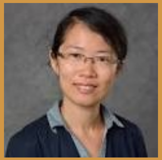
A Dual Formulation of Wavefield Reconstruction Inversion for Large-Scale Seismic Inversion

Gabrio Rizzuti[1], Mathias Louboutin[2], Rongrong Wang[3], and Felix J. Herrmann[2]

[1] Georgia Institute of Technology, now Utrecht University
[2] Georgia Institute of Technology
[3] Michigan State University

SLIM

ML4Seismic
11/23/2021

Agenda

- Motivations:
  robustness and complexity
- Theory:
  FWI vs WRI vs WRI*
- Numerical experiments:
  acoustic, TTI, small 3D
- Discussion and conclusions

Full Waveform Inversion

$$\min_{\mathbf{m}} \mathcal{J}_{\mathrm{FWI}}(\mathbf{m}) = \frac{1}{2}||F(\mathbf{m})\mathbf{q} - \mathbf{d}||^2, \quad F(\mathbf{m}) = RA(\mathbf{m})^{-1}$$

[Tarantola, A., ‘84; Haber, E., et al, 2000; Epanomeritakis, I., et al, 2008]

✔ **3D** computations are affordable
via large HPC systems (or cloud computing) [Witte et al., 2019]

✗ (Effectively) **multimodal** problem:
it needs a good starting model!

Wavefield Reconstruction Inversion

$$\min_{\mathbf{m}} \mathcal{J}_{\mathrm{WRI}}(\mathbf{m}, \bar{\mathbf{u}}(\mathbf{m})) = \frac{1}{2}||R\bar{\mathbf{u}} - \mathbf{d}||^2 + \frac{\lambda^2}{2}||A(\mathbf{m})\bar{\mathbf{u}} - \mathbf{q}||^2$$

[van den Berg, P. M., and Kleinman, R. E., 1997; van Leeuwen, T. and Herrmann, F. J., 2013]

✔Better **conditioning**

Wavefield Reconstruction Inversion

$$\min_{\mathbf{m}} \mathcal{J}_{\mathrm{WRI}}(\mathbf{m}, \bar{\mathbf{u}}(\mathbf{m})) = \frac{1}{2}||R\bar{\mathbf{u}} - \mathbf{d}||^2 + \frac{\lambda^2}{2}||A(\mathbf{m})\bar{\mathbf{u}} - \mathbf{q}||^2$$

[van den Berg, P. M., and K… and Herrmann, F. J., 2013]

$$\begin{bmatrix} R \\ \lambda A(\mathbf{m}) \end{bmatrix} \bar{\mathbf{u}} = \begin{bmatrix} \mathbf{d} \\ \lambda \mathbf{q} \end{bmatrix}$$

augmented wave equation

✔ Better **conditioning**

✗ Augmented solver: hard to scale to **3D**
(explicit time-marching schemes?)

Prior art: extended-source formulation

$$\min_{\mathbf{m}} \mathcal{J}_{\mathrm{WRI}}(\mathbf{m}, \bar{\mathbf{u}}) = \frac{1}{2}||R\bar{\mathbf{u}} - \mathbf{d}||^2 + \frac{\lambda^2}{2}||A(\mathbf{m})\bar{\mathbf{u}} - \mathbf{q}||^2$$

[van den Berg, P. M., and Kleinman, R. E., 1997; van Leeuwen, T. and Herrmann, F. J., 2013]

=

$$\min_{\mathbf{m}} \mathcal{J}_{\mathrm{WRI-q}}(\mathbf{m}, \bar{\mathbf{q}}) = \frac{1}{2}||F(\mathbf{m})\bar{\mathbf{q}} - \mathbf{d}||^2 + \frac{\lambda^2}{2}||\bar{\mathbf{q}} - \mathbf{q}||^2$$

[Wang et al, 2016, Huang et al, 2018]

Prior art: augmented state approximation

$$\min_{\mathbf{m}} \mathcal{J}_{\mathrm{WRI-q}}(\mathbf{m}, \bar{\mathbf{q}}) = \frac{1}{2}||F(\mathbf{m})\bar{\mathbf{q}} - \mathbf{d}||^2 + \frac{\lambda^2}{2}||\bar{\mathbf{q}} - \mathbf{q}||^2$$

[Wang et al, 2016]

$$\begin{bmatrix} F(\mathbf{m}) \\ \lambda I \end{bmatrix} \bar{\mathbf{q}} = \begin{bmatrix} \mathbf{d} \\ \lambda \mathbf{q} \end{bmatrix} \quad \Longrightarrow \quad \bar{\mathbf{q}} \approx \mathbf{q} + F(\mathbf{m})^{\mathrm{H}}(\mathbf{d} - F(\mathbf{m})\mathbf{q})/\lambda^2$$

✔(Approximated) augmented solver: scale to **3D**

Prior art: augmented state approximation

$$\min_{\mathbf{m}} \mathcal{J}_{\mathrm{WRI-q}}(\mathbf{m}, \bar{\mathbf{q}}) = \frac{1}{2}||F(\mathbf{m})\bar{\mathbf{q}} - \mathbf{d}||^2 + \frac{\lambda^2}{2}||\bar{\mathbf{q}} - \mathbf{q}||^2$$

[Wang et al, 2016]

$$\nabla_{\mathbf{m}} \mathcal{J}_{\mathrm{WRI-q}} = \nabla_{\mathbf{m}} \mathcal{J}_{\mathrm{WRI-q}}(\mathbf{m}, \bar{\mathbf{q}}) + \frac{\partial \bar{\mathbf{q}}}{\partial \mathbf{m}}^H \nabla_{\mathbf{q}} \mathcal{J}_{\mathrm{WRI-q}}(\mathbf{m}, \bar{\mathbf{q}})$$

✔ (Approximated) augmented solver: scale to **3D**

✘ Gradient computation inconsistency

WRI*: denoising reformulation

$$\min_{\mathbf{m},\mathbf{u}} \frac{1}{2}||A(\mathbf{m})\mathbf{u} - \mathbf{q}||^2 \quad \text{s.t.} \quad ||R\mathbf{u} - \mathbf{d}|| \leq \varepsilon$$

[Wang, R., and Herrmann, F. J., 2017]

WRI*: Lagrangian

$$\min_{\mathbf{m},\mathbf{u}} \frac{1}{2}||A(\mathbf{m})\mathbf{u} - \mathbf{q}||^2 \quad \text{s.t.} \quad ||R\mathbf{u} - \mathbf{d}|| \leq \varepsilon$$

[Wang, R., and Herrmann, F. J., 2017]

$$\max_{\mathbf{y}} \min_{\mathbf{m},\mathbf{u}} \mathcal{L}(\mathbf{m},\mathbf{u},\mathbf{y}) = \frac{1}{2}||A(\mathbf{m})\mathbf{u} - \mathbf{q}||^2 + \mathbf{y} \cdot (R\mathbf{u} - \mathbf{d}) - \varepsilon||\mathbf{y}||$$

WRI*: Lagrangian

$$\min_{\mathbf{m},\mathbf{u}} \frac{1}{2}||A(\mathbf{m})\mathbf{u}-\mathbf{q}||^2 \quad \text{s.t.} \quad ||R\mathbf{u}-\mathbf{d}|| \leq \varepsilon$$

[Wang, R., and Herrmann, F. J., 2017]

$$\max_{\mathbf{y}} \min_{\mathbf{m}} \bar{\mathcal{L}}(\mathbf{m},\mathbf{y}) = \frac{1}{2}||A(\mathbf{m})\boxed{\bar{\mathbf{u}}}-\mathbf{q}||^2 + \mathbf{y}\cdot(R\boxed{\bar{\mathbf{u}}}-\mathbf{d}) - \varepsilon||\mathbf{y}||$$

$$A(\mathbf{m})\bar{\mathbf{u}} = \mathbf{q} + F(\mathbf{m})^H\mathbf{y}$$

WRI*: augmented state approximation

$$\min_{\mathbf{m},\mathbf{u}} \frac{1}{2}||A(\mathbf{m})\mathbf{u}-\mathbf{q}||^2 \quad \text{s.t.} \quad ||R\mathbf{u}-\mathbf{d}|| \leq \varepsilon$$

