Rank minimization based seismic data processing & inversion

Rajiv Kumar

SLIM
University of British Columbia

Seismic data Interpolation

Rajiv Kumar

SLIM
University of British Columbia

SLIM

University of British Columbia

Motivation

- acquisition challenges
  - missing data
  - irregular acquisition grid
- fully sampled data
  - simultaneous shot based FWI & migration
  - estimation of primaries by sparse inversion & SRME
- regularization
  - imaging and inversion algorithm require equi-spaced grid
- exploit *low-rank* structure of seismic data
  - *SVD-free* matrix factorization

[Candes and Plan 2010, Oropeza and Sacchi 2011]

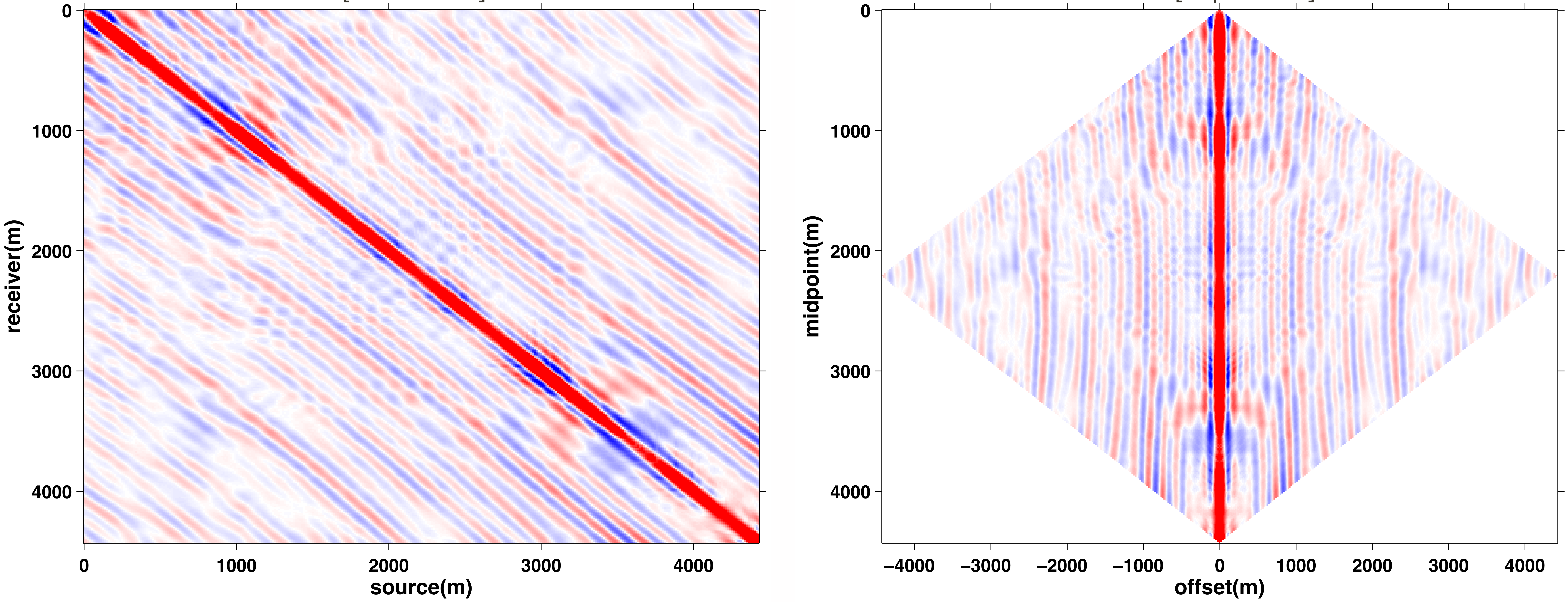
Matrix completion

- signal structure
  - *low rank/fast decay* of singular values
- sampling scheme
  - missing data *increase* rank in "transform domain"
- recovery using *rank penalization* scheme

Low-rank structure
2-D acquisition

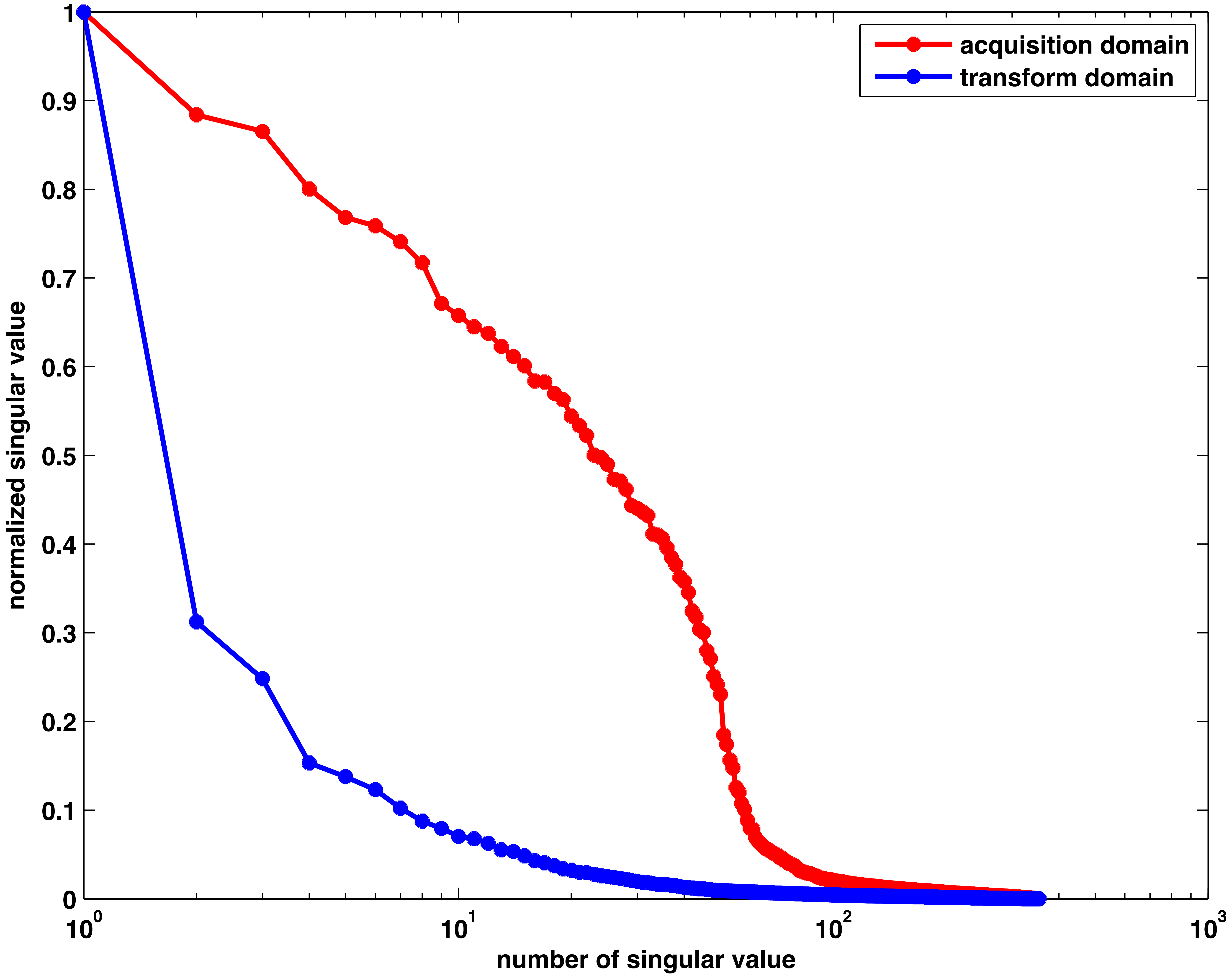
acquisition domain [source-receiver]

transform domain [midpoint-offset]



Singular value decay

2-D acquisition



Matrix completion

- signal structure
  - *low rank/fast decay* of singular values
- sampling scheme
  - missing data *increase* rank in "transform domain"
- recovery using *rank penalization* scheme

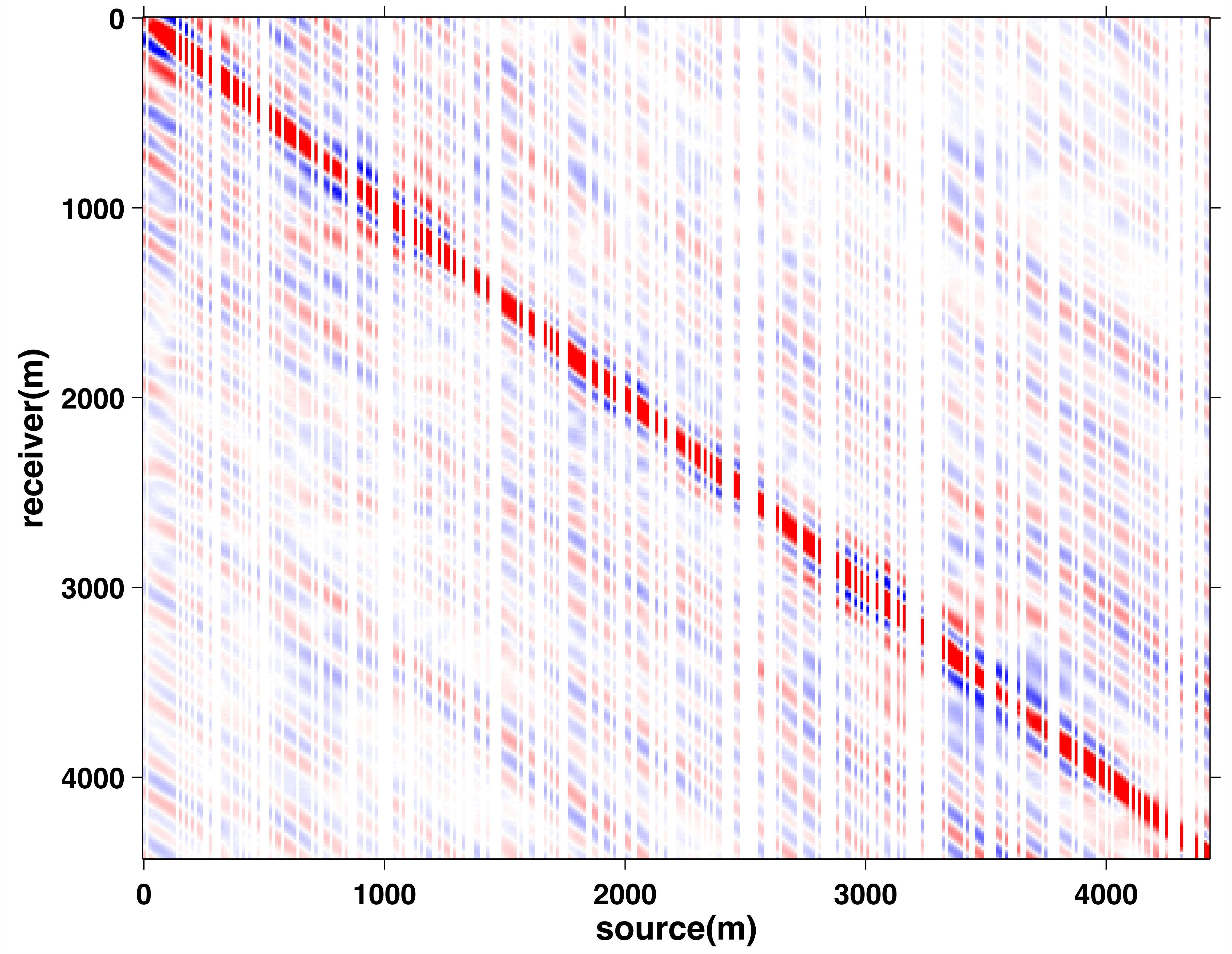
2-D acquisition

uniform-random sampling

SLIM

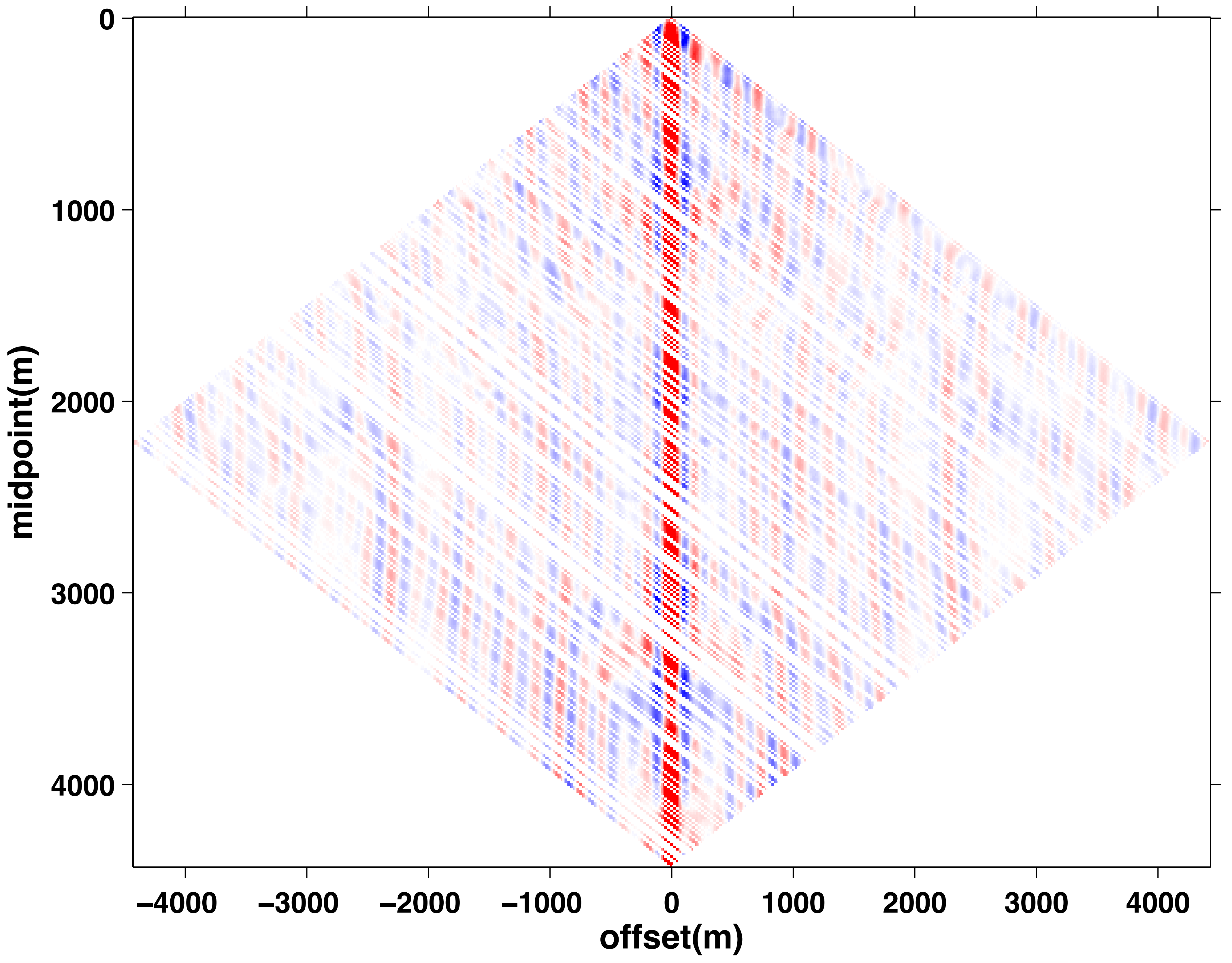
acquisition domain

missing columns *do not* increase rank

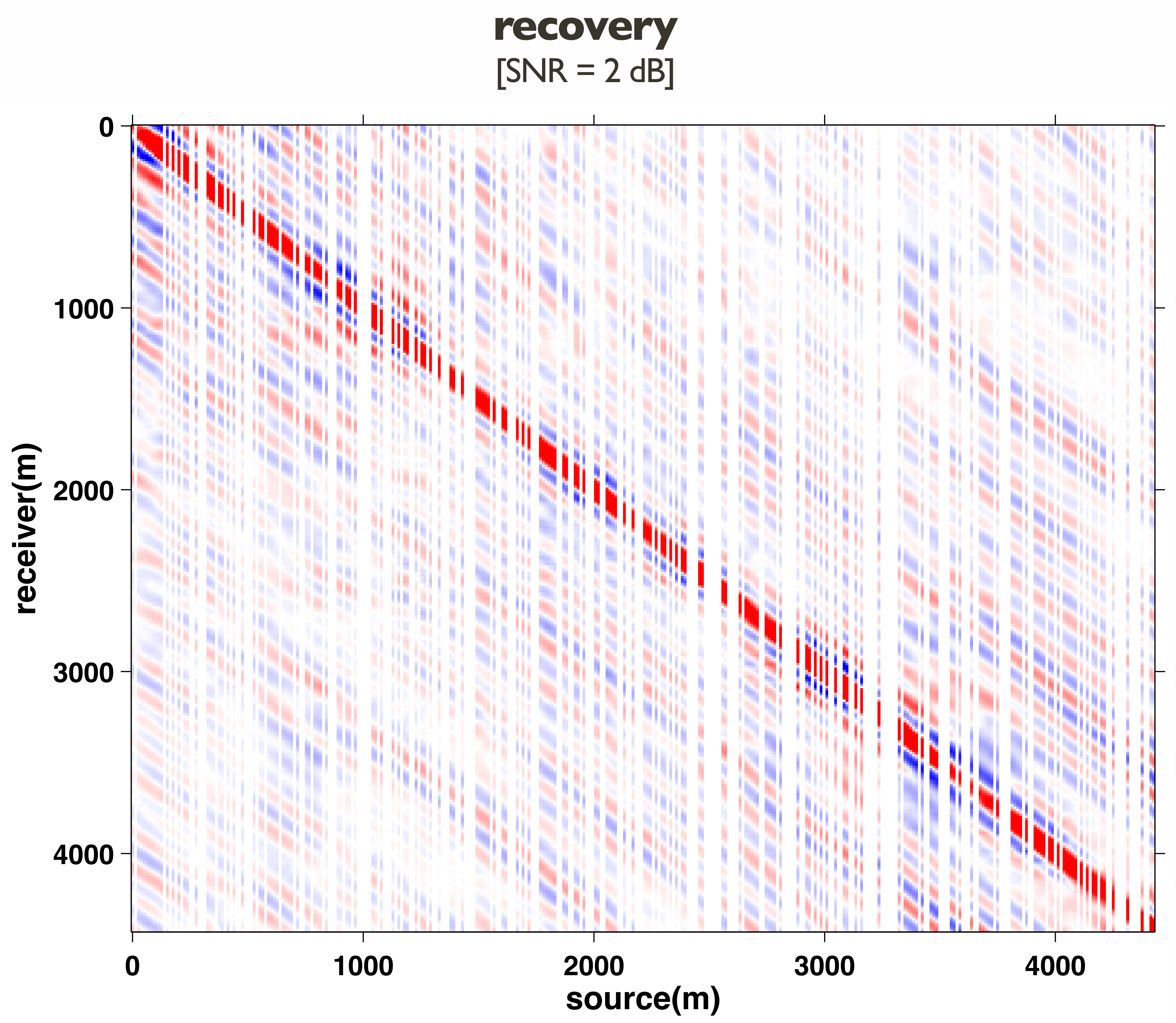


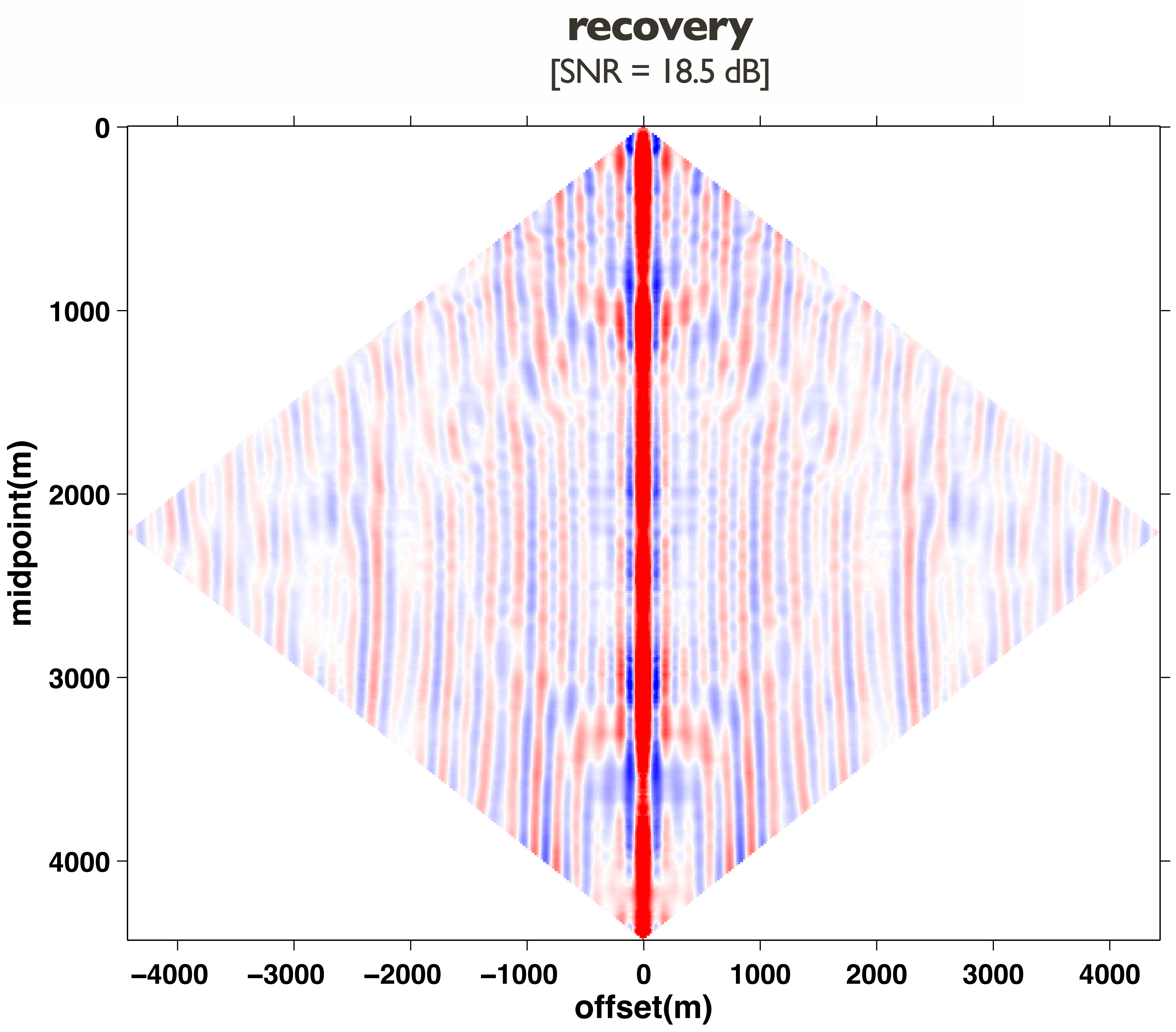
transform domain

missing columns *do* increase rank



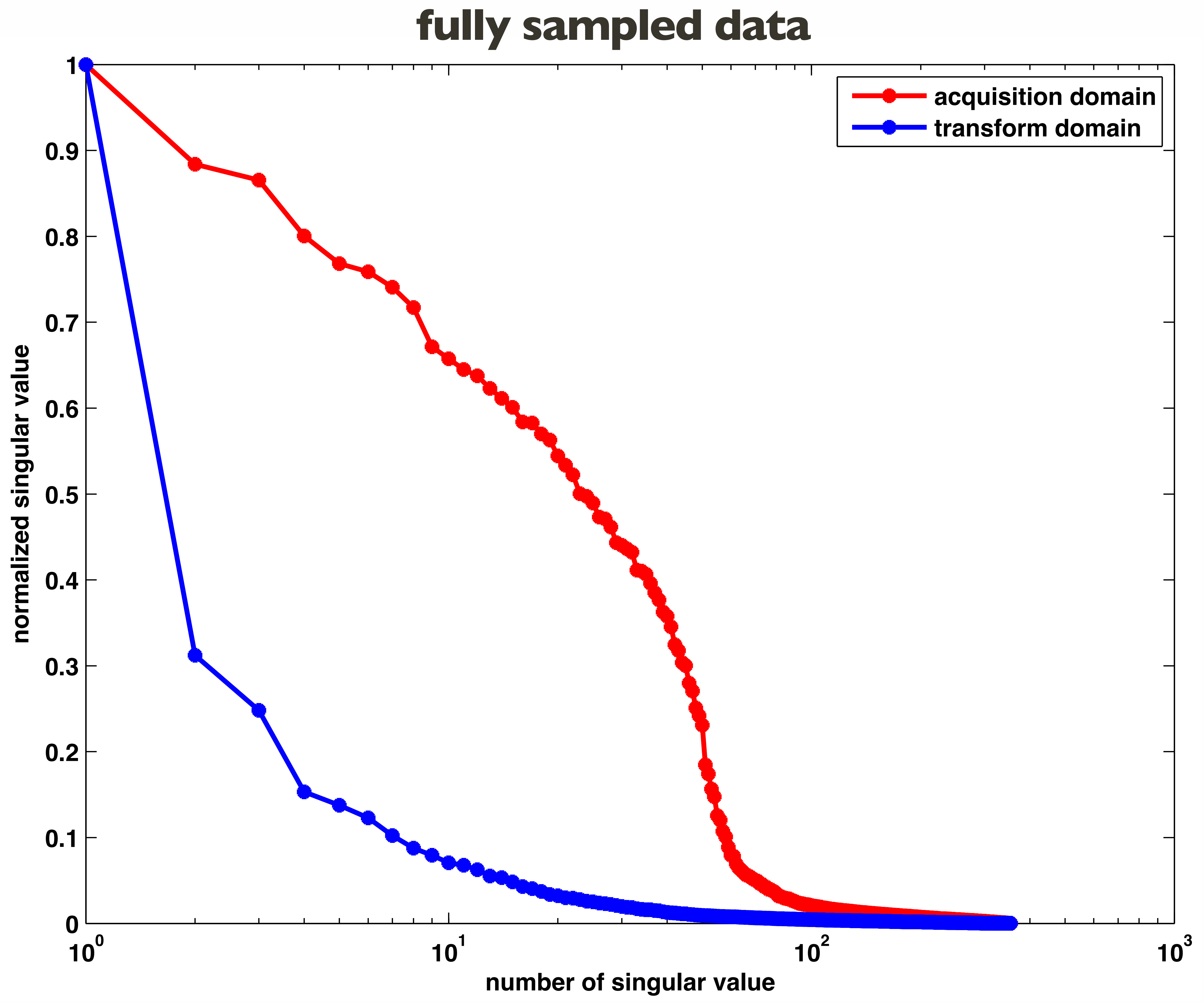
Low-rank interpolation

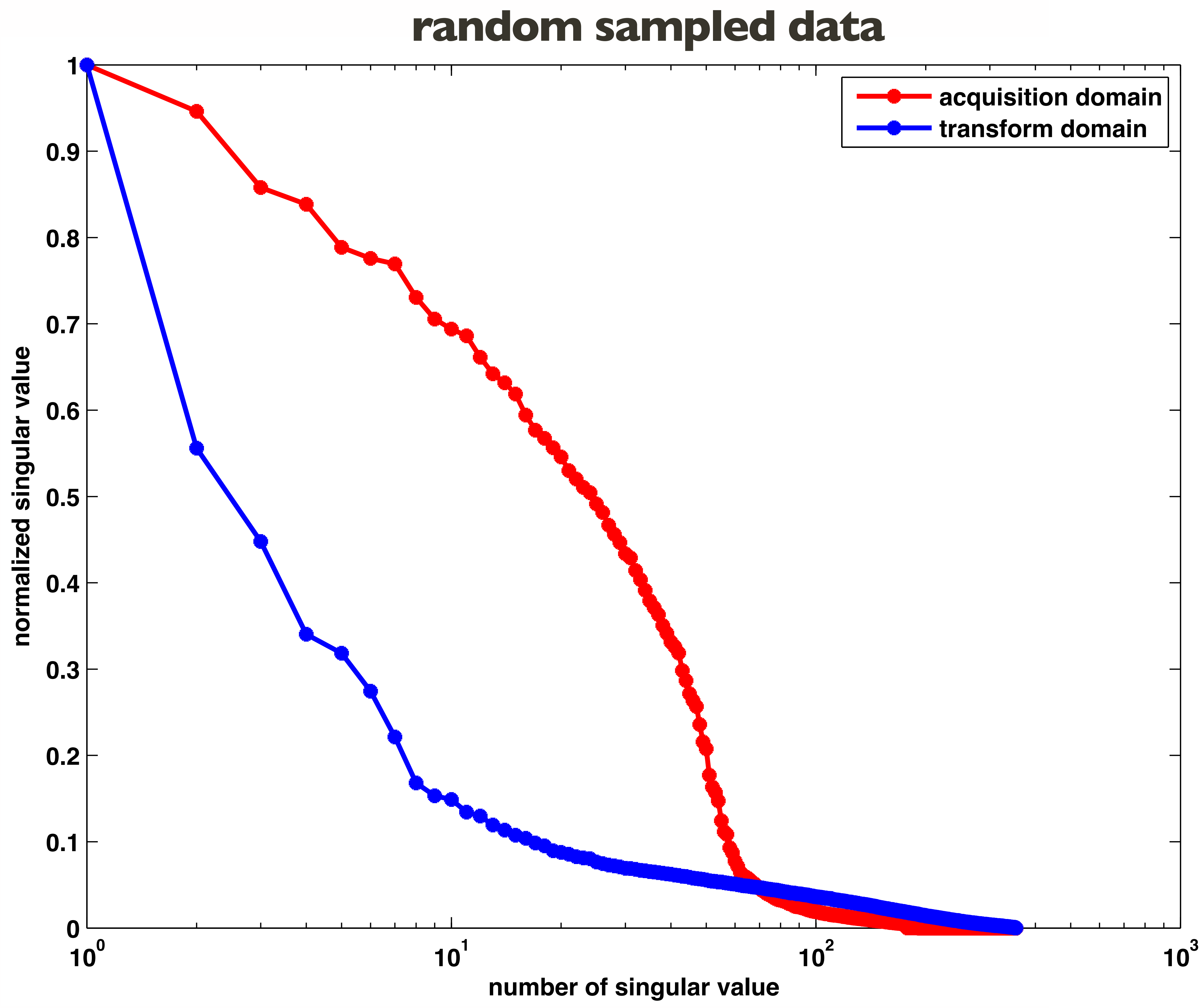




Randomized sampling

singular value decay





Observations

- sampling become *incoherent* in “transform” domain
- *slow decay* of singular values in “transform” domain

