Extending the search space of time-domain adjoint-state FWI w/ randomized implicit time shifts

Mathias Louboutin
SINBAD Fall consortium meeting 2015

SLIM
University of British Columbia

Extending the search space of time-domain adjoint-state FWI w/ randomized implicit time shifts (confidential)

Mathias Louboutin
SINBAD Fall consortium meeting 2015

SLIM
University of British Columbia

Motivations

Sensitivity to cycle skipping

Memory cost

- storing 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"

Motivations

Global methods have shown good results

- low-rank extension
- full-space

New way to extend the research space for time-domain.

Gradient Sampling Algorithm

Designed for Non-Smooth Non-Convex problems:

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

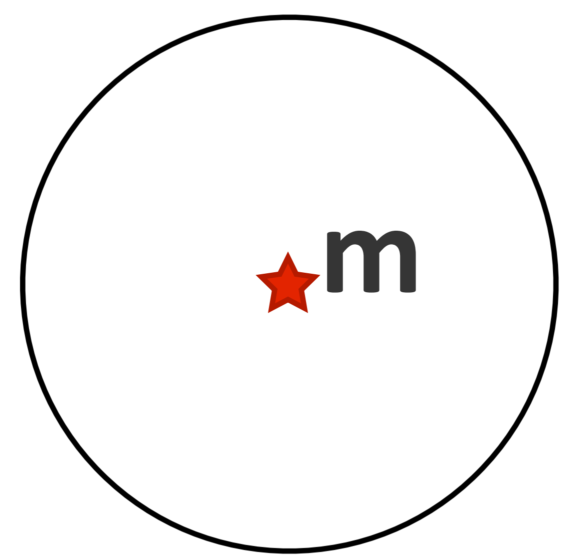
