

A new take on FWI: Wavefield Reconstruction Inversion

Felix J. Herrmann

A new take on FWI: Wavefield Reconstruction Inversion

Tristan van Leeuwen* and Bas Peters



* now @ CWI Amsterdam

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

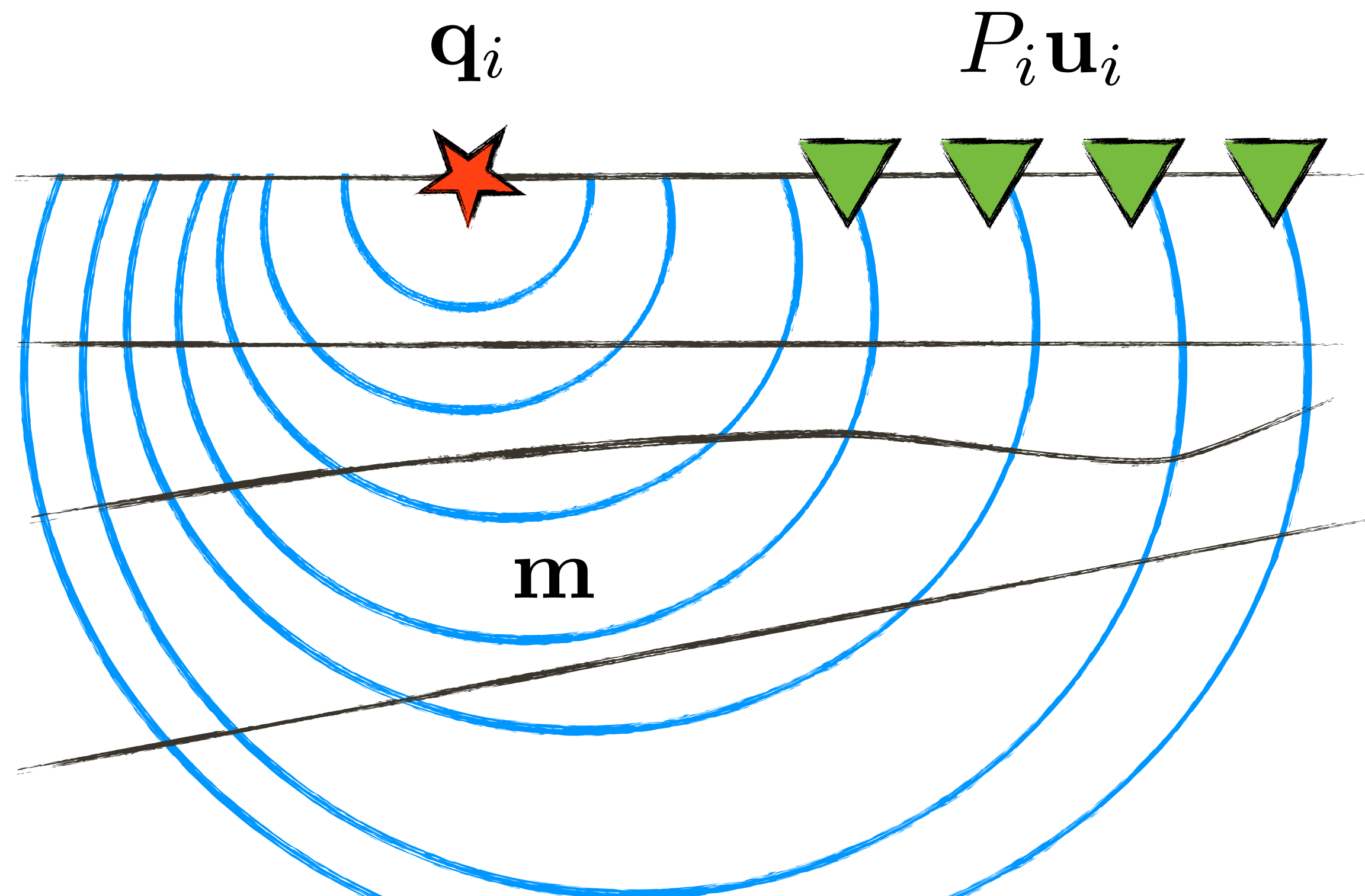
Full-waveform inversion is plagued with *local* minima

Derive an *alternative* 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*: $A(\mathbf{m})\mathbf{u}_i = \mathbf{q}_i$



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 are *consistent* with the wave equation.

Waveform inversion

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

$$\begin{array}{ccc} \text{"known" physics} & \text{"known" source} & \\ \downarrow & \downarrow & \\ A(\mathbf{m})\mathbf{u}_i = \mathbf{q}_i & & \\ \uparrow & & \\ \text{"unknown" wavefield} & & \end{array}$$

- ▶ holds *exactly* for each source i
- ▶ *differ* on *insisting* wave equations to *hold* for *each* iteration
- ▶ *different* unknowns: $\mathbf{m} \longleftrightarrow \mathbf{m} \ \& \ \mathbf{u}$

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

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

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

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*

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

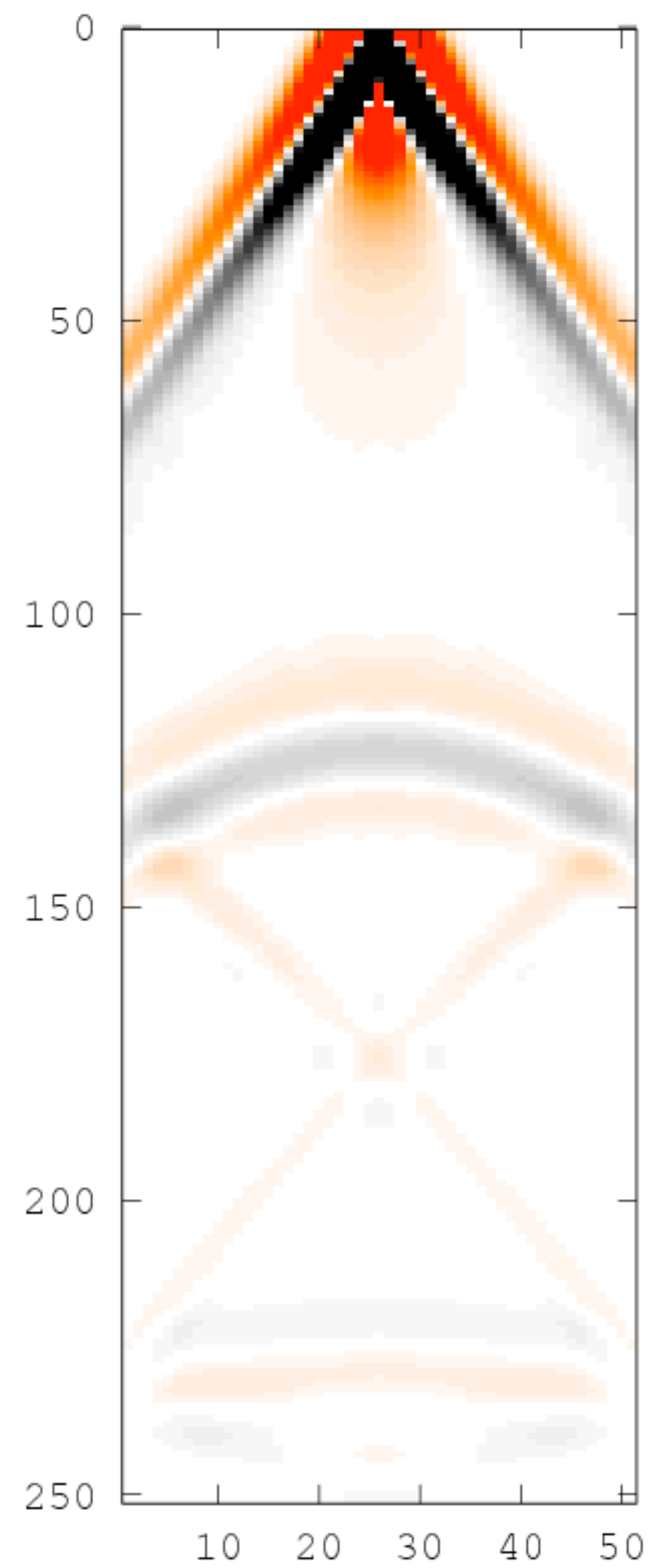
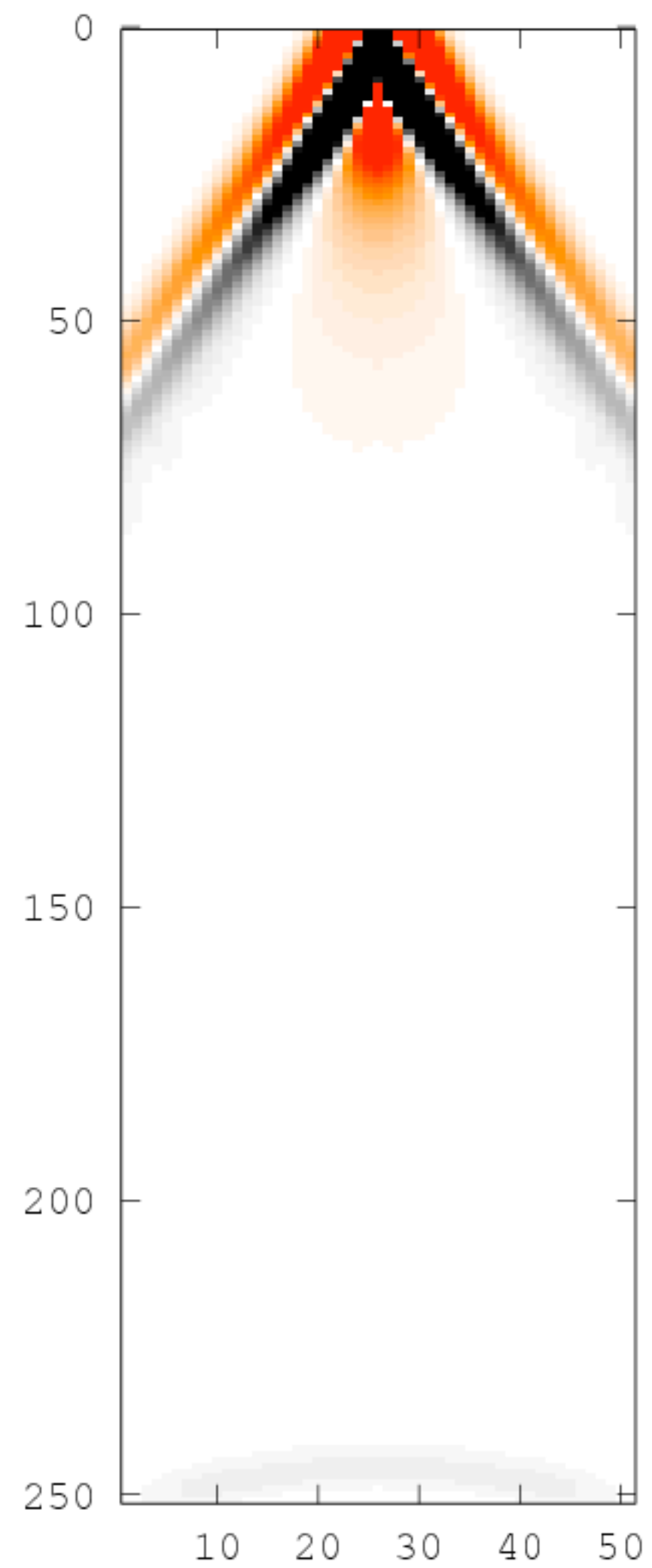
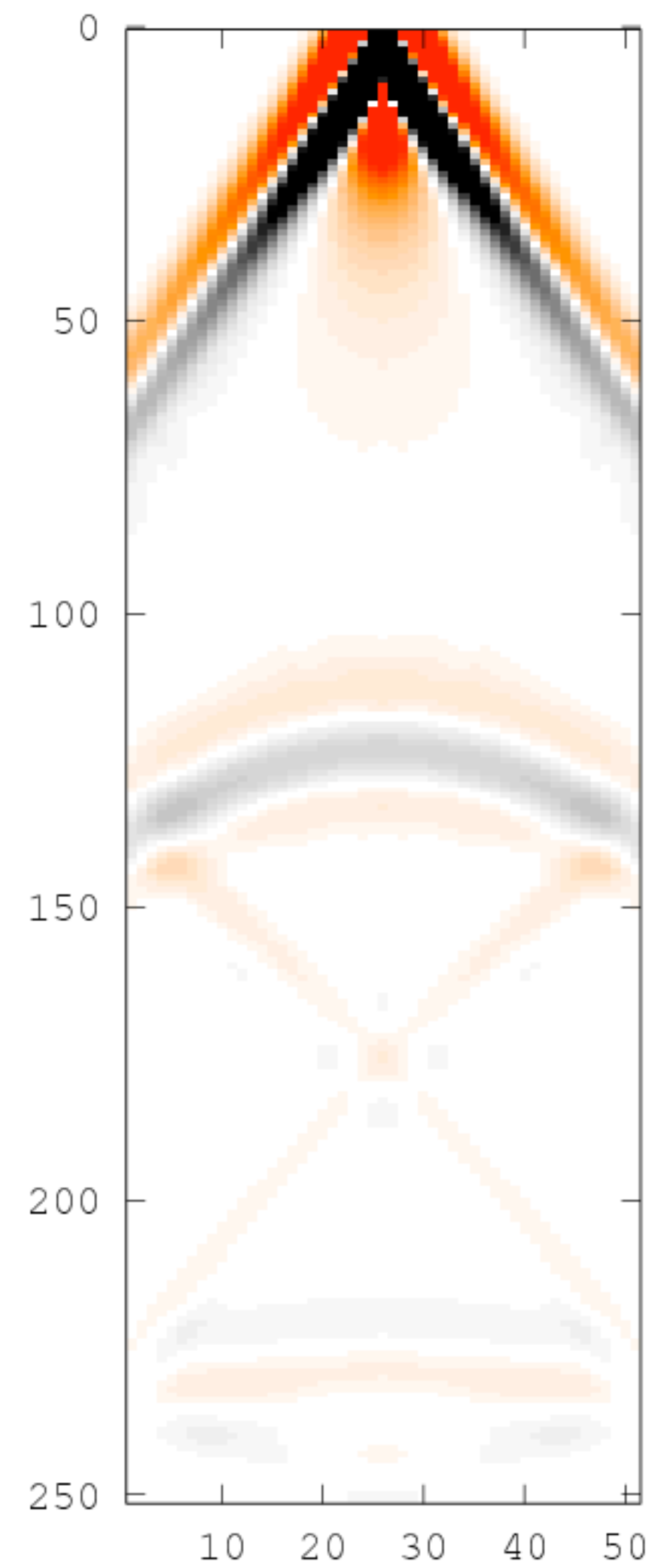
followed by *fixing* \mathbf{u}_i and solving

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

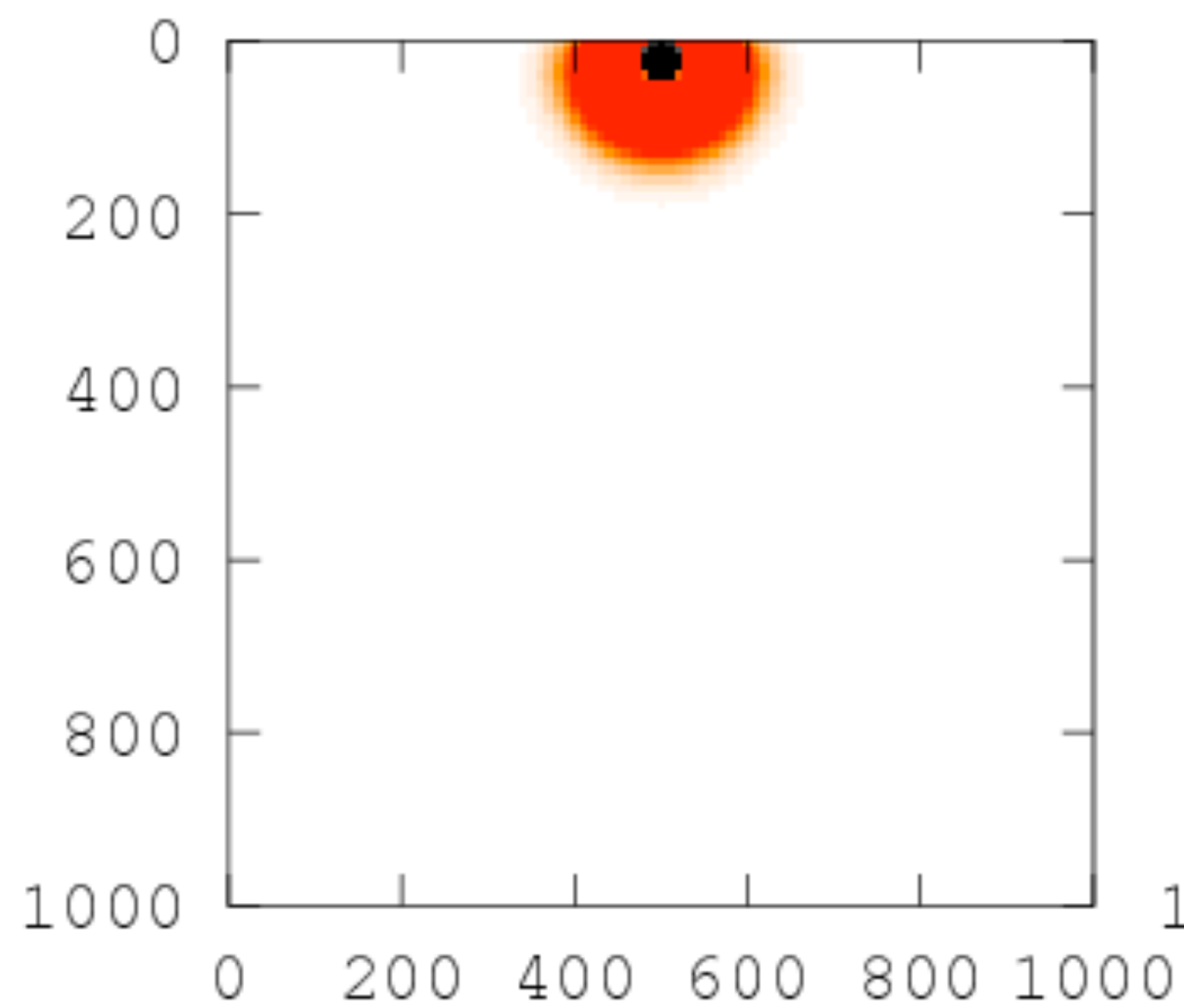
wave-equation \times wavefield = source

versus

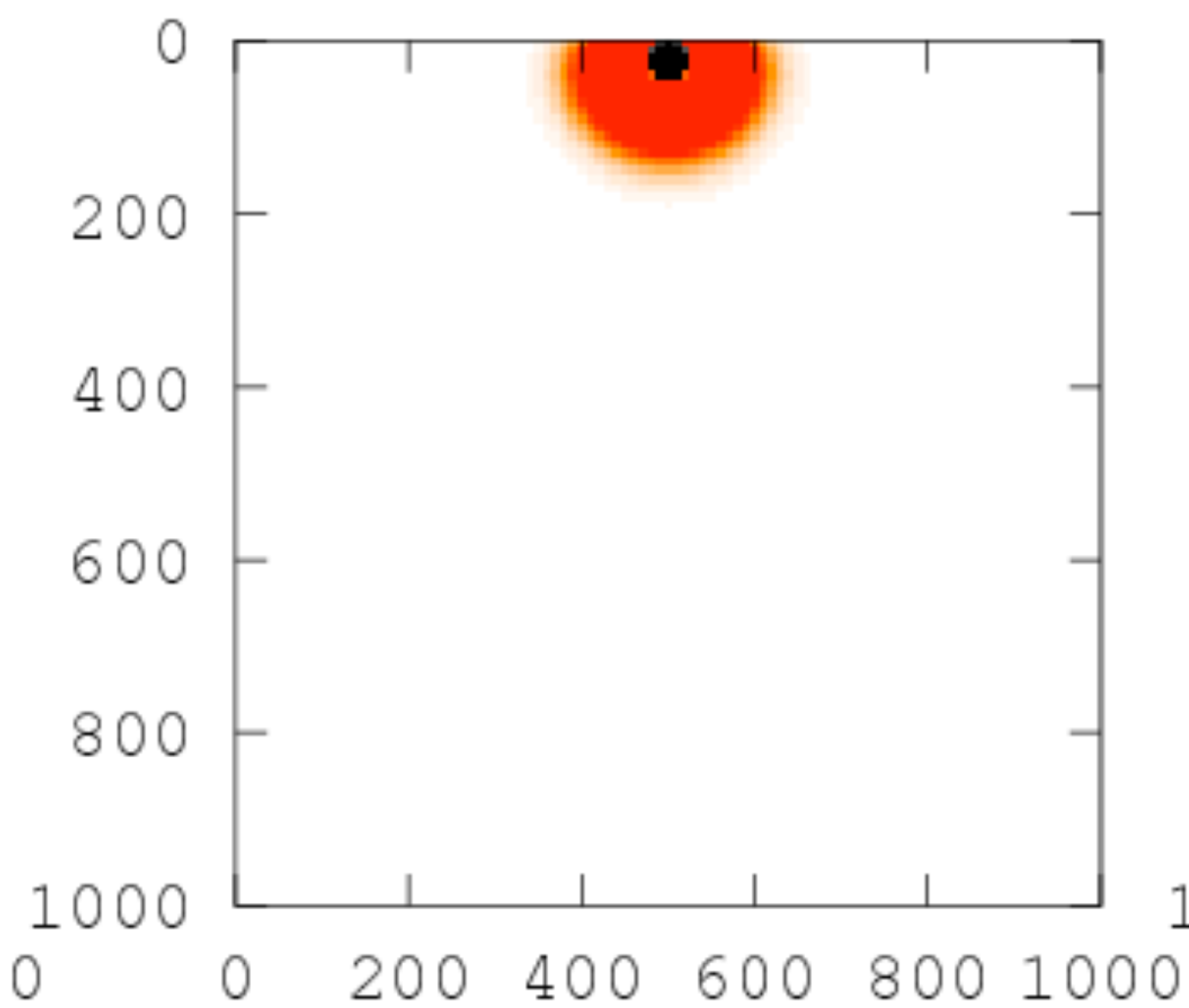
$\left(\begin{array}{c} \text{wave-equation} \\ \text{-----} \\ \text{sampling operator} \end{array} \right) \times \text{wavefield} = \left(\begin{array}{c} \text{source} \\ \text{-----} \\ \text{data} \end{array} \right)$

observed data**initial data****data-augmented solution**

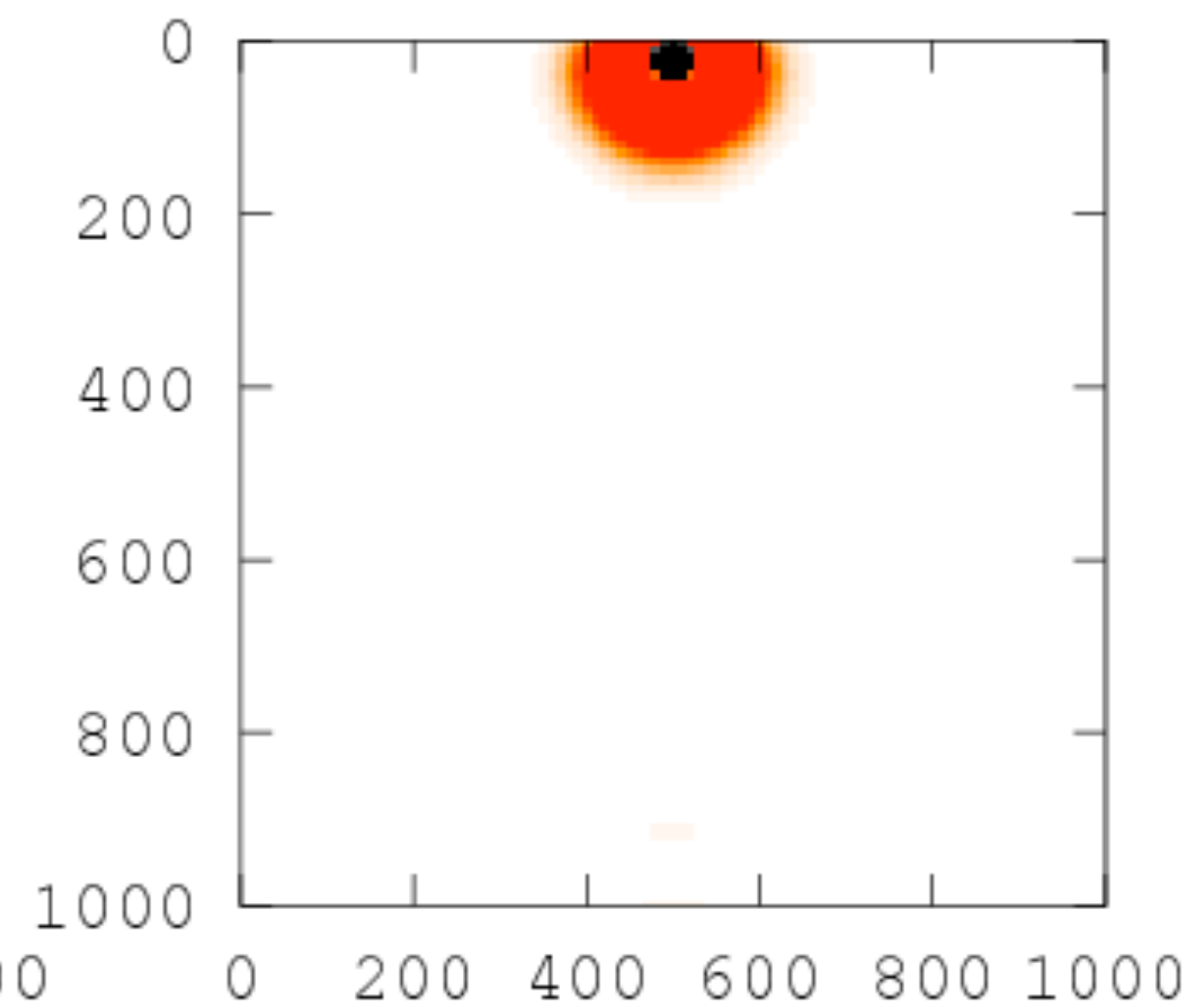
wavefield in *true* model



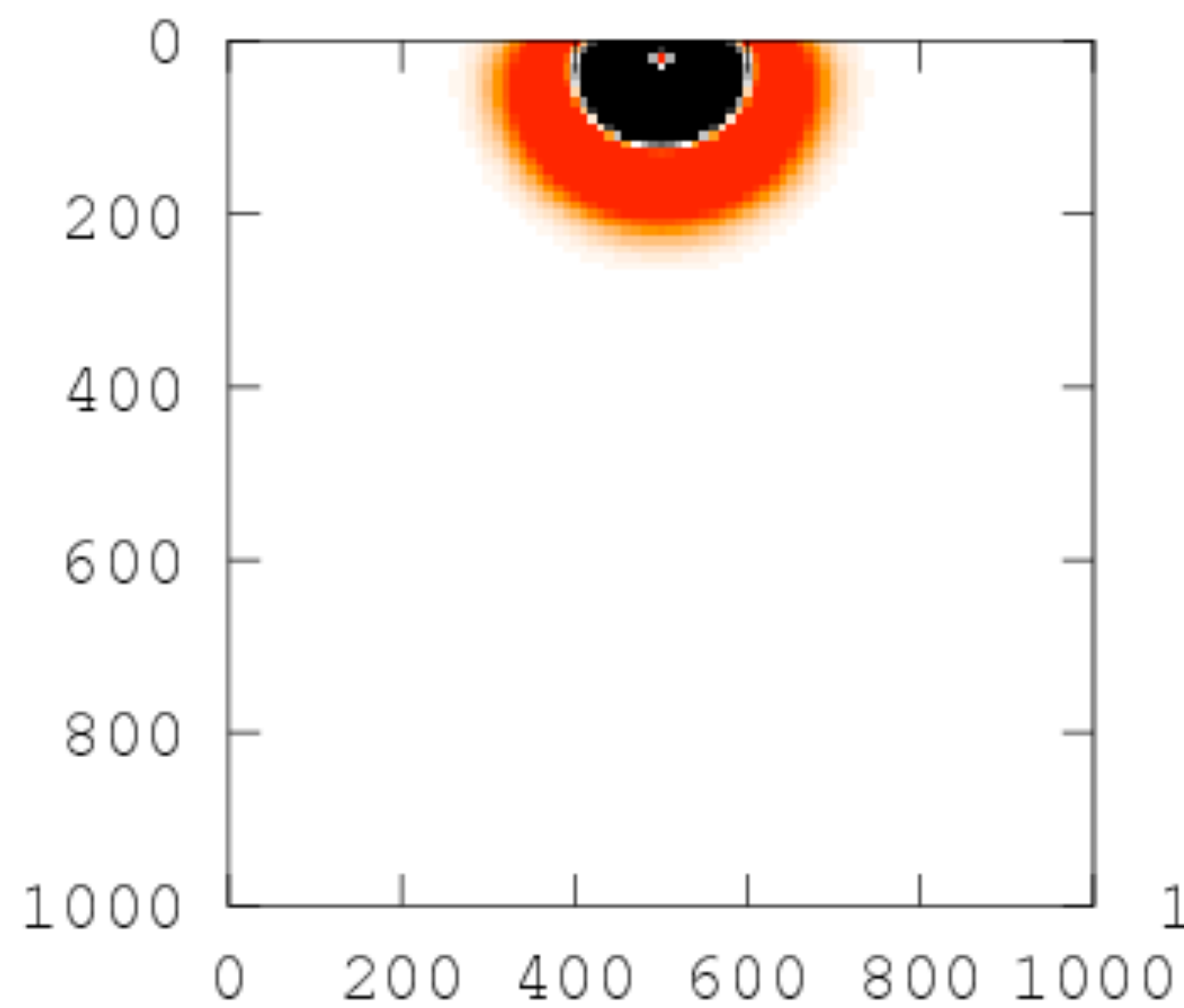
wavefield in *constant* model



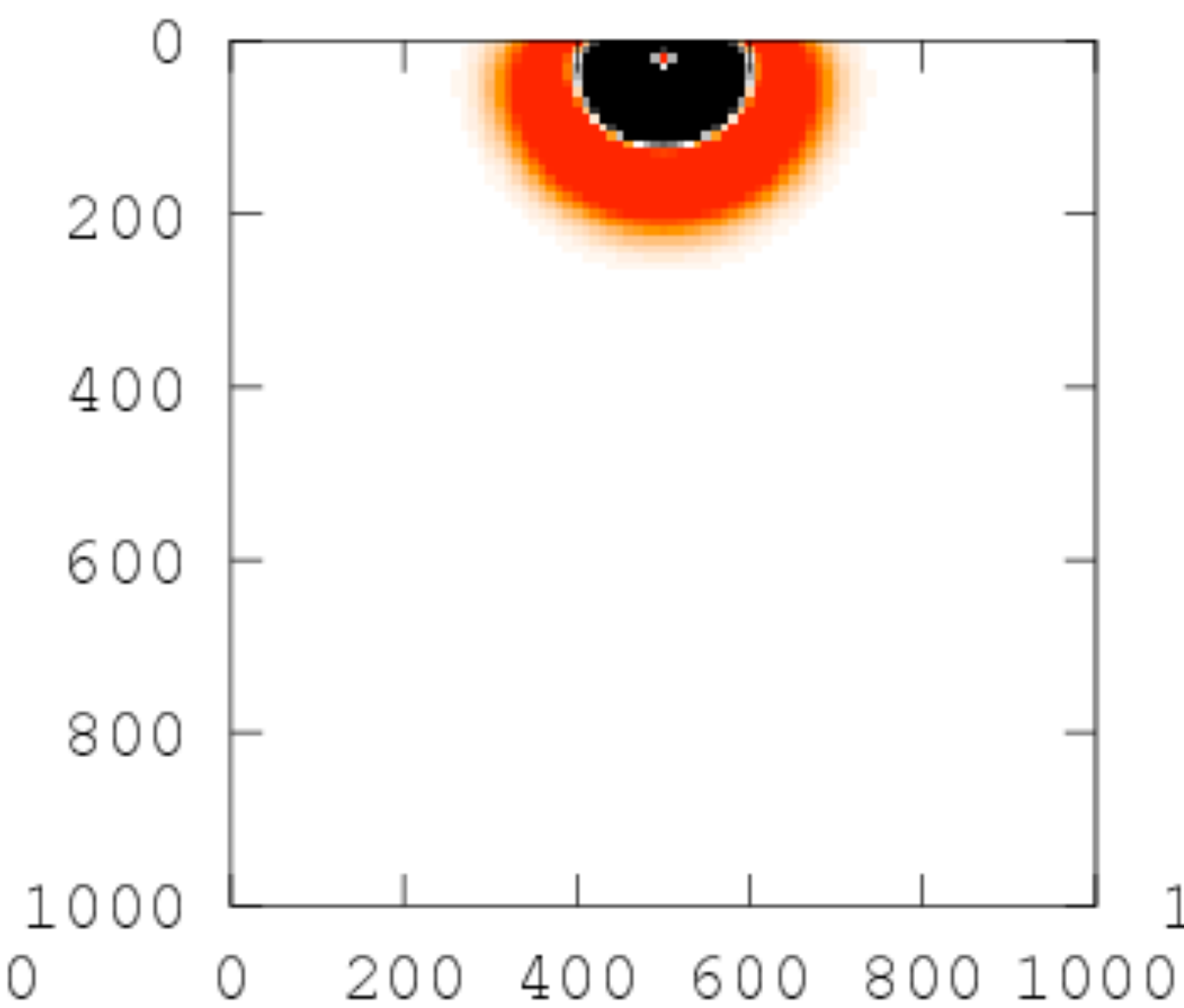
**data-augmented
wavefield in *constant* model**



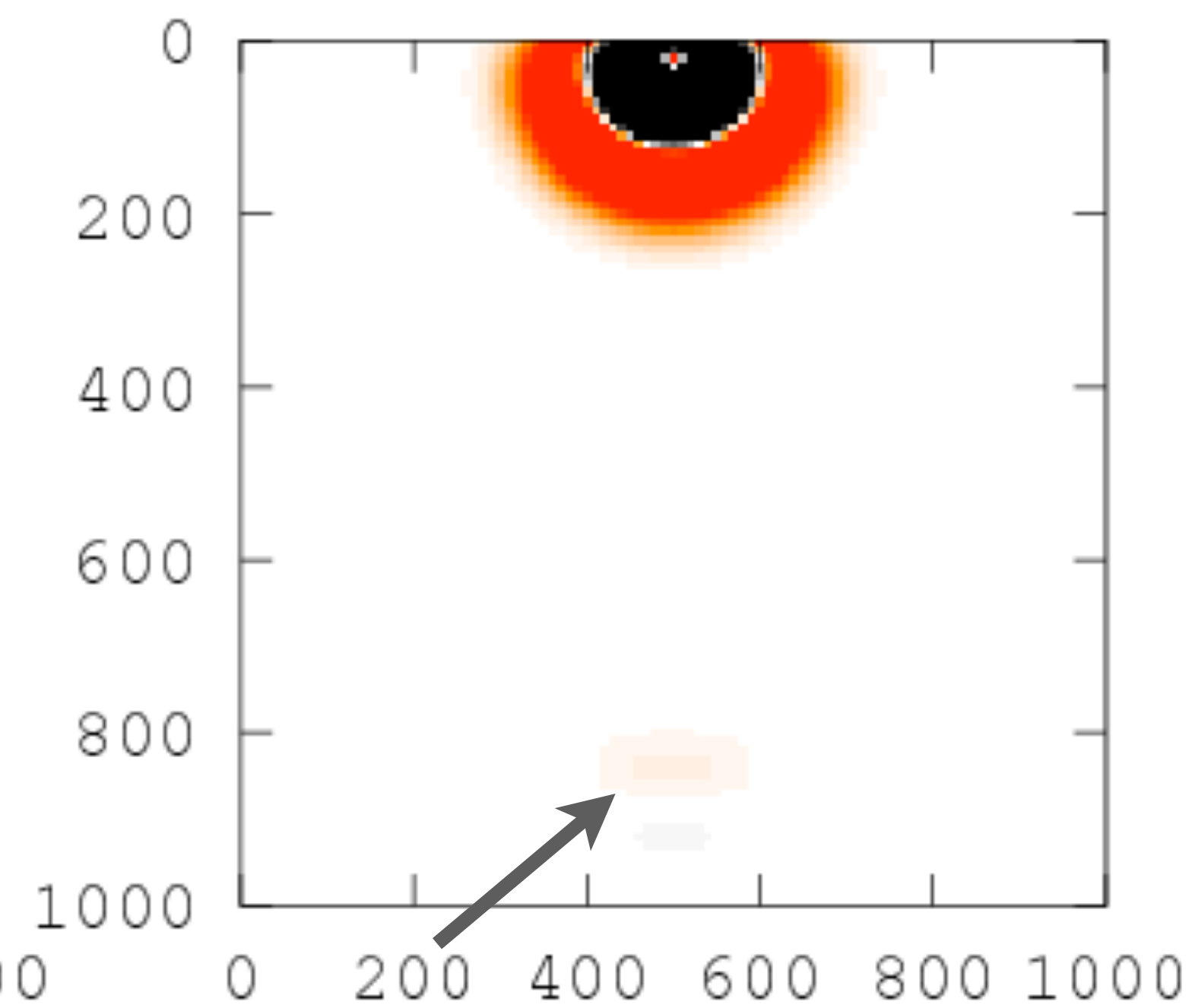
wavefield in *true* model



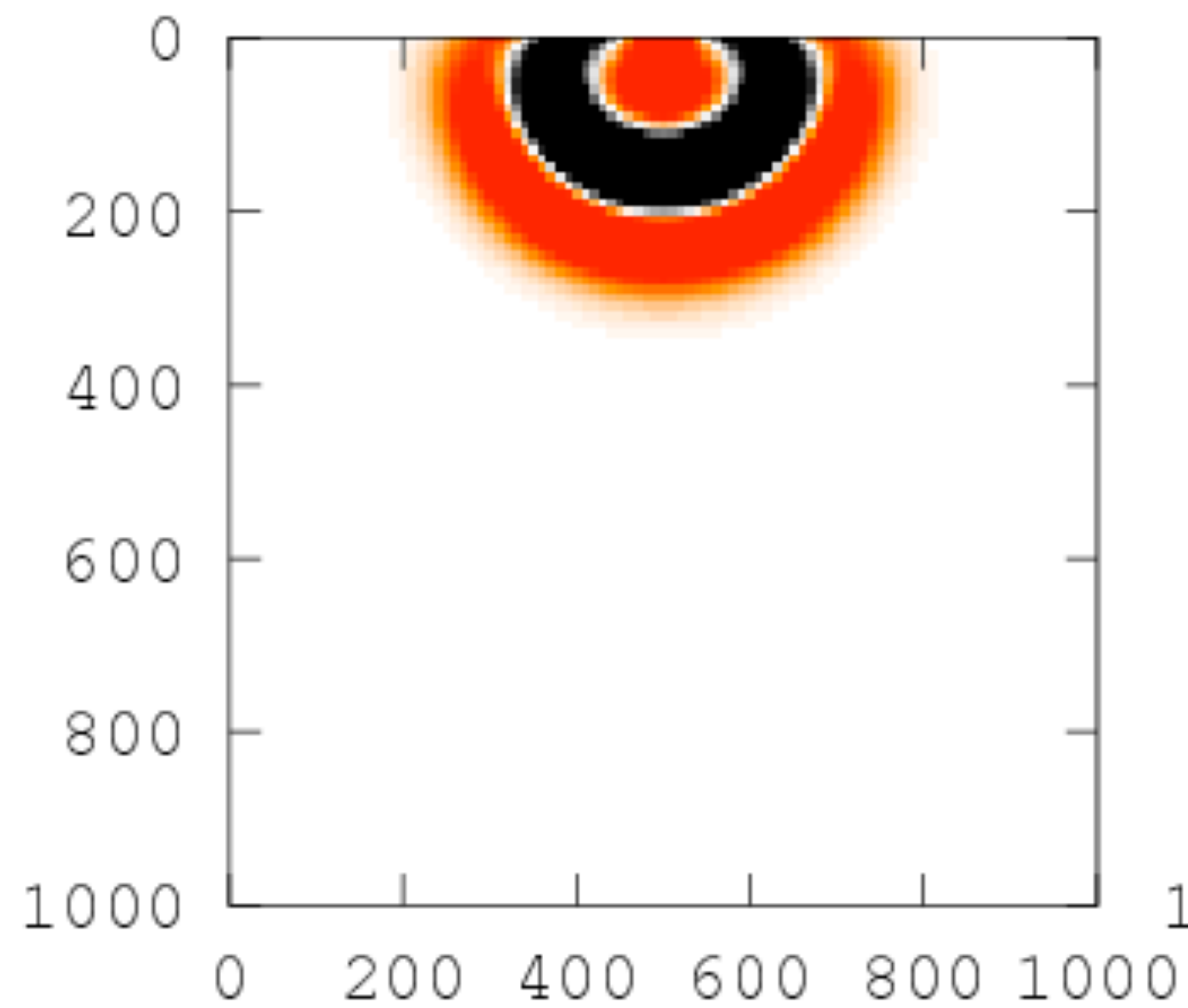
wavefield in *constant* model



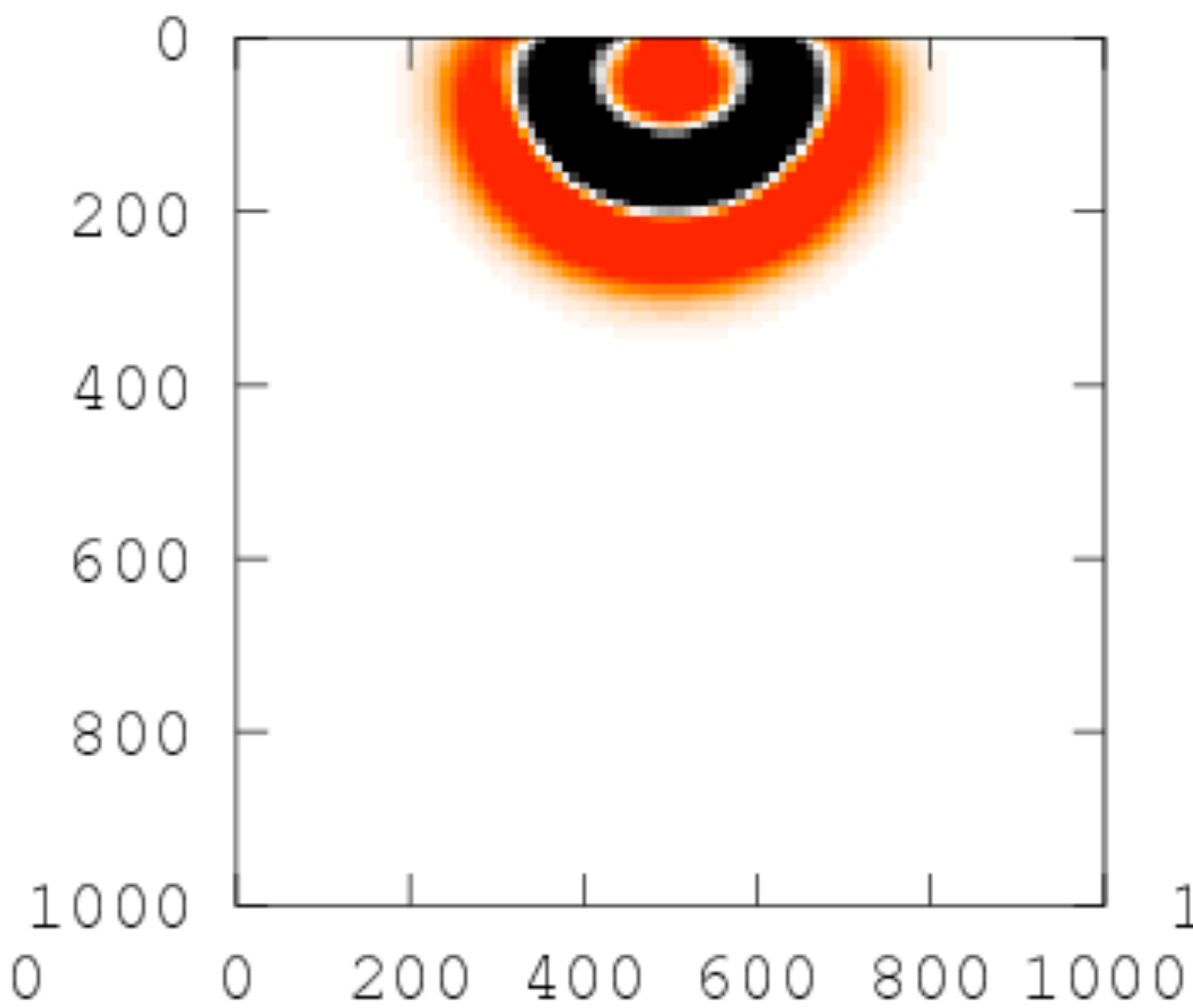
**data-augmented
wavefield in *constant* model**