Matrix completion

- signal structure
  - *low rank/fast decay* of singular values
- sampling scheme
  - missing data *increase* rank in "transform domain"
- recovery using *rank penalization* scheme

Rank minimization

- given a set of measurements $\mathbf{b}$, aim is to solve

$$(BPDN_{\sigma}) \qquad \min_{\mathbf{X}} \quad \mathrm{rank}(\mathbf{X}) \quad \text{s.t.} \;\; ||\mathcal{A}(\mathbf{X}) - \mathbf{b}||_2^2 \leq \sigma$$

where $\mathrm{rank}(\mathbf{X}) =$ number of singular values of $\mathbf{X}$

- $\mathcal{A}$ is the transform-sampling operator defined as

$$\mathcal{A} = \mathbf{RM}\mathcal{S}^H$$

where

$\mathbf{R}$: restriction operator
$\mathbf{M}$: measurement operator
$\mathcal{S}^H$: transform operator

Rank minimization

- prohibitively *expensive*
  - do not know rank value in advance
  - search over all possible values of rank

- instead solve nuclear-norm minimization
  - convex relaxation of rank-minimization **[Recht et. al. 2010]**

[Recht et. al. 2010]

Nuclear-norm minimization

- we want to solve

$$(BPDN_\sigma) \qquad \min_{\mathbf{X}} ||\mathbf{X}||_* \quad \text{s.t.} \quad ||\mathcal{A}(\mathbf{X}) - \mathbf{b}||_2^2 \leq \sigma$$

where

$$\|\mathbf{X}\|_* = \sum_{i=1}^{m} \lambda_i = \|\lambda\|_1$$

where λ_i are the *singular* values

Challenges

- requires repeated application of *SVD* for projections

- expensive to compute for large system
  - curse of dimensionality

- can we exploit rank structure "*SVD* free"

[Rennie and Srebro 2005, Lee et. al. 2010, Recht and Re 2011]

Factorized formulation

$\mathbf{X} \in \mathbb{R}^{n \times m}$ = $\mathbf{L} \in \mathbb{R}^{n \times k}$ $\mathbf{R}^H \in \mathbb{R}^{k \times m}$

$$\mathbf{X} = \mathbf{L}\mathbf{R}^H$$

[Berg and Friedlander 2008, Aravkin et al. 2012b]

Factorized formulation

- reformulate $(BPDN_\sigma)$ formulation

$$\min_{\mathbf{L},\mathbf{R}} ||\mathbf{L}\mathbf{R}^H||_* \quad \text{s.t.} \quad ||\mathcal{A}(\mathbf{L}\mathbf{R}^H) - \mathbf{b}||_2^2 \leq \sigma$$

- approximately solve a series of $LASSO_\tau$ formulation

$$v(\tau) = \min_{\mathbf{L},\mathbf{R}} ||\mathcal{A}(\mathbf{L}\mathbf{R}^H) - \mathbf{b}||_2^2 \quad \text{s.t.} \quad \|\mathbf{L}\mathbf{R}^H\|_* \leq \tau$$

where τ is a rank regularization parameter

[Rennie and Srebro 2005]

Factorized formulation

- Upper-bound on nuclear norm is defined as

$$\|\mathbf{L}\mathbf{R}^H\|_* \leq \frac{1}{2}\left\|\begin{bmatrix}\mathbf{L}\\ \mathbf{R}\end{bmatrix}\right\|_F^2$$

where $\|\cdot\|_F^2$ is sum of squares of all entries

- choose k *explicitly* & avoid costly *SVD's*

Computational cost
with and without SVD

		50.0%		75.0%	
	σ	0.1	0.1	0.1	0.1
Matrix completion w/ SVD	SNR (dB)	17.3	18.3	11.6	11.5
	time (sec)	**812**	**937**	**790**	**765**
Matrix completion w/o SVD	SNR (dB)	17.6	18.4	12.6	13.3
	time (sec)	**8**	**10**	**8**	**7**

SLIM

Upcoming paper
Check https://www.slim.eos.ubc.ca soon!

SLIM

Computational cost

matrix completion v/s curvelet-based methods

		50.0%		75.0%	
	σ	0.1	0.1	0.1	0.1
Matrix completion w/ SVD	SNR (dB)	17.3	18.3	11.6	11.5
	time (sec)	812.0	937.0	790.0	765.0
Matrix completion w/o SVD	SNR (dB)	17.6	18.4	12.6	13.1
	time (sec)	**8**	**10**	**8**	**7**
Curvelet-based sparsity promotion	SNR (dB)	17.4	18.6	12.5	12.8
	time (sec)	**879**	**989**	**817**	**1010**

Observation

matrix completion v/s curvelet-based methods

	Low-rank	Curvelet
computational time	O(*minutes*)	O(*hours*)
storage	$k \times (n+m)$	$8 \times nm$

SLIM

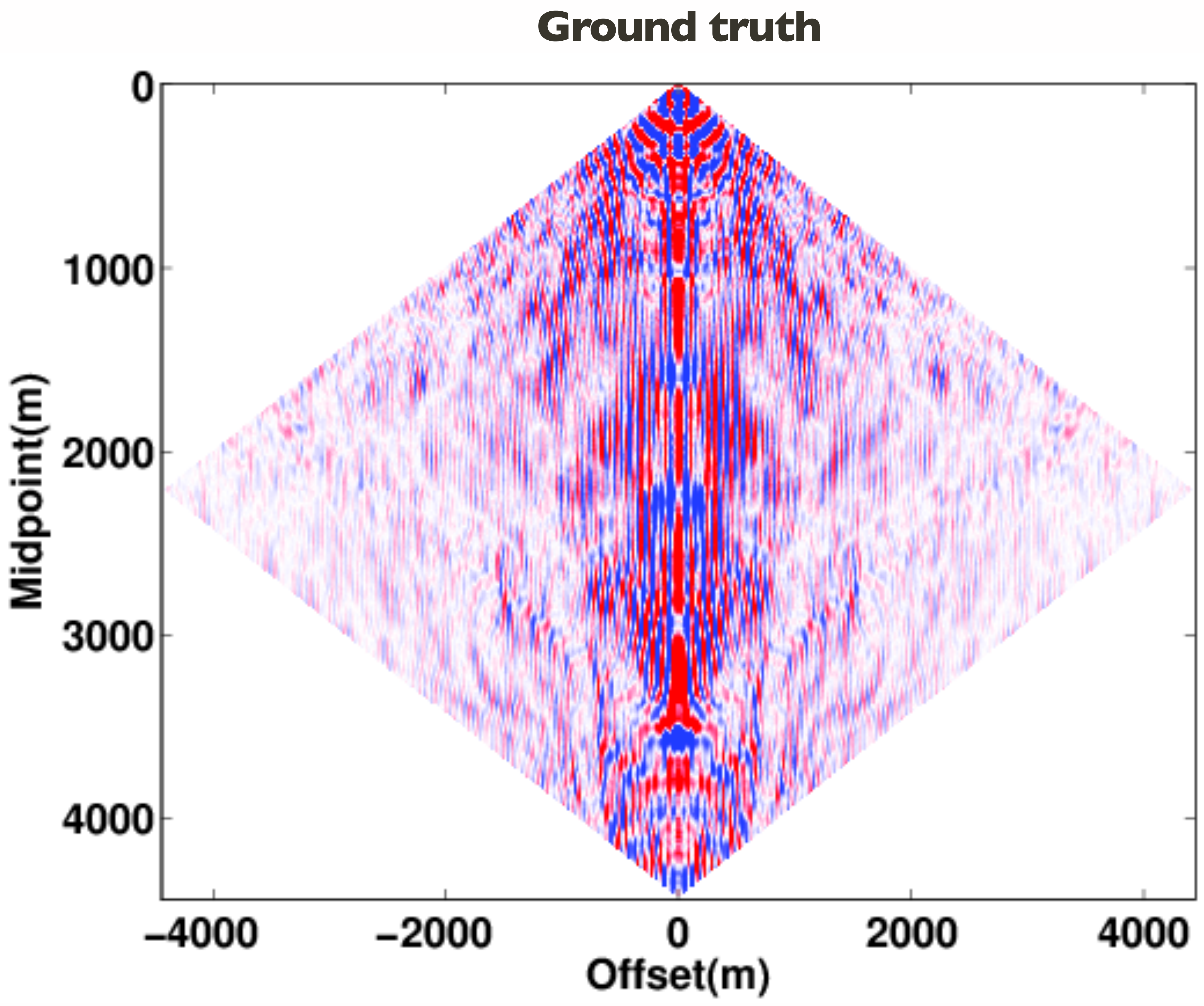
Take-away message

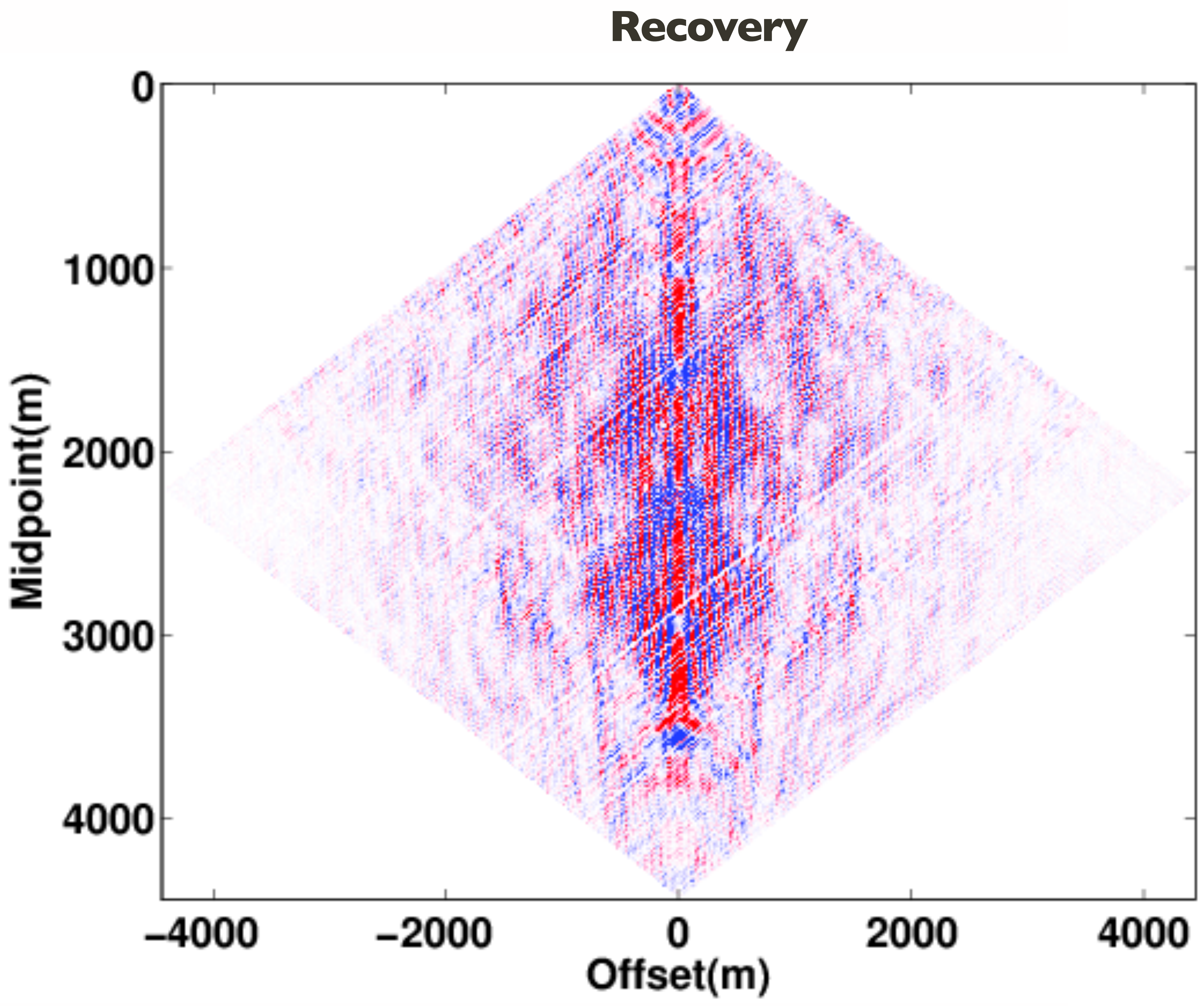
- can avoid “*SVD*”
- faster compare to curvelet-based sparsity promotion techniques
- memory efficient compare to curvelet-based techniques

Regularization

- unstructured acquisition grid
- imaging and inversion algorithm
  - regularly sampled data
- binning
  - does not preserve the data-structure

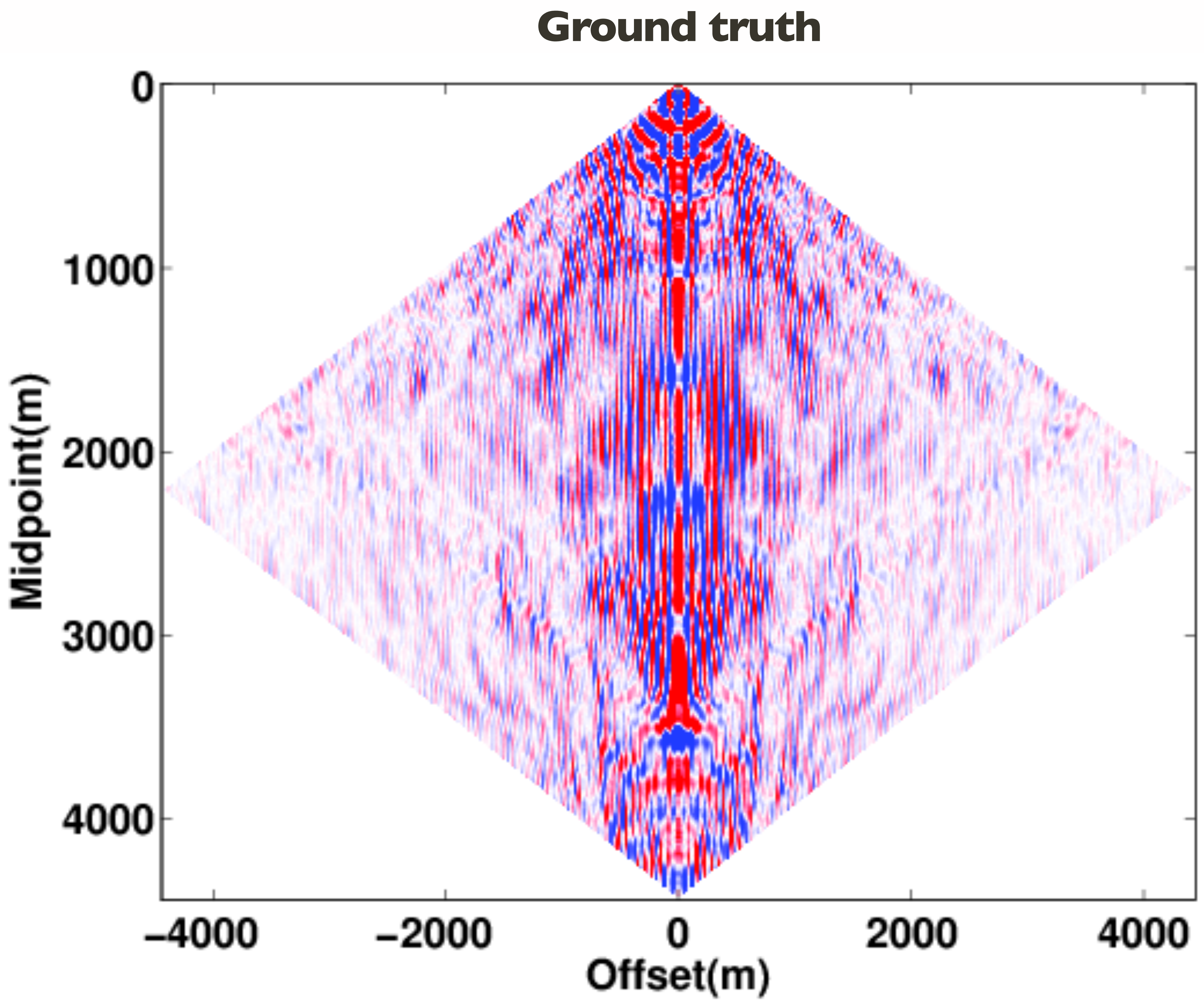
Low-rank structure
binning, midpoint-offset domain

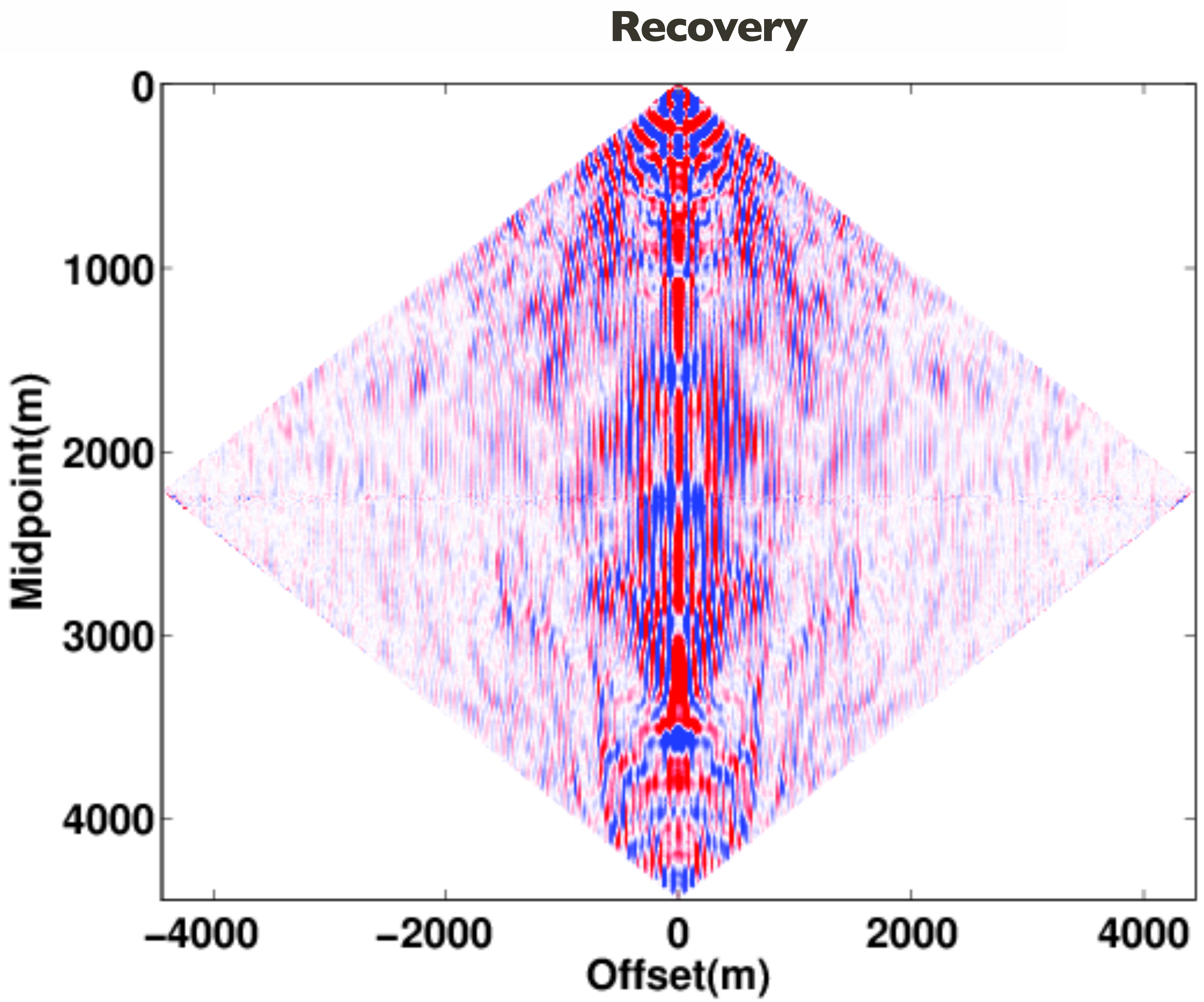




Regularization

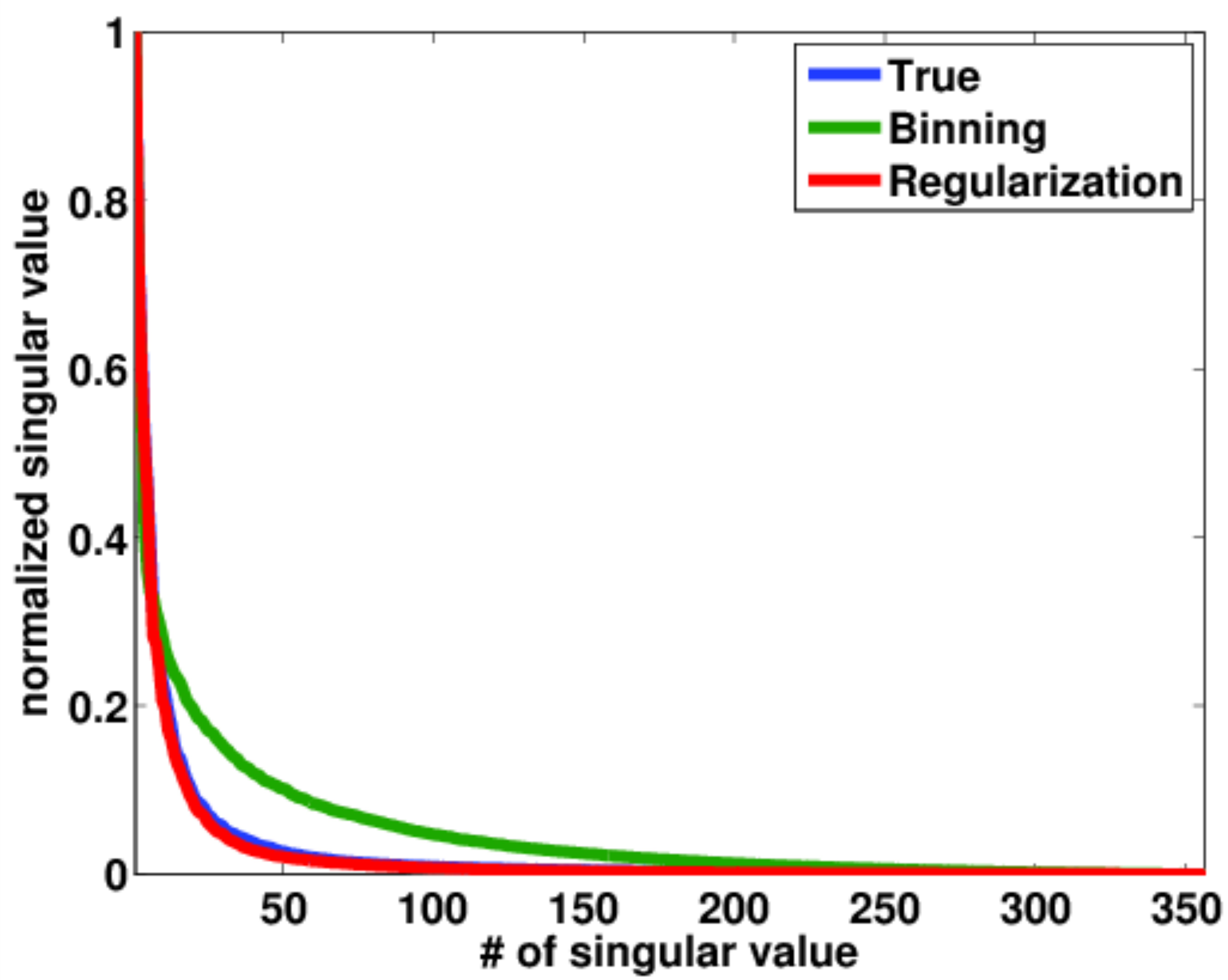
matrix completion, midpoint-offset domain





Singular value decay

regularization v/s binning



Methodology

matrix completion

- given a regularization operator $\mathbf{N} : \mathbb{C}^{n\times m} \rightarrow \mathbb{C}^{n\times m} \quad \text{so that} \quad \mathbf{N}(\mathbf{X}_r) = (\mathbf{X}_{ir})\,,$ transform-sampling operator is redefine as

$$\mathcal{A} = \mathbf{RMN}^H \mathcal{S}^H$$

where

$\mathbf{R}$: restriction operator
$\mathbf{M}$: measurement operator
$\mathbf{N}^H$: regularization operator
$\mathcal{S}^H$: transform operator

