Breaking *structure* - why *randomized* sampling matters

Felix J. Herrmann

SLIM
University of British Columbia

SLIM

Carry home message

Randomization of field-data *acquisition*

- can lead to *improved* wavefield reconstruction
- *new* insights how to *acquire* & *process* data

Randomization of wave-equation based *inversion*

- can lead to *fast* RTM & FWI
- *new* insights how to *invert* large data volumes

Incorporation of more *wave* physics

- can lead to *hifi* & *fast* inversions from sparsely *sampled* marine data

SINBAD Consortium

Research group of 25-30 people w/ $1.5M annual budget

- support from 13 oil & gas majors & including contractors
- NSERC matching industrial contributions
- long-term *fundamental* research with direct involvement from *faculty* in math & computer science

Recognized world leaders in

- seismic data acquisition w/ compressive sensing
- wave-equation based imaging and inversion
- large-scale (convex & stochastic) optimization w/ seismic applications

Research team

- (under)graduate students,
- PDFs, visitors, faculty, and
- staff...

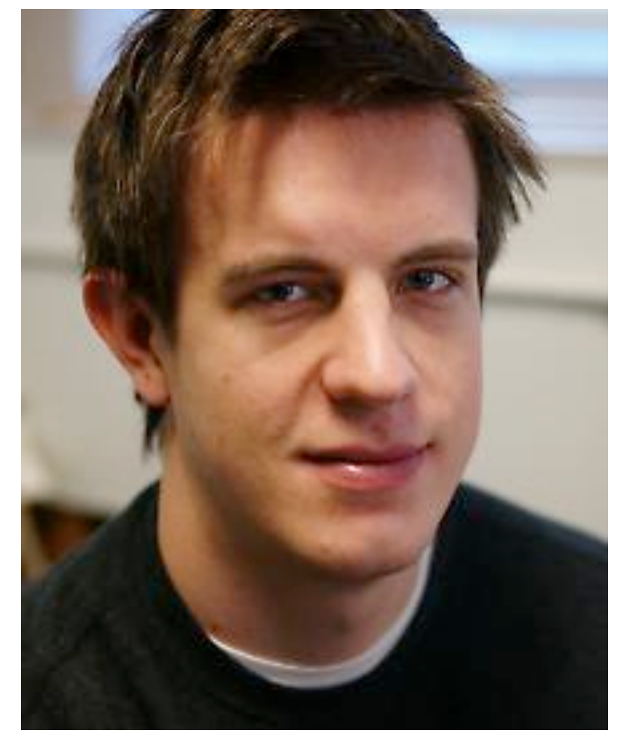
Highly *interdisciplinary* and able to *leverage* recent developments in

- exploration geophysics
- computer science (convex/stochastic optimization)
- mathematics (compressive sensing)

