Data-space affordable Gradient Sampling for Seismic Inversion

Mathias Louboutin

SINBAD consortium meeting
Wednesday October 4th

SLIM
University of British Columbia

Motivations

Sensitivity to cycle skipping

Memory cost due to storage of time history of the wavefield

Computationally expensive

- checkpointing
- random boundaries
- wavelet compression
- ......

Rajiv Kumar, Curt Da Silva, Oscar Lopez, Aleksandr Y. Aravkin, Hassan Mansour, Haneet Wason, Ernie Esser, and Felix J. Herrmann, "Rank minimization based seismic data processing and inversion"

Bas Peters and Felix J. Herrmann,"A quadratic-penalty full-space method for waveform inversion"

SLIM

Motivations

Global methods have shown good results

- low-rank extension
- full-space

New way to extend the search space for time-domain FWI.

Tristan van Leeuwen, "A correlation-based misfit criterion for wave-equation traveltime tomography", in *ICIAM*, 2011

Guanghui Huang, William Symes, and Rami Nammour Matched source waveform inversion: Space-time extension SEG Technical Program Expanded Abstracts 2016. September 2016, 1426-1431

Tristan van Leeuwen, Rajiv Kumar, and Felix J. Herrmann, "Affordable full subsurface image volume–-an application to WEMVA", in *EAGE Annual Conference Proceedings*, 2015

SLIM

Related work

Stochastic gradient

Subsurface image volumes

Correlation-based misfit

Extended sources

SLIM

Gradient Sampling Algorithm

Designed for Non-Smooth Non-Convex problems:

- global method
- use information from many “nearby” models
- simple & computationally cheap implementation

