Released to public domain under Creative Commons license type BY (https://creativecommons.org/licenses/by/4.0).
Copyright (c) 2018 SINBAD consortium - SLIM group @ The University of British Columbia.

Automatic salt delineation — Wavefield Reconstruction Inversion with convex constraints

Felix J. Herrmann

SLIM
University of British Columbia

Automatic salt delineation — Wavefield Reconstruction Inversion with 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.”

Challenge

Salt has

- sharp edges
- strong-velocity contrasts w.r.t. sedimentary layers
- high velocities

Major challenges for (automatic) velocity building...

SLIM

Strategy

Extend the search space

- "less" nonlinear
- ensures data fit & avoids cycle skips

"Squeeze" the extension by

- enforcing the wave equation to compute model updates
- imposing *asymmetric* constraints that encode "rudimentary" properties of the geology
- relaxing the constraints to allow data fits & details to enter the solution

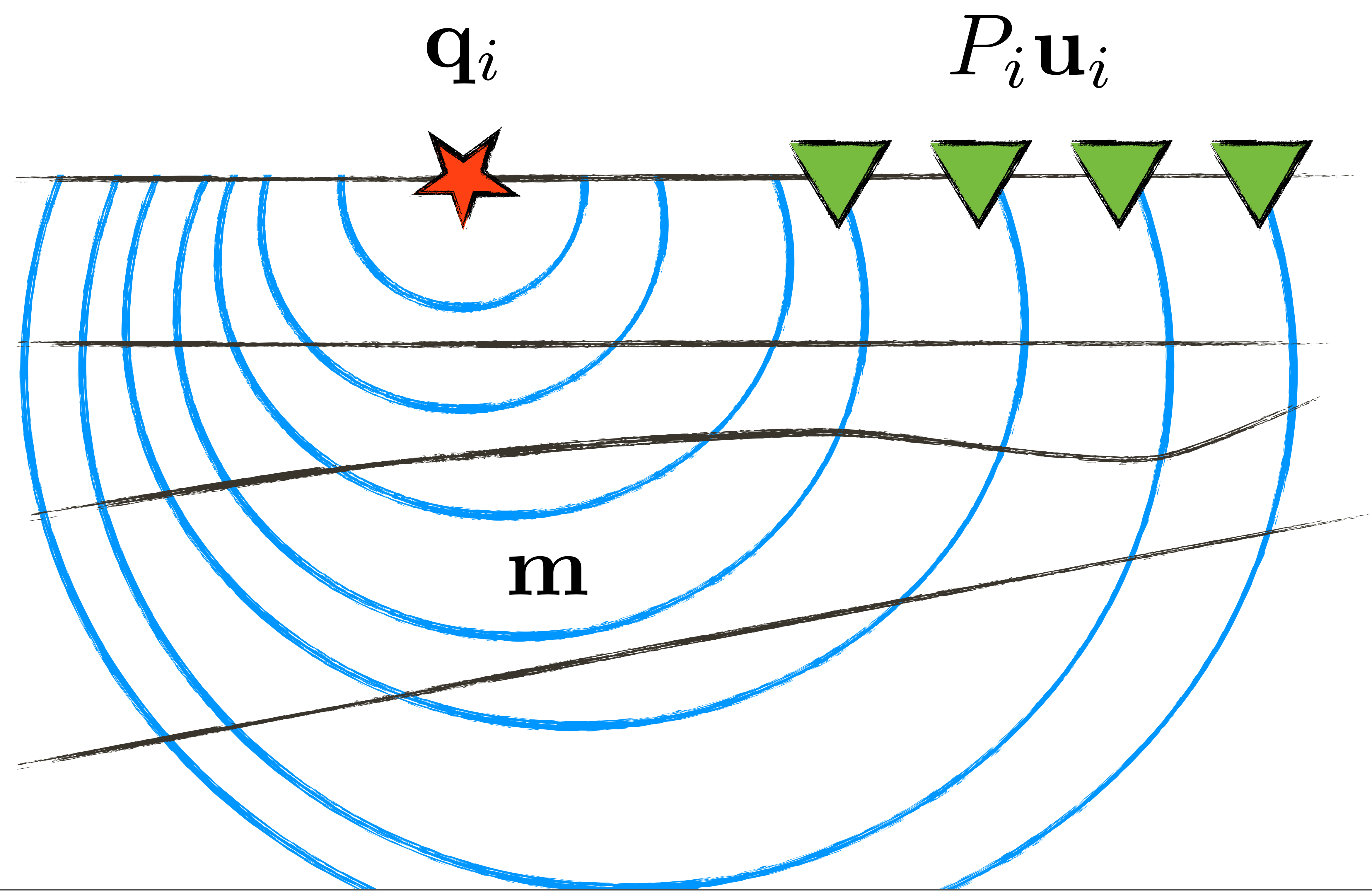
Leverage frequency continuation & warm starts where

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

Outcome: an automatic multi-cycle optimization-driven workflow

Waveform inversion

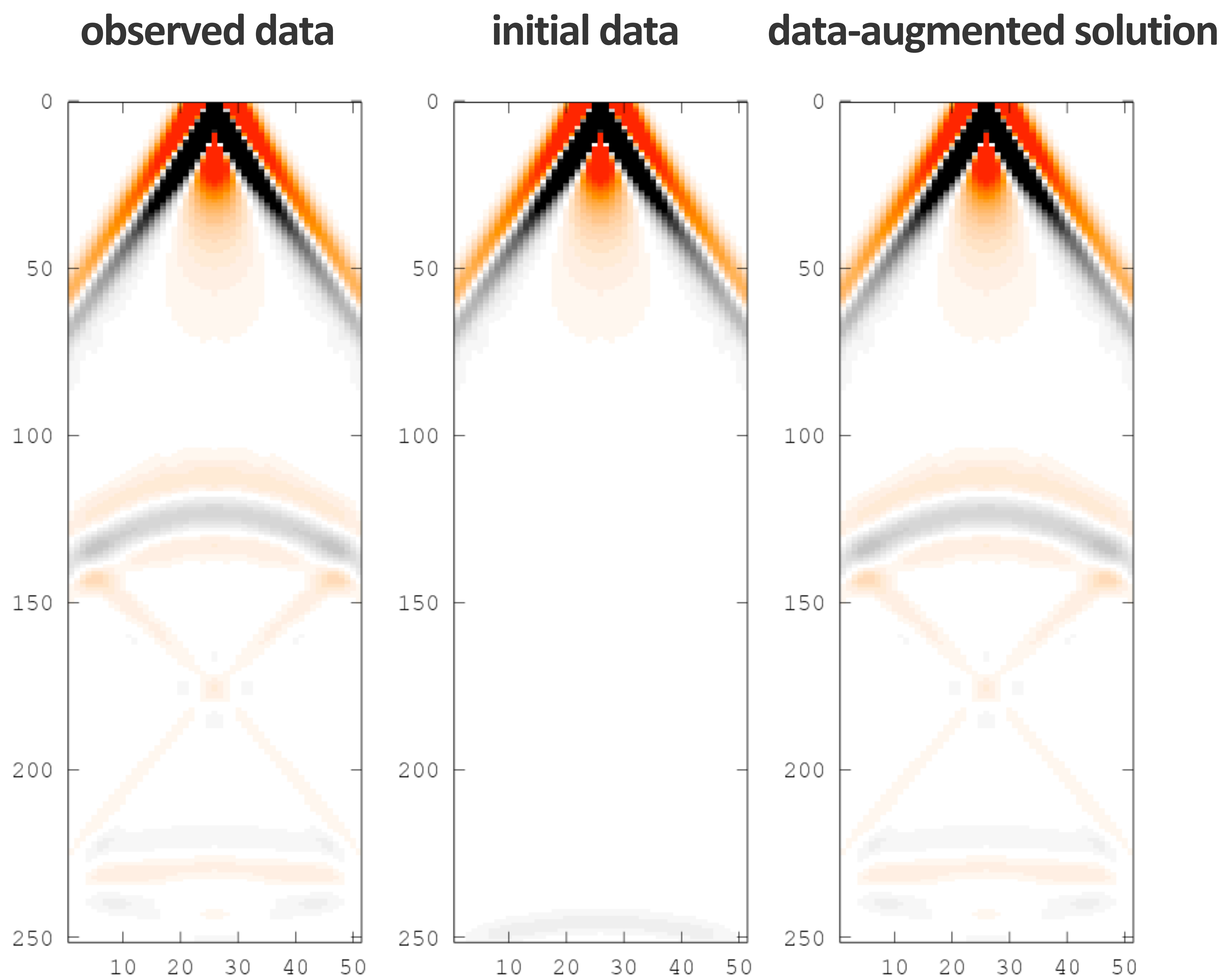
Retrieve the medium parameters from partial measurements of the solution of the wave-equation: $A(\mathbf{m})\mathbf{u}_i = \mathbf{q}_i$



wave-equation X **wavefield** = **source**

versus

$$\begin{pmatrix}\textbf{wave-equation} \\ \text{-----------} \\ \textbf{sampling operator}\end{pmatrix} \times \textbf{wavefield} = \begin{pmatrix}\textbf{source} \\ \text{-----------} \\ \textbf{data}\end{pmatrix}$$


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

[van Leeuwen & FJH, 2013]

US PCT Patent Application: PCT/CA2014/000386. A Penalty Method for PDE-Constrained Optimization

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

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

[van Leeuwen & FJH, ’13; ’15, Peters et. al. ’14]

SLIM

WRI – iterations

WRI method

for each source i

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

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

end

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

correlation proxy wavefield & PDE residual

Conventional method

for each source i

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

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

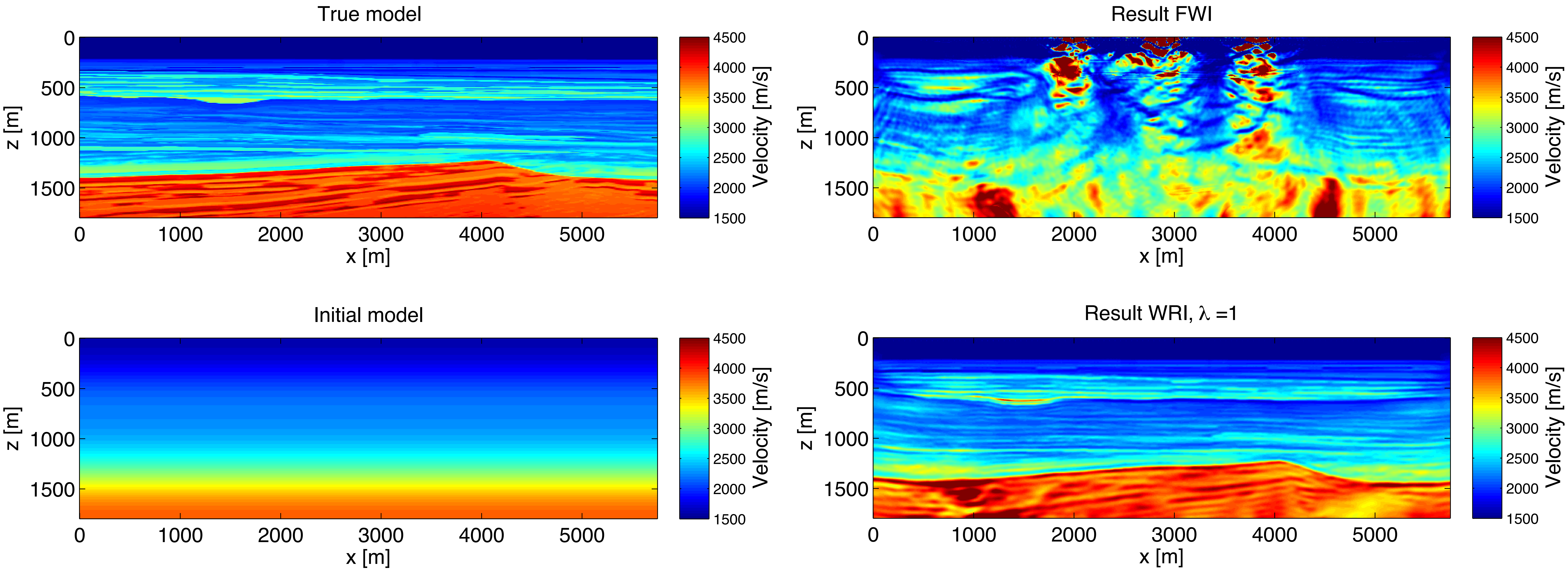
$\mathbf{g} = \mathbf{g} + \omega^2 \mathsf{diag}(\mathbf{u}_i)^*\mathbf{v}_i$

end

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

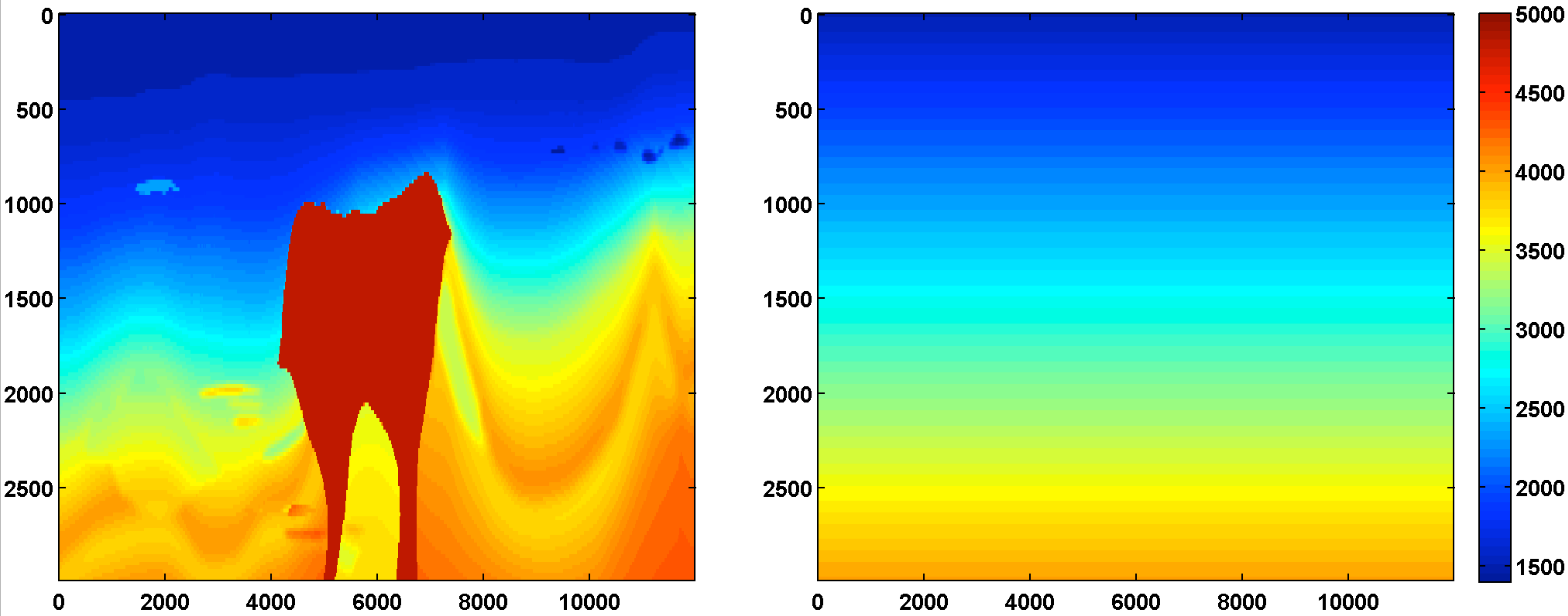
correlation wavefield & data residual

Wavefield Reconstruction Inversion (WRI)
– poor starting model



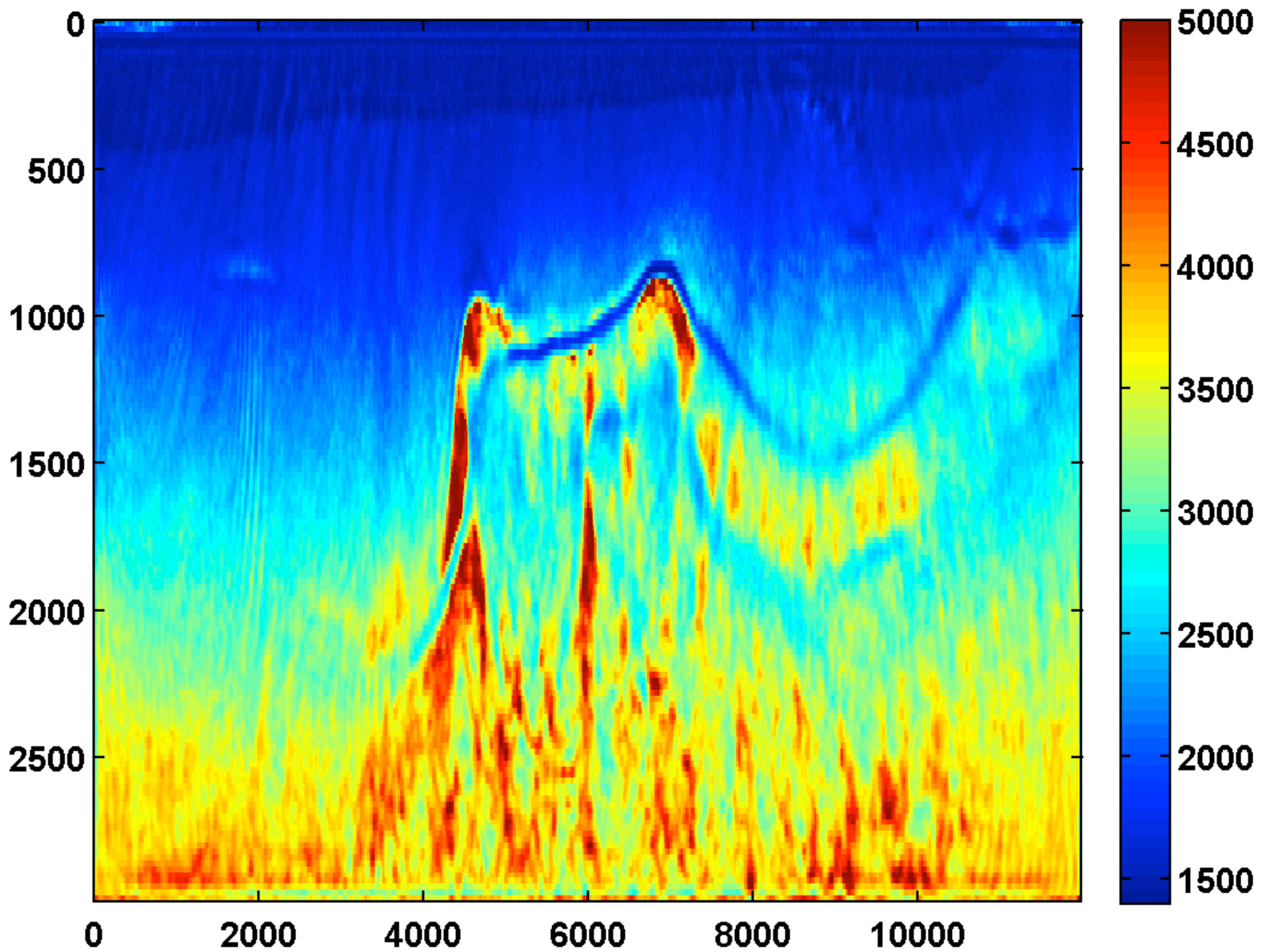
Example from **[Peters et al. 2013]**

Waveform inversion – poor starting model

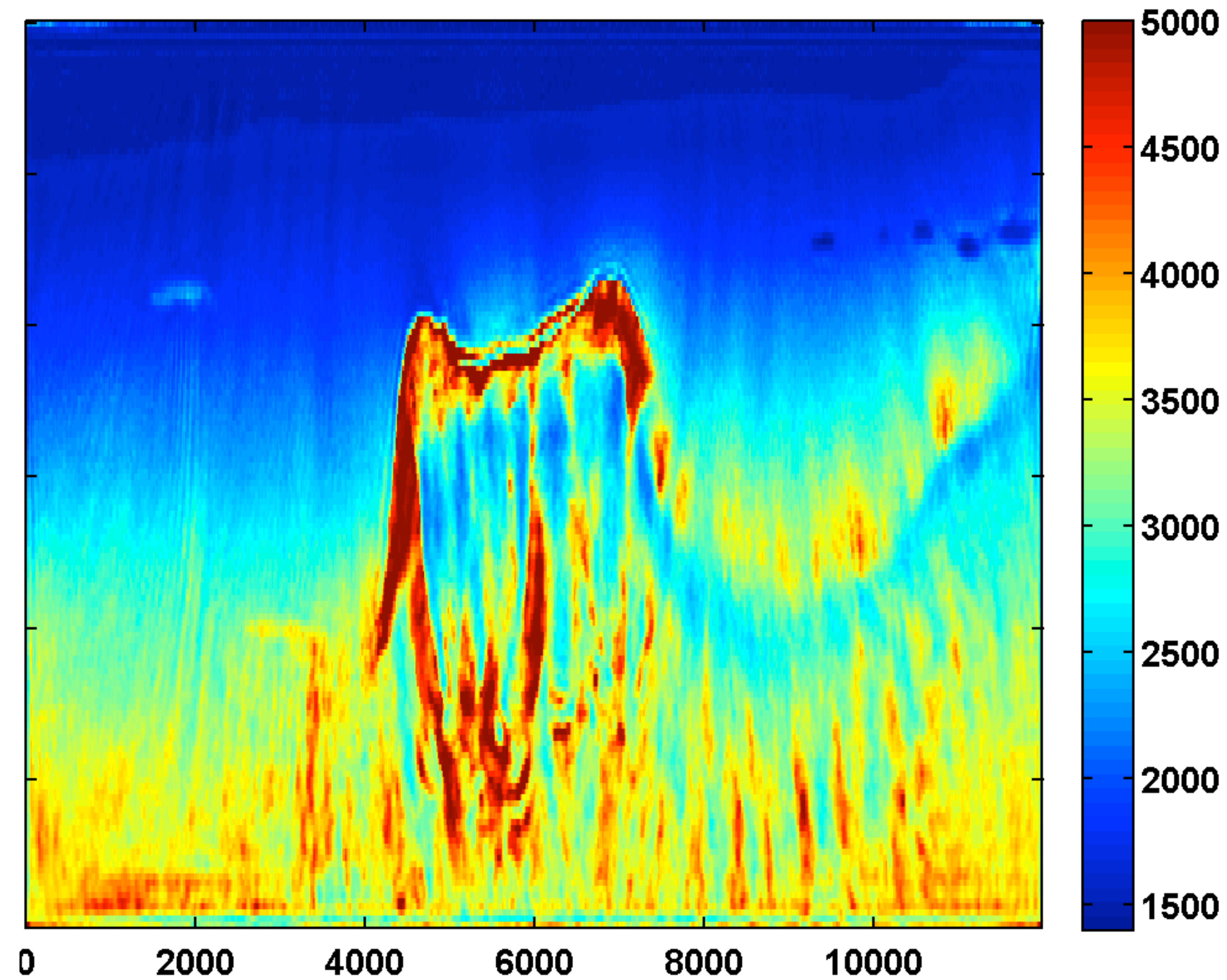


WRI results w/o TV

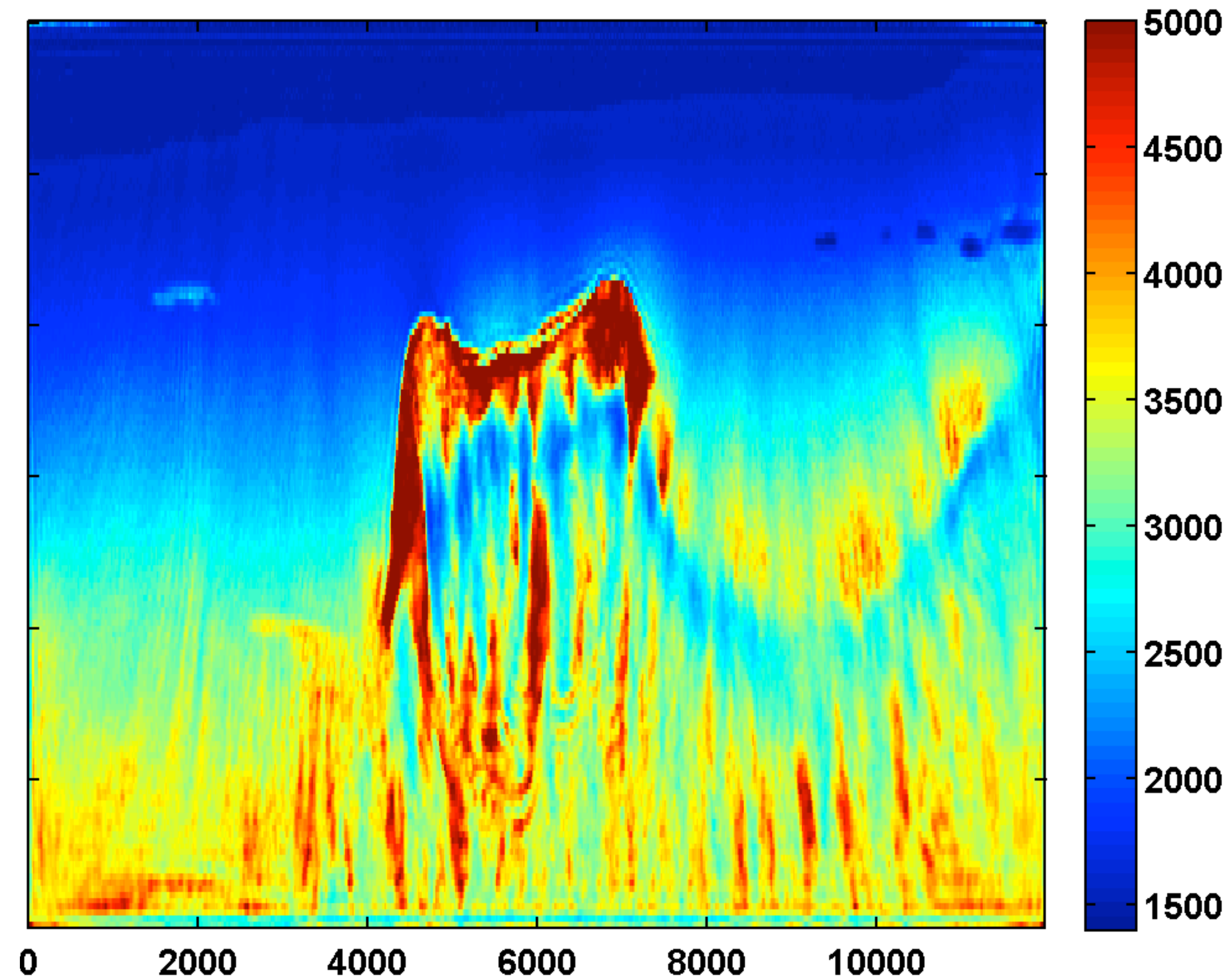
after one cycle through the frequencies



after two cycles through the frequencies



after three cycles through the frequencies



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

Including convex constraints

Aim: constrain inverted models to a certain set of models, e.g. via

- box constraints – ensure point-wise values to be w/i specified interval
- TV-norm constraints – promote sharp edges & remove unwanted “noise”

Challenge: optimization of large-scale problems

- intersections of convex sets
- computational costs

Solution: leverage developments in optimization w/ “simple” constraints

- saddle-point formulation – primal-dual hybrid gradients
- leverage WRI’s cheap GN Hessians in inner-/outer-loop structure

[Heinkenschloss, '98 , Haber, '00]

Full-space / "all at once"

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

SLIM

Adjoint-state/reduced-space formulation

By eliminating the constraint

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

- 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

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

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

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

Use 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]

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

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

SLIM

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

Including convex constraints

expensive but fixed — cheap — damped

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

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

- 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]

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.

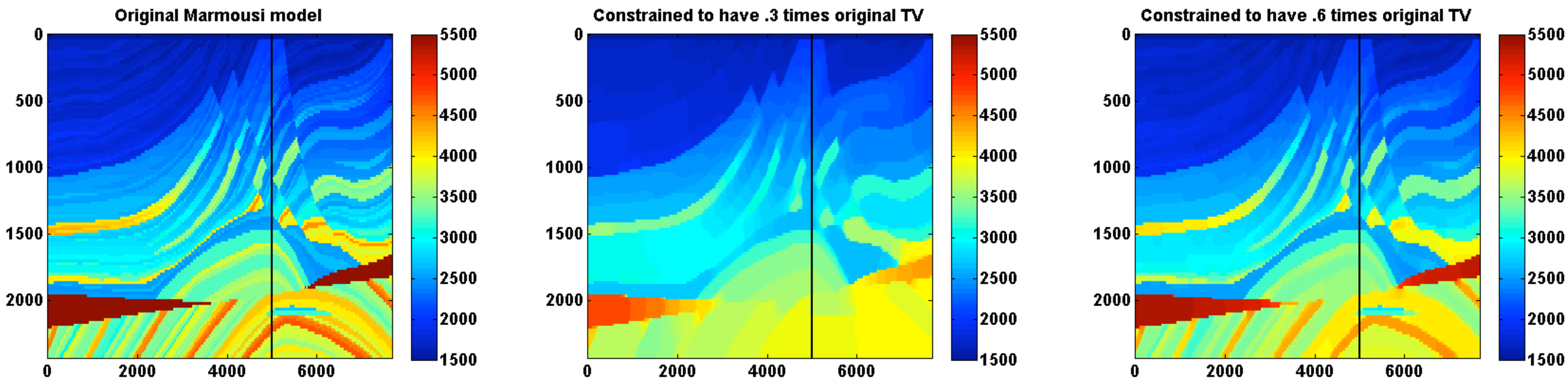
[Oldenburg ’83; Akçelik ’08; Anagaw ’11; Maharramov ’14; Esser & FJH ’14]

SLIM

Edge-preserving Total-Variation projections

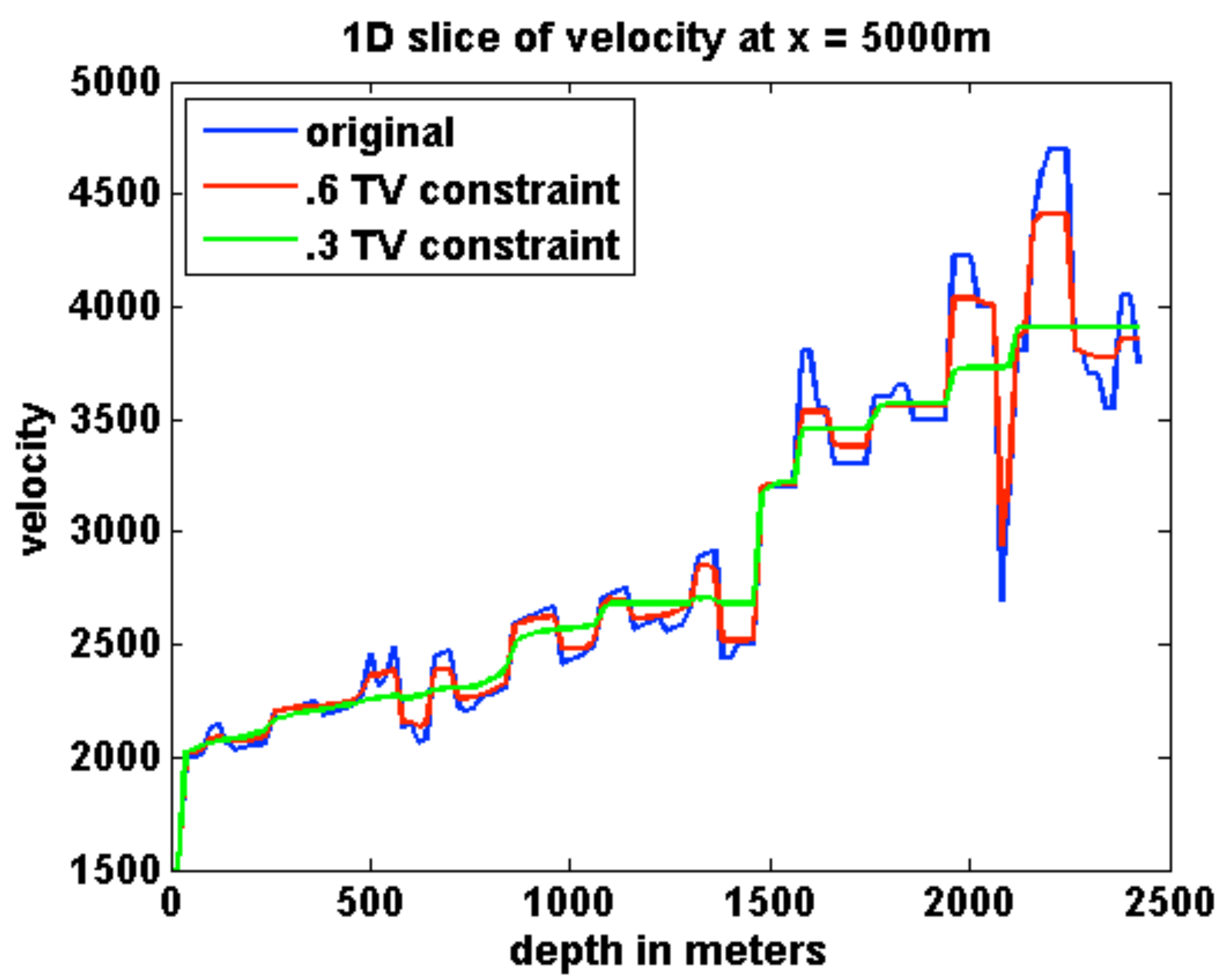
$v_{\min} = 1500,\ v_{\max} = 5500,\ \text{and}\ \tau = \{0.3\tau_0,\ 0.6\tau_0\}$

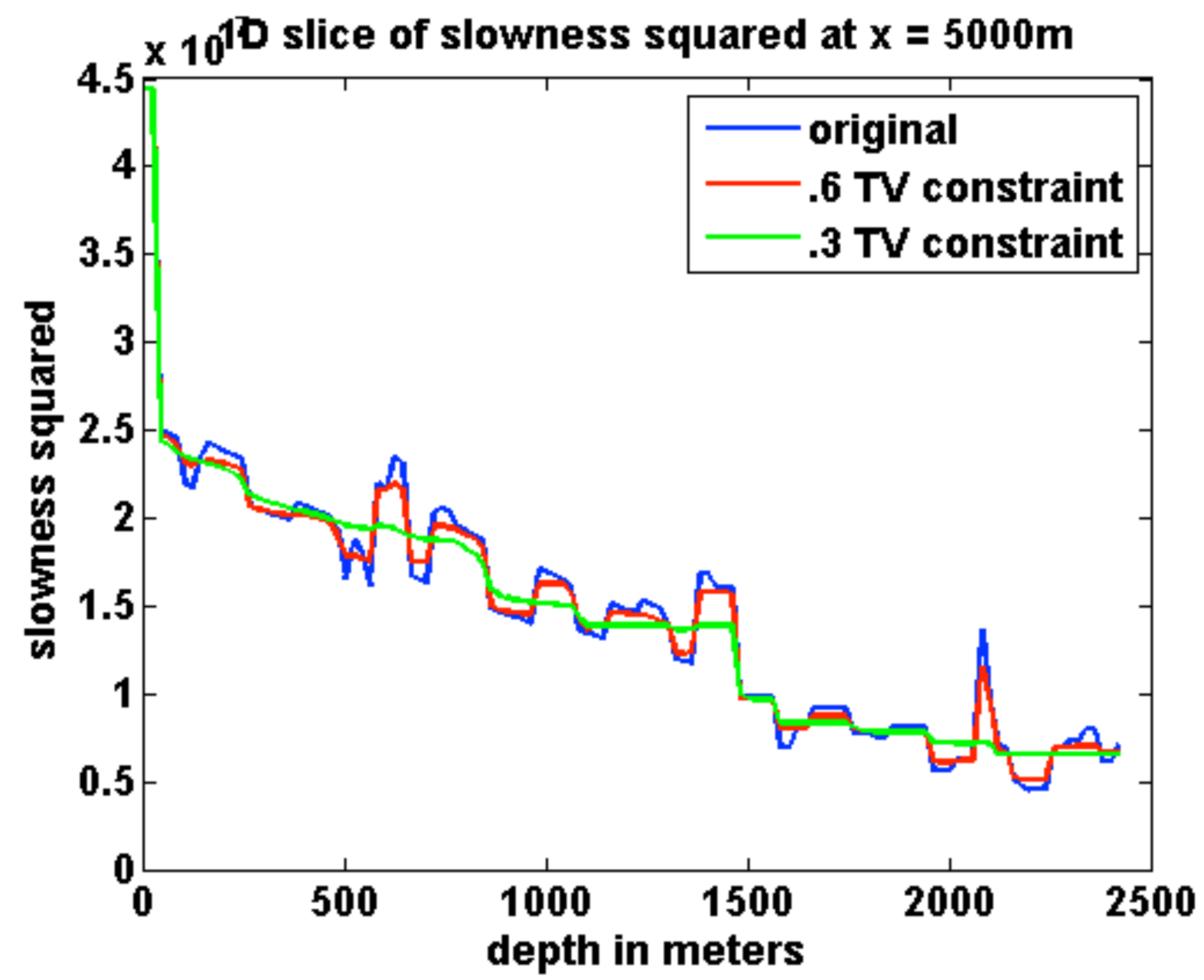
$$\Pi_C(\mathbf{m}_0) = \arg\min_{\mathbf{m}} \frac{1}{2}\|\mathbf{m} - \mathbf{m}_0\|^2 \quad \text{subject to} \quad \mathbf{m}_i \in [B_i^l, B_i^u] \text{ and } \|\mathbf{m}\|_{TV} \leq \tau$$



Edge-preserving Total-Variation projections

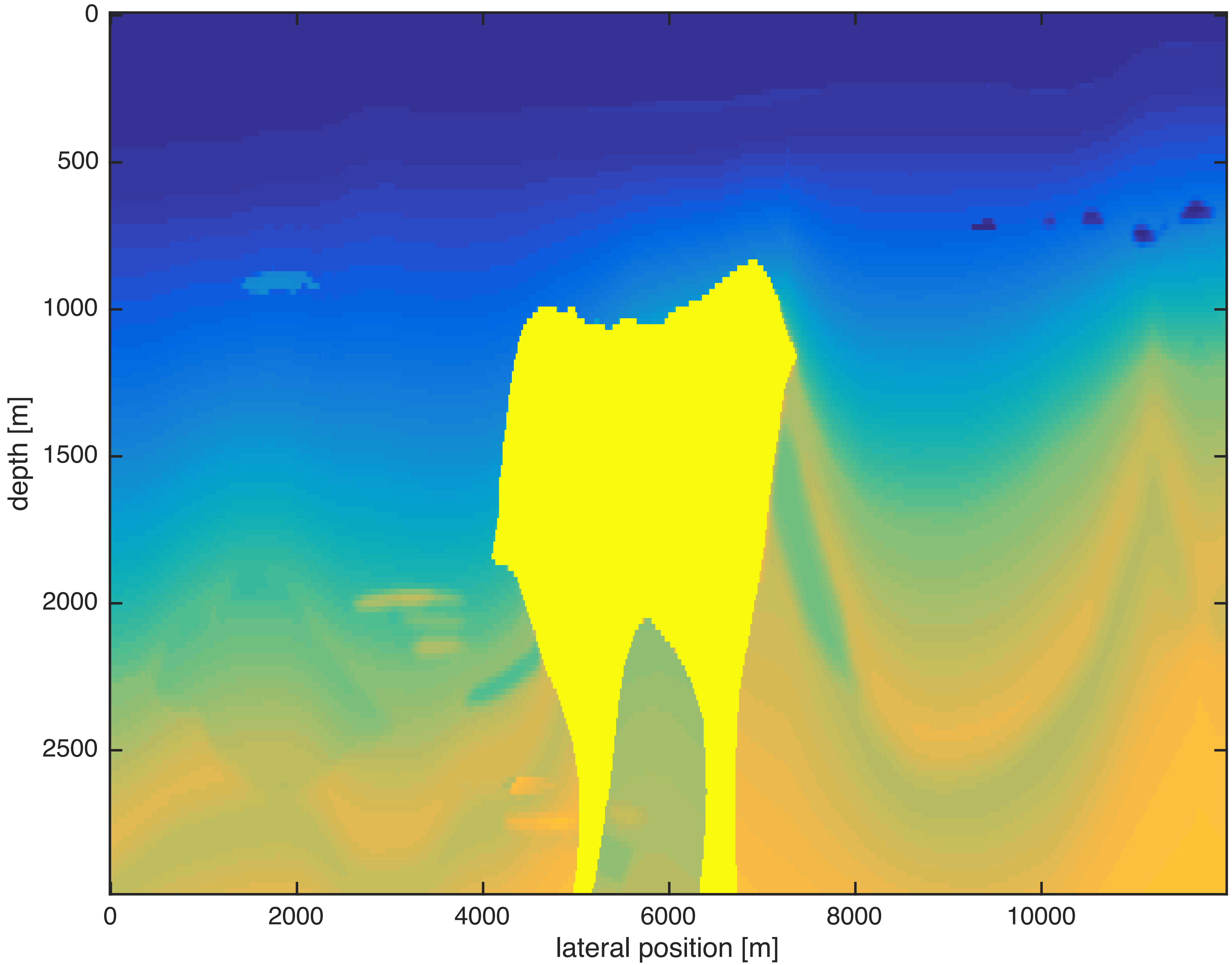
$v_{\min} = 1500,\ v_{\max} = 5500,$ and $\tau = \{0.3\tau_0,\ 0.6\tau_0\}$



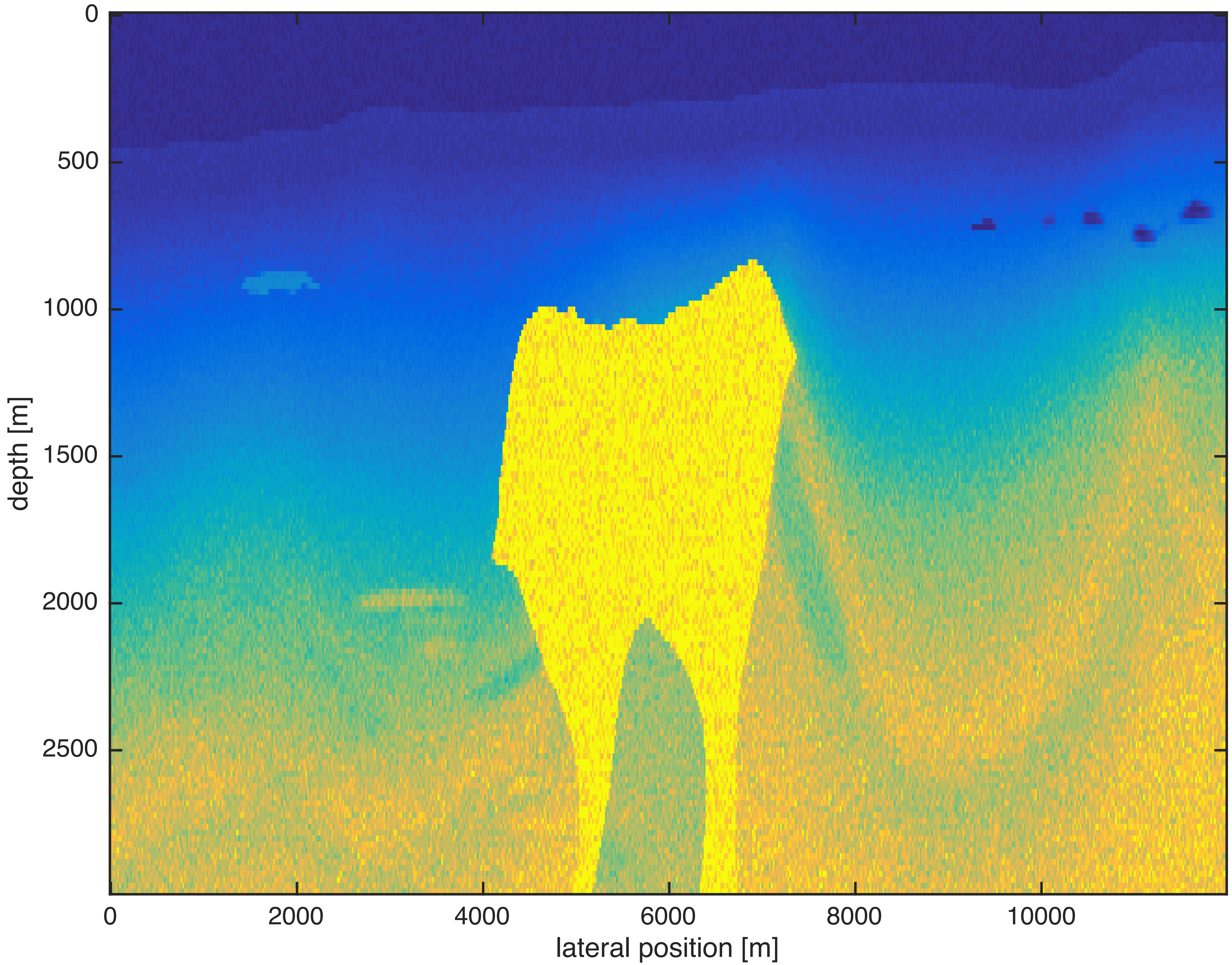


- reduces complexity of the model
- preserves blockiness of reflectors
- relaxing the constraint allows for more layers

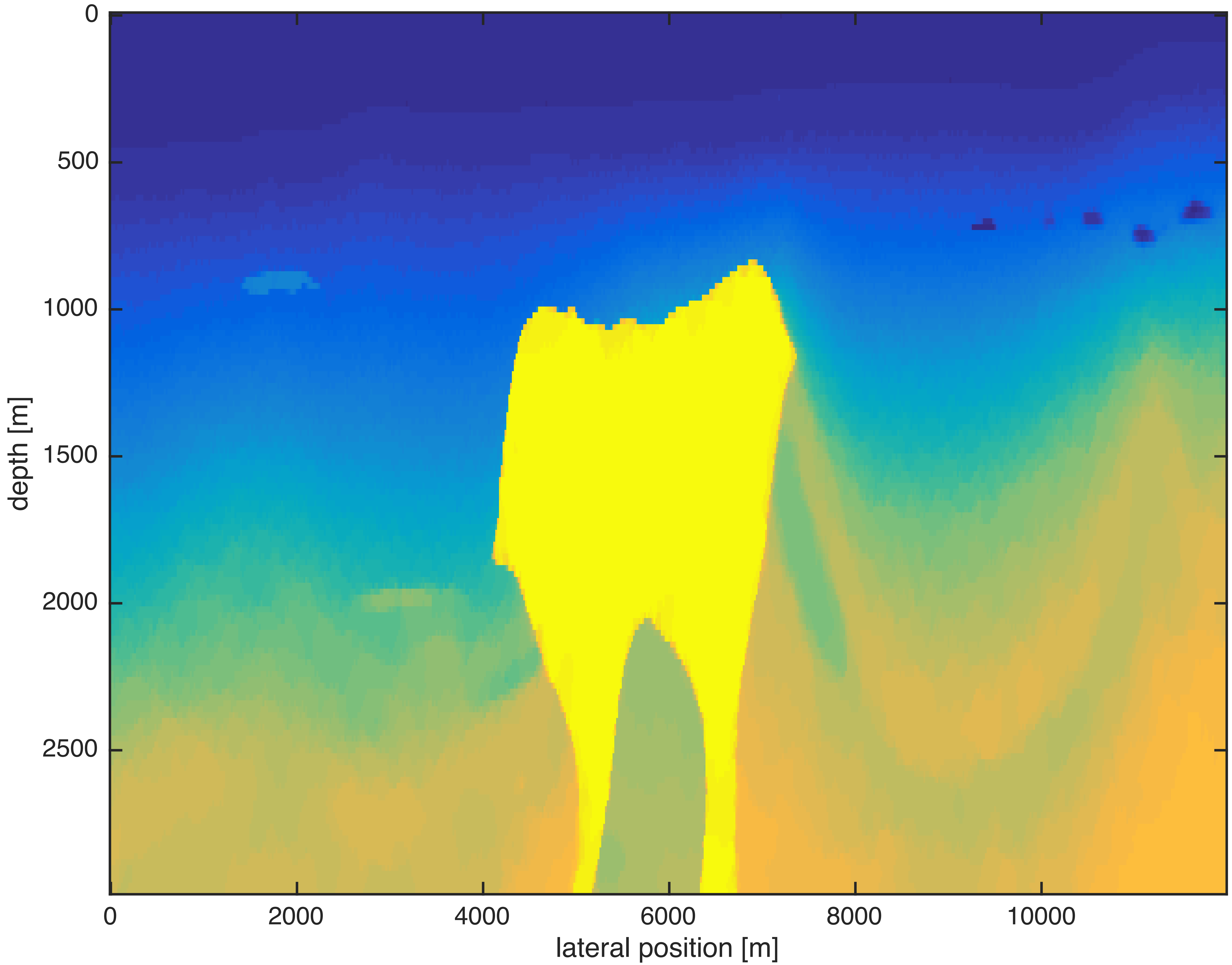
Subset BP 2004 Salt model – original



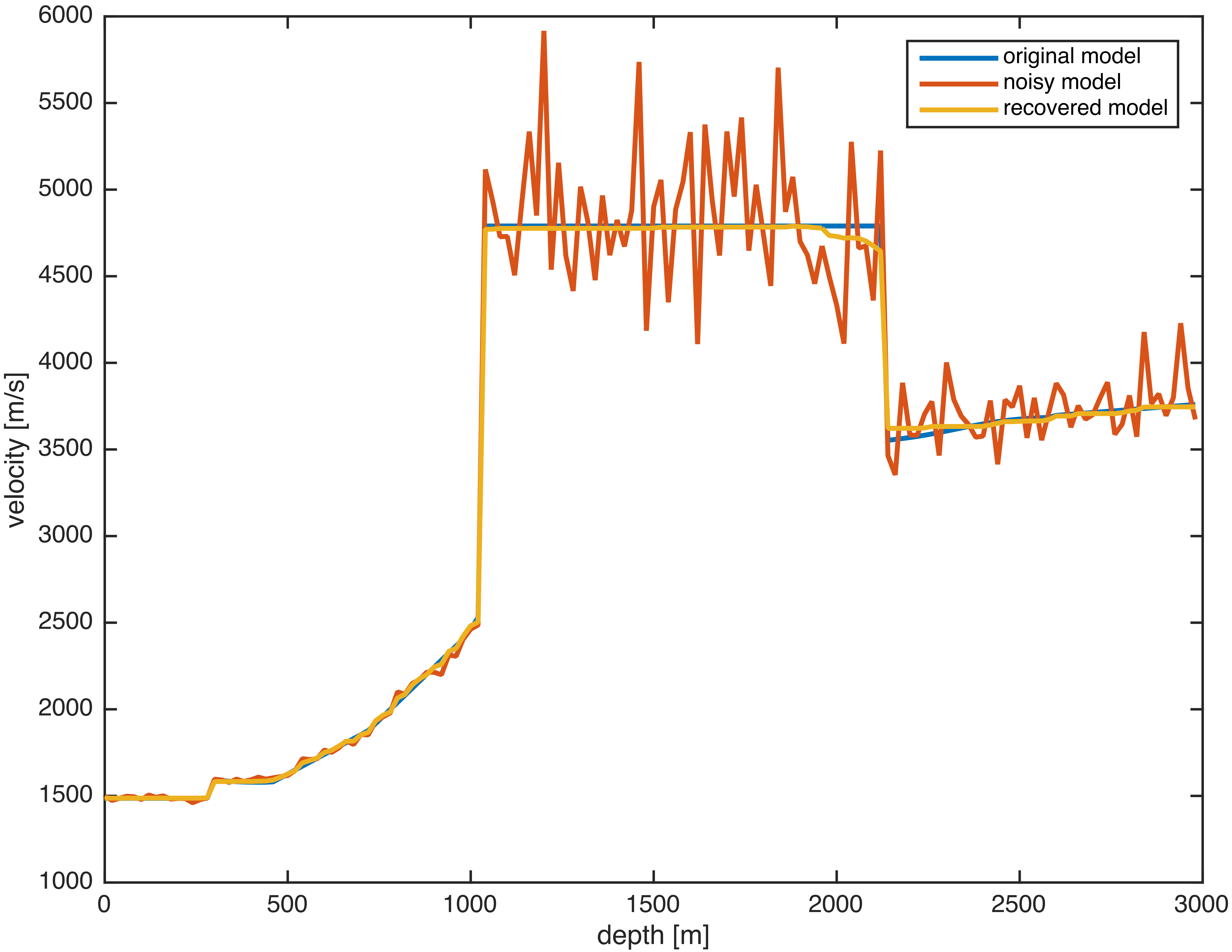
Subset BP 2004 Salt model – noisy



Subset BP 2004 Salt model – TV-constraint imposed



Subset BP 2004 Salt model – TV-constraint imposed



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

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

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

$$\mathbf{\Delta m} = \arg\min_{\mathbf{\Delta m}} \mathbf{\Delta m}^T \mathbf{g}^n + \frac{1}{2}\mathbf{\Delta m}^T (H^n + c_n \mathrm{I})\mathbf{\Delta m}$$

$$\text{subject to} \quad \mathbf{m}_i^n + \mathbf{\Delta m}_i \in [B_i^l, B_i^u] \text{ and } \|\mathbf{m}^n \mathbf{\Delta m}\|_{TV} \leq \tau$$

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

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

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

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

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

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

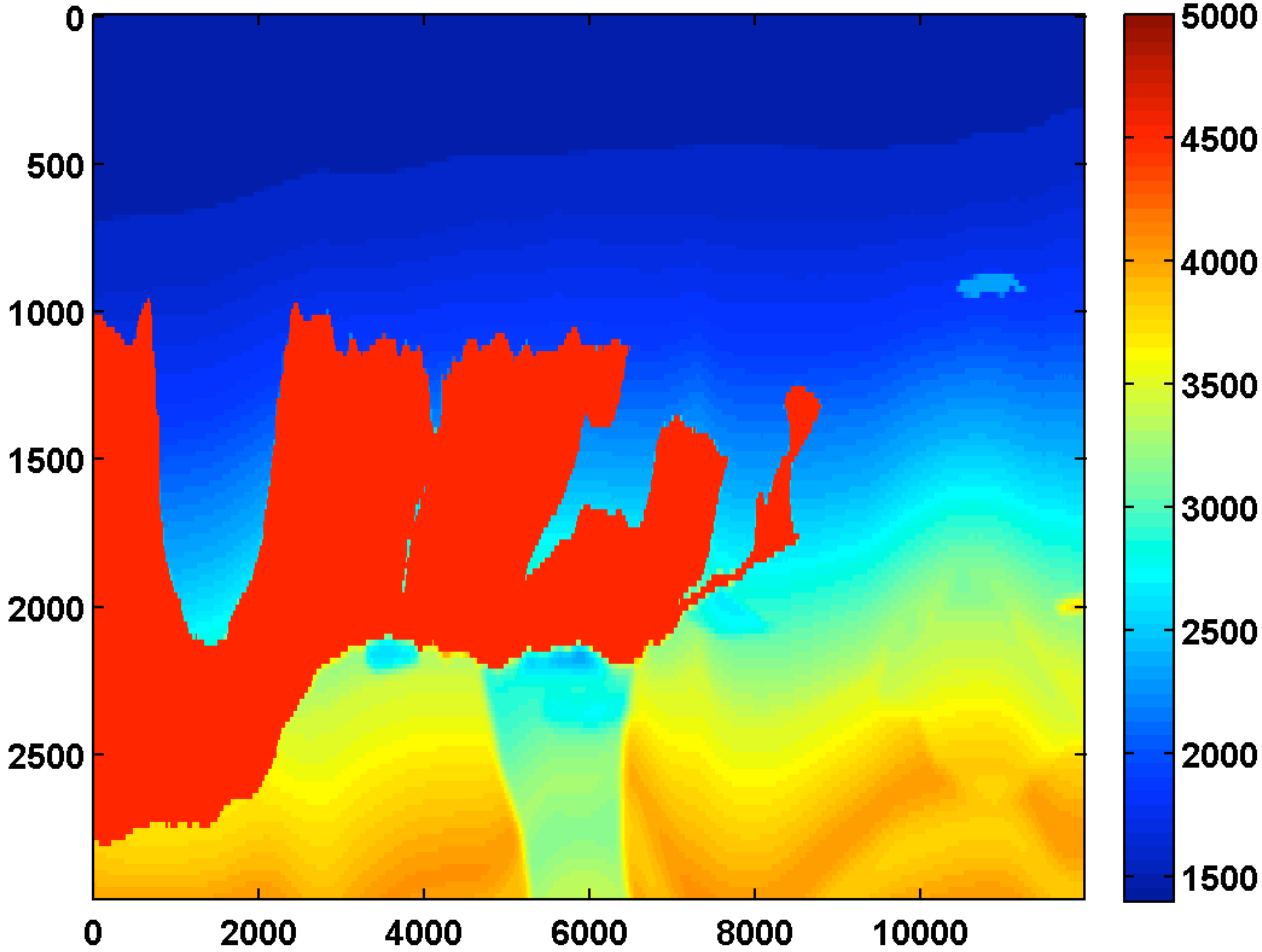
$$B_i = \min\left((B_i^u - \mathbf{m}_i^n), [(H^n + (c_n + \frac{1}{\alpha})\mathrm{I})^{-1}(-\mathbf{g}^n + \frac{\Delta\mathbf{m}^k}{\alpha} - D^T(2\mathbf{p}^{k+1} - \mathbf{p}^k)]_i\right)$$

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

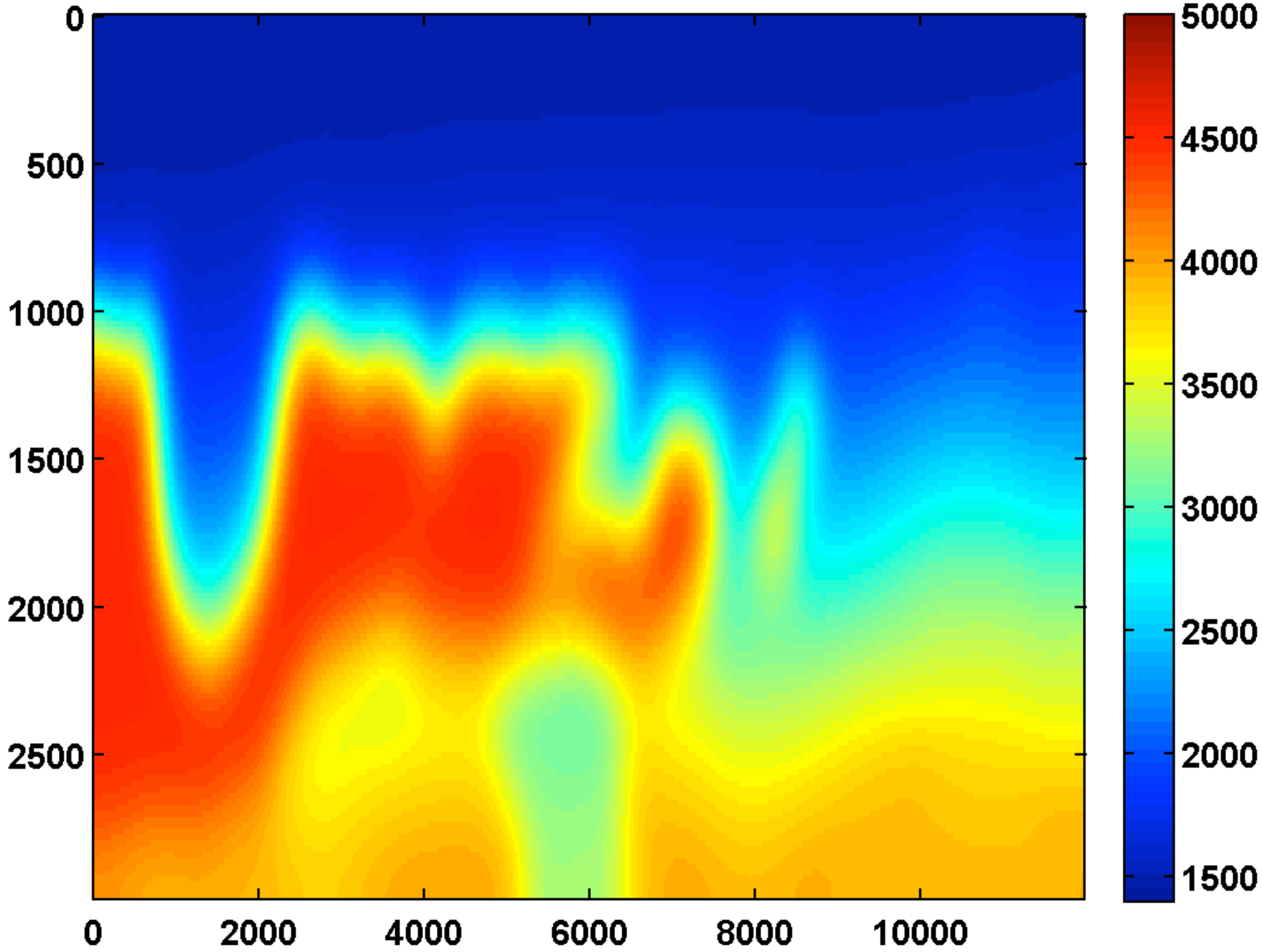
- do not involve solutions of (data-augmented) wave equations
- allows for data-dependent stepsizes

BP model

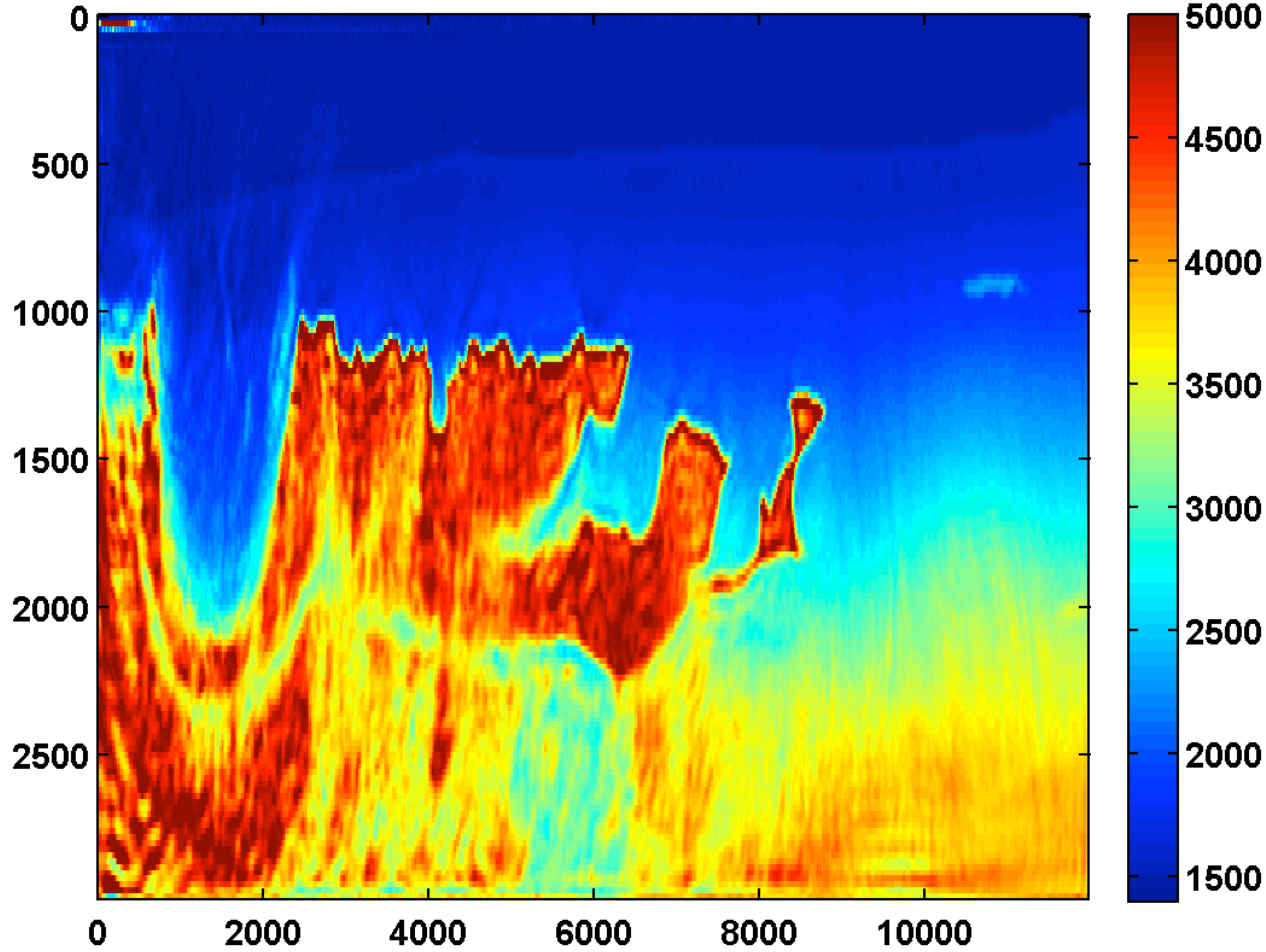
- number of sources: 126
- number of receivers: 299
- frequency continuation over 3-20Hz in overlapping batches of 2
- maximum number of outer iterations per frequency batch: 25
- maximum number of inner iterations for convex subproblems: 2000
- known Ricker wavelet sources with 15Hz peak frequency
-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


0
500
1000
1500
2000
2500
0
2000
4000
6000
8000
10000
5000
4500
4000
3500
3000
2500
2000
1500

