathbf{m})\mathbf{u}_s - \mathbf{b}_s\|_2^2 \quad \text{(amount to set } \mathbf{v} = \lambda\,[\mathbf{A}(\mathbf{m})\mathbf{u} - \mathbf{b}]\text{)}.$$

1. **Wave-equation relaxation** with a feedback term to the data
   → foster data fidelity and prevent cycle skipping accordingly.

$$\begin{pmatrix} \sqrt{\lambda}\mathbf{A}(\mathbf{m}_0) \\ \mathbf{P} \end{pmatrix} \mathbf{u}_s = \begin{pmatrix} \sqrt{\lambda}\mathbf{b}_s \\ \mathbf{d}_s \end{pmatrix},$$

2. Parameter estimation by **minimization of the source residuals** the relaxation generated
   → Push back the reconstructed wavefield toward the wave equation constraint.

$$\mathbf{m}^* = \arg\min_{\mathbf{m}} \sum_s \|\mathbf{A}(\mathbf{m})\mathbf{u}_s^* - \mathbf{b}_s\|_2^2$$

$$\nabla_m C(\mathbf{m}) = \sum_s \left(\frac{\partial \mathbf{A}(\mathbf{m})}{\partial m} \mathbf{u}_s^* \right)^T \left(\mathbf{A}(\mathbf{m})\mathbf{u}_s^* - \mathbf{b}_s \right).$$

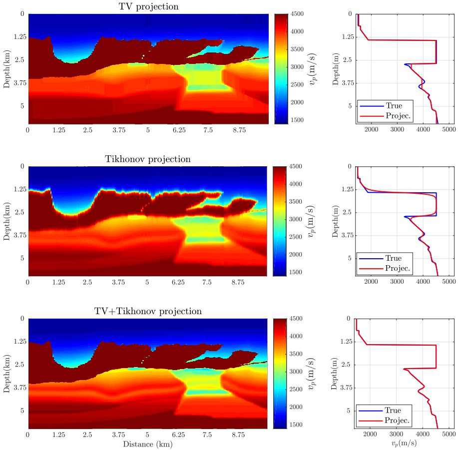
- Steps 1 and 2 are solved in **alternating mode** or through **variable projection**. The former breaks down FWI into a sequence of two linear subproblems (**FWI is a biconvex problem**).
- The issue of the **dynamic control of the penalty parameter** (Fu and Symes, 2017).

- FWI ill-posedness $\rightarrow$ need prior implemented with **regularization**.
  1. Noise
  2. Approximate wave physics
  3. Incomplete subsurface illumination from the surface
  4. Parameter cross-talk in multi-parameter reconstruction
  5. Large contrasts (salt, basalt, ...)
- Two popular regularizations in geophysics and image denoising:
  1. Second-order Tikhonov regularization
  $$\|\mathbf{m}\|_{Tikh} = \sum \|\nabla_x^2 \mathbf{m}\|_2^2 + \|\nabla_y^2 \mathbf{m}\|_2^2 + \|\nabla_z^2 \mathbf{m}\|_2^2.$$
  Drive inversion toward smooth reconstruction.
  2. Blockiness-promoting Isotropic Total Variation (TV) regularization
  $$\|\mathbf{m}\|_{TV} = \sum \sqrt{|\nabla_x \mathbf{m}|^2 + |\nabla_y \mathbf{m}|^2 + |\nabla_z \mathbf{m}|^2}.$$
  Drive inversion toward piecewise homogeneous (blocky) reconstruction.
  ∇_i and ∇_i^2: first and second-order difference operators in the i direction ($i \in \{x, y, z\}$).
- Regularizations implemented in FWI with penalty method (Askan et al., 2007; Anagaw and Sacchi, 2011; Brandsberg-Dahl et al., 2017; Kazei et al., 2017) or as a hard constraint (Peters and Herrmann, 2017; Esser et al., 2018).

Interfacing compound regularizations with ADMM in WRI:

Combining smoothness and blockiness (Gholami and Hosseini, 2013)

- The subsurface as a **piecewise smooth medium** (a stack of layers in which properties vary smoothly) → a single regularization cannot account for the different statistical properties of the subsurface → combine Tikhonov and TV regularizations.



Outline

Introduction

Regularized ADMM-based Wavefield Reconstruction Inversion (WRI)

General framework of ADMM-based WRI

Models of compound regularizers

Implementing compound regularizers in ADMM-based WRI

Numerical example: The 2004 BP salt model

Conclusions

References

Outline

Introduction

Regularized ADMM-based Wavefield Reconstruction Inversion (WRI)

General framework of ADMM-based WRI

Models of compound regularizers

Implementing compound regularizers in ADMM-based WRI

Numerical example: The 2004 BP salt model

Conclusions

References

- Regularized constrained problem

$$\min_{\mathbf{u}_s, \mathbf{m}} \mathcal{C}(\mathbf{u}, \mathbf{m}) = \text{Reg}(\mathbf{m}) \text{ subject to } \mathbf{P}\mathbf{u}_s = \mathbf{d}_s \text{ and } \mathbf{A}(\mathbf{m})\mathbf{u}_s = \mathbf{b}_s, s \in [1; N_s].$$

- Tackled with Augmented Lagrangian method

$$\begin{aligned} \min_{\mathbf{u}_s, \mathbf{m} \in \mathcal{C}} \max_{\mathbf{v}, \mathbf{w}} \text{Reg}(\mathbf{m}) &+ \sum_{s=1}^{N_s} \mathbf{v}_s^T [\mathbf{P}\mathbf{u}_s - \mathbf{d}_s] + \sum_{s=1}^{N_s} \mathbf{w}_s^T [\mathbf{A}(\mathbf{m})\mathbf{u}_s - \mathbf{b}_s] \\ &+ \frac{\lambda}{2} \sum_{s=1}^{N_s} \|\mathbf{P}\mathbf{u}_s - \mathbf{d}_s\|_2^2 + \frac{\gamma}{2} \sum_{s=1}^{N_s} \|\mathbf{A}(\mathbf{m})\mathbf{u}_s - \mathbf{b}_s\|_2^2, \end{aligned} \quad (3)$$

where $\mathbf{v}_s$ and $\mathbf{w}_s$ are the dual variables (the Lagrangian multipliers).