[Wang, R., and Herrmann, F. J., 2017]

$$\min_{\mathbf{m}} \tilde{\bar{\mathcal{L}}}(\mathbf{m}) = \frac{1}{2}||A(\mathbf{m})\bar{\mathbf{u}}-\mathbf{q}||^2 + \tilde{\mathbf{y}} \cdot (R\bar{\mathbf{u}}-\mathbf{d}) - \varepsilon||\tilde{\mathbf{y}}||$$

$$\tilde{\mathbf{y}} \propto F(\mathbf{m})\mathbf{q} - \mathbf{d}$$

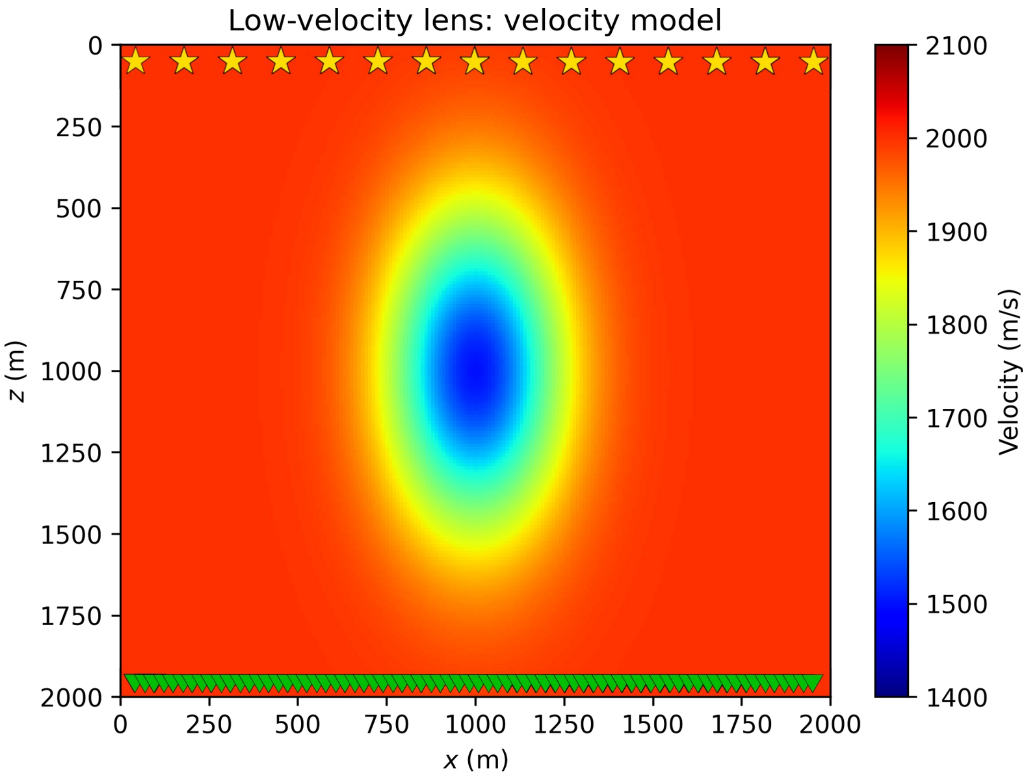
can be differentiated through...

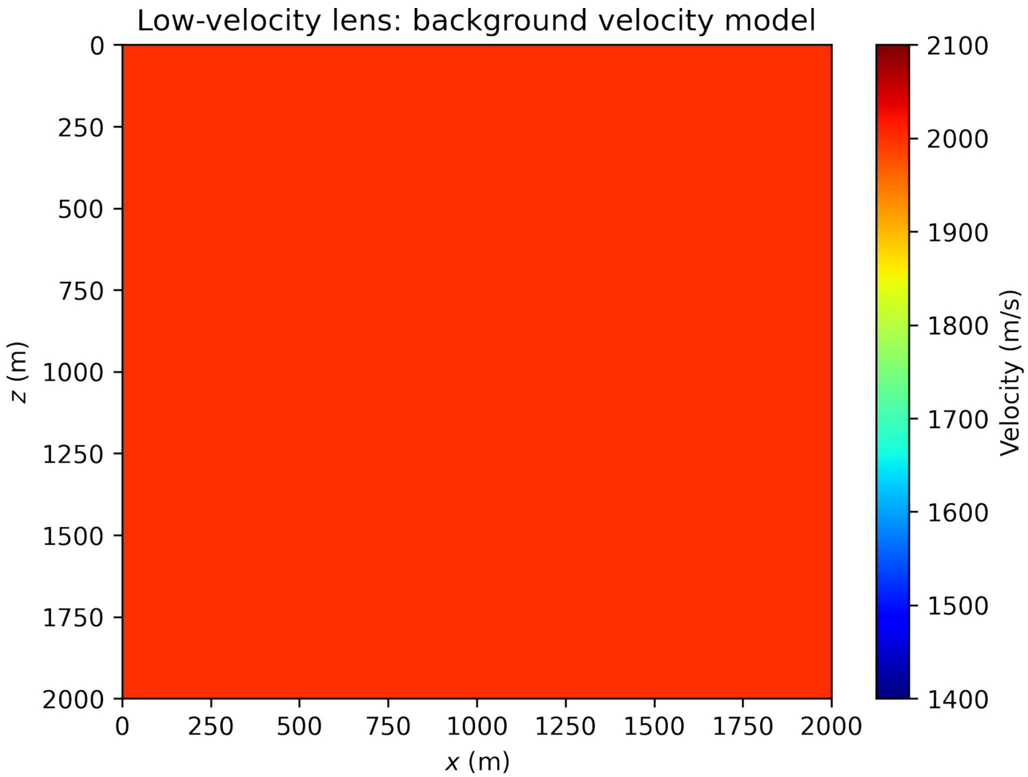
WRI*: reduced formulation

$$\min_{\mathbf{m}} \tilde{\bar{\mathcal{L}}}(\mathbf{m}) = \frac{1}{2}||A(\mathbf{m})\bar{\mathbf{u}} - \mathbf{q}||^2 + \tilde{\mathbf{y}} \cdot (R\bar{\mathbf{u}} - \mathbf{d}) - \varepsilon||\tilde{\mathbf{y}}||$$

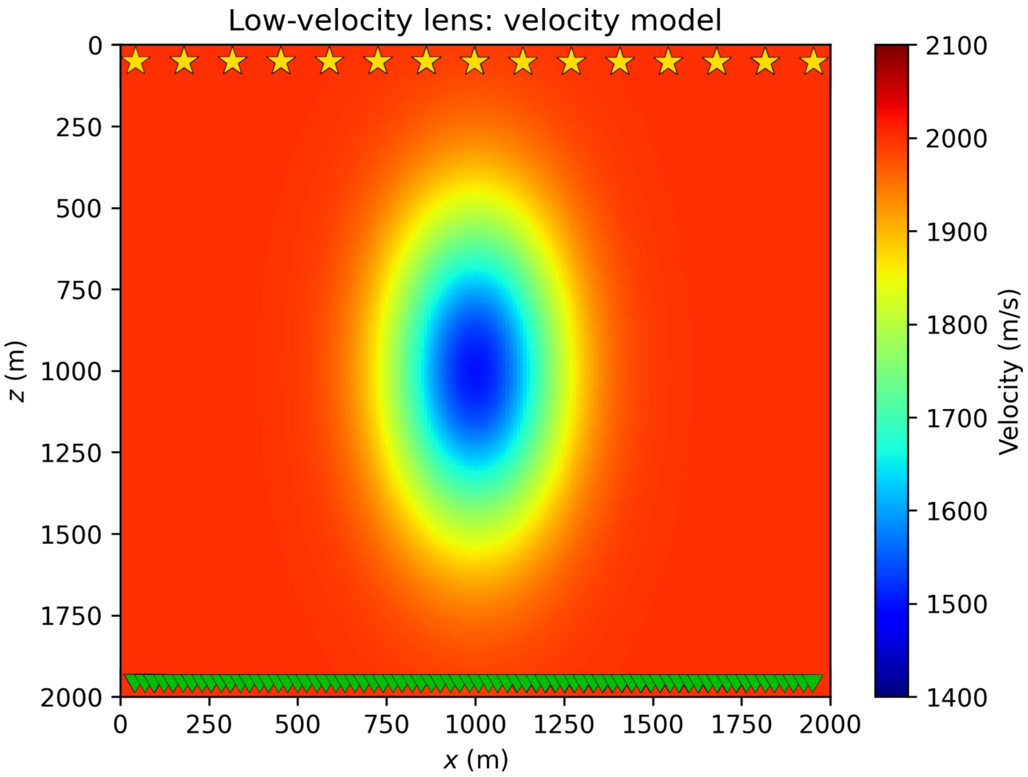
✔ Requires standard wave equation solver: scale to 3D
✔ No gradient computation inconsistency
✔ Retains WRI robustness
✔ Roughly equivalent to 2x the computational cost of FWI
✘ Roughly equivalent to 2x the computational cost of FWI
✘ Model resolution generally inferior to WRI