Gradient sampling

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**

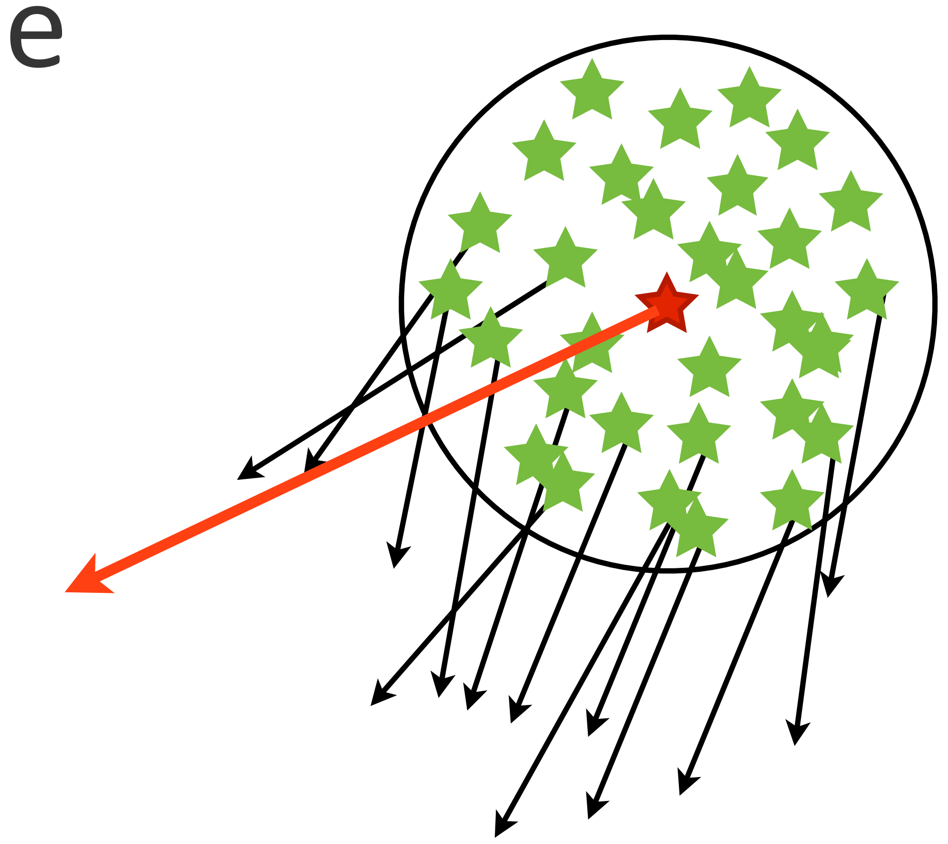
2- Take *p* sample inside the ball

Gradient sampling

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

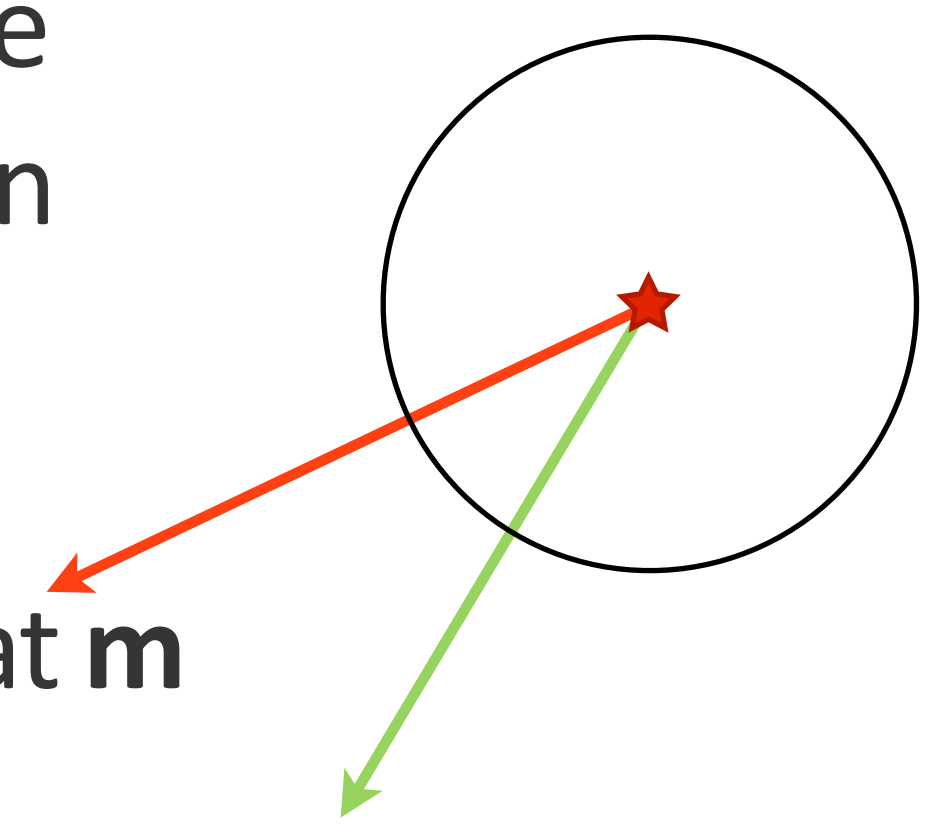
2- Take *p* sample inside the ball

3 - Compute direction for each sample



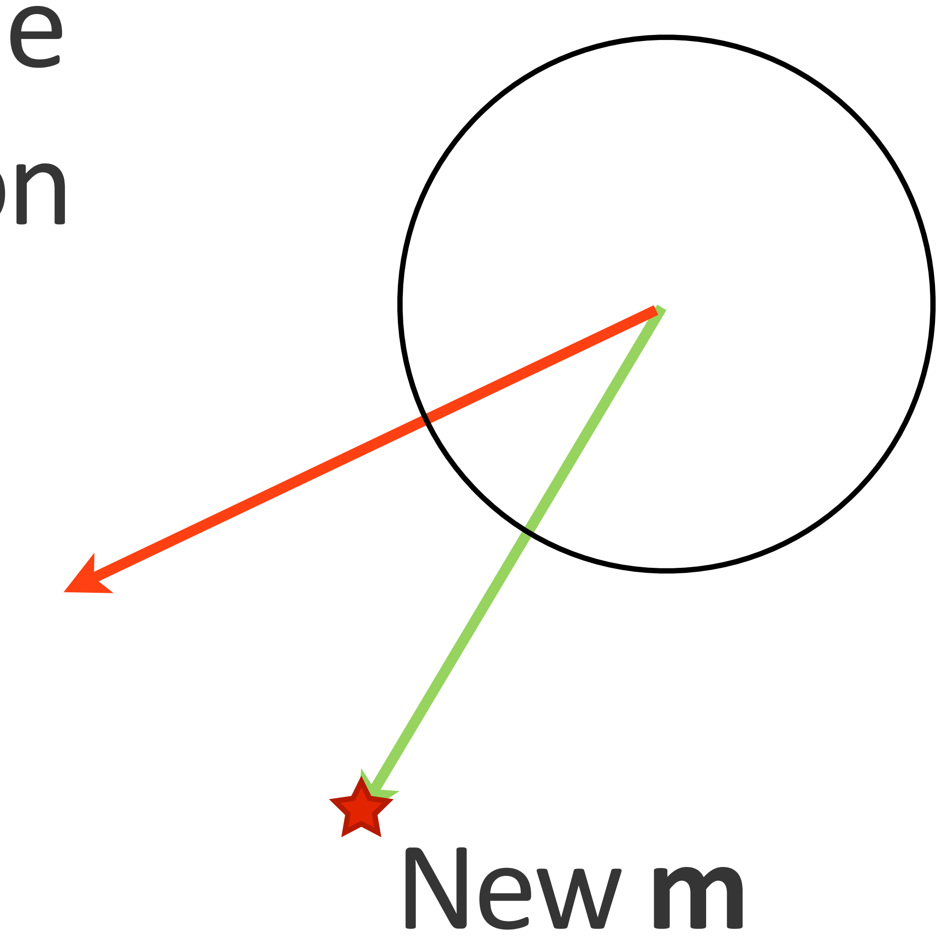
Gradient sampling

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**

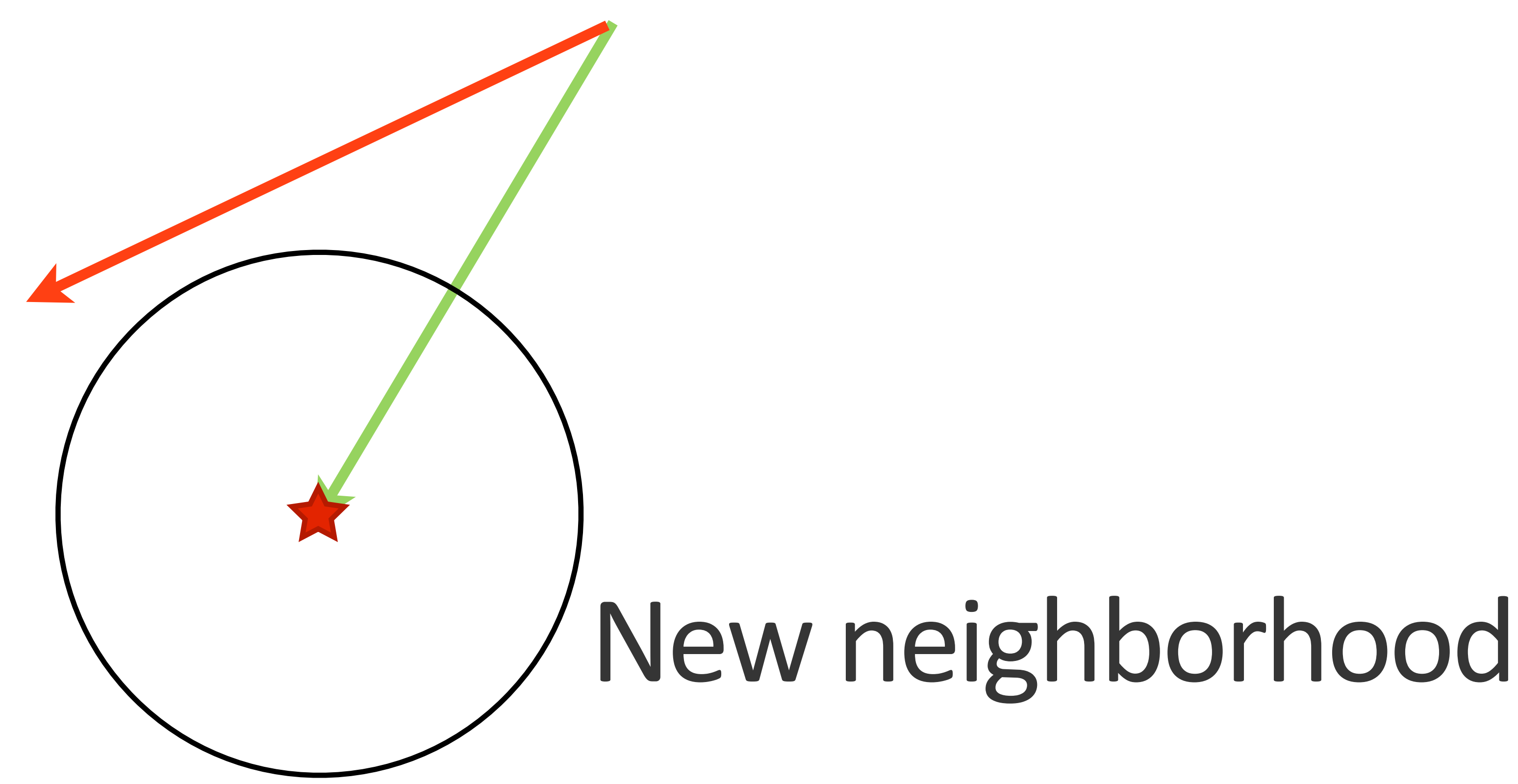
2- Take *p* sample inside the ball

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 iteration
- *p* times more expensive than FWI

Shortcomings

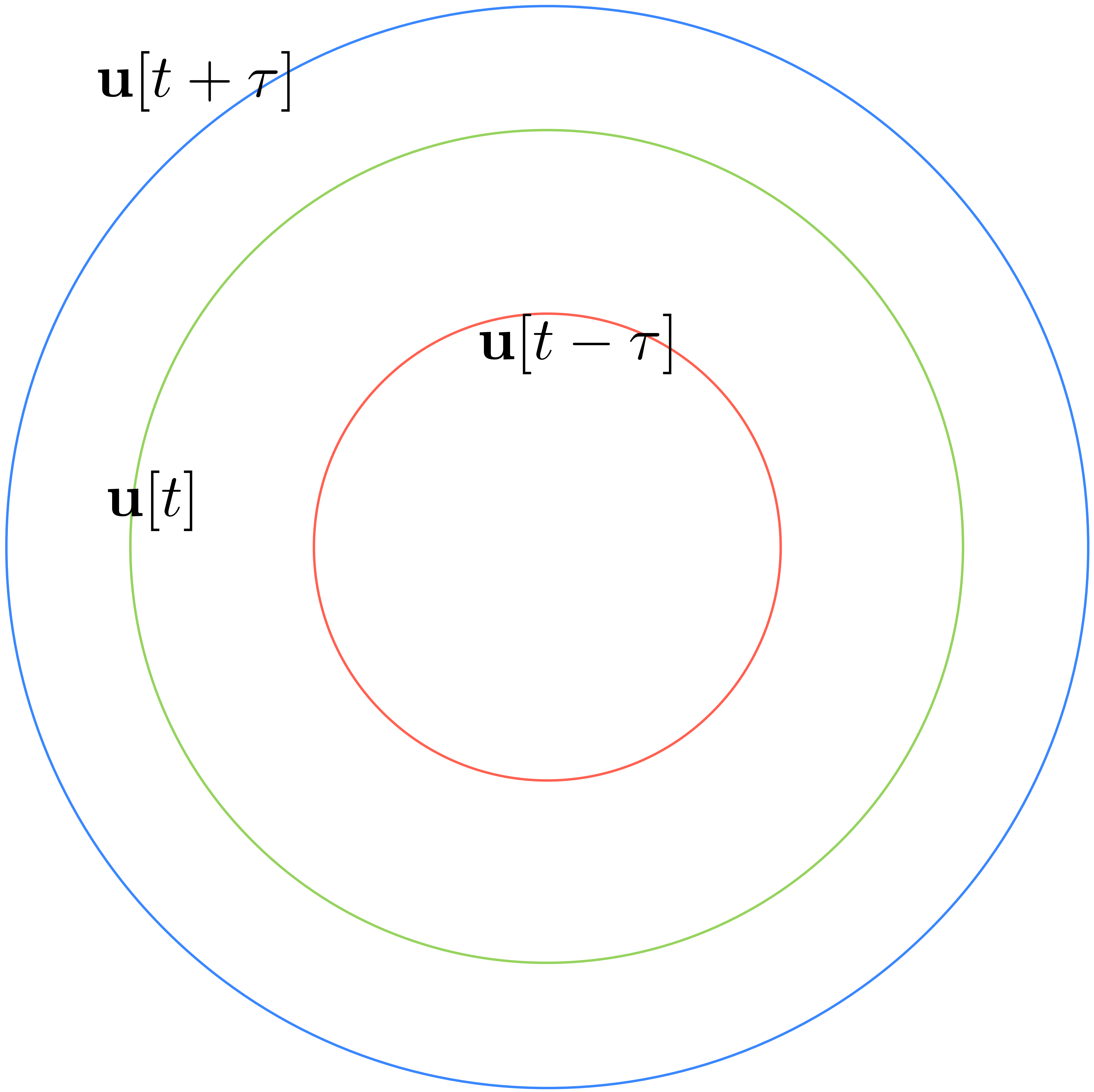
Needs to compute p gradients independently

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

Redefine the neighborhood...