- The augmented Lagrangian, eq. (3), can also be written in scaled compact form as

$$\begin{aligned} \min_{\mathbf{u}_s,\mathbf{m}\in\mathcal{C}} \max_{\mathbf{v}_s,\mathbf{w}_s} \; & \mathrm{Reg}(\mathbf{m}) + \frac{\lambda}{2}\sum_{s=1}^{N_s}\|\mathbf{P}\mathbf{u}_s - \mathbf{d}_s + \frac{1}{\lambda}\mathbf{v}_s\|_2^2 - \frac{\lambda}{2}\sum_{s=1}^{N_s}\|\mathbf{v}_s\|_2^2 \\ & + \frac{\gamma}{2}\sum_{s=1}^{N_s}\|\mathbf{A}(\mathbf{m})\mathbf{u}_s - \mathbf{b}_s + \frac{1}{\gamma}\mathbf{w}_s\|_2^2 - \frac{\gamma}{2}\sum_{s=1}^{N_s}\|\mathbf{w}_s\|_2^2, \end{aligned} \quad (4)$$

- Method of multiplier (Primal descent / Dual ascent) after change of variables $\mathbf{d}_s^k = -\mathbf{v}_s^k/\lambda$ and $\mathbf{b}_s^k = -\mathbf{w}_s^k/\gamma$:

$$\begin{aligned} & \min_{\mathbf{u}_s,\mathbf{m}\in\mathcal{C}} \; \mathrm{Reg}(\mathbf{m}) + \frac{\lambda}{2}\sum_{s=1}^{N_s}\|\mathbf{P}\mathbf{u}_s - \mathbf{d}_s - \mathbf{d}_s^k\|_2^2 + \frac{\gamma}{2}\sum_{s=1}^{N_s}\|\mathbf{A}(\mathbf{m})\mathbf{u}_s - \mathbf{b}_s - \mathbf{b}_s^k\|_2^2, \quad \text{(Primal d} \\ & \mathbf{d}_s^{k+1} = \mathbf{d}_s^k + \mathbf{d}_s - \mathbf{P}\mathbf{u}_s, \quad \text{(Dual ascent)} \\ & \mathbf{b}_s^{k+1} = \mathbf{b}_s^k + \mathbf{b}_s - \mathbf{A}(\mathbf{m})\mathbf{u}_s, \quad \text{(Dual ascent)} \end{aligned} \quad (5)$$

beginning with $\mathbf{d}_s^0 = 0$ and $\mathbf{b}_s^0 = 0$.

- ADMM breaks down the primal problem into two linear sub-problems (biconvex problem)

$$\mathbf{u}_s^{k+1} = \arg\min_{\mathbf{u}} \left\| \begin{bmatrix} \sqrt{\frac{\lambda}{\gamma}}\mathbf{P} \\ \mathbf{A}(\mathbf{m}^k) \end{bmatrix} \mathbf{u}_s - \begin{bmatrix} \sqrt{\frac{\lambda}{\gamma}}(\mathbf{d}_s + \mathbf{d}_s^k) \\ \mathbf{b}_s + \mathbf{b}_s^k \end{bmatrix} \right\|_2^2 \quad \text{(Primal descent)} \tag{6a}$$

$$\mathbf{m}^{k+1} = \arg\min_{\mathbf{m}\in\mathcal{C}} \text{Reg}(\mathbf{m}) + \frac{\gamma}{2}\sum_{s=1}^{N_s} \|\mathbf{A}(\mathbf{m})\mathbf{u}_s^{k+1} - \mathbf{b}_s - \mathbf{b}_s^k\|_2^2, \quad \text{(Primal descent)} \tag{6b}$$

$$\mathbf{d}_s^{k+1} = \mathbf{d}_s^k + \mathbf{d}_s - \mathbf{P}\mathbf{u}_s^{k+1}, \quad \text{(Dual ascent)} \tag{6c}$$

$$\mathbf{b}_s^{k+1} = \mathbf{b}_s^k + \mathbf{b}_s - \mathbf{A}(\mathbf{m}^{k+1})\mathbf{u}_s^{k+1}, \quad \text{(Dual ascent)} \tag{6d}$$

- The first subproblem (wavefield reconstruction) has a closed-form solution

$$\left(\mathbf{A}^T(\mathbf{m}_0)\mathbf{A}(\mathbf{m}_0) + \frac{\lambda}{\gamma}\mathbf{P}^T\mathbf{P}\right)\mathbf{u}_s^* = \left(\mathbf{A}^T(\mathbf{m}_0)\mathbf{b}_s + \frac{\lambda}{\gamma}\mathbf{P}^T\mathbf{d}_s\right).$$

- The second subproblem is more complex. We discuss first the choice of Reg before presenting its solution.

Outline

Introduction

Regularized ADMM-based Wavefield Reconstruction Inversion (WRI)

General framework of ADMM-based WRI

Models of compound regularizers

Implementing compound regularizers in ADMM-based WRI

Numerical example: The 2004 BP salt model

Conclusions

References

- **Convex Combination (CC)**: The solution is forced to satisfy the individual priors simultaneously.

$$\Phi_\alpha(\mathbf{x}) = \alpha_1 \Phi_1(\mathbf{x}) + ... + \alpha_r \Phi_r(\mathbf{x}), \tag{7}$$

where weights α_i satisfy $\alpha_i \geq 0$ and

$$\alpha_1 + \alpha_2, ..., +\alpha_r = 1. \tag{8}$$

As an example, CC of ℓ_1- and ℓ_2-norms ($\ell_1 + \ell_2$-norm) (Gholami, 2013)

$$\Phi_\alpha(\mathbf{x}) = (1-\alpha)\|\mathbf{x}\|_2^2 + \alpha\|\mathbf{x}\|_1, \tag{9}$$

with $0 \leq \alpha \leq 1$, which is called Elastic net (Zou and Hastie, 2005).

- **Infimal Convolution (IC)**: The solution is explicitly decomposed into simple components, each of them being regularized by an appropriate prior.

$$\Phi_\alpha(\mathbf{x}) = \min_{\mathbf{x}=\mathbf{x}_1+\ldots+\mathbf{x}_r} \{\alpha_1 \Phi_1(\mathbf{x}_1) + \ldots + \alpha_r \Phi_r(\mathbf{x}_r)\}. \tag{10}$$

In the case of two functionals,

$$\Phi_\alpha(\mathbf{x}) = \min_{\mathbf{z}} \{(1-\alpha)\Phi_1(\mathbf{x}-\mathbf{z}) + \alpha\Phi_2(\mathbf{z})\}. \tag{11}$$

The IC of ℓ_1- and ℓ_2-norms ($\ell_1 \oplus \ell_2$-norm) leads to the following denoising problem

$$\Phi_\alpha(\mathbf{x}) = \min_{\mathbf{z}} \{(1-\alpha)\|\mathbf{x}-\mathbf{z}\|_2^2 + \alpha\|\mathbf{z}\|_1\}, \tag{12}$$

which reduces to soft-thresholding (Donoho, 1995):

$$\mathbf{z} = \max\left(1 - \frac{\mu}{|\mathbf{x}|}, 0\right) \circ \mathbf{x}, \tag{13}$$

where $\mu = \frac{\alpha}{2(1-\alpha)}$. Plugging $\mathbf{z}$ from (13) into (12) gives

$$\Phi_\mu(\mathbf{x}) = \begin{cases} \frac{1}{2\mu}|\mathbf{x}|^2 & \text{if } |\mathbf{x}| \leq \mu \\ |\mathbf{x}| - \frac{\mu}{2} & \text{if } |\mathbf{x}| > \mu \end{cases}, \tag{14}$$

which is nothing other than the Huber function (Huber, 1973).