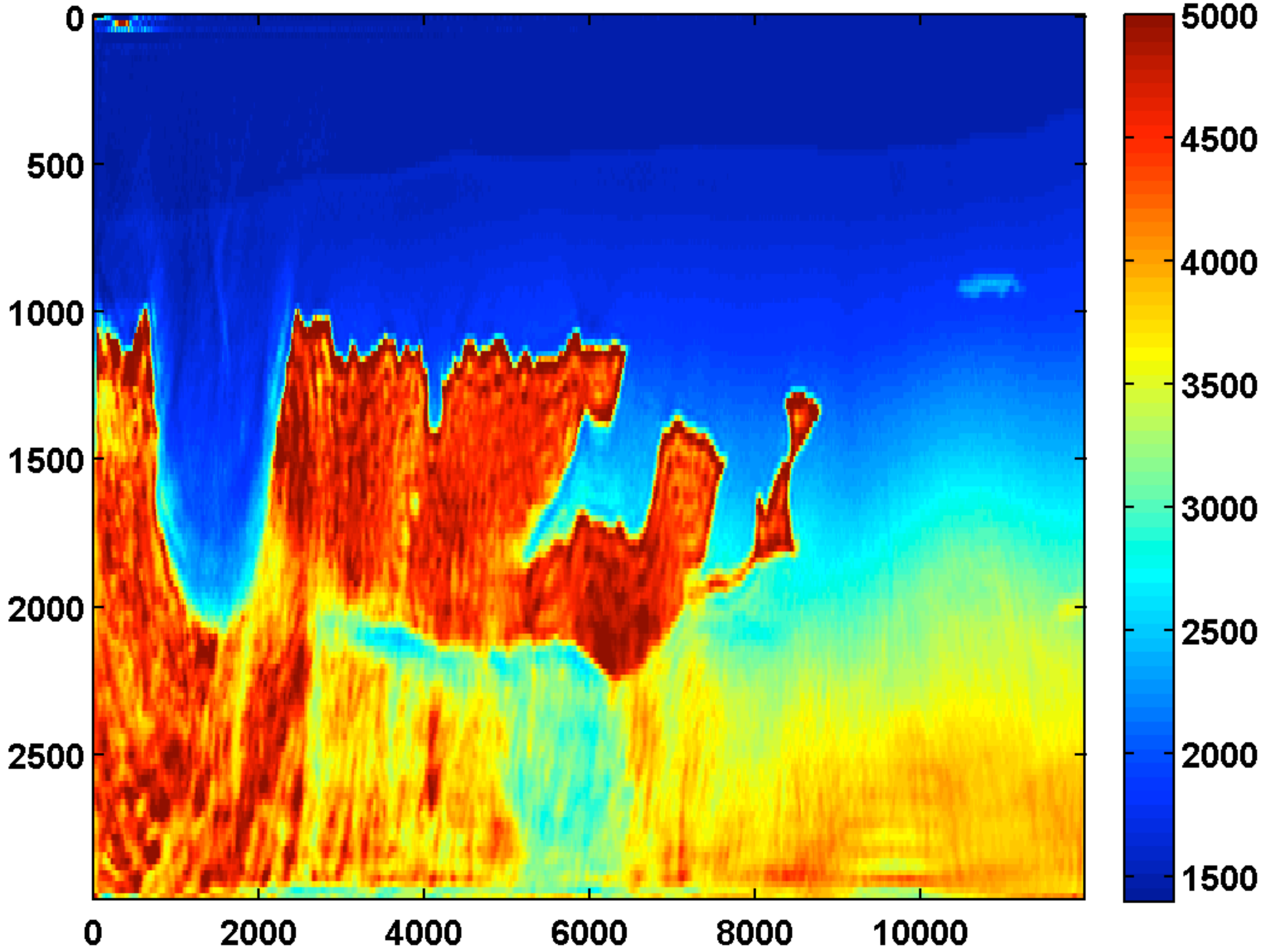
Salt – good starting model



w/o TV – sweep 1

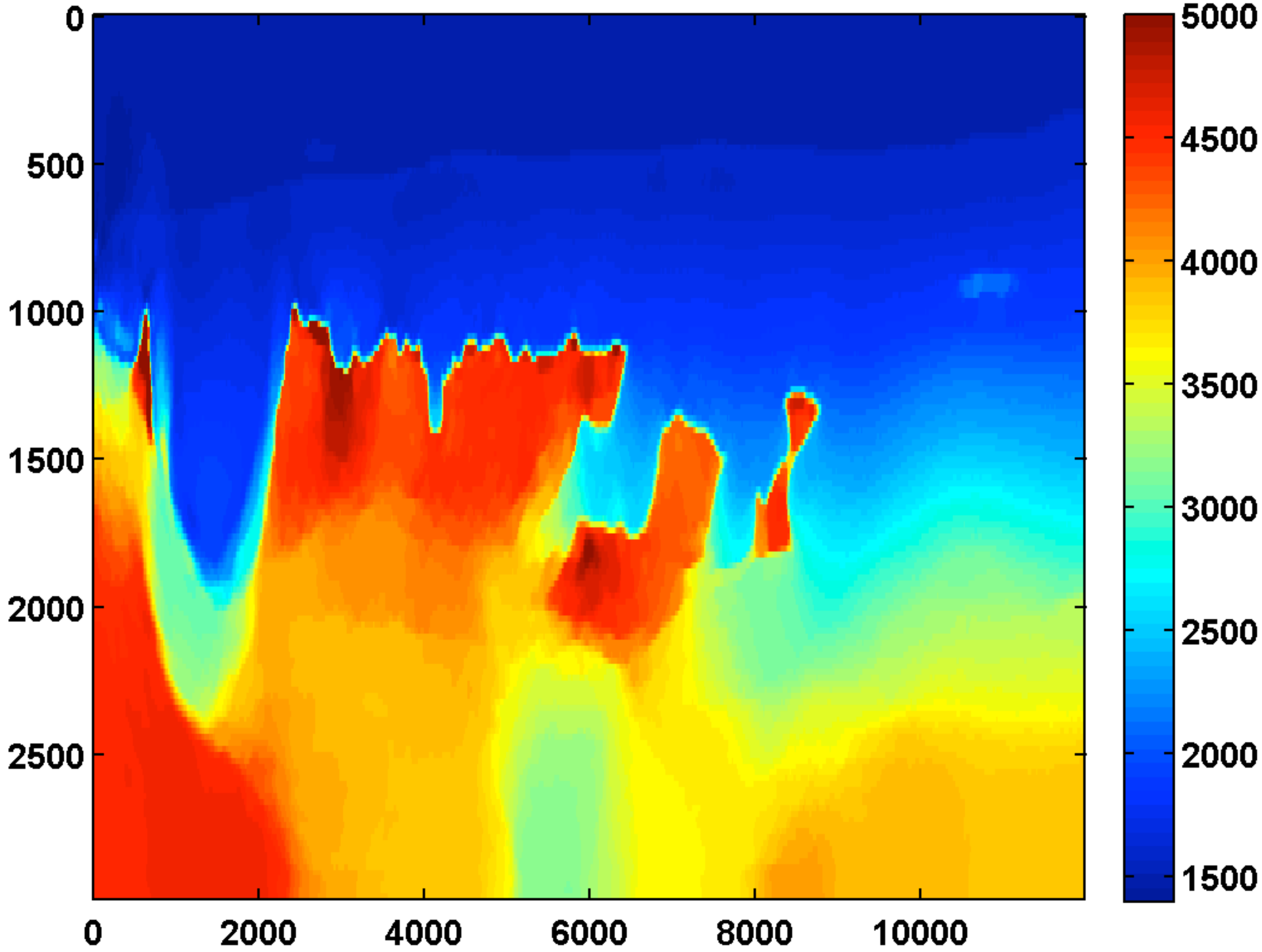

0
500
1000
1500
2000
2500
0
2000
4000
6000
8000
10000
5000
4500
4000
3500
3000
2500
2000
1500

w/o TV – sweep 2



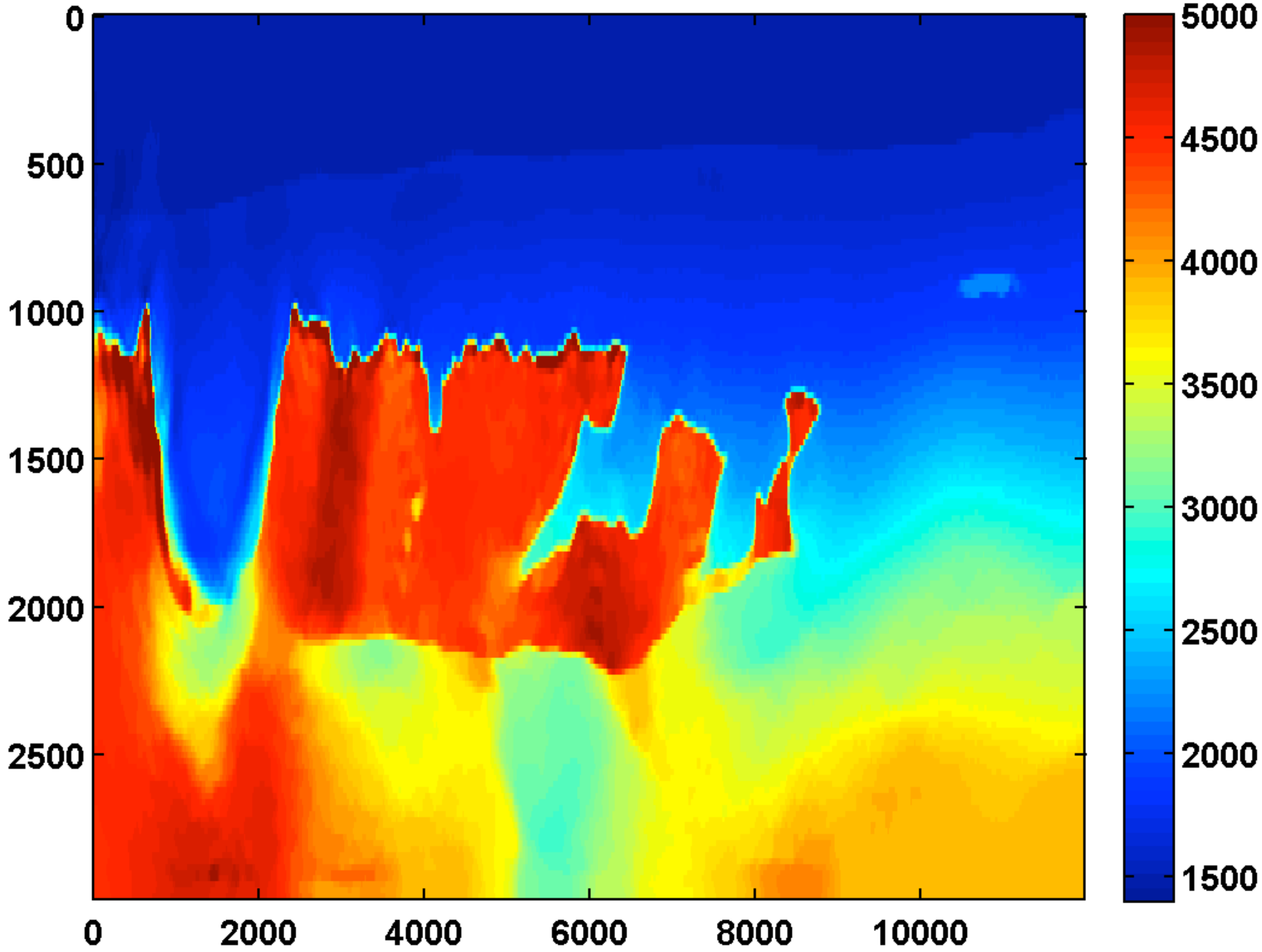
0
500
1000
1500
2000
2500
0
2000
4000
6000
8000
10000
5000
4500
4000
3500
3000
2500
2000
1500

TV – sweep 1

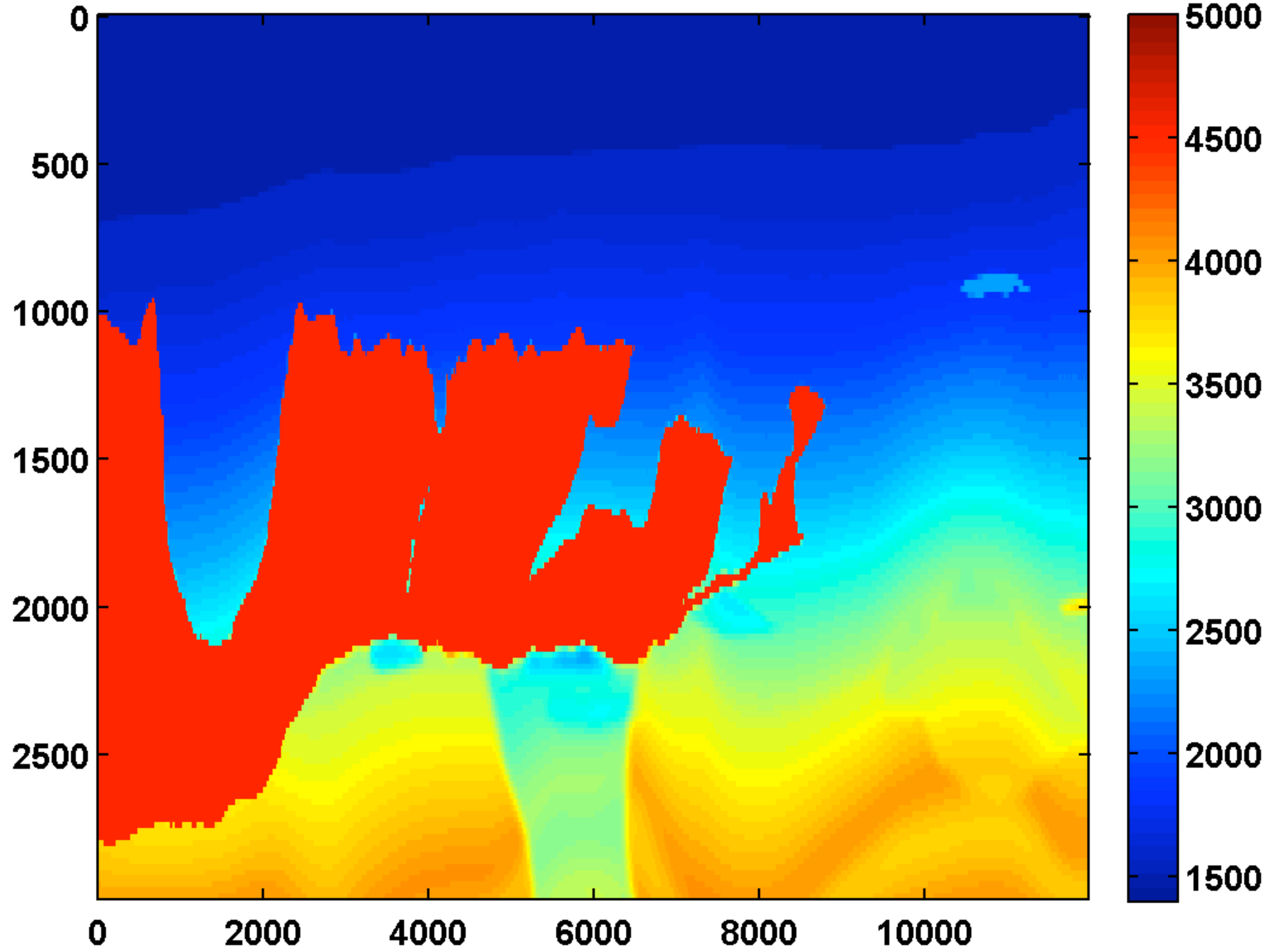


0
500
1000
1500
2000
2500
0
2000
4000
6000
8000
10000
5000
4500
4000
3500
3000
2500
2000
1500

w TV – sweep 2

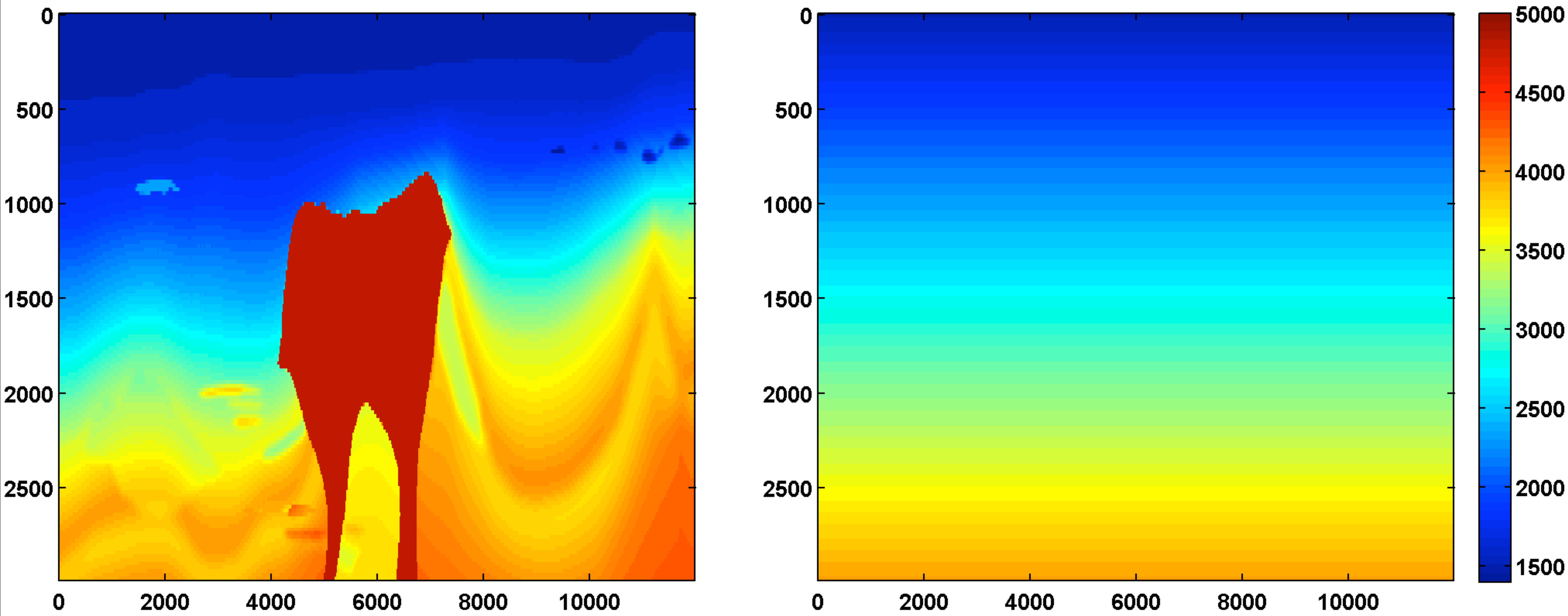


0
500
1000
1500
2000
2500
0
2000
4000
6000
8000
10000
5000
4500
4000
3500
3000
2500
2000
1500



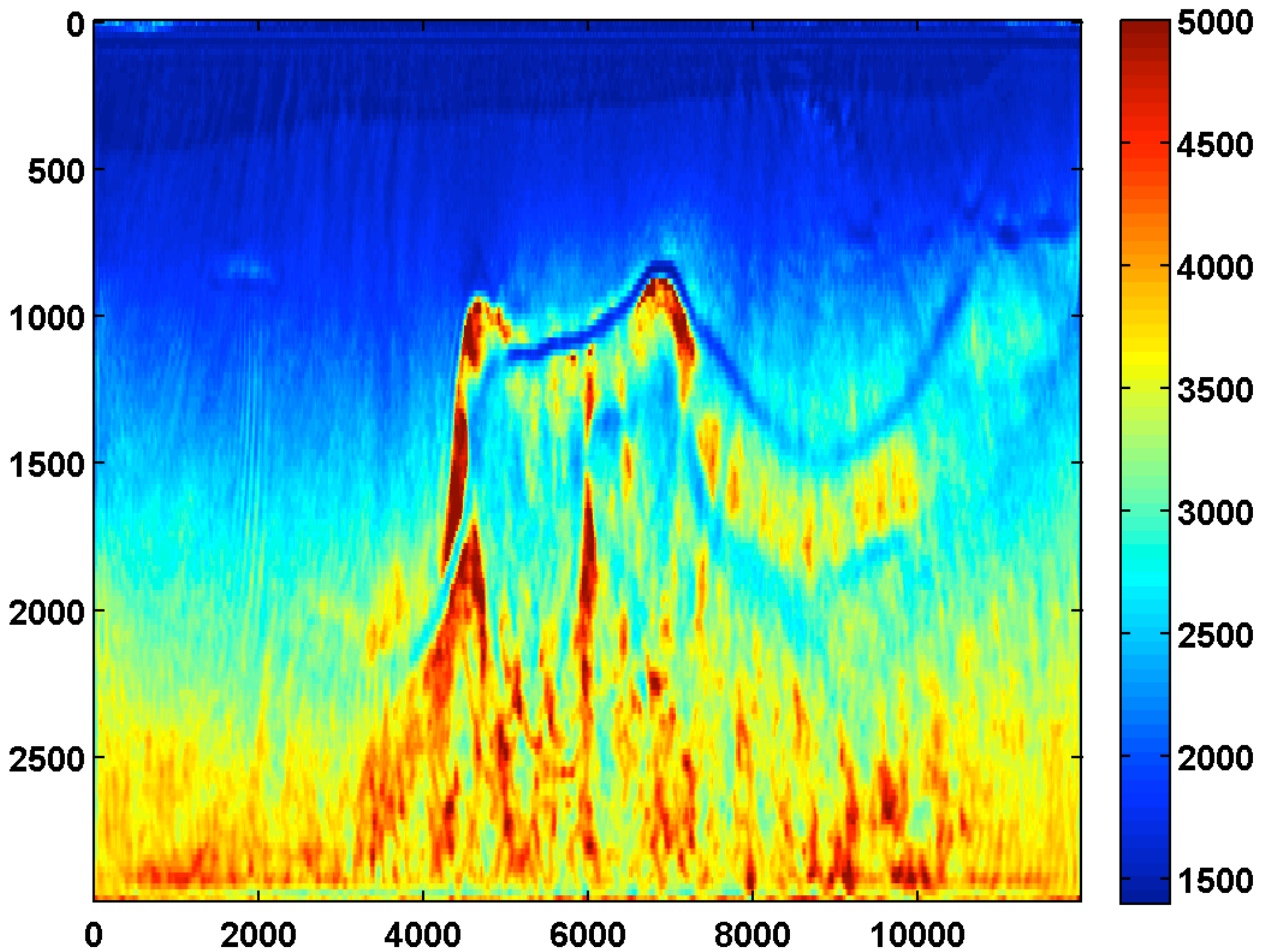
0
500
1000
1500
2000
2500
0
2000
4000
6000
8000
10000
5000
4500
4000
3500
3000
2500
2000
1500

True velocity & poor starting model

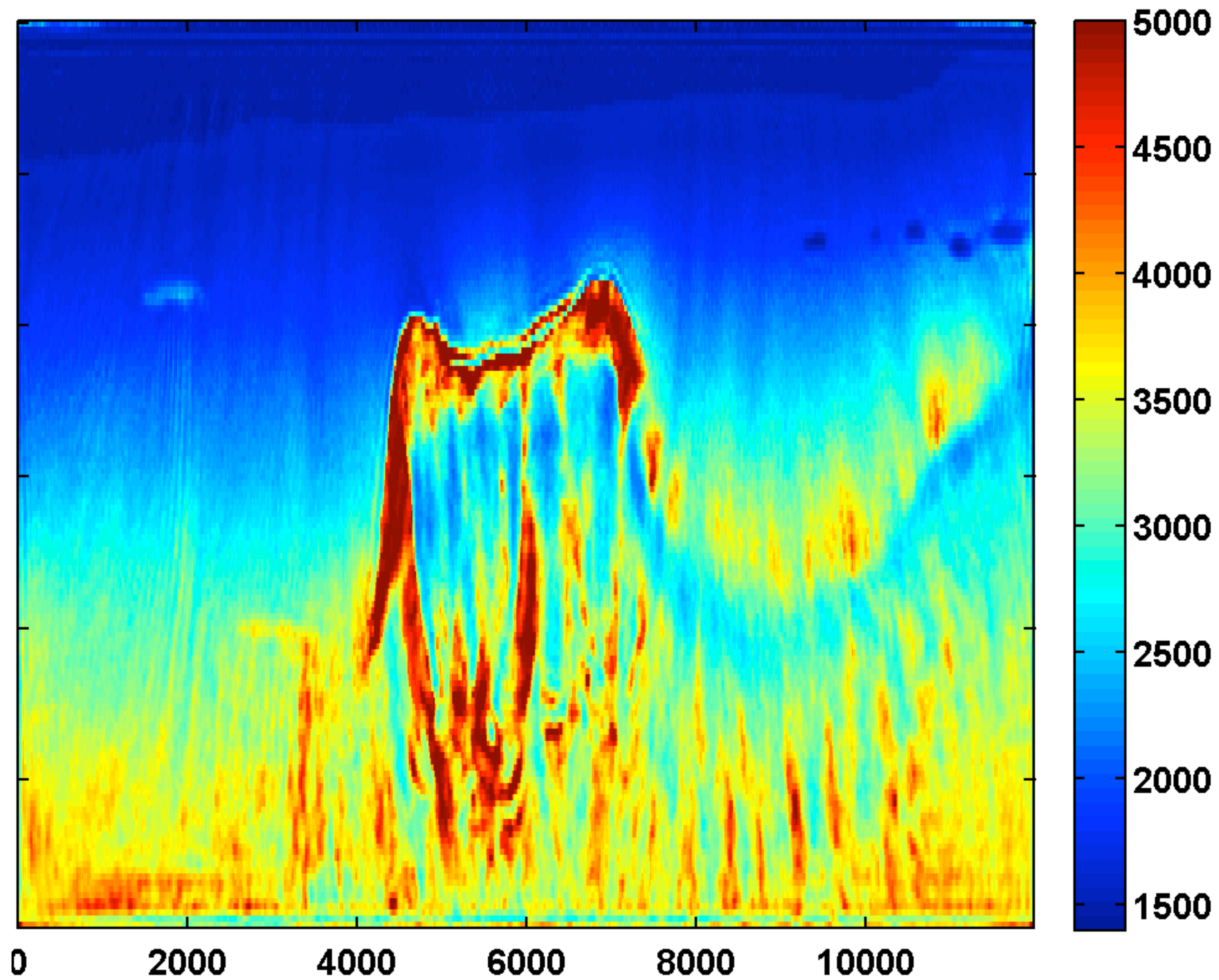


Results w/o TV

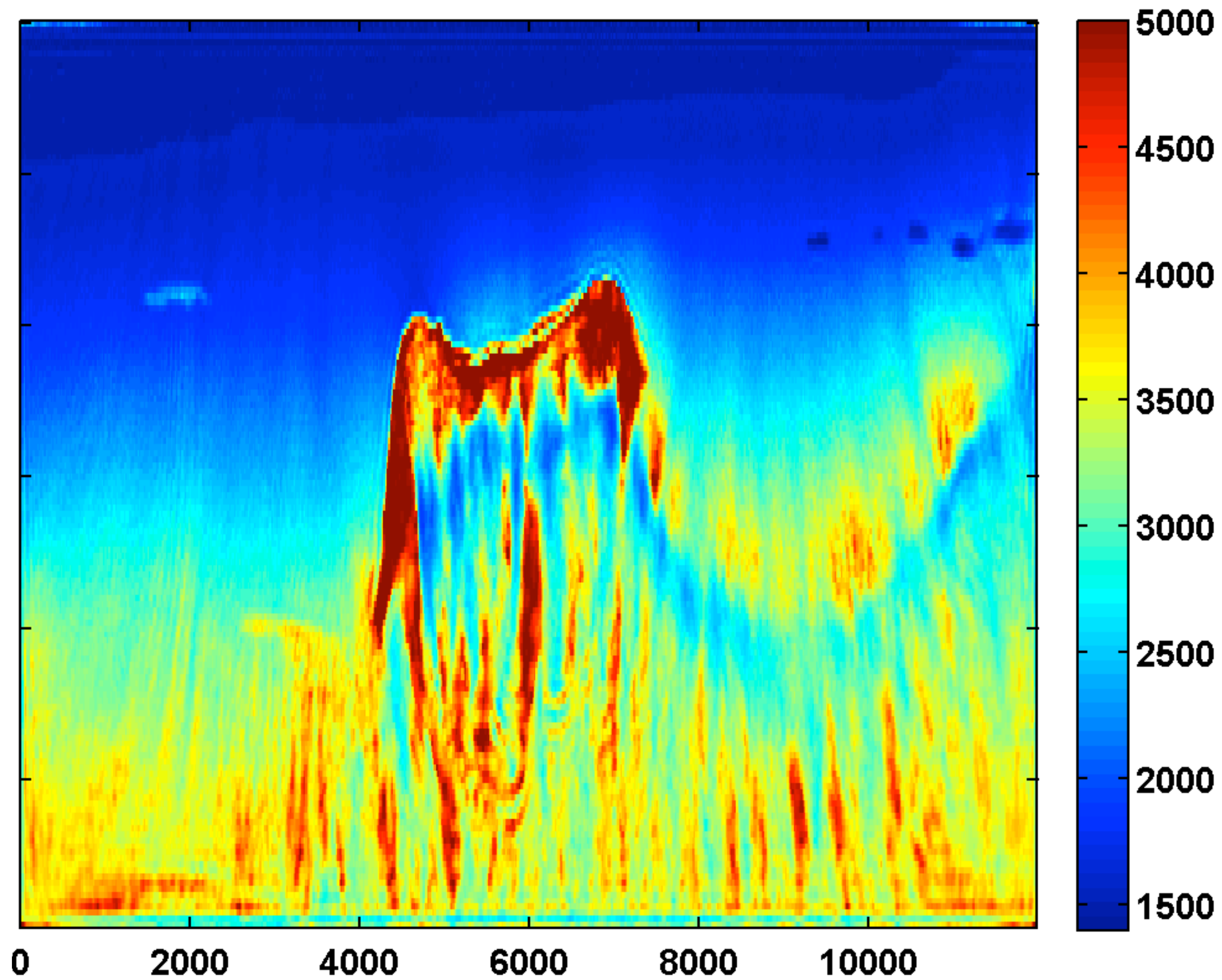
after one cycle through the frequencies



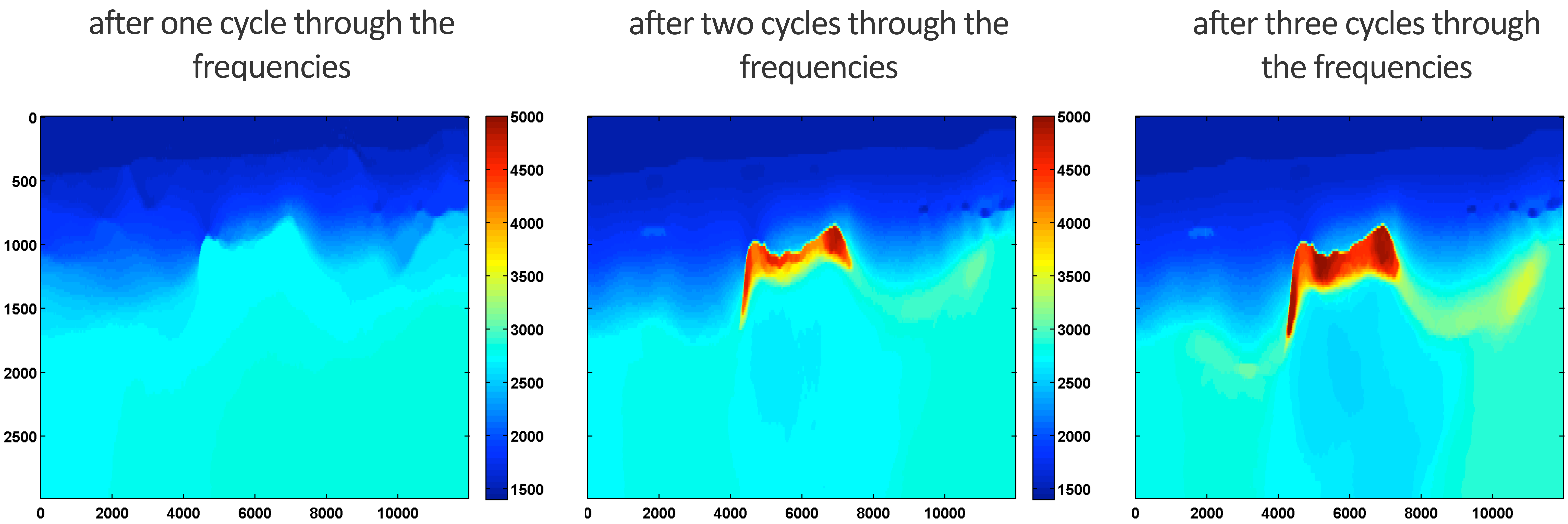
after two cycles through the frequencies



after three cycles through the frequencies



Results w/ TV



Total Variation Regularized Wavefield Reconstruction Inversion. This code implements a scaled gradient projection method to minimize the wavefield reconstruction inversion (WRI) objective subject to total variation and spatially varying bound constraints. For questions contact Ernie Esser. [Read more] [GitHub]

Joint work w/ Sub Salt Solutions Ltd. Patent filed GB1509337.0

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
- making it less probable that velocities step 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.

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

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

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

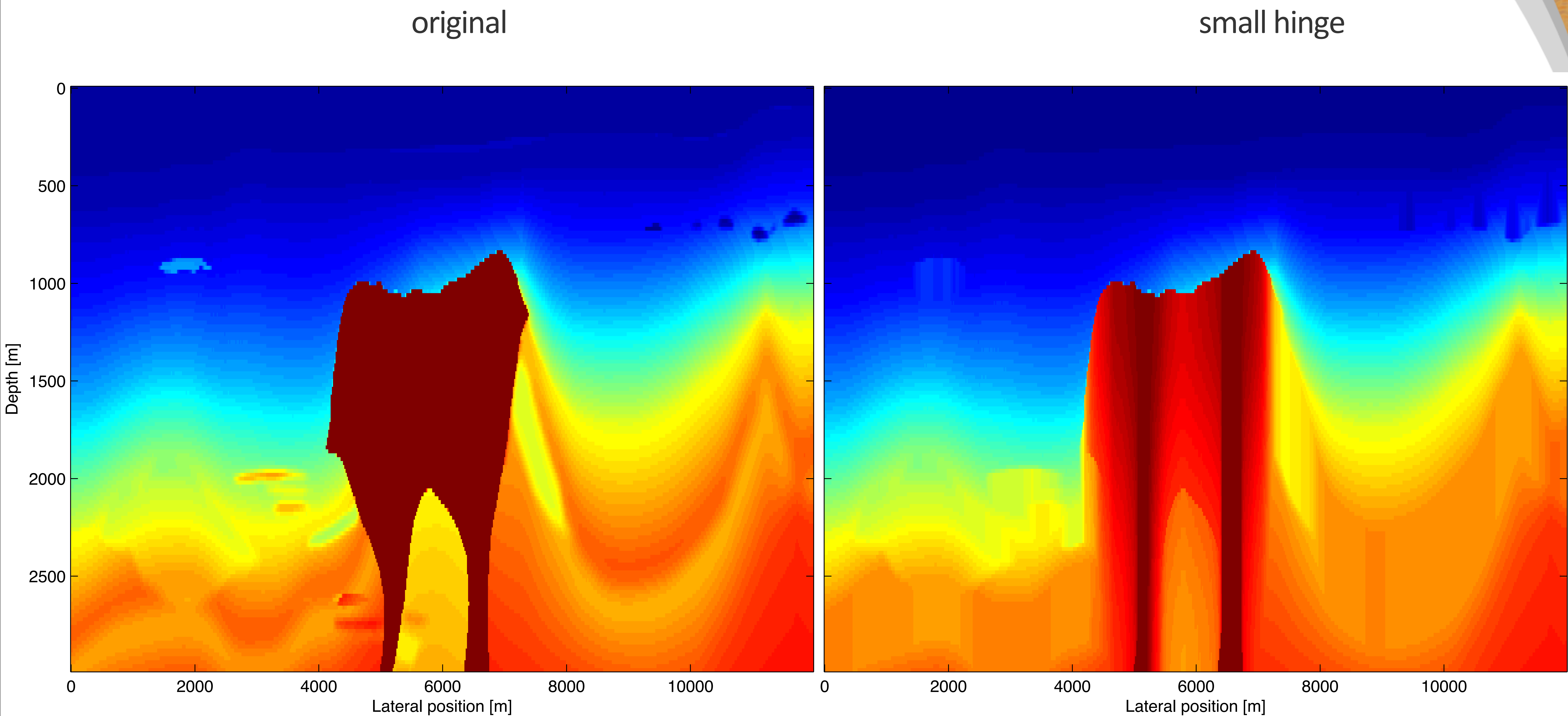
with

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

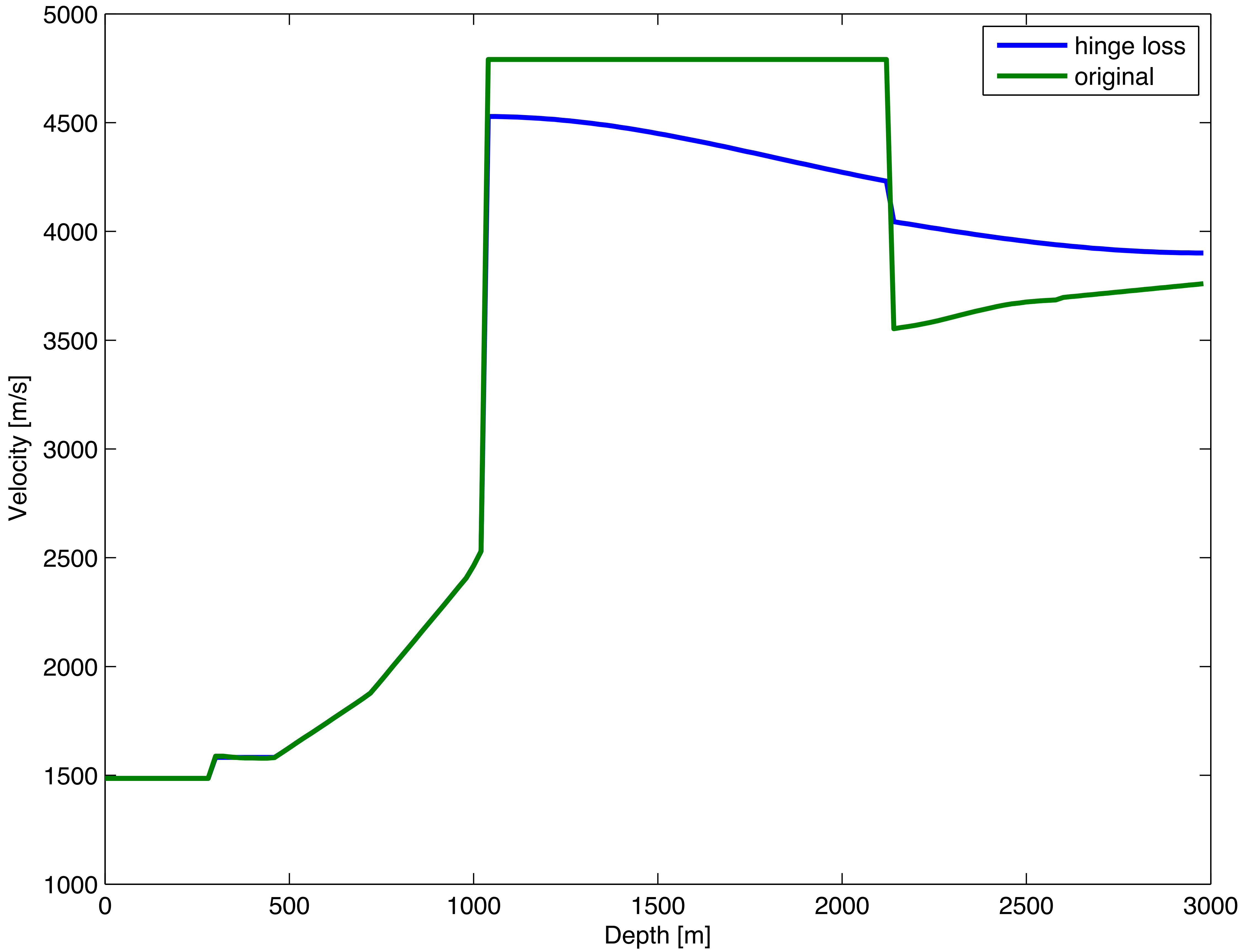
and

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

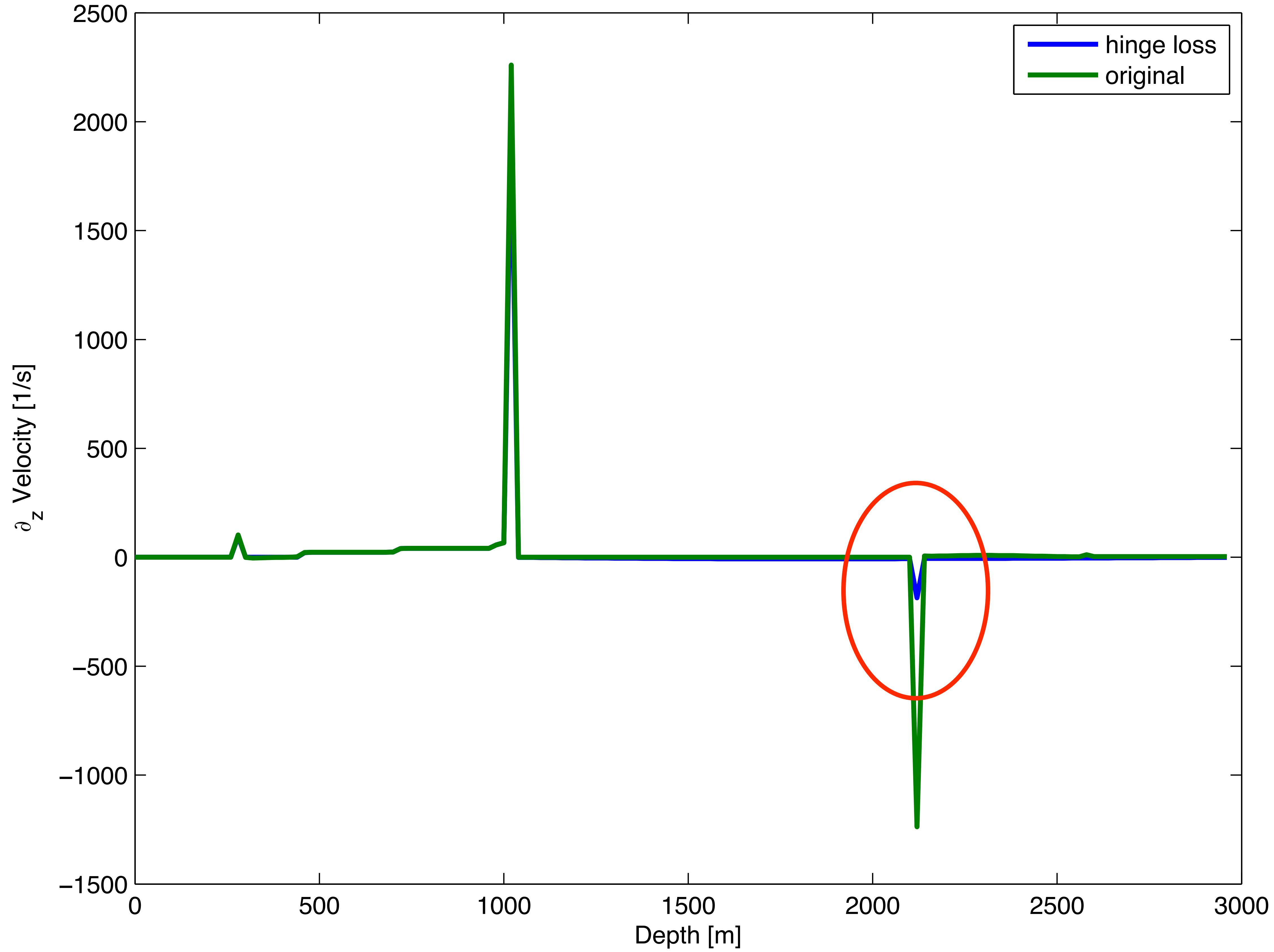
Hinge loss


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

Hinge loss



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



hinge loss
original
∂z Velocity [1/s]
Depth [m]
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

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

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

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

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

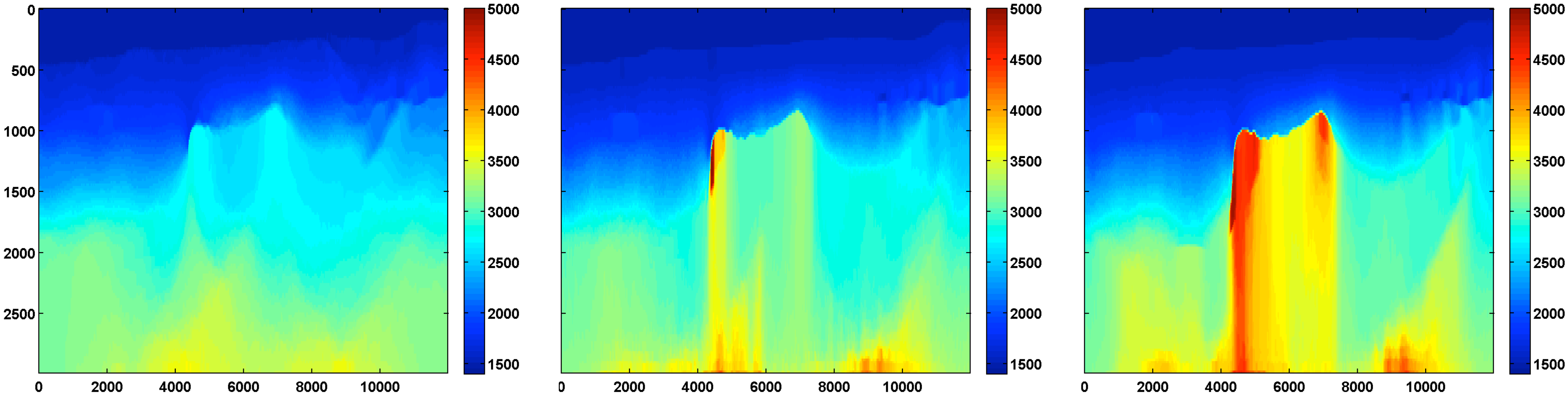
Results w/ hinge loss continuation

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

after one cycle through the frequencies

after two cycles through the frequencies

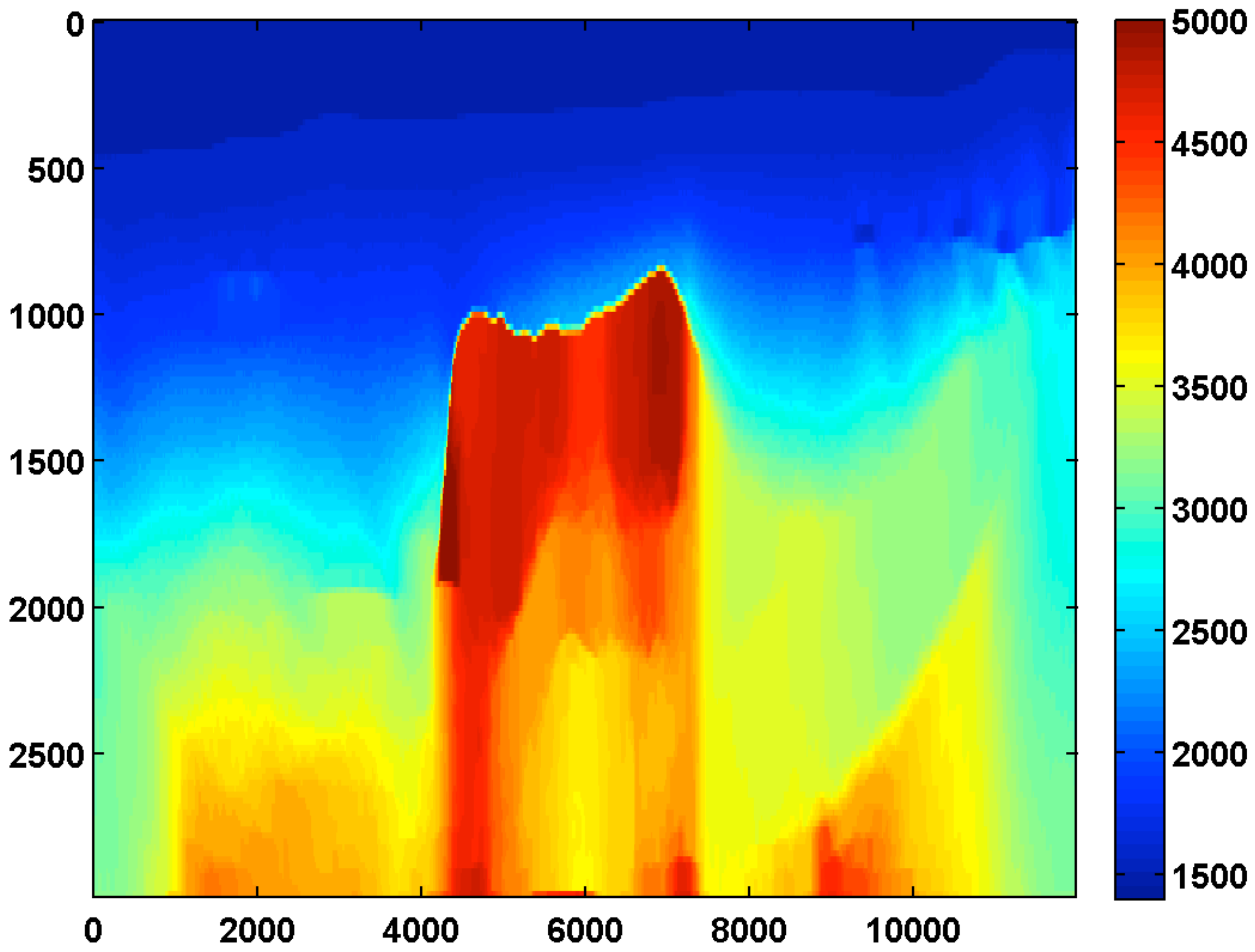
after three cycles through the frequencies



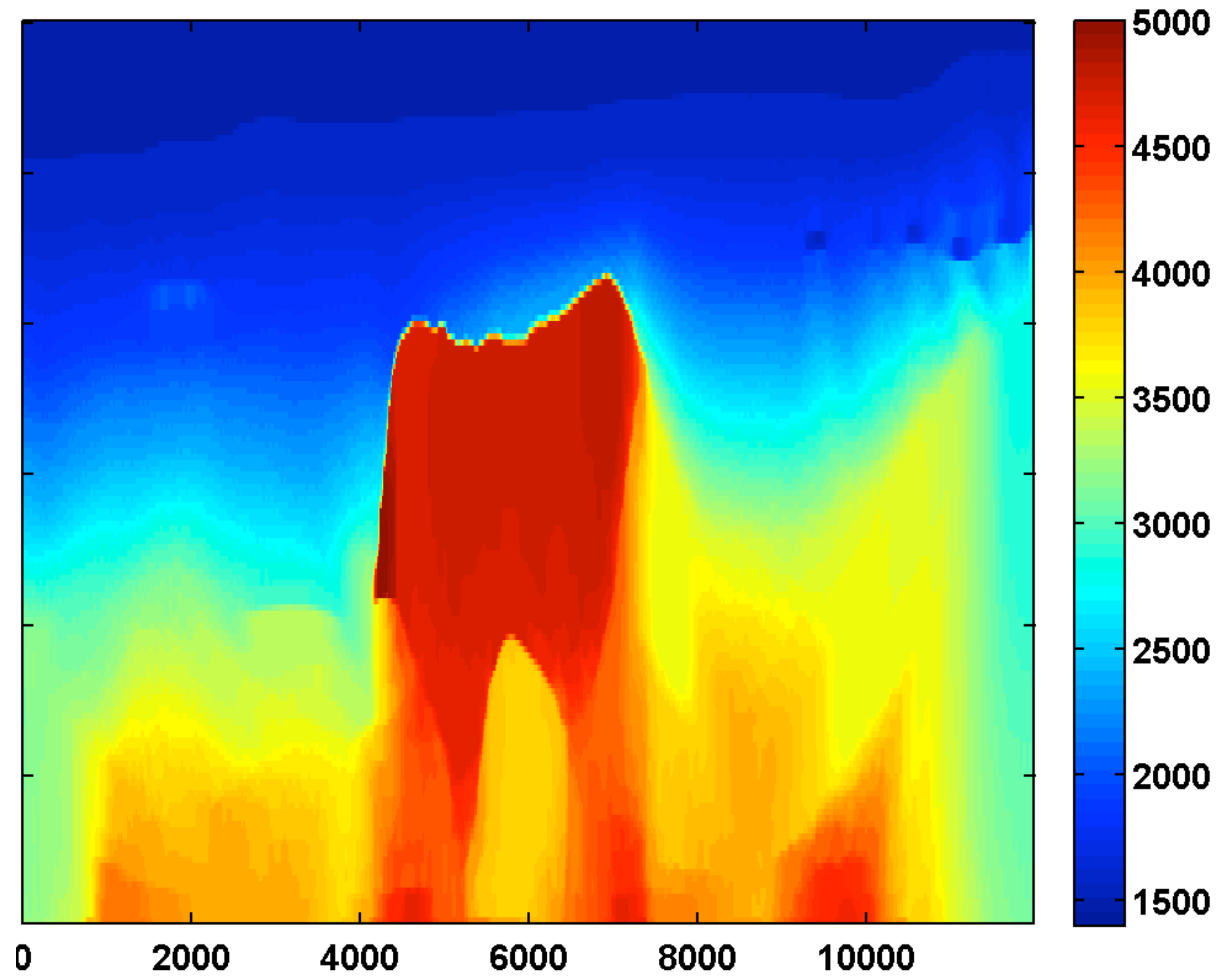
Results w/ hinge loss continuation

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

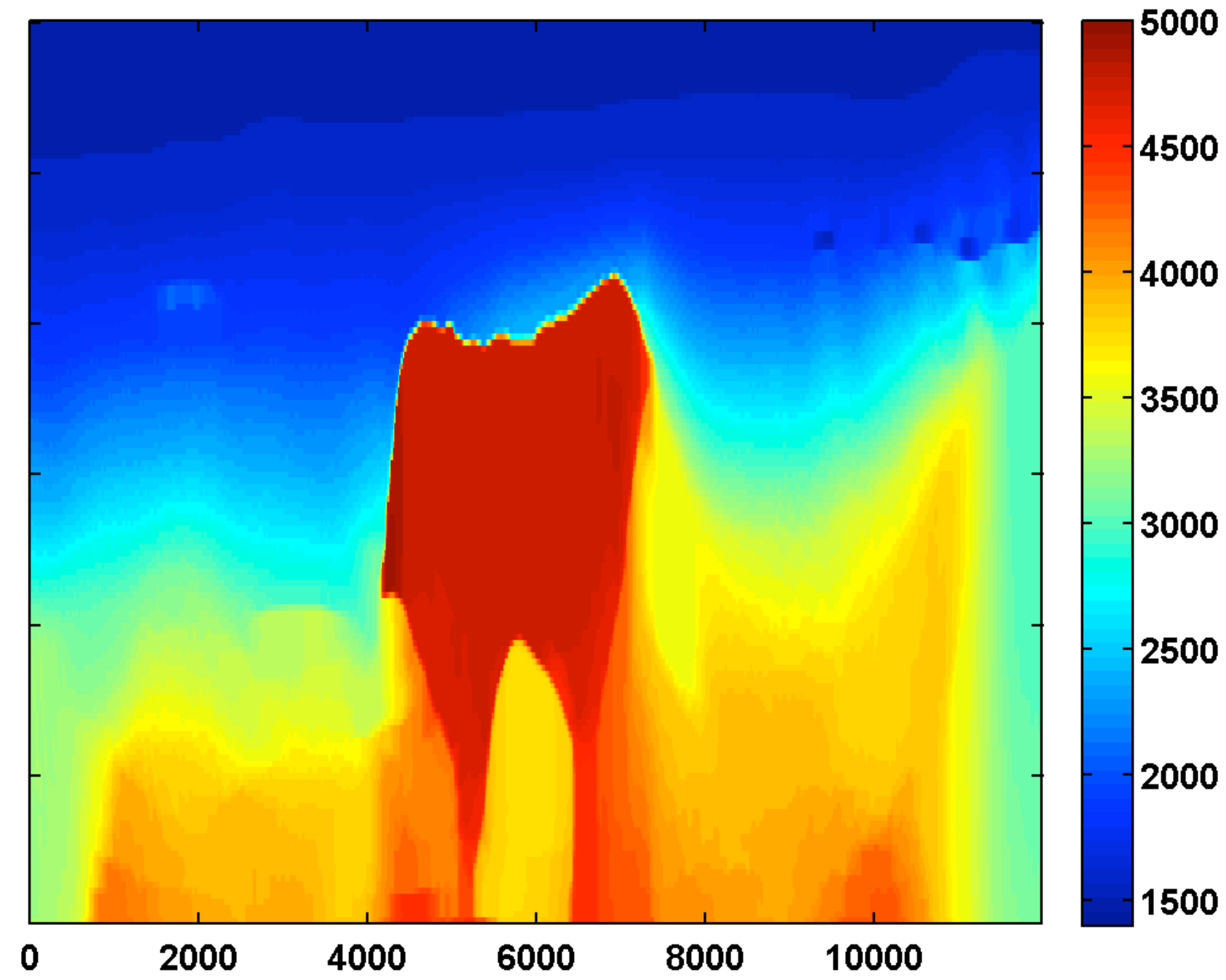
after four cycles through the frequencies