Theorem

matrix completion

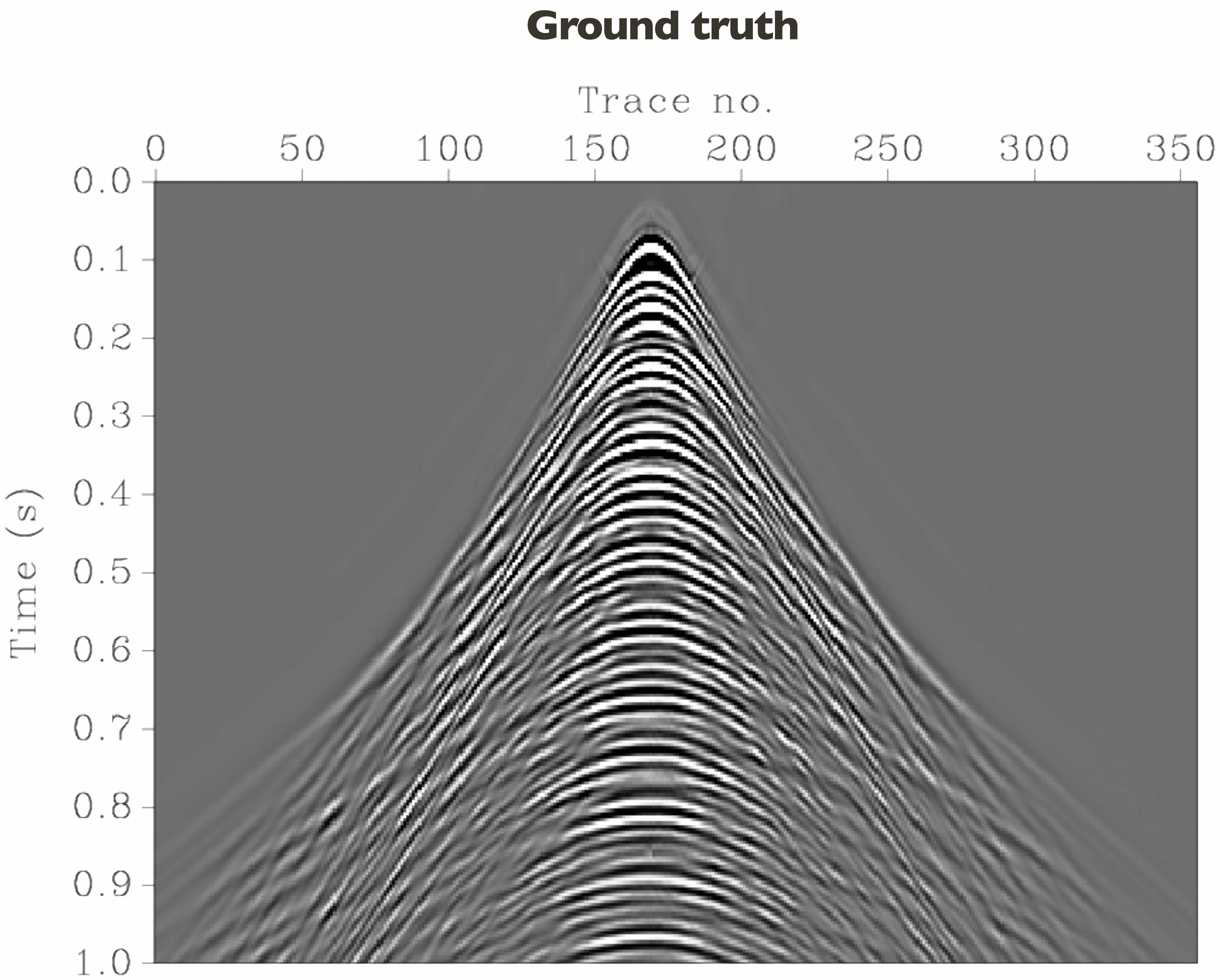
Let $\mathbf{X_r} \in \mathbb{C}^{n\times m}, \hat{\mathbf{X}}_r \in \mathcal{S}$ and $\mathbf{b} = \mathbf{RM}(\mathbf{X}_{ir}) + e$ with $\|e\| \leq \eta$. Let $\tilde{\mathbf{X}}$ be the solution of $\text{BPDN}_\sigma,$ then

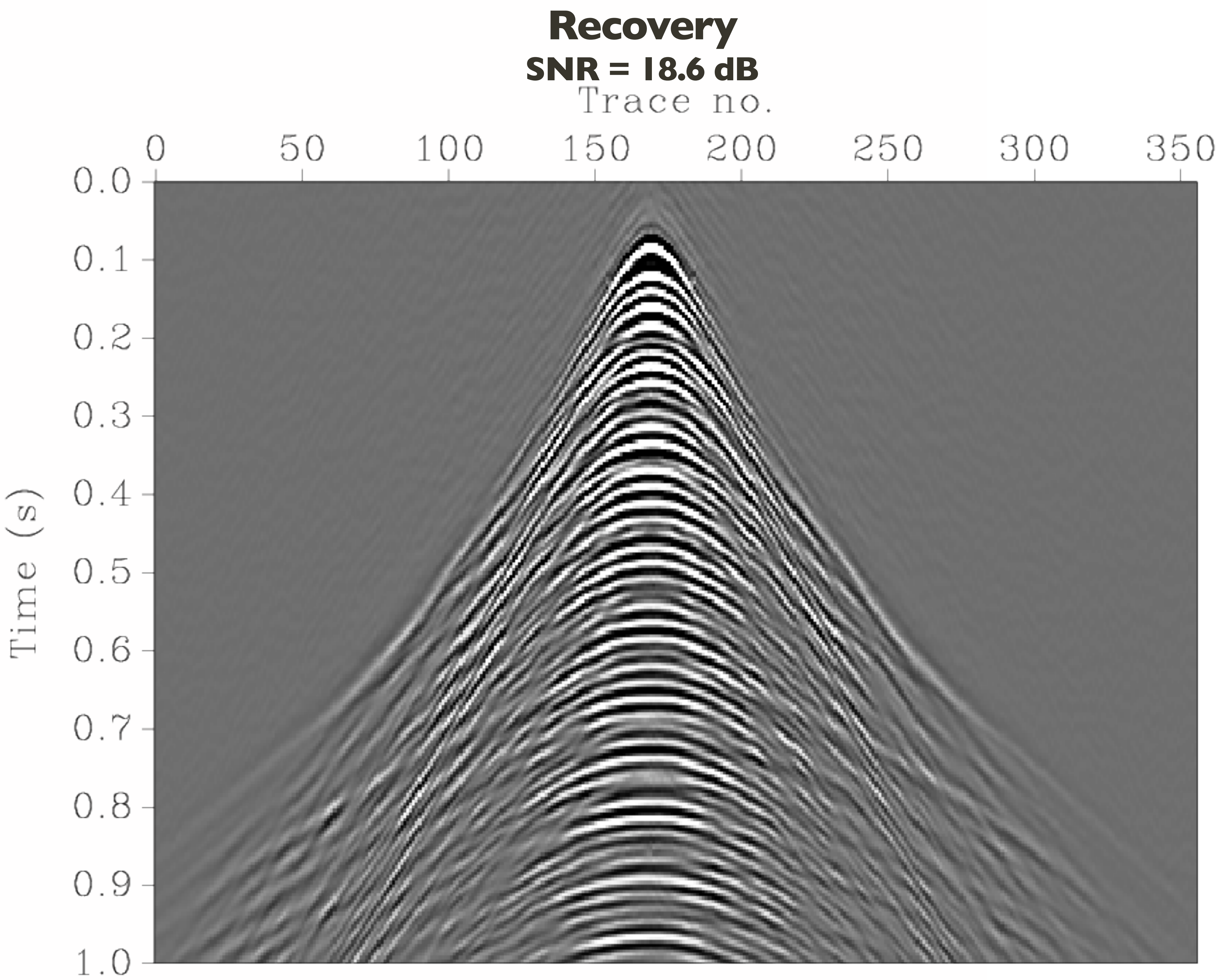
$$\|\mathcal{S}(\mathbf{X}_r - \tilde{\mathbf{X}})\| \leq \underbrace{\frac{C_1}{\sqrt{k}} \sum_{j=k+1}^{l} \sigma_j(\hat{\mathbf{X}}_r)}_{\text{interpolation error}} + \underbrace{(\frac{C_1}{\sqrt{k}} B_2 + 1)\|P\|_F}_{\text{regularization error}} + \underbrace{C_2\eta}_{\text{noise}}$$

where

$$P = \mathbf{N}^{-1}(\mathbf{X}_{ir}) - \mathbf{X}_r$$
$$l = min\{n, m\}$$
$$B_2 = (1 - \frac{k}{l})\sqrt{l}$$
$$C_1 \text{ and } C_2 > 0$$