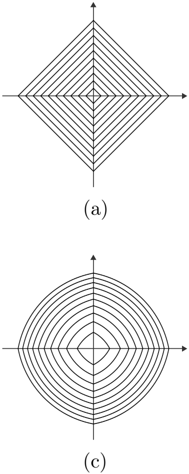


Figure 1: Geometrical illustration of different regularizers. (a) the ℓ_1-norm, (b) the ℓ_2-norm, (c) the $(\ell_1 + \ell_2)$-norm, and (d) the $(\ell_1 \oplus \ell_2)$-norm.

We test two infimal-convolution compound regularizer

- **IC-based TT regularizer**: Tikhonov + first-order Total Variation

$$\Phi_\alpha^{\mathsf{TT}}(\mathbf{x}) = \min_{\mathbf{x}=\mathbf{x}_1+\mathbf{x}_2} (1-\alpha)\|\nabla^2 \mathbf{x}_2\|_2^2 + \alpha\|\nabla \mathbf{x}_1\|_1. \tag{15}$$

- **Total Generalized Variation (TGV) regularizer**: first-order + second-order Total Variation (Bredies et al., 2010; Setzer et al., 2011)

$$\Phi_\alpha^{\mathsf{TGV}}(\mathbf{x}) = \min_{\mathbf{x}=\mathbf{x}_1+\mathbf{x}_2} (1-\alpha)\|\nabla^2 \mathbf{x}_2\|_1 + \alpha\|\nabla \mathbf{x}_1\|_1. \tag{16}$$

Introduction

Regularized ADMM-based Wavefield Reconstruction Inversion (WRI)

General framework of ADMM-based WRI

Models of compound regularizers

Implementing compound regularizers in ADMM-based WRI

Numerical example: The 2004 BP salt model

Conclusions

References

- IC-TT regularized parameter estimation subproblem

$$\mathbf{m}^{k+1} = \underset{\substack{\mathbf{m}=\mathbf{m}_1+\mathbf{m}_2 \\ \mathbf{m}\in\mathcal{C}}}{\arg\min} \; \Phi_{\alpha}^{\mathsf{TT}}(\mathbf{m}_1,\mathbf{m}_2) + \frac{\gamma}{2}\sum_{s=1}^{N_s} \|\mathbf{A}(\mathbf{m})\mathbf{u}_s^{k+1} - \mathbf{b}_s - \mathbf{b}_s^k\|_2^2, \qquad (17)$$

where $\mathbf{m}_1$ and $\mathbf{m}_2$ denote the blocky and the smooth components of the subsurface, respectively.

- Capitalizing on the bilinearity of the wave equation,

$$\mathbf{m}^{k+1} = \underset{\substack{\mathbf{m}=\mathbf{m}_1+\mathbf{m}_2 \\ \mathbf{m}\in\mathcal{C}}}{\arg\min} \; \Phi_{\alpha}^{\mathsf{TT}}(\mathbf{m}_1,\mathbf{m}_2) + \frac{\gamma}{2}\sum_{s=1}^{N_s} \|\mathbf{L}_s\mathbf{m} - \mathbf{y}_s\|_2^2, \qquad (18)$$

where $\mathbf{L}_s = \omega^2 \mathrm{diag}\left(\mathbf{u}_s^{k+1}\right)$ and $\mathbf{y}_s = \mathbf{b}_s + \mathbf{b}_s^k - \Delta\mathbf{u}_s^{k+1}$.

- In the sequel, we introduce auxiliary primal variable $\mathbf{p} = \nabla\mathbf{m}_1$ to decouple the ℓ_1 and the ℓ_2 minimization problems and solve the former ones with proximal algorithms following the split Bregman method (Goldstein and Osher, 2009).

Applying ADMM to (18) breaks down the multivariate primal problem into three subproblems

- Primal descent

$$\begin{bmatrix}\mathbf{m}_1^{k+1}\\ \mathbf{m}_2^{k+1}\end{bmatrix} = \underset{\mathbf{m}_1,\mathbf{m}_2}{\arg\min}\; C(\mathbf{m}_1,\mathbf{m}_2,\mathbf{p}^k,\mathbf{m}^k,\tilde{\mathbf{p}}^k,\tilde{\mathbf{m}}^k), \tag{19a}$$

$$\mathbf{p}^{k+1} = \underset{\mathbf{p}}{\arg\min}\; \alpha\|\mathbf{p}\|_1 + \frac{\zeta}{2}\|\nabla\mathbf{m}_1^{k+1} - \mathbf{p} - \tilde{\mathbf{p}}^k\|_2^2, \tag{19b}$$

$$\mathbf{m}^{k+1} = \underset{\mathbf{m}\in\mathcal{C}}{\arg\min}\; \frac{\eta}{2}\|\mathbf{m}_1^{k+1} + \mathbf{m}_2^{k+1} - \mathbf{m} - \tilde{\mathbf{m}}^k\|_2^2, \tag{19c}$$

where

$$\begin{aligned} C(\mathbf{m}_1,\mathbf{m}_2,\mathbf{p}^k,\mathbf{m}^k,\tilde{\mathbf{p}}^k,\tilde{\mathbf{m}}^k) = {} & \frac{\gamma}{2}\sum_{s=1}^{N_s}\|\mathbf{L}_s[\mathbf{m}_1+\mathbf{m}_2] - \mathbf{y}_s\|_2^2 + (1-\alpha)\|\nabla^2\mathbf{m}_2\|_2^2 \\ & + \frac{\zeta}{2}\|\nabla\mathbf{m}_1 - \mathbf{p}^k - \tilde{\mathbf{p}}^k\|_2^2 + \frac{\eta}{2}\|\mathbf{m}_1 + \mathbf{m}_2 - \mathbf{m}^k - \tilde{\mathbf{m}}^k\|_2^2, \end{aligned} \tag{20}$$

- Dual ascent

$$\tilde{\mathbf{p}}^{k+1} = \tilde{\mathbf{p}}^k + \mathbf{p}^{k+1} - \nabla\mathbf{m}_1^{k+1}, \tag{21a}$$

$$\tilde{\mathbf{m}}^{k+1} = \tilde{\mathbf{m}}^k + \mathbf{m}^{k+1} - (\mathbf{m}_1^{k+1} + \mathbf{m}_2^{k+1}). \tag{21b}$$

- ($\mathbf{m}_1$,$\mathbf{m}_2$) are solution of the following system

$$\begin{bmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{bmatrix}, \tag{22}$$

with

$$\begin{cases} \mathbf{G}_{11} = \gamma \sum_{s=1}^{N_s} \mathbf{L}_s^T \mathbf{L}_s + \zeta \nabla^T \nabla + \eta \mathbf{I} \\ \mathbf{G}_{12} = \mathbf{G}_{21} = \gamma \sum_{s=1}^{N_s} \mathbf{L}_s^T \mathbf{L}_s + \eta \mathbf{I} \\ \mathbf{G}_{22} = \gamma \sum_{s=1}^{N_s} \mathbf{L}_s^T \mathbf{L}_s + (1-\alpha)(\nabla^2)^T \nabla^2 + \eta \mathbf{I} \end{cases},$$

and

$$\begin{cases} \mathbf{h}_1 = \gamma \sum_{s=1}^{N_s} \mathbf{L}_s^T \mathbf{y}_s + \zeta \nabla^T [\mathbf{p}^k + \tilde{\mathbf{p}}^k] + \eta [\mathbf{m}^k + \tilde{\mathbf{m}}^k] \\ \mathbf{h}_2 = \gamma \sum_{s=1}^{N_s} \mathbf{L}_s^T \mathbf{y}_s + \eta [\mathbf{m}^k + \tilde{\mathbf{m}}^k] \end{cases},$$

where $\mathbf{I}$ is the identity matrix.

