Wavefield Reconstruction Inversion (WRI) – a new take on wave-equation based inversion

Felix J. Herrmann

SLIM
University of British Columbia

van Leeuwen, T and Herrmann, F J (2013). Mitigating local minima in full-waveform inversion by expanding the search space. Geophysical Journal International.
van Leeuwen, T and Herrmann, F J (2013). A penalty method for PDE-constrained optimization. Submitted for publication
van Leeuwen, T and Herrmann, F J (2013). US Provisional Patent Application No. 61/815,533. A Penalty Method for PDE-Constrained Optimization with Applications to Wave-Equation Based Seismic Inversion

Wavefield Reconstruction Inversion (WRI) – a new take on wave-equation based inversion

Ernie Esser, Tristan van Leeuwen*, and Bas Peters

* now @ Mathematical Institute at Utrecht University

SLIM
University of British Columbia

van Leeuwen, T and Herrmann, F J (2013). Mitigating local minima in full-waveform inversion by expanding the search space. Geophysical Journal International.
van Leeuwen, T and Herrmann, F J (2013). A penalty method for PDE-constrained optimization. Submitted for publication
van Leeuwen, T and Herrmann, F J (2013). US Provisional Patent Application No. 61/815,533. A Penalty Method for PDE-Constrained Optimization with Applications to Wave-Equation Based Seismic Inversion

Motivation

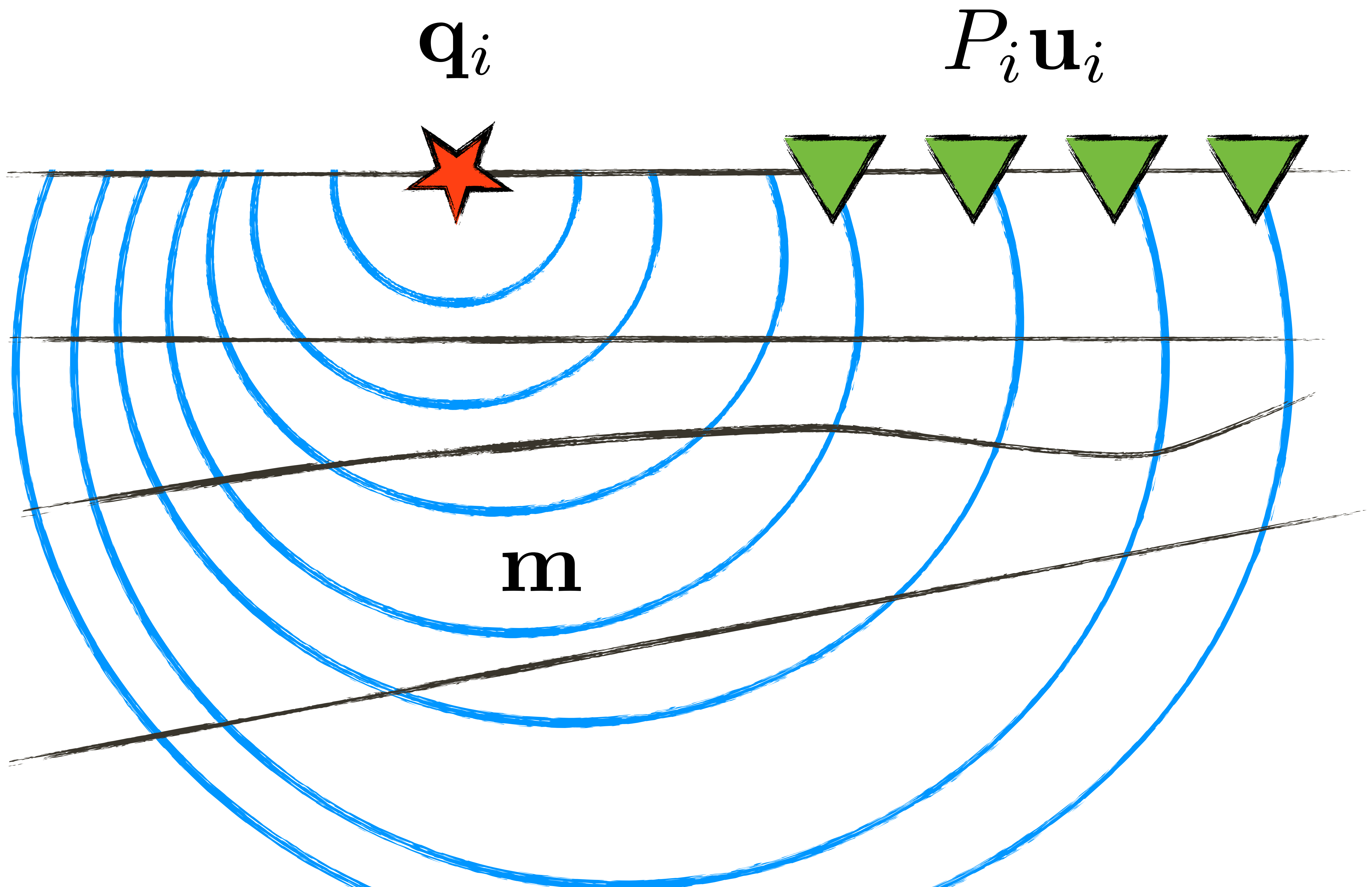
Full-waveform inversion is plagued with local minima

Derive an alternative extended formulation

- less prone to local minima
- computationally feasible
- relaxes the physics while staying solidly grounded

Waveform inversion

Retrieve the medium parameters from partial measurements of the solution of the wave-equation: $A(\mathbf{m})\mathbf{u}_i = \mathbf{q}_i$



[Tarantola, ’84; Pratt, ’98; Haber, ’00; Plessix, ’06]

Waveform inversion

Adjoint-state/reduced-space methods:

- Optimize over earth models to minimize the misfit between observed and simulated data while solving the wave equation exactly for each earth model.

Full-space or all-at-once methods:

- Optimize over earth models & wavefields jointly to minimize the misfit between observed and simulated data subject to wavefields that satisfy the wave equation.

SLIM

Waveform inversion

Both approaches assume *flawless* wave physics–i.e.,

"known" physics → $A(\mathbf{m})$; "known" source → $\mathbf{q}_i$; "unknown" wavefield → $\mathbf{u}_i$

$$A(\mathbf{m})\mathbf{u}_i = \mathbf{q}_i$$

- holds exactly for each source *i*
- differ on insisting wave equations to hold for each iteration
- different unknowns: **m** ⟷ **m** & **u**

[Richter, '81]

Equation error approach

If we "know" the wavefields everywhere, we solve for $\mathbf{m}$ from

$$A(\mathbf{m})\mathbf{u}_i = \mathbf{q}_i$$

via

$$\min_{\mathbf{m}} \|A(\mathbf{m})P_i^{-1}\mathbf{d}_i - \mathbf{q}_i\|_2^2 \qquad \left(\text{cf. } \min_{\mathbf{m}} \|P_i A(\mathbf{m})^{-1}\mathbf{q}_i - \mathbf{d}_i\|_2^2\right)$$

The challenge is to reconstruct wavefields from partial measurements...

[van Leeuwen & FJH, 2013]

SLIM

WRI – Wavefield Reconstruction Inversion

For $\mathbf{m}$ fixed, reconstruct wavefields by jointly fitting observed shots

$$P\mathbf{u}_i \approx \mathbf{d}_i$$

and wave-equations

$$A(\mathbf{m})\mathbf{u}_i \approx \mathbf{q}_i$$

via least-squares solutions of the data-augmented wave-equation

$$\min_{\mathbf{u}_i} \left\| \begin{pmatrix} P_i \\ A(\mathbf{m}) \end{pmatrix} \mathbf{u}_i - \begin{pmatrix} \mathbf{d}_i \\ \mathbf{q}_i \end{pmatrix} \right\|_2^2$$

followed by fixing $\mathbf{u}_i$ and solving

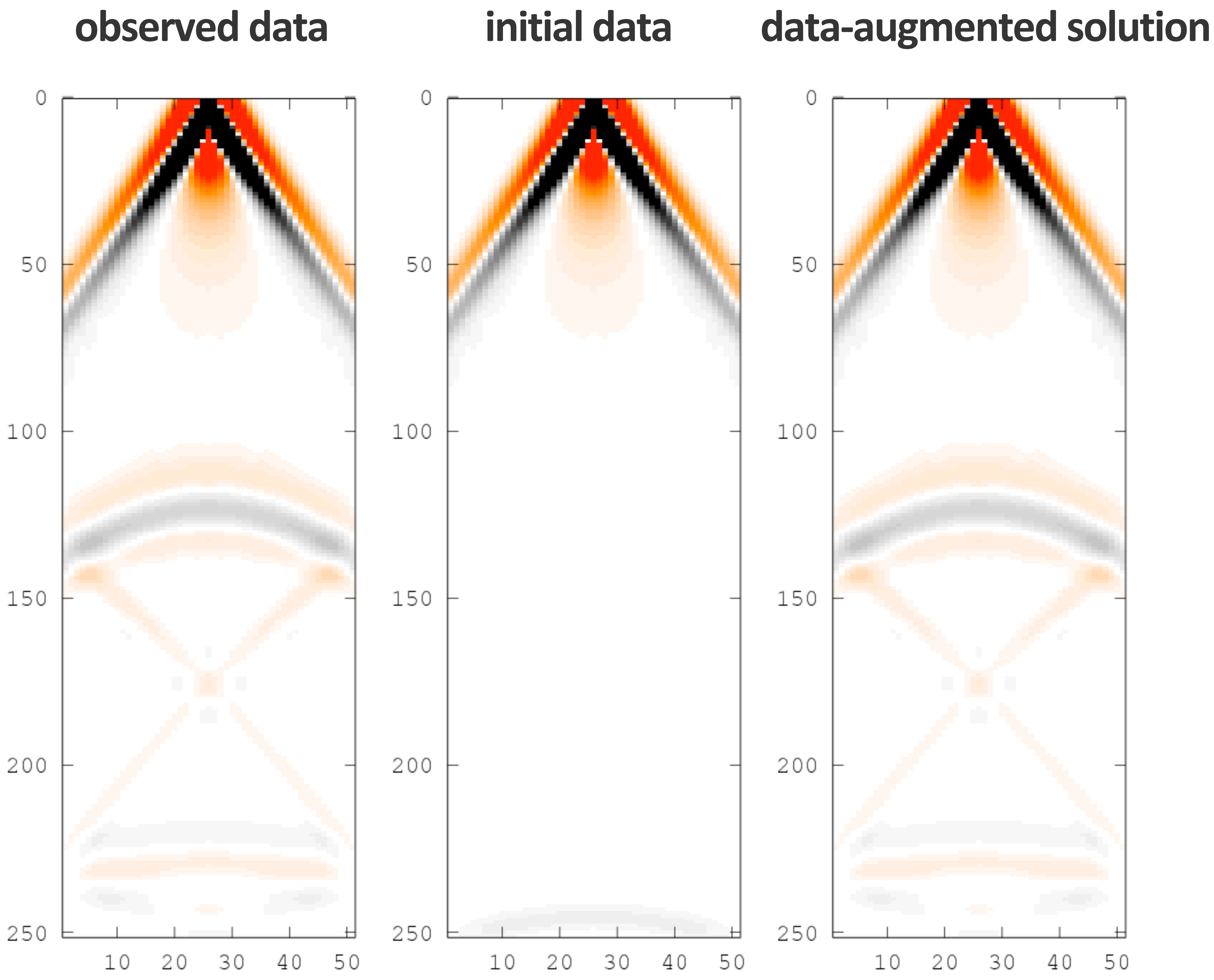
$$\min_{\mathbf{m}} \|A(\mathbf{m})\mathbf{u}_i - \mathbf{q}_i\|_2^2$$

