Wavefield Reconstruction Inversion w/ convex constraints

Felix J. Herrmann

SLIM

University of British Columbia

Wavefield Reconstruction Inversion w/ (*asymmetric*) convex constraints

Ernie Esser, Lluis Guash*, and Felix J. Herrmann

*Sub Salt Solutions Ltd

SLIM

University of British Columbia

John "Ernie" Esser (May 19, 1980 – March 8, 2015)

In memory of Ernie Esser, the UW Math Department, with additional generous funding from Ernie's family and friends and Sub Salt Solution, has created the **Ernie Esser Undergraduate Support Fund**. Gifts to the fund will support undergraduate students who are engaged in research with faculty. The UW Math Department plans to increase the fund with further contributions from Ernie's friends and others who share Ernie's passion for enlarging the mathematical research community. For more information about supporting the Ernie Esser Undergraduate Support Fund, contact Alexandra Haslam, Associate Director of Advancement, Natural Sciences, at alexeck3@uw.edu • (206) 616-1989, Or, to make your gift online, please visit www.washington.edu/giving and search for "Ernie Esser Undergraduate Award."

Motivation

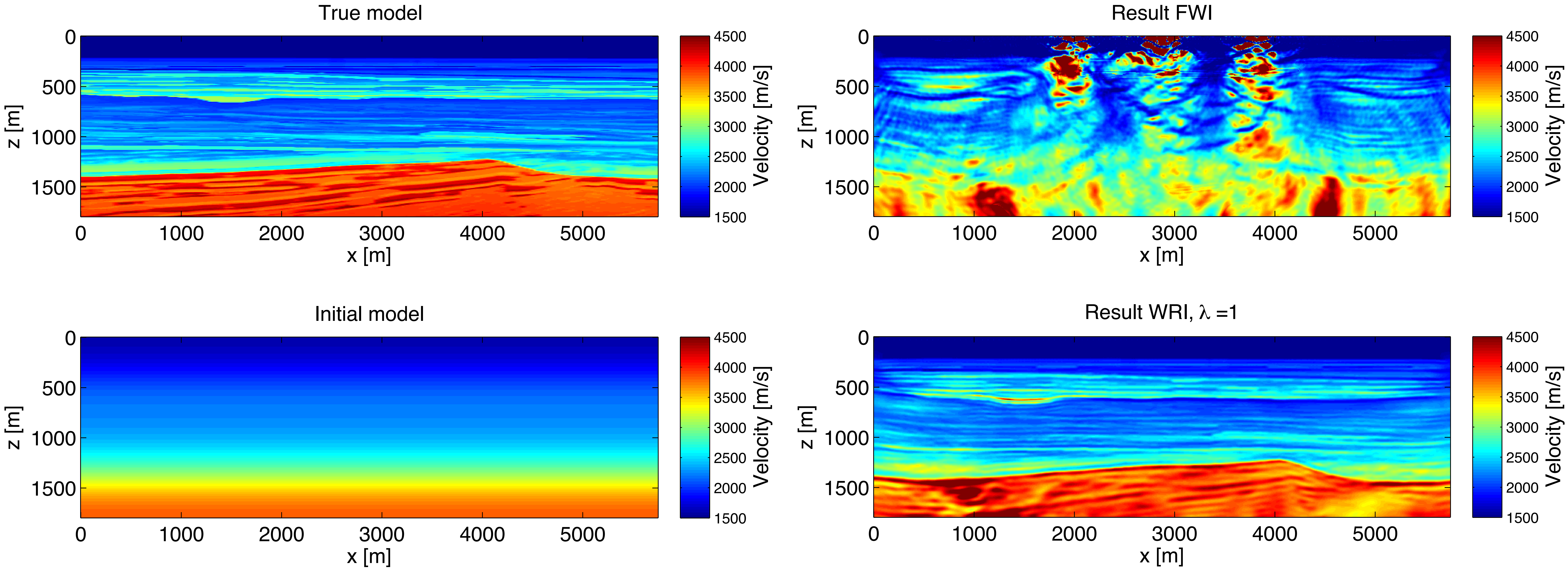
Wave-equation based inversions are non-convex & suffer from local minima (cycle skips)

- for poor starting models
- especially detrimental for high-contrast & high-velocity unconformities (salt & basalt)

Borrow ideas from

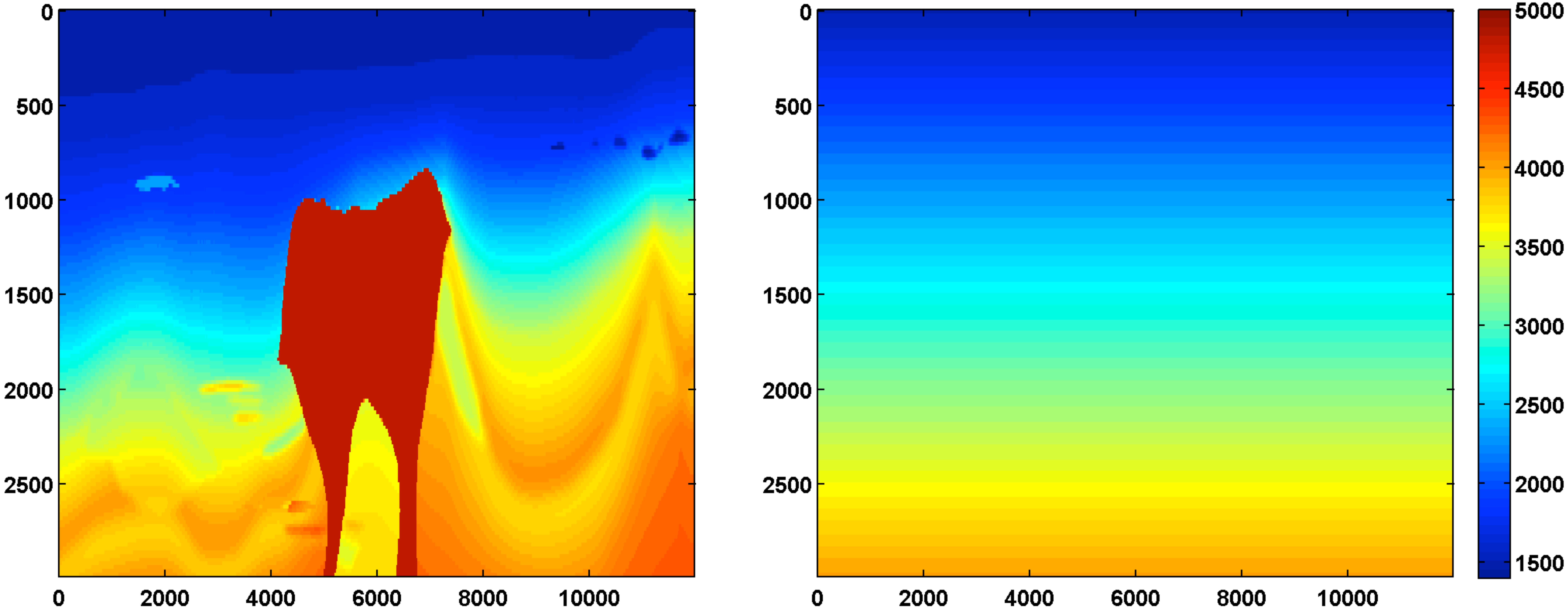
- wave-equation based inversions w/ extensions
- edge-preserving regularization in image processing & compressive sensing
- hinge-loss functions in machine learning
- continuation strategies from (convex) constrained optimization