- From the first equation of (22), we find that

$$\mathbf{m}_2 = \mathbf{G}_{12}^{-1} [\mathbf{h}_1 - \mathbf{G}_{11} \mathbf{m}_1] \tag{23}$$

and plugging this into the second equation of (22) we get the following

$$\mathbf{m}_1 = (\mathbf{G}_{11} - \mathbf{G}_{22} \mathbf{G}_{12}^{-1} \mathbf{G}_{11})^{-1} [\mathbf{h}_2 - \mathbf{G}_{22} \mathbf{G}_{12}^{-1} \mathbf{h}_1]. \tag{24}$$

Interestingly, $\mathbf{L}$ is diagonal, implying that $\mathbf{G}_{12}$ is also diagonal. Thus we only need to solve an $n \times n$ system to estimate $\mathbf{m}_1$, from which $\mathbf{m}_2$ easily follows.

- $\mathbf{p} = [\mathbf{p}_x \mathbf{p}_z]^T$ estimated with a generalized proximity operator (Combettes and Pesquet, 2011)

$$\mathbf{p}^{k+1} = \text{prox}_{\zeta/\alpha}(\mathbf{z}) = \begin{bmatrix} \xi \circ \mathbf{z}_x \\ \xi \circ \mathbf{z}_z \end{bmatrix}, \tag{25}$$

where

$$\mathbf{z} = \nabla \mathbf{m}_1^{k+1} - \tilde{\mathbf{p}}^k = \begin{bmatrix} \mathbf{z}_x \\ \mathbf{z}_z \end{bmatrix}, \tag{26}$$

and

$$\xi = \max\left(1 - \frac{\zeta}{\alpha\sqrt{\mathbf{z}_x^2 + \mathbf{z}_z^2}}, 0\right). \tag{27}$$

- The subproblem for $\mathbf{m}$ has also a component-wise solution given by

$$\mathbf{m}^{k+1} = \text{proj}_{\mathcal{C}}(\mathbf{m}_1^{k+1} + \mathbf{m}_2^{k+1} - \tilde{\mathbf{m}}^k), \tag{28}$$

where the projection operator projects its argument onto the desired box $[\mathbf{m}_l.\mathbf{m}_u]$ according to $\text{proj}_{\mathcal{C}}(\bullet) = \min(\max(\bullet, \mathbf{m}_l), \mathbf{m}_u)$.

Introduction

Regularized ADMM-based Wavefield Reconstruction Inversion (WRI)

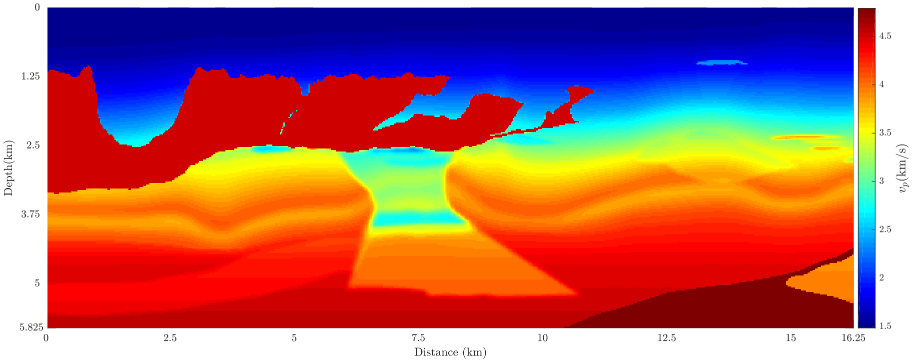
Numerical example: The 2004 BP salt model

Conclusions

References

Application to the BP salt model (left target):

True model



- Fixed-spread surface acquisition.
- Frequency bandwidth: 3-13 Hz.
- Frequency continuation: Batches of 3 frequencies with a 0.5Hz spacing. Three paths over batches.
- Noiseless.
- Stopping criterion of iteration:

$$k_{max} = 20 \quad \text{or} \quad (\|\mathbf{A}(\mathbf{m}^k)\mathbf{u}^k - \mathbf{b}\|_F \leq \delta \quad \text{and} \quad \|\mathbf{P}\mathbf{u}^k - \mathbf{d}\|_F \leq \epsilon_n), \tag{29}$$

  with δ=1e-3/1e-3 and ϵ_n=1e-5/noise level and k_{max} is the maximum number of iterations.
- Setting the penalty parameter λ
  λ = 1e-3/2e-2 μ where μ = largest eigenvalue of $\mathbf{A}^{-T}\mathbf{P}^T\mathbf{P}\mathbf{A}^{-1}$ (van Leeuwen and Herrmann, 2016).
- **Tested regularizers**: (a) Damping (DMP); (b) Tikhonov; (c) TV; (d) Convex combination of Tikhonov and TV (JTT); (e) Infimal convolution of Tikhonov and TV (TT); (f) Total Generalized Variation (TGV).

Penalty param.	α	λ	γ	ζ	η
Constraints	TT weight	Obs. Eq.	Wave Eq.	TV weight	Bounds

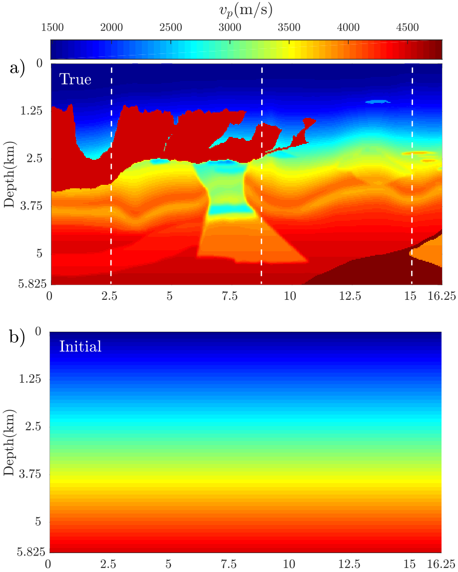
Table 1: α: balance Tikhonov and TV regularization. λ, γ, ζ, η: weights of the observation equation, wave equation, auxiliary TV term,bound constraint .wrt. regularization term, respectively.