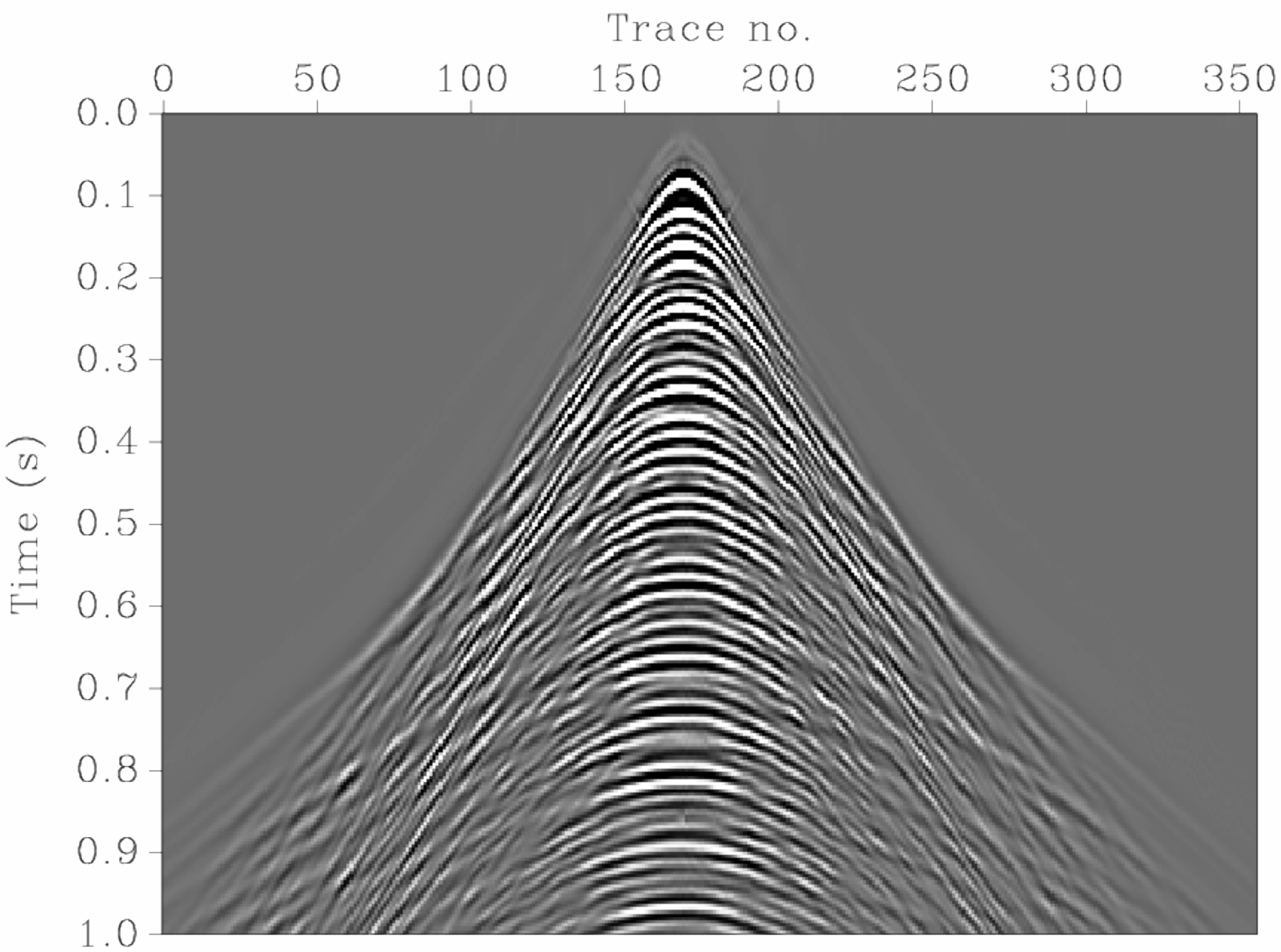
Regularization
matrix completion



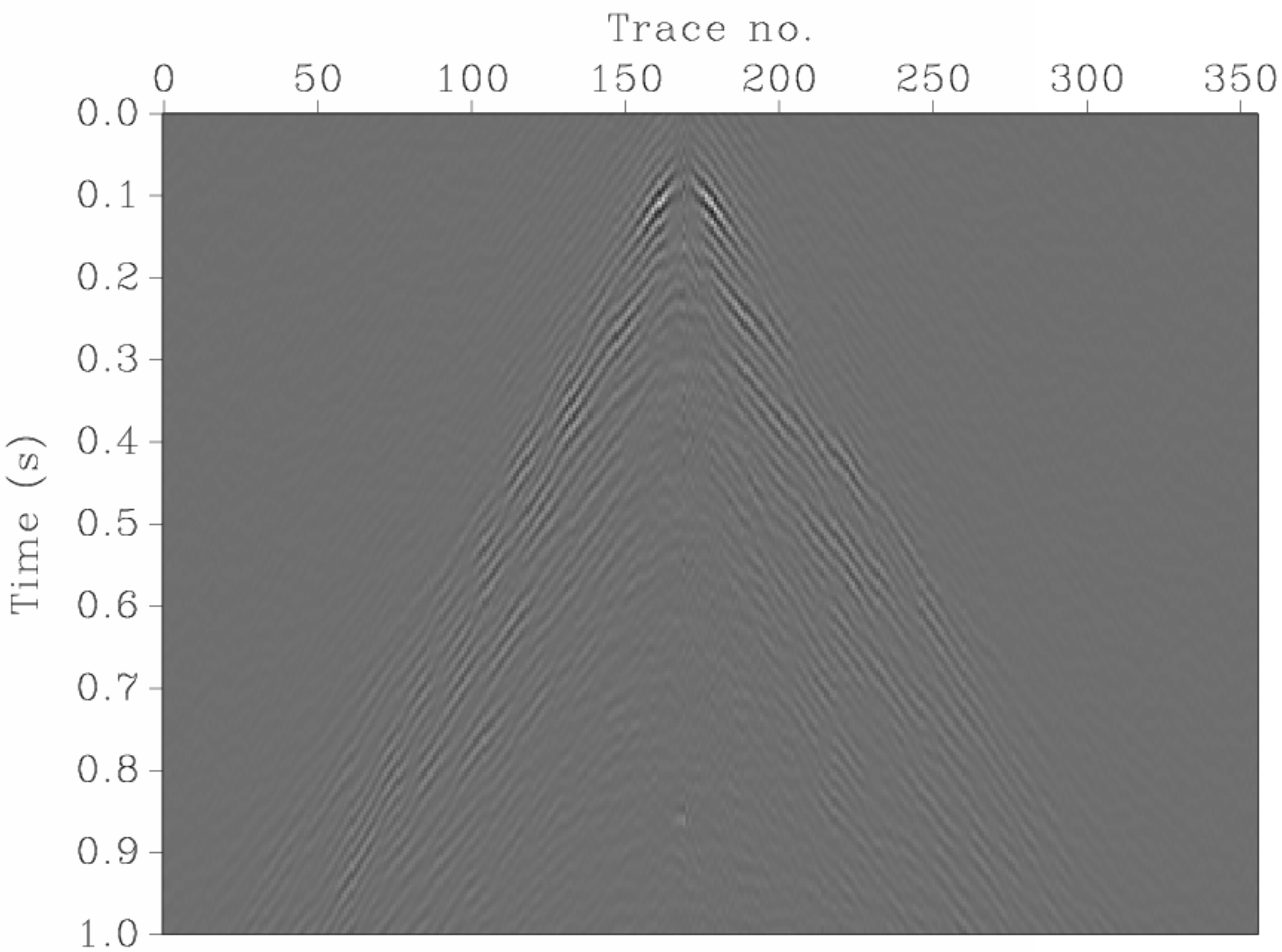


Regularization
matrix completion

Ground truth

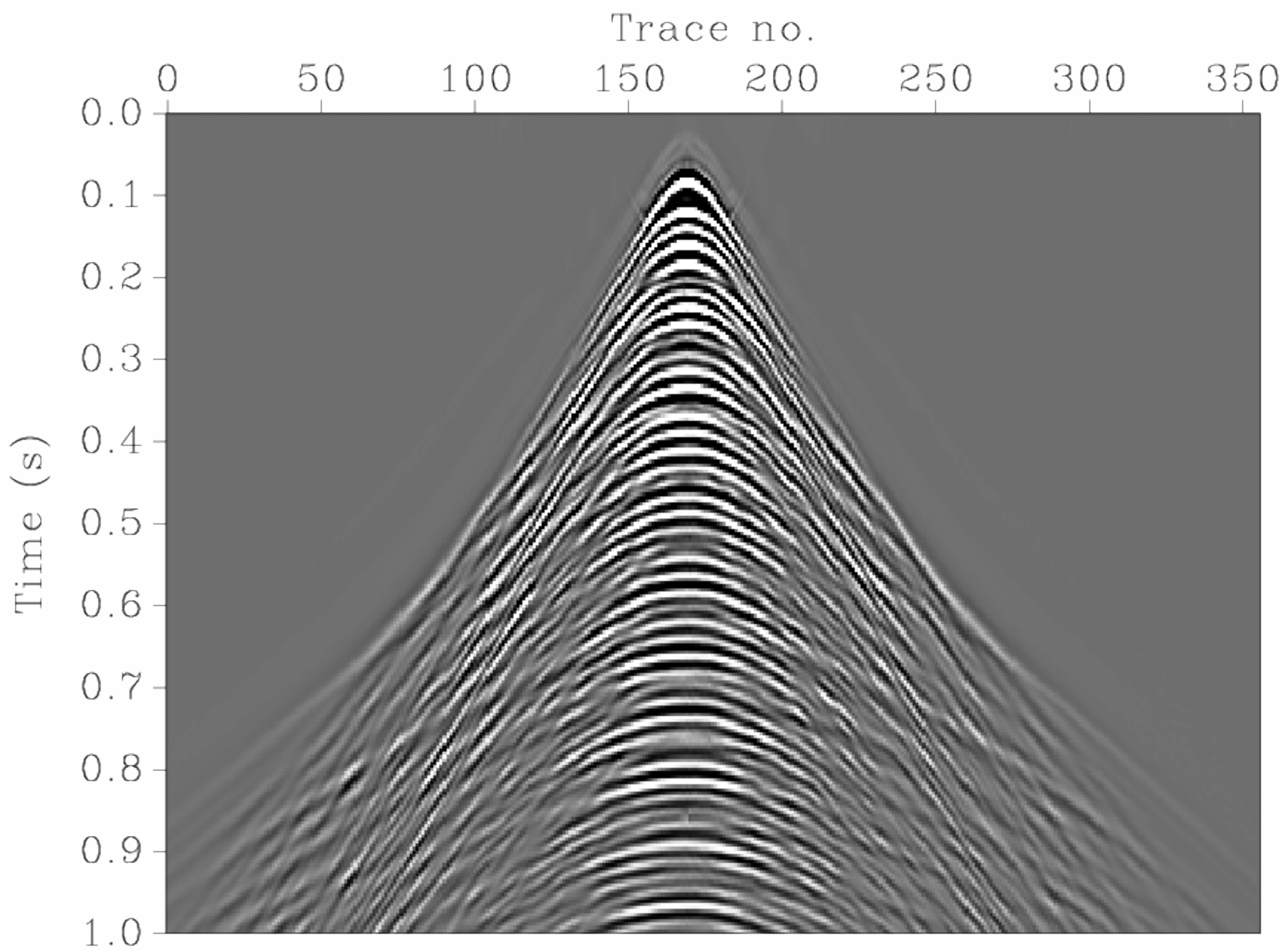


Residual

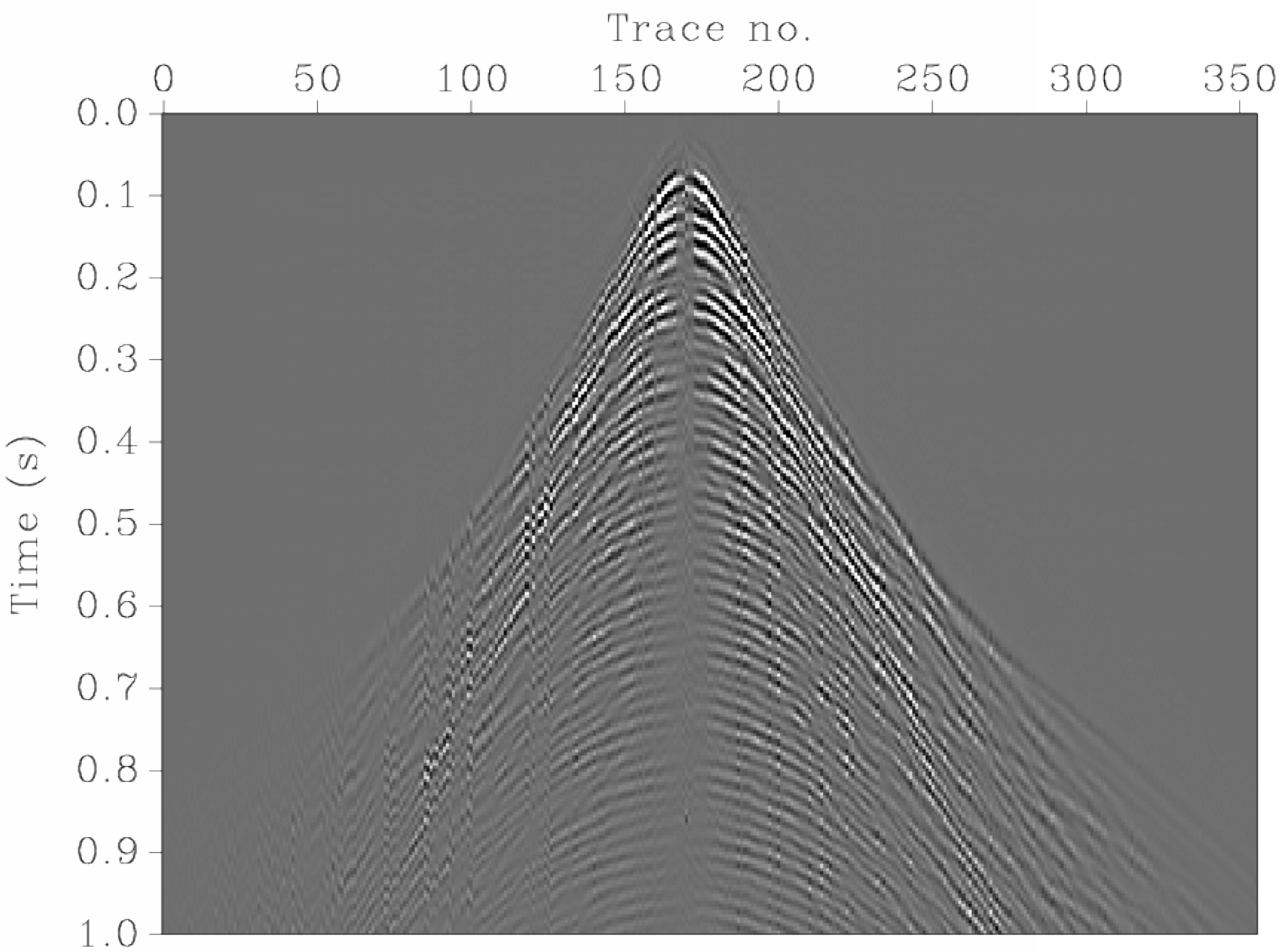


Regularization
binning

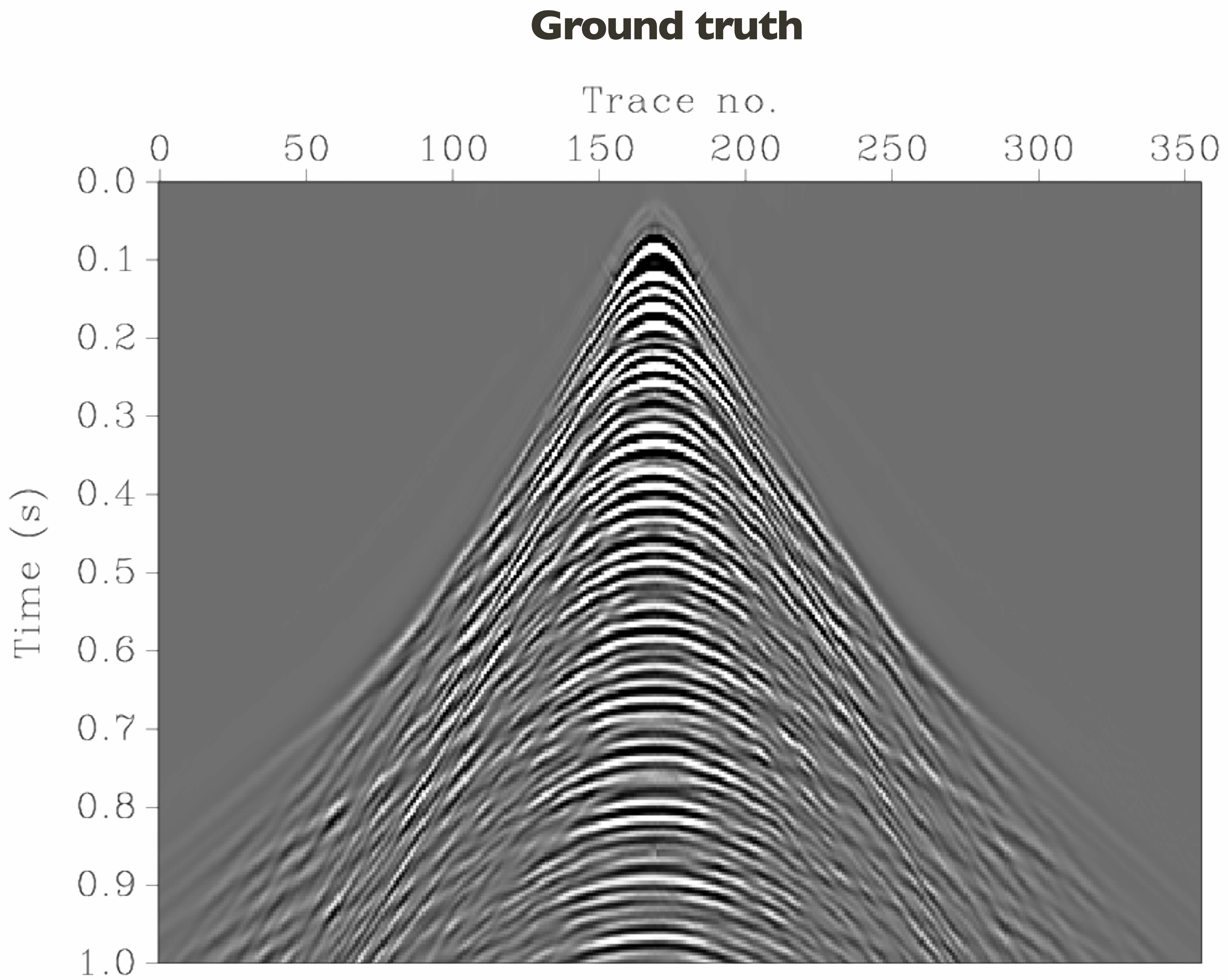
Ground truth

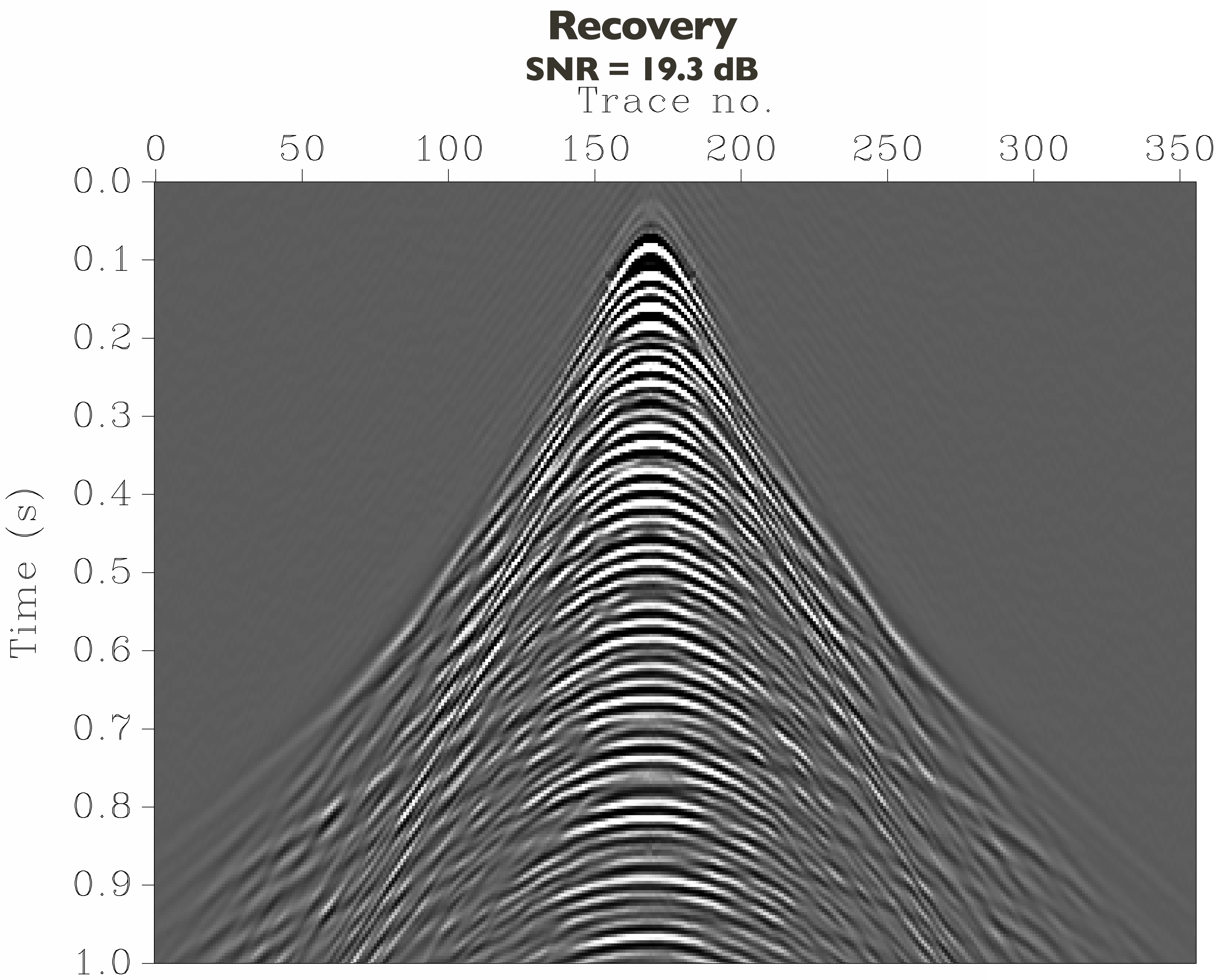


Residual



Regularization & Interpolation
matrix completion

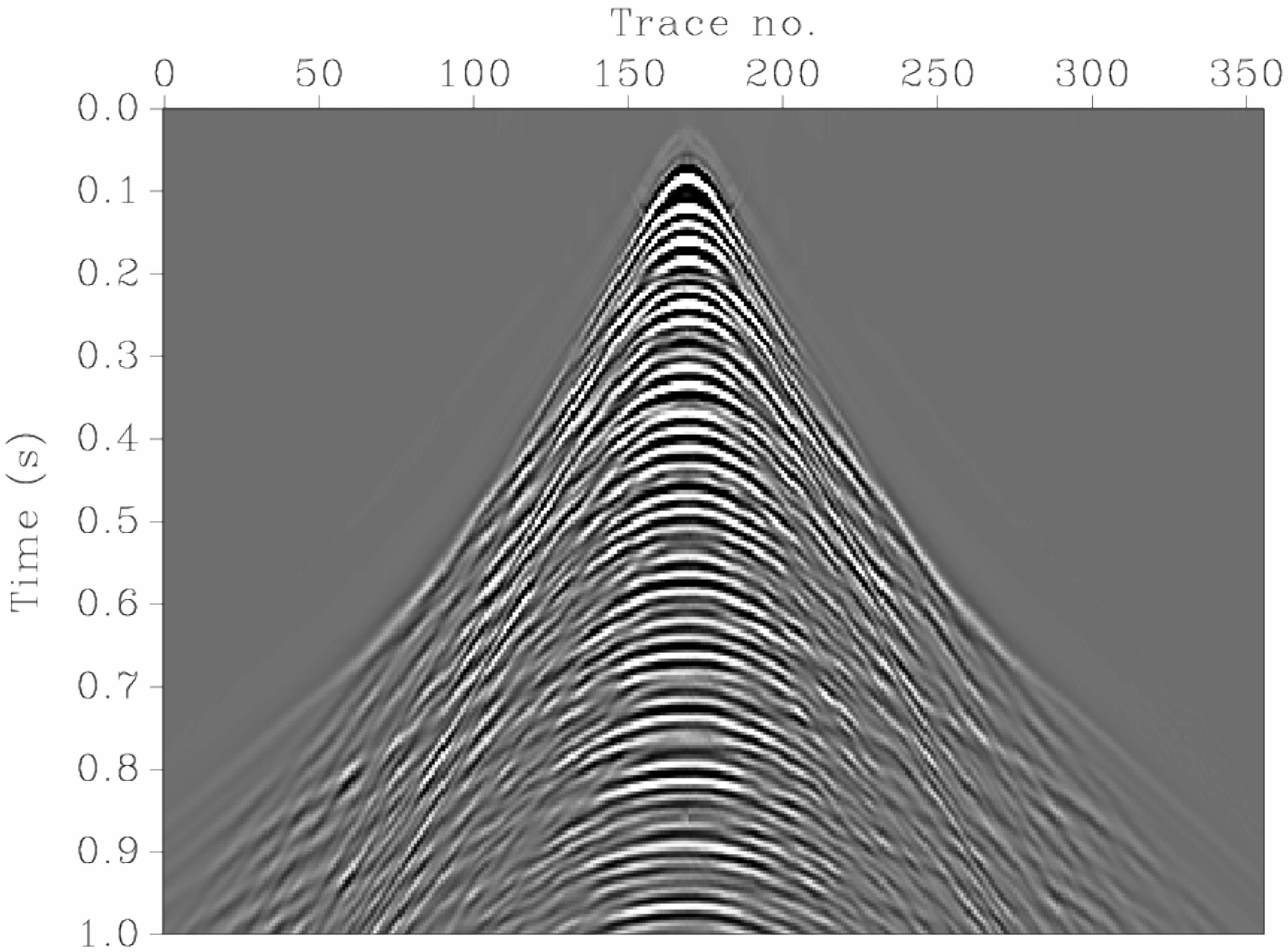




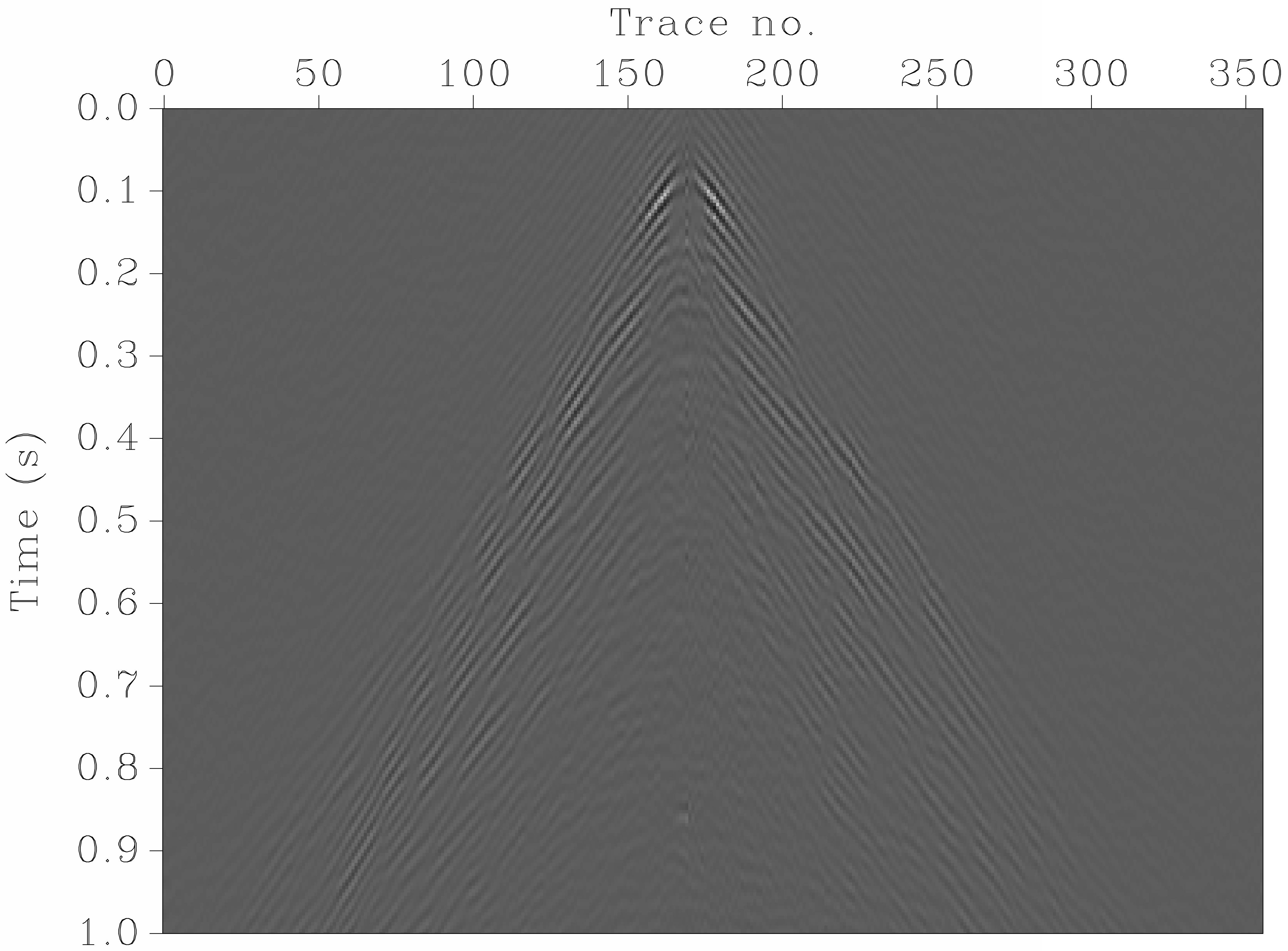
Regularization & Interpolation

matrix completion

Ground truth



Residual



Conclusion

- matrix factorization allows SVD-free low-rank methods that work fast on large data
- reconstruction quality is as good as curvelet-based techniques but computationally more feasible then curvelet
- matrix-factorization promise more compact representation

Source separation for simultaneous towed-streamer acquisition - sparsity vs. rank

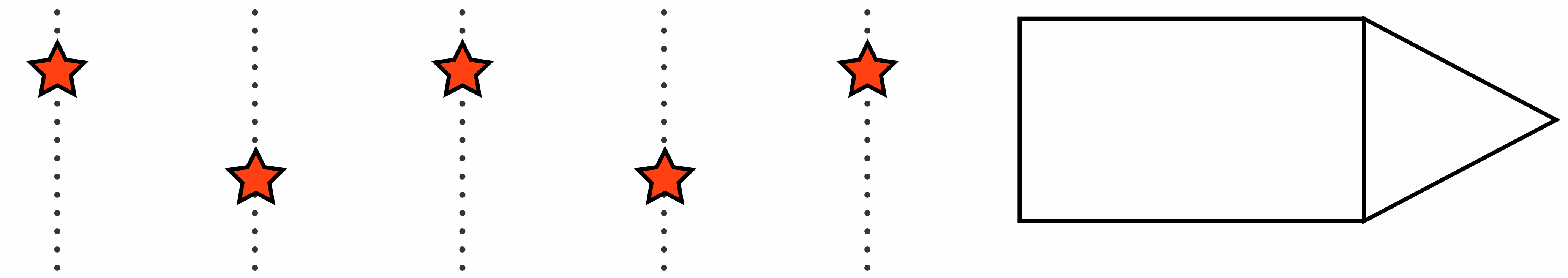
Collaborators: Haneet Wason and Felix Herrmann

SLIM

University of British Columbia

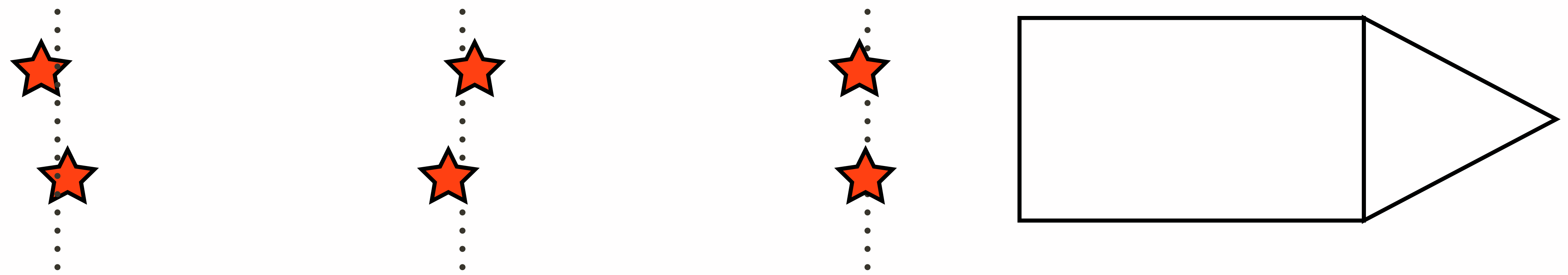
Periodic vs. jittered marine acquisition

periodically sampled spatial grid



almost periodically sampled spatial grid

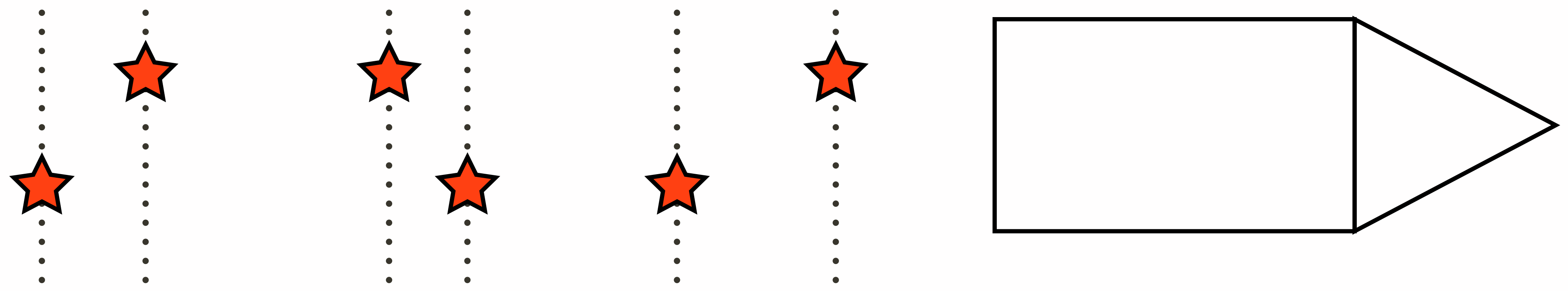
(over/under acquisition)



randomly jittered sampled spatial grid

(Time-jittered acquisition)

[Wason and Herrmann, 2013]
[Mansour et. al., 2012]

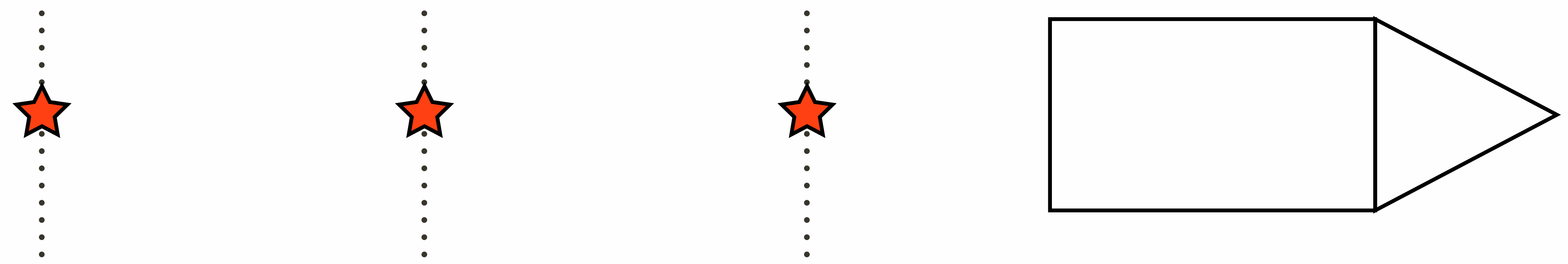


Conventional marine acquisition

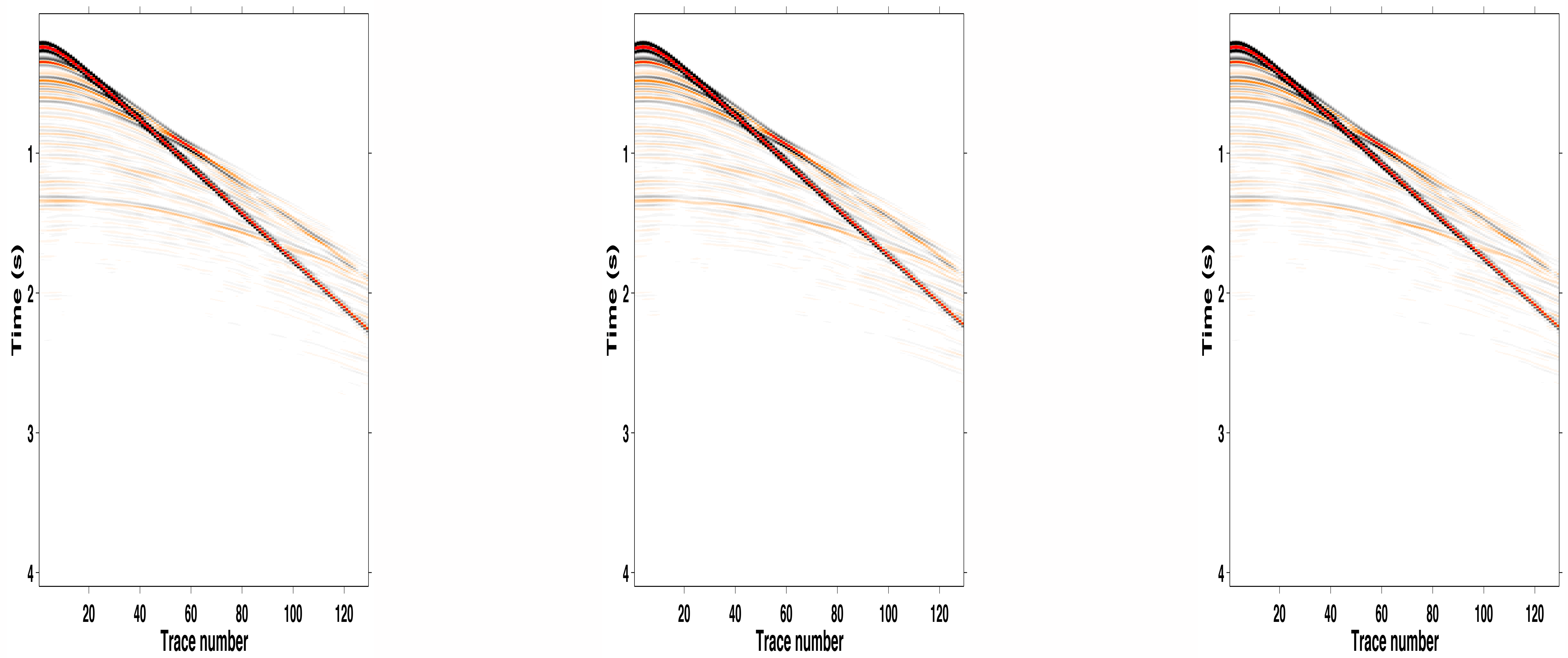
source depth = 6 m

shot 1 shot 2 shot 3

periodically sampled spatial grid →



shot 1 shot 2 shot 3



Blended/Simultaneous marine acquisition

[over/under acquisition]

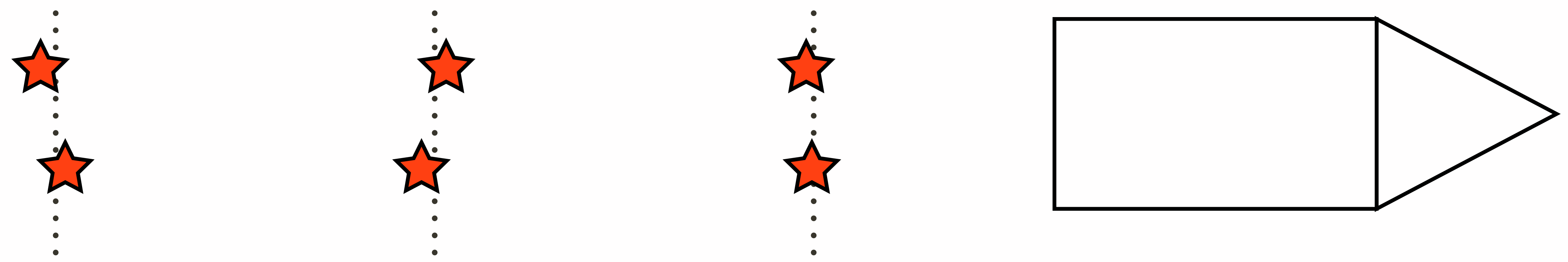
source1 depth = 6 m

source2 depth = 12 m

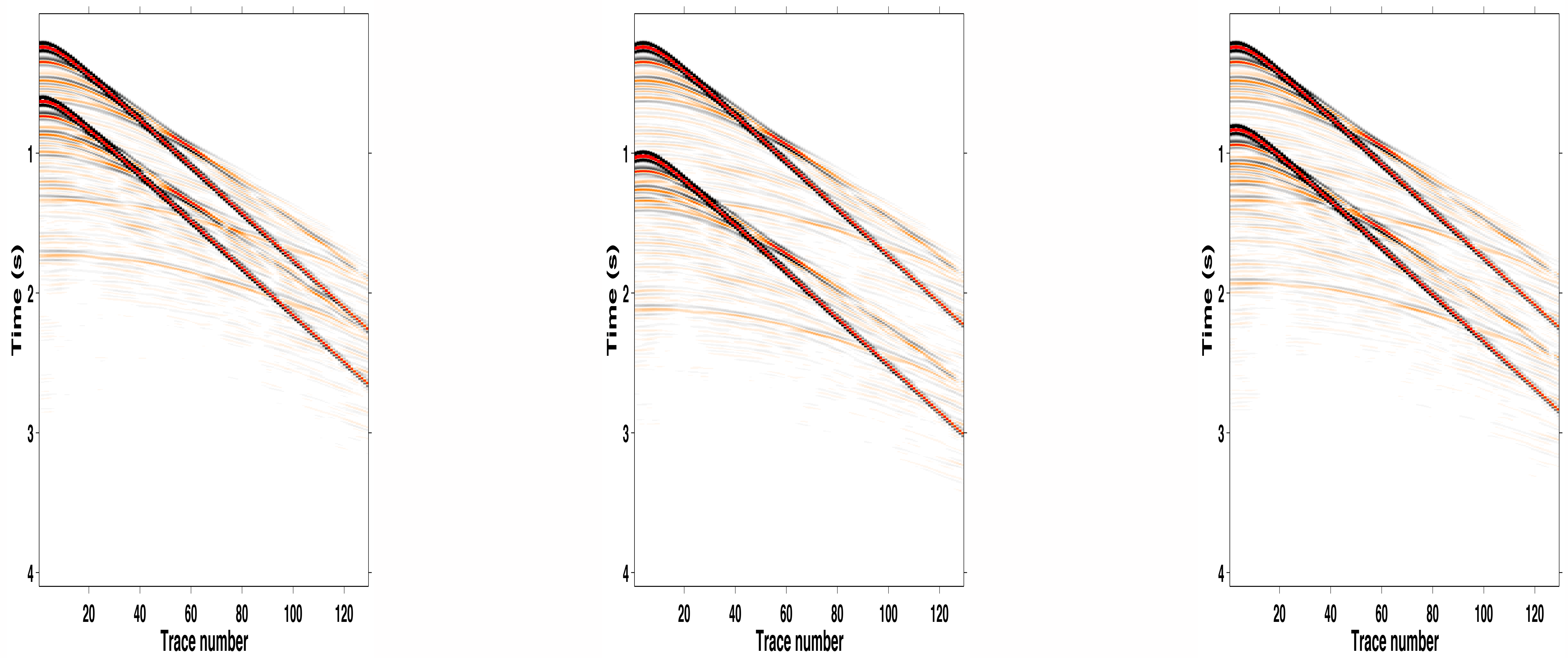
shot 1 shot 2 shot 3

periodically sampled spatial grid →
(almost)

shot-time randomness - ***LOW***



shot 1 shot 2 shot 3



Challenges

- Source separation (or *deblending*)
  - recover individual datasets
- Shot-time randomness
  - low

[Candès and Donoho, 2000; Herrmann, 2008]

Compressed sensing

Successful sampling & reconstruction scheme

- exploit *structure* via *sparsifying* transform
  - *fast decay* of "transform domain" coefficients
- sampling
  - randomly blended data *decreases* sparsity in "transform domain"
- optimization
  - via *sparsity-promotion*

[Candès and Plan, 2010, Oropeza and Sacchi, 2011]

SLIM

Matrix completion

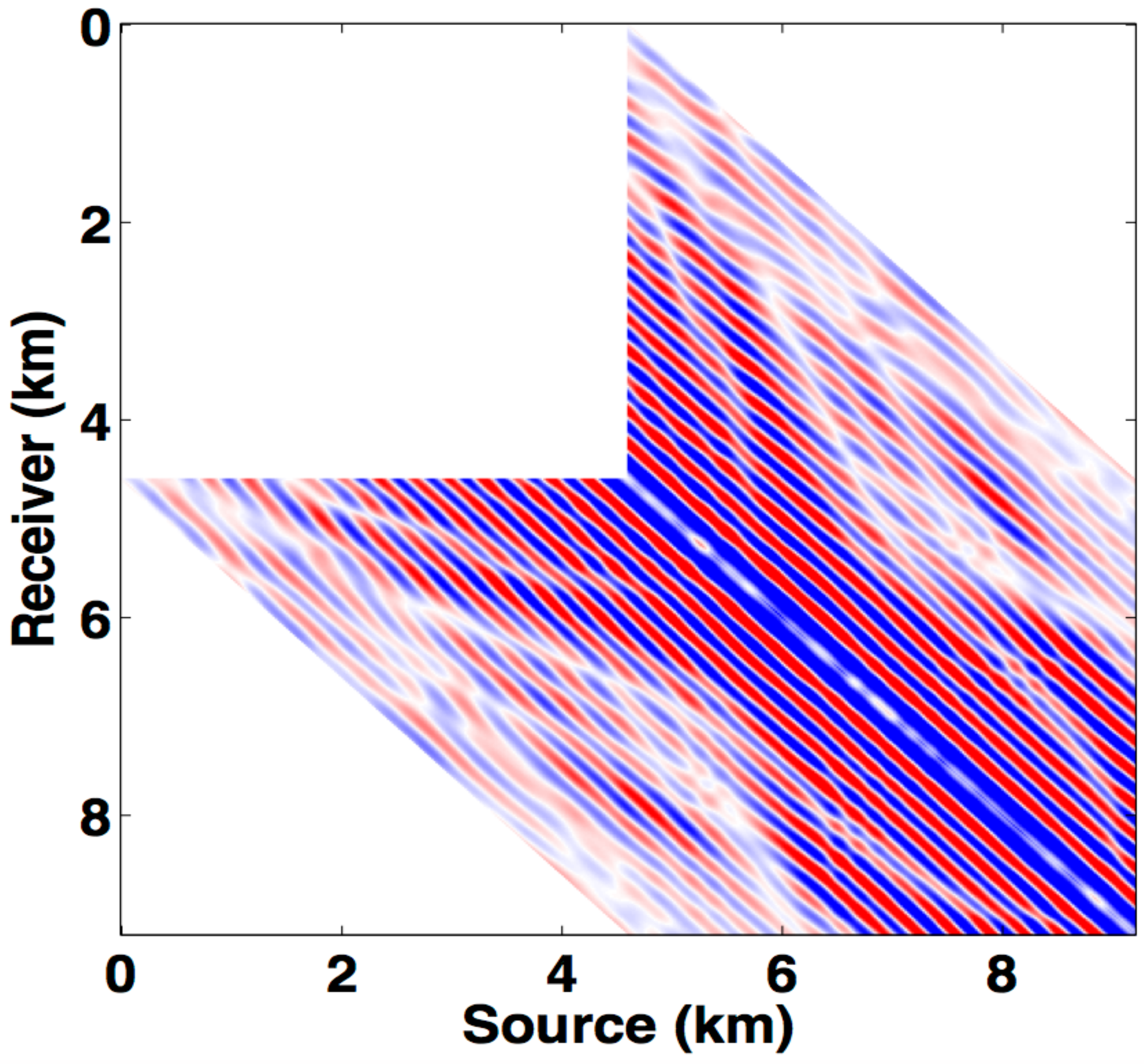
Successful reconstruction scheme

- exploit *structure*
  - *low-rank / fast decay* of singular values
- sampling
  - randomly blended data *increases* rank in "transform domain"
- optimization
  - via *rank-minimization (nuclear norm-minimization)*