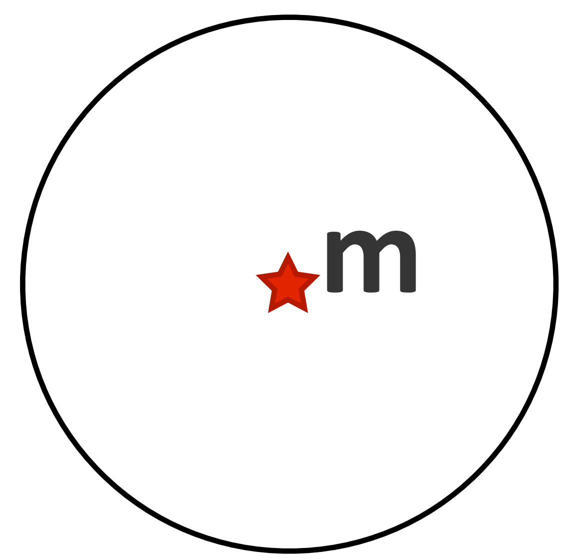
Gradient sampling

Current model **m**

m is the square slowness

Gradient sampling

1- Define a ball around current point **m**



Gradient sampling

1- Define a ball around current point **m**

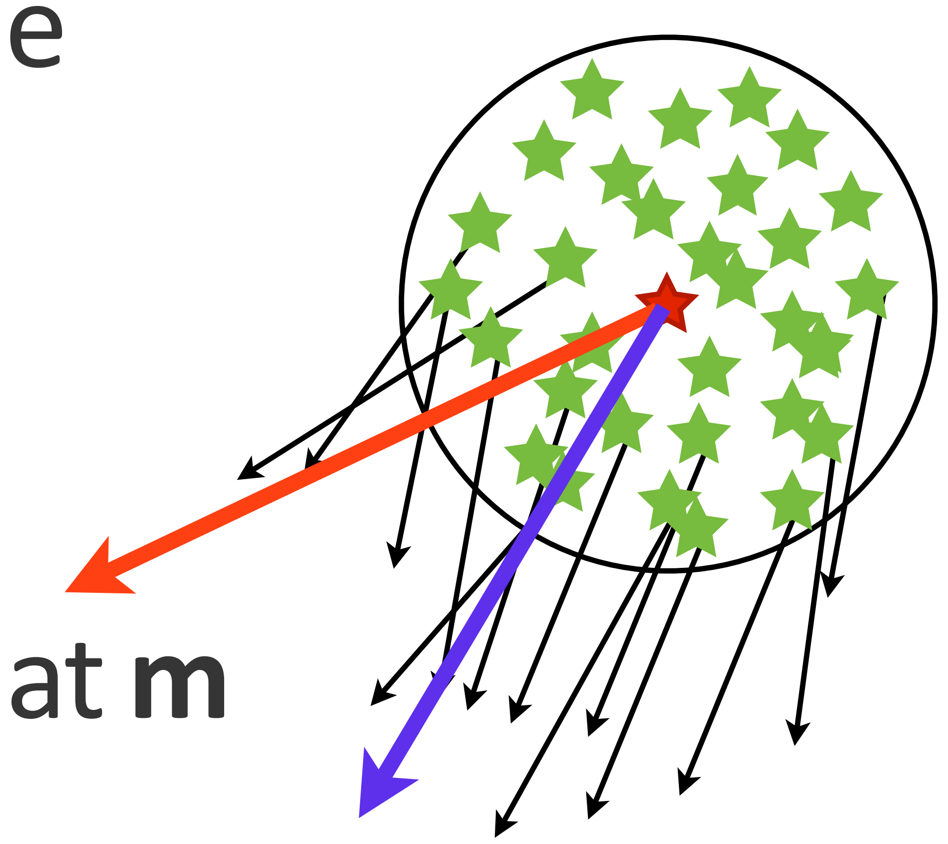
2- Take *p* sample inside the ball

Gradient sampling

1- Define a ball around current point **m**

2- Take *p* sample inside the ball

3 - Compute direction for each sample

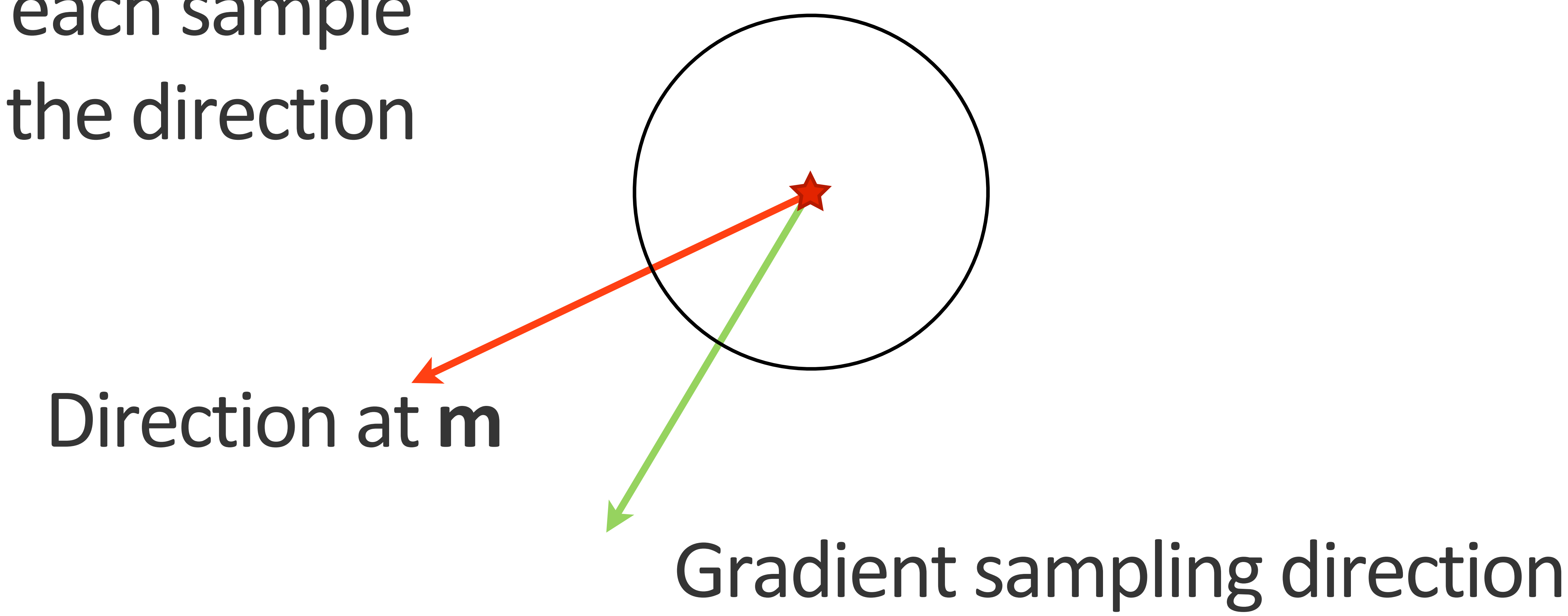


Direction at **m**

Gradient sampling direction

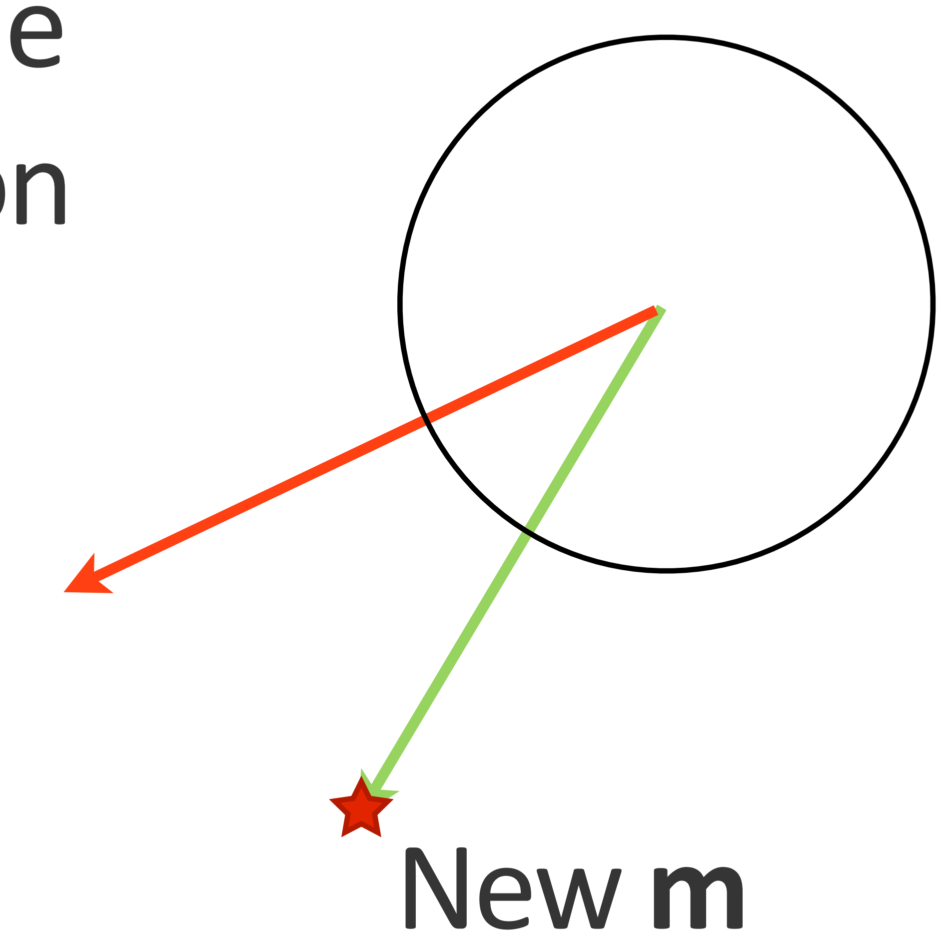
Gradient sampling

1- Define a ball around current point **m**
2- Take *p* sample inside the ball
3 - Compute direction for each sample
4 - Take weighted sum of the direction



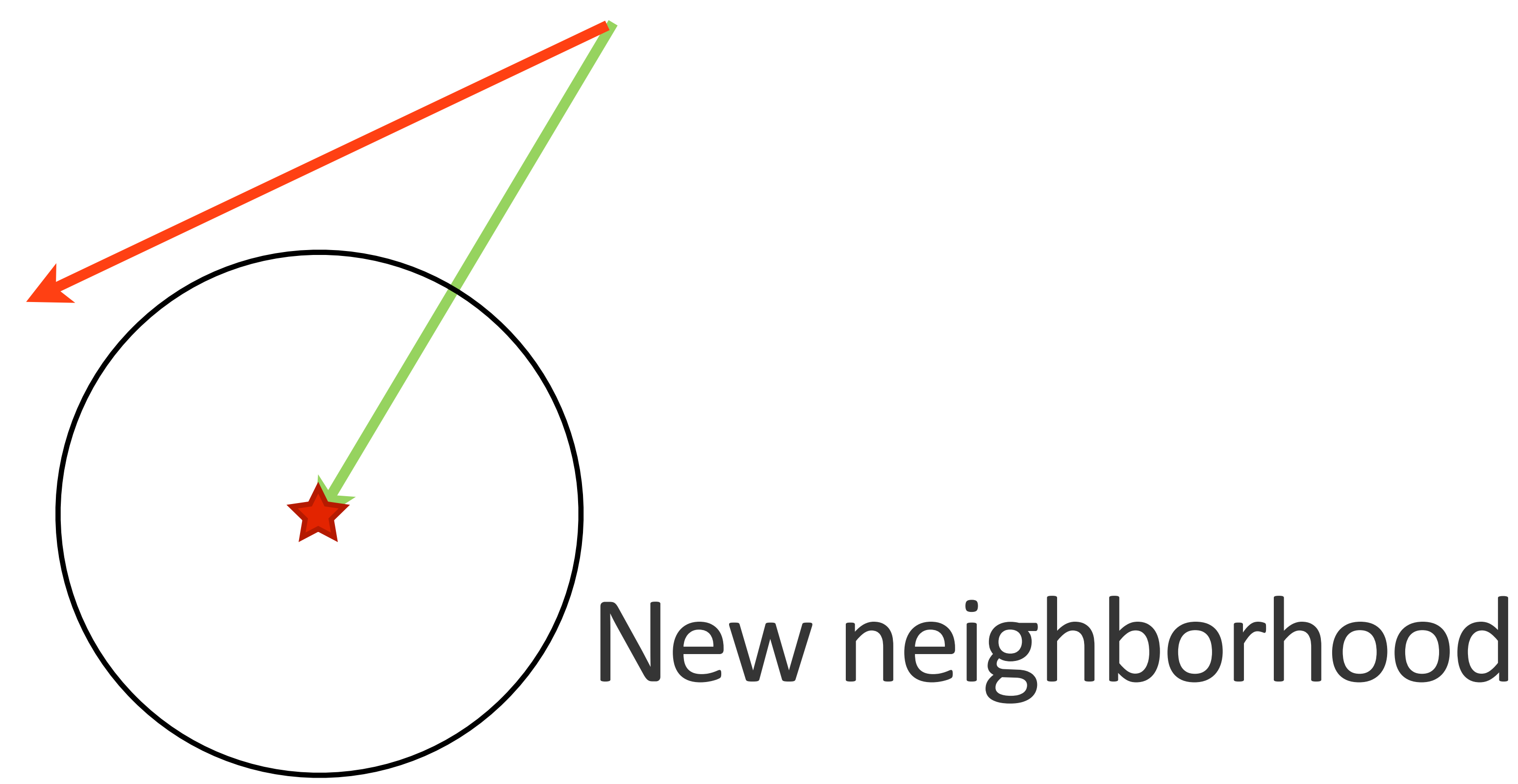
Gradient sampling

1- Define a ball around current point **m**
2- Take *p* sample inside the ball
3 - Compute direction for each sample
4 - Take weighted sum of the direction
5 - Update in this direction



Gradient sampling

1- Define a ball around current point **m**
2- Take *p* sample inside the ball
3 - Compute direction for each sample
4 - Take weighted sum of the direction
5 - Update in this direction
6 - Back to step 1