Some guidelines to select penalty parameters (Aghamiry et al., 2019a)

- We found that α=0.7 was a good pragmatical value.
- We use $\zeta = \eta$.
- We found that $\zeta/\alpha = 2\% \max \|(\mathbf{z}_x \ \mathbf{z}_x)\|$ was a good pragmatical value.
- ζ/γ: small percentage of mean absolute value of the diagonal coefficients of $\sum_{i=1}^{N_s} \mathbf{L}_s^T \mathbf{L}_s$.
- λ/γ: small percentage of of the highest eigenvalue of $\mathbf{A}(\mathbf{m})^{-T} \mathbf{P}^T \mathbf{P} \mathbf{A}(\mathbf{m})^{-1}$ (van Leeuwen and Herrmann, 2016; Aghamiry et al., 2019b)

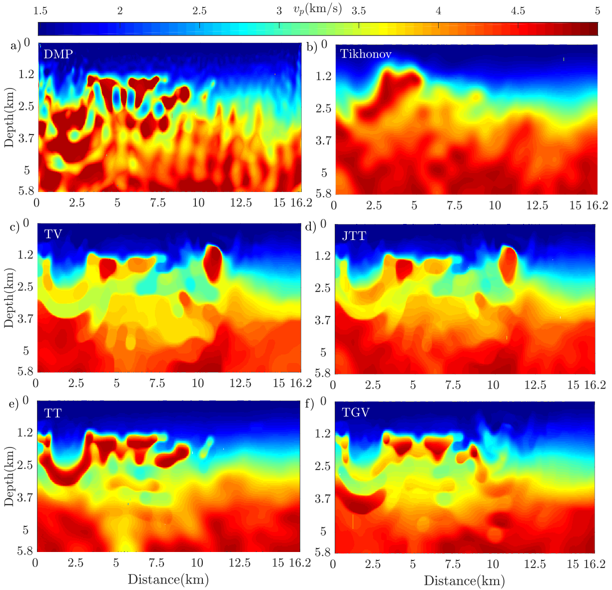
Regularized ADMM-based WRI Application to the BP salt model (left target):

True & Initial models



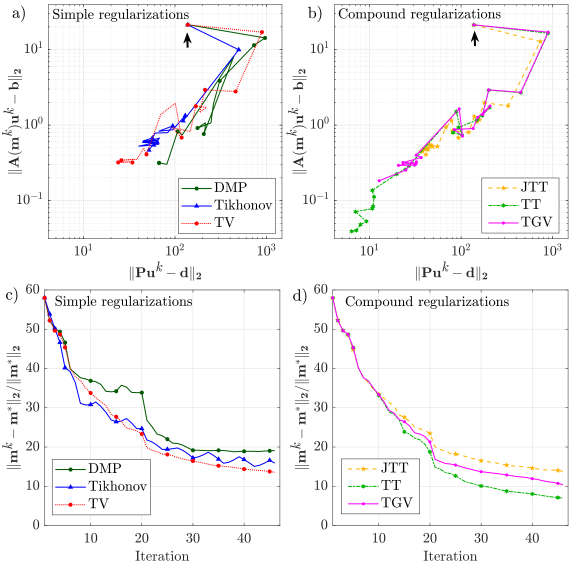
Application to the BP salt model (left target):

First frequency batch



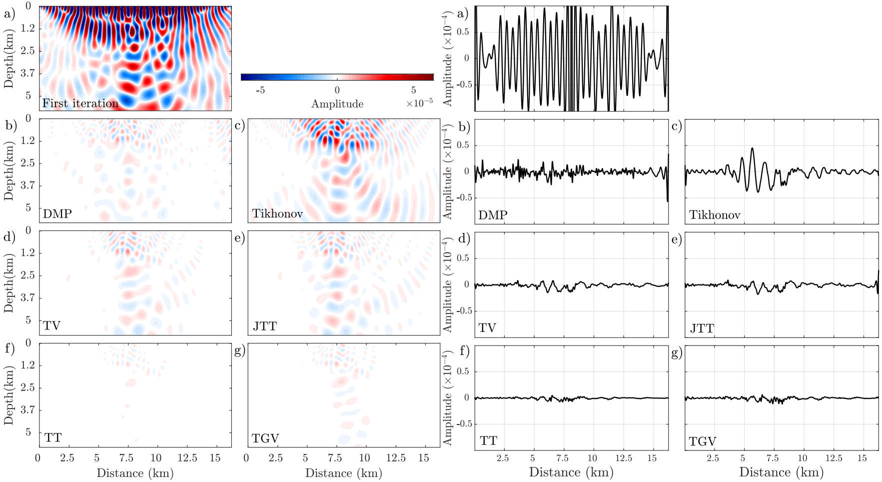
Application to the BP salt model (left target):

Convergence history



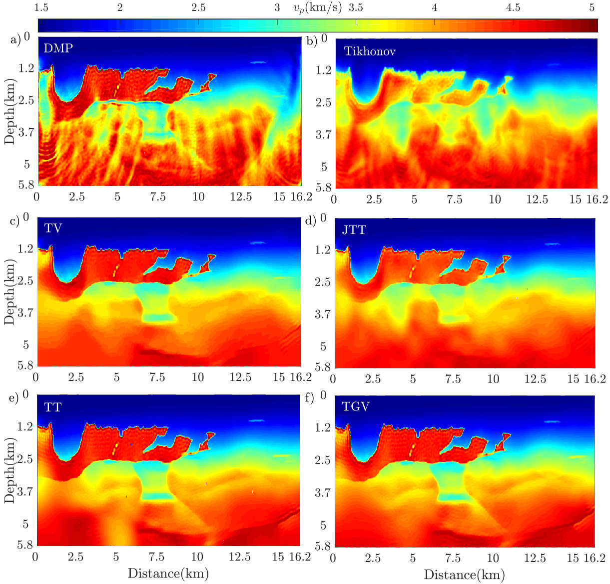
Application to the BP salt model (left target):

Source (left) and data (right) residuals at first iteration (top) and at convergence point (bottom)



Application to the BP salt model (left target):

Final models



Application to the BP salt model (left target):

Convergence speed

Regularizer	DMP	Tikhonov	TV	JTT	TT	TGV
# iteration	426	448	399	415	361	394

Table 2: Number of iterations of IR-WRI for each regularizer.

Introduction

Regularized ADMM-based Wavefield Reconstruction Inversion (WRI)

Numerical example: The 2004 BP salt model

Conclusions

References

Conclusions

- We have proposed a versatile recipe to cascade bound constraints and various regularizations in ADMM-based WRI (IR-WRI).
- Nonsmooth regularization are easily implemented with the so-called split Bregman method and proximal algorithms.
- The subsurface is formed by different components of different statistical properties. Need to combine different regularizations.
- These regularizations should be combined by infimal convolution rather than by convex combination.
- When infimal convolution is used, the different subsurface components can be jointly updated through variable projection.
- Infimal convolution of Tikhonov and TV regularizers perform the most reliable results. However, TGV is also a relevant alternative for piecewise linear models.
- Bound-constrained TT-regularized IR-WRI allows for the reconstruction of large-contrast media starting from scratch.
- Further assessment on real data collected by ultra-long offset sparse stationary-recording acquisitions is scheduled.

- For more details, see

1. H. Aghamiry, A. Gholami and S. Operto, Improving full-waveform inversion by wavefield reconstruction with alternating direction method of multipliers, Geophysics, 84(1), R139-R162, 2019.