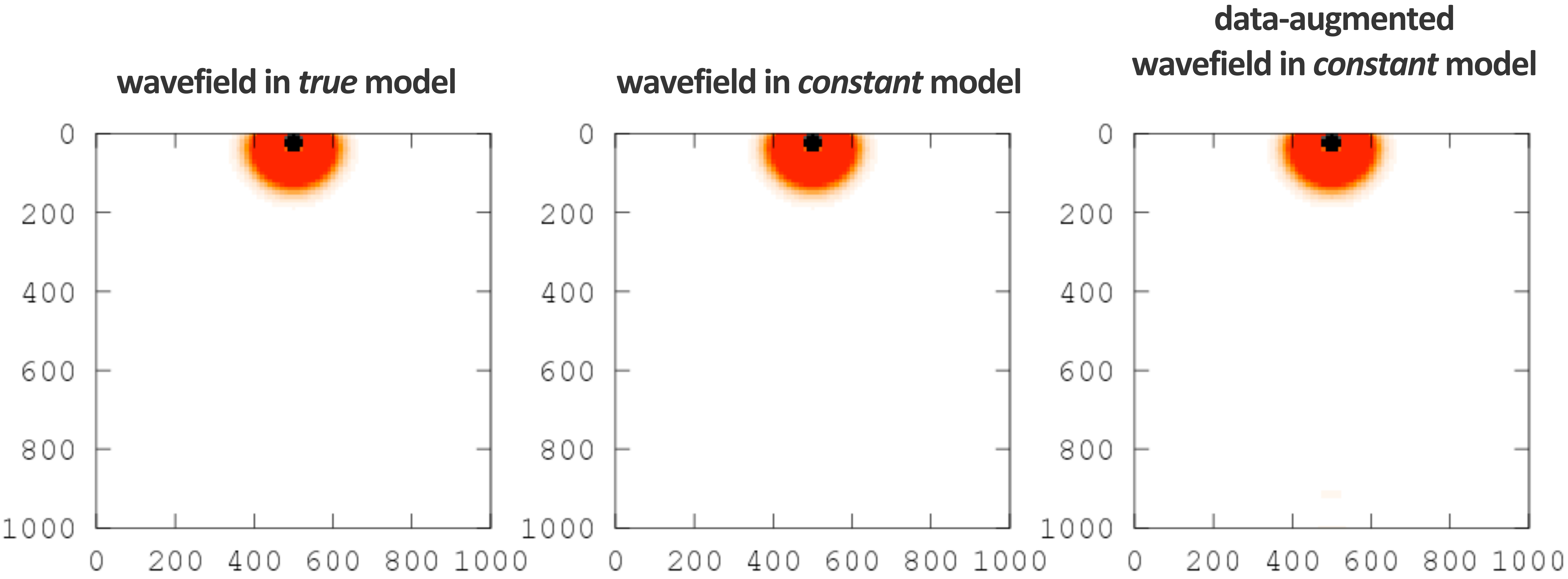
SLIM

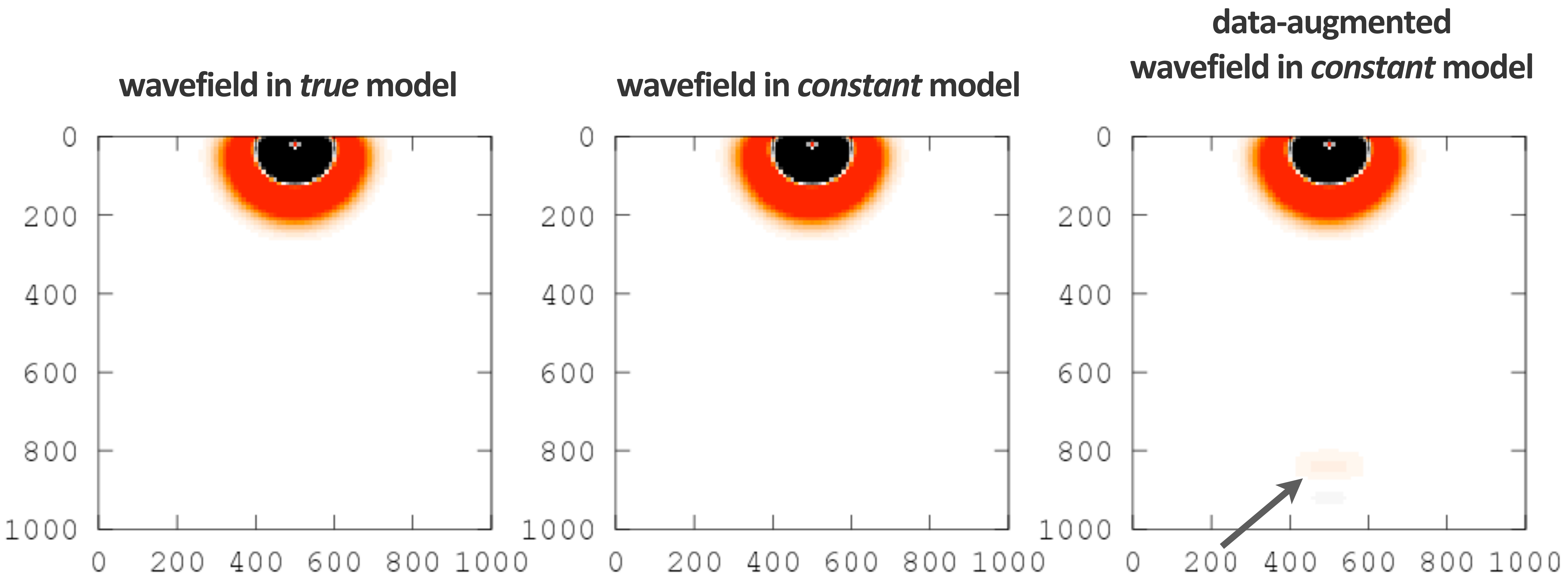
wave-equation X wavefield = source

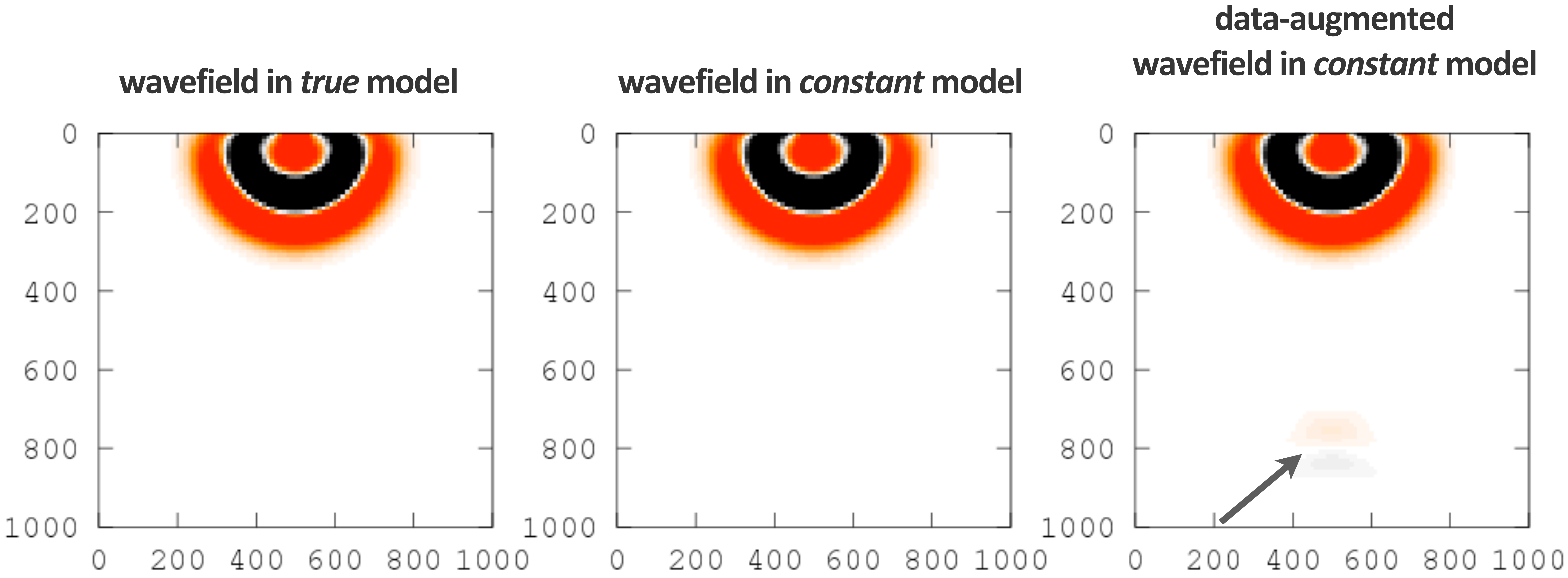
versus

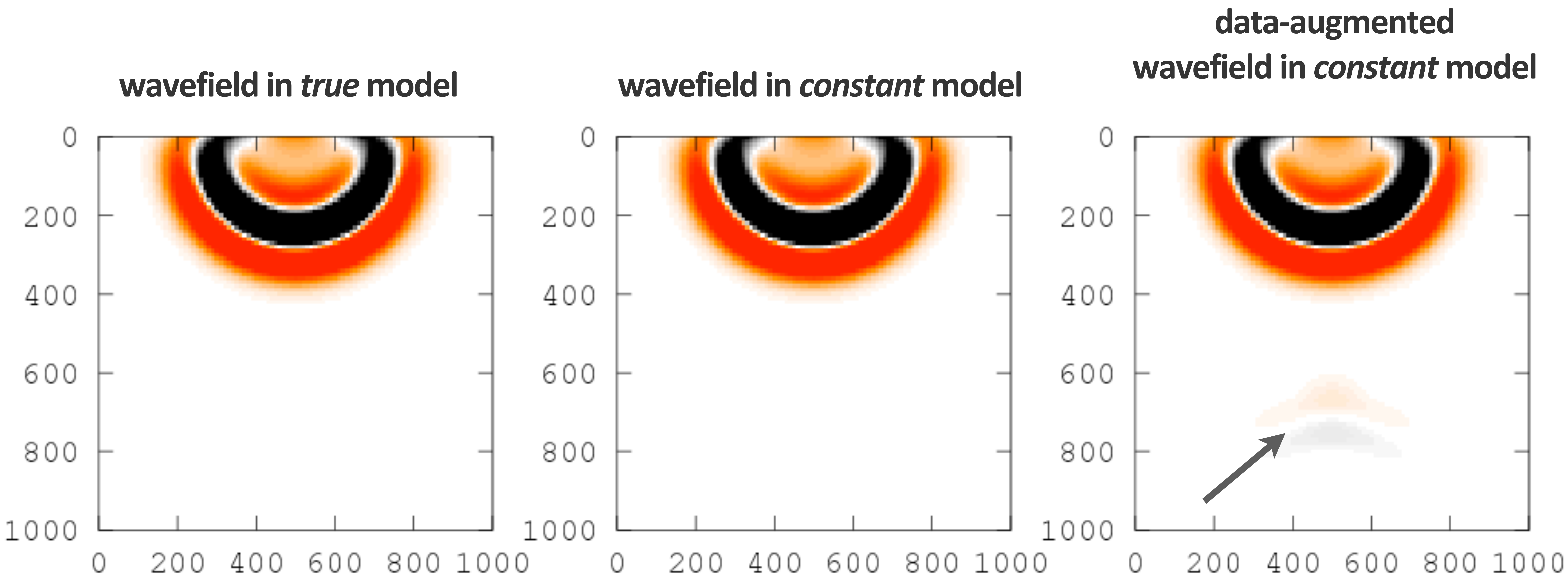
(wave-equation / sampling operator) X wavefield = (source / data)


observed data
initial data
data-augmented solution
0
50
100
150
200
250
10
20
30
40
50

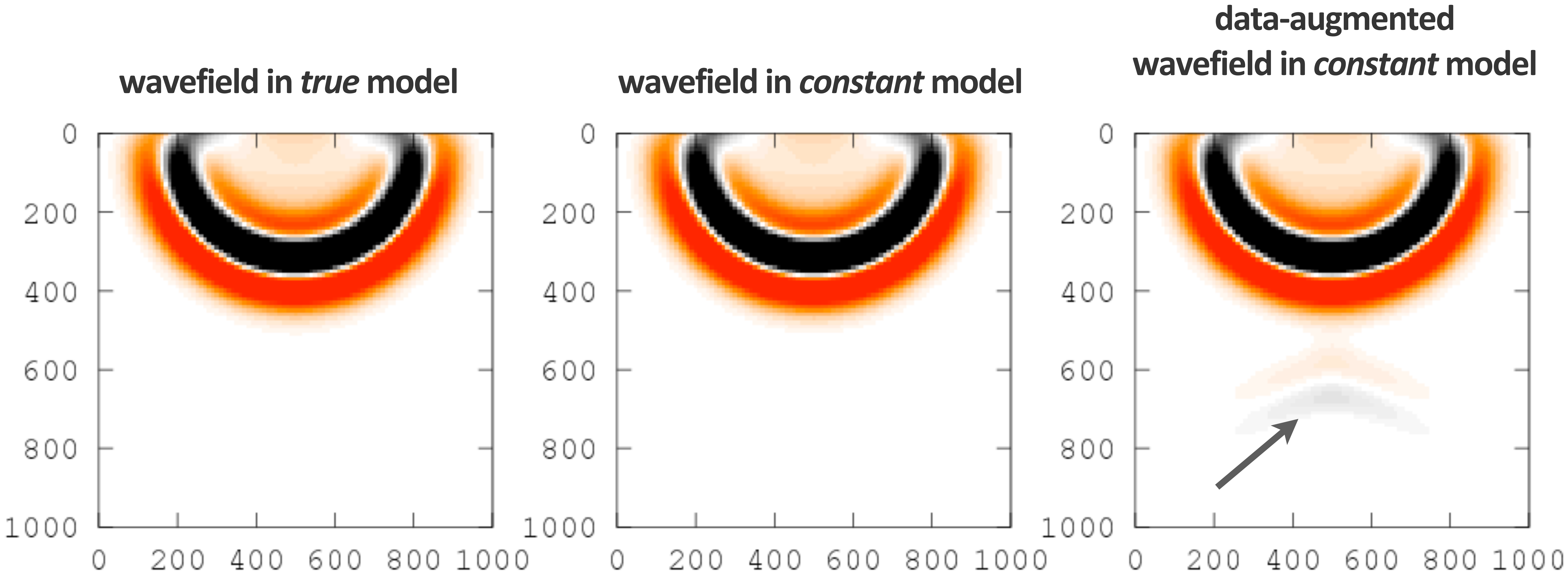

wavefield in *true* model
wavefield in *constant* model
data-augmented
wavefield in *constant* model
0
200
400
600
800
1000
0 200 400 600 800 1000

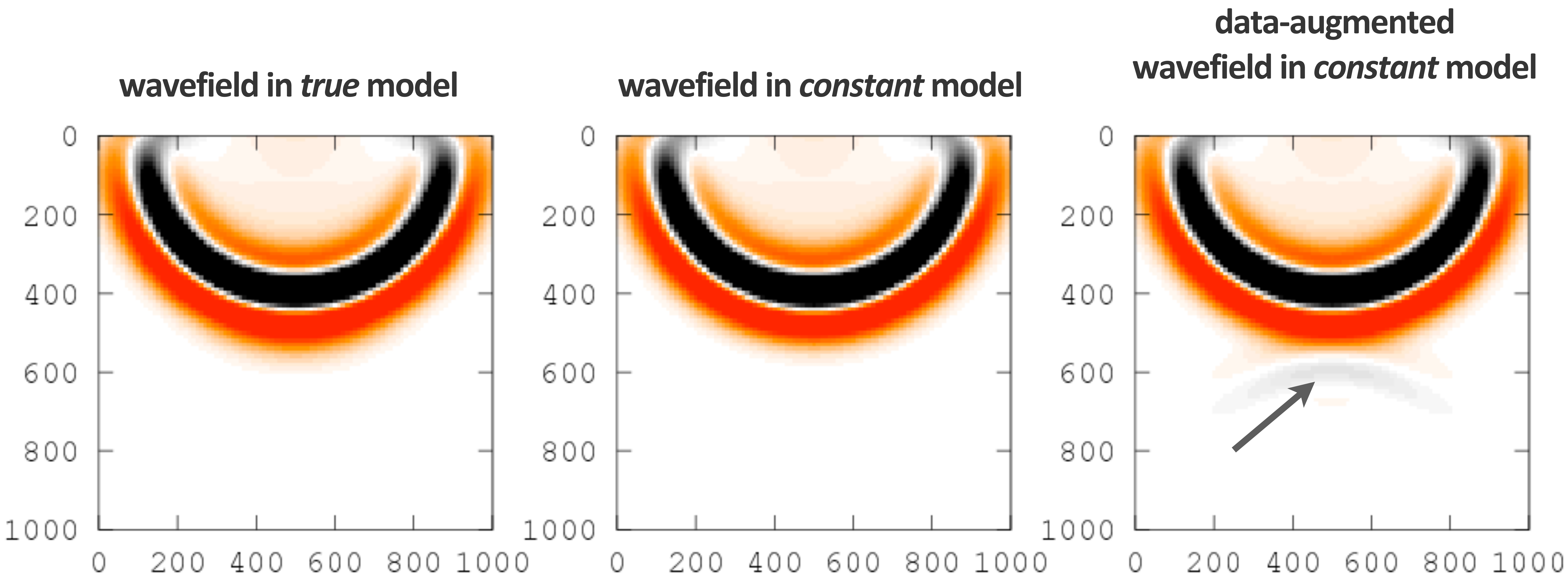

wavefield in *true* model
wavefield in *constant* model
data-augmented
wavefield in *constant* model
0
200
400
600
800
1000
0
200
400
600
800
1000


wavefield in *true* model
wavefield in *constant* model
data-augmented
wavefield in *constant* model



wavefield in *true* model
wavefield in *constant* model
data-augmented
wavefield in *constant* model
0
200
400
600
800
1000
0 200 400 600 800 1000

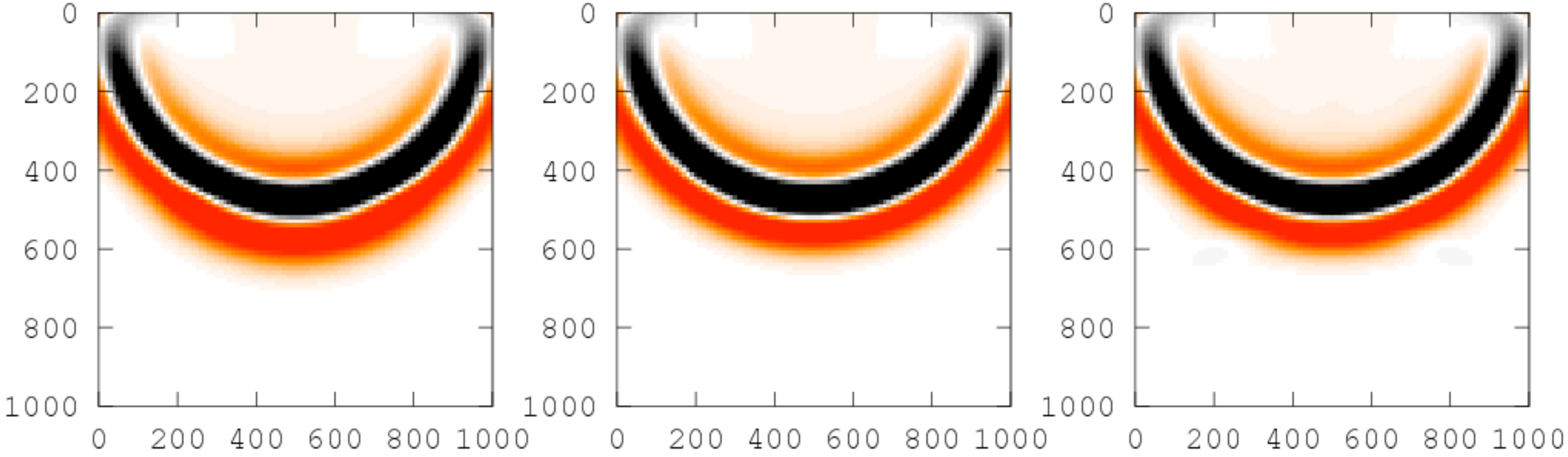

wavefield in *true* model
wavefield in *constant* model
data-augmented
wavefield in *constant* model
0
200
400
600
800
1000
0 200 400 600 800 1000

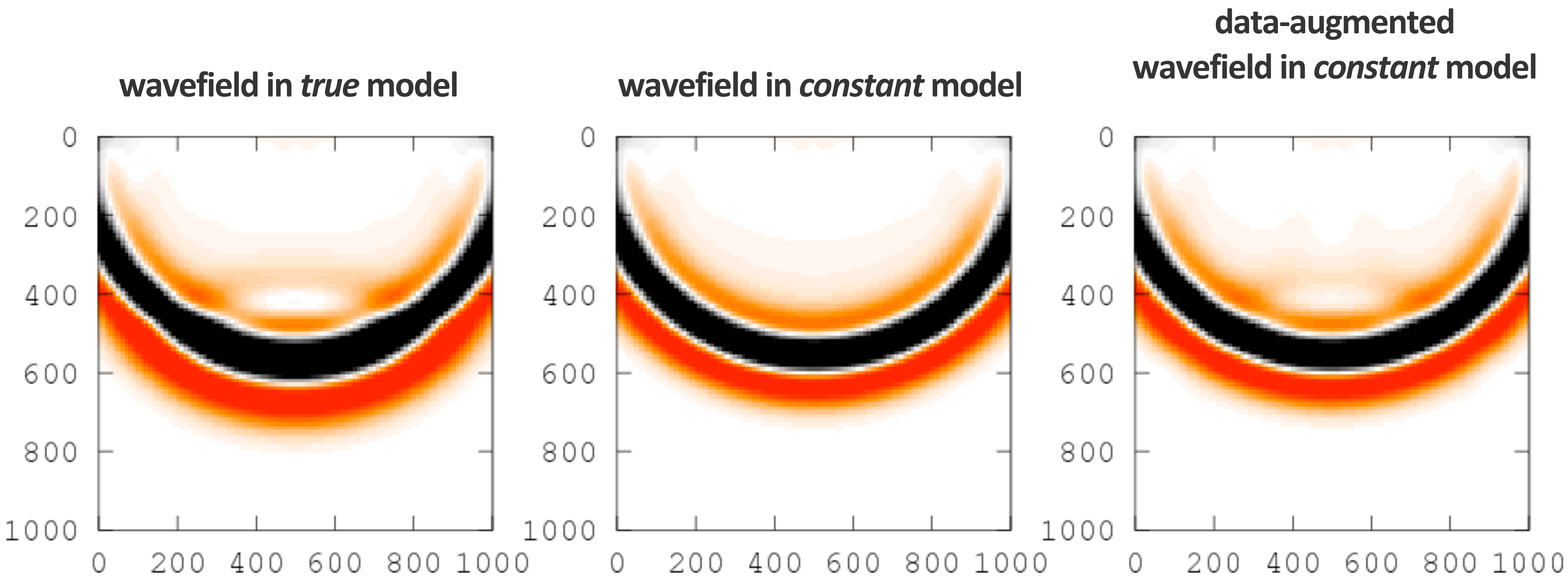

wavefield in *true* model
wavefield in *constant* model
data-augmented wavefield in *constant* model
0
200
400
600
800
1000

wavefield in *true* model

wavefield in *constant* model

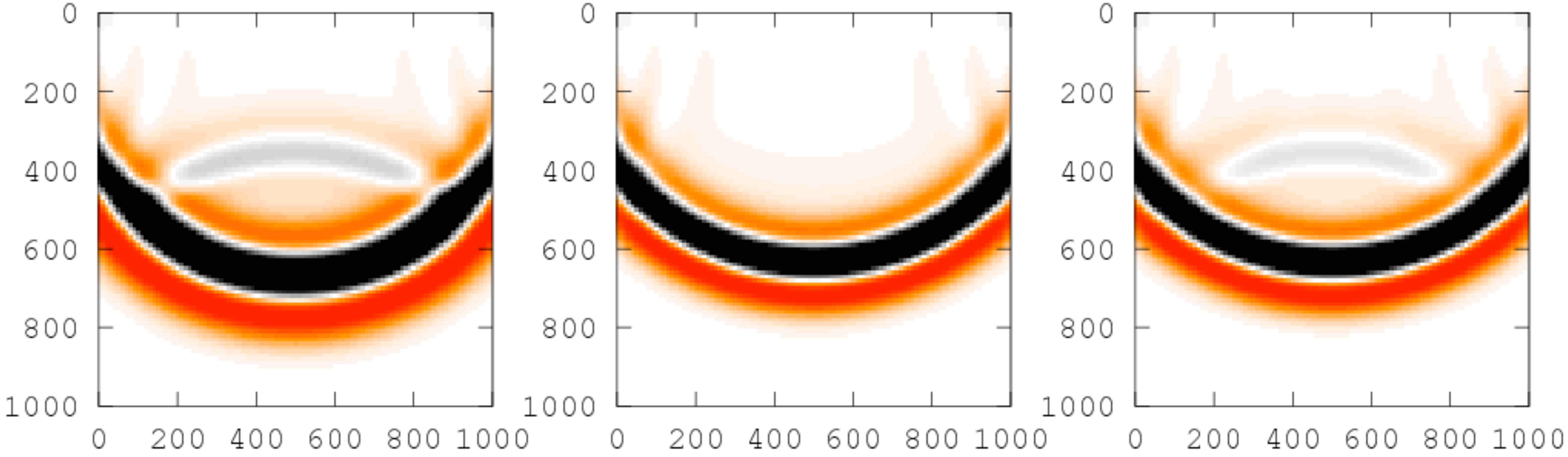
data-augmented wavefield in *constant* model