Wavefield Reconstruction Inversion (WRI)
– poor starting model



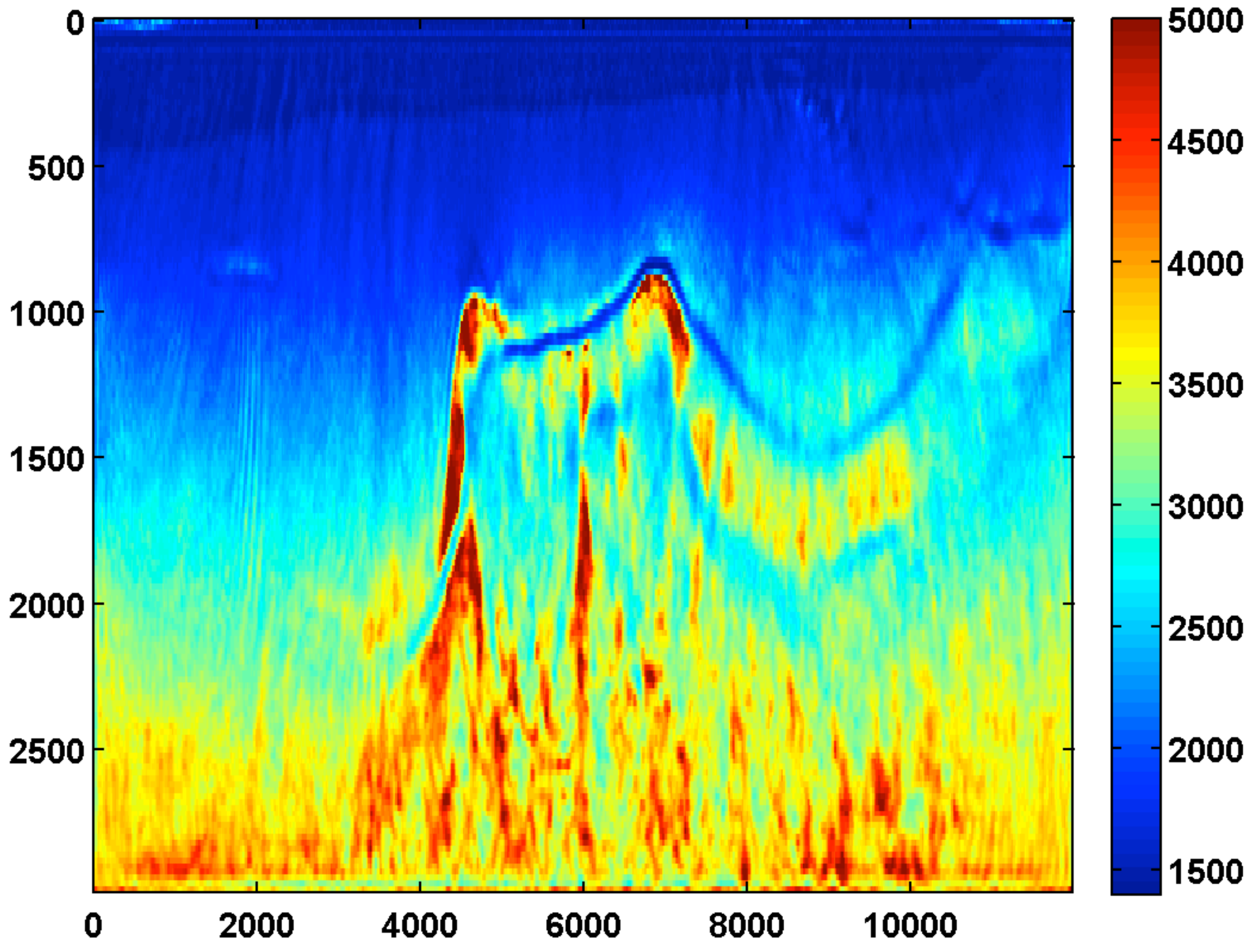
Example from **[Peters et al. 2013]**

Waveform inversion – poor starting model

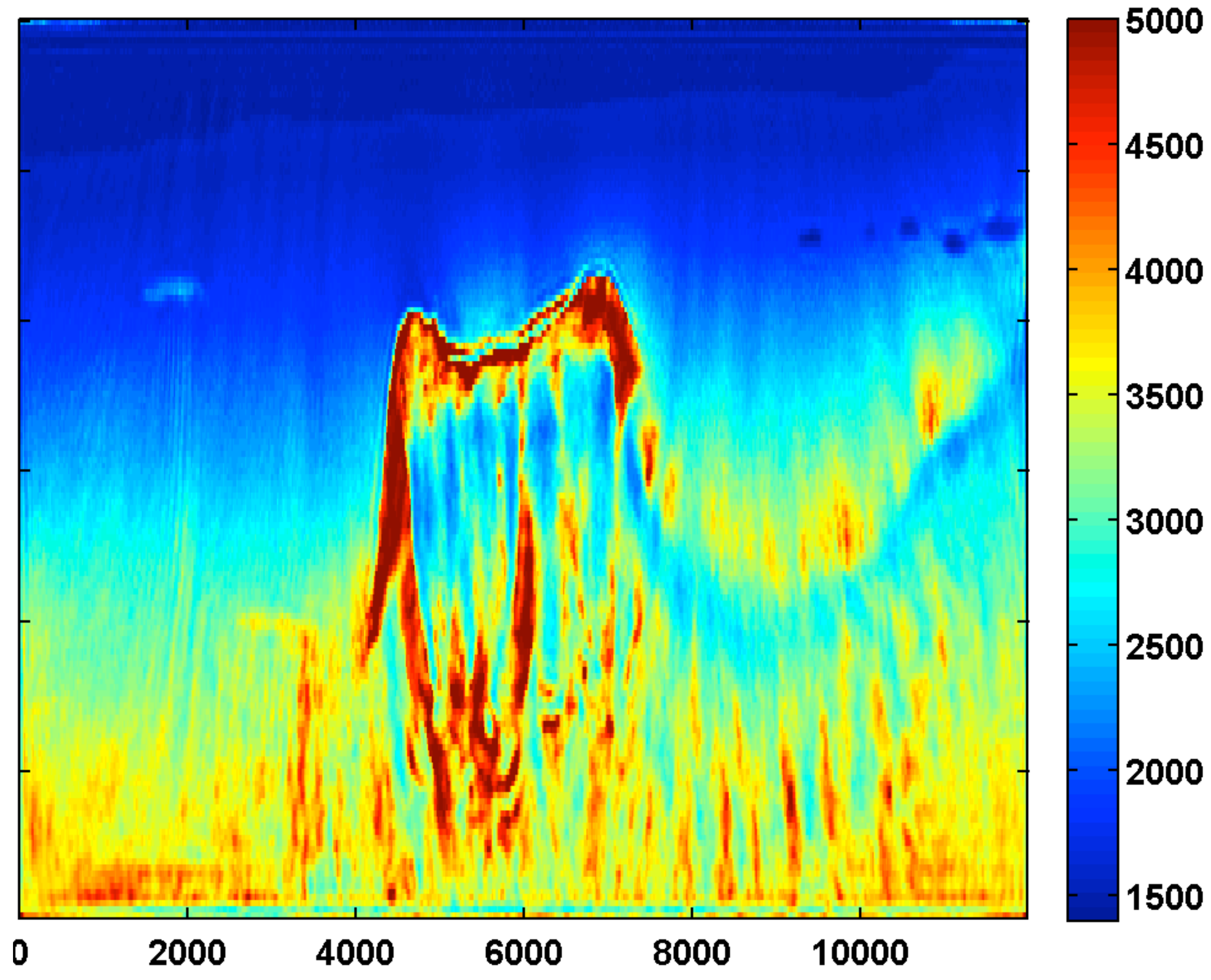


WRI results w/o TV

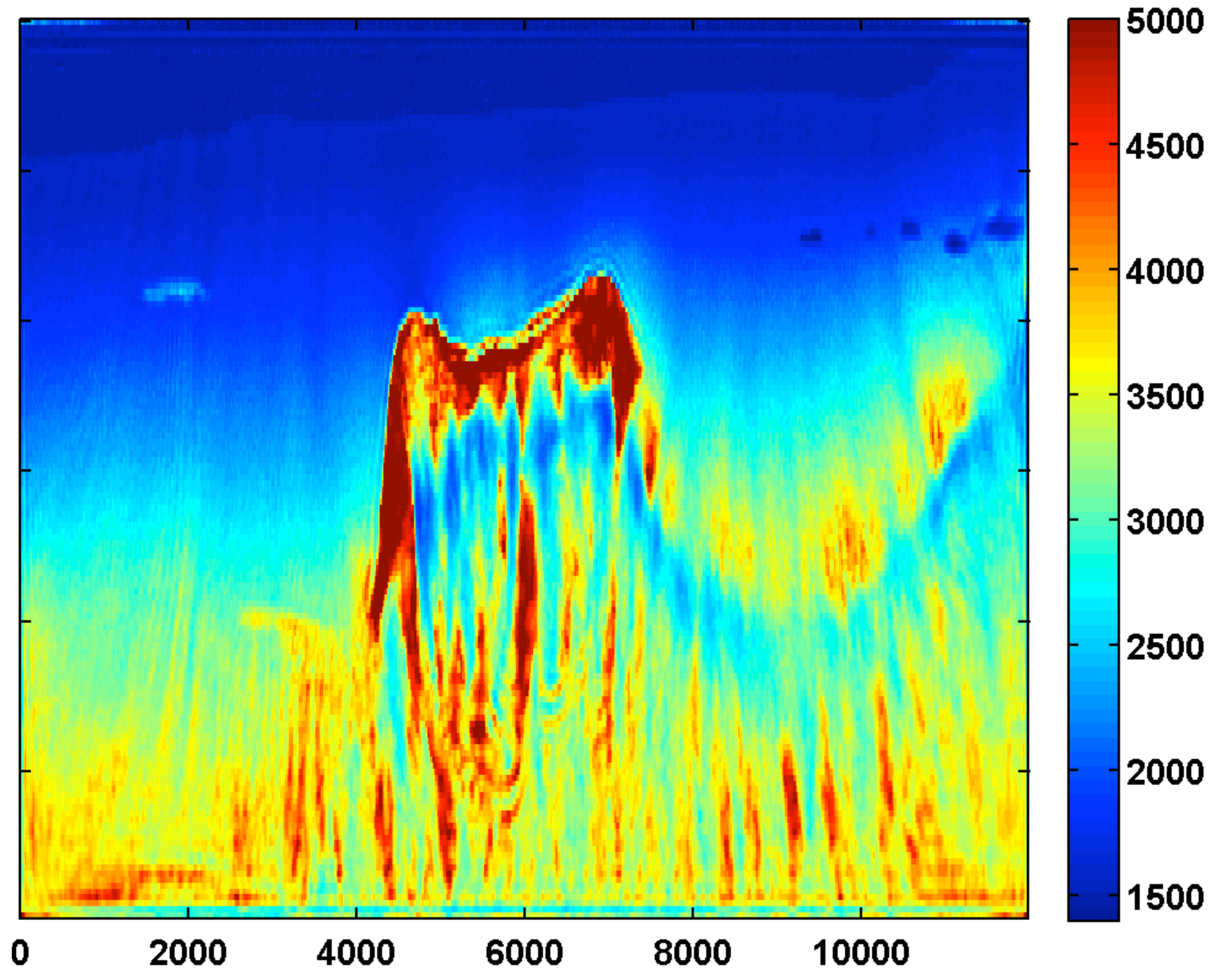
after one cycle through the frequencies



after two cycles through the frequencies

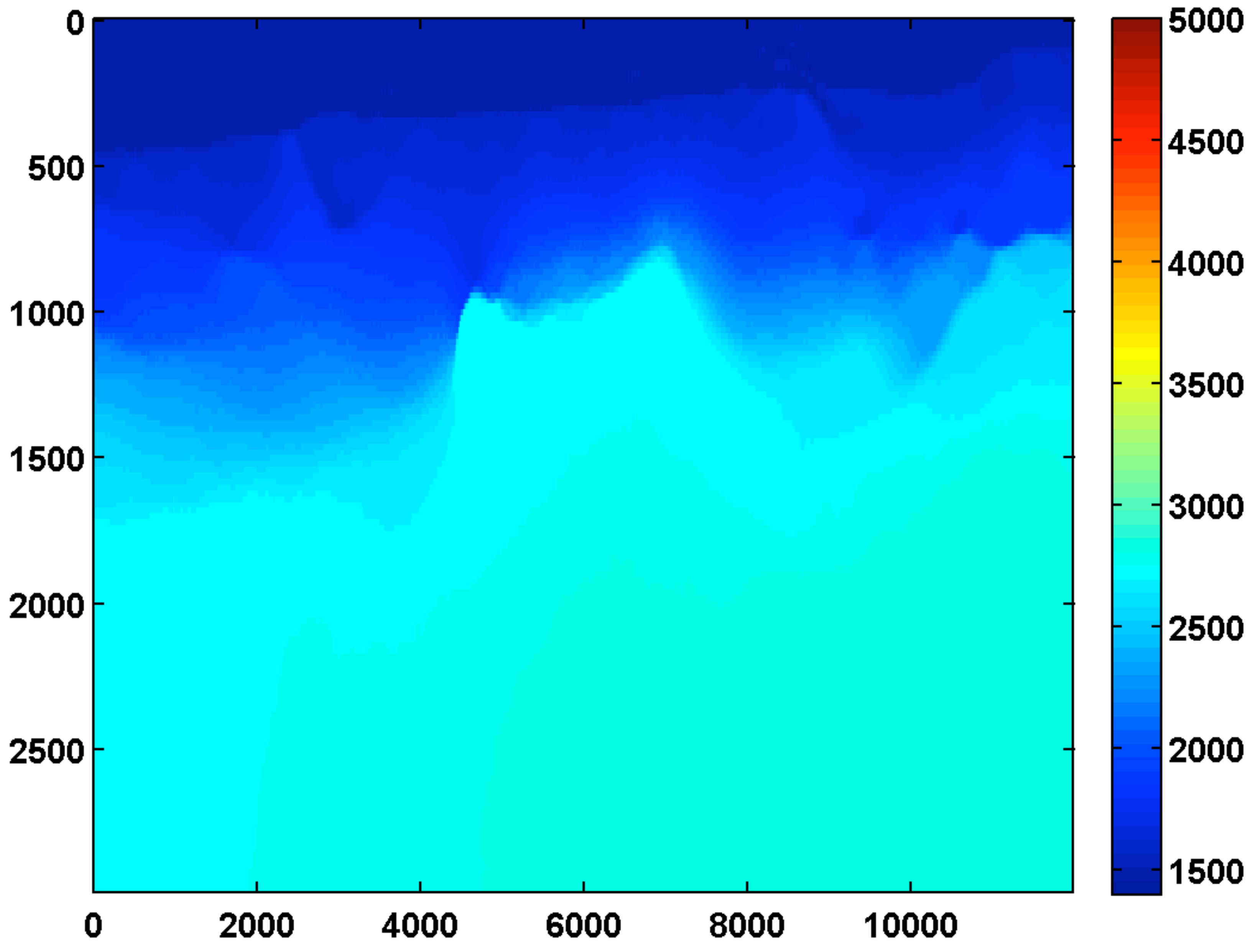


after three cycles through the frequencies

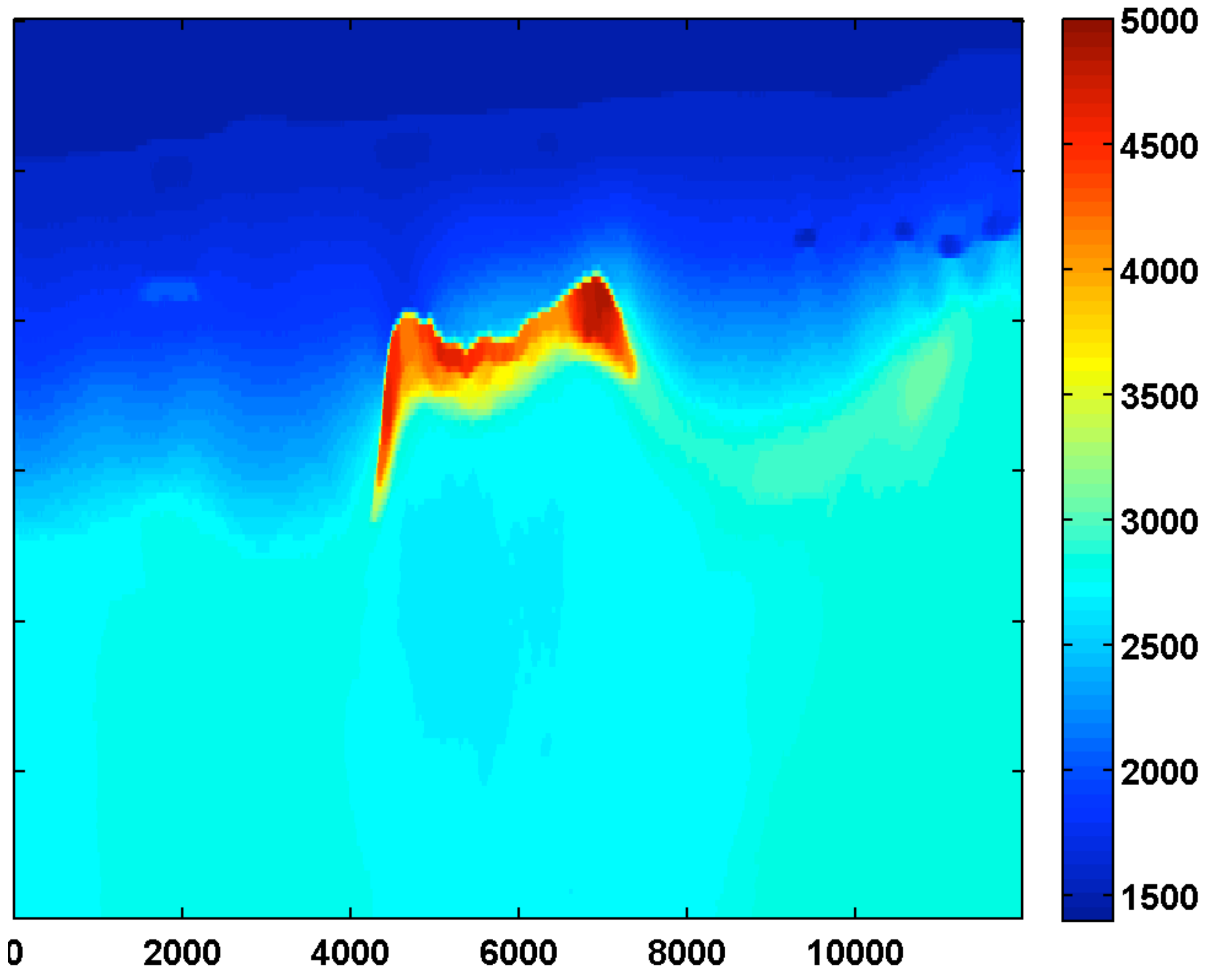


Results w/ TV

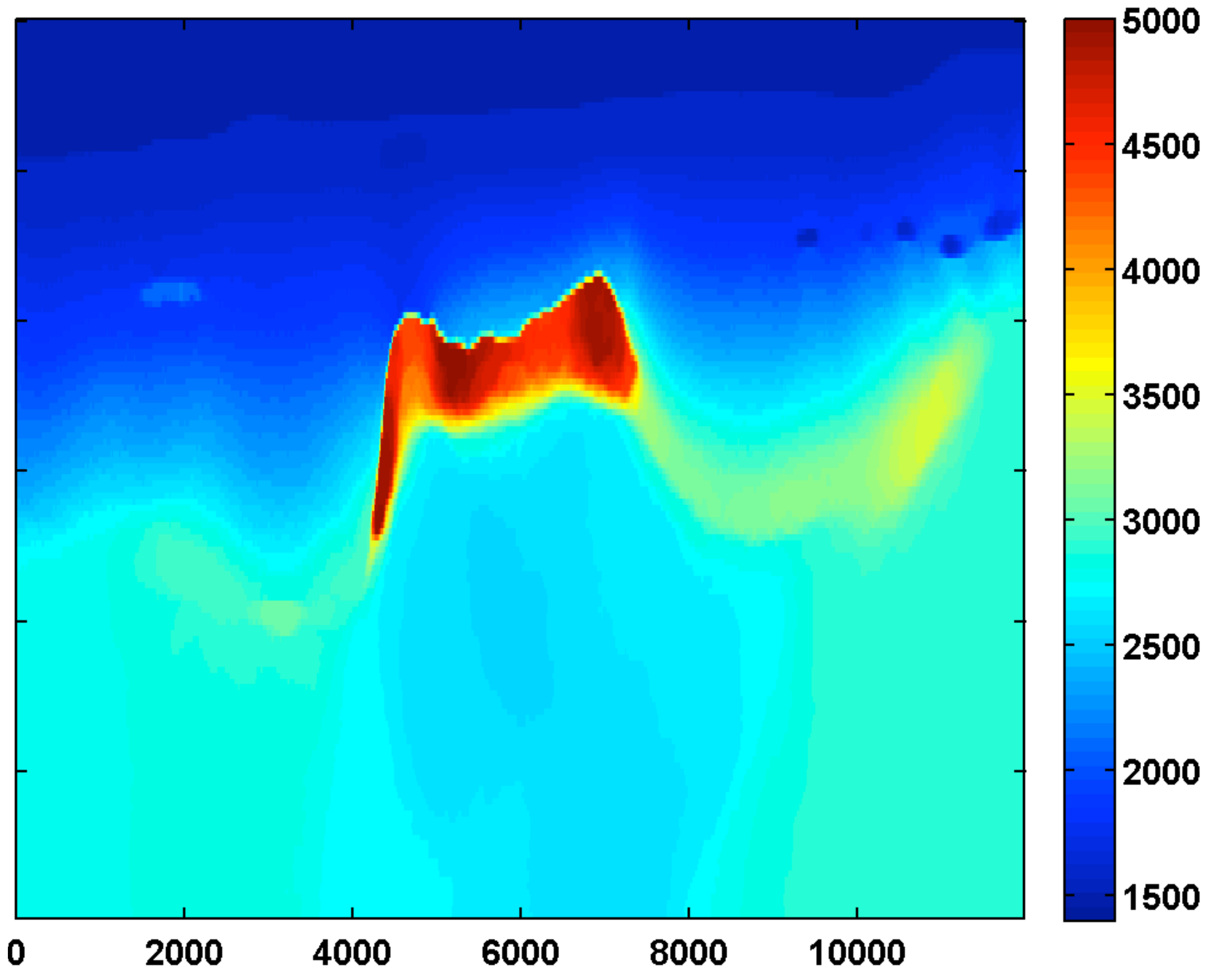
after one cycle through the frequencies



after two cycles through the frequencies



after three cycles through the frequencies



Strategy

Extend the search space

- “less” nonlinear (bi-convex)
- ensures data fit & avoids cycle skips

“Squeeze” the extension by

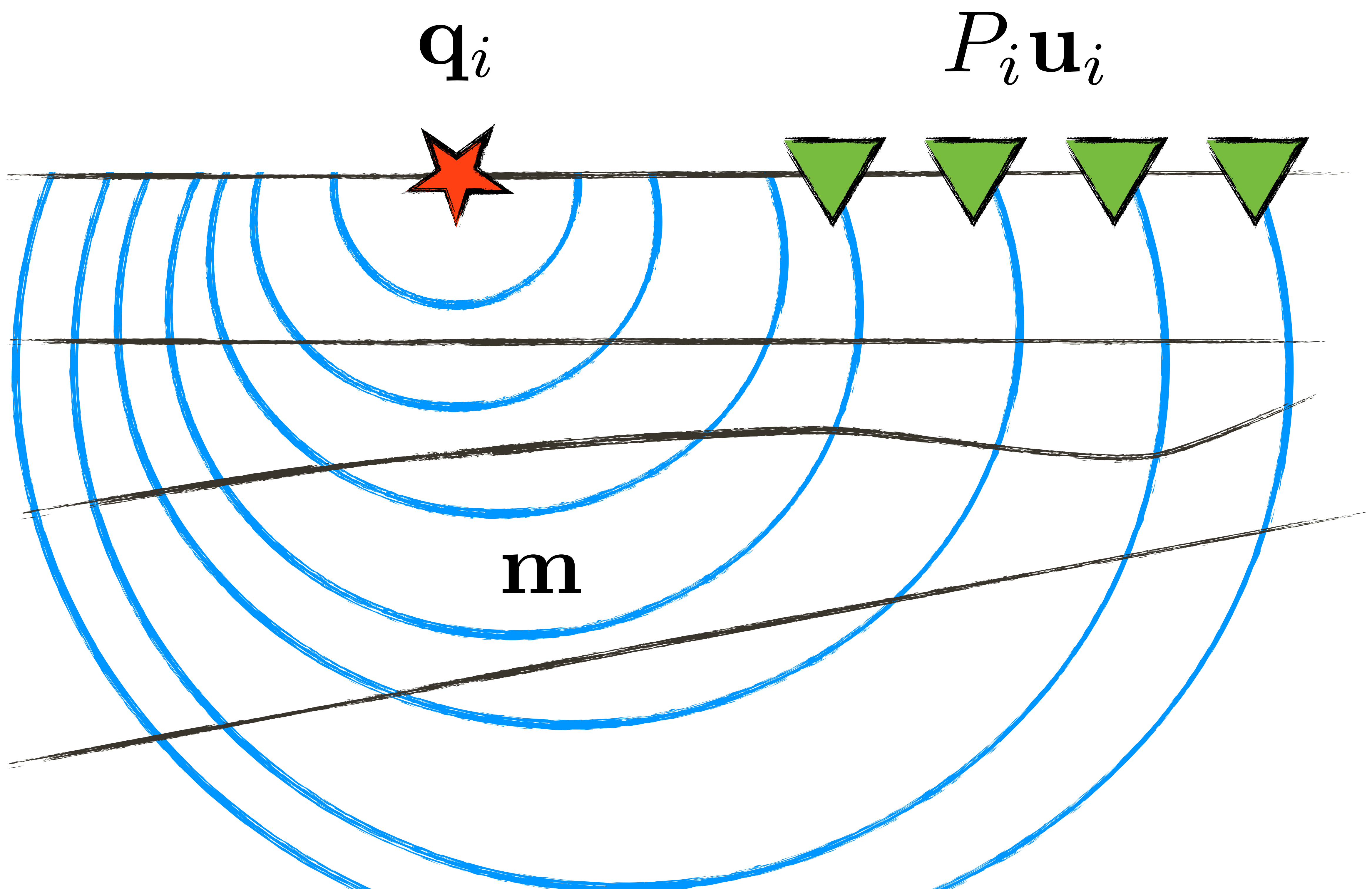
- enforcing the wave equation to compute model updates
- imposing *asymmetric* convex constraints that encode “rudimentary” properties of the geology
- relaxing the convex constraints starts while stressing wave physics

Leverage frequency continuation & warm starts where

- *sparsity-promoting asymmetric* convex constraints limit adverse affects of local minima
- there is hope as long progress towards the solution is made in each sweep

Waveform inversion

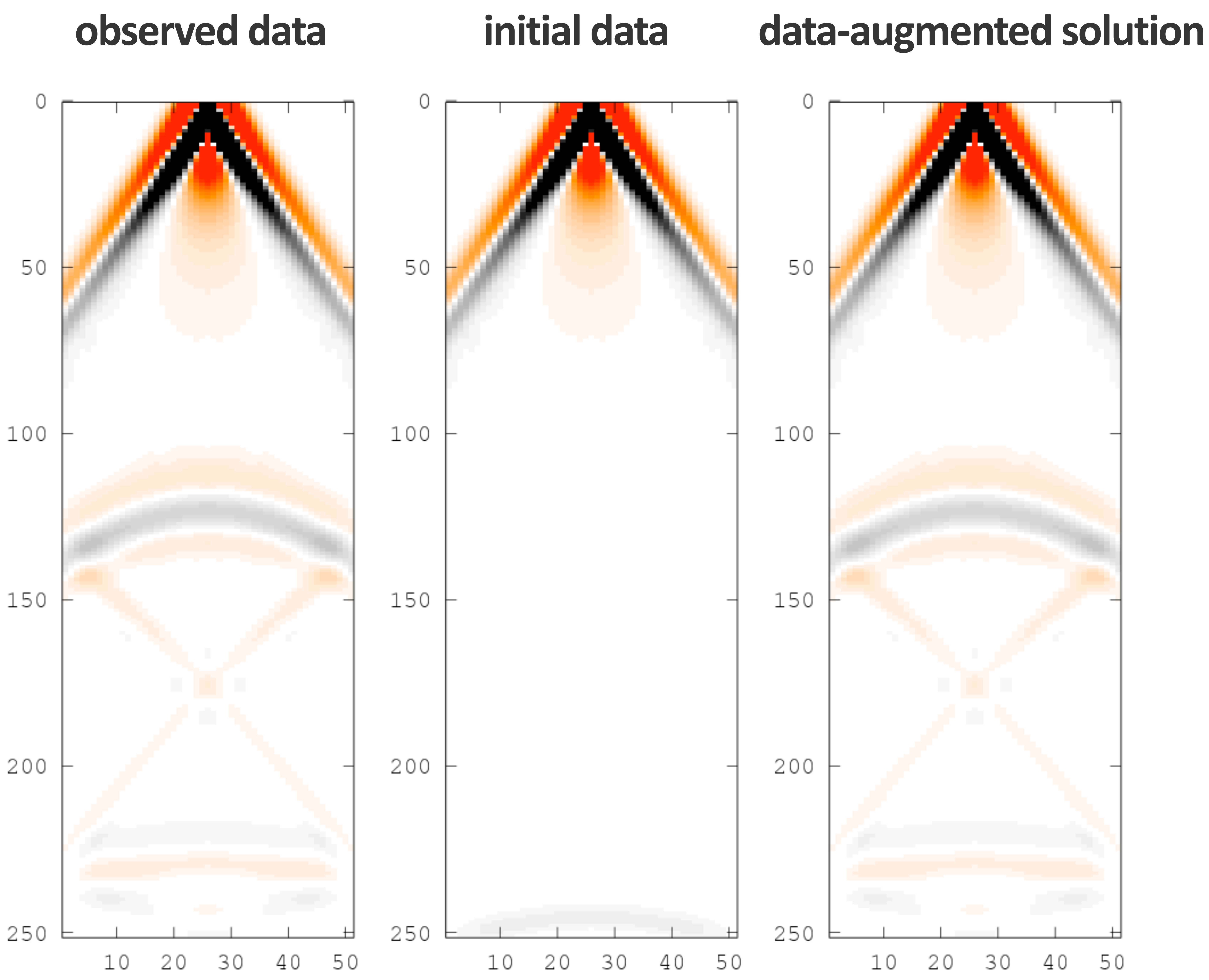
Retrieve the medium parameters from partial measurements of the solution of the wave-equation: $A(\mathbf{m})\mathbf{u}_i = \mathbf{q}_i$



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

[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

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

[Heinkenschloss, ’98 , Haber, ’00]

SLIM

Wavefield-reconstruction Inversion – WRI

Replace PDE-constrained formulation for FWI:

simulated data — simulated wavefield

$$\min_{\mathbf{m},\mathbf{u}} \sum_{sv} \frac{1}{2}\|P\mathbf{u}_{sv} - \mathbf{d}_{sv}\|^2 \quad \text{such that} \quad A_v(\mathbf{m})\mathbf{u}_{sv} = \mathbf{q}_{sv}$$

observed data — Helmholtz equation — source

- 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

by eliminating the constraint

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

- 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 dependence on earth model **(m)**
- dense (GN) Hessian, involves additional PDE solves
- **reliance on accurate starting models to avoid cycle skipping**

SLIM

WRI

or by a penalty formulation

$$\min_{\mathbf{m},\mathbf{u}} \sum_{sv} \frac{1}{2}\|P\mathbf{u}_{sv} - \mathbf{d}_{sv}\|^2 + \frac{\lambda^2}{2}\|A_v(\mathbf{m})\mathbf{u}_{sv} - \mathbf{q}_{sv}\|^2$$

and solve at the n-th iteration for proxy wavefields (for fixed $\mathbf{m}^n$)

$$\bar{\mathbf{u}}_{sv} = \arg\min_{\mathbf{u}_{sv}} \frac{1}{2}\|P\mathbf{u}_{sv} - \mathbf{d}_{sv}\|^2 + \frac{\lambda^2}{2}\|A_v(\mathbf{m}^n)\mathbf{u}_{sv} - \mathbf{q}_{sv}\|^2$$

followed by computing the gradient for the model

$$\mathbf{g}^n = \sum_{sv} \mathrm{Re}\left\{\lambda^2\omega_v^2\, \mathrm{diag}(\bar{\mathbf{u}}_{sv})^* \left(A_v(\mathbf{m}^n)\bar{\mathbf{u}}_{sv} - \mathbf{q}_{sv}\right)\right\}$$

Ernie Esser, Lluís Guasch, Tristan van Leeuwen, Aleksandr Y. Aravkin, and Felix J. Herrmann, “**Total variation regularization strategies in full waveform inversion for improving robustness to noise, limited data and poor initializations**”. 2015.

SLIM

WRI

and reduced diagonal Gauss-Newton Hessian

$$H^n_{sv} \approx \sum_{sv} \mathrm{Re}\left\{\lambda^2 \omega_v^4 \,\mathrm{diag}(\bar{\mathbf{u}}_{sv}(\mathbf{m}^n))^* \,\mathrm{diag}(\bar{\mathbf{u}}_{sv}(\mathbf{m}^n))\right\}$$

to minimize the reduced objective

$$\Phi(\mathbf{m}) = \sum_{sv} \frac{1}{2}\|P\bar{\mathbf{u}}_{sv}(\mathbf{m}) - \mathbf{d}_{sv}\|^2 + \frac{\lambda^2}{2}\|A_v(\mathbf{m})\bar{\mathbf{u}}_{sv}(\mathbf{m}) - \mathbf{q}_{sv}\|^2$$

via scaled gradient descents [Bertsekas ’99]

$$\Delta\mathbf{m} = \arg\min_{\Delta\mathbf{m}\in\mathbb{R}^N} \Delta\mathbf{m}^T\mathbf{g}^n + \frac{1}{2}\Delta\mathbf{m}^T H^n \Delta\mathbf{m} + c_n \Delta\mathbf{m}^T\Delta\mathbf{m}$$

$$\mathbf{m}^{n+1} = \mathbf{m}^n + \Delta\mathbf{m} \text{ with } c_n \geq 0$$

WRI – outer iterations

WRI method

for each source i

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

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

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

end

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

diagonal Hessian = pseudo Hessian

replace by inner loop that imposes convex constraints

Conventional method

for each source i

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

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

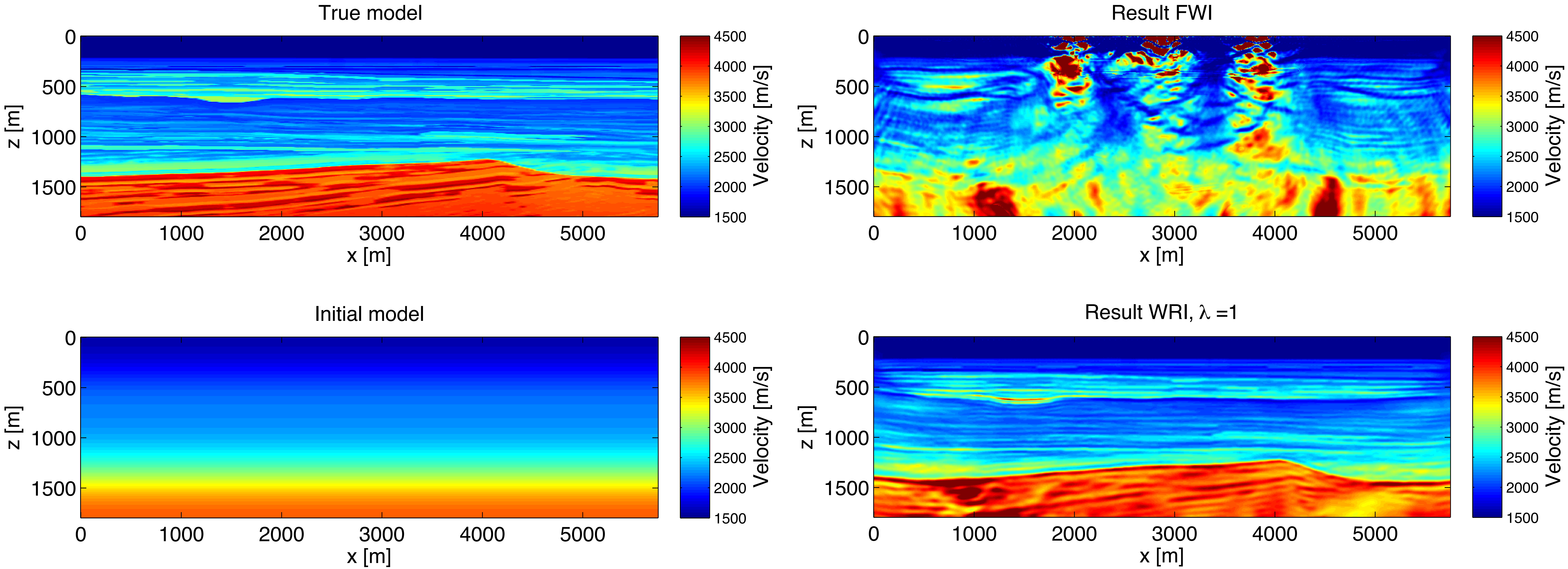
$\mathbf{g} = \mathbf{g} + \omega^2 \text{diag}(\mathbf{u}_i)^*\mathbf{v}_i$

end

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

dense Hessian & too expensive

Waveform inversion – poor starting model



Example from [Peters et al. 2013]