Summary

Update direction

- use information from “nearby” samples
- global direction instead of local
- proven to be robust for non-convex problems

Shortcomings

Needs to compute p gradients independently

- at each iterations
- for every source
- thousand times more expensive than FWI

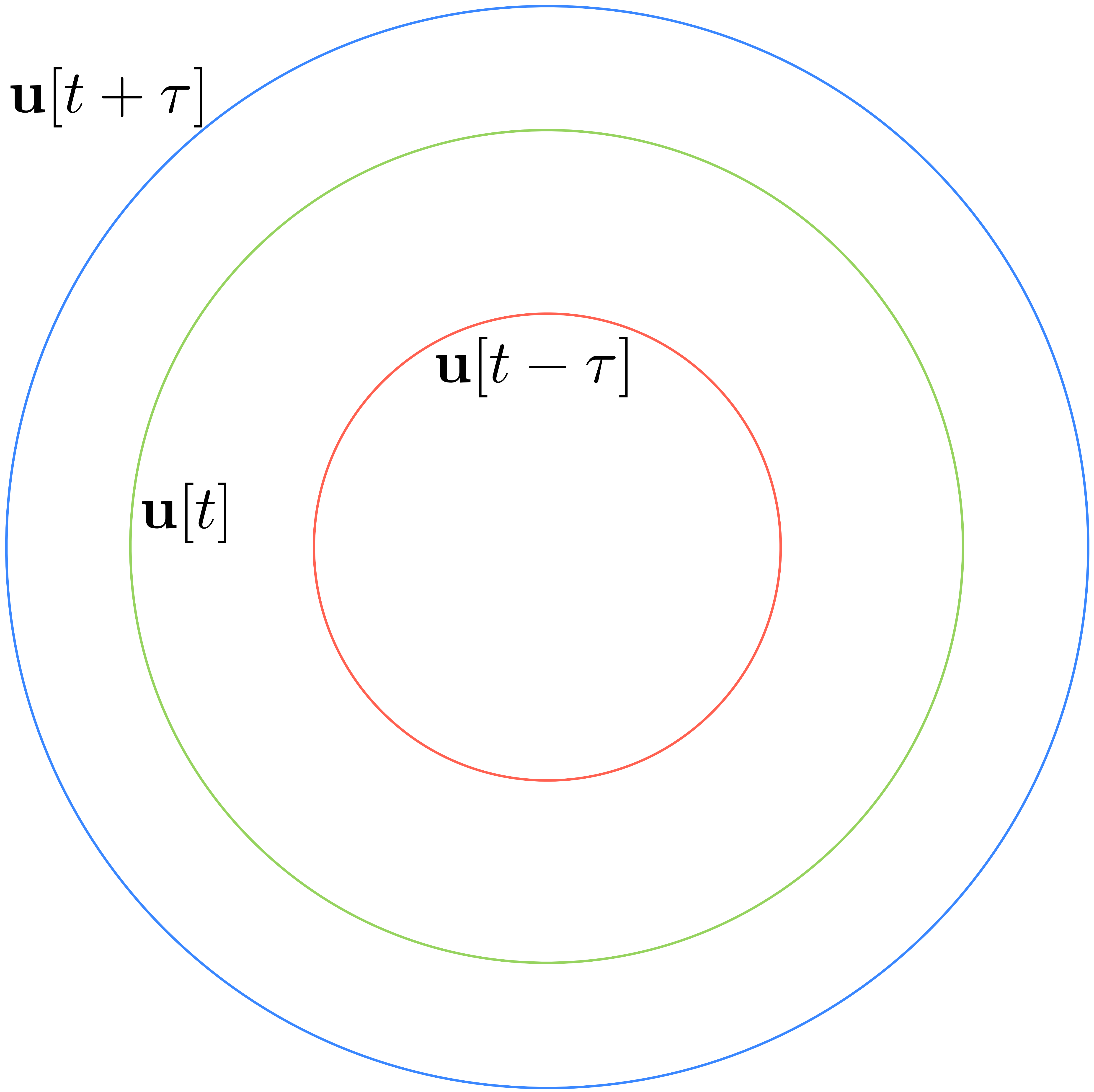
Redefine the neighborhood...

Small velocity change correspond to a time delay

Mathias Louboutin and Felix J. Herrmann, "Extending the search space of time-domain adjoint-state FWI with randomized implicit time shifts", in *EAGE Annual Conference Proceedings*, 2017

SLIM

Constant velocity model example



$\mathbf{u}[t+\tau]$ wavefield at t for a faster velocity

$\mathbf{u}[t-\tau]$ wavefield at t for a slower velocity

SLIM

Local update direction

Update direction for model $\mathbf{m}$ is

$$\nabla\Phi(\mathbf{m}) = -\sum_{t=0}^{n_t}\left[\mathrm{diag}(\mathbf{u}[t])(\mathbf{D}^T\mathbf{v}[t])\right]$$

where

$\mathbf{u}$ is the source wavefield for model $\mathbf{m}$

$\mathbf{v}$ is the receiver wavefield for model $\mathbf{m}$

$\Phi(\mathbf{m})$ is the FWI objective for model $\mathbf{m}$

SLIM

Neighbors update direction

Update direction for model $\mathbf{m} + \delta \mathbf{m}$ (slower)

$$\nabla \Phi(\mathbf{m} + \delta \mathbf{m}) = - \sum_{t=0}^{n_t} \left[\text{diag}(\mathbf{u}[t-\tau])(\mathbf{D}^T \mathbf{v}[t]) \right]$$

where

$\mathbf{u}$ is the source wavefield for model $\mathbf{m}$

$\mathbf{v}$ is the receiver wavefield for model $\mathbf{m}$

$\Phi(\mathbf{m} + \delta \mathbf{m})$ is the FWI objective for model $\mathbf{m} + \delta \mathbf{m}$

SLIM

Neighbors update direction

Update direction for model $\mathbf{m} - \delta\mathbf{m}$ (faster)

$$\nabla\Phi(\mathbf{m} - \delta\mathbf{m}) = -\sum_{t=0}^{n_t} \left[\mathrm{diag}(\mathbf{u}[t+\tau])(\mathbf{D}^T\mathbf{v}[t])\right]$$

where

$\mathbf{u}$ is the source wavefield for model $\mathbf{m}$

$\mathbf{v}$ is the receiver wavefield for model $\mathbf{m}$

$\Phi(\mathbf{m} - \delta\mathbf{m})$ is the FWI objective for model $\mathbf{m} - \delta\mathbf{m}$

GS subproblem

Quadratic subproblem

$$\arg\min_{\omega} \frac{1}{2}\omega^T G^T G\omega$$
$$\text{s.t. } \omega_i > 0 \text{ for all } i \in [0, p],$$
$$\omega^T \mathbf{1} = 1.$$

- where

  G is the matrix of the gradients $G = [\mathbf{g}_1; ...; \mathbf{g}_p]$
  and $\mathbf{1}$ is a vector of ones.