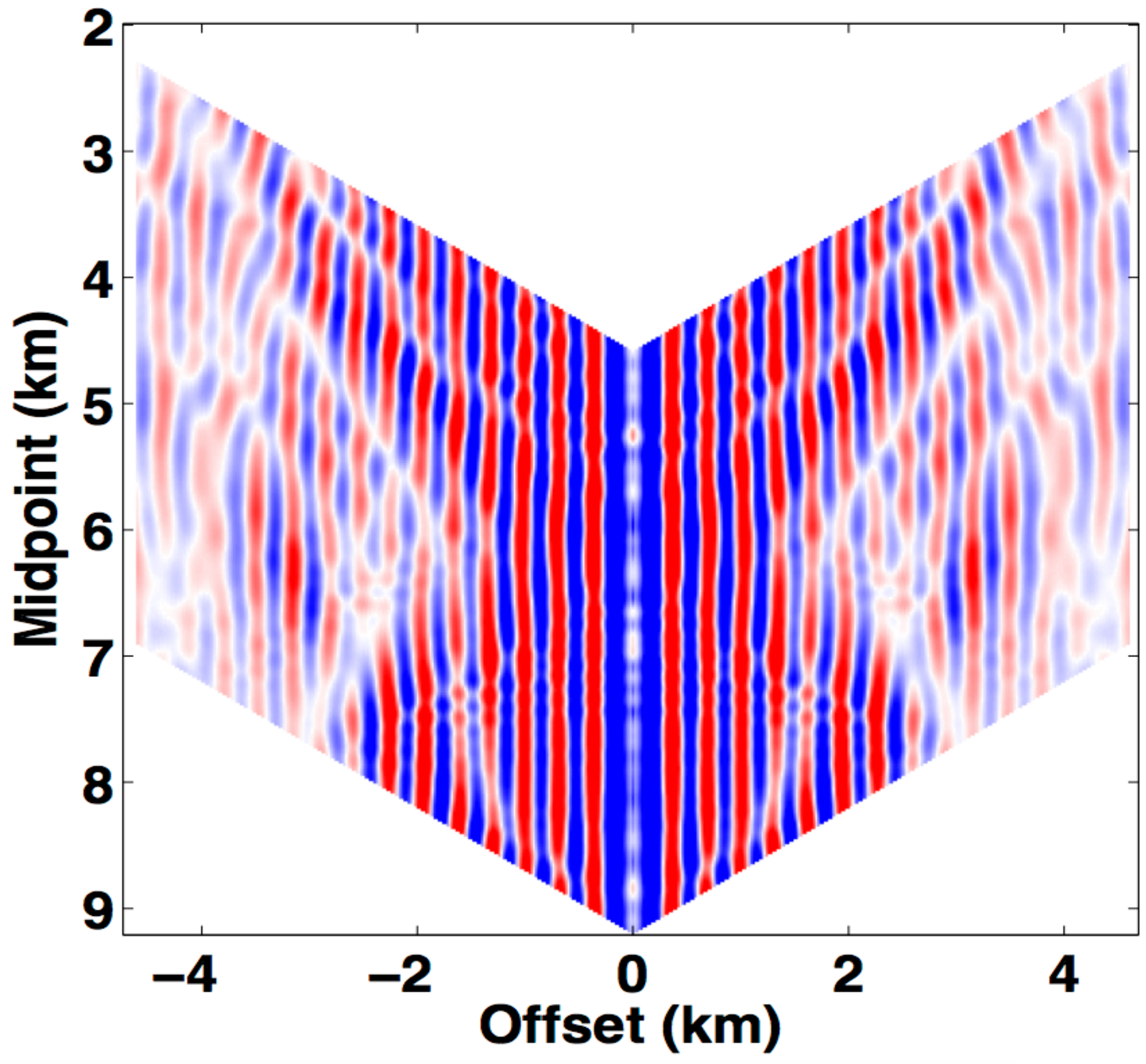
Low-rank structure in which domain?

- frequency slice at 5 Hz

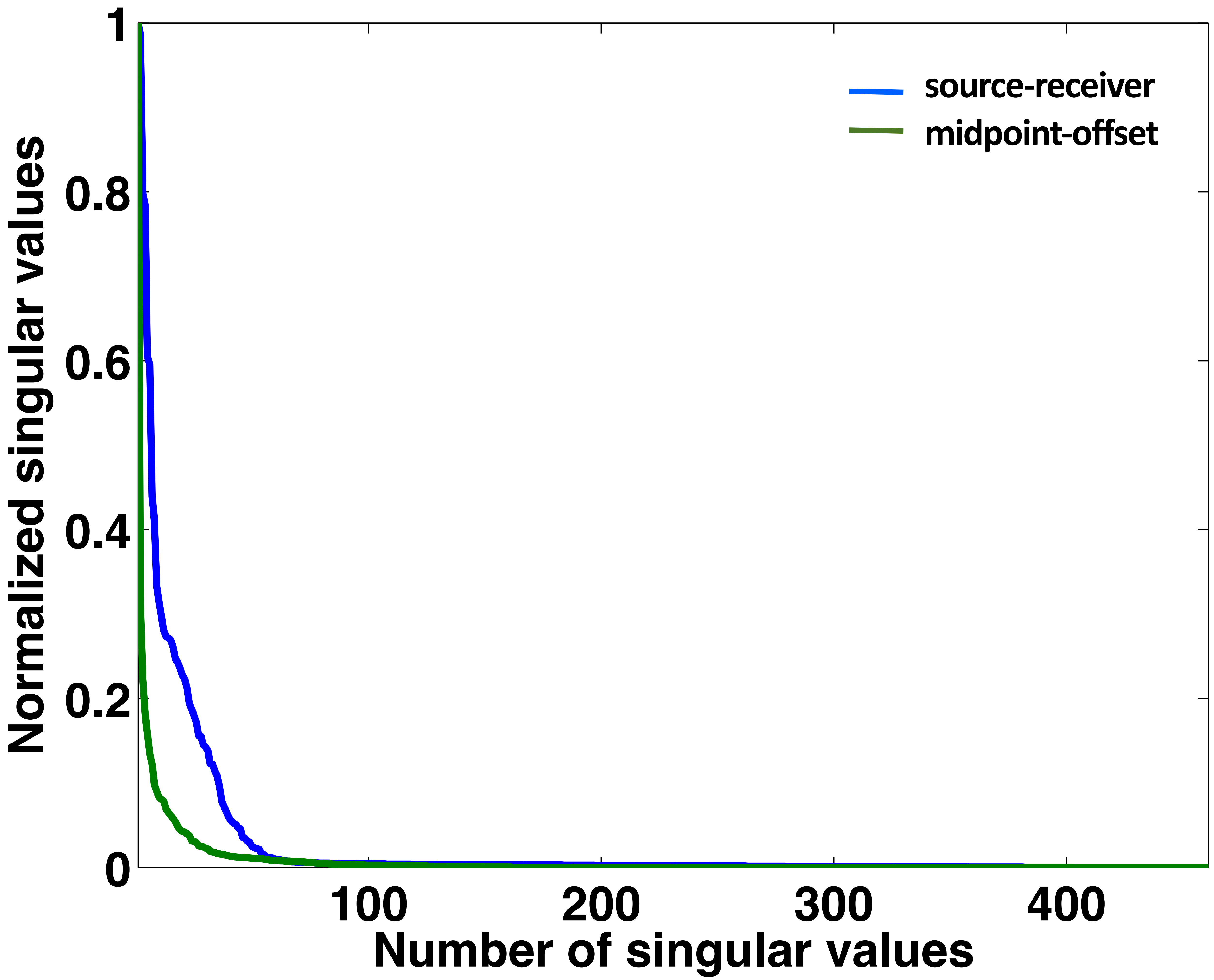
source-receiver domain
(with reciprocity)



midpoint-offset domain
(with reciprocity)



Decay of singular values

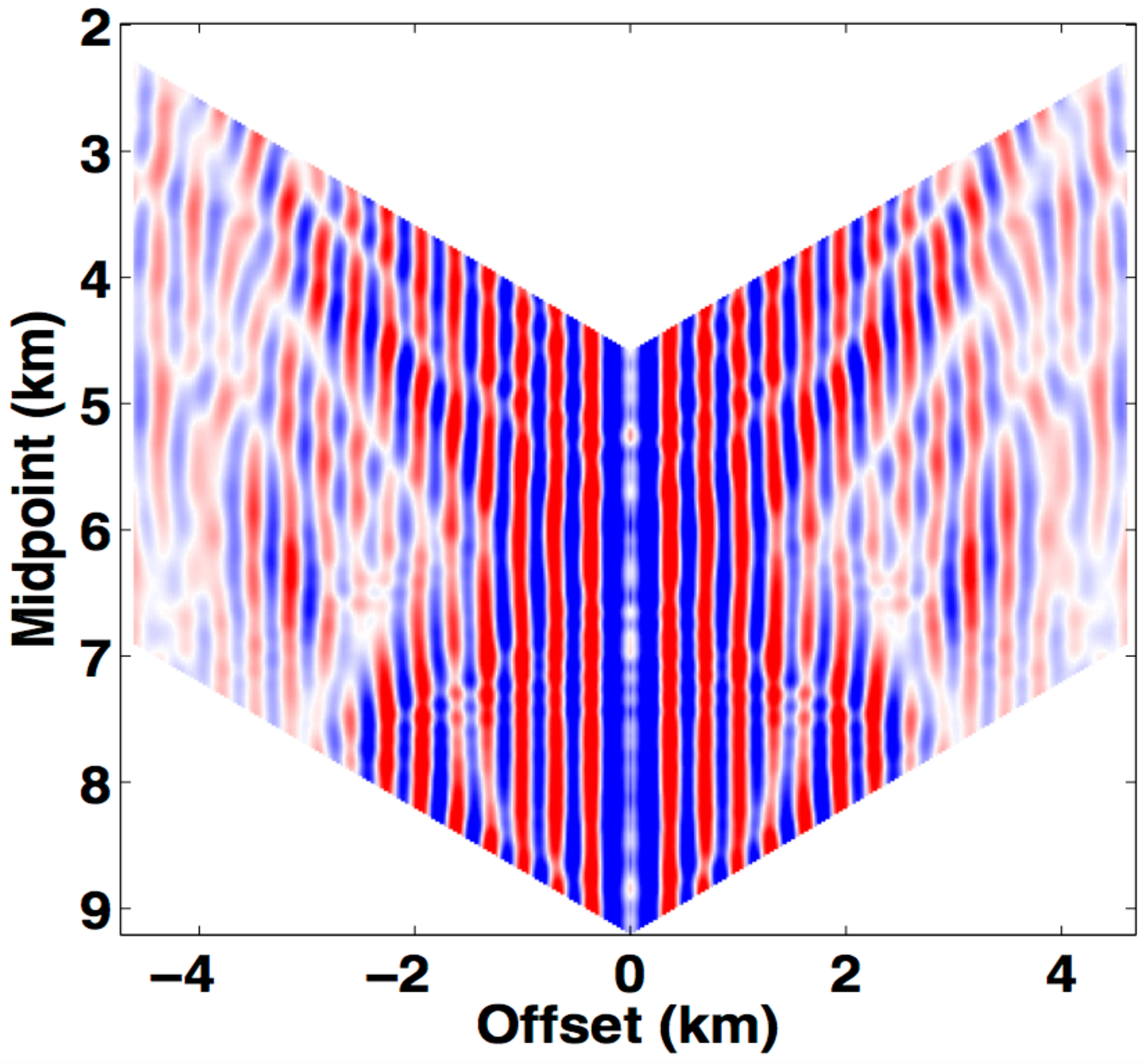


low-rank in midpoint-offset domain

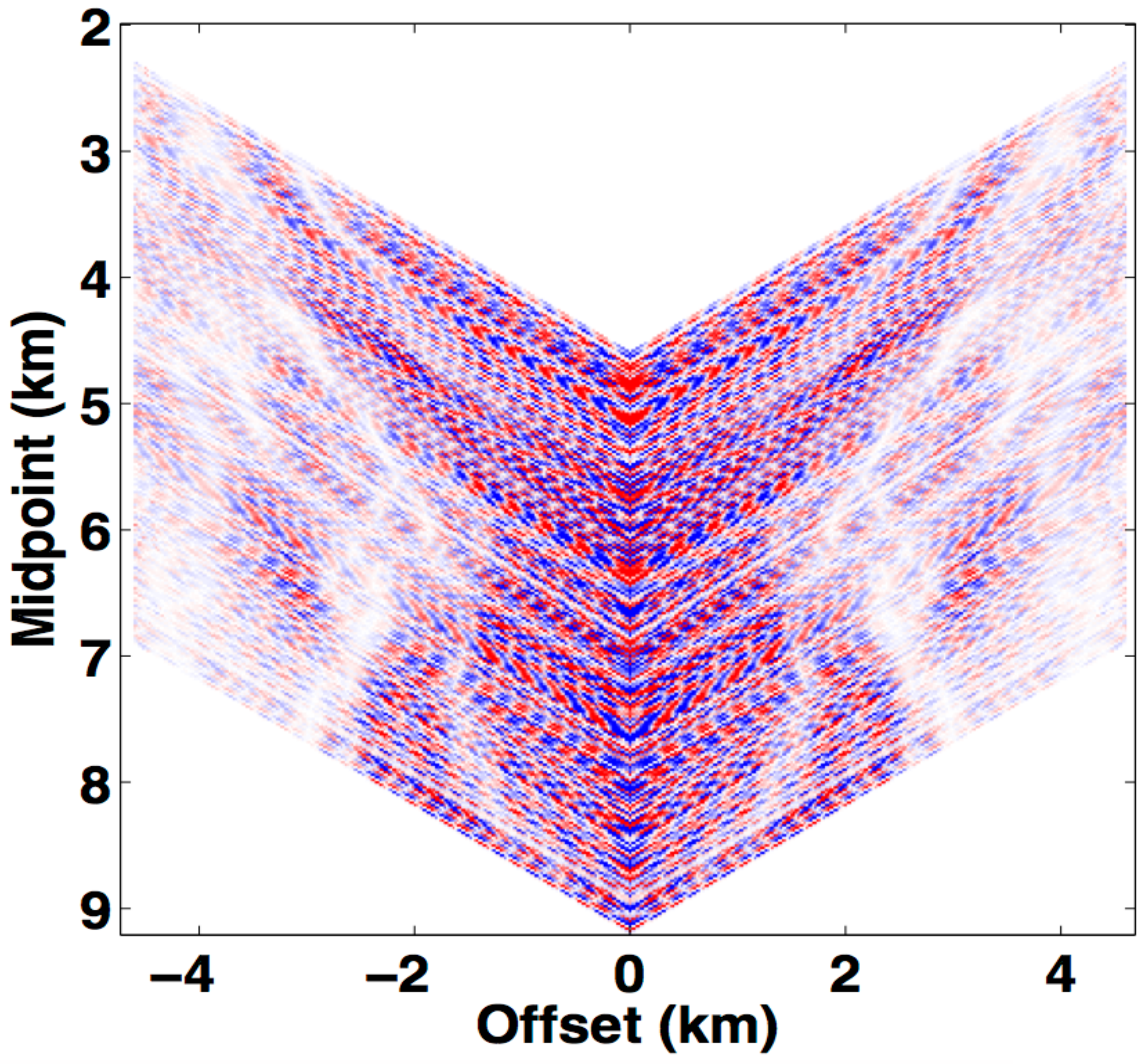
How to destroy the structure?

- add random time delays

without delays

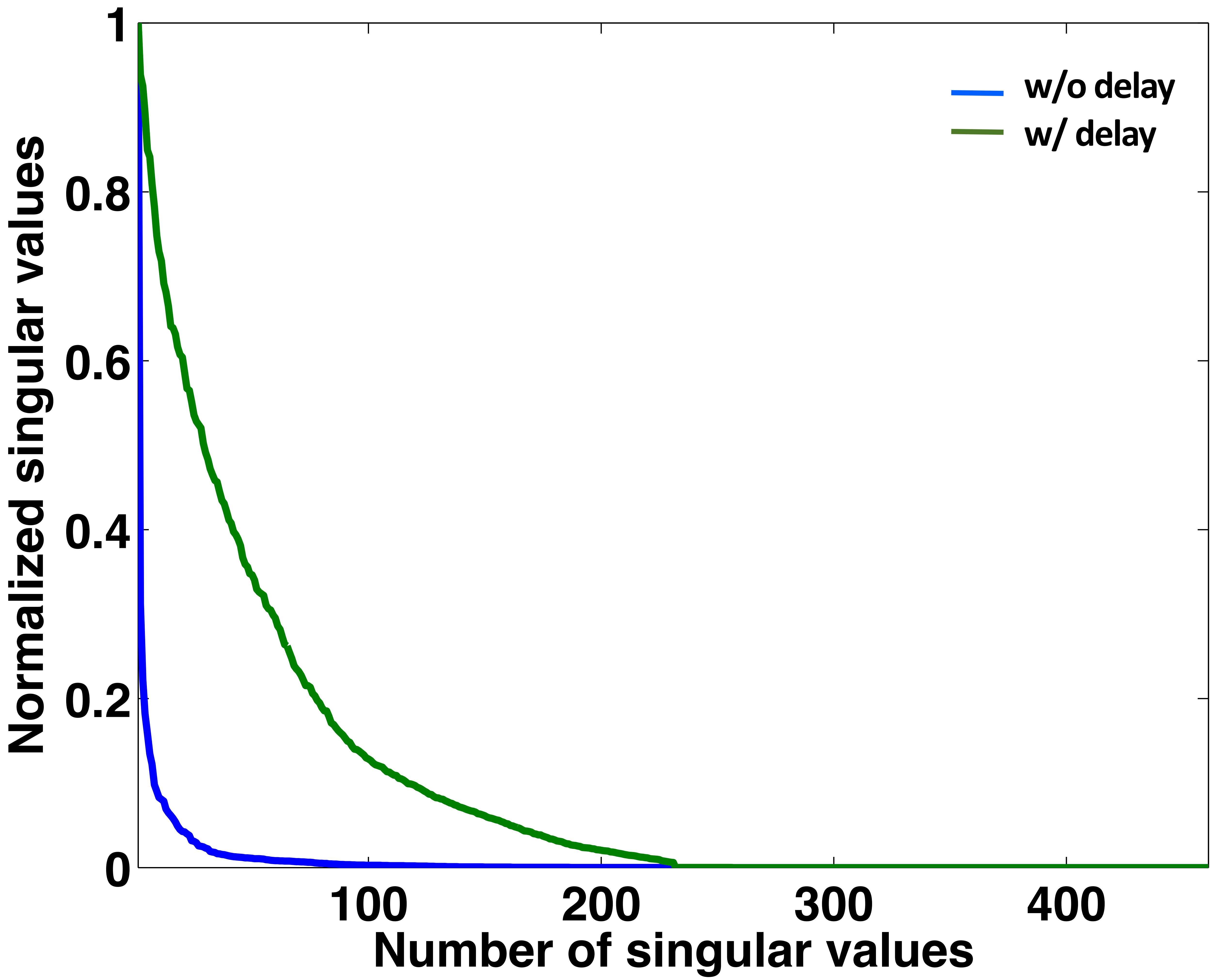


with random delays (< 1s)



Decay of singular values

- midpoint-offset domain



random time delays increase the rank

Rank-minimization

$$\min_{\mathbf{X}} \quad \underbrace{\operatorname{rank}(\mathbf{X})} \quad \text{s.t.} \quad \|\mathcal{A}(\mathbf{X}) - \mathbf{b}\|_2 \leq \epsilon$$

number of singular values of $\mathbf{X}$

for blended acquisition:

unblended data matrix

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix}$$

$\mathbf{X}_1$ ⟵ source 1

$\mathbf{X}_2$ ⟵ source 2

$\mathbf{b}$: blended data

$$\mathcal{A} := \begin{bmatrix} \mathbf{M}\mathbf{T_1}\mathbf{S^H} & \mathbf{M}\mathbf{T_2}\mathbf{S^H} \end{bmatrix}$$

↑ $\mathbf{T_1}$, $\mathbf{T_2}$: time delay matrices

Factorized formulation ("*SVD-free*")

[Rennie and Srebro, 2005; Lee et. al., 2010; Recht and Re, 2011]

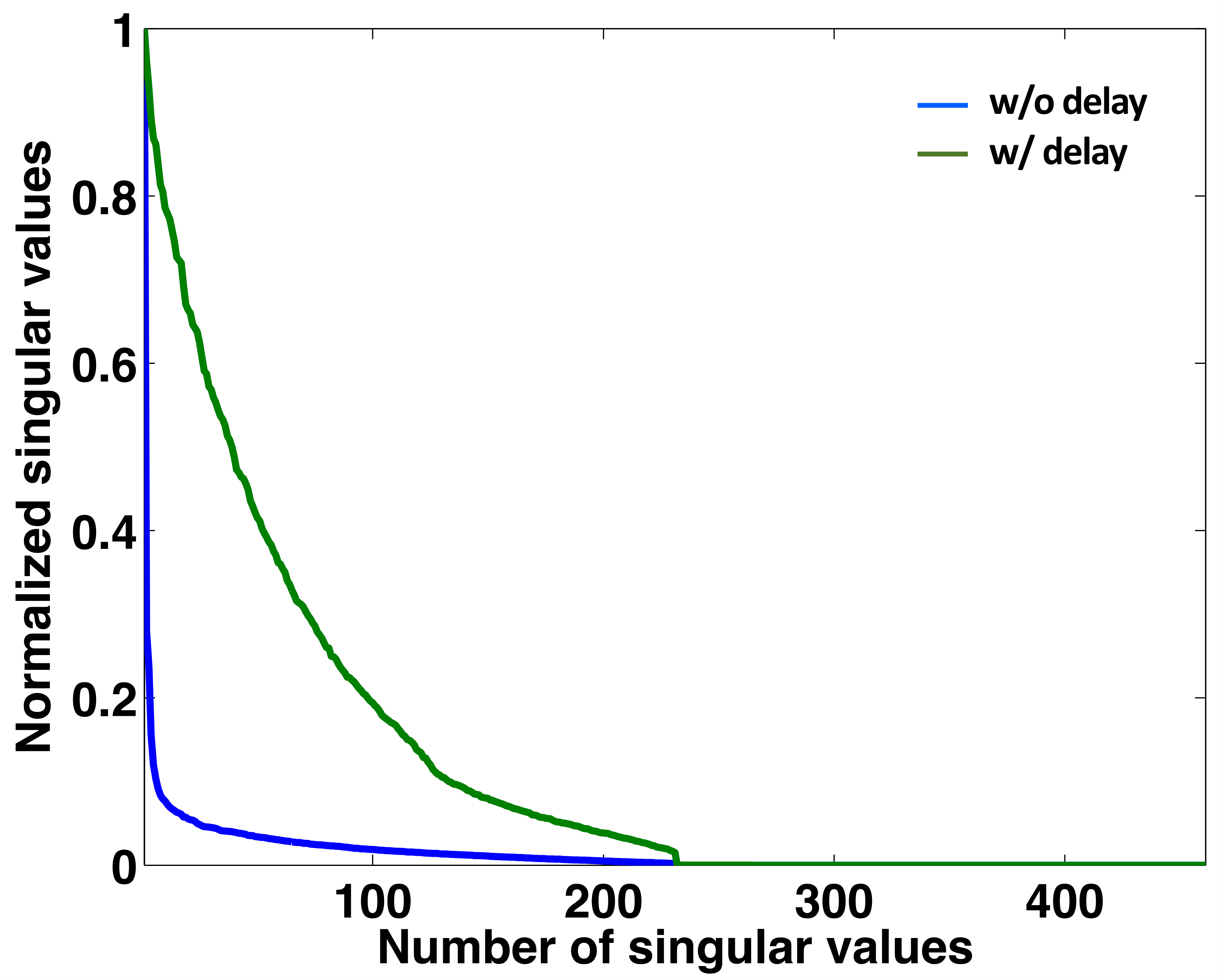
$$\mathbf{X} \in \mathbb{R}^{n \times m} = \mathbf{L} \in \mathbb{R}^{n \times k} \quad \mathbf{R}^{\mathbf{H}} \in \mathbb{R}^{k \times m}$$

Upper-bound on nuclear norm:

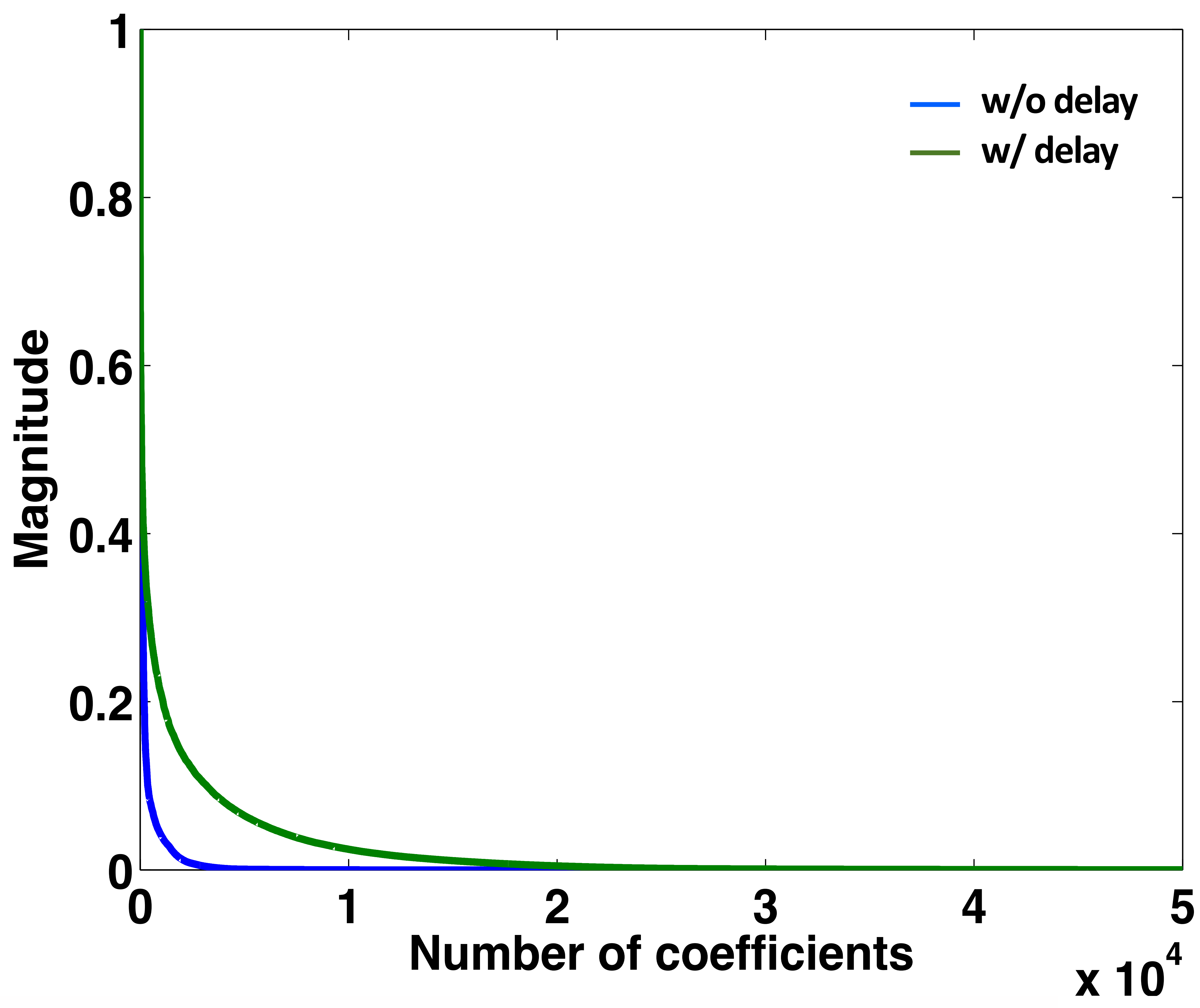
$$\|\mathbf{X}\|_* \leq \frac{1}{2} \left\| \begin{bmatrix} \mathbf{L}_1 \\ \mathbf{R}_1 \end{bmatrix} \right\|_F^2 + \frac{1}{2} \left\| \begin{bmatrix} \mathbf{L}_2 \\ \mathbf{R}_2 \end{bmatrix} \right\|_F^2 =: \Phi(\mathbf{L}_1, \mathbf{R}_1, \mathbf{L}_2, \mathbf{R}_2)$$