Tristan van Leeuwen and Felix J. Herrmann, "A penalty method for PDE-constrained optimization in inverse problems". 2015. Abstract

A note on choosing λ

Low-noise case:

$$\lambda \sim \mu_1(A^{-*}P^*PA^{-1})$$

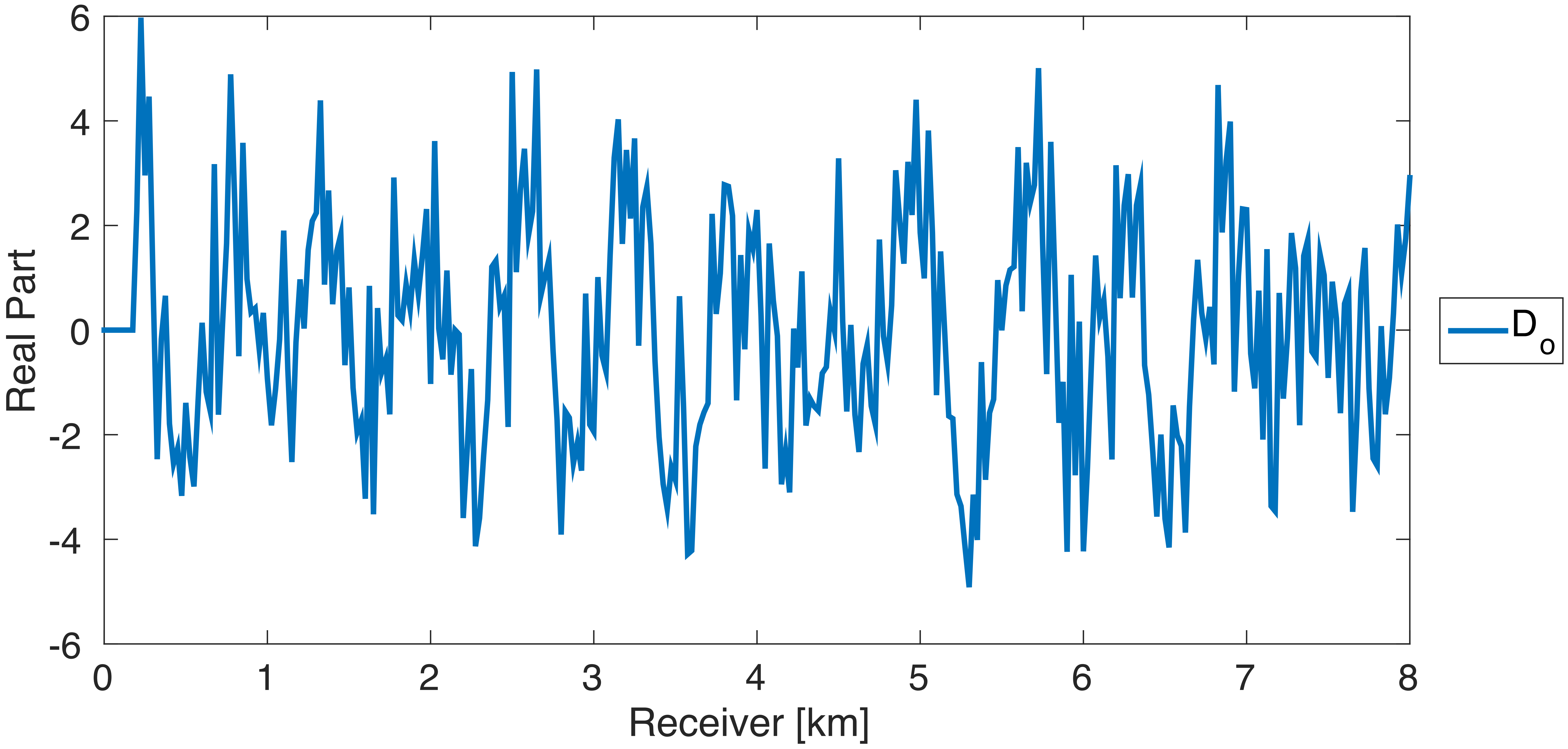
High-noise case:

Select by striking a balance between

- sufficient data fit to avoid cycle skipping
- sufficient "smoothing" to avoid fitting the noise

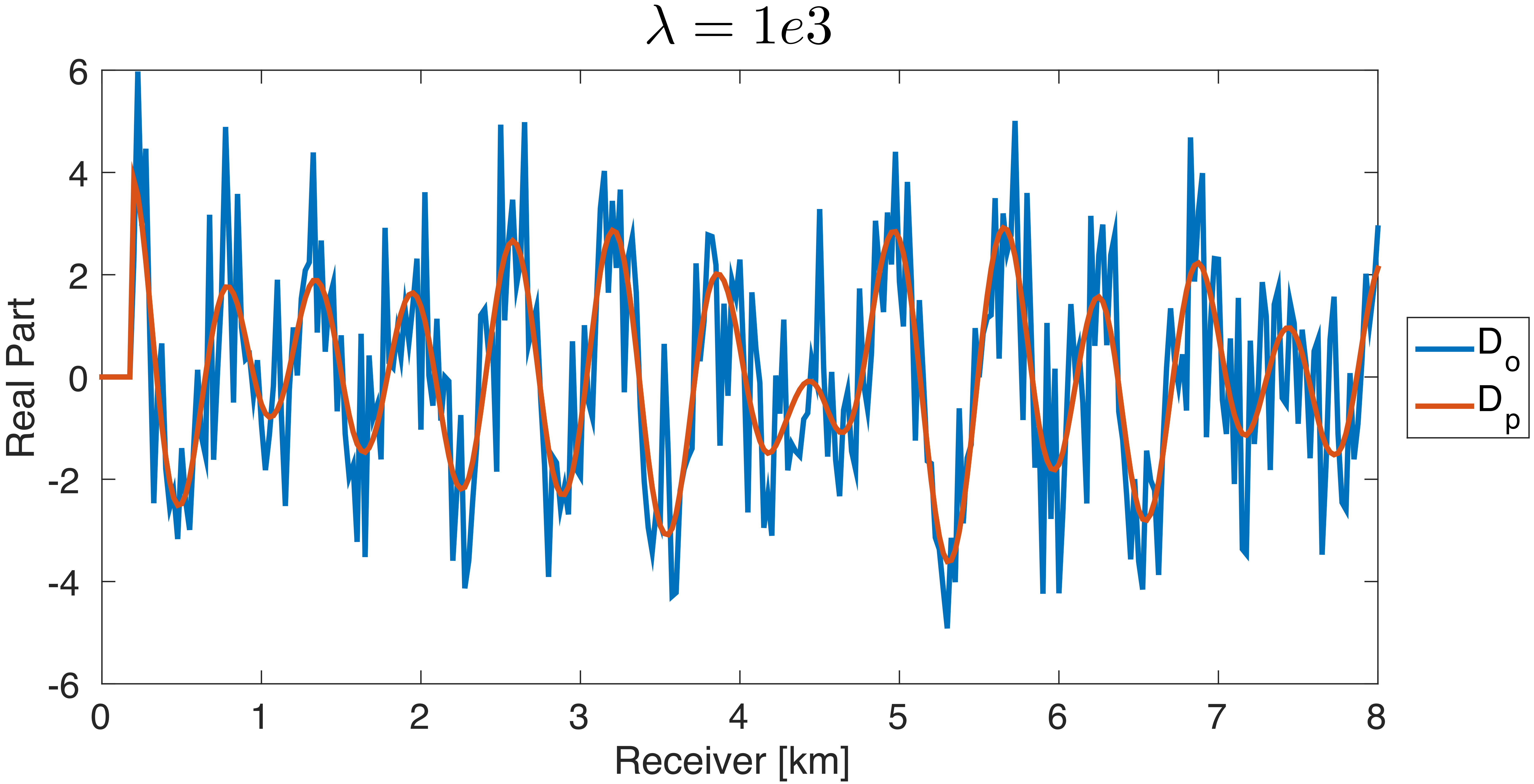
WRI's penalty formulation can be interpreted as a "denoiser"…

Noisy data
— 3 Hz



Noisy data + fit

— 3 Hz



SLIM

Including convex constraints

Wave-equation based inversions call for regularization, e.g. via convex constraints

$$\Delta \mathbf{m} = \underset{\Delta \mathbf{m} \in R^N}{\arg\min}\, \Delta \mathbf{m}^T \mathbf{g}^n + \frac{1}{2} \Delta \mathbf{m}^T H^n \Delta \mathbf{m} + c_n \Delta \mathbf{m}^T \Delta \mathbf{m}$$

$$\text{such that} \quad \mathbf{m}^n + \Delta \mathbf{m} \in C$$

- guarantees $\mathbf{m}^{n+1} \in C$
- more difficult to compute
- feasible if it is easy to project onto
- naive projections $\mathbf{m}^{m+1} = \Pi_C \left(\mathbf{m}^n - (H^n)^{-1} \mathbf{g}^n \right)$ are not guaranteed to converge [Bertsekas '99]

Scaled Gradient Projections

Algorithm 1 A Scaled Gradient Projection Algorithm :

$n = 0$; $m^0 \in C$; $\rho > 0$; $\epsilon > 0$; $\sigma \in (0, 1]$;
H symmetric with eigenvalues between $\lambda_H^{\min}$ and $\lambda_H^{\max}$;
$\xi_1 > 1$; $\xi_2 > 1$; $c_0 > \max(0, \rho - \lambda_H^{\min})$;
while $n = 0$ or $\frac{\|m^n - m^{n-1}\|}{\|m^n\|} > \epsilon$
  $\Delta m = \arg\min_{\Delta m \in C - m^n} \Delta m^T \nabla F(m^n) + \frac{1}{2} \Delta m^T (H^n + c_n \mathrm{I}) \Delta m$
  if $F(m^n + \Delta m) - F(m^n) > \sigma(\Delta m^T \nabla F(m^n) + \frac{1}{2} \Delta m^T (H^n + c_n \mathrm{I}) \Delta m)$
    $c_n = \xi_2 c_n$
  else
    $m^{n+1} = m^n + \Delta m$
    $c_{n+1} = \begin{cases} \frac{c_n}{\xi_1} & \text{if } \frac{c_n}{\xi_1} > \max(0, \rho - \lambda_H^{\min}) \\ c_n & \text{otherwise} \end{cases}$
    Define H^{n+1} to be symmetric Hessian approximation
      with eigenvalues between $\lambda_H^{\min}$ and $\lambda_H^{\max}$
    $n = n + 1$
  end if
end while

Bound constraints
– via scaled gradient projections

For strictly positive diagonal Gauss-Newton Hessians:

$$\Delta\mathbf{m} = \arg\min_{\Delta\mathbf{m}} \Delta\mathbf{m}^T\mathbf{g}^n + \frac{1}{2}\Delta\mathbf{m}^T(H^n + c_n\mathrm{I})\Delta\mathbf{m}$$
$$\text{subject to } \mathbf{m}_i^n + \Delta\mathbf{m}_i \in [B_i^l, B_i^u], \;\; i = 1 \cdots N$$

for which there exists a closed form solution

$$\Delta\mathbf{m}_i = \max\left(B_i^l - \mathbf{m}_i^n, \min\left(B_i^u - \mathbf{m}_i^n, -[(H^n + c_n I)^{-1}\mathbf{g}^n]_i\right)\right)$$

that is computationally affordable.

[Oldenburg ’83; Akçelik ’08; Anagaw ’11; Maharramov ’14; Esser & FJH ’14]

SLIM

Total-variation regularization
– w/ bound constraints

Promote models w/ sharp boundaries via

$$\mathbf{m}^{n+1} = \mathbf{m}^n + \Delta\mathbf{m} \quad \text{subject to} \quad \mathbf{m}^{n+1} \in C_{\text{box}} \cap C_{\text{TV}}$$

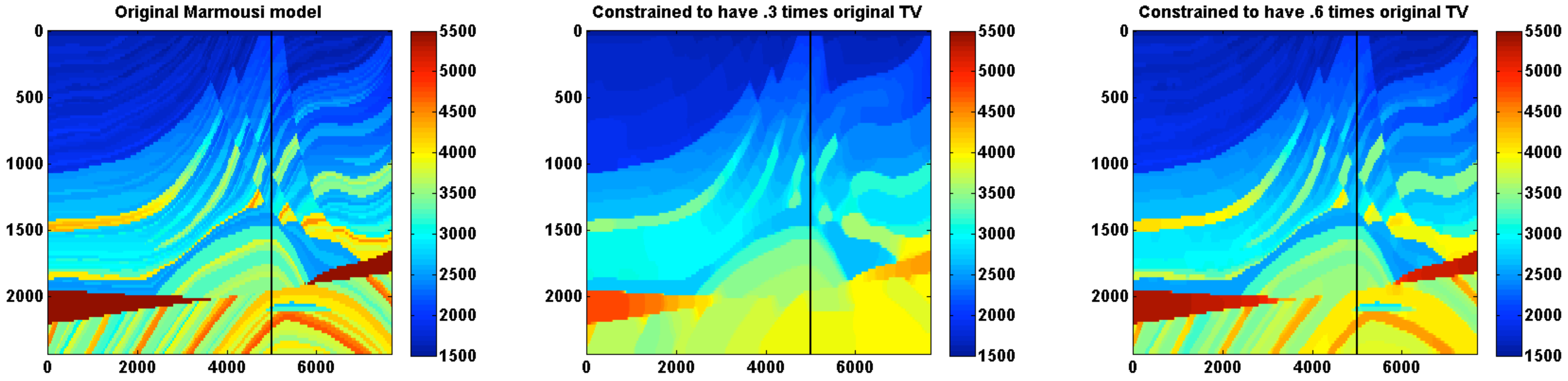
where $C_{\text{TV}} = \{\mathbf{m} : \|\mathbf{m}\|_{\text{TV}} \leq \tau\}$ and

$$\begin{aligned}\|\mathbf{m}\|_{TV} &= \frac{1}{h}\sum_{ij}\sqrt{(m_{i+1,j}-m_{i,j})^2 + (m_{i,j+1}-m_{i,j})^2} \\ &= \sum_{ij}\frac{1}{h}\left\|\begin{bmatrix}(m_{i,j+1}-m_{i,j}) \\ (m_{i+1,j}-m_{i,j})\end{bmatrix}\right\| \\ &= \|D\mathbf{m}\|_{1,2} := \sum_{l=1}^{N}\|(D\mathbf{m})_l\| \ .\end{aligned}$$

Projections onto convex sets

$v_{\min} = 1500,\ v_{\max} = 5500,$ and $\tau = \{0.3\tau_0,\ 0.6\tau_0\}$

$$\Pi_C(\mathbf{m}_0) = \arg\min_{\mathbf{m}} \frac{1}{2}\|\mathbf{m} - \mathbf{m}_0\|^2 \quad \text{subject to} \quad \mathbf{m}_i \in [B_i^l, B_i^u] \text{ and } \|\mathbf{m}\|_{TV} \leq \tau$$



Proposed algorithm

Solve

$$\underset{\mathbf{m}}{\text{minimize}}\, \Phi(\mathbf{m}) \quad \text{subject to} \quad \mathbf{m}^{n+1} \in C_{\text{box}} \cap C_{\text{TV}}$$

by iterating

$$\begin{aligned}\Delta\mathbf{m} = \underset{\Delta\mathbf{m}}{\arg\min}\, & \Delta\mathbf{m}^T \mathbf{g}^n + \frac{1}{2} \Delta\mathbf{m}^T (H^n + c_n \mathrm{I}) \Delta\mathbf{m} \\ & \text{subject to} \quad \mathbf{m}_i^n + \Delta\mathbf{m}_i \in [B_i^l, B_i^u] \text{ and } \|\mathbf{m}^n \Delta\mathbf{m}\|_{TV} \leq \tau \end{aligned}$$

$$\mathbf{m}^{n+1} = \mathbf{m}^n + \Delta\mathbf{m}$$

Solving the convex subproblems

Find saddle point of

$$\mathcal{L}(\Delta\mathbf{m},\mathbf{p}) = \Delta\mathbf{m}^T\mathbf{g}^n + \frac{1}{2}\Delta\mathbf{m}^T(H^n + c_n\mathrm{I})\Delta\mathbf{m} + g_B(\mathbf{m}^n + \Delta\mathbf{m}) + \mathbf{p}^T D(\mathbf{m}^n + \Delta\mathbf{m}) - \tau\|\mathbf{p}\|_{\infty,2}$$

with indicator functions for

Bound constraint

$$g_B(\mathbf{m}) = \begin{cases} 0 & \text{if } m_i \in [B_i^l, B_i^u] \\ \infty & \text{otherwise} \end{cases}$$

TV-norm constraint

$$\sup_{\mathbf{p}} +\mathbf{p}^T D(\mathbf{m}^n + \Delta\mathbf{m}) - \tau\|\mathbf{p}\|_{\infty,2} = \begin{cases} 0 & \text{if } \|D(\mathbf{m}^n + \Delta\mathbf{m})\|_{1,2} \leq \tau \\ \infty & \text{otherwise} \end{cases}$$

Ernie Esser, Xiaoqun Zhang, and Tony F. Chan. A General Frame- work for a Class of First Order Primal-Dual Algorithms for Convex Optimization in Imaging Science. SIAM Journal on Imaging Sciences, 3(4):1015–1046, 2010.

SLIM

Iterations
– primal dual hybrid gradient (PDHG)

projection onto TV ball

$$\mathbf{p}^{k+1} = \mathbf{p}^k + \delta D(\mathbf{m}^n + \Delta\mathbf{m}^k) - \Pi_{\|\cdot\|_{1,2}\leq\tau\delta}(\mathbf{p}^k + \delta D(\mathbf{m}^n + \Delta\mathbf{m}^k))$$

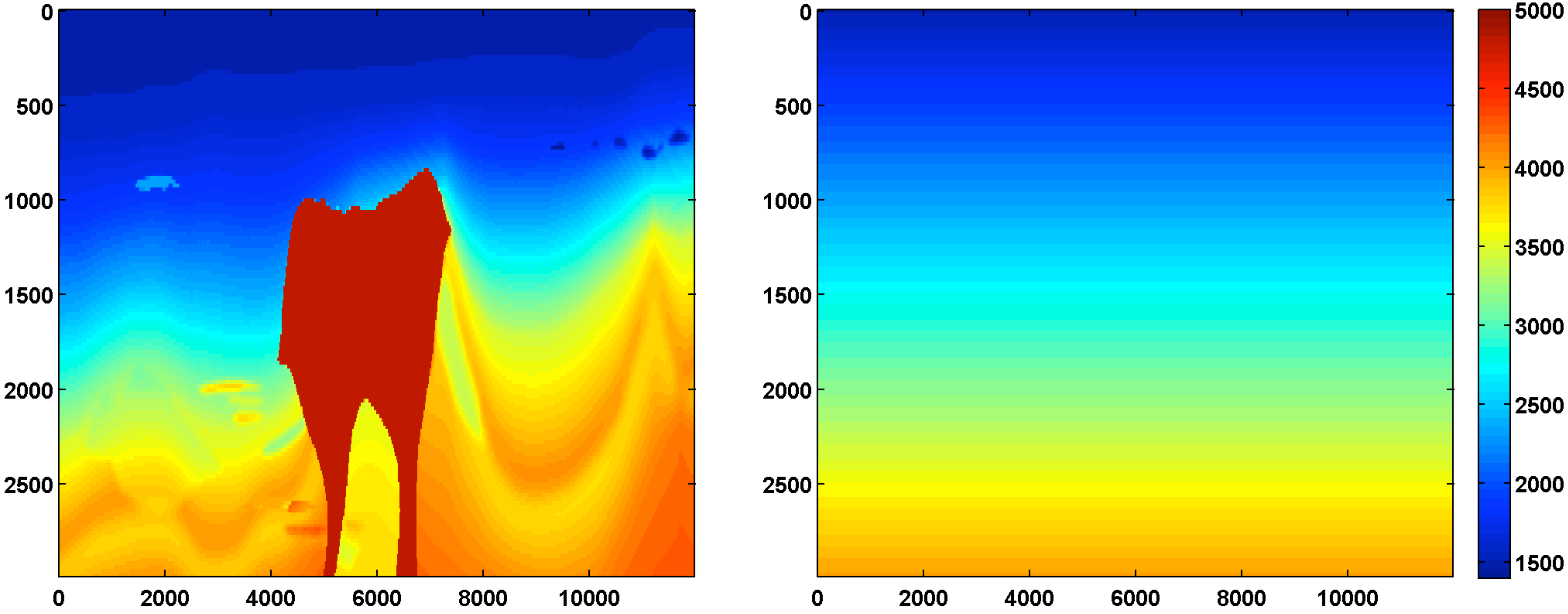
$$\Delta\mathbf{m}_i^{k+1} = \max\left((B_i^l - \mathbf{m}_i^n), B_i\right)$$

$$B_i = \min\left((B_i^u - \mathbf{m}_i^n), [(H^n + (c_n + \frac{1}{\alpha})\mathrm{I})^{-1}(-\mathbf{g}^n + \frac{\Delta\mathbf{m}^k}{\alpha} - D^T(2\mathbf{p}^{k+1} - \mathbf{p}^k)]_i\right)$$

for steplengths $\alpha\delta \leq \frac{1}{\|D^T D\|}$ and $\alpha = \frac{1}{\max(H^n + c_n\mathrm{I})}$

- do not involve solutions of (data-augmented) wave equations
- allows for data-dependent stepsizes