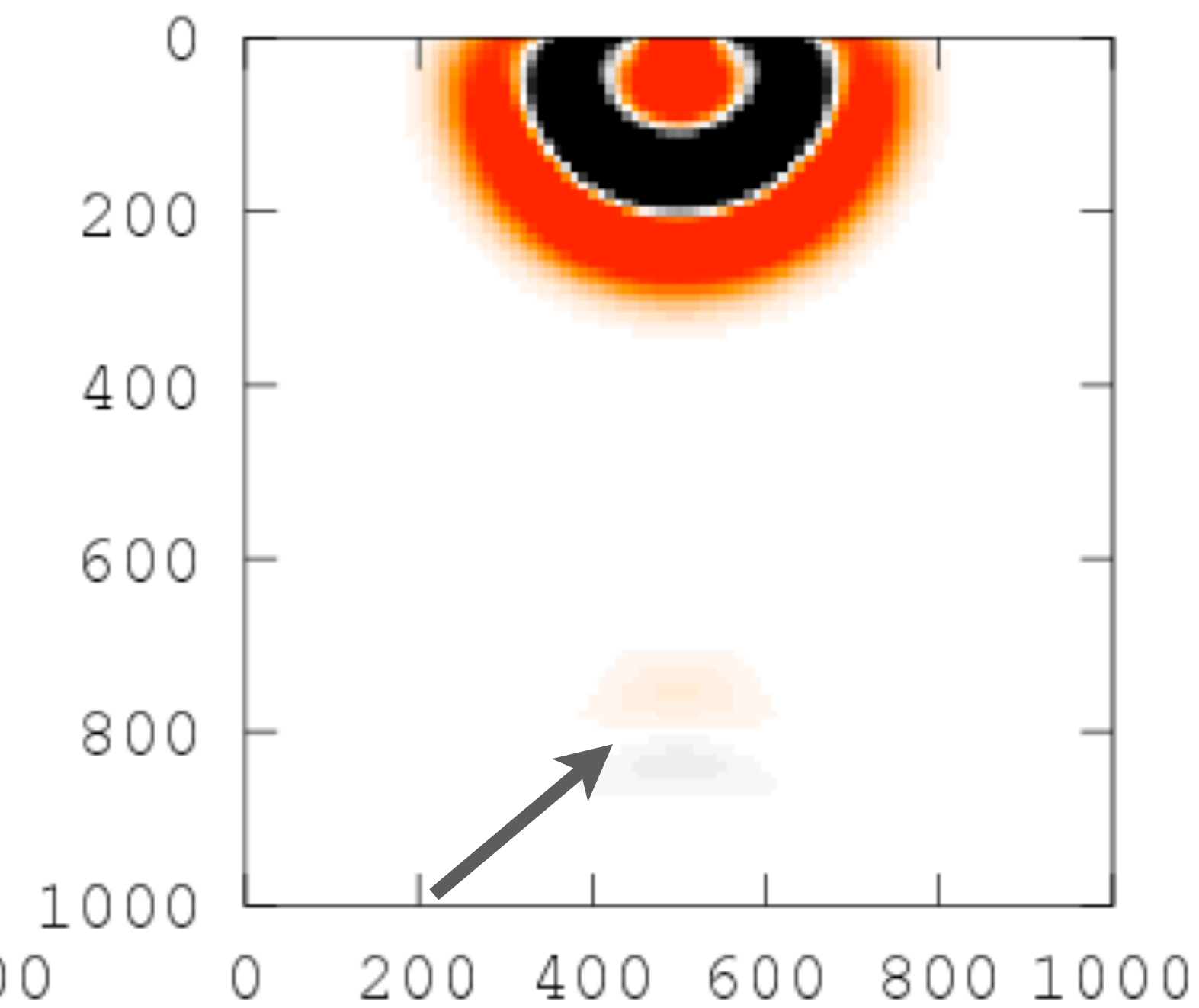
wavefield in *true* model



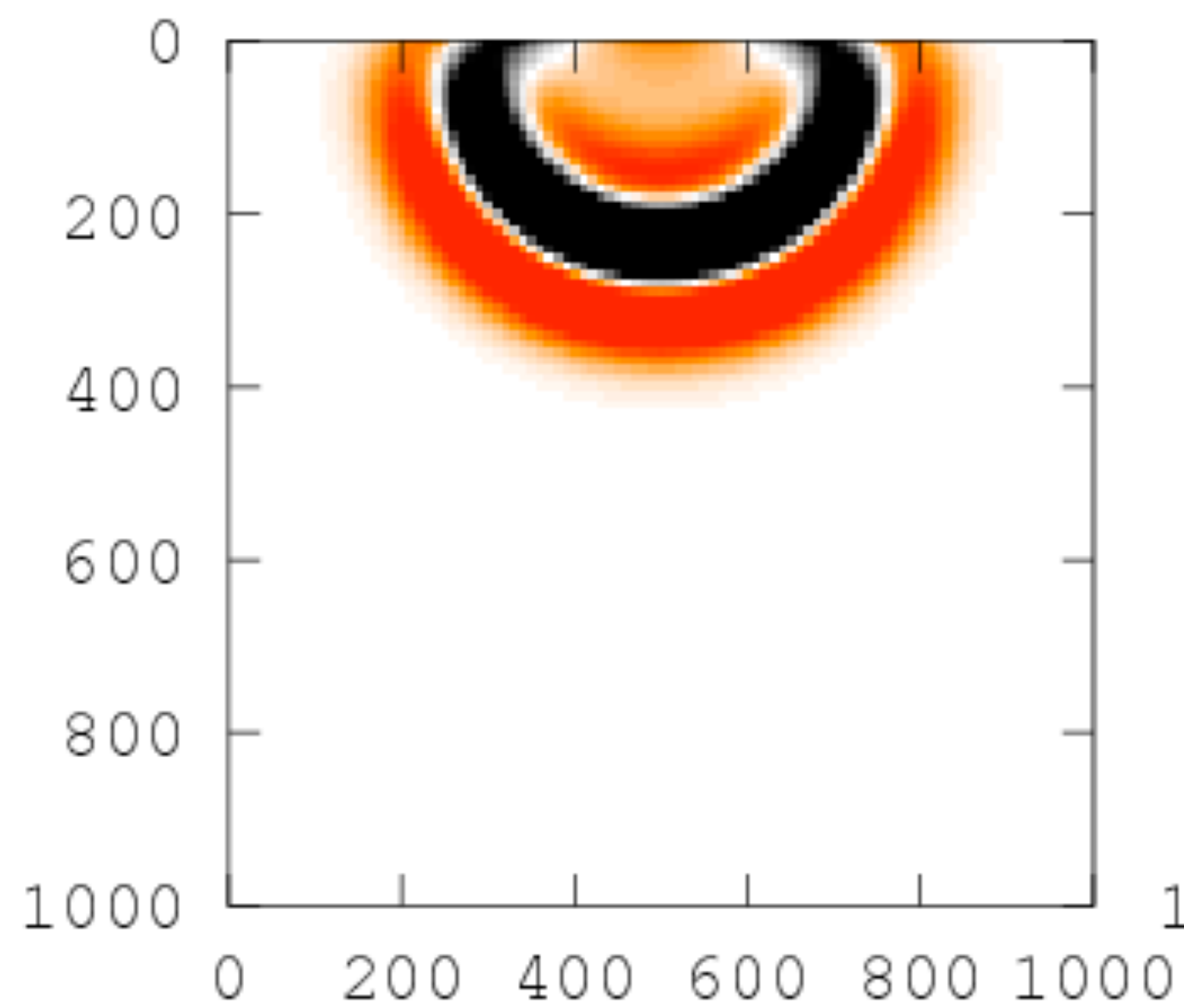
wavefield in *constant* model



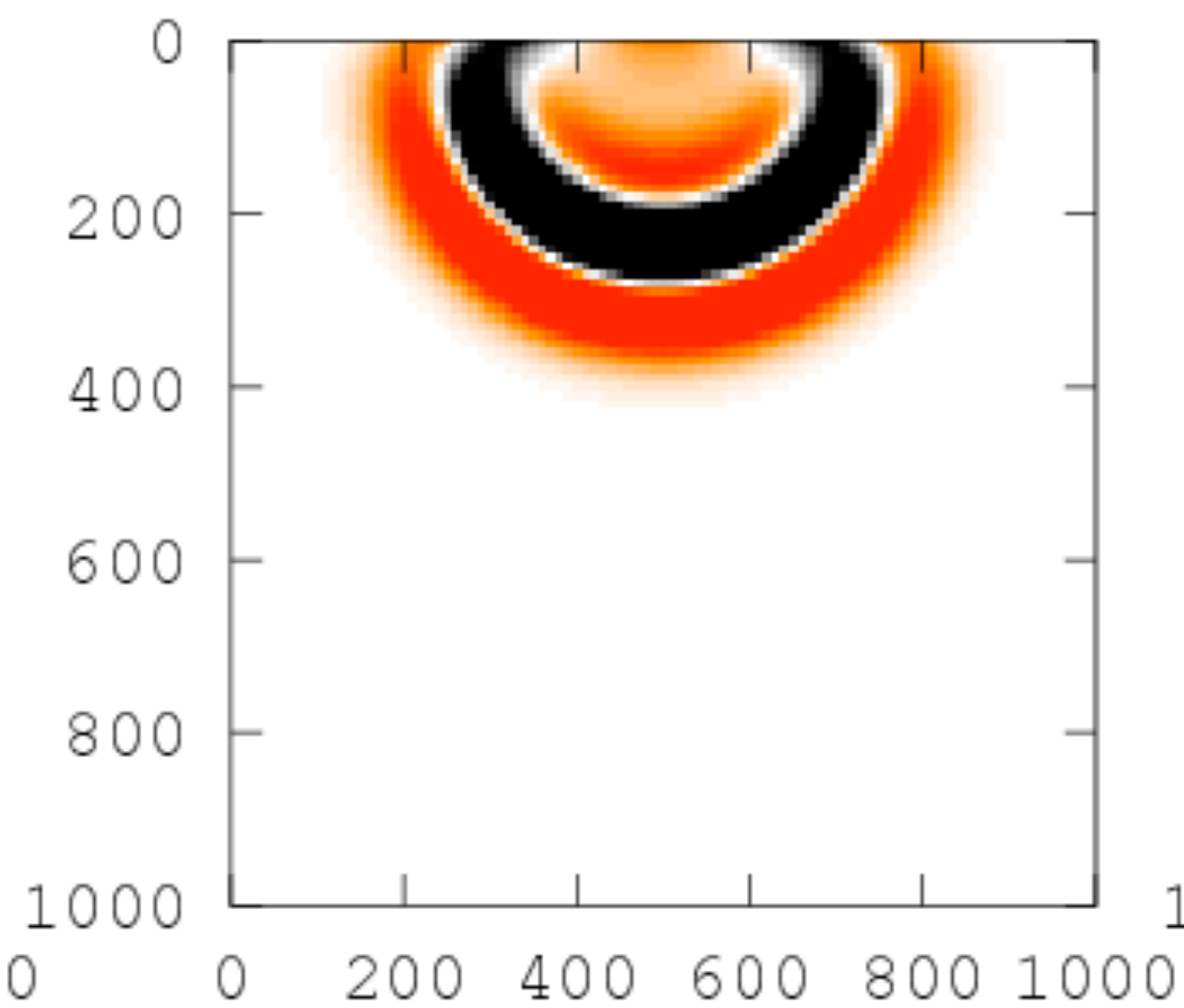
**data-augmented
wavefield in *constant* model**



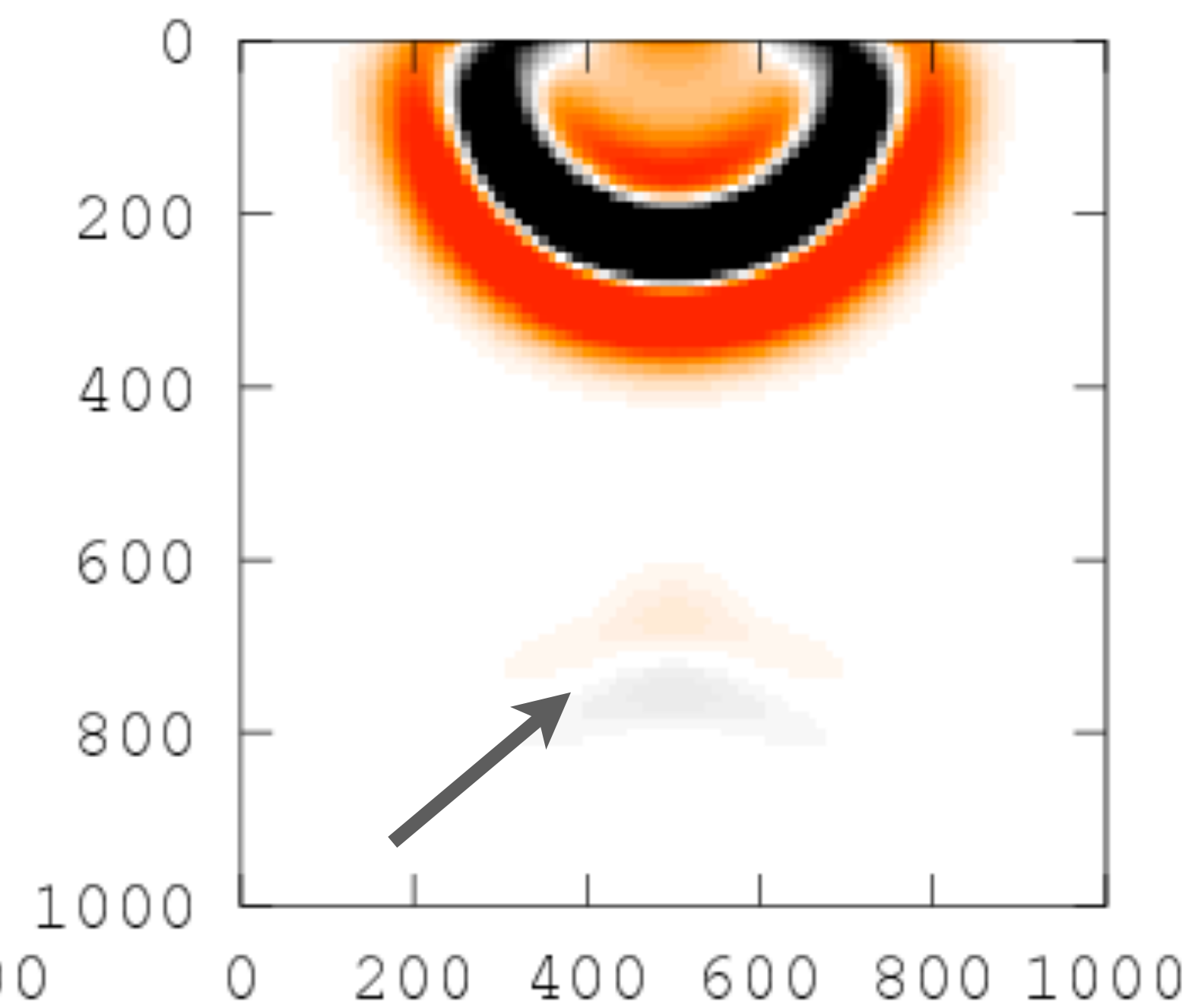
wavefield in *true* model



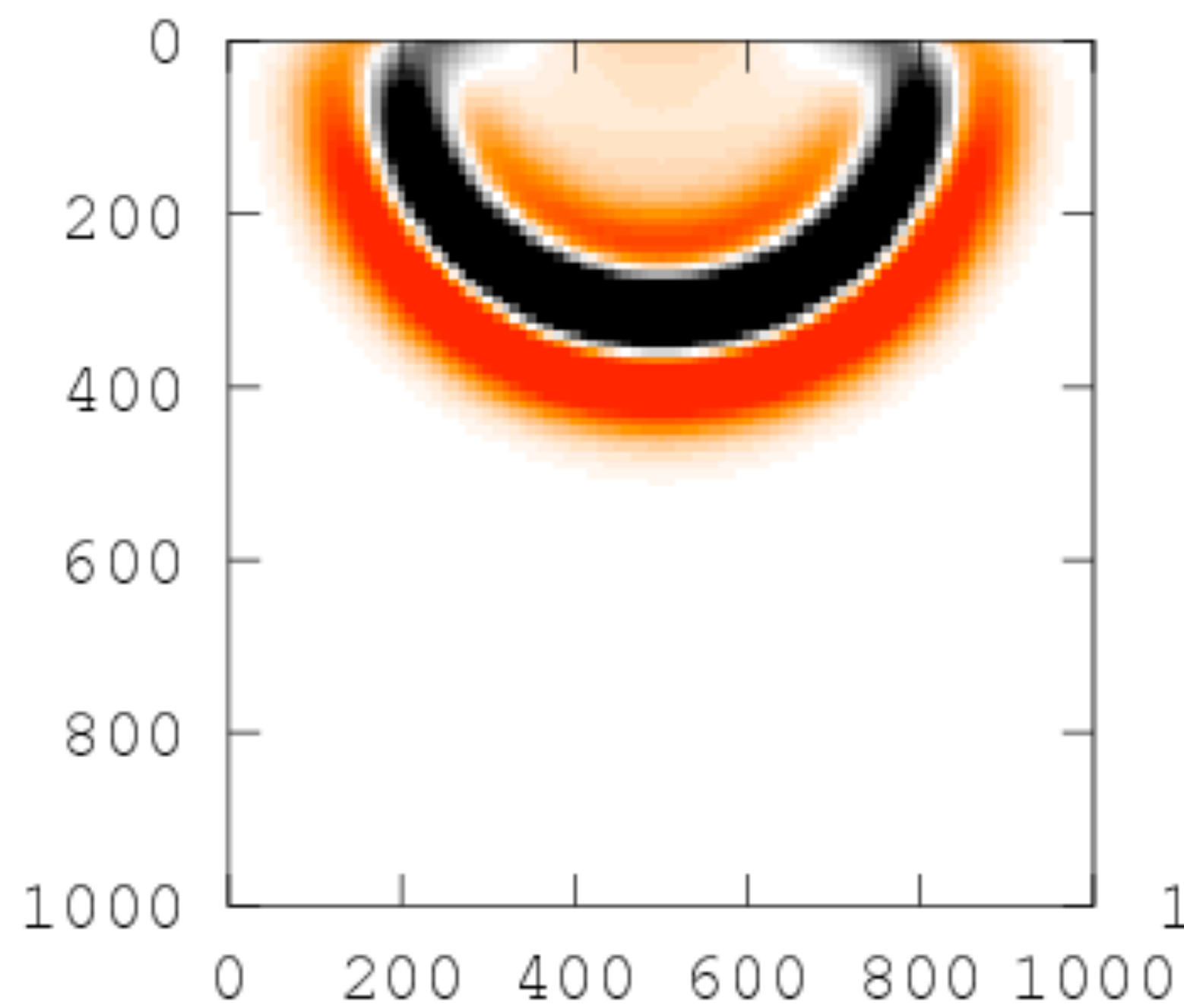
wavefield in *constant* model



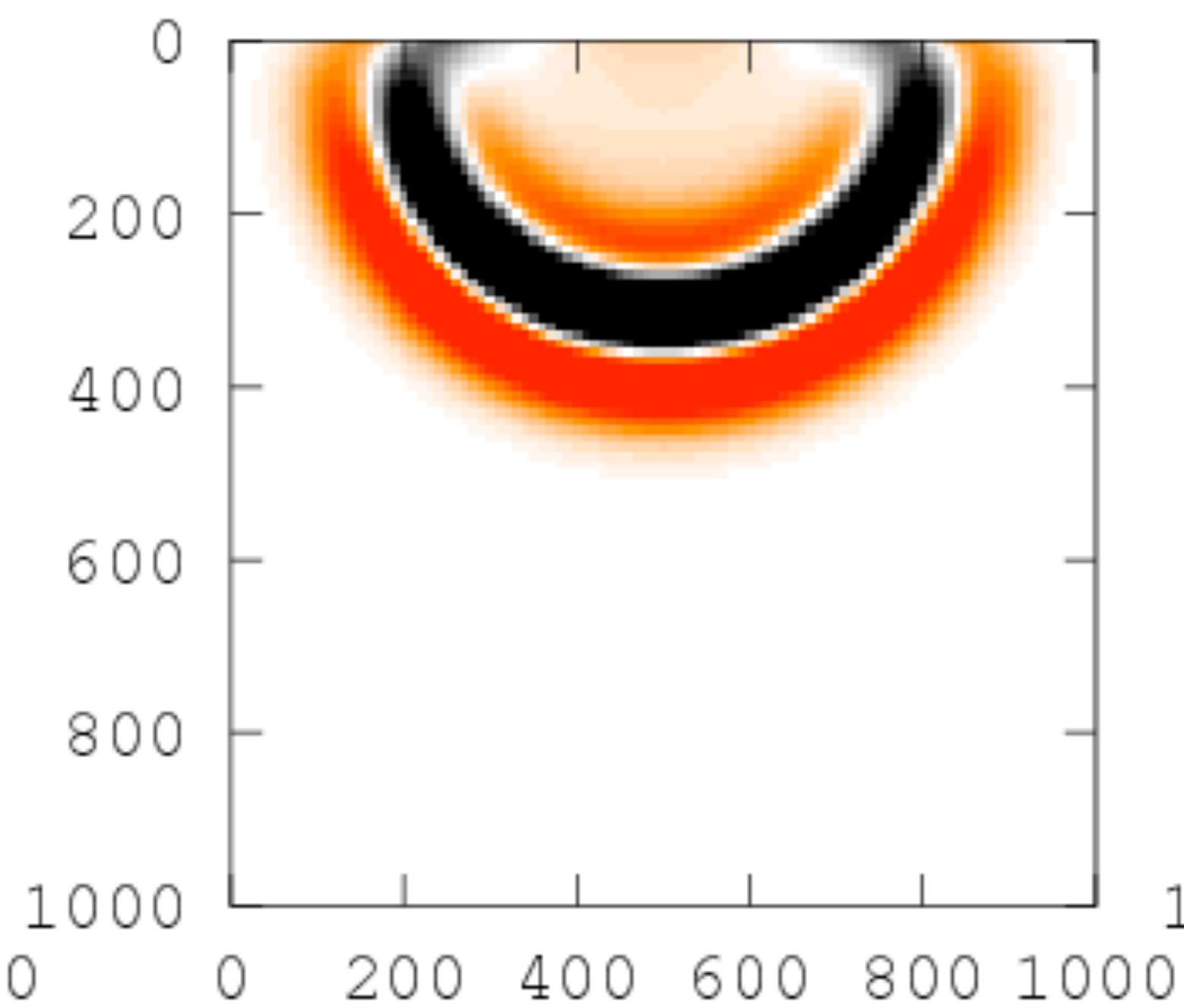
**data-augmented
wavefield in *constant* model**



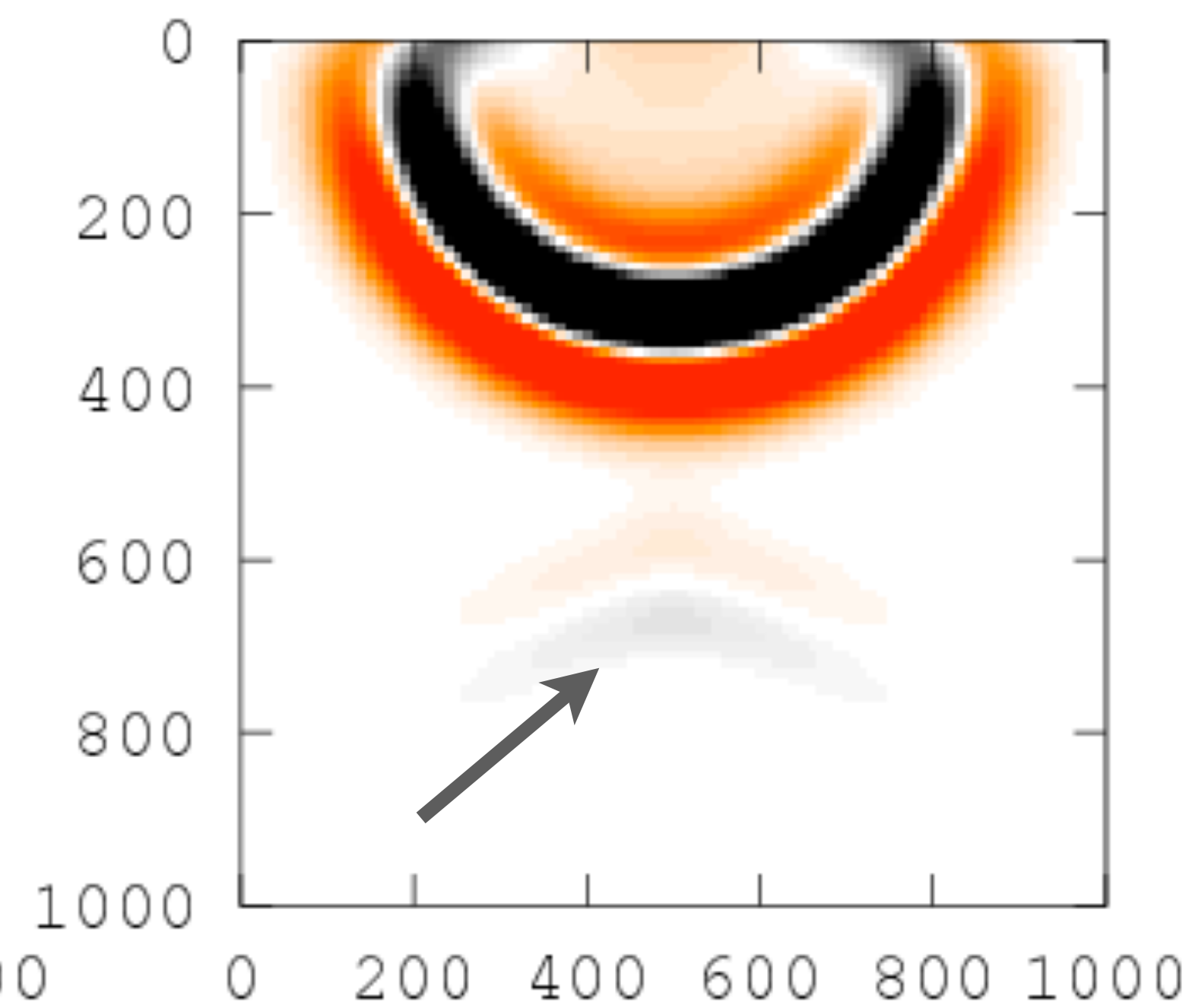
wavefield in *true* model



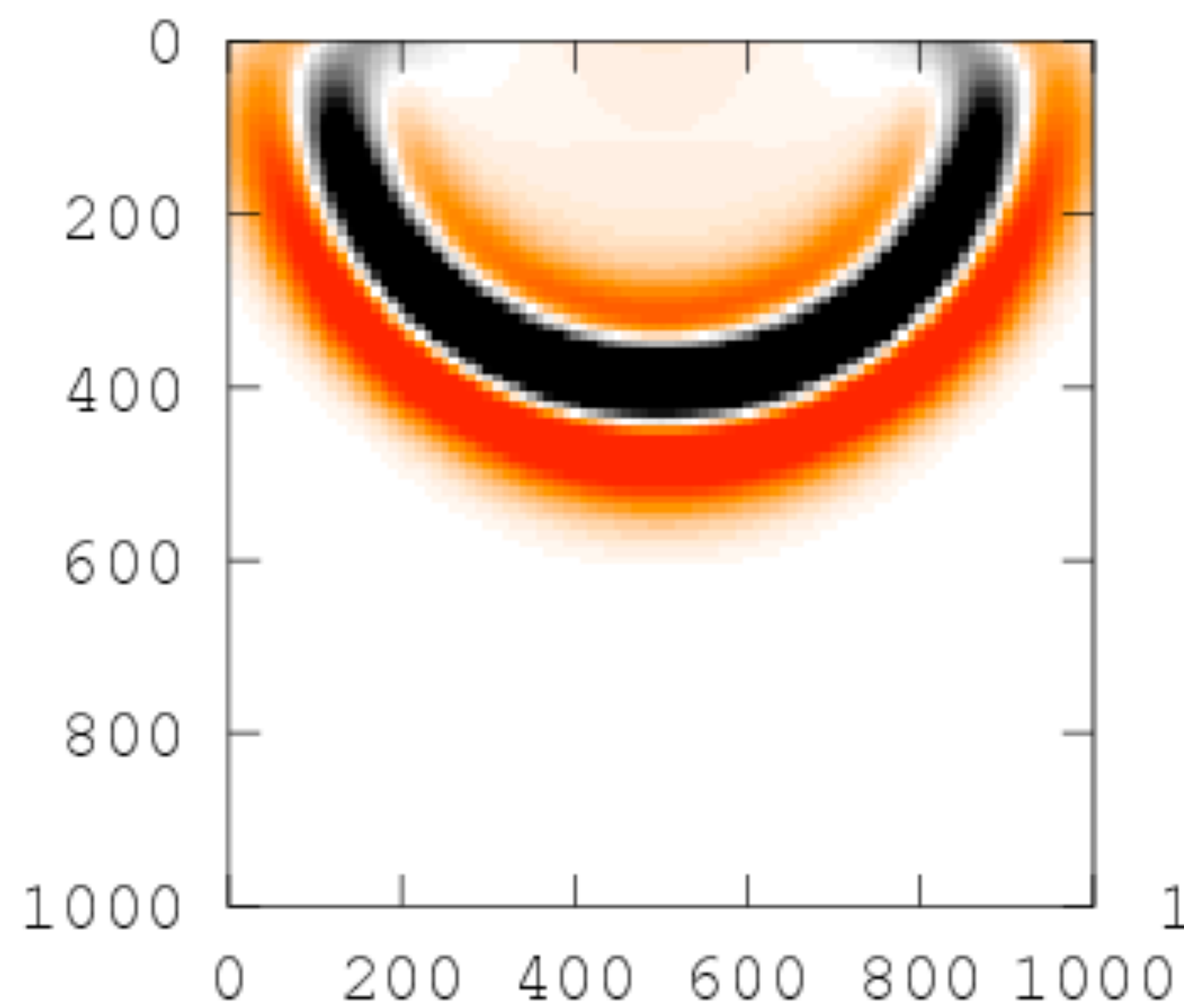
wavefield in *constant* model



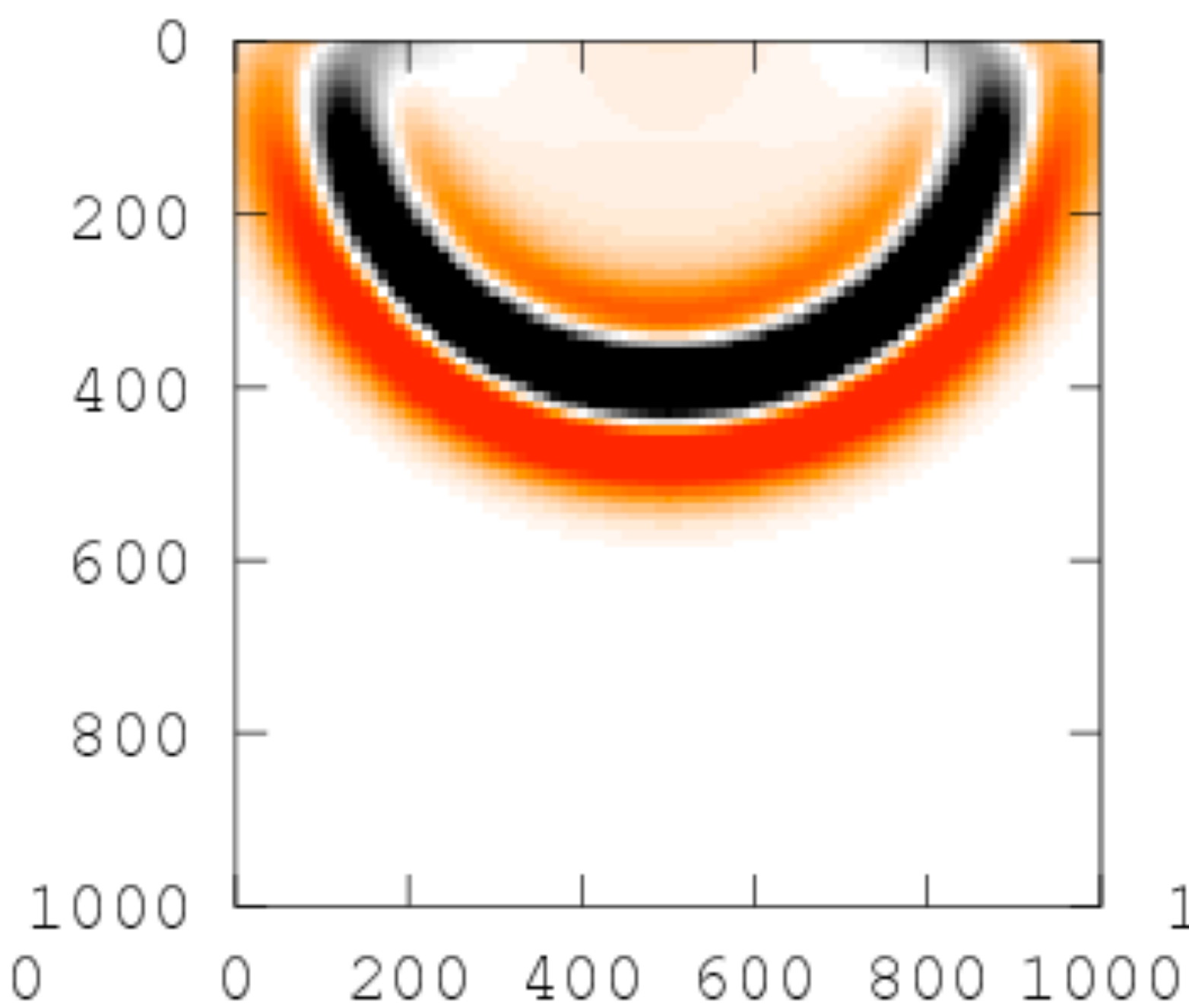
**data-augmented
wavefield in *constant* model**