Small velocity changes correspond to a time delay

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

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 $\mathbf{m}$

Neighbors update direction

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

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

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 $\mathbf{m} + \delta\mathbf{m}$

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 $\mathbf{m} - \delta\mathbf{m}$

Weighted sum of the gradients

Gradient sampling direction becomes

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

where

$$\bar{\mathbf{v}}[t] = \sum_{\tau=0}^{\epsilon} \alpha_\tau \mathbf{v}[t-\tau]$$

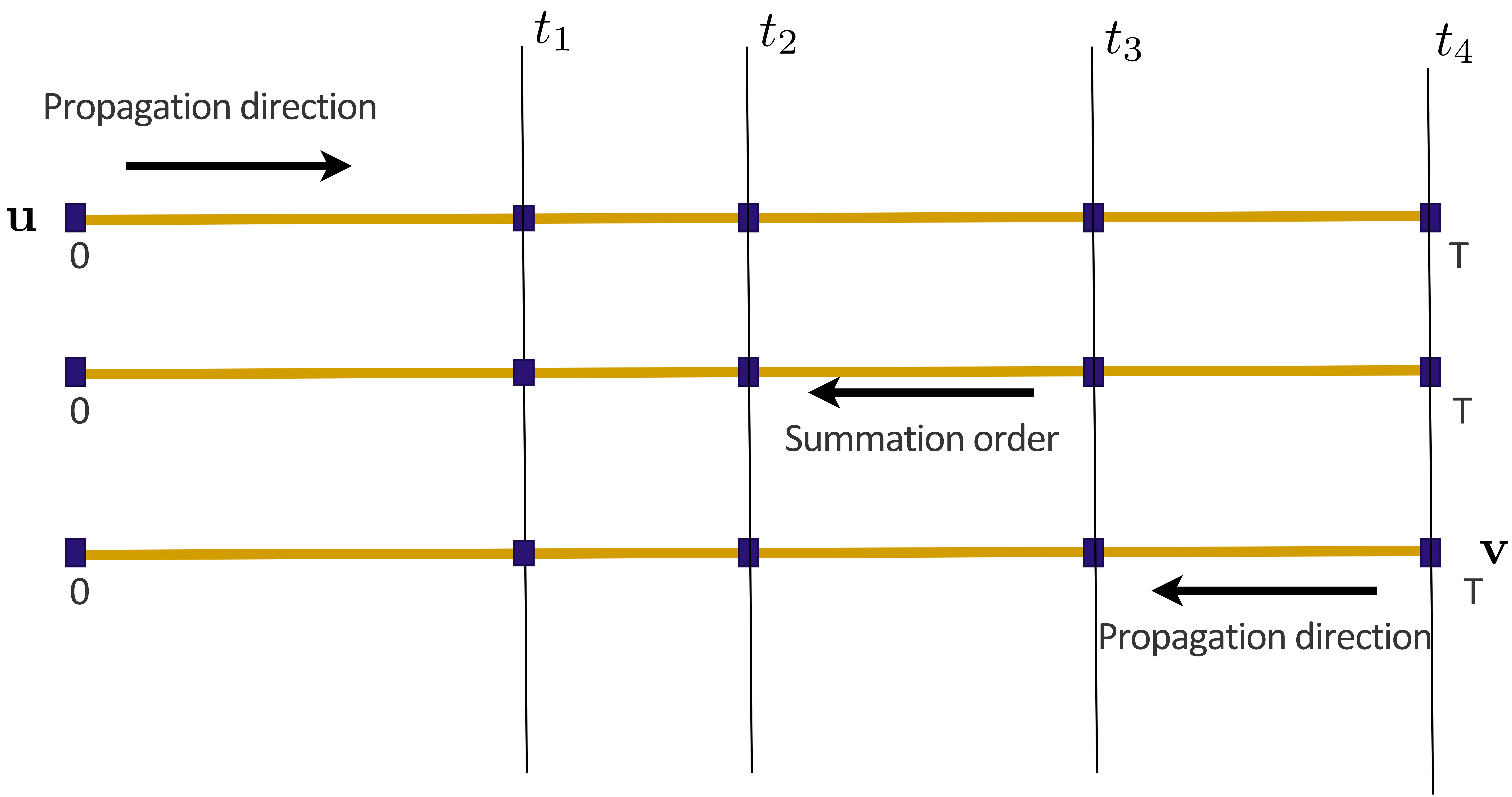
$$\bar{\mathbf{u}}[t] = \sum_{\tau=0}^{\epsilon} \alpha_\tau \mathbf{u}[t-\tau]$$

