Compound Regularization of Full-waveform Inversion for Imaging Piecewise Media

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

[1] Univ. of Tehran, Institute of Geophysics, Tehran, Iran
[2] Univ. Côte d'Azur, CNRS (IRD), Géoazur, Valbonne, France,

Workshop on Machine-Learning-Assisted Image Formation
10-12 Jul 2019 Nice (France)

What is Full waveform inversion (FWI)?

FWI → Build high resolution earth parameters models (velocity, density, anisotropy, quality factor) with recorded seismograms (Virieux and Operto, 2009).

What is Full waveform inversion (FWI)?

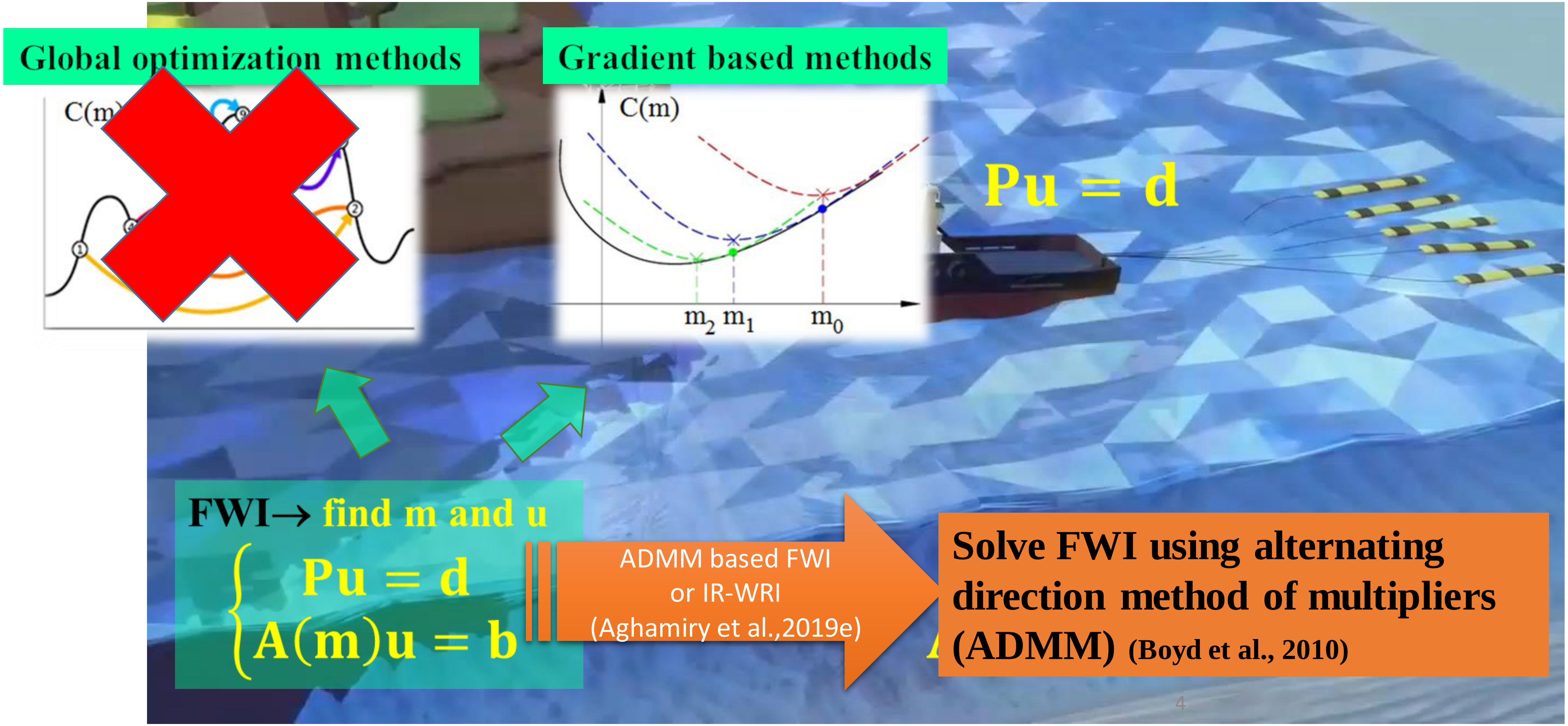
FWI → Build **high resolution earth parameters** models (velocity, density, anisotropy, quality factor) **with recorded seismograms** (Virieux and Operto, 2009).

	Seismic exploration	Medical ultrasound (soft tissues/cortical bone)
Wave-speed	3000 m/s	1540 / 3500 m/s
Frequency	2.5-25 Hz	1-5 MHZ
Investigated depth	10 Km	150 mm in soft tissues
Attenuation	0.5 dB per wavelength	0.1 / 5 dB per wavelength

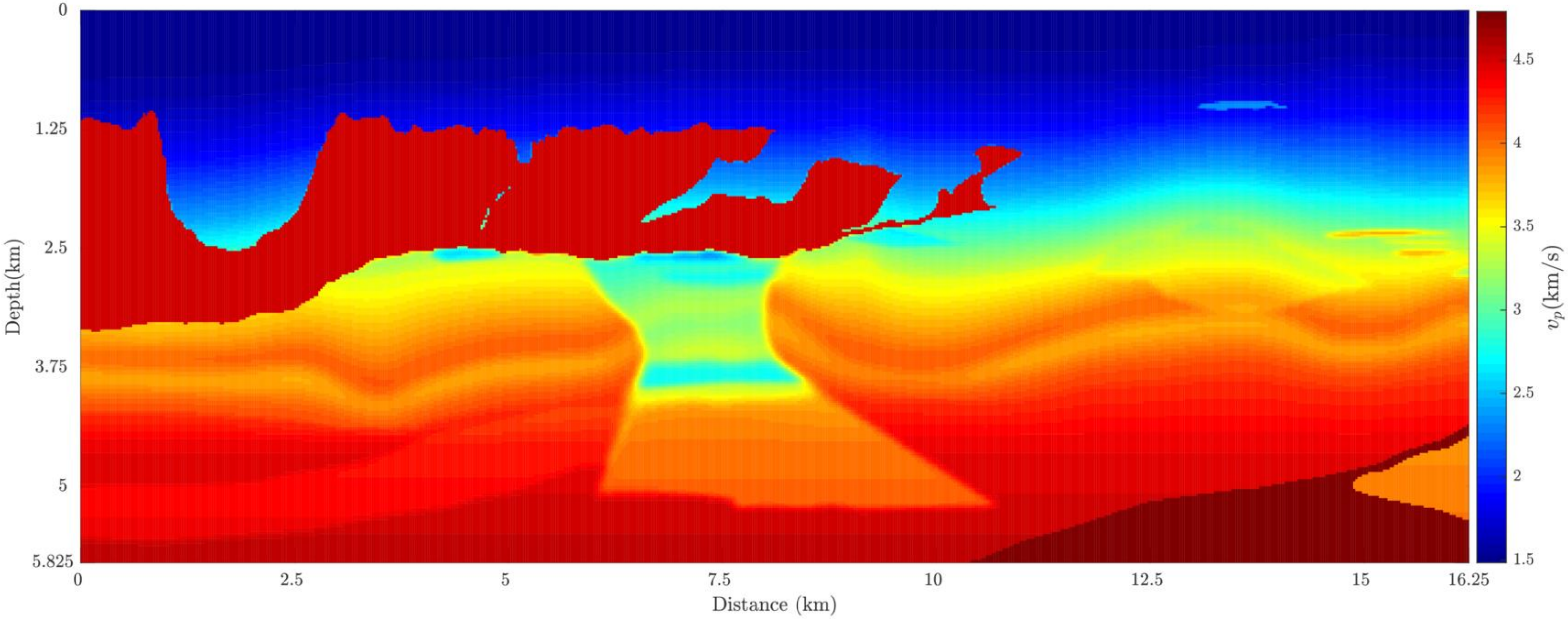
FWI from mathematical point of view

Unknowns $\rightarrow \begin{cases} u \rightarrow \text{wavefield} \\ m \rightarrow \text{model parameters} \end{cases}$

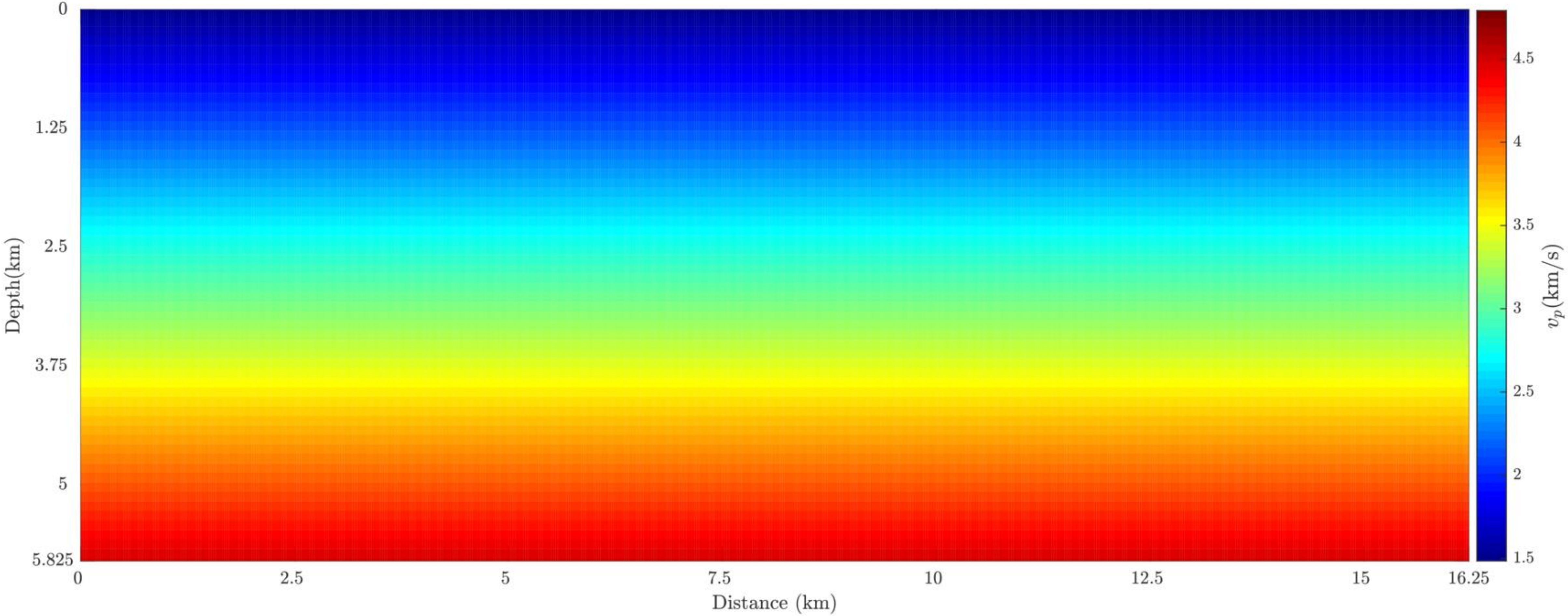
Knowns $\rightarrow \begin{cases} d \rightarrow \text{recorded data} \\ b \rightarrow \text{source} \end{cases}$