True velocity & poor starting model

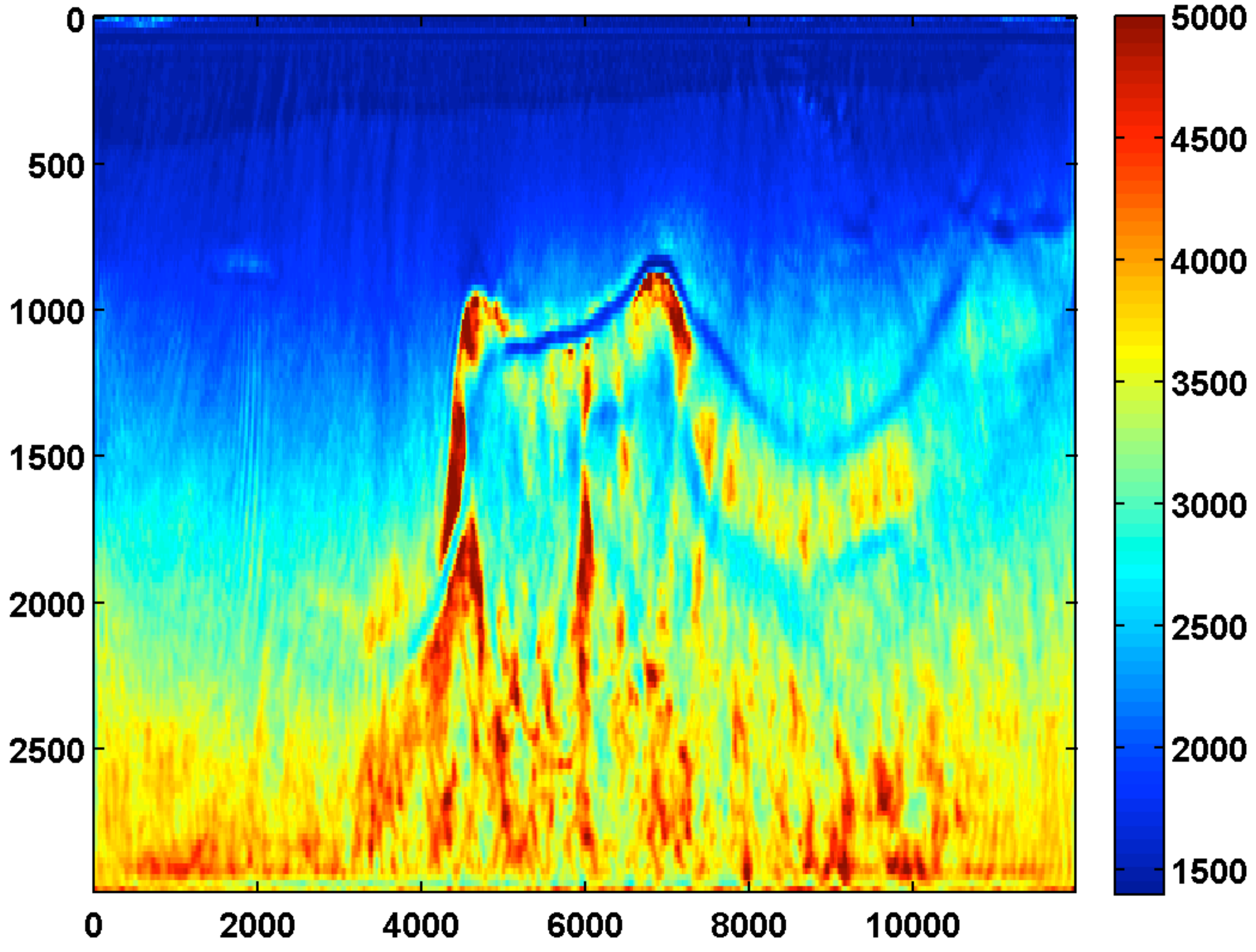


BP model

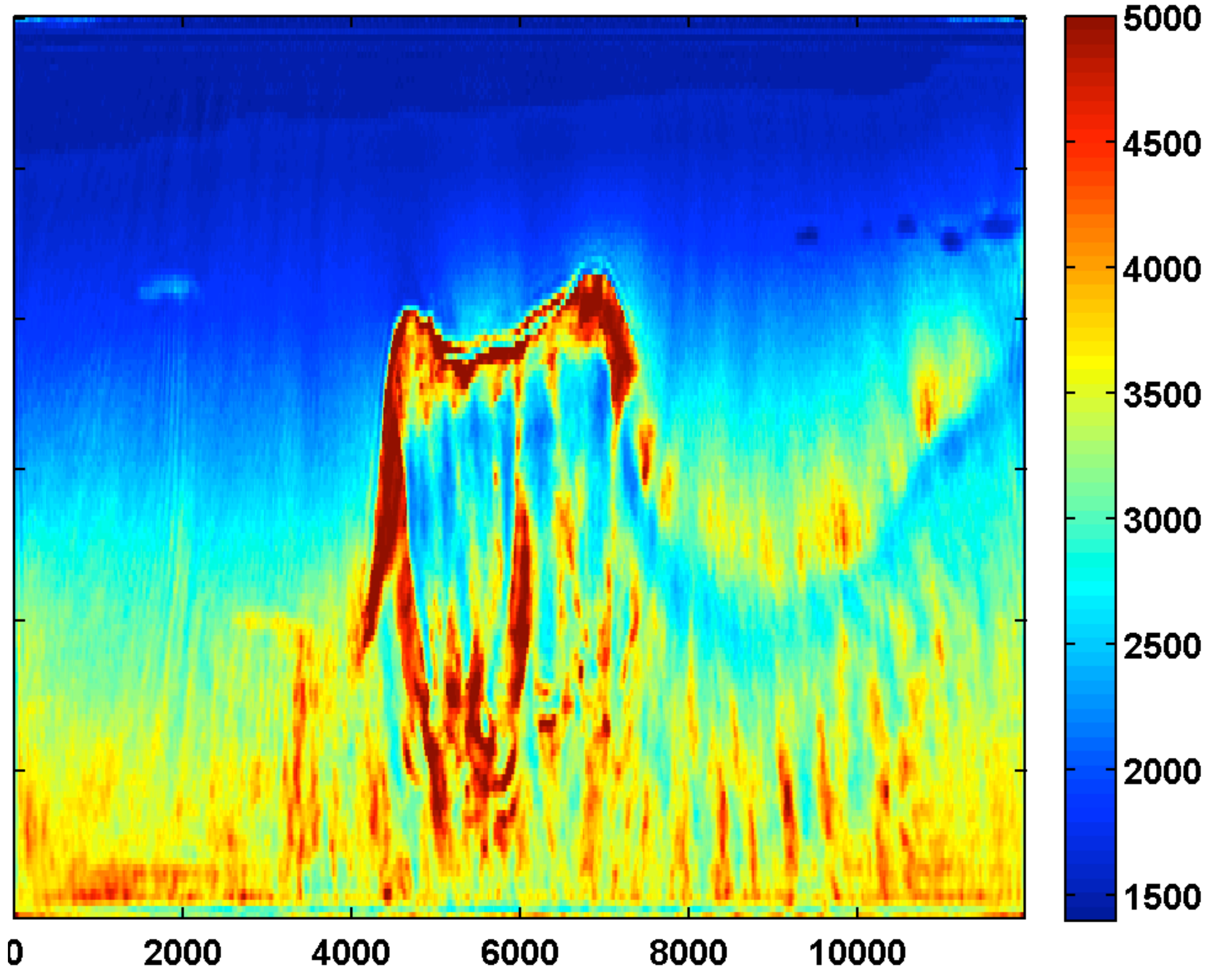
- number of sources: 126
- number of receivers: 299
- frequency continuation over 3-20Hz in overlapping batches of 2
- maximum number of outer iterations per frequency batch: 25
- maximum number of inner iterations for convex subproblems: 2000
- known Ricker wavelet sources with 15Hz peak frequency
-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

Results w/o TV

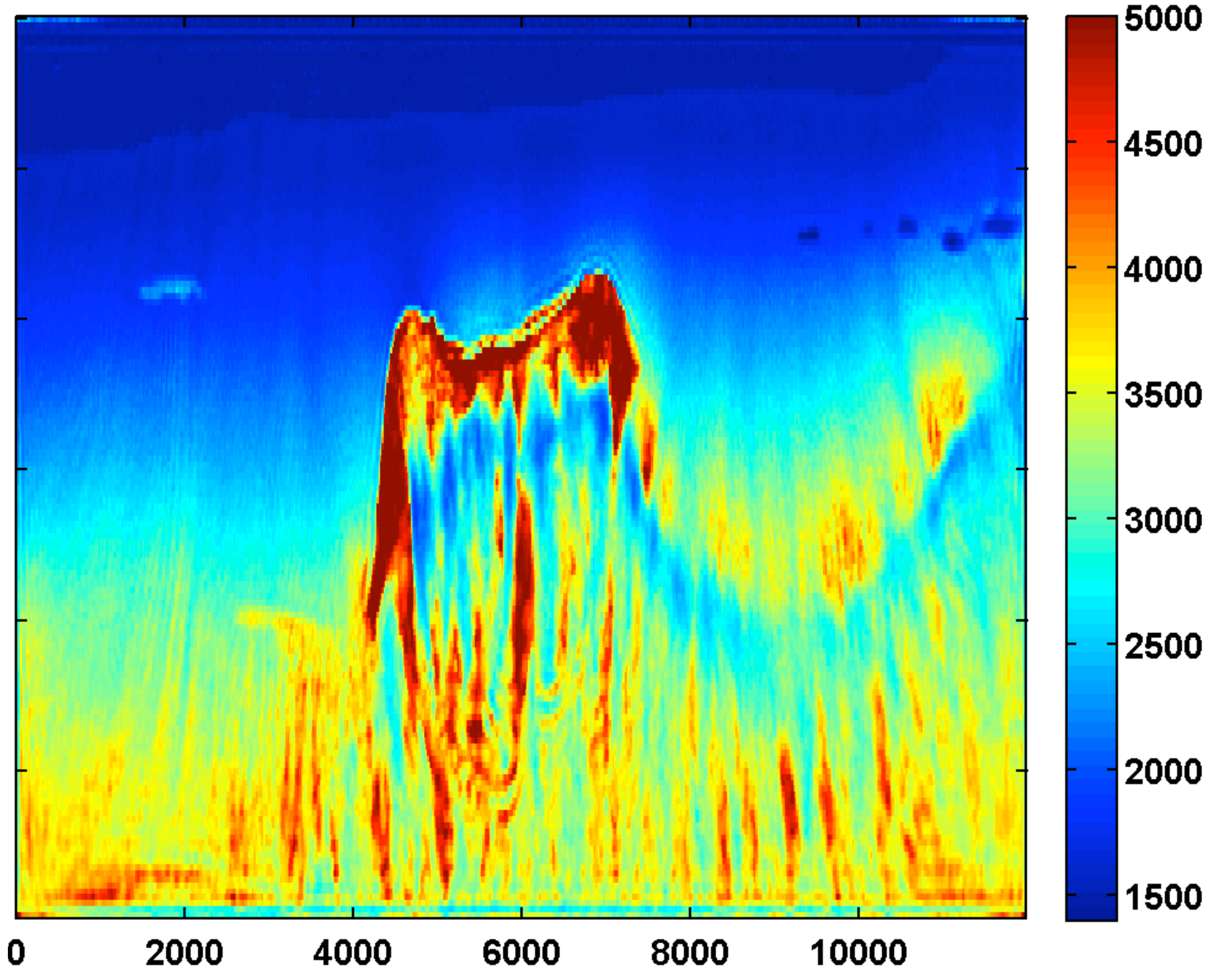
after one cycle through the frequencies



after two cycles through the frequencies

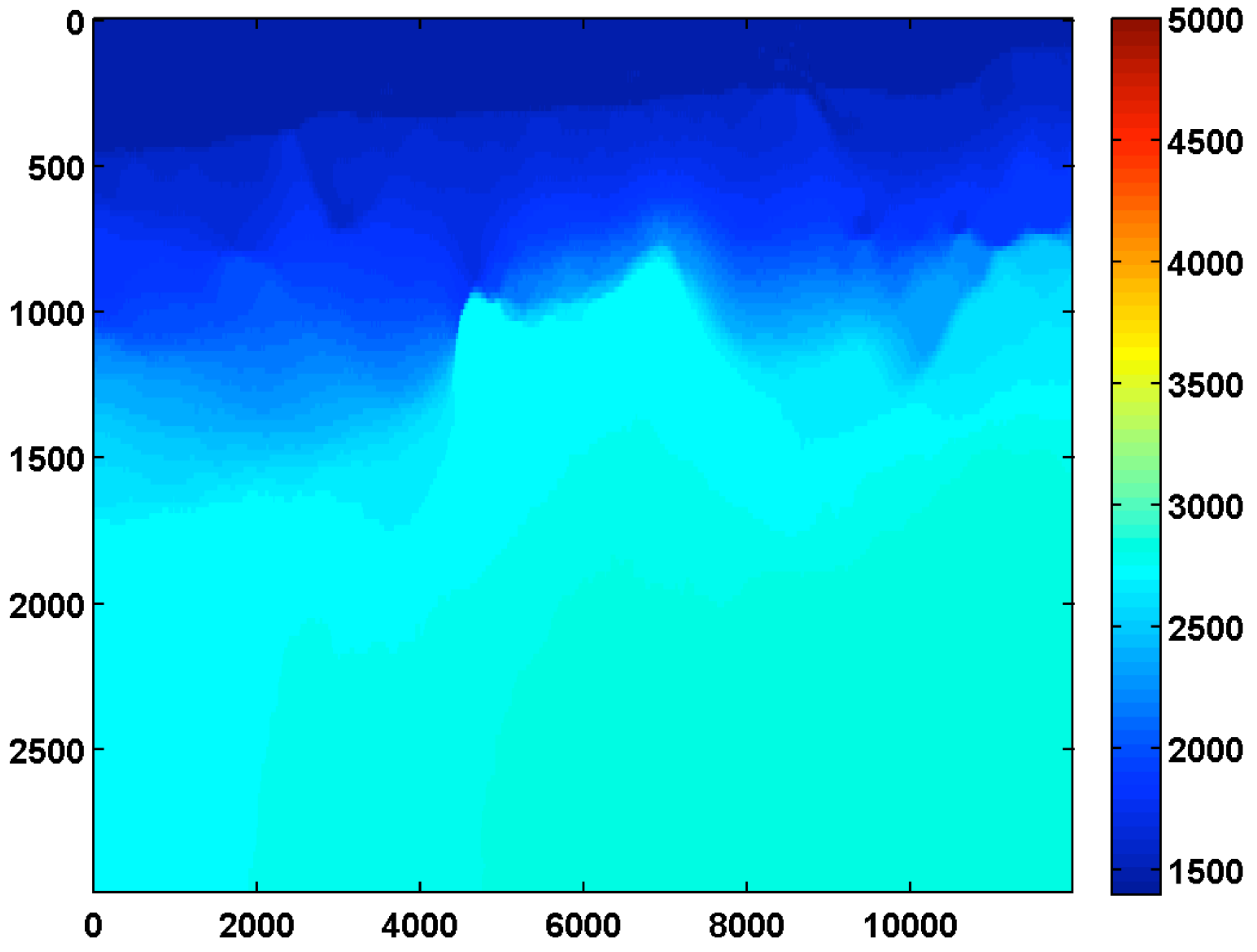


after three cycles through the frequencies

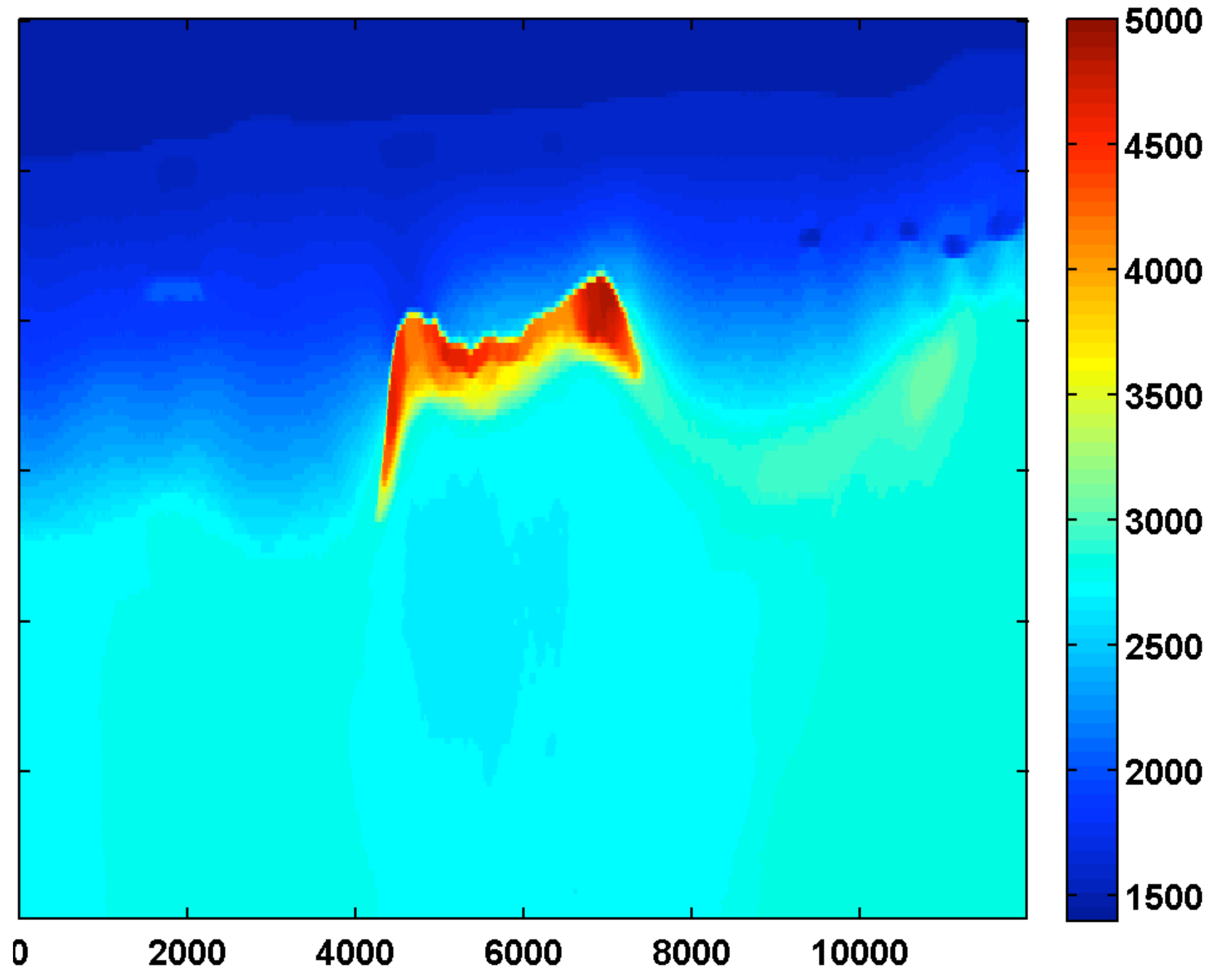


Results w/ TV

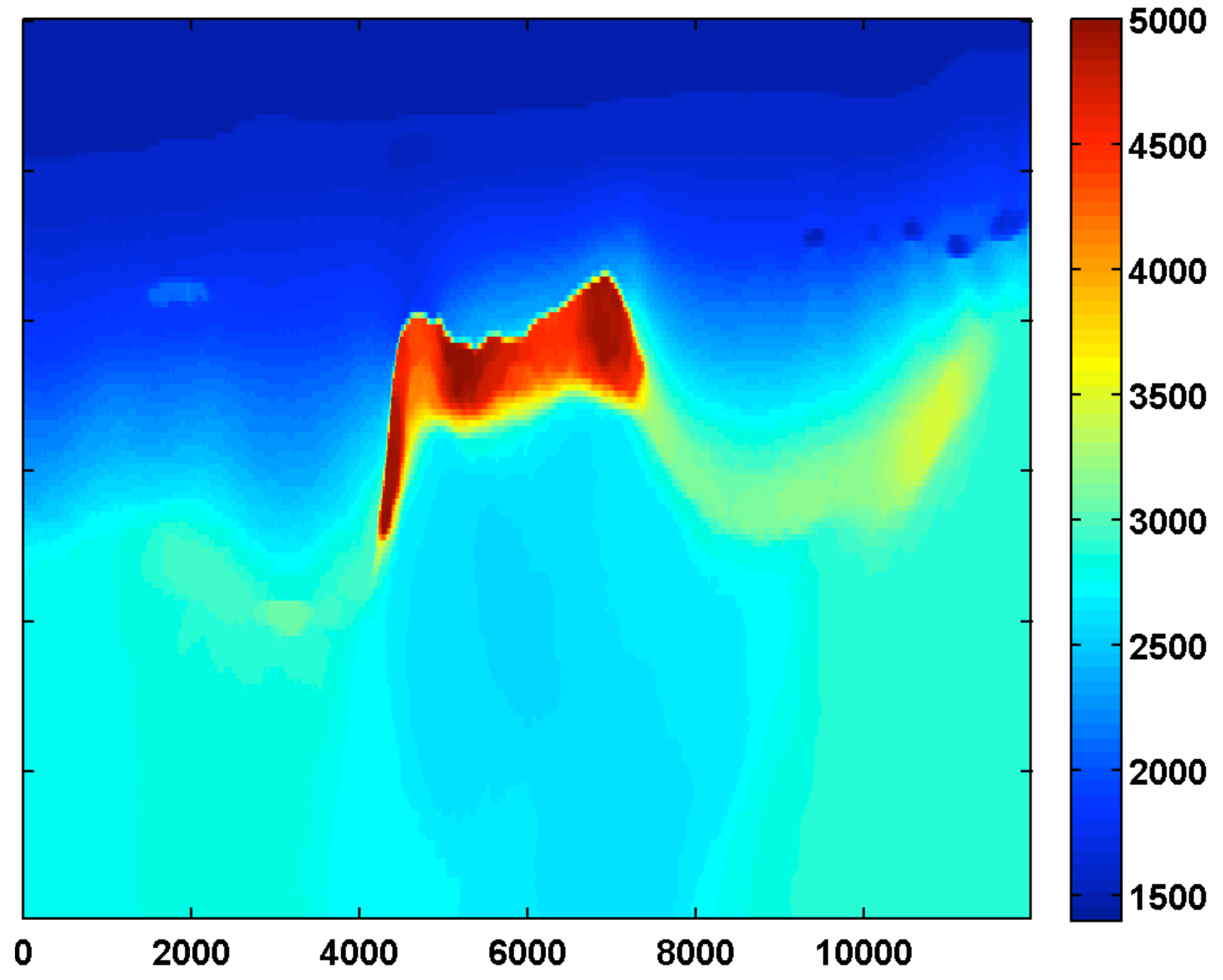
after one cycle through the frequencies



after two cycles through the frequencies



after three cycles through the frequencies



SLIM

Hinge loss
– one-sided TV constraint

Mitigate erroneous velocity model updates by using the fact that

- vertical slowness profiles tend to decrease w/ depth
- makes it less probable that velocities jump down along the vertical

Mathematically expressed as the one-norm of a hinge-loss function

$$\|\max(0, D_z\mathbf{m})\|_1 \leq \xi$$

- for ξ small slowness is unlikely to step up
- extended to a weighted directional gradient
- combined w/ omni-directional TV and bound constraints

Ernie Esser, Lluís Guasch, Tristan van Leeuwen, Aleksandr Y. Aravkin, and Felix J. Herrmann, "Total variation regularization strategies in full waveform inversion for improving robustness to noise, limited data and poor initializations". 2015.