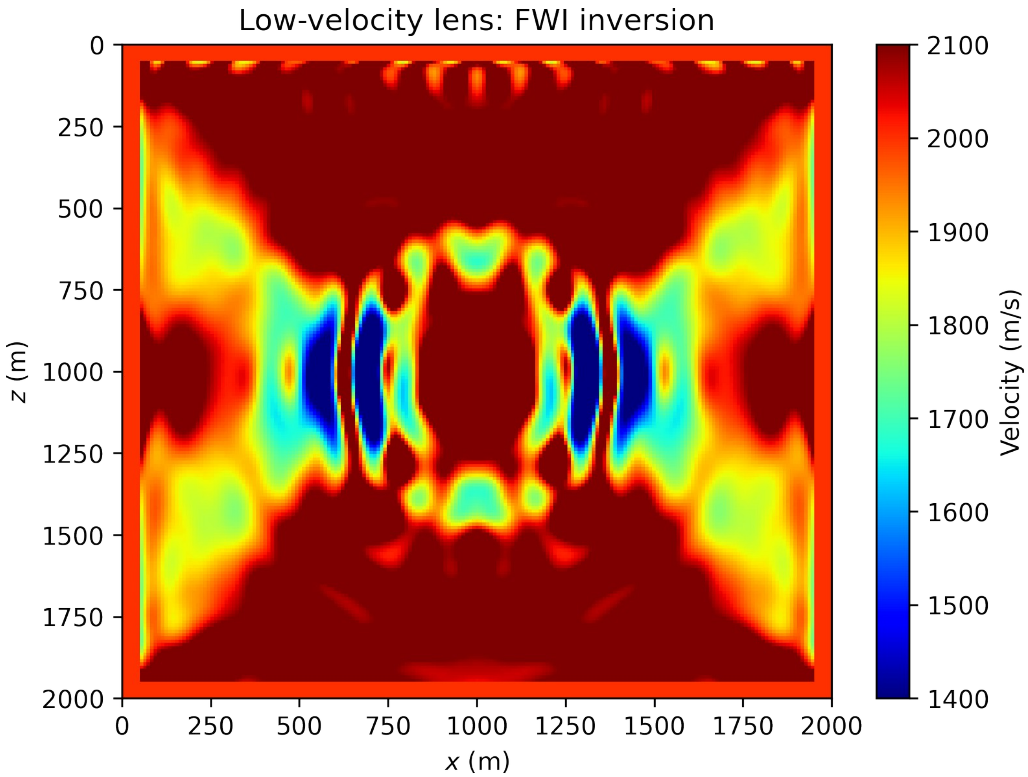
Numerical examples: transmission



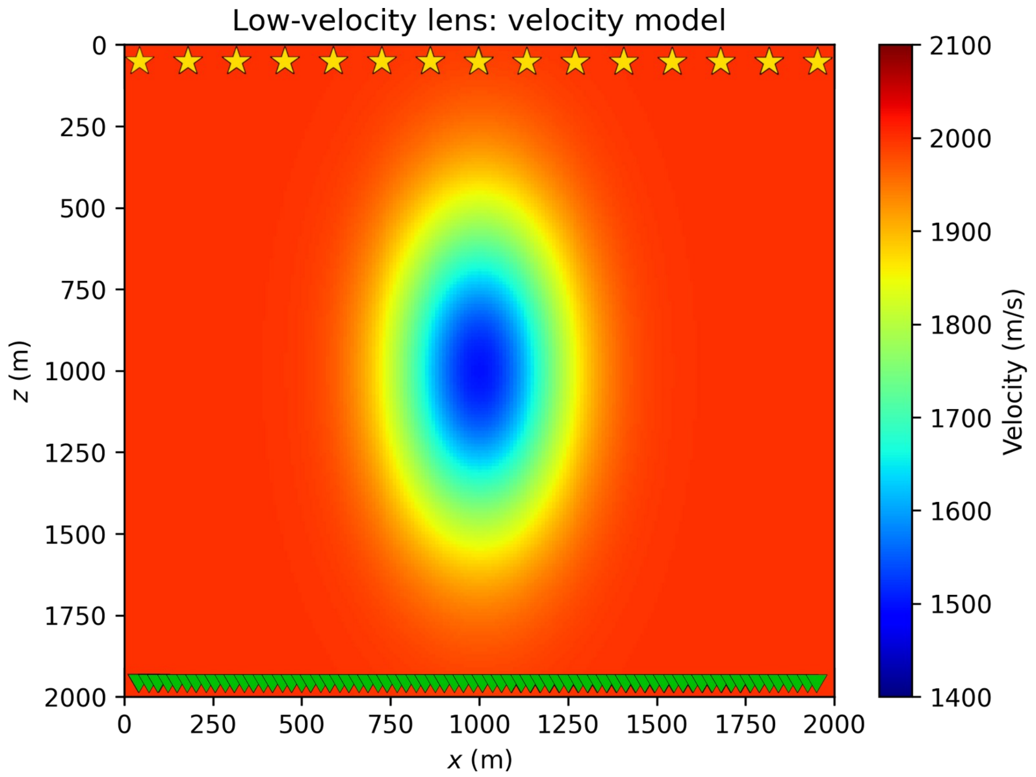


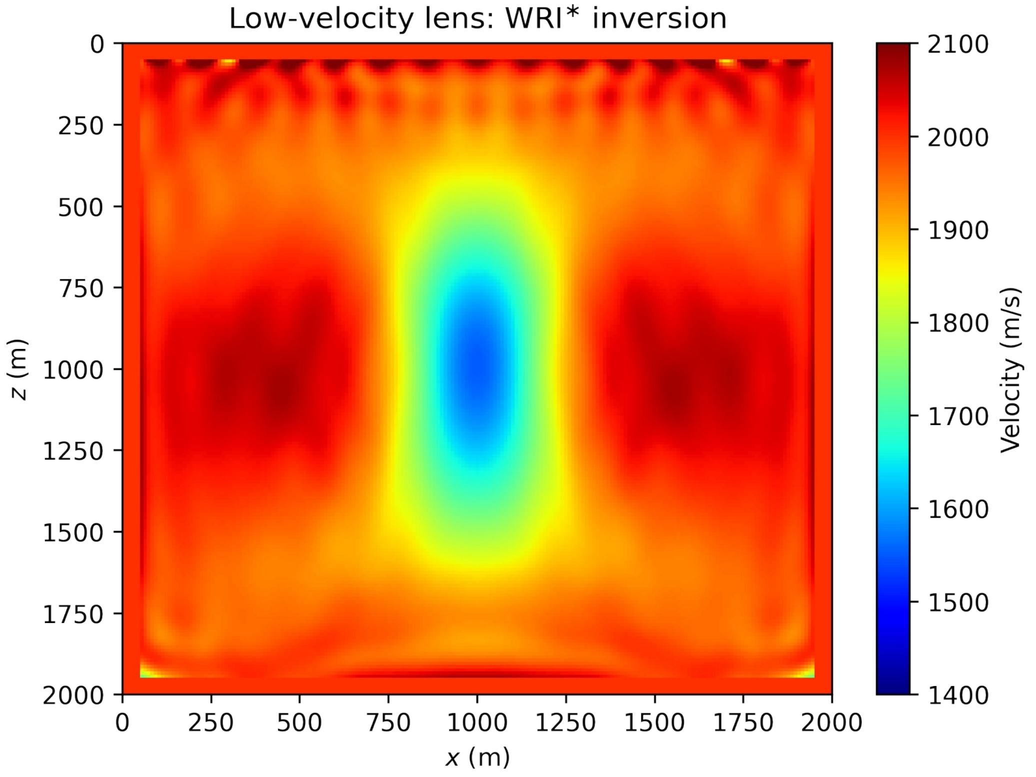
Numerical examples: transmission