2. H. Aghamiry, A. Gholami and S. Operto, Implementing bound constraints and total-variation regularization in extended full waveform inversion with the alternating direction method of multiplier: application to large contrast media, arXiv:1902.02744, 2019.

3. H. Aghamiry, A. Gholami and S. Operto, Compound Regularization in Full-waveform Inversion for Imaging Piecewise Media, arXiv:1903.04405, 2019.

Introduction

Regularized ADMM-based Wavefield Reconstruction Inversion (WRI)

Numerical example: The 2004 BP salt model

Conclusions

References

Aghamiry, H., Gholami, A., and Operto, S. (2019a). Implementing bound constraints and total-variation regularization in extended full waveform inversion with the alternating direction method of multiplier: application to large contrast media. *arXiv:1902.02744*.

Aghamiry, H., Gholami, A., and Operto, S. (2019b). Improving full-waveform inversion by wavefield reconstruction with alternating direction method of multipliers. *Geophysics*, 84(1):R139–R162.

Anagaw, A. and Sacchi, M. (2011). Regularized 2D acoustic full waveform inversion. In *Expanded Abstracts*. EAGE.

Askan, A., Akcelik, V., Bielak, J., and Ghattas, O. (2007). Full waveform inversion for seismic velocity and anelastic losses in heterogeneous structures. *Bulletin of the Seismological Society of America*, 97(6):1990–2008.

Bertsekas, D. P. (2016). *Nonlinear programming, Third edition*. Athena Scientific.

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

Brandsberg-Dahl, S., Chemingui, N., Valenciano, A., Ramos-Martinez, J., and Qiu, L. (2017). FWI for model updates in large-contrast media. *The Leading Edge*, Special section: Full Waveform Inversion Part II(1):81–87.

Bredies, K., Kunisch, K., and Pock, T. (2010). Total generalized variation. *SIAM Journal on Imaging Sciences*, 3(3):492–526.

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., editors, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, volume 49 of *Springer Optimization and Its Applications*, pages 185–212. Springer New York.

Donoho, D. L. (1995). De-noising by soft-thresholding. *IEEE transactions on information theory*, 41(3):613–627.

Esser, E., Guasch, L., van Leeuwen, T., Aravkin, A. Y., and Herrmann, F. J. (2018). Total variation regularization strategies in Full-Waveform Inversion. *SIAM Journal Imaging Sciences*, 11(1):376–406.

Fu, L. and Symes, W. W. (2017). A discrepancy-based penalty method for extended waveform inversion. *Geophysics*, R282-R298:78–82.

Gholami, A. (2013). Sparse time–frequency decomposition and some applications. *IEEE Transactions on Geoscience and Remote Sensing*, 51(6):3598–3604.

Gholami, A. and Hosseini, S. M. (2013). A balanced combination of tikhonov and total variation regularizations for reconstruction of piecewise-smooth signals. *Signal Processing*, 93:1945–1960.

Goldstein, T. and Osher, S. (2009). The split Bregman method for L1-regularized problems. *SIAM Journal on Imaging Sciences*, 2(2):323–343.

Huber, P. J. (1973). Robust regression: Asymptotics, conjectures, and Monte Carlo. *The Annals of Statistics*, 1(5):799–821.

Kazei, V., Kalita, M., and Alkhalifah, T. (2017). Salt-body inversion with minimum gradient support and sobolev space norm regularizations. In *Expanded Abstracts, 79th Annual EAGE Meeting (Paris)*.

Peters, B. and Herrmann, F. J. (2017). Constraints versus penalties for edge-preserving full-waveform inversion. *The Leading Edge*, 36(1):94–100.

Setzer, S., Steidl, G., and Teuber, T. (2011). Infimal convolution regularizations with discrete ℓ1-type functionals. *Communications in Mathematical Sciences*, 9(3):797–827.

van Leeuwen, T. and Herrmann, F. (2016). A penalty method for PDE-constrained optimization in inverse problems. *Inverse Problems*, 32(1):1–26.

van Leeuwen, T. and Herrmann, F. J. (2013). Mitigating local minima in full-waveform inversion by expanding the search space. *Geophysical Journal International*, 195(1):661–667.

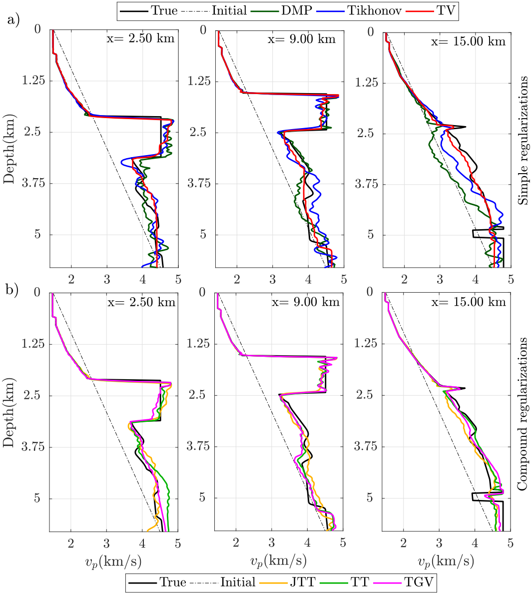
Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, 67(2):301–320.

- **Definition**: A function (here, $\mathbf{b}$) of two variables (here, $\mathbf{u}$ and $\mathbf{m}$) is bilinear if it is linear with respect to each of its variables.
- Illustration with the scalar Helmholtz equation

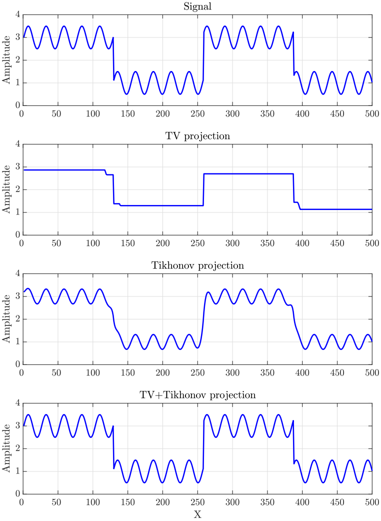
$$\begin{aligned}
&\mathbf{A}(\mathbf{m})\mathbf{u} = \mathbf{b}.\\
&\omega^2 \text{diag}\left(\mathbf{m}\right)\mathbf{u} + \triangle\mathbf{u} = \mathbf{b}.\\
&\omega^2 \text{diag}\left(\mathbf{u}\right)\mathbf{m} + \triangle\mathbf{u} = \mathbf{b}.\\
&\mathbf{L}(\mathbf{u})\mathbf{m} = \mathbf{y},\\
&\text{with } \mathbf{L}(\mathbf{u}) = \omega^2 \text{diag}\left(\mathbf{u}\right) \text{ and } \mathbf{y} = \mathbf{b} - \triangle\mathbf{u}.
\end{aligned} \tag{30}$$

Application to the BP salt model (left target):