SLIM

Scaled-gradient projections
– w/ convex total-variation, box, & hinge-loss constraints

Solve for given $\bar{\mathbf{u}}_\lambda$

$$\min_{\mathbf{m}} \phi(\mathbf{m}, \bar{\mathbf{u}}_\lambda) \quad \text{subject to} \quad \begin{cases} m_i \in [B_1, B_2] \\ \|\mathbf{m}\|_{\mathrm{TV}} \leq \tau \\ \|\mathbf{m}\|_{\mathrm{Hinge}} \leq \xi \end{cases}$$

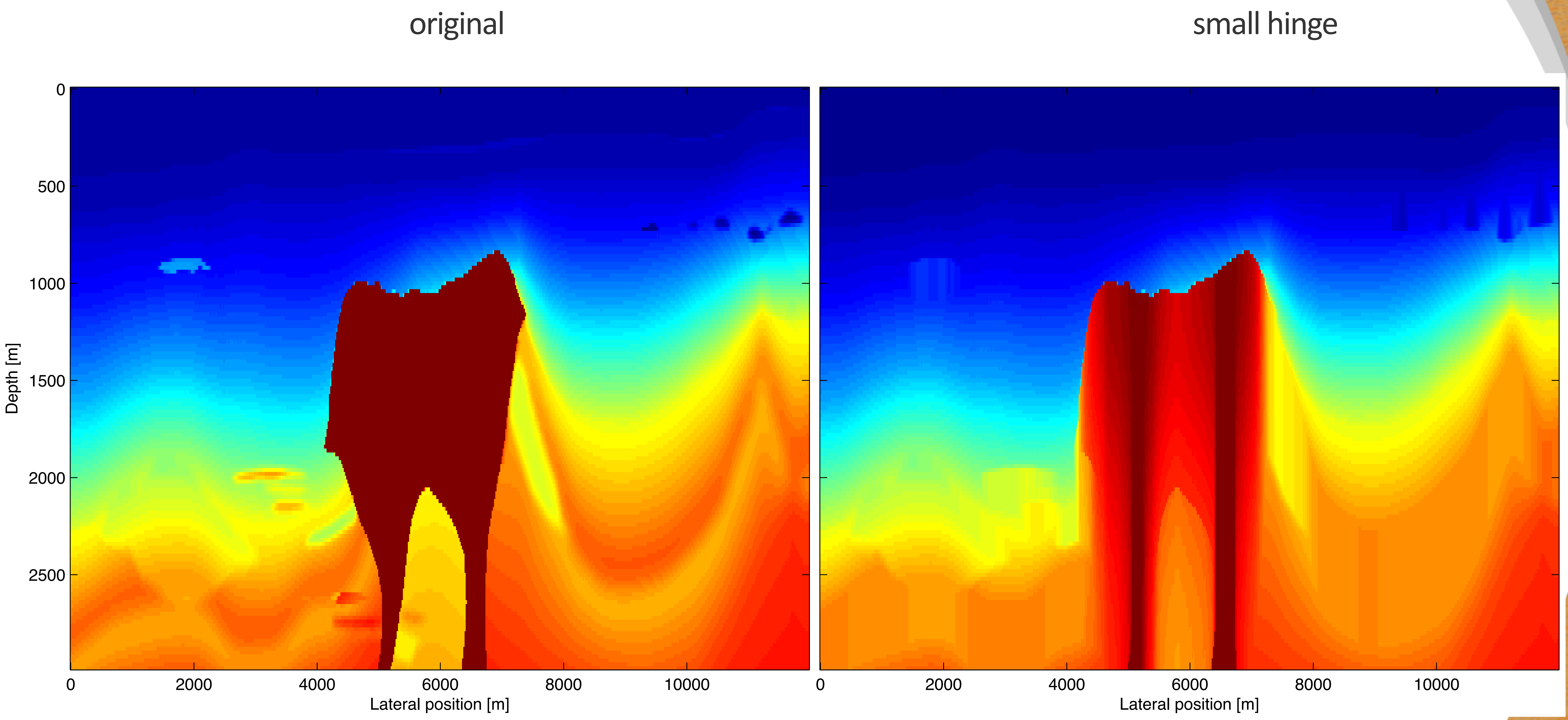
with

$$\|\mathbf{m}\|_{TV} = \sum_{ij} \frac{1}{h} \left\| \begin{bmatrix} (m_{i,j+1} - m_{i,j}) \\ (m_{i+1,j} - m_{i,j}) \end{bmatrix} \right\|$$

and

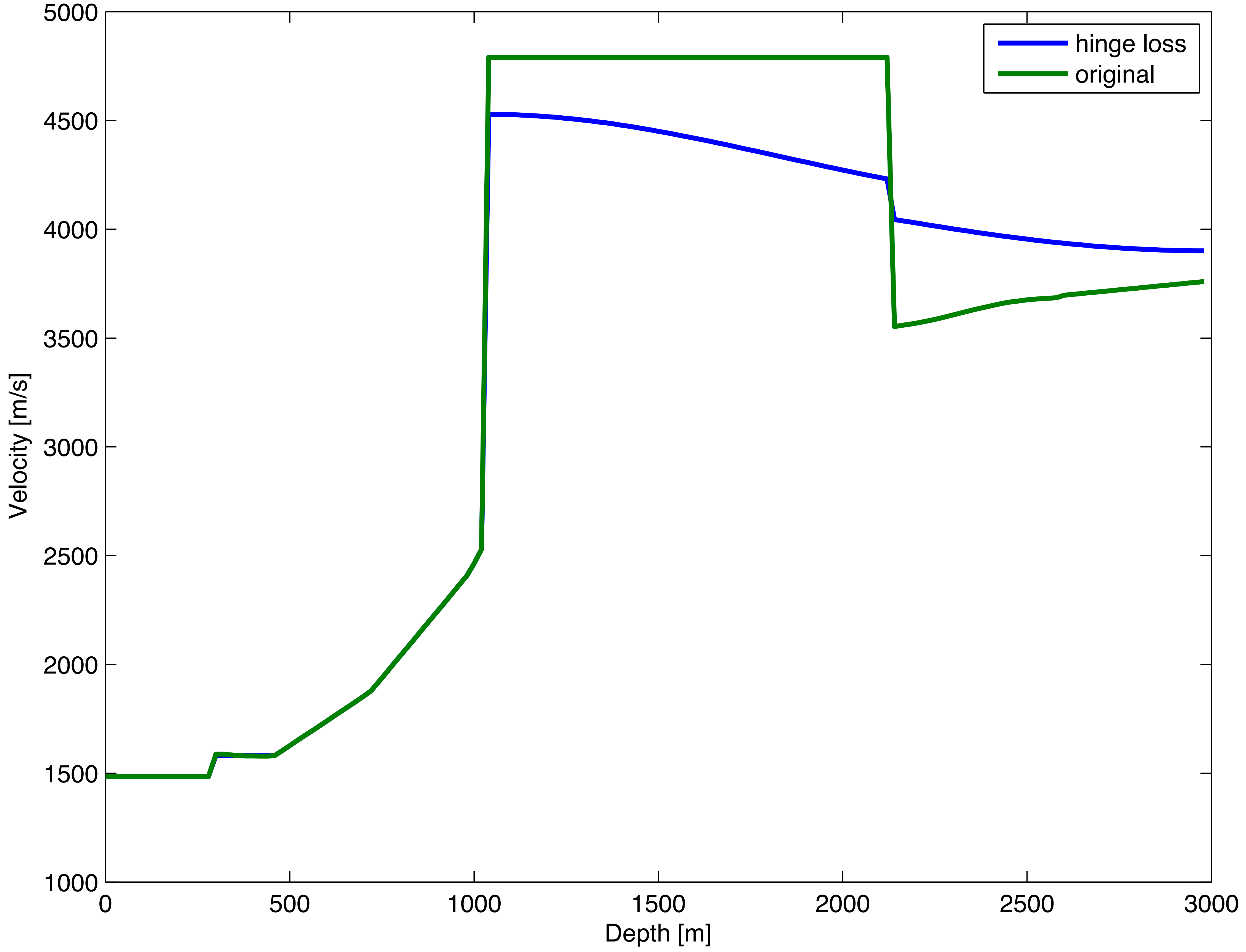
$$\|\mathbf{m}\|_{\mathrm{Hinge}} = \|\max(0, D_z\mathbf{m})\|_1$$

Hinge loss

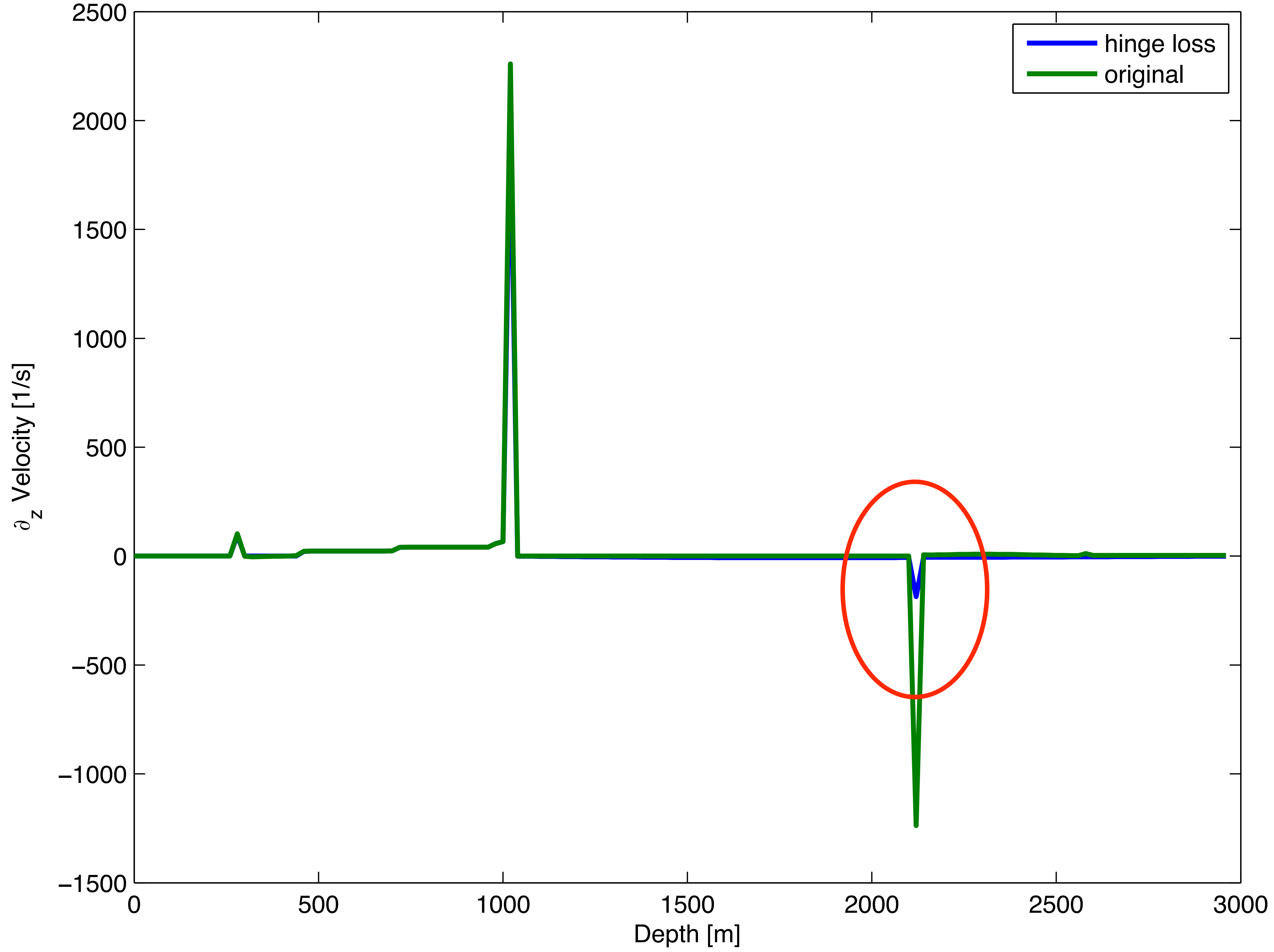


original
small hinge
Depth [m]
0
500
1000
1500
2000
2500
0
2000
4000
6000
8000
10000
Lateral position [m]
0
2000
4000
6000
8000
10000
Lateral position [m]

Hinge loss



Velocity [m/s]
Depth [m]
hinge loss
original
5000
4500
4000
3500
3000
2500
2000
1500
1000
0
500
1000
1500
2000
2500
3000



∂_z Velocity [1/s]
Depth [m]
hinge loss
original
2500
2000
1500
1000
500
0
−500
−1000
−1500
0
500
1000
1500
2000
2500
3000

Ernie Esser, Lluís Guasch, Tristan van Leeuwen, Aleksandr Y. Aravkin, and Felix J. Herrmann, "Total variation regularization strategies in full waveform inversion for improving robustness to noise, limited data and poor initializations". 2015.

Proposed algorithm

Solve

$$\underset{\mathbf{m}}{\text{minimize}}\, \Phi(\mathbf{m}) \quad \text{subject to} \quad \mathbf{m}^{n+1} \in C_{\text{box}} \cap C_{\text{TV}} \cap C_{\text{Hinge}}$$

by iterating

$$\mathbf{p}_1^{k+1} = \mathbf{p}_1^k + \delta D(\mathbf{m}^n + \Delta\mathbf{m}^k) - \Pi_{\|\cdot\|_{1,2}\leq\tau\delta}(\mathbf{p}_1^k + \delta D(\mathbf{m}^n + \Delta\mathbf{m}^k))$$

$$\mathbf{p}_2^{k+1} = \mathbf{p}_2^k + \delta D_z(\mathbf{m}^n + \Delta\mathbf{m}^k) - \Pi_{\|\max(0,\cdot)\|_1\leq\xi\delta}(\mathbf{p}_2^k + \delta D_z(\mathbf{m}^n + \Delta\mathbf{m}^k))$$

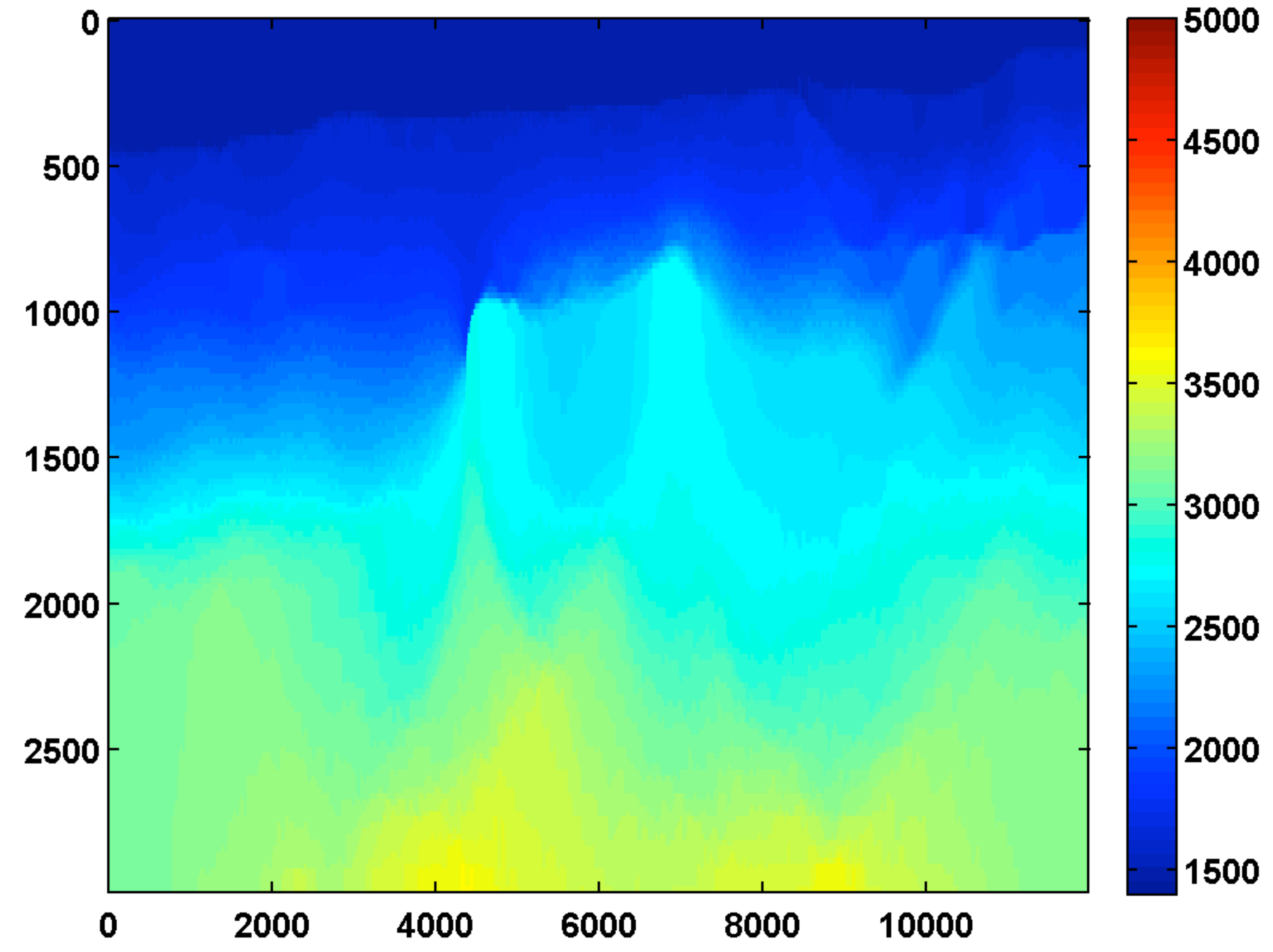
$$B_i = \min\left((B_i^u - \mathbf{m}_i^n), [(H^n + (c_n + \frac{1}{\alpha})\mathrm{I})^{-1}(-\mathbf{g}^n + \frac{\Delta\mathbf{m}^k}{\alpha} - D^T(2\mathbf{p}_1^{k+1} - \mathbf{p}_1^k) - D_z^T(2\mathbf{p}_2^{k+1} - \mathbf{p}_2^k))]_i\right)$$