after five cycles through the frequencies

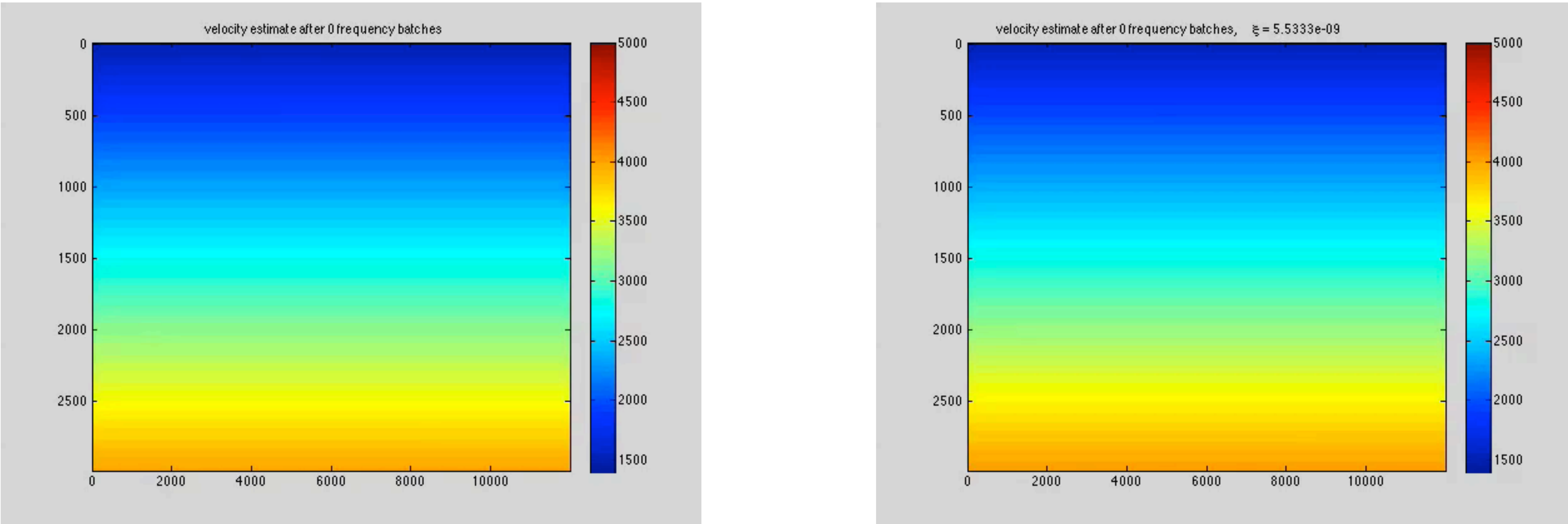


after six cycles through the frequencies



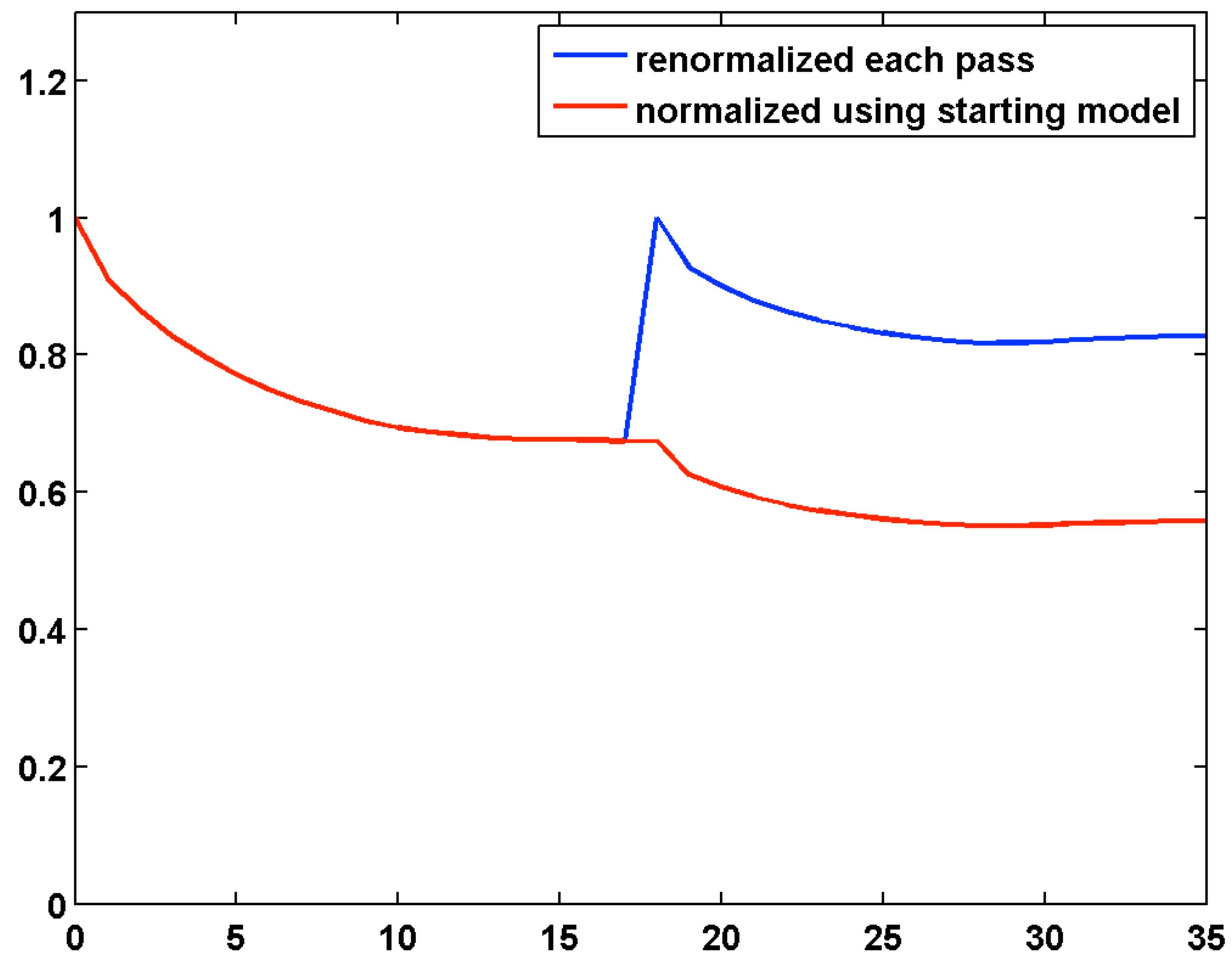
WRI

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

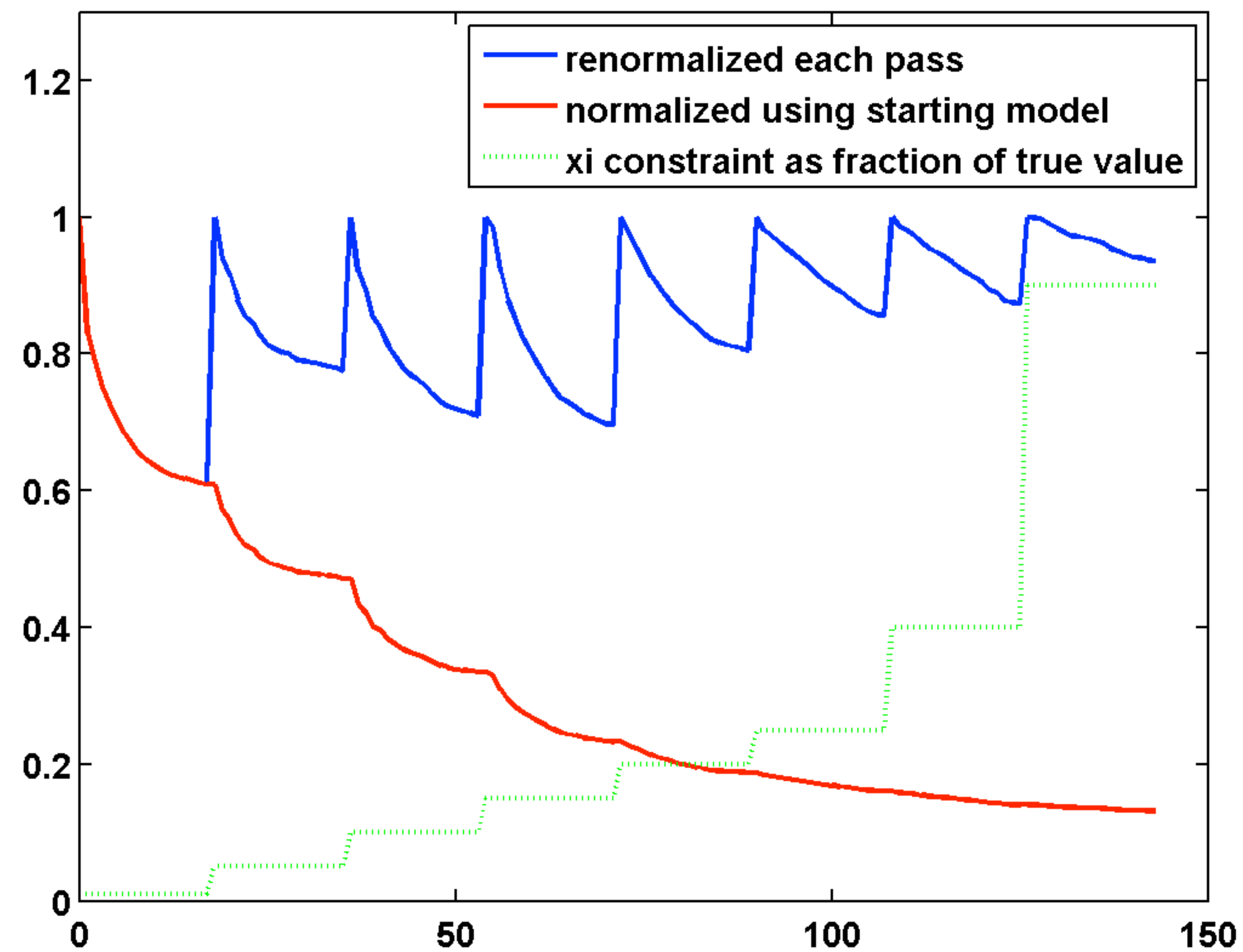


Relative model errors

w/o TV

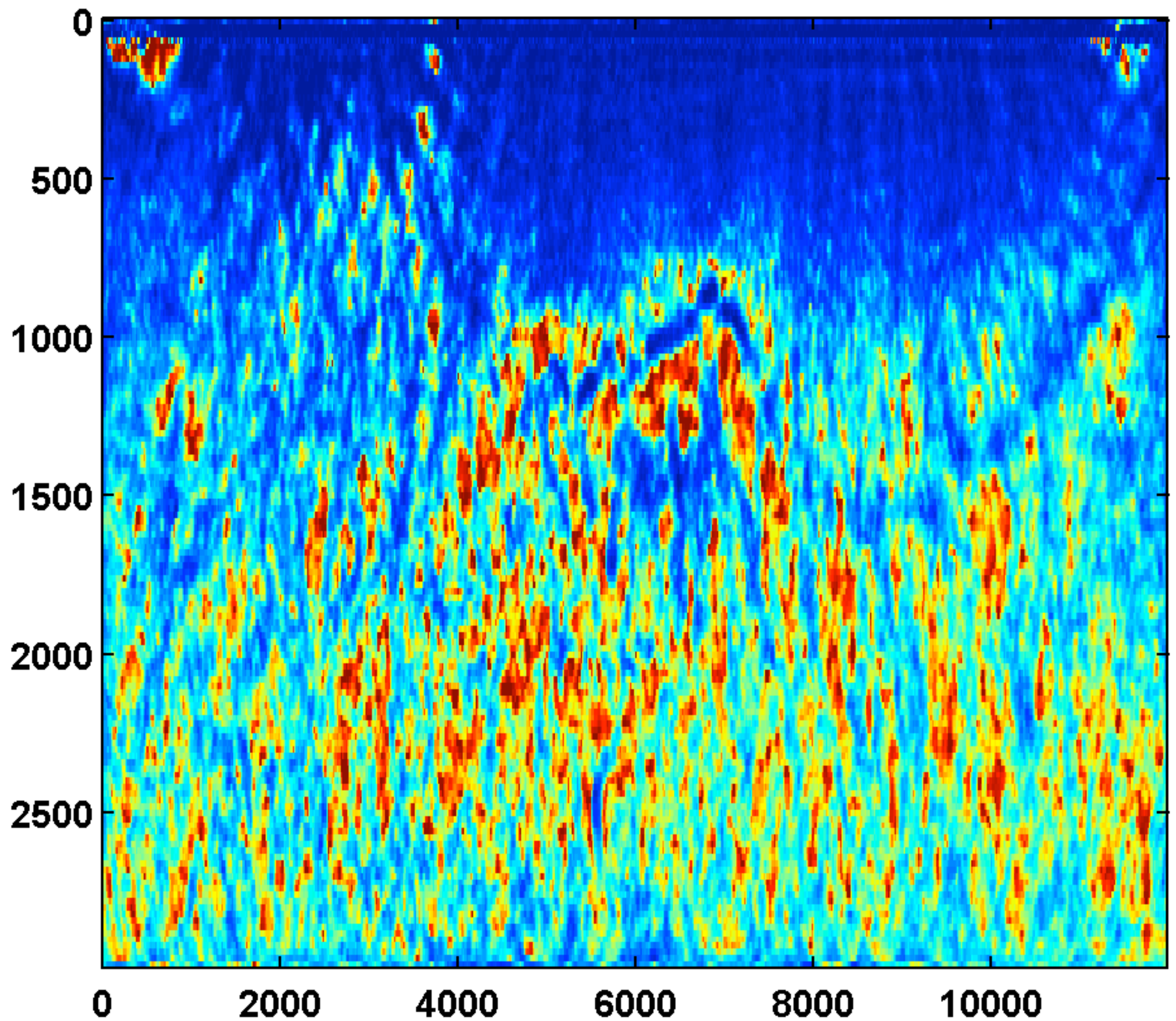


w/ TV & hinge

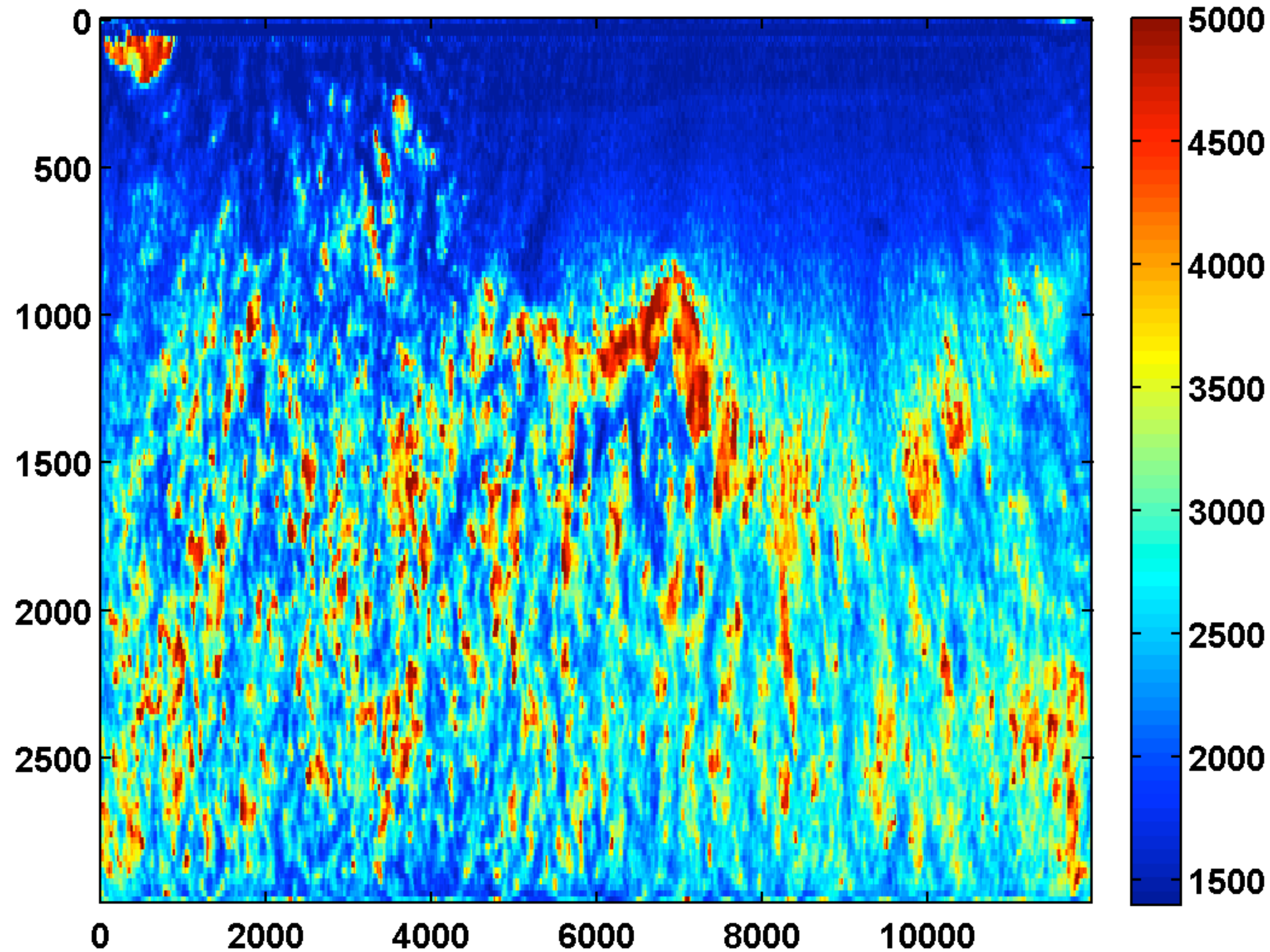


Adjoint-state w/o TV

After one cycle through the frequencies



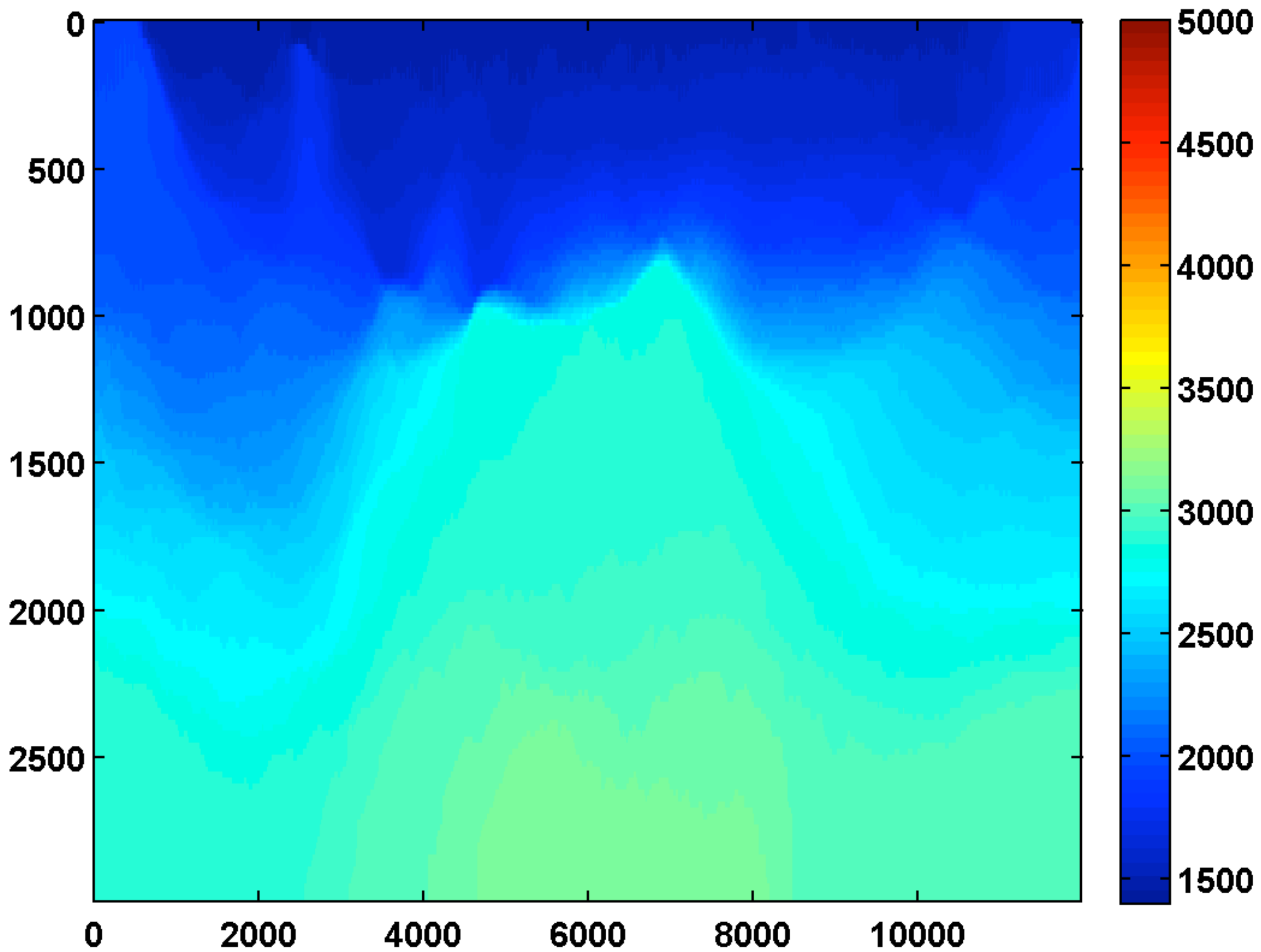
After two cycles through the frequencies



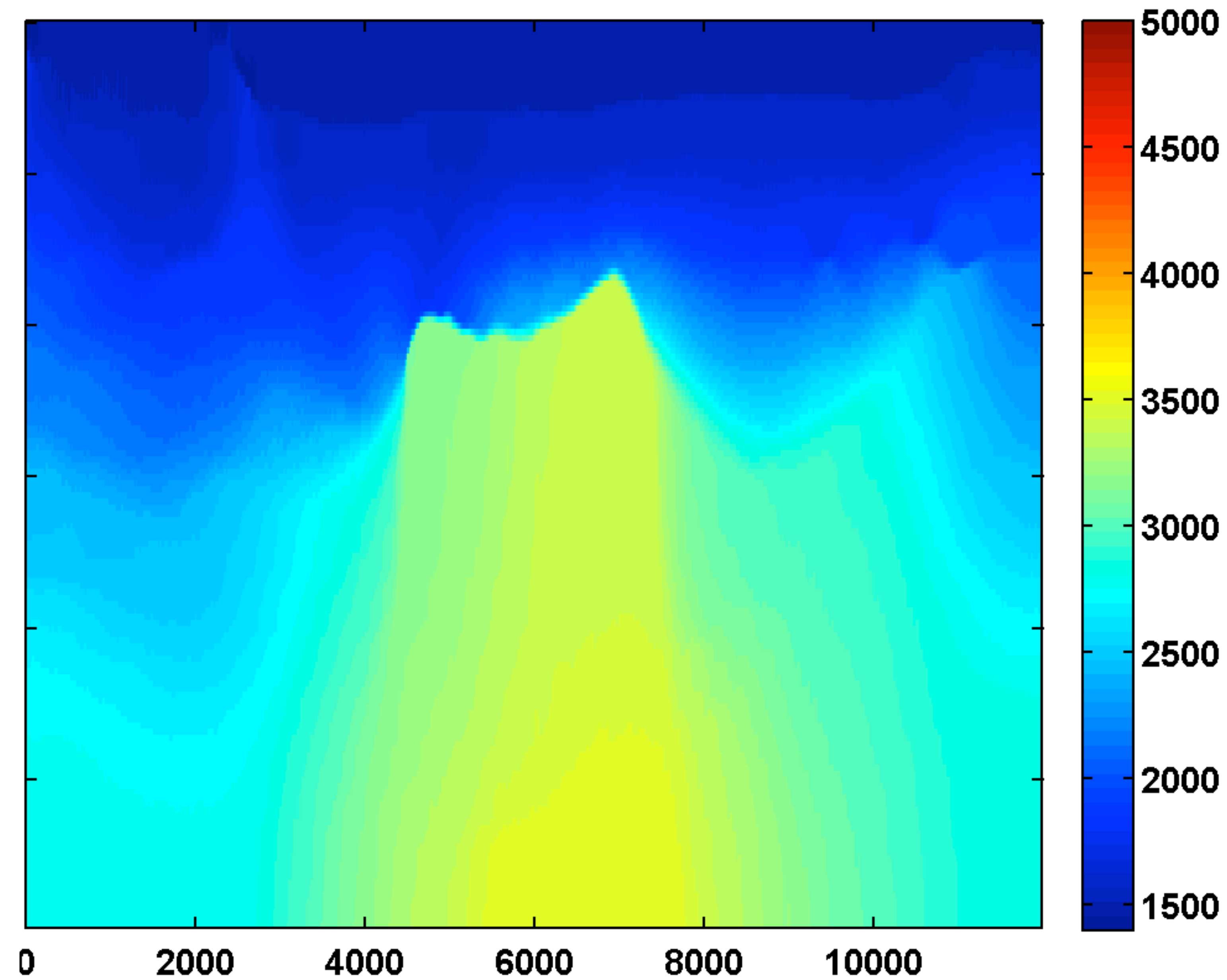
Adjoint-state w/ hinge loss continuation

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

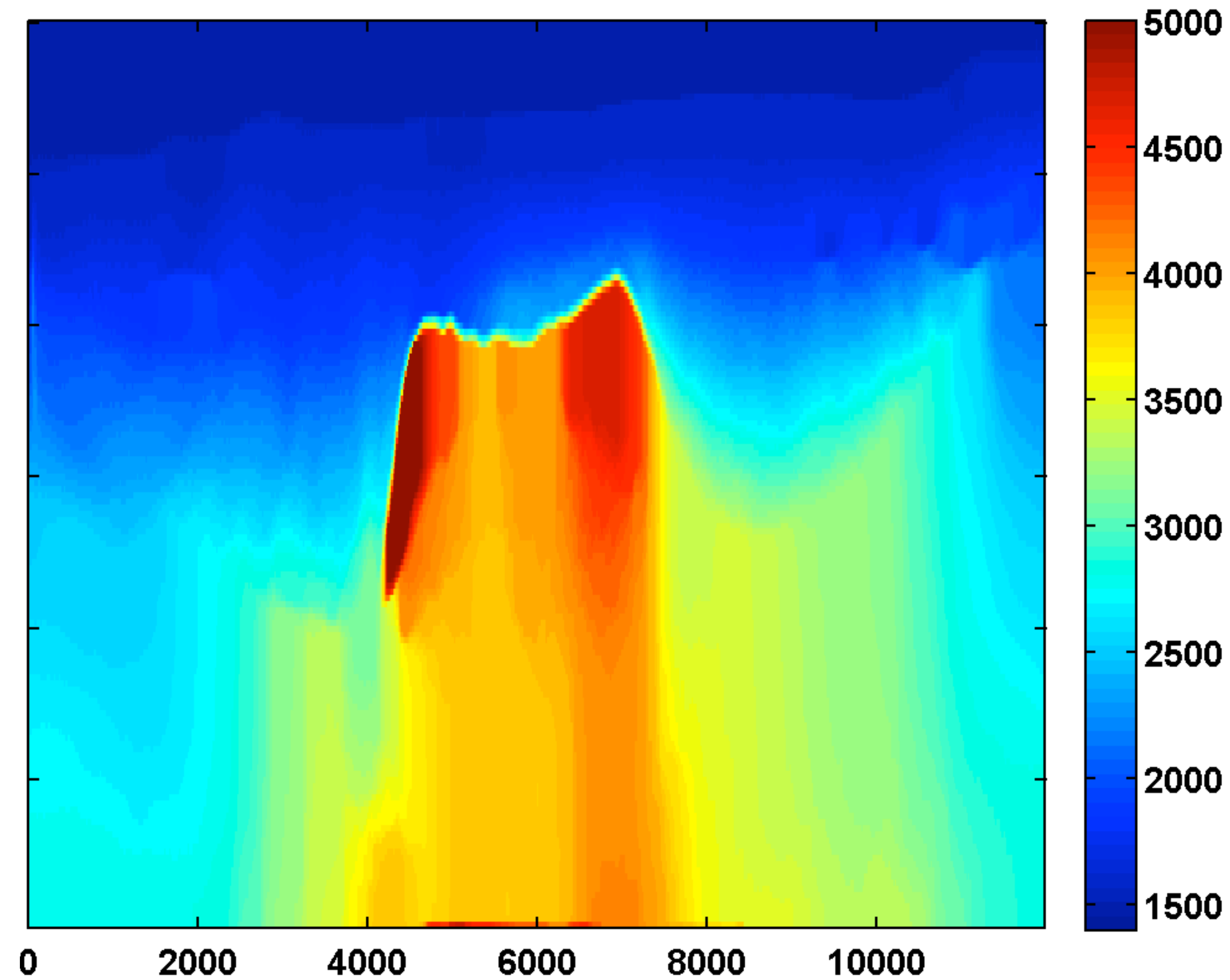
after one cycle through the frequencies



after two cycles through the frequencies



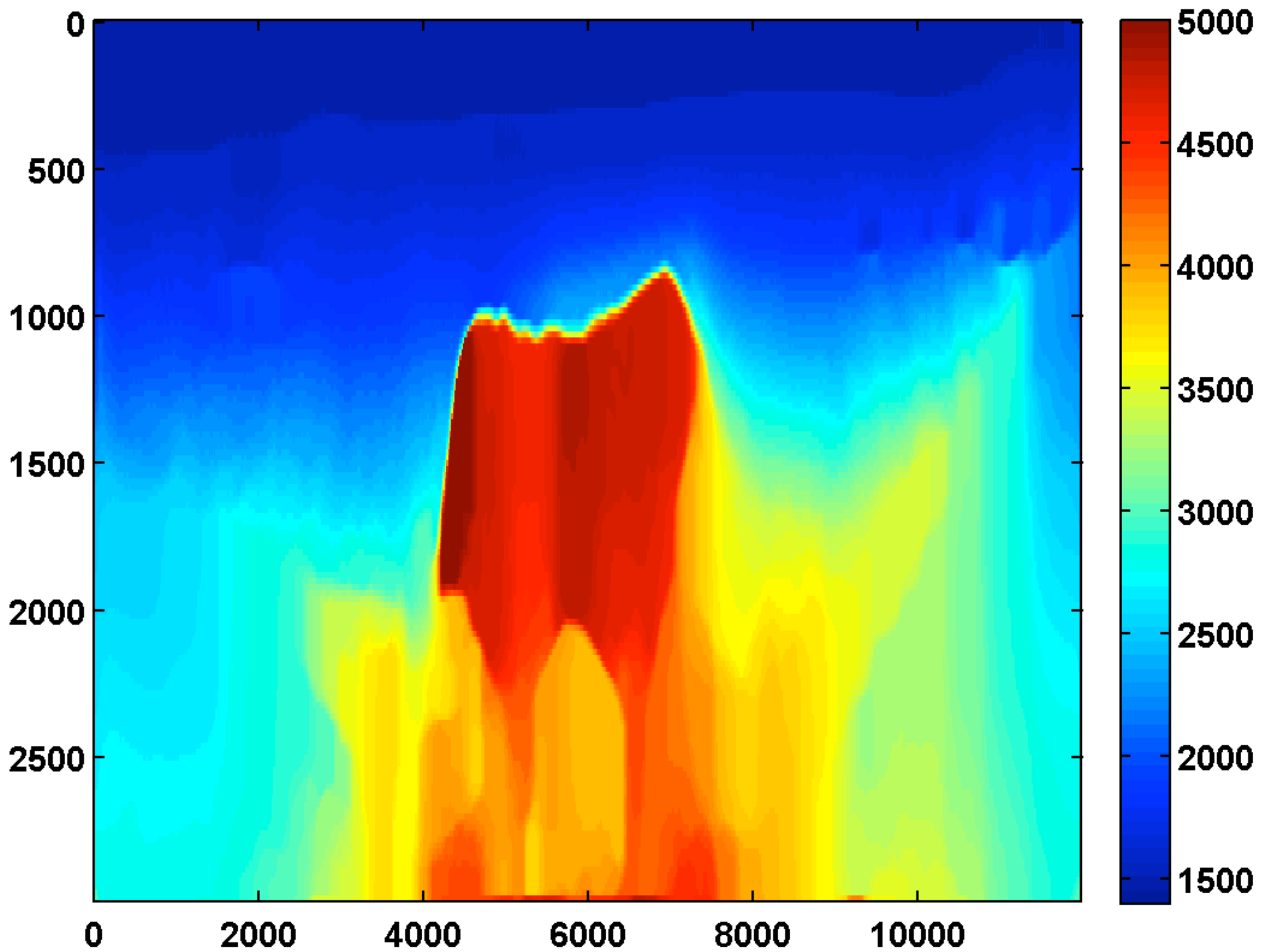
after three cycles through the frequencies



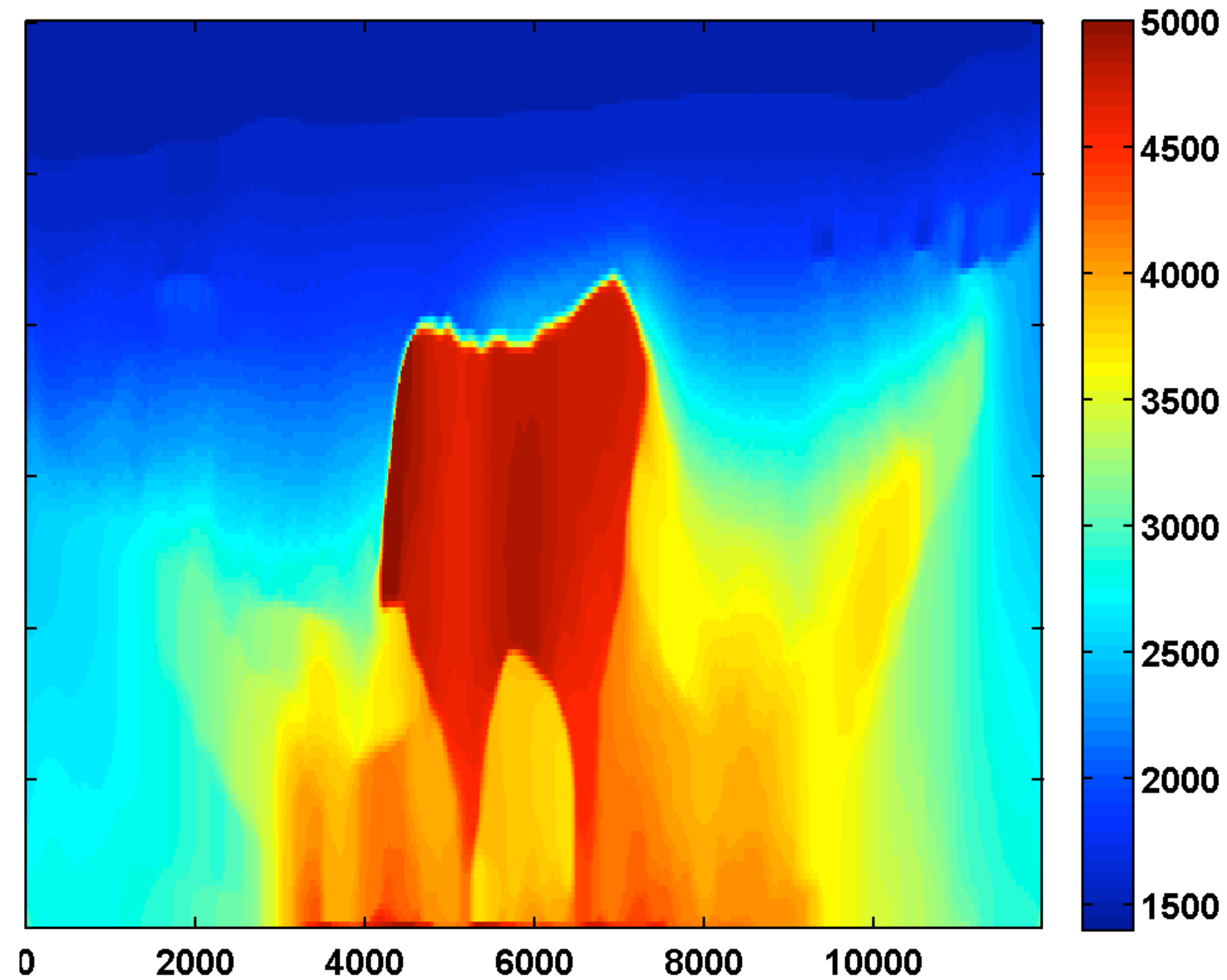
Adjoint-state w/ hinge loss continuation

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

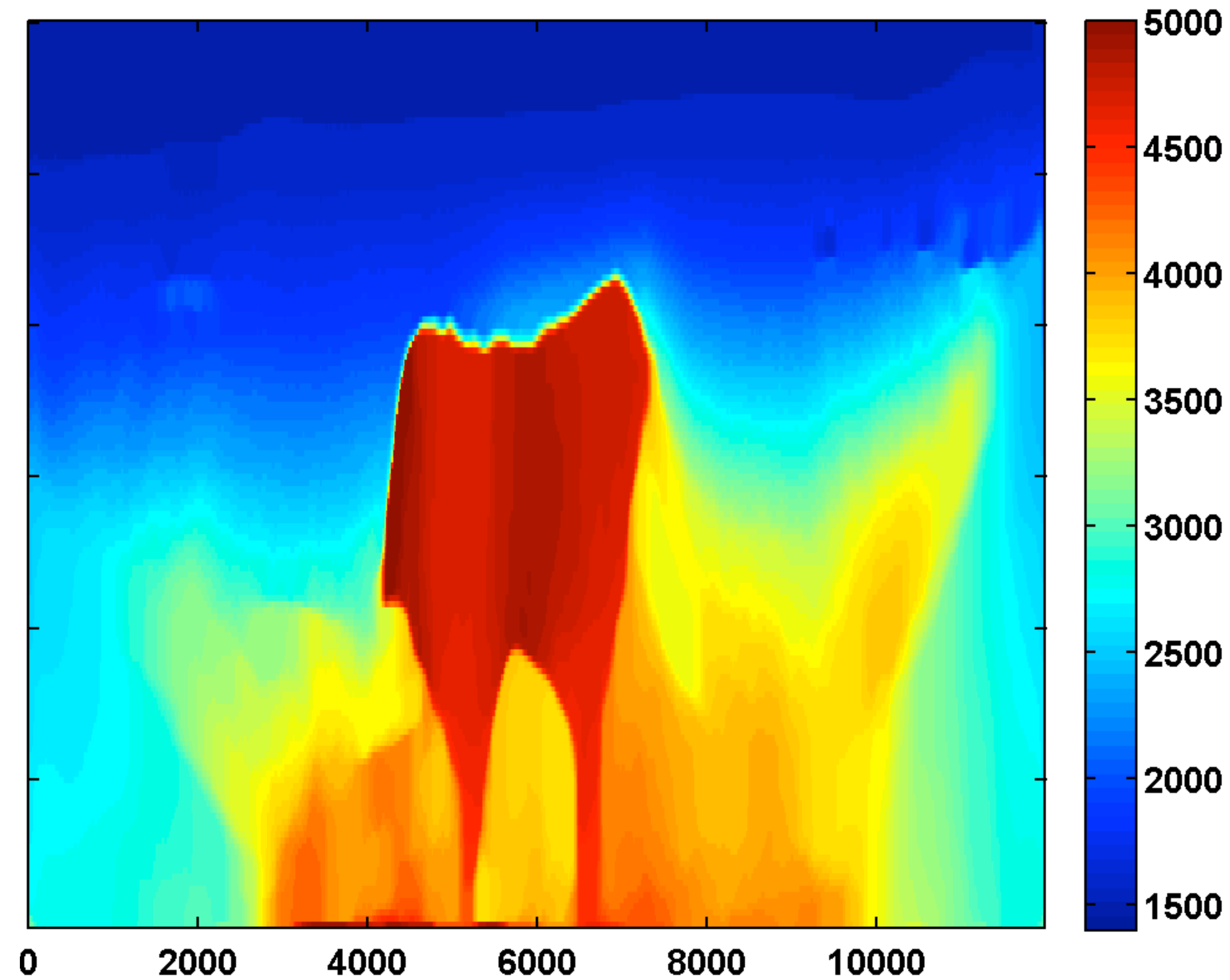
after four cycles through the frequencies



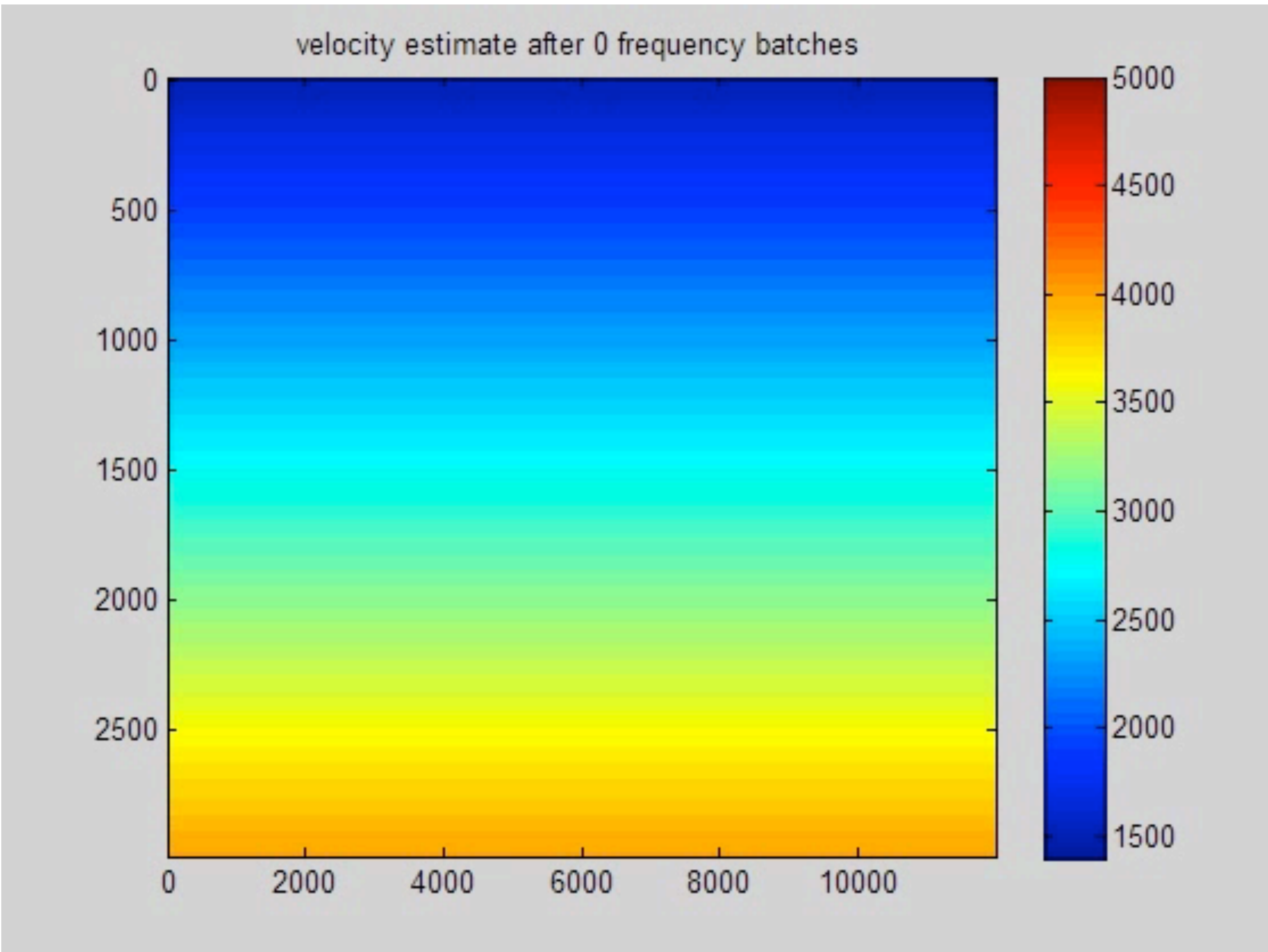
after five cycles through the frequencies

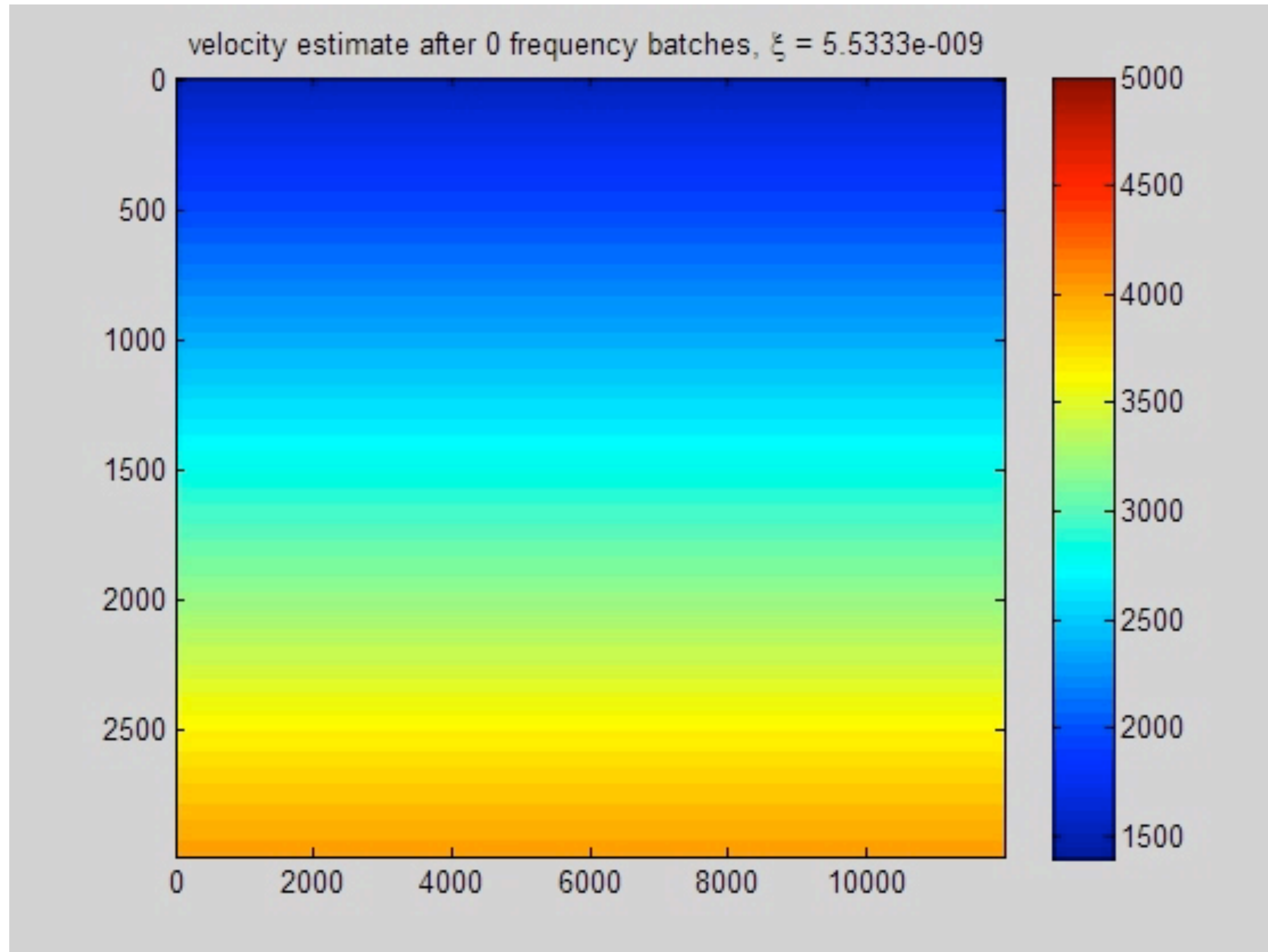


after six cycles through the frequencies

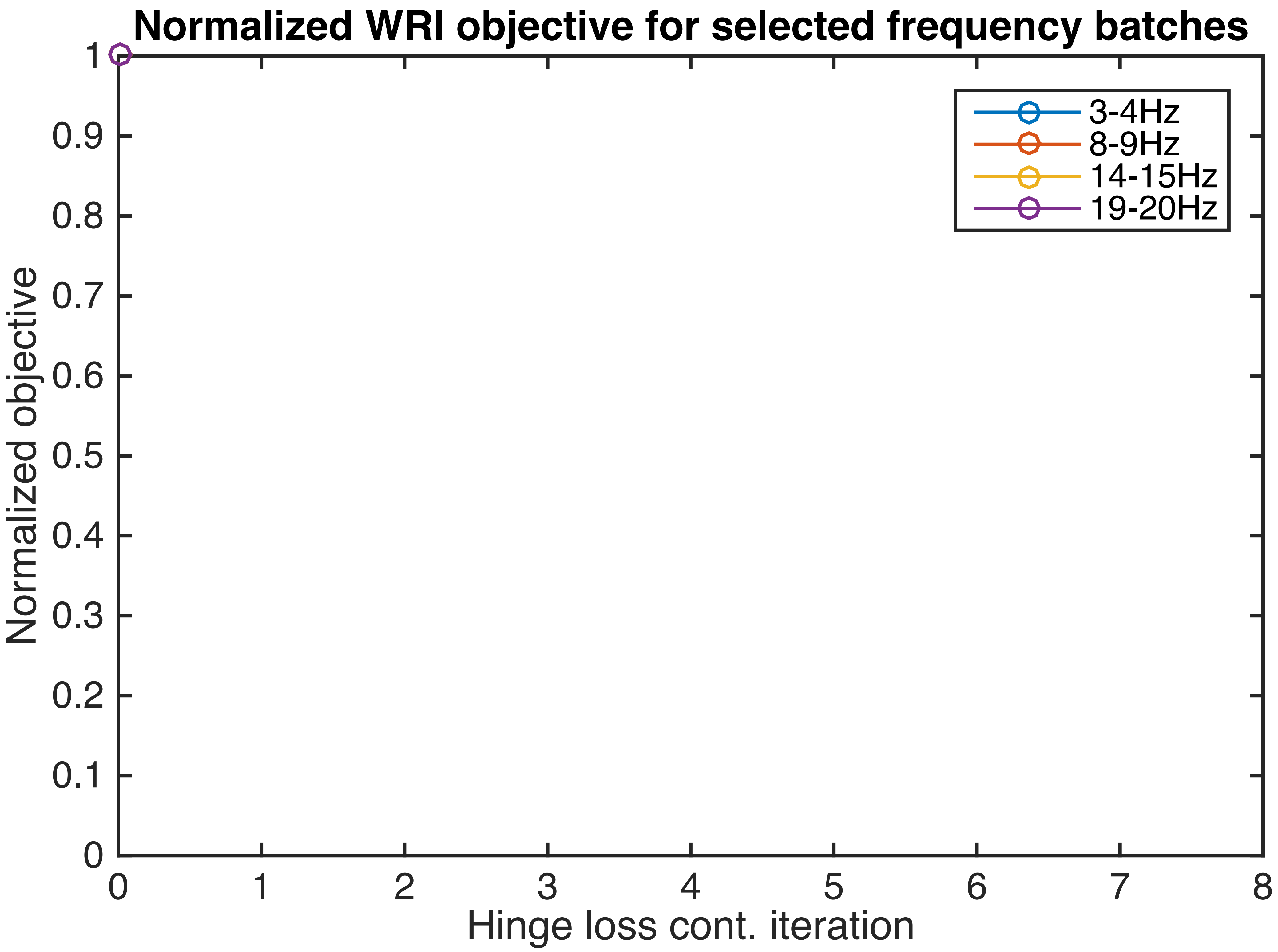


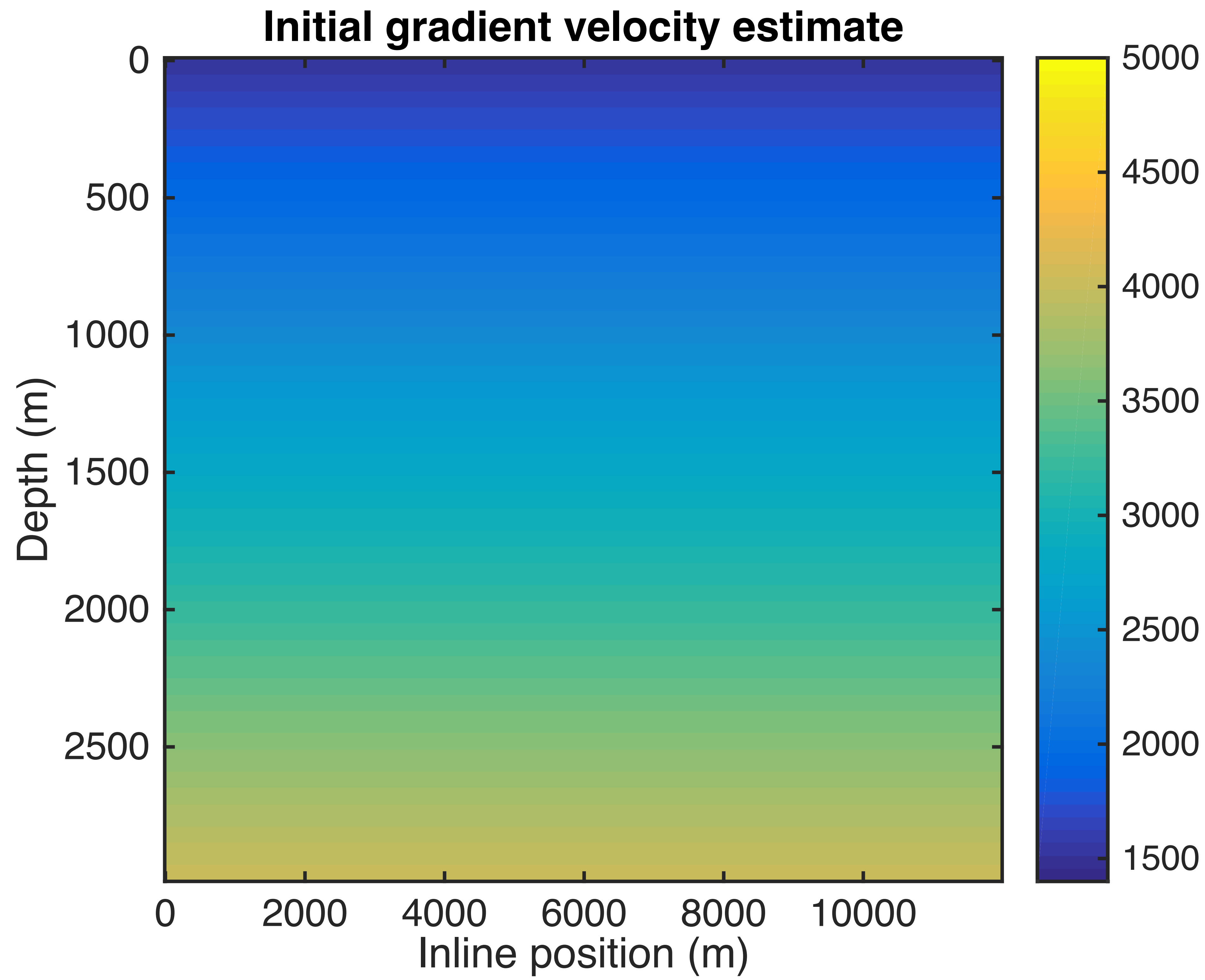
Adjoint-state FWI
w/ or w/o TV-norm & hinge-loss projections & poor starting model





Objective convergence with increasing HL constraint

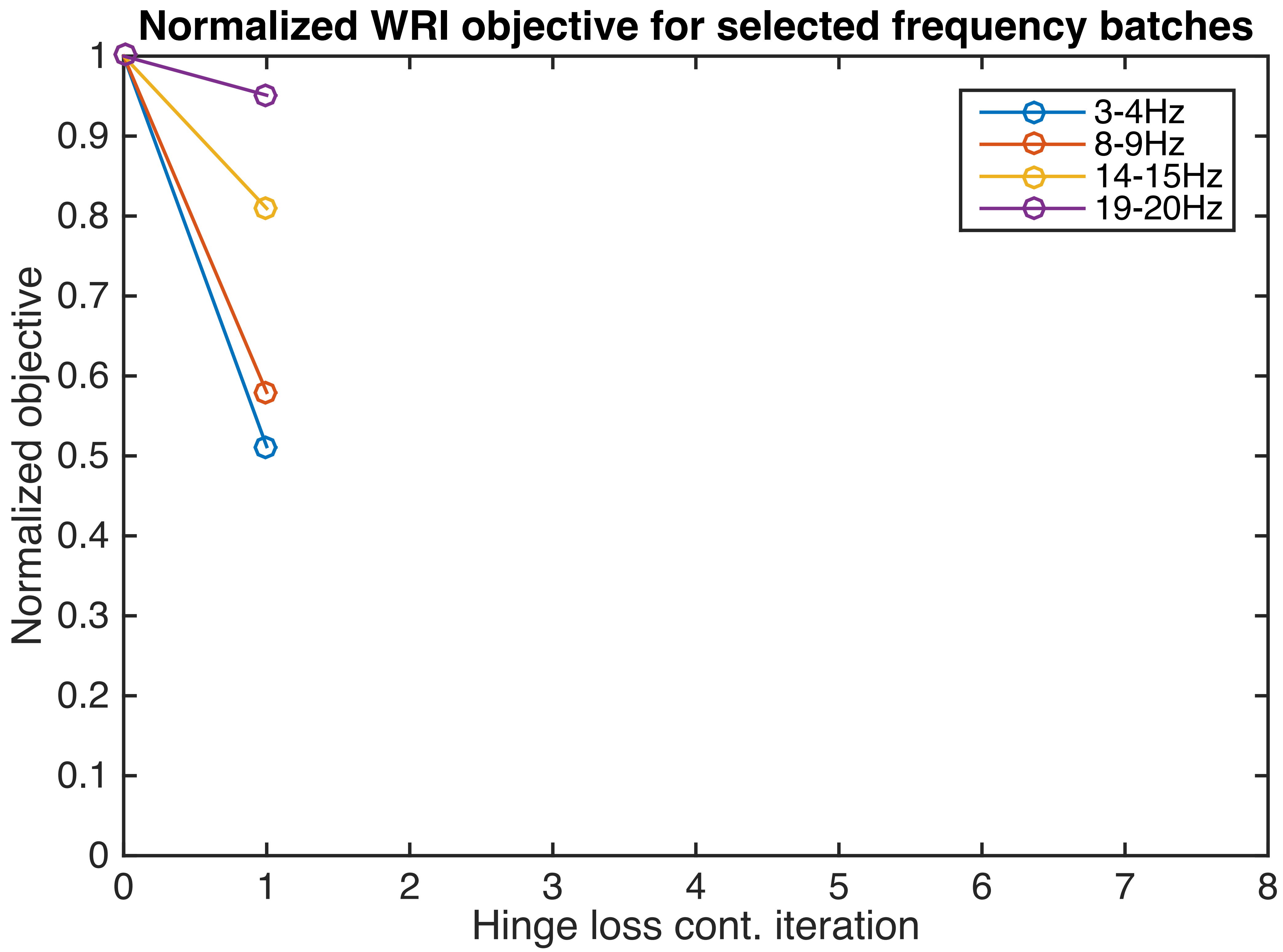


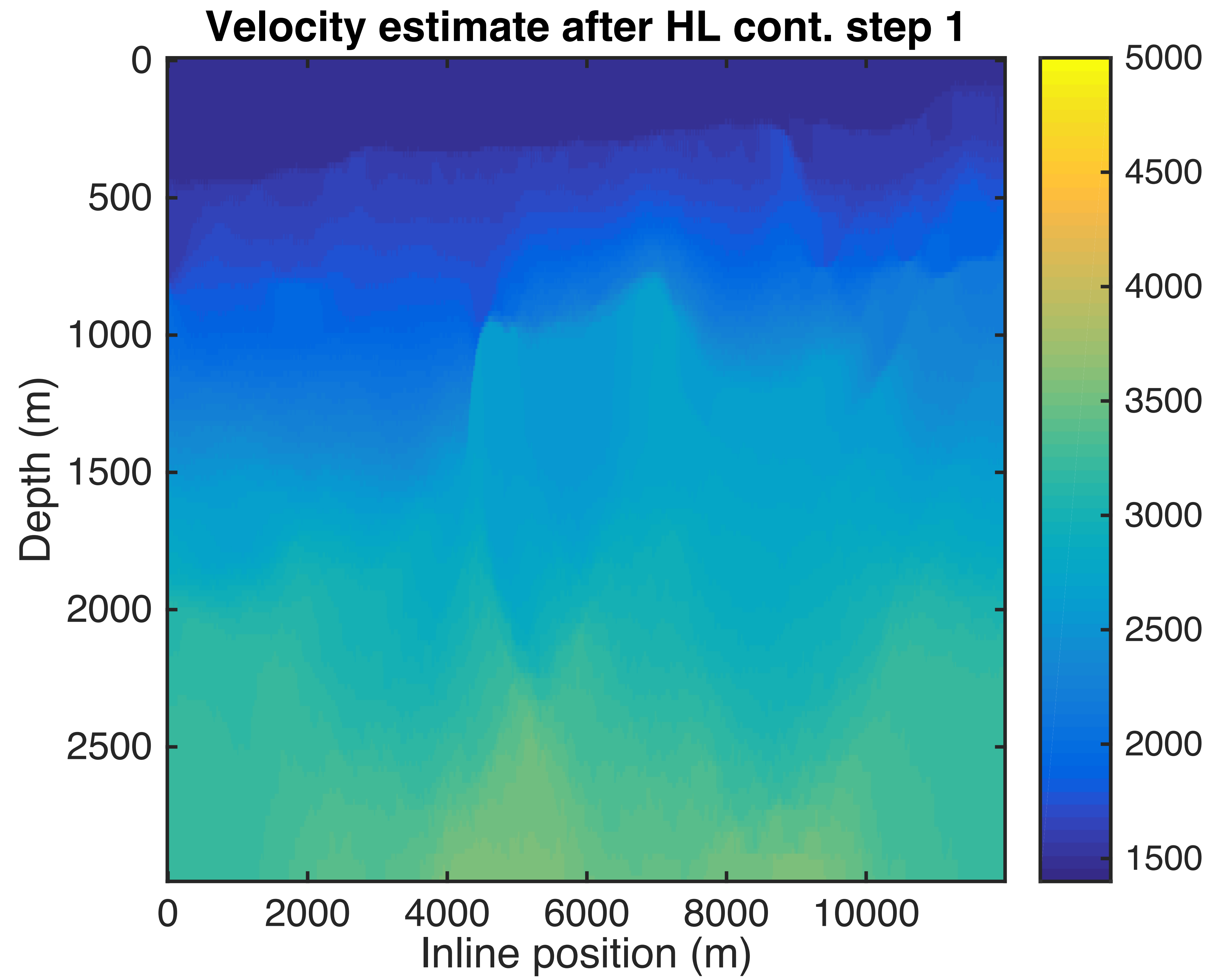


Objective convergence with increasing HL constraint

HL constraint: 0.01 of true

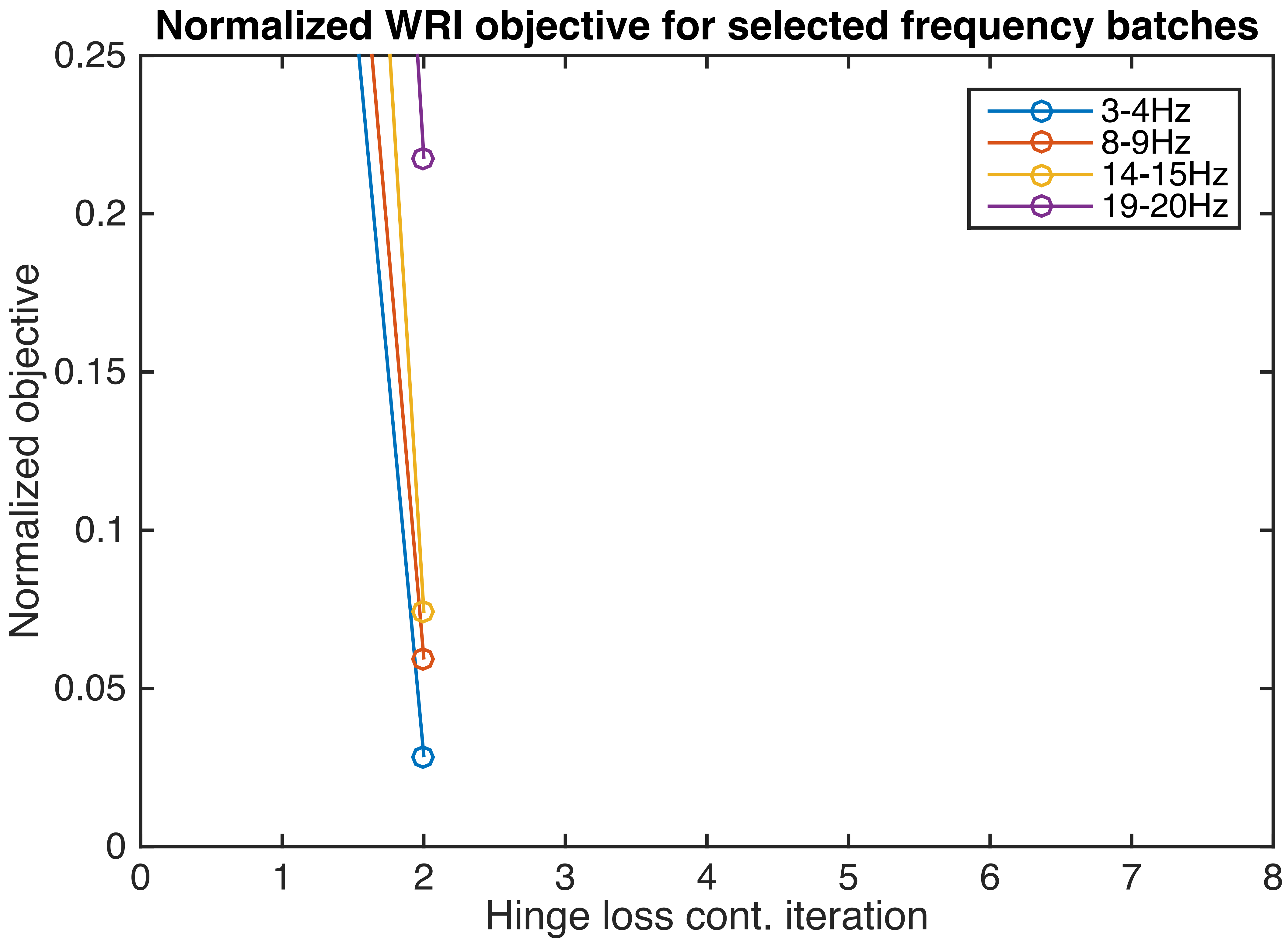
TV constraint: 0.90 of true

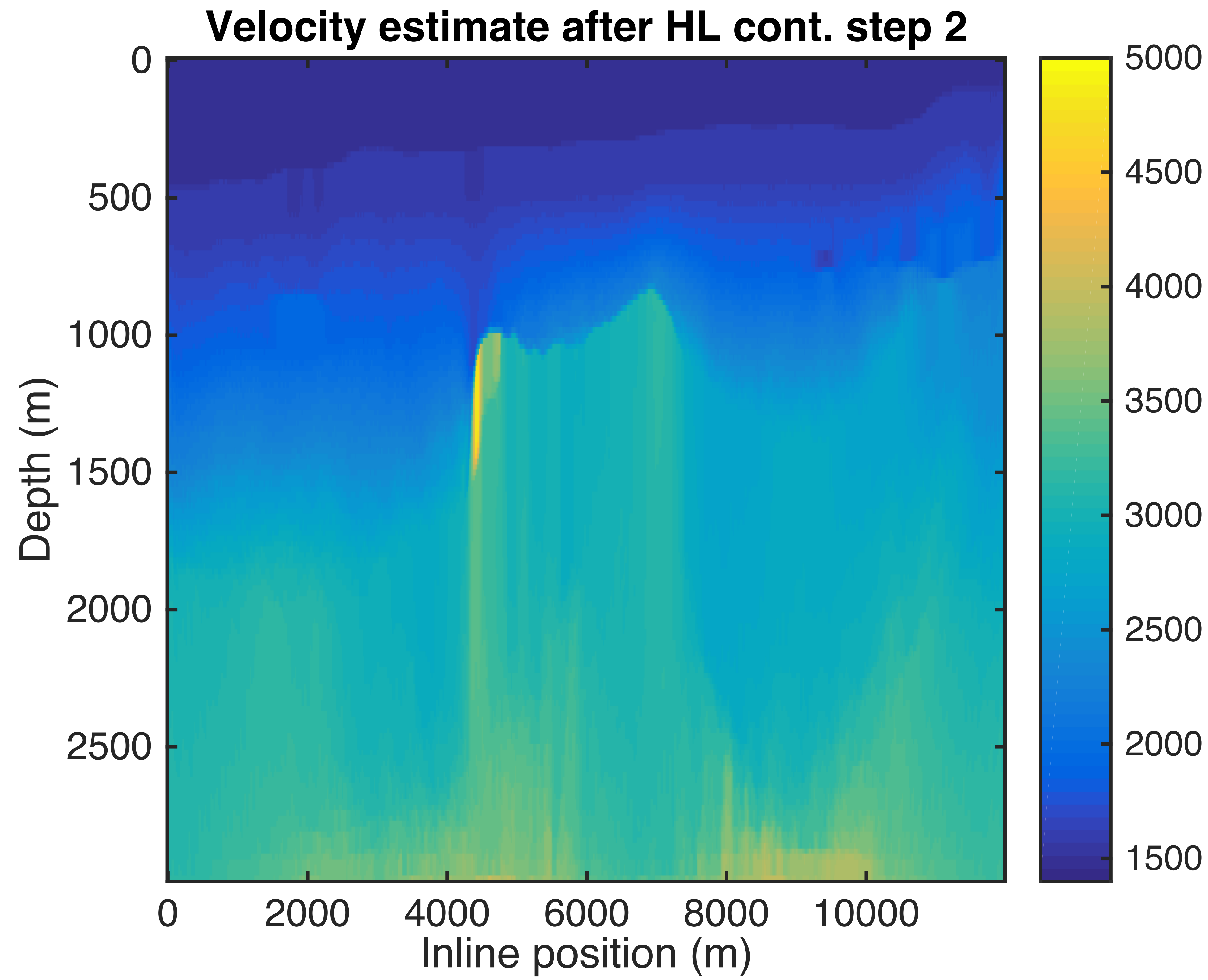




Objective convergence with increasing HL constraint

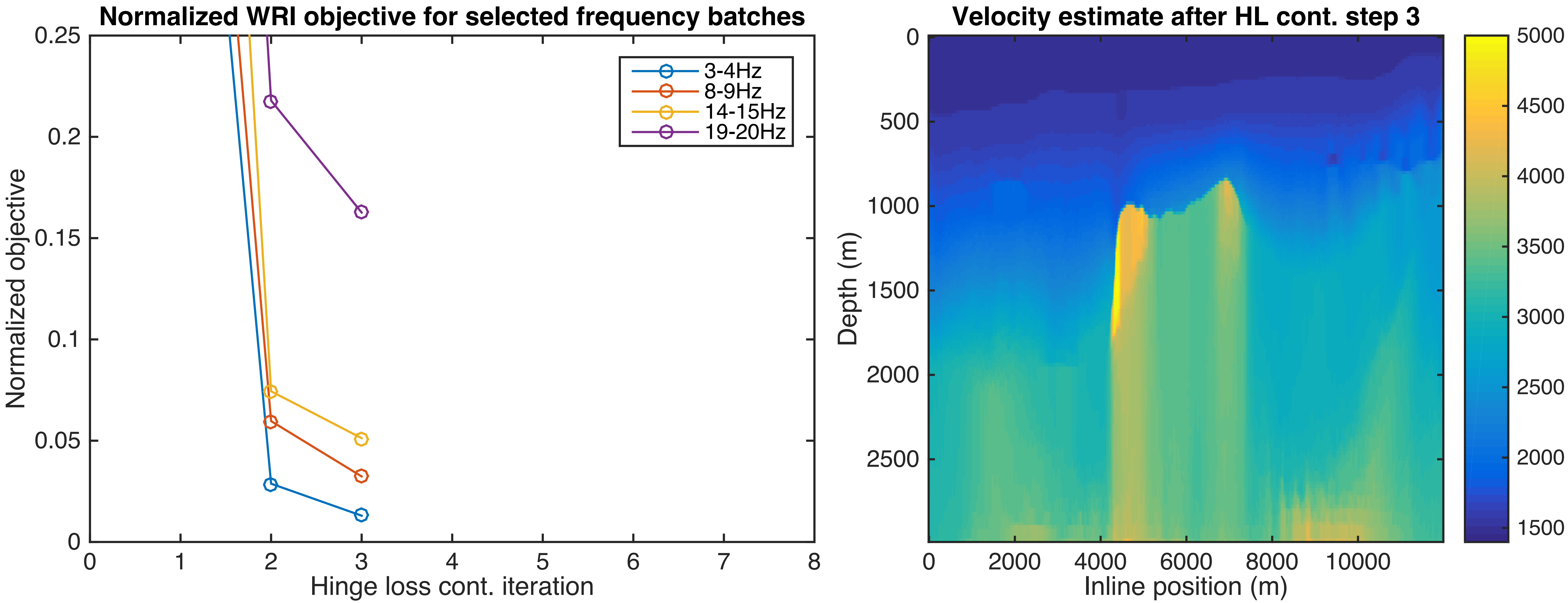
HL constraint: 0.05 of true
TV constraint: 0.90 of true