Alumni

Gilles Hennenfent (Chevron) – Karcher & van Weelden Award

James Johnson (Schlumberger) – local celebrity



Carson Yarham (Schlumberger) – local celebrity

Sponsors

NSERC CRSNG

BG GROUP

CNPC BGP

bp

Chevron

ConocoPhillips

BR PETROBRAS

PGS

TOTAL

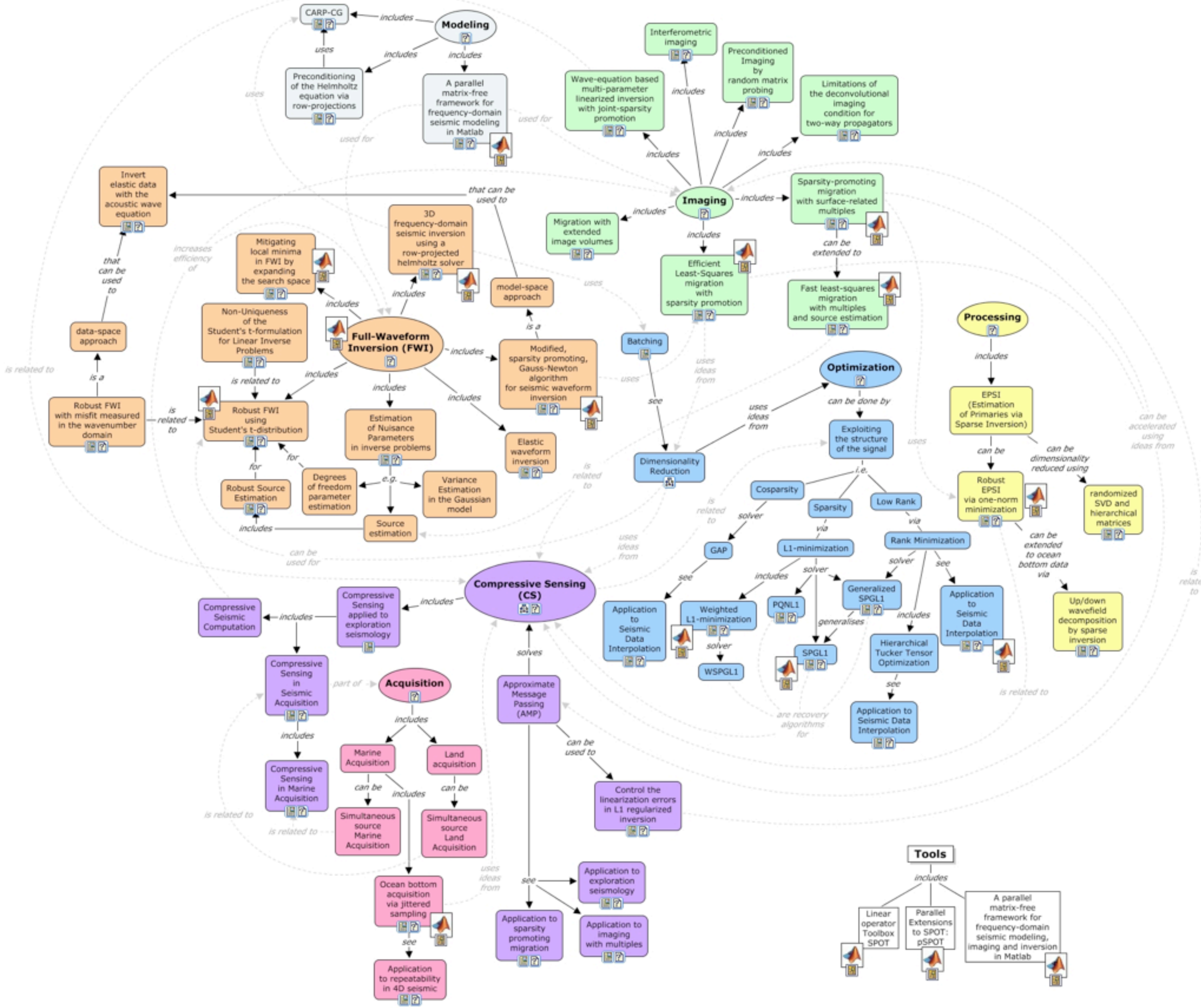
WesternGeco

woodside

ion

CGG

Statoil


Modeling
CARP-CG
Preconditioning of the Helmholtz equation via row-projections
A parallel matrix-free framework for frequency-domain seismic modeling in Matlab
Imaging
Interferometric imaging
Wave-equation based multi-parameter linearized inversion with joint-sparsity promotion
Preconditioned Imaging by random matrix probing
Limitations of the deconvolutional imaging condition for two-way propagators
Sparsity-promoting migration with surface-related multiples
Migration with extended image volumes
Efficient Least-Squares migration with sparsity promotion
Fast least-squares migration with multiples and source estimation
Invert elastic data with the acoustic wave equation
3D frequency-domain seismic inversion using a row-projected helmholtz solver
Mitigating local minima in FWI by expanding the search space
model-space approach
Non-Uniqueness of the Student's t-formulation for Linear Inverse Problems
Full-Waveform Inversion (FWI)
Modified, sparsity promoting, Gauss-Newton algorithm for seismic waveform inversion
data-space approach
Robust FWI with misfit measured in the wavenumber domain
Robust FWI using Student's t-distribution
Estimation of Nuisance Parameters in inverse problems
Elastic waveform inversion
Robust Source Estimation
Degrees of freedom parameter estimation
Variance Estimation in the Gaussian model
Source estimation
Batching
Dimensionality Reduction
Optimization
Exploiting the structure of the signal
Cosparsity
Sparsity
Low Rank
GAP
L1-minimization
Rank Minimization
Application to Seismic Data Interpolation
Weighted L1-minimization
PQNL1
Generalized SPGL1
SPGL1
WSPGL1
Hierarchical Tucker Tensor Optimization
Application to Seismic Data Interpolation
Application to Seismic Data Interpolation
Processing
EPSI (Estimation of Primaries via Sparse Inversion)
Robust EPSI via one-norm minimization
randomized SVD and hierarchical matrices
Up/down wavefield decomposition by sparse inversion
Compressive Sensing (CS)
Compressive Sensing applied to exploration seismology
Compressive Seismic Computation
Compressive Sensing in Seismic Acquisition
Compressive Sensing in Marine Acquisition
Acquisition
Marine Acquisition
Land acquisition
Simultaneous source Marine Acquisition
Simultaneous source Land Acquisition
Ocean bottom acquisition via jittered sampling
Application to repeatability in 4D seismic
Approximate Message Passing (AMP)
Control the linearization errors in L1 regularized inversion
Application to exploration seismology
Application to sparsity promoting migration
Application to imaging with multiples
Tools
Linear operator Toolbox SPOT
Parallel Extensions to SPOT: pSPOT
A parallel matrix-free framework for frequency-domain seismic modeling, imaging and inversion in Matlab
includes
uses
used for
that can be used to
can be extended to
can be done by
uses ideas from
increases efficiency of
is related to
is a
for
e.g.
see
solver
via
i.e.
generalises
solves
part of
can be
can be used to
can be dimensionality reduced using
can be extended to ocean bottom data via
can be accelerated using ideas from
are recovery algorithms for

SLIM

HPC–initiatives

$600k grant from NSERC:

- replaces & extends our current HPC capacity (0.01 petaflop)

$30M International Inversion Initiative in Brazil:

- BG Group, Imperial College London (Mike Warner) & 2 Brazilian Institutions
- semi-exclusive access to $10M facility (0.2 petaflop, 16k cores)
- true-cost of research model (Scientists w/o Borders)

Outcome:

- increased research capacity
- industrialization & development of 3D workflows on field data
- testing & evaluation of SINBAD algorithms

Industry uptake

Seismic data acquisition:

- BP uses our *sparse* recovery technology in *simultaneous* marine acquisition
- ConocoPhilips & WesternGeco employ our *randomized* sampling ideas

Seismic data processing:

- Shell implemented our *robust* estimation of primaries by sparse inversion

Full waveform inversion:

- Total & WesternGeco implemented our *robust* FWI & *source* estimation
- WesternGeco implemented our *batching* FWI in their *production* codes