Rank vs. sparsity

rank-minimization
(midpoint-offset domain)



sparsity-promotion
(source-receiver domain)

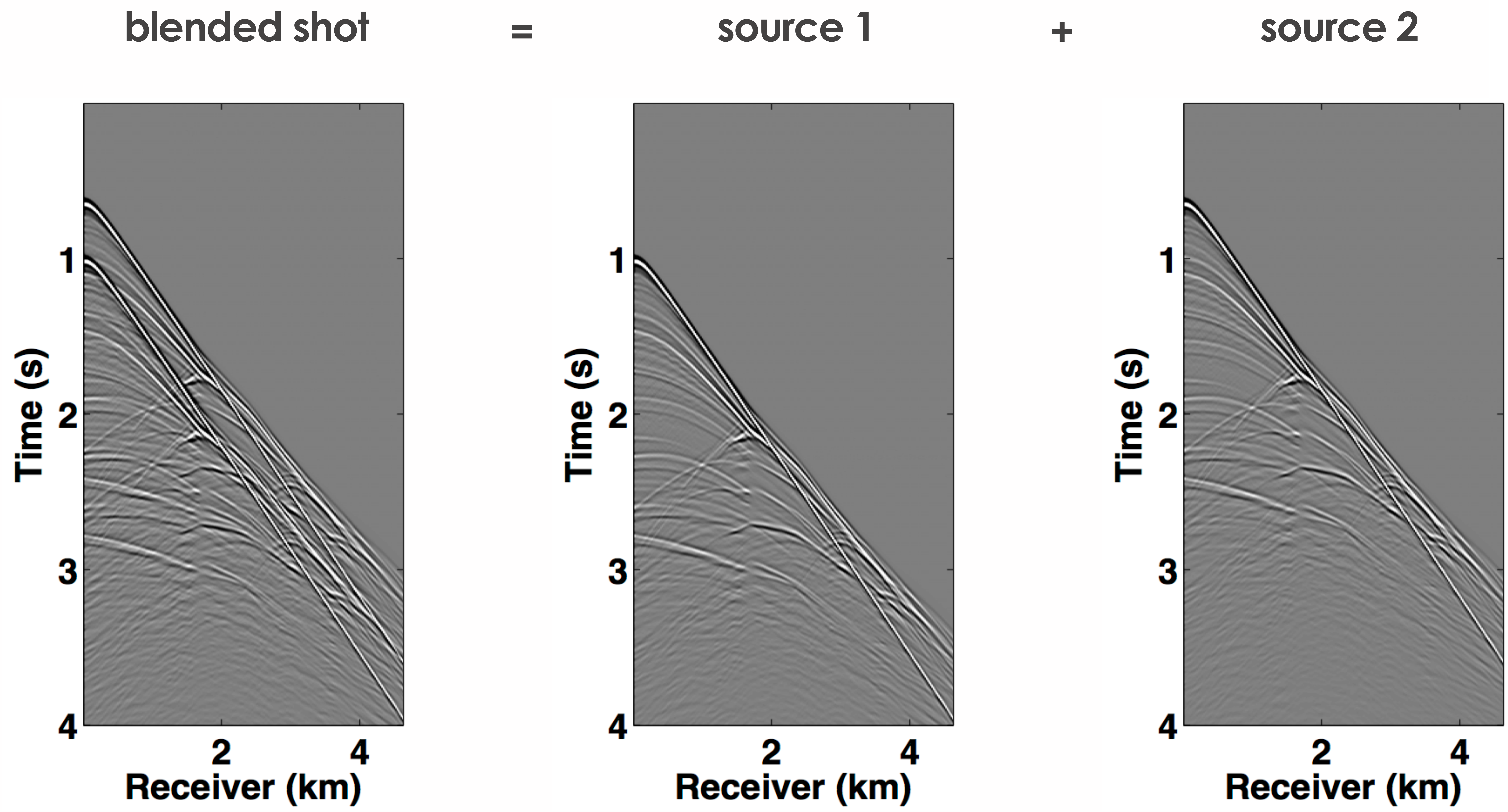


Source separation results

Rank-minimization vs. sparsity-promotion

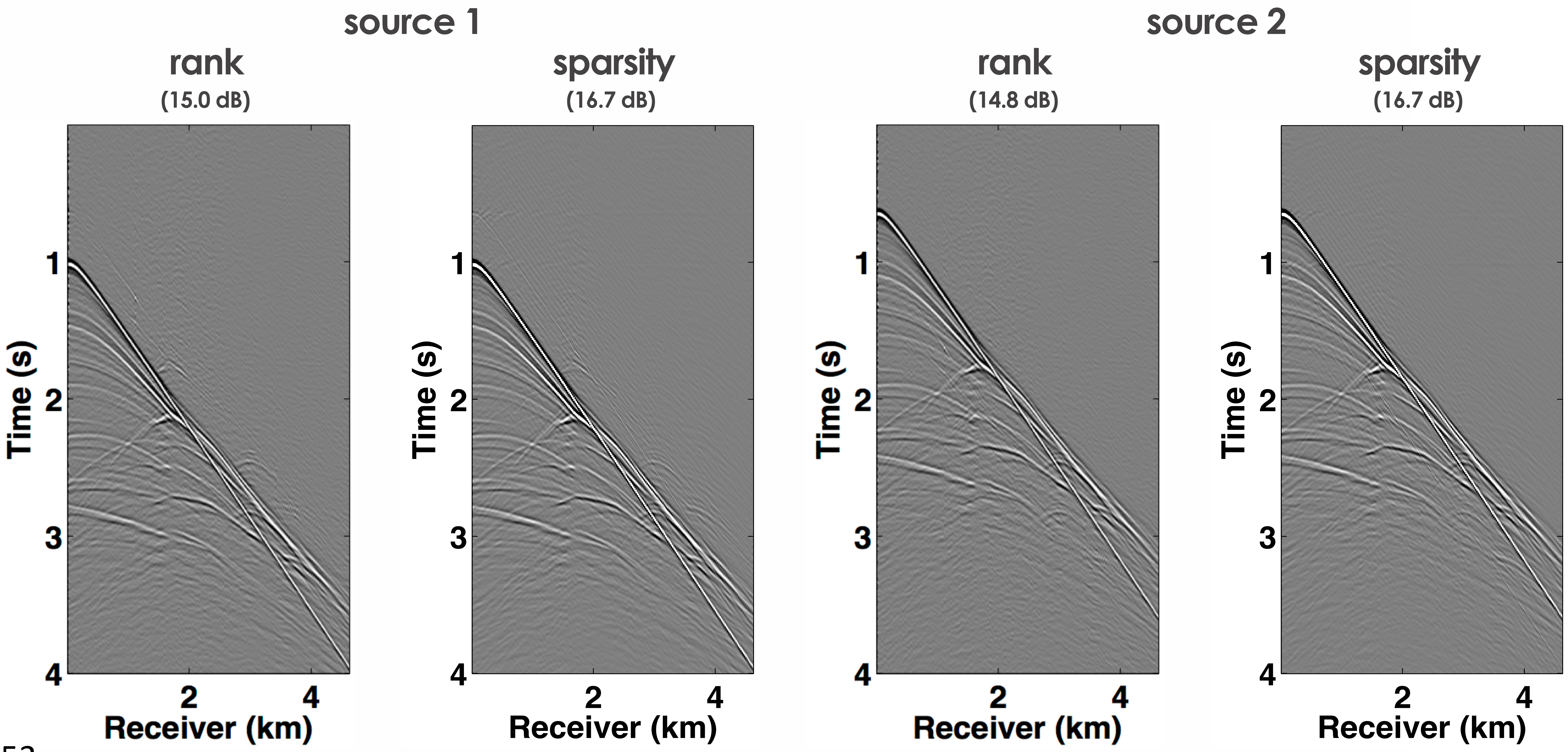
Blended data (w/ delay)

- random time delays (< 1 sec) applied to both sources



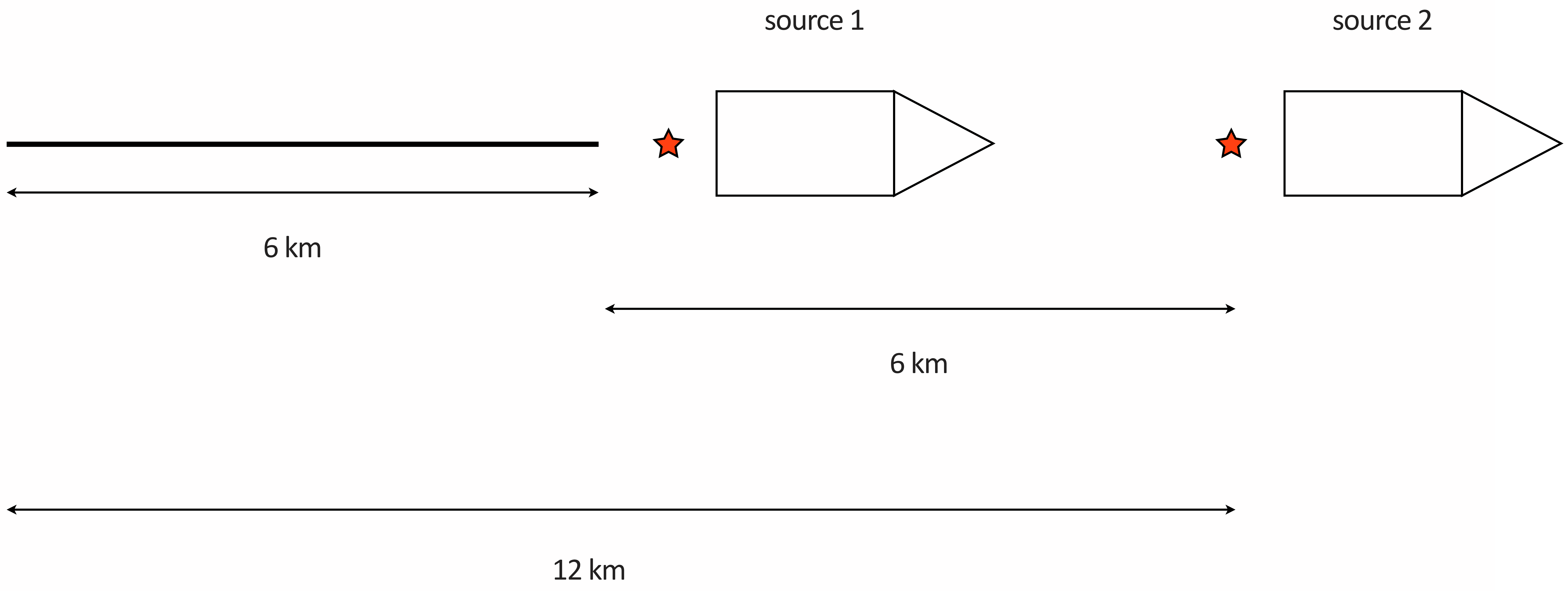
Source separation - rank vs. sparsity

computation time = 5 vs. 62 hours; memory usage = 2.8 vs. 7.0 GB;



Simultaneous long offset acquisition

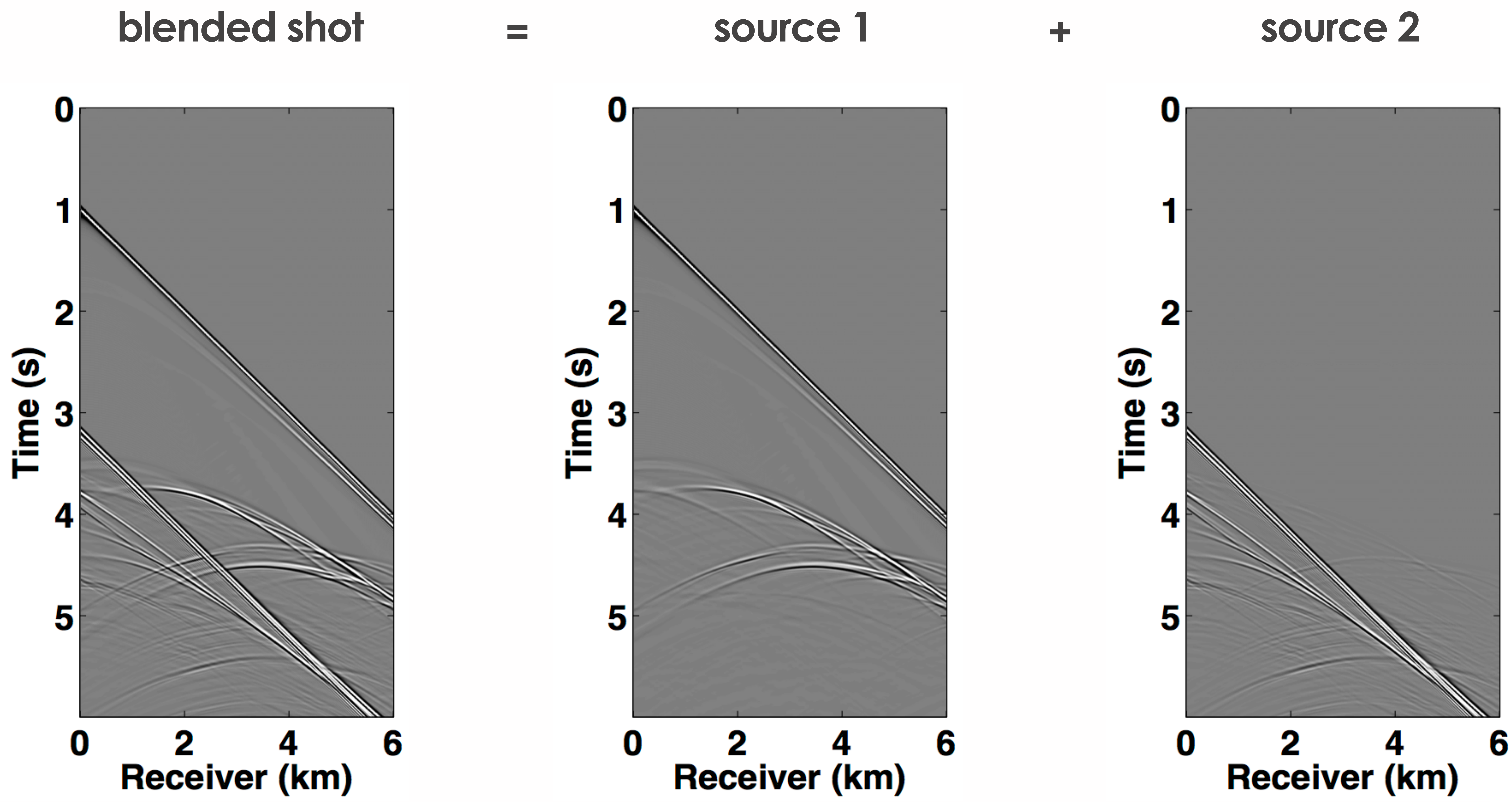
- adapted from Long, et. al., 2013



A. S. Long, et. al., "**Simultaneous long offset (SLO) towed streamer seismic acquisition**", presented at the *75th EAGE Conference and Exhibition*, June 2013.

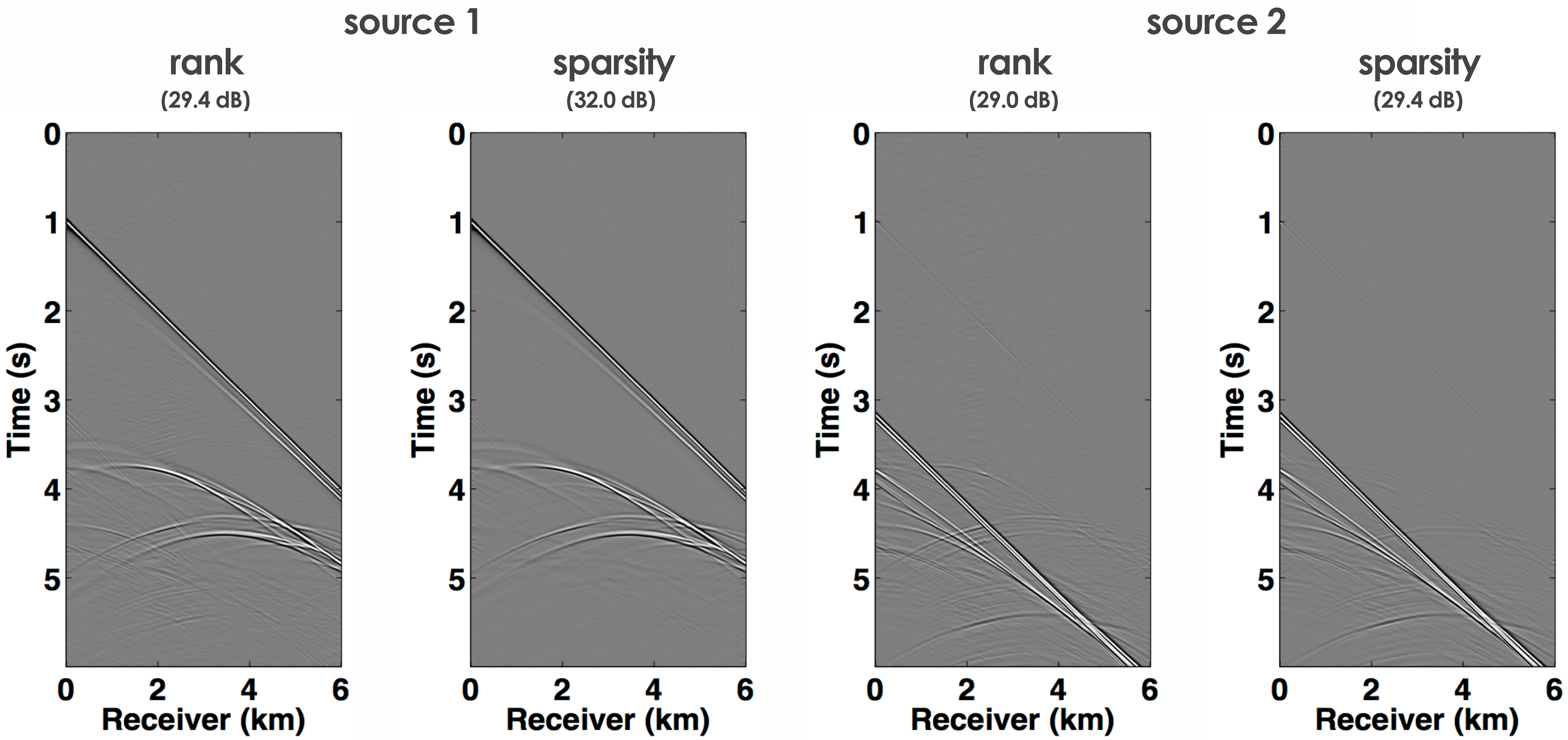
Blended data (w/ delay)

- random time delays (< 1 sec) applied to both sources



Source separation - rank vs. sparsity

computation time = 19 vs. 183 hours; memory usage = 6 vs. 12 GB



Rajiv Kumar, Haneet Wason and Felix J. Herrmann, "Source separation for simultaneous towed-streamer marine acquisition---a compressed sensing approach", *Geophysics*

SLIM

Summary

– time (in hours), memory (in GB), average SNR (in dB)

	Over/under acquisition			Simultaneous long offset acquisition		
	time	memory	SNR*	time	memory	SNR*
sparsity	62	7	17	183	12	32.0, 29.4
rank	5	3	15.0, 14.8	19	6	29.4, 29.0

* average SNR for source 1, source 2

Full-waveform inversion

Ernie Esser, Felix Herrmann

SLIM
University of British Columbia

FWI issues

- *cycle skipping*
- *accurate* starting model
- *longer* offsets
- *low frequency*

SLIM

WRI

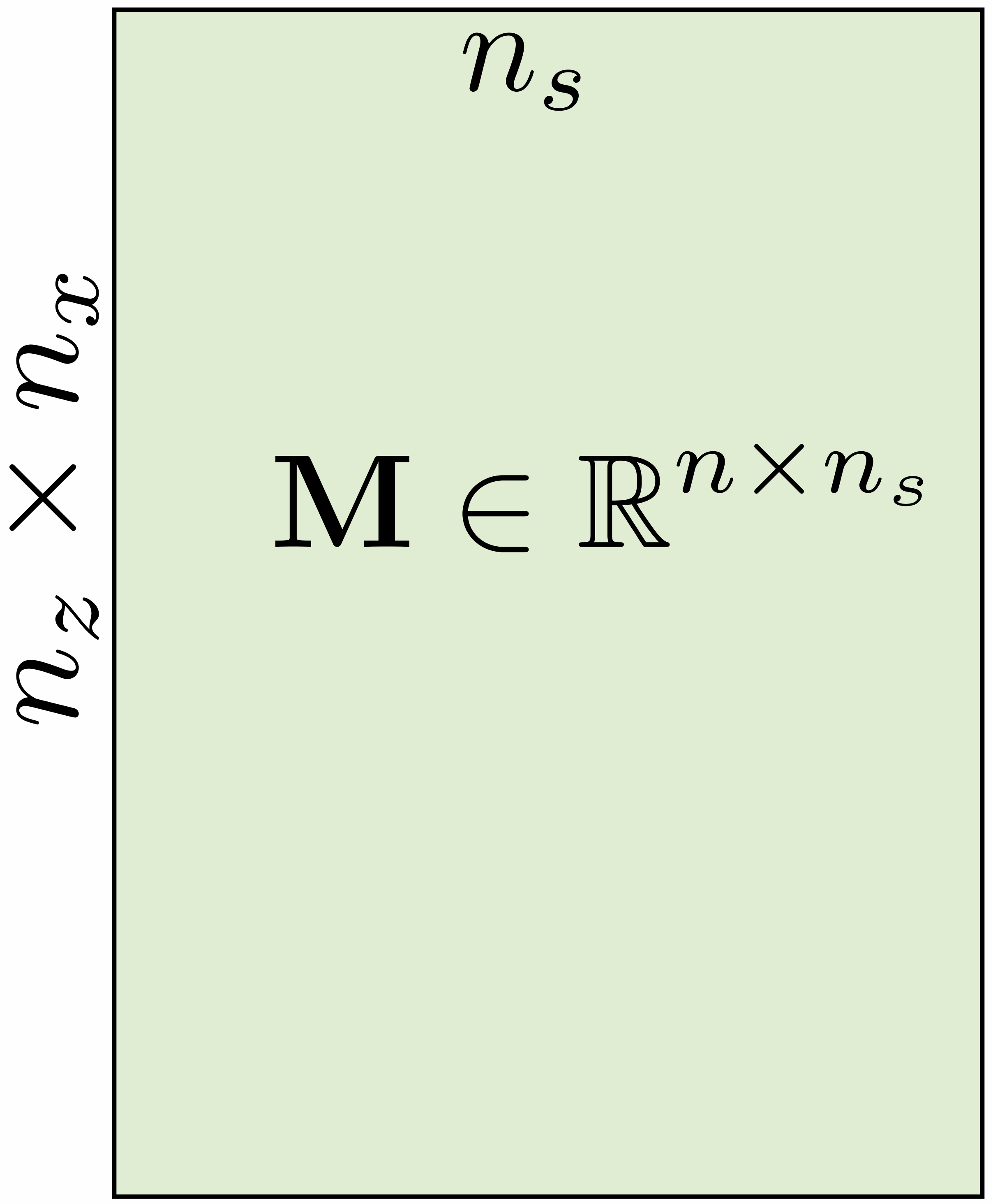
$$\bar{\phi}_\lambda(\mathbf{m}) = \frac{1}{2}\sum_{i}^{n_s} \|P\bar{\mathbf{u}}_i - \mathbf{d}_i\|_2^2 + \frac{\lambda^2}{2}\|A(\mathbf{m})\bar{\mathbf{u}}_i - \mathbf{q}_i\|_2^2$$

where

$$\bar{\mathbf{u}}_i = \arg\min_{\mathbf{u}_i} \left\| \begin{pmatrix} \lambda A(\mathbf{m}) \\ P \end{pmatrix} \mathbf{u}_i - \begin{pmatrix} \lambda \mathbf{q}_i \\ \mathbf{d}_i \end{pmatrix} \right\|_2$$

Low-rank extension

$$\bar{\phi}_\lambda(\mathbf{M}) = \frac{1}{2}\sum_{i}^{n_s} \|P\bar{\mathbf{u}}_i - \mathbf{d}_i\|_2^2 + \frac{\lambda^2}{2}\|A(\mathbf{M}_i)\bar{\mathbf{u}}_i - \mathbf{q}_i\|_2^2$$