Still requires p gradients => gradient free in data-space

Quadratic subproblem in data-space

- One entry of the subproblem matrix

$$(G^T G)_{i,j} = \delta \mathbf{d}(\tau_{ki})^T \underbrace{\mathbf{J}(\tau_{ki})\mathbf{J}(\tau_{kj})^T}_{I} \delta \mathbf{d}(\tau_{kj}).$$

- Correlation of a residual with a migrated-demigrated residual.

$$(G^T G)_{i,j} \simeq \delta \mathbf{d}(\tau_{ki})^T \delta \mathbf{d}(\tau_{kj}).$$

Only need to compute correlation of shifted version of the residual

FInal direction

- Boils down to a modified adjoint source

$$\mathbf{g}_k = \sum_{t=1}^{n_t} \frac{\partial \mathbf{u}(\mathbf{t})}{\partial t^2} \tilde{\mathbf{v}}(\mathbf{t})$$

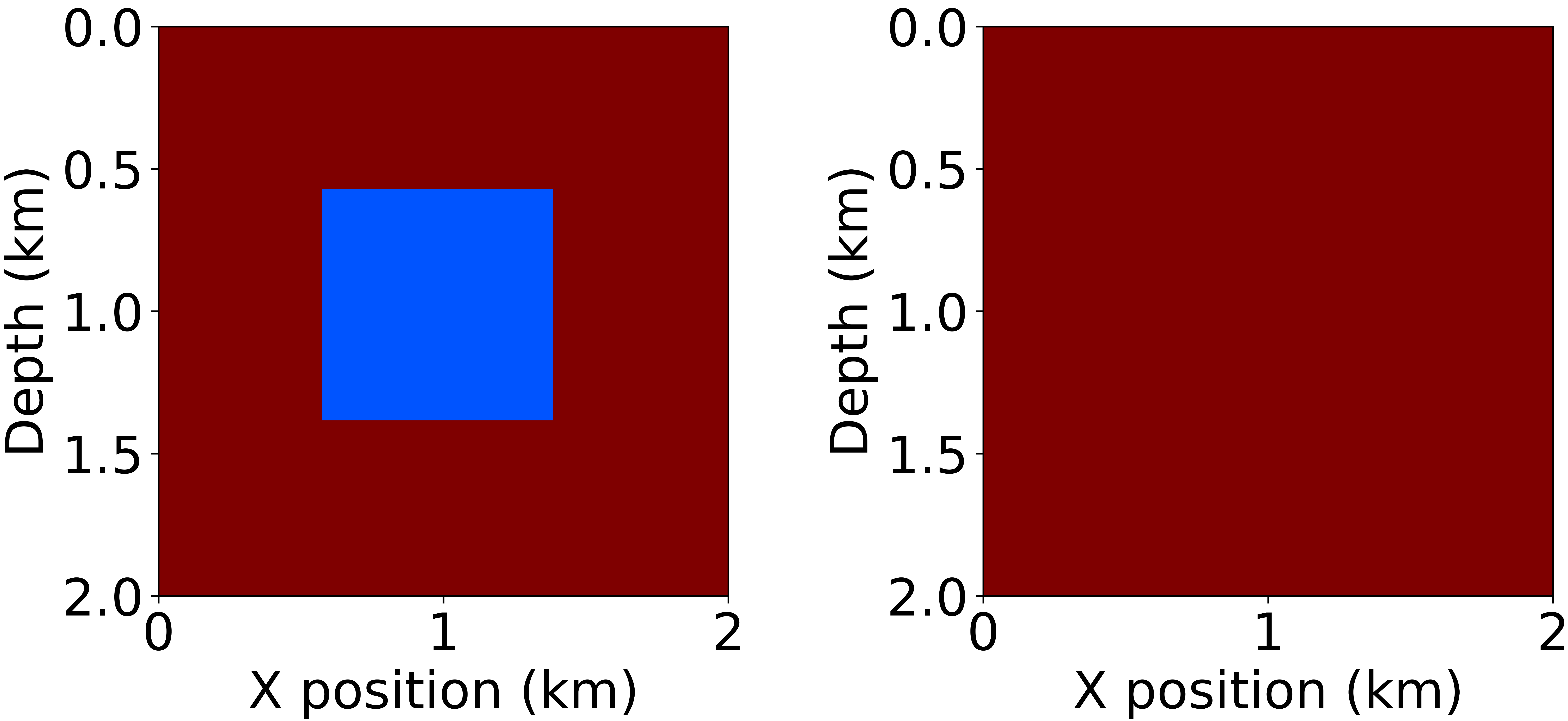
$$\tilde{\mathbf{v}} = \mathbf{A}^{-T} \mathbf{P}_r^T \delta \mathbf{d}_{GS}$$

$$\delta \mathbf{d}_{GS} = \sum_{i=1}^{p} \omega_i (\mathbf{d}_{syn}(\mathbf{t} - \tau_{ki}) - \mathbf{d}_{obs})\big|_{t=t+2\tau_i}$$

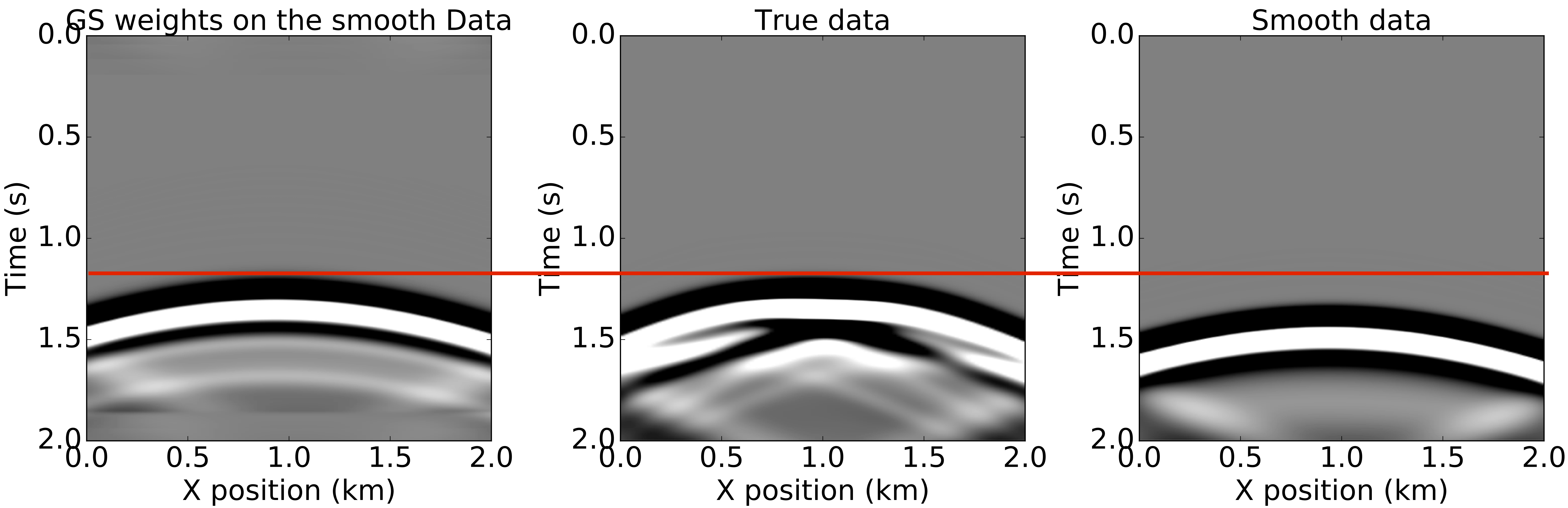
- No extra computation, only need the conventional forward wavefield