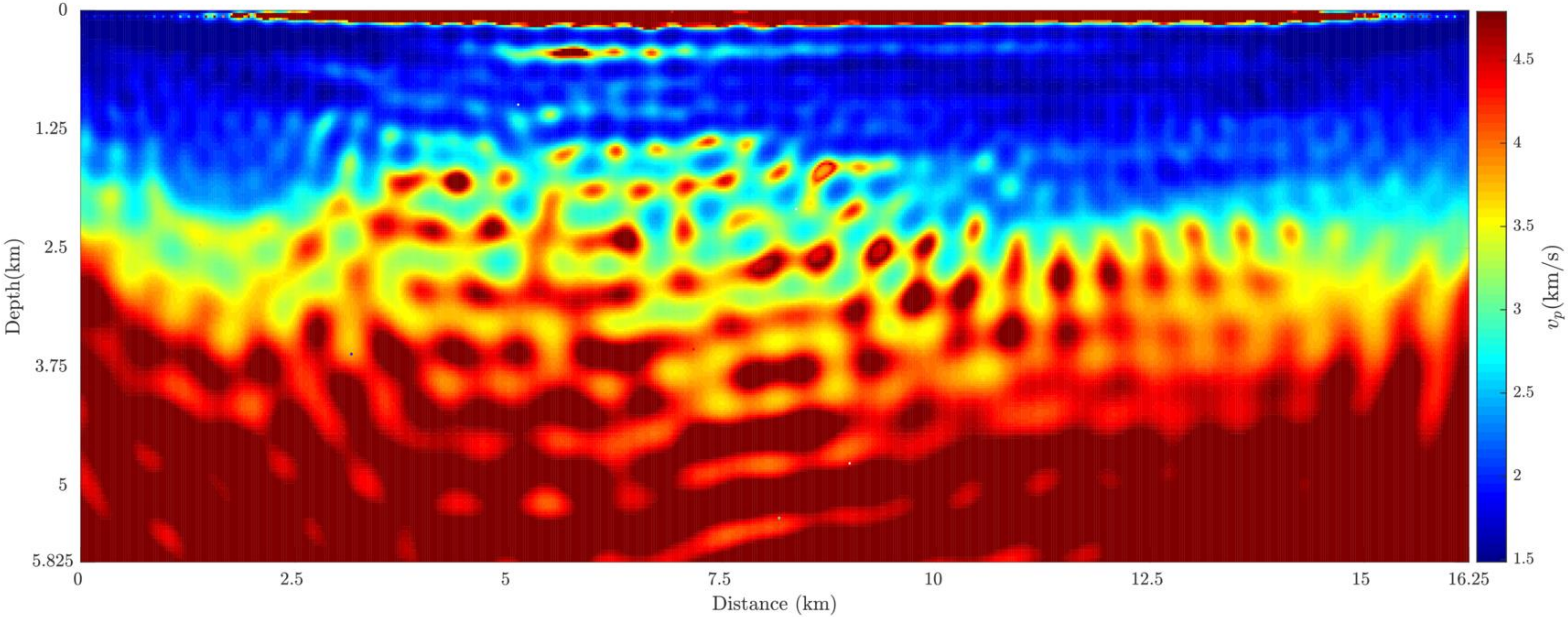
2004 BP model: *True model*

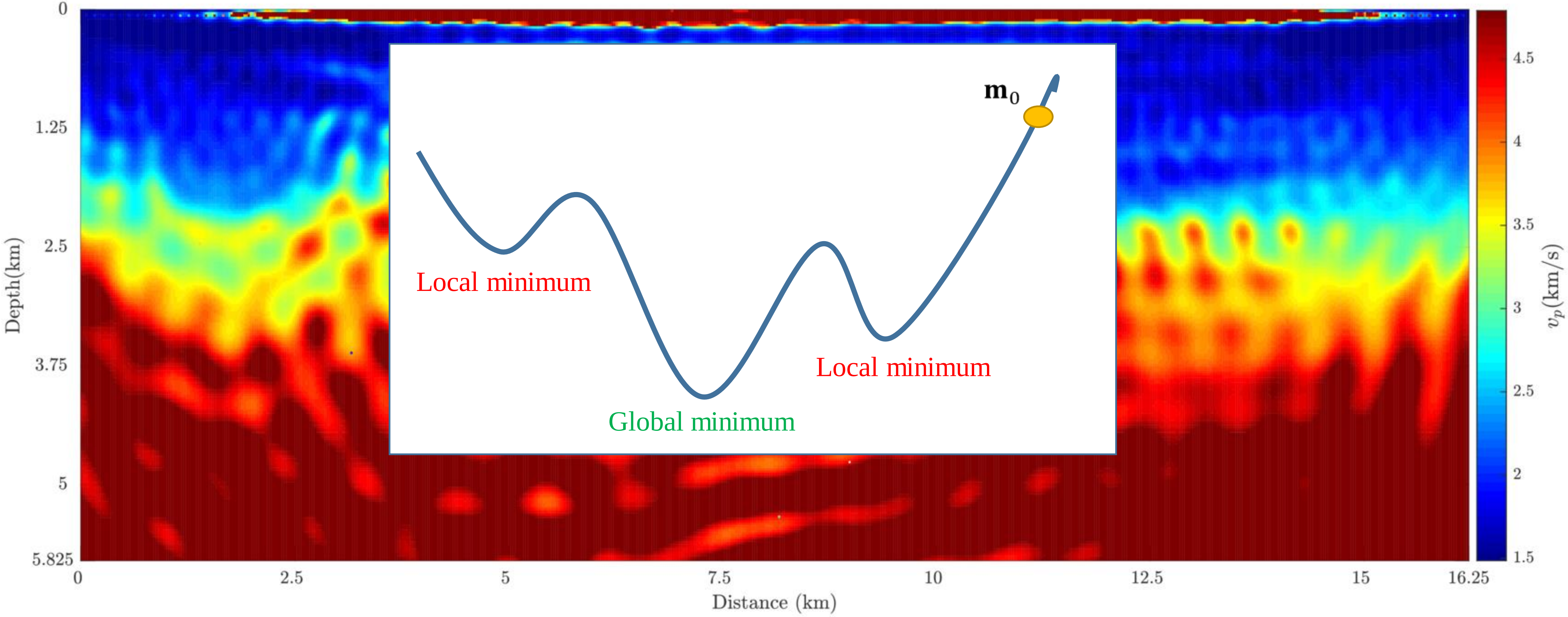


Initial model

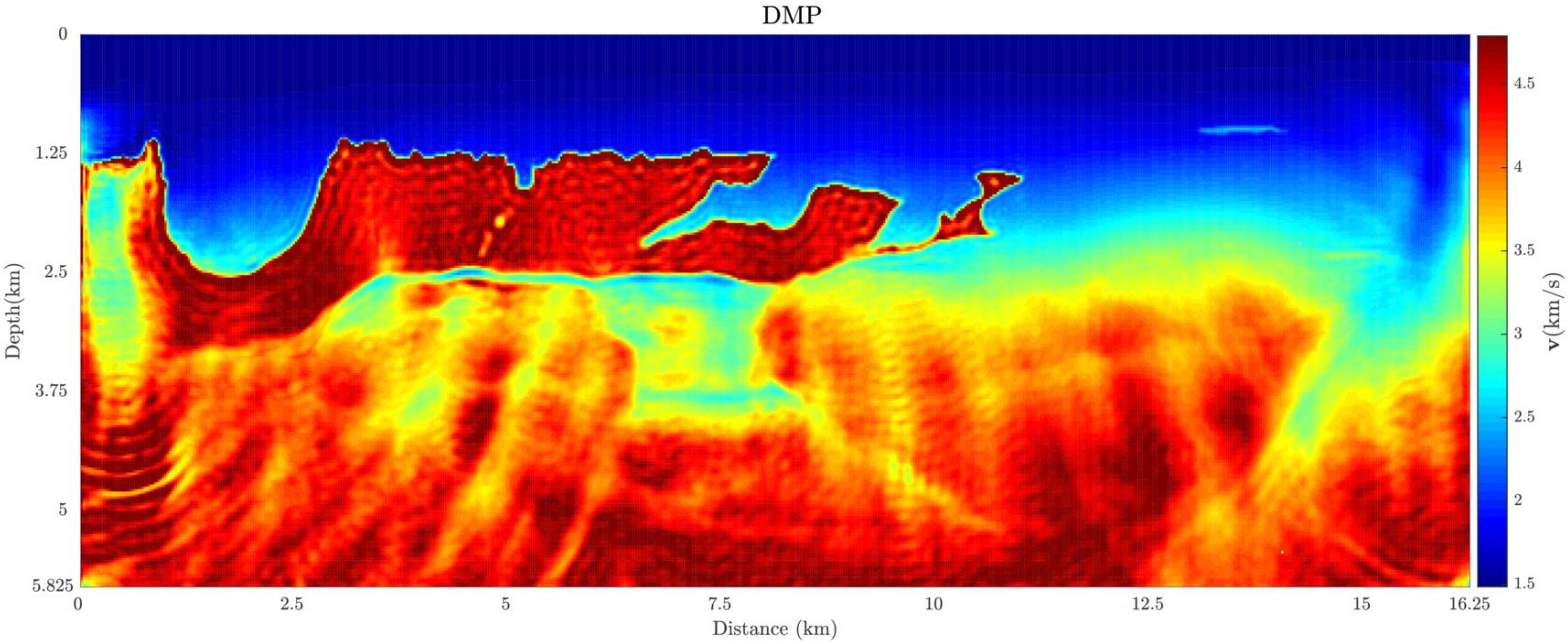


Classical approach: *Estimated model*

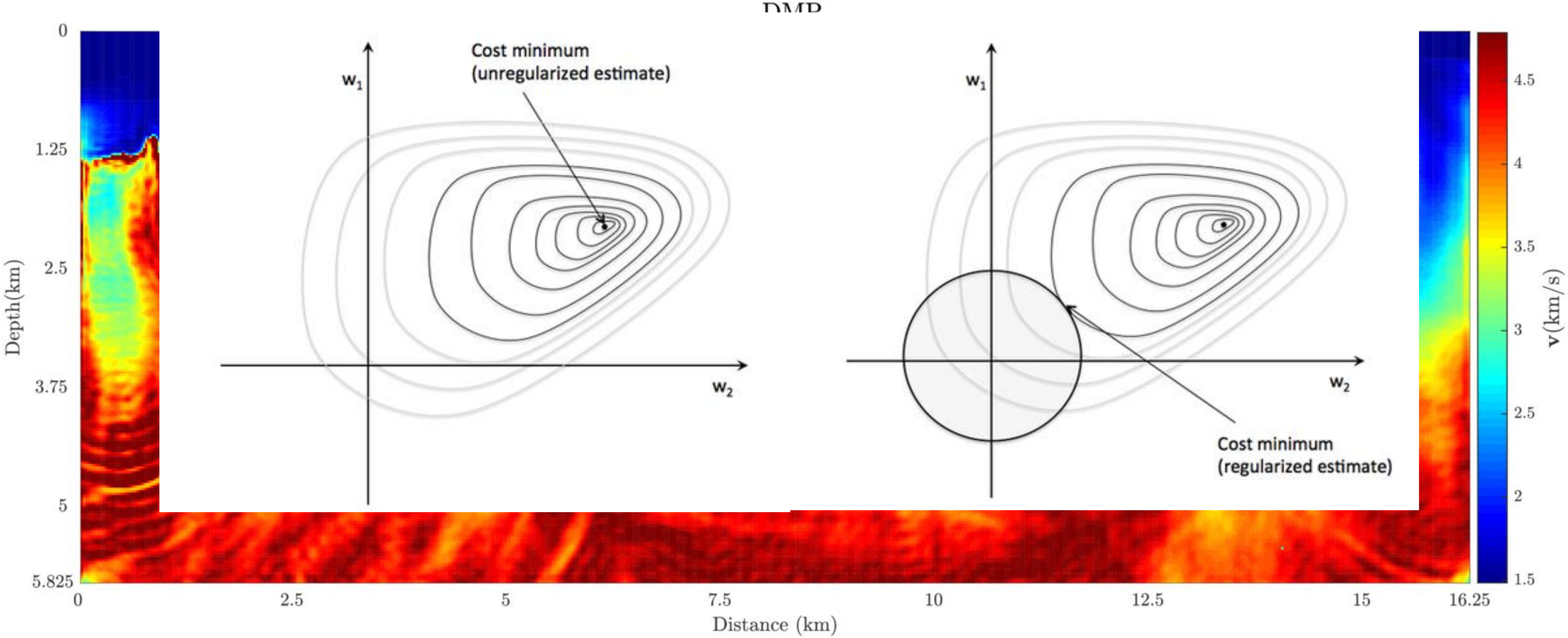




0
1.25
2.5
3.75
5
5.825
Depth(km)
0
2.5
5
7.5
10
12.5
15
16.25
Distance (km)
4.5
4
3.5
3
2.5
2
1.5
v_p(km/s)
$\mathbf{m}_0$
Local minimum
Global minimum
Local minimum


DMP
0
1.25
2.5
3.75
5
5.825
Depth(km)
0
2.5
5
7.5
10
12.5
15
16.25
Distance (km)
4.5
4
3.5
3
2.5
2
1.5
v(km/s)

IR-WRI: *Estimated model*



- Due to the ill-posedness of FWI, regularization is necessary.
- It is used to inject prior knowledge and statistical properties to the inversion.
- But, how to choose regularization?

Two popular regularizations in geophysics and image denoising:

1. Second-order Tikhonov regularization

$$\|\mathbf{m}\|_{Tikh2} = \sum \|\nabla_x^2 \mathbf{m}\|_2^2 + \|\nabla_z^2 \mathbf{m}\|_2^2.$$

   Drive inversion toward smooth reconstruction.