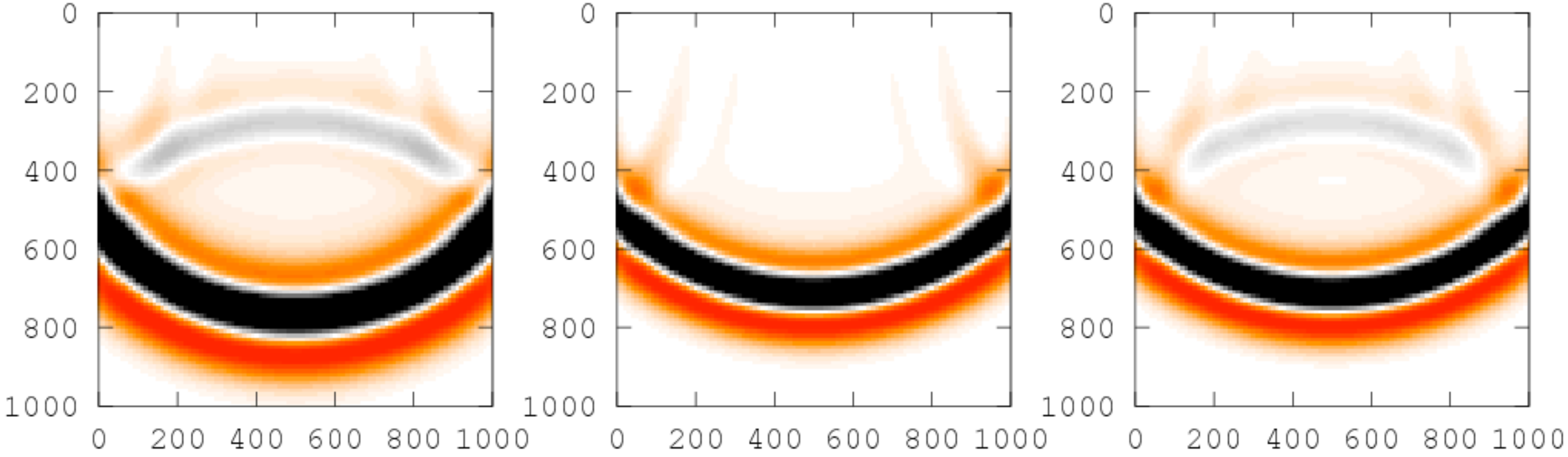


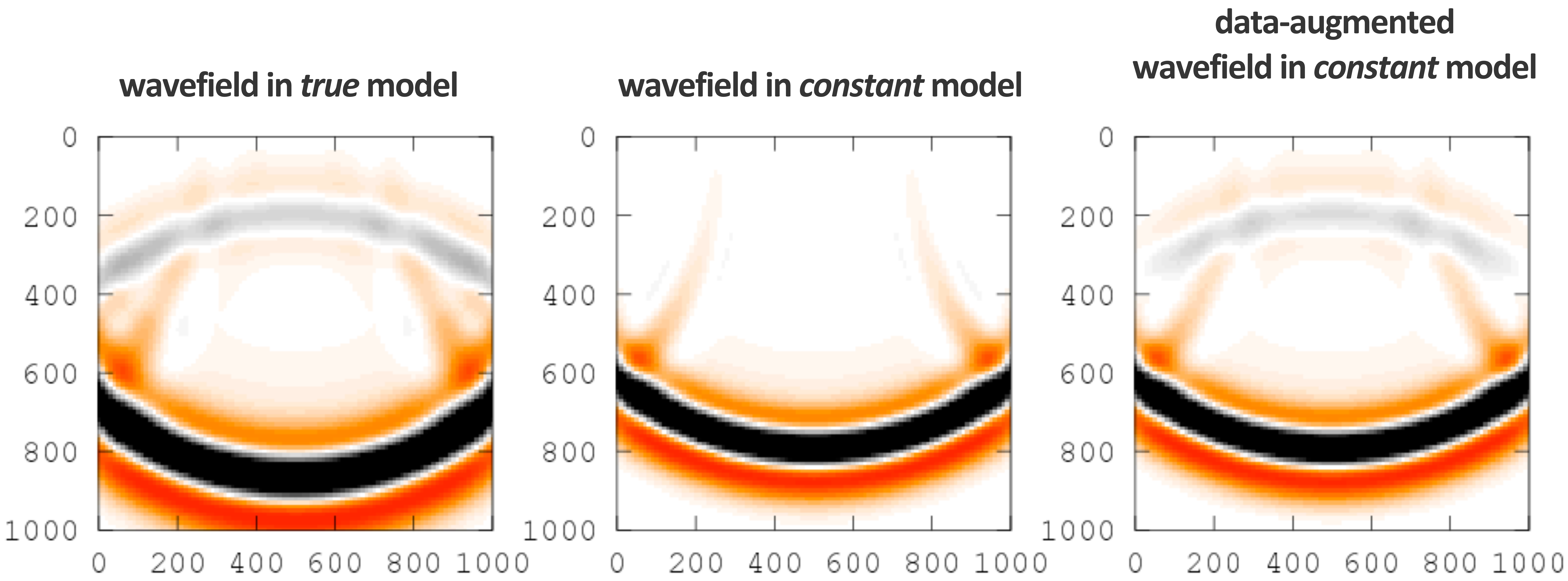
wavefield in *true* model
wavefield in *constant* model
data-augmented
wavefield in *constant* model
0
200
400
600
800
1000
0 200 400 600 800 1000

wavefield in *true* model | **wavefield in *constant* model** | **data-augmented wavefield in *constant* model**



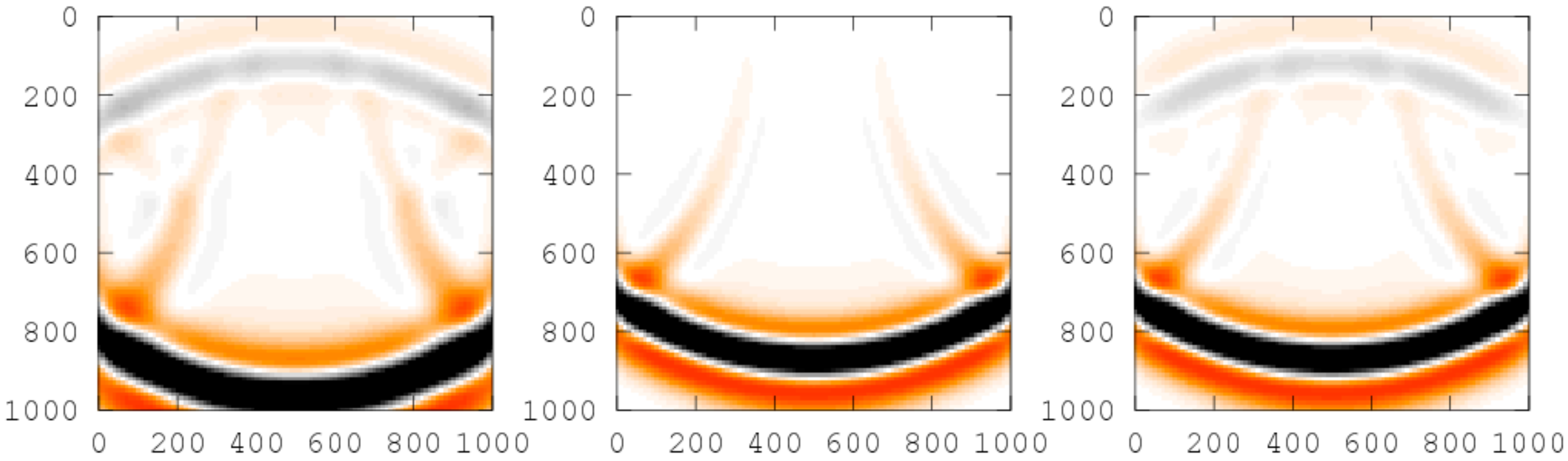
wavefield in *true* model | **wavefield in *constant* model** | **data-augmented wavefield in *constant* model**

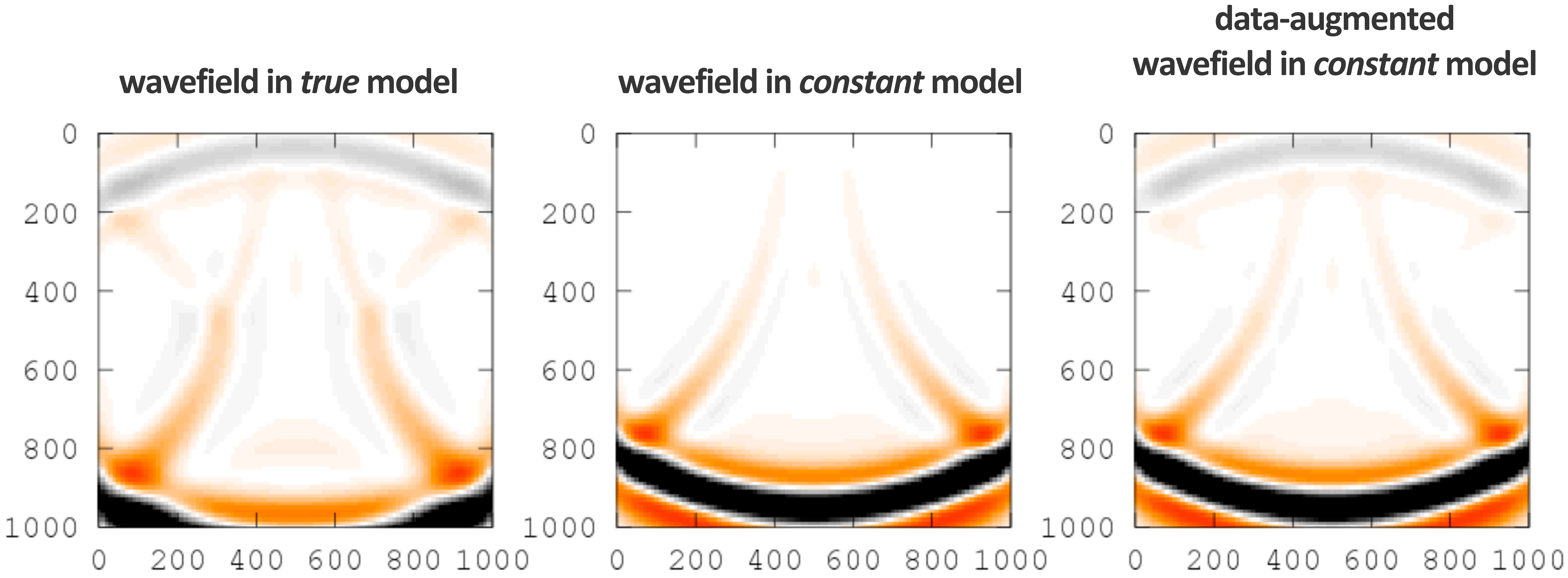


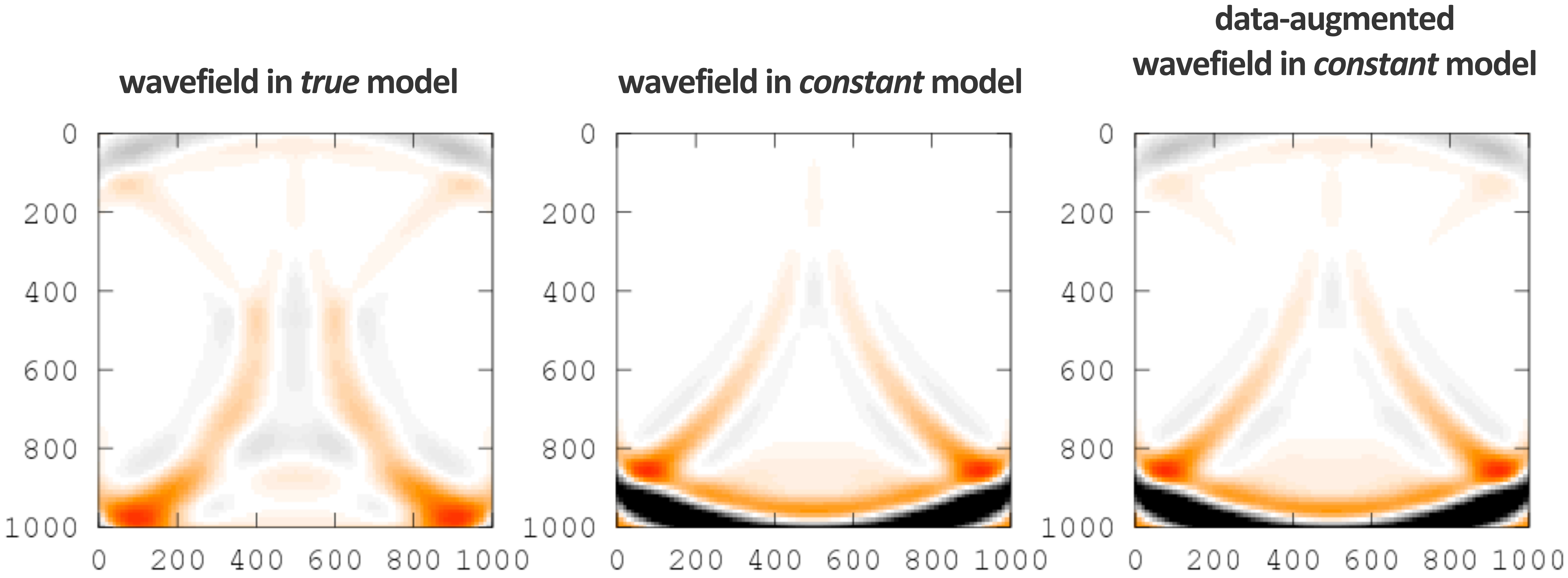


wavefield in *true* model
wavefield in *constant* model
data-augmented
wavefield in *constant* model
0
200
400
600
800
1000
0 200 400 600 800 1000

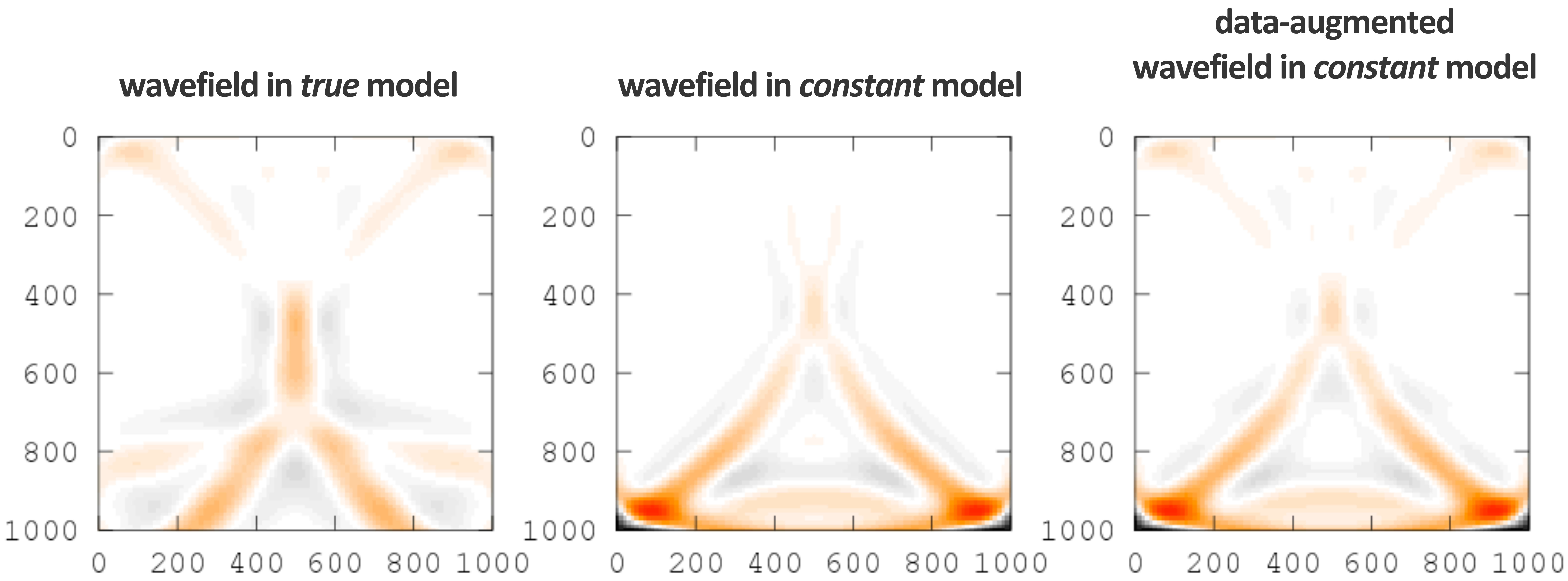
wavefield in *true* model | wavefield in *constant* model | data-augmented wavefield in *constant* model