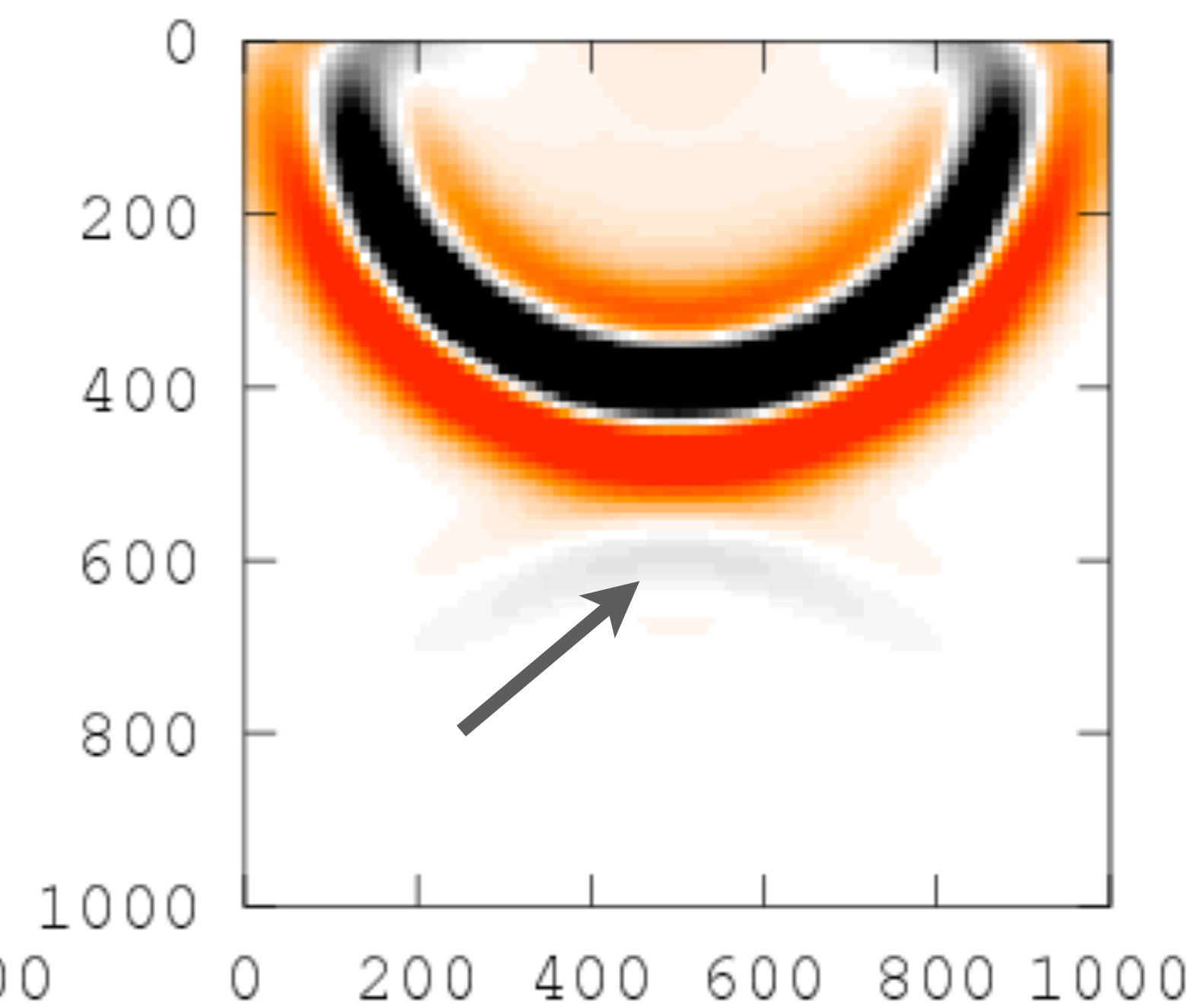
wavefield in *true* model



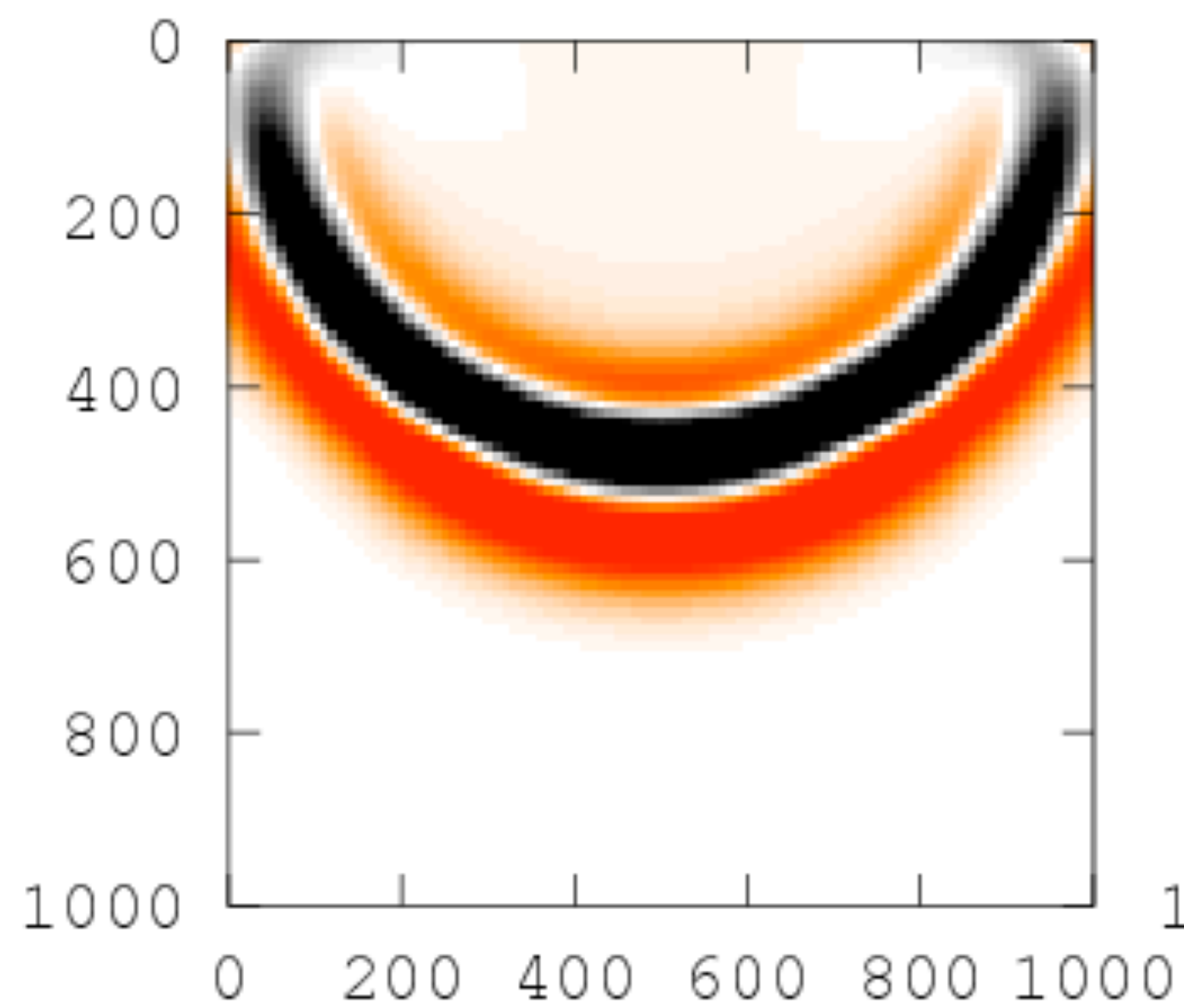
wavefield in *constant* model



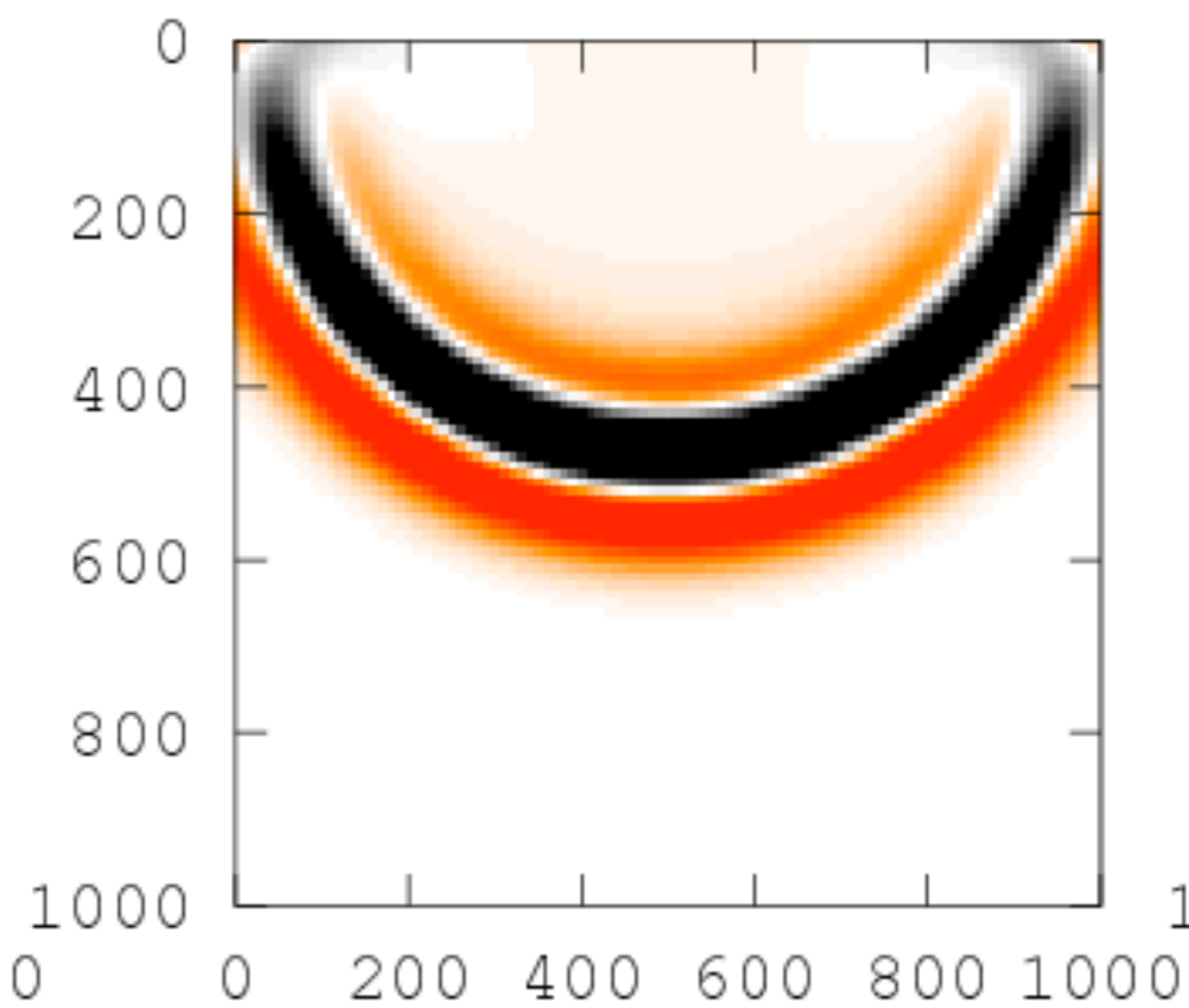
**data-augmented
wavefield in *constant* model**



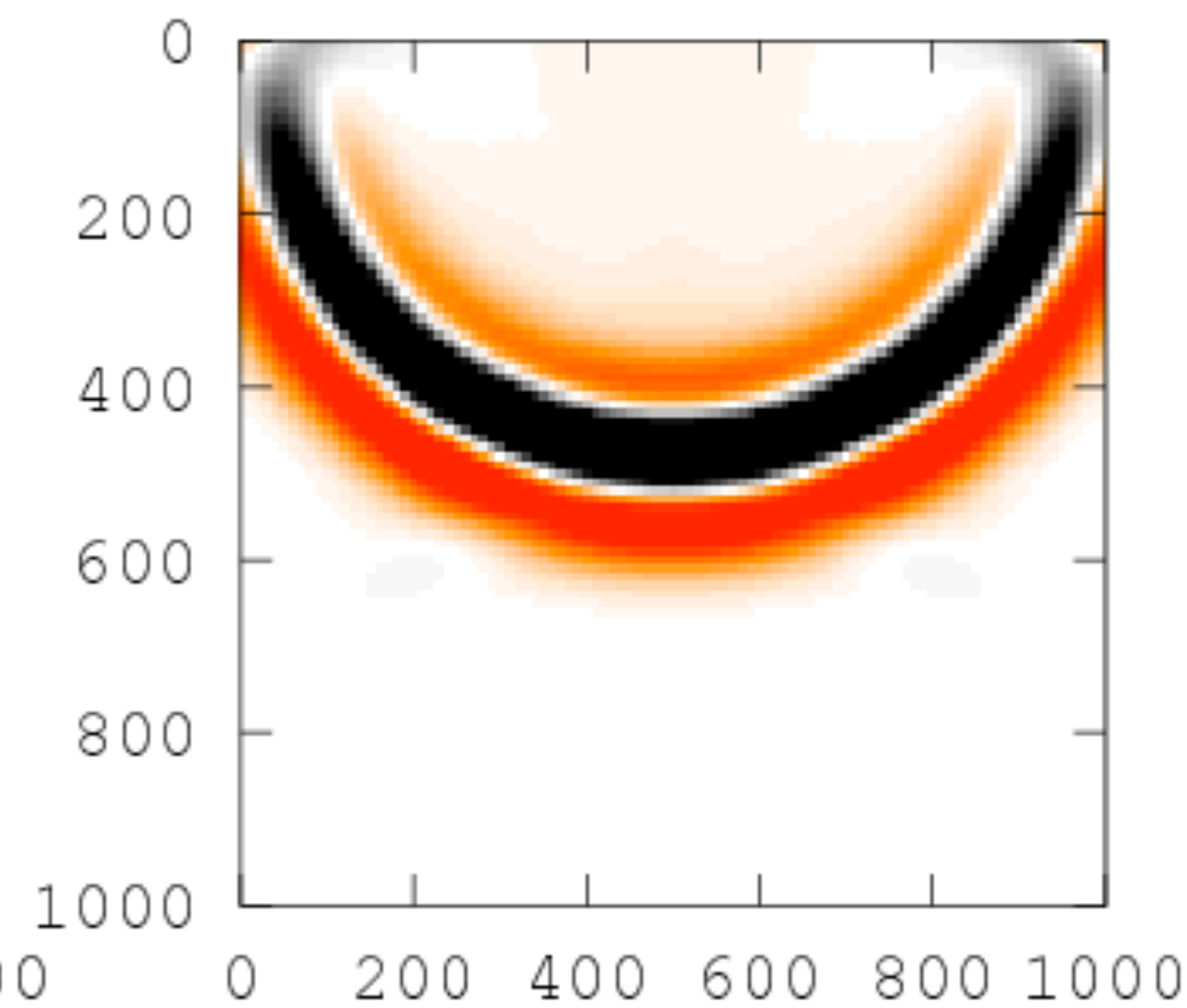
wavefield in *true* model



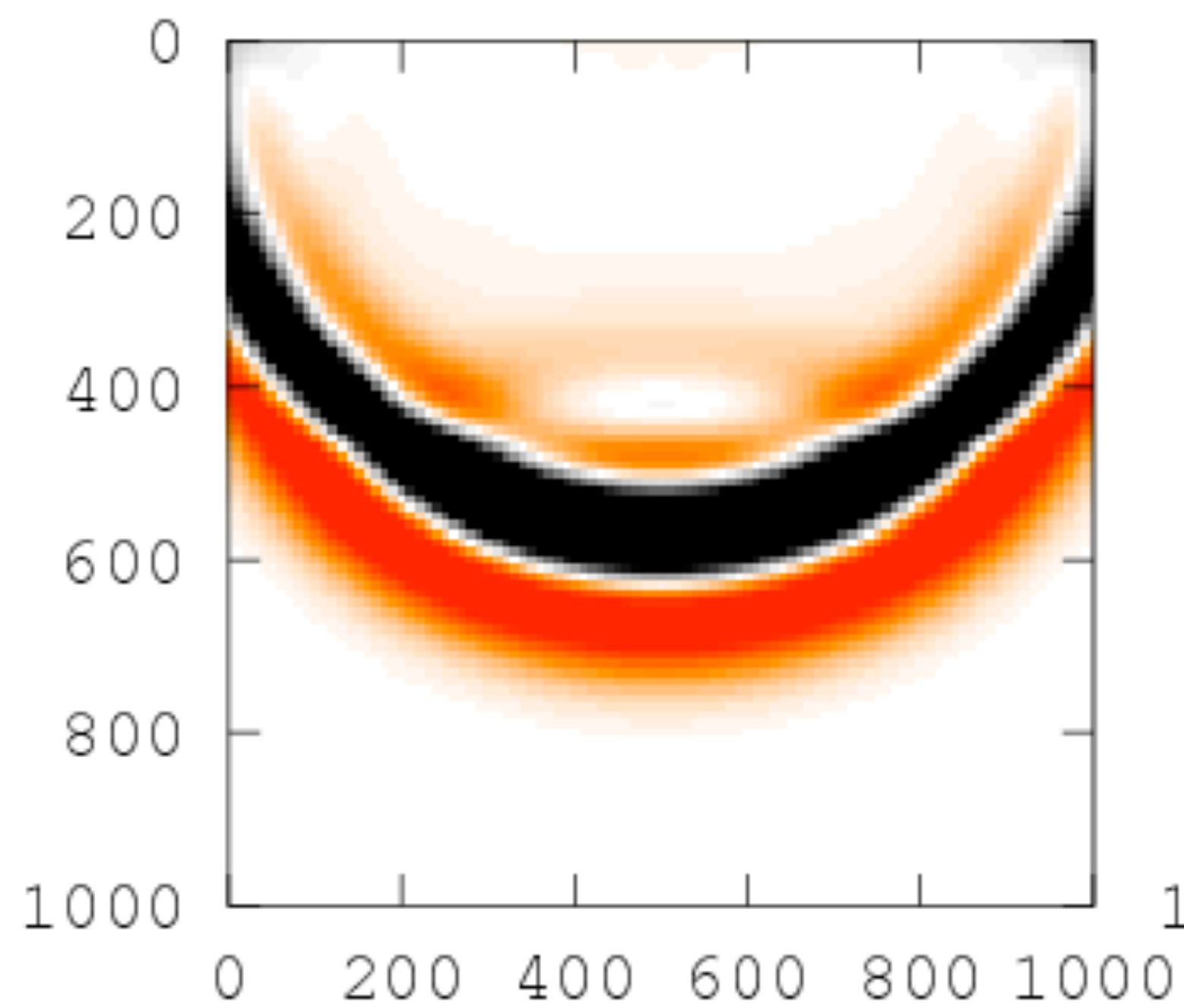
wavefield in *constant* model



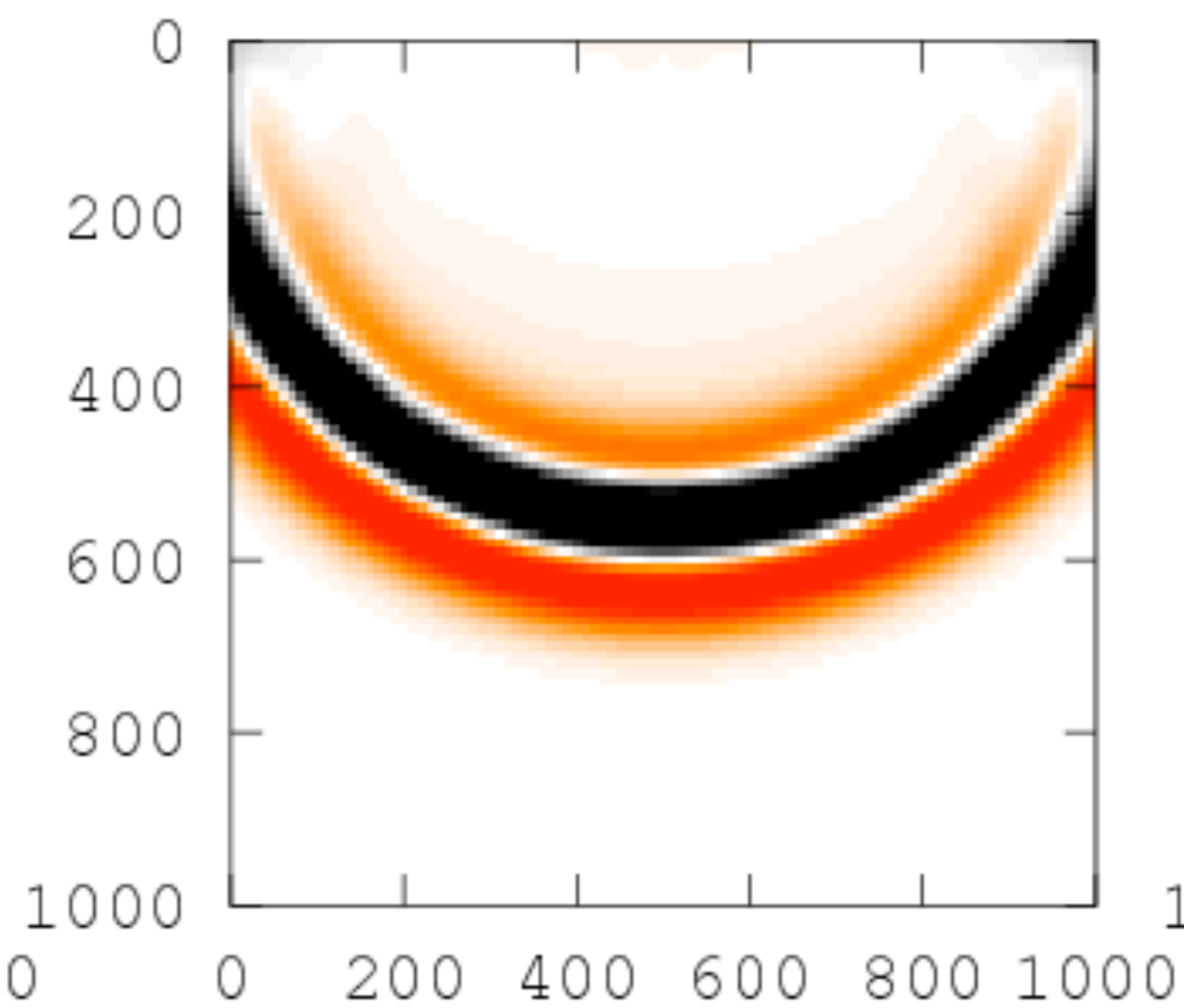
**data-augmented
wavefield in *constant* model**



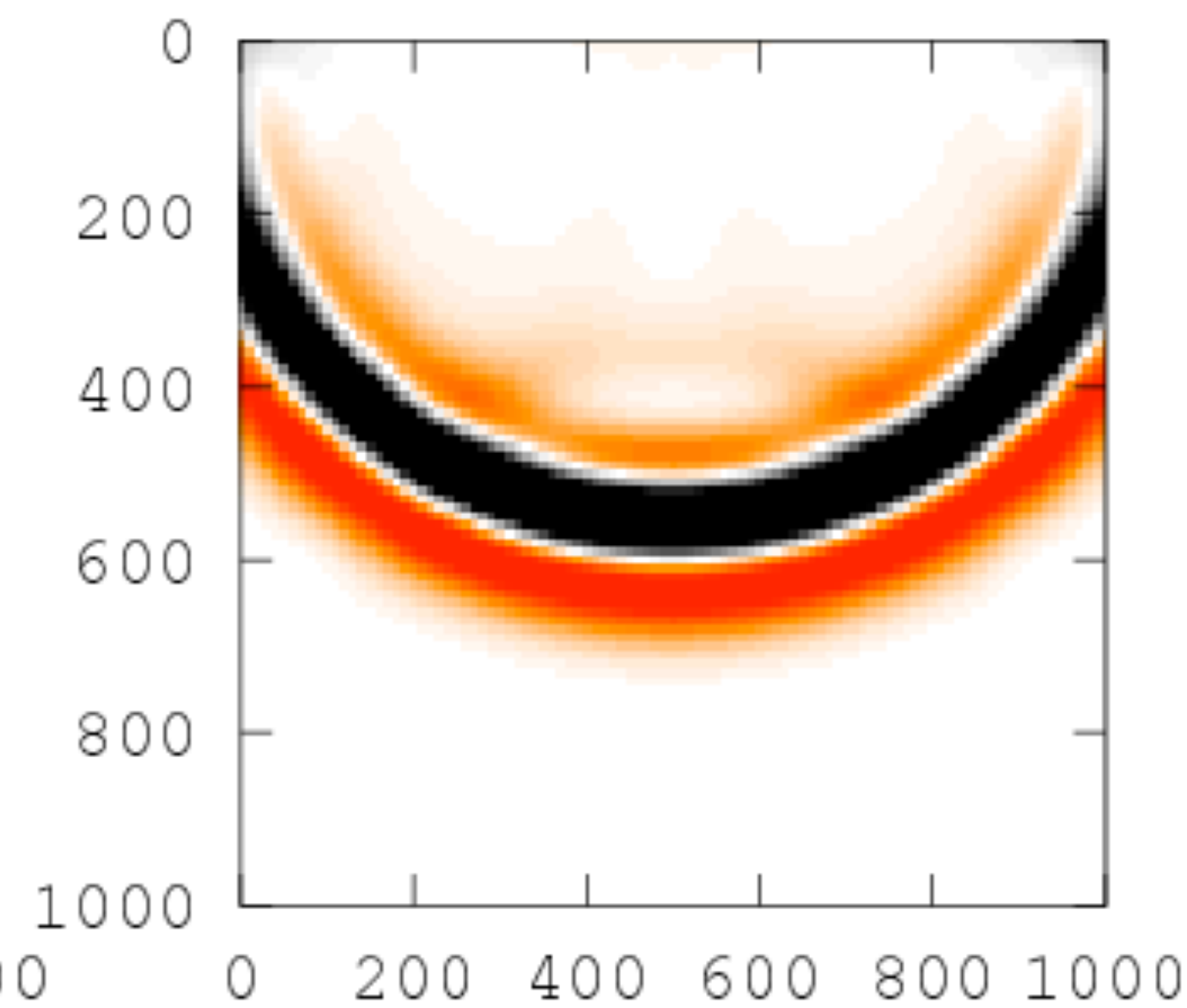
wavefield in *true* model



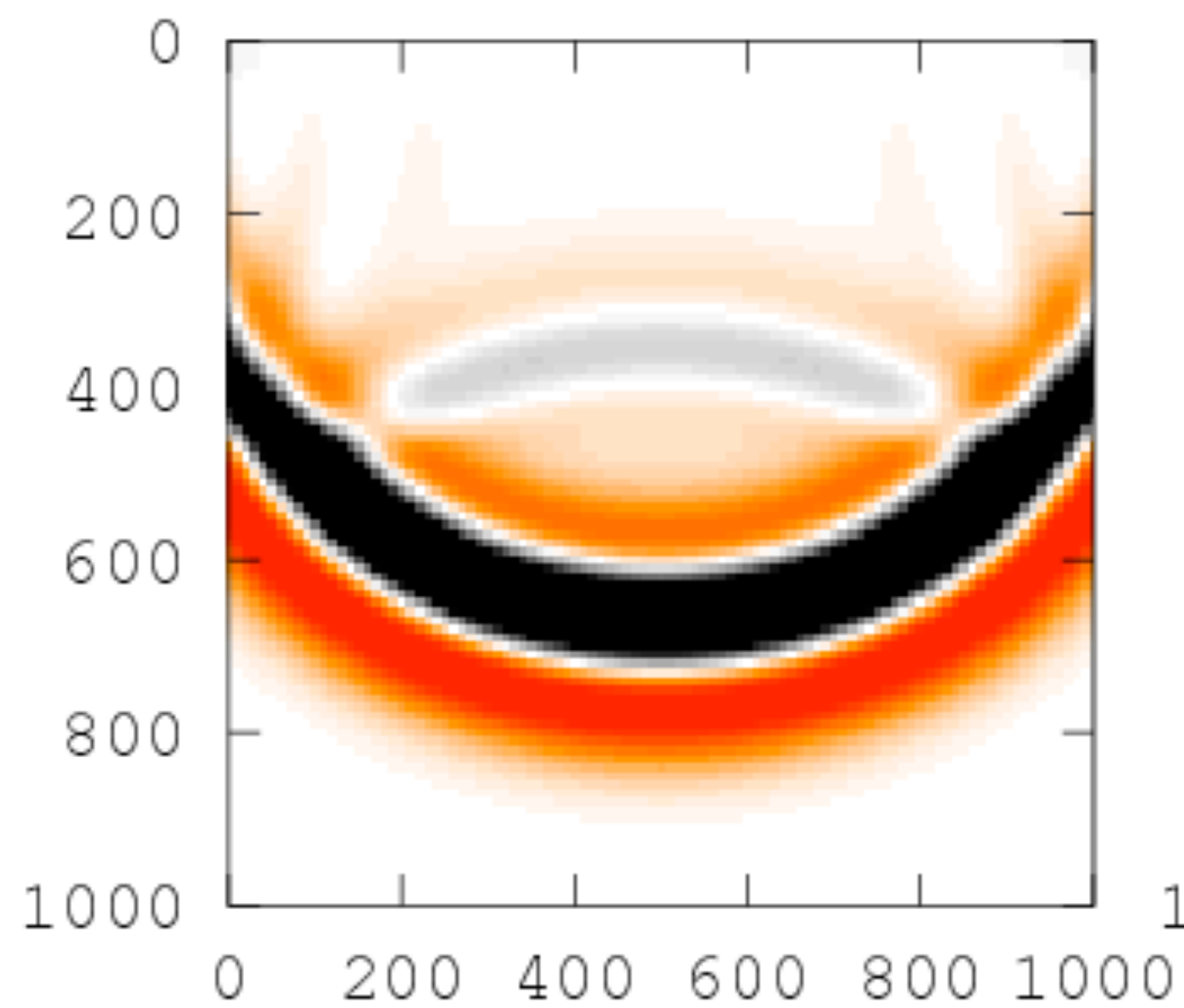
wavefield in *constant* model



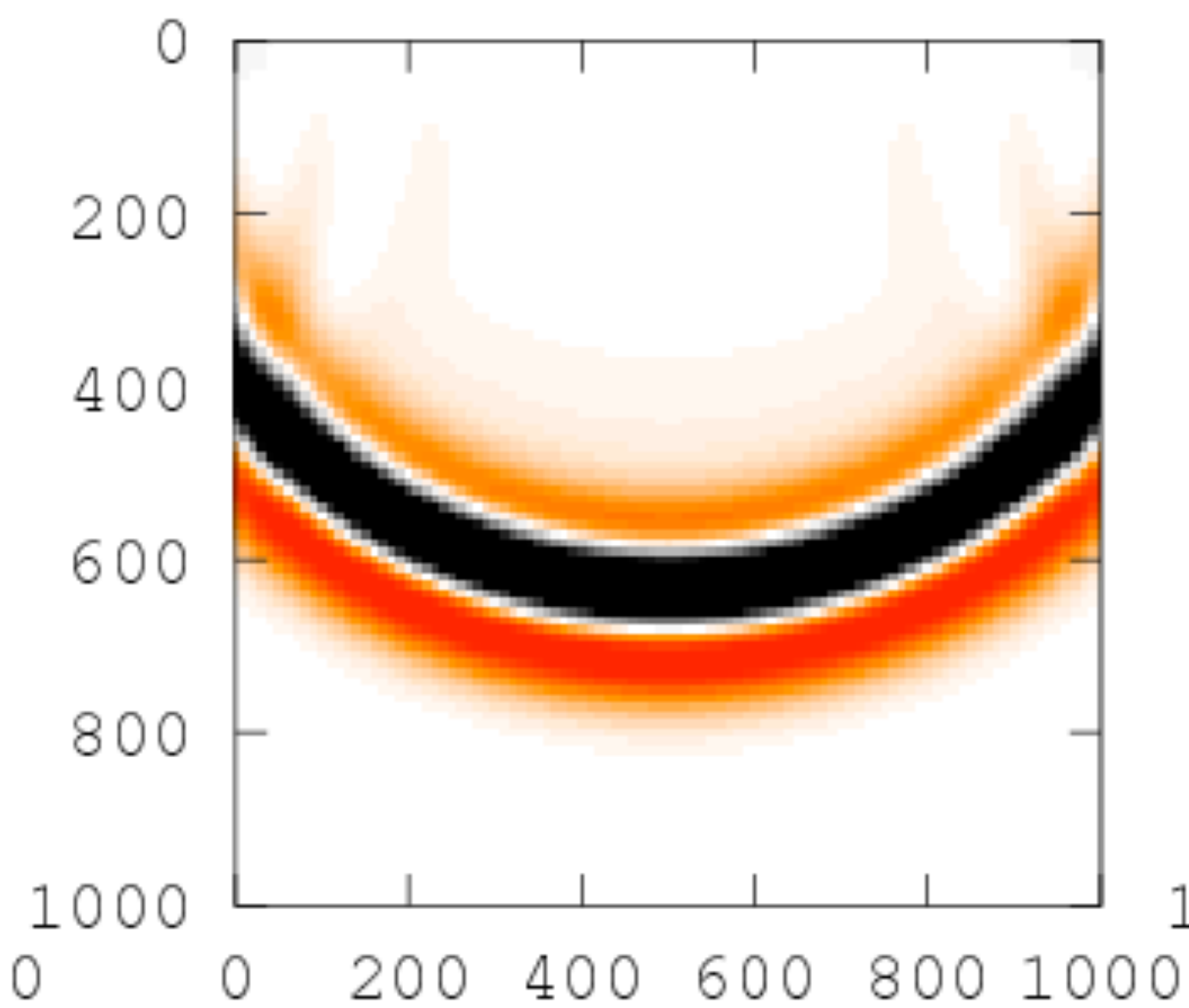
**data-augmented
wavefield in *constant* model**



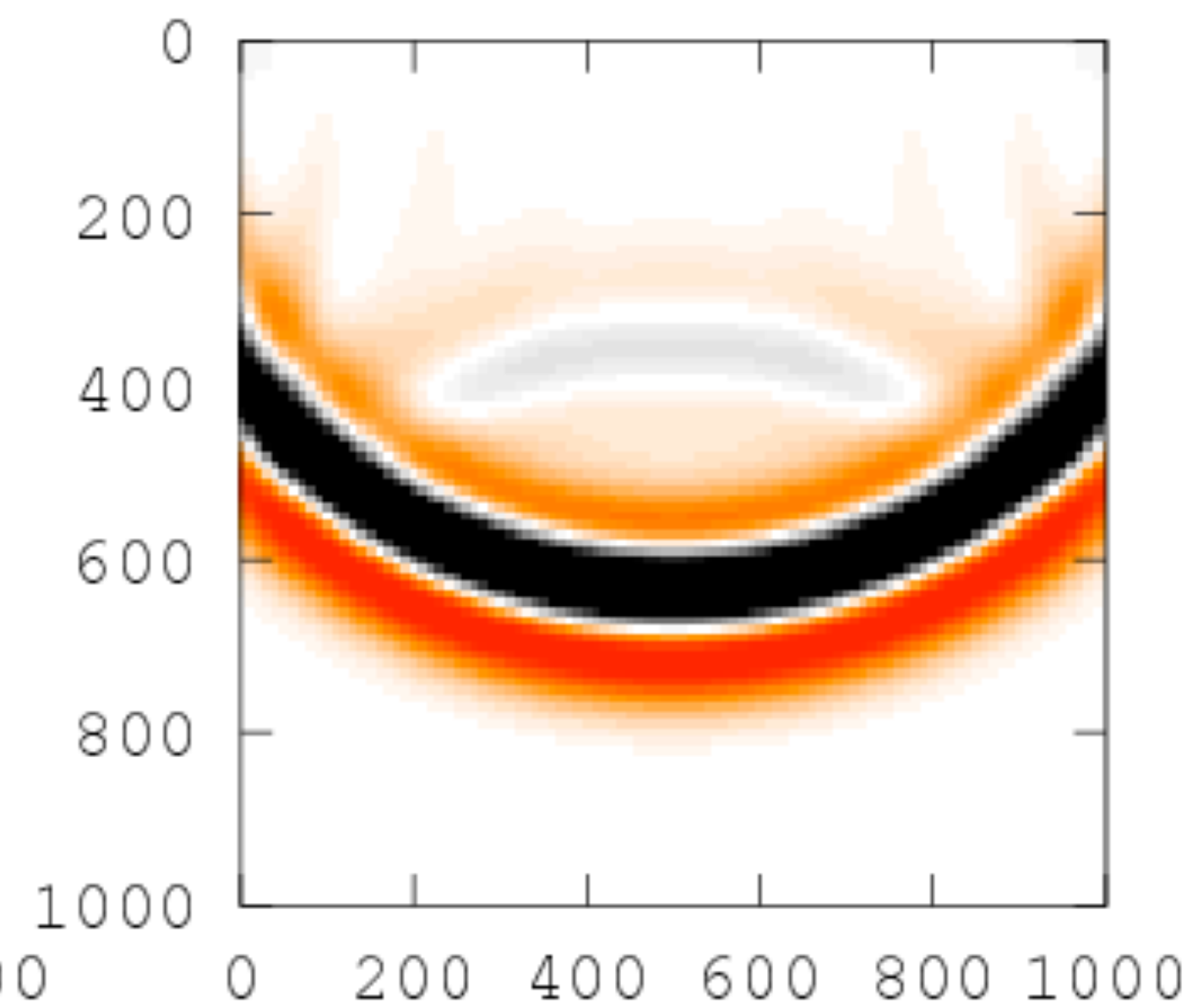
wavefield in *true* model



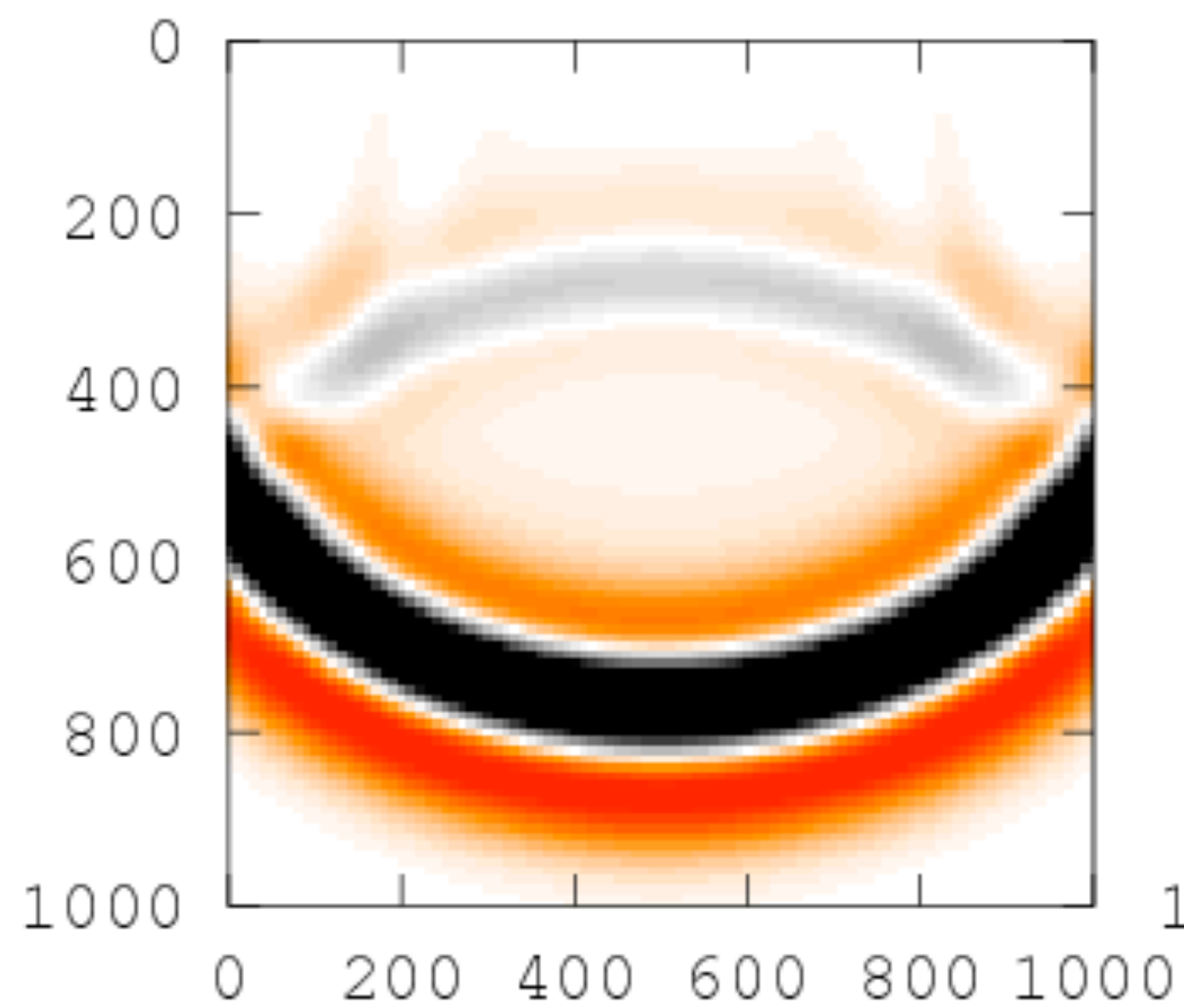
wavefield in *constant* model



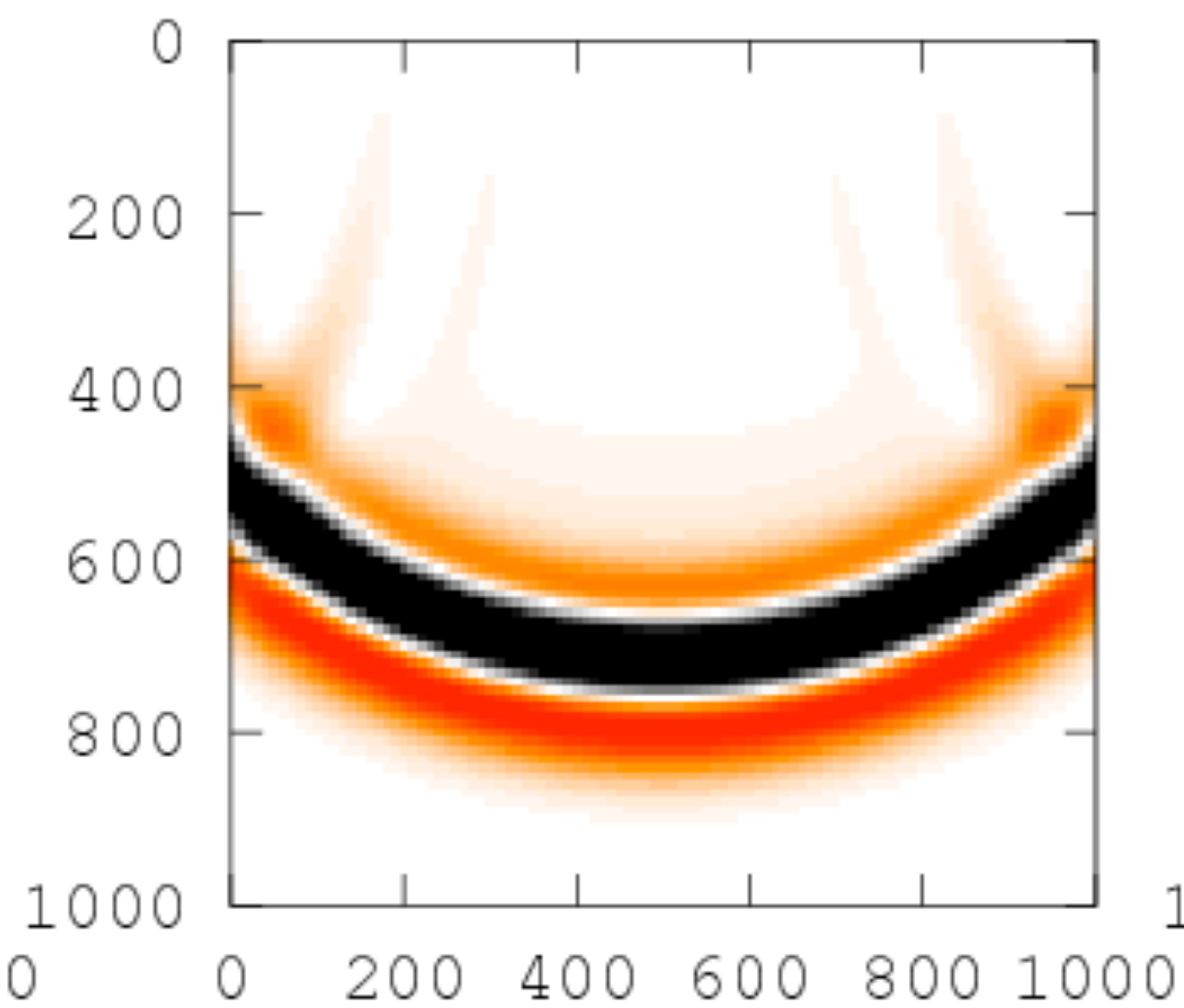
**data-augmented
wavefield in *constant* model**