Objective convergence with increasing HL constraint

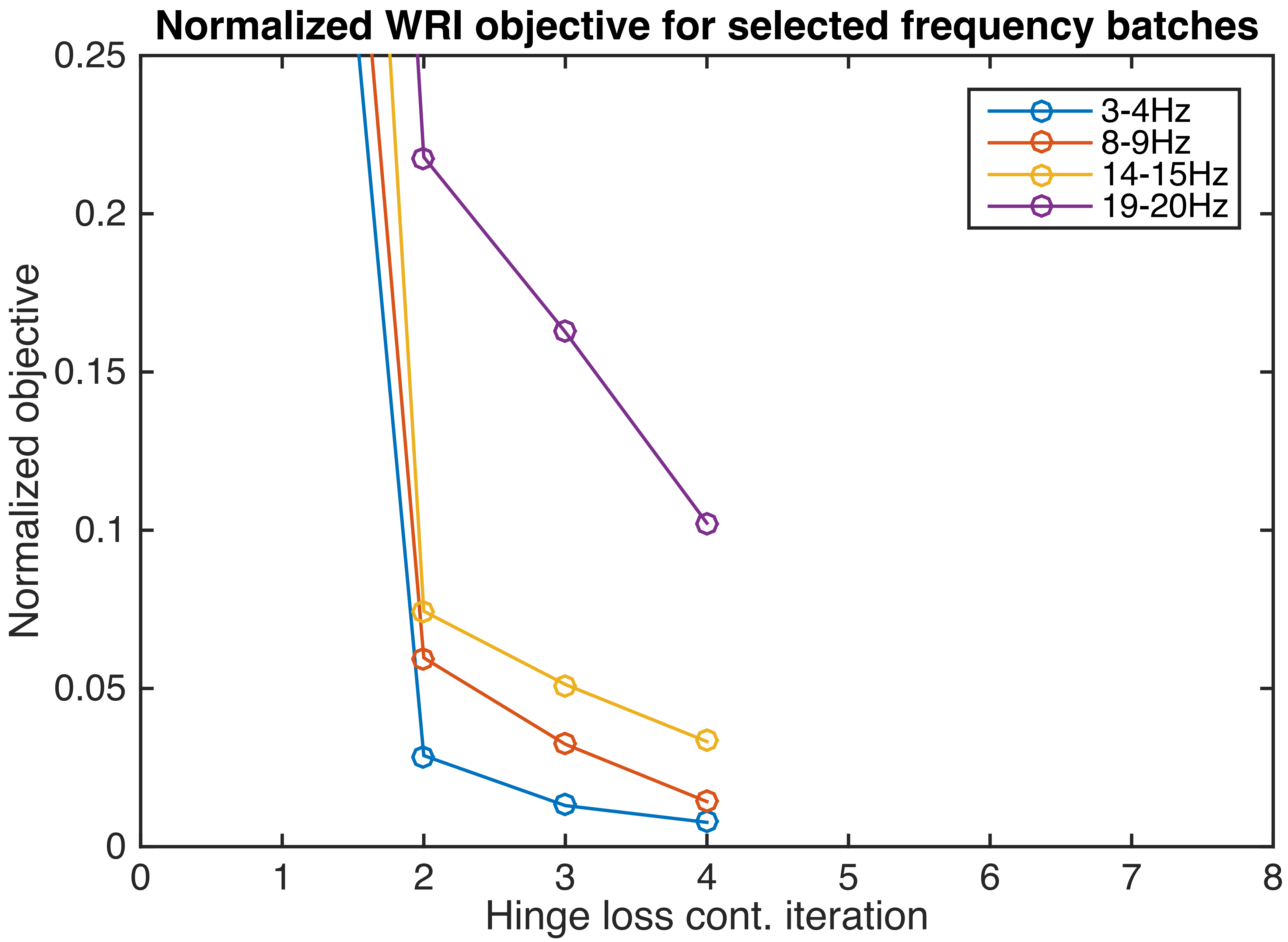
HL constraint: 0.10 of true
TV constraint: 0.90 of true

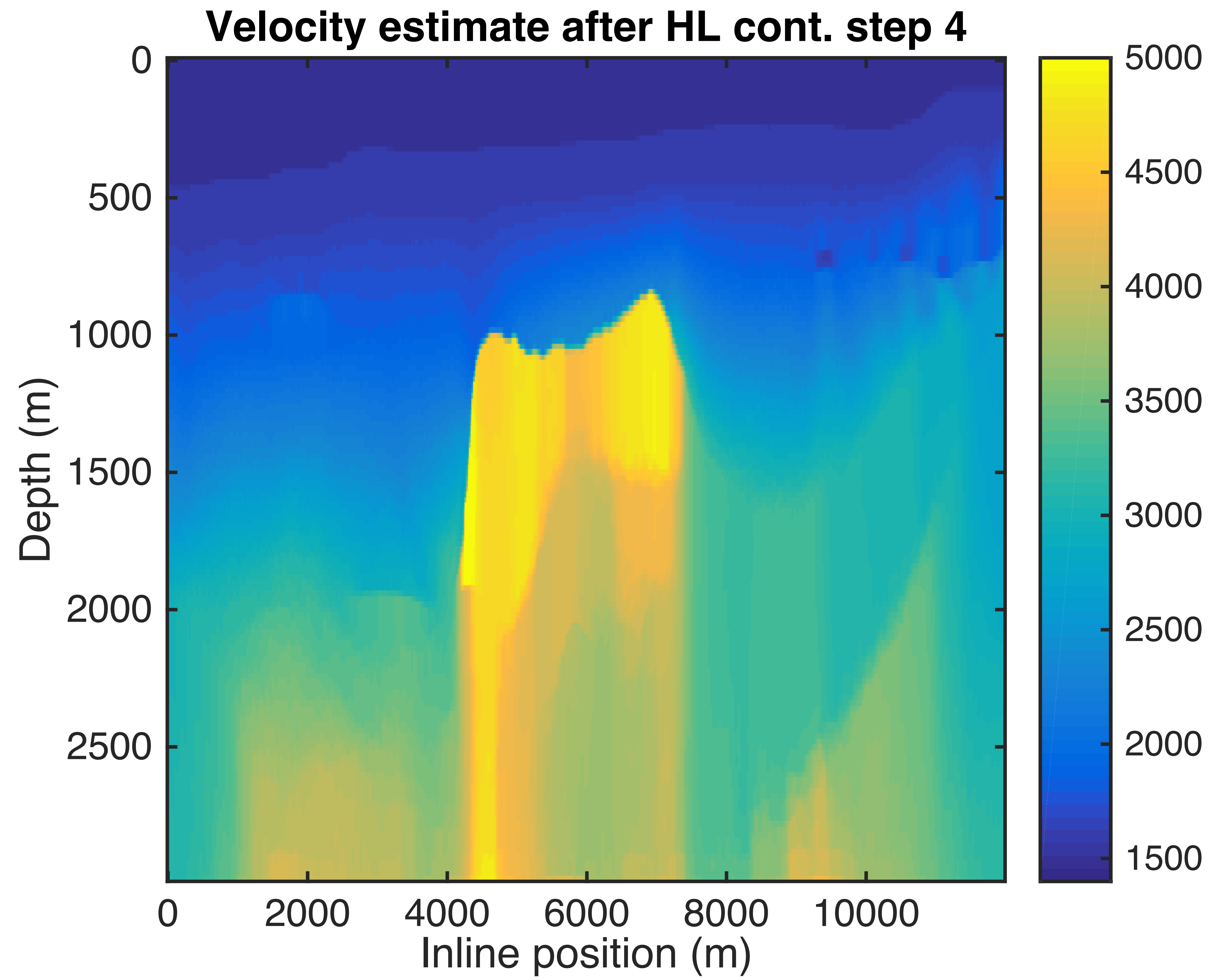


Objective convergence with increasing HL constraint

HL constraint: 0.15 of true

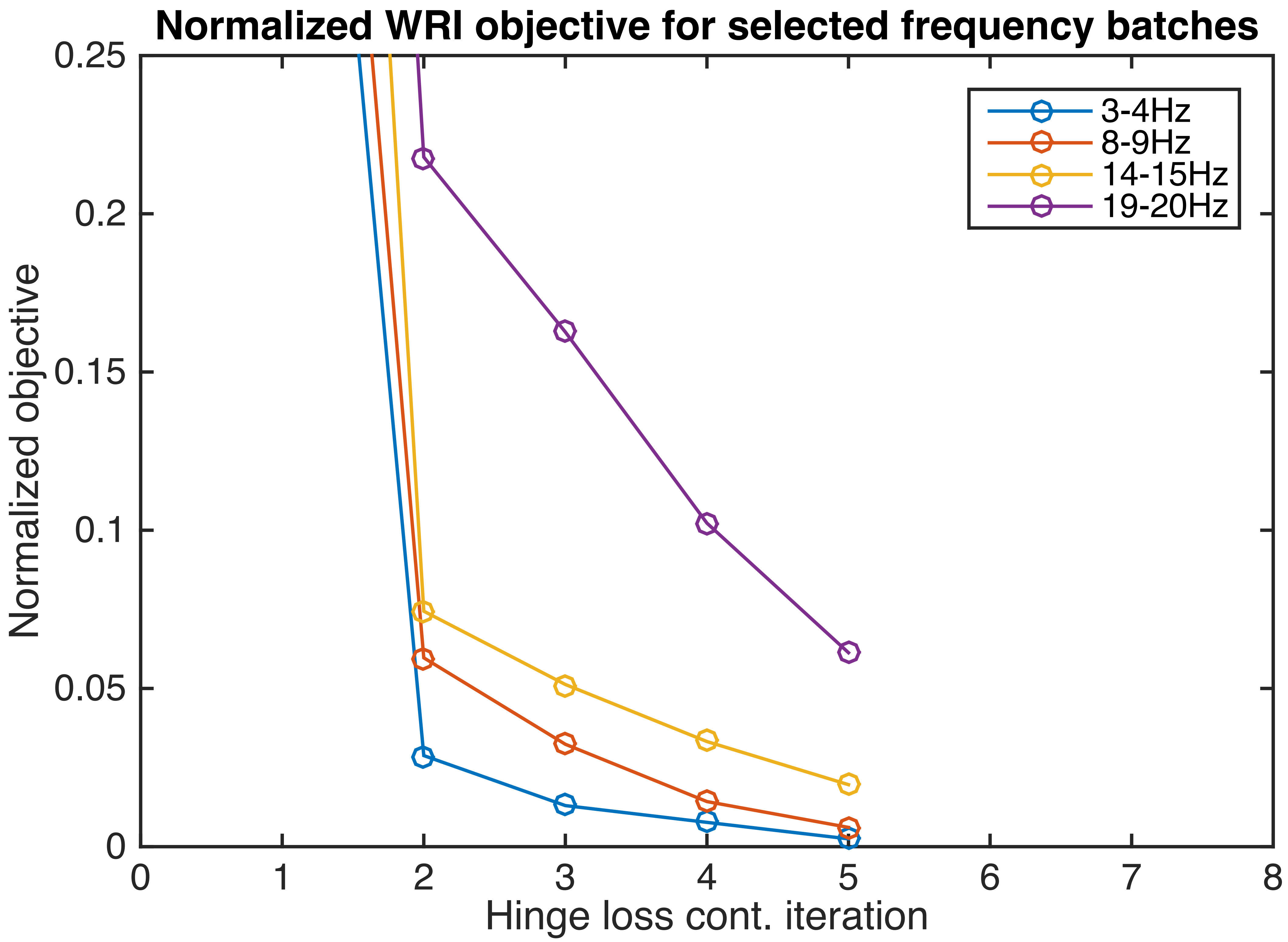
TV constraint: 0.90 of true

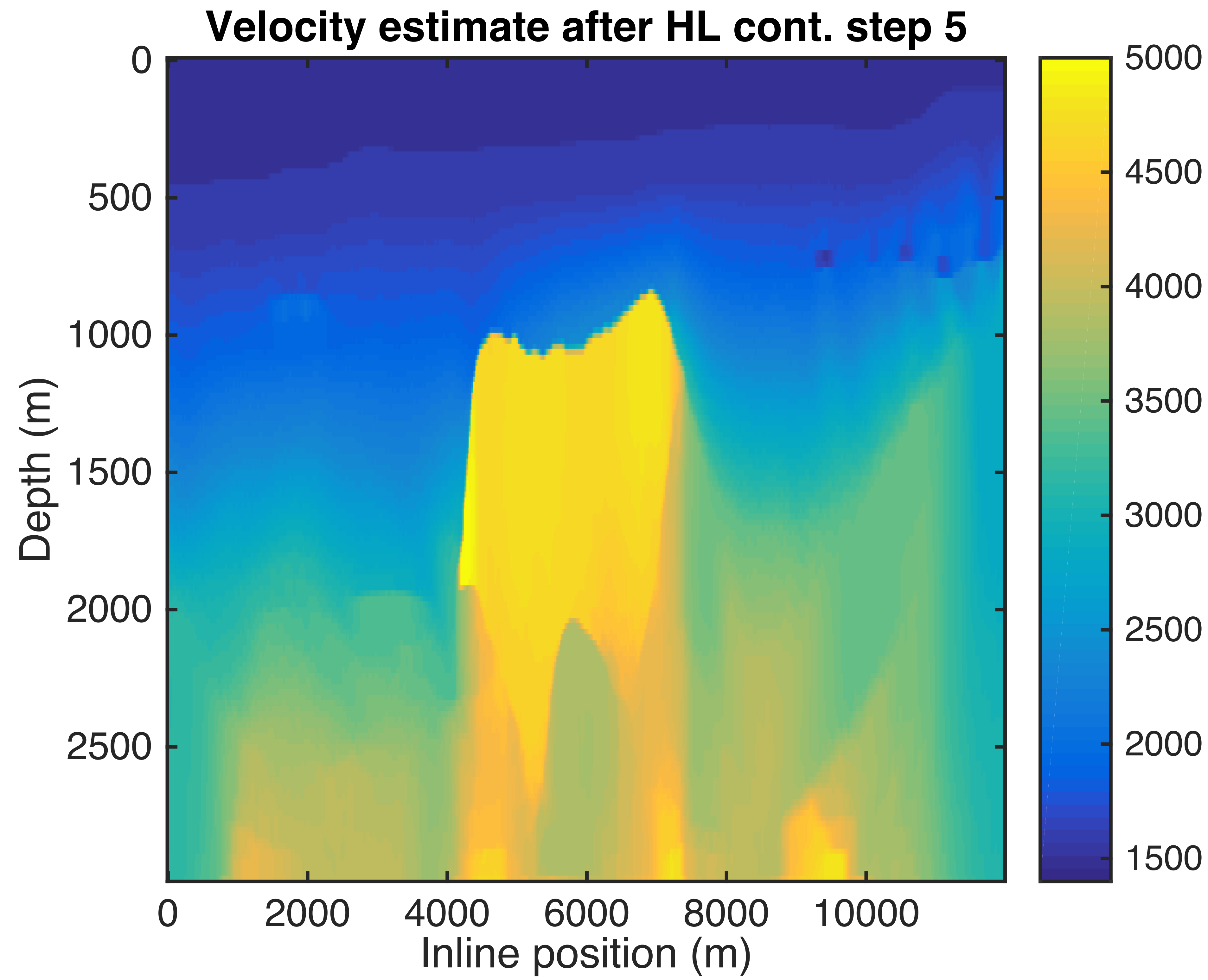




Objective convergence with increasing HL constraint

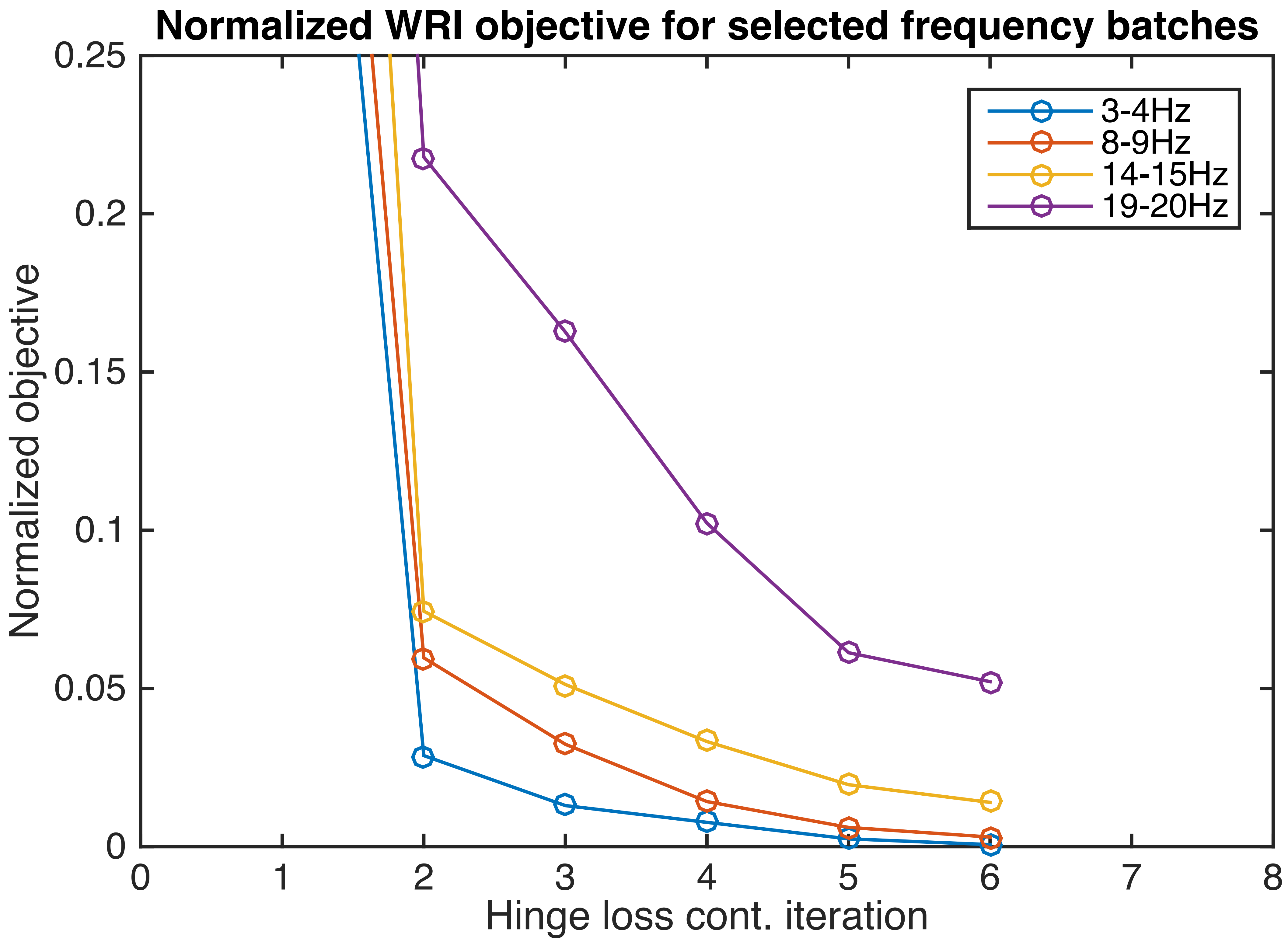
HL constraint: 0.20 of true
TV constraint: 0.90 of true

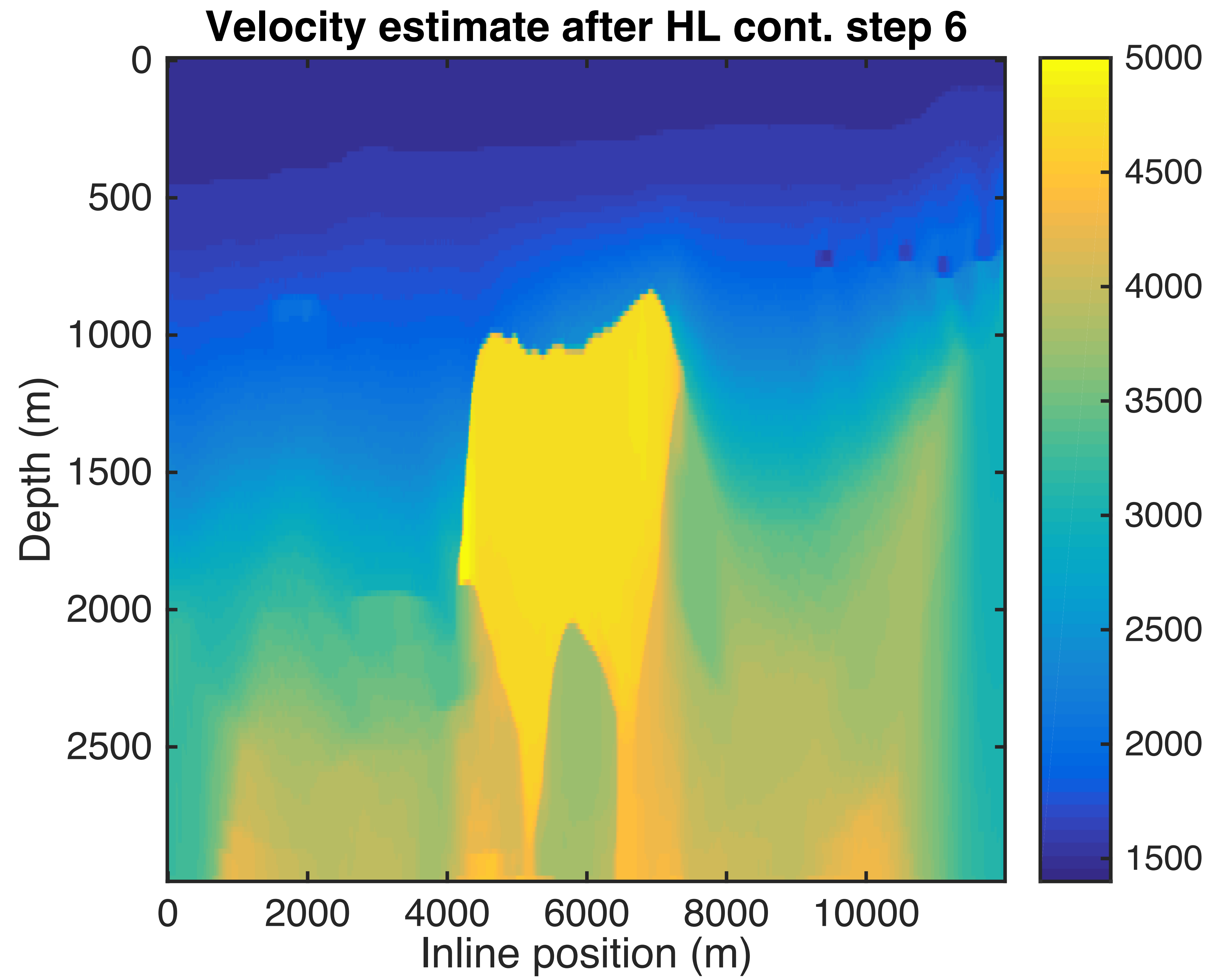




Objective convergence with increasing HL constraint

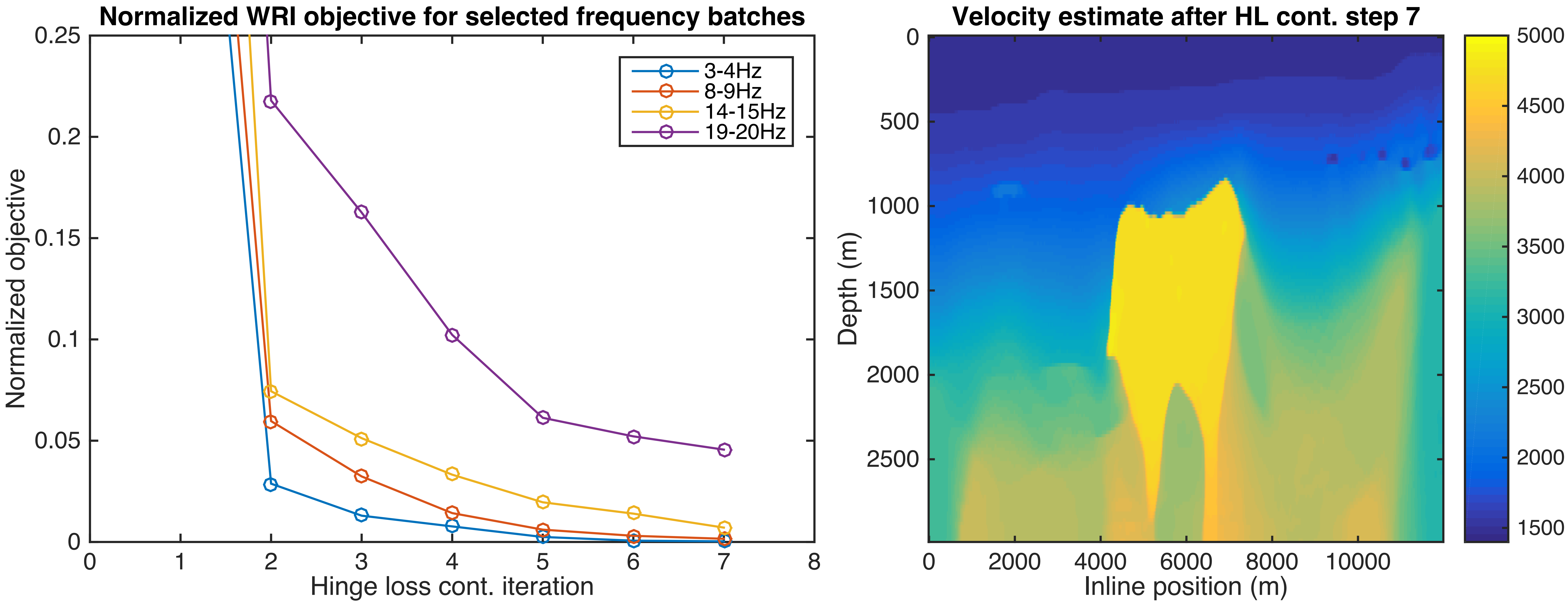
HL constraint: 0.25 of true
TV constraint: 0.90 of true





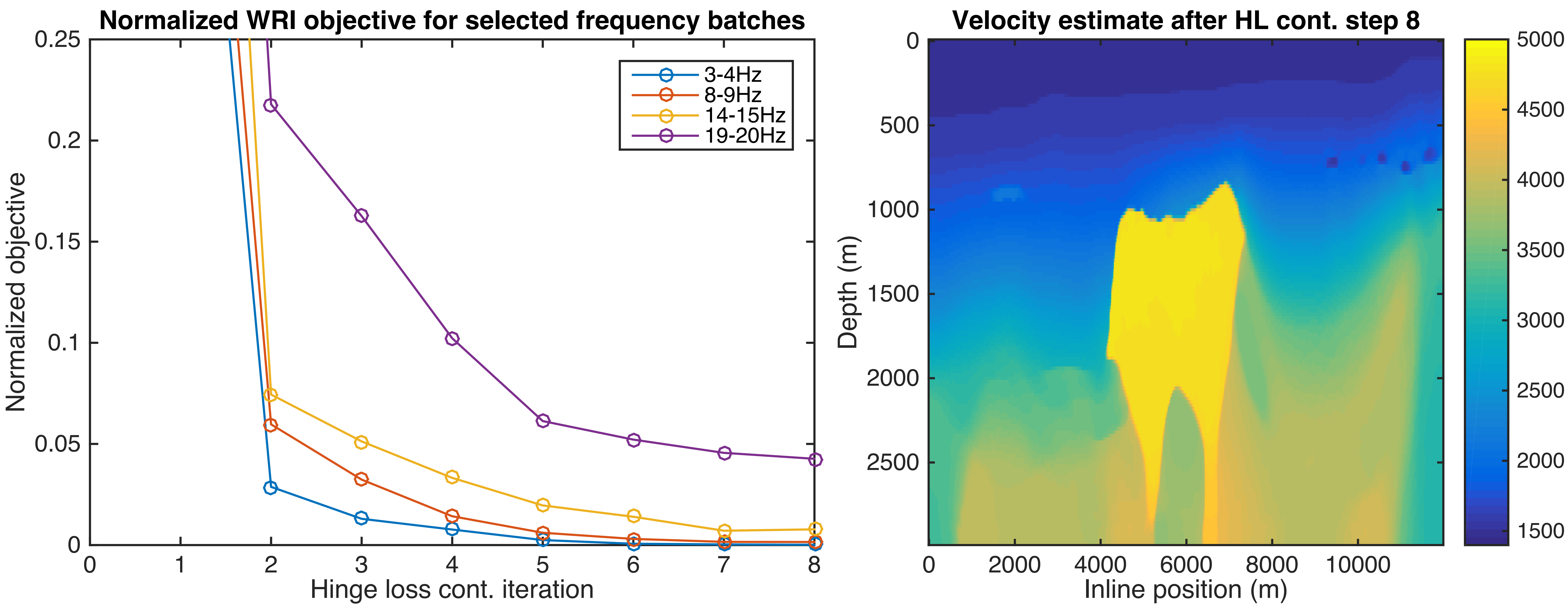
Objective convergence with increasing HL constraint

HL constraint: 0.40 of true
TV constraint: 0.90 of true

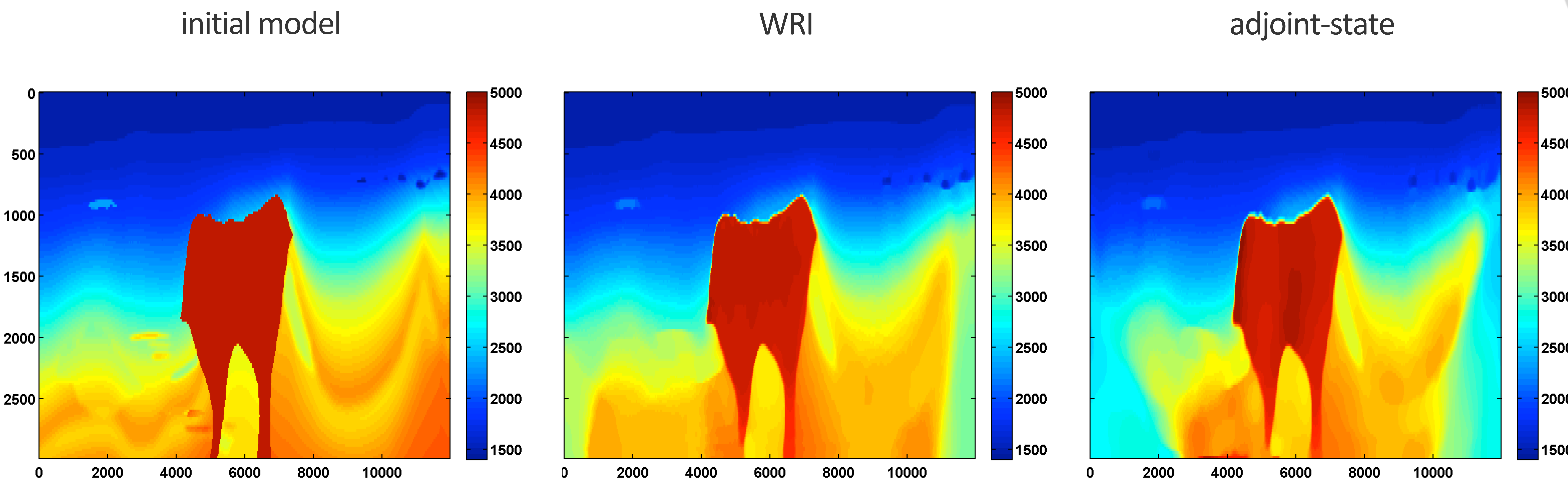


Objective convergence with increasing HL constraint

HL constraint: 0.90 of true
TV constraint: 0.90 of true



WRI vs adjoint-state



Why this works?

Combination of

- multiscale frequency cycles
- relaxation of *asymmetric* constraints

works when

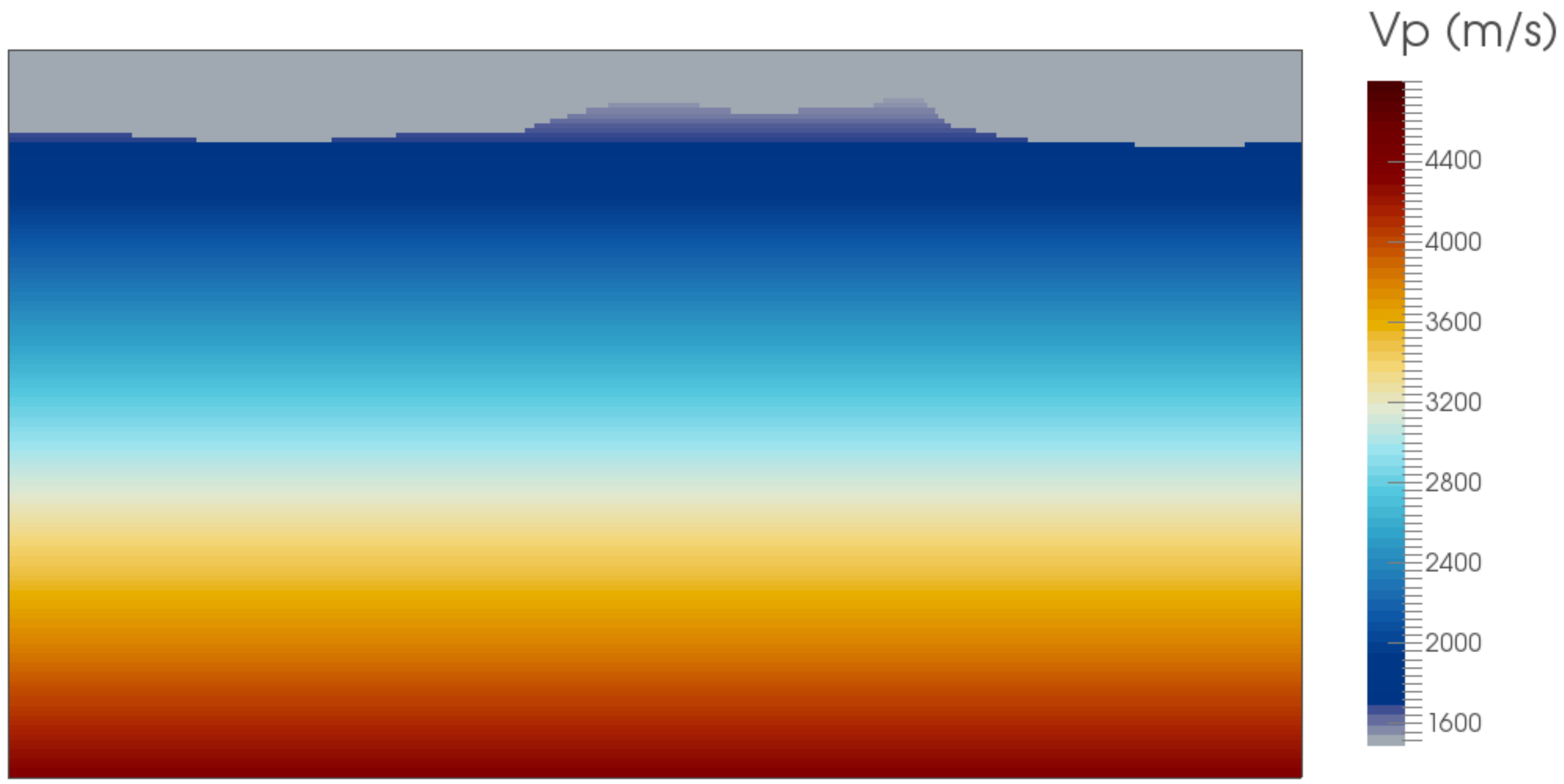
- sharp reflectors are introduced that are of the correct sign
- progress is made during previous cycles
- adverse affects of local minima are controlled by constraints
- “fine-scales” contribute to “coarse-scales” of the next cycle

“Deep” SEAM model – inversion parameters

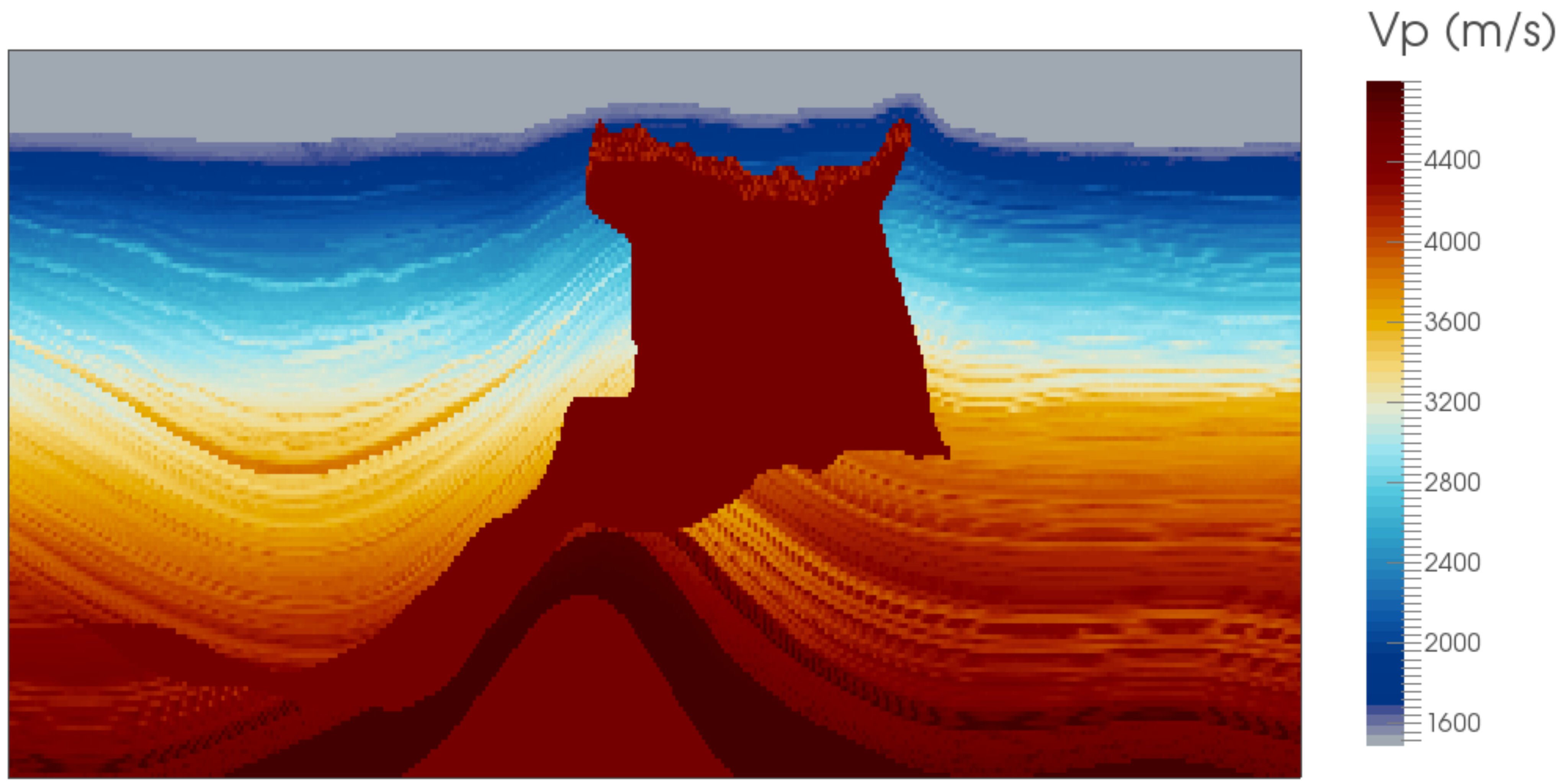
- model size 401x151
- grid size: 20 m
- frequencies: 3 – 12 Hz
- source: 10 Hz Ricker wavelet
- 98 sources
- 196 receivers
- two simultaneous shots with Gaussian weights w/ redraws
- noise-free & inversion crime

"Deep" SEAM model

starting model

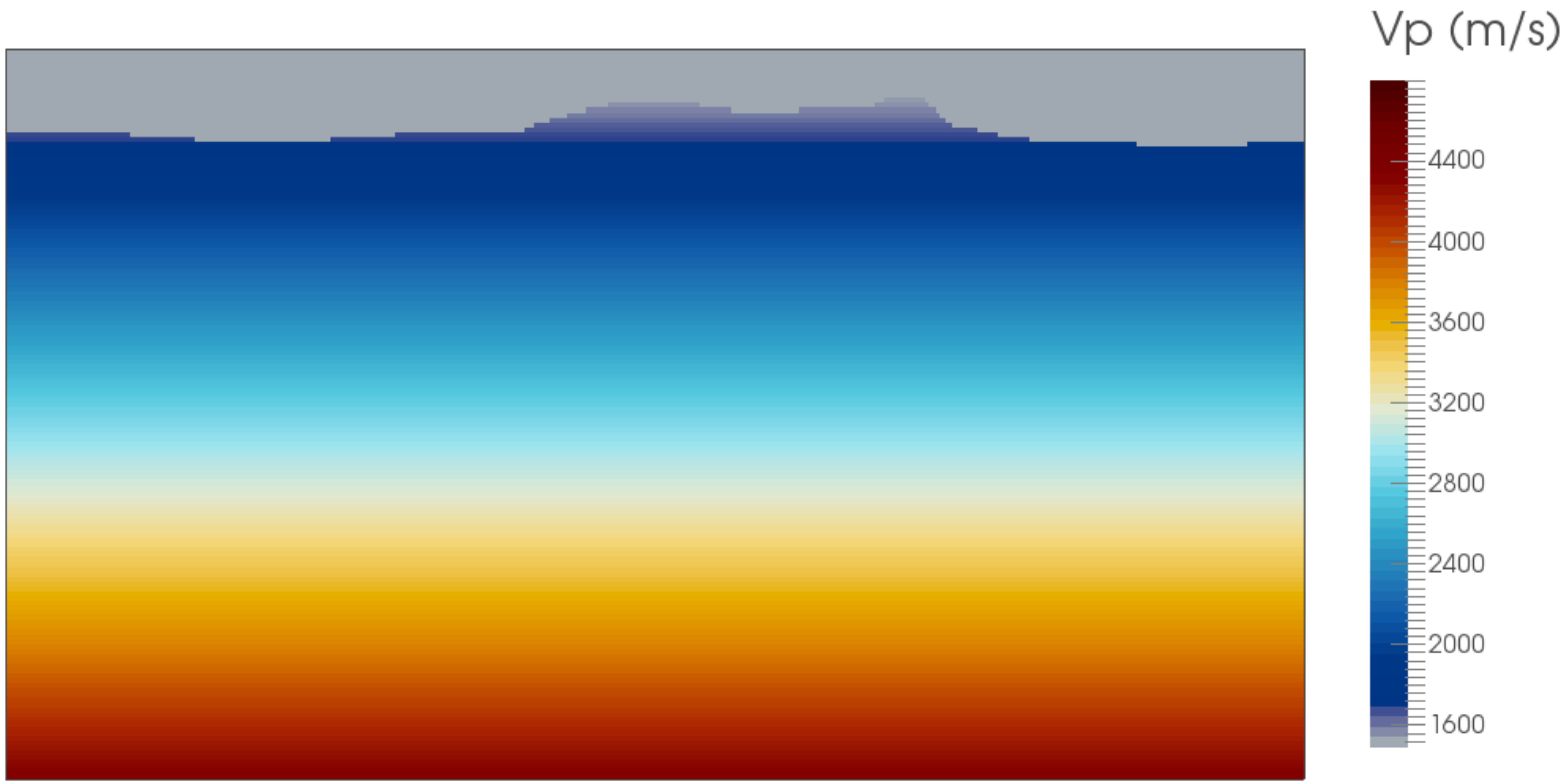


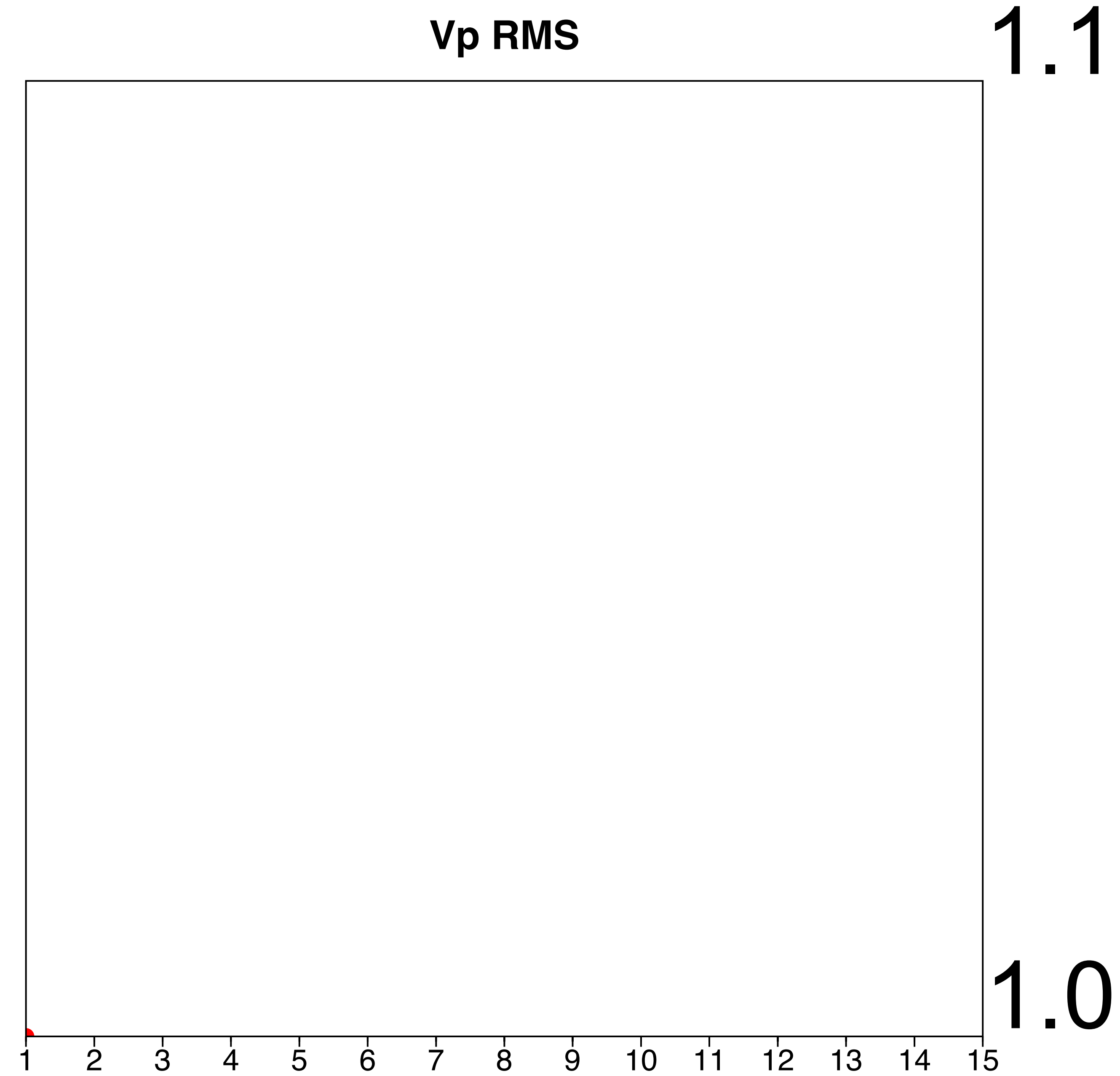
true model



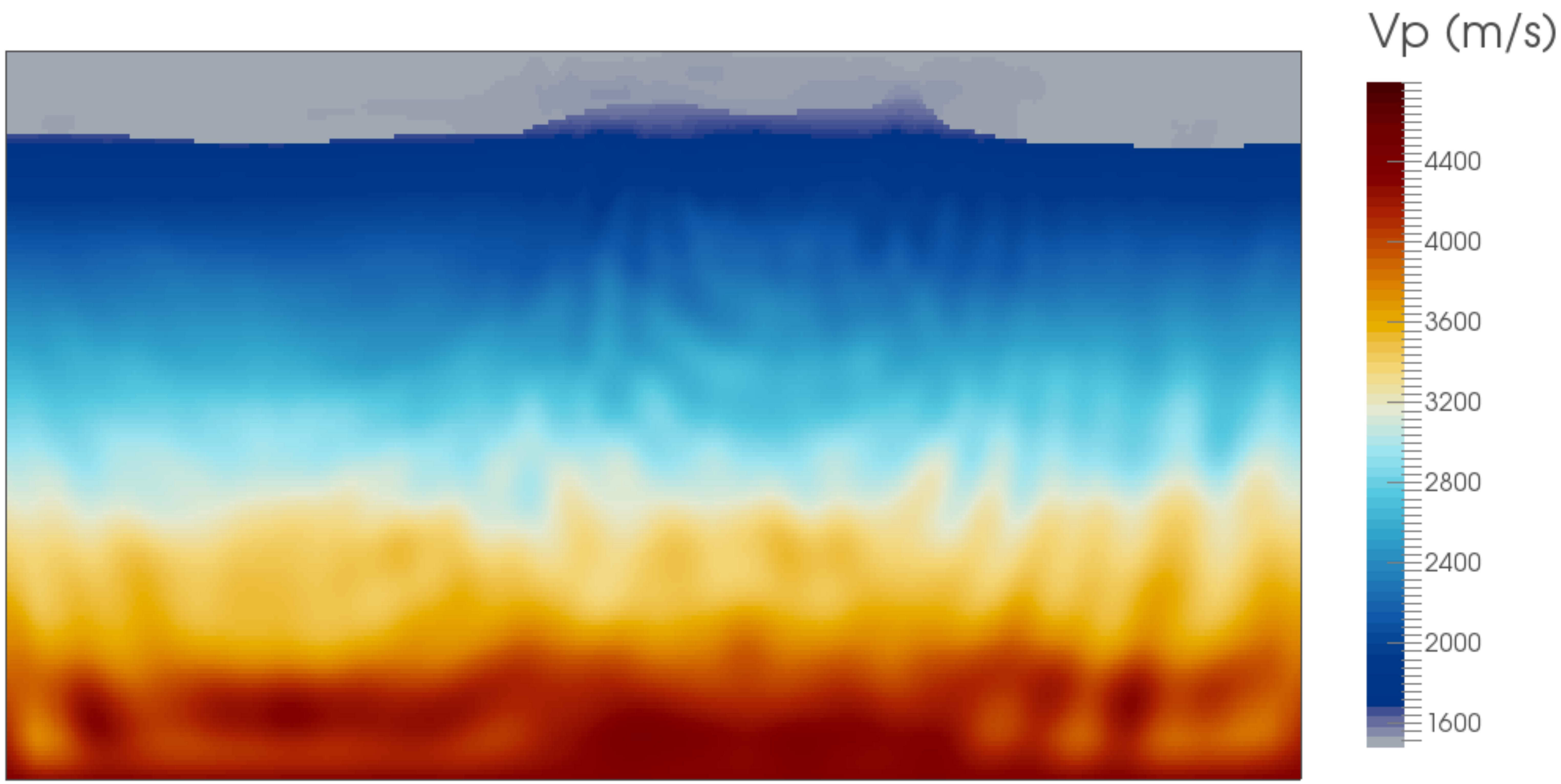
Models from SEG Advanced Modeling Corporation (SEAM)