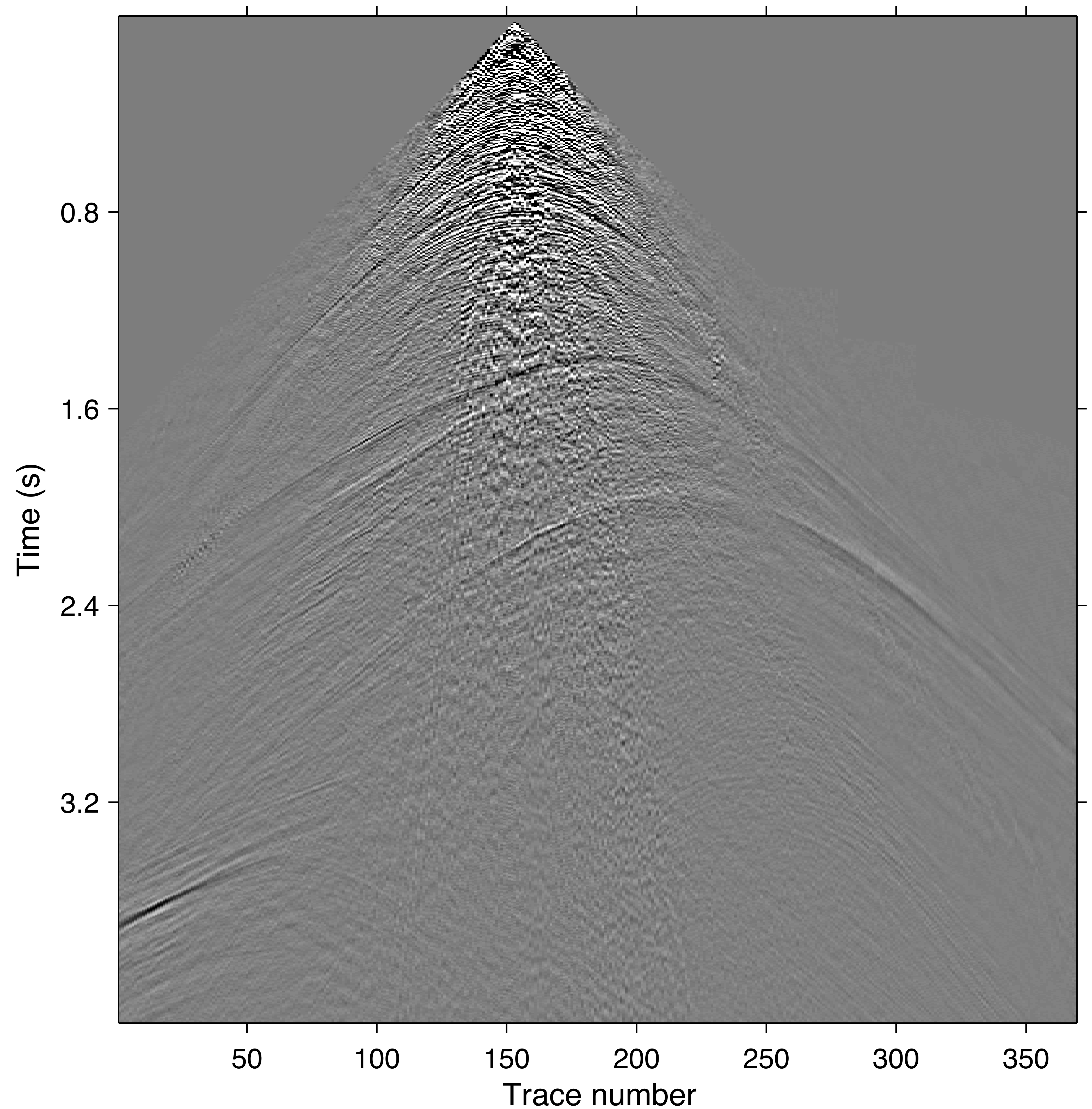
A teaser: SLIM's findings on BP's Machar dataset