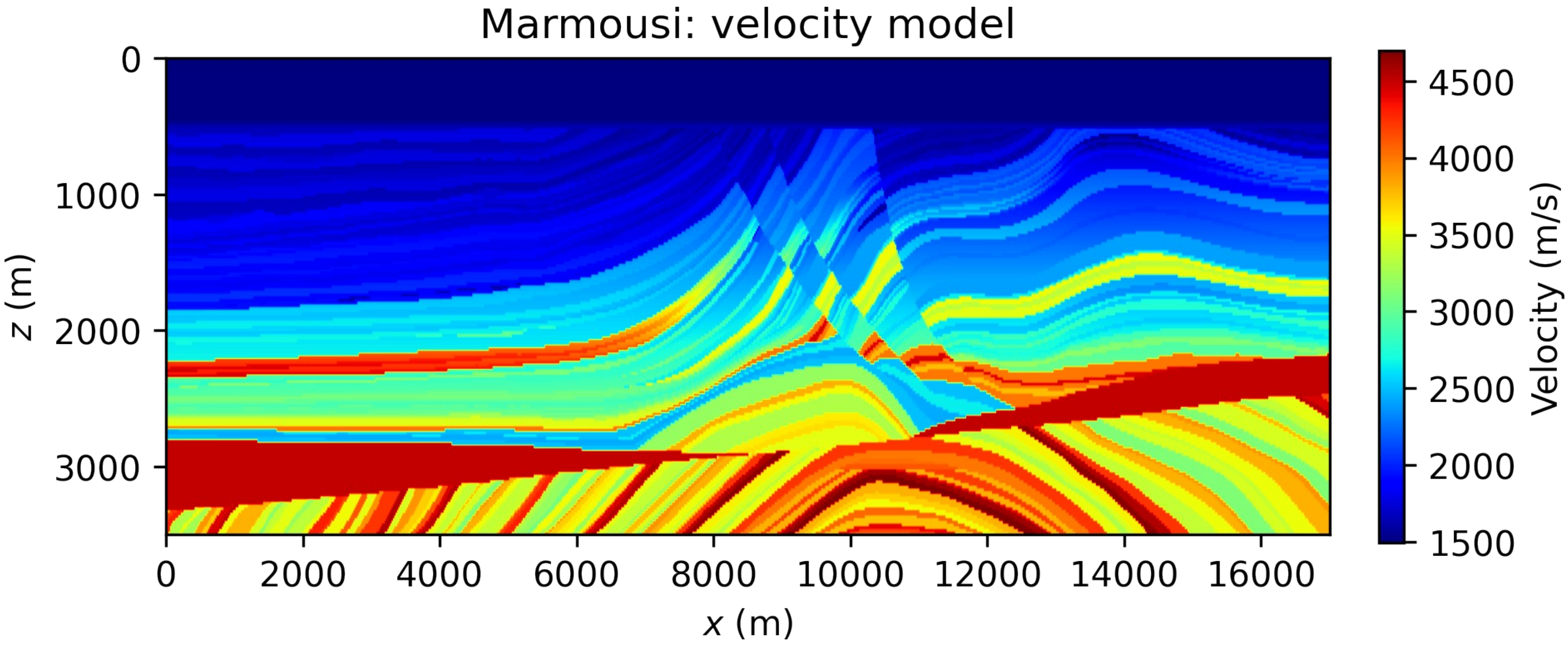


Numerical examples: transmission

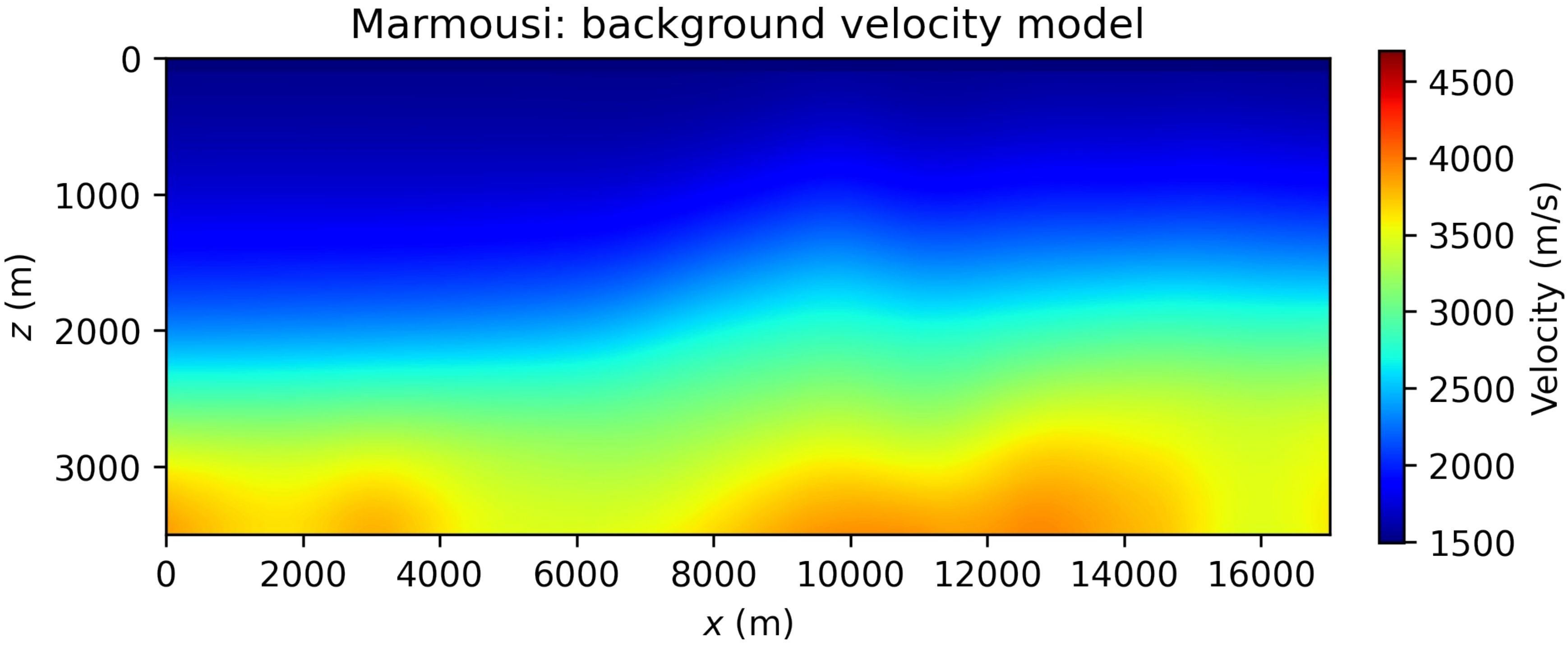




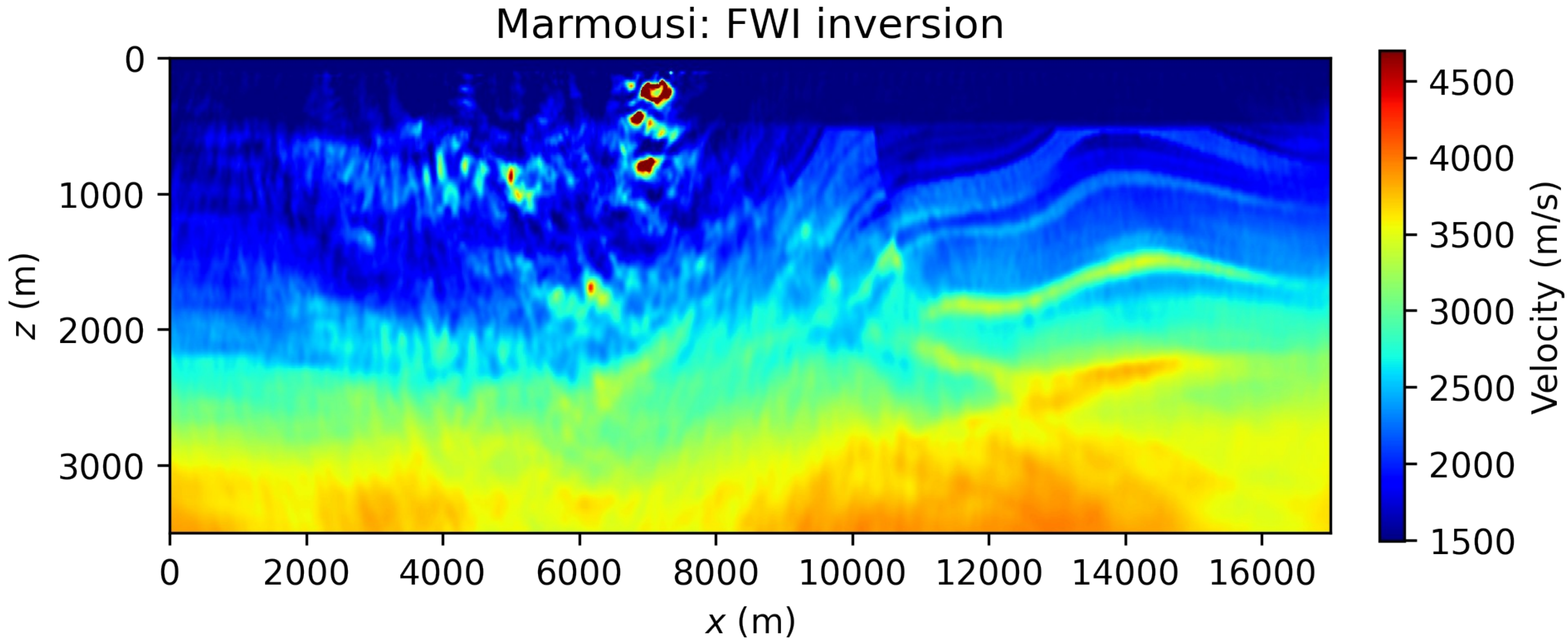
Numerical examples: Marmousi (freq. domain)