$$\Delta\mathbf{m}_i^{k+1} = \max\left((B_i^l - \mathbf{m}_i^n), B_i\right)$$

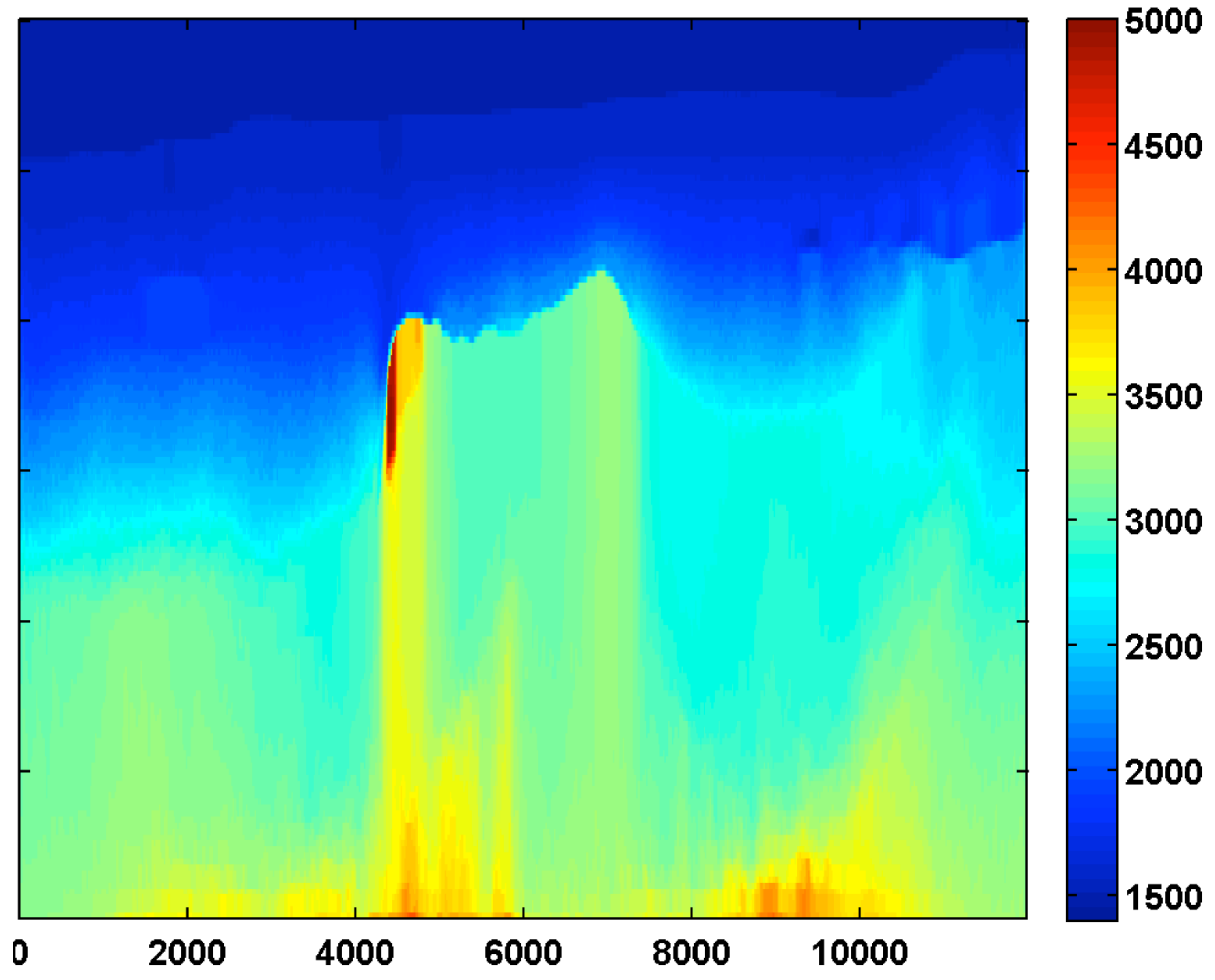
Results w/ hinge loss continuation

$$\frac{\xi}{\xi_{\text{true}}} = \{.01, .05, .10\}$$

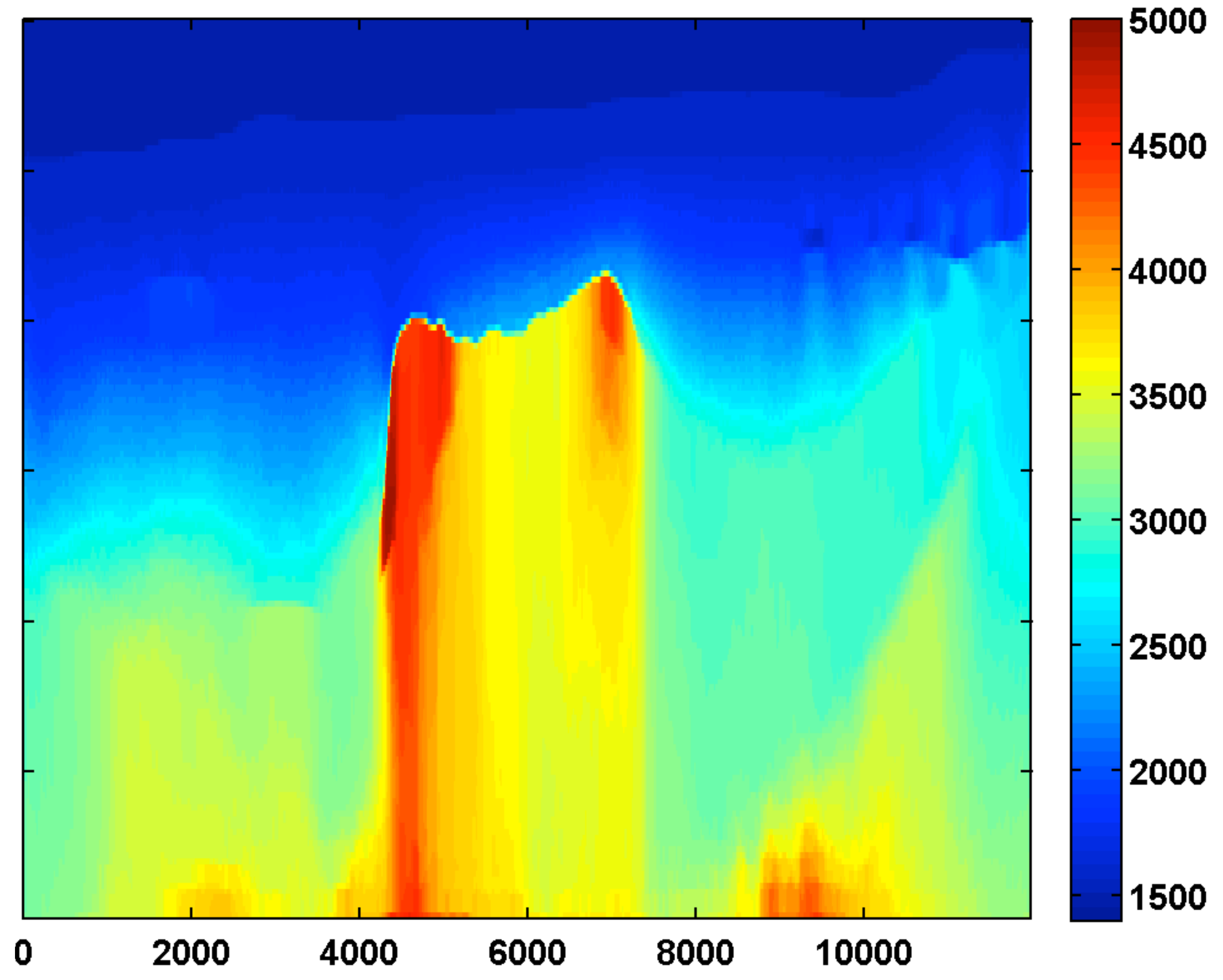
after one cycle through the frequencies



after two cycles through the frequencies



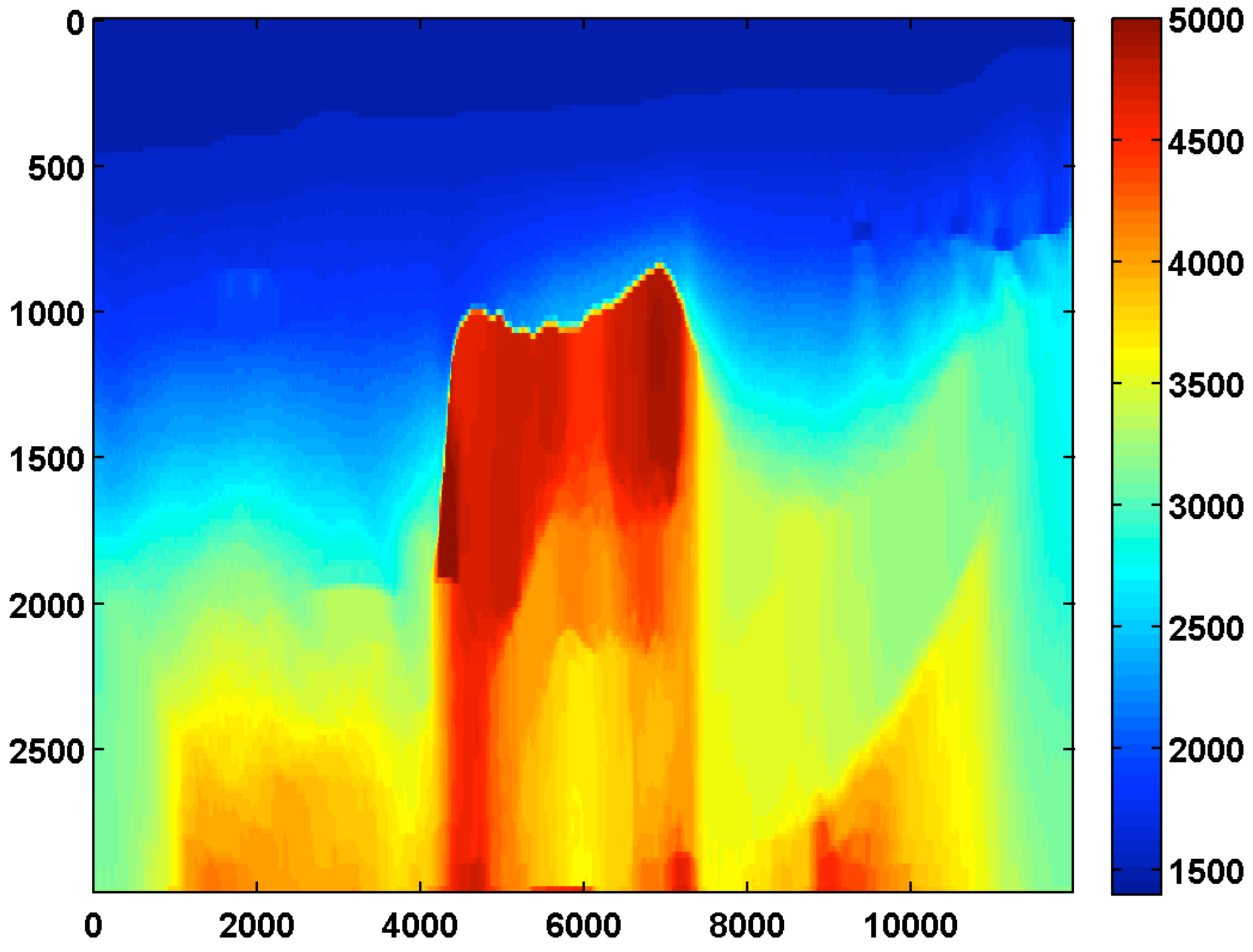
after three cycles through the frequencies



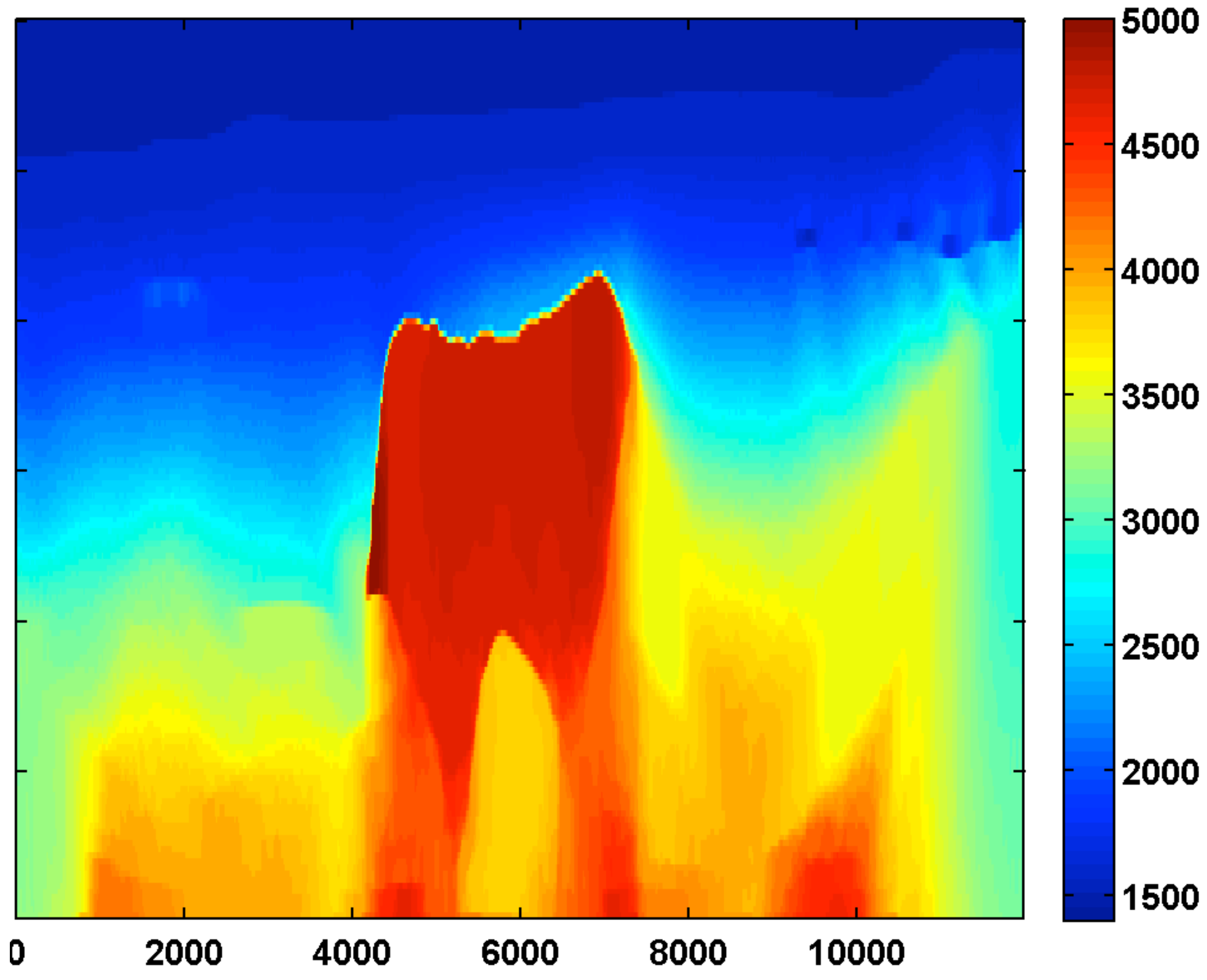
Results w/ hinge loss continuation

$$\frac{\xi}{\xi_{\text{true}}} = \{.15, .20, .25\}$$

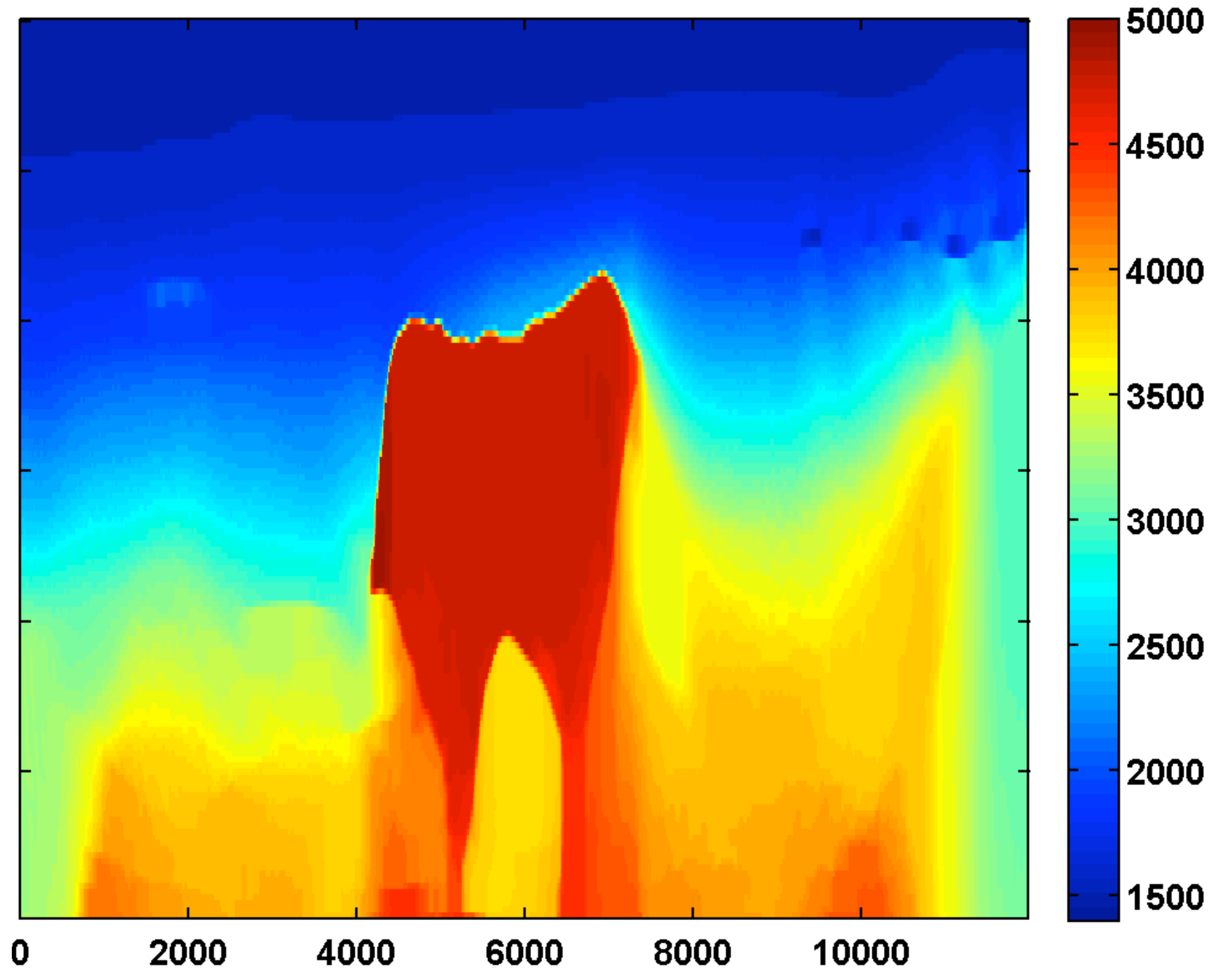
after four cycles through the frequencies



after five cycles through the frequencies

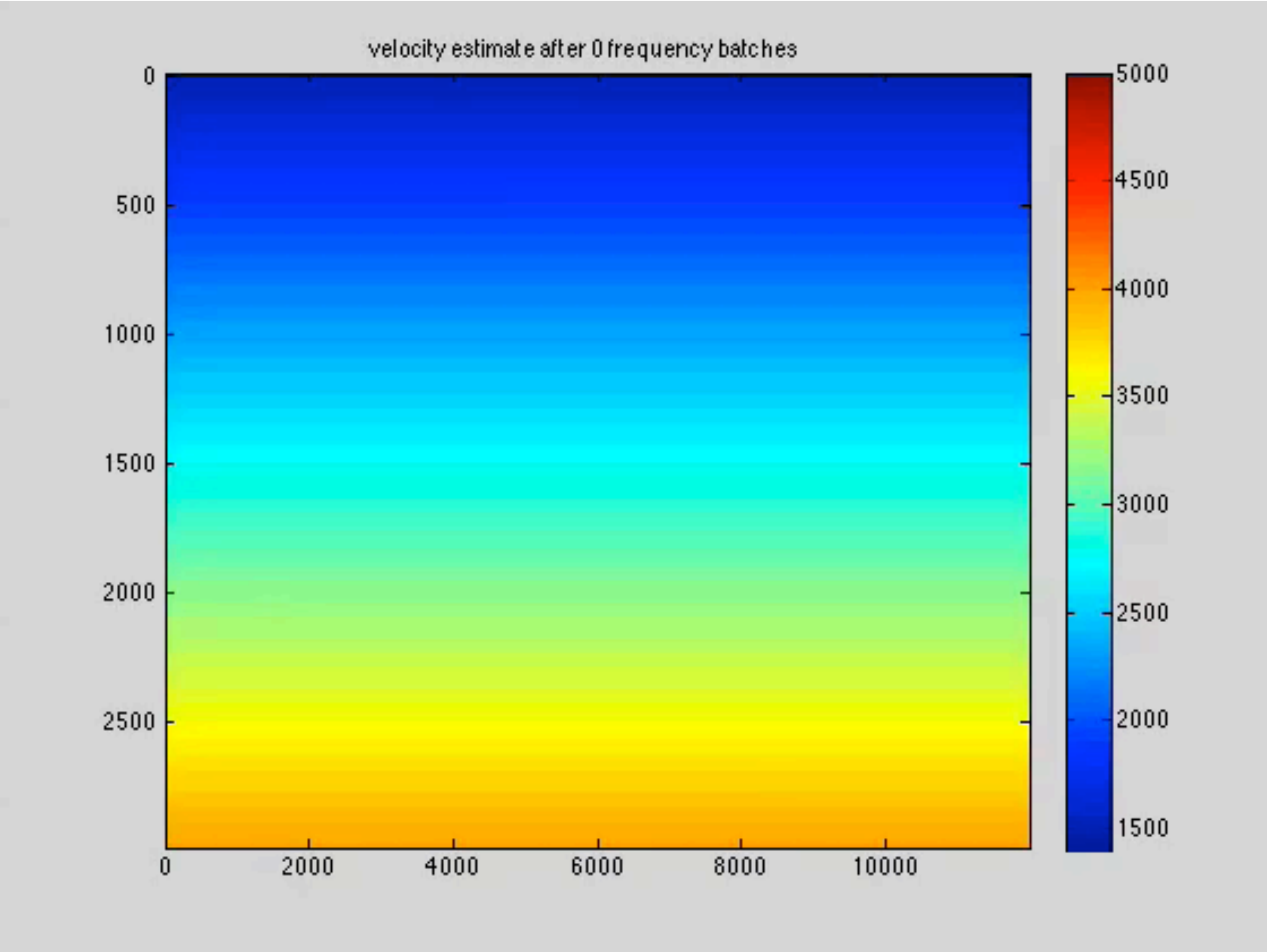


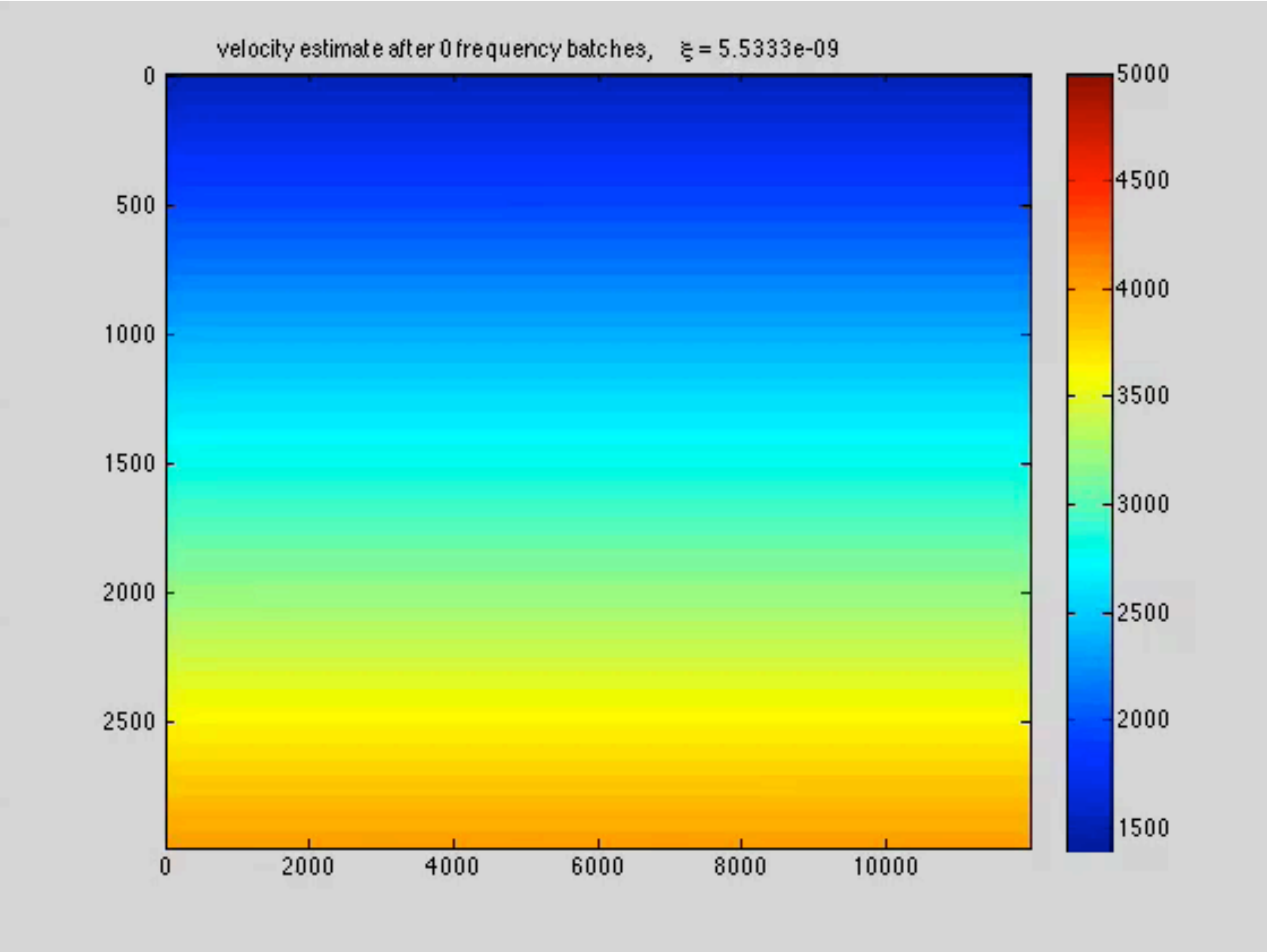
after six cycles through the frequencies