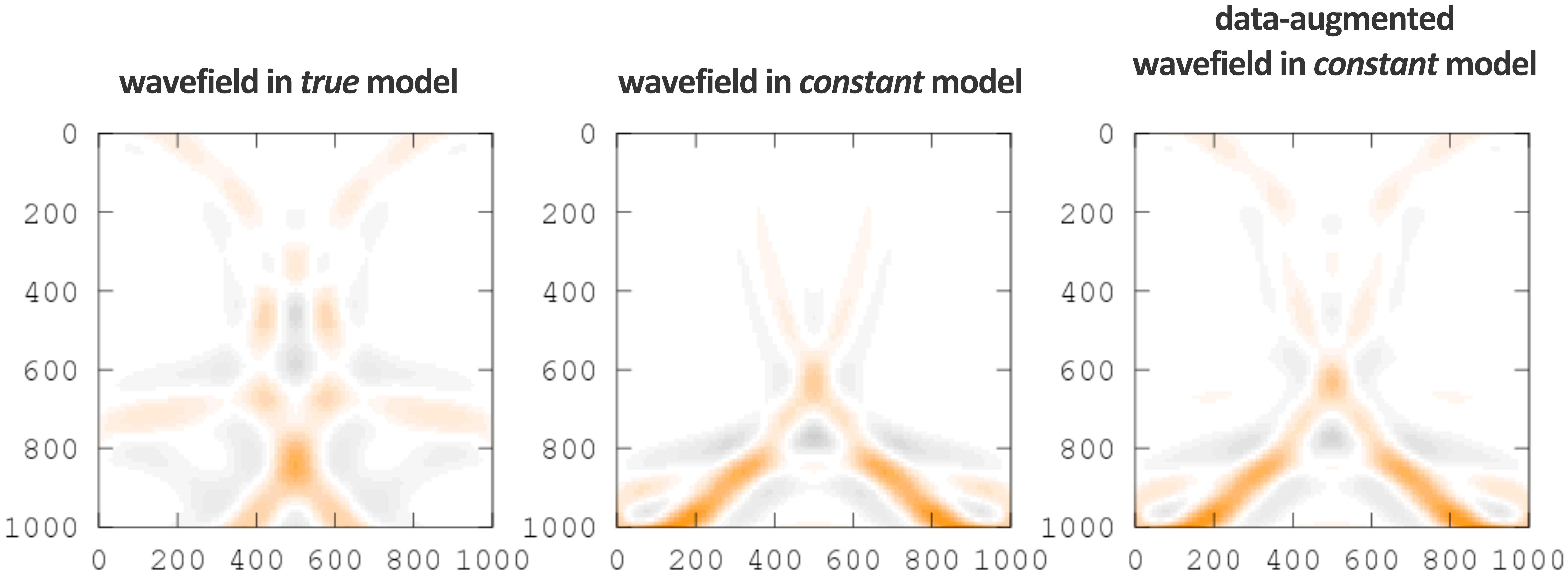



wavefield in *true* model
wavefield in *constant* model
data-augmented wavefield in *constant* model
0
200
400
600
800
1000
0 200 400 600 800 1000



wavefield in *true* model
wavefield in *constant* model
data-augmented
wavefield in *constant* model
0
200
400
600
800
1000
0 200 400 600 800 1000


wavefield in *true* model
wavefield in *constant* model
data-augmented
wavefield in *constant* model
0
200
400
600
800
1000


wavefield in *true* model
wavefield in *constant* model
data-augmented
wavefield in *constant* model
0
200
400
600
800
1000

[Heinkenschloss, ’98 , Haber, ’00]

PDE-constrained optimization
all-at-once full-space approach

simulated data (↓ $P_i\mathbf{u}_i$), simulated wavefield (↓ $\mathbf{u}_i$)

$$\min_{\mathbf{m},\mathbf{u}} \sum_{i=1}^{M} \|P_i\mathbf{u}_i - \mathbf{d}_i\|_2^2 \quad \text{s.t.} \quad A_i(\mathbf{m})\mathbf{u}_i = \mathbf{q}_i$$

observed data (↑ $\mathbf{d}_i$), Helmholtz equation (↑ A_i), source (↑ $\mathbf{q}_i$)

- avoids having to solve the PDE explicitly
- sparse (GN) Hessian
- requires storing all variables **(m,u)**
- does **not** scale to industry-scale seismic problems

[Tarantola ’84; Pratt, ’98; Plessix, ’06]

SLIM

Adjoint-state/reduced-space formulation

Elimination of the constraint leads for all sources to

$$\min_{\mathbf{m}} \phi_{\text{red}}(\mathbf{m}) = \sum_{i=1}^{M} \|P_i A_i(\mathbf{m})^{-1}\mathbf{q}_i - \mathbf{d}_i\|_2^2$$

- no need to store all wavefields (block-elimination)
- suitable for black-box optimization (e.g., l-BFGS)
- need to solve forward & adjoint PDEs
- very non-linear in earth model (m)
- dense (GN) Hessian, involves additional PDE solves
- paints you in a corner by insisting on the physics...

SLIM

WRI – penalty formulation

Instead of eliminating, we add constraints as penalties–i.e.,

$$\min_{\mathbf{m},\mathbf{u}} \phi_\lambda(\mathbf{m},\mathbf{u}) = \sum_{i=1}^{M} \|P\mathbf{u}_i - \mathbf{d}_i\|_2^2 + \lambda^2 || A_i(\mathbf{m})\mathbf{u}_i - \mathbf{q}_i\|_2^2$$

coincides with original problem when $\lambda \uparrow \infty$

[Aravkin & van Leeuwen, ’12; van Leeuwen & FJH, ’13]

SLIM

Variable projection

Solve data-augmented wave equation for each source

$$\begin{pmatrix} P_i \\ \lambda A_i(\mathbf{m}) \end{pmatrix} \mathbf{u}_{i,\lambda} \approx \begin{pmatrix} \mathbf{d}_i \\ \lambda \mathbf{q}_i \end{pmatrix}$$

Define reduced objective with proxy wavefields

$$\phi_\lambda(\mathbf{m}) = \phi_\lambda(\mathbf{m}, \bar{\mathbf{u}}_\lambda) = \|P\bar{\mathbf{u}}_\lambda - \mathbf{d}\|_2^2 + \lambda^2 \|A(\mathbf{m})\bar{\mathbf{u}}_\lambda - \mathbf{q}\|_2^2$$

[van Leeuwen & FJH, ’13]

SLIM

Wavefield Reconstruction Inversion

WRI method

for each source i

solve $\begin{pmatrix} P_i \\ \lambda A_i(\mathbf{m}) \end{pmatrix} \mathbf{u}_{\lambda,i} \approx \begin{pmatrix} \mathbf{d}_i \\ \lambda \mathbf{q}_i \end{pmatrix}$

$\mathbf{g} = \mathbf{g} + \lambda^2 \omega^2 \mathsf{diag}(\bar{\mathbf{u}}_{i,\lambda})^* (A(\mathbf{m}) \bar{\mathbf{u}}_{i,\lambda} - \mathbf{q}_i)$

$\mathbf{m} = \mathbf{m} - \alpha \mathbf{g}$

end

correlation proxy wavefield & PDE residual

Conventional method

for each source i

solve $A(\mathbf{m}) \mathbf{u}_i = \mathbf{q}_i$

solve $A(\mathbf{m})^* \mathbf{v}_i = P_i^* (P_i \mathbf{u}_i - \mathbf{d}_i)$

$\mathbf{g} = \mathbf{g} + \omega^2 \mathsf{diag}(\mathbf{u}_i)^* \mathbf{v}_i$

$\mathbf{m} = \mathbf{m} - \alpha \mathbf{g}$

end

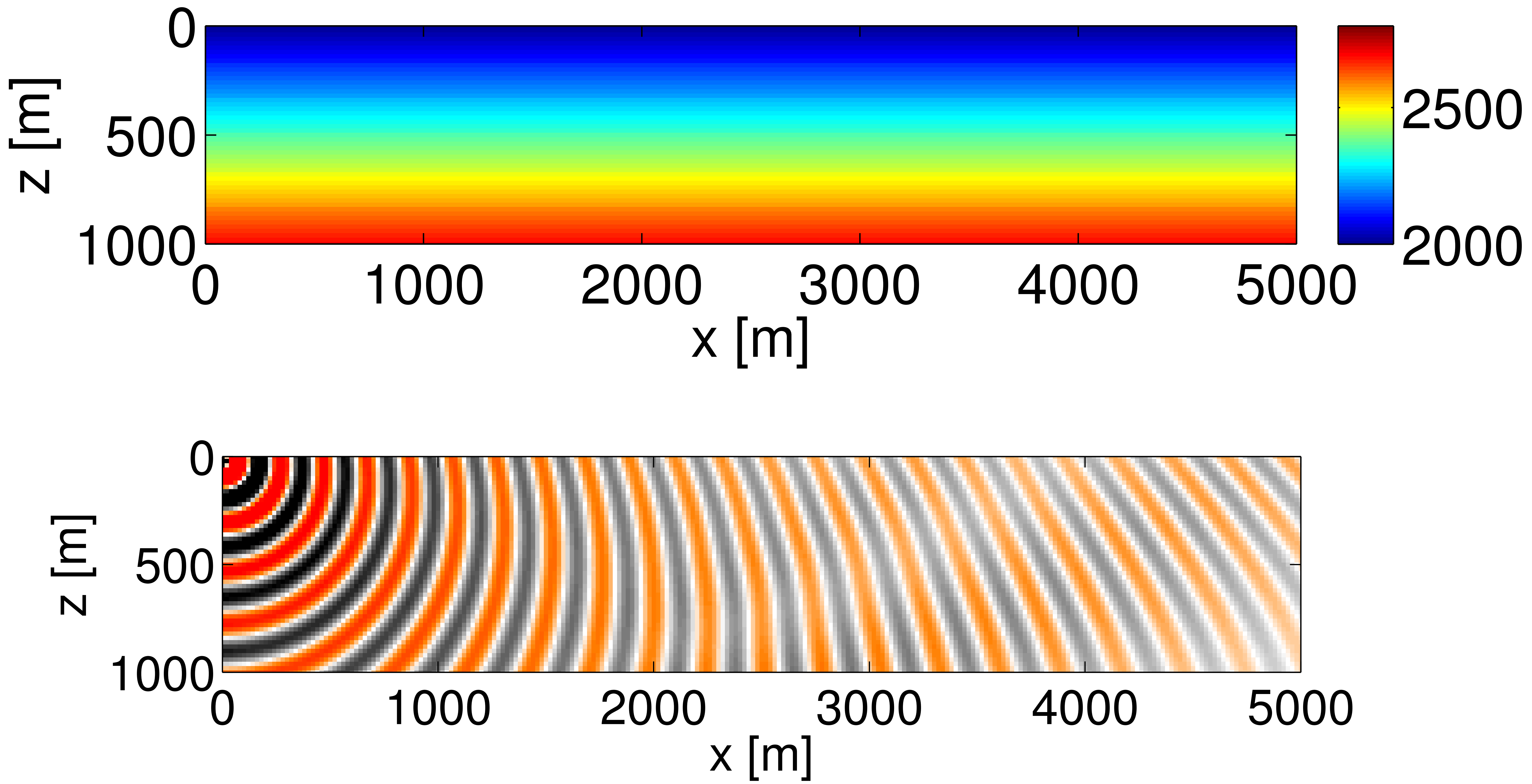
correlation wavefield & data residual

Wavefield Reconstruction Inversion

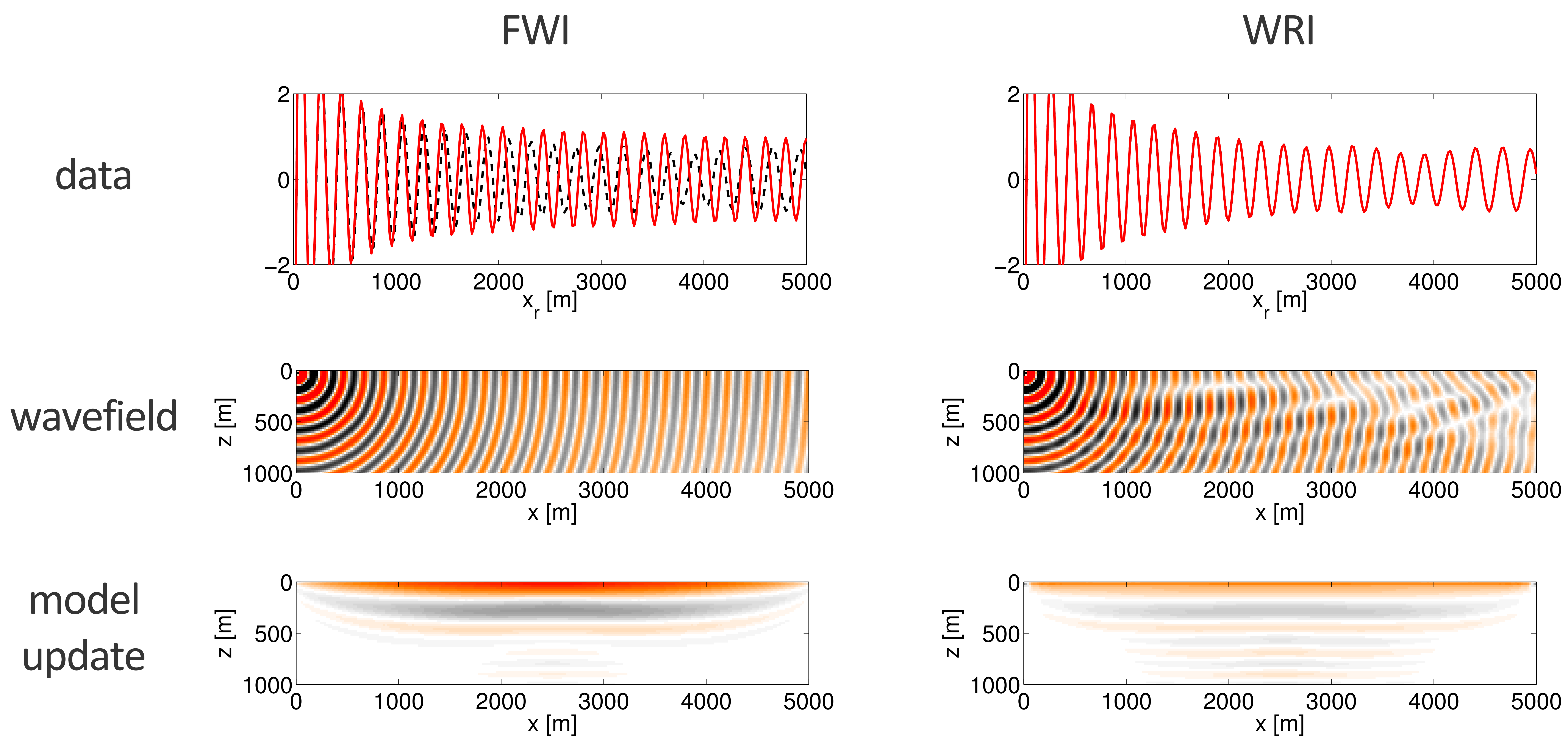
- **no** need to store all the fields (**u**)
- **no** adjoint solves
- sparse approximation of GN Hessian for small λ
- less non-linear in **m**
- need to solve overdetermined PDEs
- ...