ϵ Maximum shift

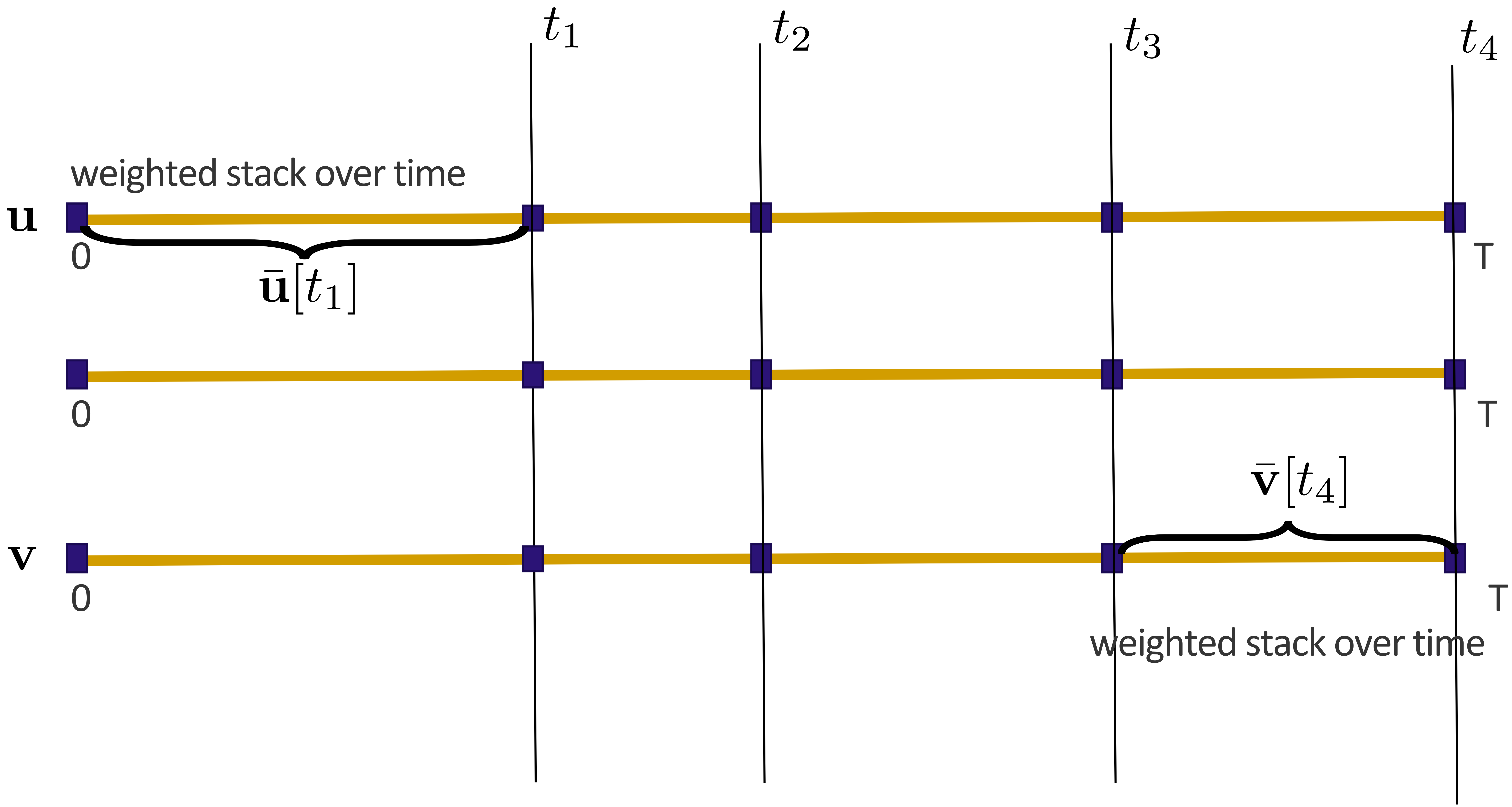
ω_t depends on α_τ

α_τ random numbers in $[0, 1]$

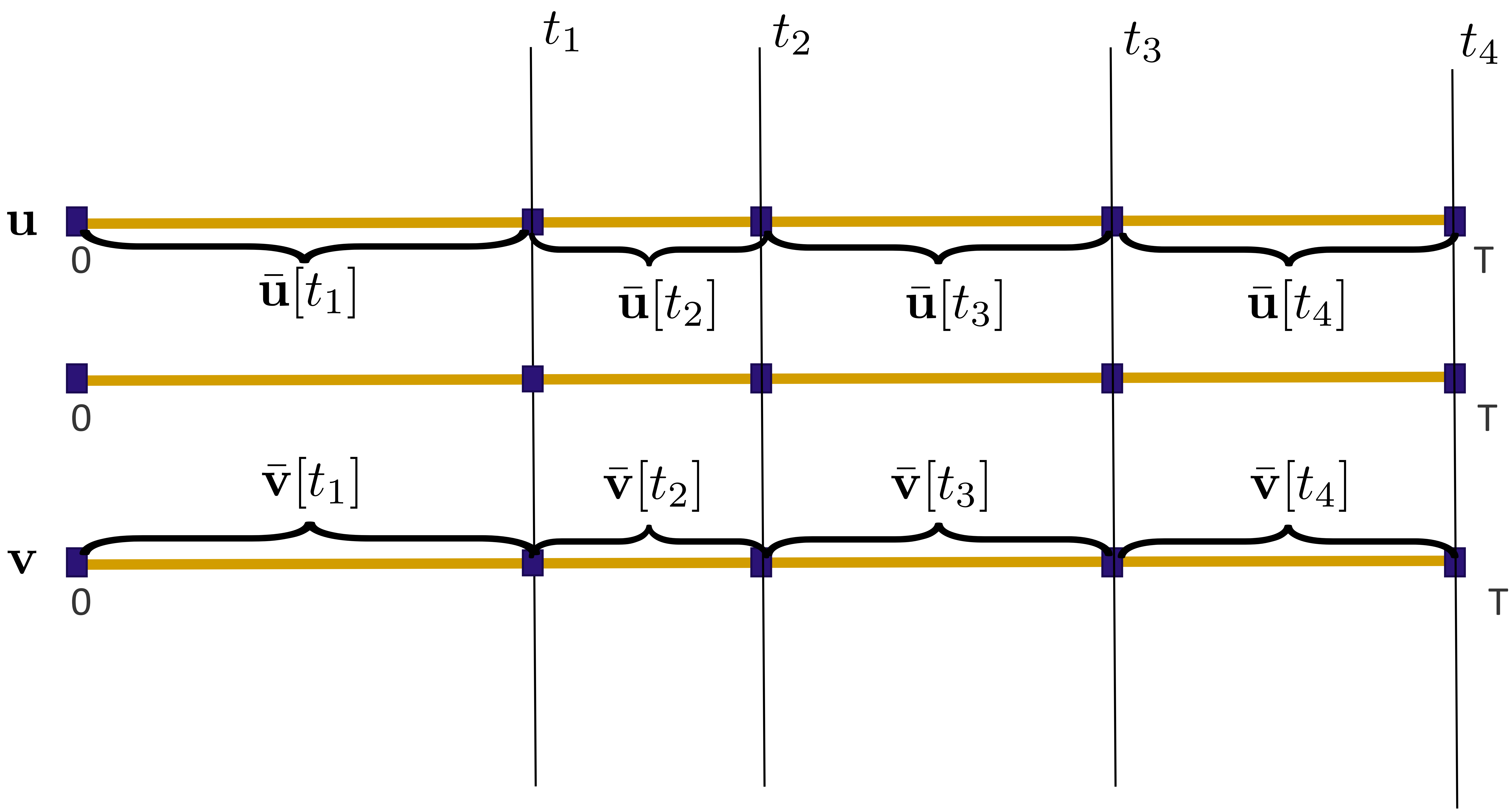
On-the-fly compressed gradient sampling



On-the-fly compressed gradient sampling



On-the-fly compressed gradient sampling



SLIM

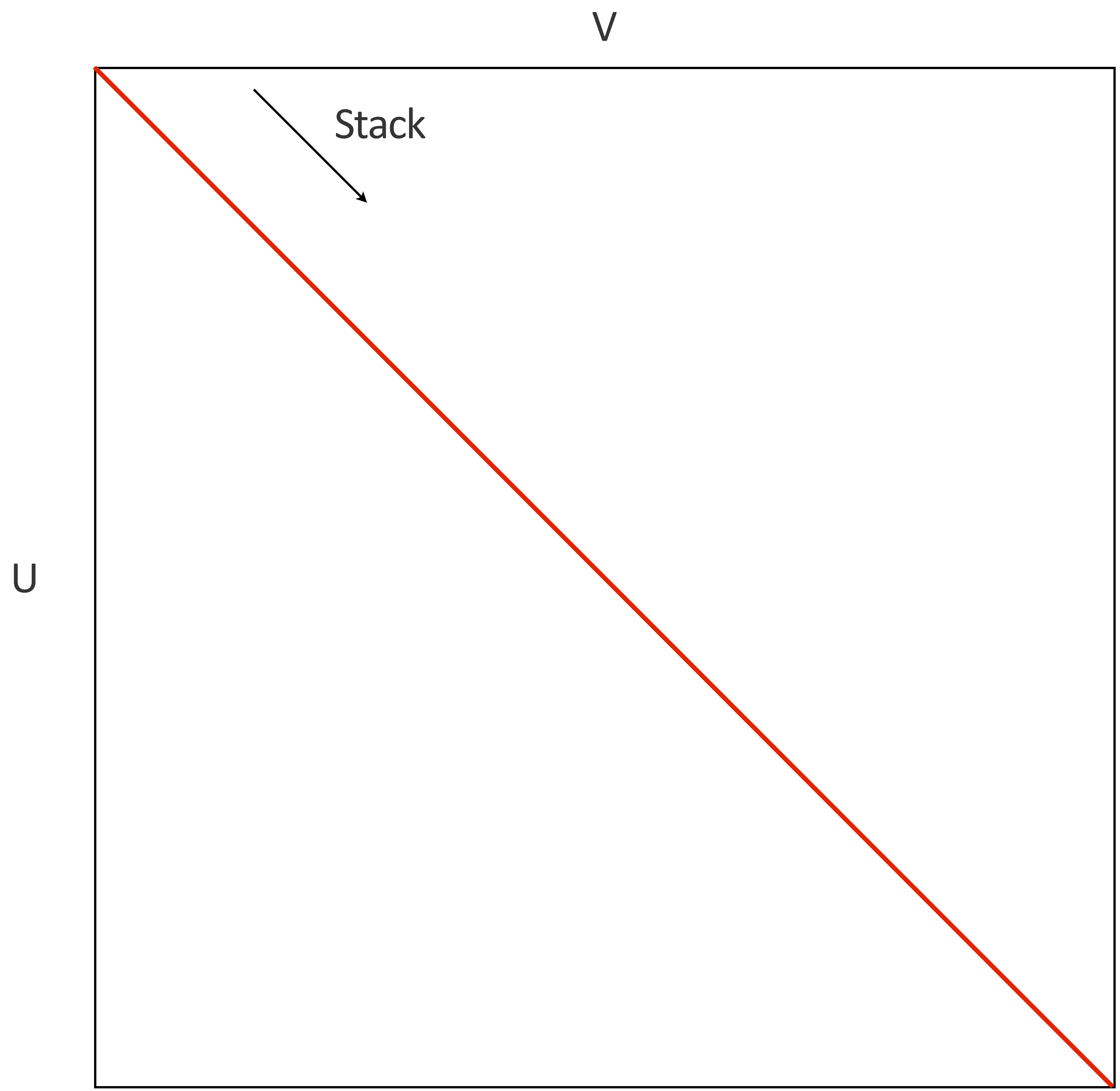
Gives a time compressibly sampled gradient sampling direction

$$\nabla\Phi_w(\mathbf{m}) = -\sum_{t\in I}\left[\mathrm{diag}(\bar{\mathbf{u}}[t])(\mathbf{D}^T\bar{\mathbf{v}}[t])\right]$$

$$I = \{t_1, t_2, t_3, t_4\}$$

In the previous cartoon

- redrawing new time indexes for each source
- redrawing new weights for each source