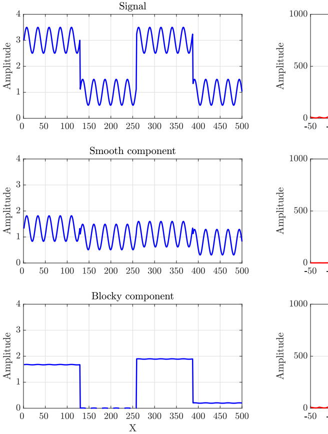
Final models (Logs)



Matching Piecewise smooth function

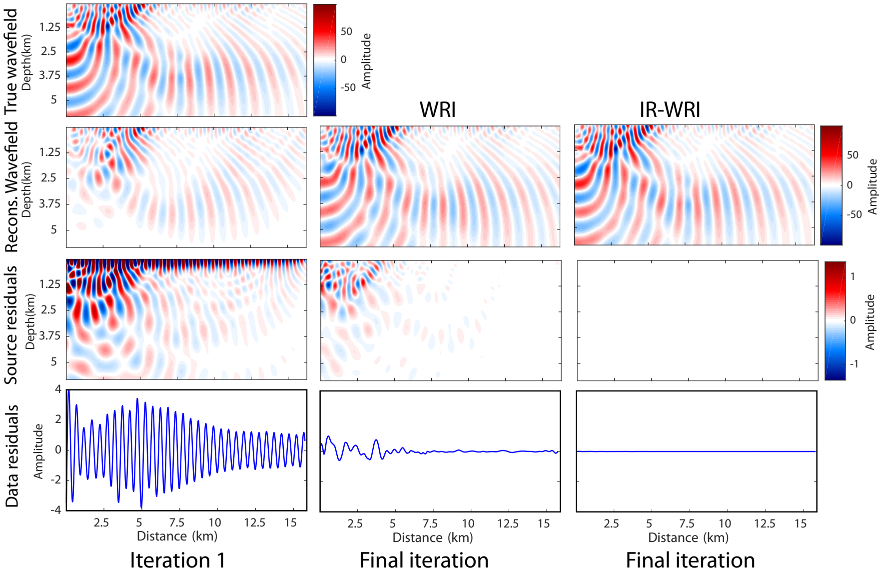


Tikhonov (middle) and TV (bo
of the Tikhonov+TV proj



Wavefield reconstruction Inversion (WRI):

Fitting the data & Satisfying the constraint at convergence point



Joint update of $\mathbf{m}_1$ and $\mathbf{m}_2$

$$
\begin{aligned}
(\mathbf{m}_1^{k+1}, \mathbf{m}_2^{k+1}) &= \arg \min_{\mathbf{m}_1, \mathbf{m}_2} \lambda \|\mathbf{L}(\mathbf{u}^{k+1})[\mathbf{m}_1 + \mathbf{m}_2] - \mathbf{y} - \mathbf{y}^k\|_2^2 + \alpha \|\nabla^2 \mathbf{m}_2\|_2^2 \\
&+ \gamma_b \|\mathbf{q} + \mathbf{q}'^k - \mathbf{m}_1 - \mathbf{m}_2\|_2^2 + \gamma_t \|\mathbf{p} + \mathbf{p}'^k - \nabla \mathbf{m}_1\|_2^2.
\end{aligned}
$$

A joint update of $\mathbf{m}_1$ and $\mathbf{m}_2$ occurs at the point where the derivatives of the functional with respect to them vanish simultaneously. It is then a solution of the following system of equations with two unknowns $\mathbf{m_1}$ and $\mathbf{m_2}$:

$$
\begin{bmatrix} \mathbf{g}_{11} & \mathbf{g}_{12} \\ \mathbf{g}_{12} & \mathbf{g}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix} \tag{31}
$$

$$
\begin{cases}
\mathbf{g}_{11} = \lambda \mathbf{L^T L} + \gamma_t \nabla^T \nabla + \gamma_b \mathbf{I}, \\
\mathbf{g}_{12} = \lambda \mathbf{L^T L} + \gamma_b \mathbf{I}, \\
\mathbf{g}_{22} = \lambda \mathbf{L^T L} + \alpha {\nabla^2}^T \nabla^2 + \gamma_b \mathbf{I}, \\
\mathbf{r}_1 = \lambda \mathbf{L^T} [\mathbf{y} + \mathbf{y}^k] + \gamma_b [\mathbf{q} + \mathbf{q}'^k] + \gamma_t \nabla^T [\mathbf{p} + \mathbf{p}'^k], \\
\mathbf{r}_2 = \lambda \mathbf{L^T} [\mathbf{y} + \mathbf{y}^k] + \gamma_b [\mathbf{q} + \mathbf{q}'^k]
\end{cases} \tag{32}
$$

where $\mathbf{I}$ is the identity matrix.

By using a variable-projection scheme, estimate $\mathbf{m}_2$ as a function of $\mathbf{m}_2$, it is possible to reduce the linear system size that we need to solve.

Solve first equation of 31 for $\mathbf{m}_2$, $\mathbf{m}_2 = \mathbf{g}_{12}^{-1} [\mathbf{r}_1 - \mathbf{g}_{11} \mathbf{m}_1]$, and injecting in the second

$$\mathbf{q}^{k+1} = \arg\min_{\mathbf{q}\in\mathcal{C}} \|\mathbf{q} + \mathbf{q}'^{k} - \mathbf{m}_1^{k+1} - \mathbf{m}_2^{k+1}\|_2^2,$$

has a closed-form solution which is projection into $\mathcal{C}$.

Projection operator $\rightarrow$ $\textbf{proj}_{\mathcal{C}}(\bullet) = \arg\min_{\mathbf{q}\in\mathcal{C}} \|\mathbf{q} - \bullet\|_2^2$ $\rightarrow$ $\textbf{proj}_{\mathcal{C}}(\bullet) = \min(\max(\bullet, \mathbf{m}_{lb}), \mathbf{m}_{ub})$.

Approximates the input point with some other point in the desired set $\mathcal{C}$ which is closest to it in the L2 sense.

Update of $\mathbf{q}^{k+1}$ $\rightarrow$

$$\mathbf{q}^{k+1} = \text{proj}_{\mathcal{C}}(\mathbf{m}_1^{k+1} + \mathbf{m}_2^{k+1} - \mathbf{q}'^{k}) = \min(\max(\mathbf{m}_1^{k+1} + \mathbf{m}_2^{k+1} - \mathbf{q}'^{k}, \mathbf{m}_{lb}), \mathbf{m}_{ub}), \quad (34)$$

Go to main

SEISCOPE

The subproblem for $\mathbf{p}_1$ and $\mathbf{p}_2$ $\rightarrow$

$$\mathbf{p}^{k+1} = \arg\min_{\mathbf{p}} \sum \sqrt{|\mathbf{p}_x|^2 + |\mathbf{p}_z|^2} + \gamma_t \|\mathbf{p} + \mathbf{p'}^k - \nabla \mathbf{m}_1^{k+1}\|_2^2,$$

has a closed-form solution via proximity.