Diving wave example

true model and wavefield

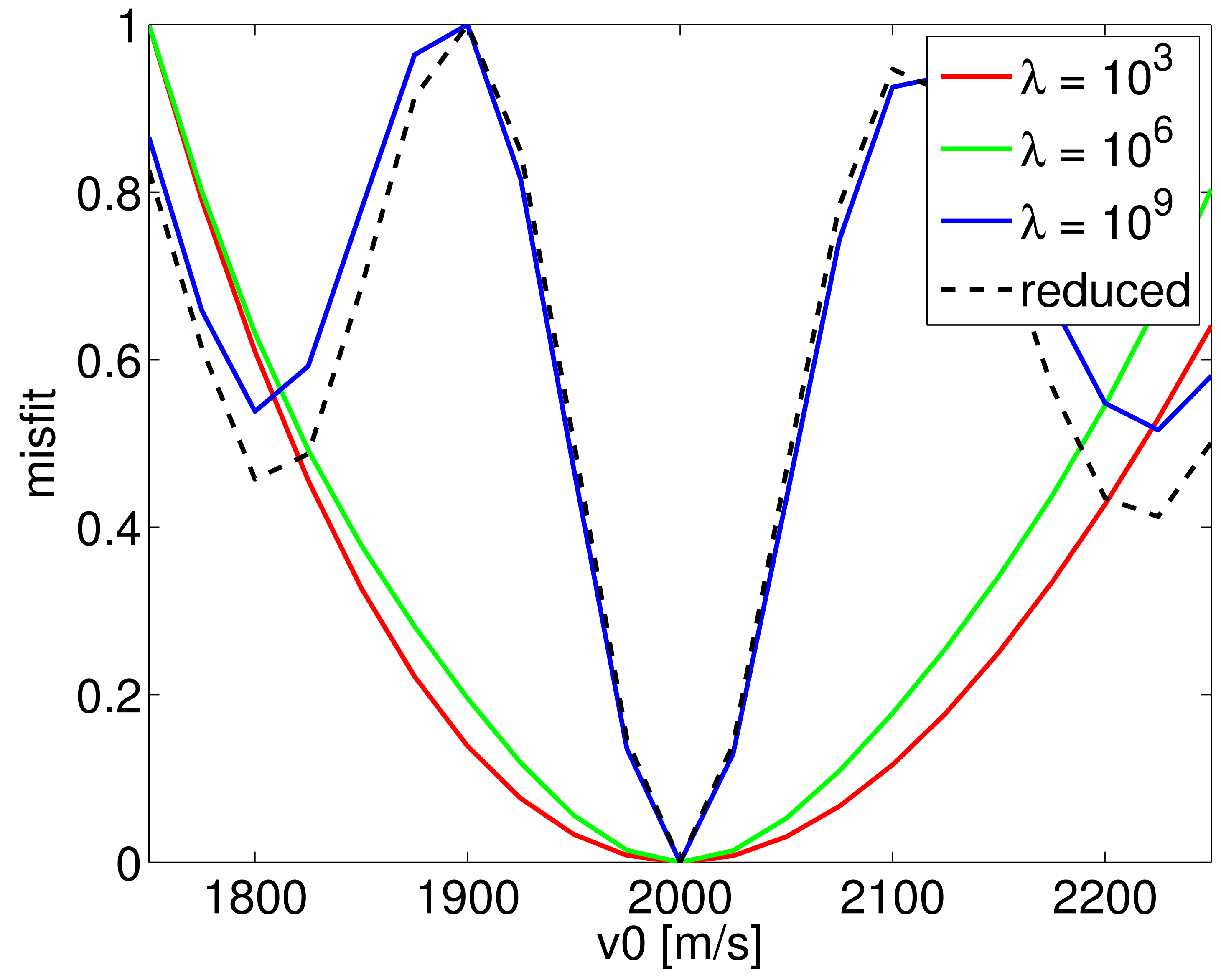


Wavefields in *homogeneous* background

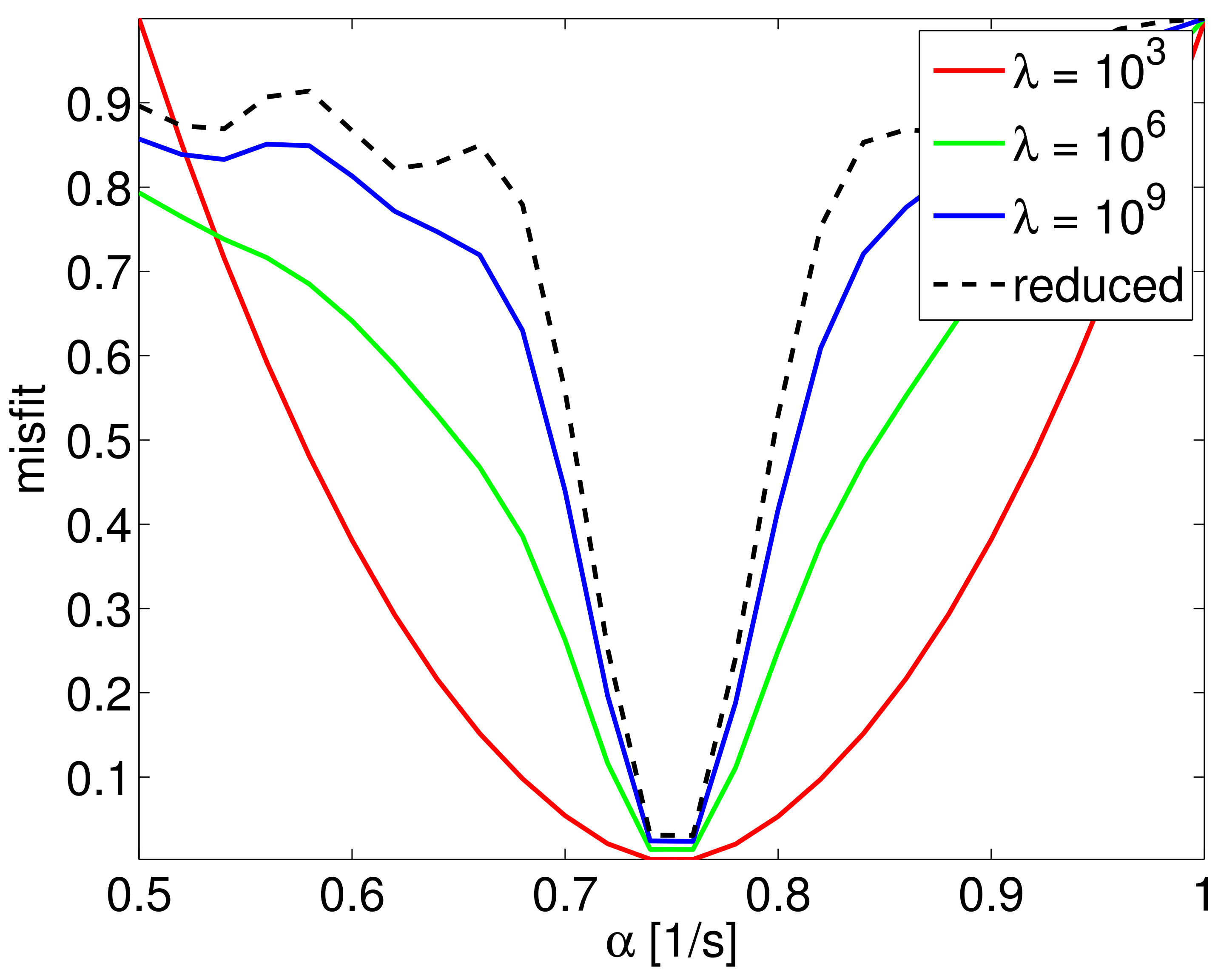


Local minima

single shot, single frequency data for *linear* velocity profile $v(z) = v_0 + \alpha z$,
misfit as function of (v_0, α)



constant



slope

Connections

[Aria Abubakar et. al. ’09]

SLIM

Related work

Contrast-source formulation

- combined objective is similar
- but does not eliminate wavefields via variable projection
- **requires storage of wavefields for all sources**

Extended modelling

The penalty formulation

$$\min_{\mathbf{m},\mathbf{u}} ||P\mathbf{u} - \mathbf{d}||_2^2 + \lambda^2 ||A(\mathbf{m})\mathbf{u} - \mathbf{q}||_2^2$$

can be interpreted as

$$\min_{\tilde{\mathbf{m}}} \text{misfit}(\tilde{\mathbf{m}}) + \text{annihilator}(\tilde{\mathbf{m}})$$

with

$$\tilde{\mathbf{m}} = (\mathbf{m}, \mathbf{u})$$

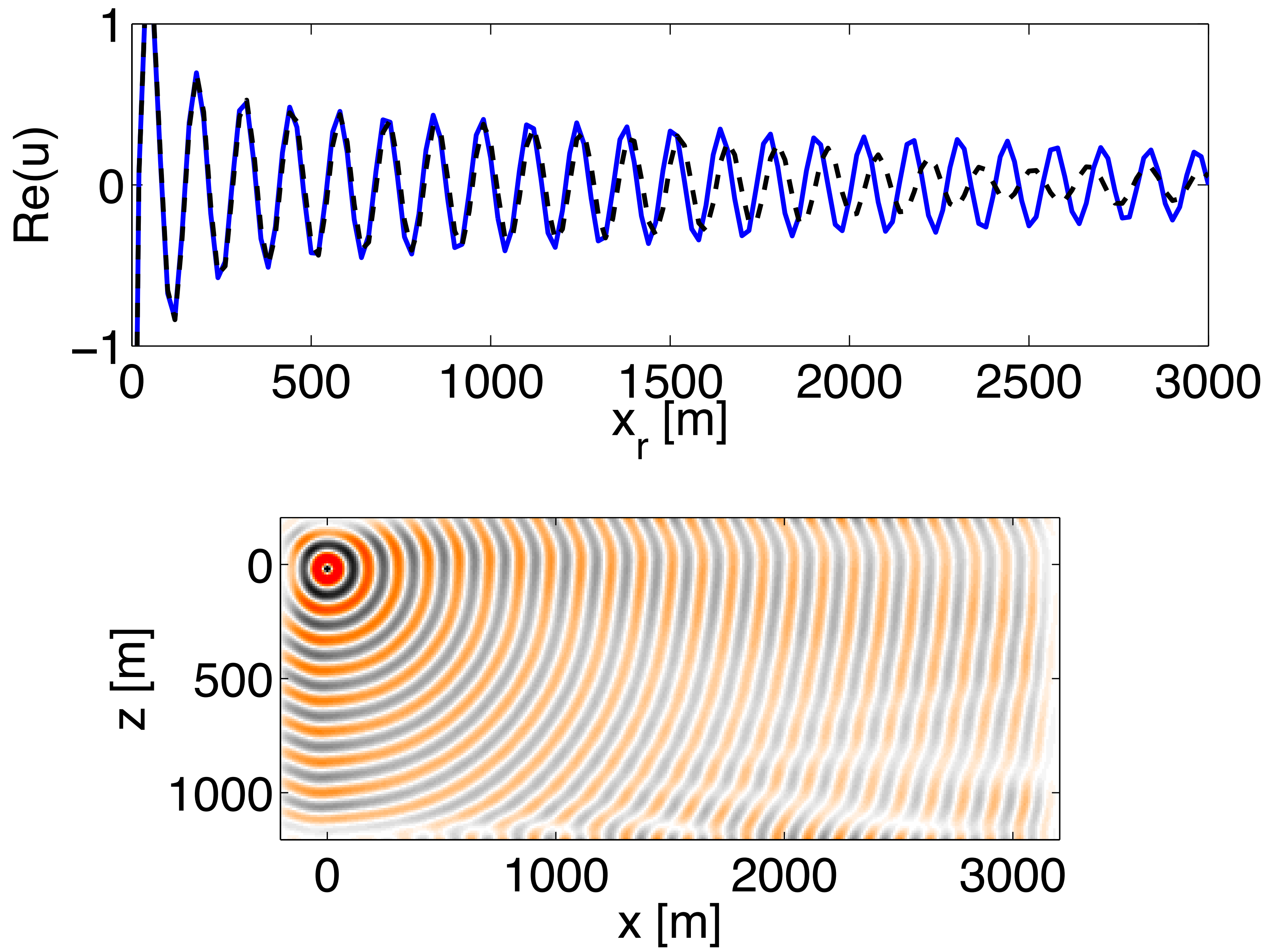
For a physically plausible model we have

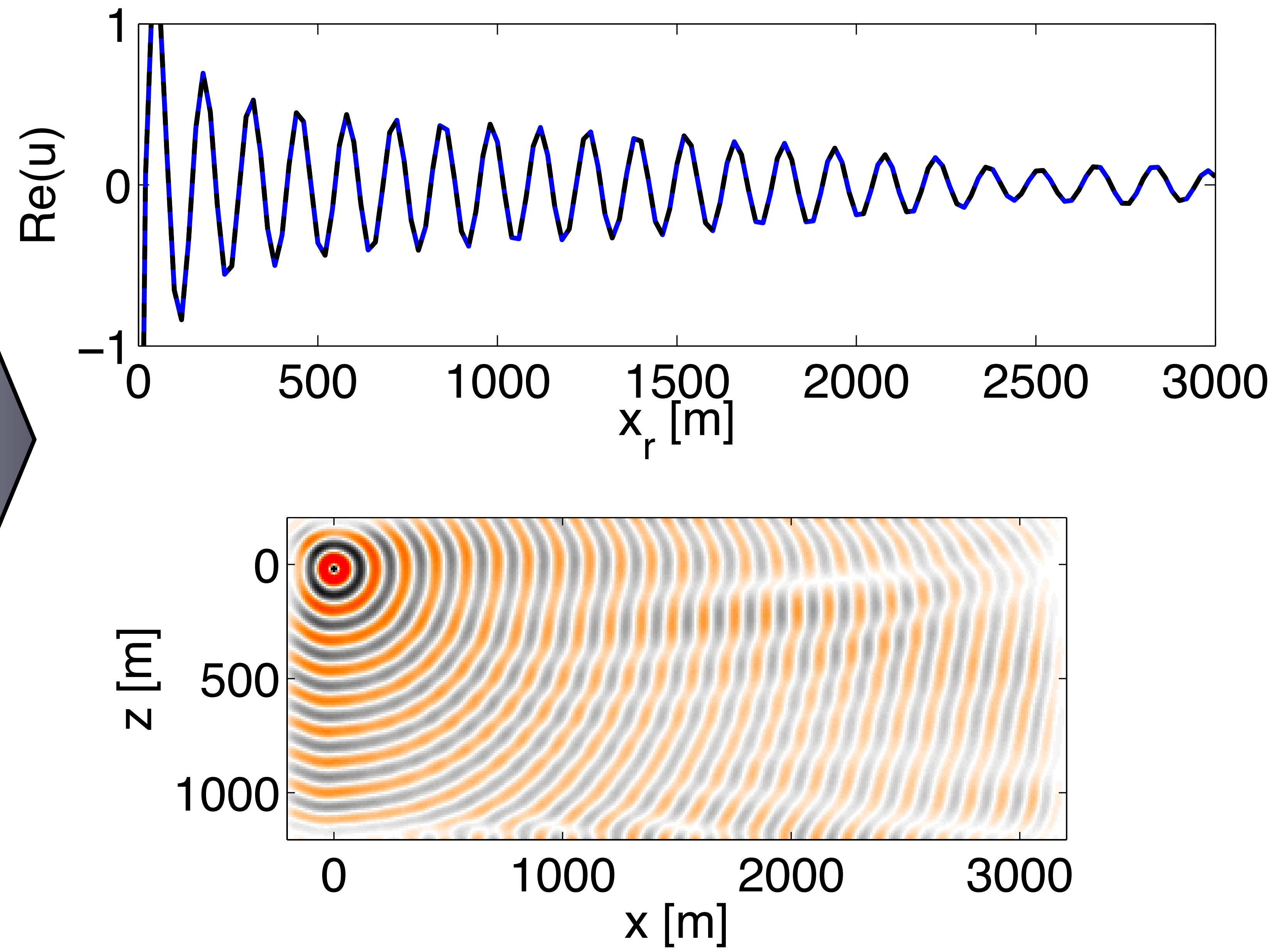
$$\text{annihilator}(\tilde{\mathbf{m}}) = 0$$

[Symes, personal communication]