WRI

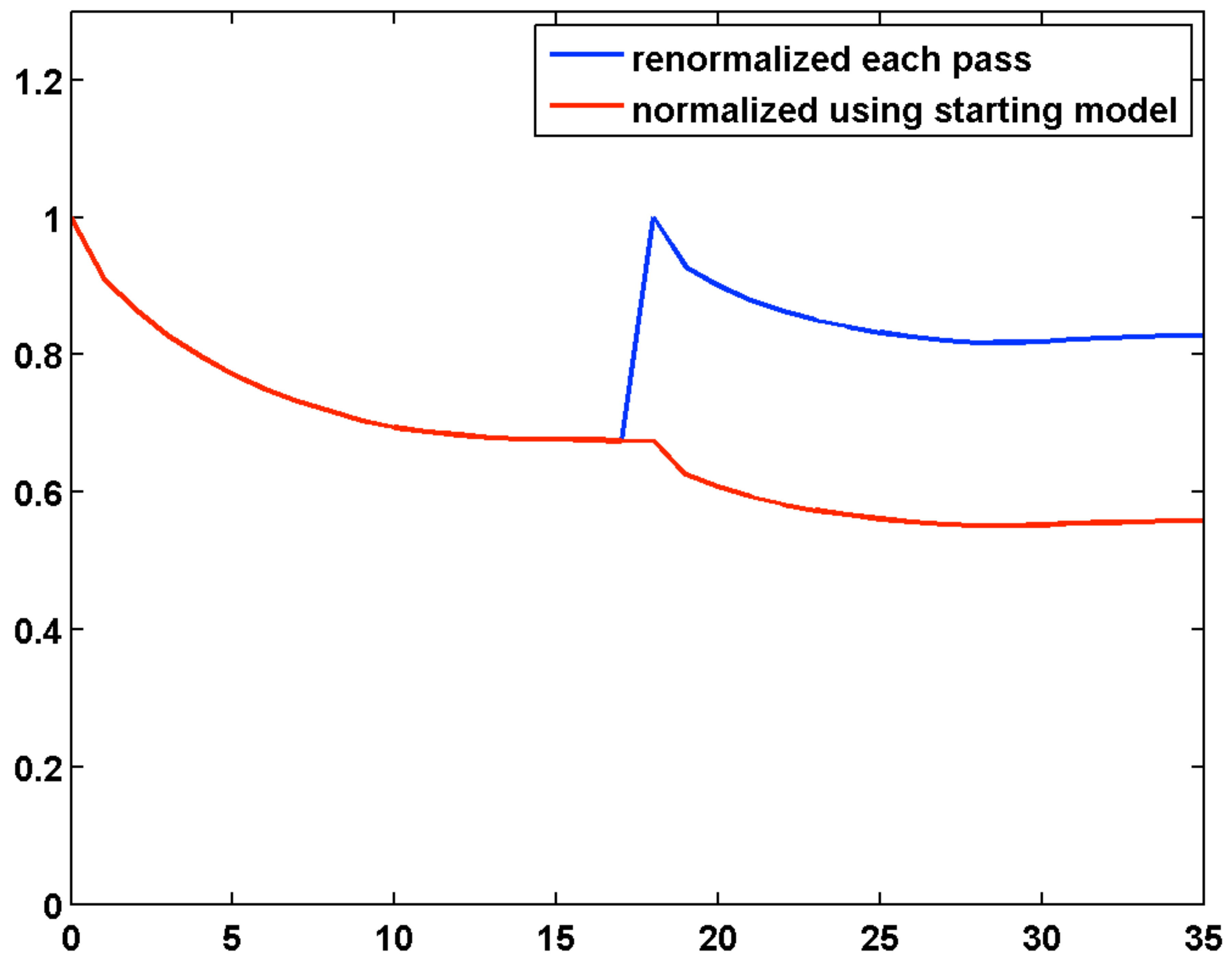
w/ or w/o TV-norm & hinge-loss projections & poor starting model



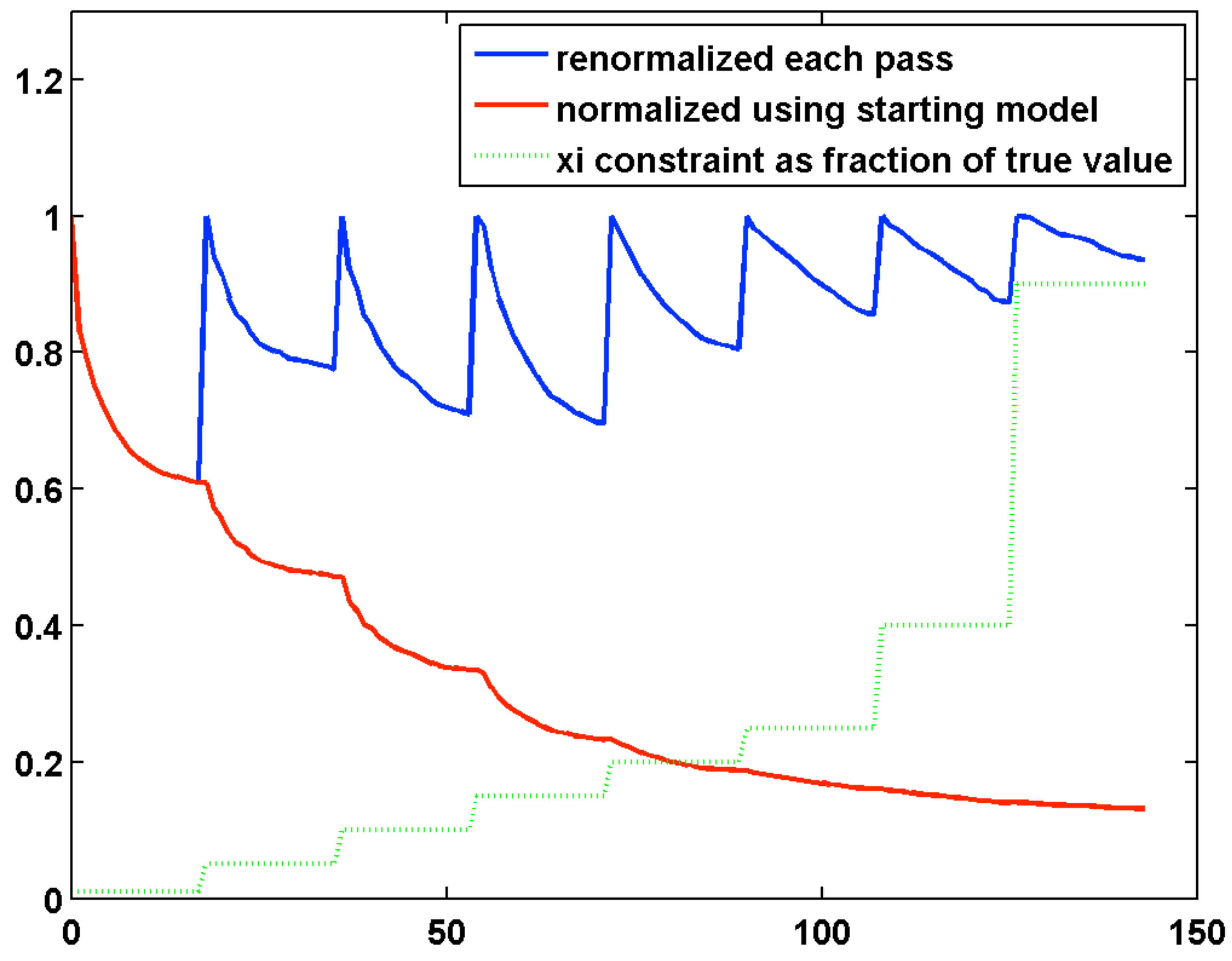


Relative model errors

w/o TV

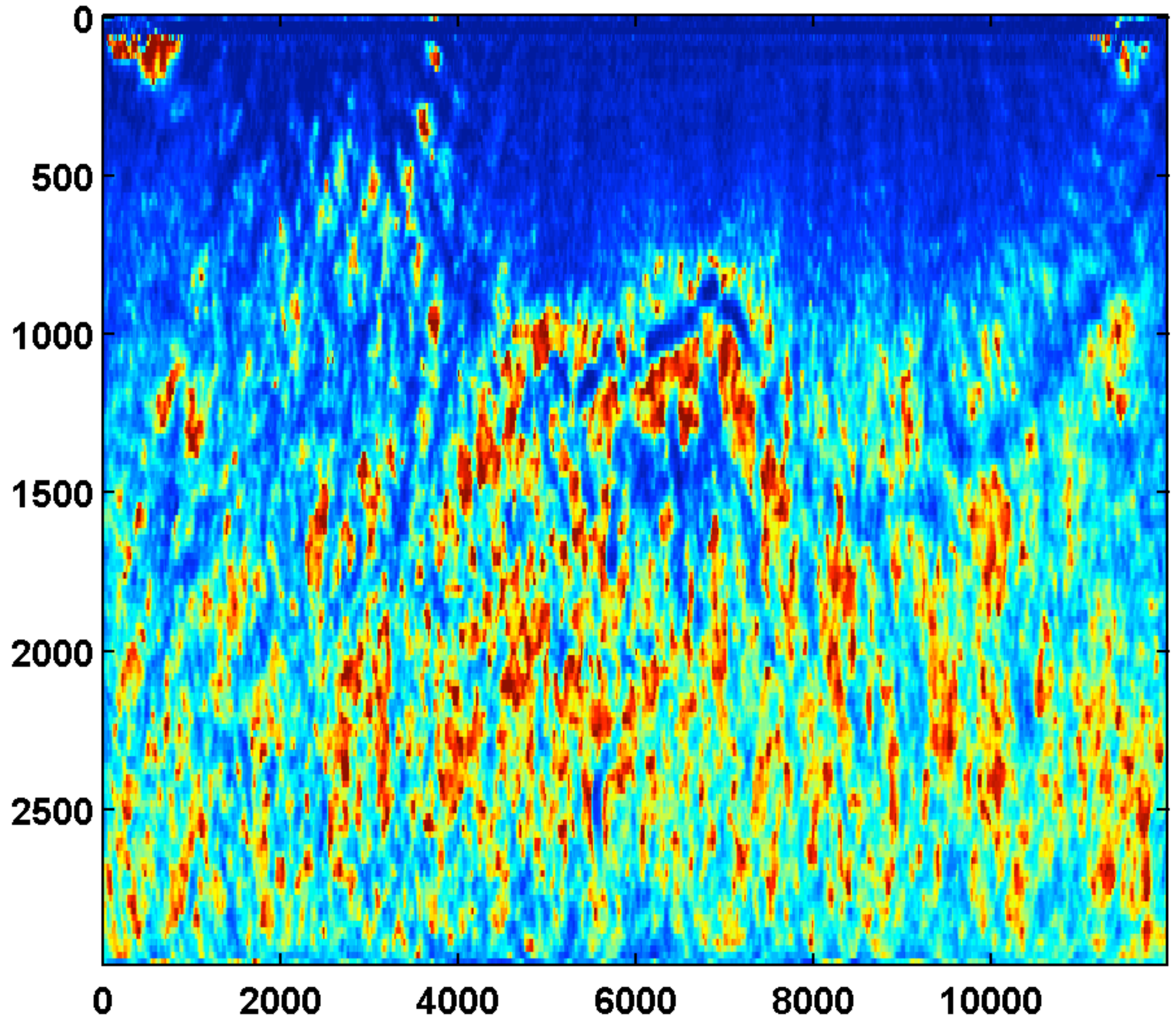


w/ TV & hinge

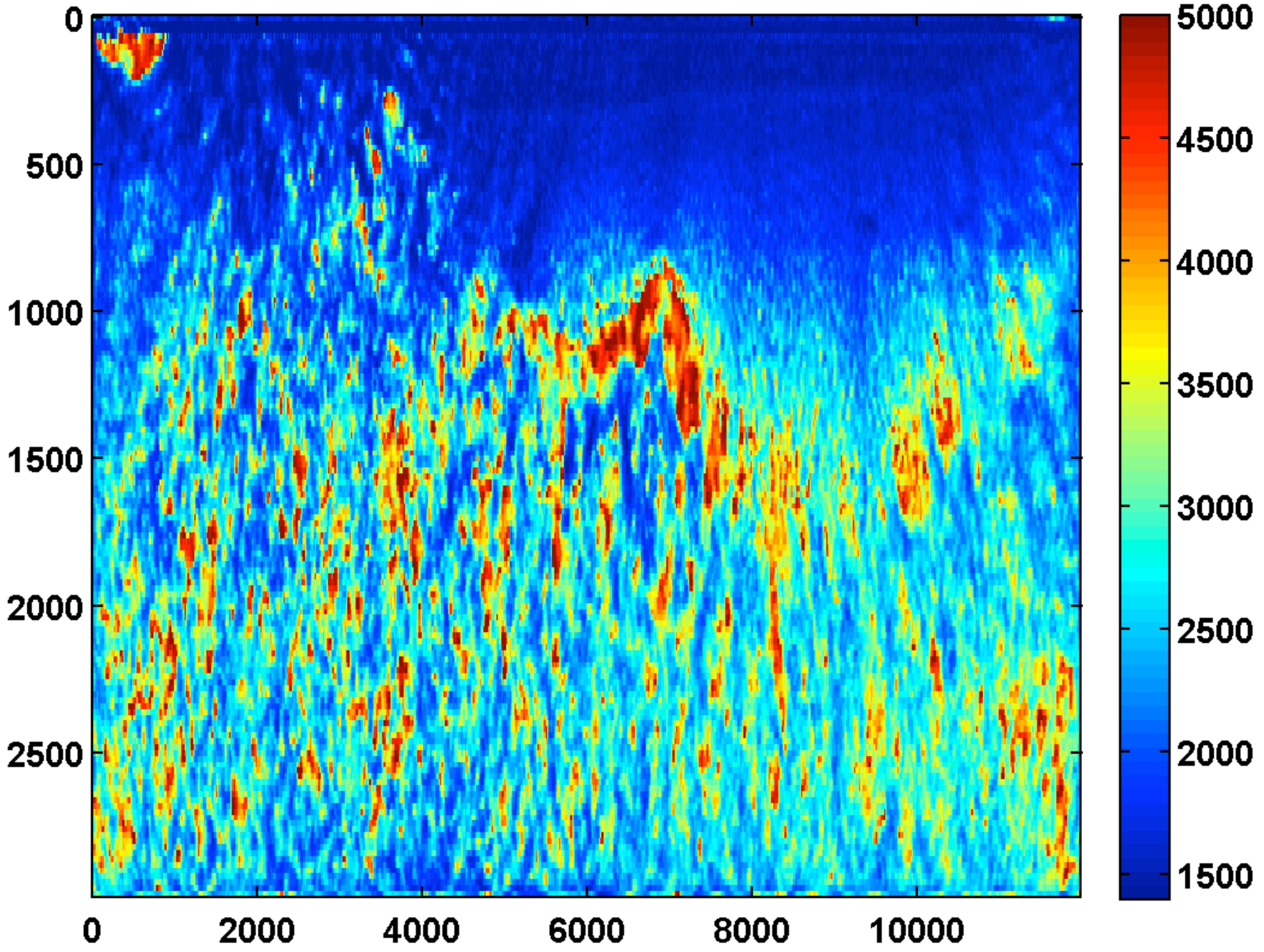


Adjoint-state w/o TV

After one cycle through the frequencies



After two cycles through the frequencies



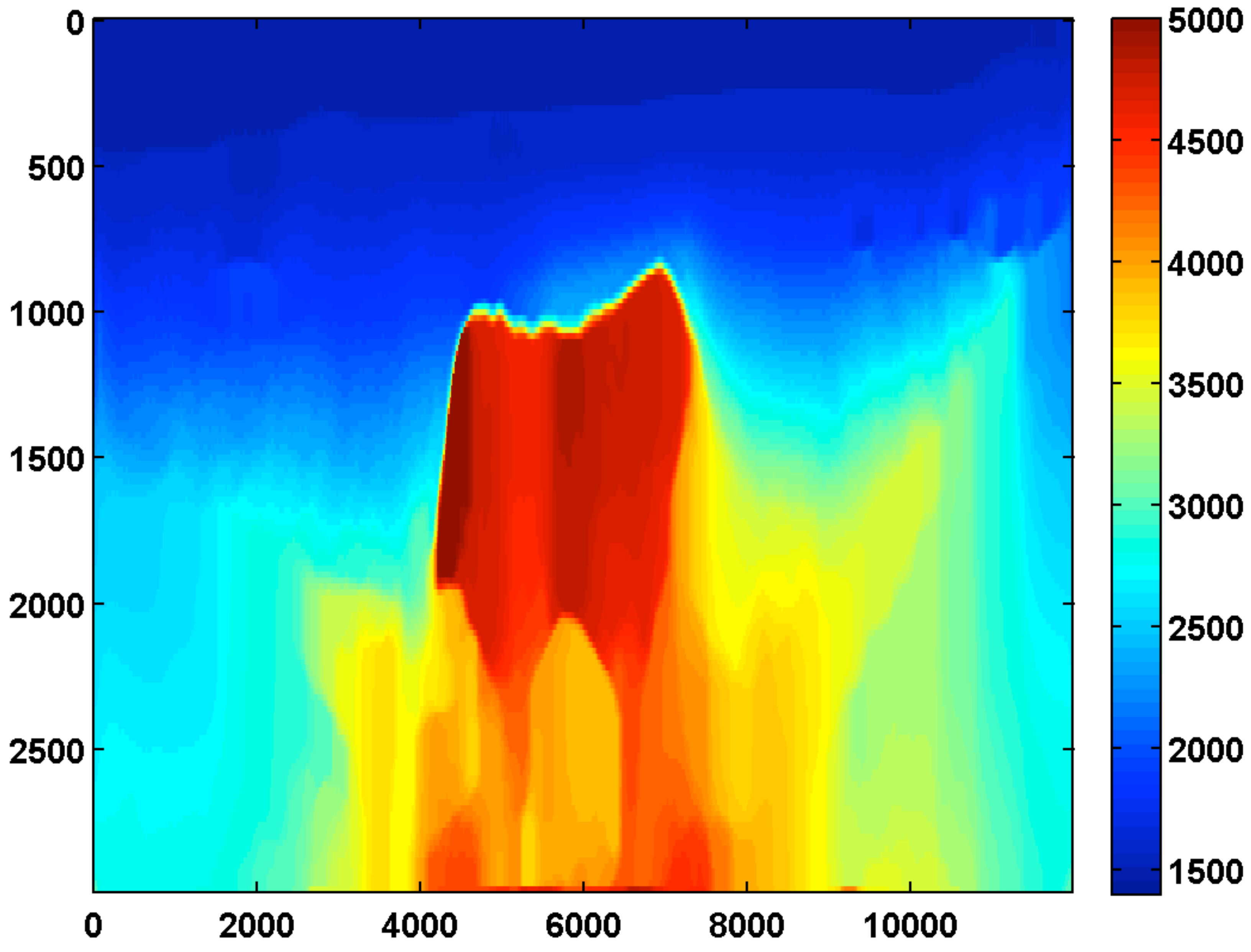
Adjoint-state w/ hinge loss continuation

$$\frac{\xi}{\xi_{\text{true}}} = \{.01, .05, .10\}$$

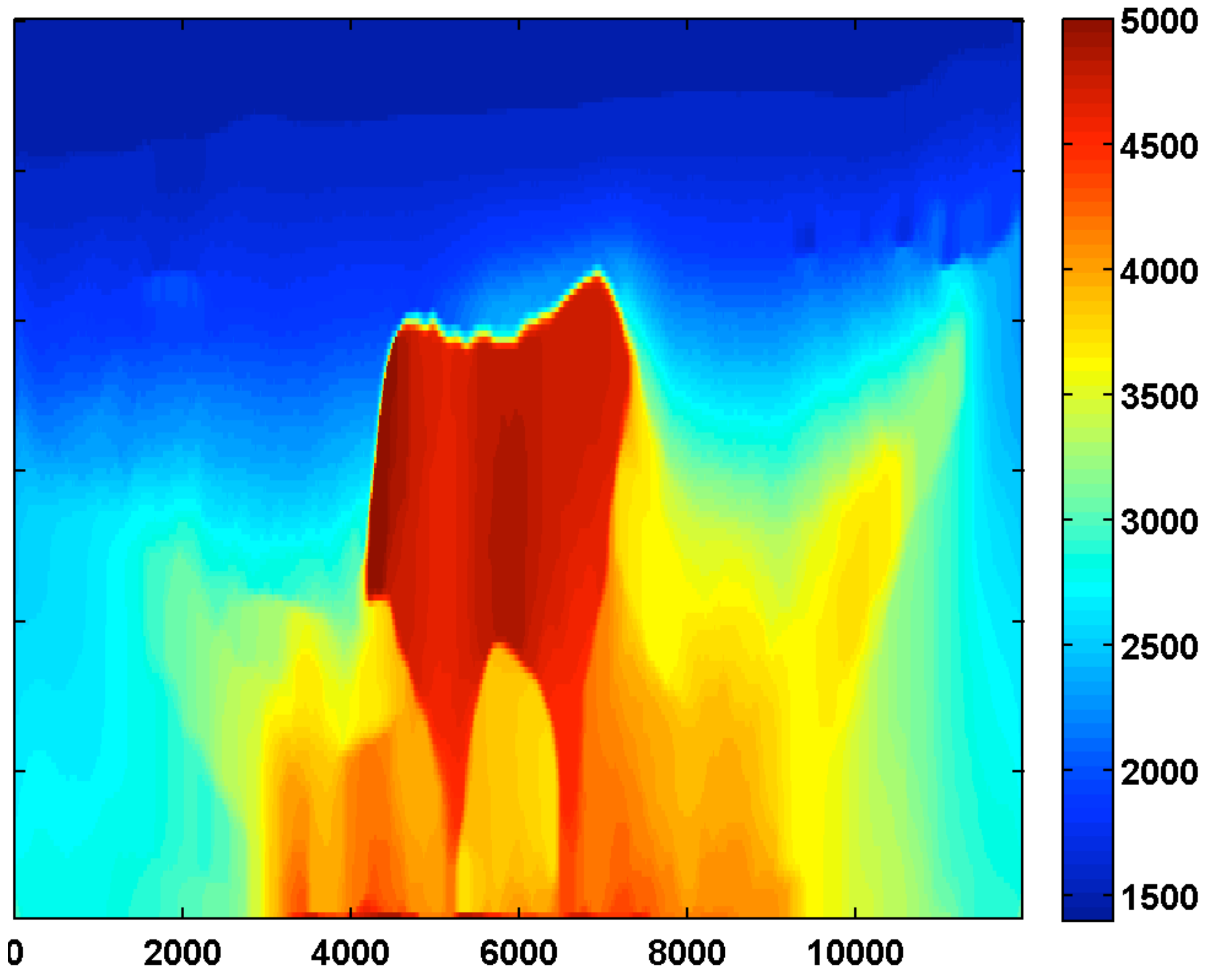
Adjoint-state w/ hinge loss continuation