Examples

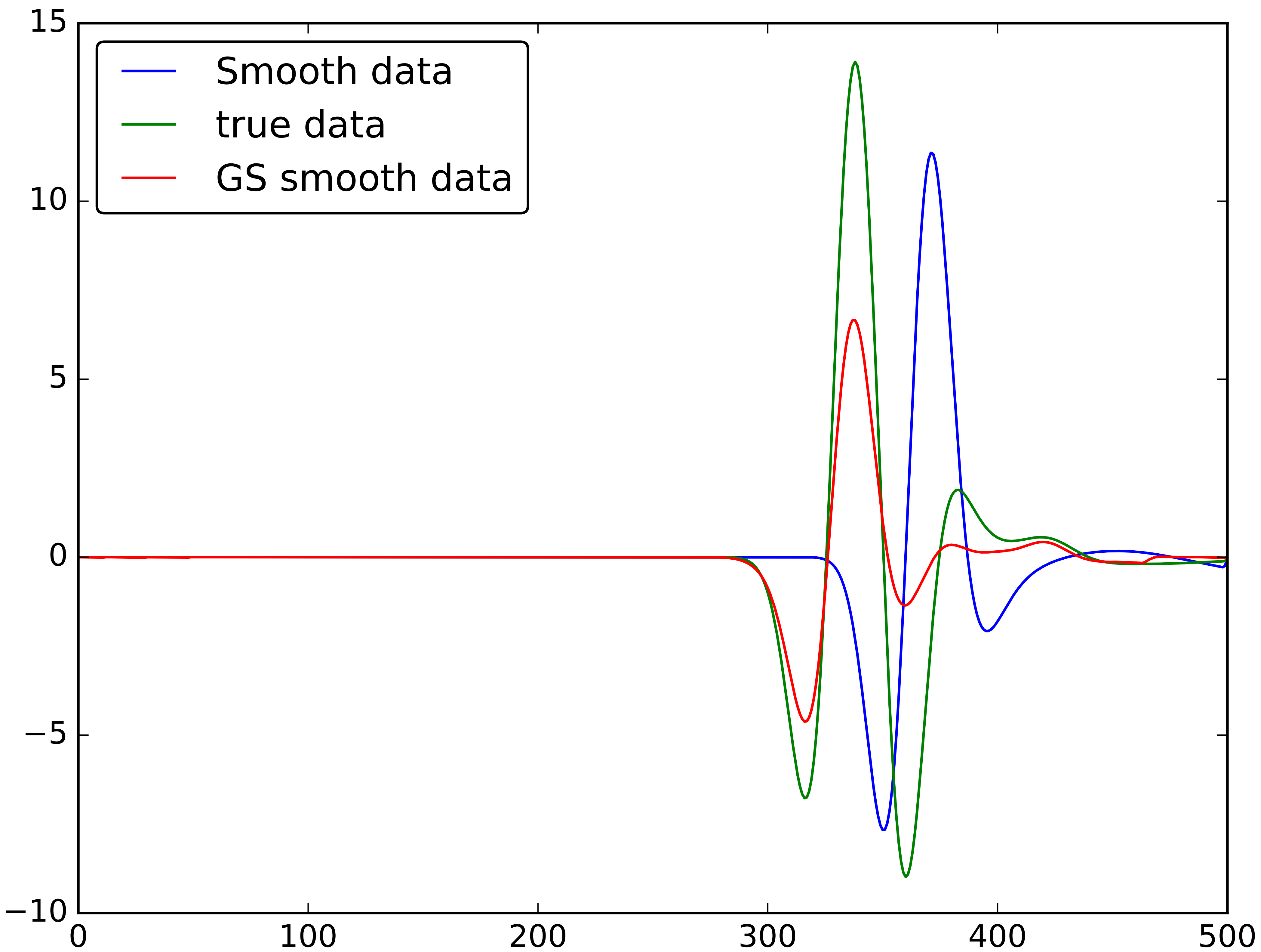
Simple transmission

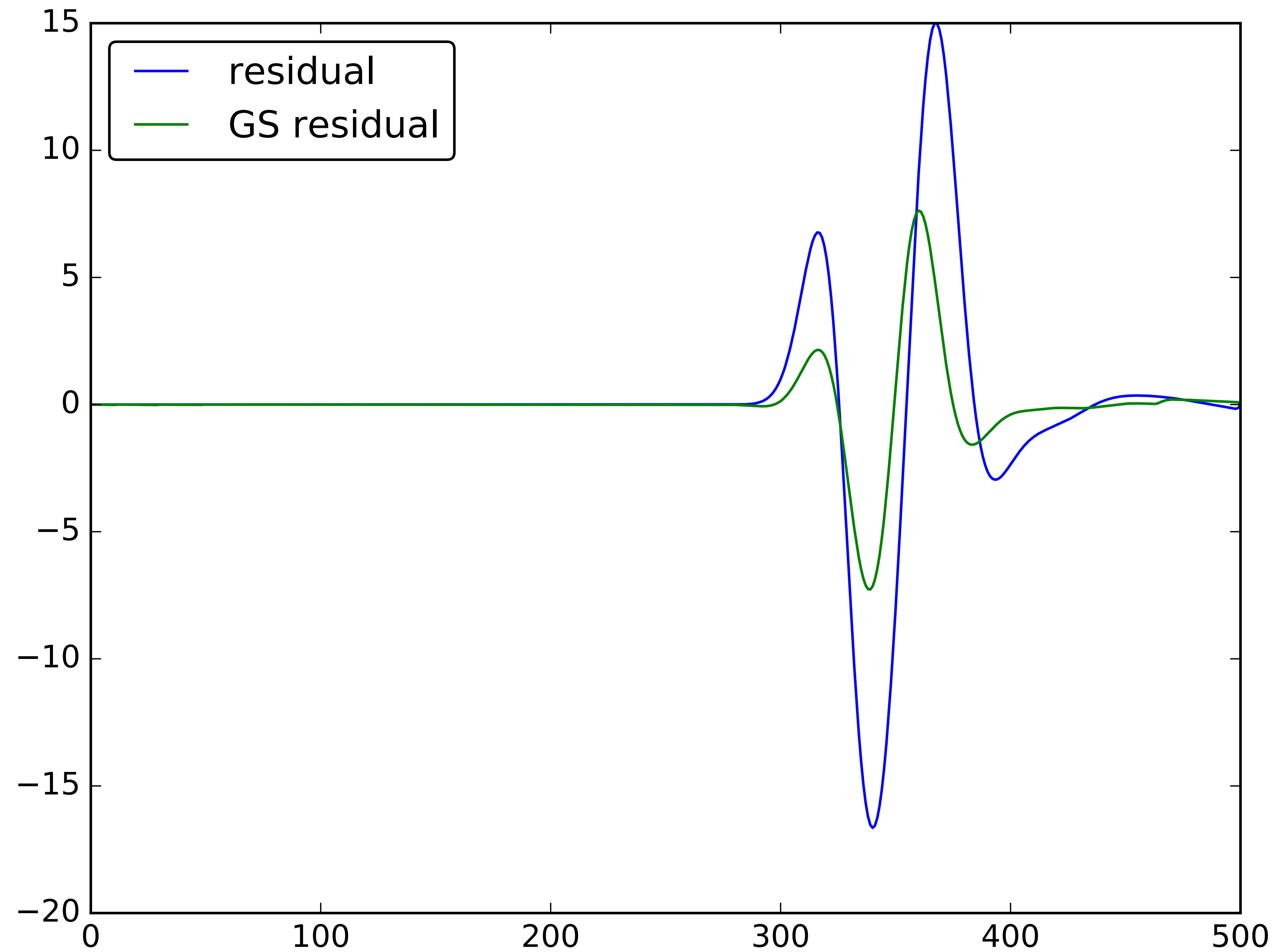

0.0
0.5
1.0
1.5
2.0
Depth (km)
0
1
2
X position (km)
0.0
0.5
1.0
1.5
2.0
Depth (km)
0
1
2
X position (km)