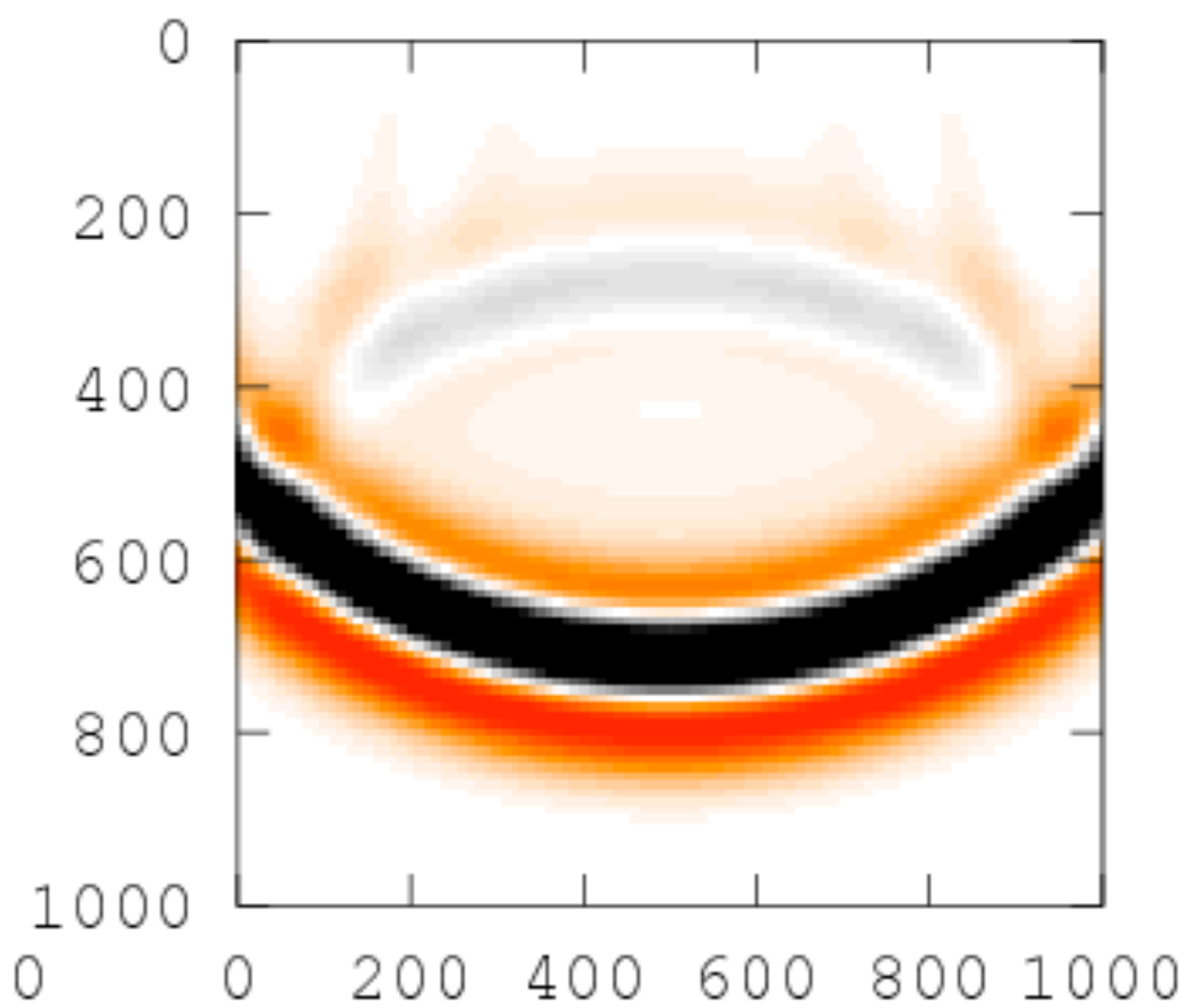
wavefield in *true* model



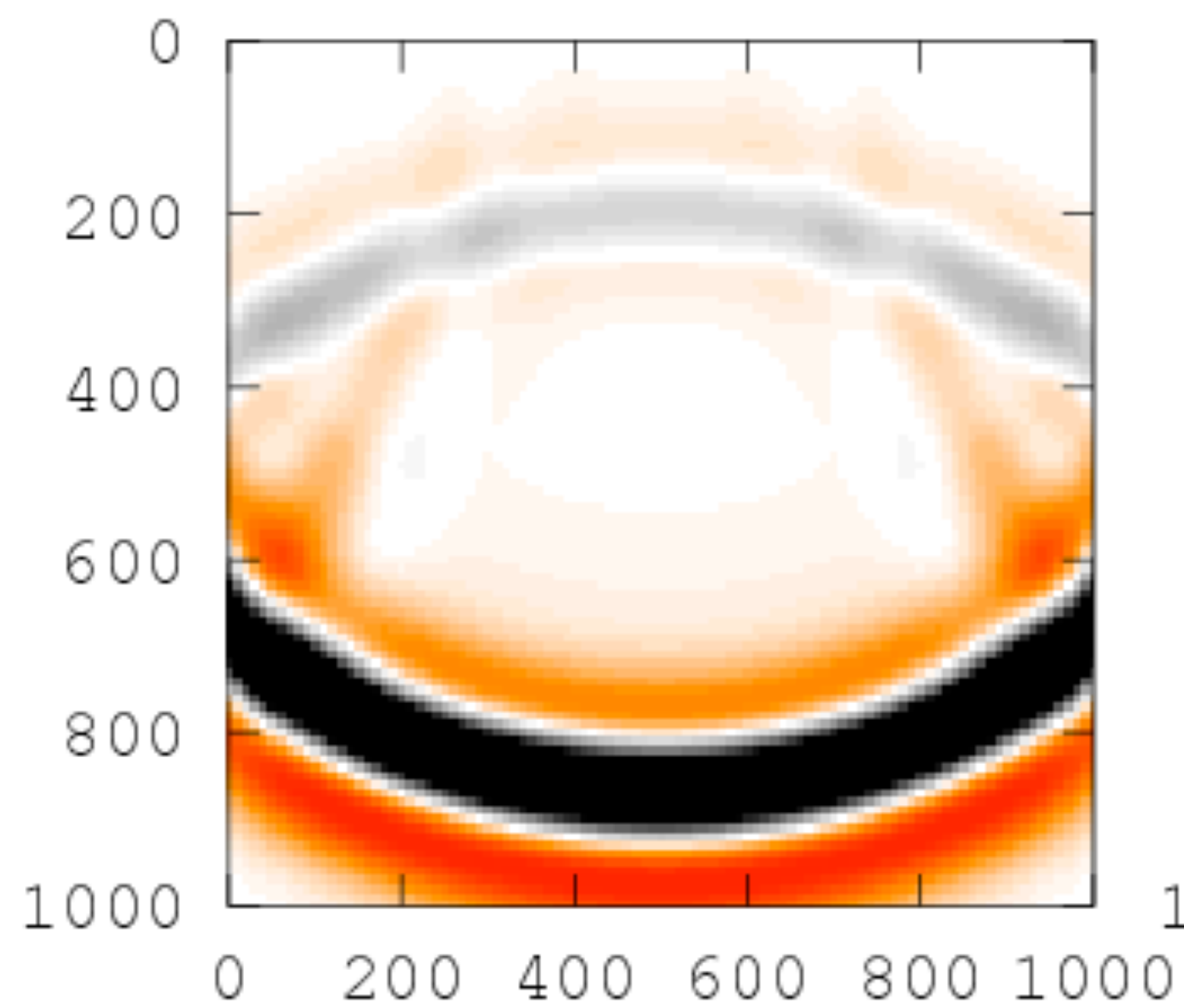
wavefield in *constant* model



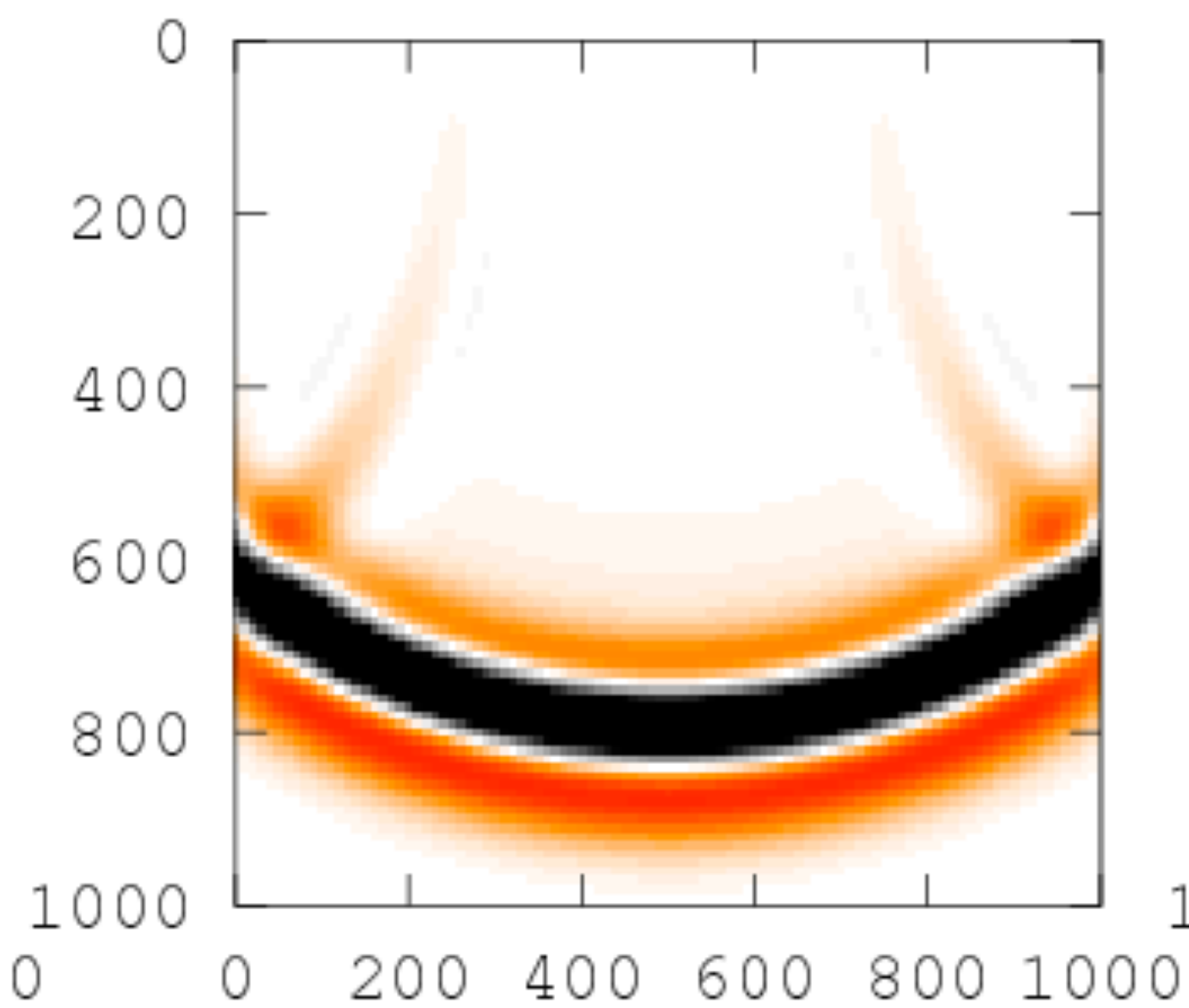
**data-augmented
wavefield in *constant* model**



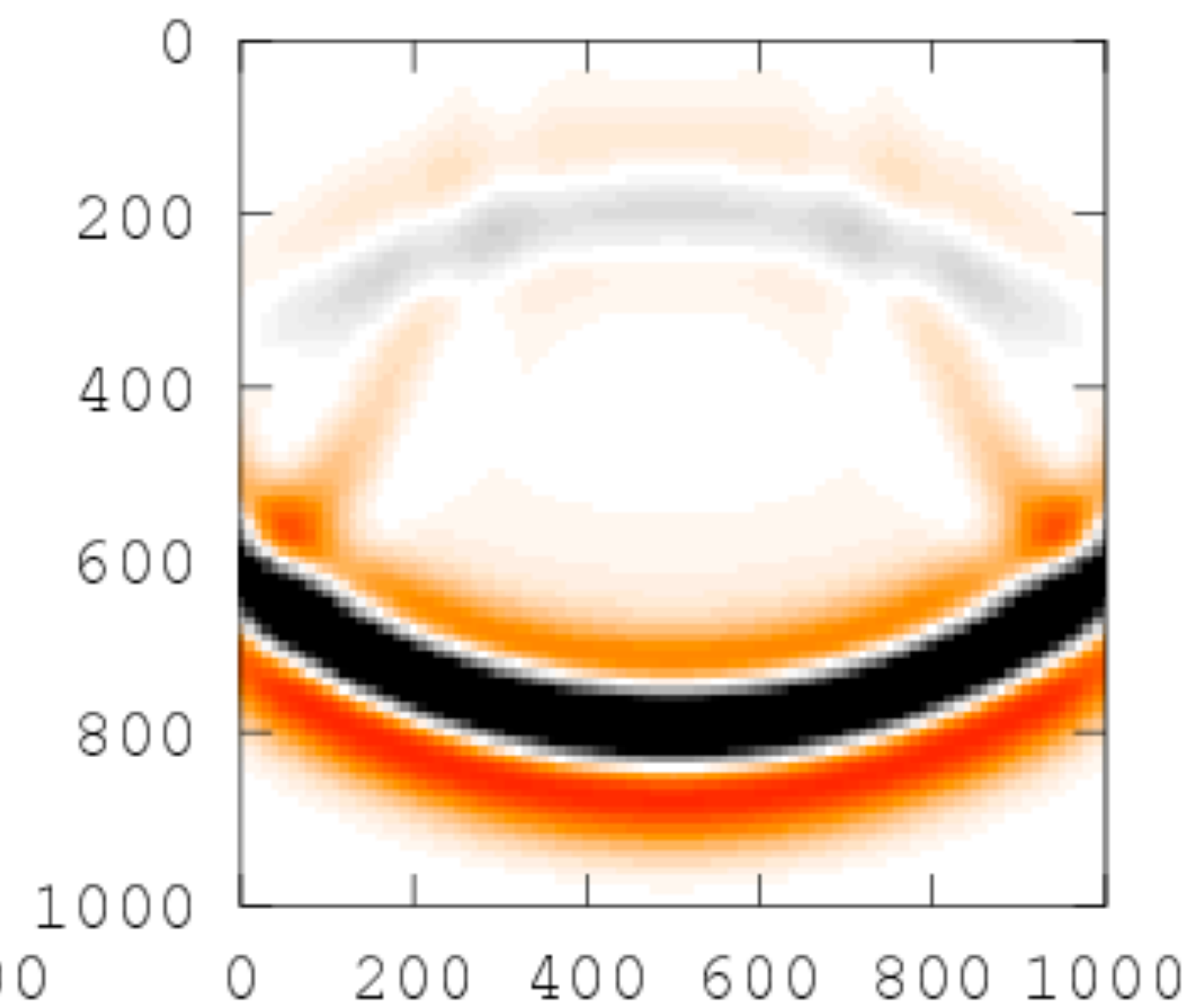
wavefield in *true* model



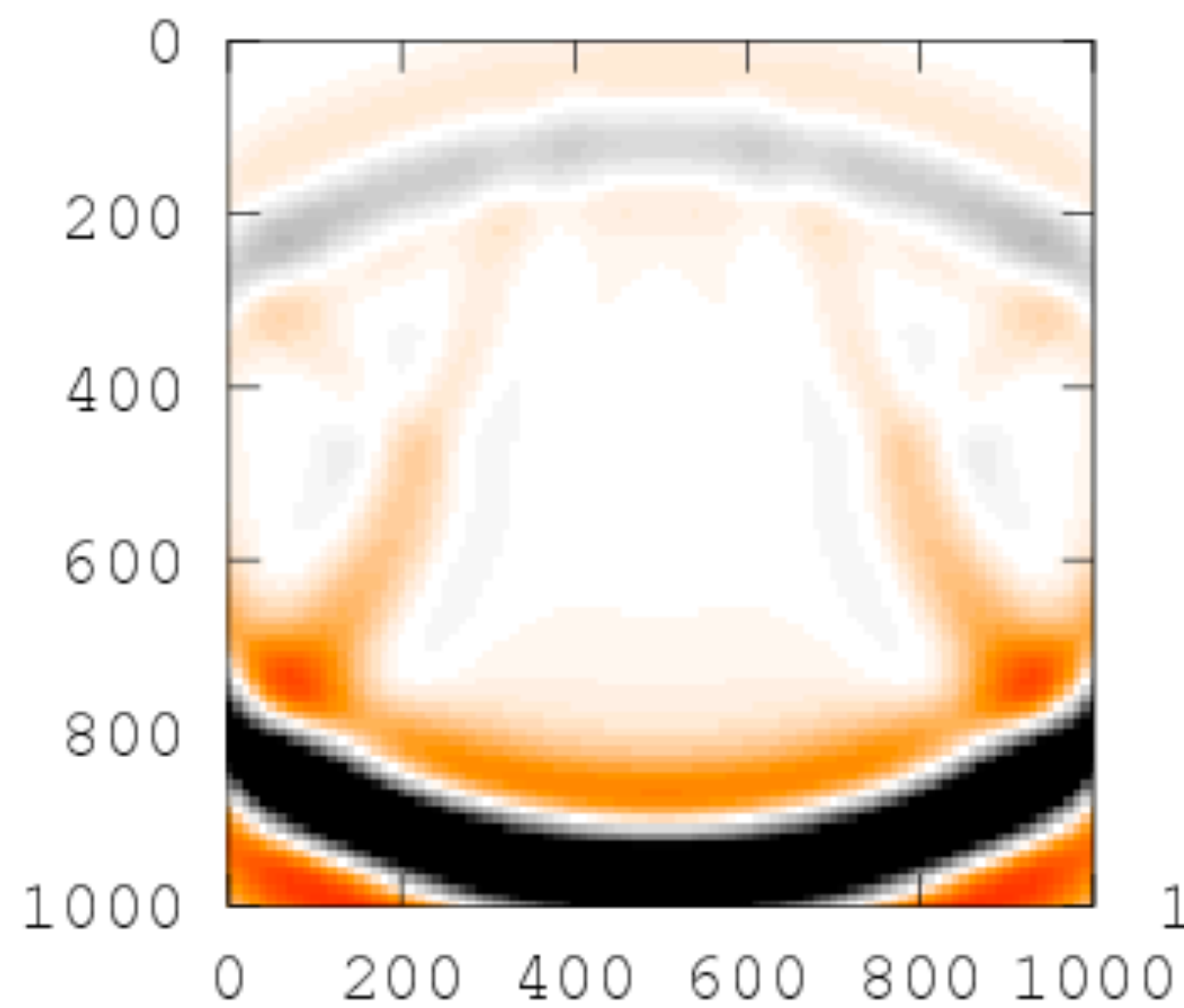
wavefield in *constant* model



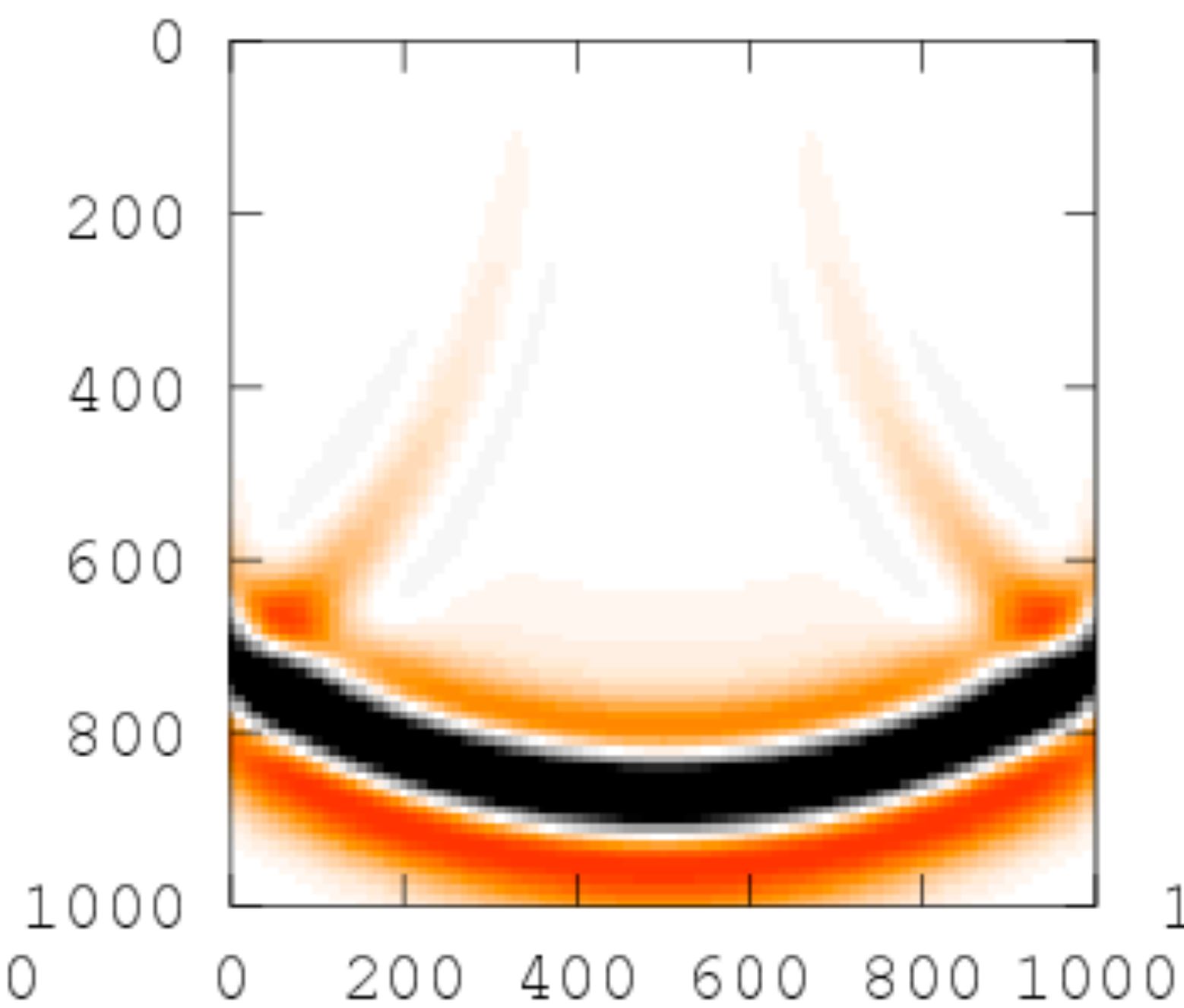
**data-augmented
wavefield in *constant* model**



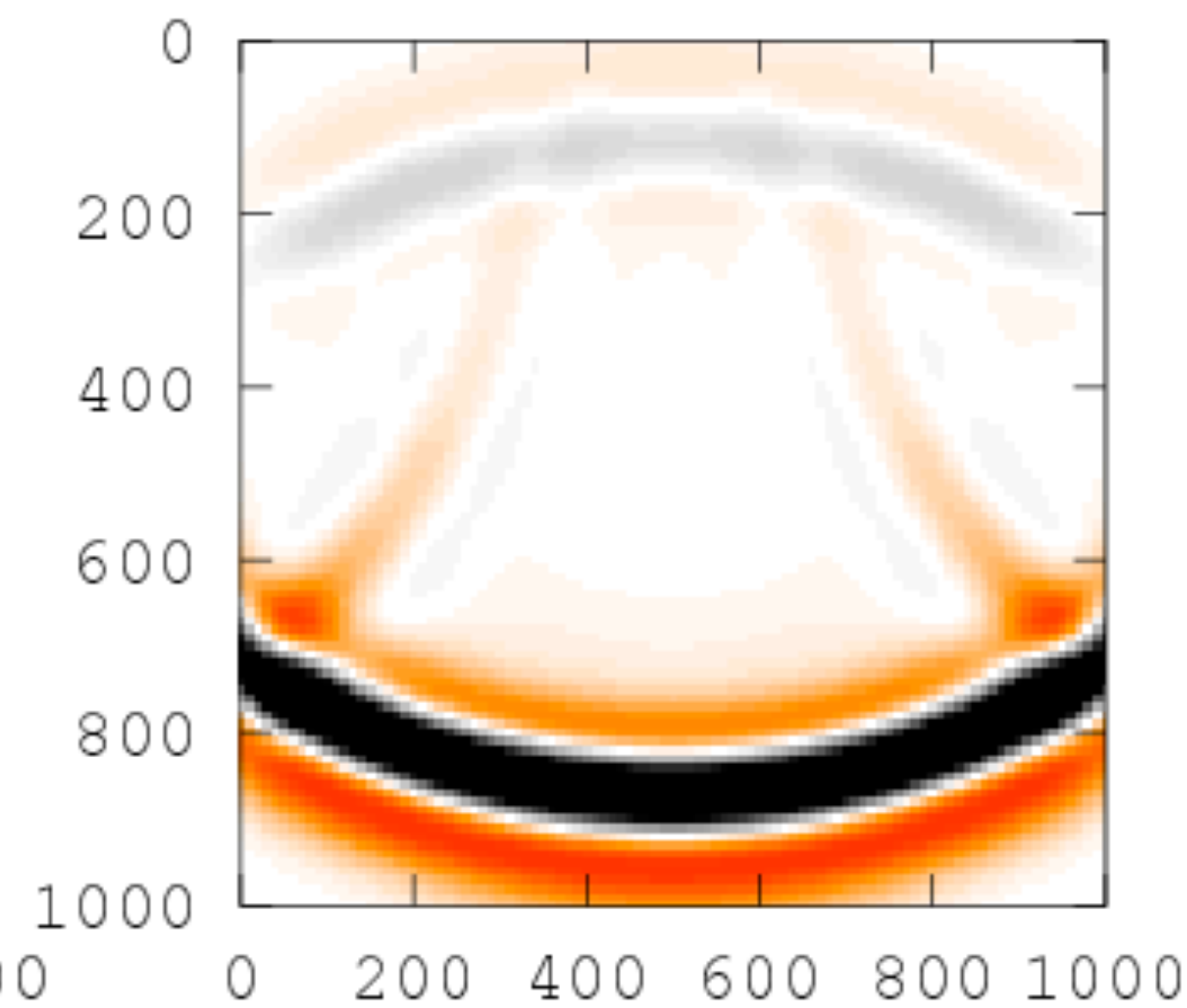
wavefield in *true* model



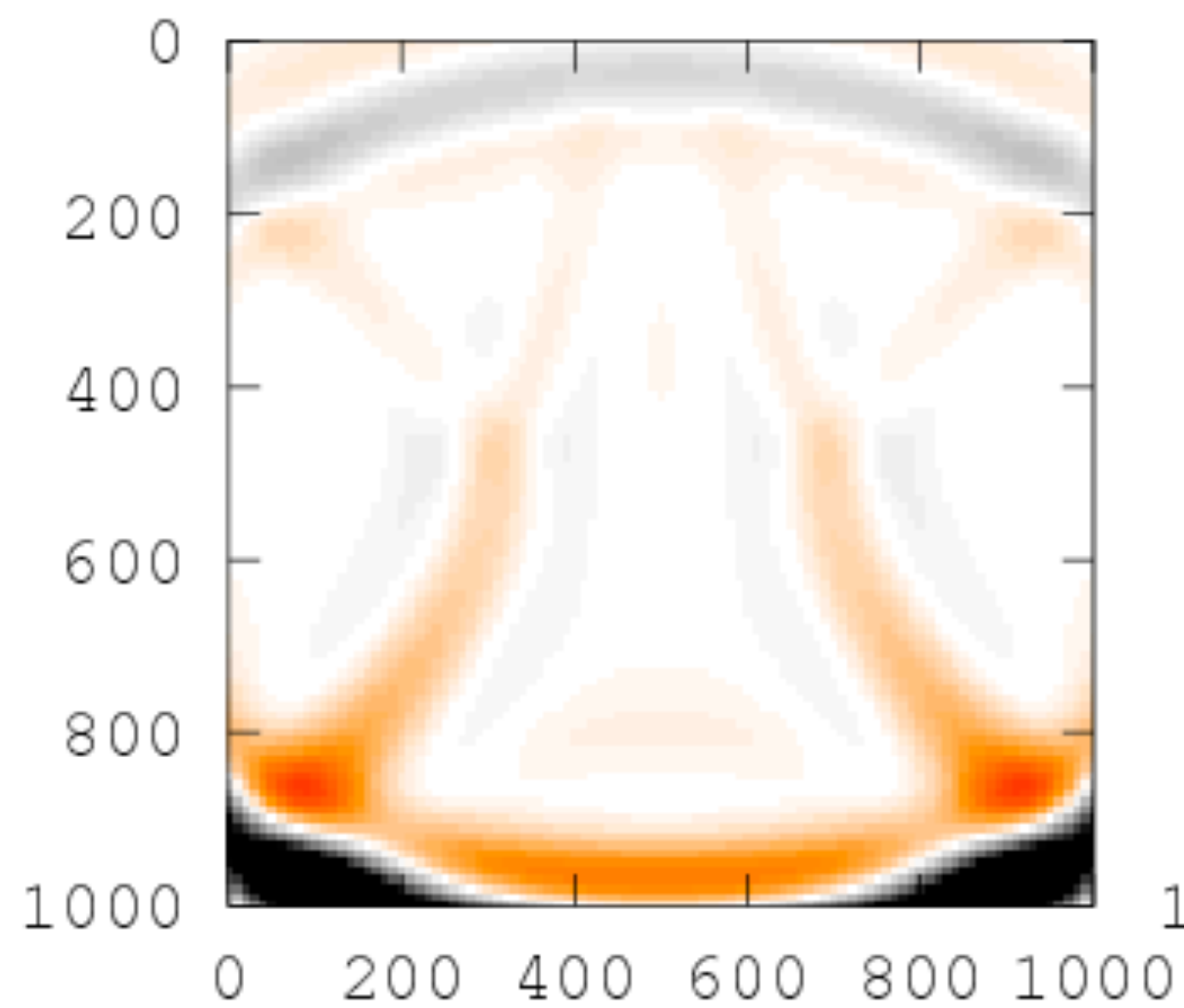
wavefield in *constant* model



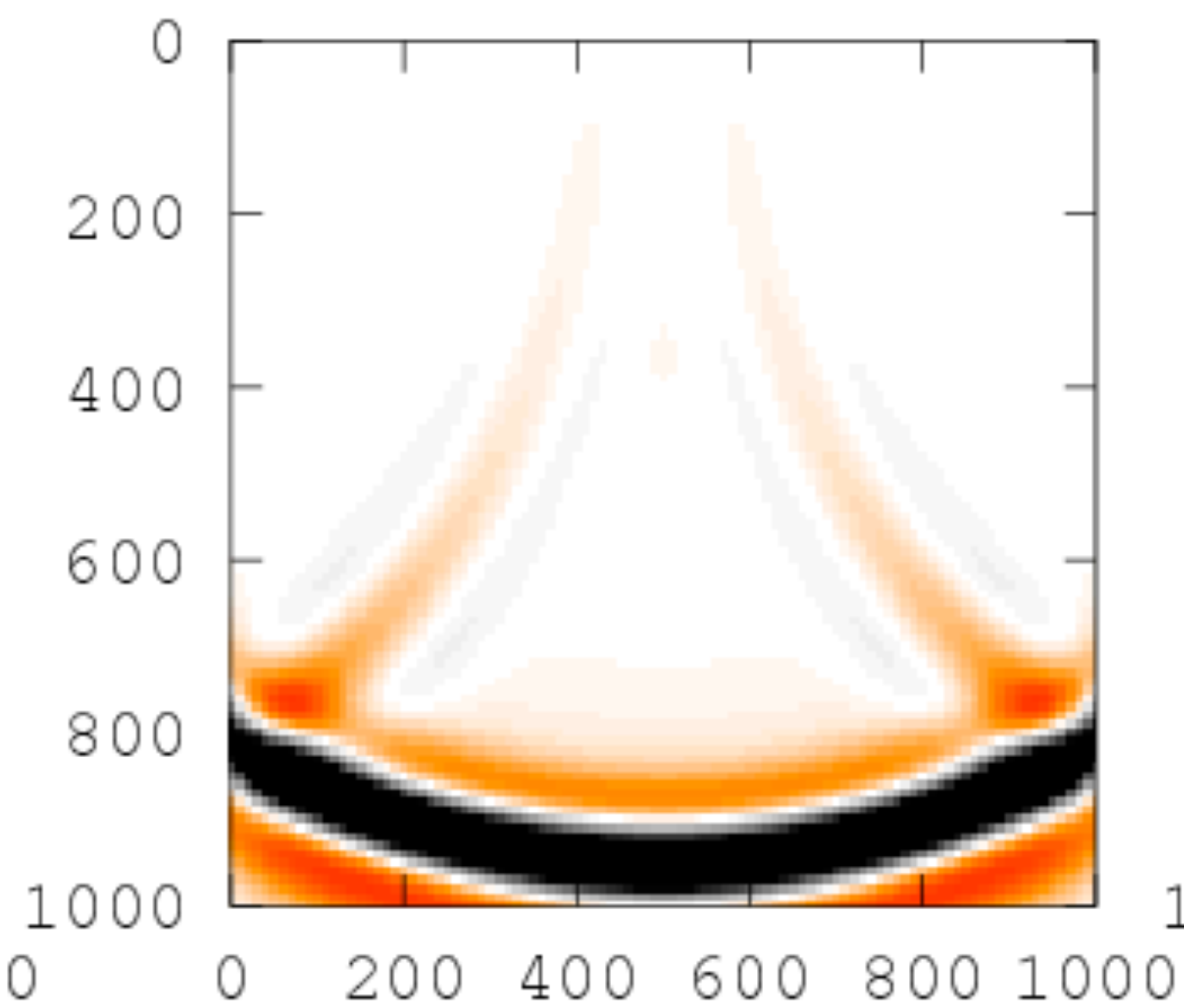
**data-augmented
wavefield in *constant* model**