Warping

The overdetermined WE is a way of warping



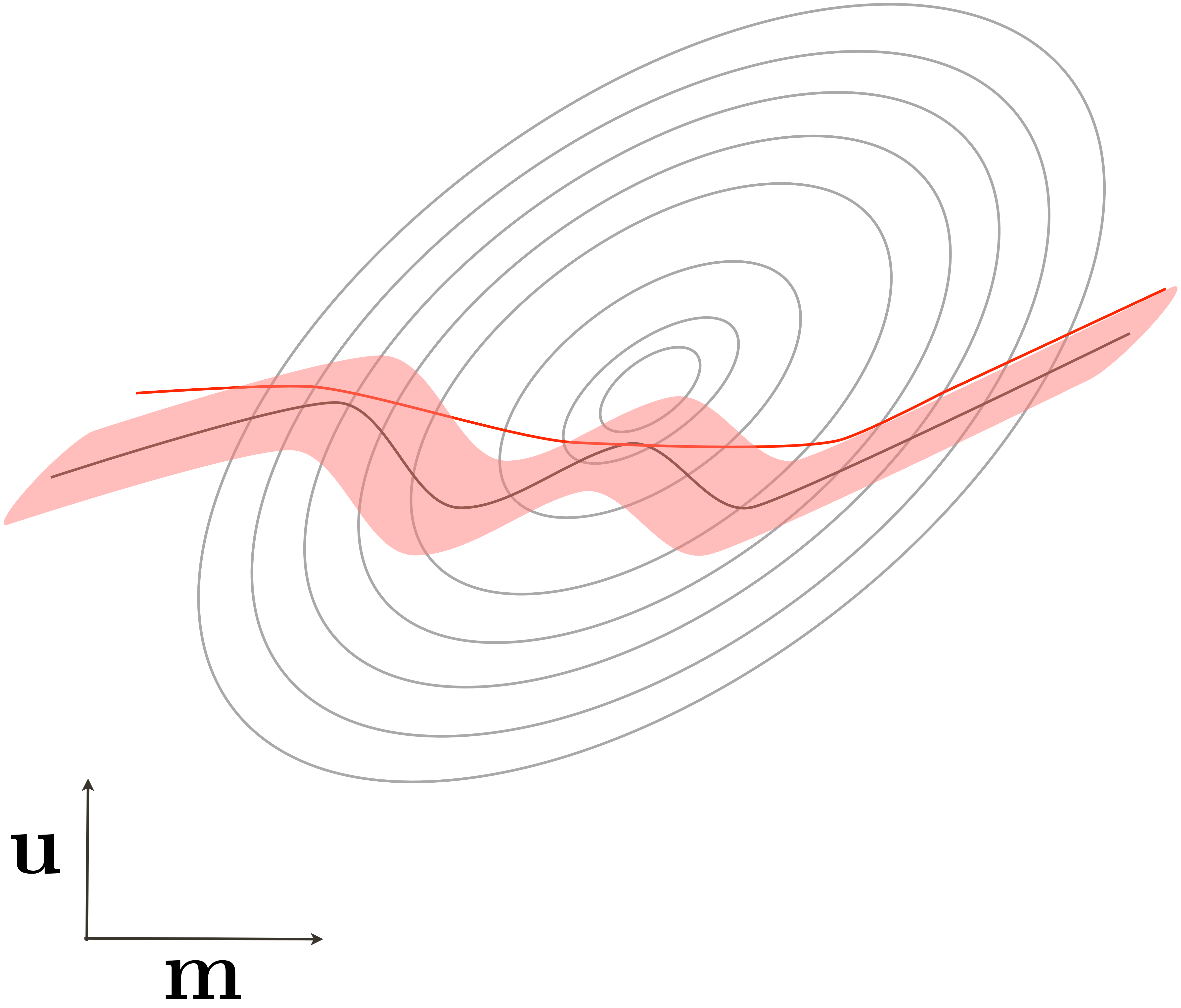


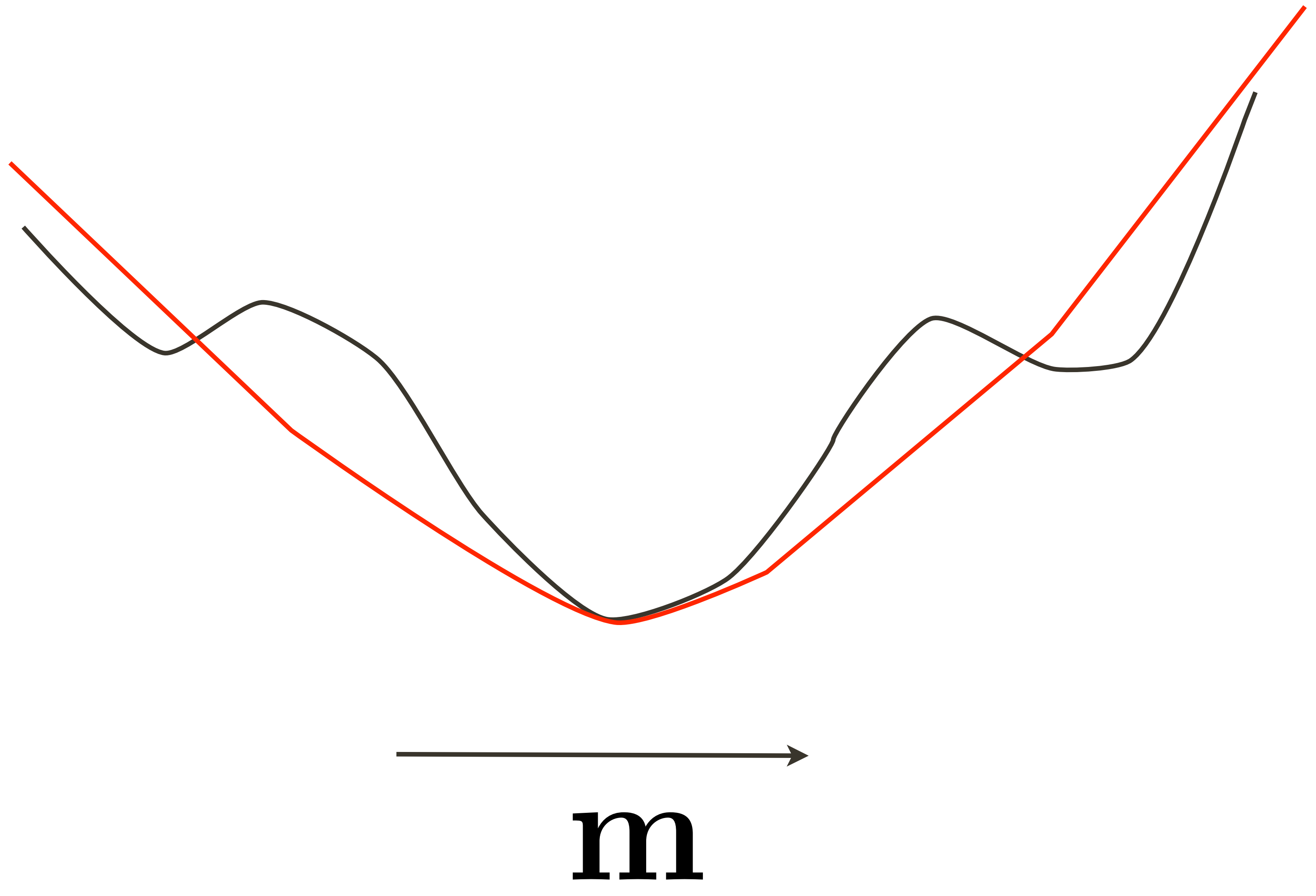
[Baek ’13, Ma ’13]

WRI vs. FWI

Larger # of degrees of freedom

“more convex”





Wavefield Reconstruction Inversion

WRI method

for each source i

solve $\begin{pmatrix} P_i \\ \lambda A_i(\mathbf{m}) \end{pmatrix} \mathbf{u}_{\lambda,i} \approx \begin{pmatrix} \mathbf{d}_i \\ \lambda \mathbf{q}_i \end{pmatrix}$

$\mathbf{g} = \mathbf{g} + \lambda^2\omega^2 \text{diag}(\bar{\mathbf{u}}_{i,\lambda})^*(A(\mathbf{m})\bar{\mathbf{u}}_{i,\lambda} - \mathbf{q}_i)$

$H_{GN} = H_{GN} + \lambda^2\omega^4 \text{diag}(\mathbf{u}_i)^* \text{diag}(\mathbf{u}_i)$

$\mathbf{m} = \mathbf{m} - \alpha H_{GN}^{-1}\mathbf{g}$

end

diagonal
=
pseudo Hessian

Conventional method

for each source i

solve $A(\mathbf{m})\mathbf{u}_i = \mathbf{q}_i$

solve $A(\mathbf{m})^*\mathbf{v}_i = P_i^*(P_i\mathbf{u}_i - \mathbf{d}_i)$

$\mathbf{g} = \mathbf{g} + \omega^2 \text{diag}(\mathbf{u}_i)^*\mathbf{v}_i$

$\mathbf{m} = \mathbf{m} - \alpha\mathbf{g}$

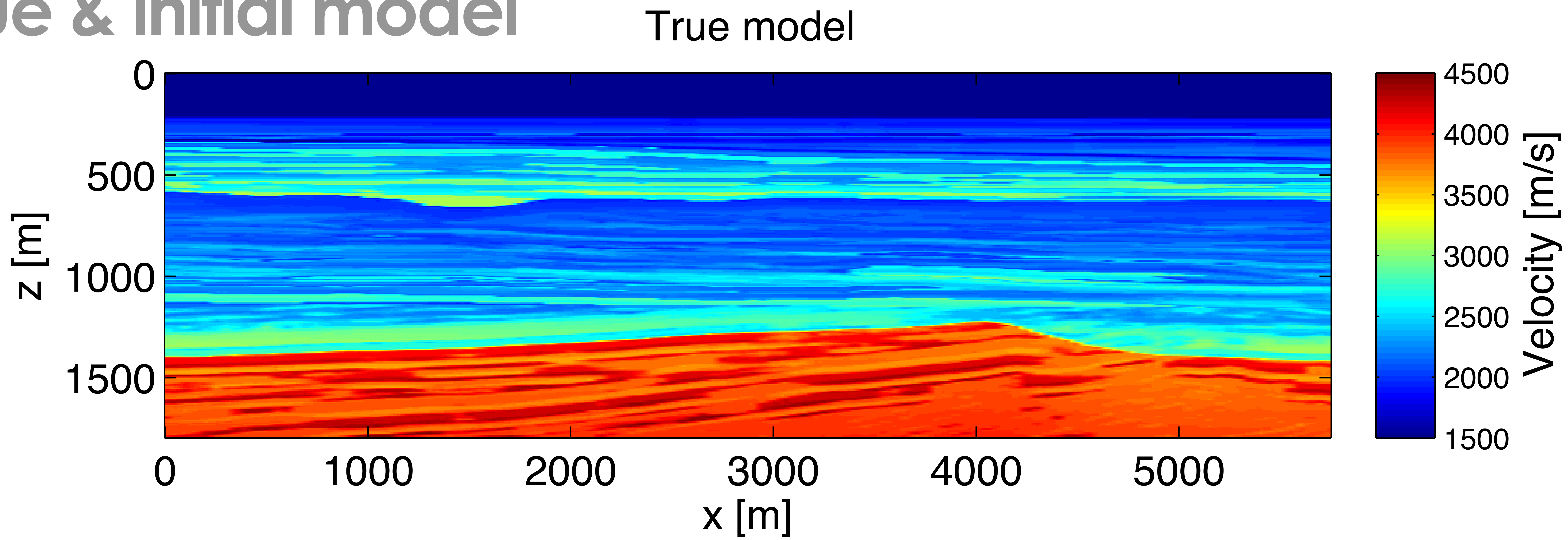
end

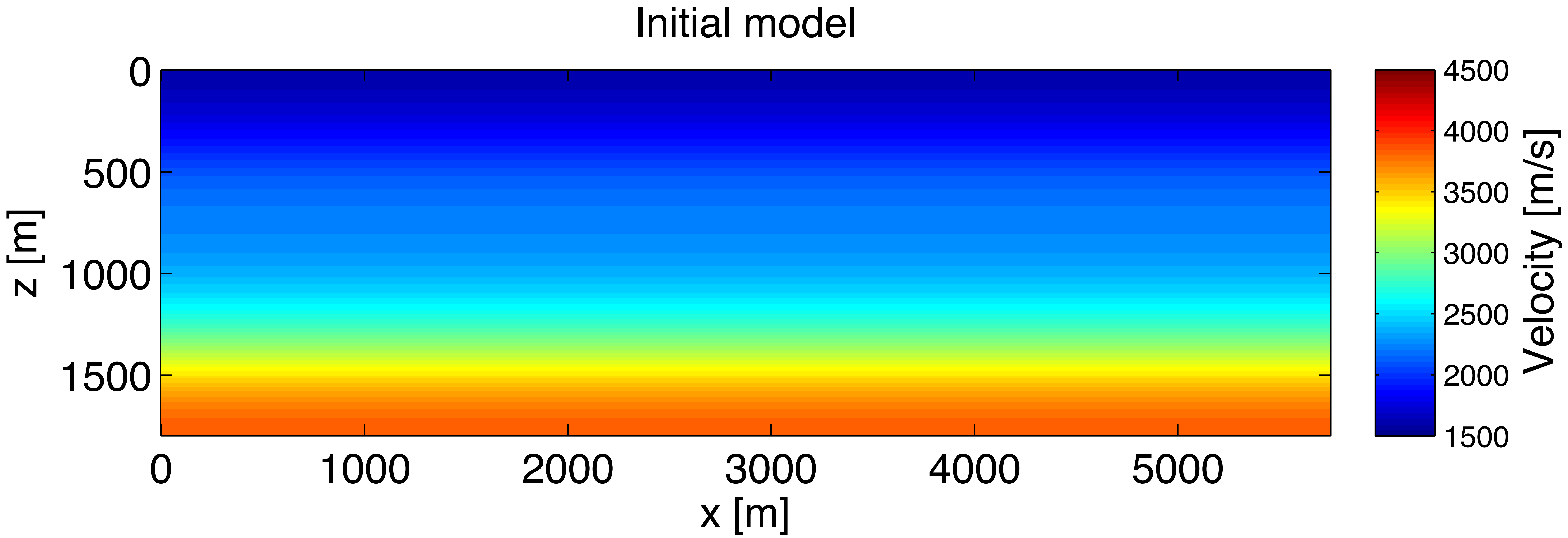
dense
&
too expsensive

Example – BG Compass model

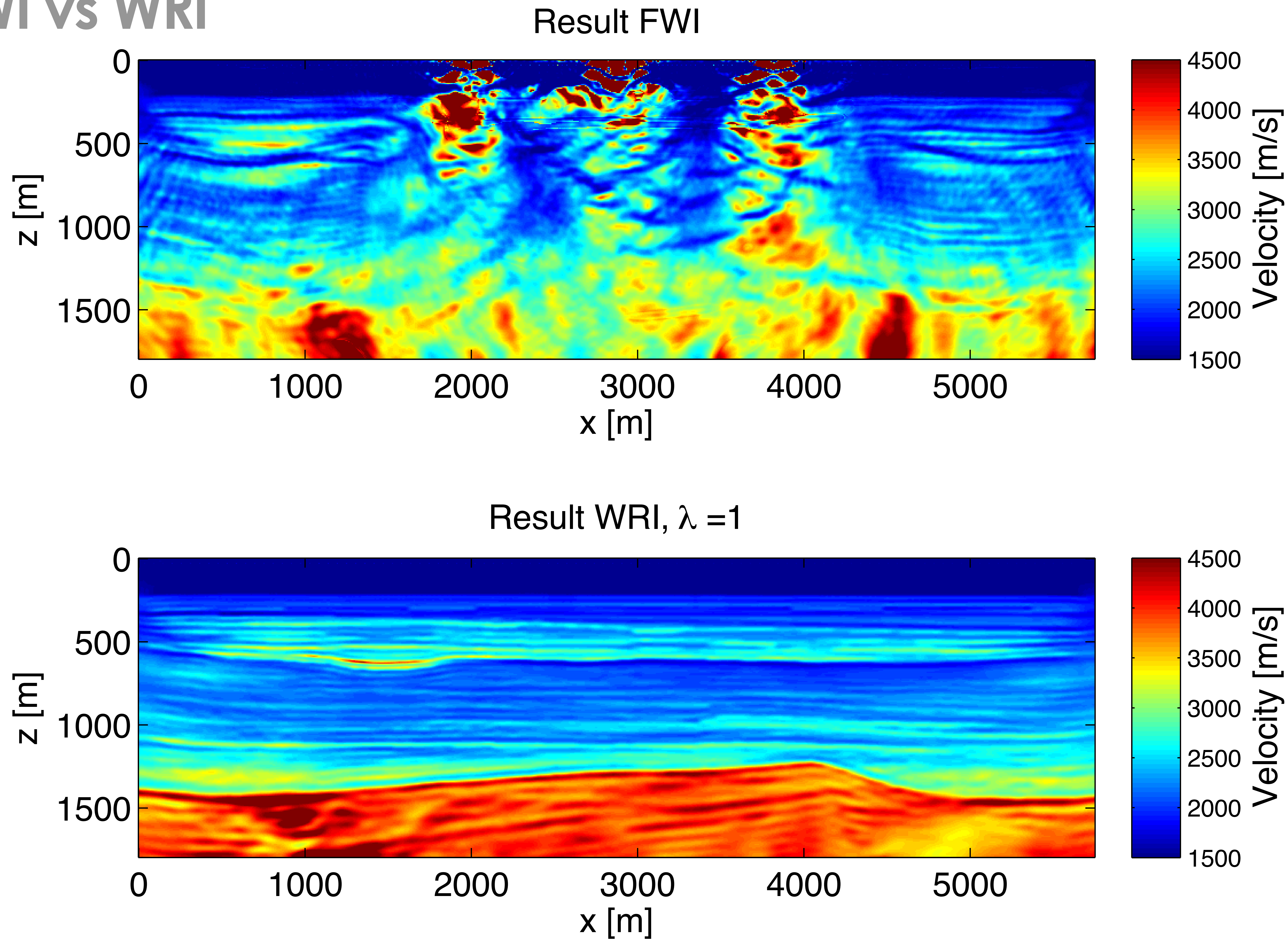
- Low frequencies missing, 24 frequency batches (15 iterations each) {5 6} ,{6 7},... ,{28 29} Hertz. Each interval contains 5 frequencies.
- 103 sources/receivers w/ 55m sample interval
- Inaccurate initial model

True & initial model



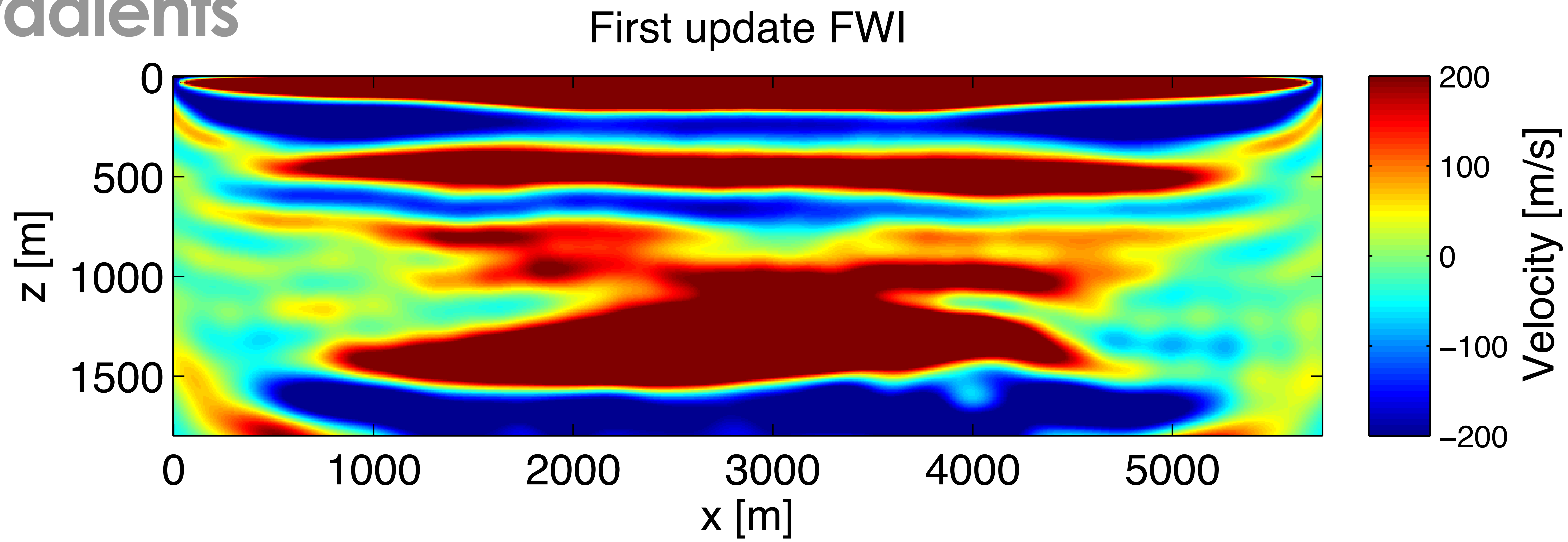


FWI vs WRI

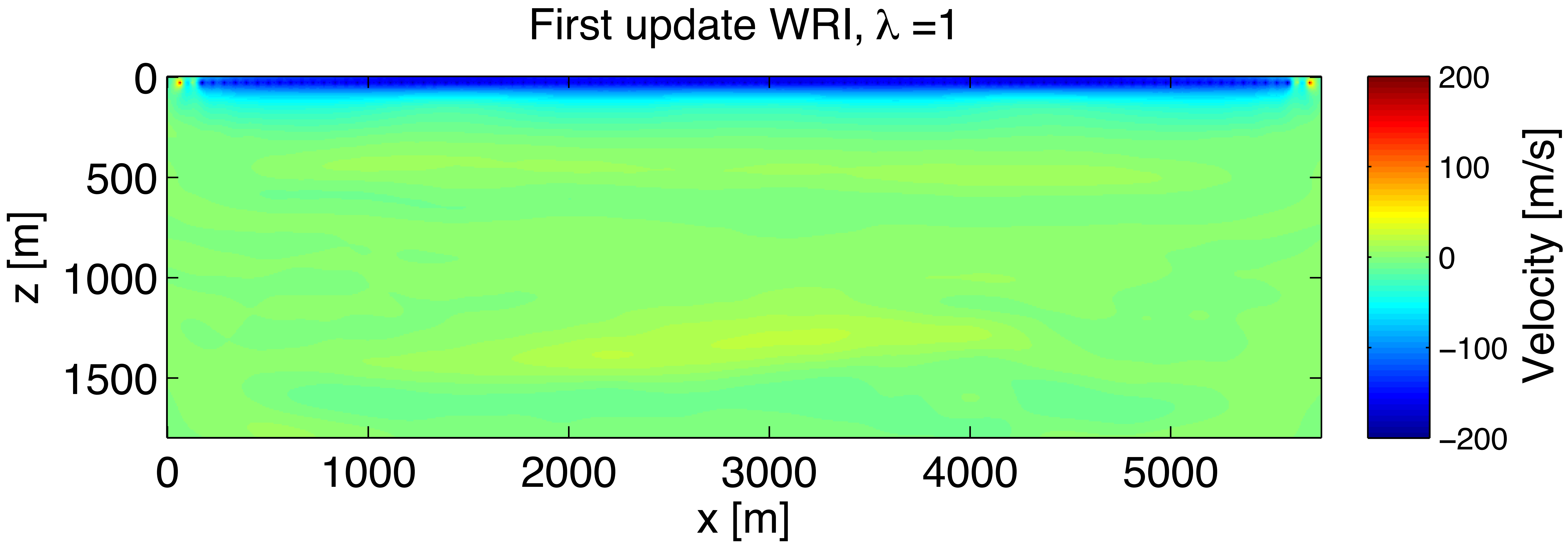


Result FWI
0
500
1000
1500
z [m]
0
1000
2000
3000
4000
5000
x [m]
4500
4000
3500
3000
2500
2000
1500
Velocity [m/s]
Result WRI, λ =1
0
500
1000
1500
z [m]
0
1000
2000
3000
4000
5000
x [m]
4500
4000
3500
3000
2500
2000
1500
Velocity [m/s]