Subproblem solution

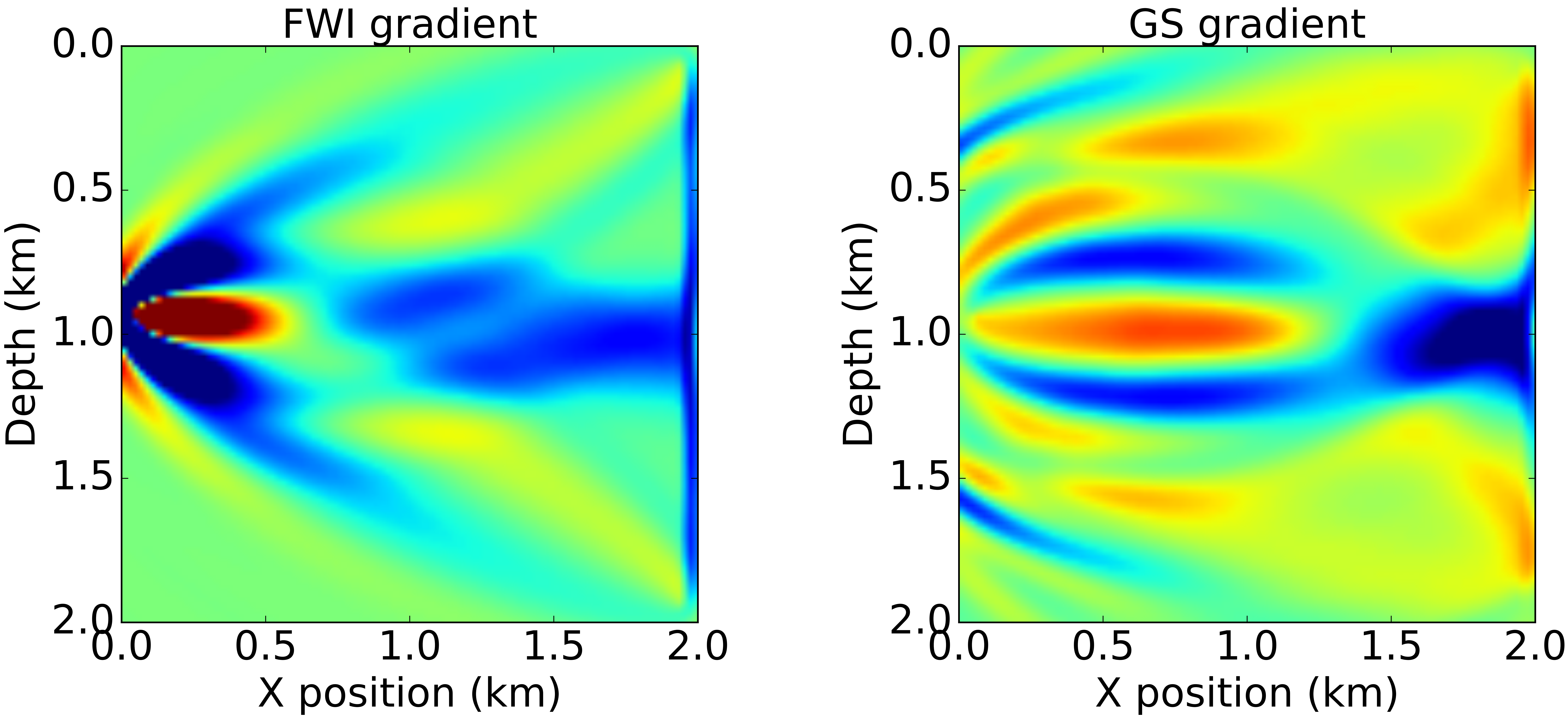


Subproblem solution

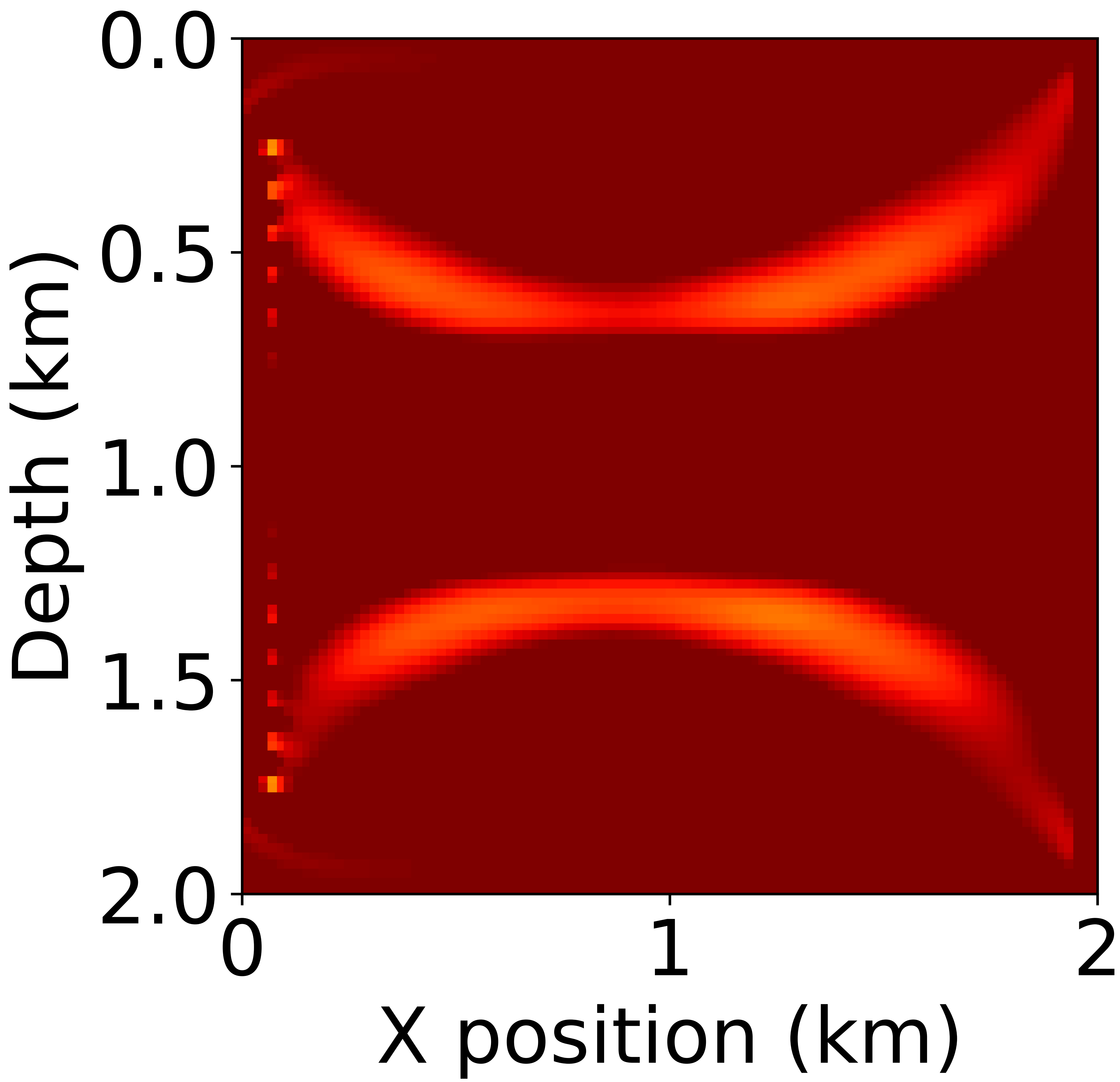


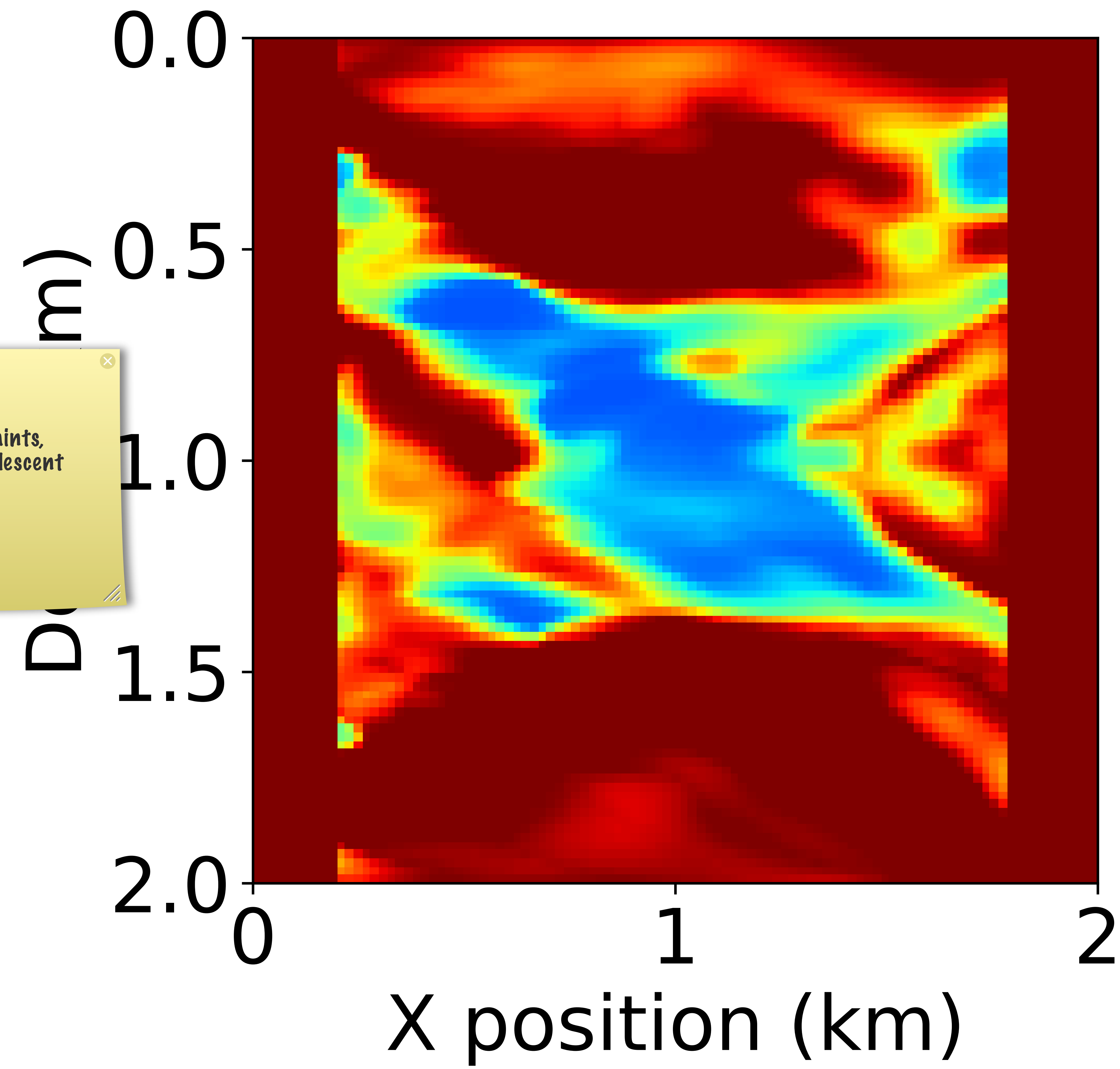


Single source gradient



Inversion results





What is this?

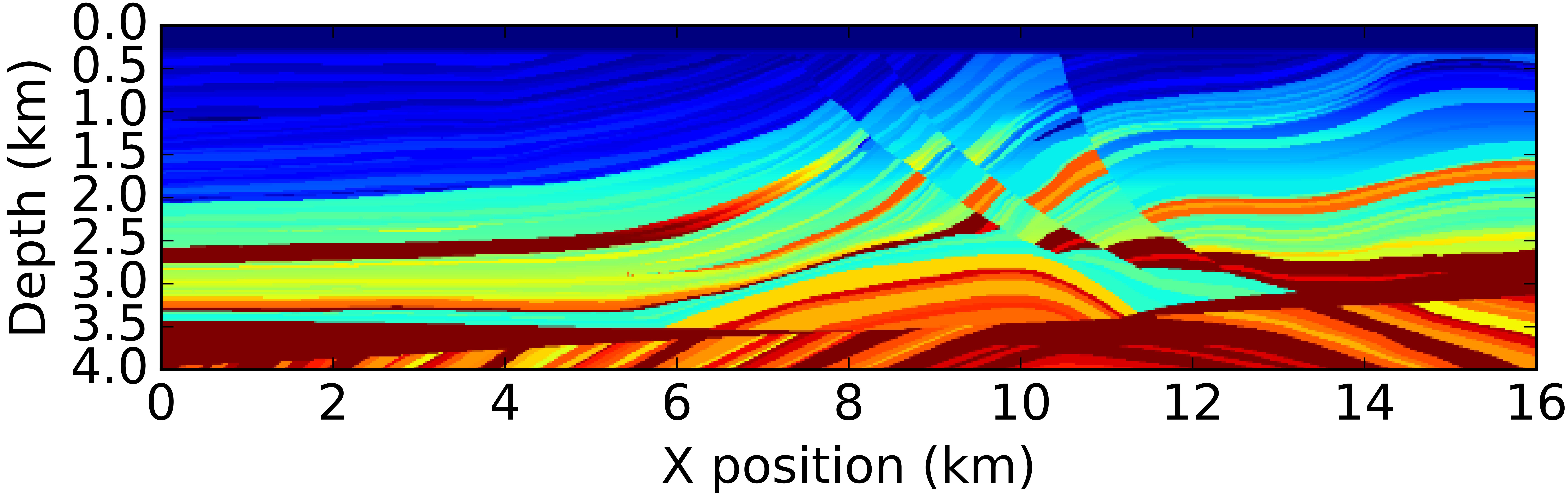