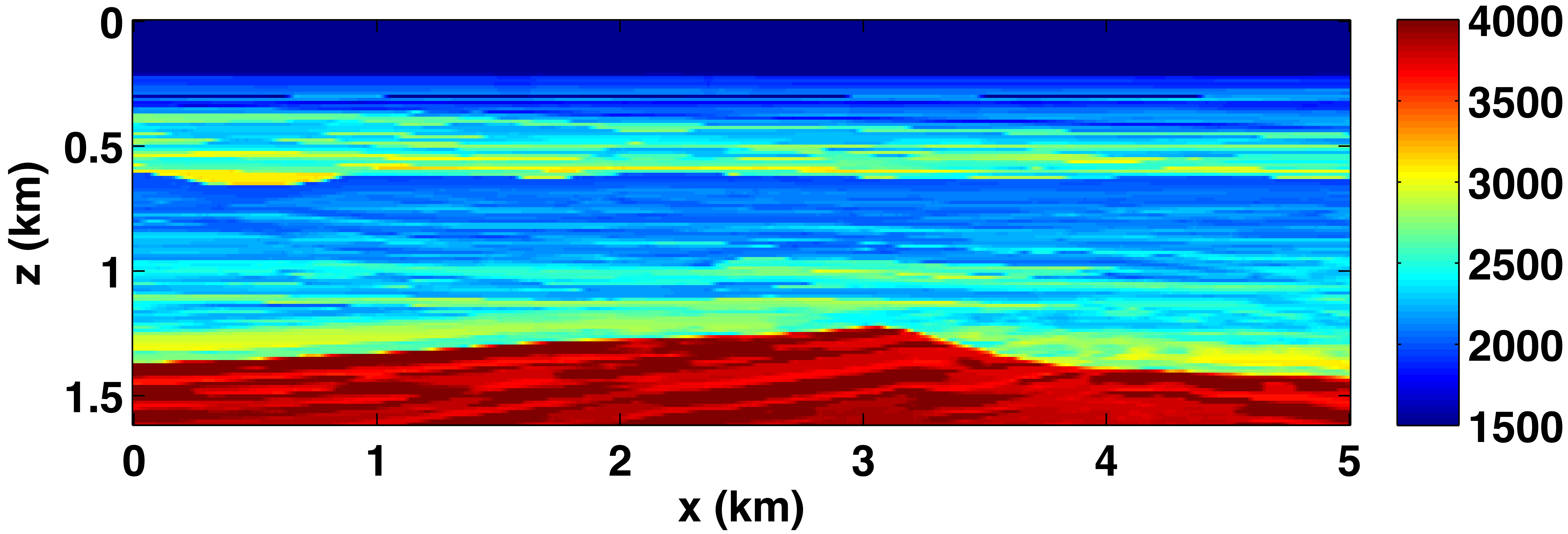


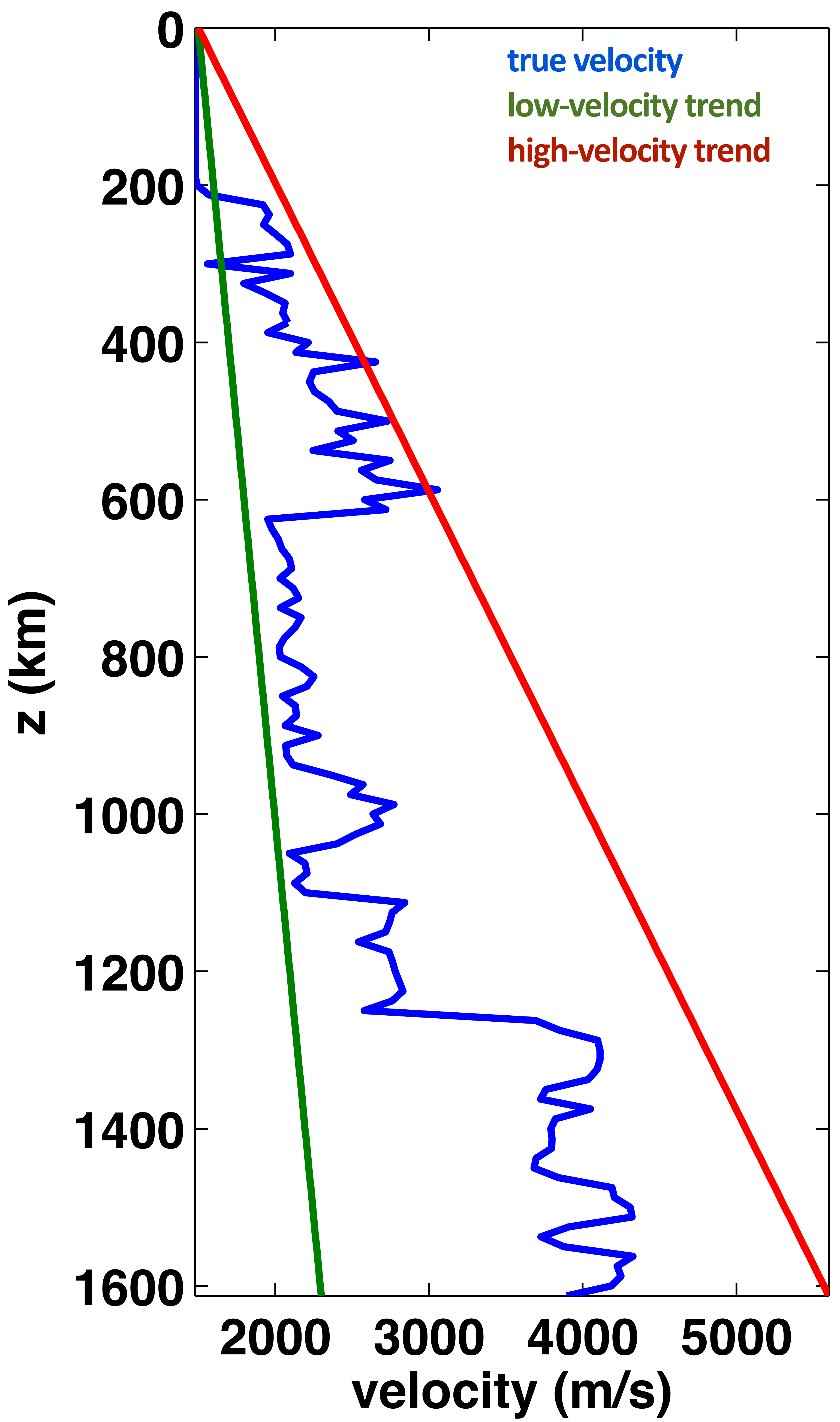
BG Model Example

BG compass model

- 25 frequency batches (20 iterations each)
  {3 4} ,{3.6 4.4},… ,{17 18} Hertz. Each interval contains 6 frequencies.
- 10 simultaneous sources
- receivers 25m sample interval
- different inaccurate initial model