Gradients

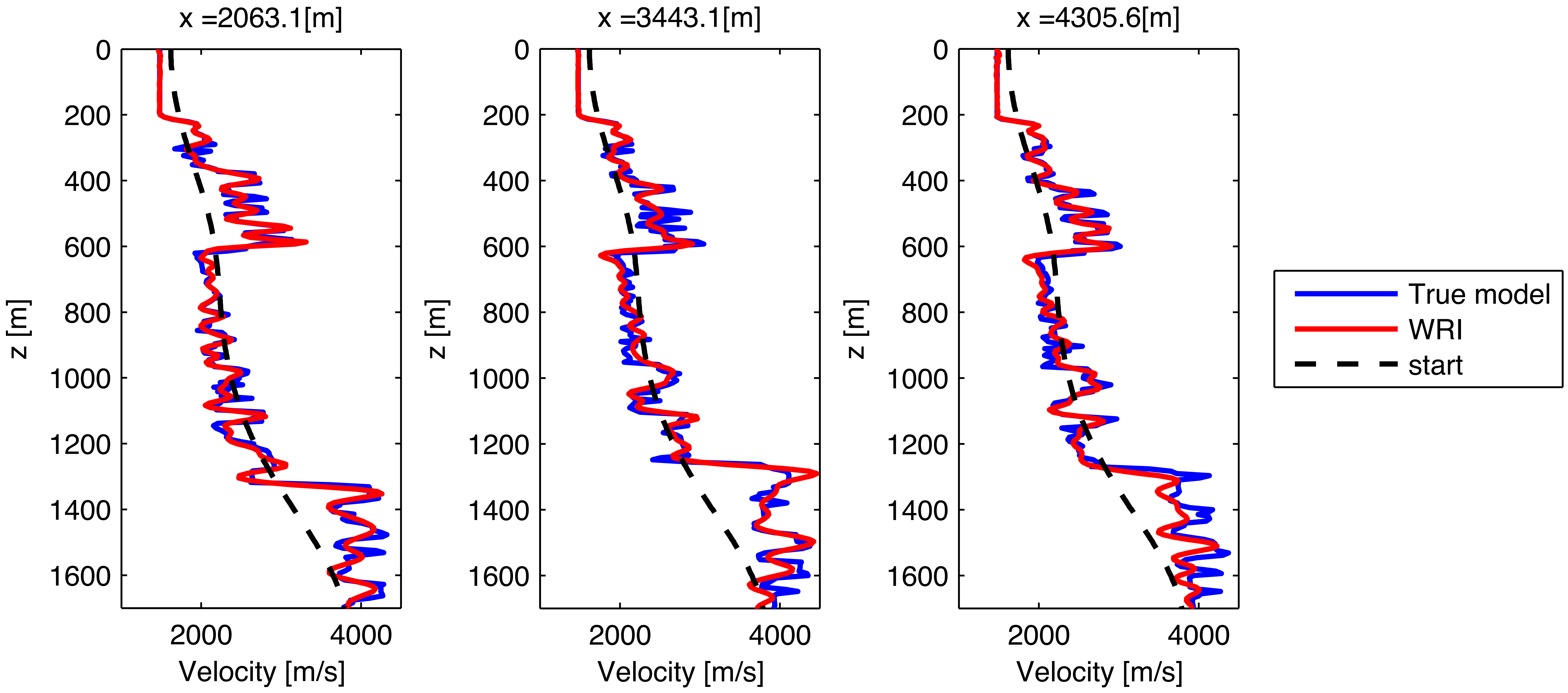


First update FWI
z [m]
0
500
1000
1500
0
1000
2000
3000
4000
5000
x [m]
200
100
0
−100
−200
Velocity [m/s]

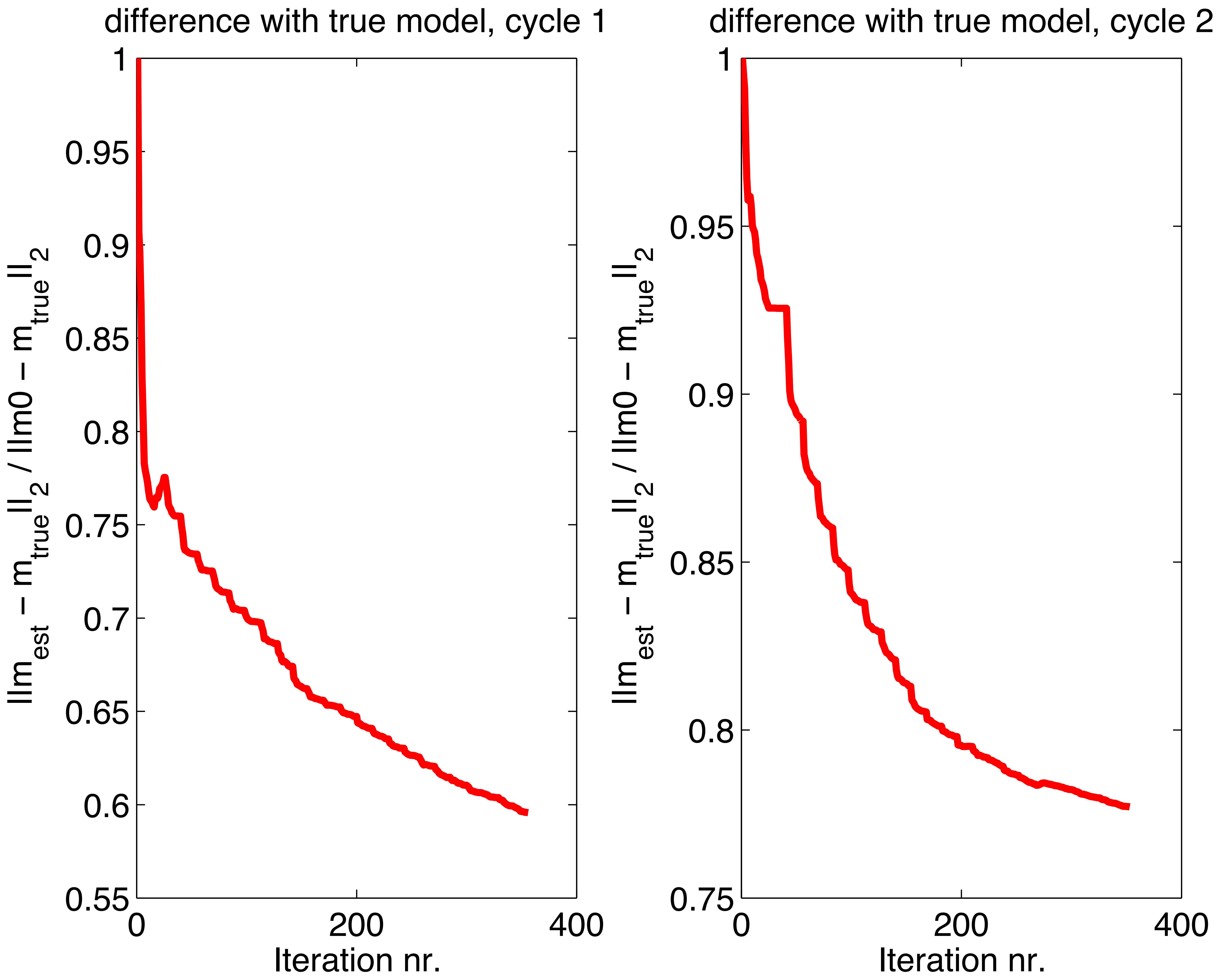


First update WRI, λ =1
z [m]
0
500
1000
1500
0
1000
2000
3000
4000
5000
x [m]
200
100
0
−100
−200
Velocity [m/s]