Proximity operators (Combettes and Pesquet, 2011) $\rightarrow$ $\text{prox}_{\gamma}(\bullet) = \arg\min_{\mathbf{p}} \; f(\mathbf{p}) + \gamma\| \bullet - \mathbf{p}\|_2^2$

Approximates the input point with some other point closest to it in the L2 distance sense under regularization implemented with the penalty term $f(\mathbf{p})$.

For homogeneous TV norm $\rightarrow$

$$\text{prox}_{\gamma_t}(\bullet) = \max\left(1 - \frac{1}{\gamma_t \sqrt{|\nabla_x \mathbf{m}_1^{k+1} - \mathbf{p'}_x^k|^2 + |\nabla_z \mathbf{m}_1^{k+1} - \mathbf{p'}_z^k|^2}}, 0\right) \bullet .$$

Update of $\mathbf{p}^{k+1}$ $\rightarrow$

$$\mathbf{p}^{k+1} = \begin{bmatrix} \text{prox}_{\gamma_t}(\nabla_x \mathbf{m}_1 - \mathbf{p'}_x^k) \\ \text{prox}_{\gamma_t}(\nabla_z \mathbf{m}_1 - \mathbf{p'}_z^k) \end{bmatrix} = \begin{bmatrix} \max(1 - \frac{1}{\gamma_t \sqrt{|\nabla_x \mathbf{m}_1^{k+1} - \mathbf{p'}_x^k|^2 + |\nabla_z \mathbf{m}_1^{k+1} - \mathbf{p'}_z^k|^2}}, 0)[\nabla_x \mathbf{m}_1 - \\ \max(1 - \frac{1}{\gamma_t \sqrt{|\nabla_x \mathbf{m}_1^{k+1} - \mathbf{p'}_x^k|^2 + |\nabla_z \mathbf{m}_1^{k+1} - \mathbf{p'}_z^k|^2}}, 0)[\nabla_z \mathbf{m}_1 - \end{bmatrix} \quad (35)$$

Go to main

- Bound constraints

$$\mathbf{q}^{k+1} = \text{proj}_{\mathcal{C}}(\mathbf{m}_1^{k+1} + \mathbf{m}_2^{k+1} - \mathbf{q'}^{k}) \;=\; \min(\max(\mathbf{m}_1^{k+1} + \mathbf{m}_2^{k+1} - \mathbf{q'}^{k}, \mathbf{m}_{lb}), \mathbf{m}_{ub})$$

- TV regularization

$$\mathbf{p}^{k+1} = \begin{bmatrix} \text{prox}_{\gamma_t}(\nabla_x \mathbf{m}_1 - \mathbf{p'}_x^k) \\ \text{prox}_{\gamma_t}(\nabla_z \mathbf{m}_1 - \mathbf{p'}_z^k) \end{bmatrix} = \left[\begin{matrix} \max\left(1 - \frac{1}{\gamma_t \sqrt{|\nabla_x \mathbf{m}_1^{k+1} - \mathbf{p'}_x^k|^2 + |\nabla_z \mathbf{m}_1^{k+1} - \mathbf{p'}_z^k|^2}}, 0\right) [\nabla \\ \max\left(1 - \frac{1}{\gamma_t \sqrt{|\nabla_x \mathbf{m}_1^{k+1} - \mathbf{p'}_x^k|^2 + |\nabla_z \mathbf{m}_1^{k+1} - \mathbf{p'}_z^k|^2}}, 0\right) [\nabla \end{matrix} \right.$$

Implementing hybrid Tikhonov + TV regularization with IR-WRI: Separate update of m_1 and m_2

The least-squares system

$$\mathbf{m}_1^{k+1} = \arg\min_{\mathbf{m}_1} \lambda\|\mathbf{L}(\mathbf{u}^{k+1})[\mathbf{m}_1 + \mathbf{m}_2^k] - \mathbf{y} - \mathbf{y}^k\|_2^2 + \gamma_b\|\mathbf{q} + \mathbf{q}'^k - \mathbf{m}_1 - \mathbf{m}_2^k\|_2^2 + \gamma_t\|\mathbf{p} +$$

has a closed-form solution as

$$\mathbf{m}_1^{k+1} = \left[\lambda\mathbf{L}^T\mathbf{L}+\gamma_t\nabla^T\nabla+\gamma_b\mathbf{I}\right]^{-1}\left[\lambda\mathbf{L}^T[\mathbf{y}+\mathbf{y}^k-\mathbf{L}\mathbf{m}_2^k]+\gamma_t\nabla^T[\mathbf{p}^k+\mathbf{q}^k]+\gamma_b[\mathbf{q}^k+\mathbf{q}'^k-\mathbf{m}_2^k]\right]. \quad (36)$$

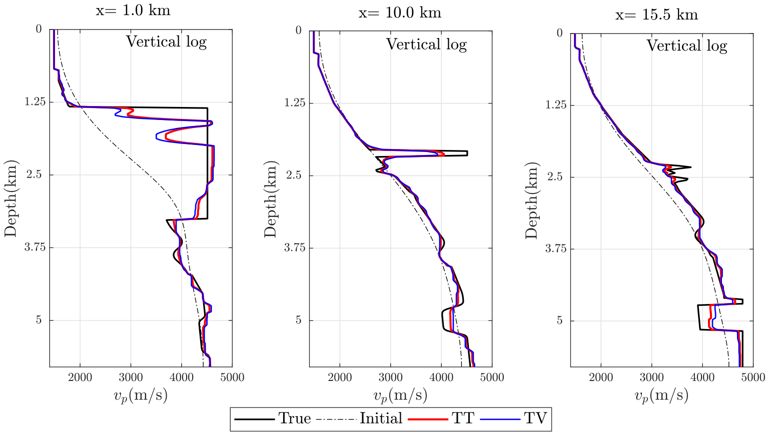
Also, the least-squares system of $\mathbf{m}_2$

$$\mathbf{m}_2^{k+1} = \arg\min_{\mathbf{m}_2} \lambda\|\mathbf{L}(\mathbf{u}^{k+1})[\mathbf{m}_2+\mathbf{m}_1^{k+1}]-\mathbf{y}-\mathbf{y}^k\|_2^2+\alpha\|\nabla^2\mathbf{m}_2\|_2^2+\gamma_b\|\mathbf{q} + \mathbf{q}'^k-\mathbf{m}_2-\mathbf{m}_1^{k+1}\|_2^2,$$

has a closed-form solution as

$$\mathbf{m}_2^{k+1} = \left[\lambda\mathbf{L}^T\mathbf{L}+\alpha{\nabla^2}^T\nabla^2+\gamma_b\mathbf{I}\right]^{-1}\left[\lambda\mathbf{L}^T[\mathbf{y}+\mathbf{y}^k-\mathbf{L}\mathbf{m}_1^{k+1}]+\gamma_b[\mathbf{q}^k+\mathbf{q}'^k-\mathbf{m}_1^{k+1}]\right]. \quad (37)$$

Go to main



x= 1.0 km
x= 10.0 km
x= 15.5 km
Vertical log
Vertical log
Vertical log
Depth(km)
0
1.25
2.5
3.75
5
2000
3000
4000
5000
v_p(m/s)
True
Initial
TT
TV

TT vs TV regularized IR-WRI