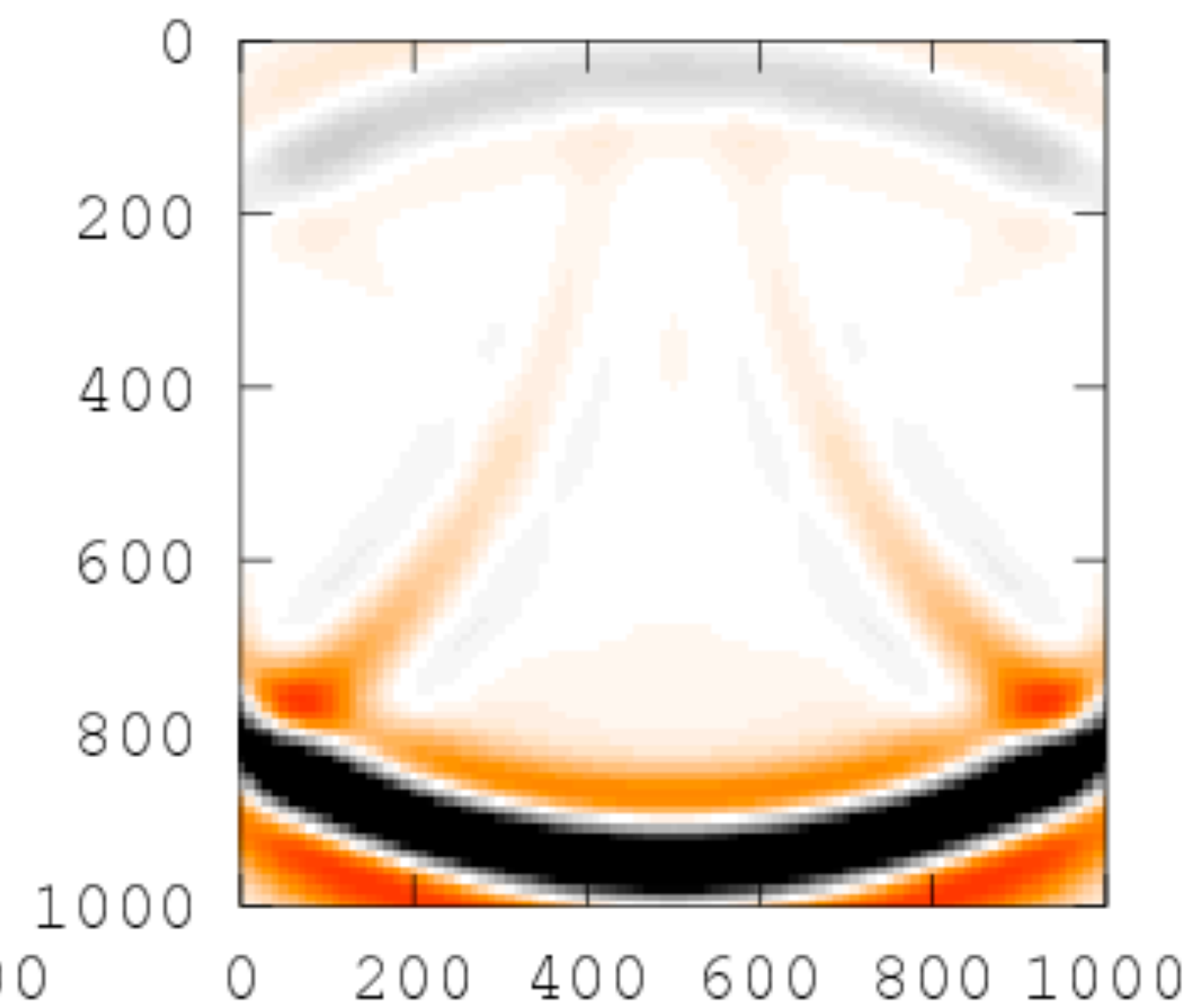
wavefield in *true* model



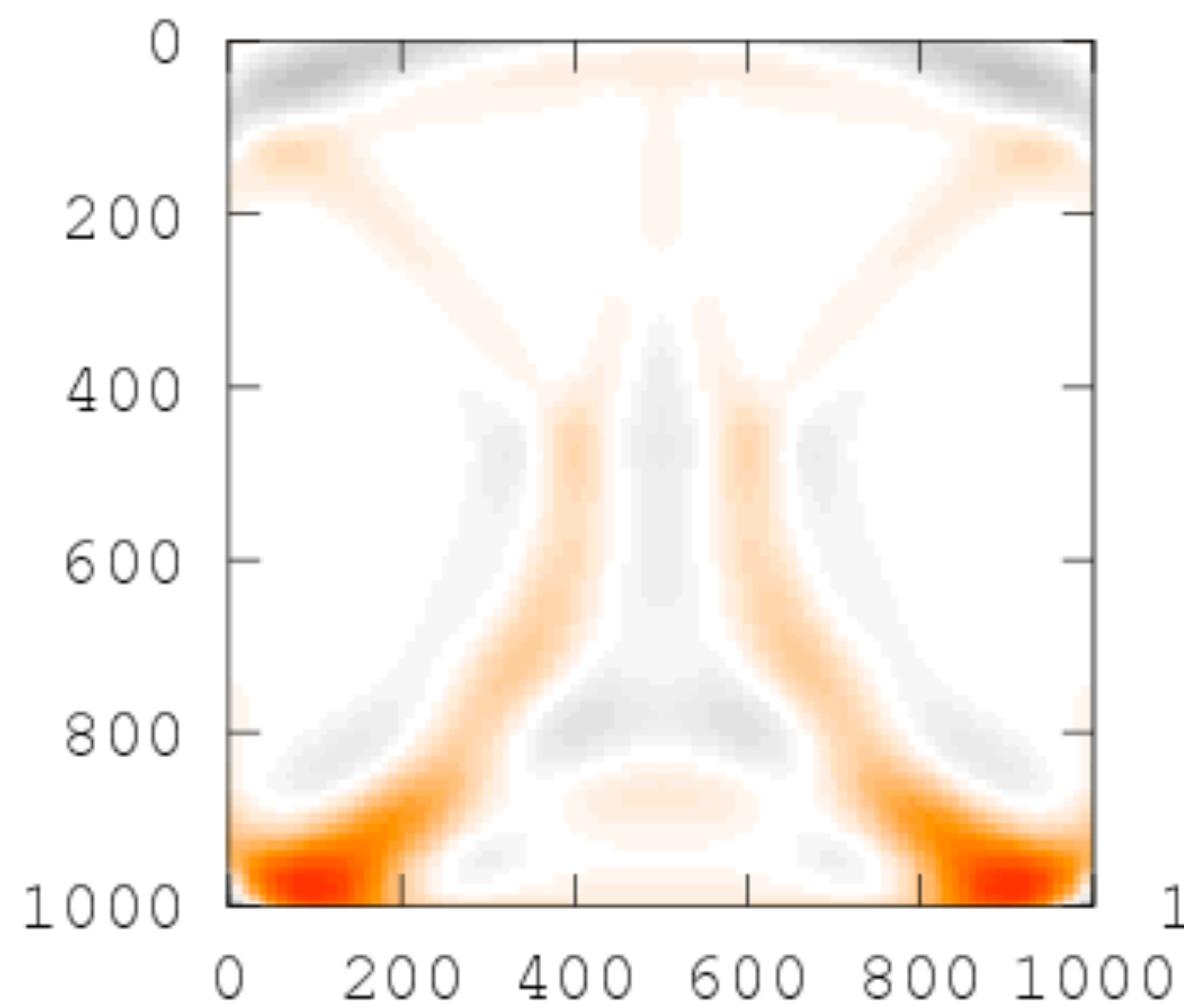
wavefield in *constant* model



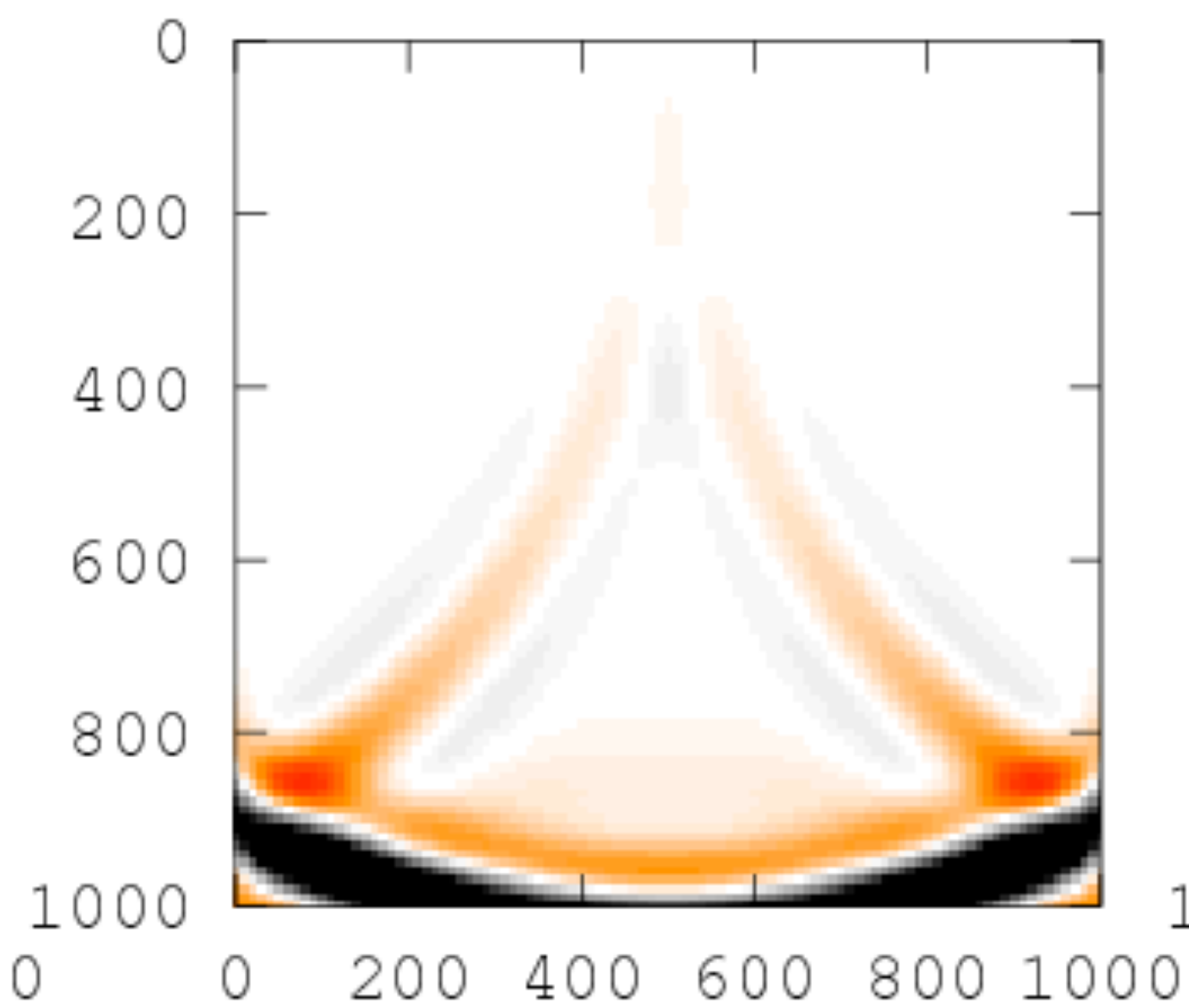
**data-augmented
wavefield in *constant* model**



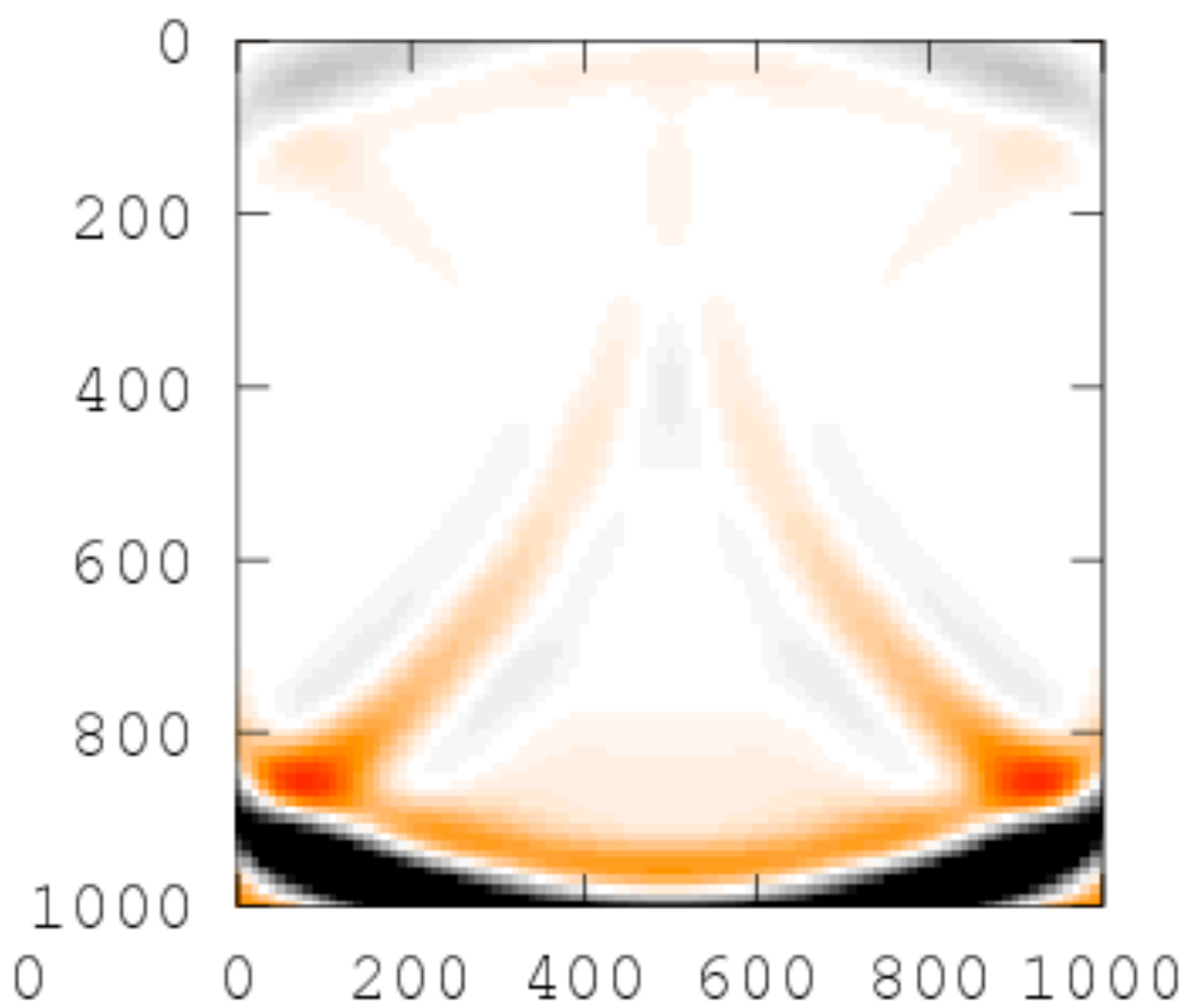
wavefield in *true* model



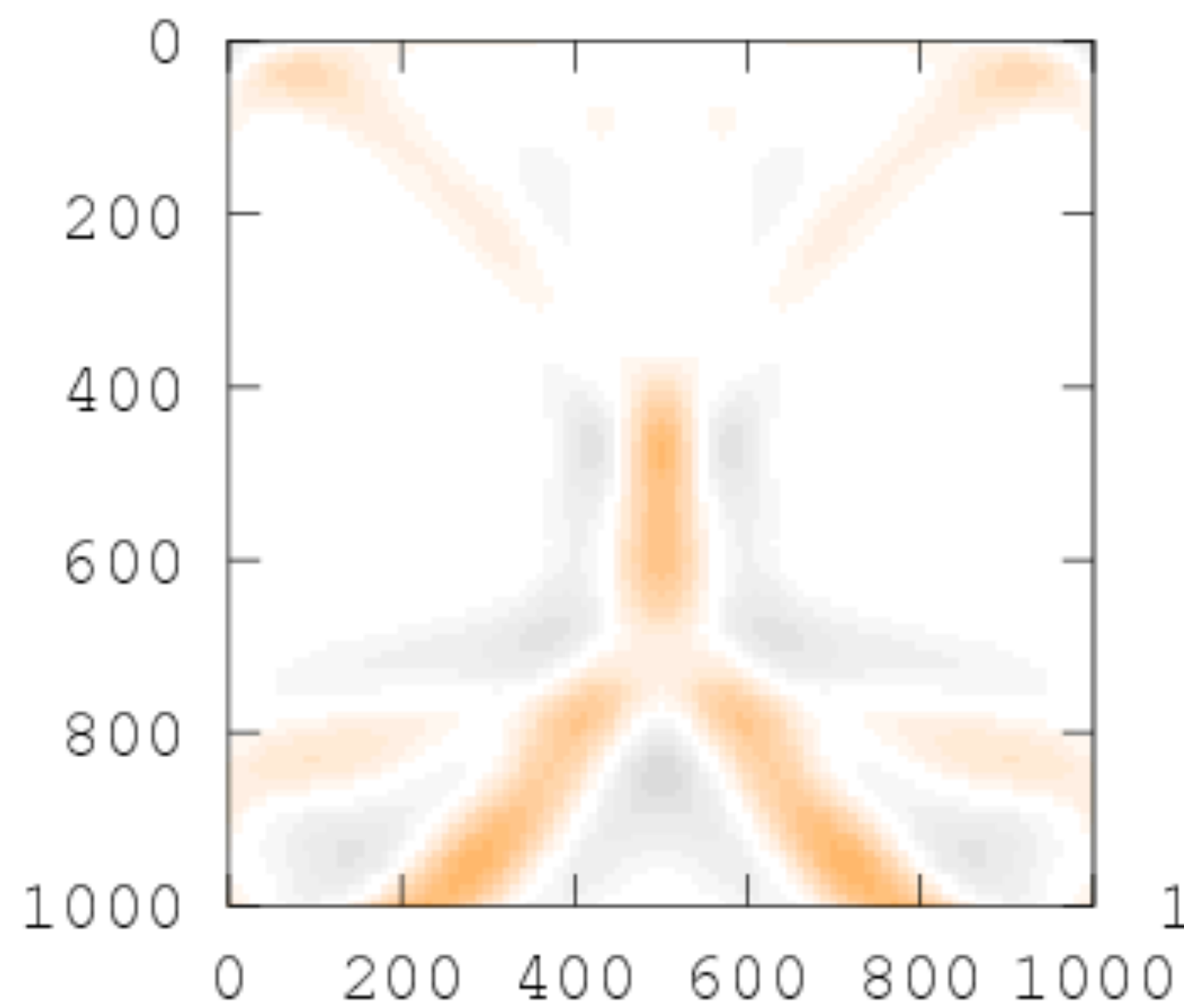
wavefield in *constant* model



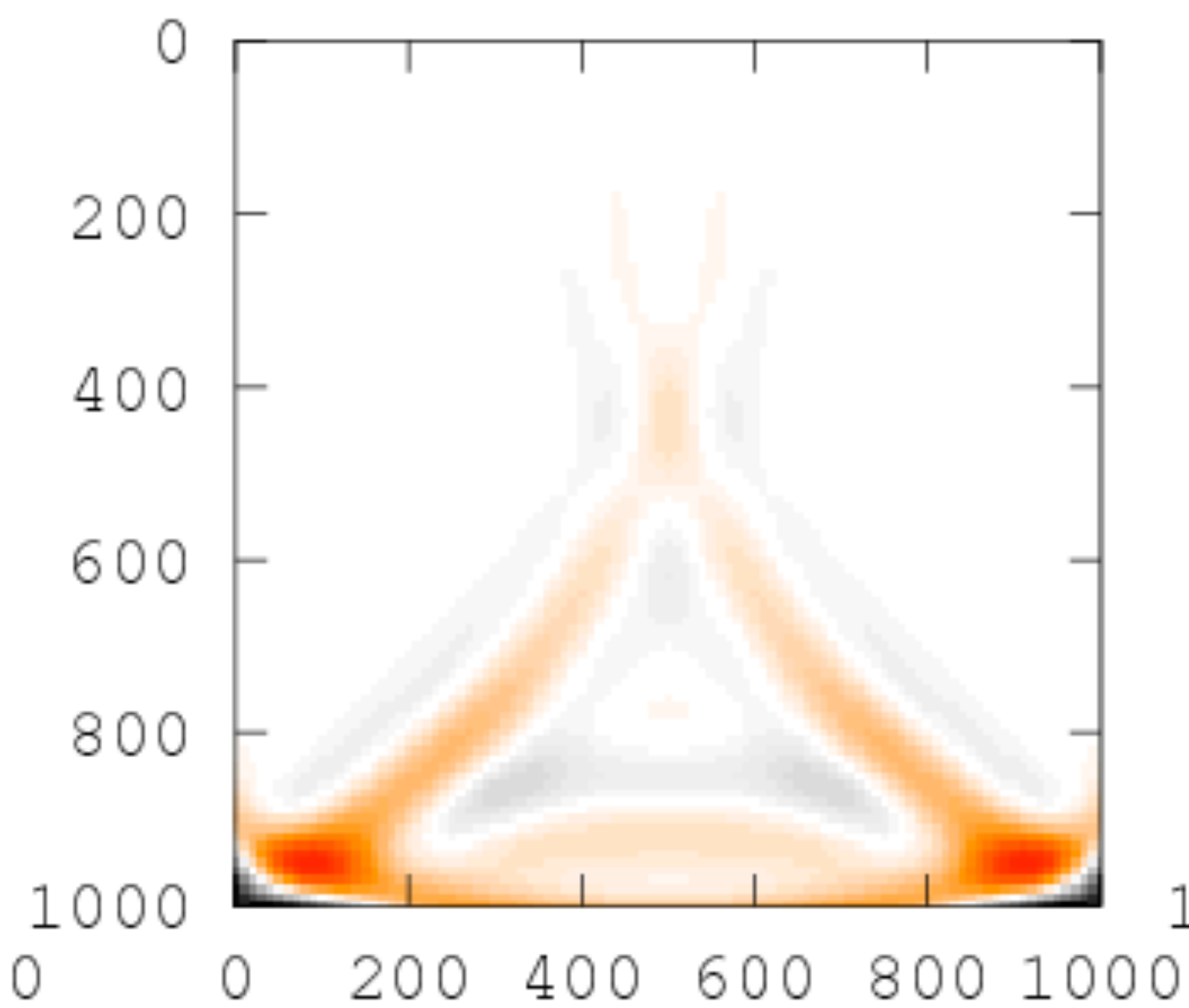
**data-augmented
wavefield in *constant* model**



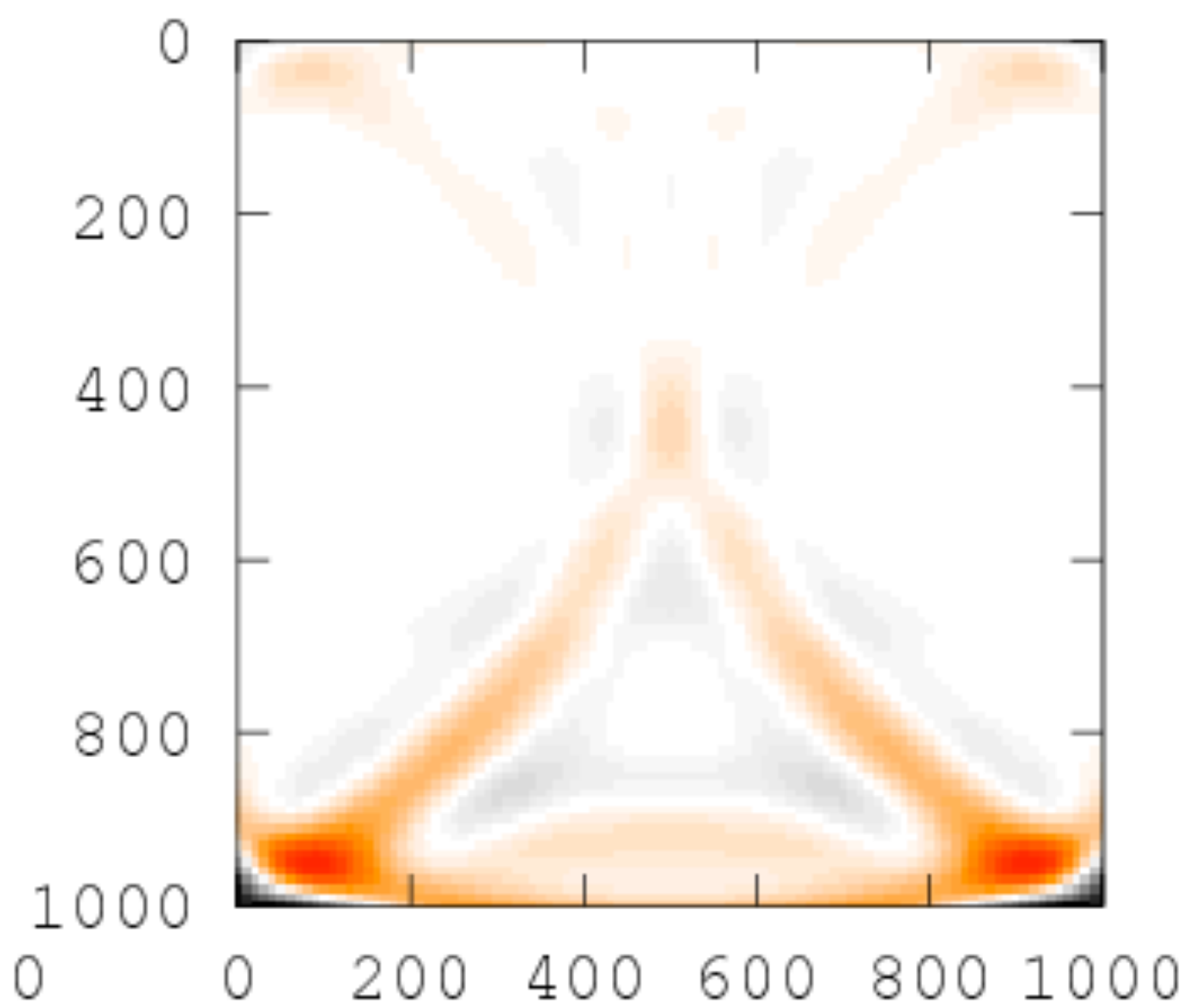
wavefield in *true* model



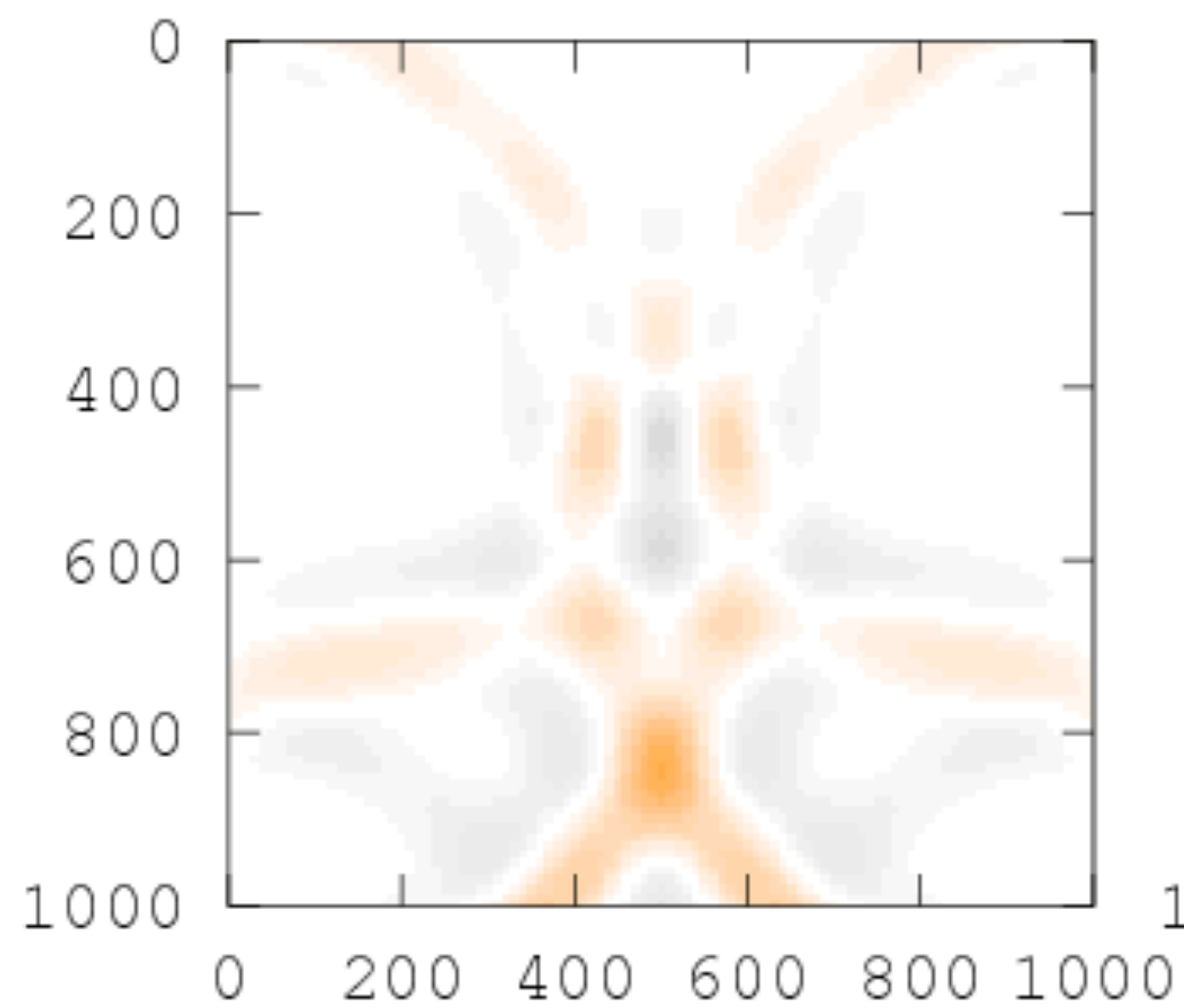
wavefield in *constant* model



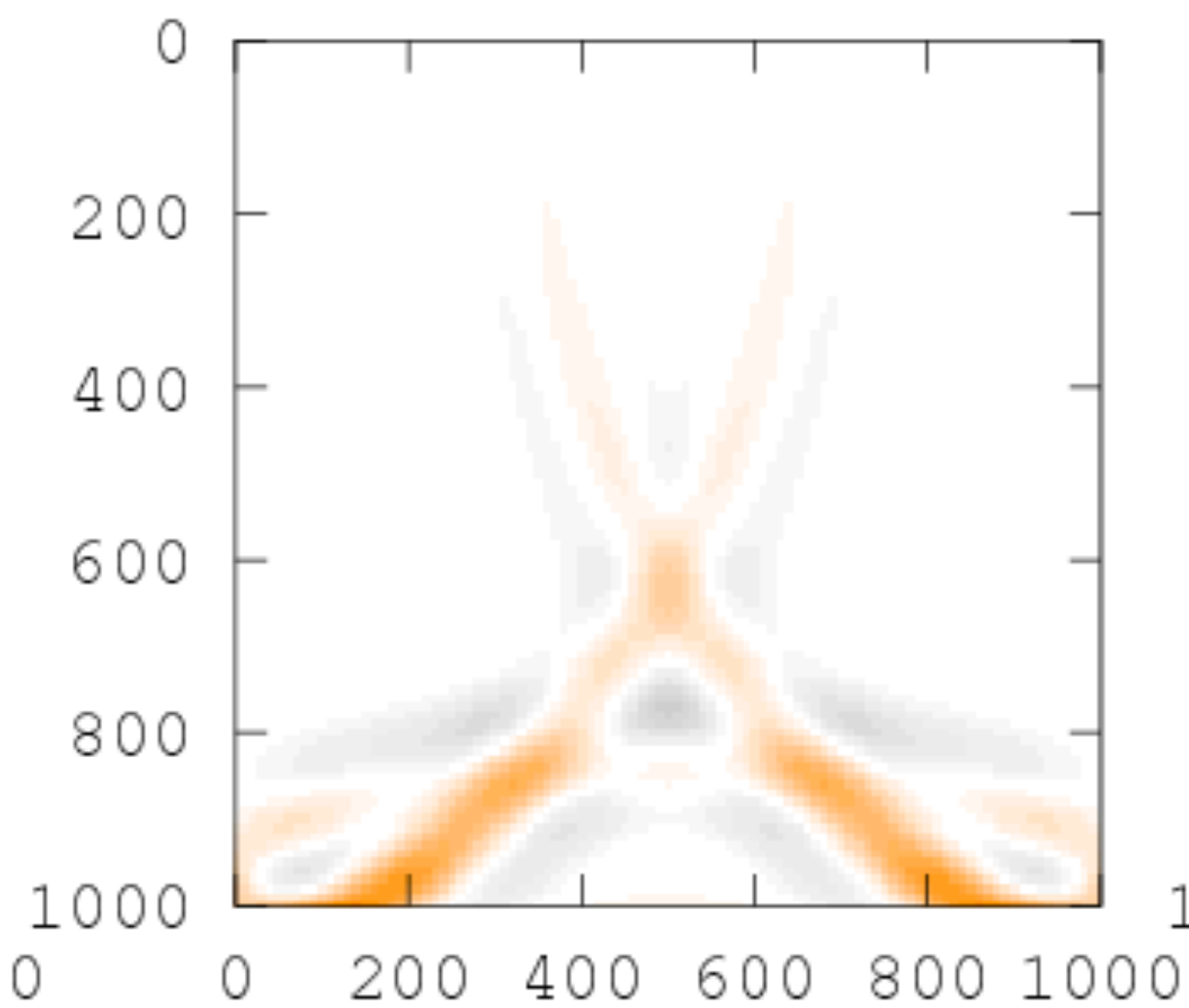
**data-augmented
wavefield in *constant* model**



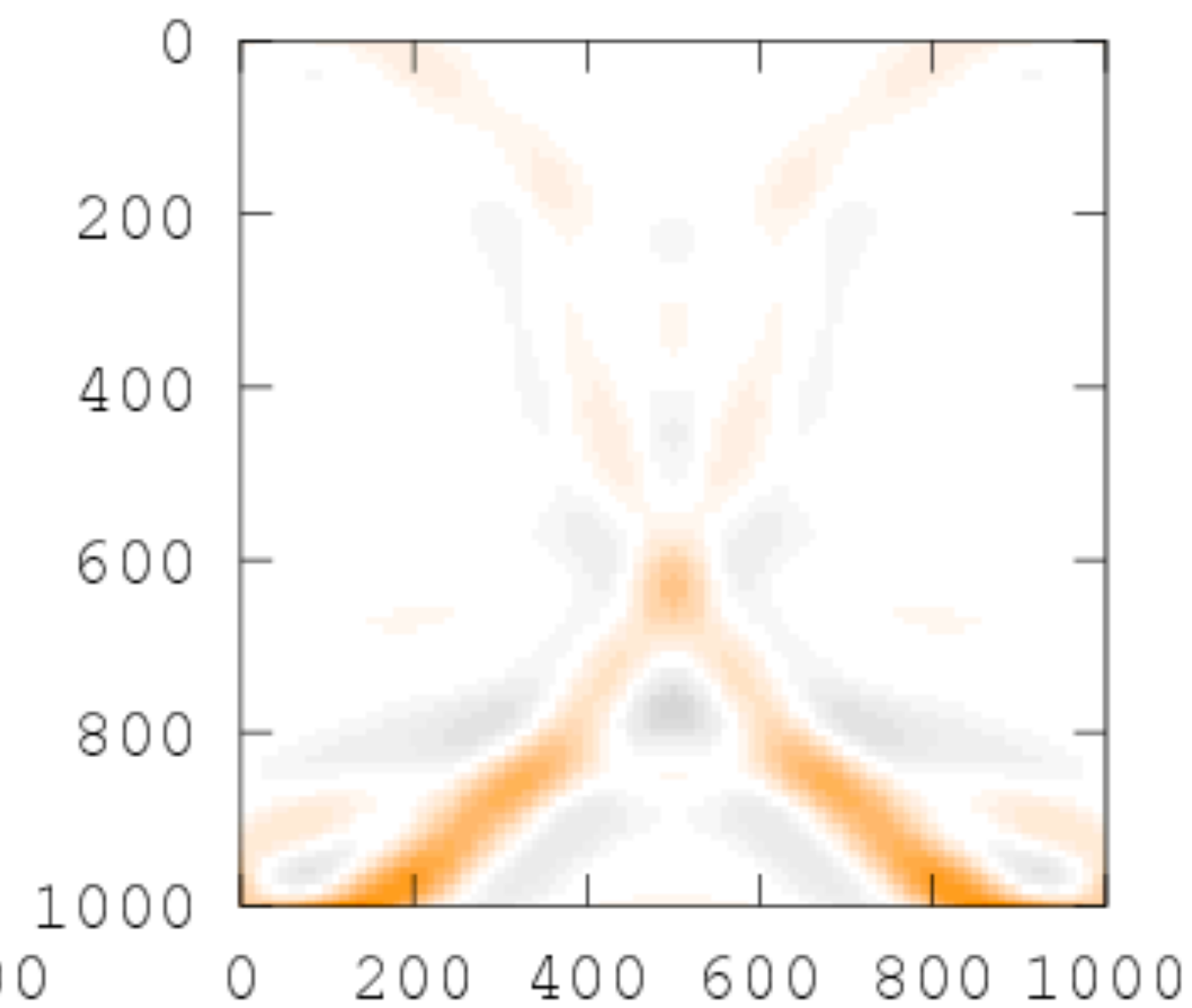
wavefield in *true* model



wavefield in *constant* model



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



[Heinkenschloss, '98 , Haber, '00]

PDE-constrained optimization

all-at-once full-space approach

$$\begin{array}{ccc}
 \text{simulated data} & & \text{simulated wavefield} \\
 \downarrow & & \downarrow \\
 \min_{\mathbf{m}, \mathbf{u}} \sum_{i=1}^M \|P_i \mathbf{u}_i - \mathbf{d}_i\|_2^2 & \text{s.t.} & A_i(\mathbf{m}) \mathbf{u}_i = \mathbf{q}_i \\
 \uparrow & & \uparrow \\
 \text{observed data} & & \text{source} \\
 & \text{Helmholtz equation} &
 \end{array}$$

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

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 (\mathbf{m})
- ▶ *dense* (GN) Hessian, involves *additional* PDE solves

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 $\lambda \uparrow \infty$

Variable projection

Solve *data-augmented* wave equation for each source

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

Define *reduced* objective

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

Gradient

Wavefield eliminated—i.e., $\nabla_{\bar{\mathbf{u}}} \phi_{\lambda}(\mathbf{m}, \bar{\mathbf{u}}) = 0$ by solving

$$(\lambda^2 A^*(\mathbf{m})A(\mathbf{m}) + P^*P) \mathbf{u} = \lambda^2 A^*(\mathbf{m})\mathbf{q} + P^*\mathbf{d}$$

yielding the gradient

Jacobian of $A(\mathbf{m})\bar{\mathbf{u}}_{\lambda}$

$$\nabla \phi_{\lambda}(\mathbf{m}) = G(\mathbf{m}, \bar{\mathbf{u}}_{\lambda})^* \bar{\mathbf{v}}_{\lambda}$$

with

$$\bar{\mathbf{v}}_{\lambda} = \lambda^2 (A(\mathbf{m})\bar{\mathbf{u}}_{\lambda} - \mathbf{q}) \quad (\text{PDE residual})$$

Wavefield Reconstruction Inversion

WRI method

for each source i

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

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

end

Conventional method

for each source i

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

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

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

end

Wavefield Reconstruction Inversion

Penalty method

for each source i

$$\text{solve } \begin{pmatrix} P \\ \lambda A(\mathbf{m}) \end{pmatrix} \mathbf{u} \approx \begin{pmatrix} \mathbf{d}_i \\ \lambda \mathbf{q}_i \end{pmatrix}$$

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

end

correlation
augmented
wavefield &
PDE residual

Conventional method

for each source i

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

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

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

end

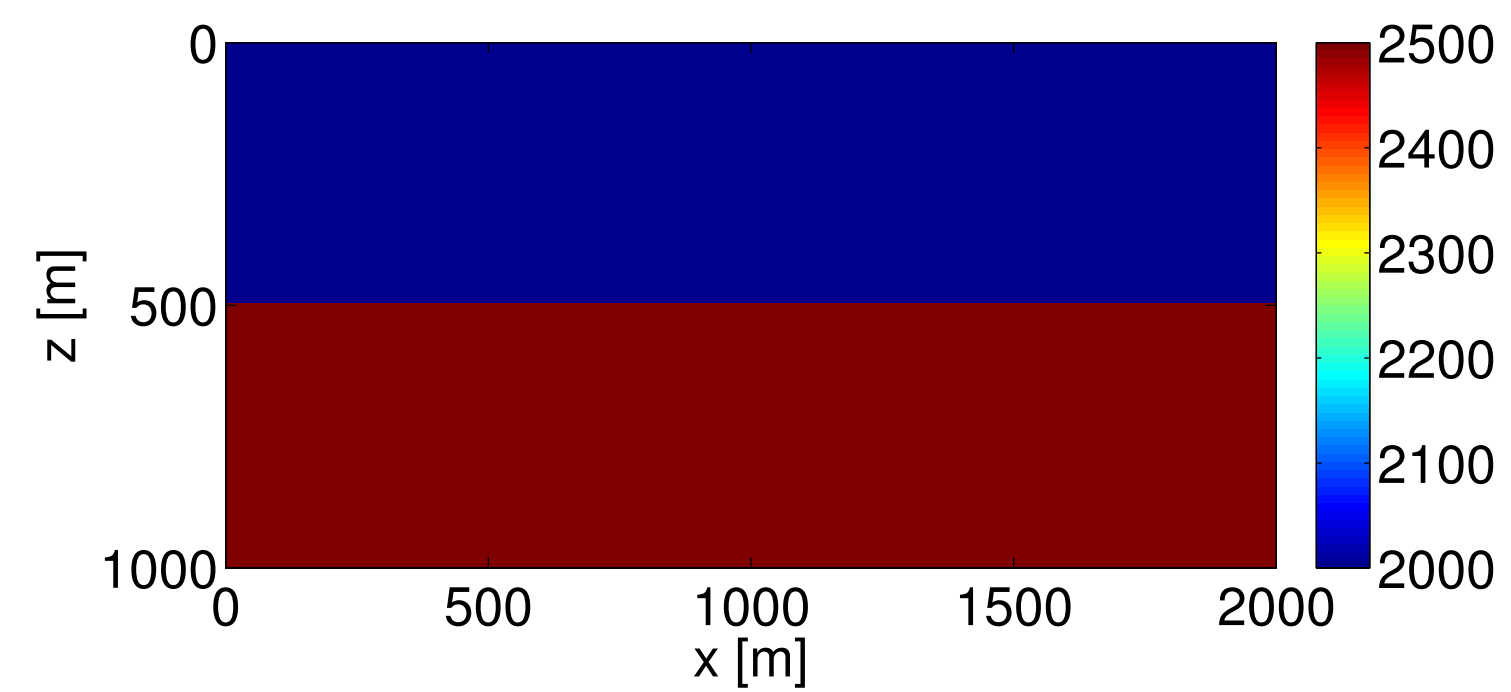
correlation
wavefield & *data*
residual

Wavefield *Reconstruction Inversion*

- ▶ *no* need to store *all* the fields (\mathbf{u})
- ▶ *no adjoint* solves
- ▶ *sparse* approximation of GN Hessian for small
- ▶ *less non-linear* in \mathbf{m}
- ▶ need to solve *overdetermined* PDE
- ▶ *not clear* how to *pick* λ
- ▶ ...