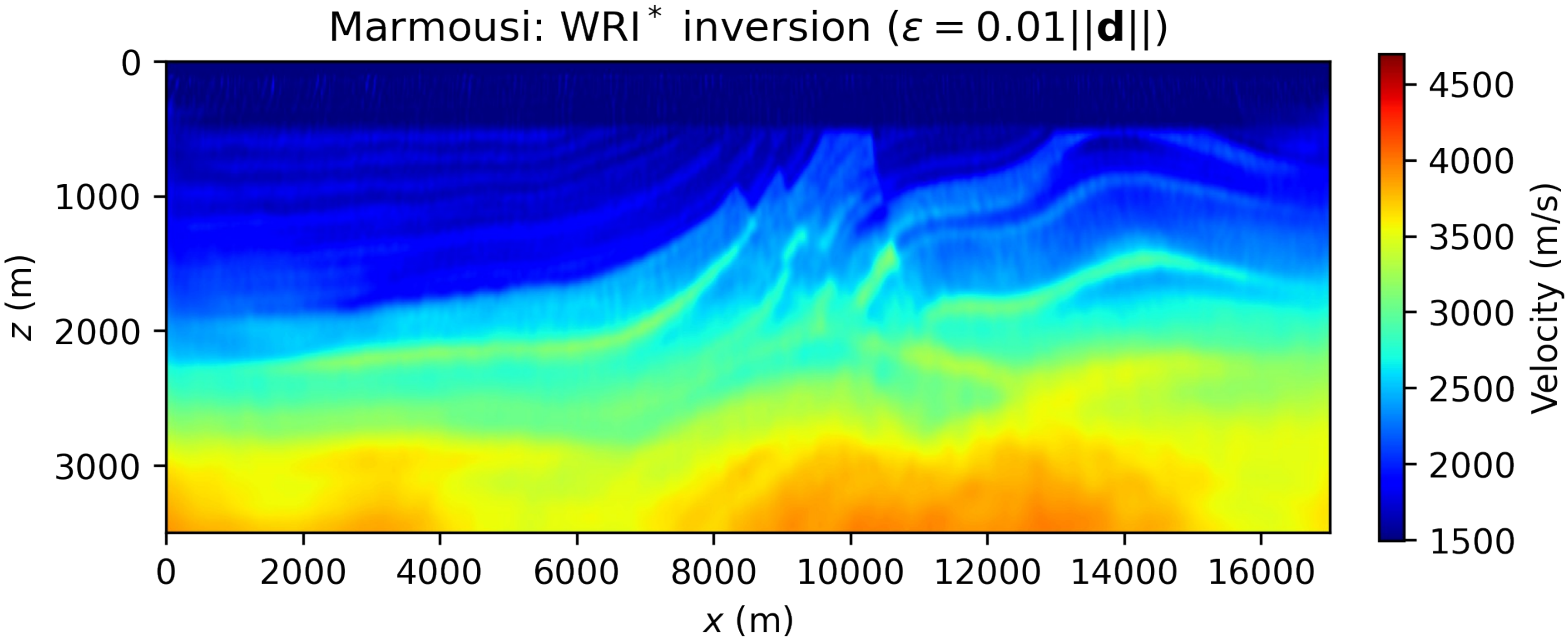
Numerical examples: Marmousi (freq. domain)



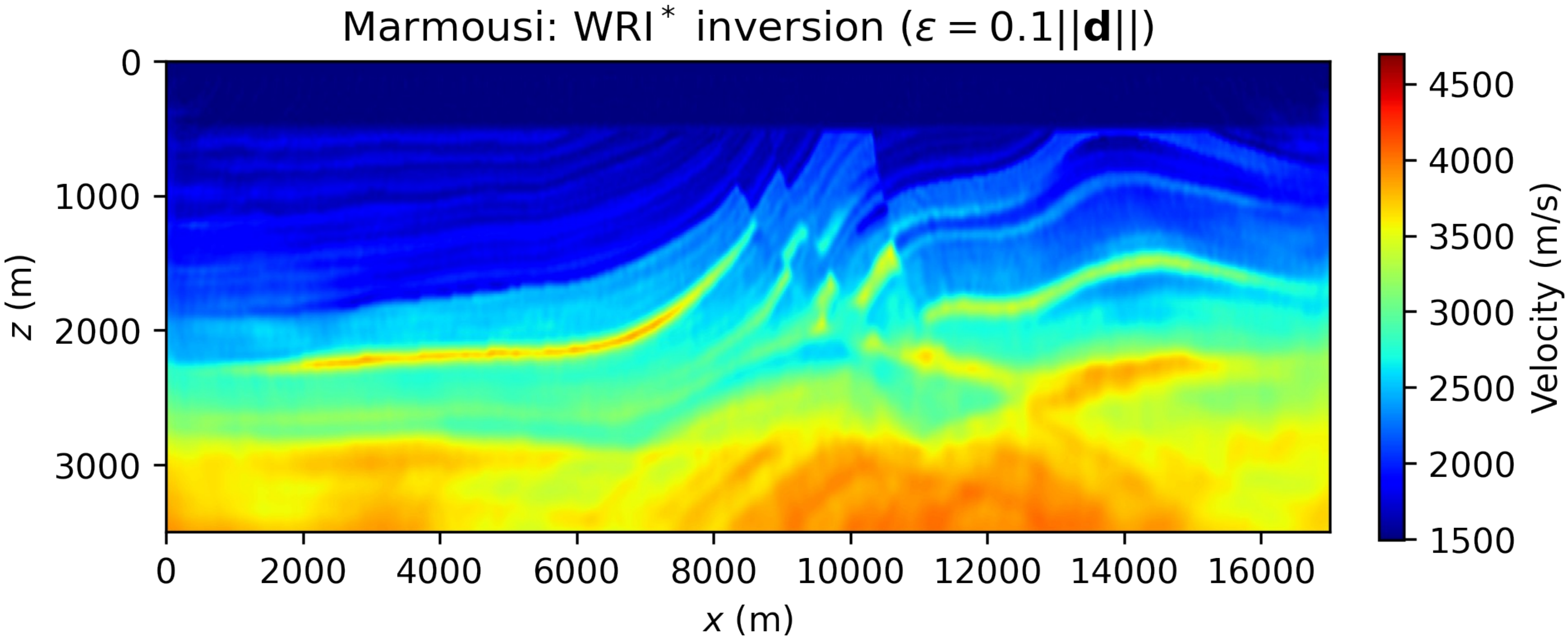
Numerical examples: Marmousi (freq. domain)



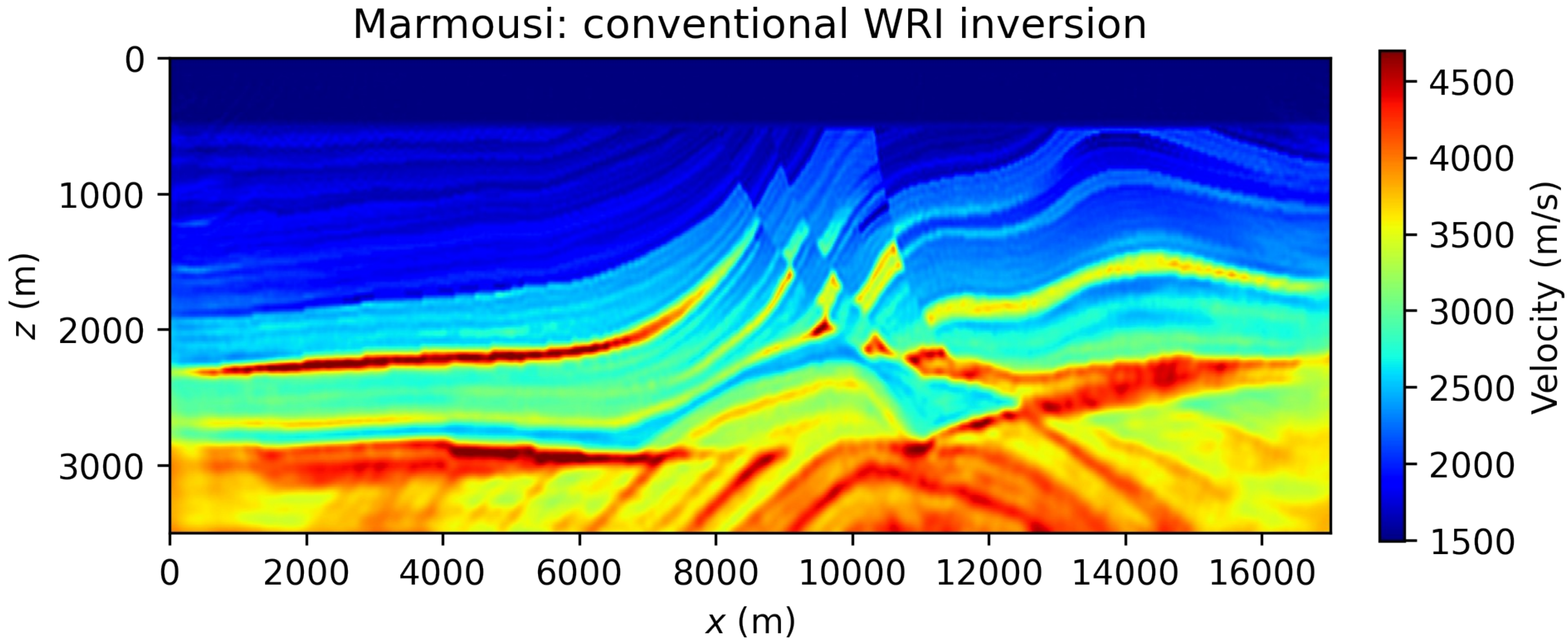
Numerical examples: Marmousi (freq. domain)