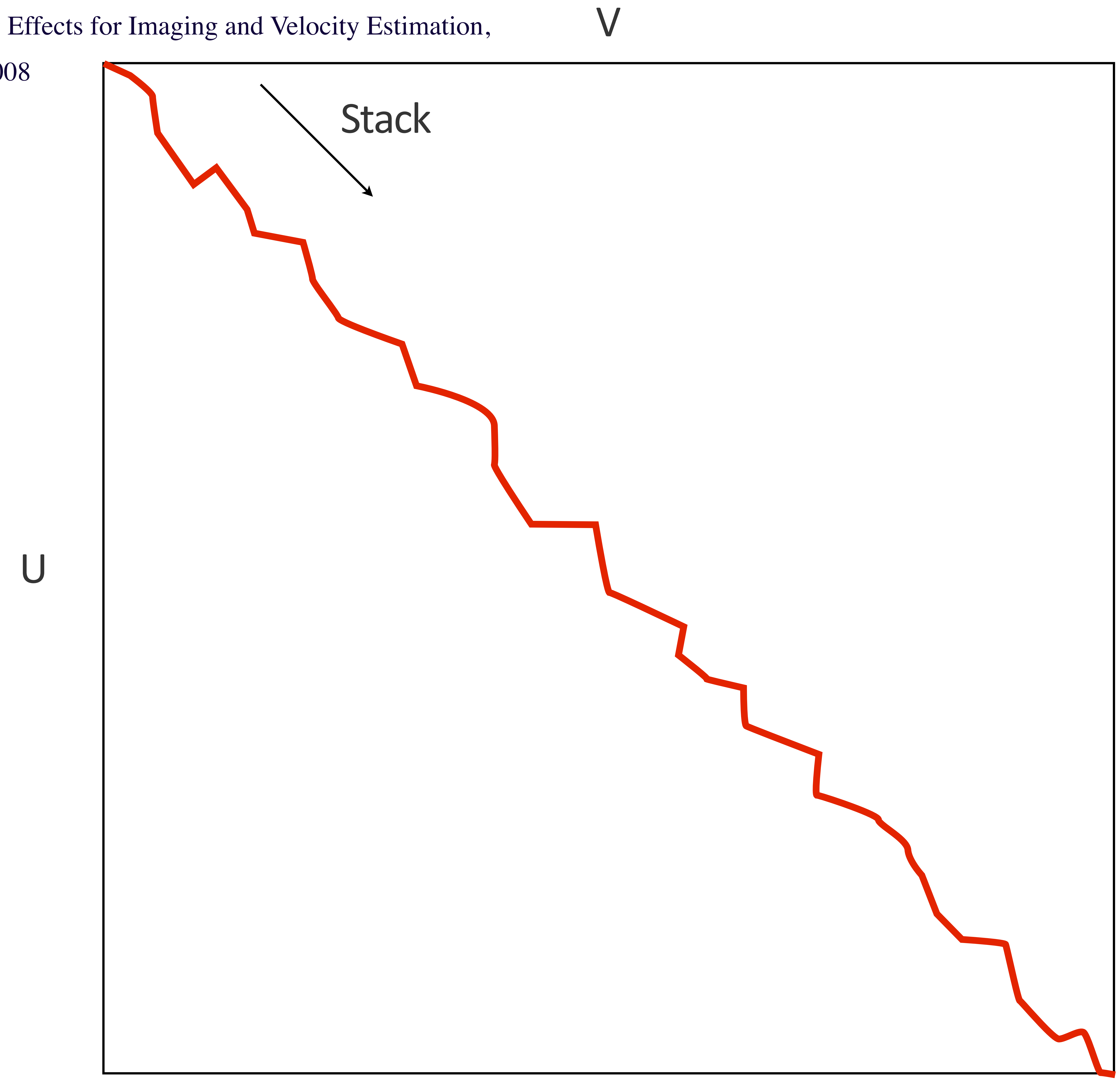
FWI

$$\sum_{t=1}^{n_t} \mathbf{u}[t]\mathbf{v}[t]$$

Time-shift imaging condition in seismic migration, Paul Sava and Sergey Fomel, GEOPHYSICS, VOL. 71, NO. 6 NOV-DEC 2006; P. S209–S217, 16 FIGS.10.1190/1.2338824

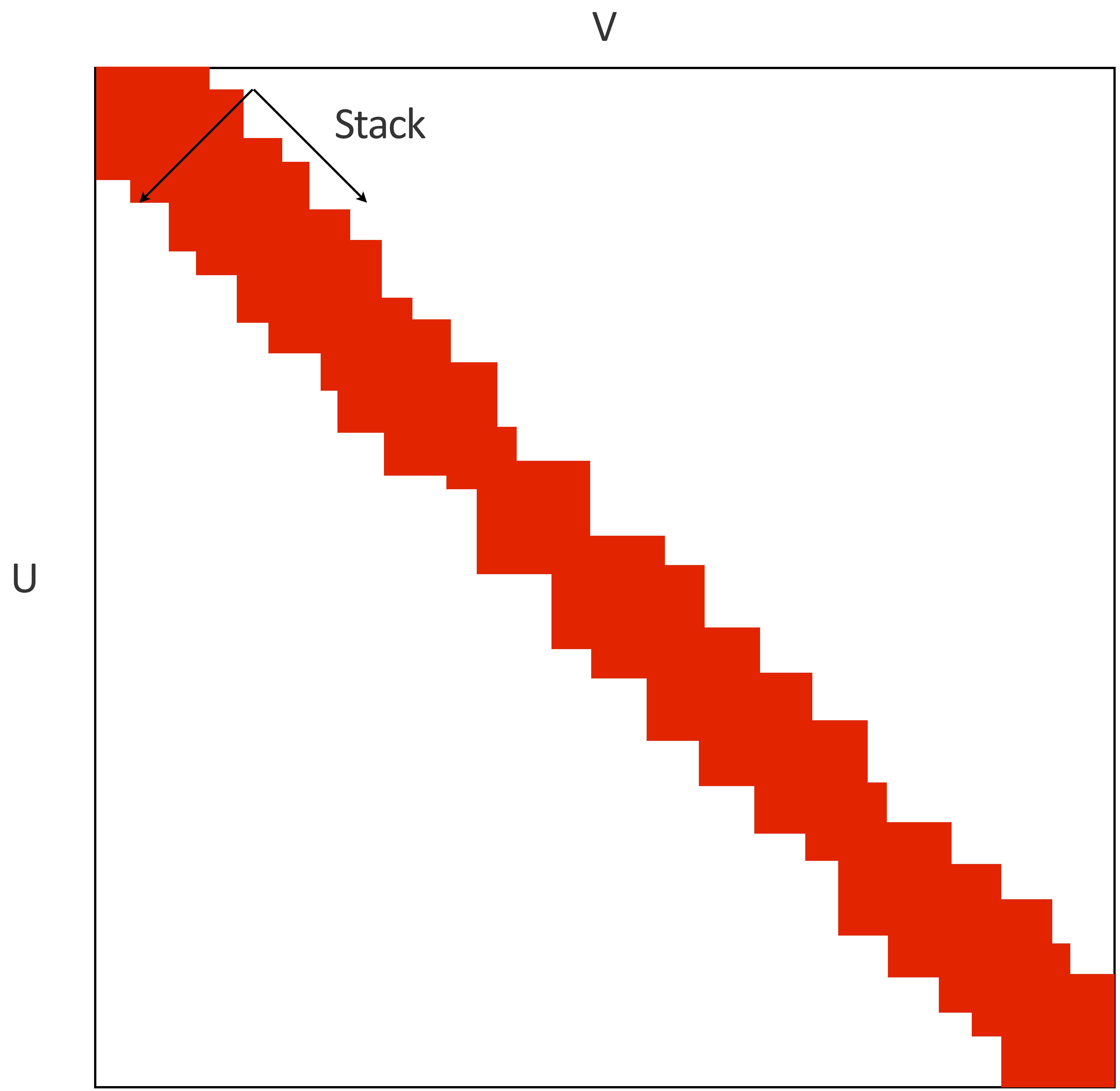
Filtering Random Layering Effects for Imaging and Velocity Estimation,