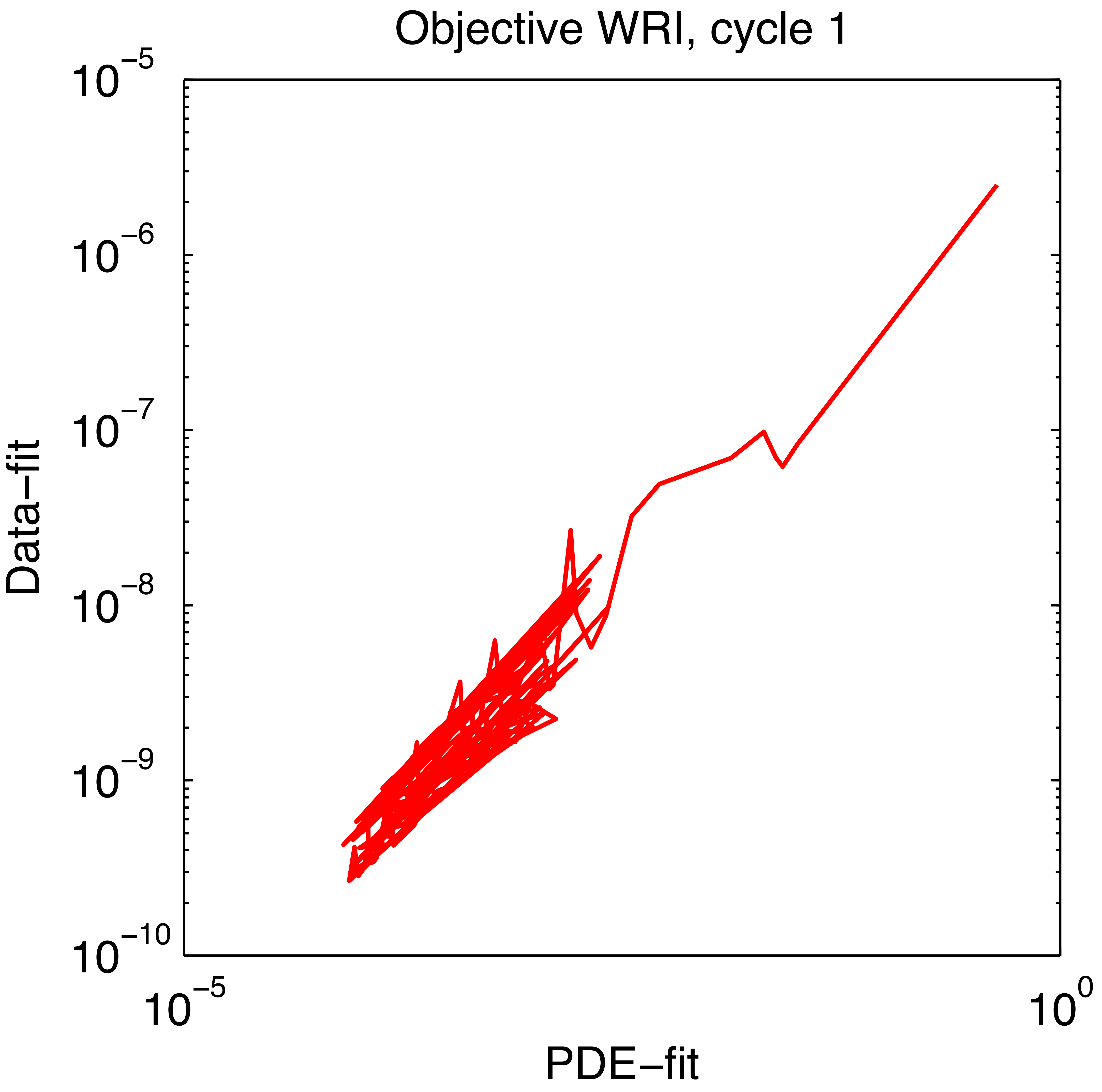
Cross sections

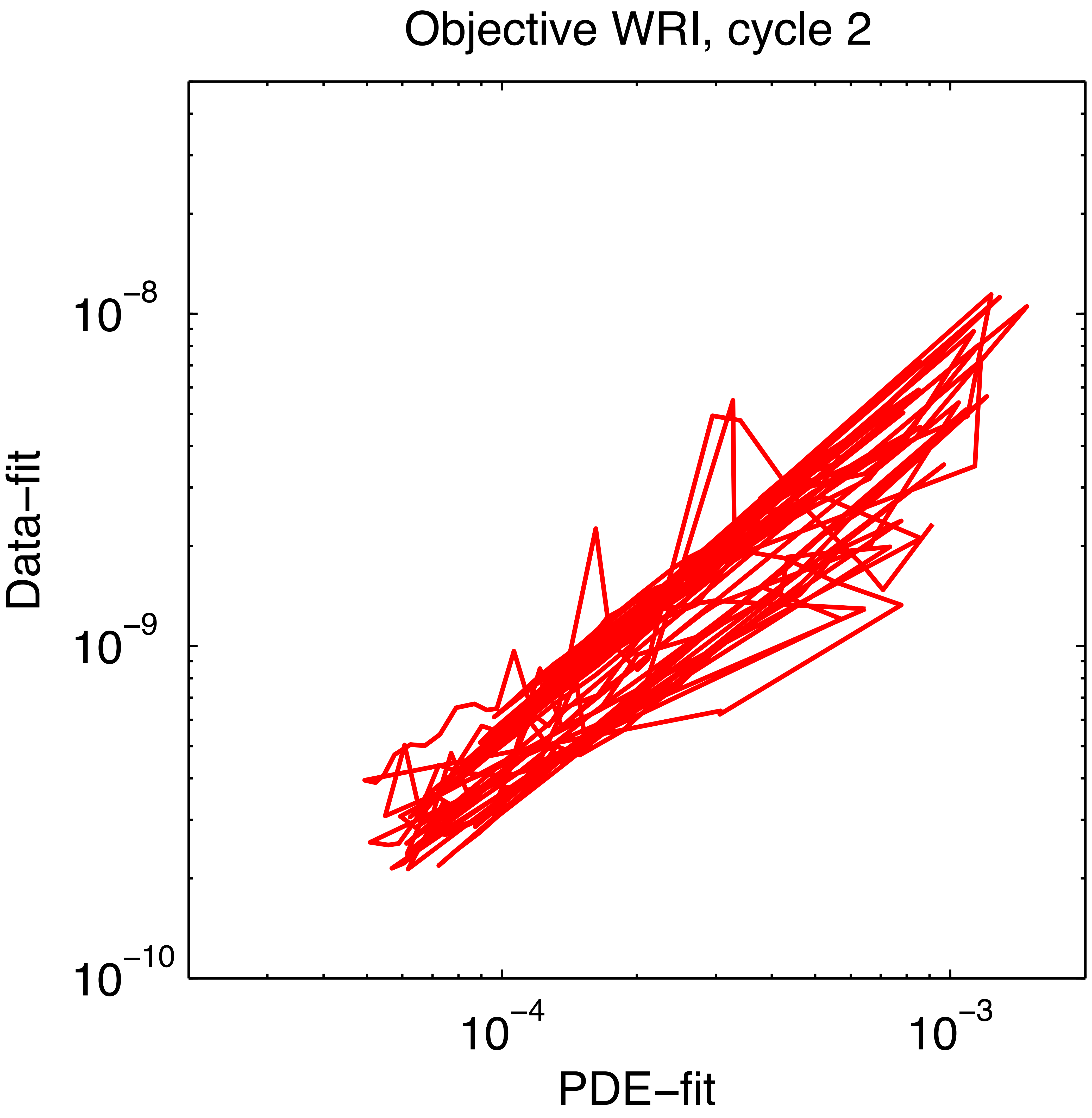


Relative model errors



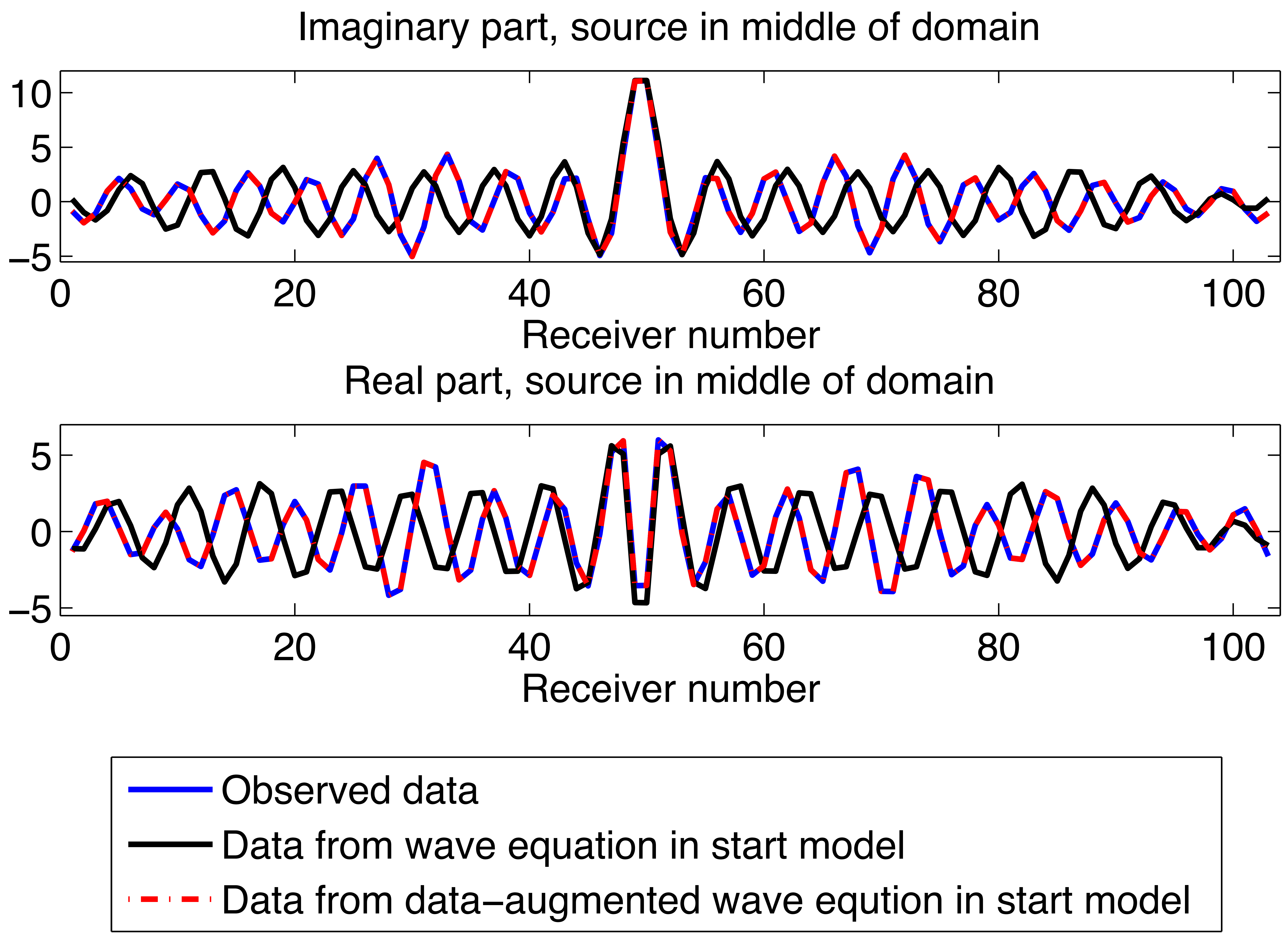
Objective function value





Data fit increases at some iterates

Data fit



Imaginary part, source in middle of domain
10
5
0
−5
0
20
40
60
80
100
Receiver number
Real part, source in middle of domain
5
0
−5
0
20
40
60
80
100
Receiver number
Observed data
Data from wave equation in start model
Data from data−augmented wave eqution in start model

Ernie Esser, Tristan van Leeuwen, Aleksandr Y. Aravkin, and Felix J. Herrmann, "**A scaled gradient projection method for total variation regularized full waveform inversion**". 2014.
Bas Peters and Felix J. Herrmann, "**A sparse reduced Hessian approximation for multi-parameter Wavefield Reconstruction Inversion**". 2014.

SLIM

Extensions

Total-variation regularization

via scaled gradient projections & bound constraints

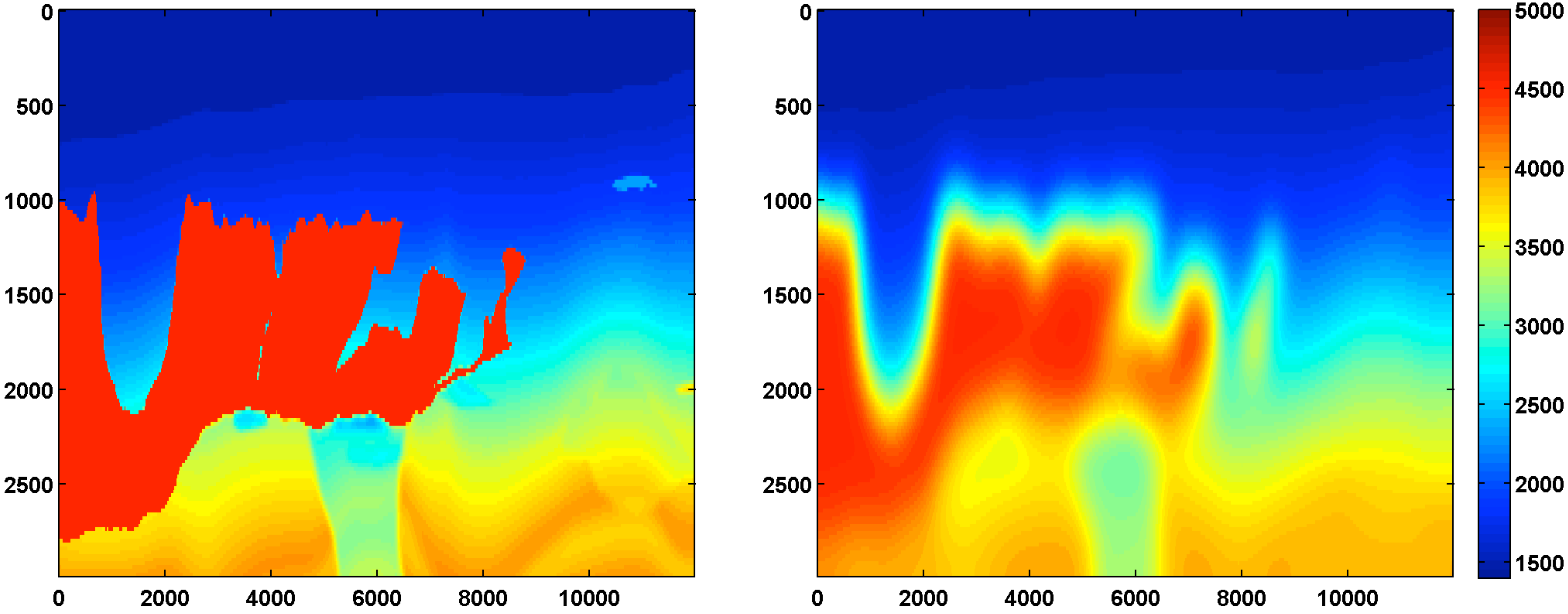
Multi-parameter case

via sparse approximate Gauss-Newton scaling

BP model

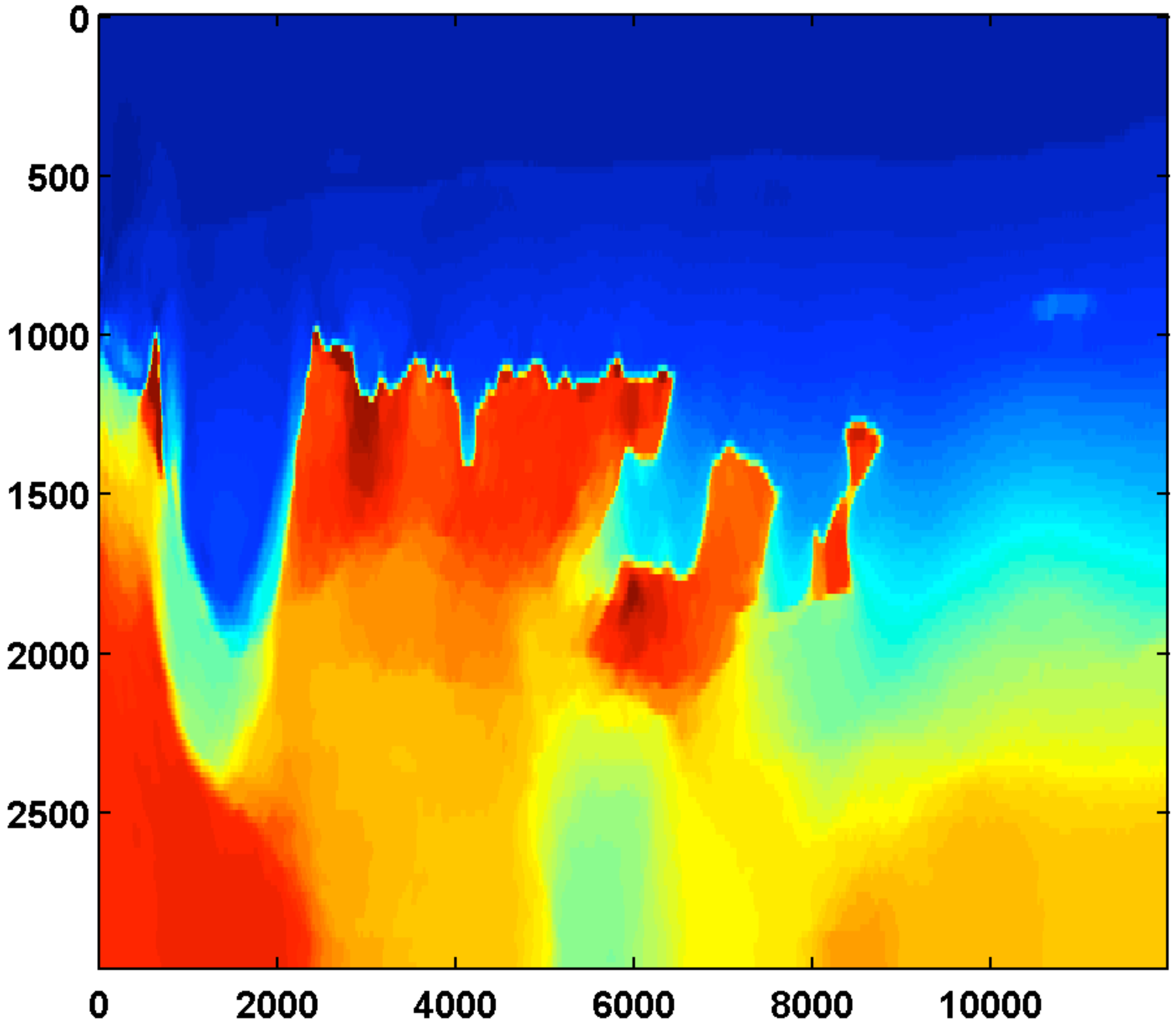
- number of sources: 126 (starting 1000m in from boundary)
- number of receivers: 299
- frequency range: 3-20Hz in overlapping batches of 2
- maximum number of outer iterations per frequency batch: 25
- maximum number of inner iterations for convex subproblems: 2000
- known Ricker wavelet sources with 15Hz peak frequency
- two simultaneous shots with Gaussian weights w/ redraws
- no added noise

True and initial velocity

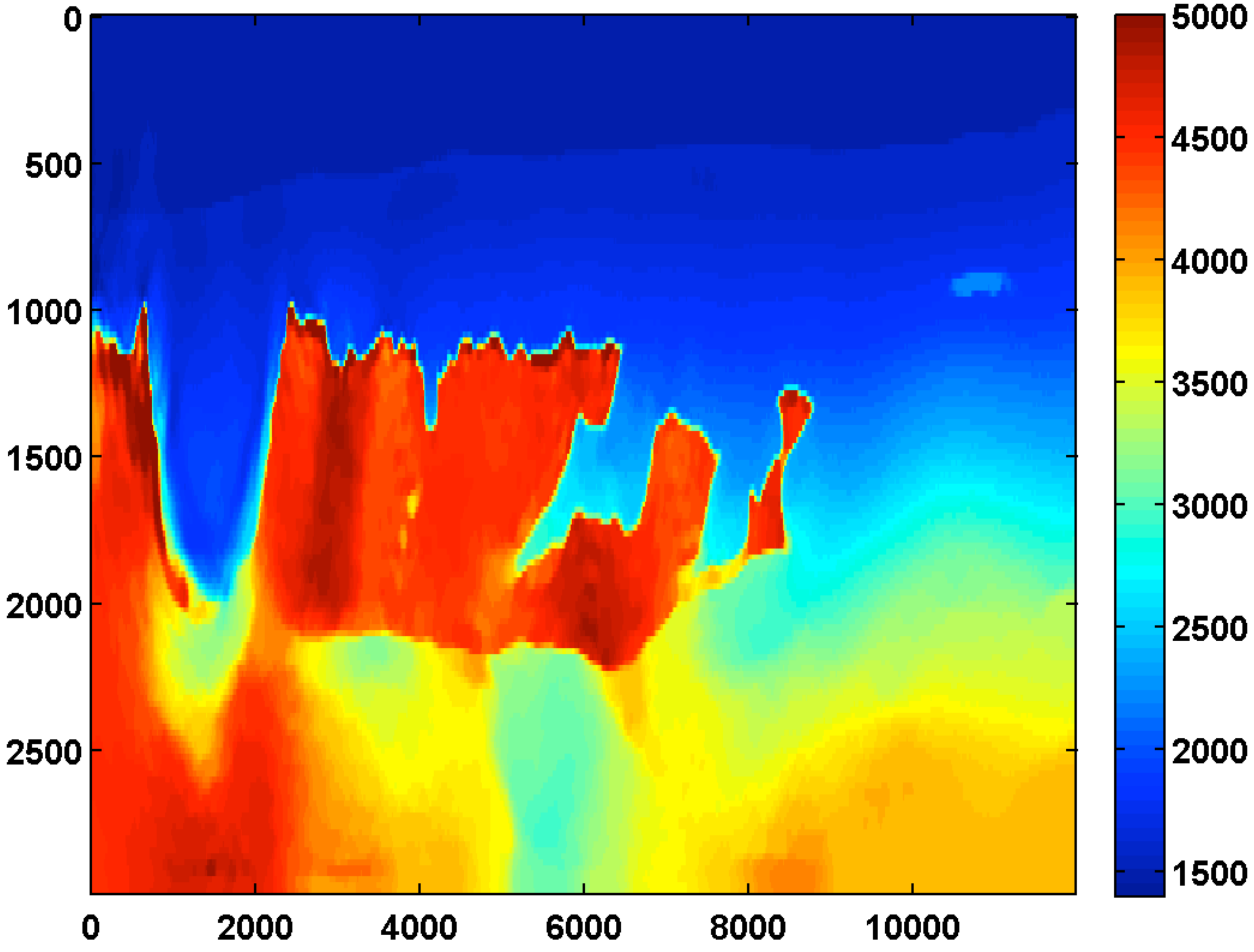


Results w/ TV

After one cycle through the frequencies



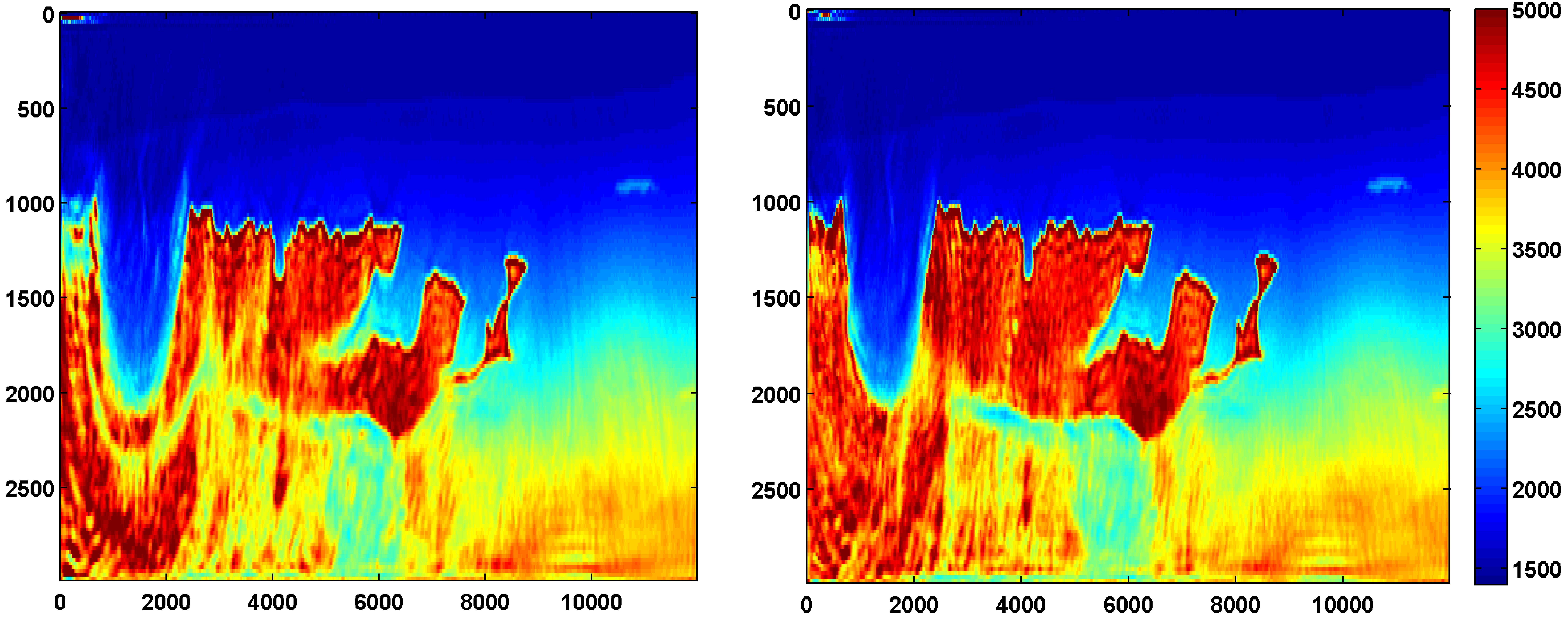
After two cycles through the frequencies



Results w/o TV

After one cycle through the frequencies

After two cycles through the frequencies



Conclusions

New alternating method for wave-equation based inversion:

- same extended search space as in all-at-once but with memory & CPU requirements as in adjoint-state approach
- no adjoints & sparse GN-Hessian approximation
- less susceptible to local minima due to data fit
- sparse GN Hessians
- bilinear

Challenge: Stationary points are not necessary global minima

SLIM

Acknowledgements

Thank you for your attention !

https://www.slim.eos.ubc.ca/

SINBAD

NSERC CRSNG

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE II (375142-08). This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BGP, CGG, Chevron, ConocoPhillips, ION, Petrobras, PGS, Statoil, Subsalt Ltd, Total SA, WesternGeco, and Woodside.