One reflector example

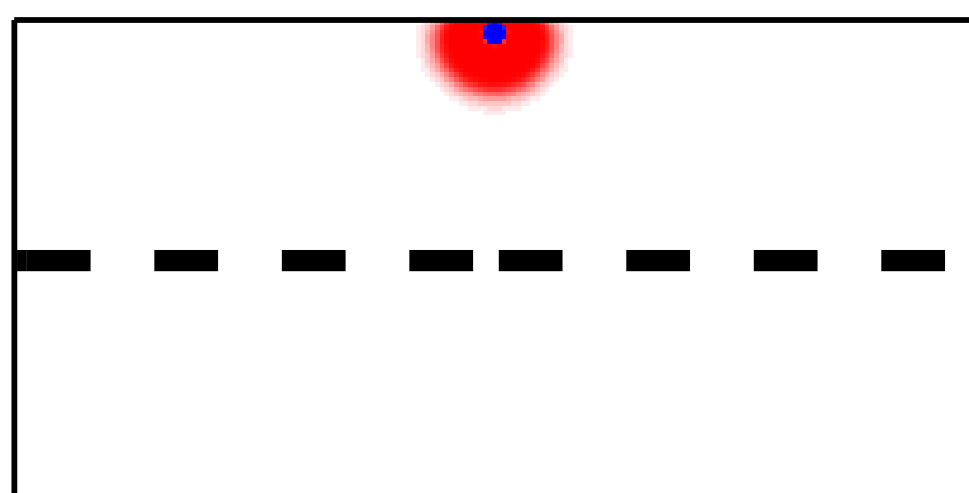
true model



Wavefields in *homogeneous* background

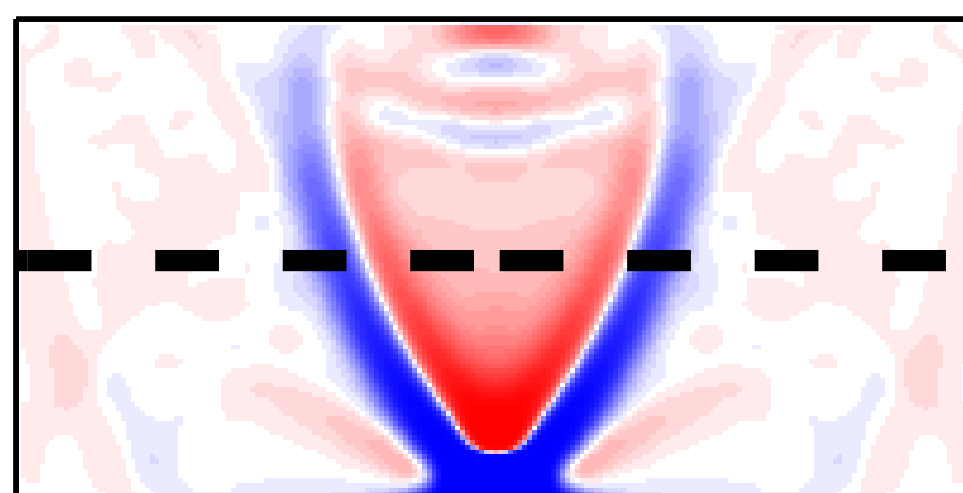
FWI

forward



\bar{u}

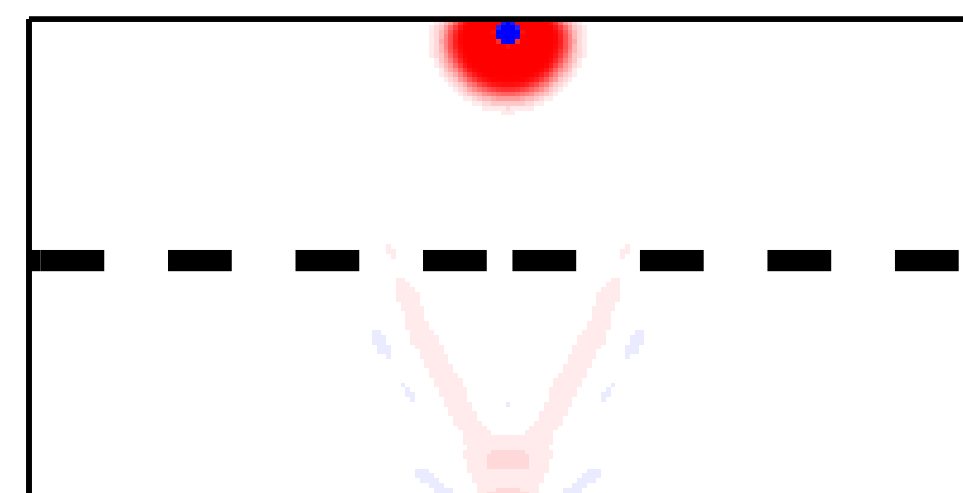
adjoint



\bar{v}

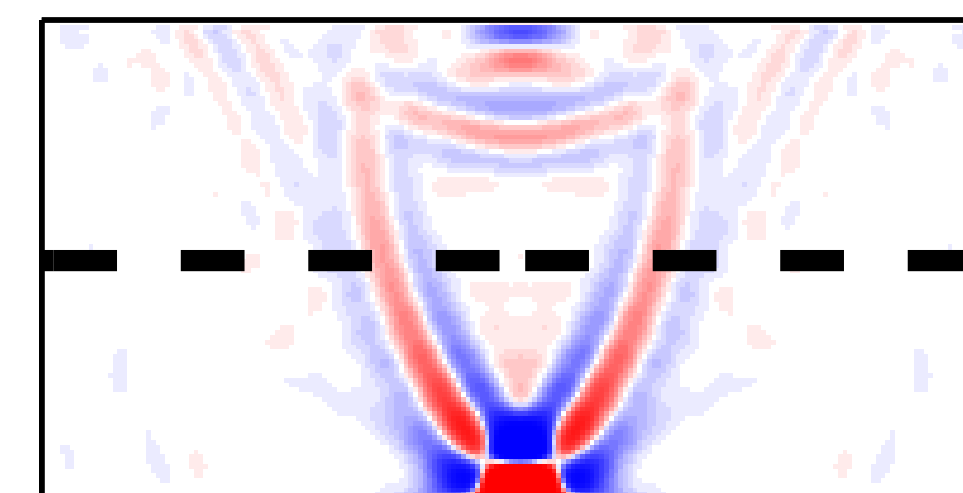
WRI

reconstructed wavefield



\bar{u}_λ

PDE residual

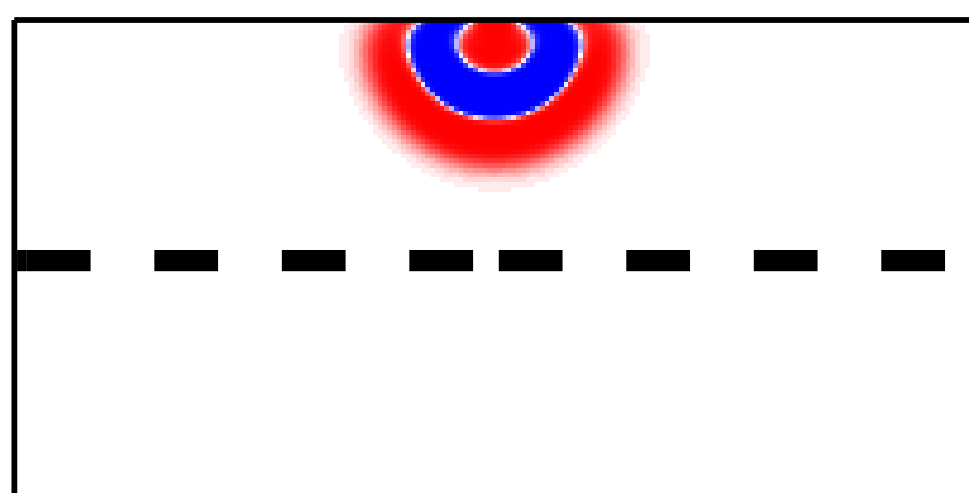


\bar{v}_λ

Wavefields in *homogeneous* background

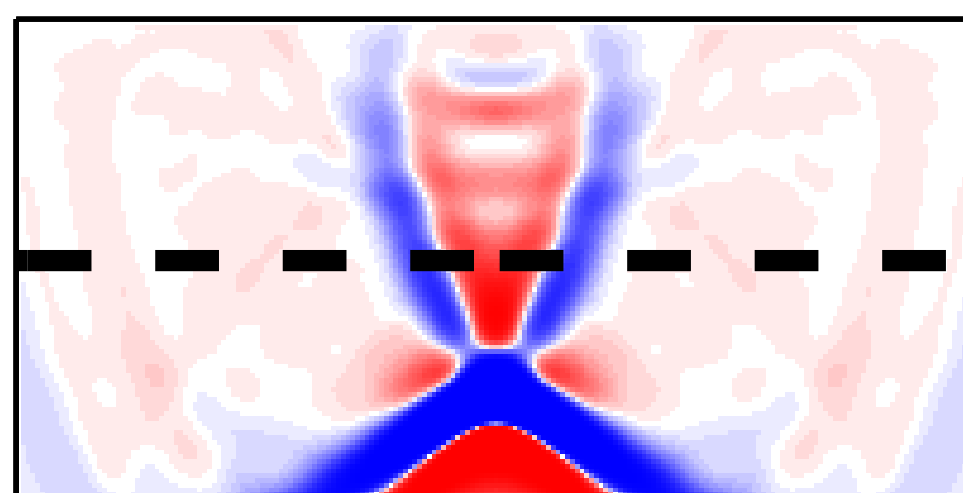
FWI

forward



\bar{u}

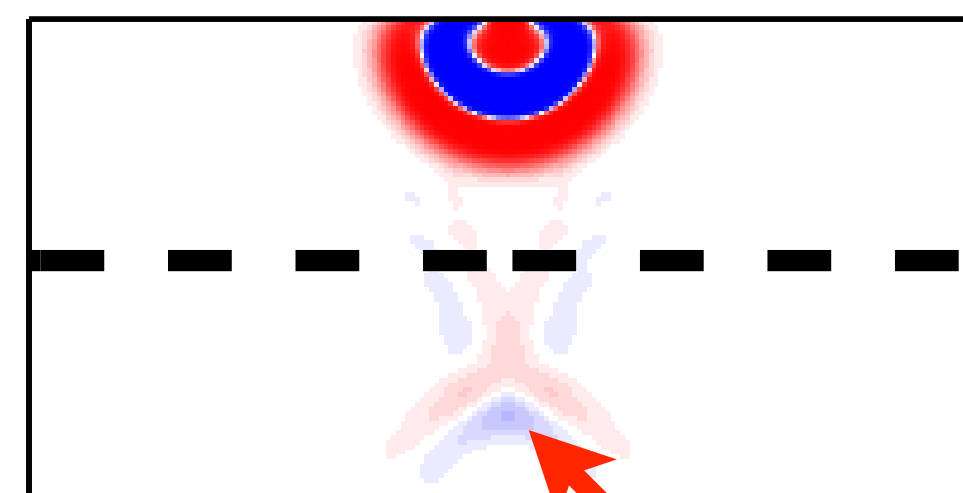
adjoint



\bar{v}

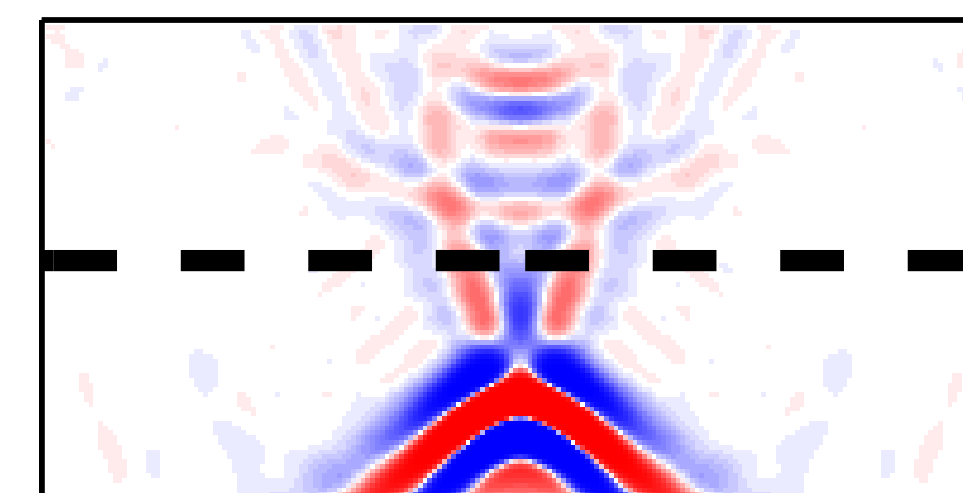
WRI

reconstructed wavefield



\bar{u}_λ

PDE residual

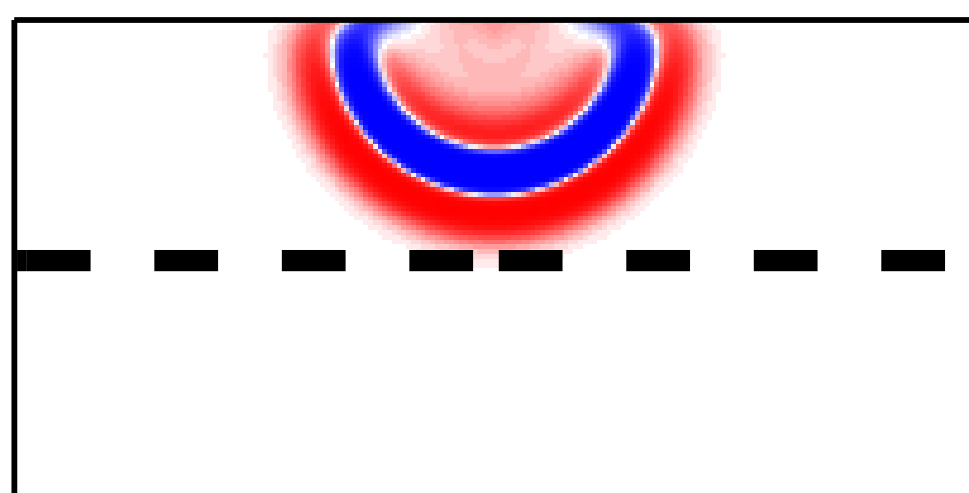


\bar{v}_λ

Wavefields in *homogeneous* background

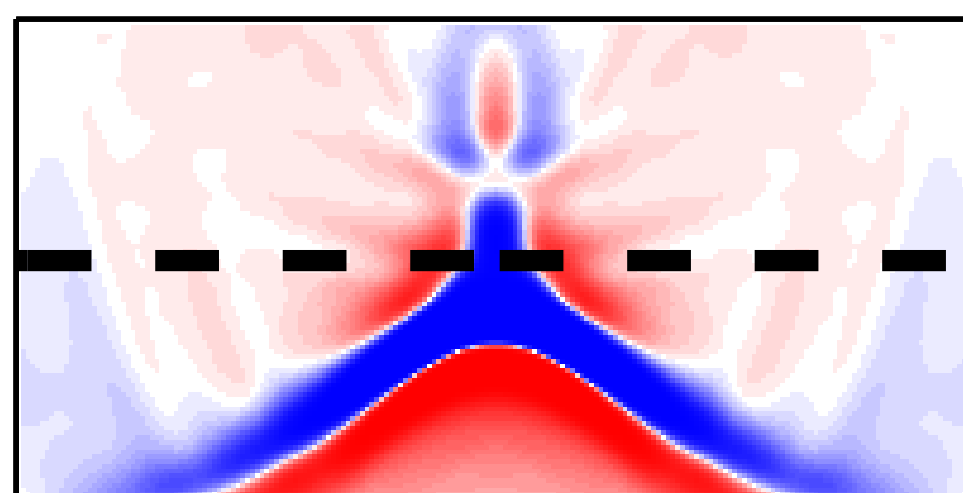
FWI

forward



\bar{u}

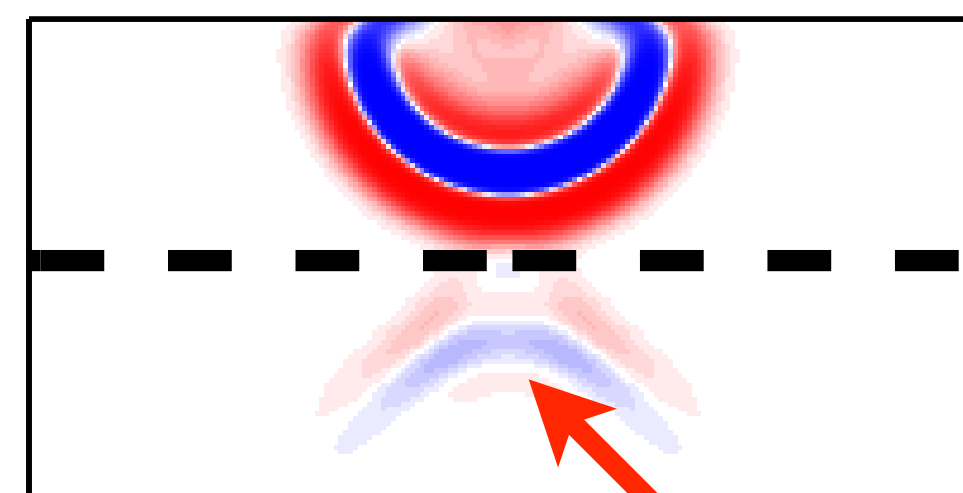
adjoint



\bar{v}

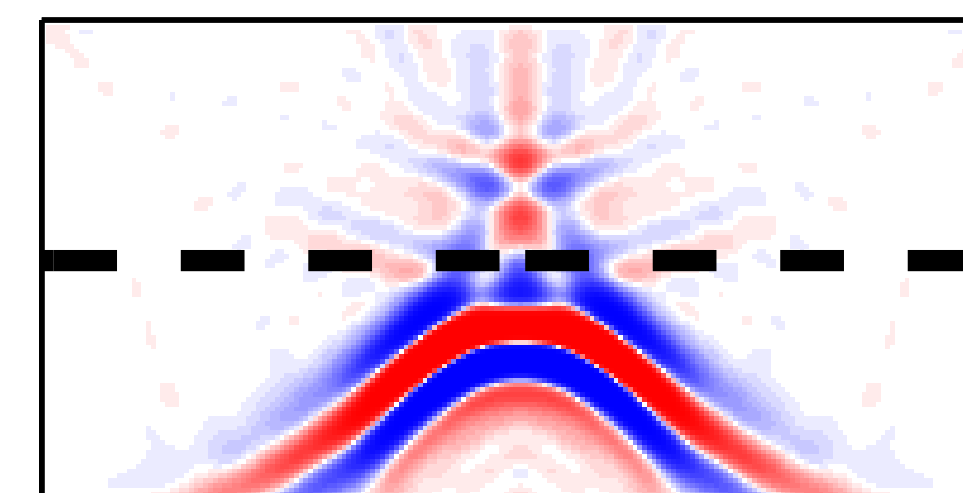
WRI

reconstructed wavefield



\bar{u}_λ

PDE residual

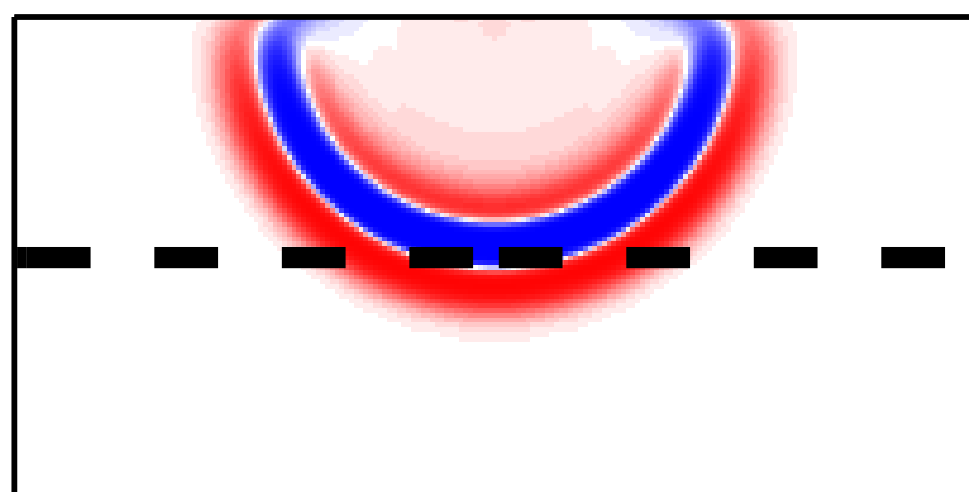


\bar{v}_λ

Wavefields in *homogeneous* background

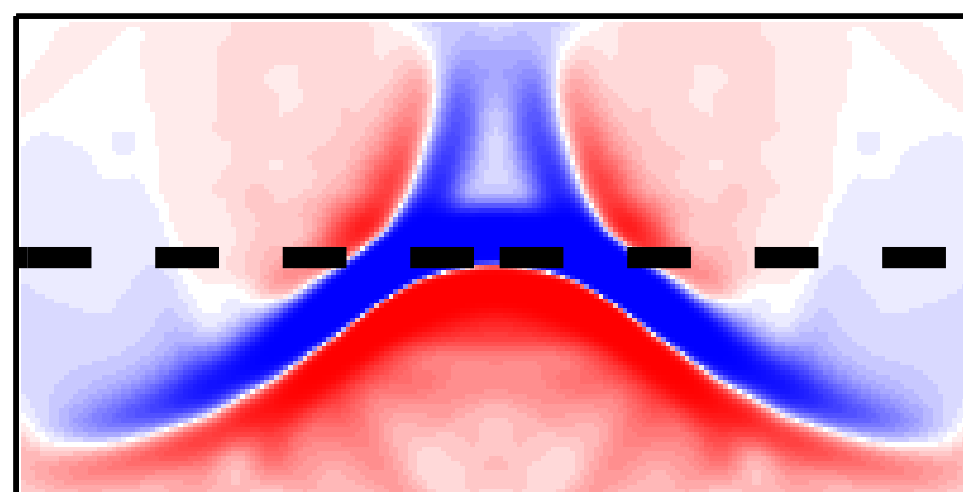
FWI

forward



\bar{u}

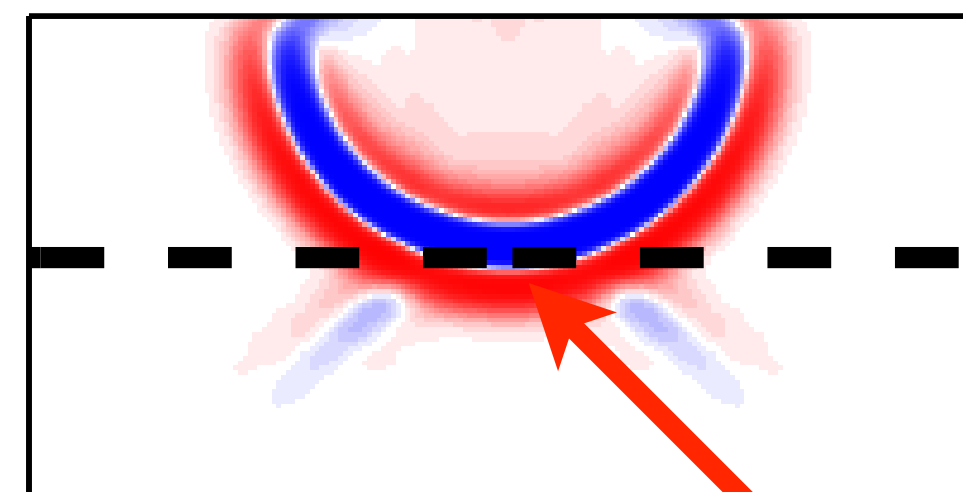
adjoint



\bar{v}

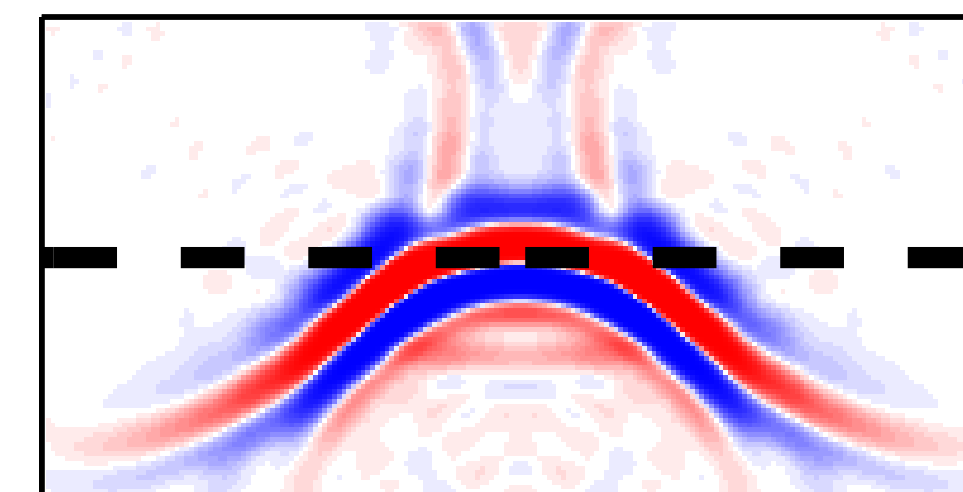
WRI

reconstructed wavefield



\bar{u}_λ

PDE residual

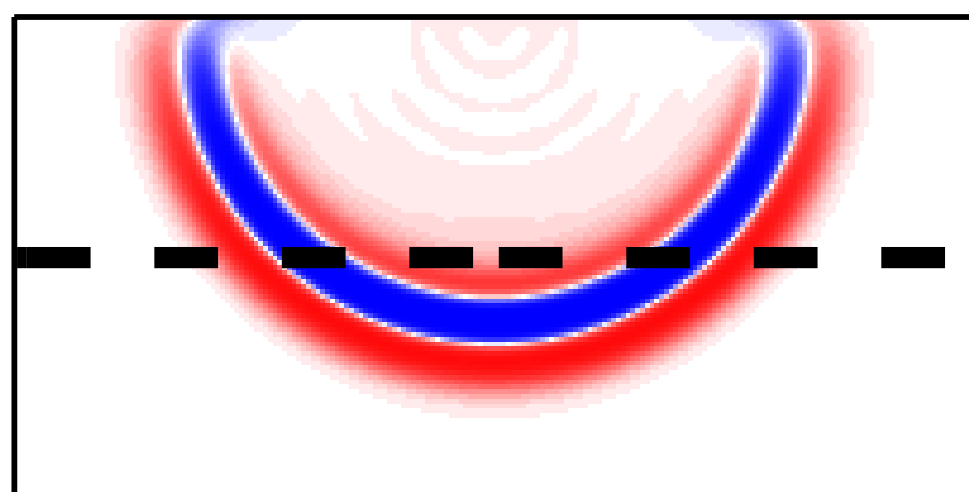


\bar{v}_λ

Wavefields in *homogeneous* background

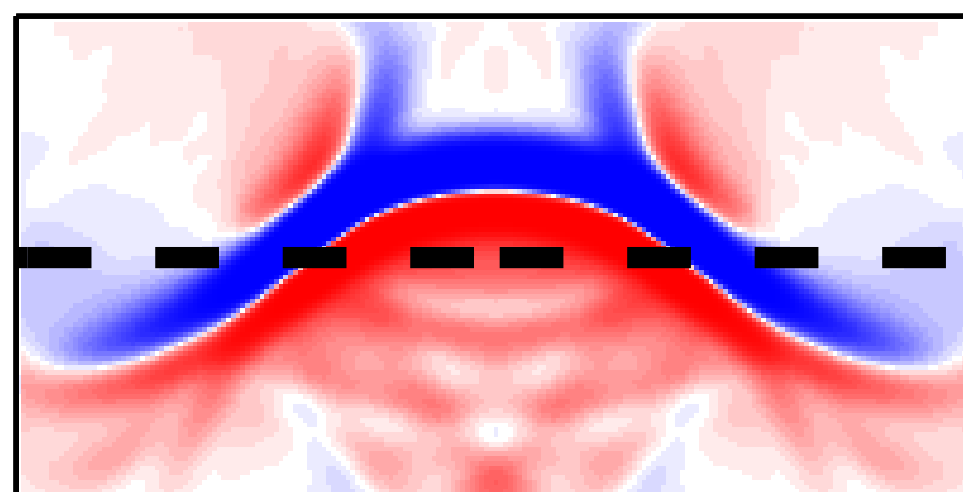
FWI

forward



\bar{u}

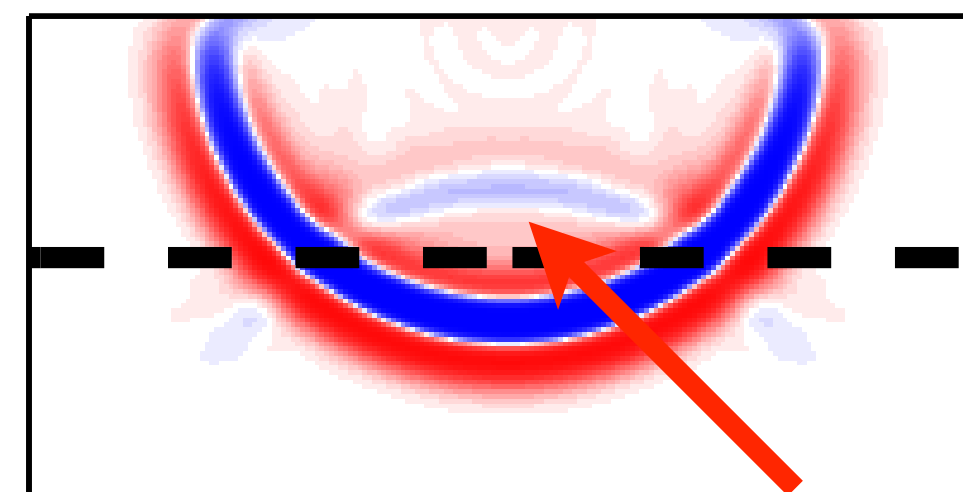
adjoint



\bar{v}

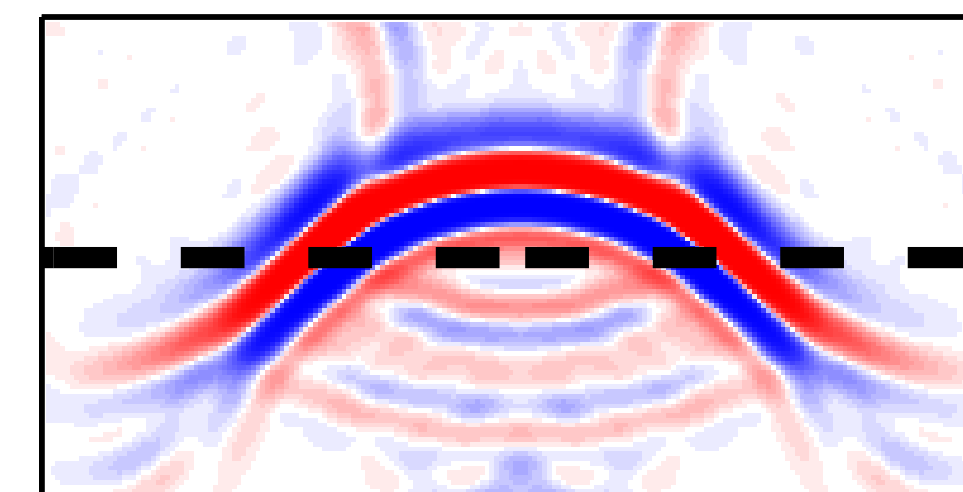
WRI

reconstructed wavefield



\bar{u}_λ

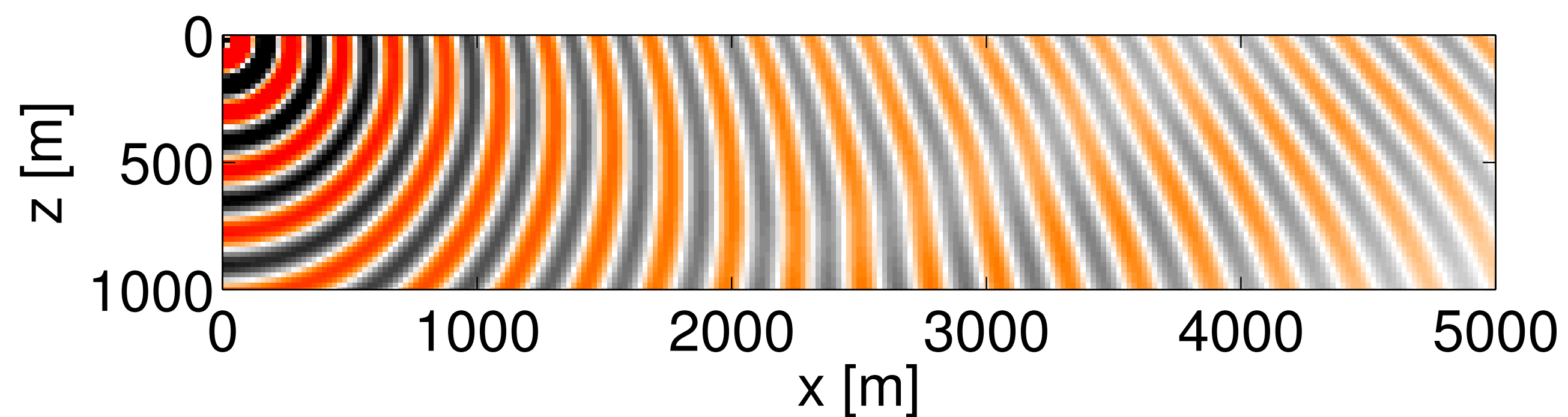
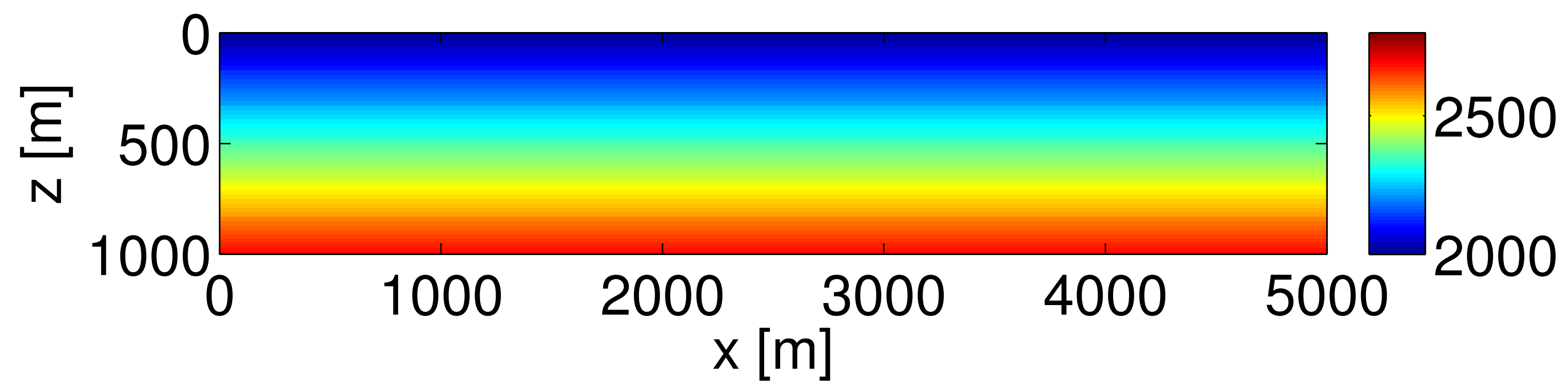
PDE residual



\bar{v}_λ

Diving wave example

true model and wavefield

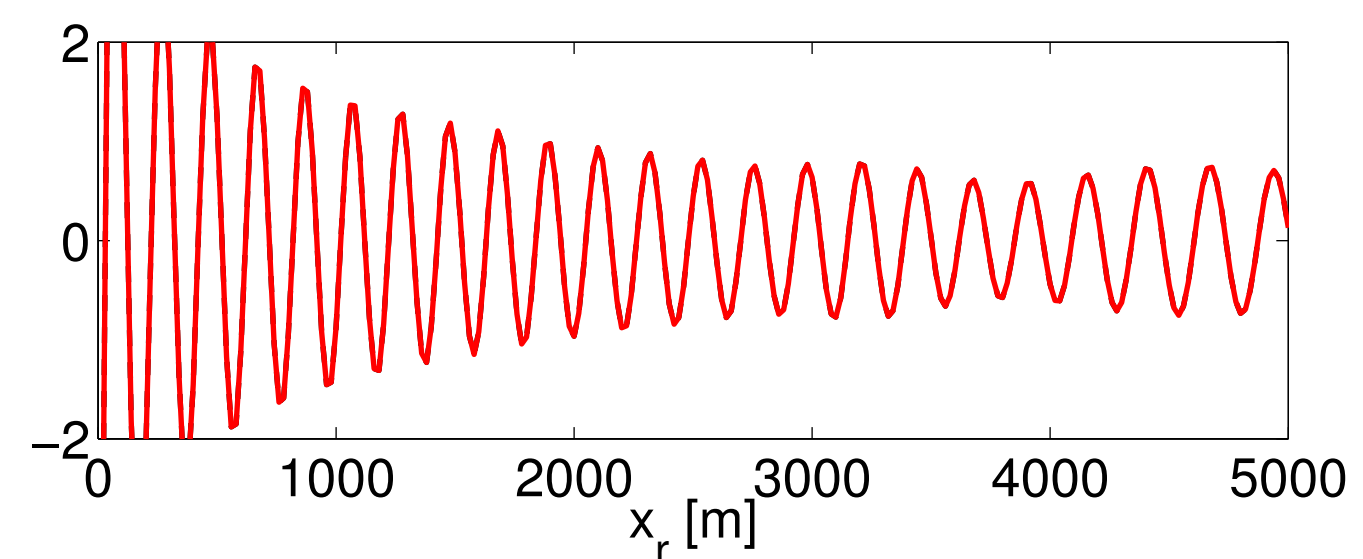
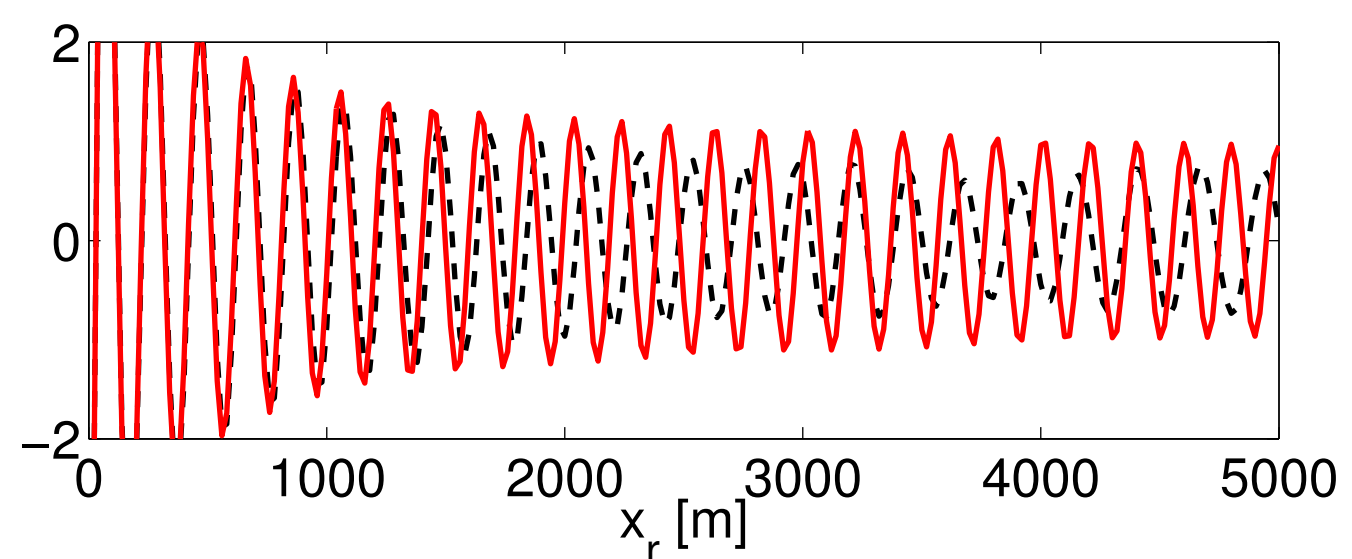