F.G delCueto, W.Symes 2008

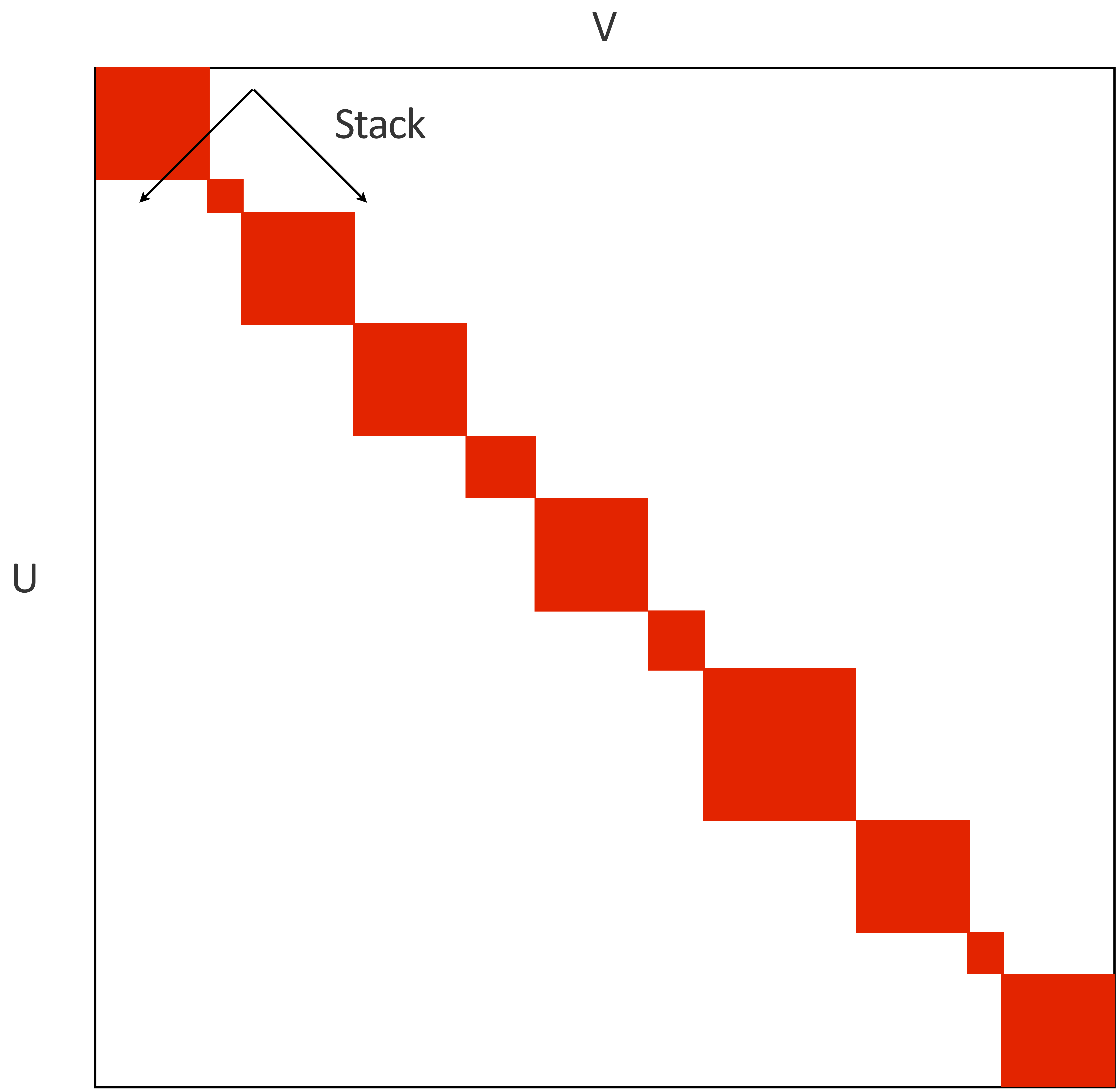


$$\sum_{t=1}^{n_t} \mathbf{u}[t+\delta t]\mathbf{v}[t+\delta t']$$

$$\mathbb{E}(\delta t) = \mathbb{E}(\delta t') = 0$$



Implicit time shift
Full history



Time compressed
implicit time shift

Summary

Time-compressed implicit gradient sampling

- uses information from "nearby models"
- for an interval of length p uses p^2 different models
- search direction is now global
- "nearby models" calculated cheaply on the fly w/ weighted stacking
- reduces memory usage