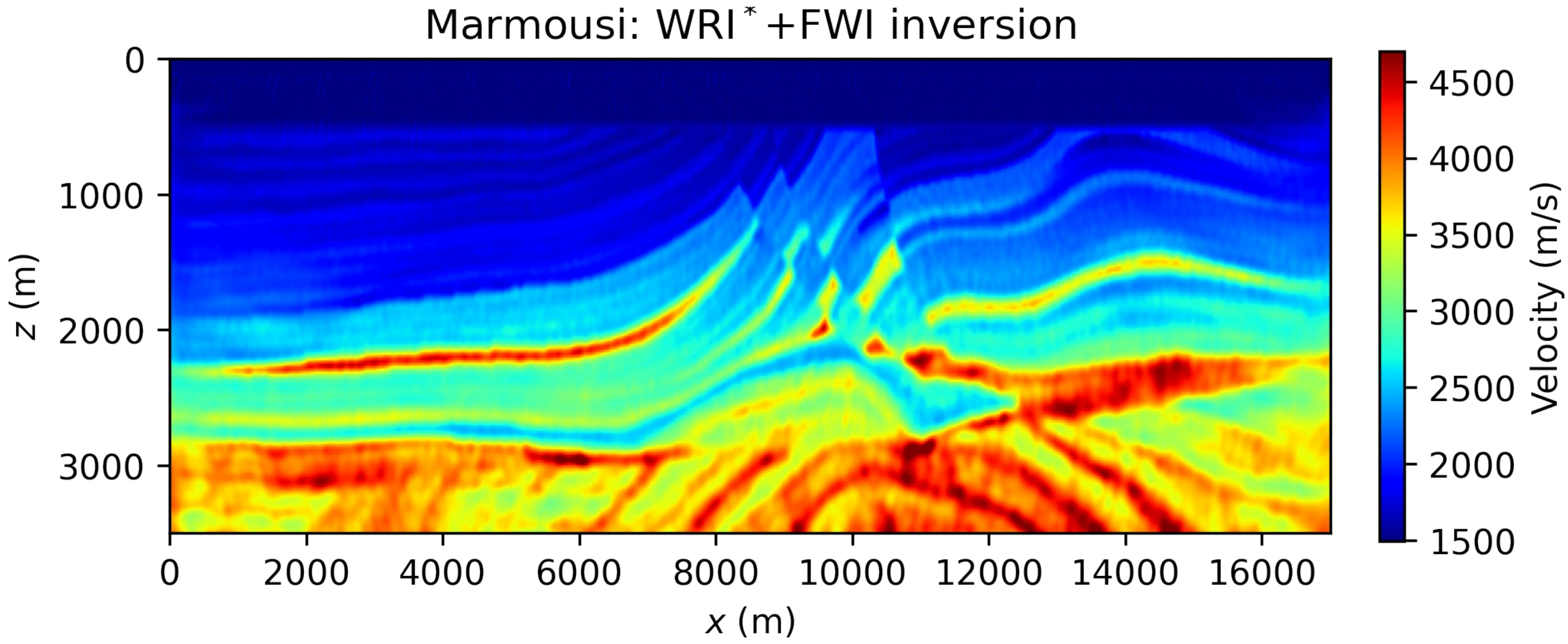
Numerical examples: Marmousi (freq. domain)



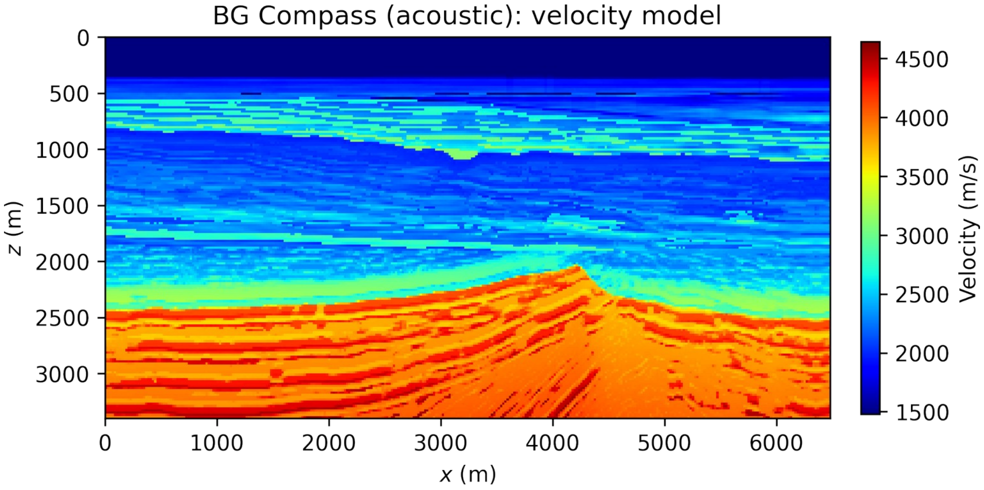
Numerical examples: Marmousi (freq. domain)

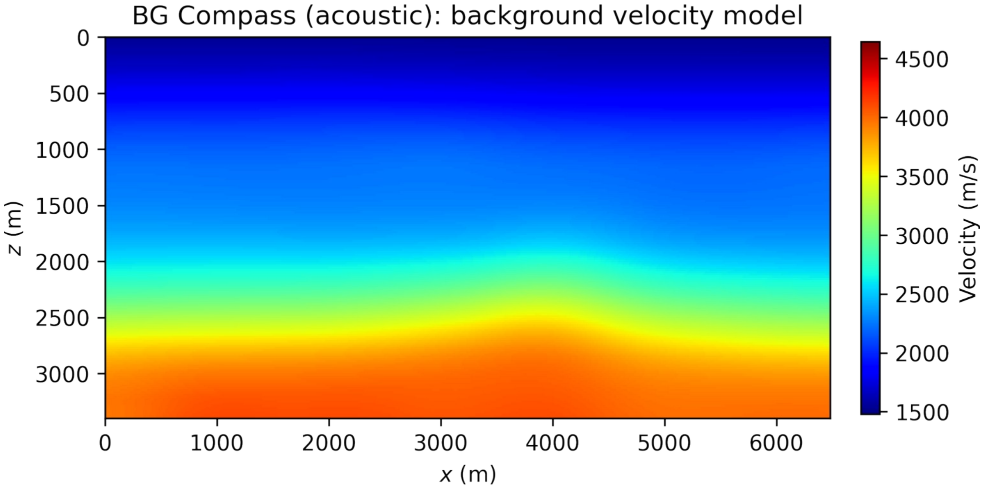


Numerical examples: Marmousi (freq. domain)

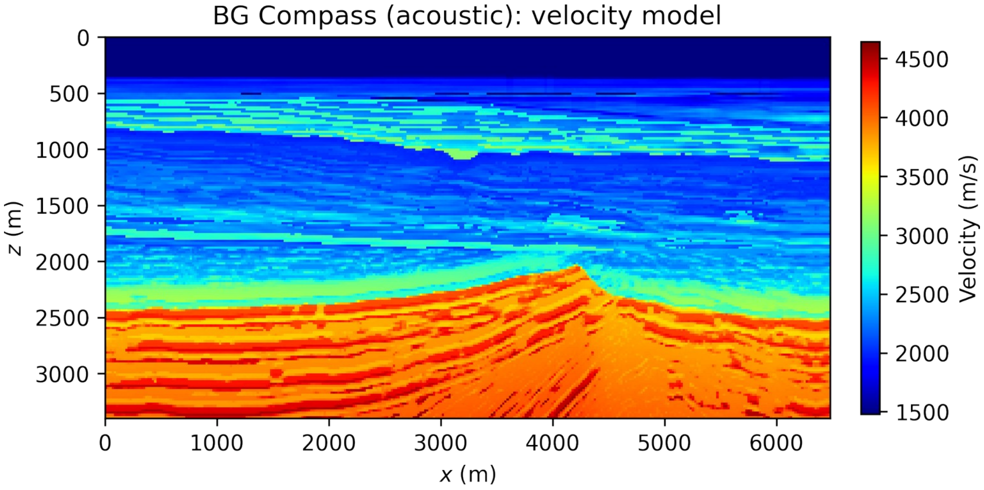


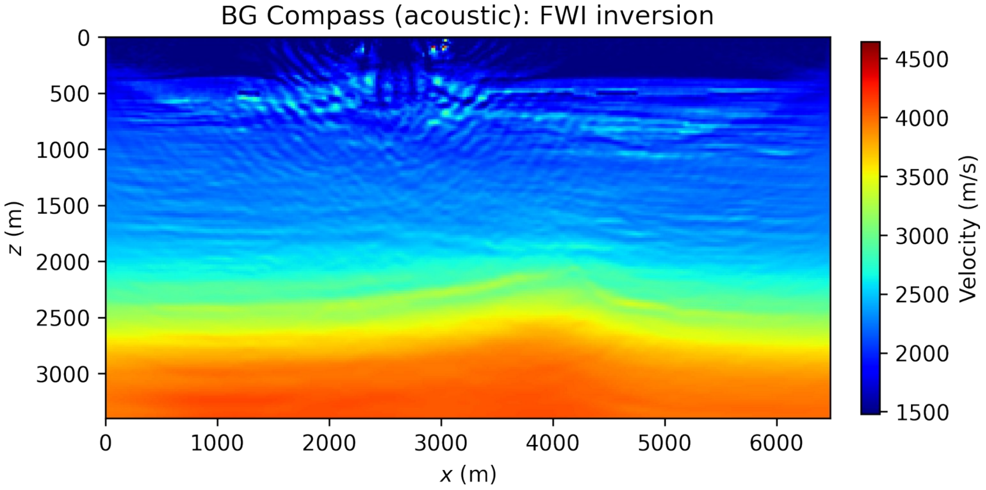
Numerical examples: BG Compass (acoustic)