Ning Tu, Tim Lin, Zhilong Fang, with contribution from many other SLIM members

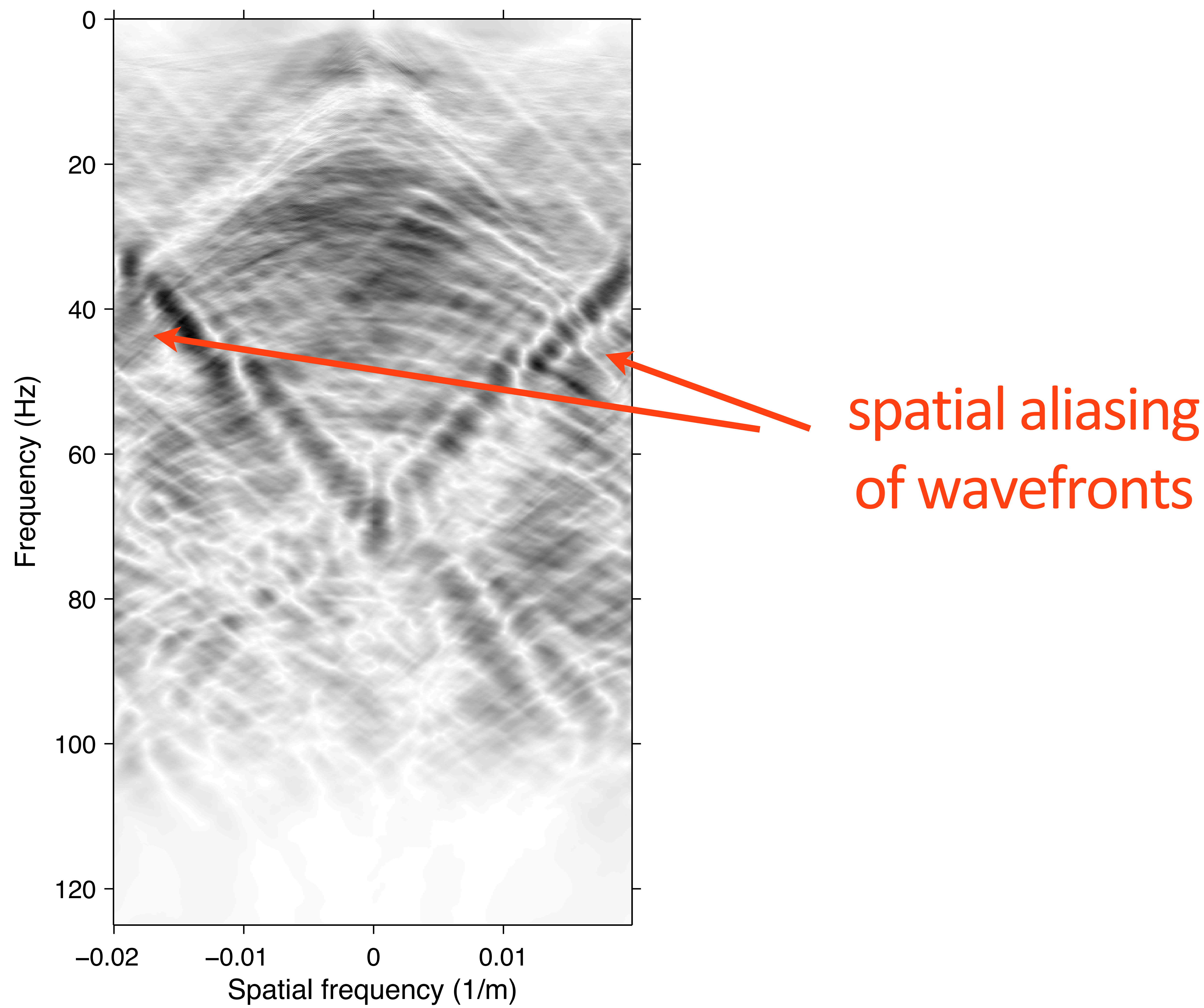
SLIM
University of British Columbia

Original seismic data

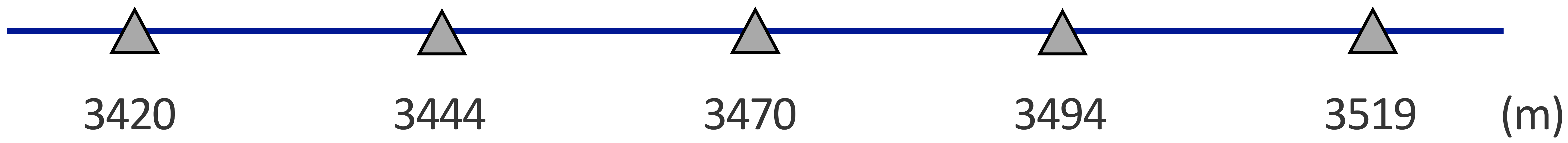
about 25m spatial sampling, 370 traces



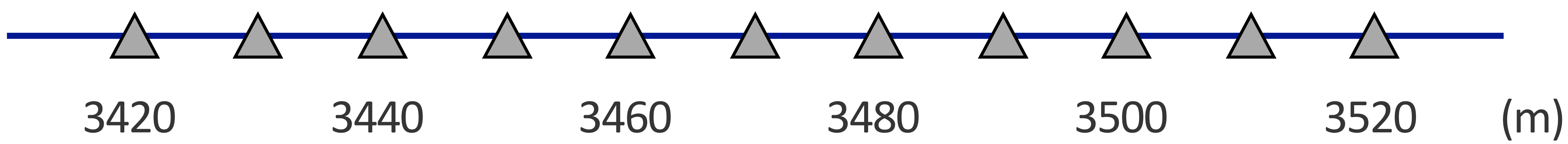
F-K (2D Fourier) spectrum



Desired sampling

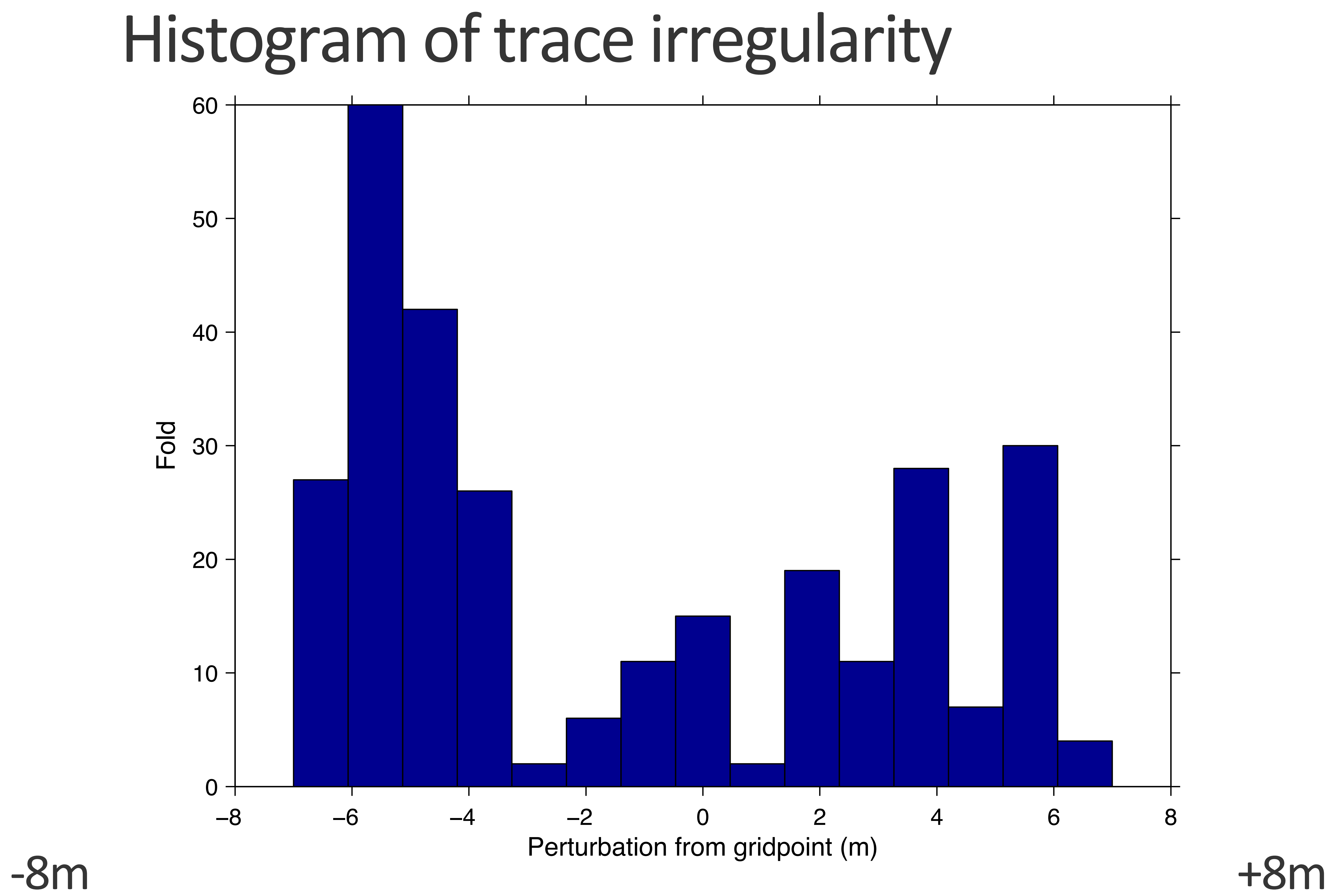


Original sampling (irregular), ~25m spacing



Desired sampling after interpolation, 10m spacing

Exploiting natural irregular sampling



SLIM

FDCT vs. NFDCT

Fast Discrete Curvelet Transform

x

t Input $\xrightarrow{\text{2-D FFT}}$

k

f curvelet tilling $\xrightarrow{\text{2-D IFFT}}$ curvelet coefficients

$\mathbf{S}$

Non-equispaced Fast Discrete Curvelet Transform

x