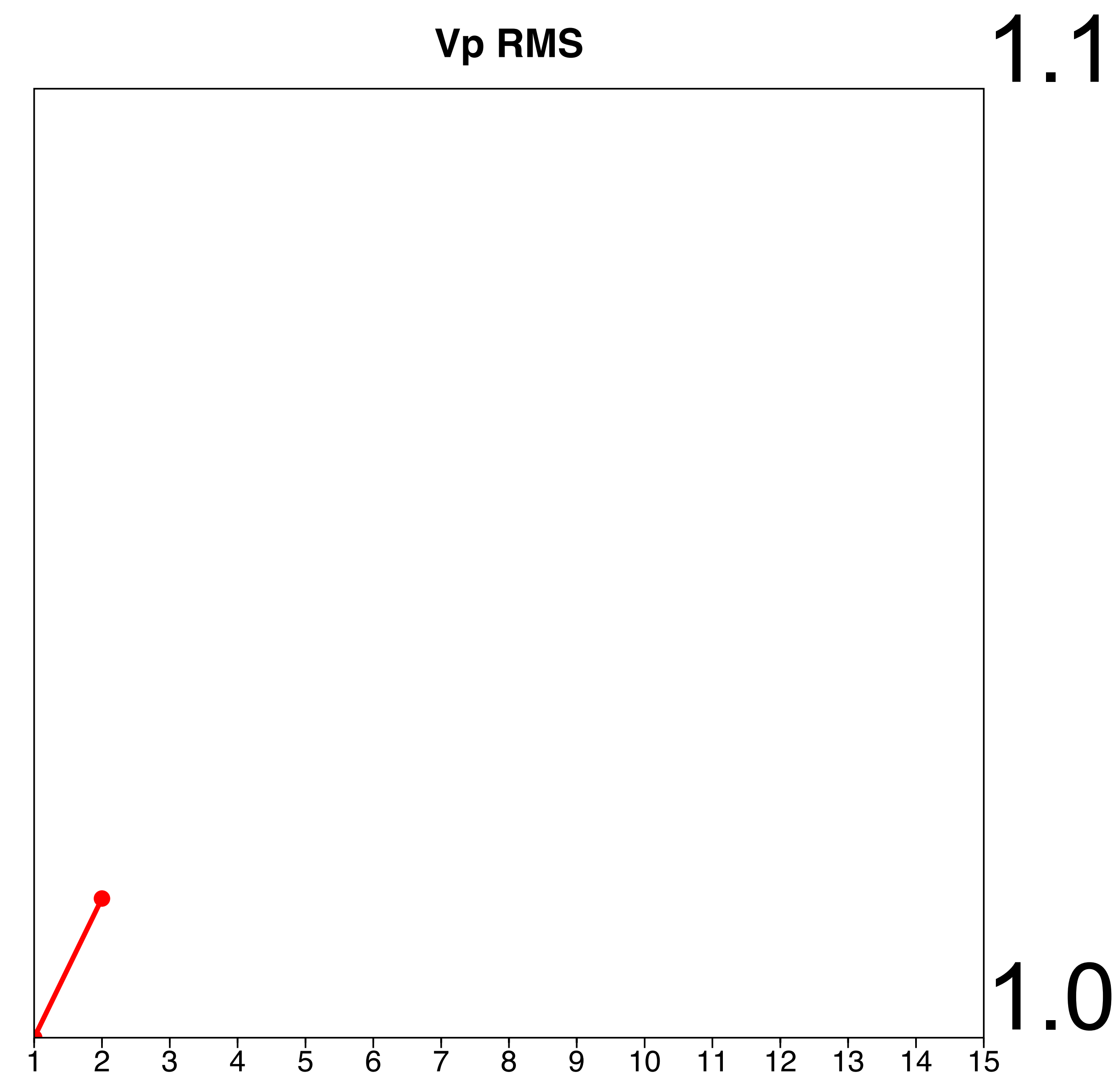
Adjoint-state – w/o constraints



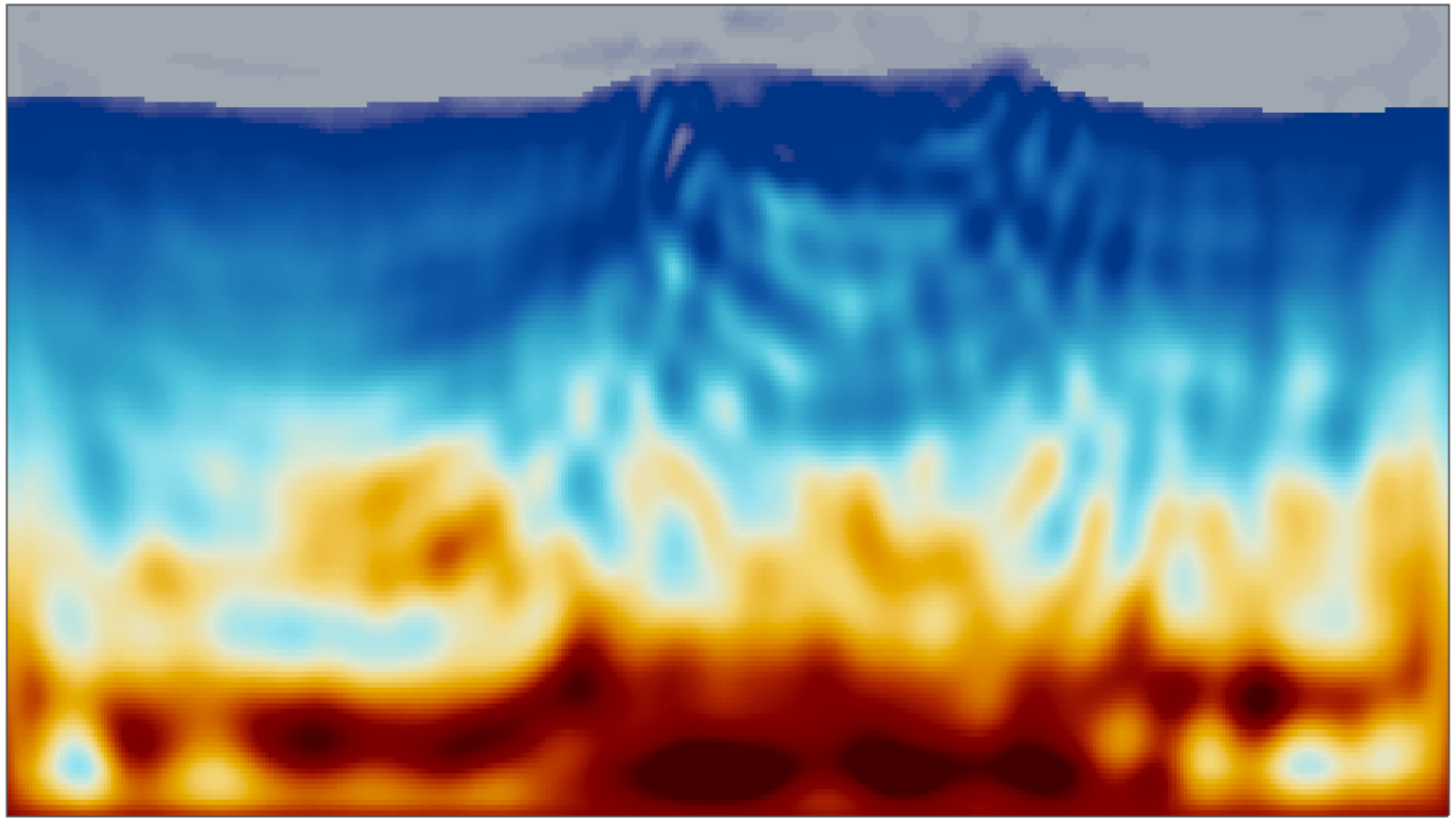


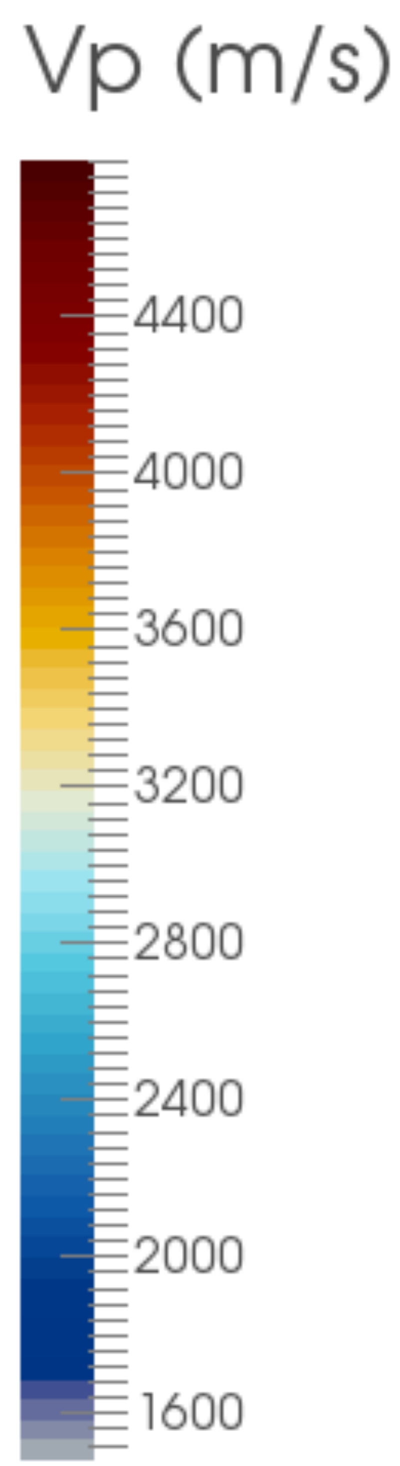
Adjoint-state – w/o constraints

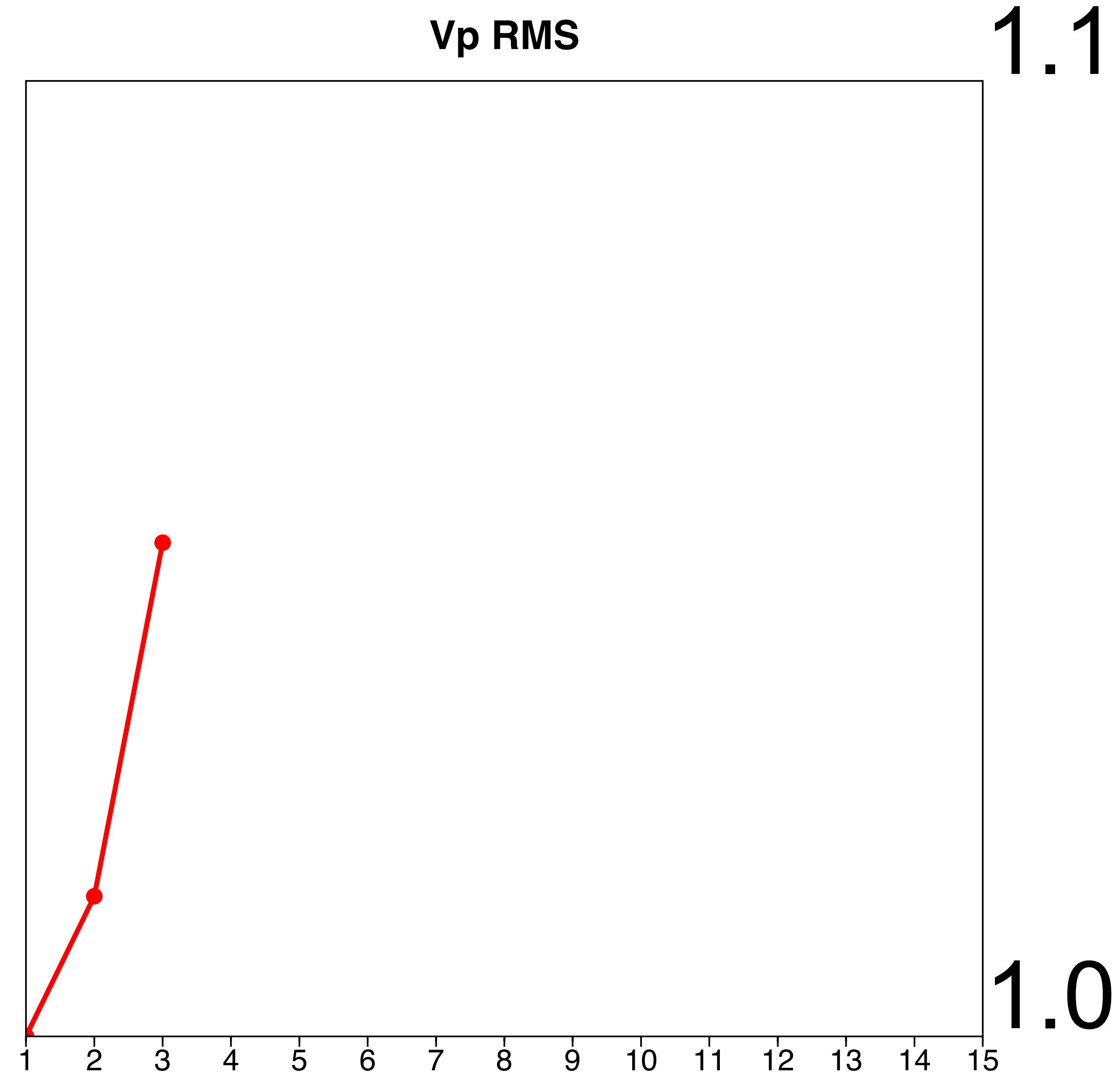




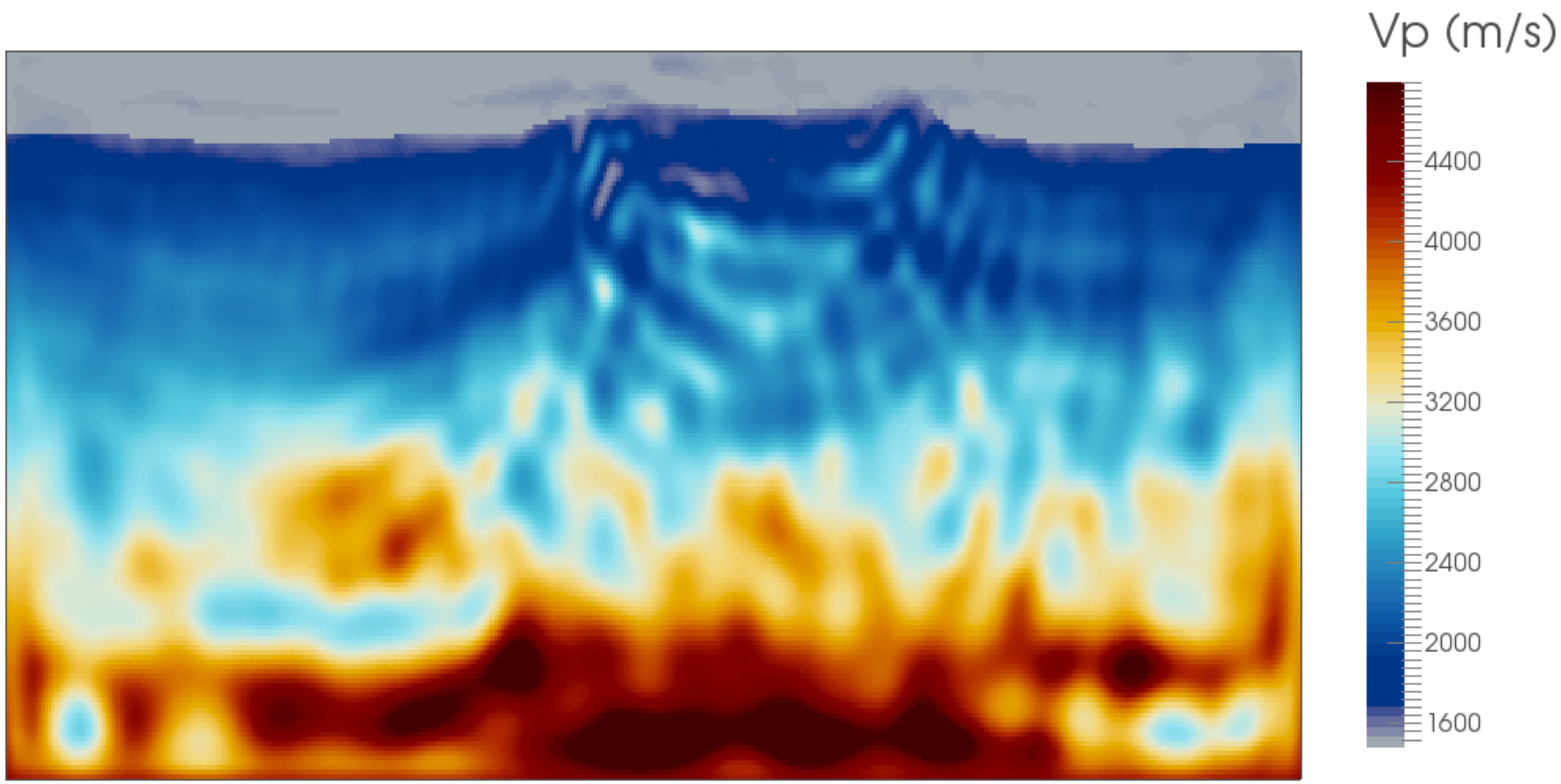
Adjoint-state – w/o constraints

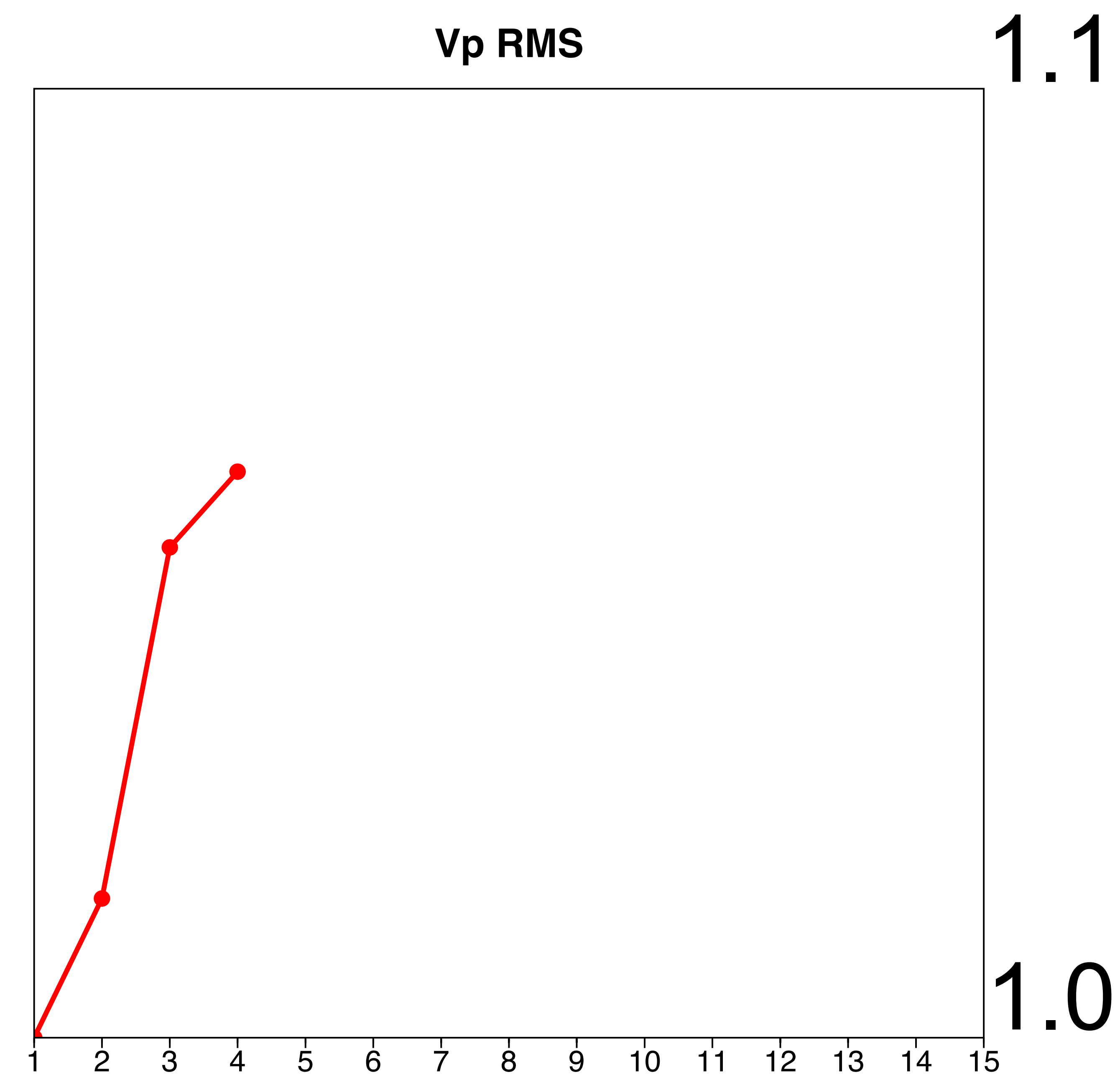




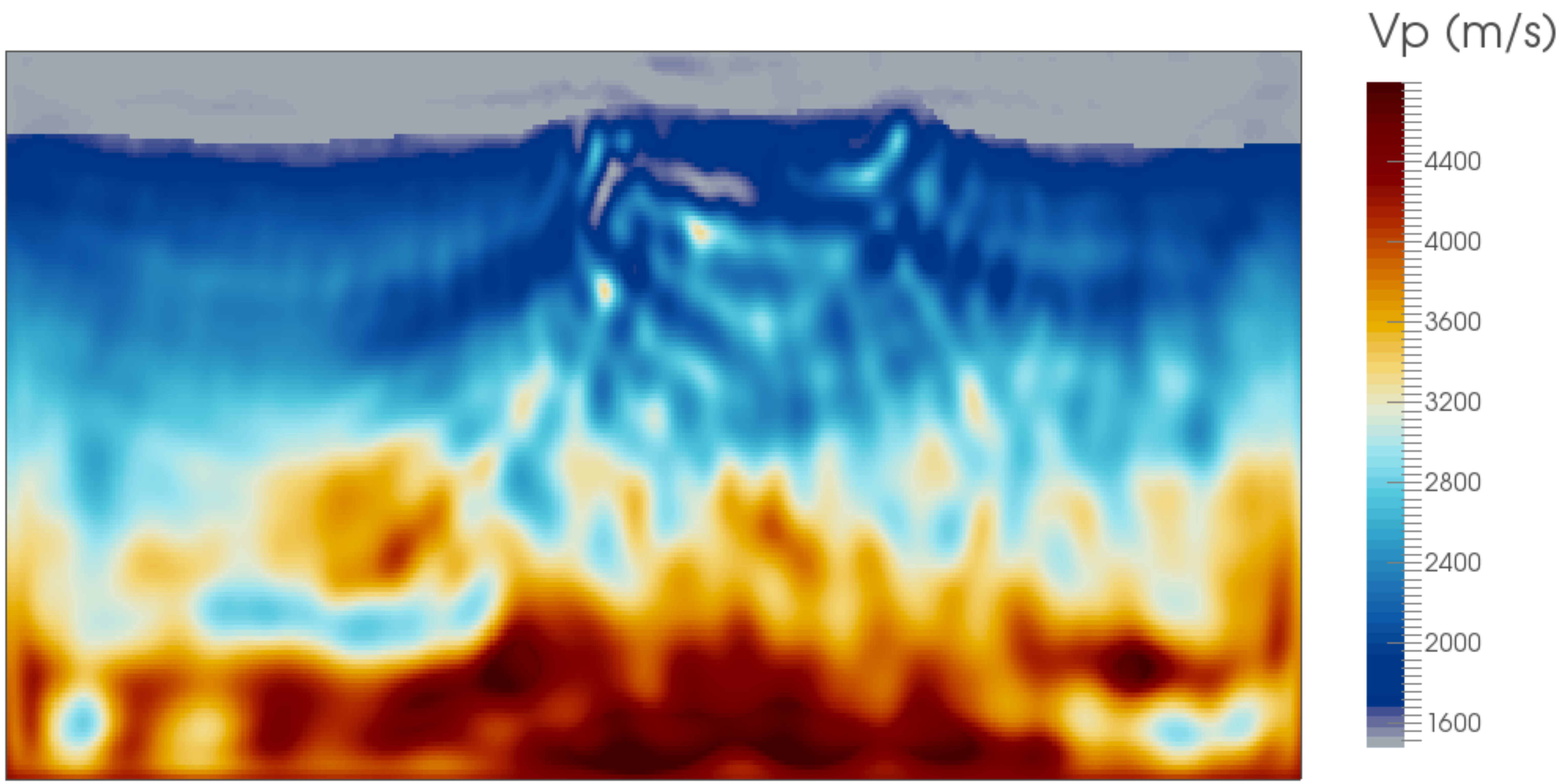


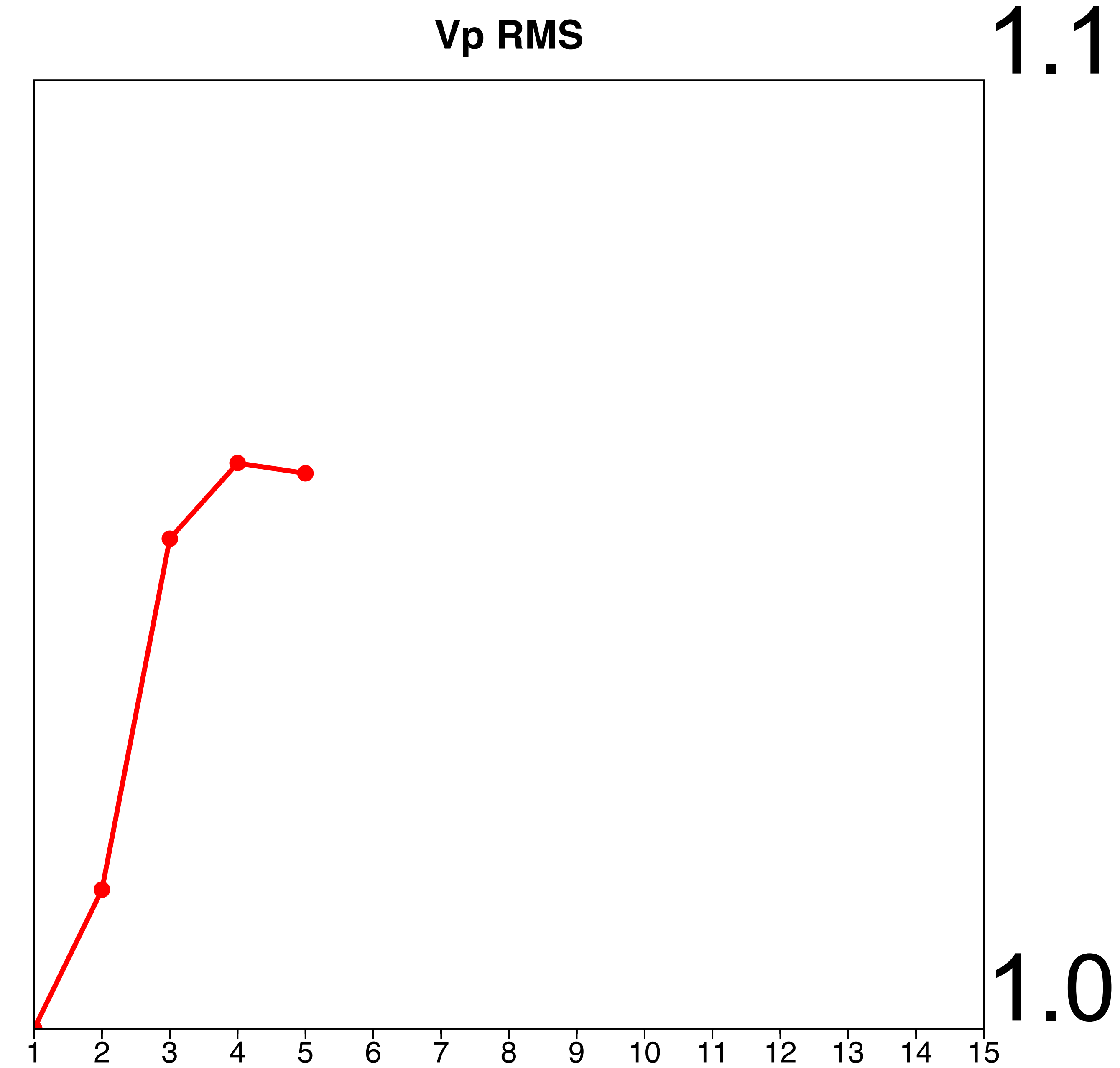
Adjoint-state – w/o constraints



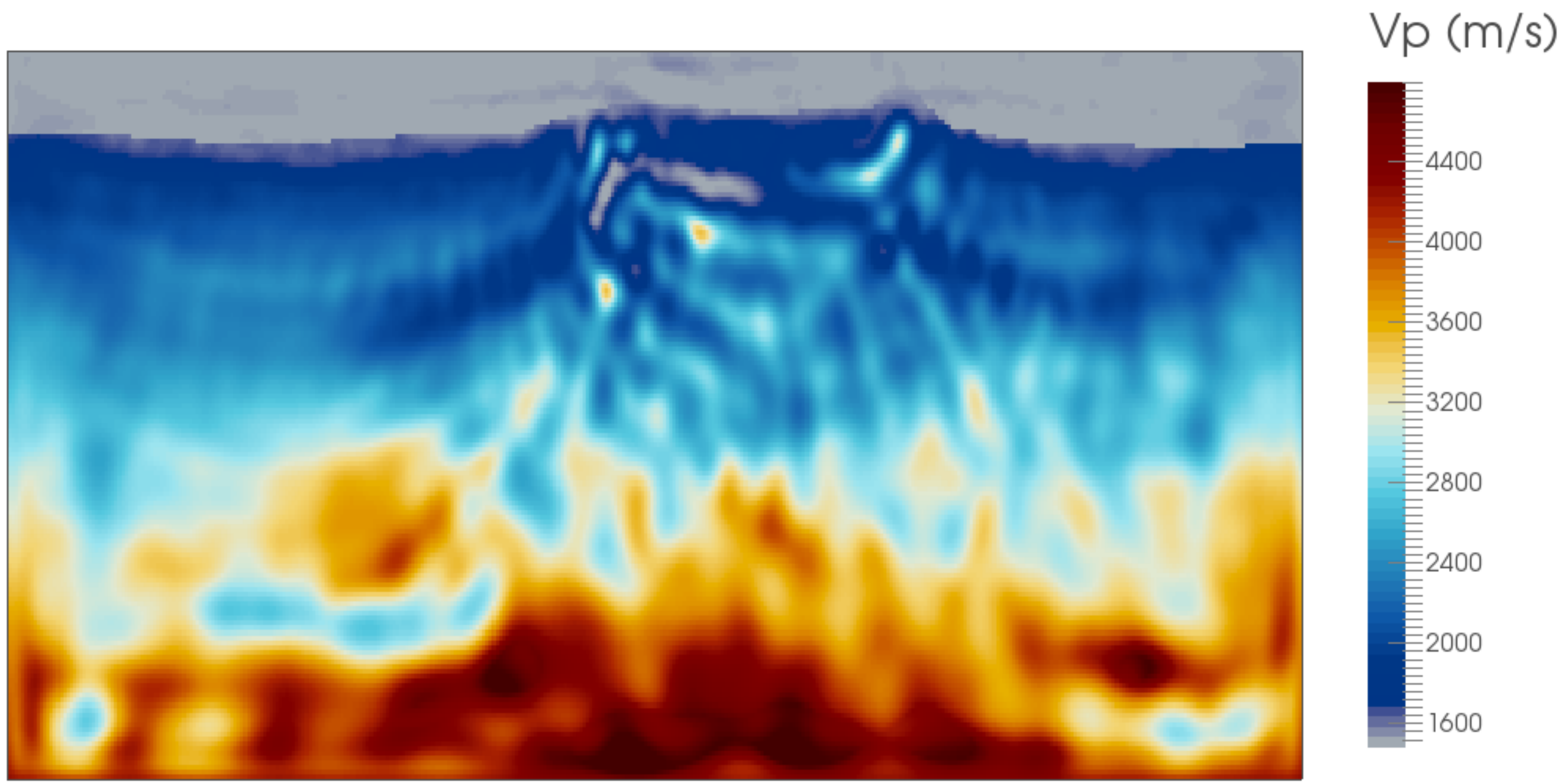


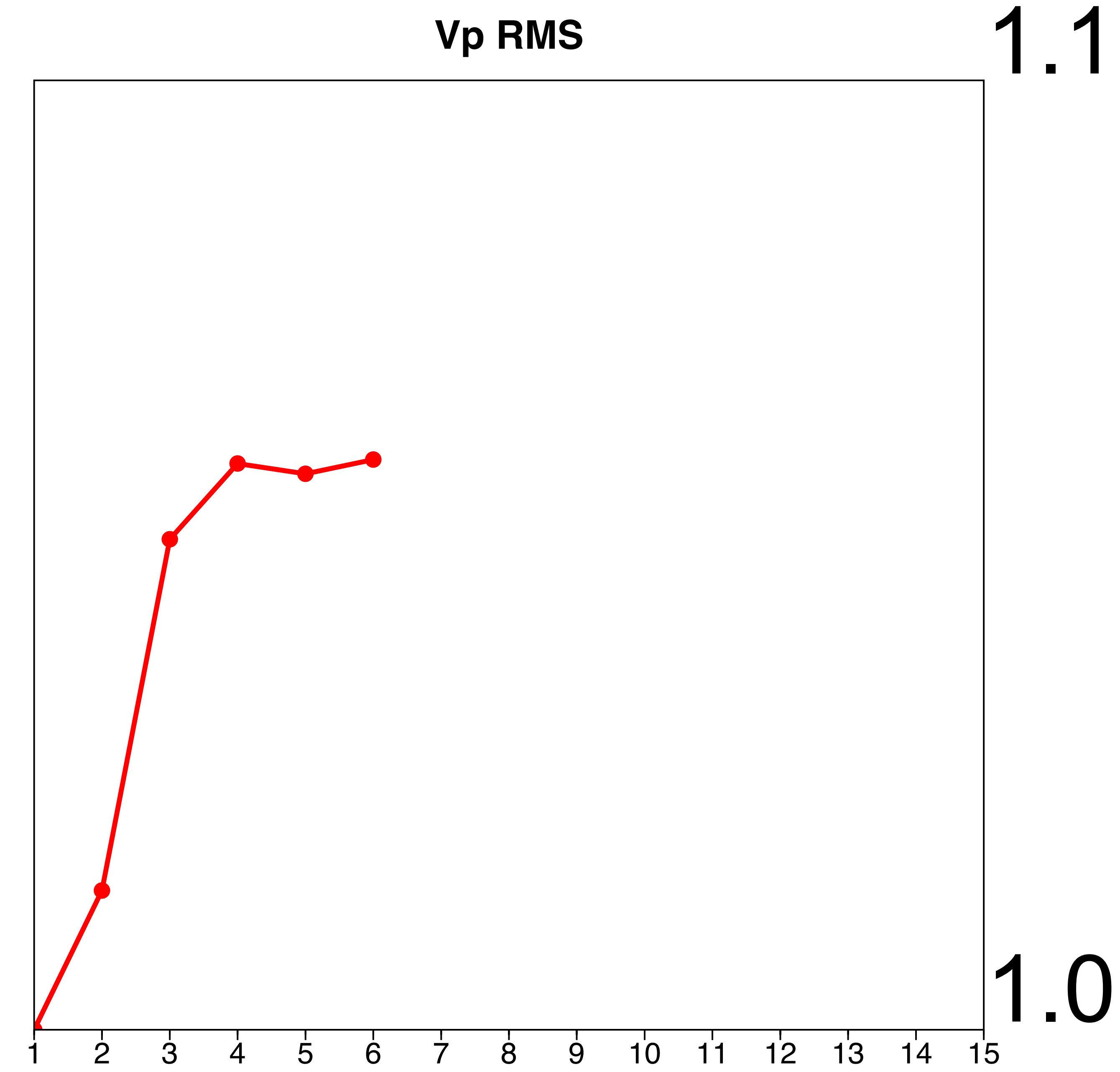
Adjoint-state – w/o constraints



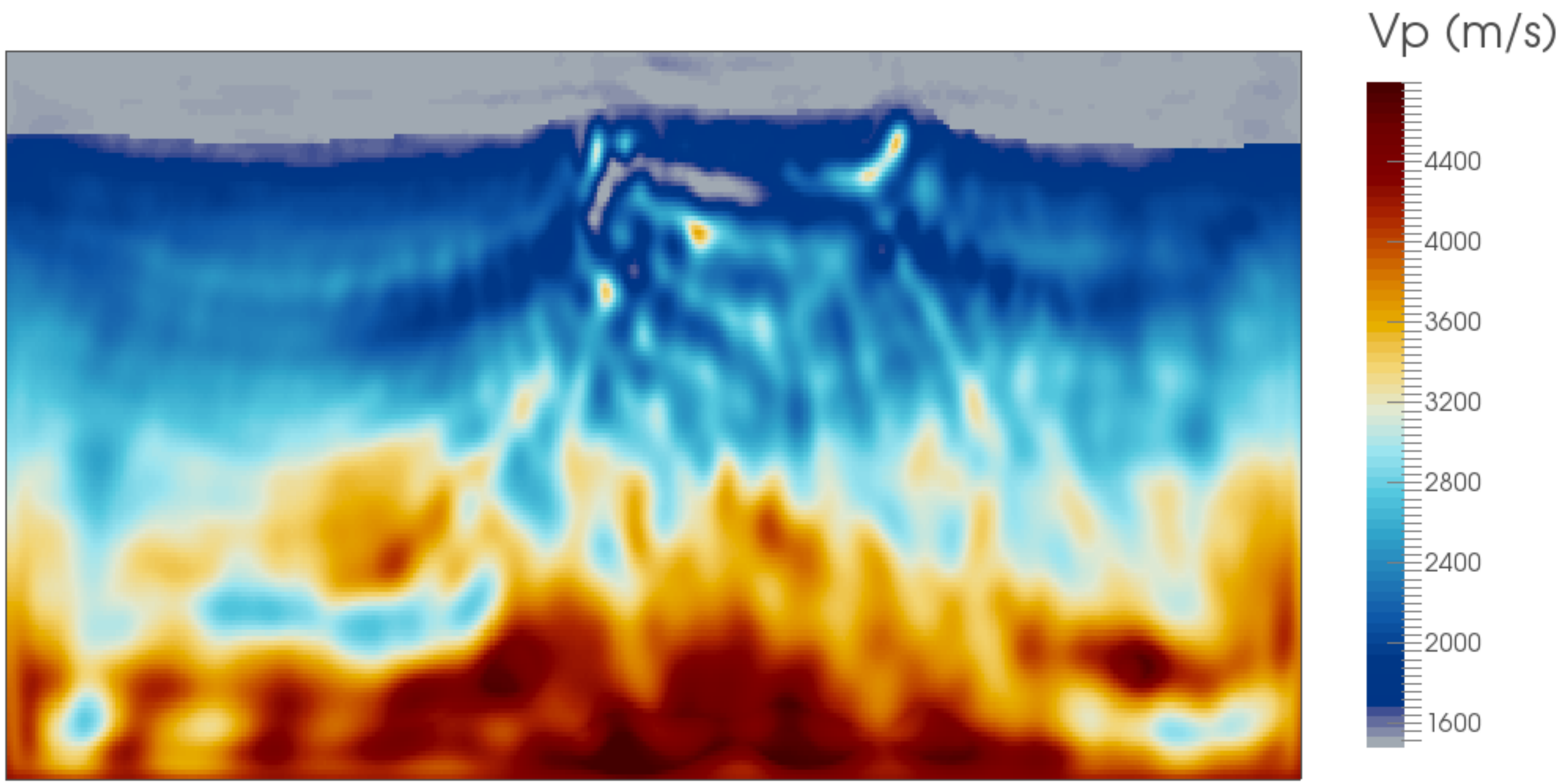


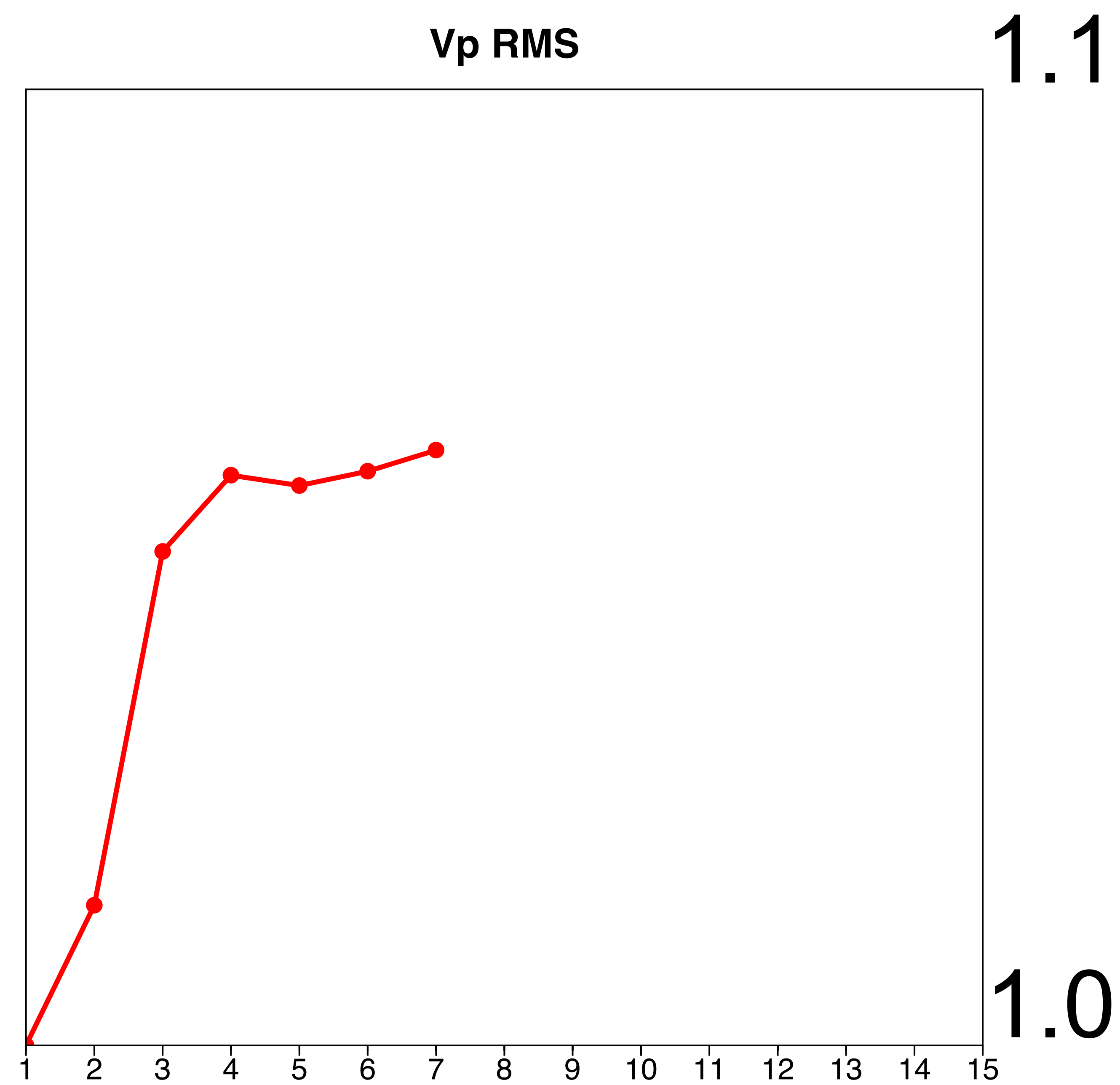
Adjoint-state – w/o constraints



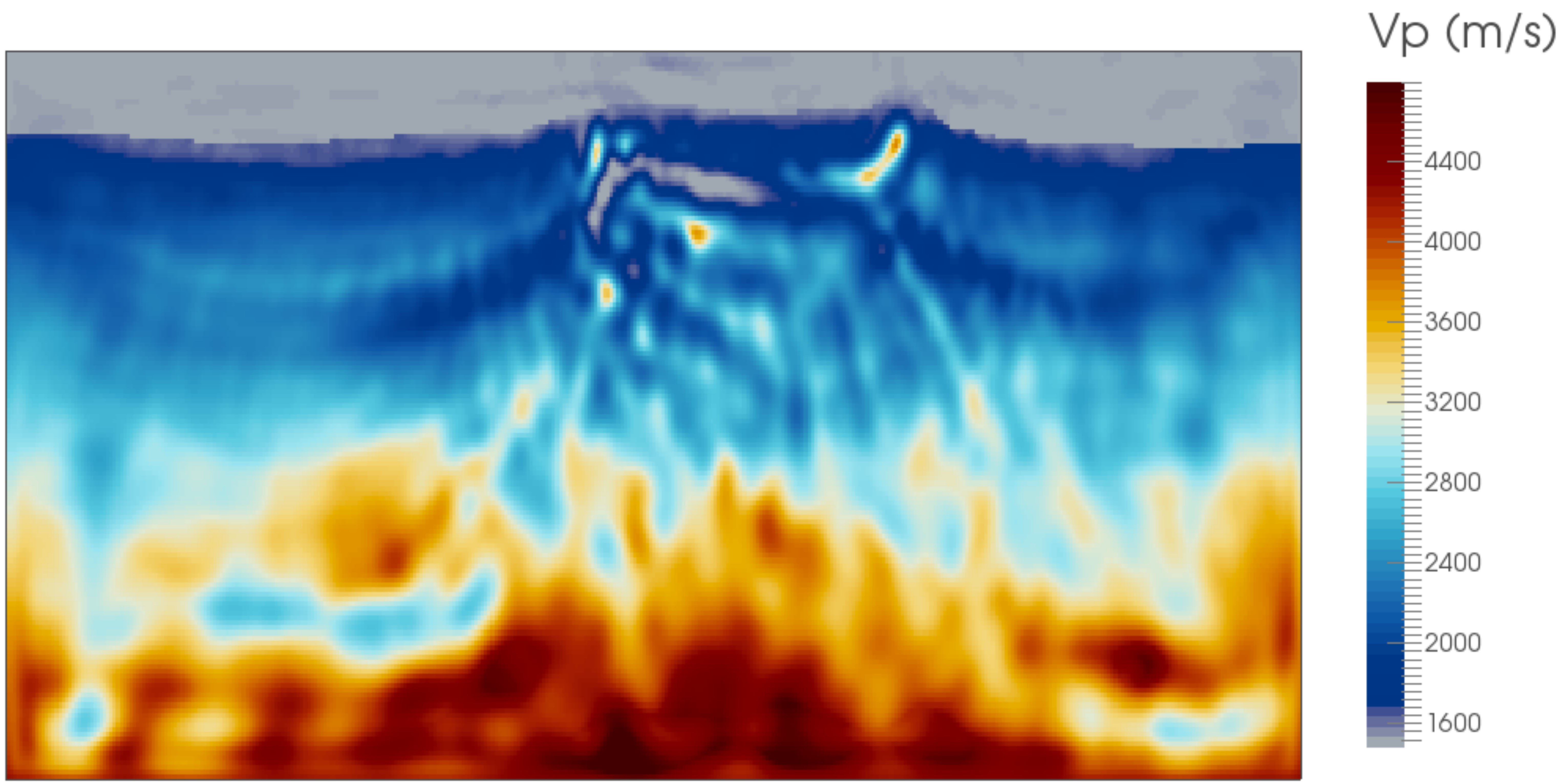


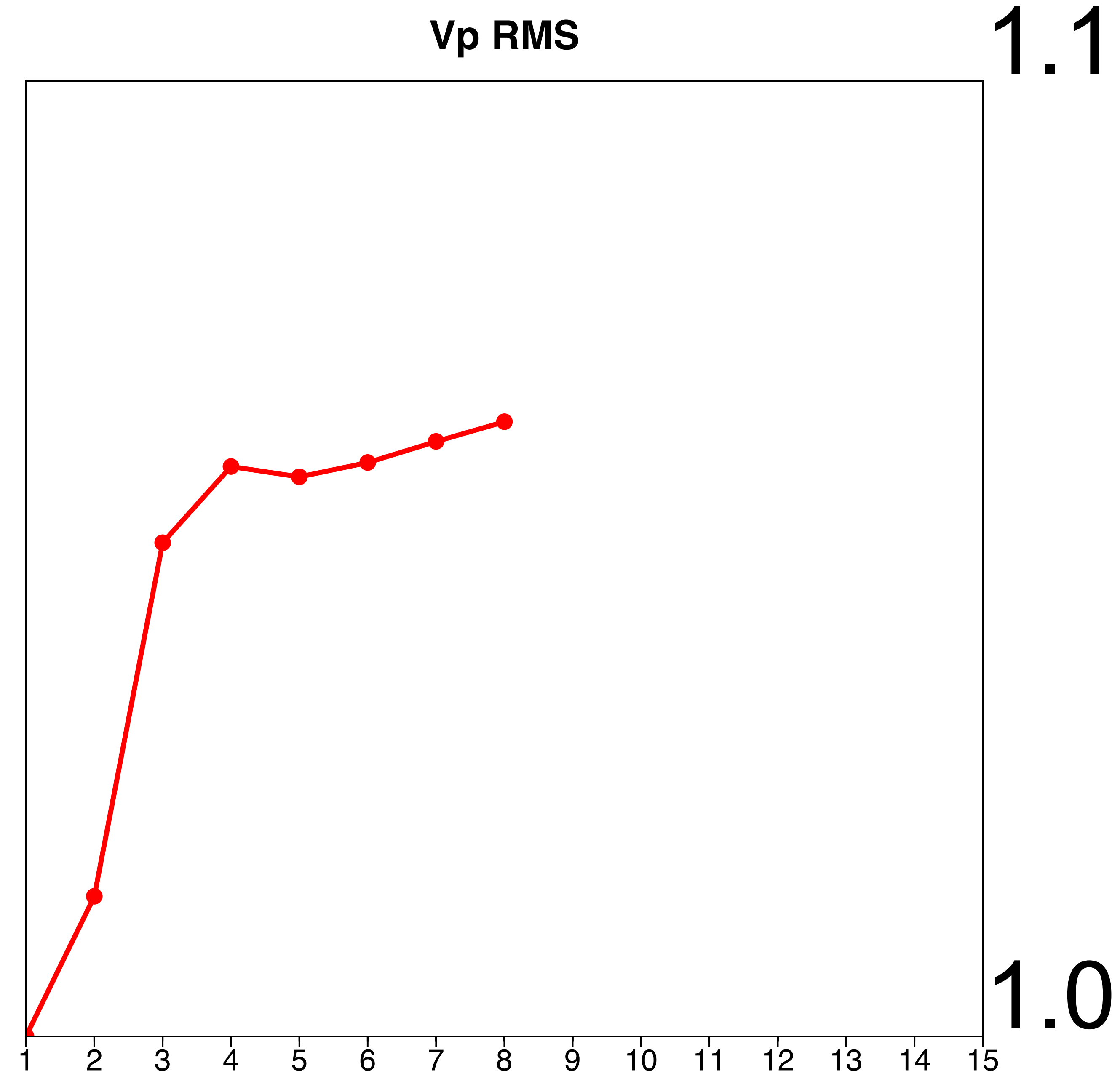
Adjoint-state – w/o constraints



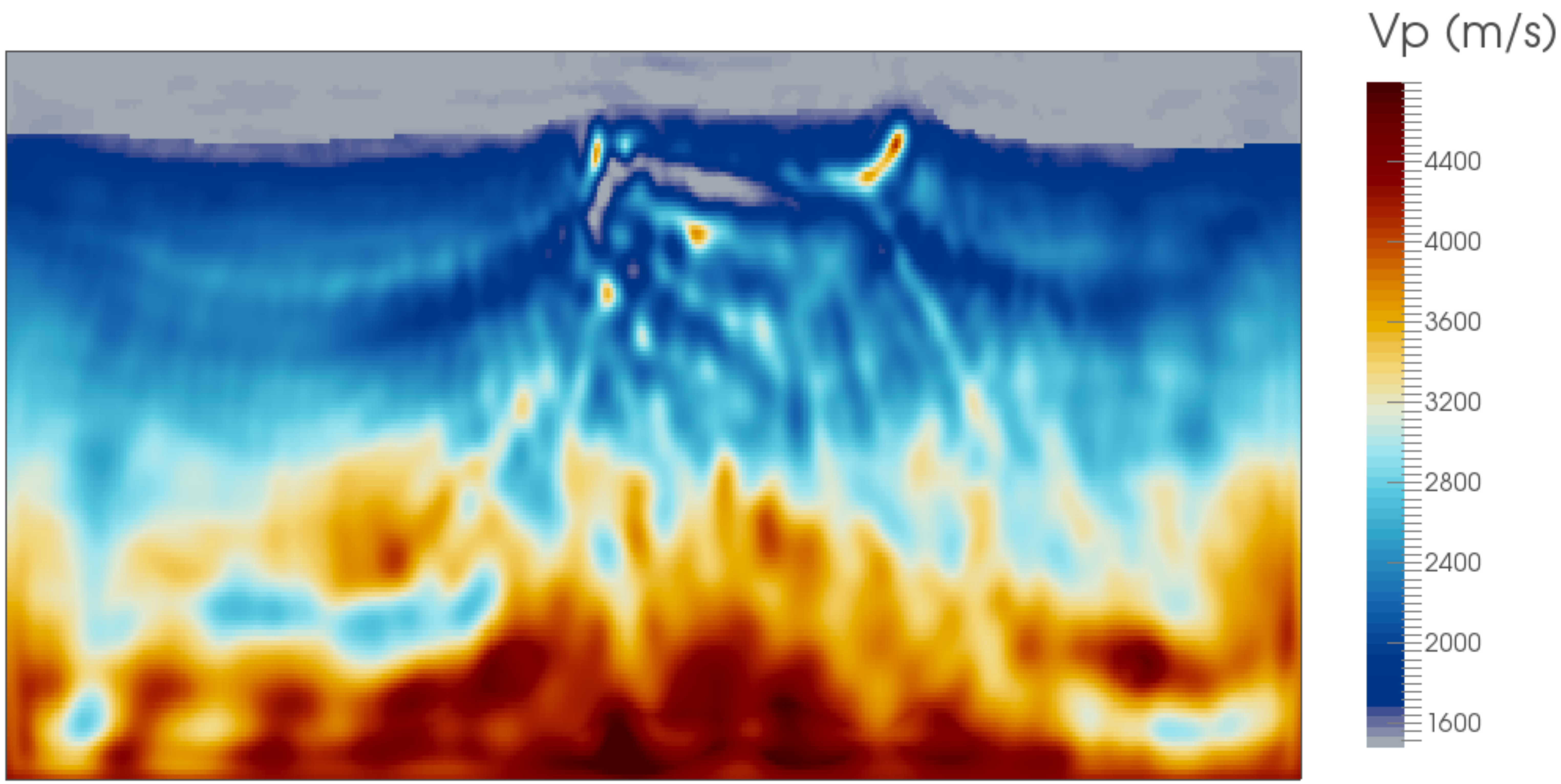


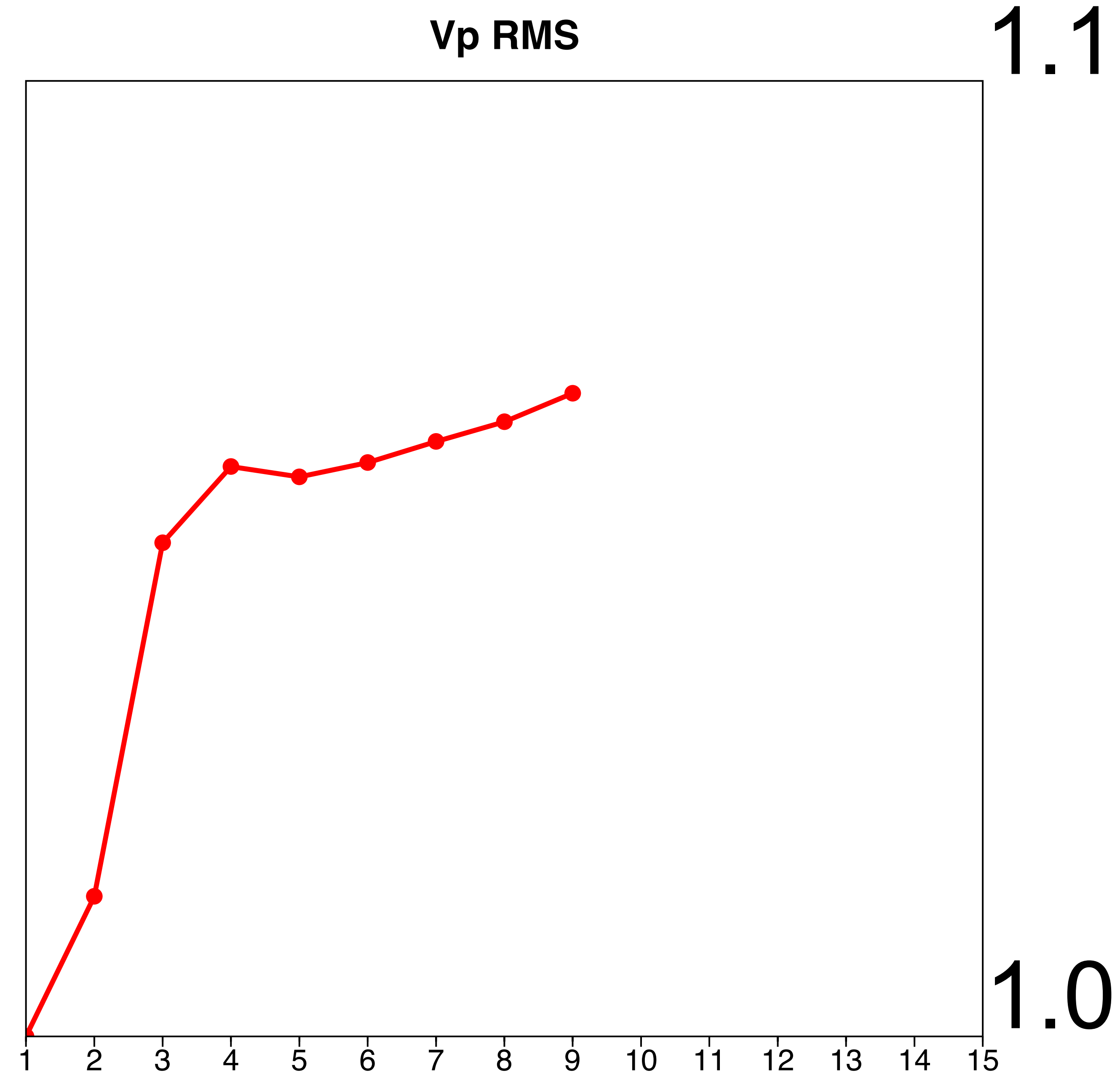
Adjoint-state – w/o constraints



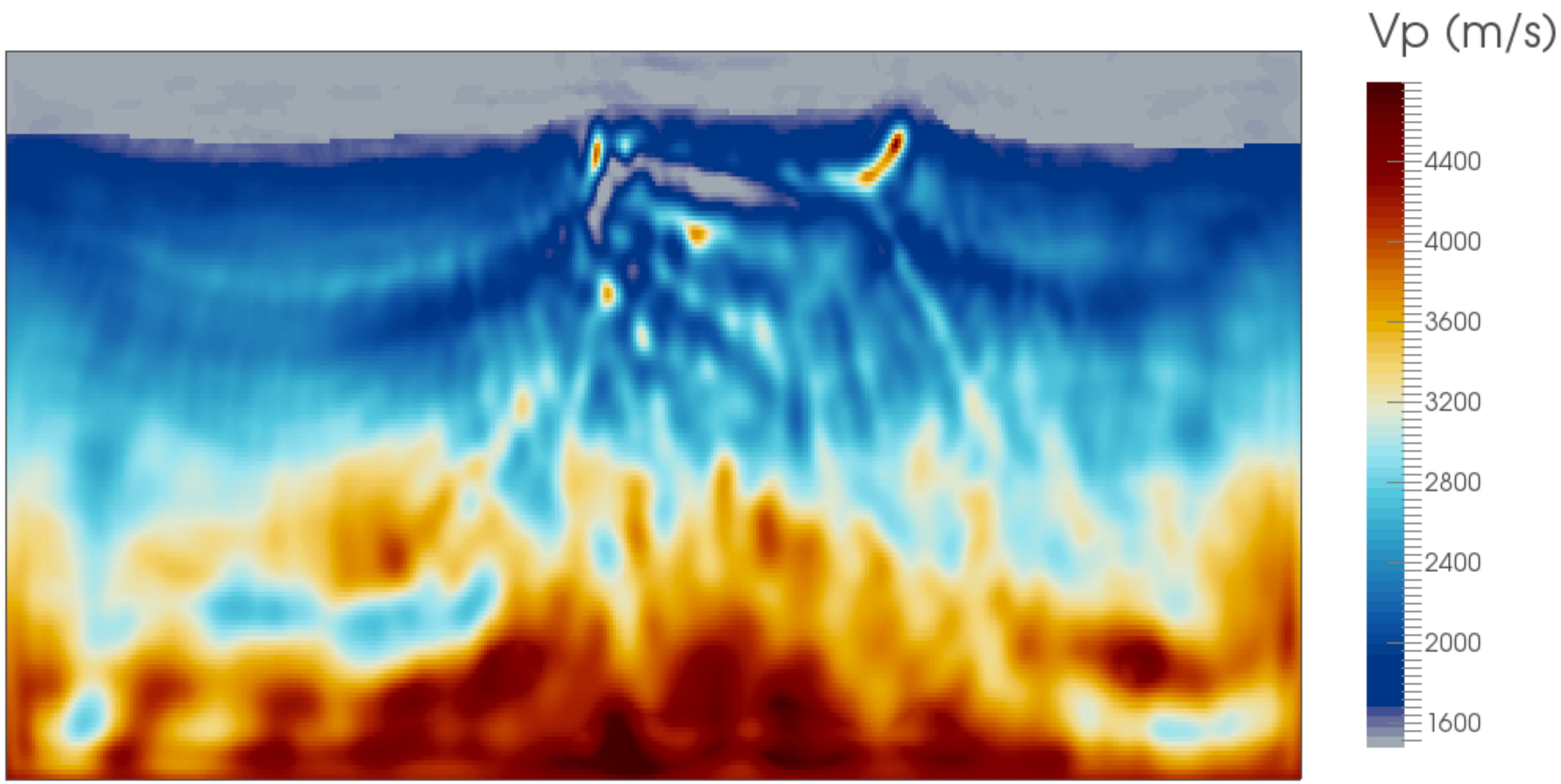


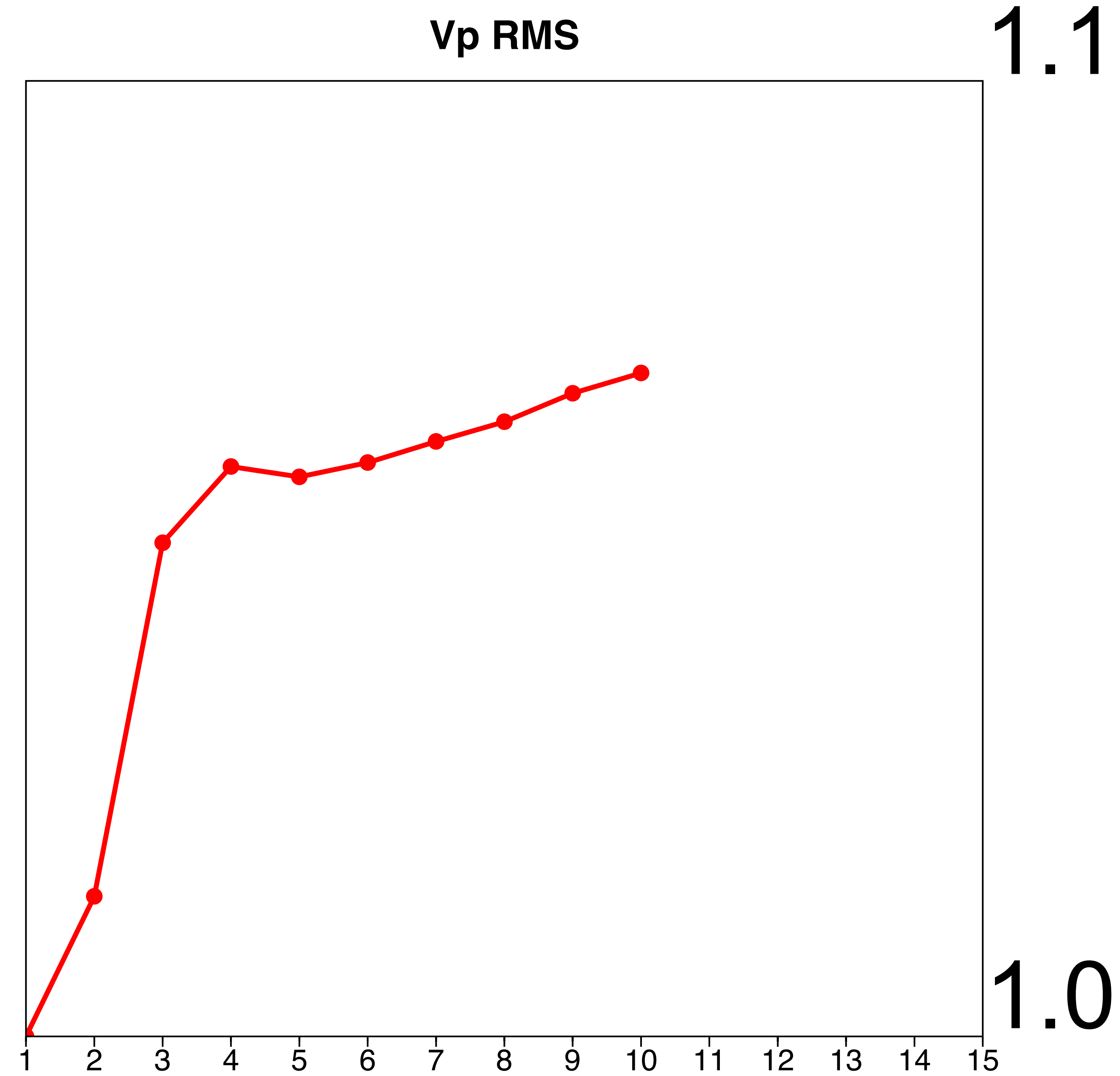
Adjoint-state – w/o constraints



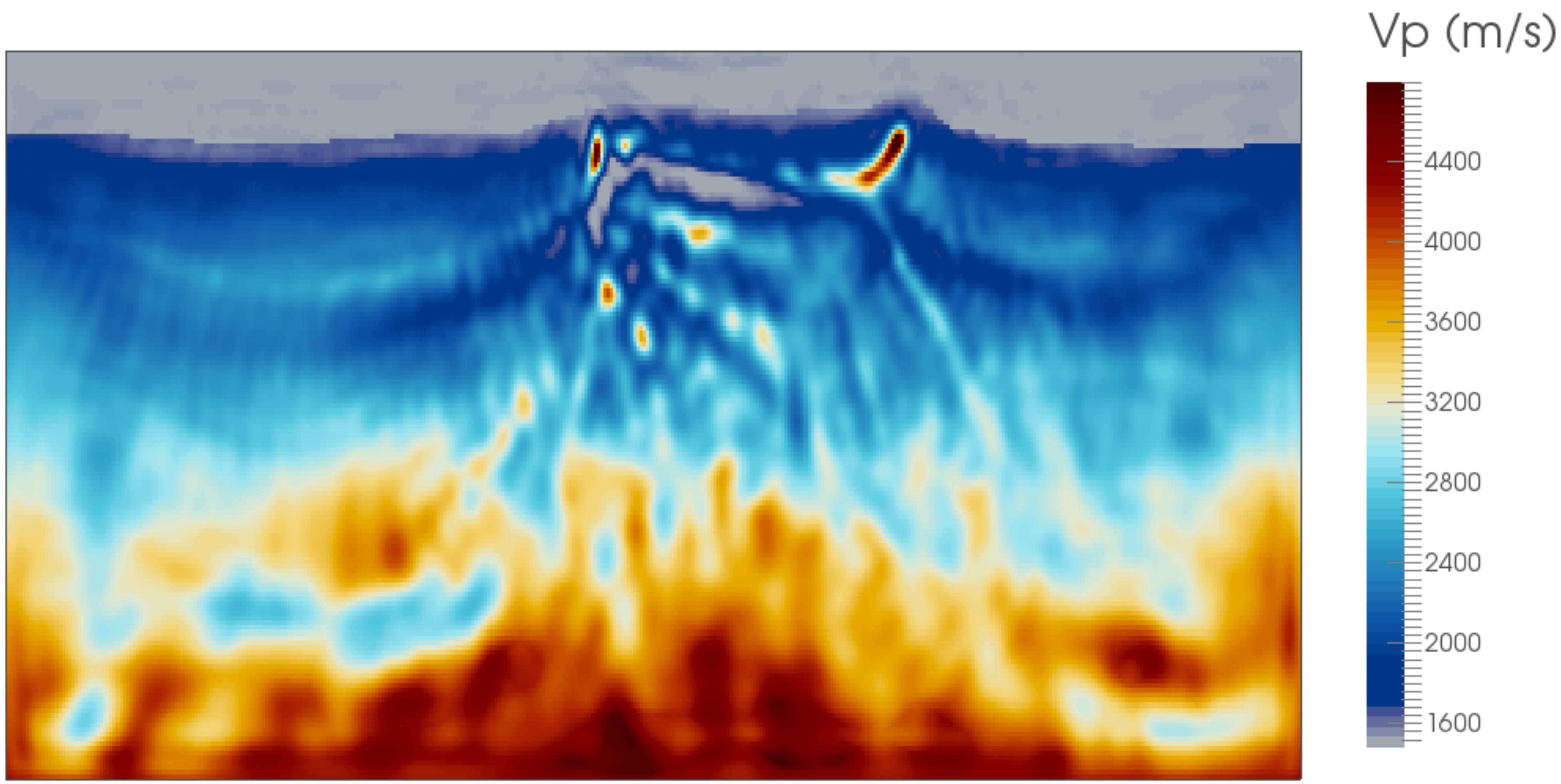


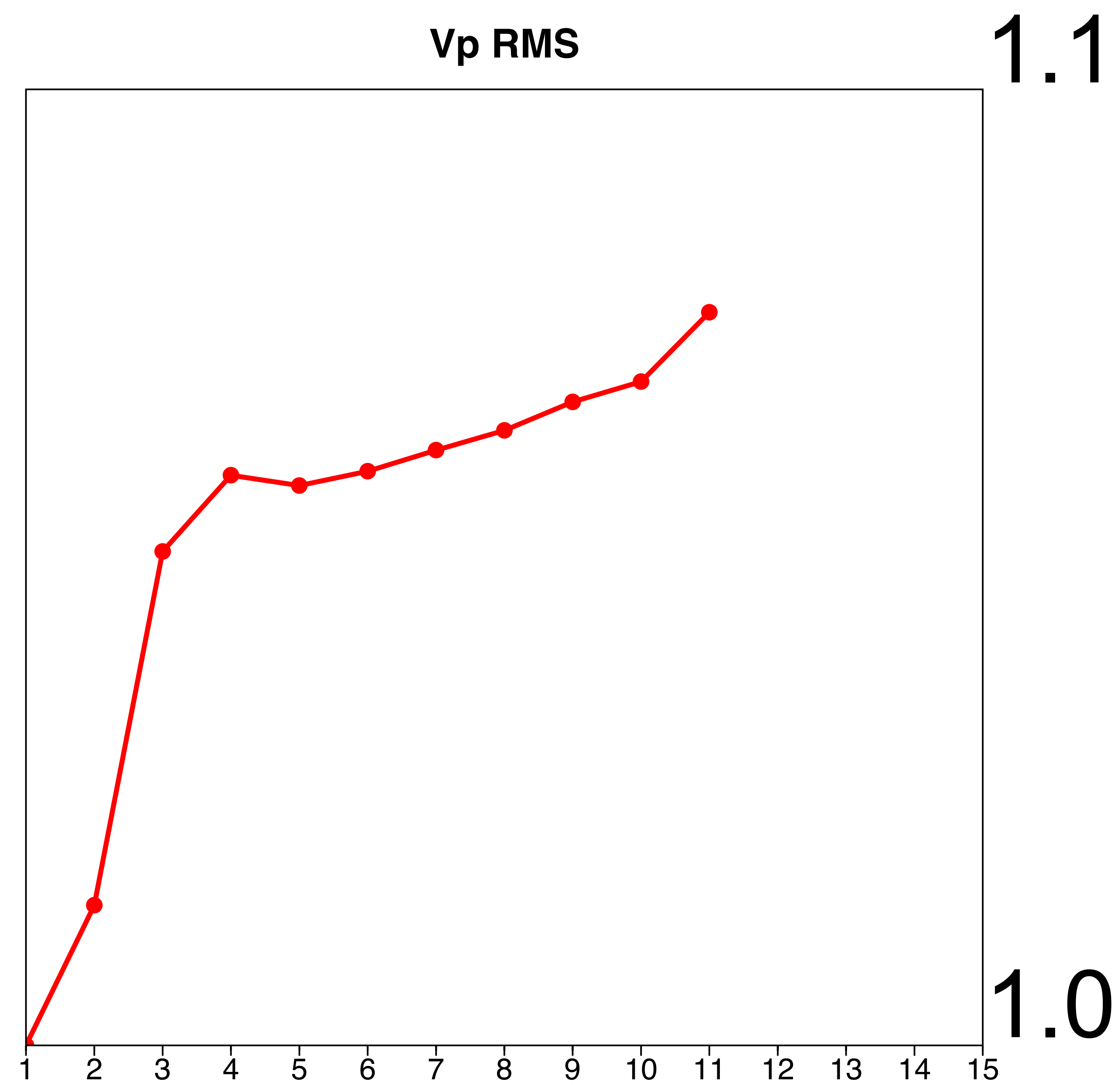