ML: No constraints, basic gradient descent

Marmousi

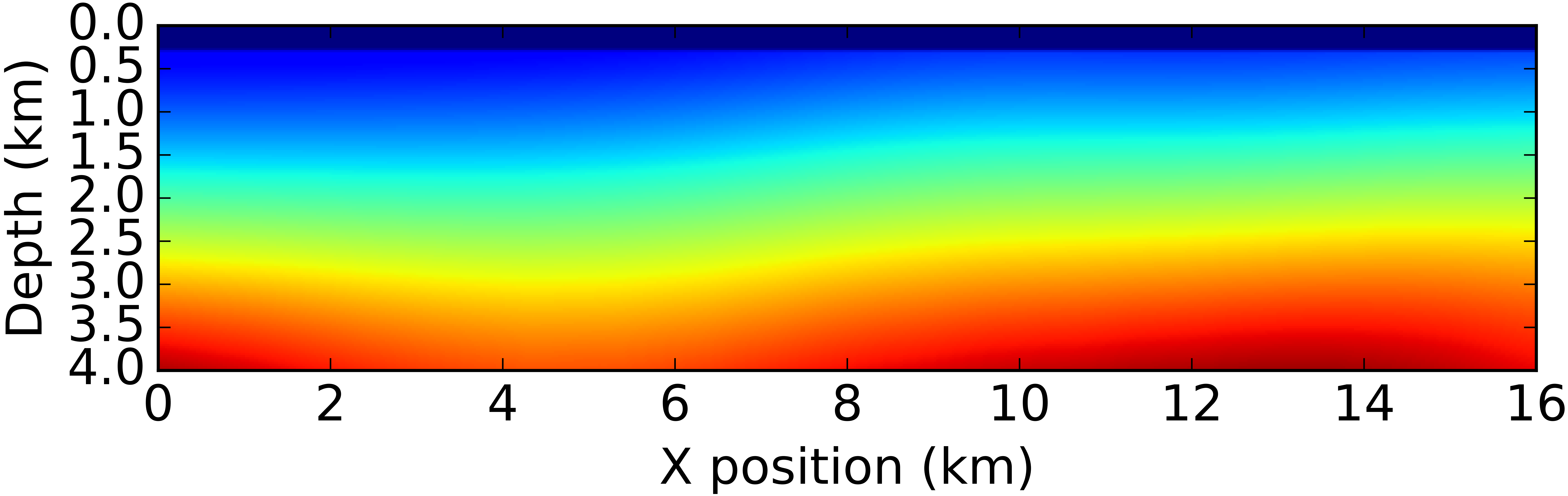
Setup

- 4km x 16 km
- 320 sources 50m deep
- 501 receivers 300m deep (ocean bottom)
- 6 seconds recording
- Source signature is a sum of a 8Hz Ricker wavelet and a 12 Hz Ricker wavelet.