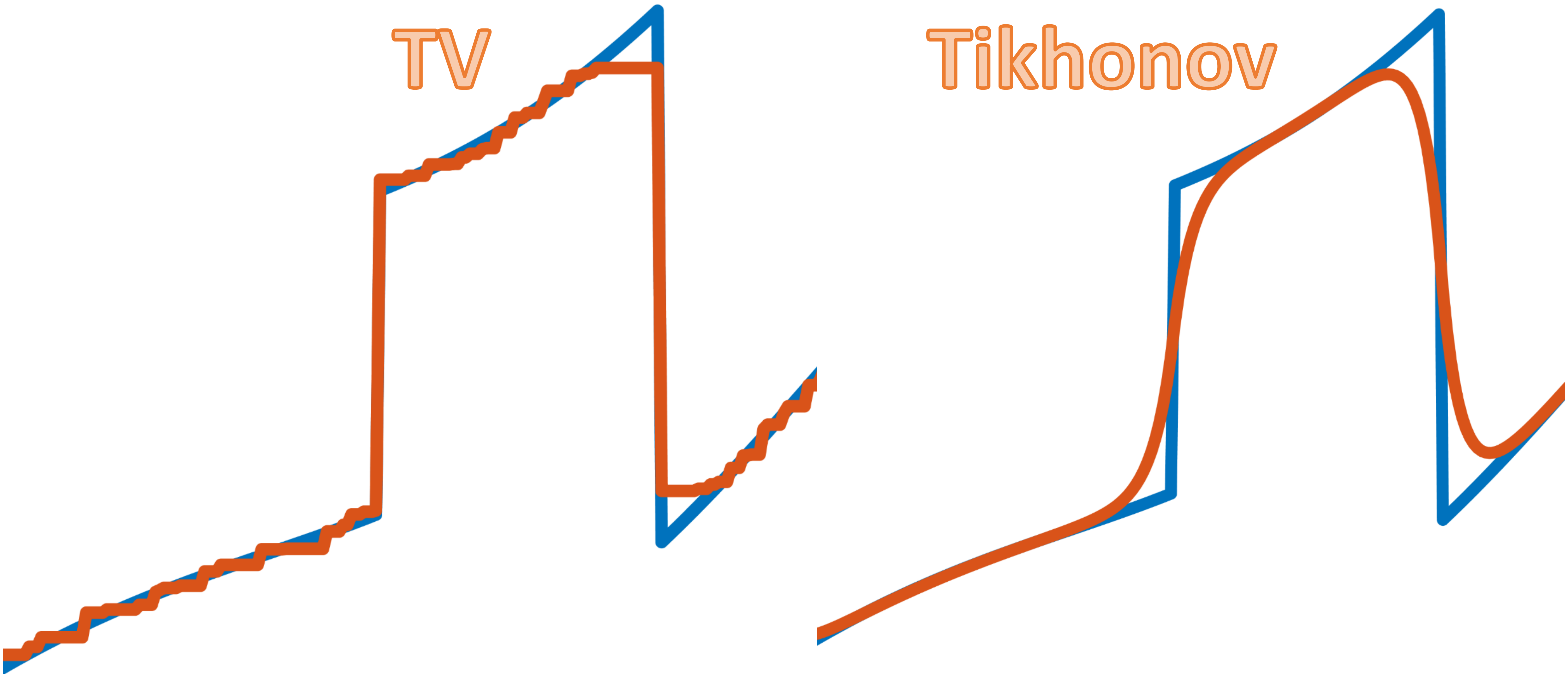
2. Blockiness-promoting Isotropic Total Variation (TV) regularization

$$\|\mathbf{m}\|_{TV} = \sum \sqrt{|\nabla_x \mathbf{m}|^2 + |\nabla_z \mathbf{m}|^2}.$$

   Drive inversion toward piecewise homogeneous (blocky) reconstruction.

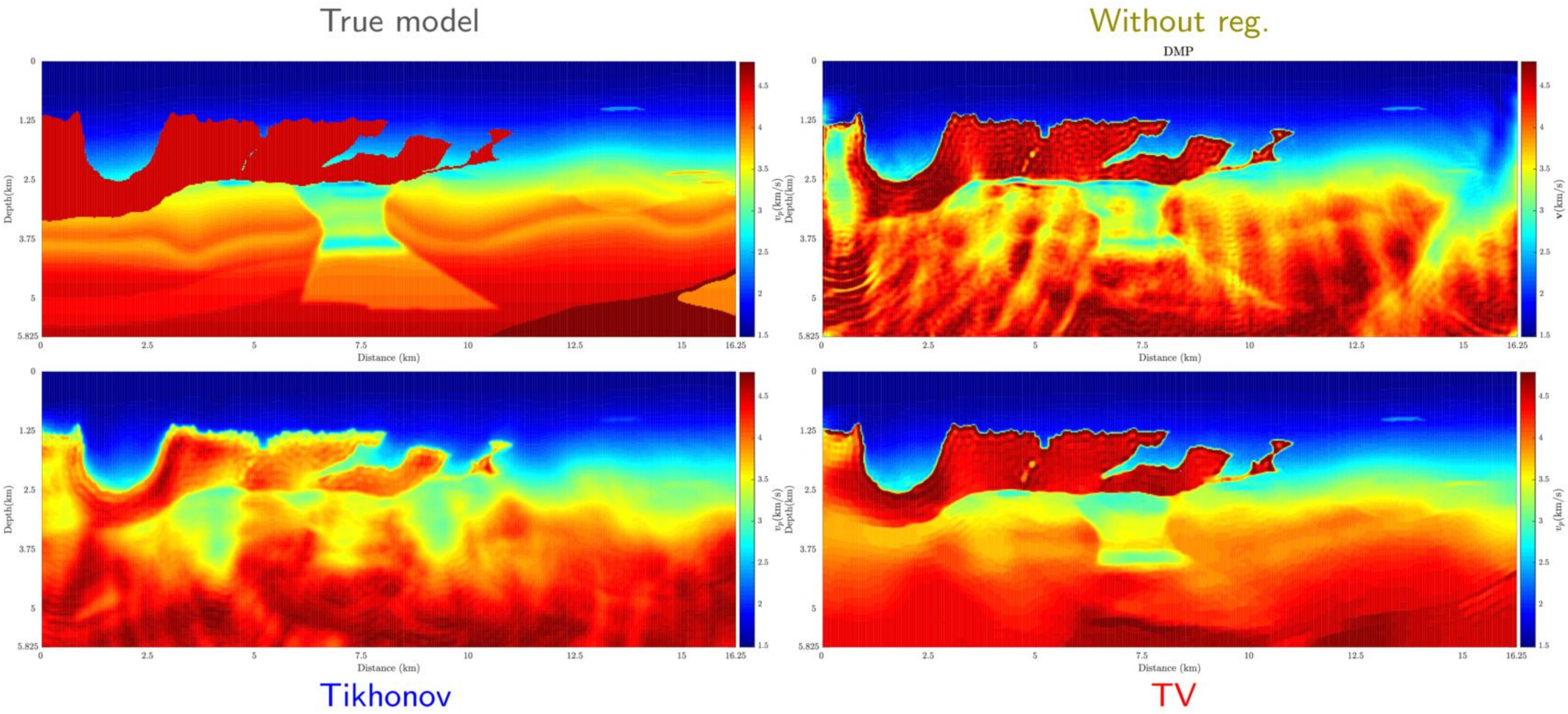
∇_i and ∇_i^2: first and second-order difference operators in the i direction ($i \in \{x, z\}$).

1D inversion test: *What is the shortage of Tikhonov and TV?*

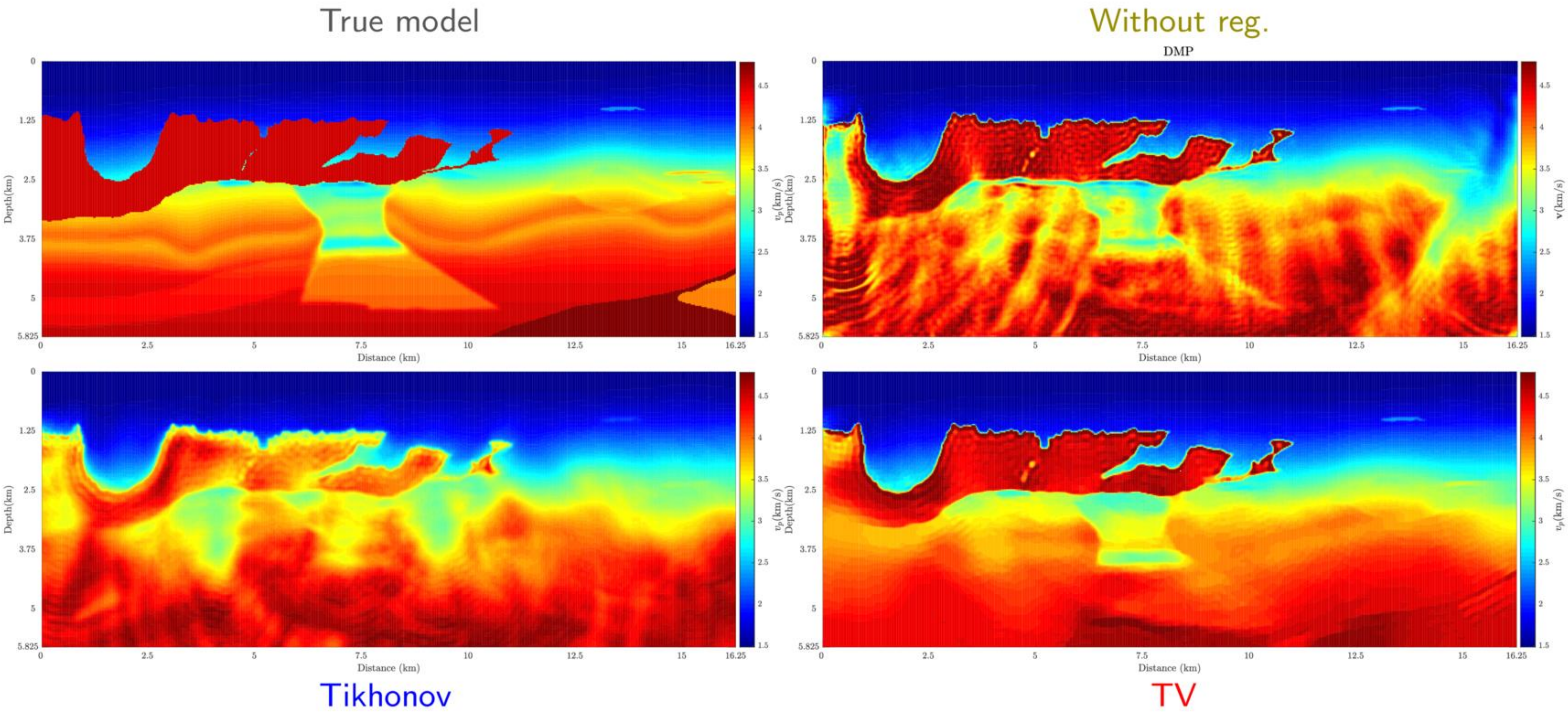


When the model become complicated, they can't recover all the elements.

2D inversion test: *What is the shortage of Tikhonov and TV?*



2D inversion test: *What is the shortage of Tikhonov and TV?*



How we can reconstruct such complicated models?

Either we can use more complicated regularization functions or combine the simple regularizations.

Compound regularization: *(Benning and Burger, 2018; Aghamiry et al., 2019b)*

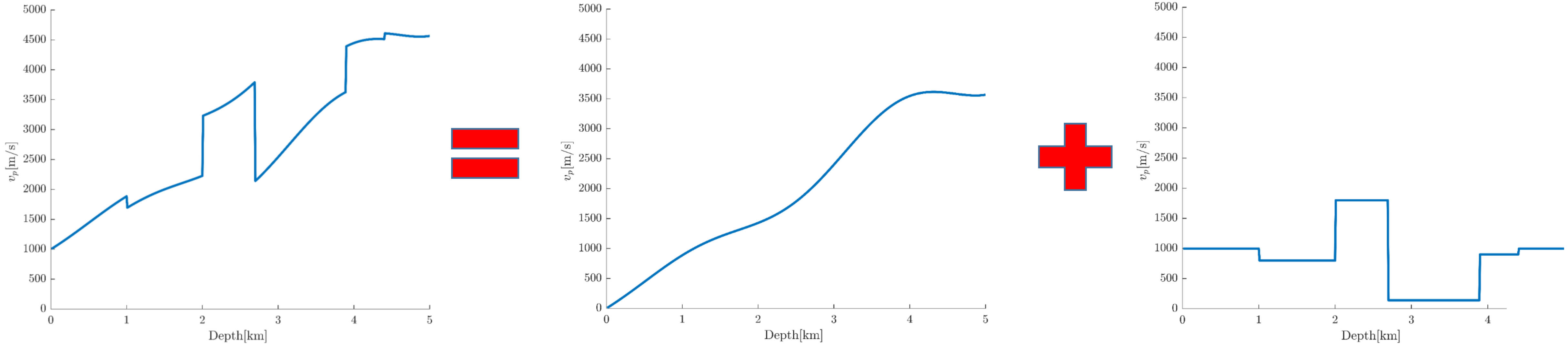
Compound regularizers are constructed by combining two or more separate simple regularizers.