t Input $\xrightarrow[\text{1-D } \textit{NFFT} \text{ on x}]{\text{1-D FFT on t}}$

k

f curvelet tilling $\xrightarrow[\text{1-D } \textit{INFFT} \text{ on k}]{\text{1-D IFFT on f}}$ curvelet coefficients

S

Gilles Hennenfent, Lloyd Fenelon, and Felix J. Herrmann, "Nonequispaced curvelet transform for seismic data reconstruction: A sparsity-promoting approach", *Geophysics*, vol. 75, p. WB203-WB210, 2010

SLIM

Sparse recovery

Exploit *curvelet*-domain *sparsity* of seismic data

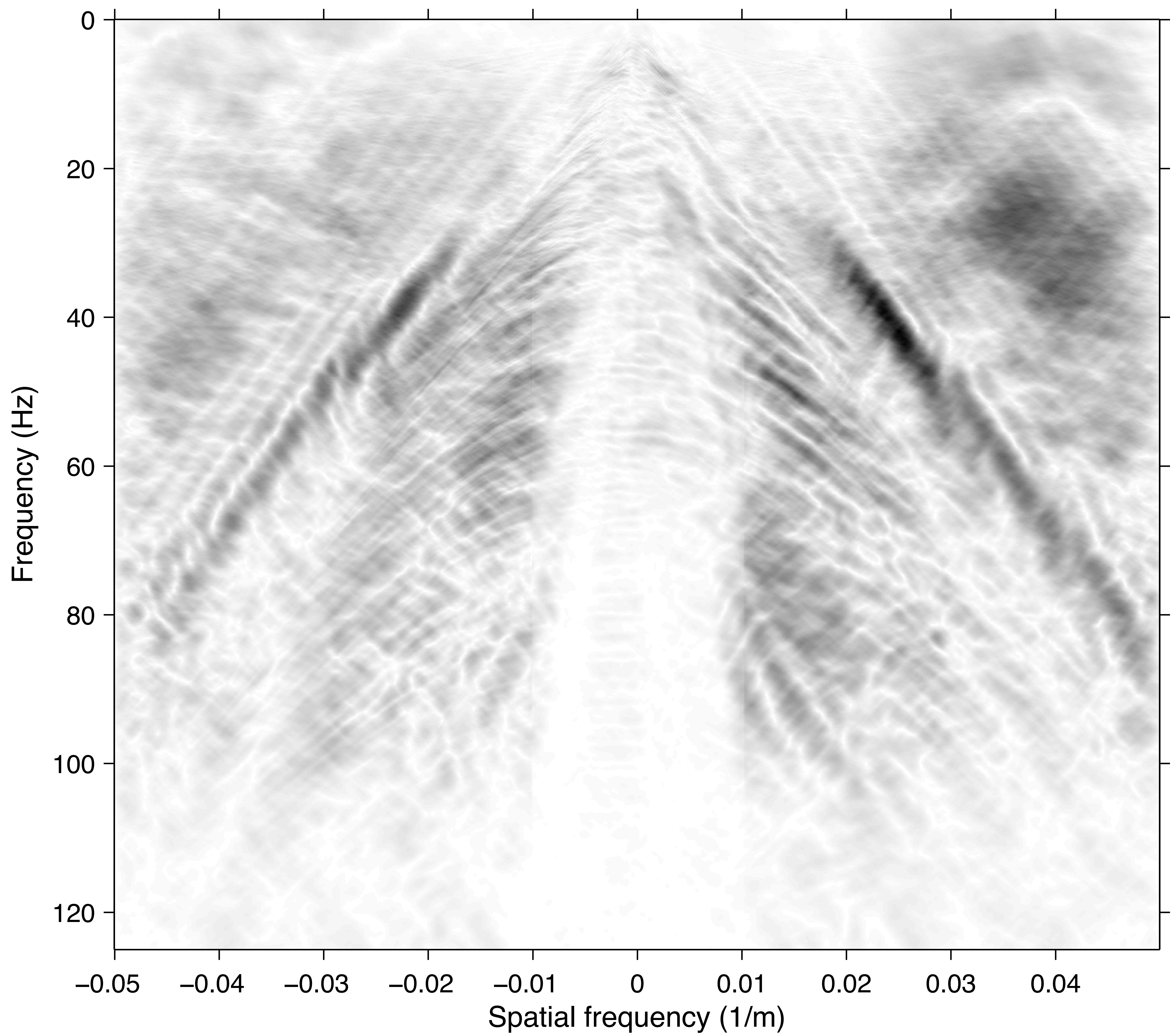
Sparsity-promoting program:

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}} \underbrace{\|\mathbf{x}\|_1}_{\text{support detection}} \quad \text{subject to} \quad \underbrace{\mathbf{A}\mathbf{x} = \mathbf{b}}_{\text{data-consistent amplitude recovery}}$$