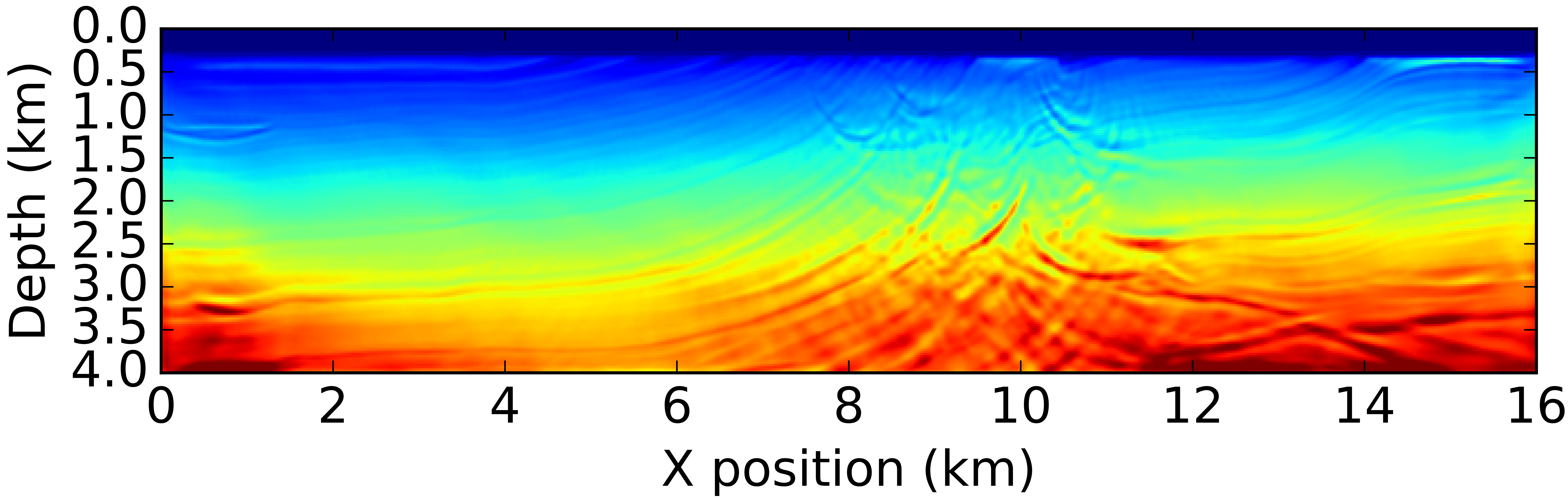
- 20 SPG iterations at 8 Hz, bounds constraints only
- 20 SPG iterations at 12 HZ, bounds constraints only



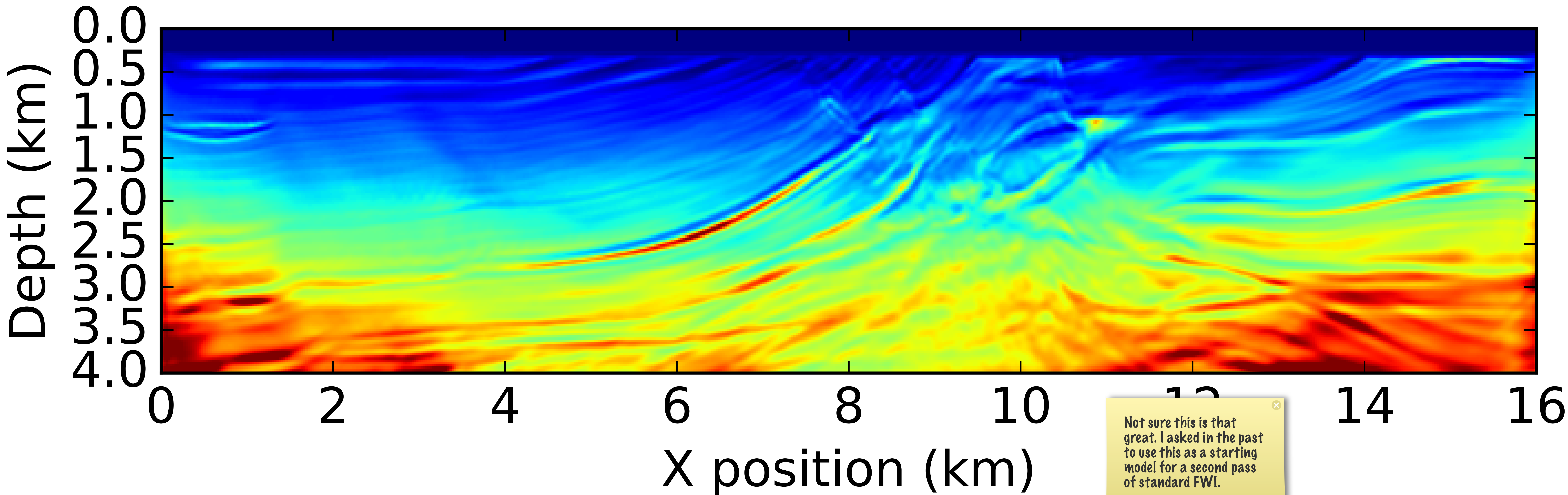
Depth (km)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
0
2
4
6
8
10
12
14
16
X position (km)


Depth (km)
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
0
2
4
6
8
10
12
14
16
X position (km)

FWI

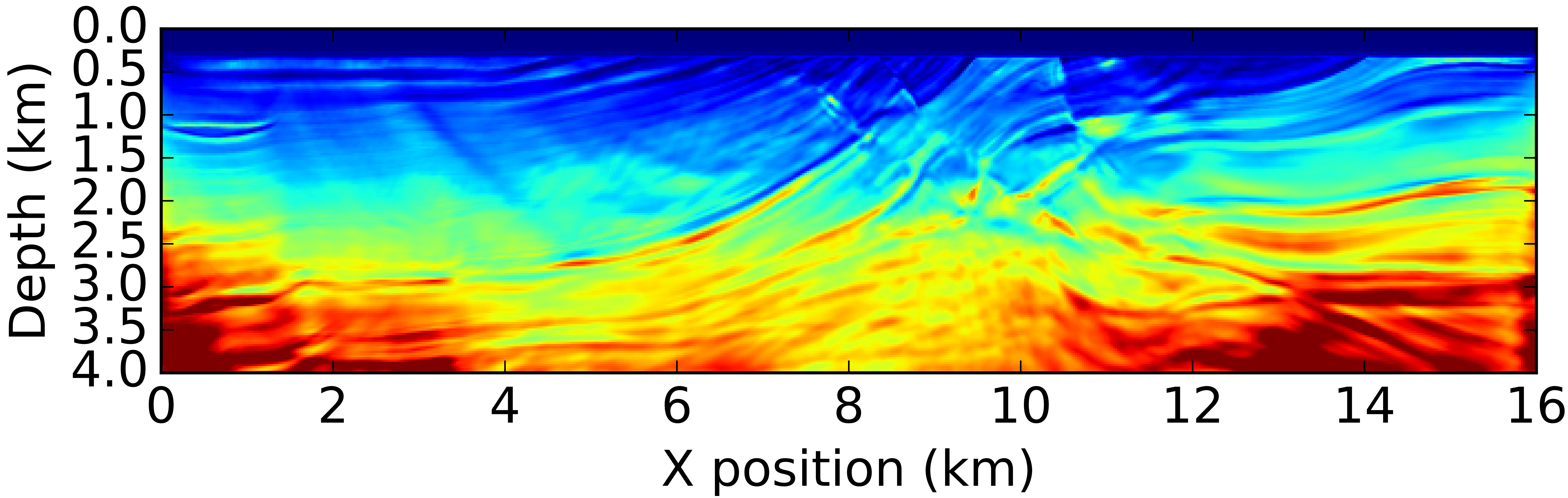


GSFWI



Not sure this is that great. I asked in the past to use this as a starting model for a second pass of standard FWI.

20 extra SPG FWI iterations at 8 Hz



SLIM

Conclusion

Implicit extension of the model space

Same computational/memory cost than FWI

Potentially more robust

Easy to implement

Future work

Explore limits of the robustness

Non linear shifts (take into account propagation time/distance)

Theoretical results?

Acknowledgements

Thank you for your attention !

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

SINBAD

NSERC CRSNG

This work was financially supported by SINBAD Consortium members BG Group, BGP, CGG, Chevron, ConocoPhillips, DownUnder GeoSolutions, Hess, Petrobras, PGS, Schlumberger, Statoil, Sub Salt Solutions and Woodside; and by the Natural Sciences and Engineering Research Council of Canada via NSERC Collaborative Research and Development Grant DNOISEII (CRDPJ 375142--08).