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

$$\mathrm{Reg}(\mathbf{m}) = \alpha_1 \Phi_1(\mathbf{m}) + \ldots + \alpha_r \Phi_r(\mathbf{m}), \tag{1}$$

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

$$\mathrm{Reg}(\mathbf{m}) = \min \{\alpha_1 \Phi_1(\mathbf{m}_1) + \ldots + \alpha_r \Phi_r(\mathbf{m}_r)\}. \tag{2}$$



Compound regularization:

Geometrical illustration of l1, l2, their CC and IC regularizations (Aghamiry et al., 2019b)

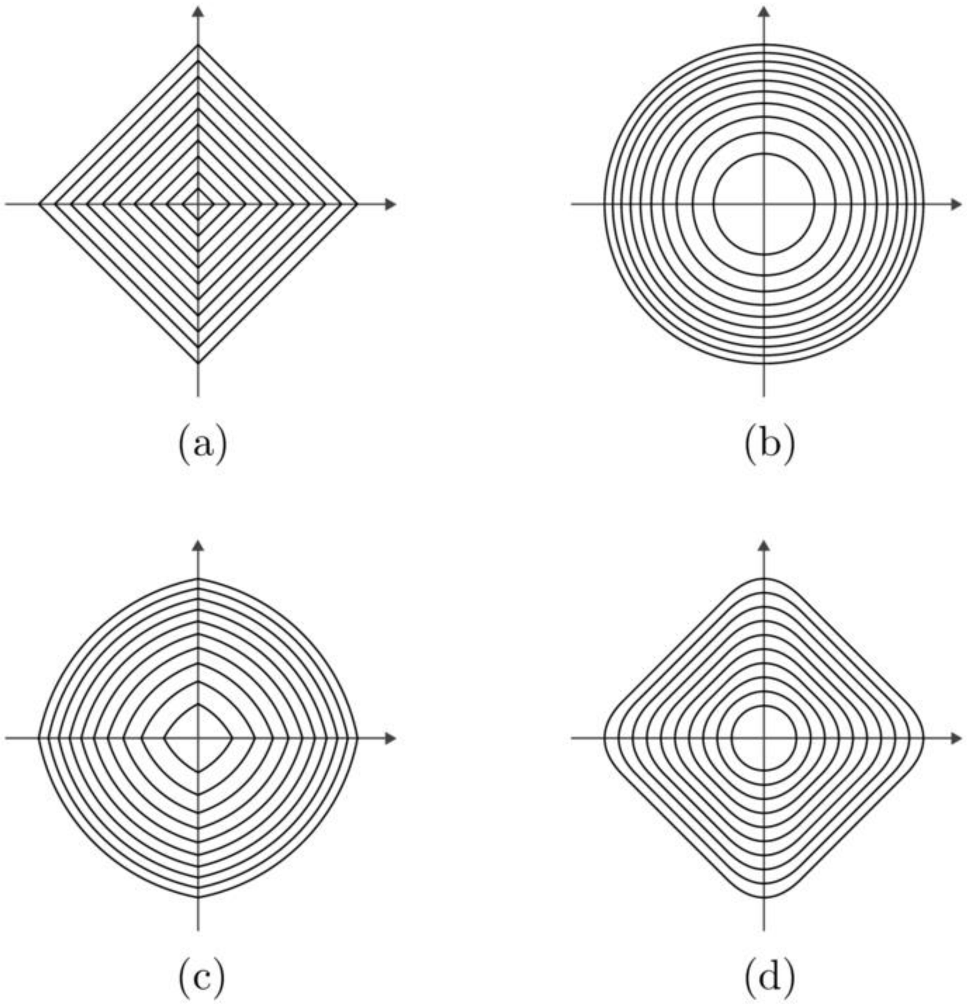


Figure 1: (a) the ℓ_1-norm, (b) the ℓ_2-norm, (c) the CC of ℓ_1 and ℓ_2-norm, and (d) IC of ℓ_1

Geometrical illustration of l1, l2, their CC and IC regularizations (Aghamiry et al., 2019b)

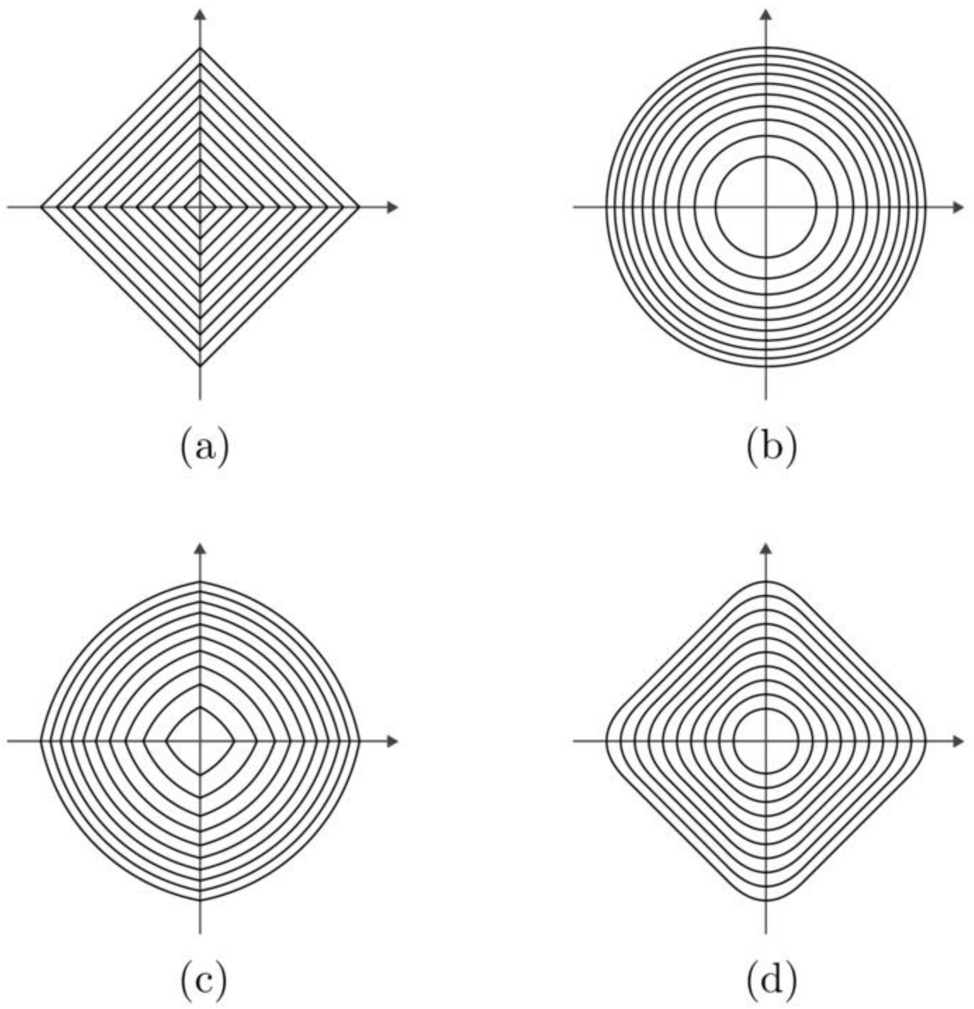


Figure 1: (a) the ℓ_1-norm, (b) the ℓ_2-norm, (c) the CC of ℓ_1 and ℓ_2-norm, and (d) IC of ℓ_1

What do we want to do?

We are going to apply combination of regularizations in the framework of ADMM-based FWI or IR-WRI (Aghamiry et al., 2019e).

Of course at the global minimum $\mathbf{u}$ and $\mathbf{m}$ satisfy the wave-equation as well as the observation equation. FWI with $\text{Reg}(\mathbf{m})$ as a regularization reads

$$\min_{\mathbf{m},\mathbf{u}} \quad \text{Reg}(\mathbf{m}) \quad \text{subject to} \quad \begin{cases} \mathbf{A}(\mathbf{m})\mathbf{u} = \mathbf{b} \\ \mathbf{P}\mathbf{u} = \mathbf{d} \end{cases} \tag{4}$$

- **IR-WRI** → Solving for primals and duals in alternating mode with ADMM.

Compound regularized IR-WRI: *(Aghamiry et al., 2018, 2019b)*

We test the following regularizers

- **Joint Tikhonov-TV regularizer (JTT)**: CC of second-order Tikhonov and first-order TV

$$\text{Reg}(\mathbf{m}) = \alpha_1 \|\mathbf{m}\|_{TV} + \alpha_2 \|\mathbf{m}\|_{Tikh2}. \tag{7}$$

- **Tikhonov-TV regularizer (TT)**: IC of second-order Tikhonov and first-order TV (Gholami and Hosseini, 2013)

$$\text{Reg}(\mathbf{m}) = \min_{\mathbf{m}=\mathbf{m}_1+\mathbf{m}_2} \{\alpha_1 \|\mathbf{m}_1\|_{TV} + \alpha_2 \|\mathbf{m}_2\|_{Tikh2}\}. \tag{8}$$

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

$$\text{Reg}(\mathbf{m}) = \min_{\mathbf{m}=\mathbf{m}_1+\mathbf{m}_2} \{\alpha_1 \|\mathbf{m}_1\|_{TV} + \alpha_2 \|\mathbf{m}_2\|_{TV2}\}. \tag{9}$$

TT regularized and bound constrained parameter estimation subproblem

$$\mathbf{m}^{k+1} = \underset{\substack{\mathbf{m}=\mathbf{m}_1+\mathbf{m}_2 \\ \underbrace{\mathbf{m} \in \mathcal{C}}_{\text{bound constraints}}}}{\arg\min} \quad \underbrace{\alpha_1 \|\mathbf{m}_1\|_{TV} + \alpha_2 \|\mathbf{m}_2\|_{Tikh2}}_{\text{TT regularization}} + \lambda_1 \|\mathbf{L}\mathbf{m} - \mathbf{y}\|_2^2, \qquad (10)$$

where $\mathbf{m} \in \mathcal{C} \rightarrow$ a bound on the lower and upper bound of the model $\rightarrow \{ \underbrace{\mathbf{m}_l}_{\text{lower bound}} \leq \mathbf{m} \leq \underbrace{\mathbf{m}_u}_{\text{upper bound}} \}$.

Splitting to solve problem 10

Defining auxiliary variables $\rightarrow$ $\mathbf{q}$ for bound constraint and $\mathbf{p}$ for TV regularization.

$$\min_{\mathbf{m}_1,\mathbf{m}_2,\mathbf{q},\mathbf{p}} \alpha_1 \|\mathbf{p}\|_{TV} + \alpha_2 \|\mathbf{m}_2\|_{Tikh2} + \lambda_1 \|\mathbf{L}[\mathbf{m}_1 + \mathbf{m}_2] - \mathbf{y}\|_2^2 \quad \text{subject to} \quad \begin{cases} \mathbf{p} = \nabla \mathbf{m}_1 \\ \mathbf{q} = \mathbf{m}_1 + \mathbf{m}_2 \\ \mathbf{m}_l \leq \mathbf{q} \leq \mathbf{m}_u \end{cases} \qquad (11)$$

where $\nabla = \begin{bmatrix} \nabla_x \\ \nabla_z \end{bmatrix}$.

Then, we build augmented Lagrangian and use ADMM to solve it.

ADMM to solve problem 11

1. For $\mathbf{m} \rightarrow \min_{\mathbf{m}_1,\mathbf{m}_2} \alpha_2\|\mathbf{m}_2\|_{Tikh2} + \lambda_1\|\mathbf{L}[\mathbf{m}_1+\mathbf{m}_2]-\mathbf{y}\|_2^2 + \gamma_0\|\mathbf{m}_1+\mathbf{m}_2-\mathbf{q}-\hat{\mathbf{q}}\|_2^2 + \gamma_1\|\nabla\mathbf{m}_1-\mathbf{p}-\hat{\mathbf{p}}\|_2^2$,

   $\mathbf{m}_1$ and $\mathbf{m}_2$ can be updated simultaneously using a variable projection scheme without extra computational burden compare to TV regularization (Aghamiry et al., 2019d).