Adjoint-state – w/o constraints



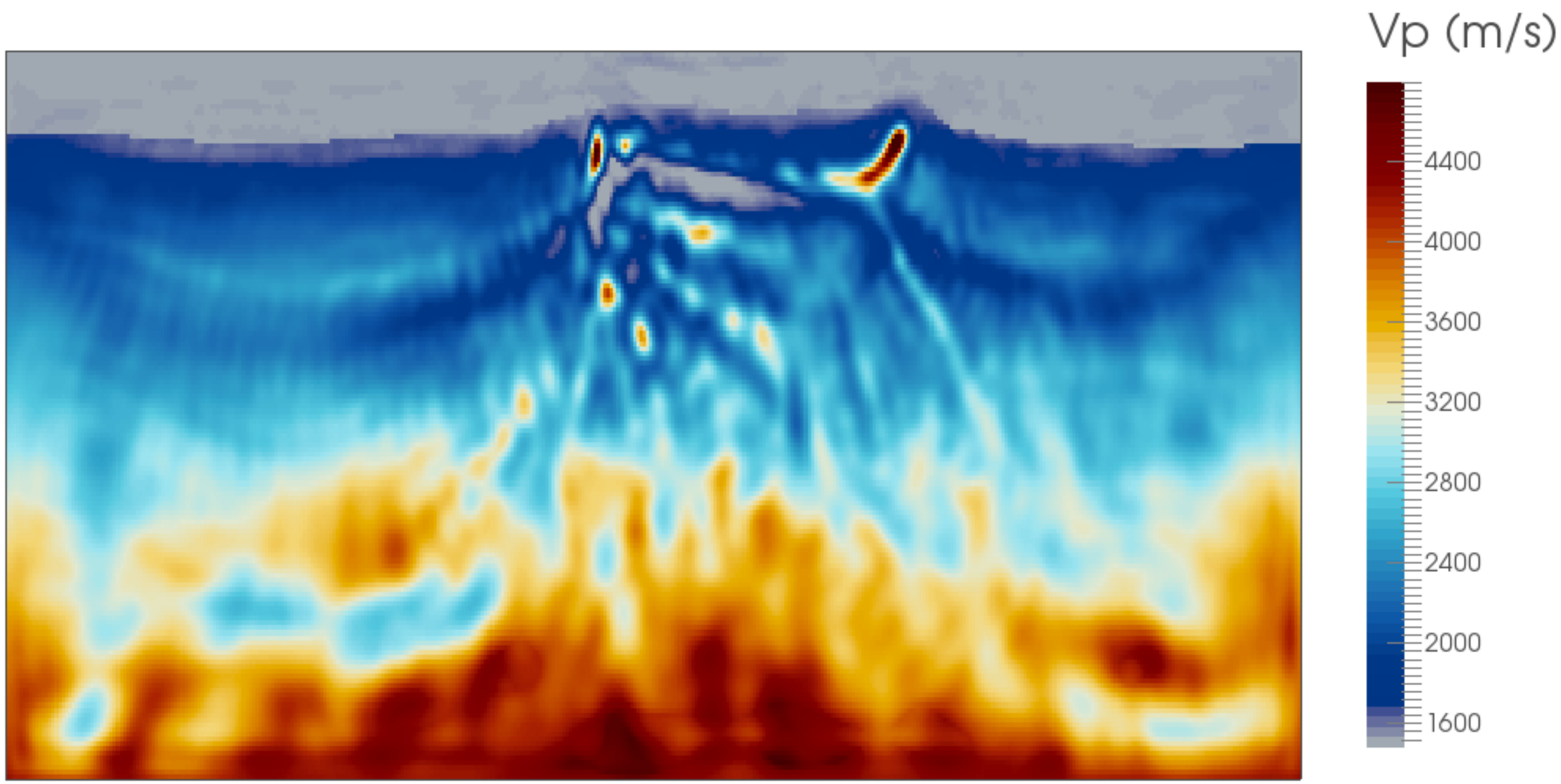


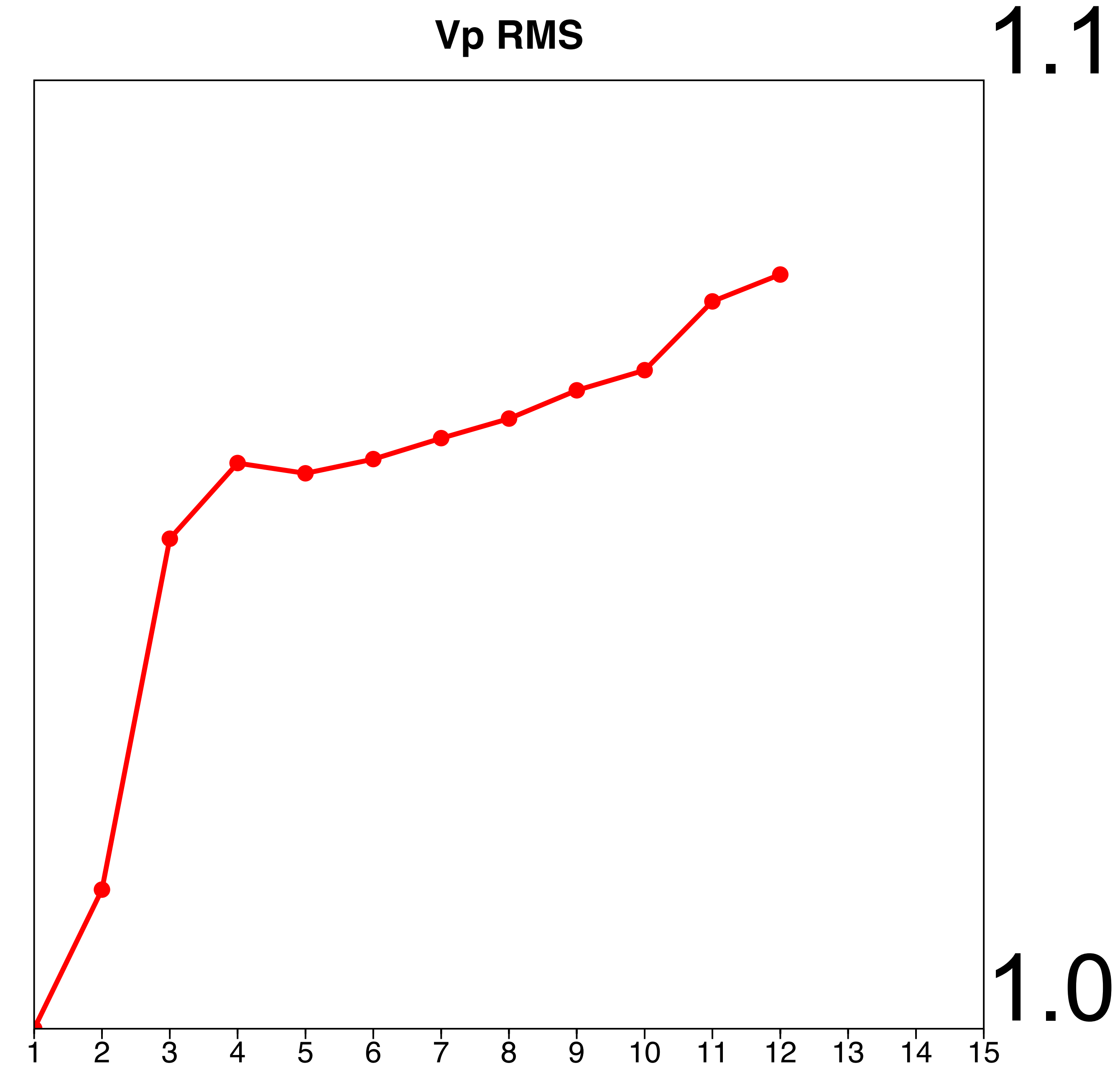
Adjoint-state – w/o constraints



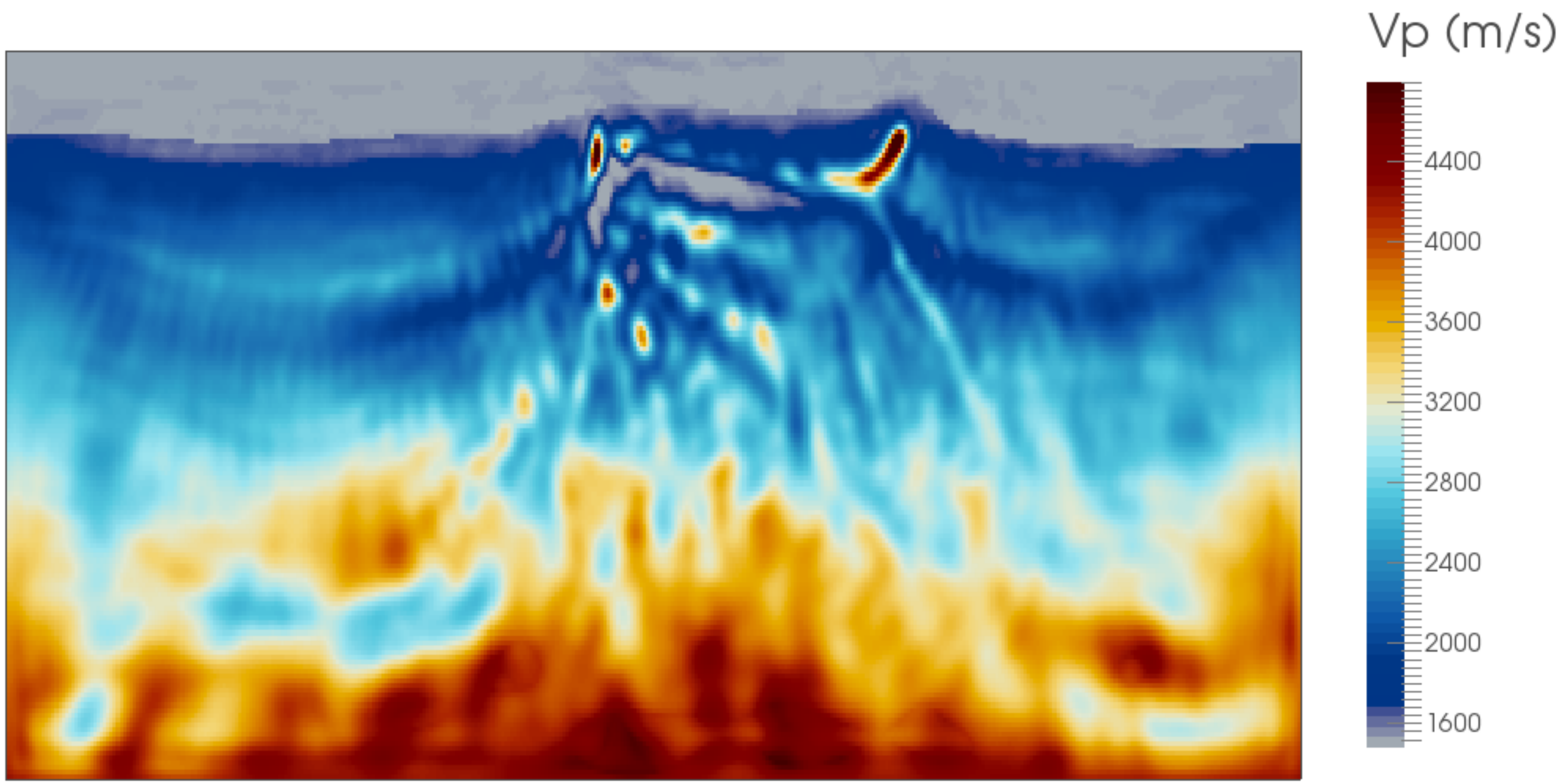


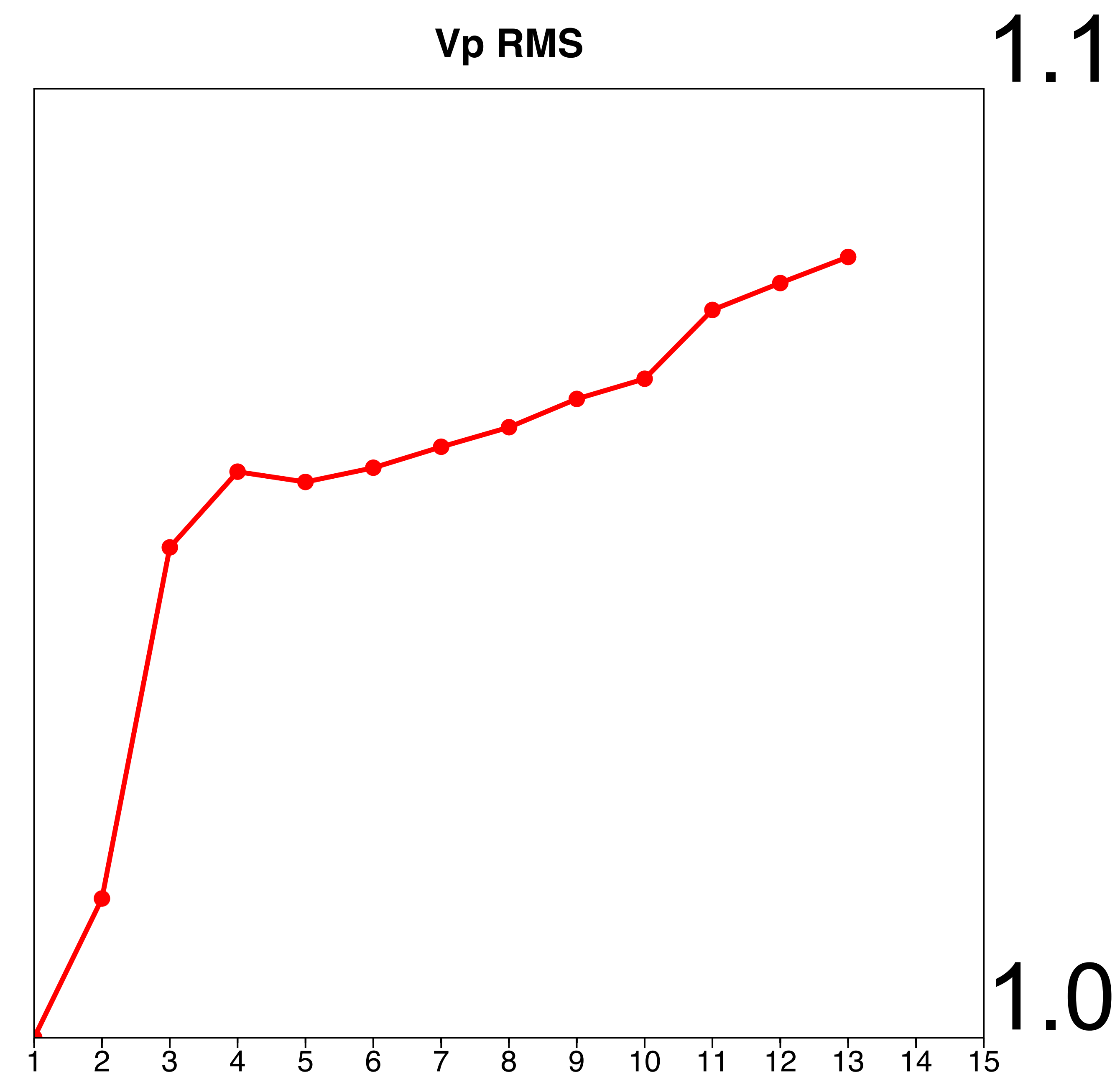
Adjoint-state – w/o constraints





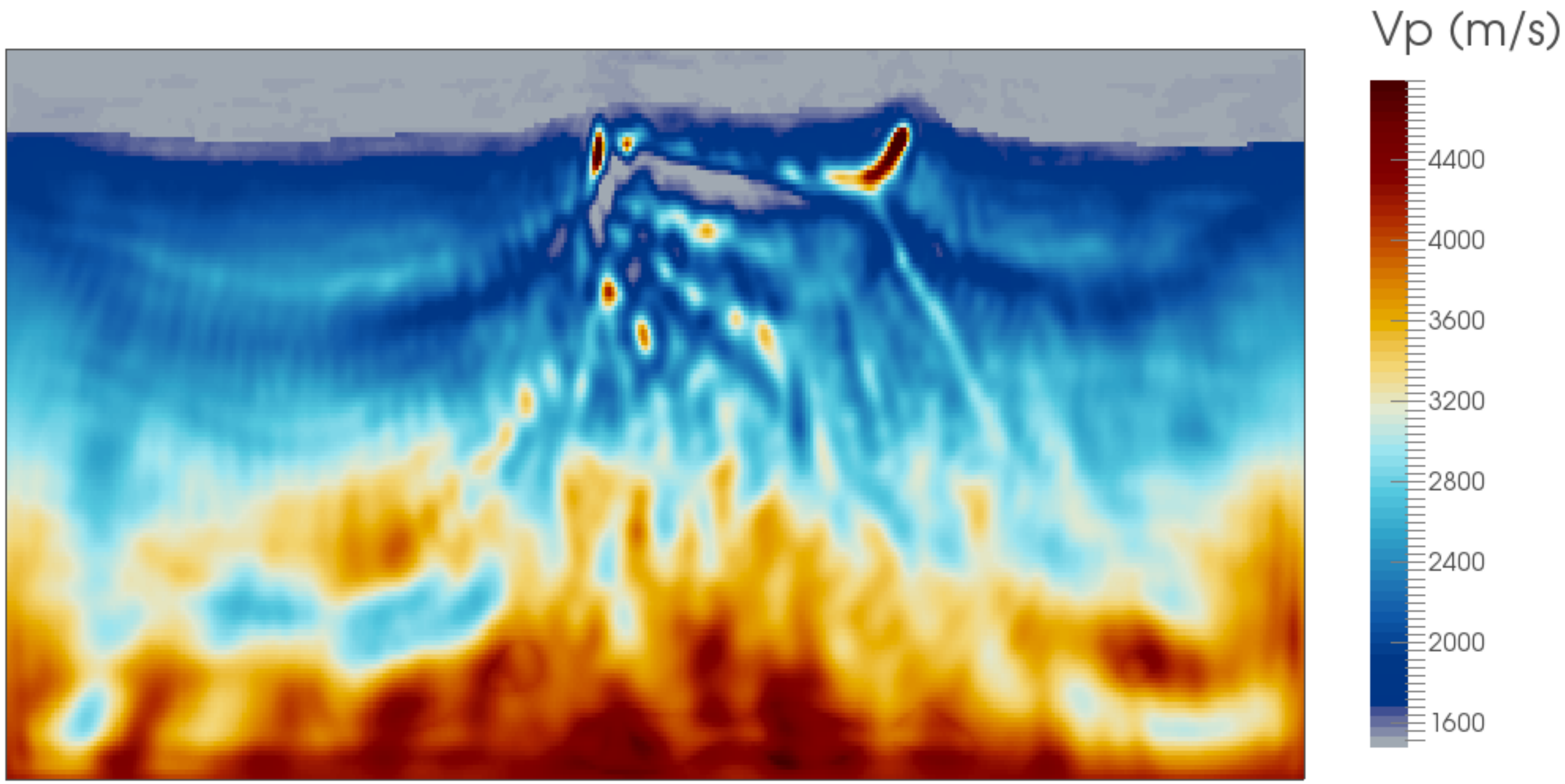
Adjoint-state – w/o constraints

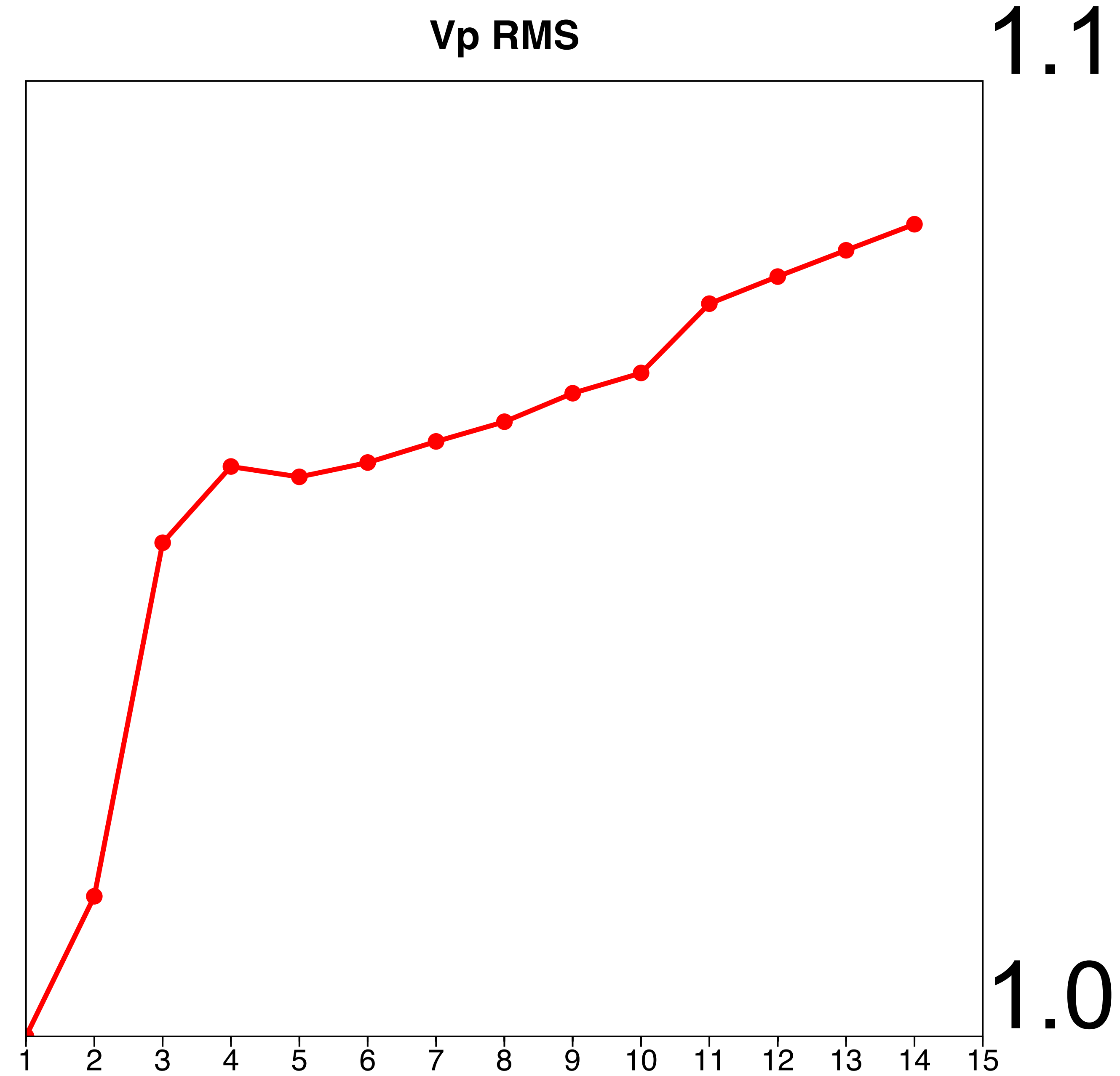




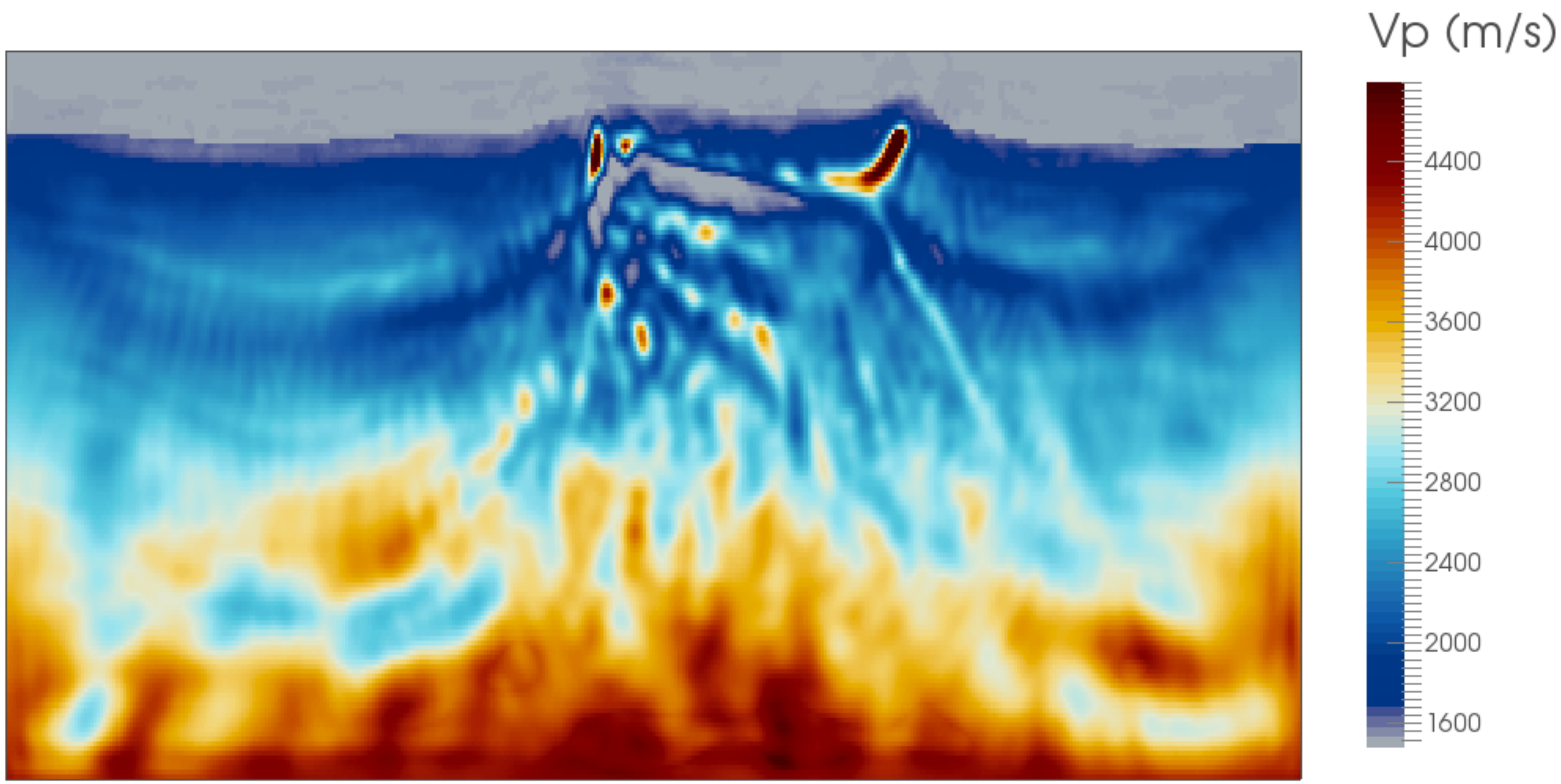
Adjoint-state – w/o constraints

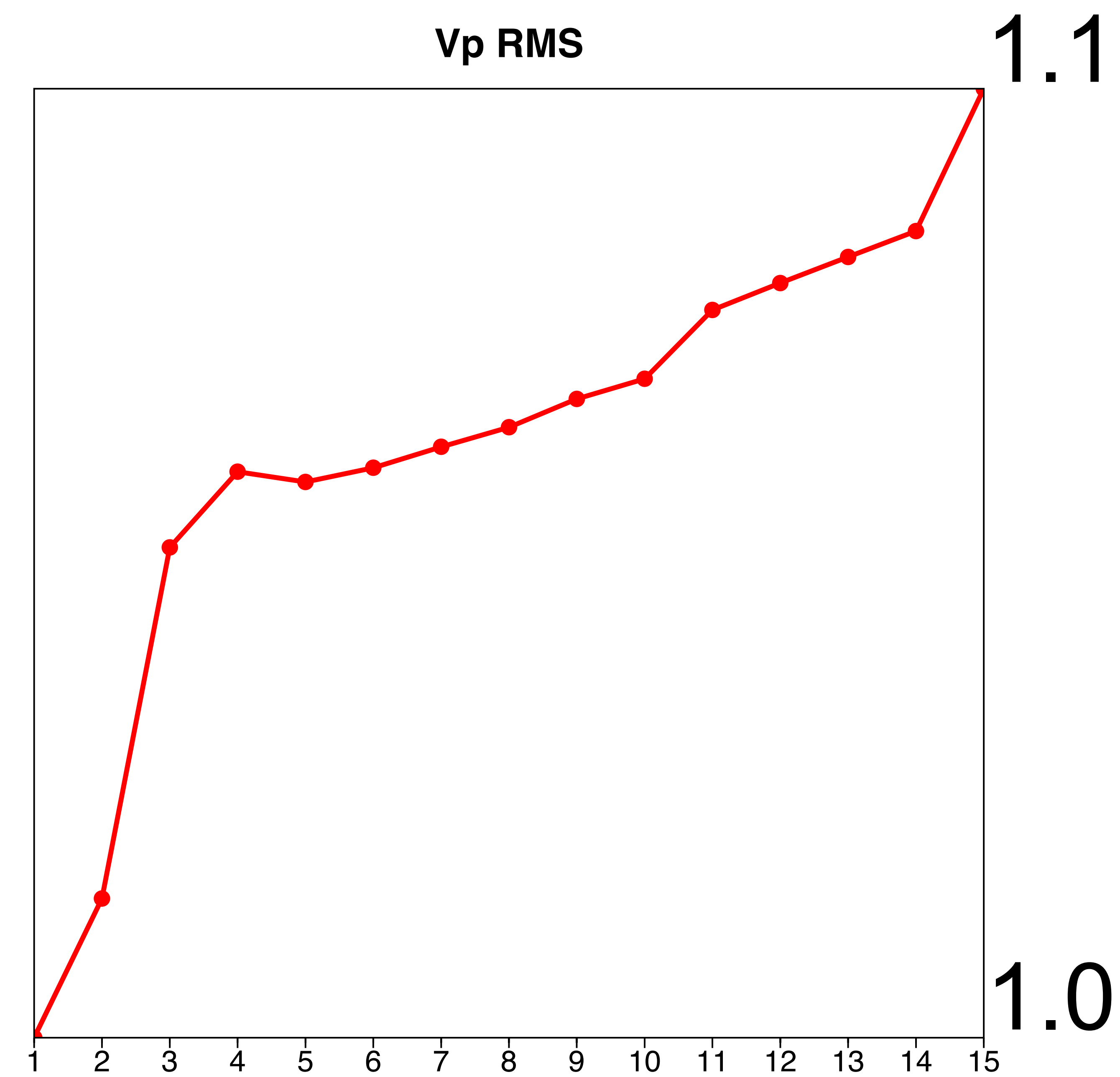
SLIM



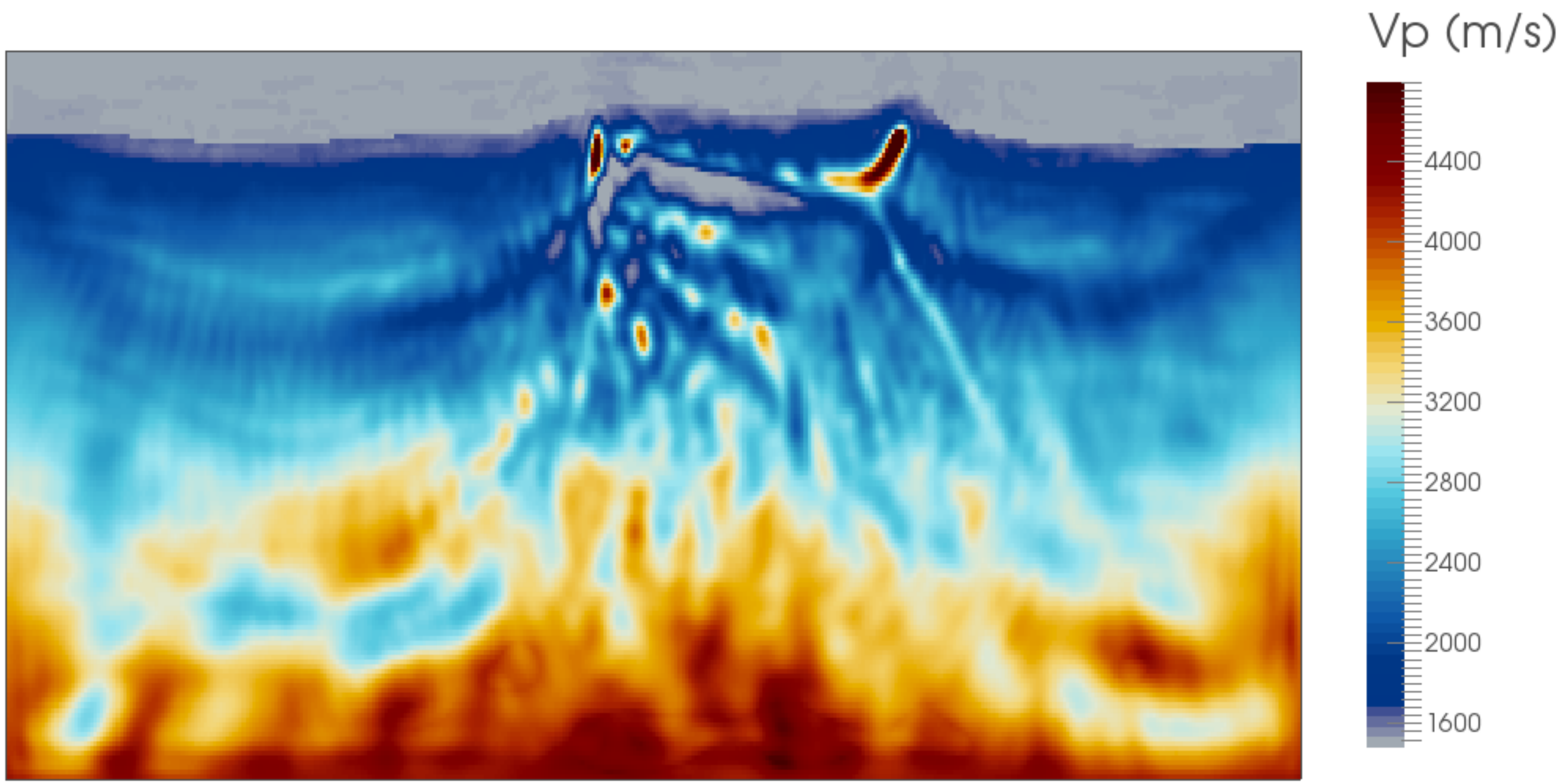


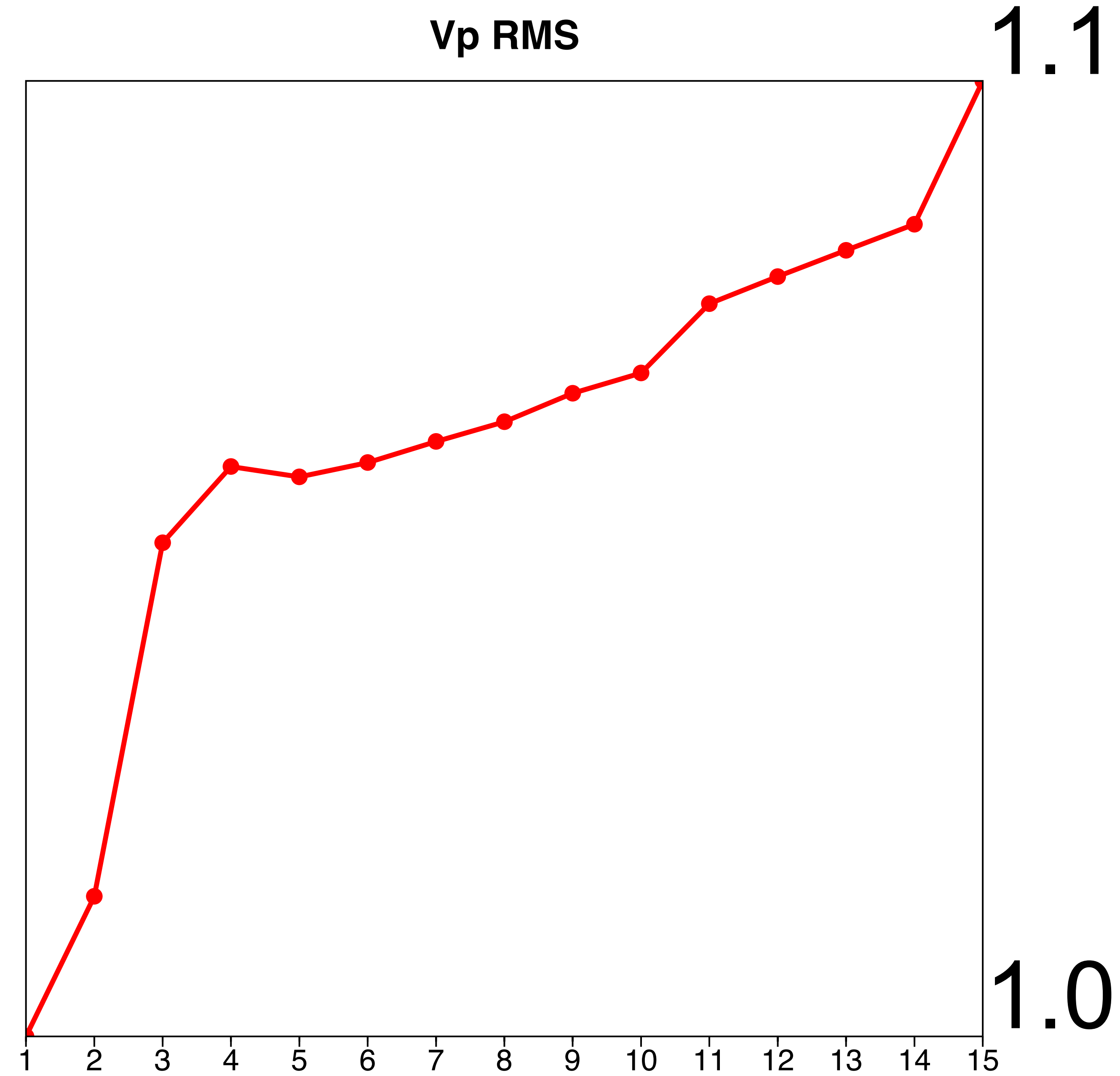
Adjoint-state – w/o constraints



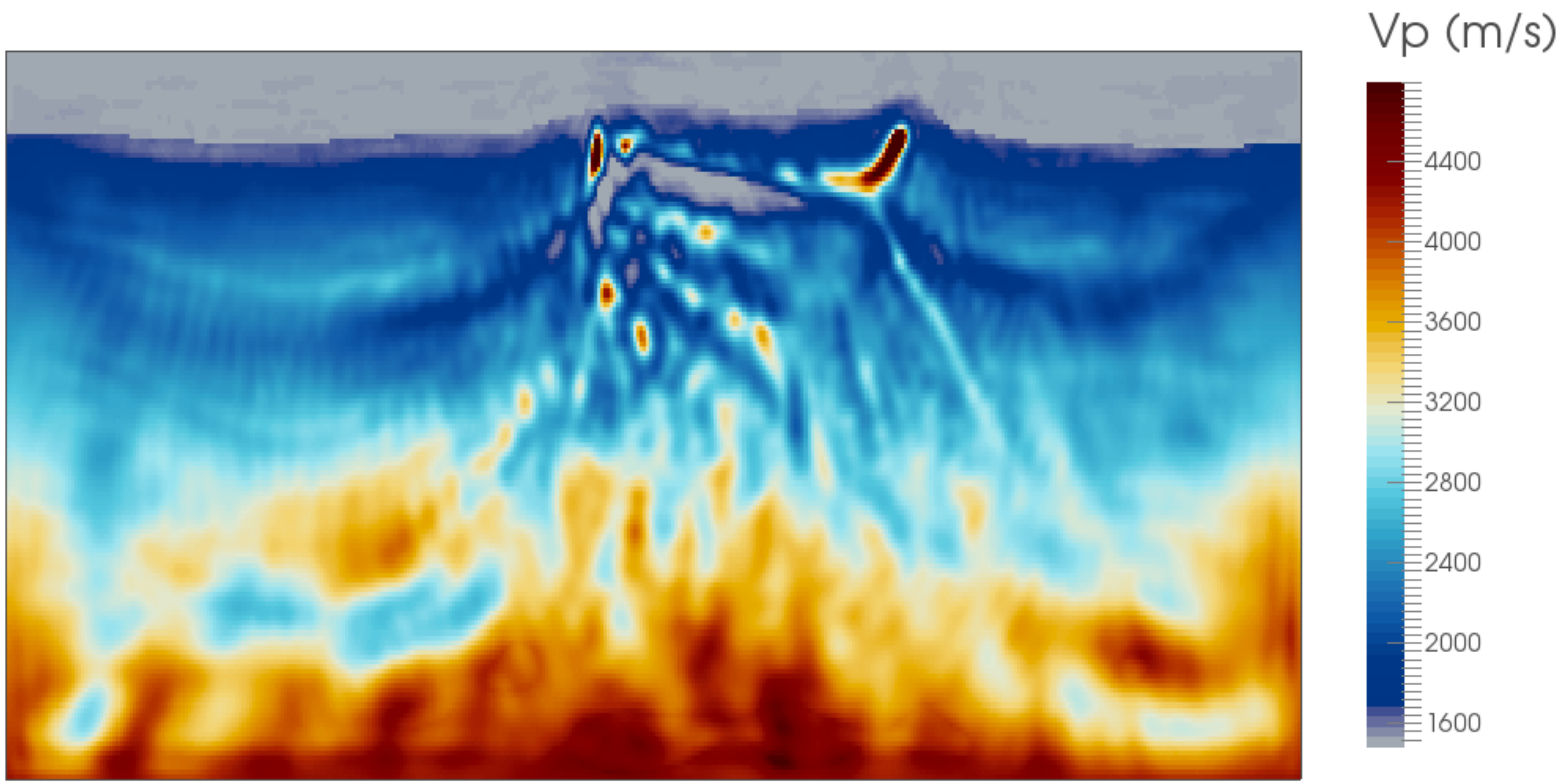


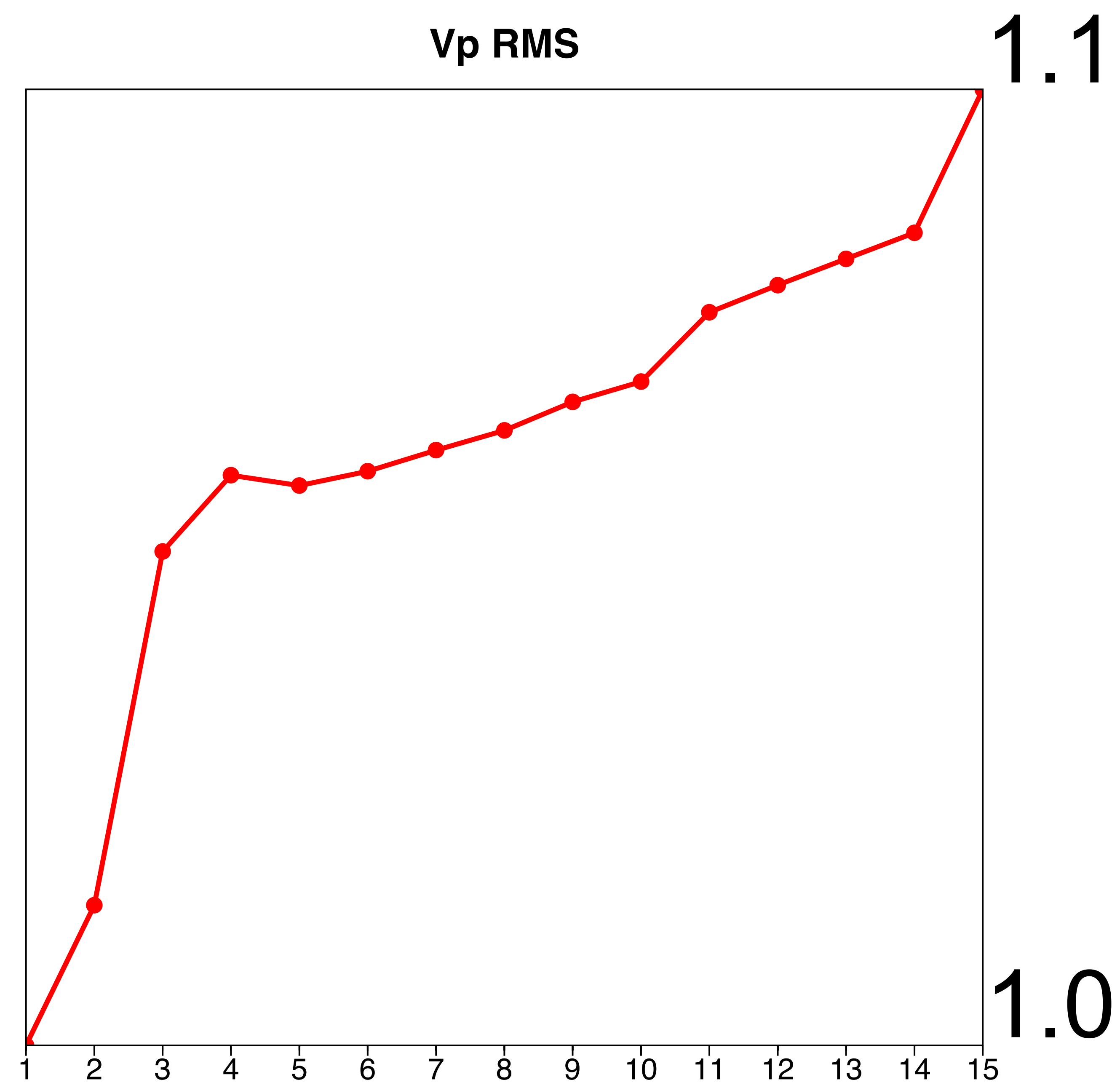
Adjoint-state – w/o constraints



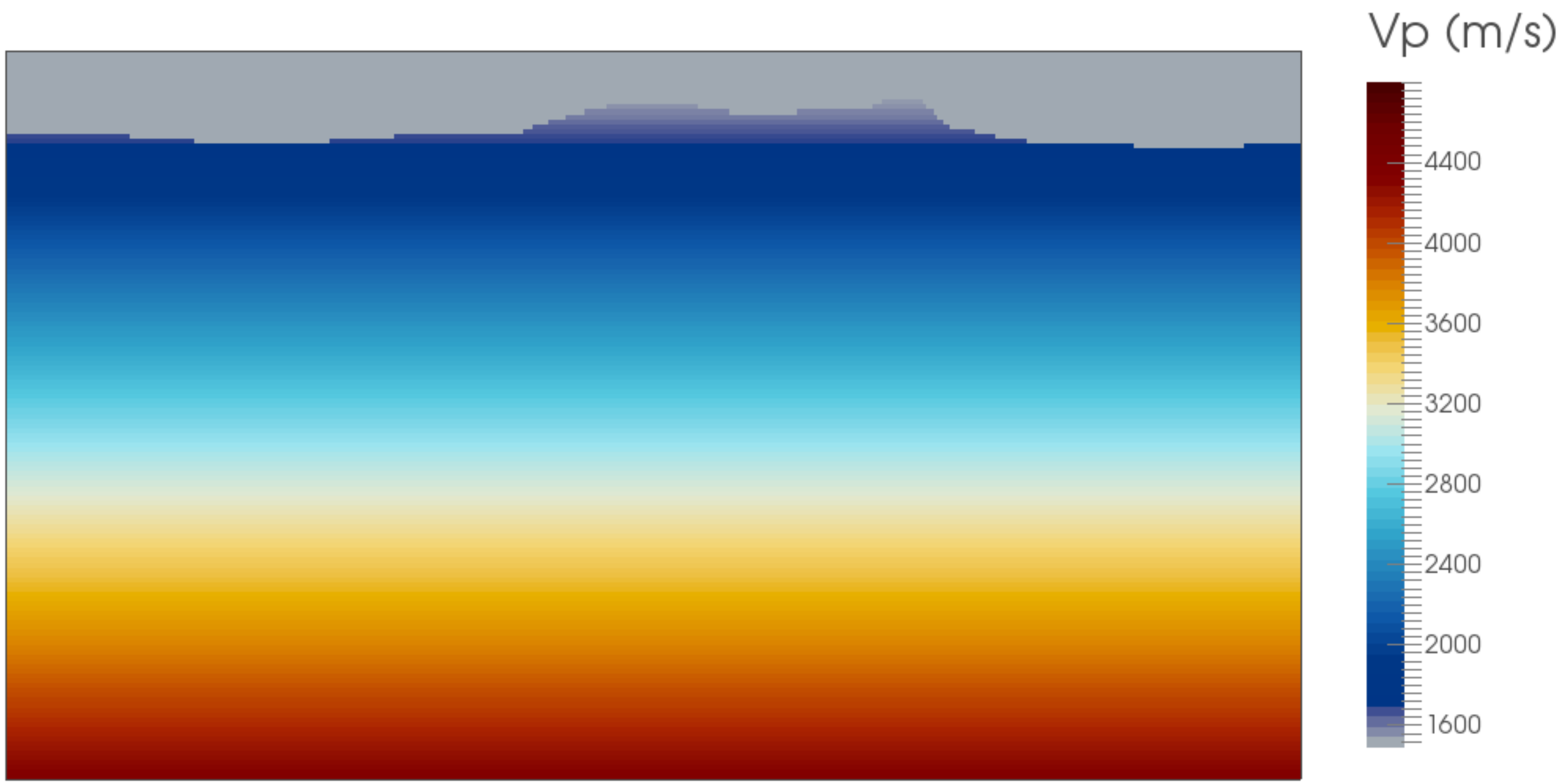


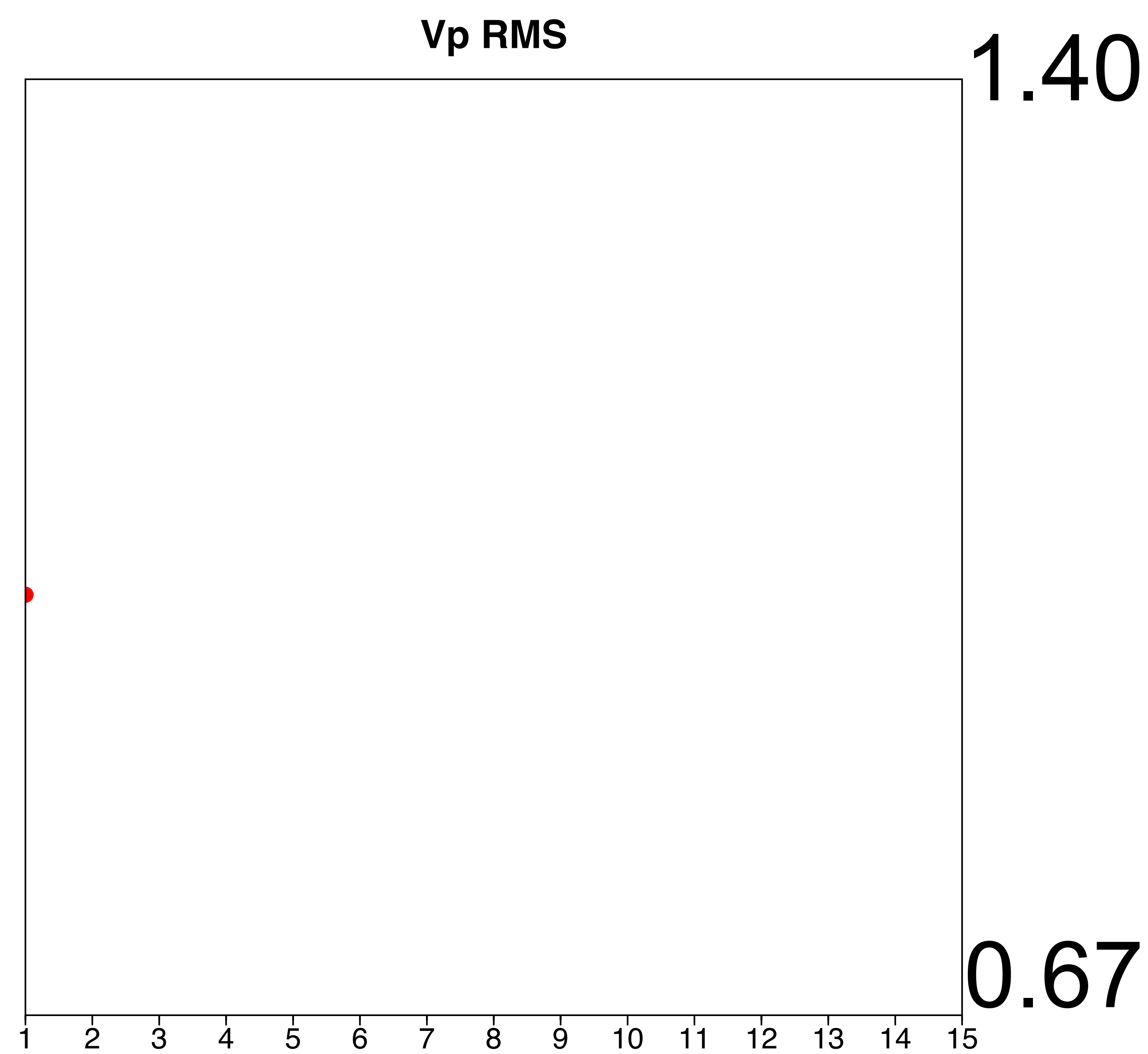
Adjoint-state – w/o constraints



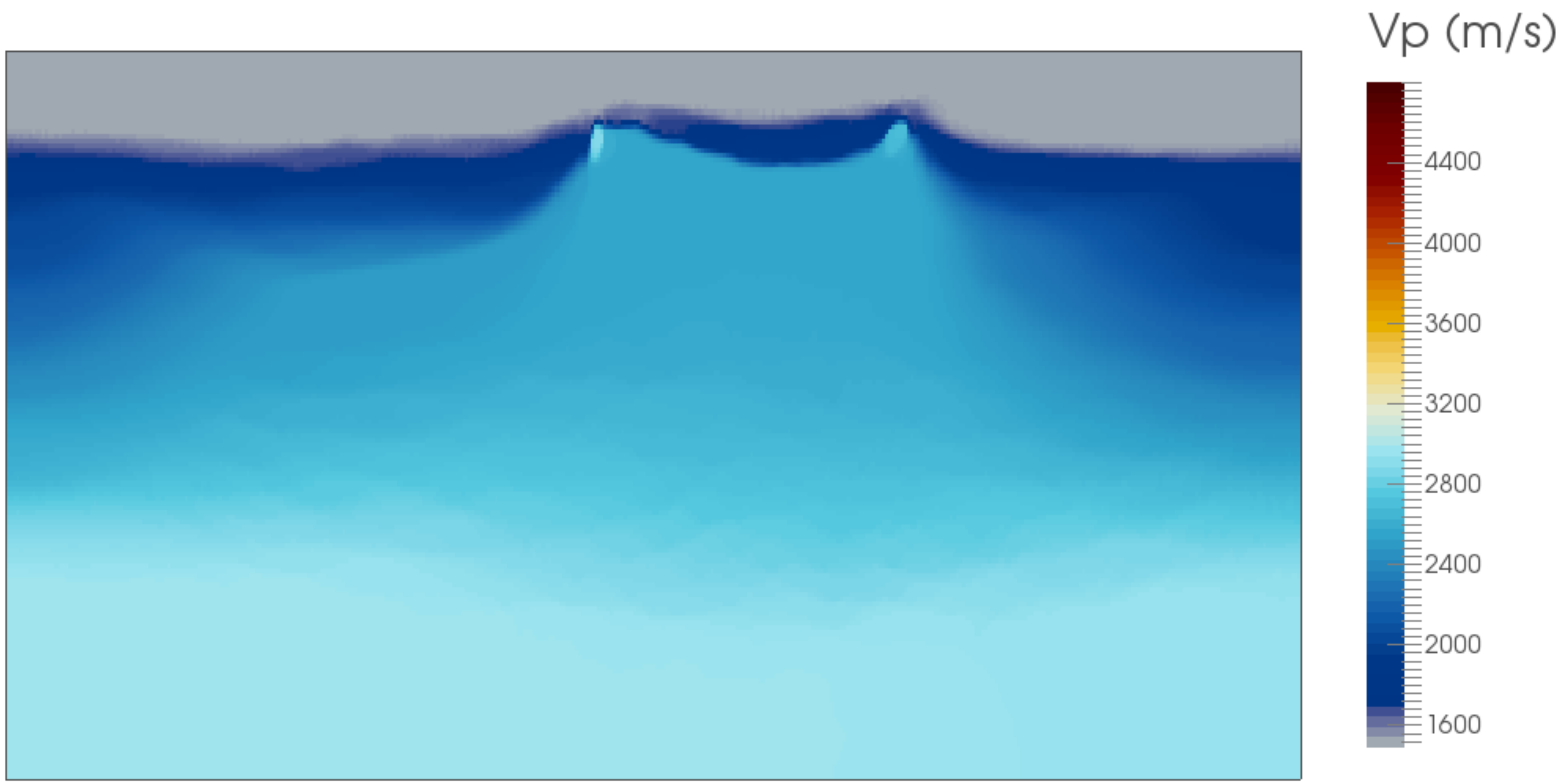


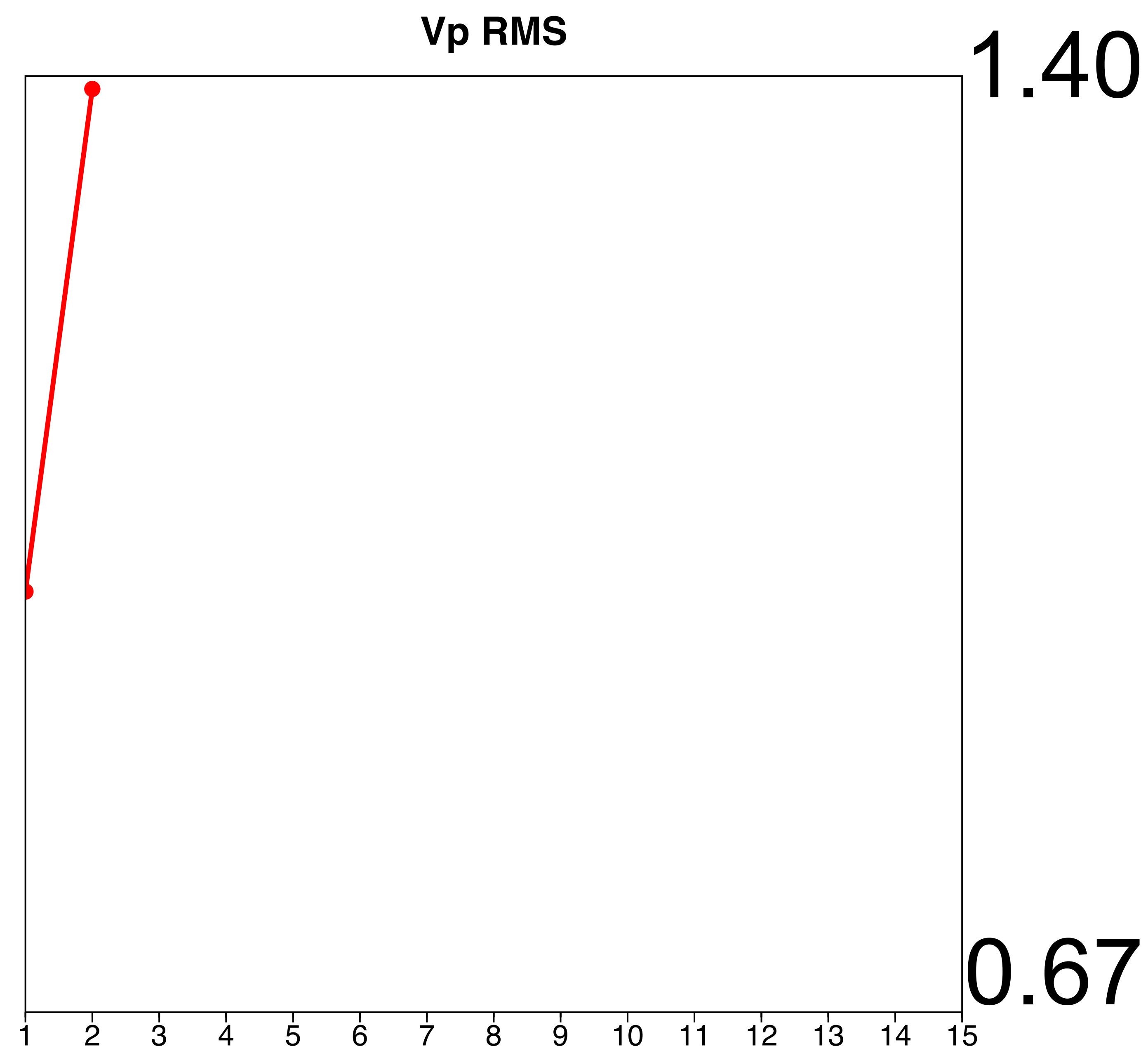
Adjoint-state – w/ TV-norm & hinge-loss projections



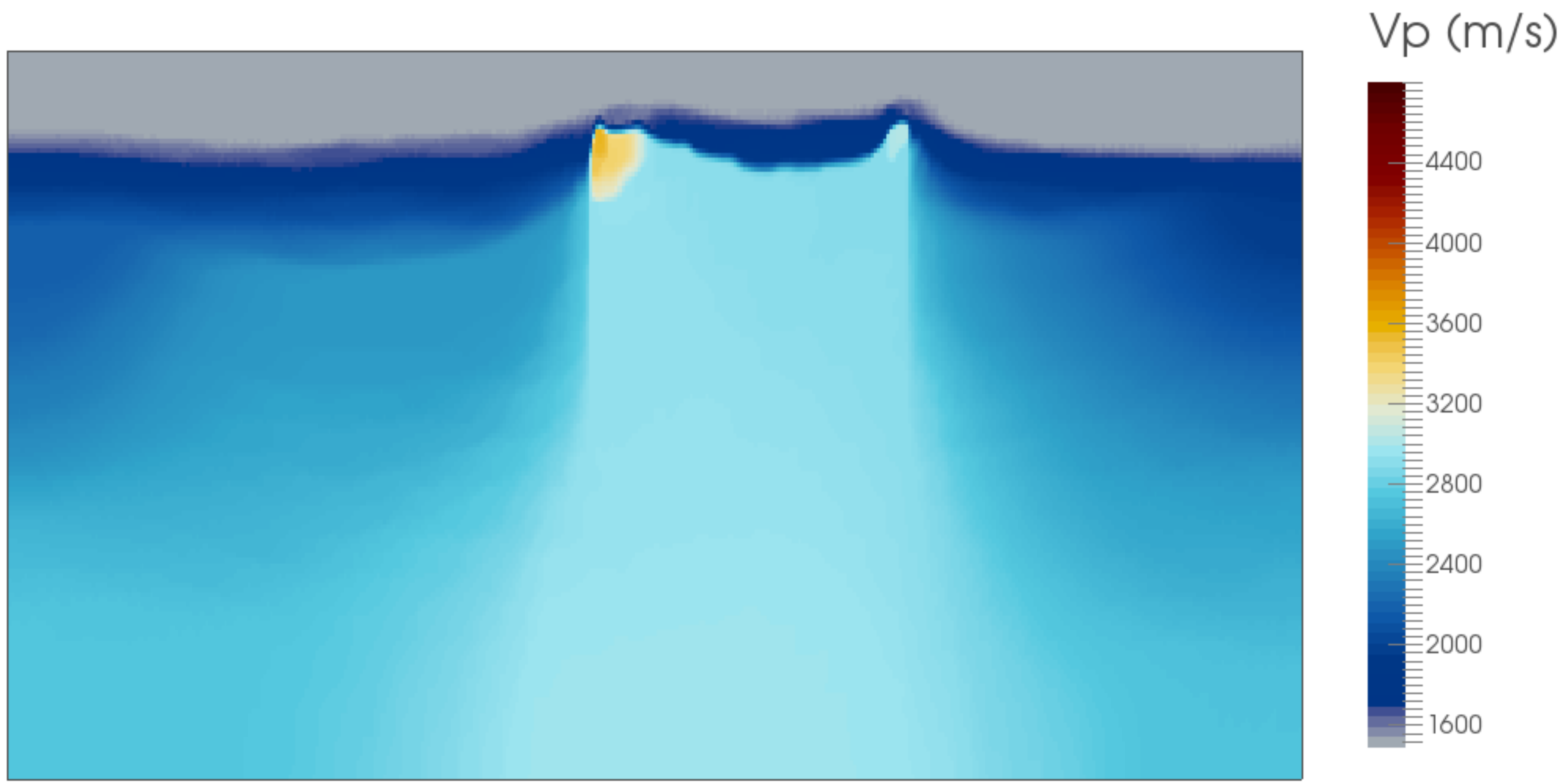


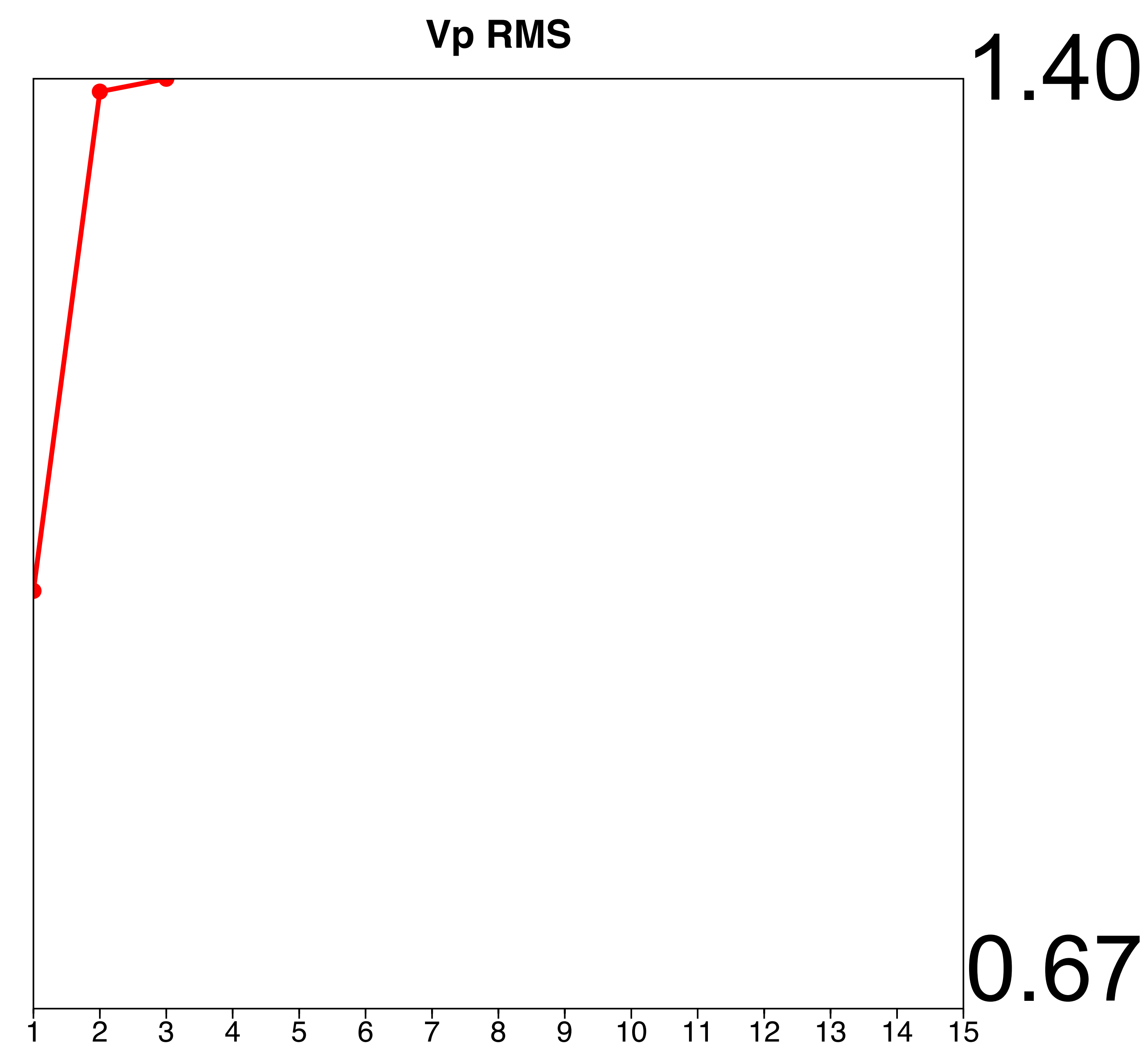
Adjoint-state – w/ TV-norm & hinge-loss projections