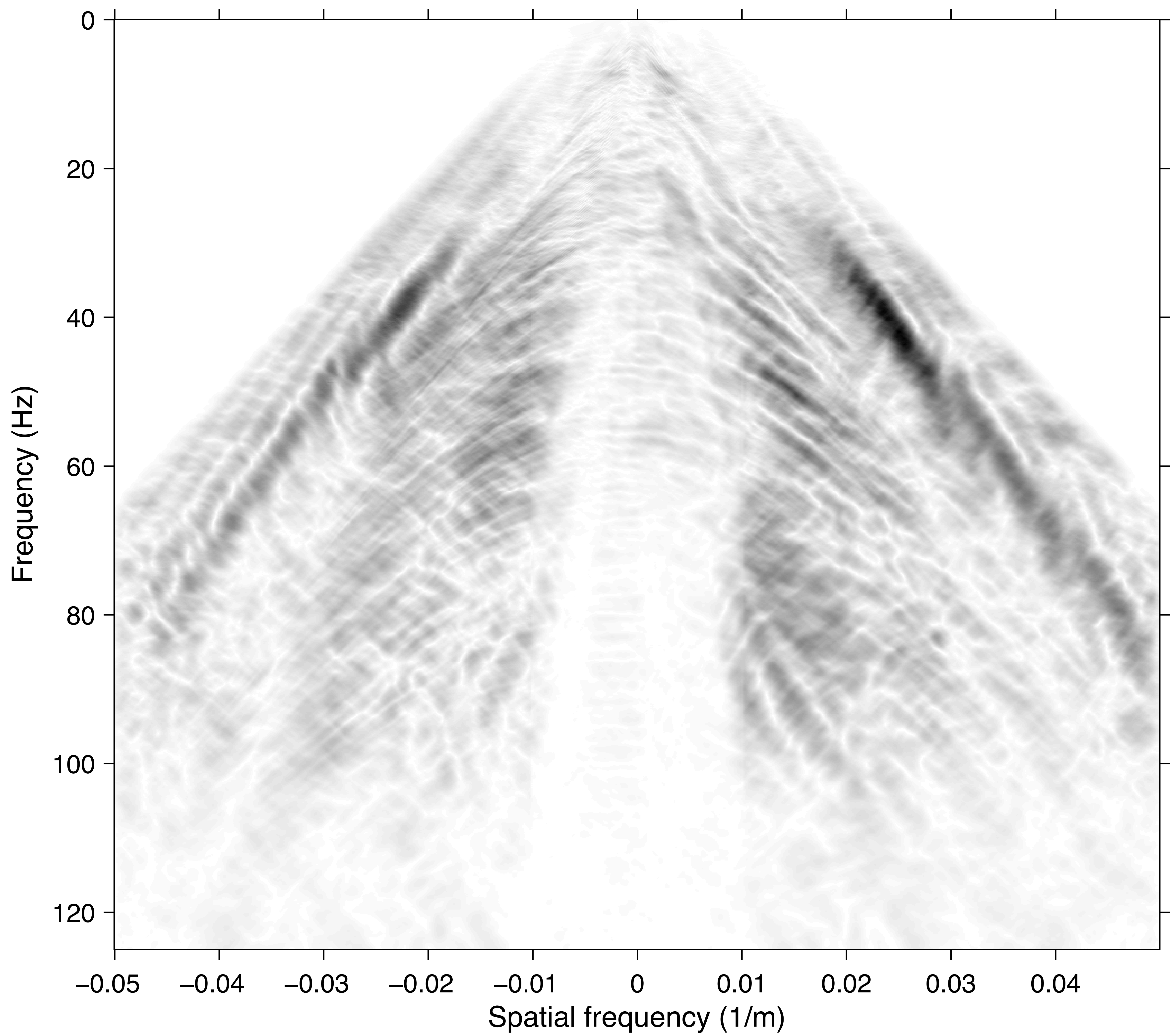
Forward model: $\underbrace{\mathbf{A} = \mathsf{S}^H}_{\text{synthesis to } \textit{irregular} \text{ grid}}$

Recovered data: $\underbrace{\tilde{\mathbf{d}} = \mathbf{S}^H \tilde{\mathbf{x}}}_{\text{synthesis to } \textit{regular} \text{ grid}}$