2. For $\mathbf{q}$ (a projection problem) $\rightarrow \min_{\mathbf{q}\in\mathcal{C}} \|\mathbf{m}_1+\mathbf{m}_2-\mathbf{q}-\hat{\mathbf{q}}\|_2^2 = \min(\max(\mathbf{m}_1+\mathbf{m}_2-\hat{\mathbf{q}},\mathbf{m}_l),\mathbf{m}_u)$,

3. For $\mathbf{p}$ (a proximity problem) $\rightarrow \min_{\mathbf{p}} \alpha_1\|\mathbf{p}\|_{TV} + \gamma_1\|\nabla\mathbf{m}_1-\mathbf{p}-\hat{\mathbf{p}}\|_2^2 = \text{prox}_{\gamma_1/\alpha_1}(\nabla\mathbf{m}_1-\hat{\mathbf{p}})$,

4. For $\hat{\mathbf{p}}\&\hat{\mathbf{q}}$ (dual ascent) $\rightarrow \begin{cases}\hat{\mathbf{q}} \leftarrow \hat{\mathbf{q}}+\mathbf{q}-\mathbf{m}_1-\mathbf{m}_2,\\ \hat{\mathbf{p}} \leftarrow \hat{\mathbf{p}}+\mathbf{p}-\nabla\mathbf{m}_1,\end{cases}$

γ_0, γ_1 and γ_2 are penalty parameters.

Penalty param.	α_1 & α_2	λ_1	λ_2	γ_1	γ_0
Constraints	Compound reg. weight	Obs. Eq.	Wave Eq.	TV weight	Bounds

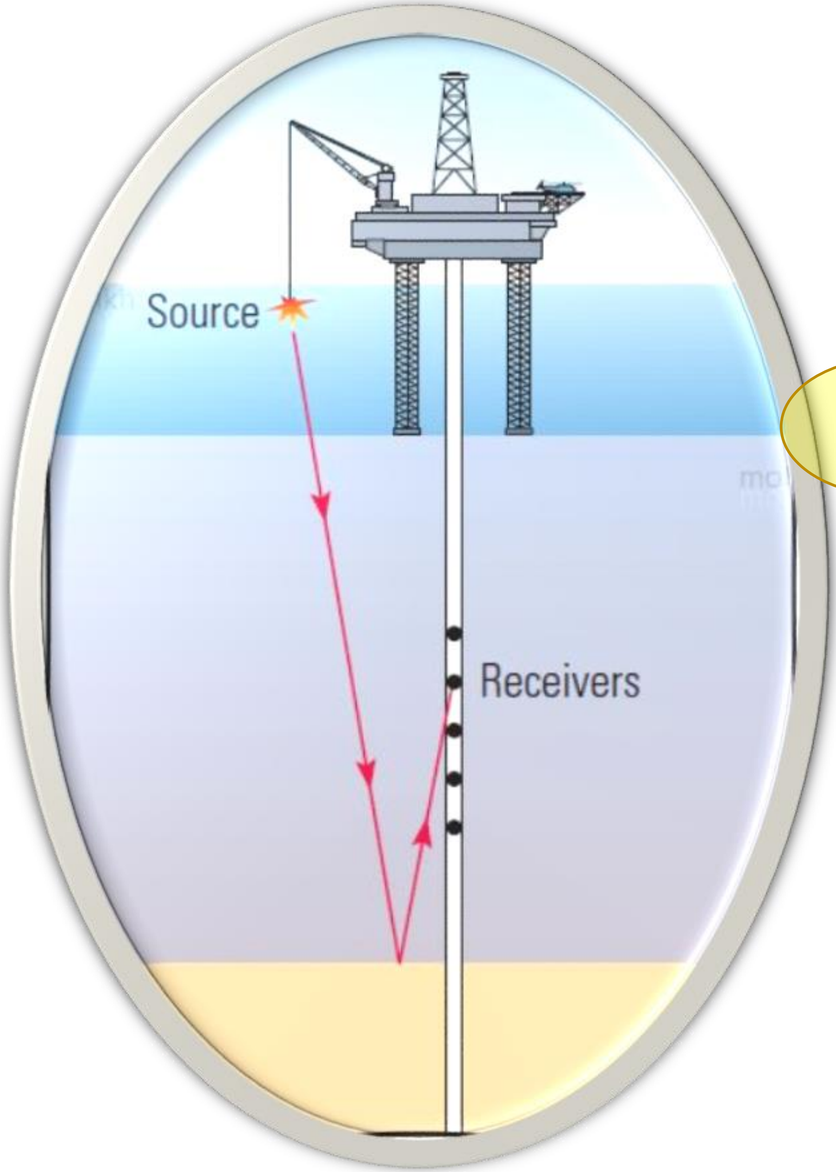
Table 1: α_1 & α_2: balance different regularization functions. λ_1, λ_2, γ_1, γ_0: weights of the observation equation, wave equation, auxiliary TV term,bound constraint .wrt. regularization term, respectively.

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

- We found that α_1=0.7 & α_2=0.3 was a good pragmatical value.
- We use $\gamma_0 = \gamma_1$.
- We found that $\gamma_1/\alpha_1 = 2\% \max \sqrt{|\nabla_x \mathbf{m}_1 - \hat{\mathbf{p}}_1|^2 + |\nabla_z \mathbf{m}_1 - \hat{\mathbf{p}}_2|^2}$ was a good pragmatical value.
- γ_1/λ_1: small percentage of mean absolute value of the diagonal coefficients of $\mathbf{L}^T\mathbf{L}$.
- λ_1/λ_2: 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., 2019e)

Zero offset VSP test:

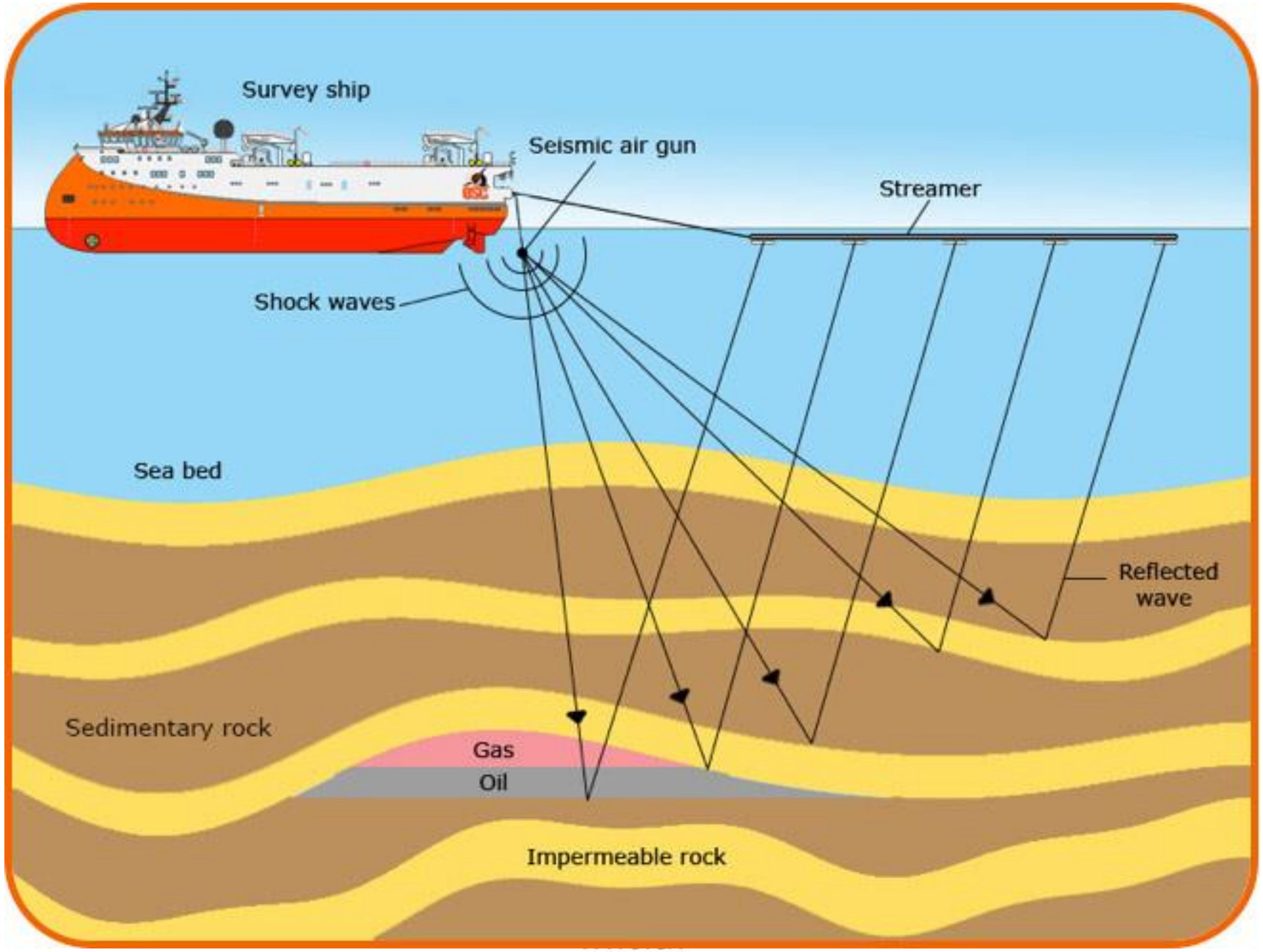
IR-WRI with Noiseless data - 100 iterations



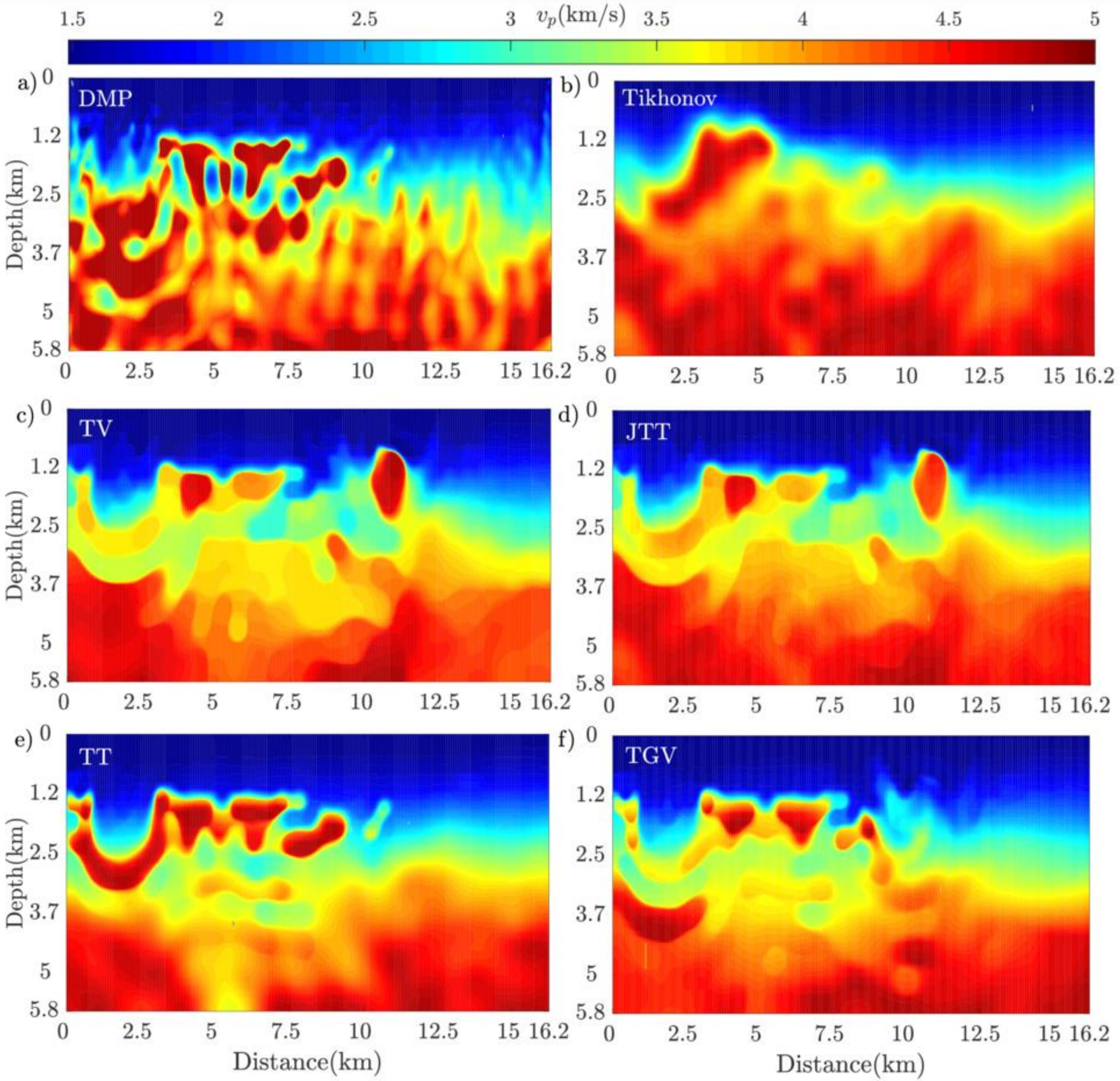
- Fixed-spread surface acquisition.
- Frequency bandwidth: 3-13 Hz.
- Frequency continuation: Batches of 3 frequencies with a 0.5Hz spacing.
- Three paths over batches.
- Noiseless data.
- Stopping criterion of iteration:

$$k_{max} = 15 \quad \text{or}$$
$$(\|\mathbf{A}(\mathbf{m})\mathbf{u} - \mathbf{b}\|_F \leq 1e-3 \quad \text{and}$$
$$\|\mathbf{P}\mathbf{u} - \mathbf{d}\|_F \leq 1e-5),$$

where k_{max} is the maximum number of iterations.