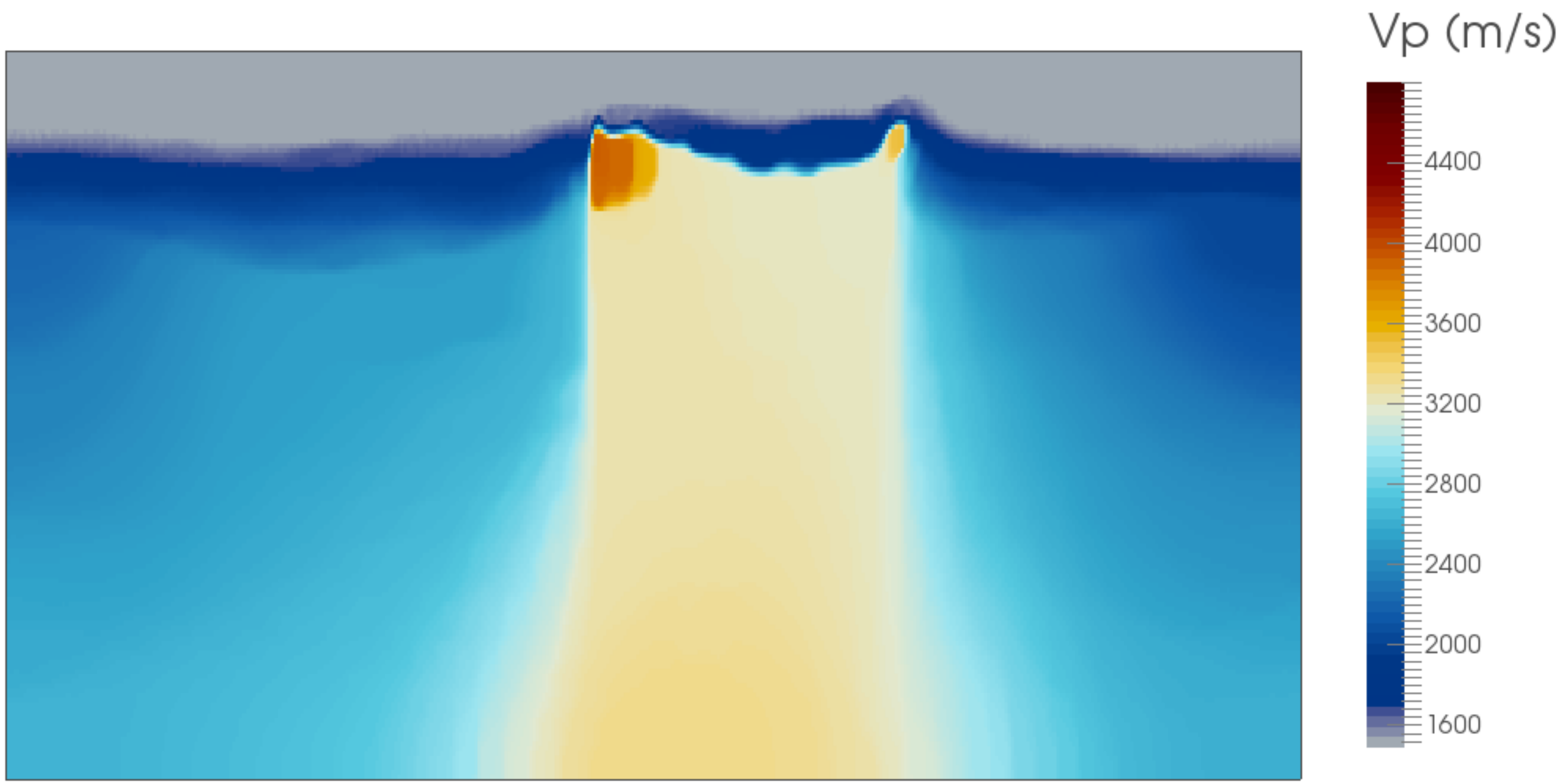


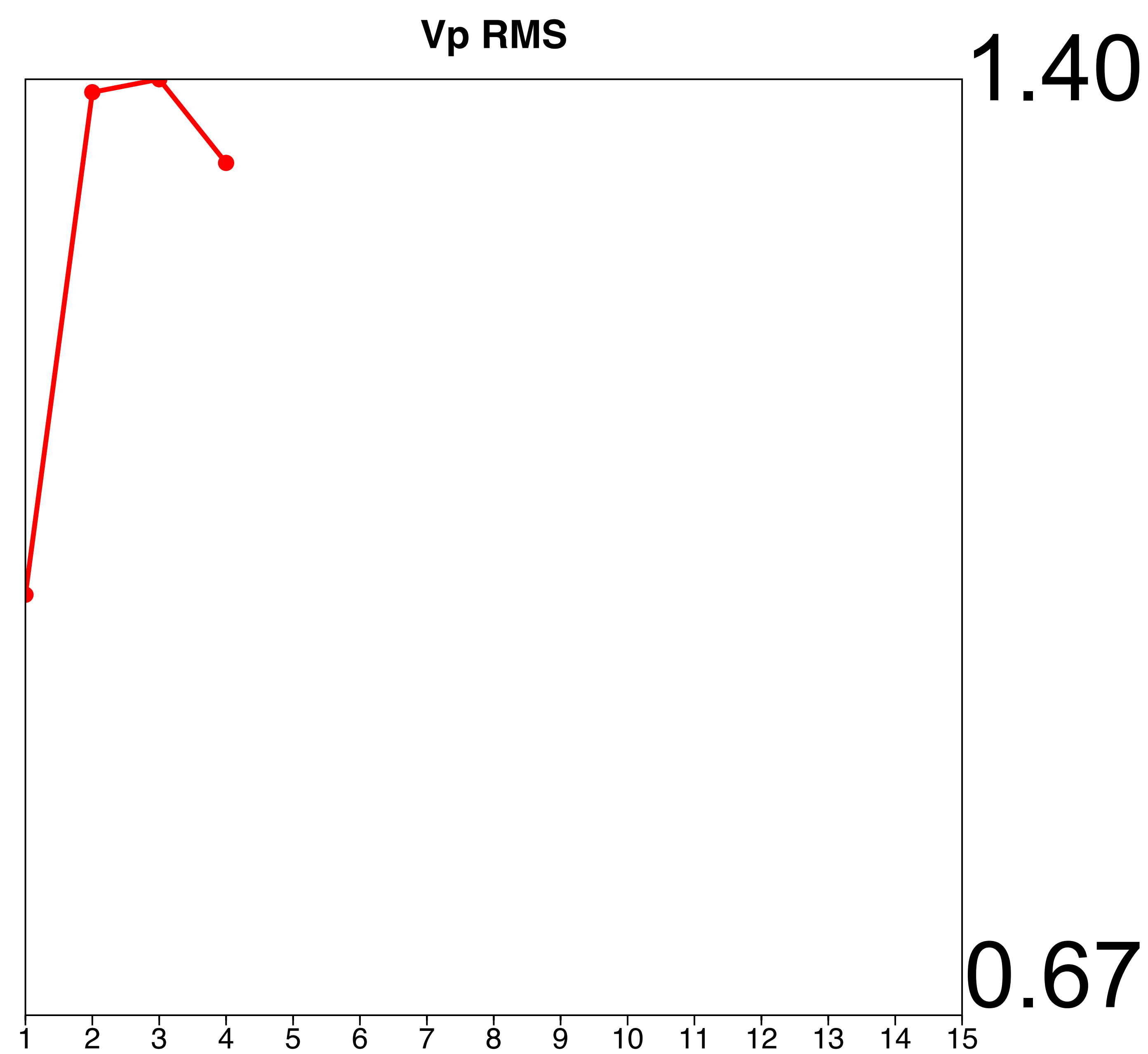
Adjoint-state – w/ TV-norm & hinge-loss projections



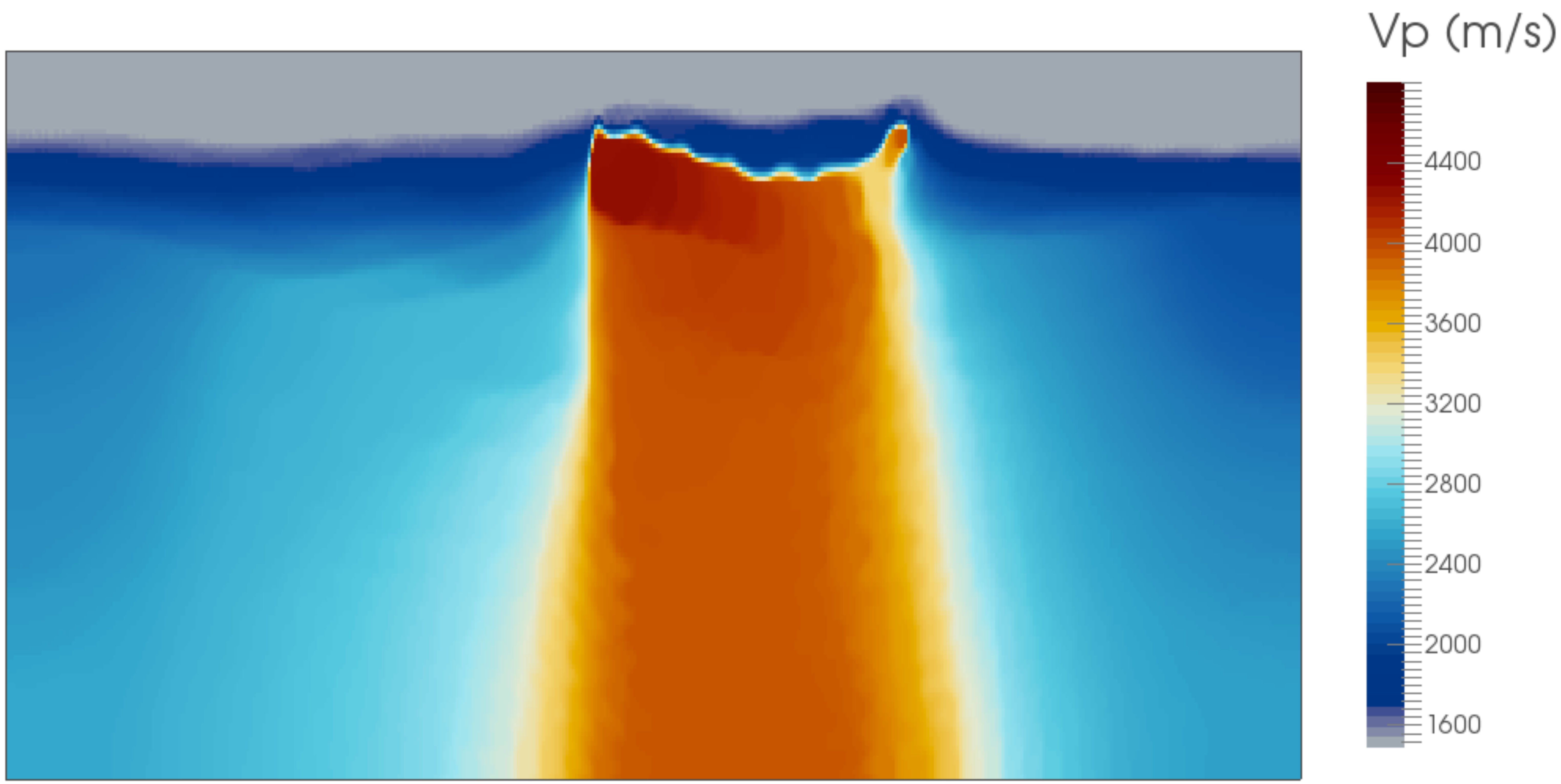


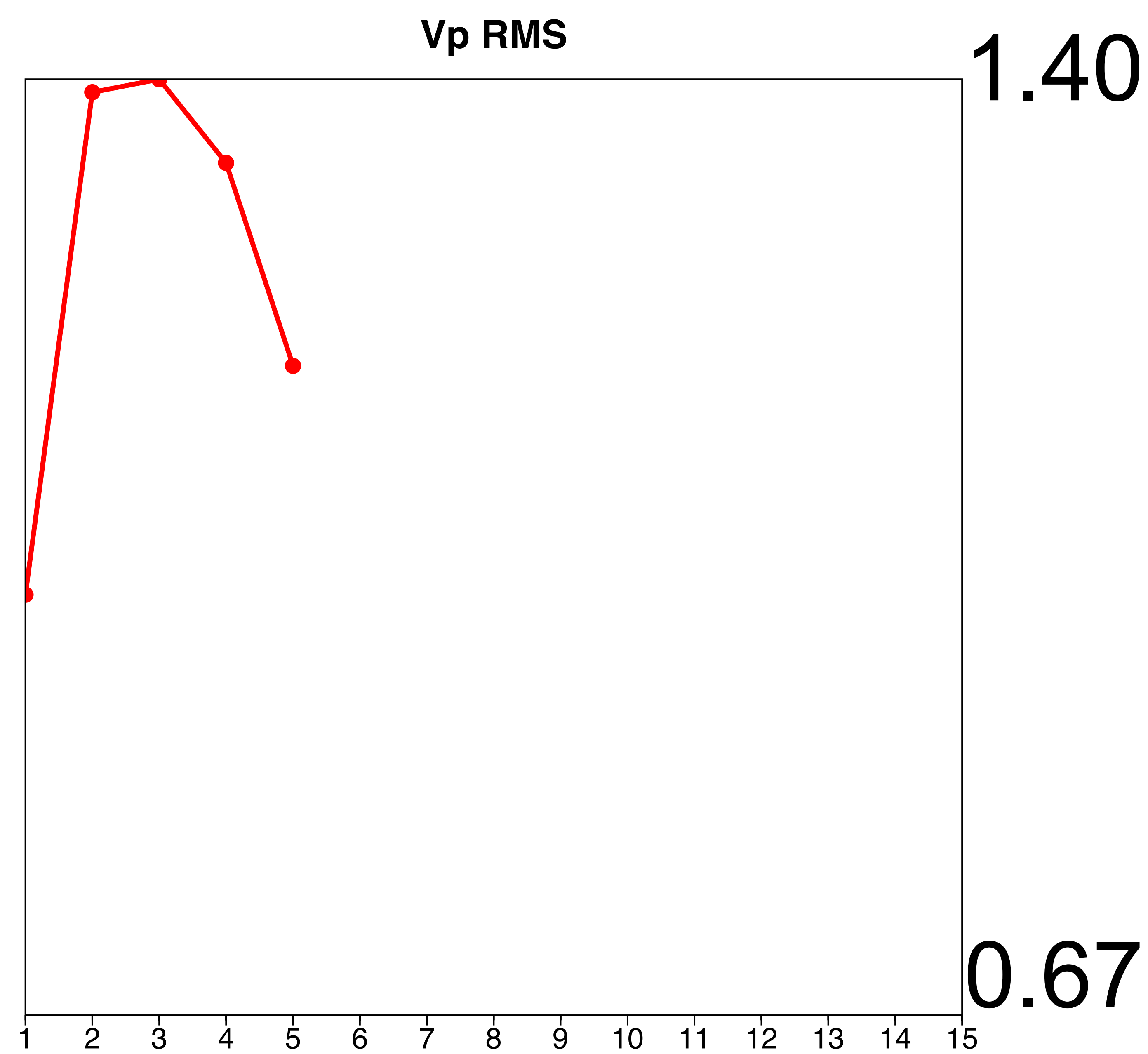
Adjoint-state – w/ TV-norm & hinge-loss projections



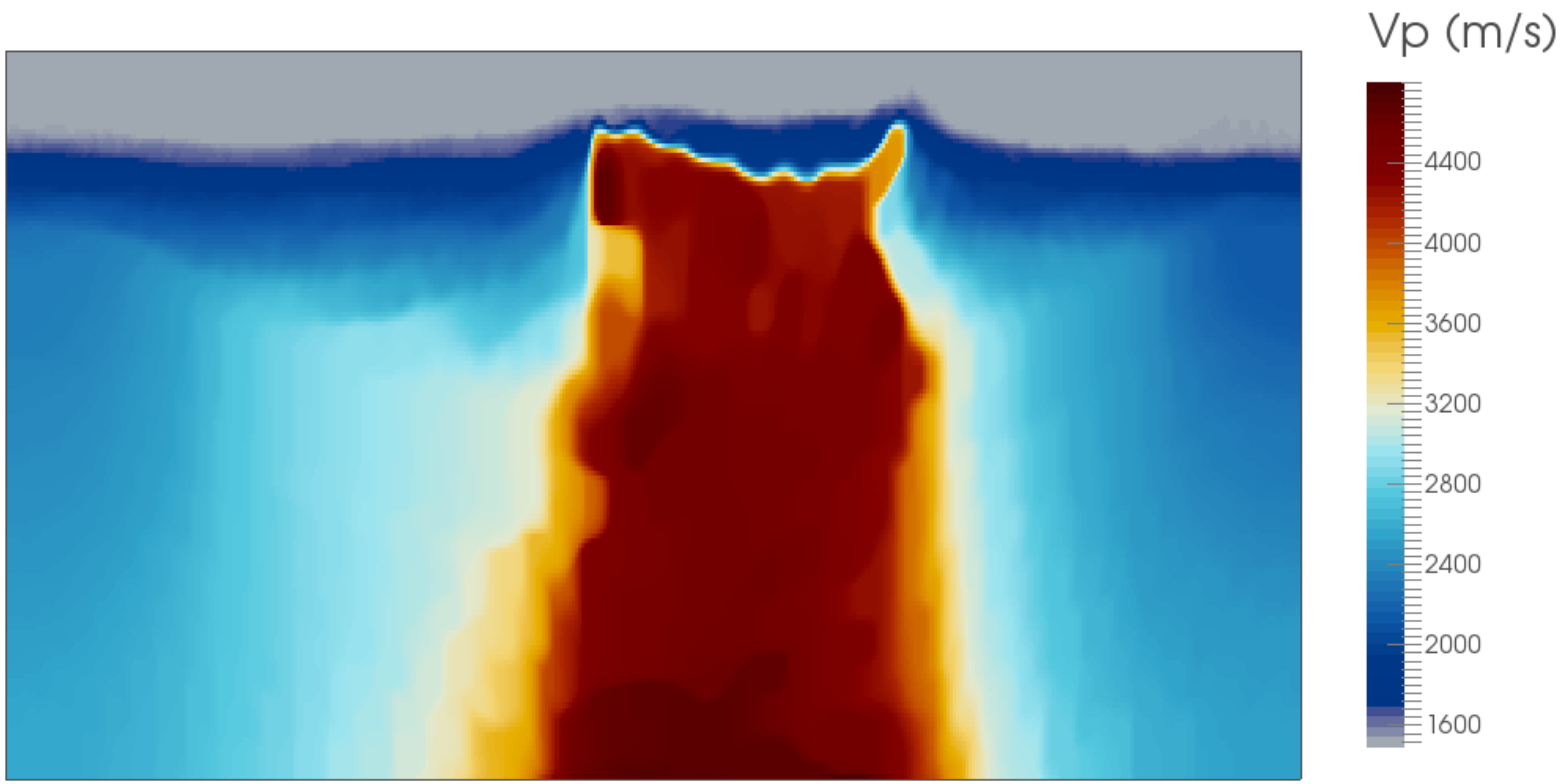


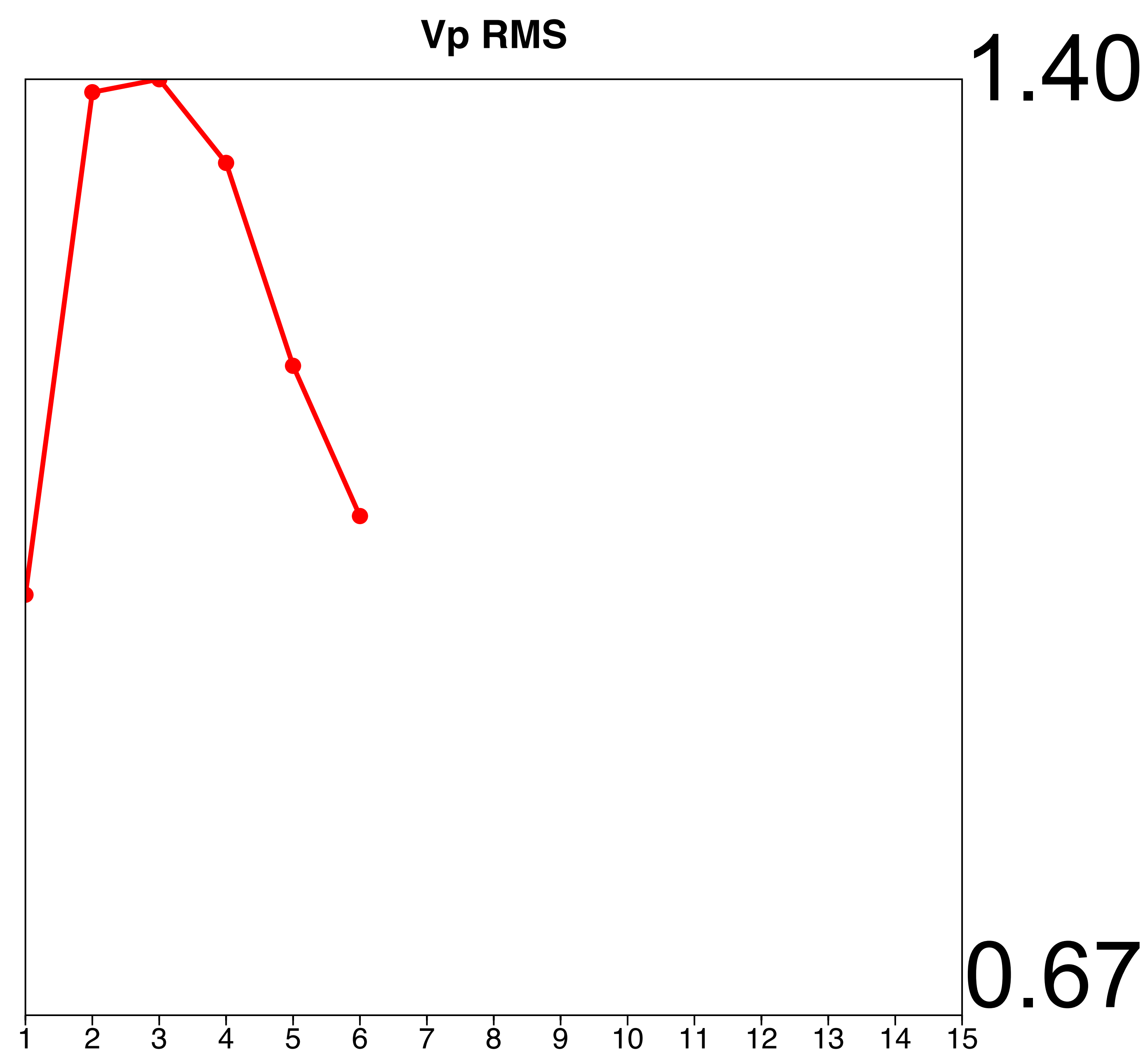
Adjoint-state – w/ TV-norm & hinge-loss projections



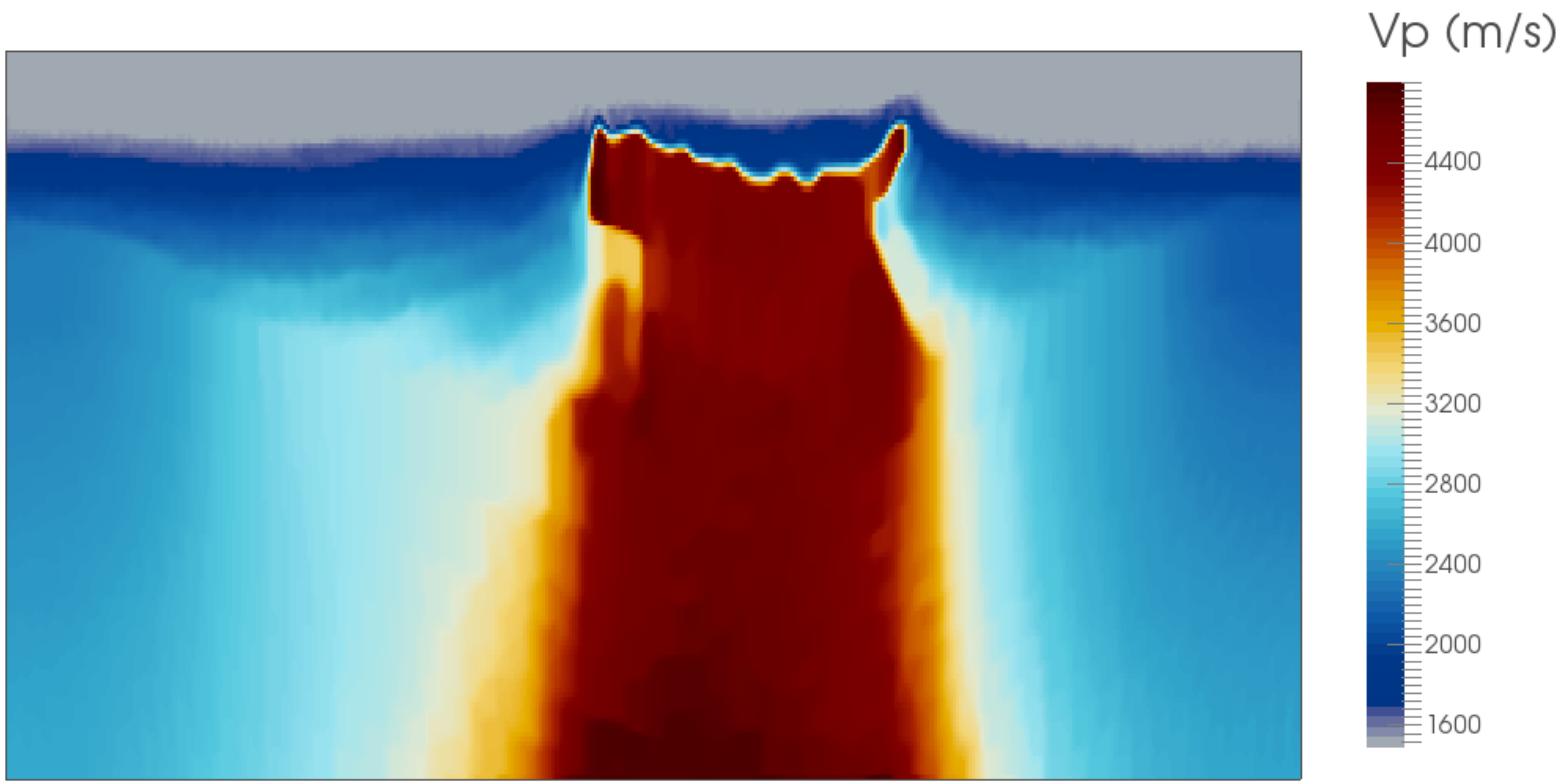


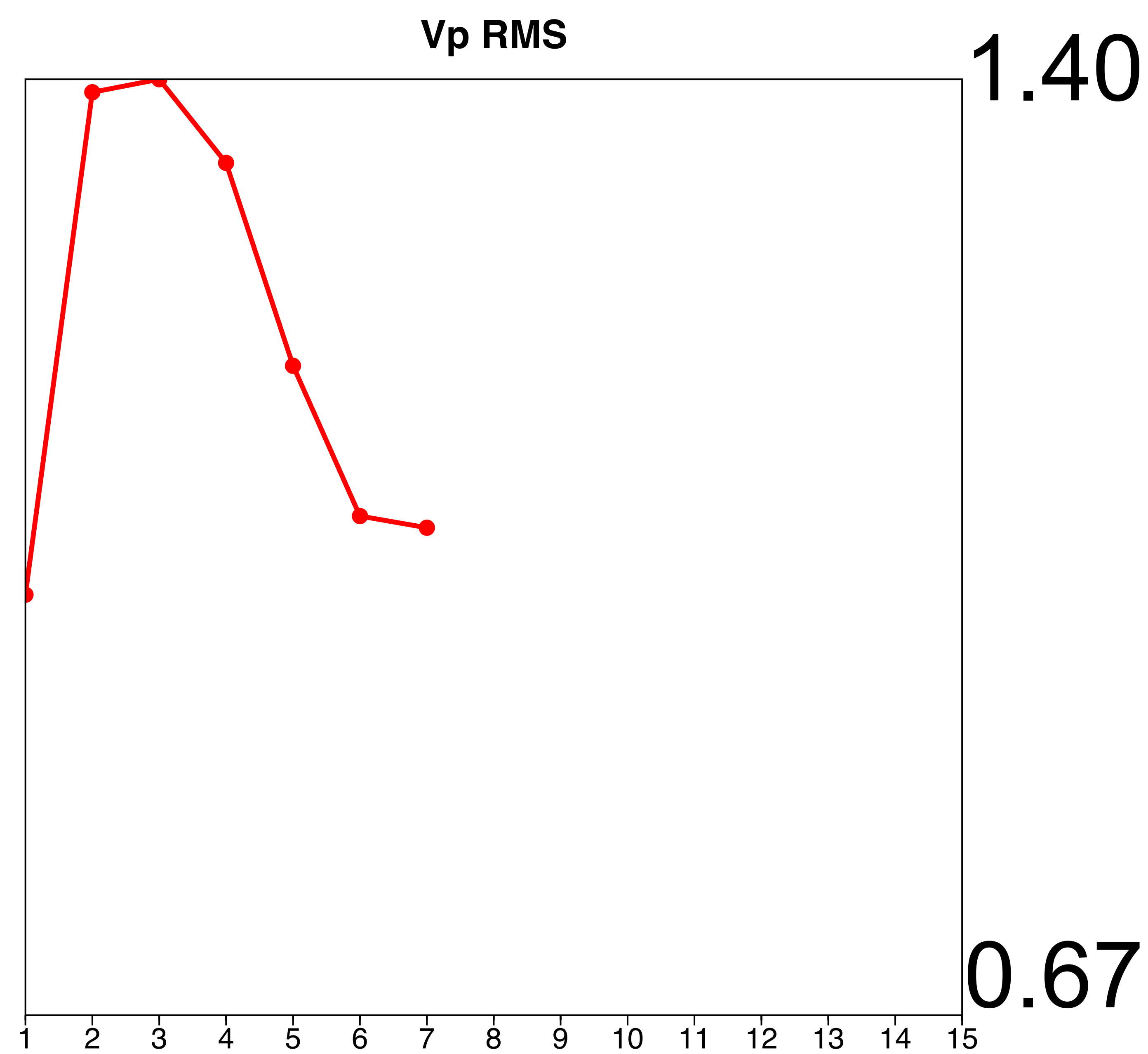
Adjoint-state – w/ TV-norm & hinge-loss projections



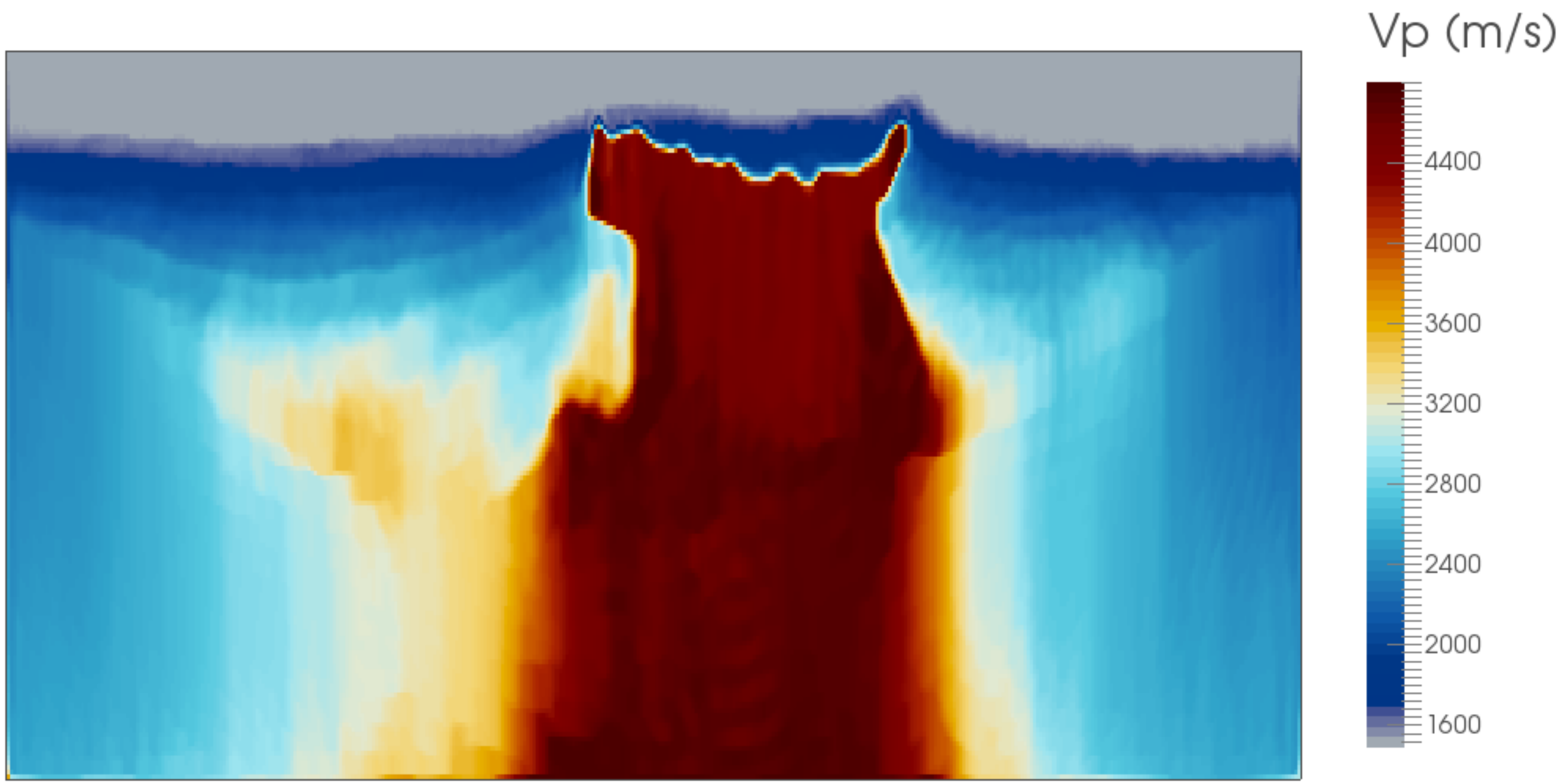


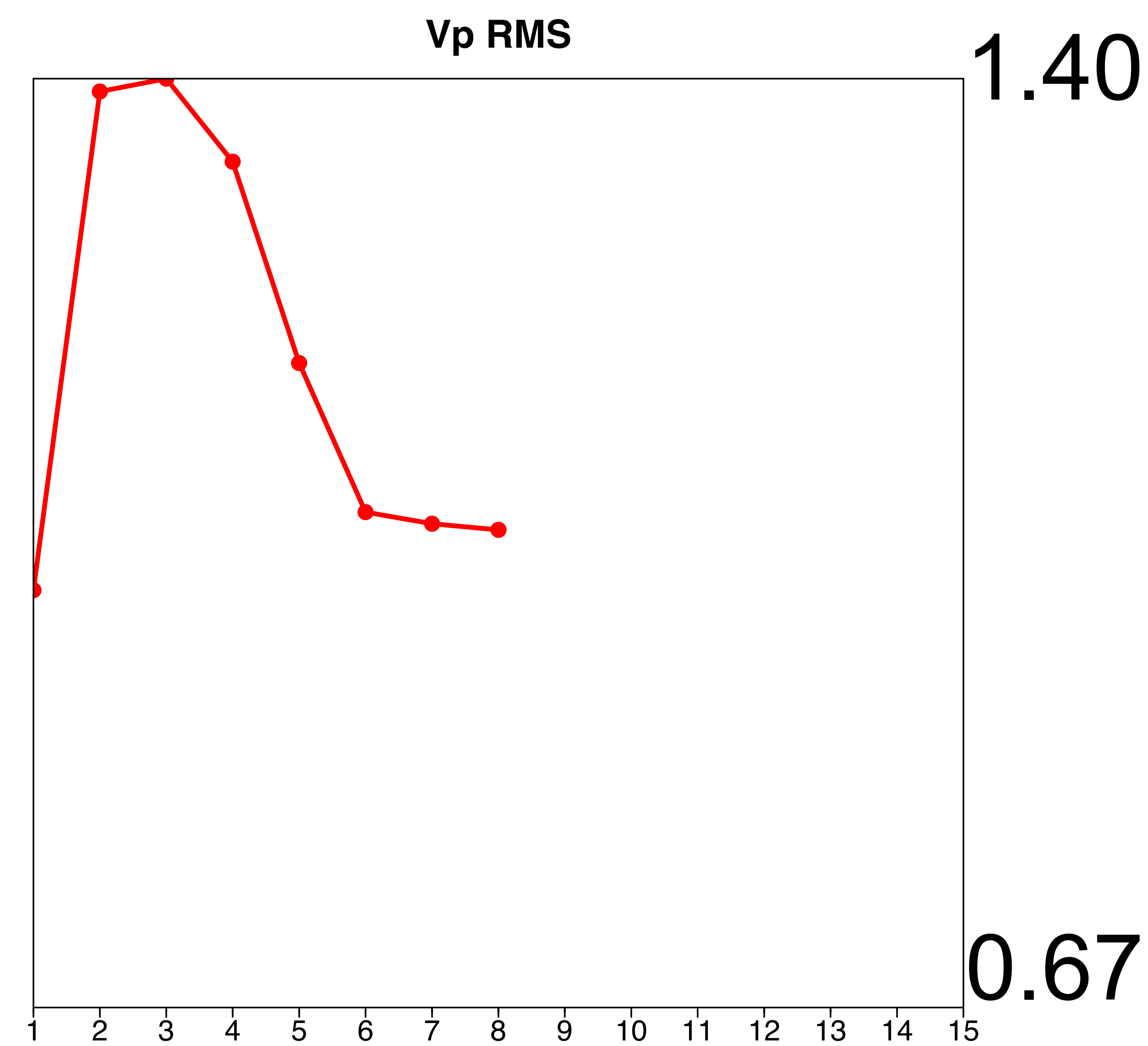
Adjoint-state – w/ TV-norm & hinge-loss projections



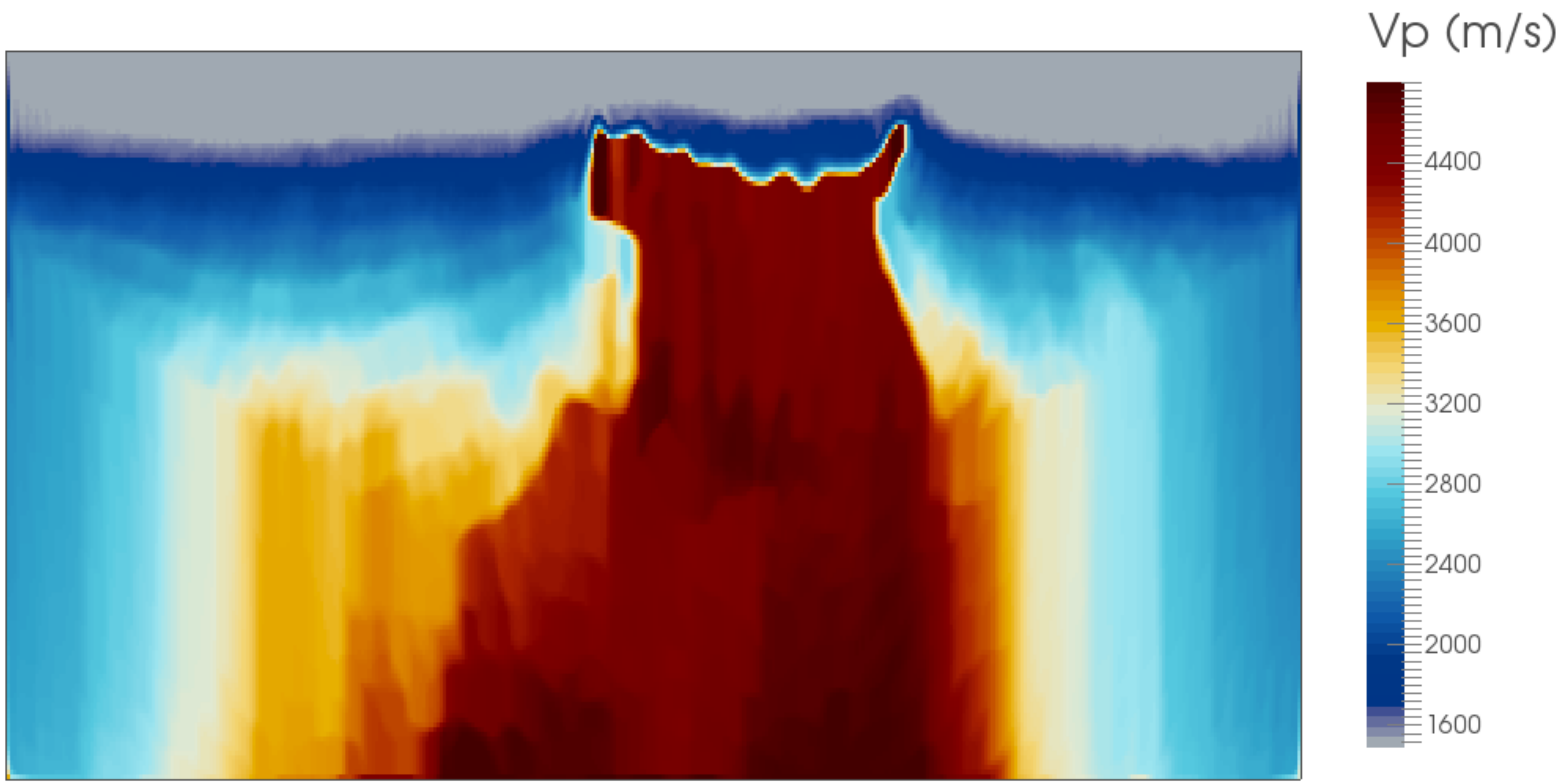


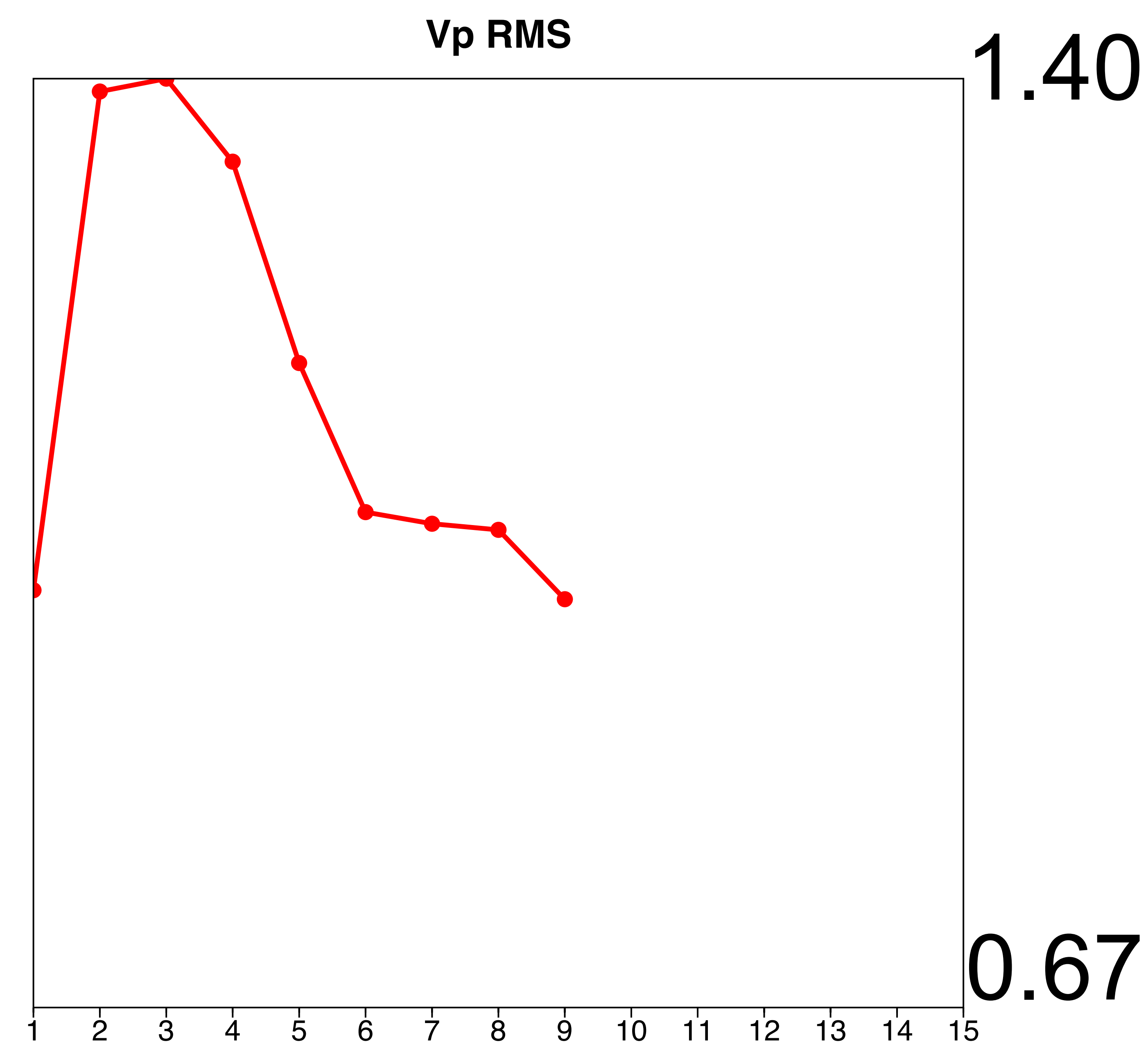
Adjoint-state – w/ TV-norm & hinge-loss projections



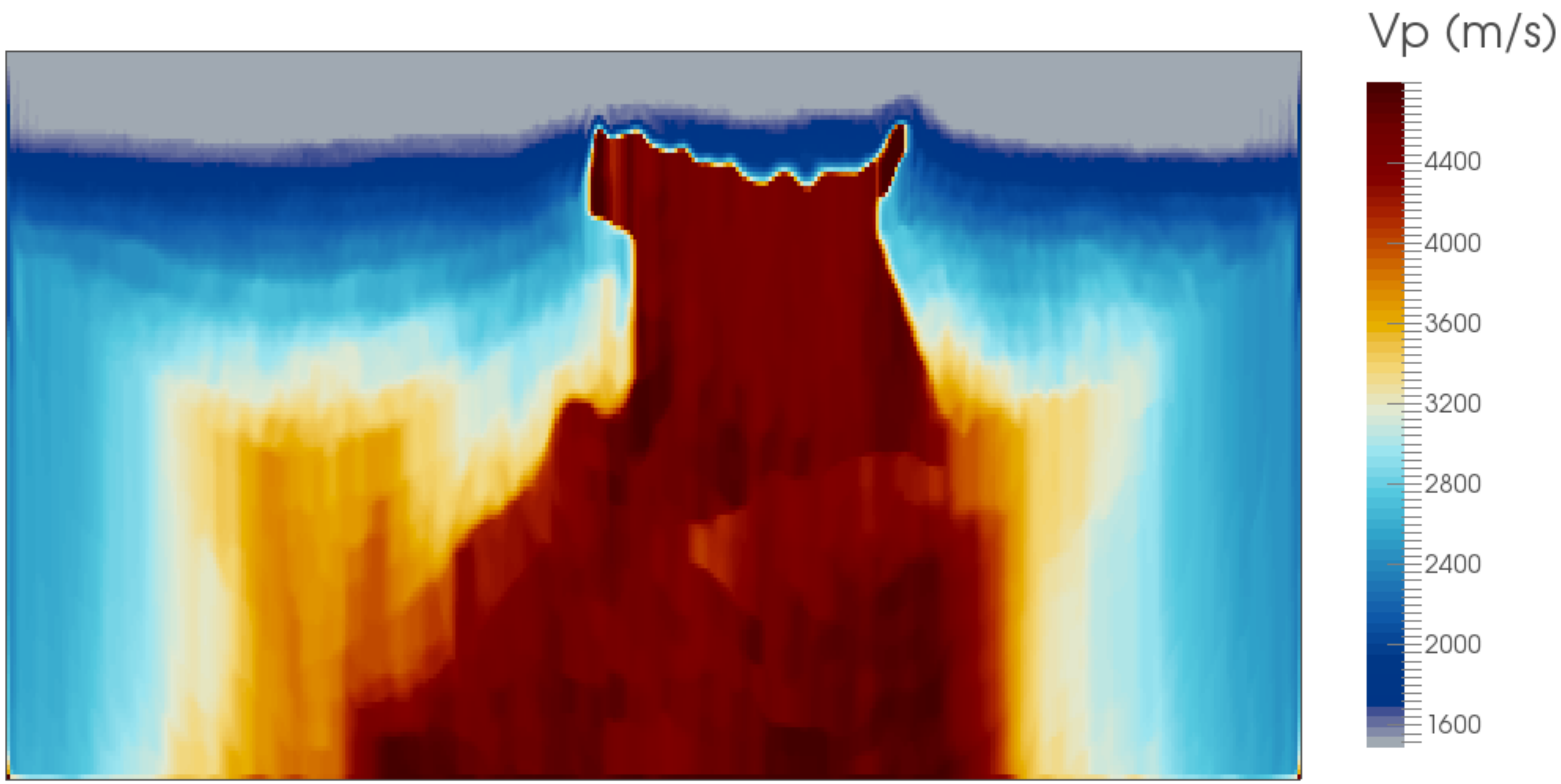


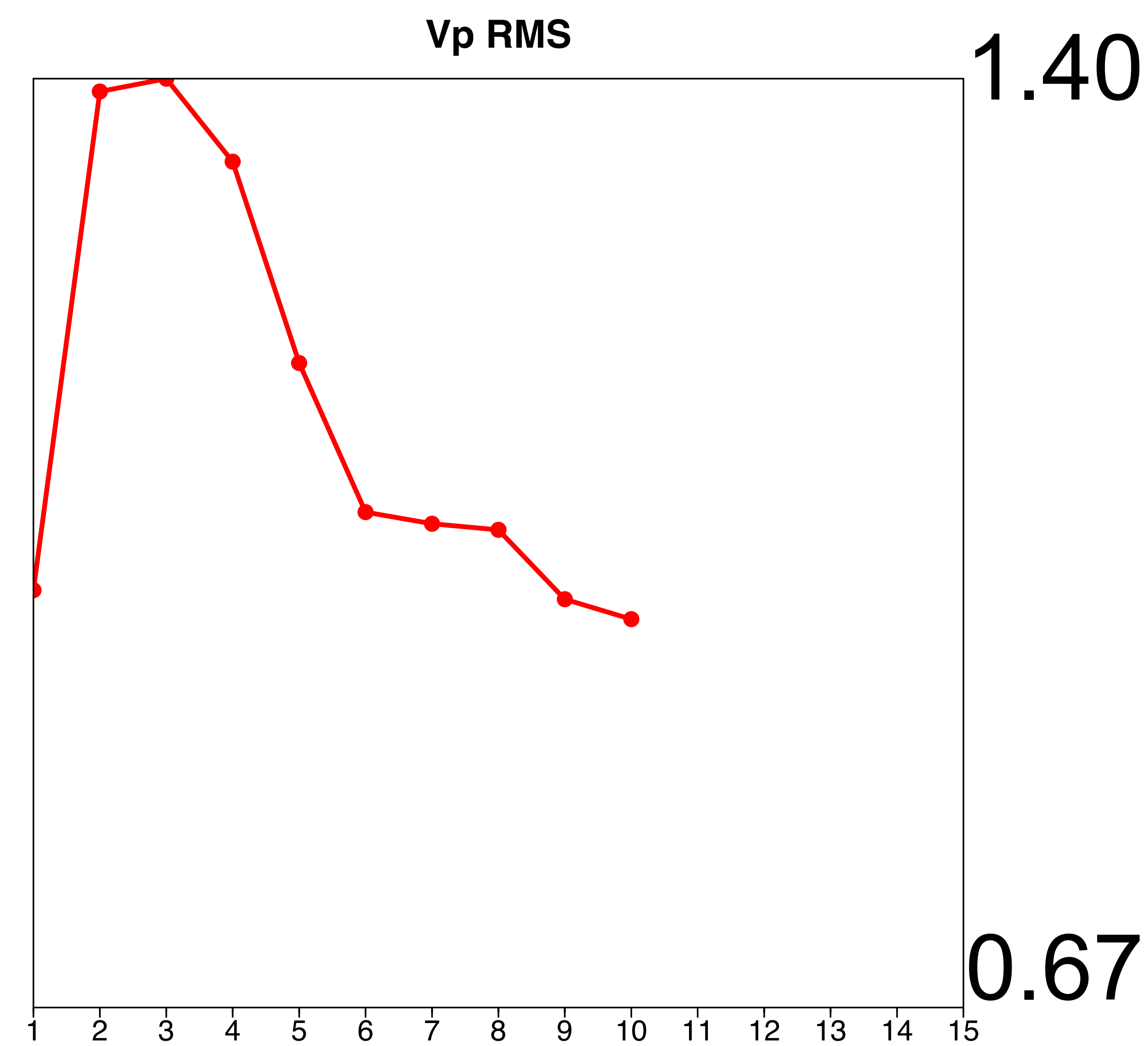
Adjoint-state – w/ TV-norm & hinge-loss projections



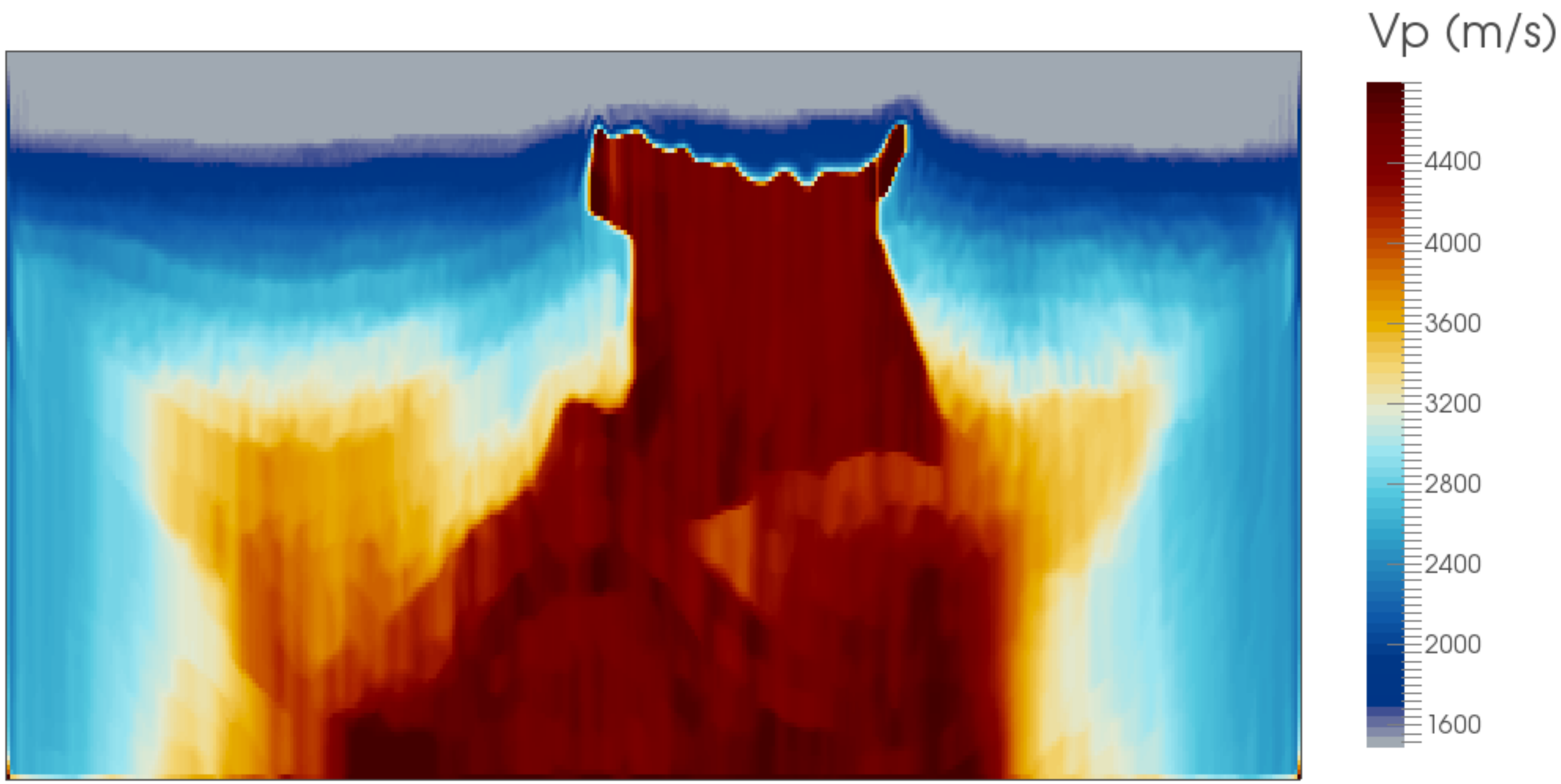


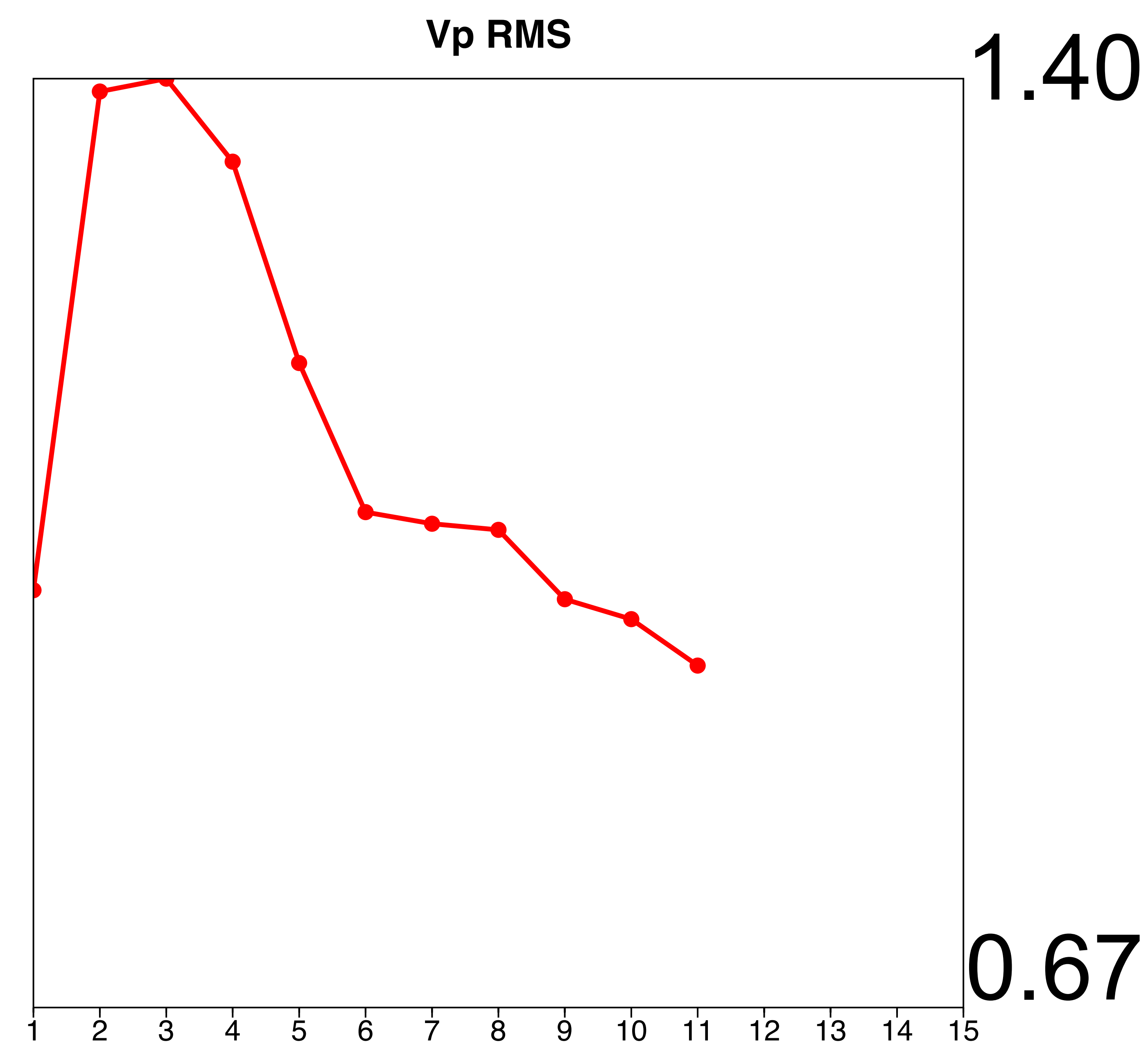
Adjoint-state – w/ TV-norm & hinge-loss projections



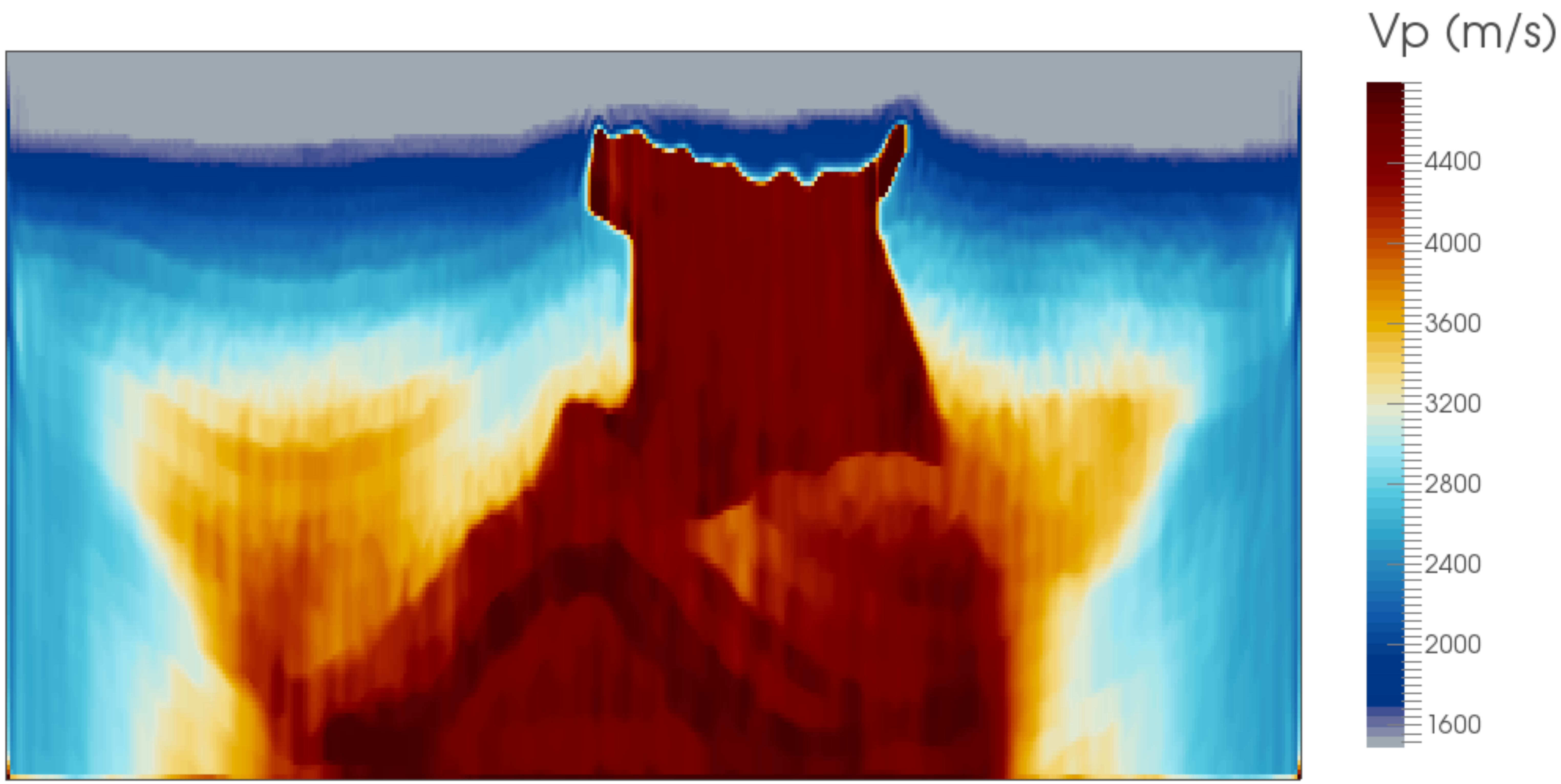


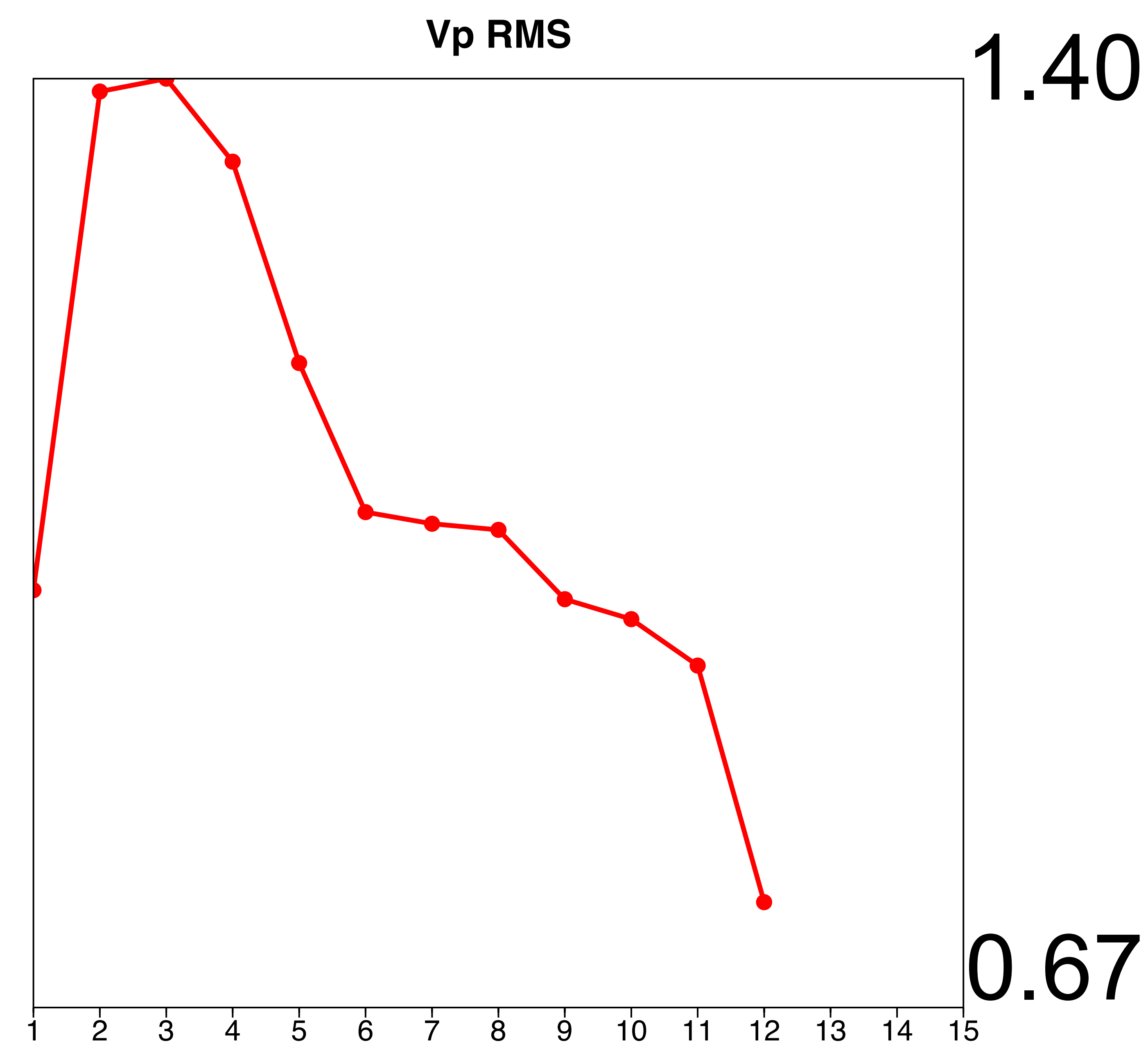
Adjoint-state – w/ TV-norm & hinge-loss projections