Wavefields in *homogeneous* background

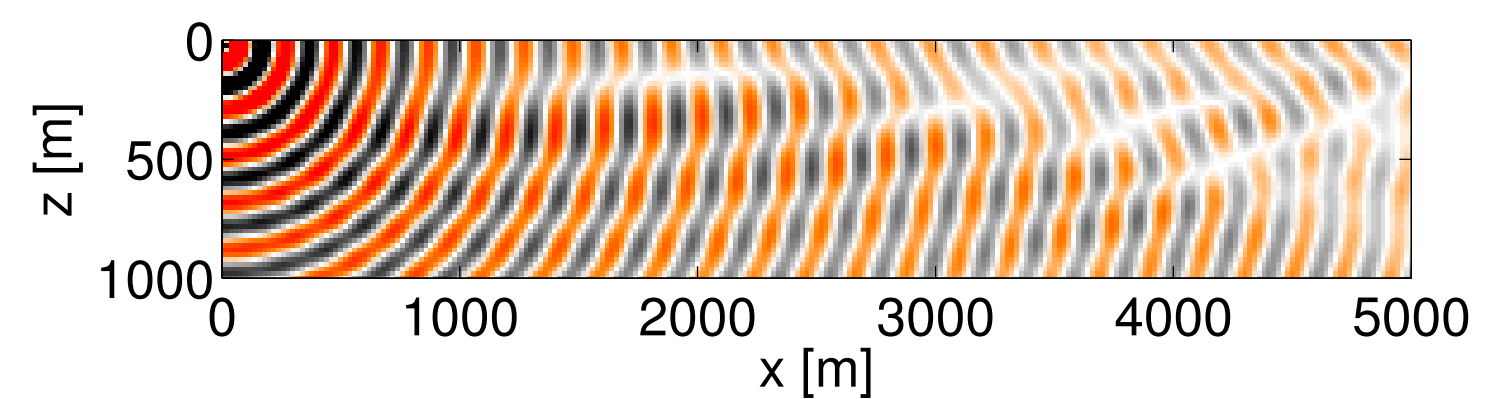
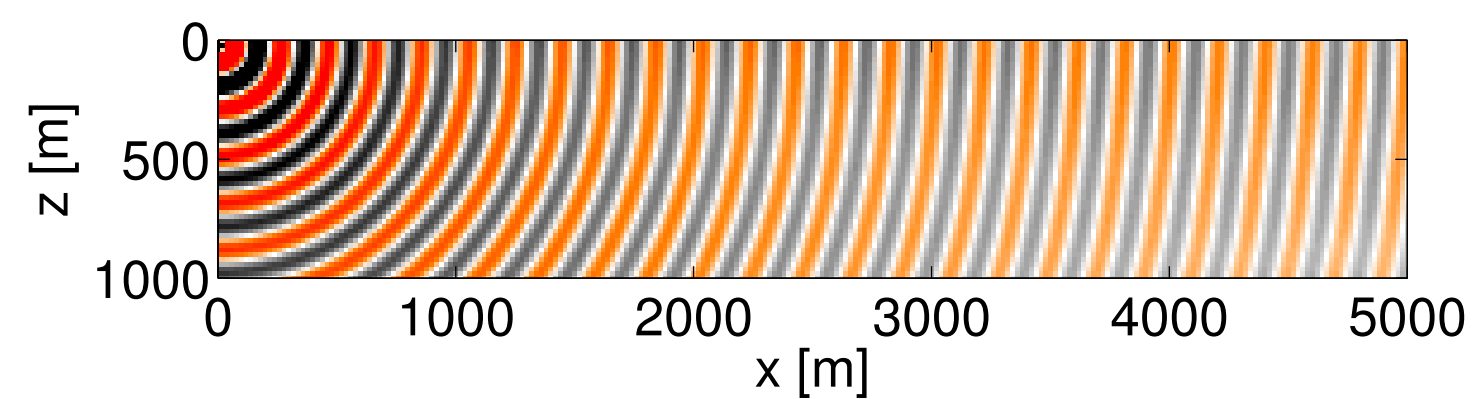
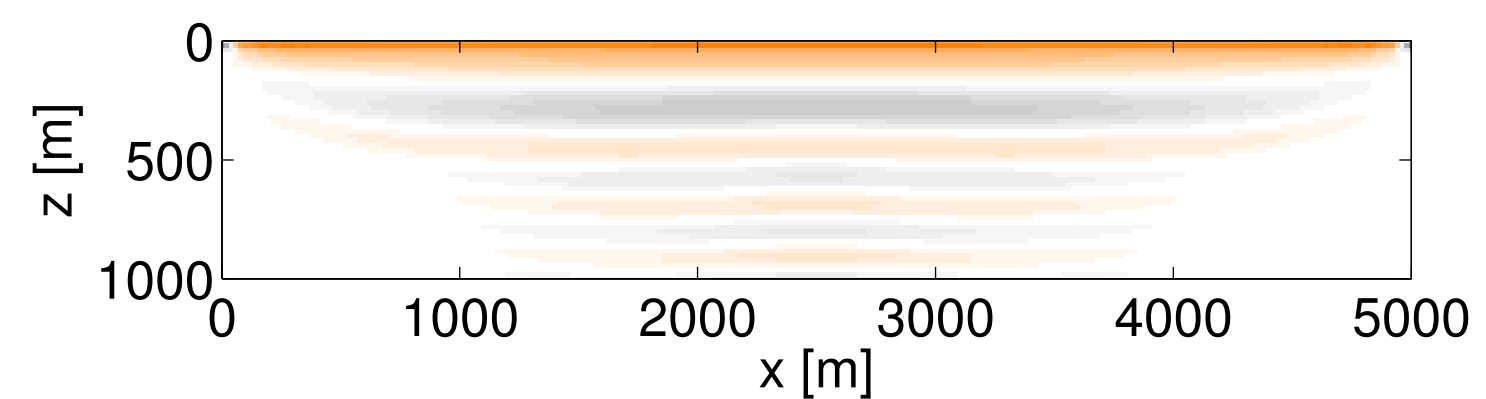
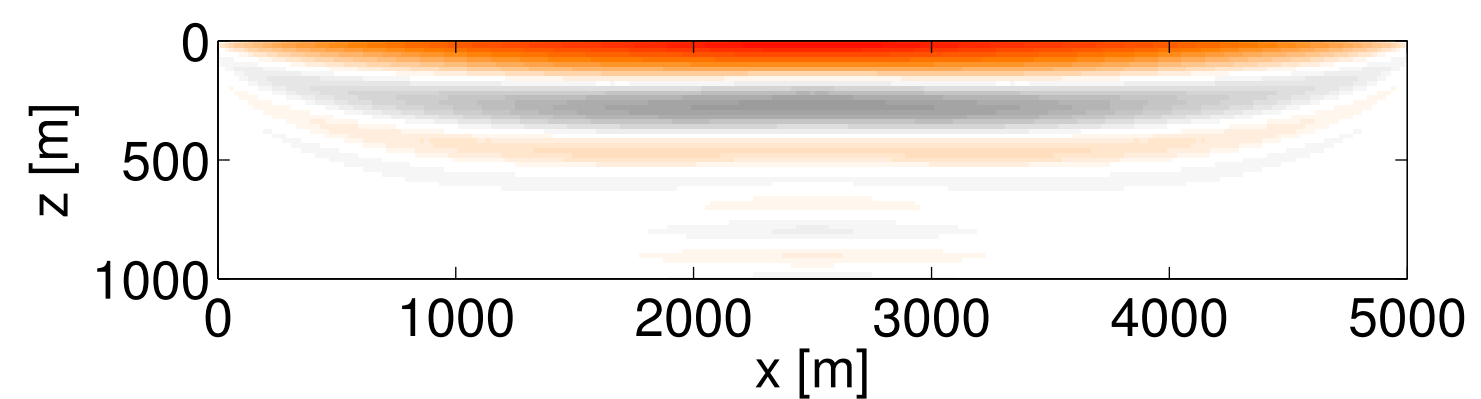
FWI

WRI

data



wavefield

model
update

Stylized examples

WRI vs. FWI

Penalty method

for each source i

$$\text{solve} \begin{pmatrix} P \\ \lambda A(\mathbf{m}) \end{pmatrix} \mathbf{u} \approx \begin{pmatrix} \mathbf{d}_i \\ \lambda \mathbf{q}_i \end{pmatrix}$$

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

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

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

end

Conventional method

for each source i

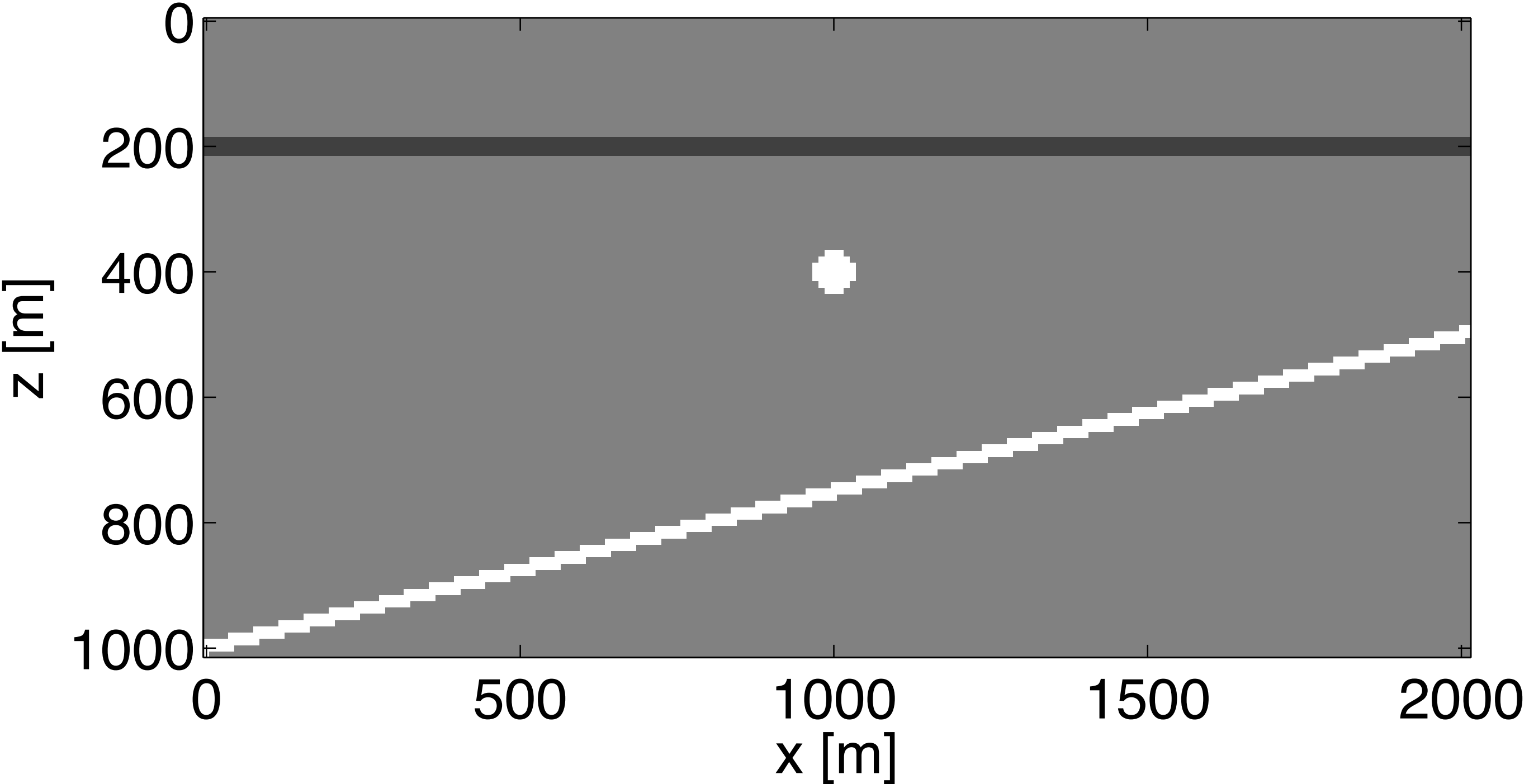
$$\text{solve } A(\mathbf{m})\mathbf{u}_i = \mathbf{q}_i$$

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

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

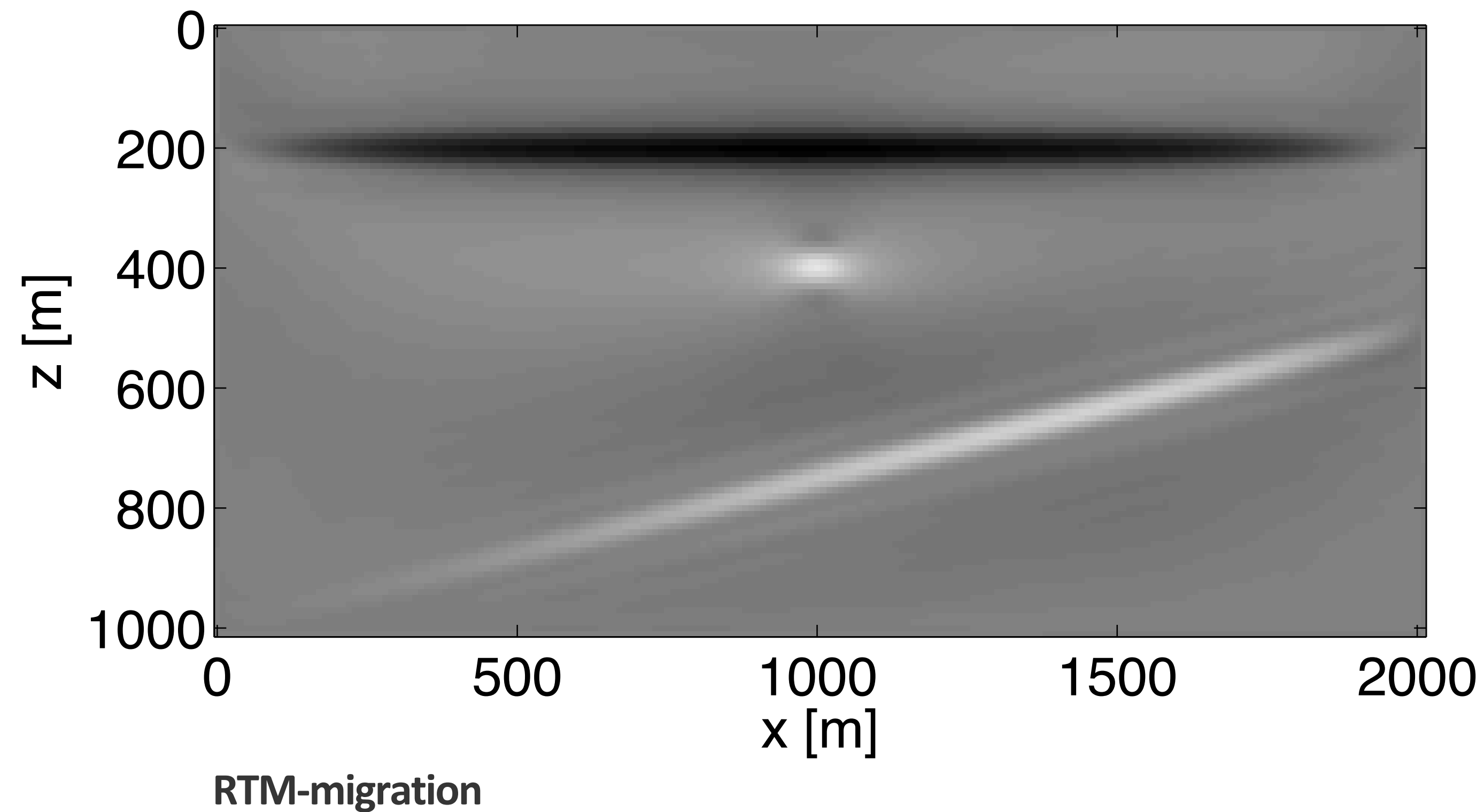
$$\mathbf{m} = \mathbf{m} - \alpha \mathbf{g}$$

end



Original image

Imaging
The *gradient* of the reduced objective yields an *image* of the subsurface..



Imaging

Conventional RTM:

- 1.solve forward wave-equation
- 2.solve adjoint wave-equation
- 3.apply `imaging condition`

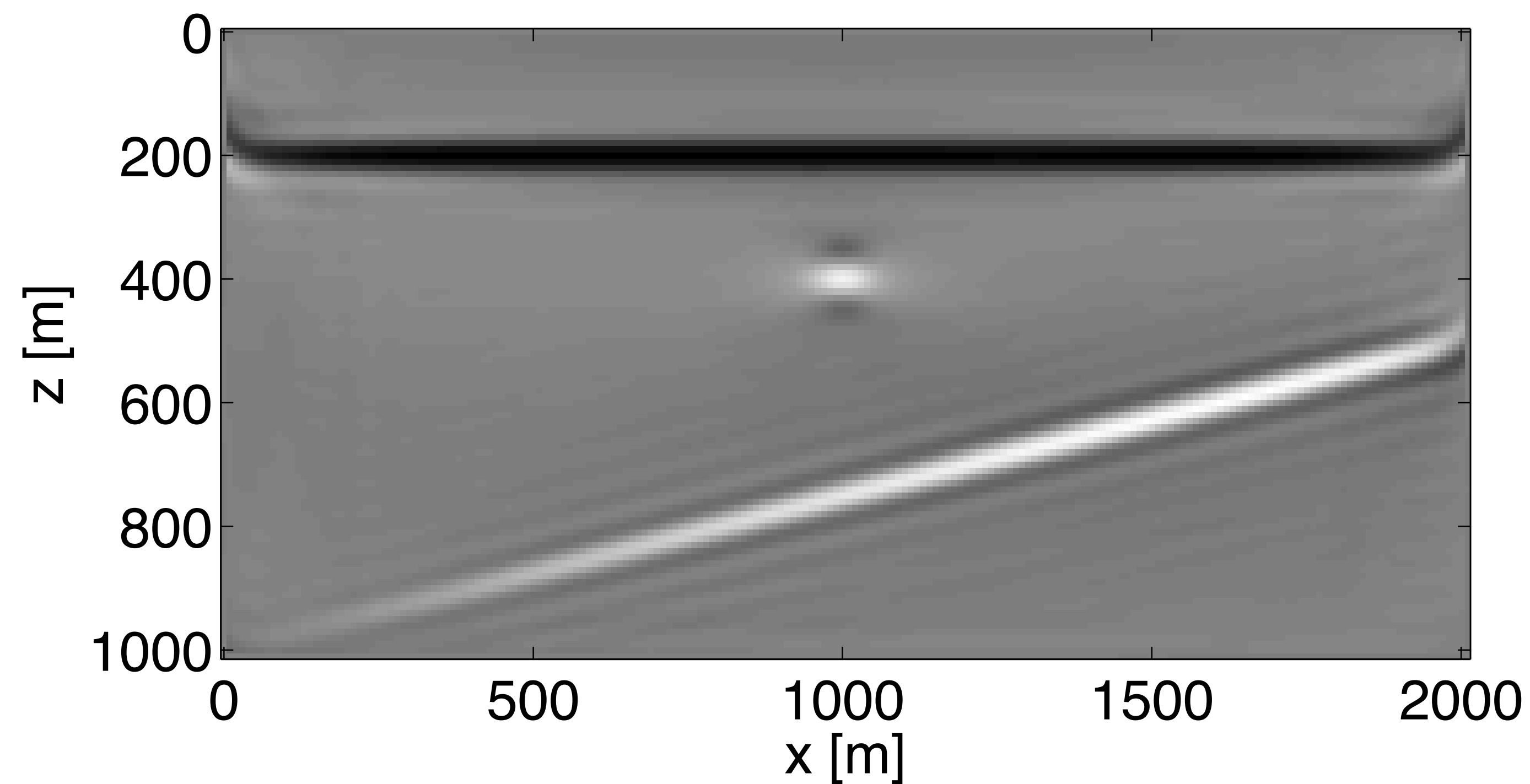


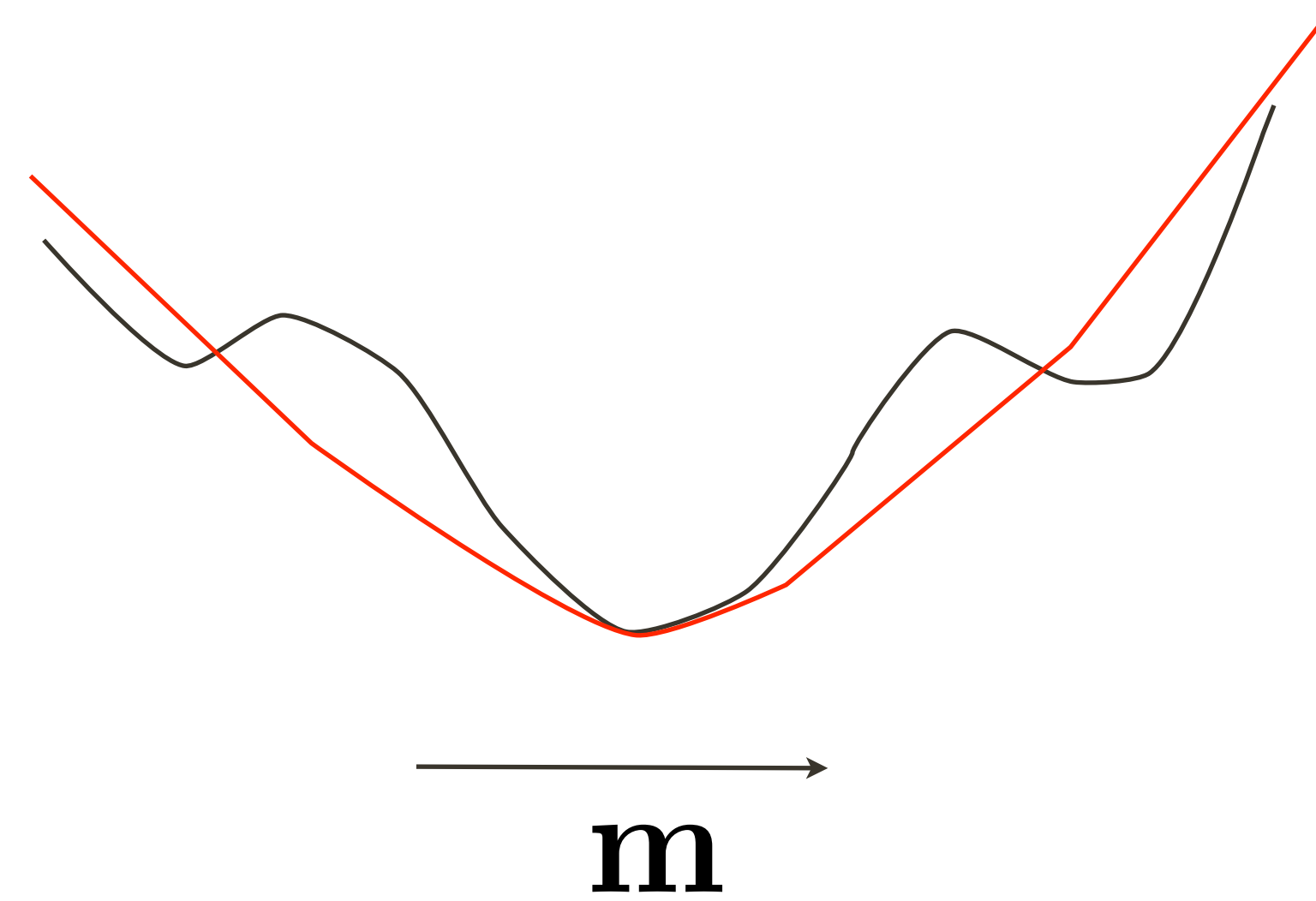
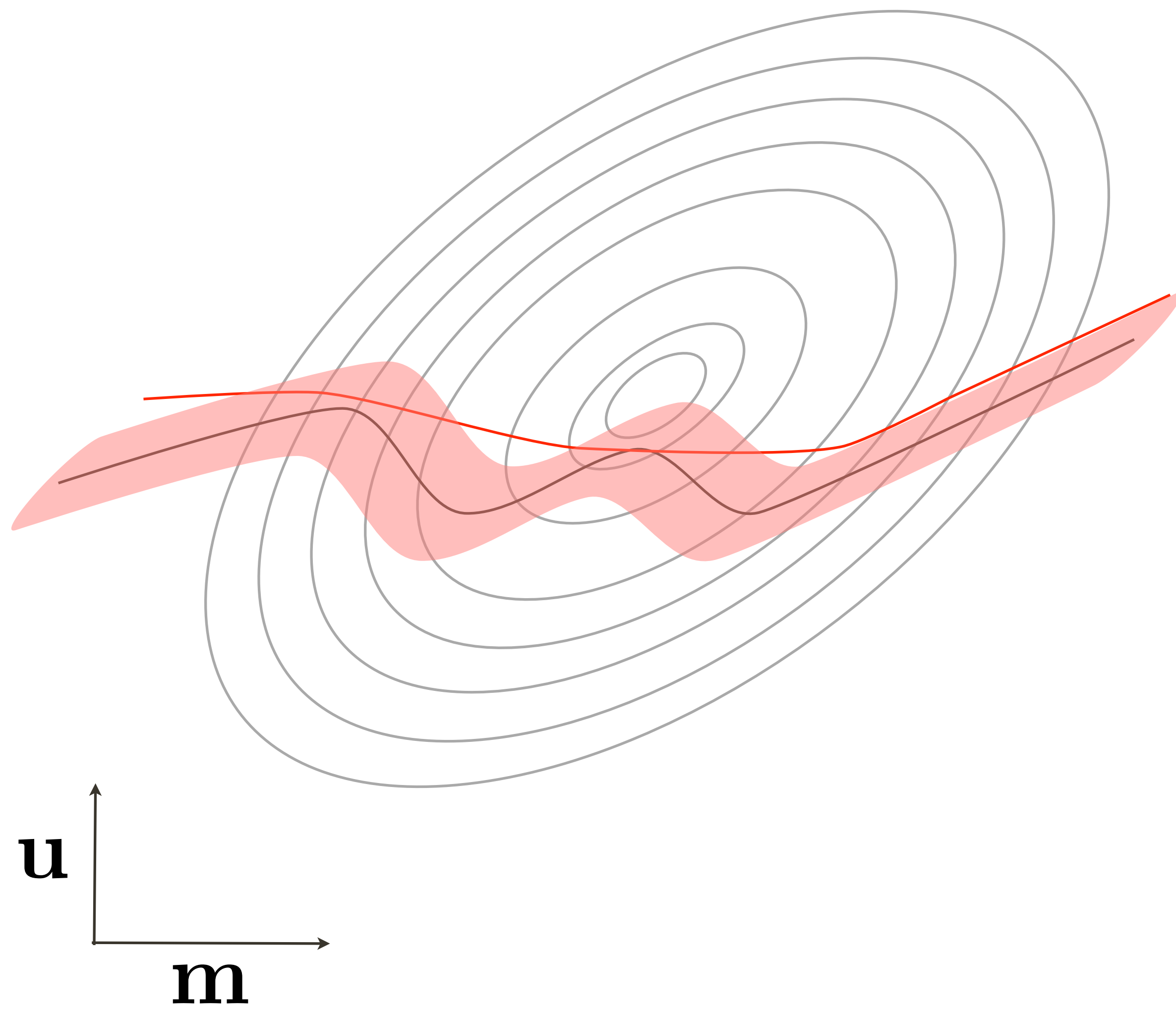
Image w/ *WRI* method

Imaging

WRI-method migration:

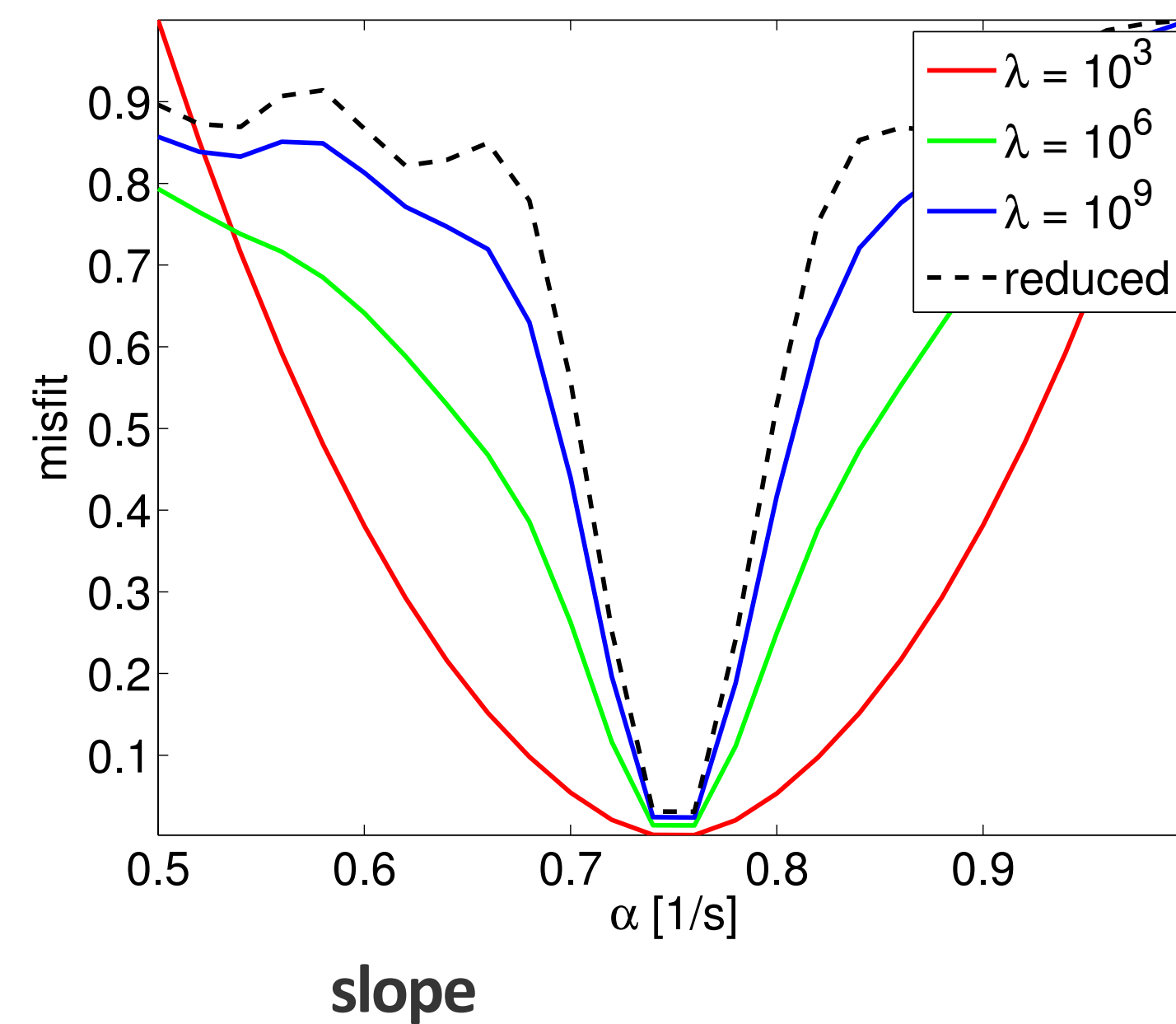
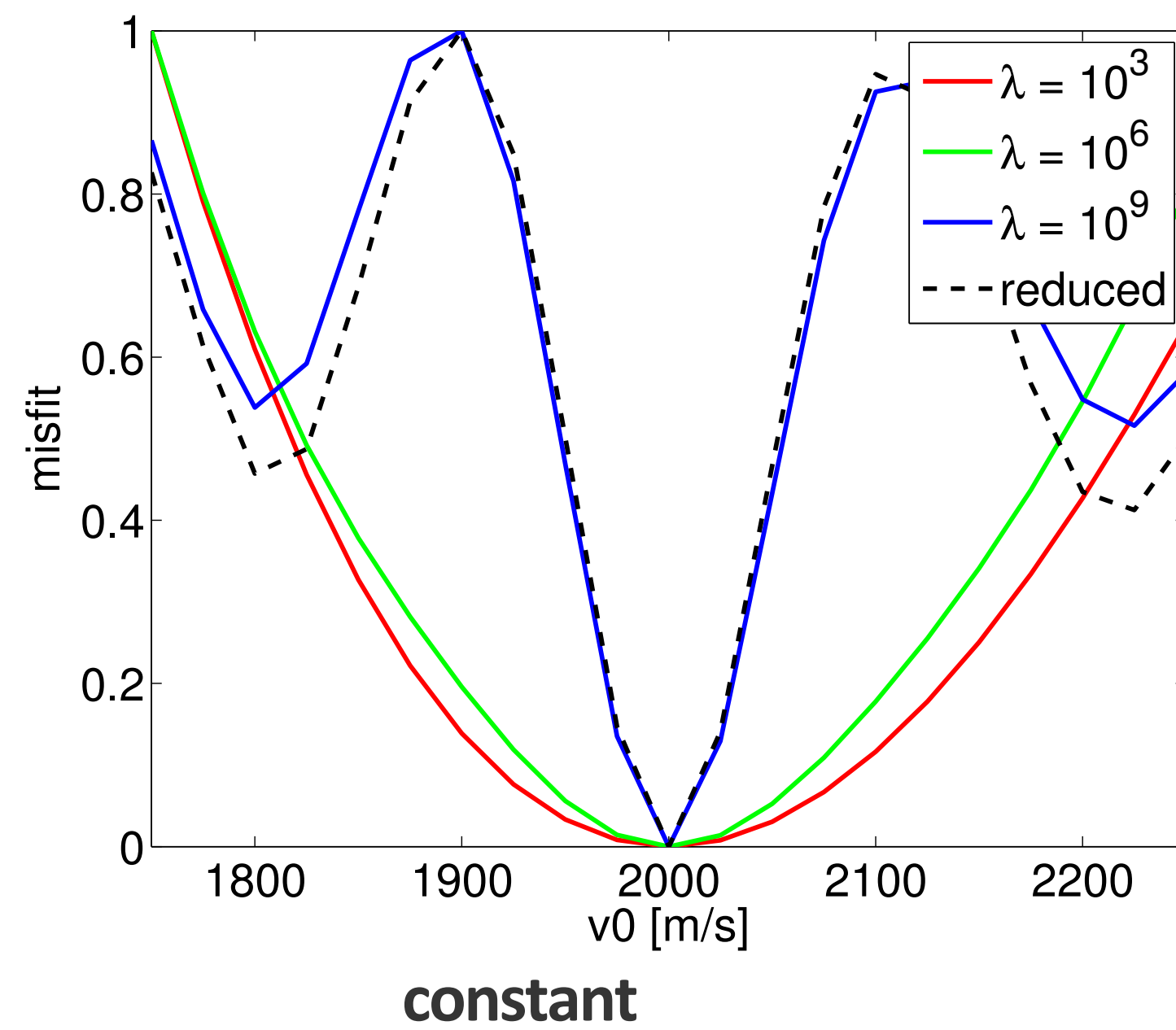
1. solve overdetermined wave-equation
2. go for lunch
3. apply 'imaging condition'

WRI vs. FWI

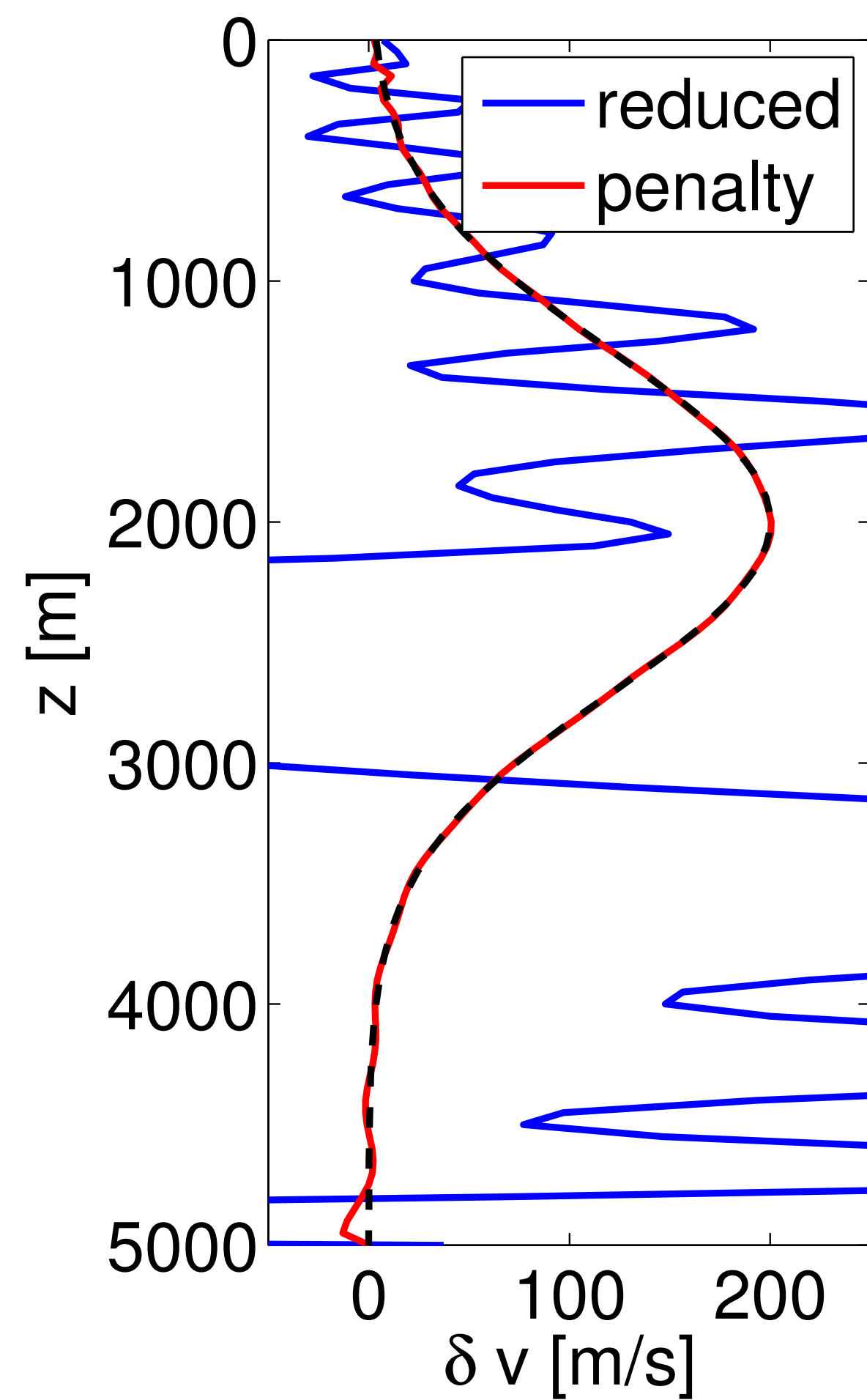


Local minima

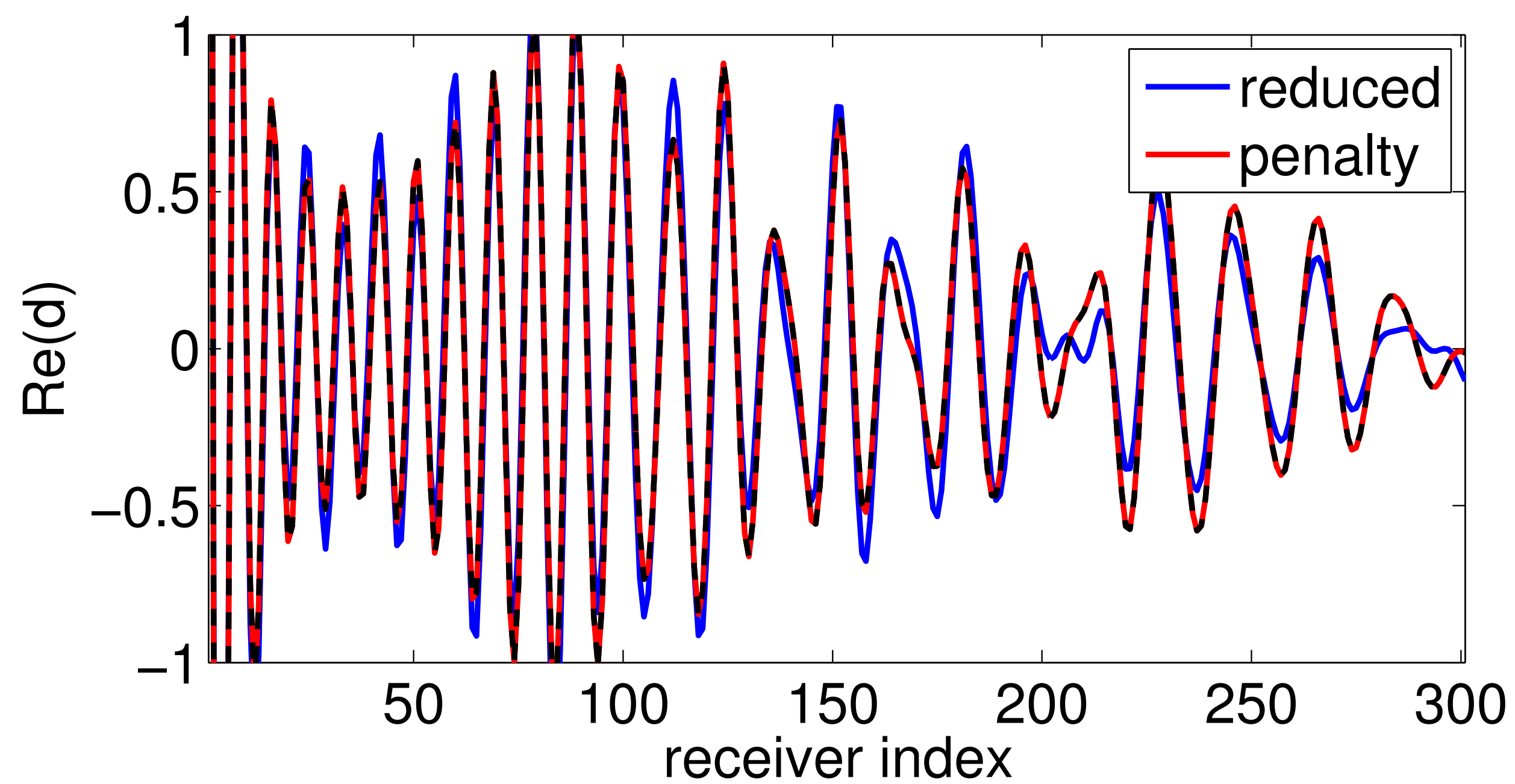
single shot, single frequency data for linear velocity profile $v(z) = v_0 + \alpha z$, misfit as function of (v_0, α)