F-K spectrum after interpolation

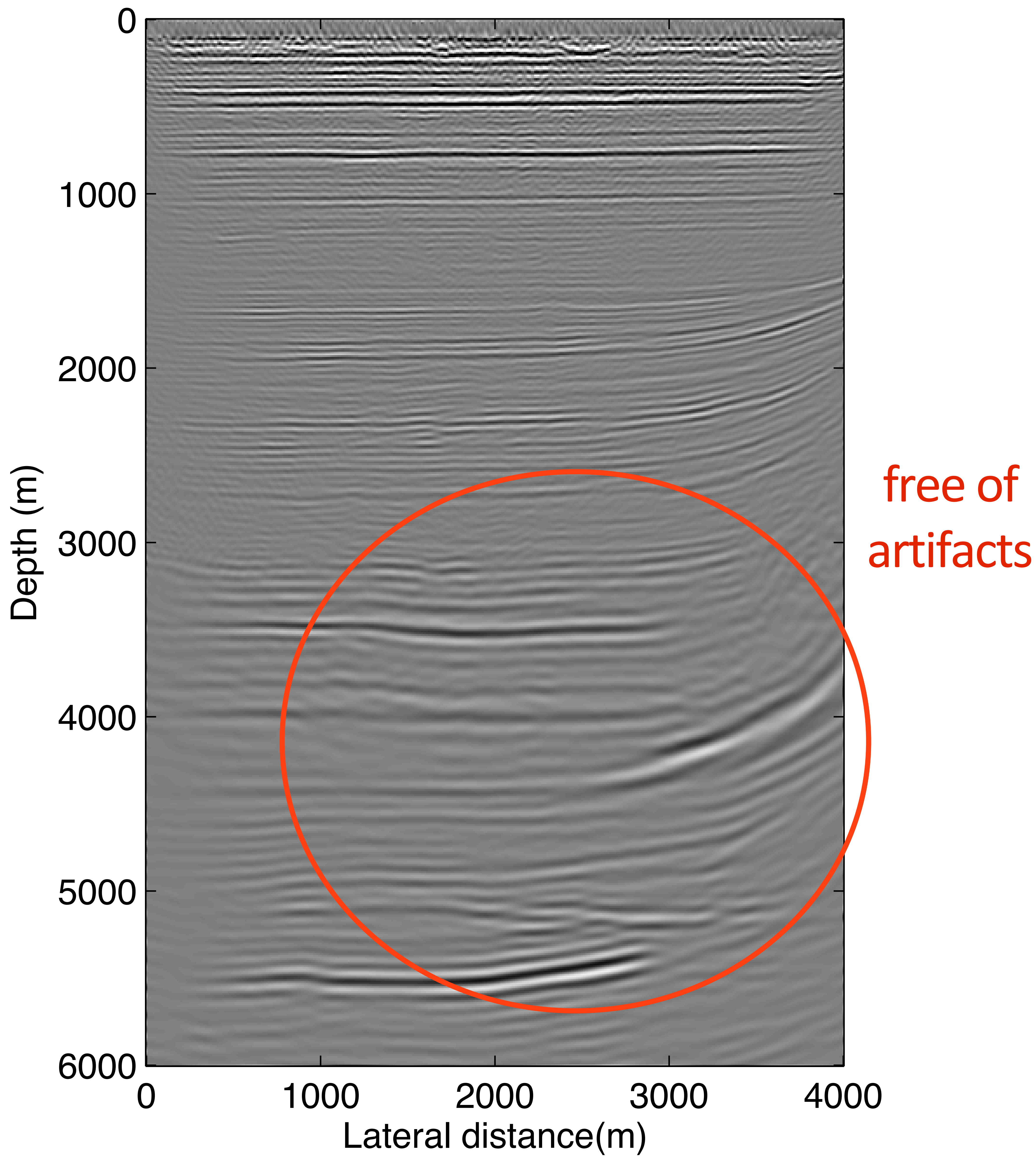


F-K spectrum masking



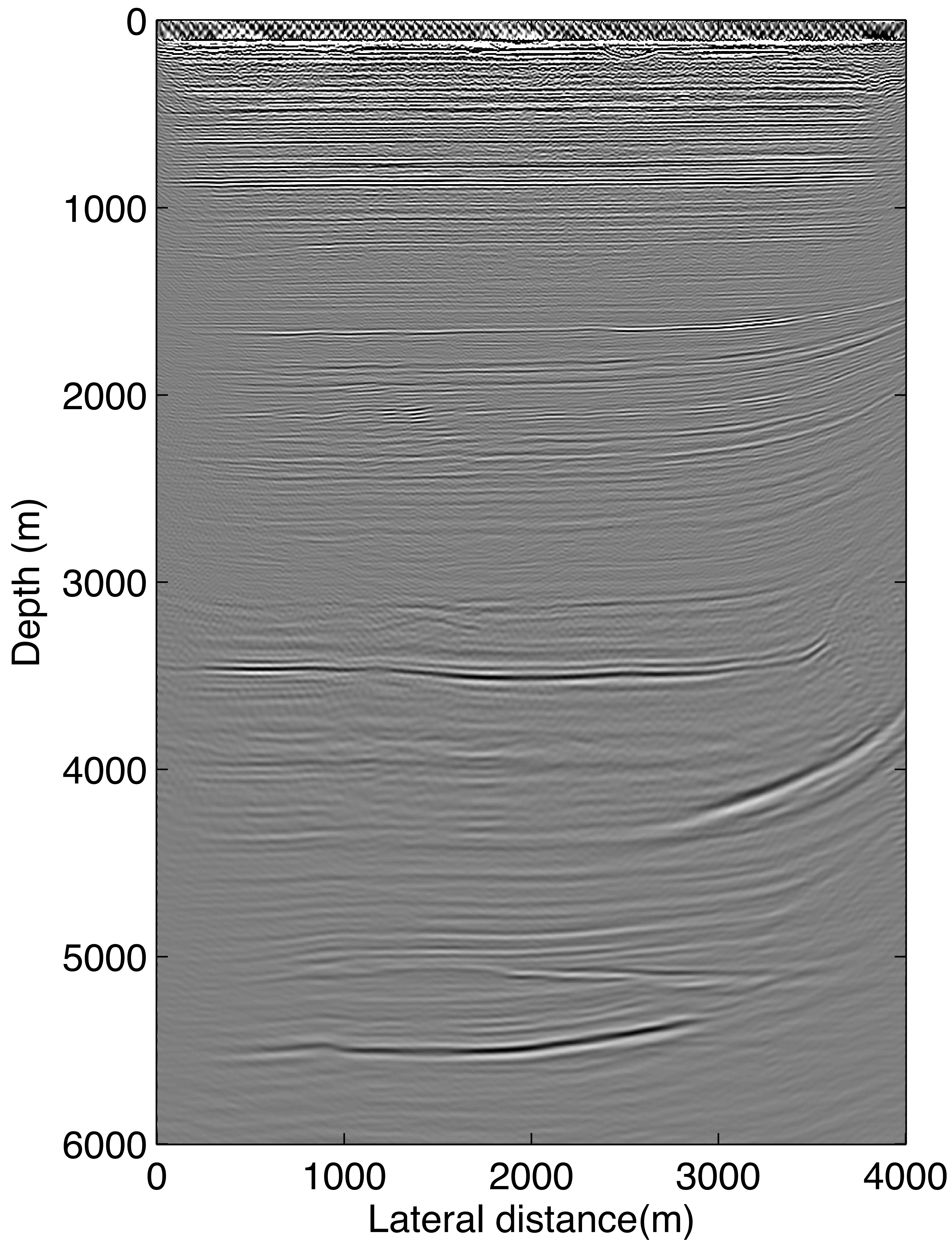
Reverse-time migrated image

[all freq. in 5-30Hz, all sources, 10m grid distance]



Reverse-time migrated image

[all freq. in 5-60Hz, all sources, 5m grid distance]



SLIM

Felix J. Herrmann, Michael P. Friedlander, and Ozgur Yilmaz, "Fighting the Curse of Dimensionality: Compressive Sensing in Exploration Seismology", *Signal Processing Magazine, IEEE*, vol. 29, p. 88-100, 2012
Felix J. Herrmann, "Randomized sampling and sparsity: Getting more information from fewer samples", *Geophysics*, vol. 75, p. WB173-WB187, 2010