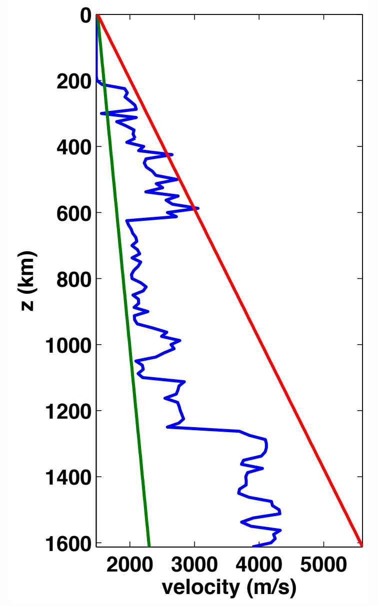
BG model

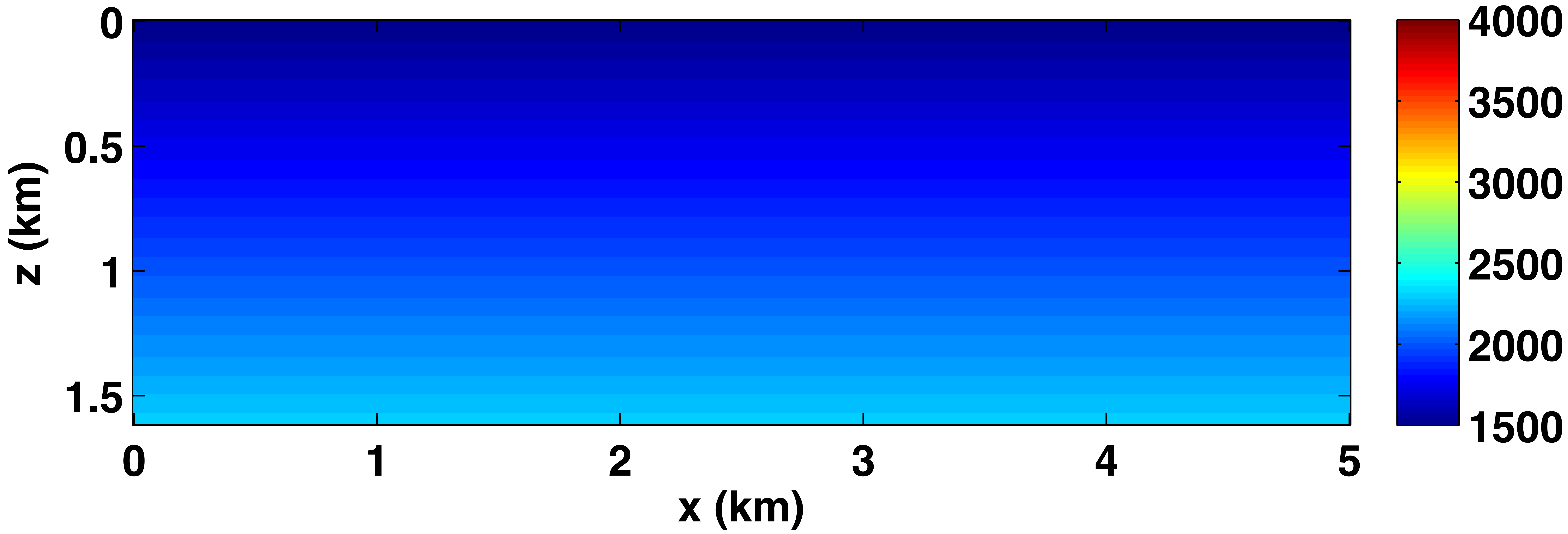




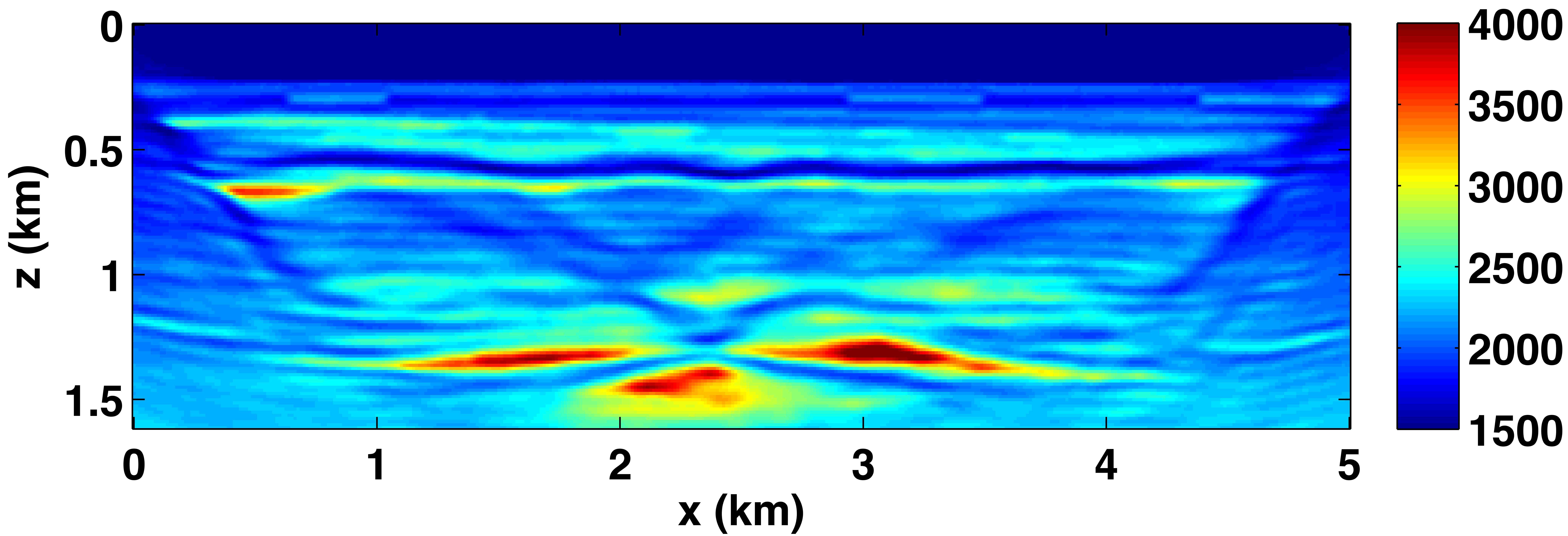
Vertical trace

Initial model
low-velocity trend

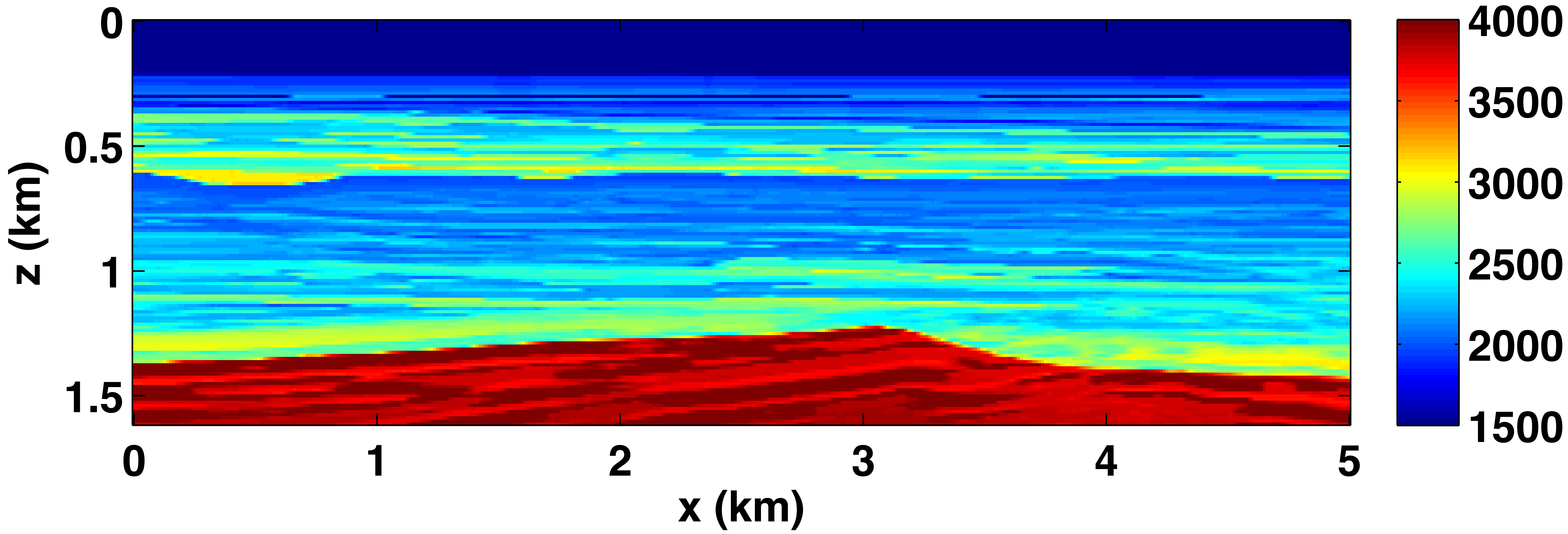




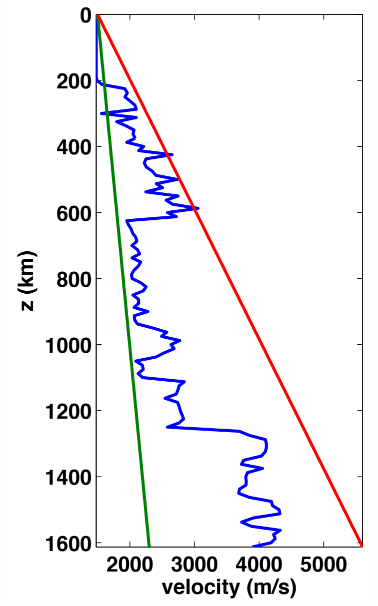
Inverted model
low-velocity trend

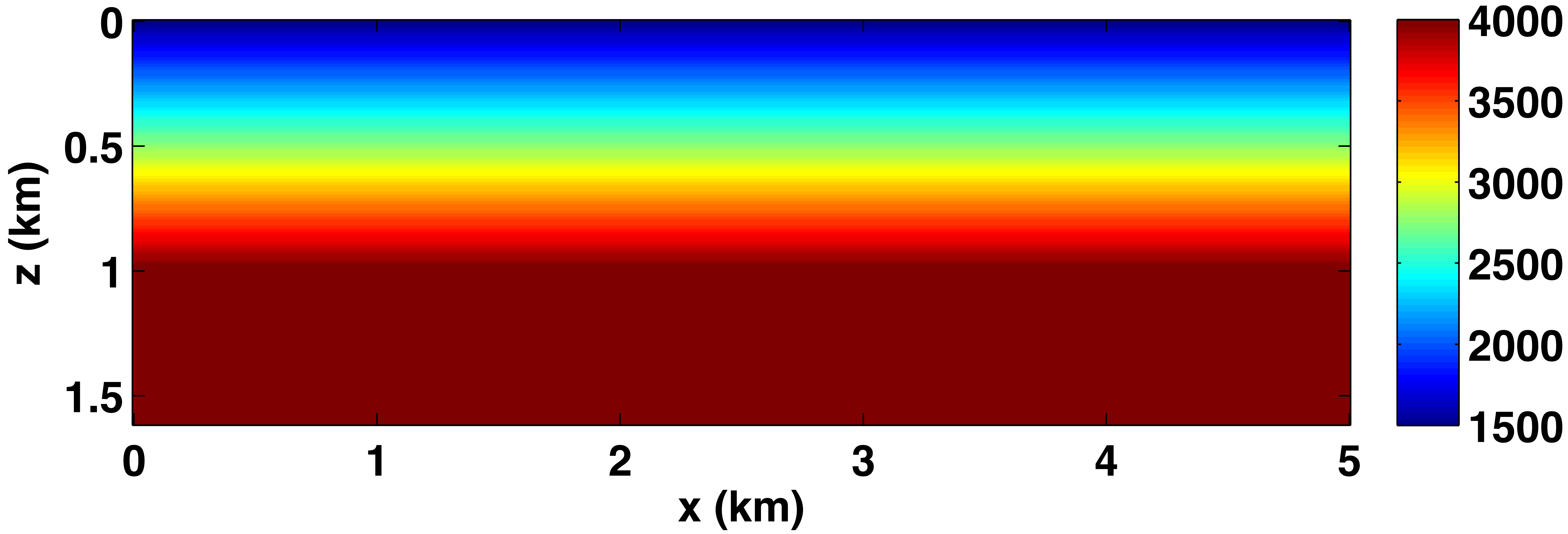


BG model

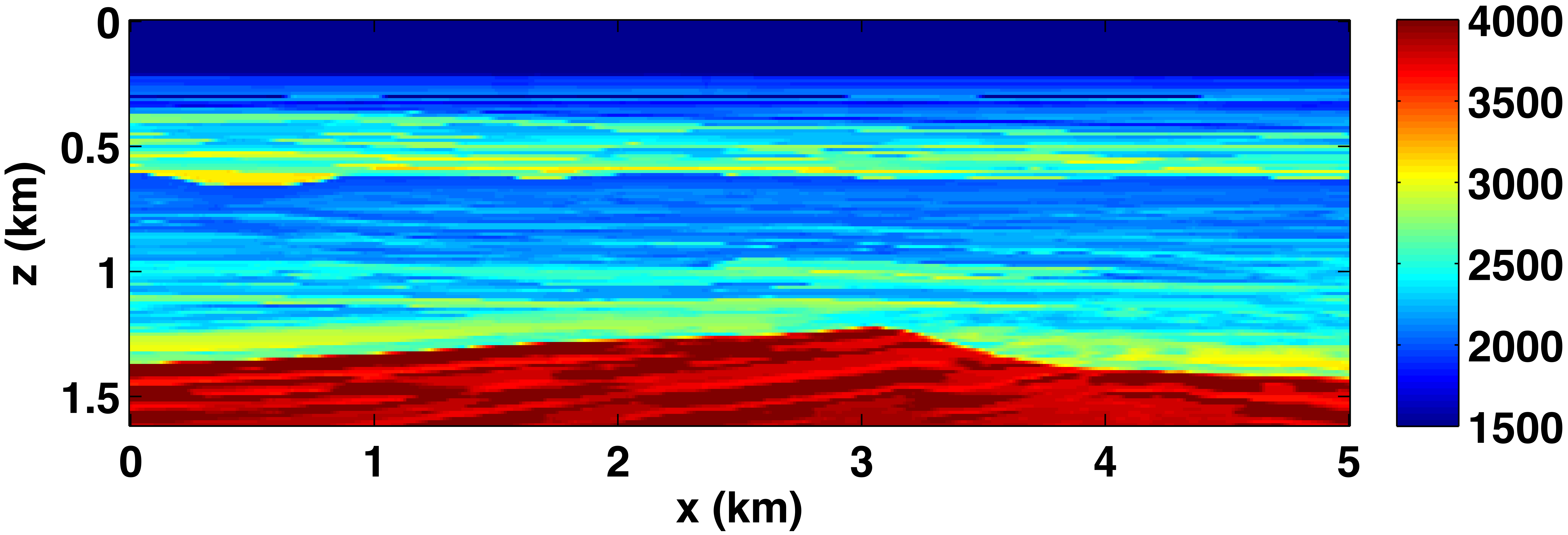


Initial model
high-velocity trend

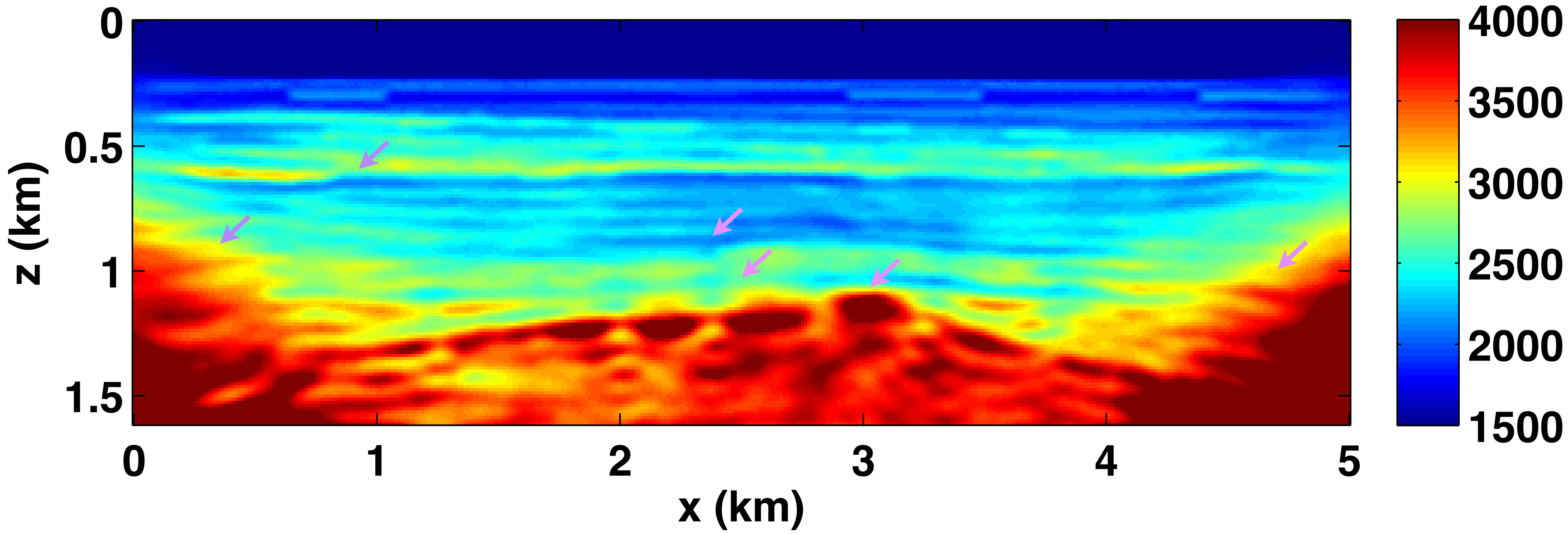


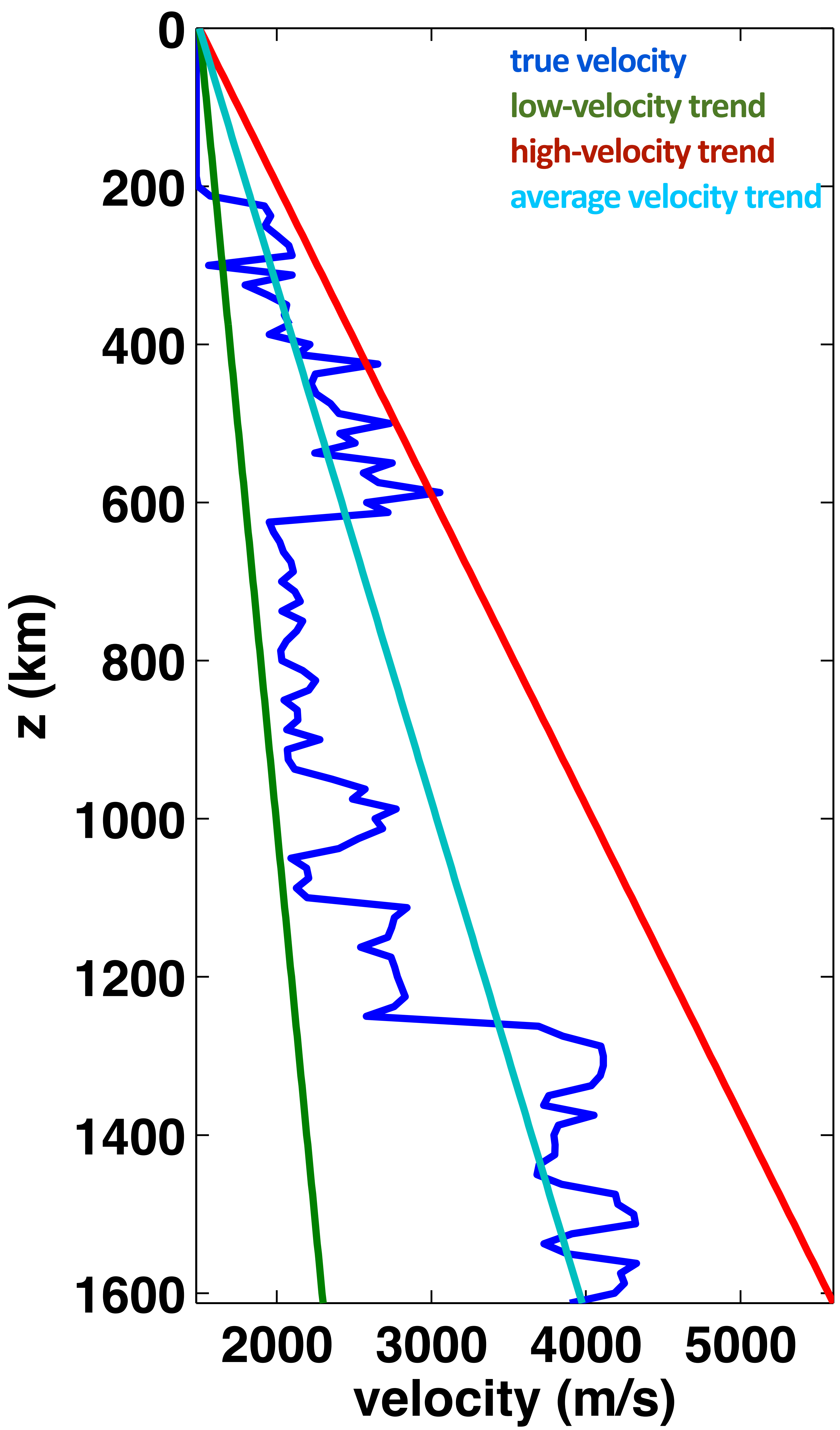


BG model



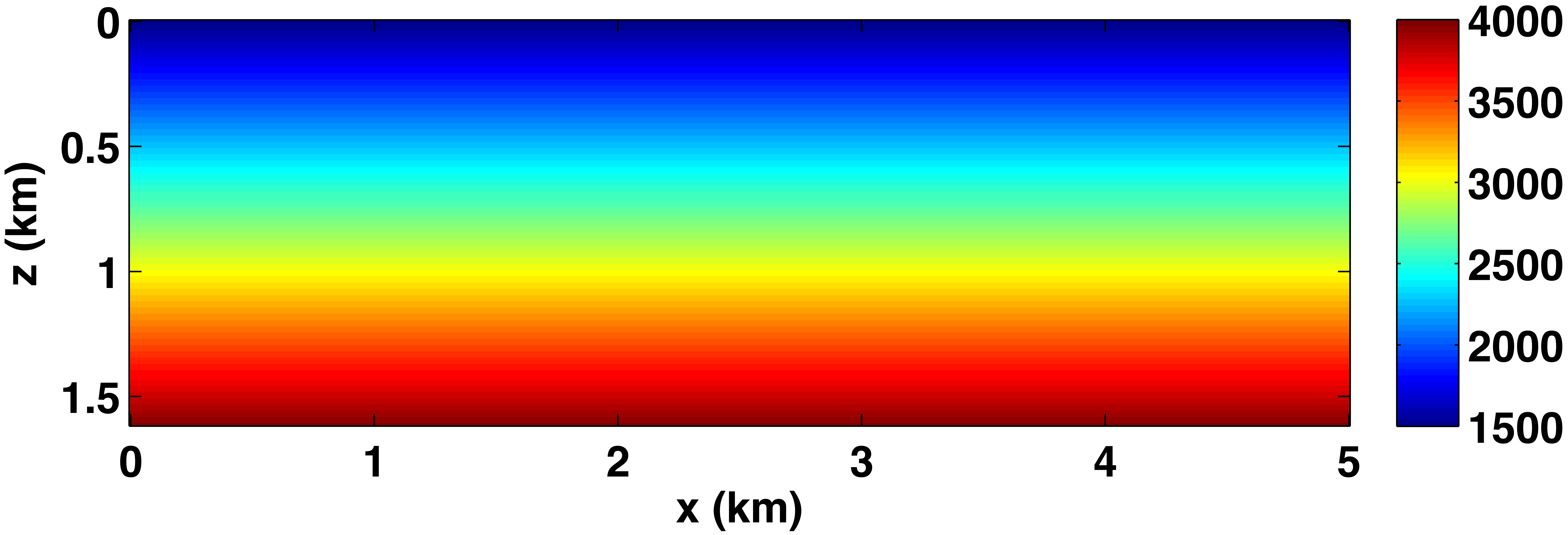
Inverted model
high-velocity trend



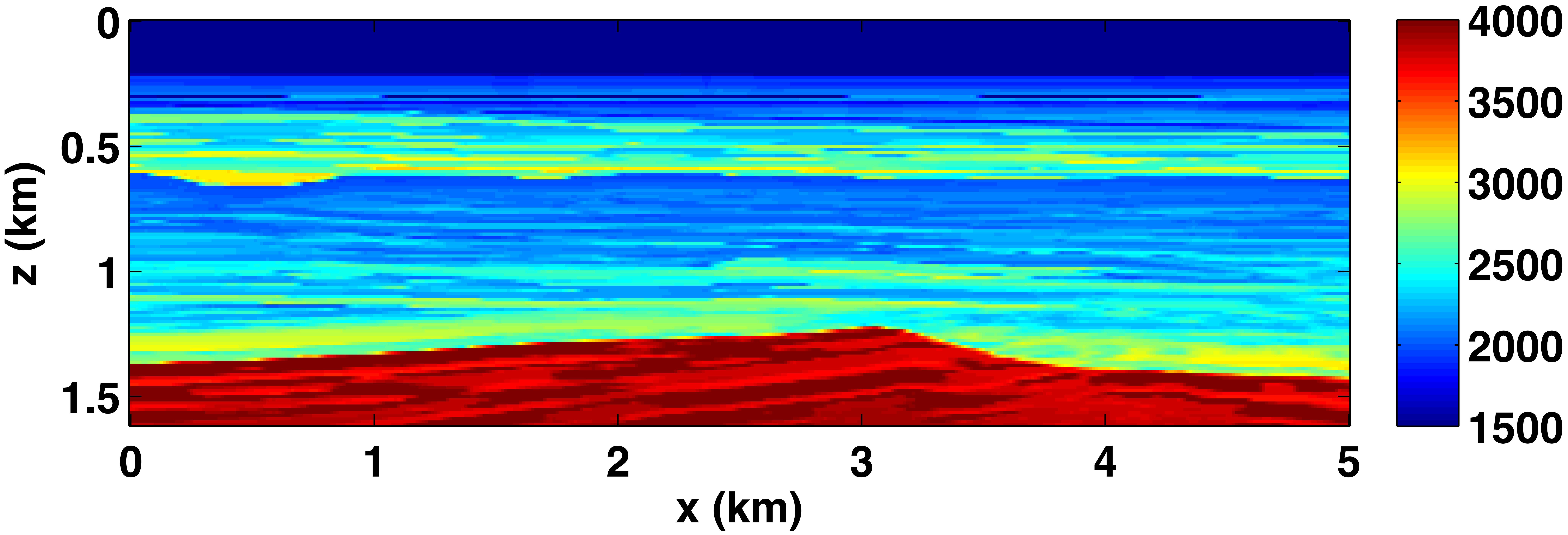


Vertical trace

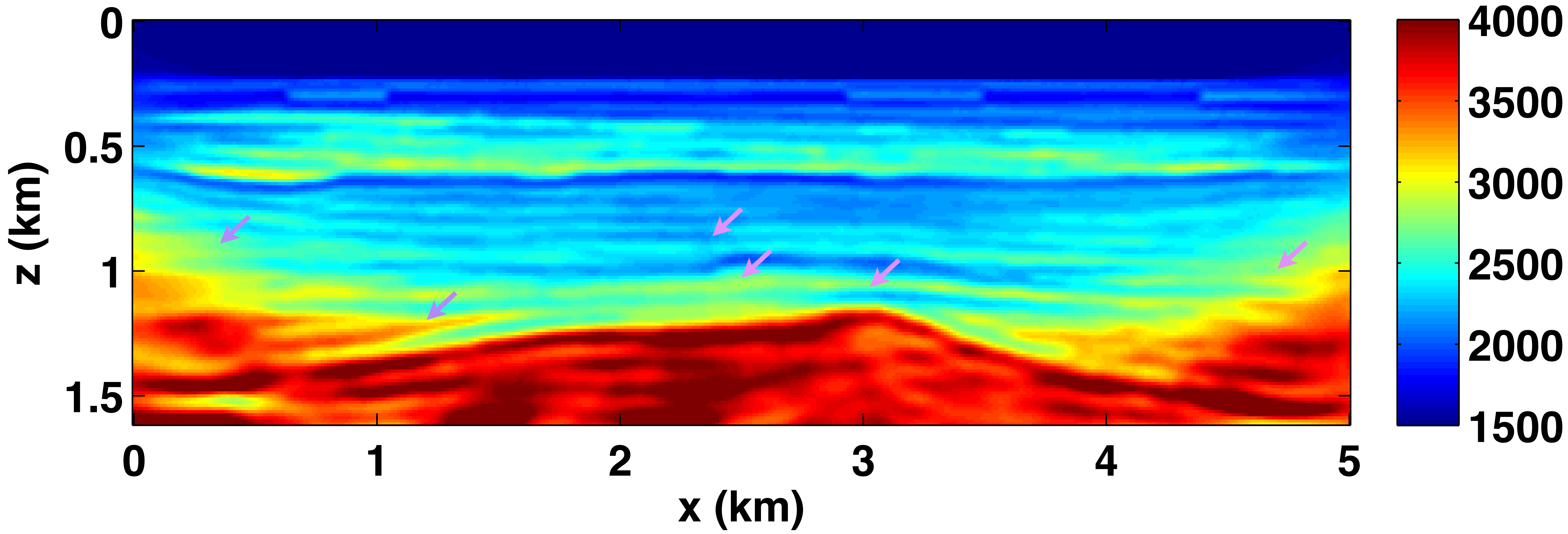
Initial model
average-velocity trend



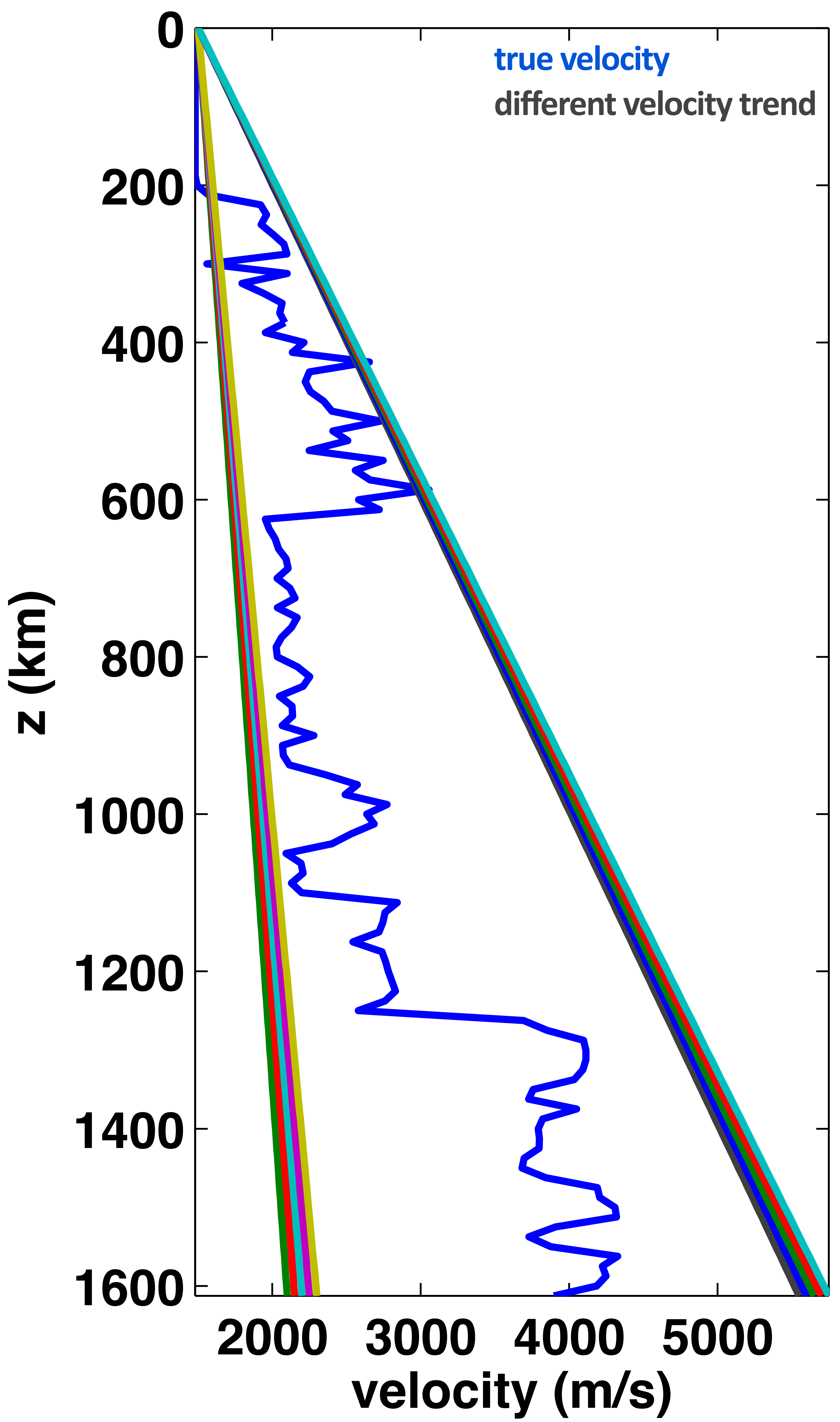
BG model



Inverted model
average-velocity trend

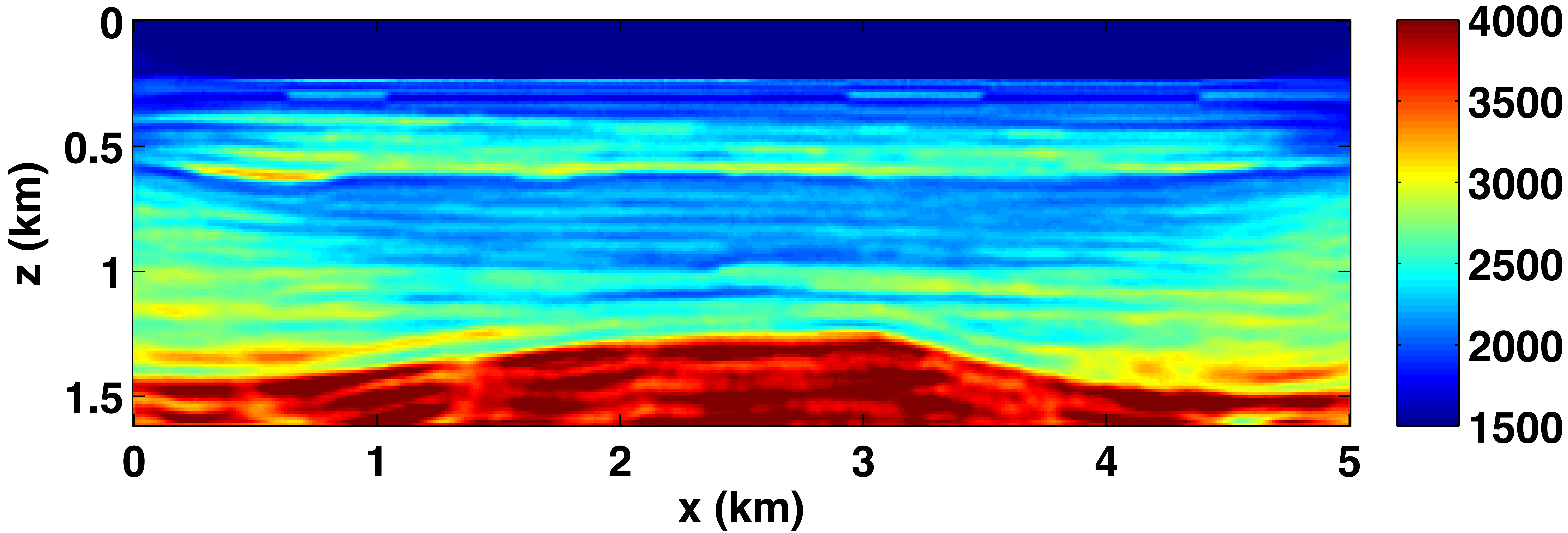


Low-rank extension

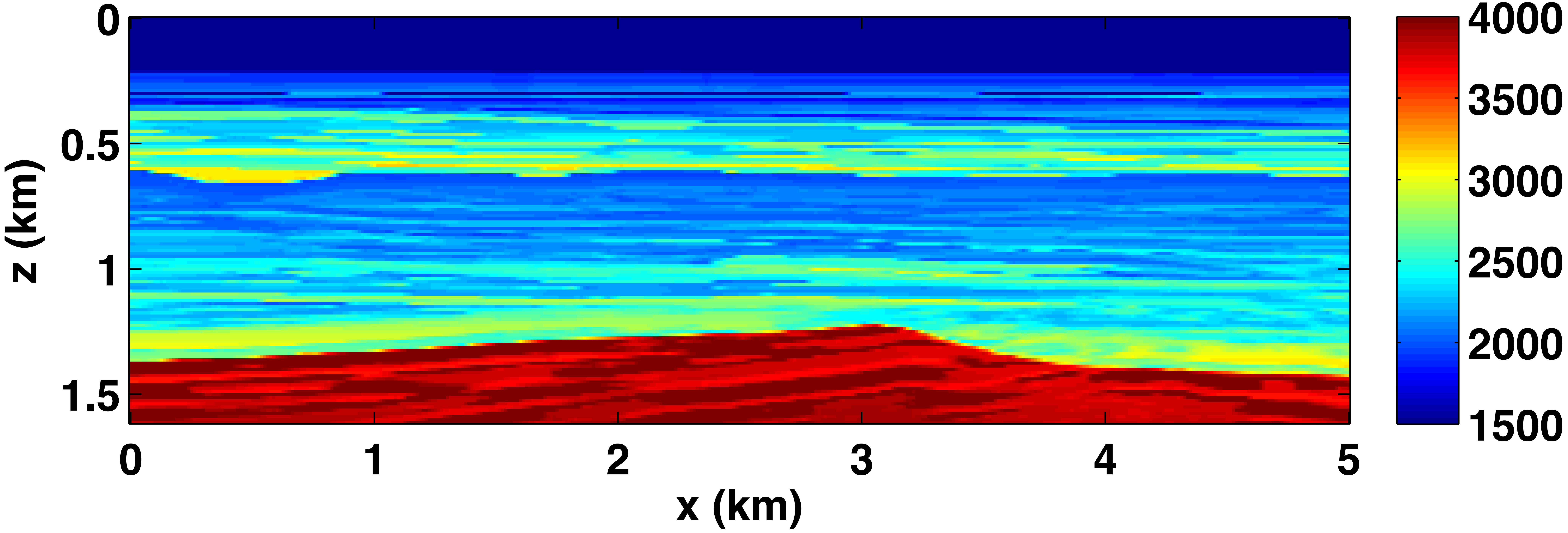


Vertical trace

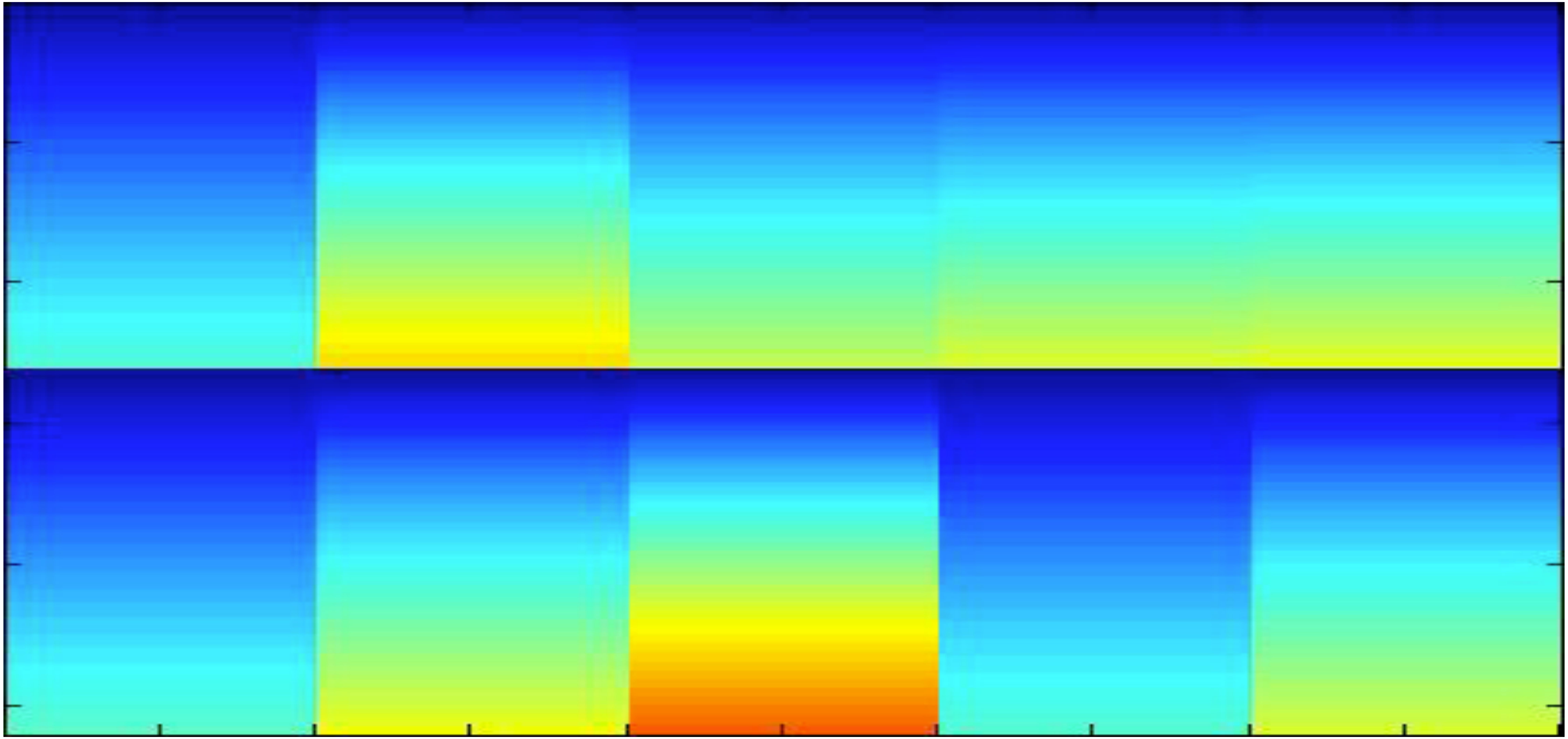
Inverted model
low-rank extension



BG model

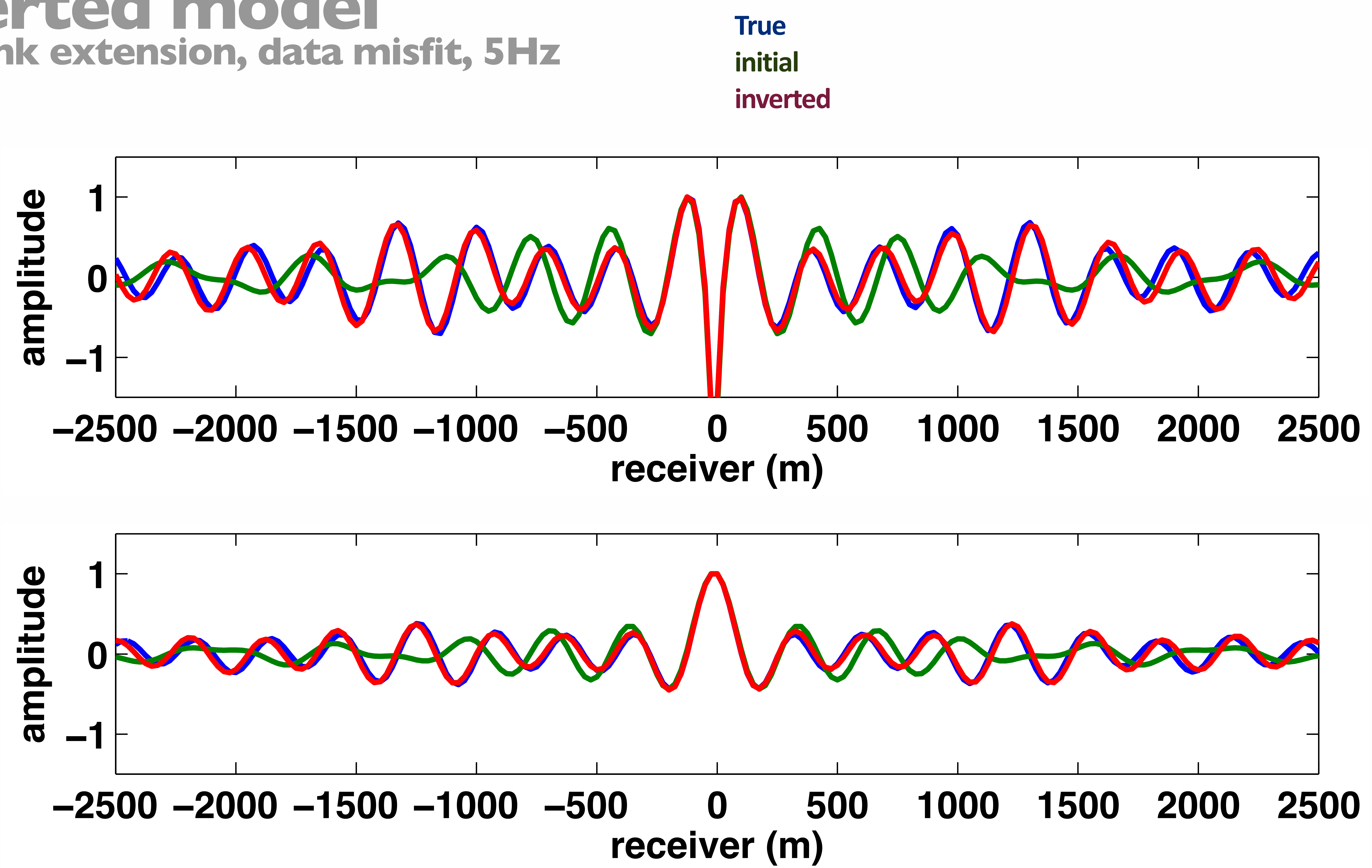


Inverted model
low-rank extension, batch1

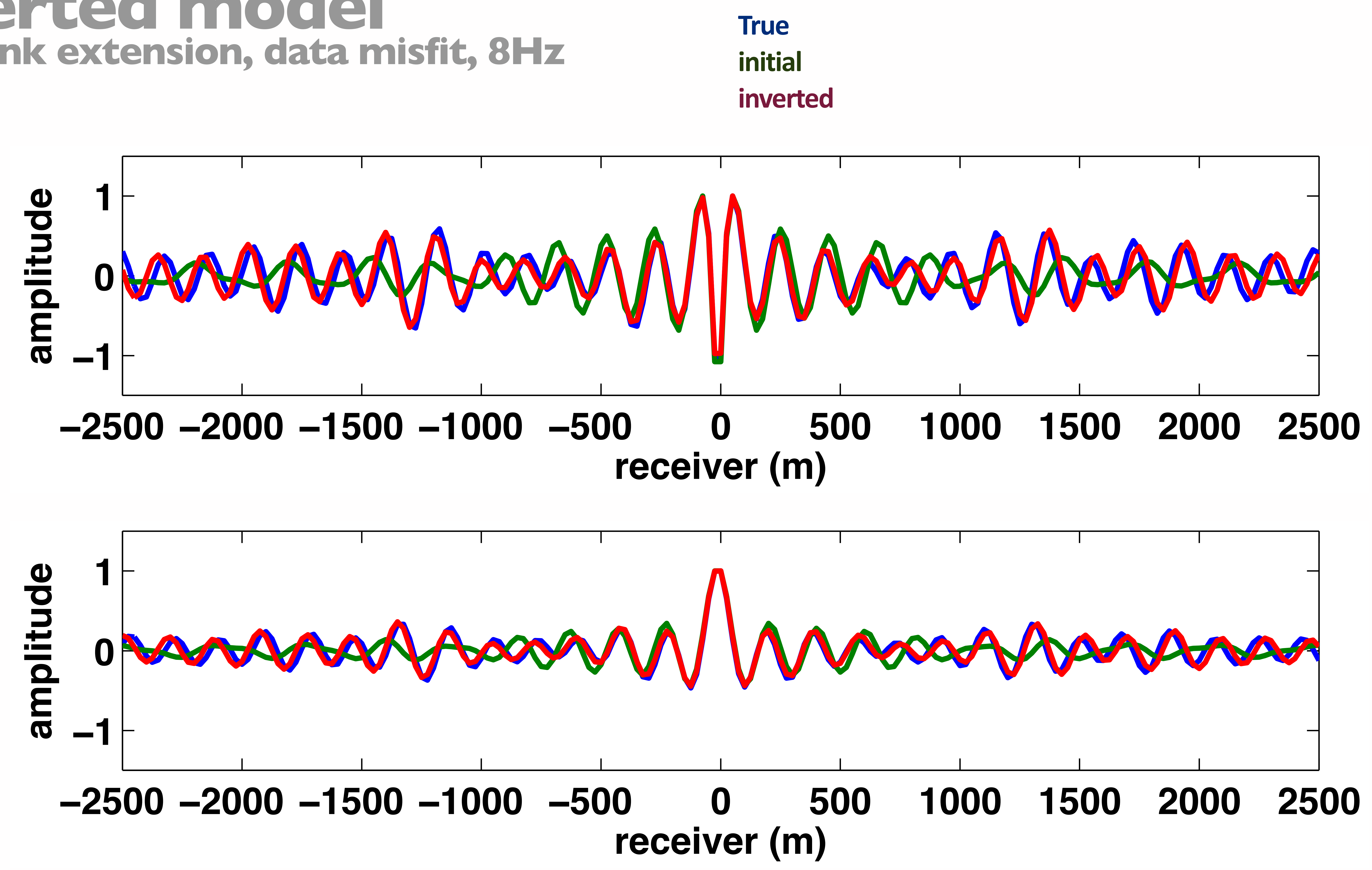


Inverted model

low-rank extension, data misfit, 5Hz



Inverted model
low-rank extension, data misfit, 8Hz



Conclusion

Low-rank extension

- extends the search space
- allows for incorporation of “starting model diversity”
- more robust w.r.t. poor starting models

Computationally feasible

Potentially a cheap way to do incorporate prior knowledge & test scenarios…

SLIM

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, ION, Petrobras, PGS, Statoil, Total SA, Sub Salt Solutions, WesternGeco, and Woodside.