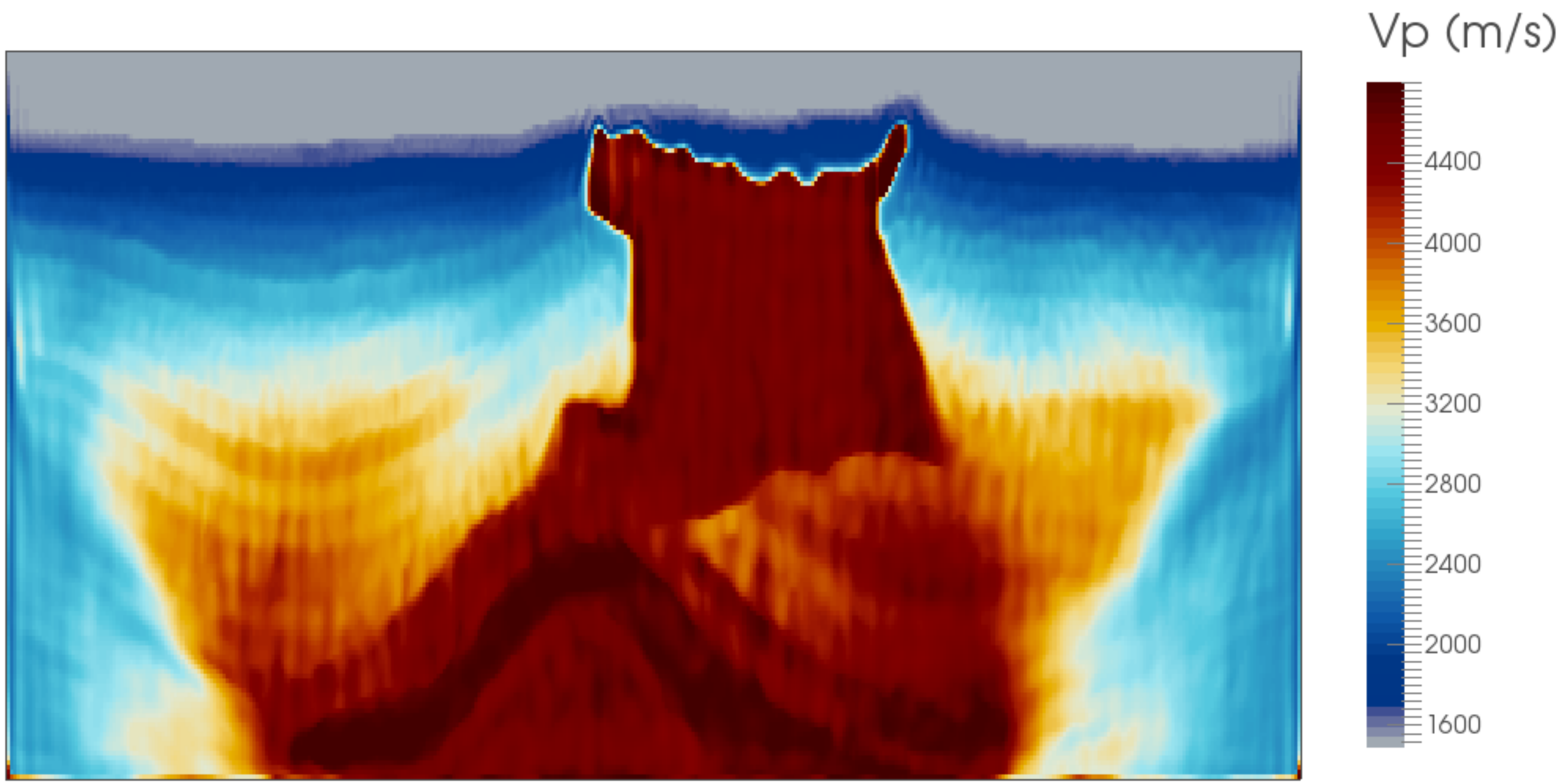


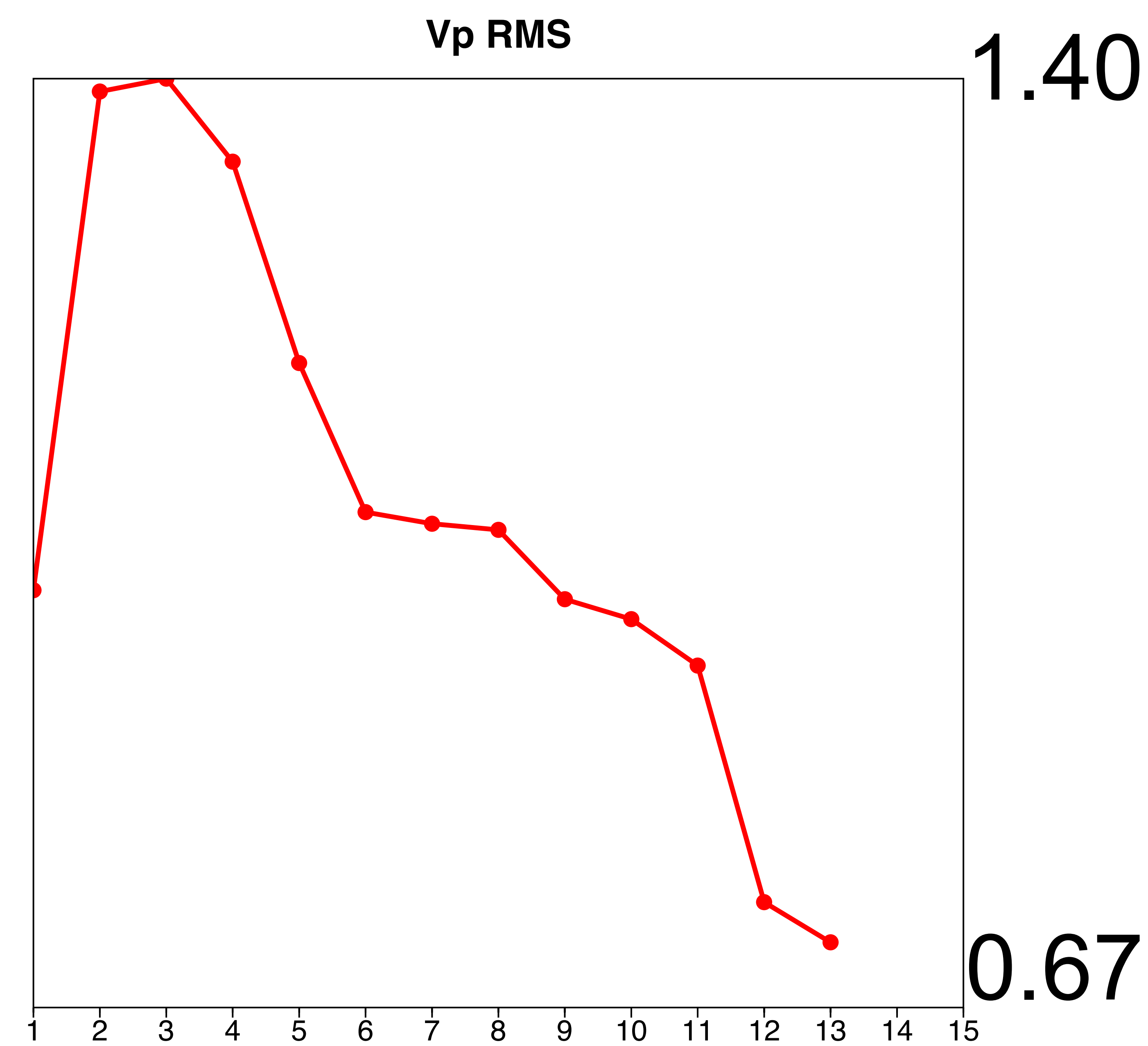
Adjoint-state – w/ TV-norm & hinge-loss projections



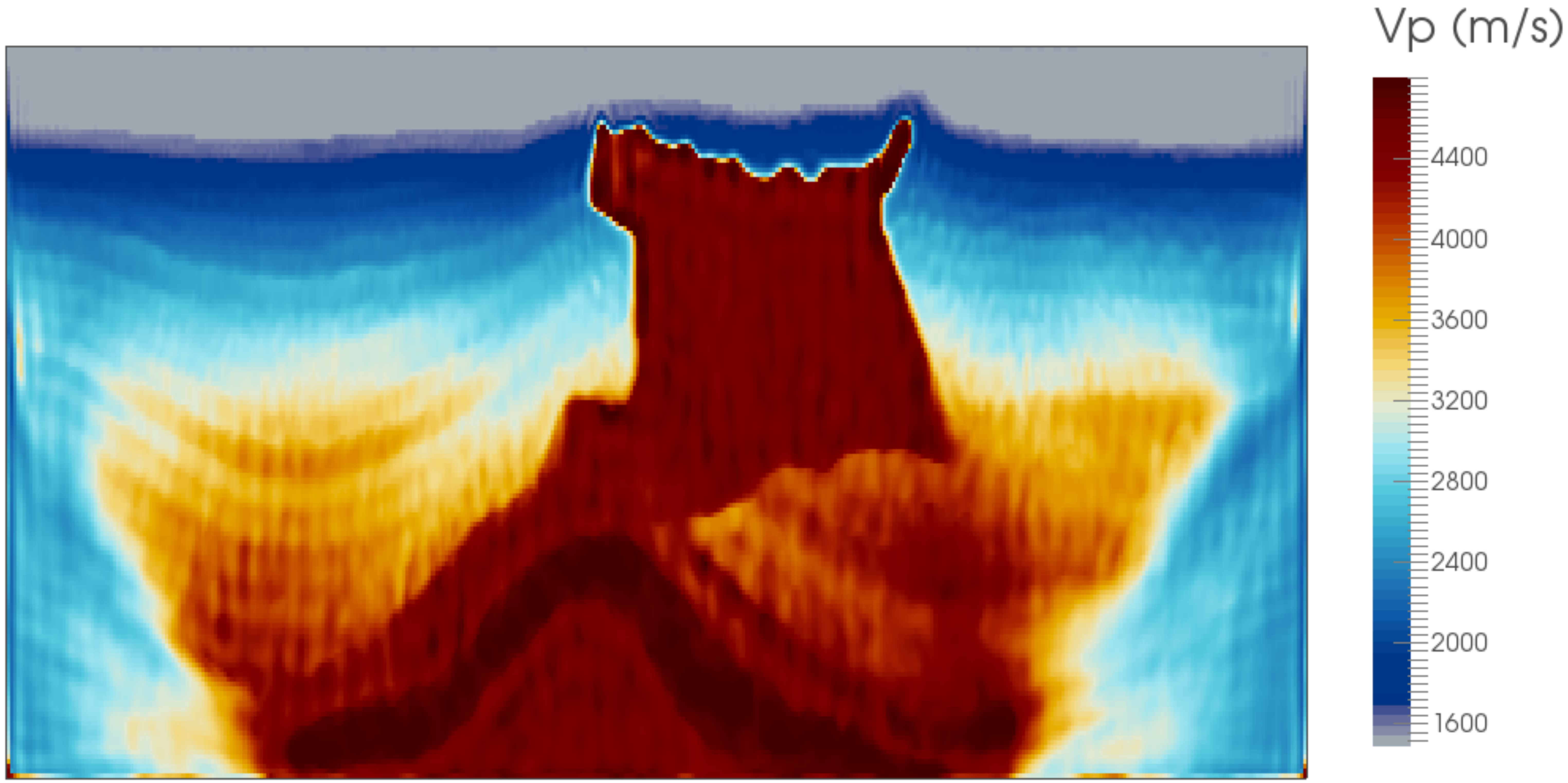


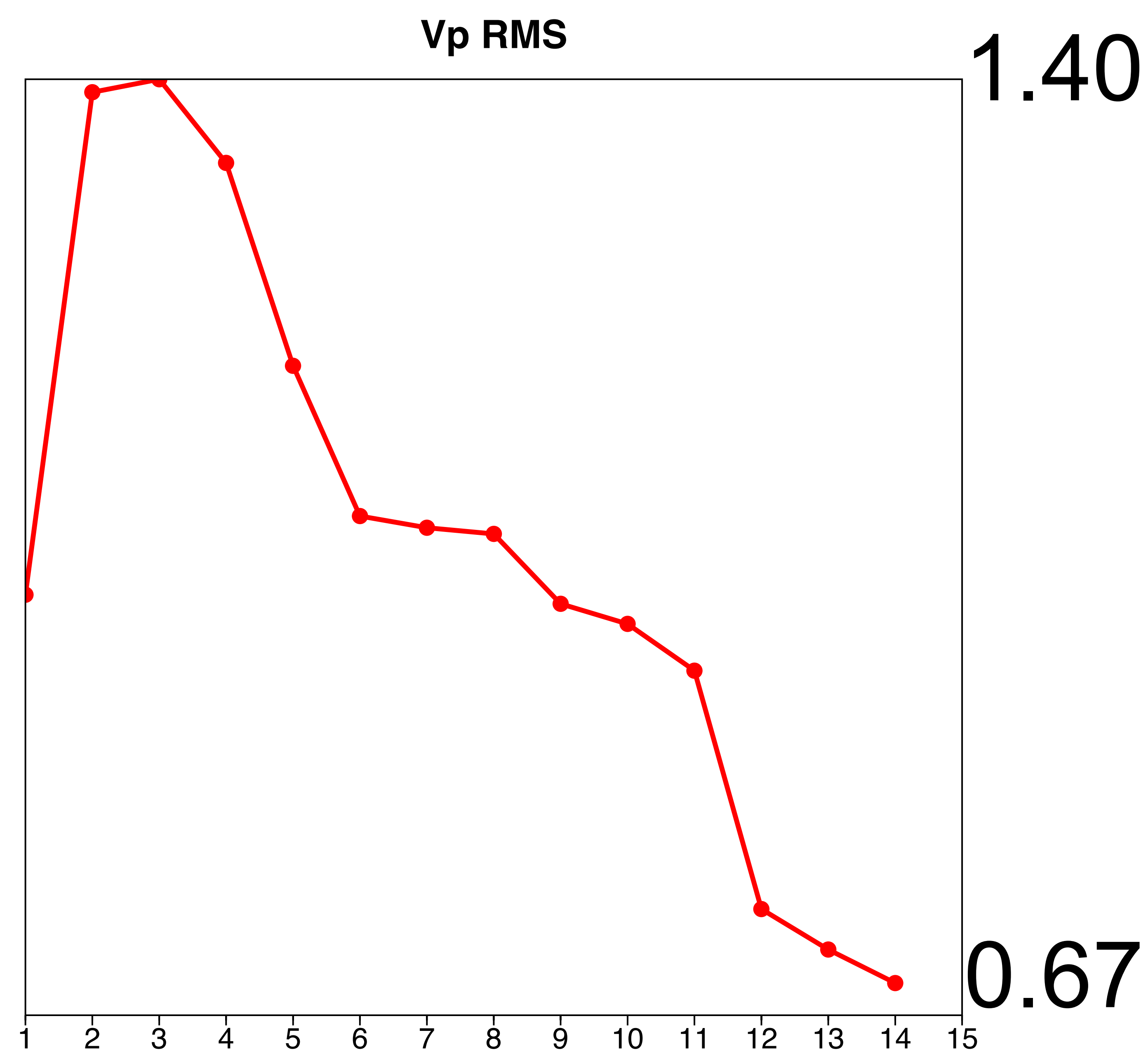
Adjoint-state – w/ TV-norm & hinge-loss projections



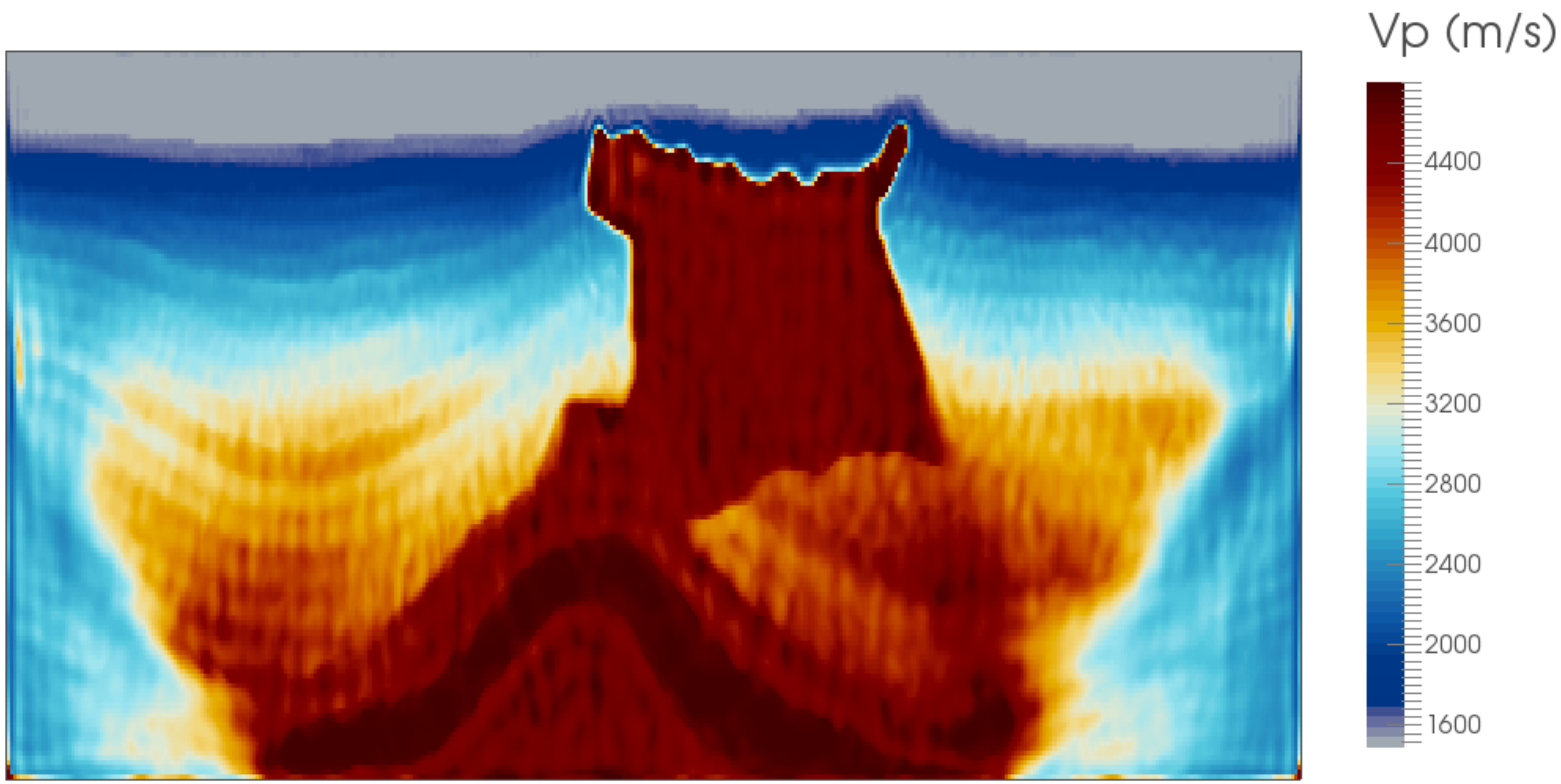


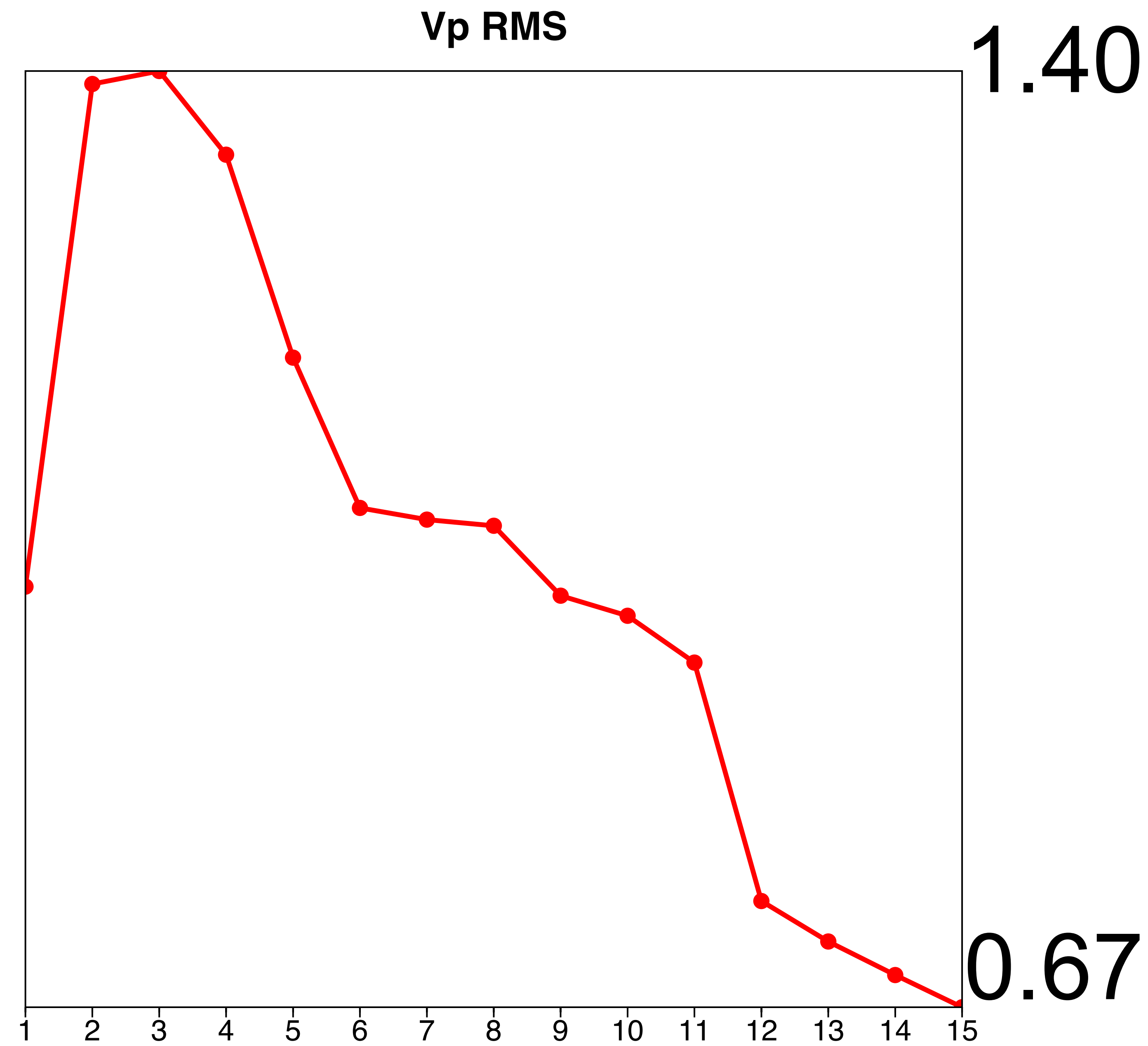
Adjoint-state – w/ TV-norm & hinge-loss projections





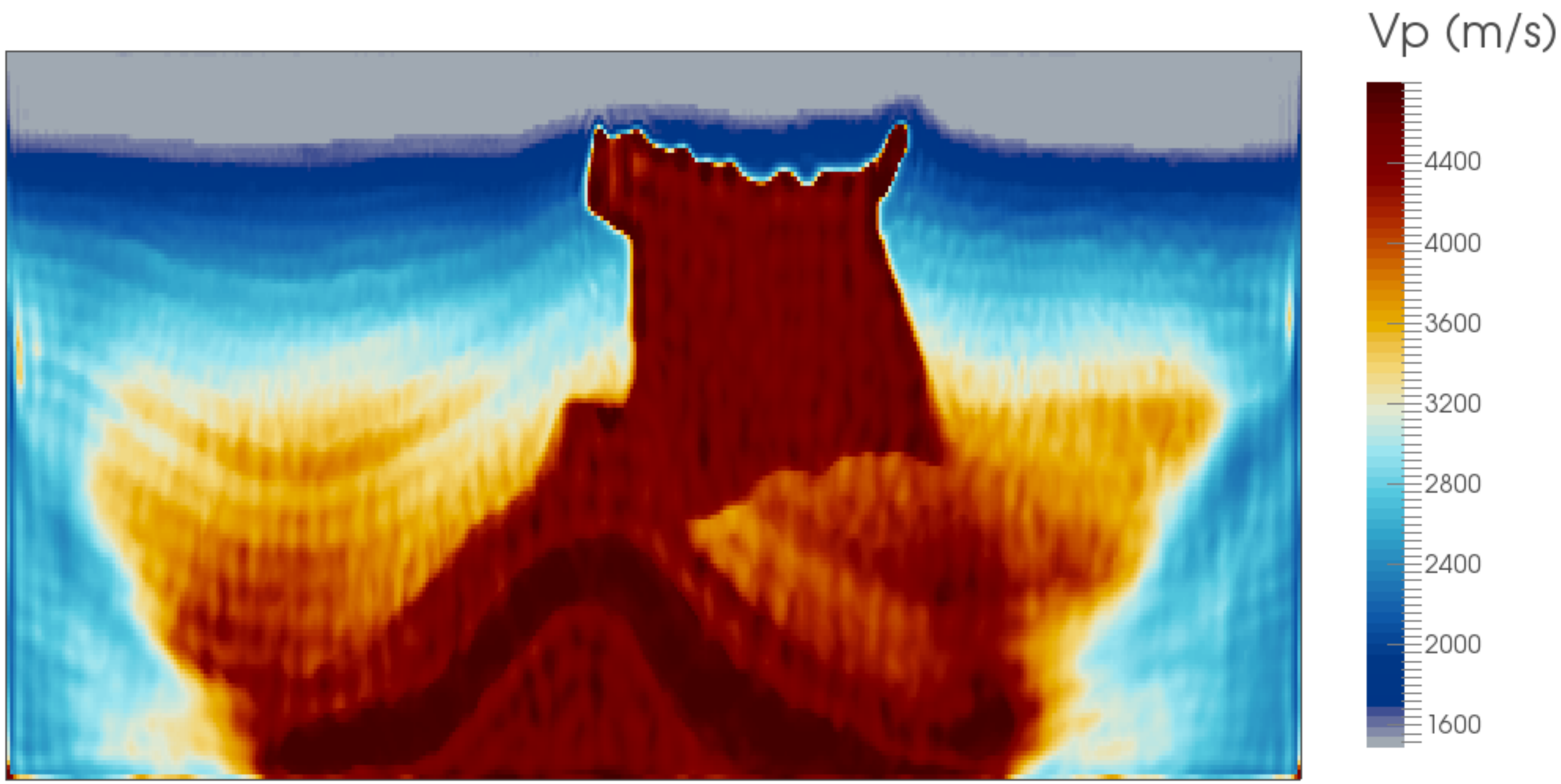
Adjoint-state – w/ TV-norm & hinge-loss projections

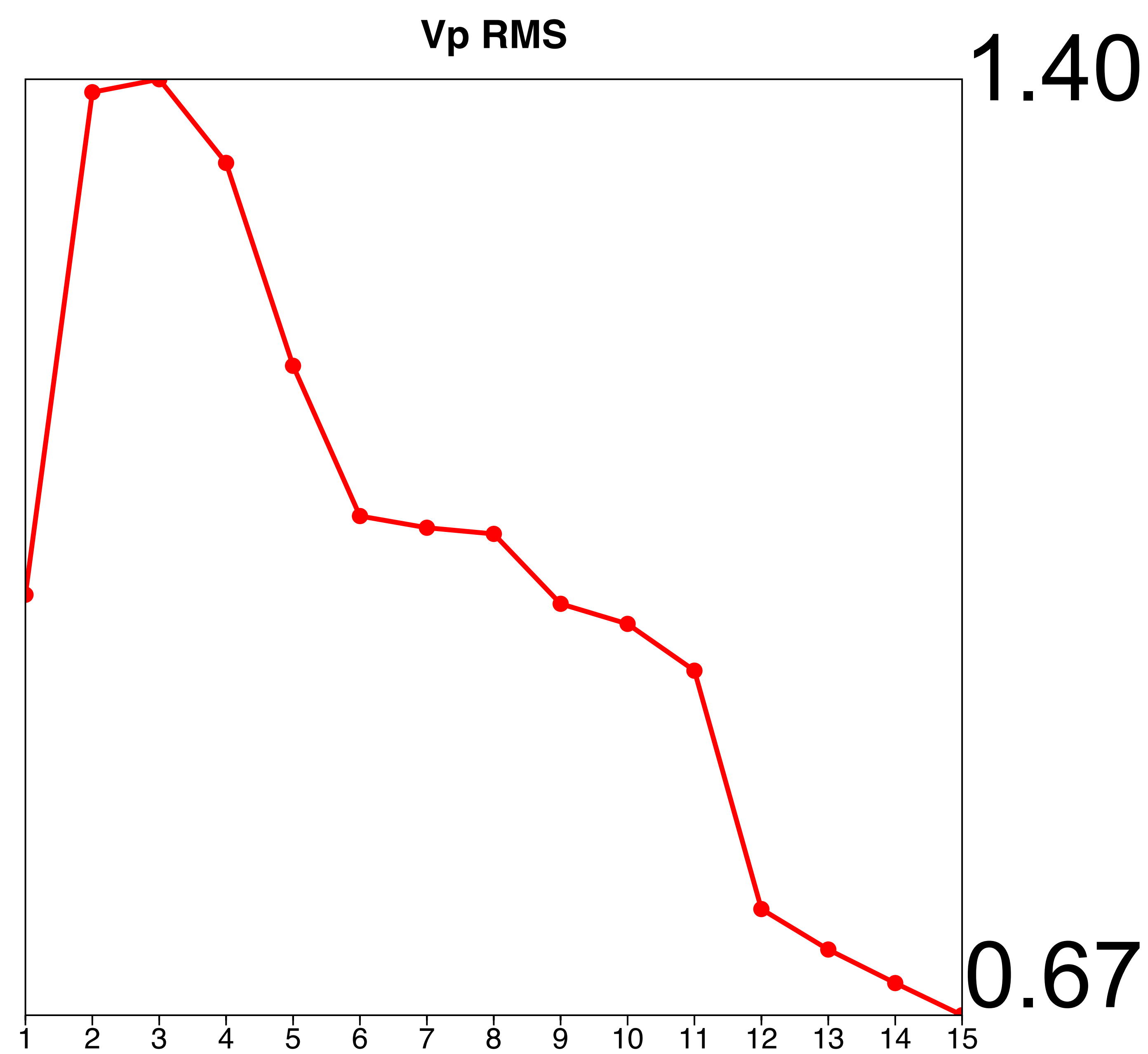




Adjoint-state – w/ TV-norm & hinge-loss projections

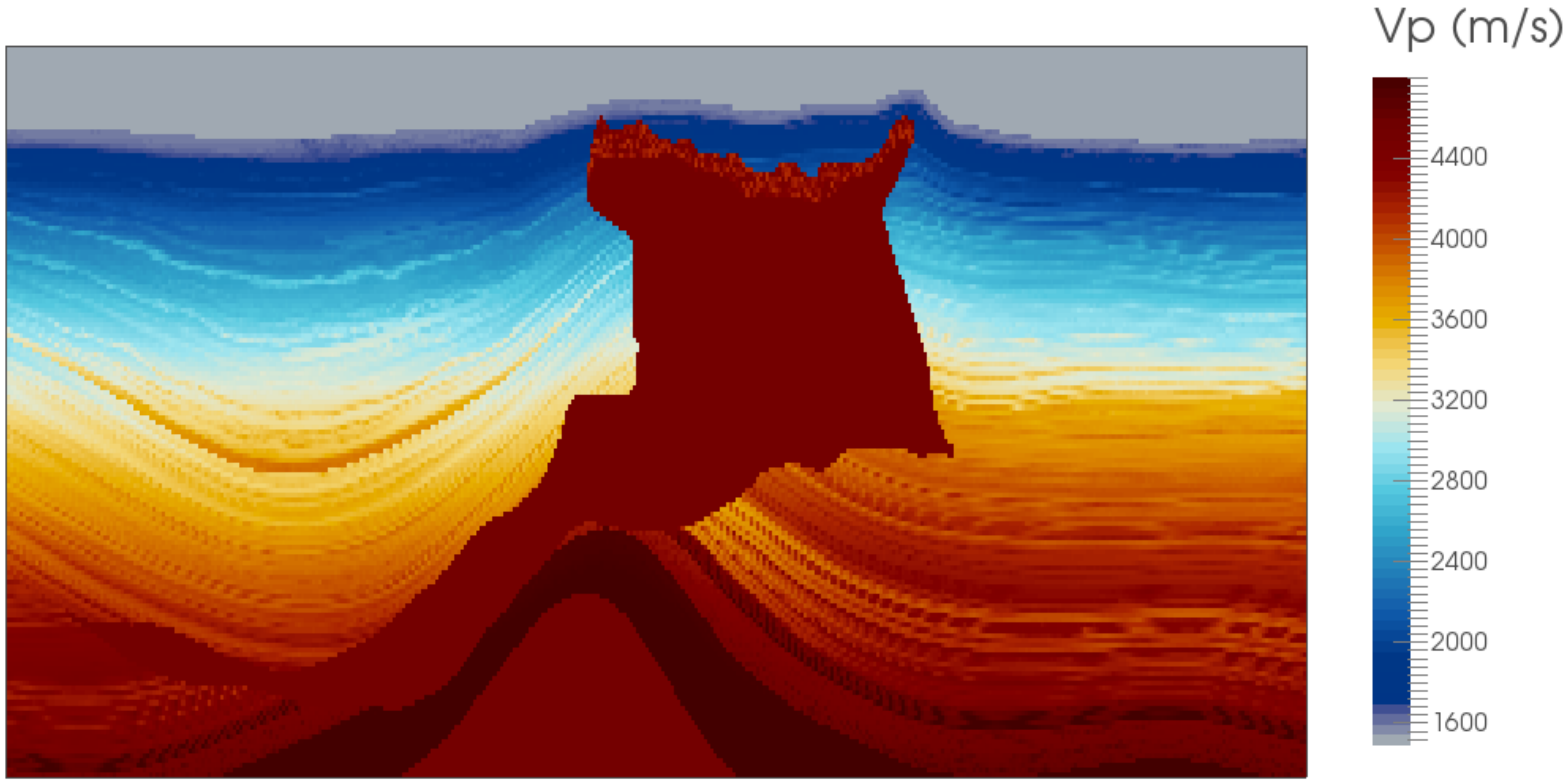
Final model





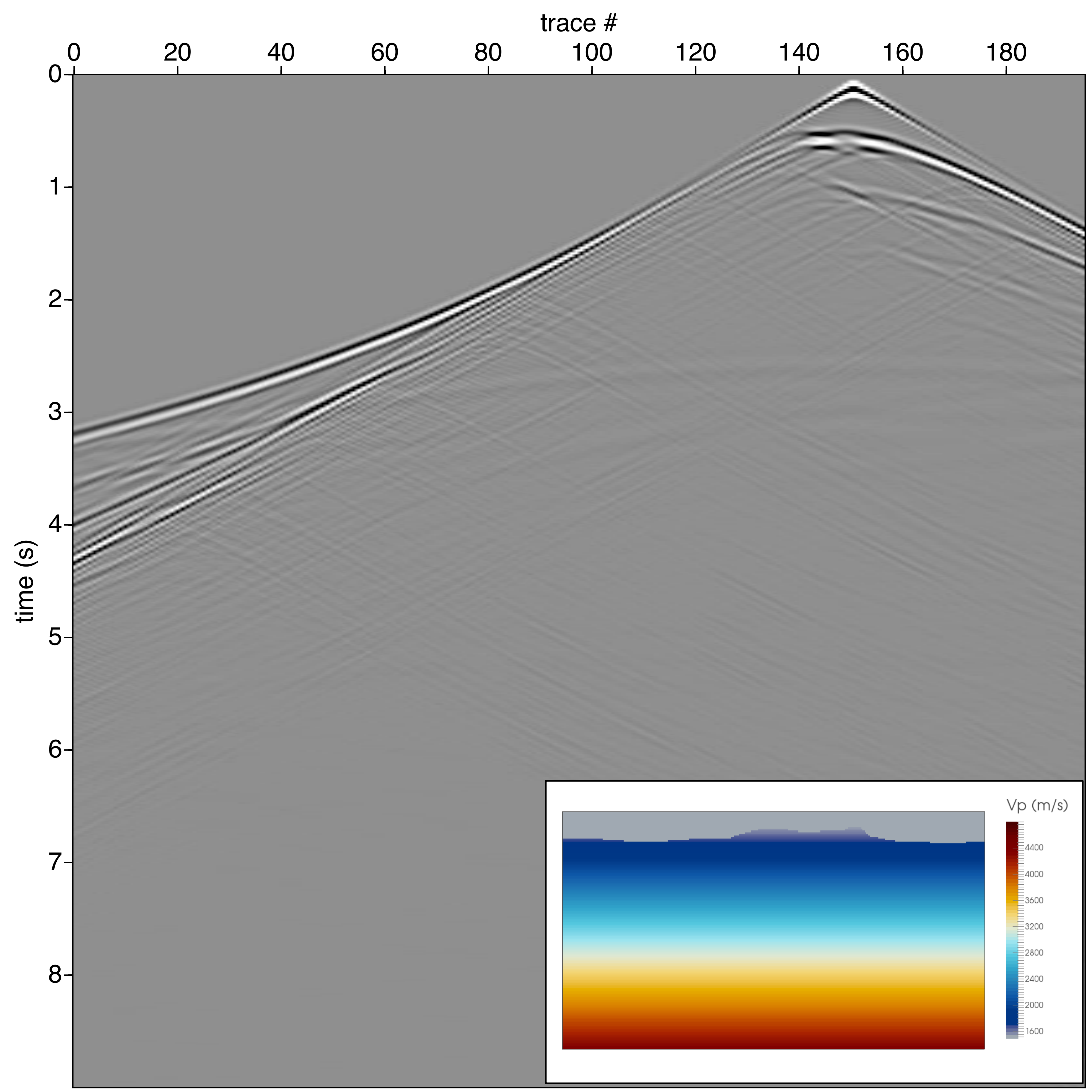
Adjoint-state – w/ TV-norm & hinge-loss projections

True model



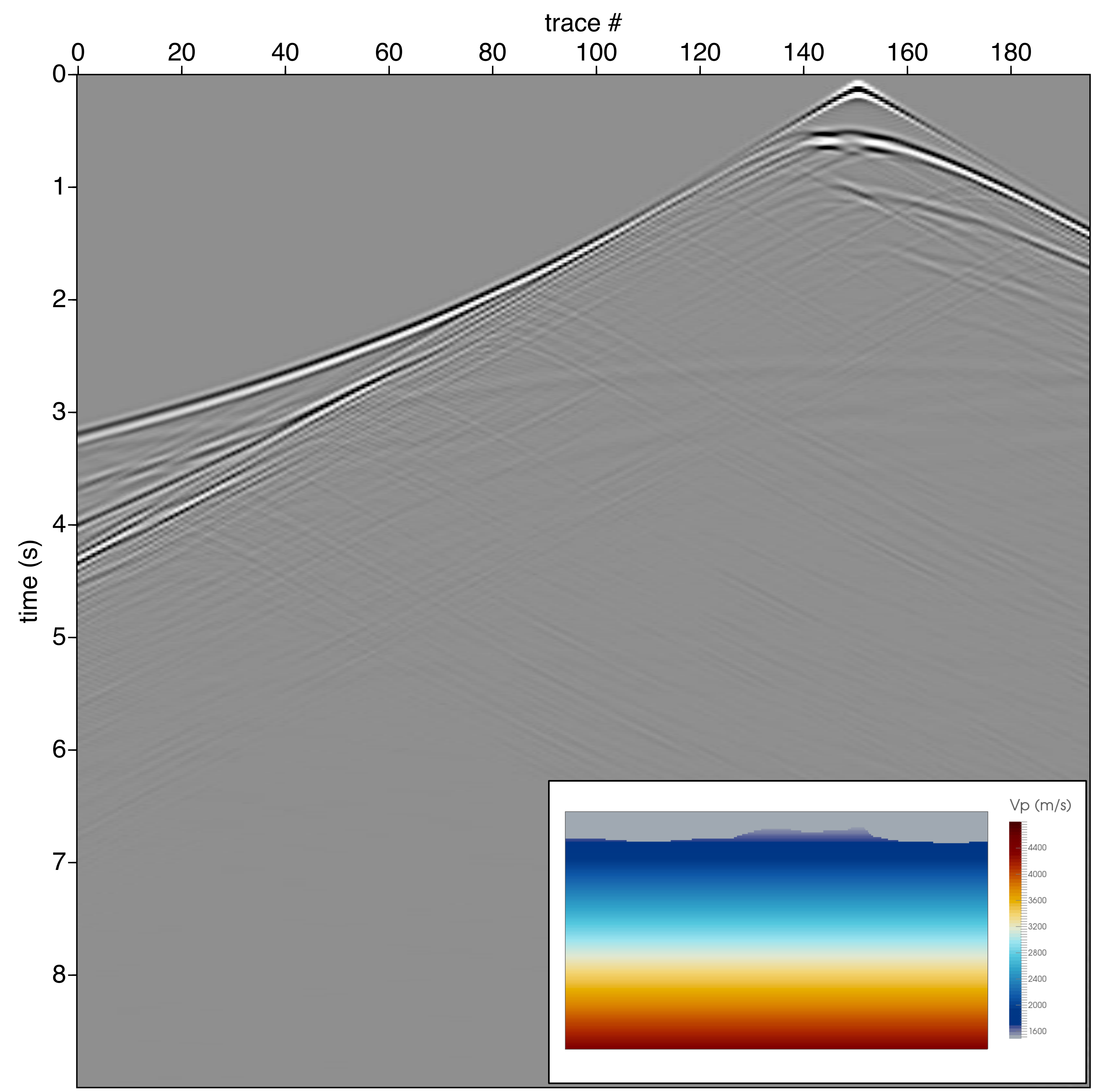
Adjoint-state – data comparisons

starting model



Starting model data

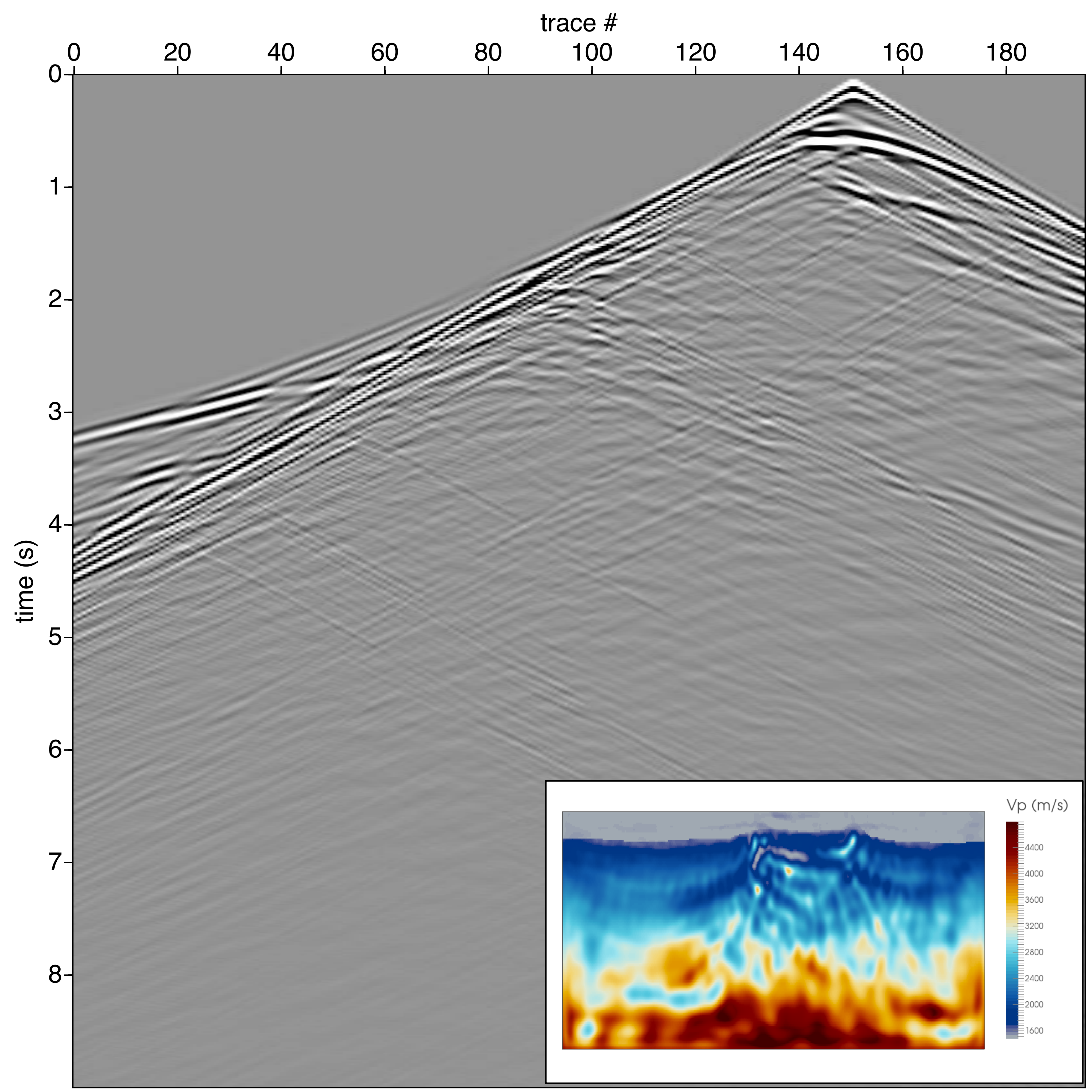
starting model



Starting model data

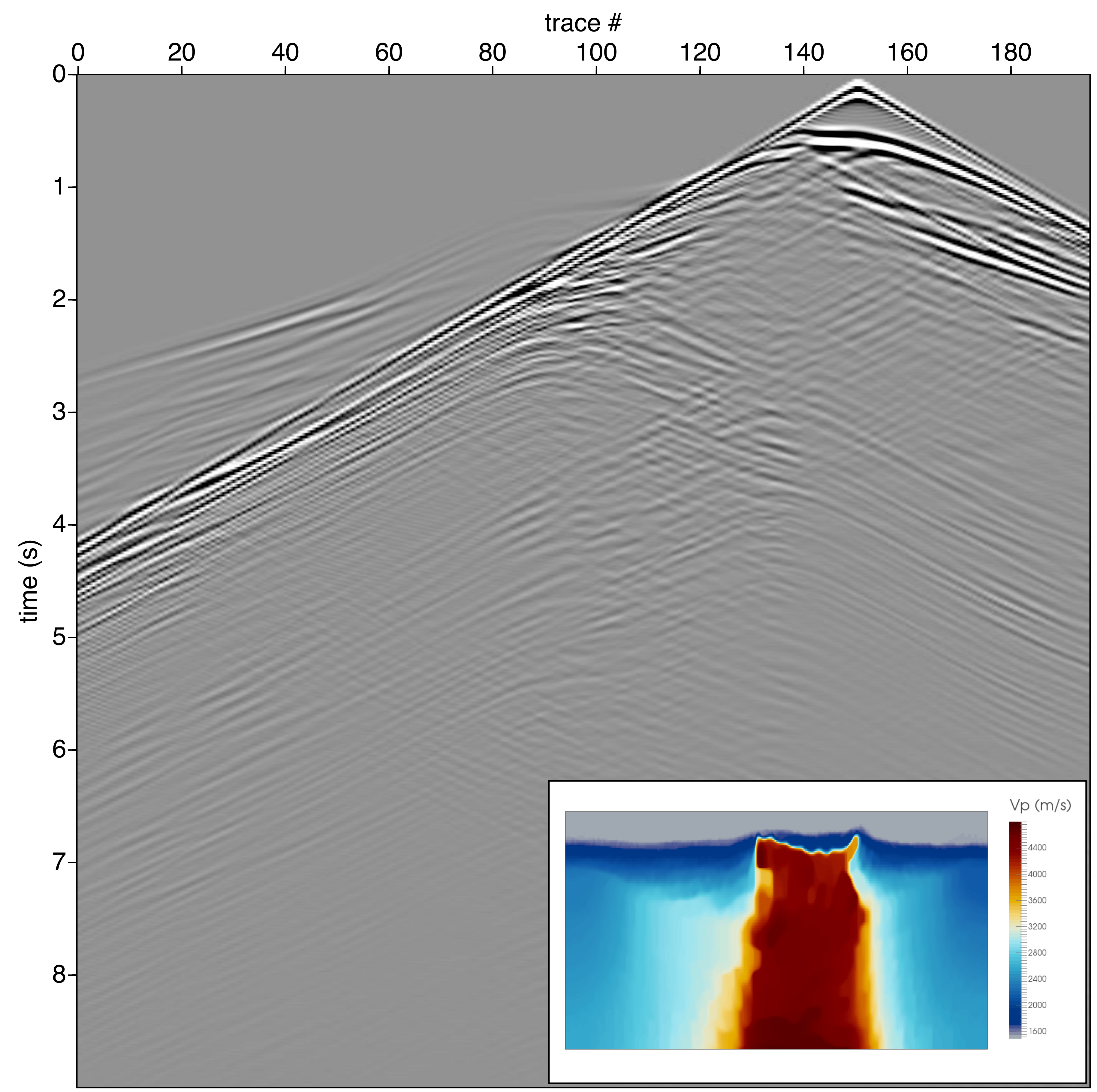
Adjoint-state – data comparisons

w/o TV & Hinge



FWI intermediate result data

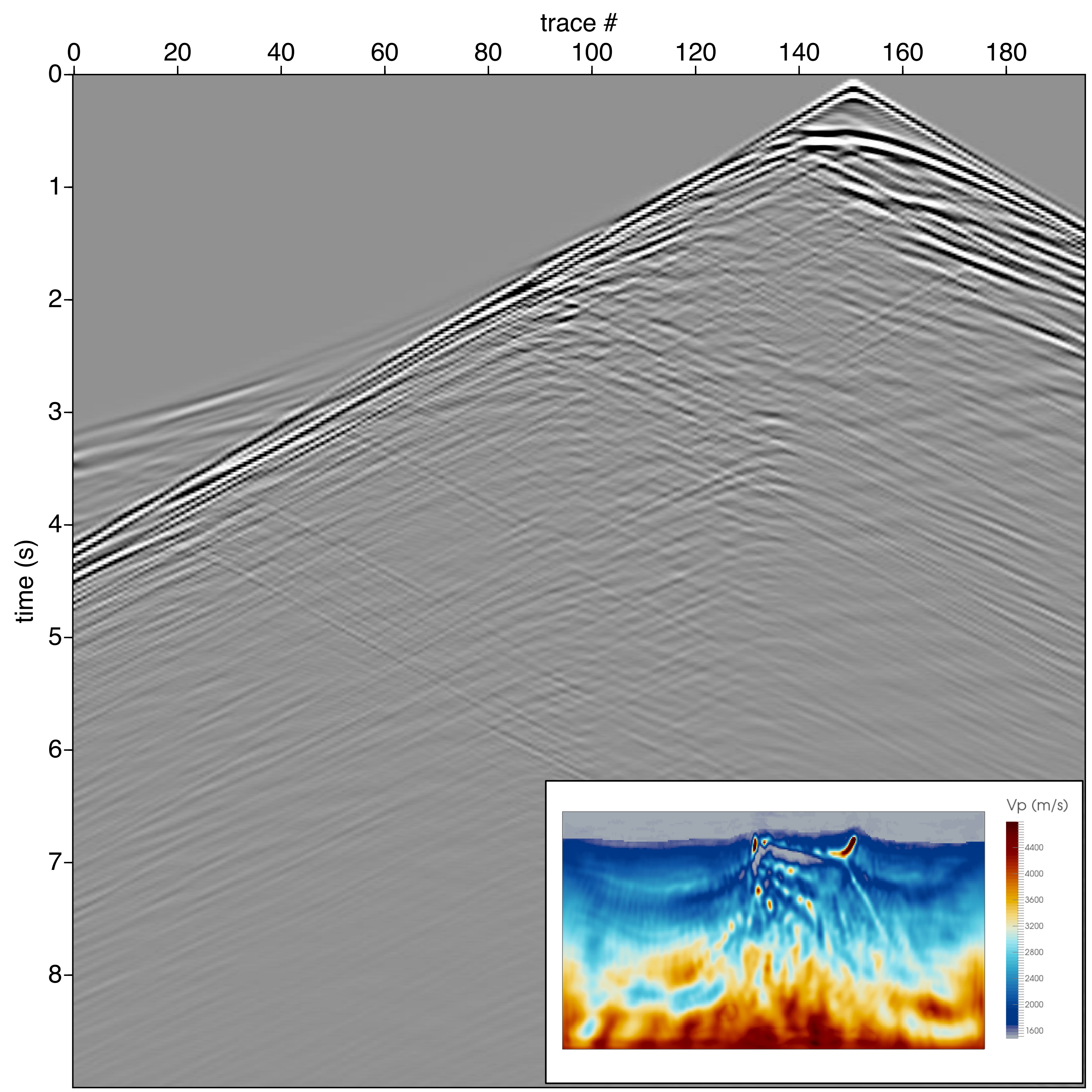
w/ TV & Hinge



FWI+TV middle result data

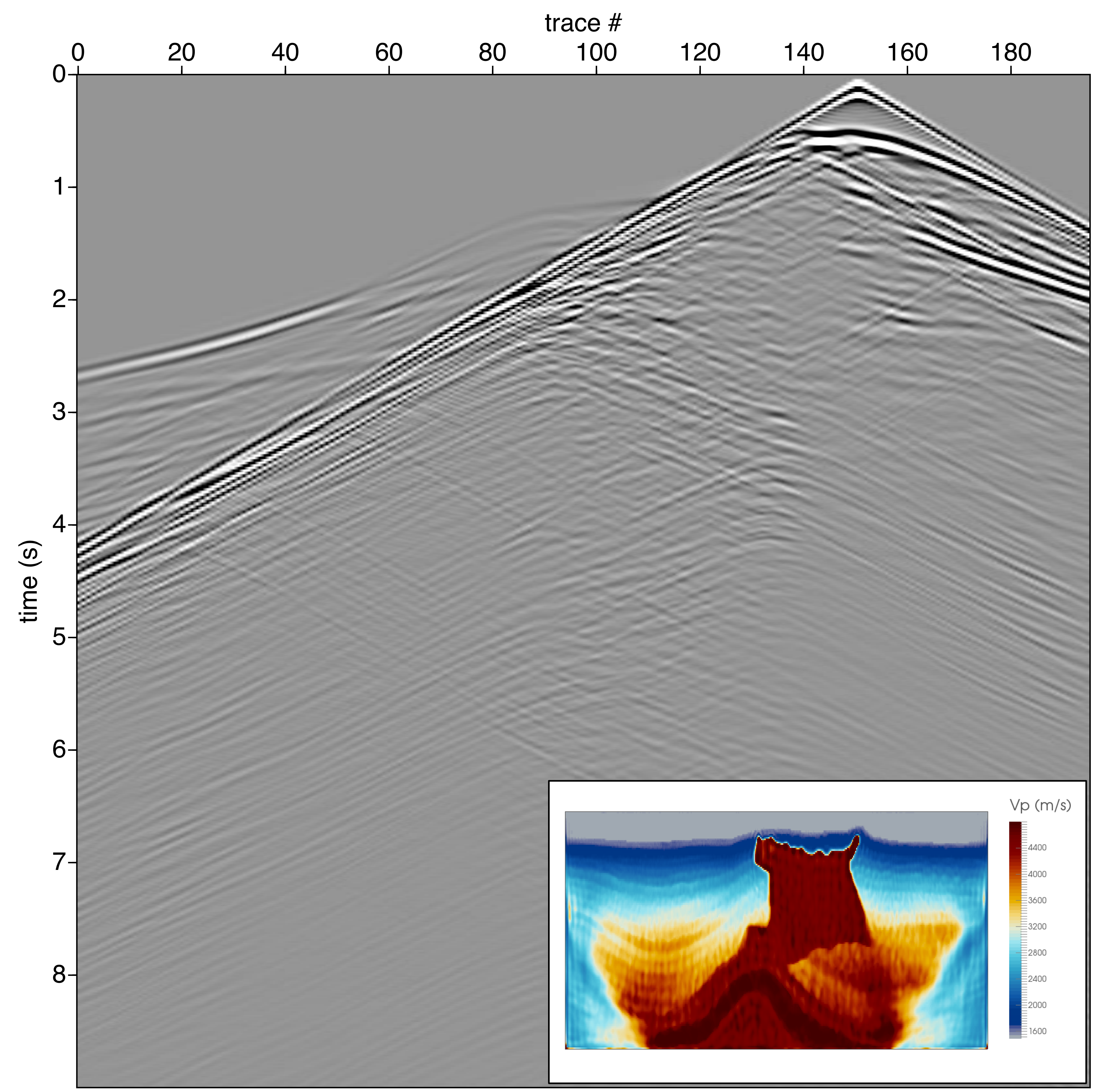
Adjoint-state – data comparisons

w/o TV & Hinge



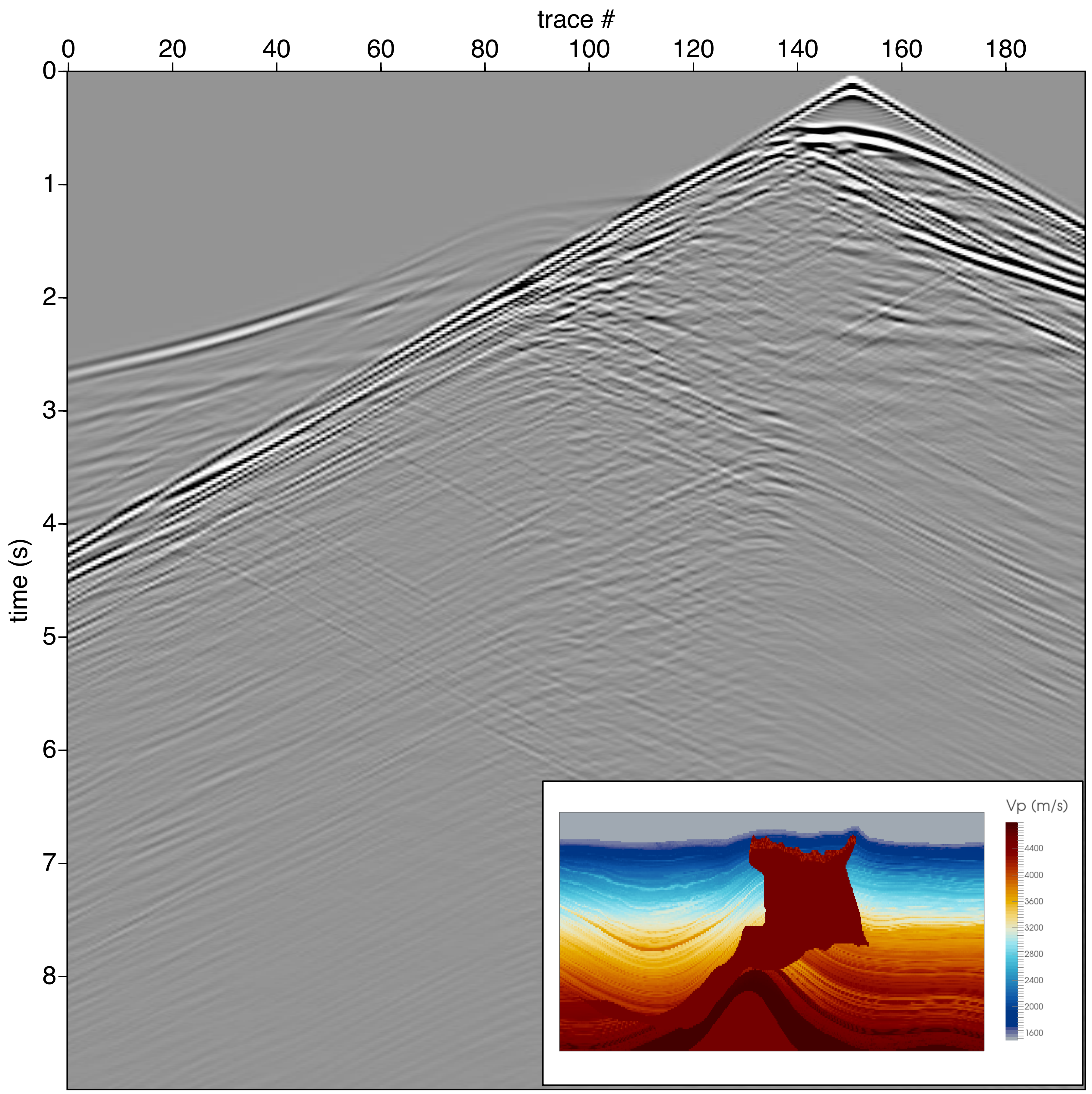
FWI final result data

w/ TV & Hinge

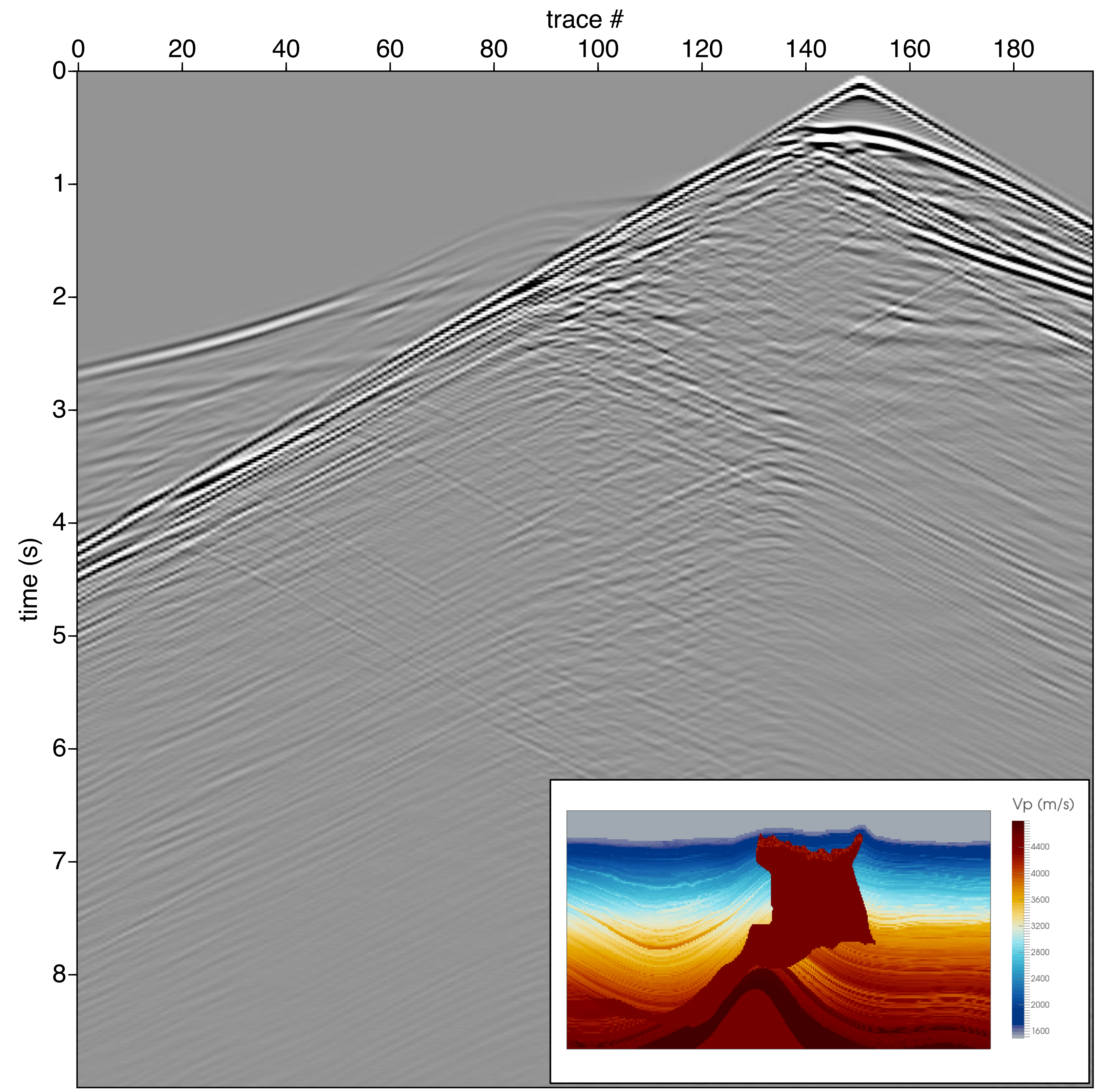


FWI+TV final result data

Adjoint-state – data comparisons



Field data



Field data

Conclusions

Constraints are enabled by WRI

- projected “second order” scheme w/ diagonal Gauss-Newton Hessian

Successful combination of

- multiple warm-started frequency cycles & relaxing constraints

Optimization framework w/ hinge-loss function on raw input data leads to

- recovery of salt, including top & bottom & sub-salt details (lows/highs)
- automatic salt flooding w/o picking & tomography
- cheap, reproducible & transparent workflows w/o handholding

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.