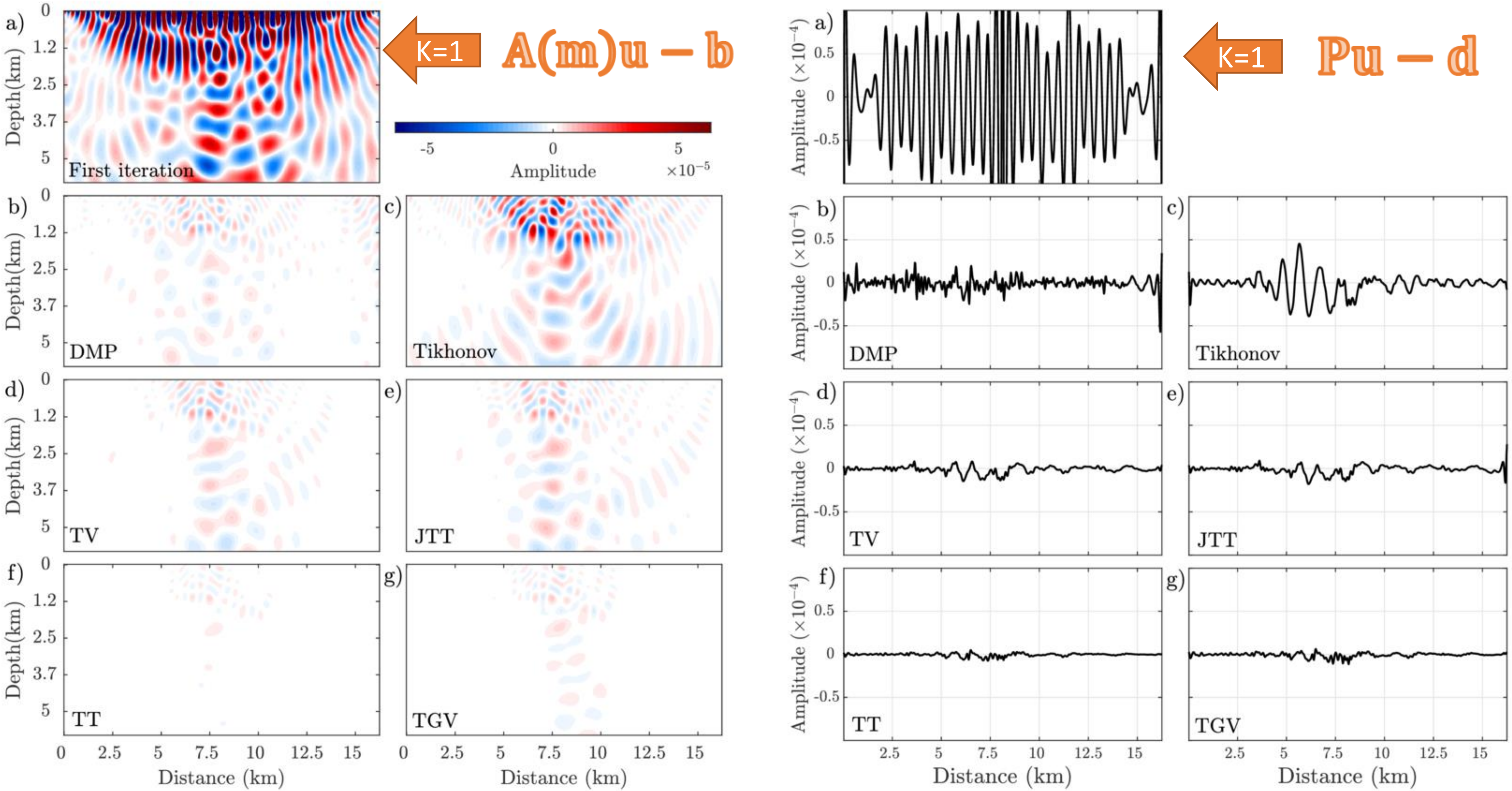


Regularized IR-WRI: *First frequency batch after 45 iterations*

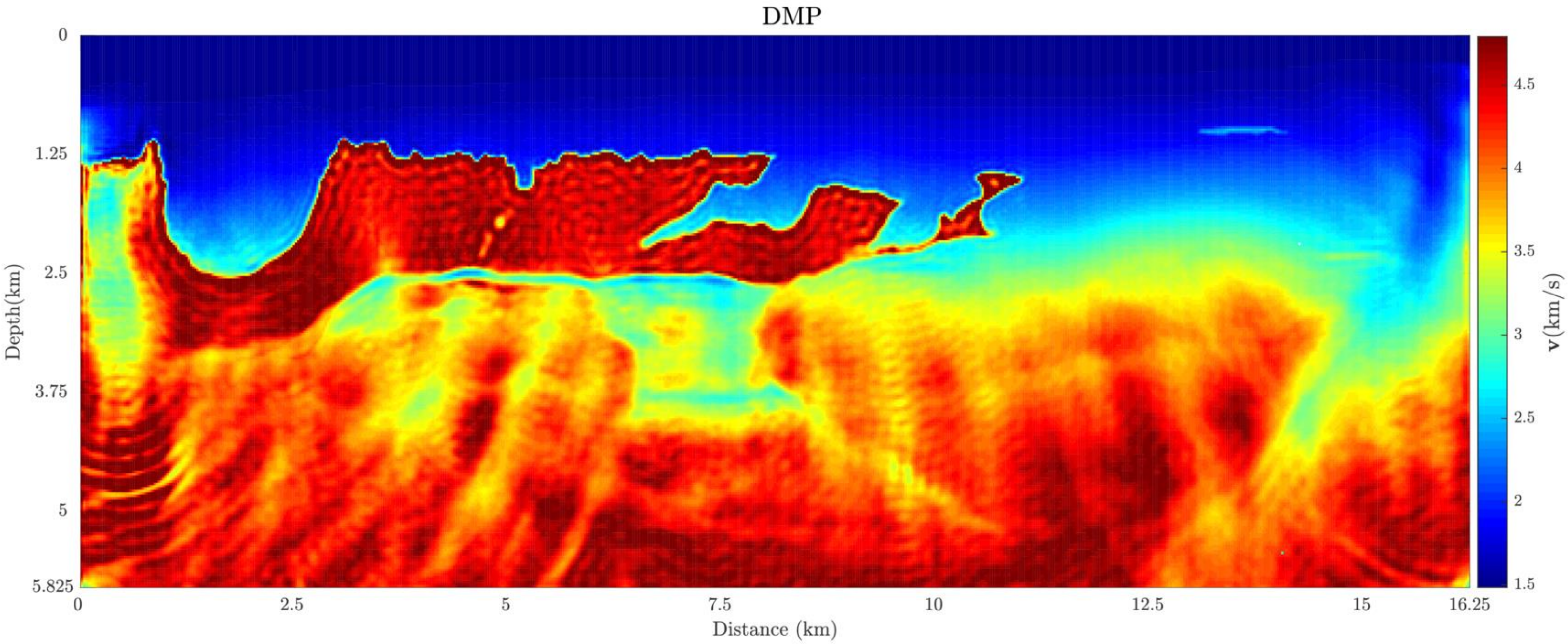


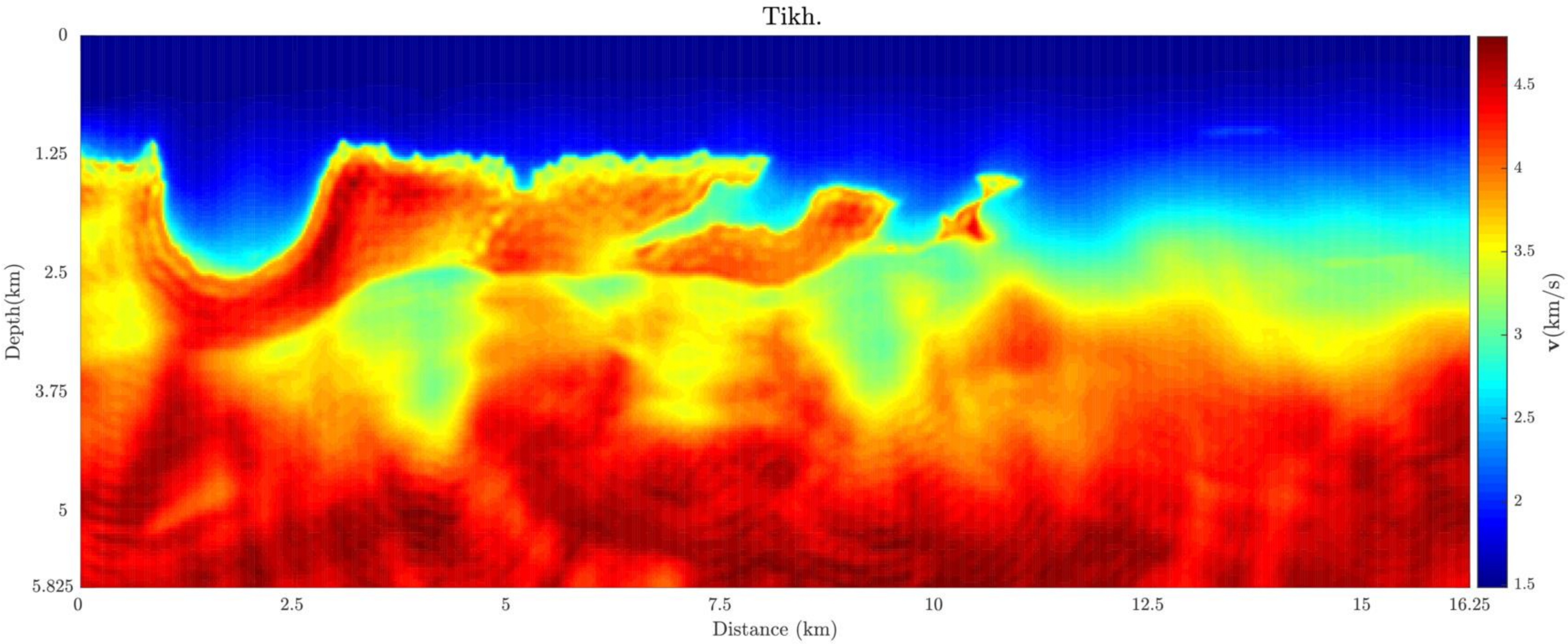
Regularized IR-WRI:

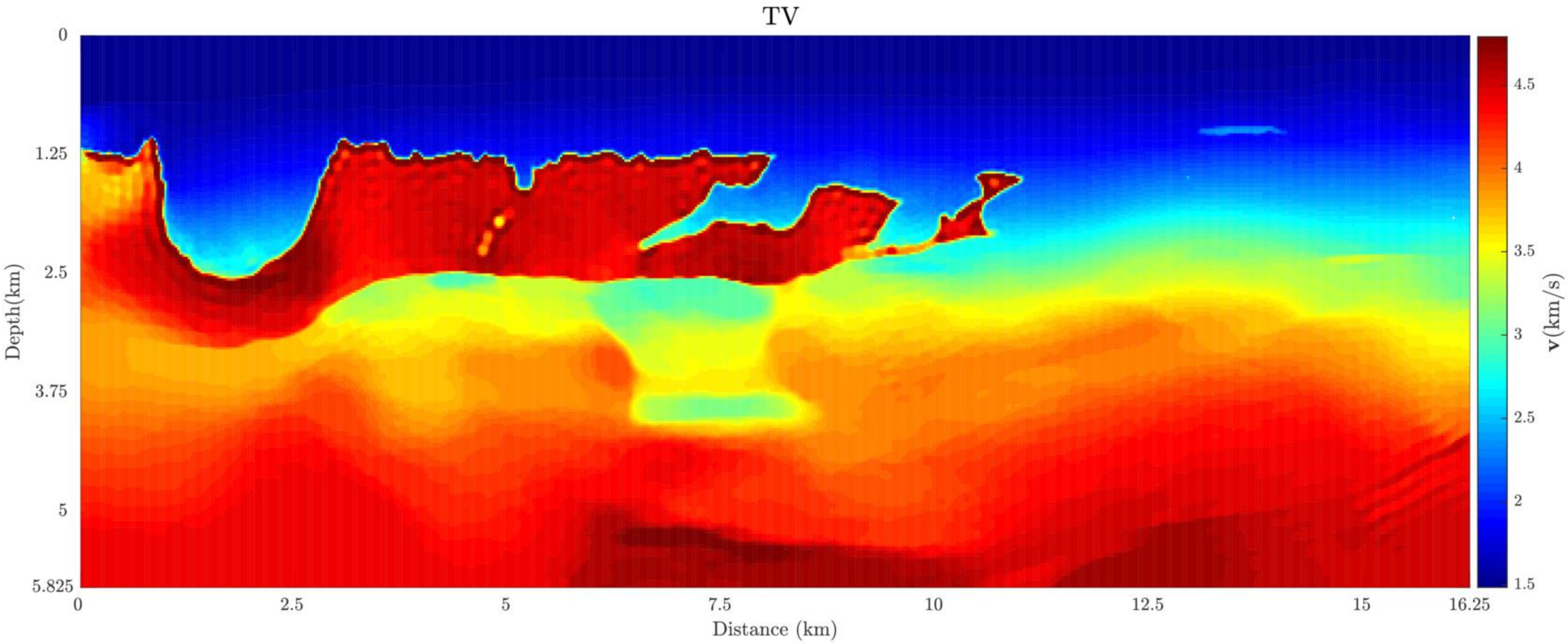
Source (left) and data (right) residuals at first iteration (top) and at convergence point (bottom)

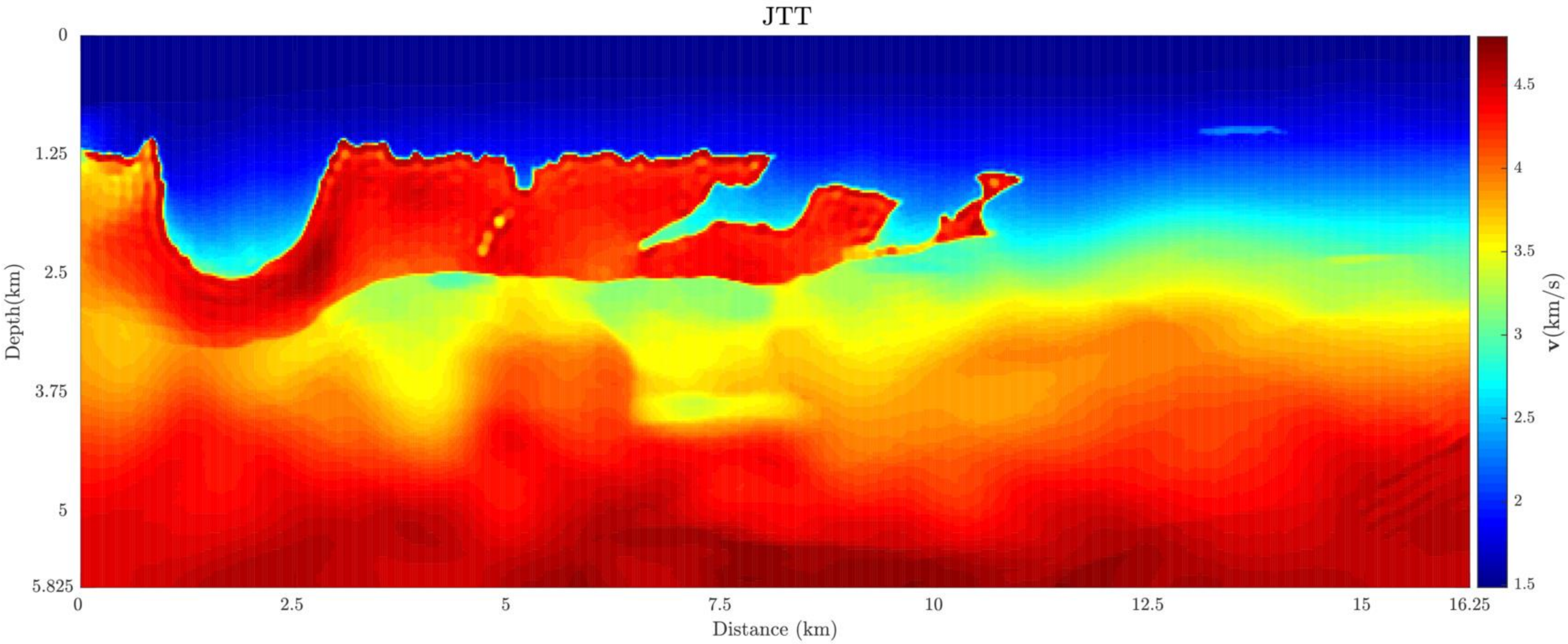


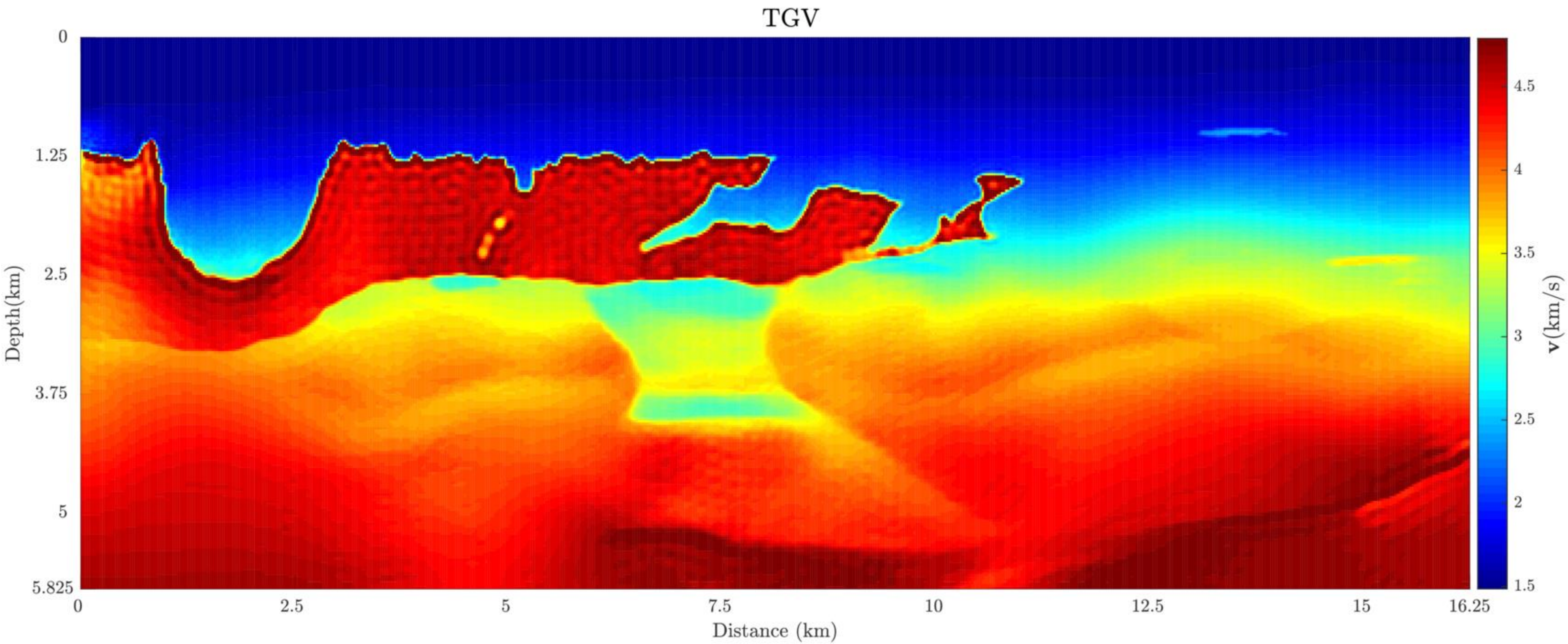
Regularized IR-WRI: *Final models*



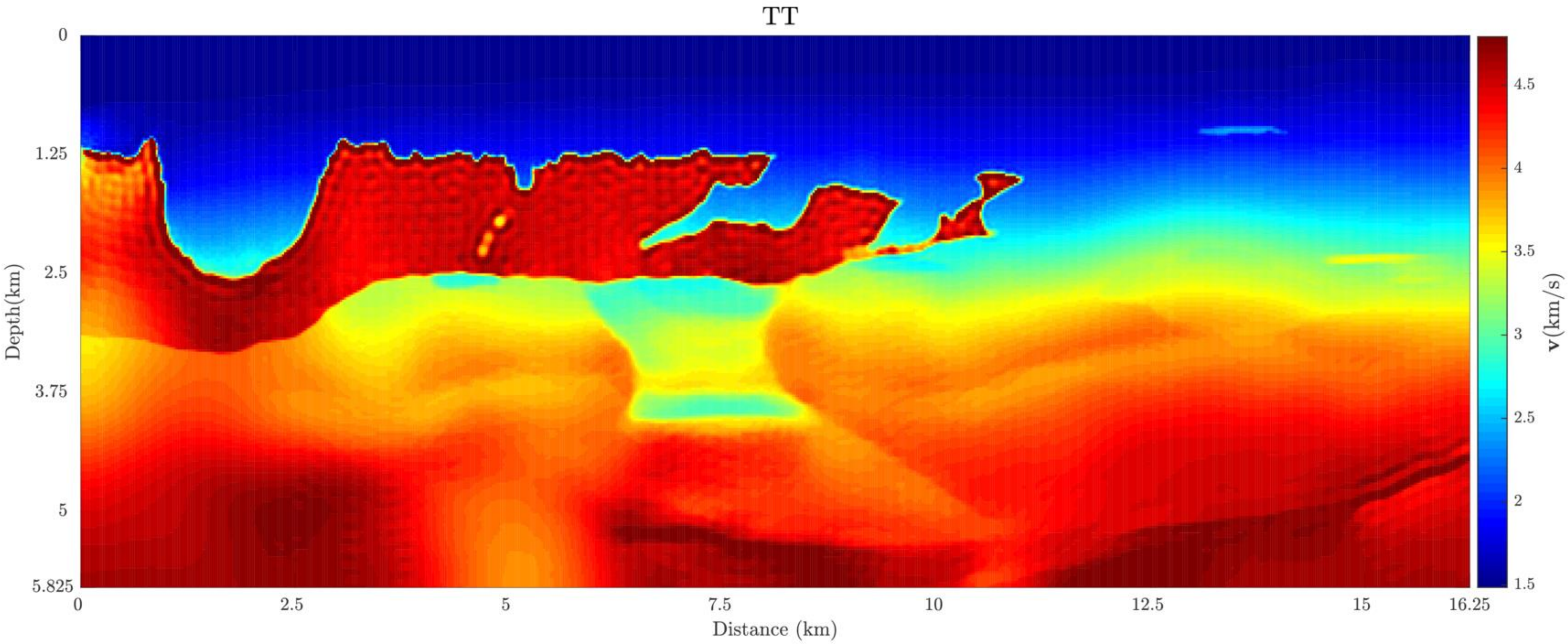

Tikh.
Depth(km)
0
1.25
2.5
3.75
5
5.825
0
2.5
5
7.5
10
12.5
15
16.25
Distance (km)
v(km/s)
4.5
4
3.5
3
2.5
2
1.5


TV
0
1.25
2.5
3.75
5
5.825
Depth(km)
0
2.5
5
7.5
10
12.5
15
16.25
Distance (km)
4.5
4
3.5
3
2.5
2
1.5
v(km/s)


JTT
Depth(km)
0
1.25
2.5
3.75
5
5.825
Distance (km)
0
2.5
5
7.5
10
12.5
15
16.25
v(km/s)
4.5
4
3.5
3
2.5
2
1.5


TGV
Depth(km)
0
1.25
2.5
3.75
5
5.825
0
2.5
5
7.5
10
12.5
15
16.25
Distance (km)
v(km/s)
4.5
4
3.5
3
2.5
2
1.5

Regularized IR-WRI: *Final models*



Regularized IR-WRI: *Final models*

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

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

Conclusions

- We have proposed a versatile recipe to cascade bound constraints and various regularizations in ADMM-based WRI (IR-WRI).
- Nonsmooth regularization is 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.