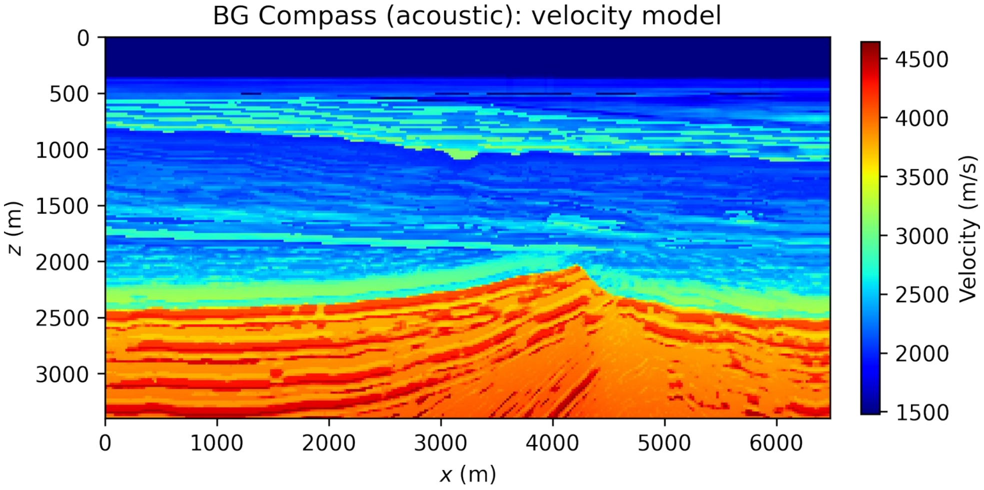


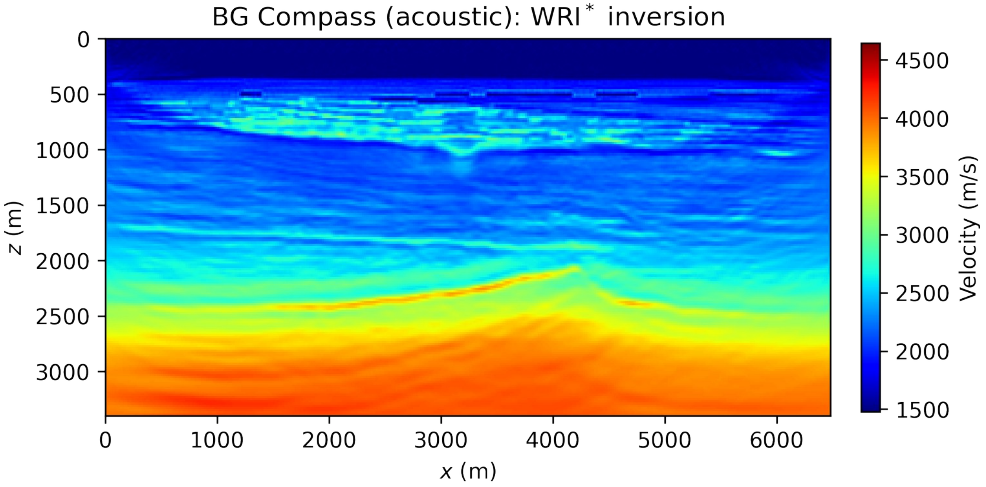
Numerical examples: BG Compass (acoustic)



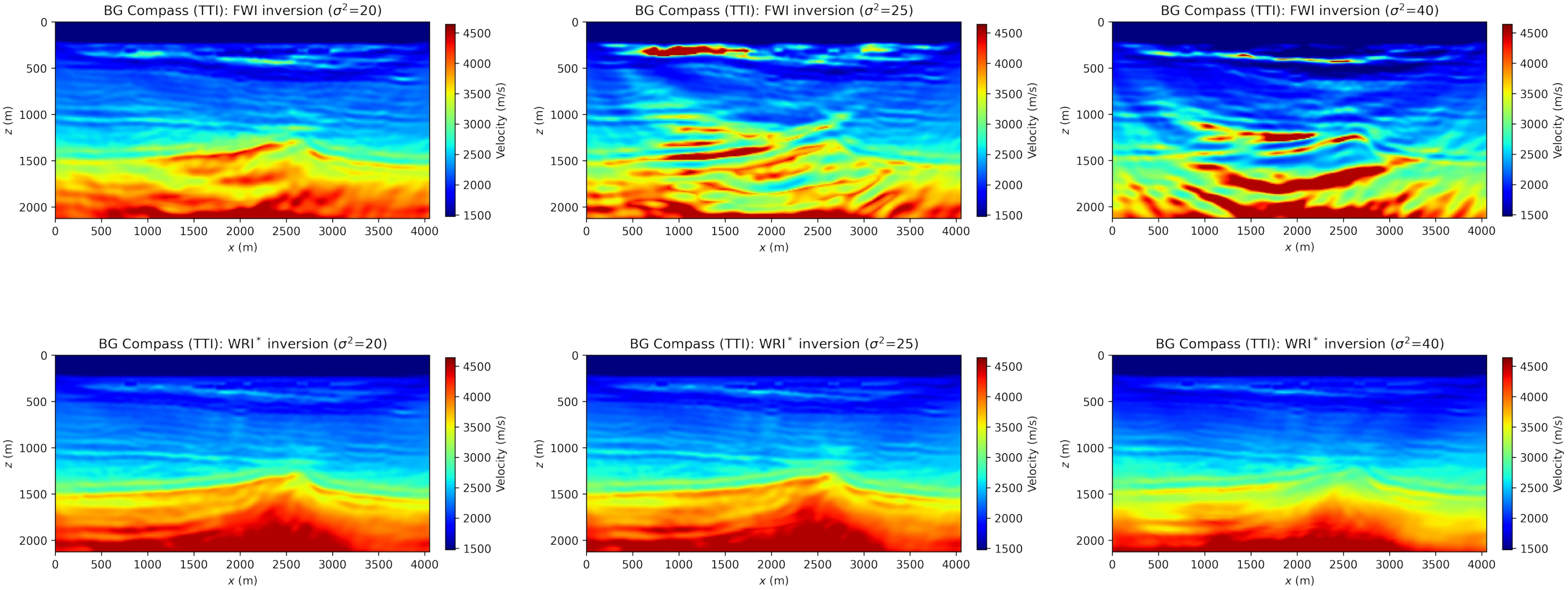


Numerical examples: BG Compass (acoustic)