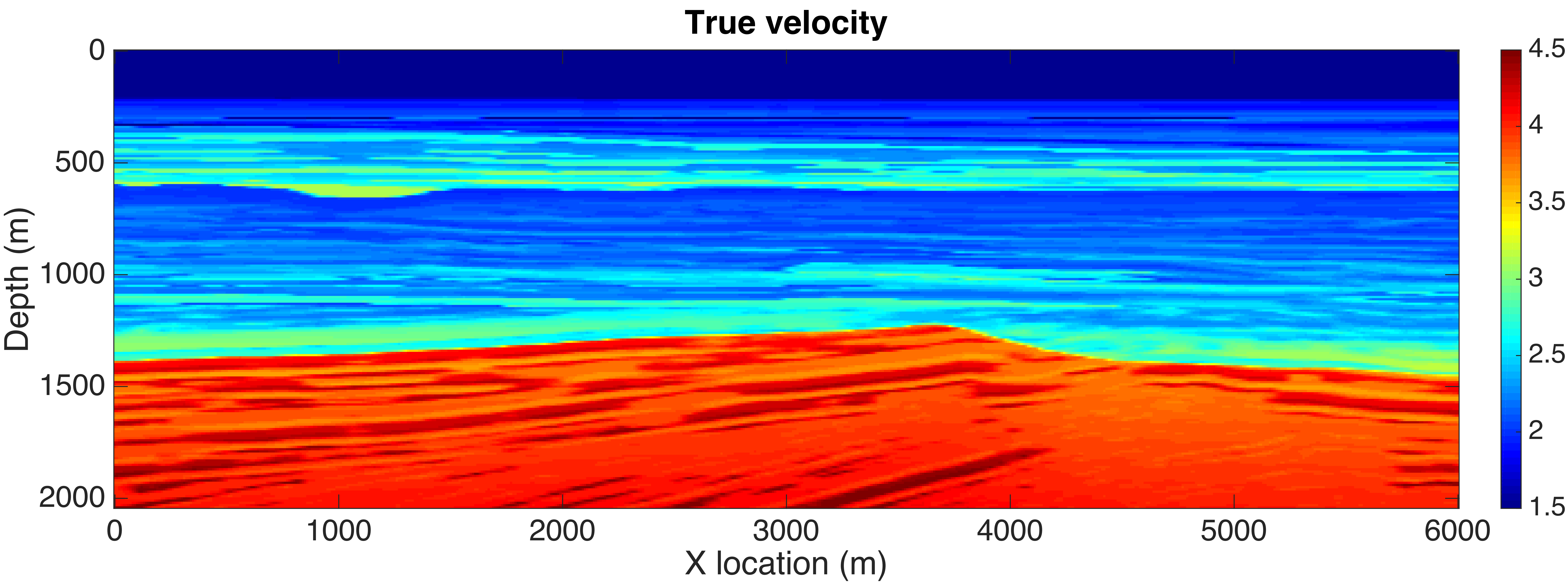
SLIM

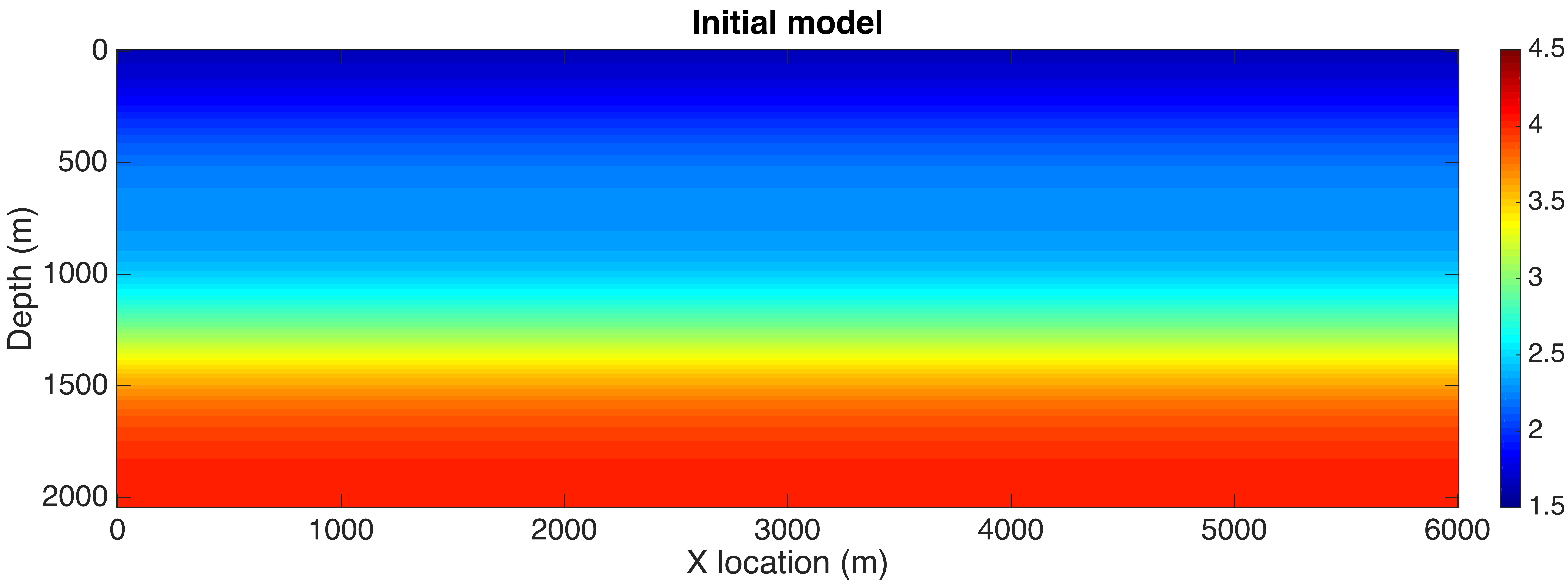
BG Compass 2D

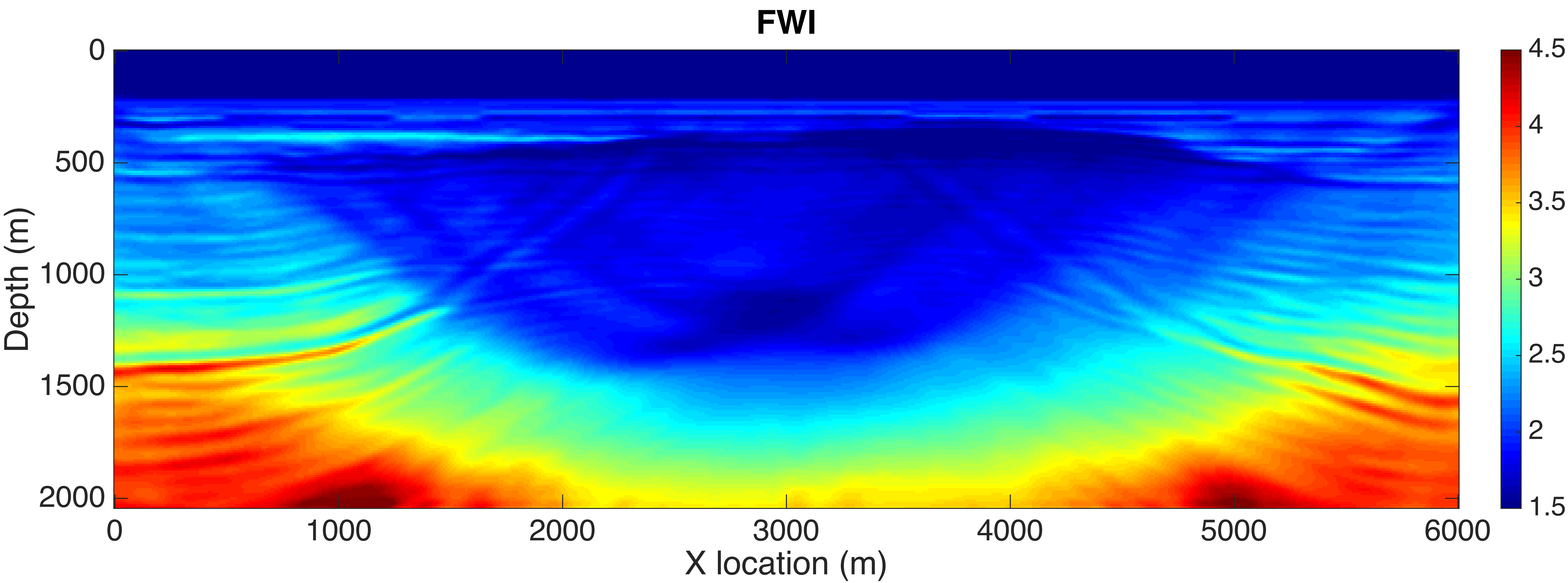
- **Data:**
  - Ricker wavelet at 15Hz, 2.4s recording
  - 61 sources at 100m interval
  - 251 receivers at 25m interval

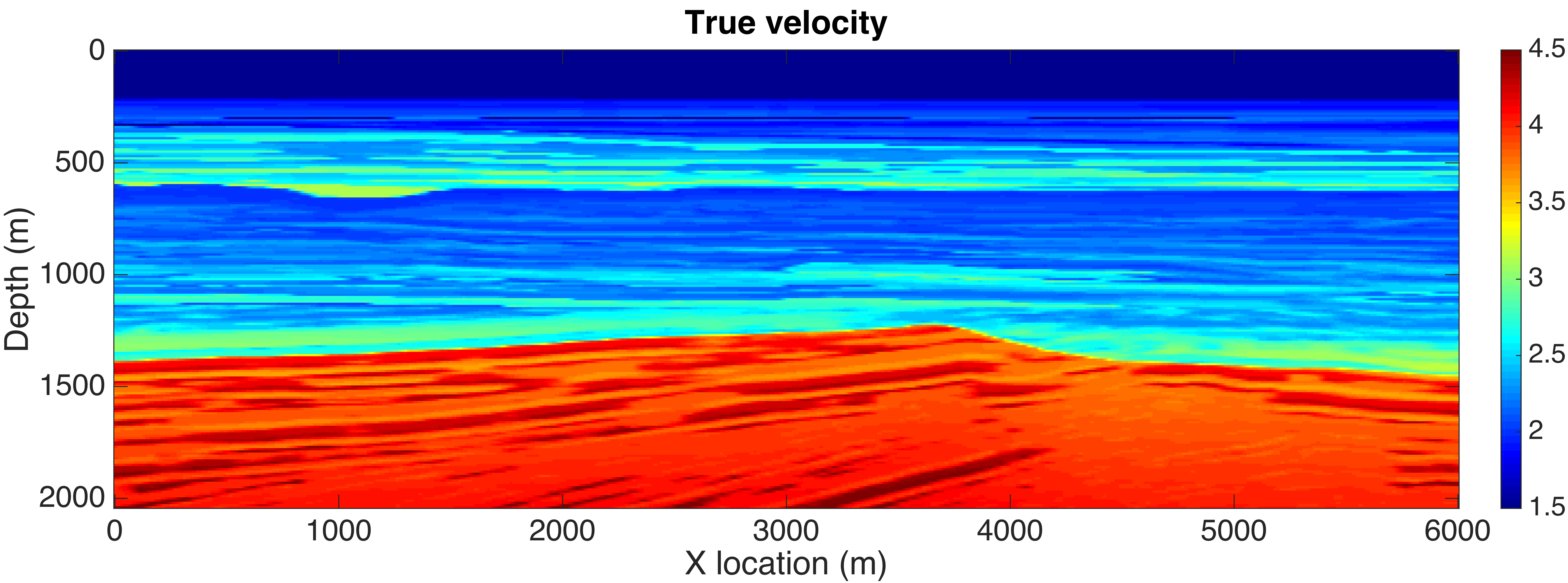
- Acoustic modelling & inversion

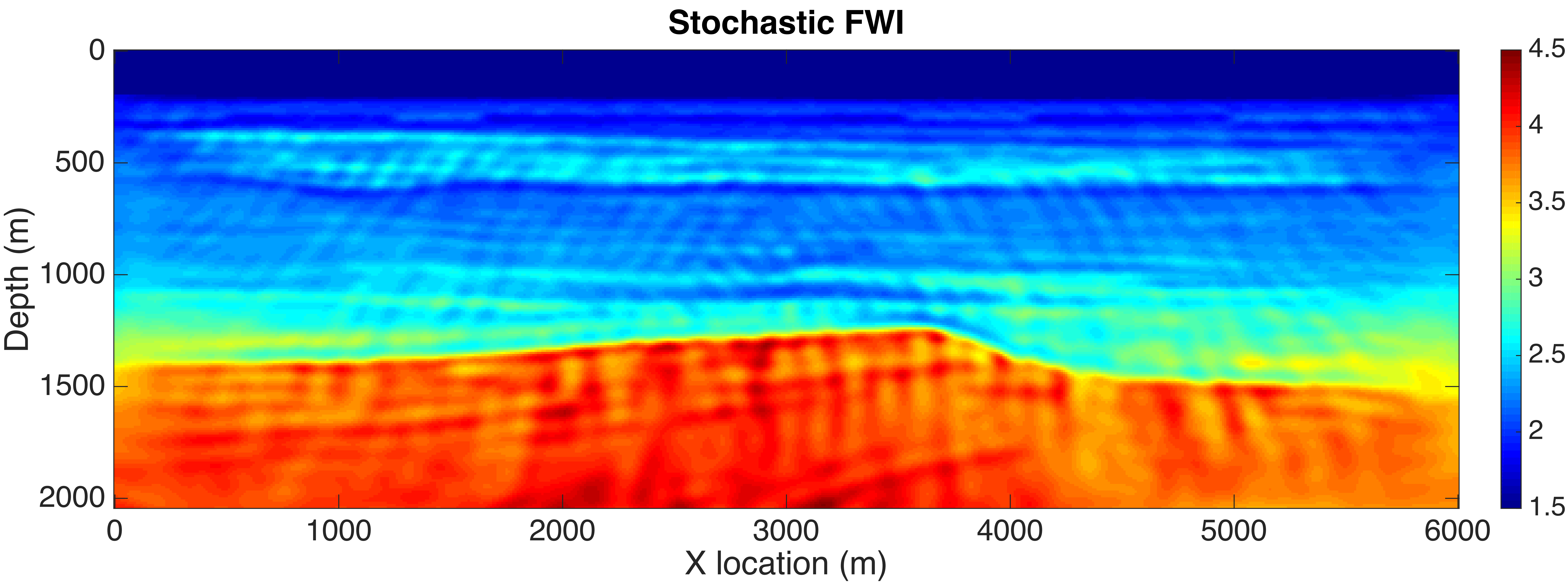
- **20 PQN iterations:**
  - bound constraints
  - minimum smoothness