We thank SEISCOPE consortium and their sponsors for their continuous support.

- This study was partially funded by the SEISCOPE consortium (*http://seiscope2.osug.fr*), sponsored by AKERBP, CGG, CHEVRON, EQUINOR, EXXON-MOBIL, JGI, PETROBRAS, SCHLUMBERGER, SHELL, SINOPEC and TOTAL.
- This study was granted access to the HPC resources of SIGAMM infrastructure (http://crimson.oca.eu), hosted by Observatoire de la Côte d'Azur and which is supported by the Provence-Alpes Côte d'Azur region, and the HPC resources of CINES/IDRIS/TGCC under the allocation 0596 made by GENCI.

Any question??

Aghamiry, H., Gholami, A., and Operto, S. (2018). Hybrid tikhonov + total-variation regularization for imaging large-contrast media by full-waveform inversion. In *Expanded Abstracts, 88th Annual SEG Meeting (Anaheim)*, pages 1253–1257.

Aghamiry, H., Gholami, A., and Operto, S. (2019a). Admm-based multi-parameter wavefield reconstruction inversion in VTI acoustic media with TV regularization. *arXiv:1905.05452*.

Aghamiry, H., Gholami, A., and Operto, S. (2019b). Compound Regularization of Full-Waveform Inversion for Imaging Piecewise Media. *arXiv:1903.04405*.

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

Aghamiry, H., Gholami, A., and Operto, S. (2019f). Wavefield reconstruction inversion with enhanced hybrid Tikhonov and TV regularization for imaging piece-wise smooth media. In *81th Annual EAGE Meeting (London)*.

Benning, M. and Burger, M. (2018). Modern regularization methods for inverse problems. *Acta Numerica*, 27:1–111.

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.

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.

Setzer, S., Steidl, G., and Teuber, T. (2011). Infimal convolution regularizations with discrete $\ell 1$-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.

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

TT regularized IR-WRI: *Application of ADMM (or Split Bregman)*

- Primal descent

$$\begin{bmatrix}\mathbf{m}_1^{k+1}\\ \mathbf{m}_2^{k+1}\end{bmatrix} = \underset{\mathbf{m}_1,\mathbf{m}_2}{\arg\min}\; C(\mathbf{m}_1,\mathbf{m}_2,\mathbf{p}^k,\mathbf{m}^k,\tilde{\mathbf{p}}^k,\tilde{\mathbf{m}}^k), \tag{12a}$$

$$\mathbf{p}^{k+1} = \underset{\mathbf{p}}{\arg\min}\; \alpha\|\mathbf{p}\|_1 + \frac{\zeta}{2}\|\nabla\mathbf{m}_1^{k+1} - \mathbf{p} - \tilde{\mathbf{p}}^k\|_2^2, \tag{12b}$$

$$\mathbf{m}^{k+1} = \underset{\mathbf{m}\in\mathcal{C}}{\arg\min}\; \frac{\eta}{2}\|\mathbf{m}_1^{k+1} + \mathbf{m}_2^{k+1} - \mathbf{m} - \tilde{\mathbf{m}}^k\|_2^2, \tag{12c}$$

where

$$\begin{aligned} C(\mathbf{m}_1,\mathbf{m}_2,\mathbf{p}^k,\mathbf{m}^k,\tilde{\mathbf{p}}^k,\tilde{\mathbf{m}}^k) = & \frac{\gamma}{2}\|\mathbf{L}[\mathbf{m}_1+\mathbf{m}_2] - \mathbf{y}\|_2^2 + (1-\alpha)\|\nabla^2\mathbf{m}_2\|_2^2 \\ & + \frac{\zeta}{2}\|\nabla\mathbf{m}_1 - \mathbf{p}^k - \tilde{\mathbf{p}}^k\|_2^2 + \frac{\eta}{2}\|\mathbf{m}_1+\mathbf{m}_2-\mathbf{m}^k-\tilde{\mathbf{m}}^k\|_2^2, \end{aligned} \tag{13}$$

- Dual ascent

$$\tilde{\mathbf{p}}^{k+1} = \tilde{\mathbf{p}}^k + \mathbf{p}^{k+1} - \nabla\mathbf{m}_1^{k+1}, \tag{14a}$$

$$\tilde{\mathbf{m}}^{k+1} = \tilde{\mathbf{m}}^k + \mathbf{m}^{k+1} - (\mathbf{m}_1^{k+1} + \mathbf{m}_2^{k+1}). \tag{14b}$$

TT regularized IR-WRI: *Subproblem ($\mathbf{m}_1$,$\mathbf{m}_2$) - Jointly updating $\mathbf{m}_1$ and $\mathbf{m}_2$ by variable projection*

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

$$\begin{bmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{bmatrix}, \tag{15}$$

with

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

and

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

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

- From the first equation of (15), we find that

$$\mathbf{m}_2 = \mathbf{G}_{12}^{-1} [\mathbf{h}_1 - \mathbf{G}_{11} \mathbf{m}_1] \tag{16}$$

and plugging this into the second equation of (15) we get the following

$$\mathbf{m}_1 = (\mathbf{G}_{11} - \mathbf{G}_{22} \mathbf{G}_{12}^{-1} \mathbf{G}_{11})^{-1} [\mathbf{h}_2 - \mathbf{G}_{22} \mathbf{G}_{12}^{-1} \mathbf{h}_1]. \tag{17}$$

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)

$$\mathbf{p}^{k+1} = \text{prox}_{\zeta/\alpha}(\mathbf{z}) = \begin{bmatrix} \xi \circ \mathbf{z}_x \\ \xi \circ \mathbf{z}_z \end{bmatrix}, \tag{18}$$

where

$$\mathbf{z} = \nabla \mathbf{m}_1^{k+1} - \tilde{\mathbf{p}}^k = \begin{bmatrix} \mathbf{z}_x \\ \mathbf{z}_z \end{bmatrix}, \tag{19}$$

and

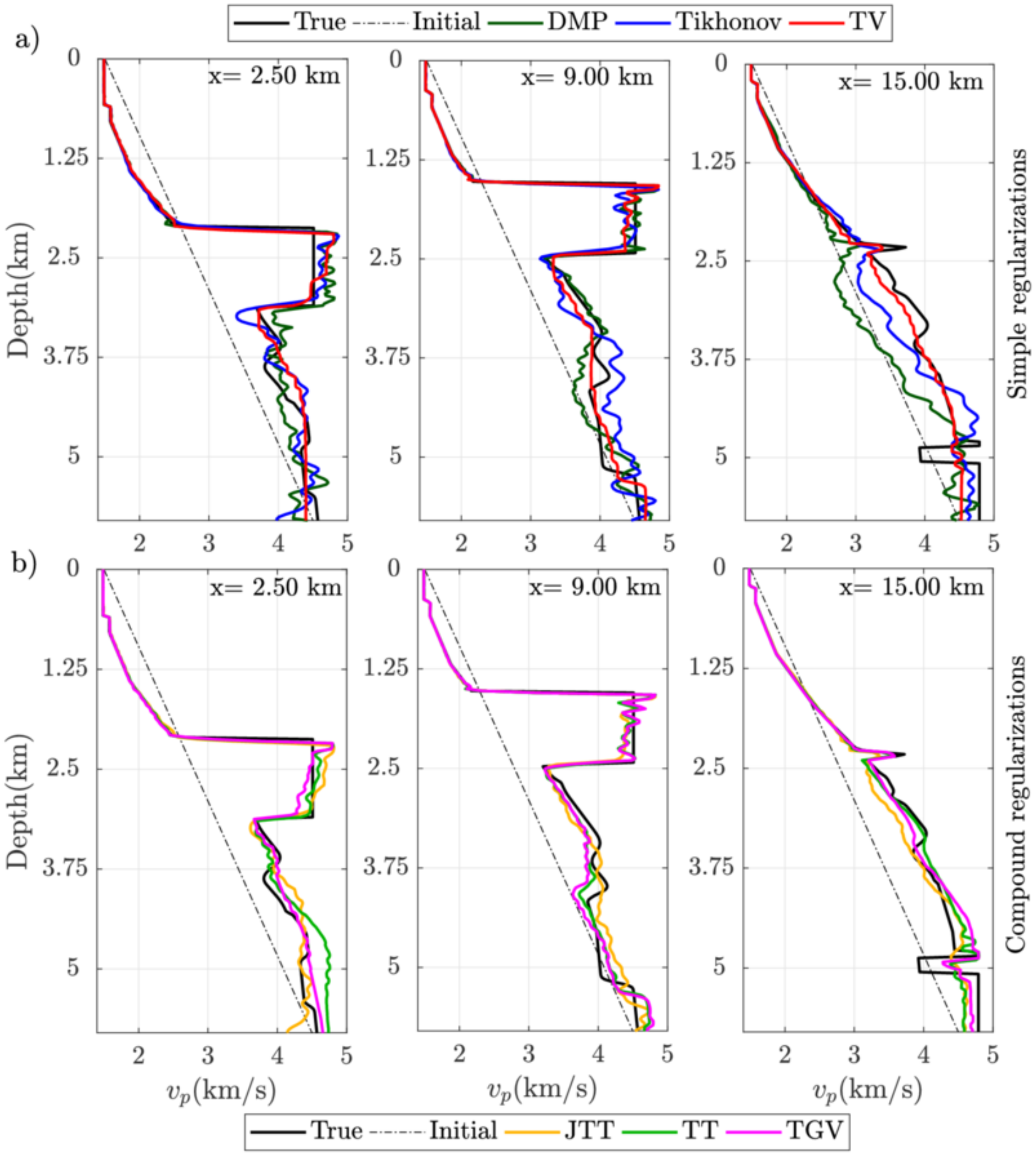
$$\xi = \max\left(1 - \frac{\zeta}{\alpha\sqrt{\mathbf{z}_x^2 + \mathbf{z}_z^2}}, 0\right). \tag{20}$$

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

$$\mathbf{m}^{k+1} = \text{proj}_{\mathcal{C}}(\mathbf{m}_1^{k+1} + \mathbf{m}_2^{k+1} - \tilde{\mathbf{m}}^k), \tag{21}$$

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

TT regularized IR-WRI: *Final models (Logs)*



TT Regularized IR-WRI:

Projection and proximity (Combettes and Pesquet, 2011)

Projection subproblem $\rightarrow$ $\mathbf{q} = \arg\min_{\mathbf{q}\in\mathcal{C}} \|\bullet - \mathbf{q}\|_2^2 = \mathbf{proj}_{\mathcal{C}}(\bullet) = \min(\max(\bullet, \mathbf{m}_l), \mathbf{m}_u)$

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

Proximity subproblem $\rightarrow$ $\mathbf{p}_1 = \arg\min_{\mathbf{p}_1} \sum \sqrt{|\mathbf{p}_1|^2 + |\mathbf{p}_2|^2} + \gamma_1 \|\nabla_x \mathbf{m}_1 - \mathbf{p}_1 - \hat{\mathbf{p}}_1\|_2^2 = \mathbf{prox}_{\gamma_1}(\nabla_x \mathbf{m}_1 - \hat{\mathbf{p}}_1) =$ $\max(1 - \frac{1}{\gamma_1 \sqrt{|\nabla_x \mathbf{m}_1 - \hat{\mathbf{p}}_1|^2 + |\nabla_z \mathbf{m}_1 - \hat{\mathbf{p}}_2|^2}}, 0)[\nabla_x \mathbf{m}_1 - \hat{\mathbf{p}}]$

Approximates the input point with some other point closest to it in the L2 distance sense under regularization implemented with the penalty term $\sum \sqrt{|\mathbf{p}_1|^2 + |\mathbf{p}_2|^2}$.