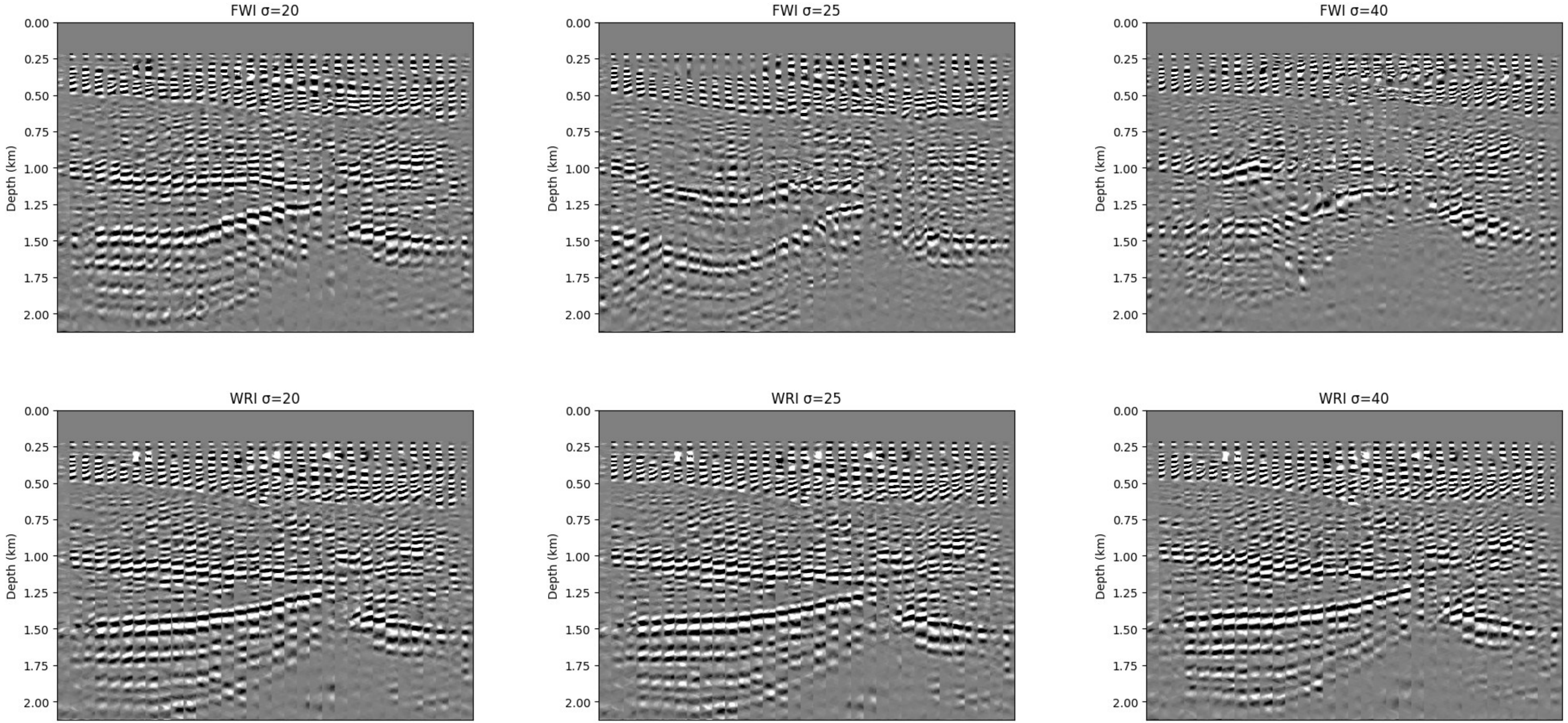




Numerical examples: BG Compass (TTI)

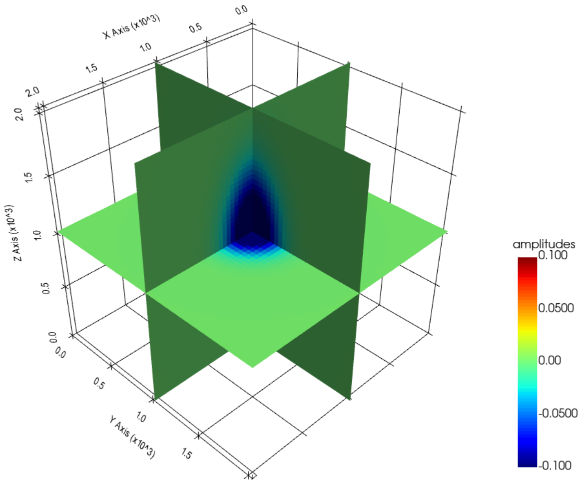


Numerical examples: BG Compass (TTI, CIGs)

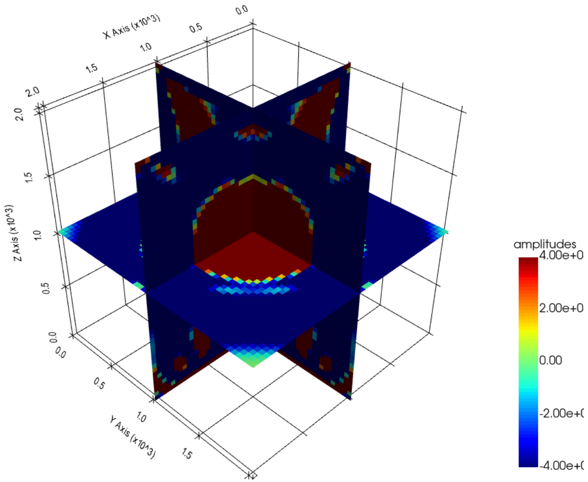


Numerical examples: small 3D

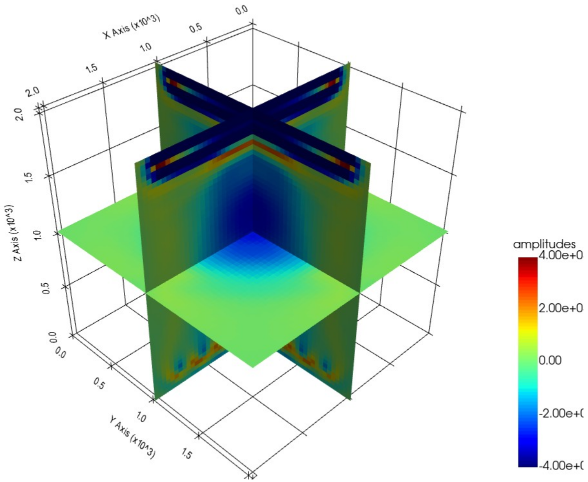
True



FWI



WRI*



Discussion

- WRI*:
  - affordable version of WRI
  - retains robustness of WRI
  - 2^{nd} order methods/hybrid schemes to improve resolution
- Source-focusing annihilator might be necessary to avoid local minima [Symes, W. W., 2020]

Discussion

- Better approximation for augmented variable? Need for automatic differentiation! [Ablin et al., 2020]

$$\nabla_{\mathbf{m}} \tilde{\bar{\mathcal{L}}} = \nabla_{\mathbf{m}} \bar{\mathcal{L}} + \frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{m}}^{H} \nabla_{\mathbf{y}} \bar{\mathcal{L}}$$

- Some special choice for covariance of data/model error avoid the need for augmented variable approximation [van Leeuwen, T., 2019]

Open source implementation

- github.com/slimgroup/JUDI.jl
- devitoproject.org
- github.com/slimgroup/ImageGather.jl

References

- Ablin, P., Peyré, G., and Moreau, T., 2020, “Super-efficiency of automatic differentiation for functions defined as a minimum”
- Biros, G., and Ghattas, O., Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization. Part i: The Krylov-Schur solver, SIAM Journal on Scientific Computing, 27 (2005)
- Epanomeritakis, I., and Akcelik, V., and Ghattas, O., and Bielak, J., A Newton-CG method for large-scale three-dimensional elastic full-waveform seismic inversion, Inverse Problems, 24(3) (2008)
- Grote, M. J., and Huber, J., and Schenk, O., Interior point methods for the inverse medium problem on massively parallel architectures, Procedia Computer Science, 4 (2011)
- Haber, E., and Ascher, U. M., and Oldenburg, D., On optimization techniques for solving nonlinear inverse problems, Inverse Problems, 16 (2000)
- Haber, E., and Ascher, U. M., Preconditioned all-at-once methods for large, sparse parameter estimation problems, Inverse Problems, 17 (2001)
- Huang, G., and Nammour, R., and Symes, W. W., Volume source-based extended waveform inversion, Geophysics (2018)
- Kleinman, R. E., and van den Berg, P. M., A modified gradient method for two-dimensional problems in tomography, Journal of Computational and Applied Mathematics, 42 (1992)
- van Leeuwen, T., and Herrmann, F. J., Mitigating local minima in full-waveform inversion by expanding the search space, Geophysical Journal International 195.1 (2013)
- van Leeuwen, T., A note on extended full waveform inversion, Geophysical Journal International (2019)
- Louboutin, M., Lange, M., Luporini, F., Kukreja, N., Witte, P. A., Herrmann, F. J., Velesko, P. and Gorman, G. J., Devito: an embedded domain-specific language for finite differences and geophysical exploration, CoRR, abs/1808.01995, , arXiv (2018)
- Luporini, F., Lange, M., Louboutin, M., Kukreja, N., Hückelheim, J., Yount, C., Witte, P., Kelly, P. H. J., Gorman, G. J. and Herrmann, F. J., Architecture and performance of Devito, a system for automated stencil computation, CoRR, abs/1807.03032, 2018, , arXiv (2018)
- Peters, B., and Herrmann, F. J., and van Leeuwen, T., Wave-equation Based Inversion with the Penalty Method-Adjoint-state Versus Wavefield-Reconstruction Inversion, 76th EAGE Conference (2014)
- Sharan, S., and Wang, R., and Herrmann, F. J., “Fast sparsity-promoting microseismic source estimation”, Geophysical Journal International (2019)
- Symes, W. W., 2020, “Full Waveform Inversion by Source Extension: Why it works”
- Symes, W. W., 2020, “Wavefield Reconstruction Inversion: an example”
- Van Leeuwen, T., 2019, “A note on extended full waveform inversion”
- Tarantola, A., Inversion of seismic reflection data in the acoustic approximation, Geophysics 49(8) (1984)
- Wang, C., and Yingst, D., and Farmer, P., and Leveille, J., Full Waveform Inversion with the Reconstructed Wavefield Method, 86th EAGE Conference (2016)
- Wang, R., and Herrmann, F. J., A denoising formulation of Full-Waveform Inversion, 87th SEG International Exposition (2017)
- Witte, P. A. and Louboutin, M. and Modzelewski, H. and Jones, C. and Selvage, J. and Herrmann, F. J., 2020, An Event-Driven Approach to Serverless Seismic Imaging in the Cloud: IEEE Transactions on Parallel and Distributed Systems, 9
- Witte, P. A., M. Louboutin, N. Kukreja, F. Luporini, M. Lange, G. J. Gorman, and F. J. Herrmann, 2019, A large-scale framework for symbolic implementations of seismic inversion algorithms in Julia: Geophysics, 84, 1–60 (2019)