$$\frac{\xi}{\xi_{\text{true}}} = \{.15, .20, .25\}$$

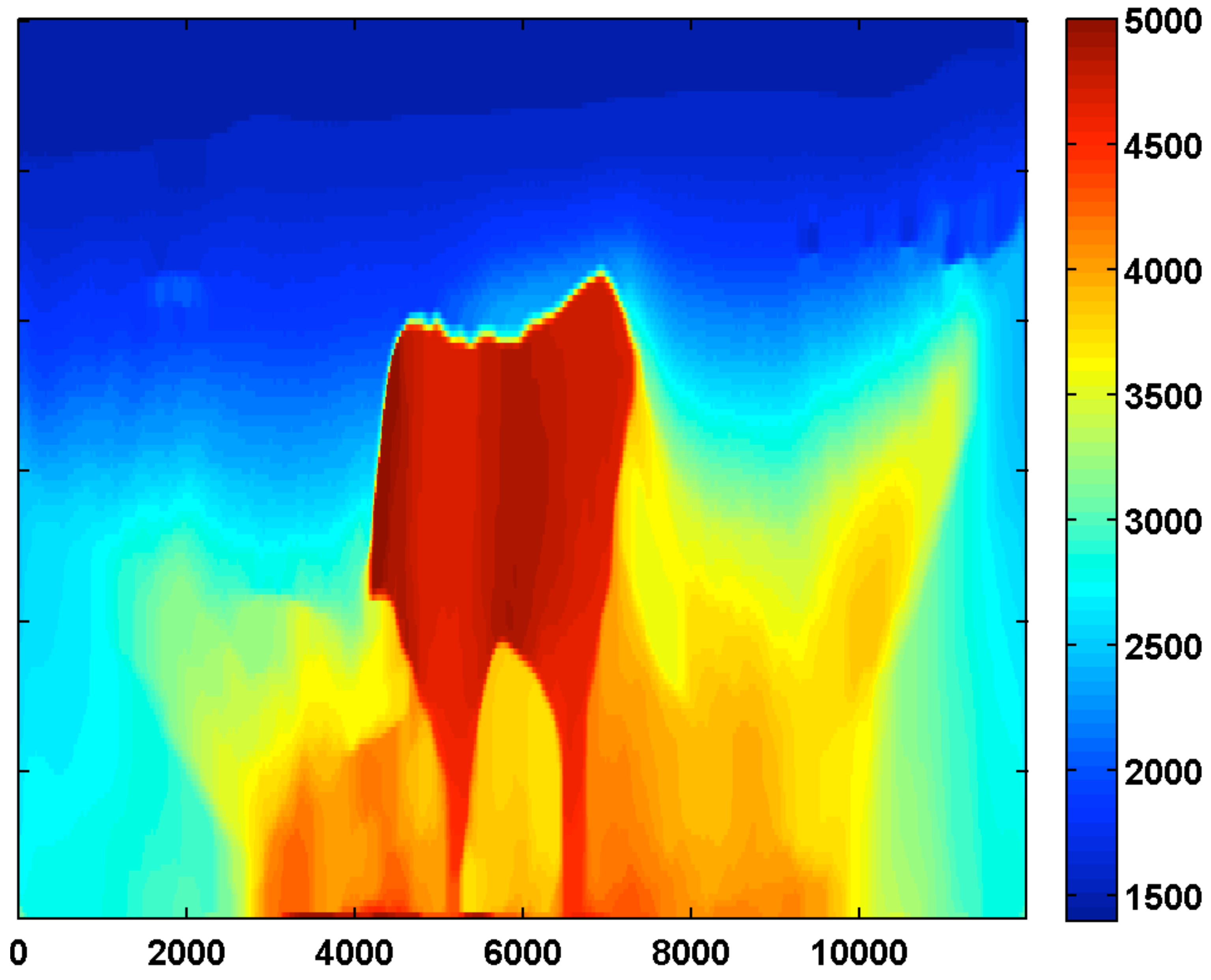
after four cycles through the frequencies



after five cycles through the frequencies

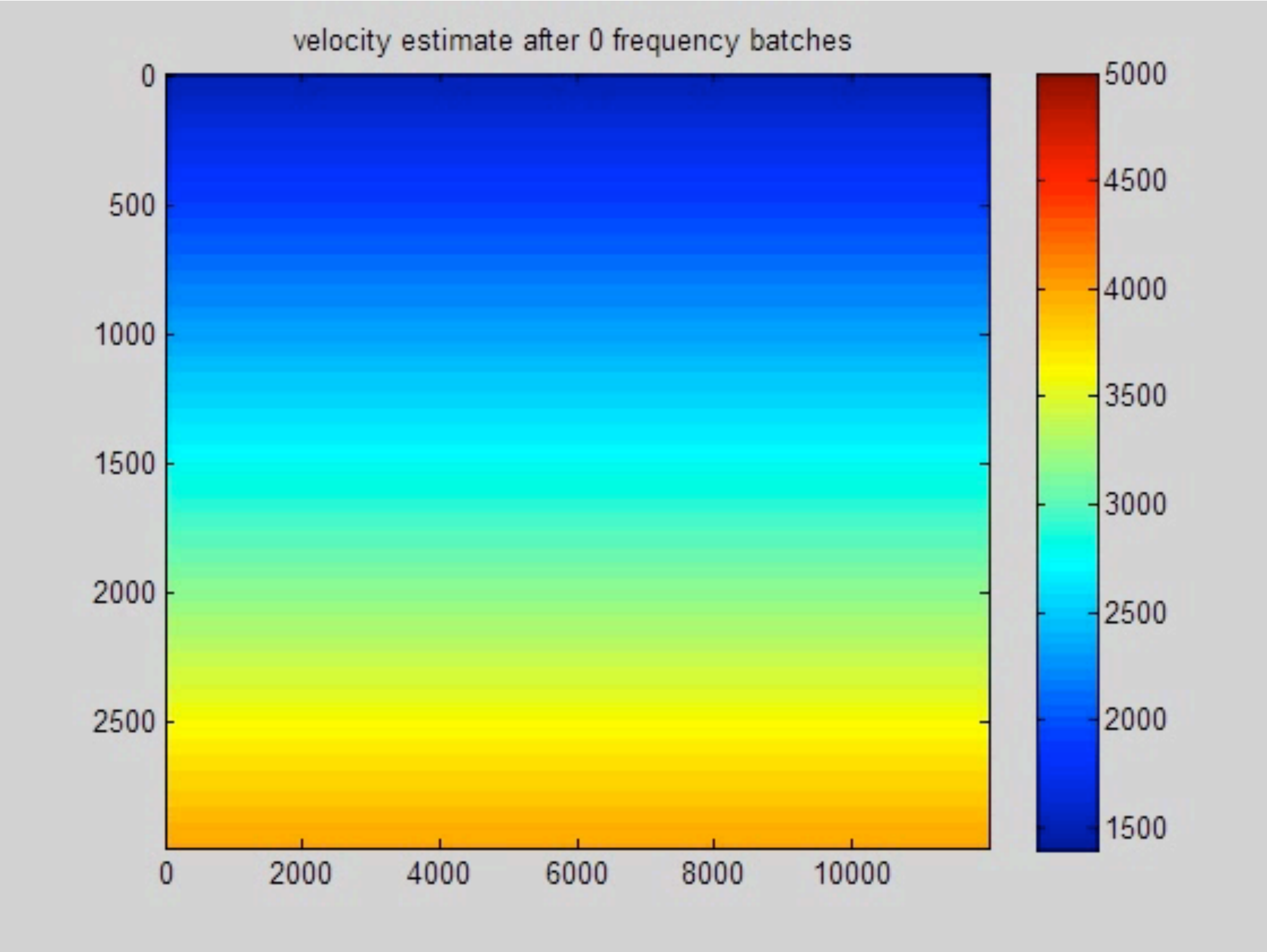


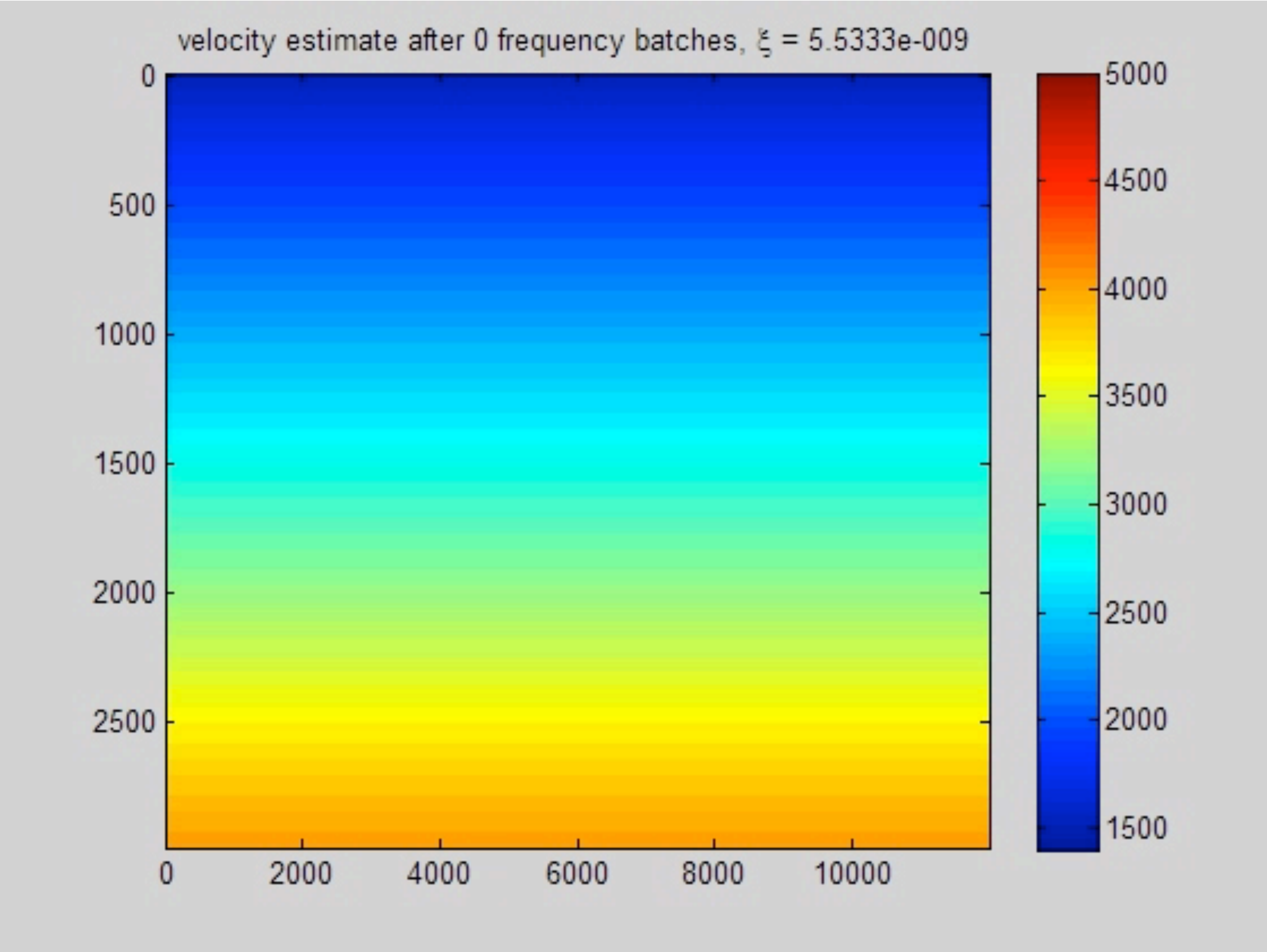
after six cycles through the frequencies



Adjoint-state FWI

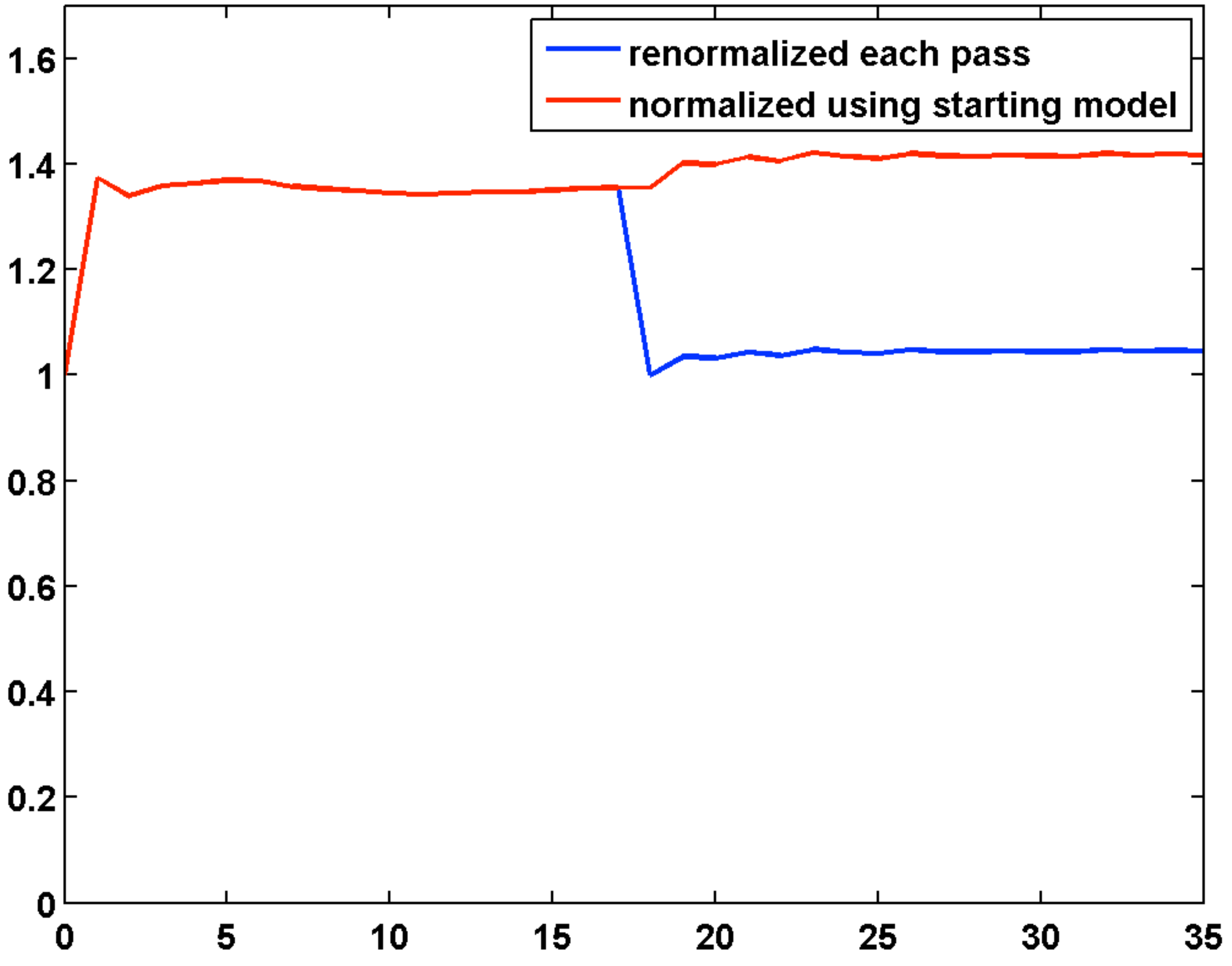
w/ or w/o TV-norm & hinge-loss projections & poor starting model



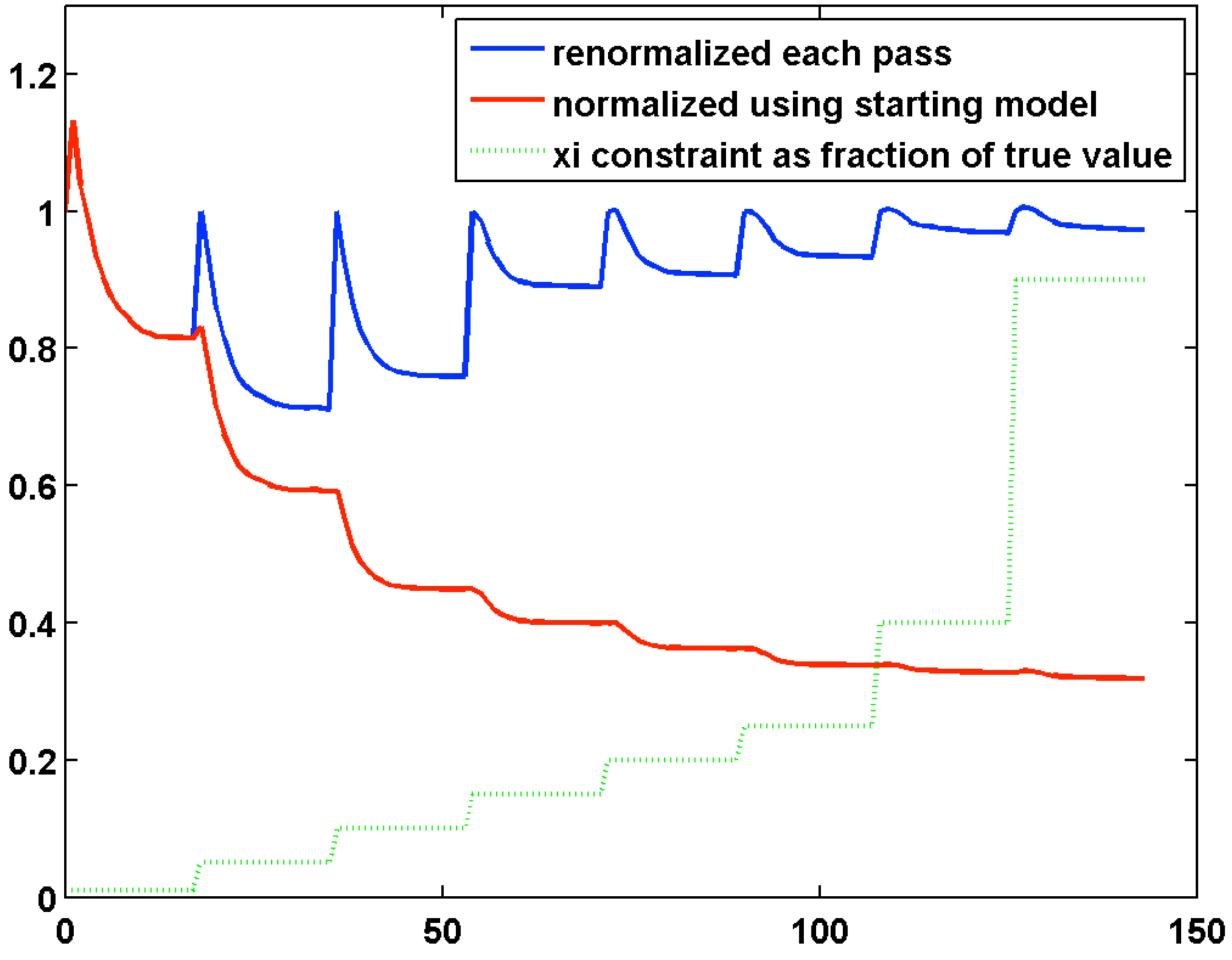


Relative model errors

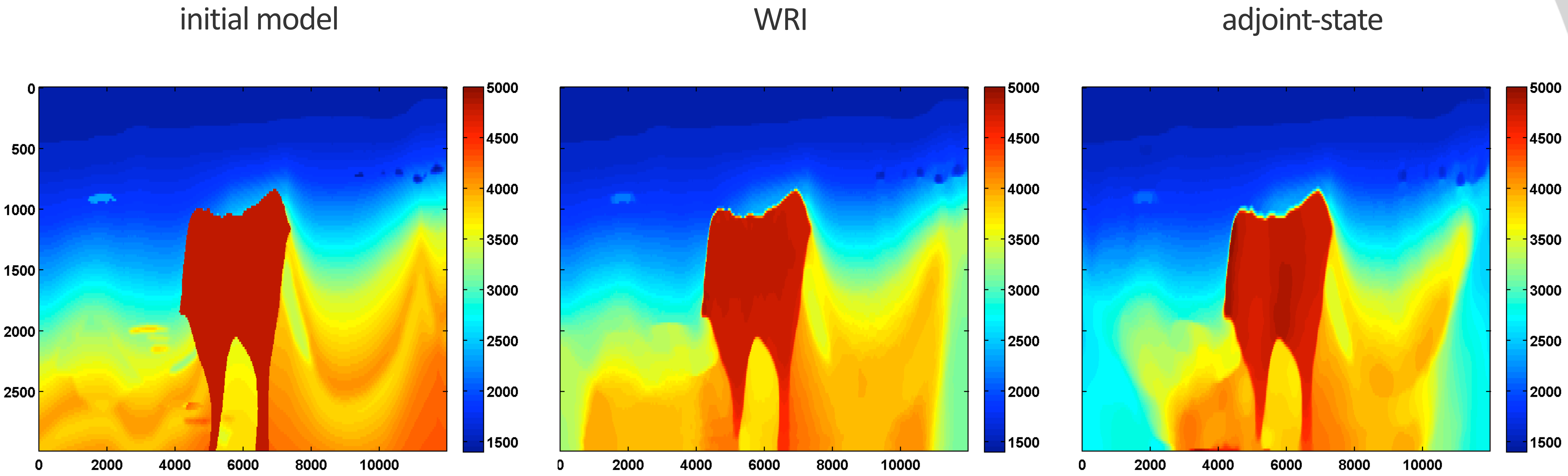
w/o TV



w/ TV & hinge



WRI vs adjoint-state



Why may this work?

Combination of

- multiscale frequency sweeps
- relaxation of the (*asymmetric*) convex constraints

work when

- progress is made during previous sweep
- adverse affects local minima are controlled by convex constraints
- “fine-scales” contribute to “coarse-scales” of the next sweep

Sounds like multi-level optimization…

Conclusions

New method for regularizing wave-equation based inversion benefits from

- combination of convex constraints
- multiple frequency sweeps w/ warm starts & relaxing of the constraints
- a hinge-loss function, which plays a critical role

Works for both WRI & adjoint-state FWI

Development of automatic continuation strategies for relaxing the constraints is ongoing.

Candidate for “automatic” salt flooding...

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, DownUnder GeoSolutions, Hess, Petrobras, PGS, Subsalt Ltd, WesternGeco, and Woodside.