True velocity
0
500
1000
1500
2000
Depth (m)
0
1000
2000
3000
4000
5000
6000
X location (m)
4.5
4
3.5
3
2.5
2
1.5

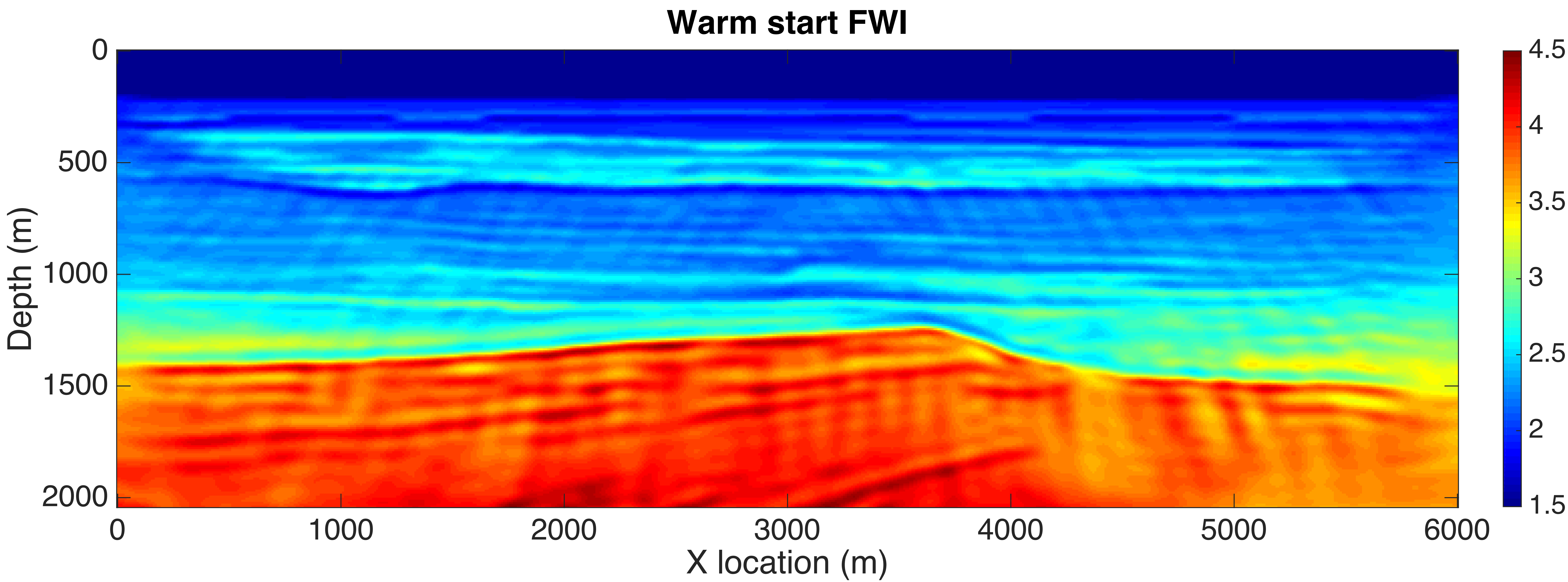

Initial model
0
500
1000
1500
2000
Depth (m)
0
1000
2000
3000
4000
5000
6000
X location (m)
4.5
4
3.5
3
2.5
2
1.5

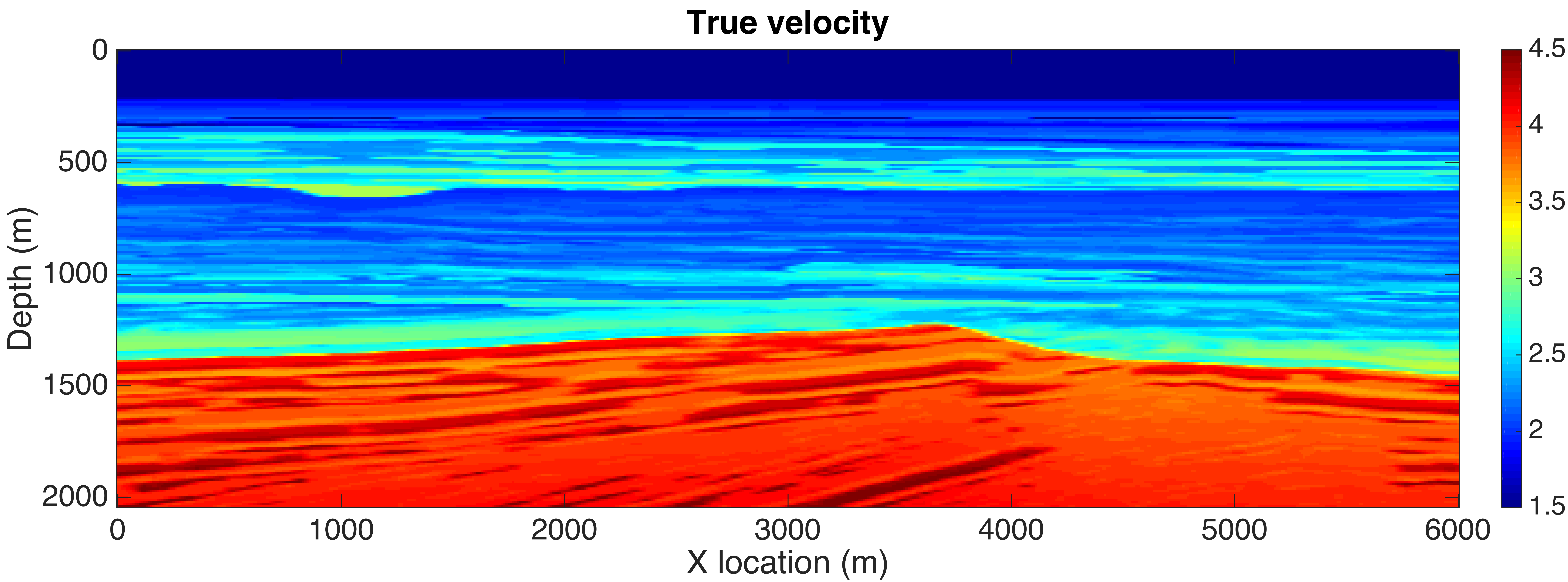

FWI
Depth (m)
0
500
1000
1500
2000
0
1000
2000
3000
4000
5000
6000
X location (m)
4.5
4
3.5
3
2.5
2
1.5


True velocity
Depth (m)
0
500
1000
1500
2000
X location (m)
0
1000
2000
3000
4000
5000
6000
4.5
4
3.5
3
2.5
2
1.5



Stochastic FWI
0
500
1000
1500
2000
Depth (m)
0
1000
2000
3000
4000
5000
6000
X location (m)
4.5
4
3.5
3
2.5
2
1.5


Warm start FWI
Depth (m)
0
500
1000
1500
2000
X location (m)
0
1000
2000
3000
4000
5000
6000
4.5
4
3.5
3
2.5
2
1.5


True velocity
Depth (m)
X location (m)
0
500
1000
1500
2000
0
1000
2000
3000
4000
5000
6000
4.5
4
3.5
3
2.5
2
1.5

SLIM

Conclusion

implicit extension of the model space

robustness increased

same or smaller computational/memory cost than FWI

easy to implement

Future work

Improve the choice of :

- the weights for the stack
- the length of the interval
- study convergence (stochastic optimization)

Explore limits of the robustness

Elastic/anisotropic

Combine with reflection FWI

Acknowledgements

Thank you for your attention !

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

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).