Local minima



After 1000 iterations



Connections

Extended modelling

The *penalty* formulation

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

can be interpreted as

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

with

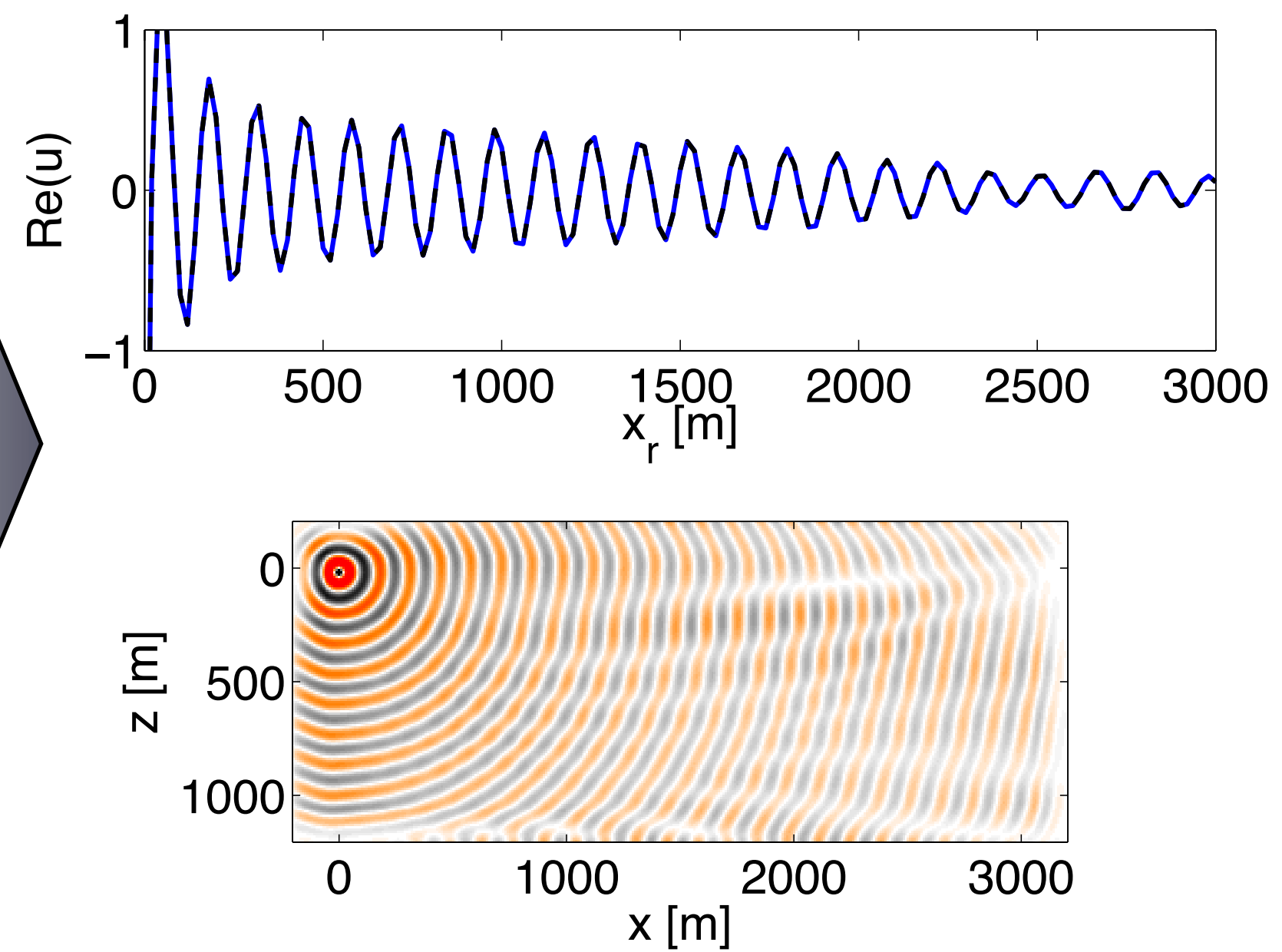
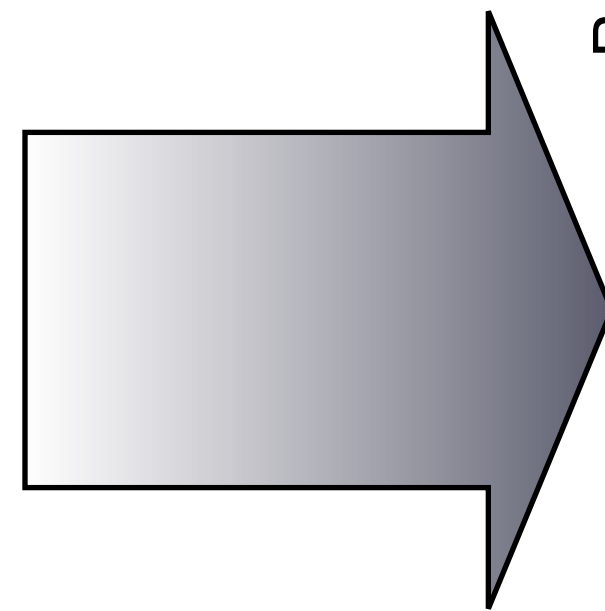
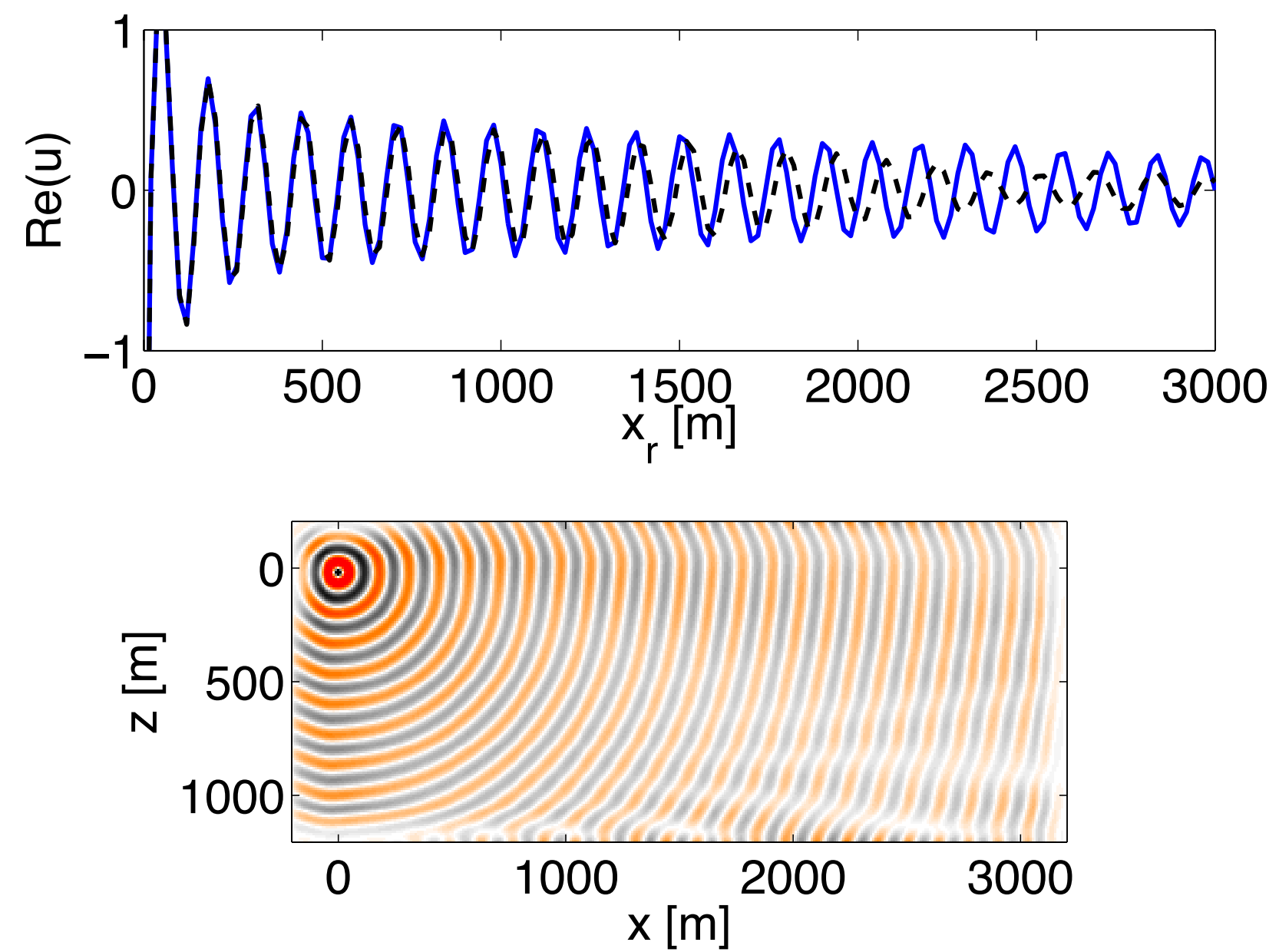
$$\tilde{\mathbf{m}} = (\mathbf{m}, \mathbf{u})$$

For a physically *plausible* model we have

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

Warping

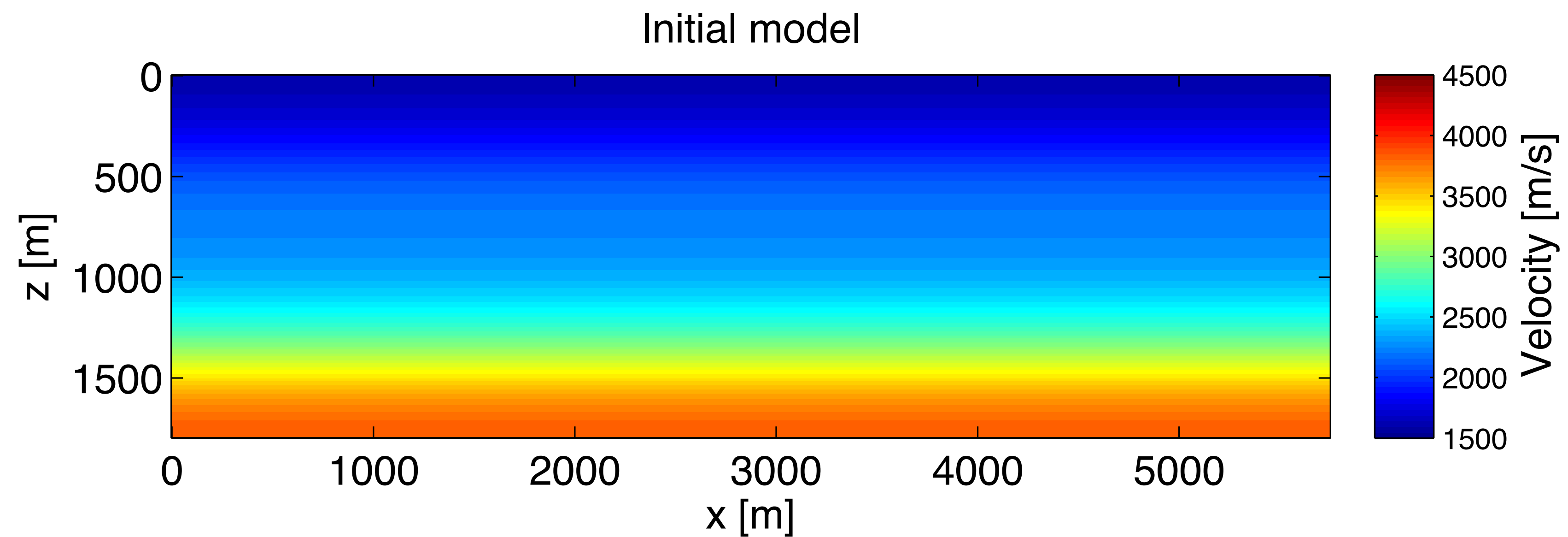
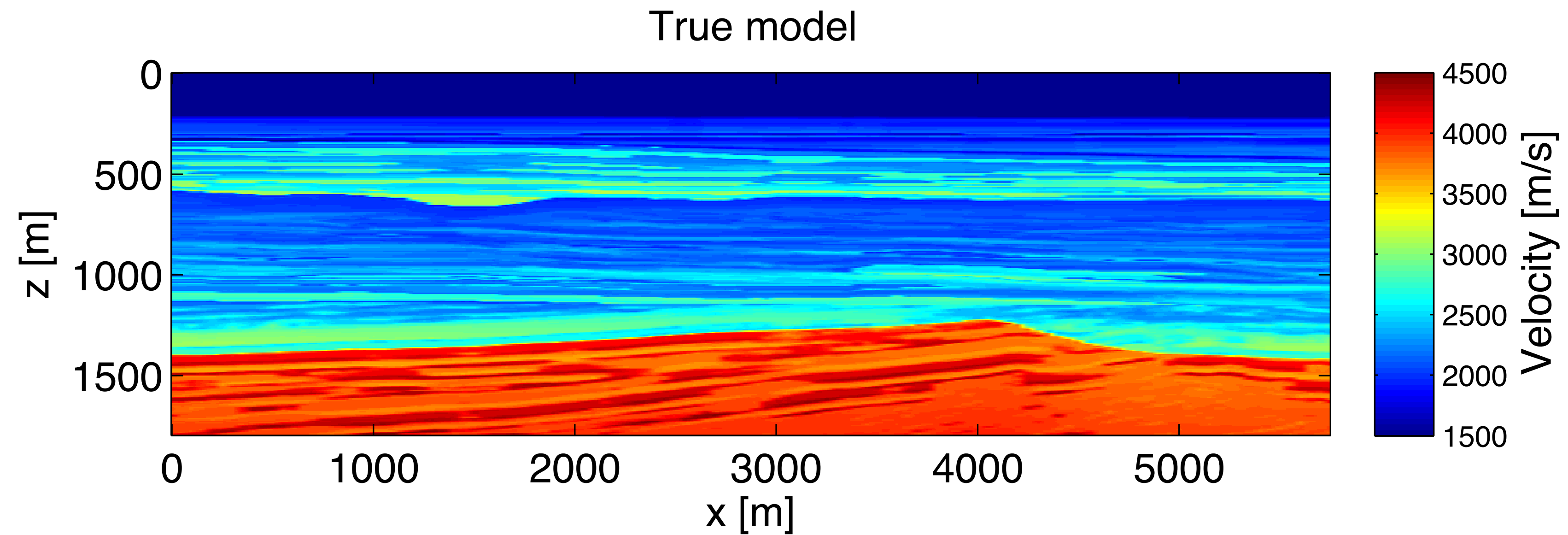
The overdetermined WE is a way of warping



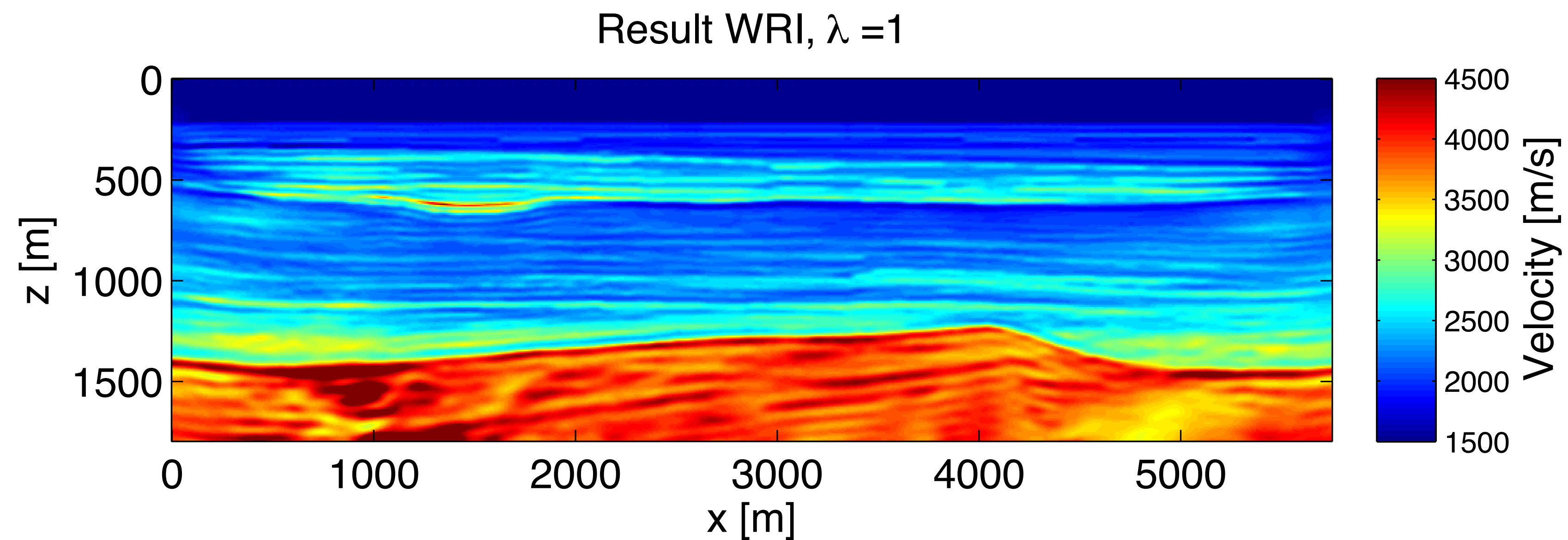
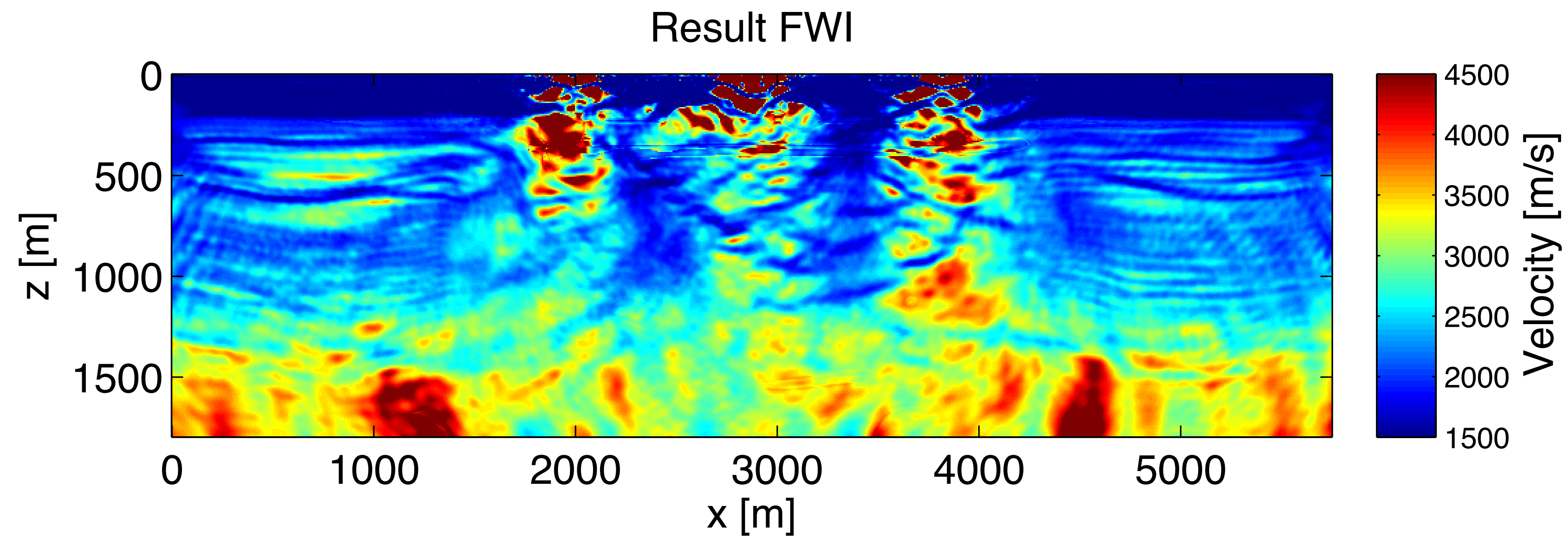
Example – BG Compass model

- *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

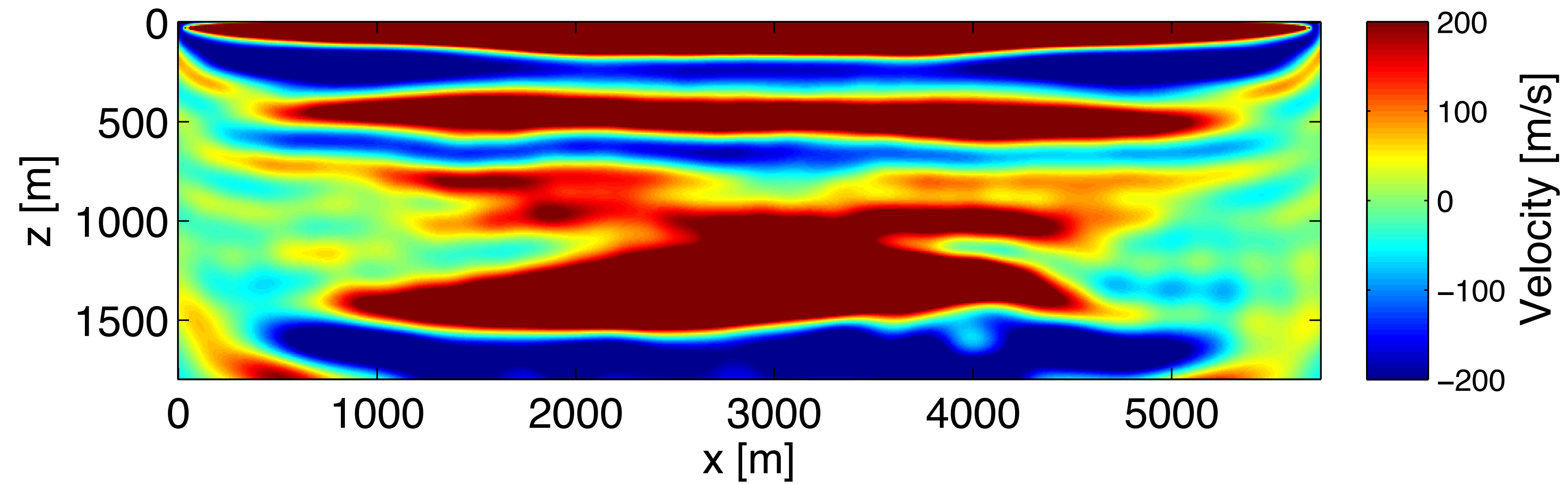


FWI vs WRI

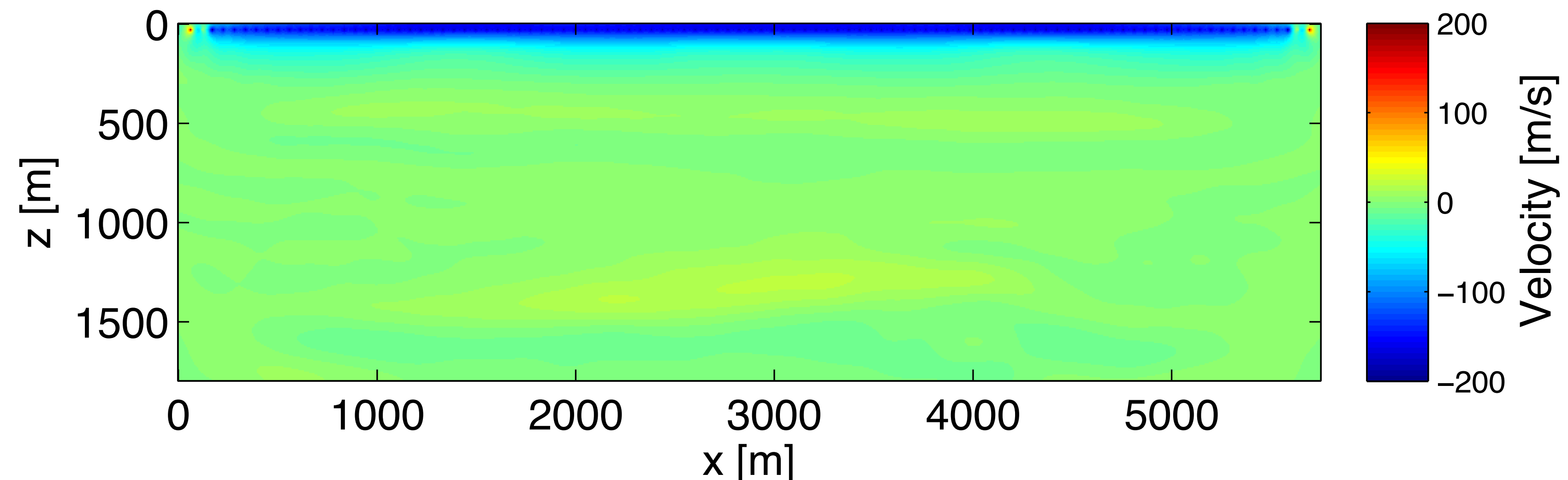


Gradients

First update FWI

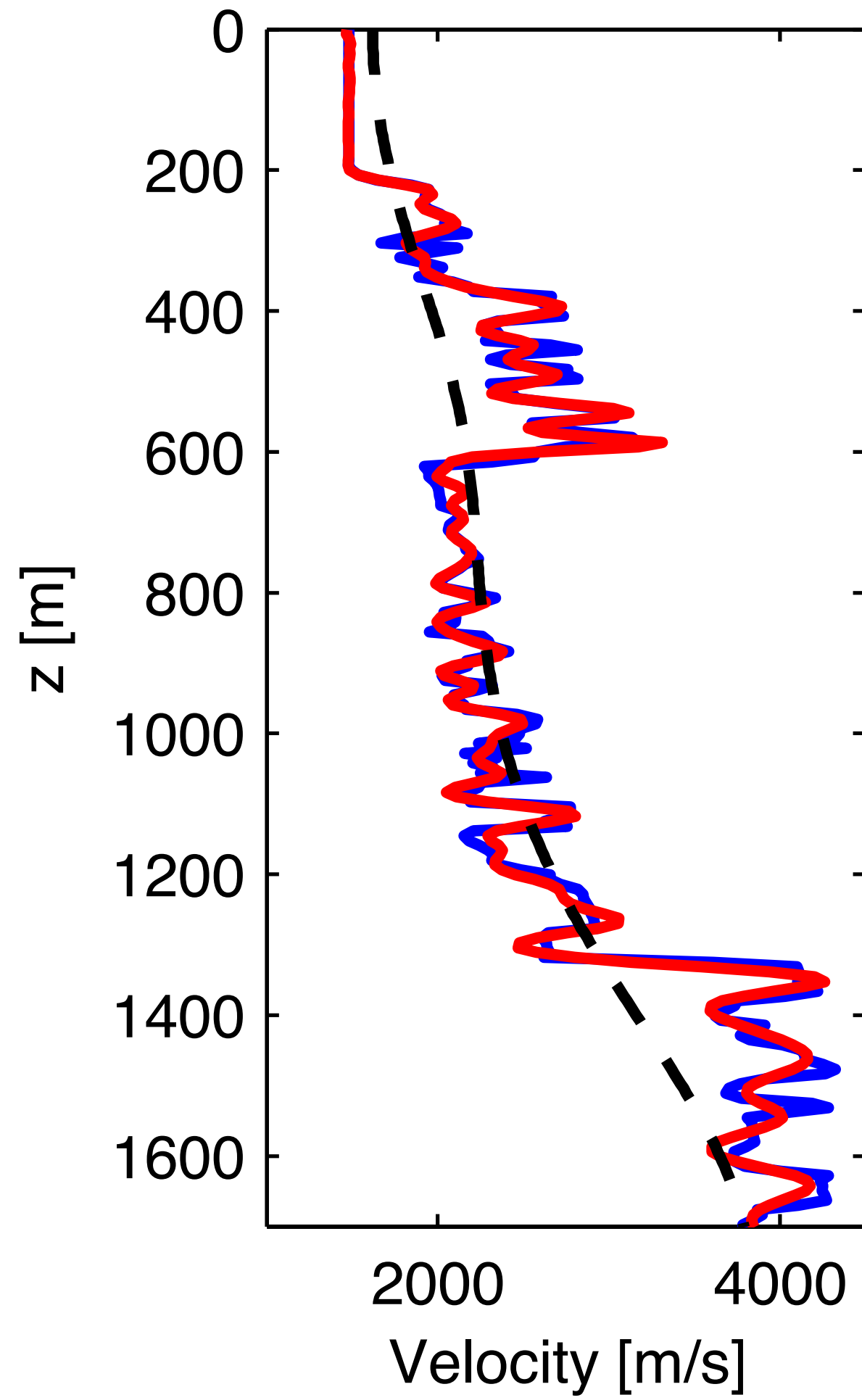


First update WRI, $\lambda = 1$

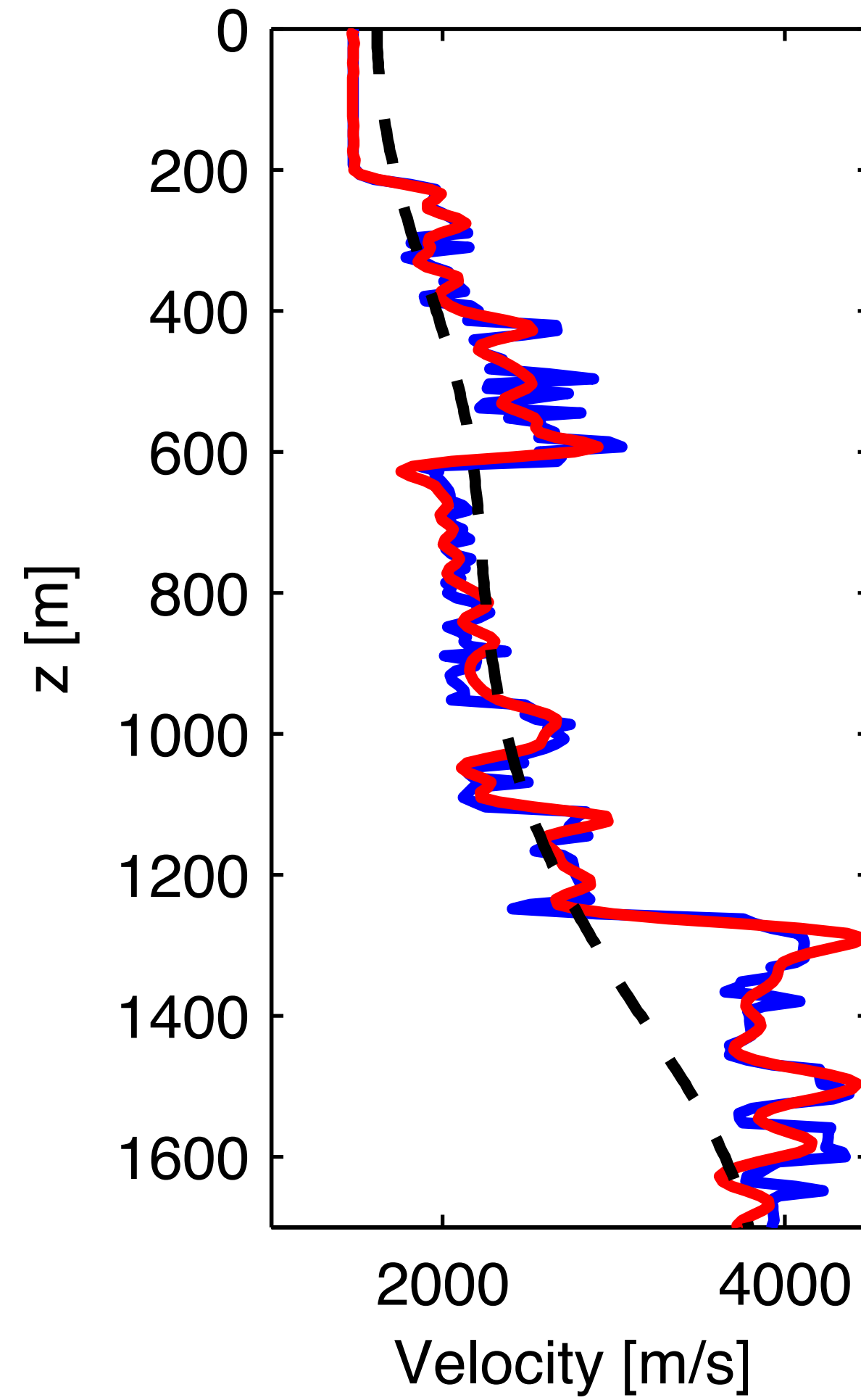


Cross sections

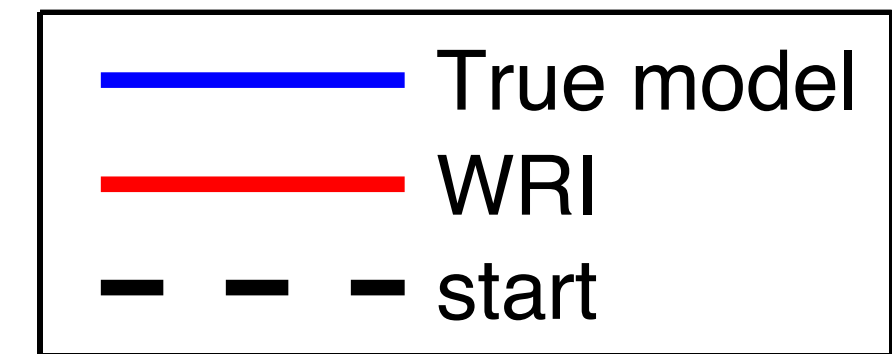
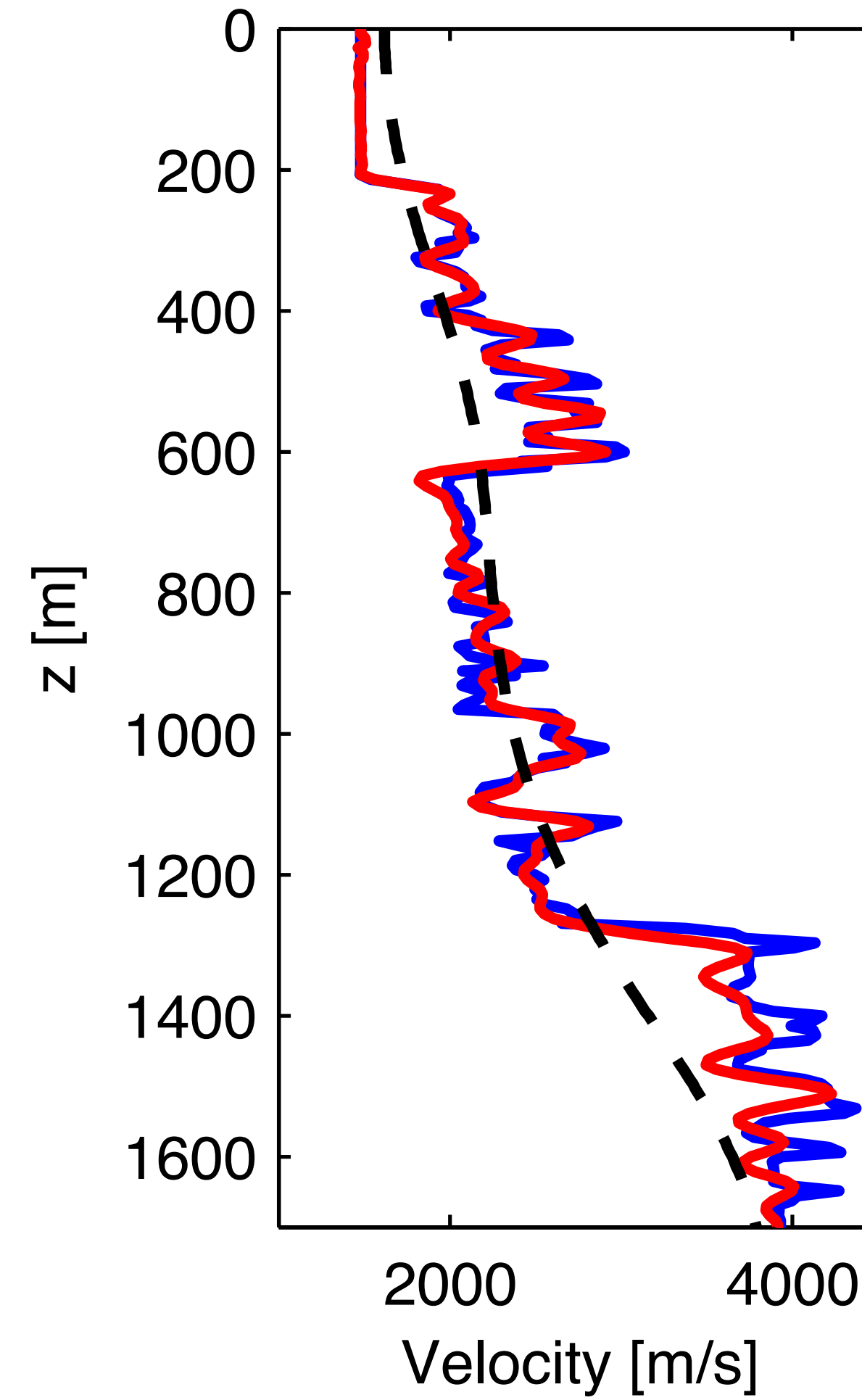
x = 2063.1[m]



x = 3443.1[m]



x = 4305.6[m]



Conclusions

New 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 non-linear*” and therefore *less* susceptible to *local* minima

Still somewhat *early* days in the development:

- ▶ *encouraging* 2-D results w/ *less* sensitivity to *initial* model
- ▶ *cheap* scalings of the *inverse* of the GN Hessian
- ▶ extension to 3D is challenging

Acknowledgements

Thank you for your attention !

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



SINBAD



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, Total SA, WesternGeco, and Woodside.