SLIM

Compressive sensing *paradigm*

Find representations that reveal _structure_

- *transform*-domain *sparsity* (e.g., Fourier, curvelets, etc.)
- *low*-rank matrices in midpoint-offset domain

Sample to _break_ the _structure_

- *randomized* acquisition (e.g., *jittered* sampling, time *dithering*, *encoding*, etc.)
- destroy *sparsity* or *low*-rank *structure*

Recover _structure_ by promoting

- *sparsity* via *one*-norm *minimization*
- *low* rank via *nuclear*-norm minimization (one-norm singular values)

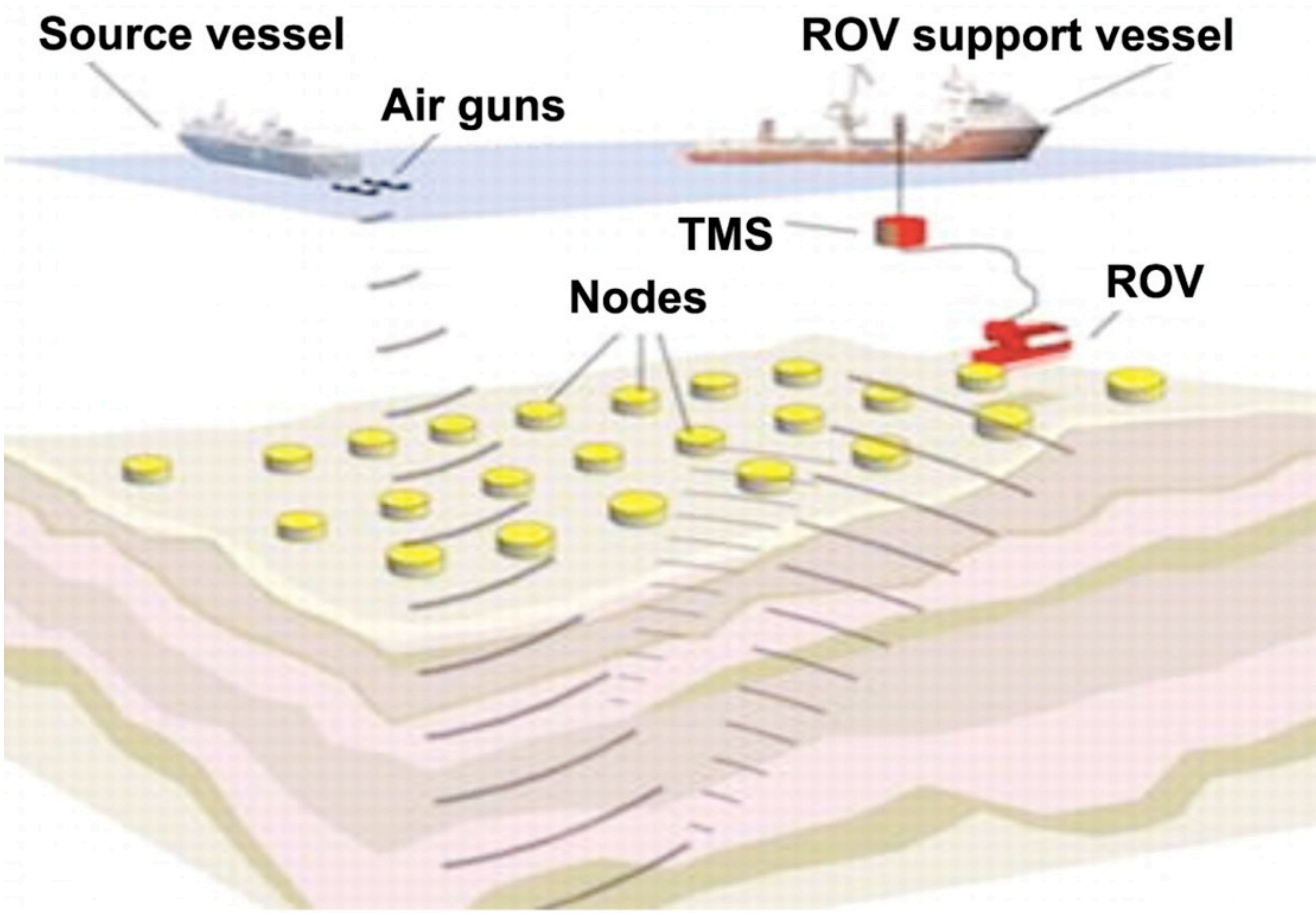
Randomized marine acquisition

Curt da Silva, Haneet Wason, and Felix Oghenekohwo

SLIM

University of British Columbia

Ocean-bottom node acquisition



http://tle.geoscienceworld.org/content/26/4/494/F4.large.jpg

Challenges

Translate *esotheric* mathematics into physically *realizable* acquisitions

Fundamentally rethink how we *sample* & *process* seismic data

Challenges

Translate *esotheric* mathematics into physically *realizable* acquisitions

Fundamentally rethink how we *sample* & *process* seismic data

- **by no longer insisting on Nyquist**

Challenges

Translate *esotheric* mathematics into physically *realizable* acquisitions

Fundamentally rethink how we *sample* & *process* seismic data

- **by no longer insisting on Nyquist**
- **by *randomizing* acquisition**

Challenges

Translate *esotheric* mathematics into physically *realizable* acquisitions

Fundamentally rethink how we *sample* & *process* seismic data

- **by no longer insisting on Nyquist**
- **by *randomizing* acquisition**

Examples in acquisition:

- *Tensor* recovery w/ Hierarchical Tucker (Curt da Silva)
- Time-*jittered* marine acquisition (Haneet Wason)
- Time *lapse* w/o repetition (Felix Oghenekohwo)

Tensor completion w/ Hierarchical Tucker

SLIM

University of British Columbia

Objective

Acquire *dense* WAZ & long-offset surveys w/ *limited* # of nodes.

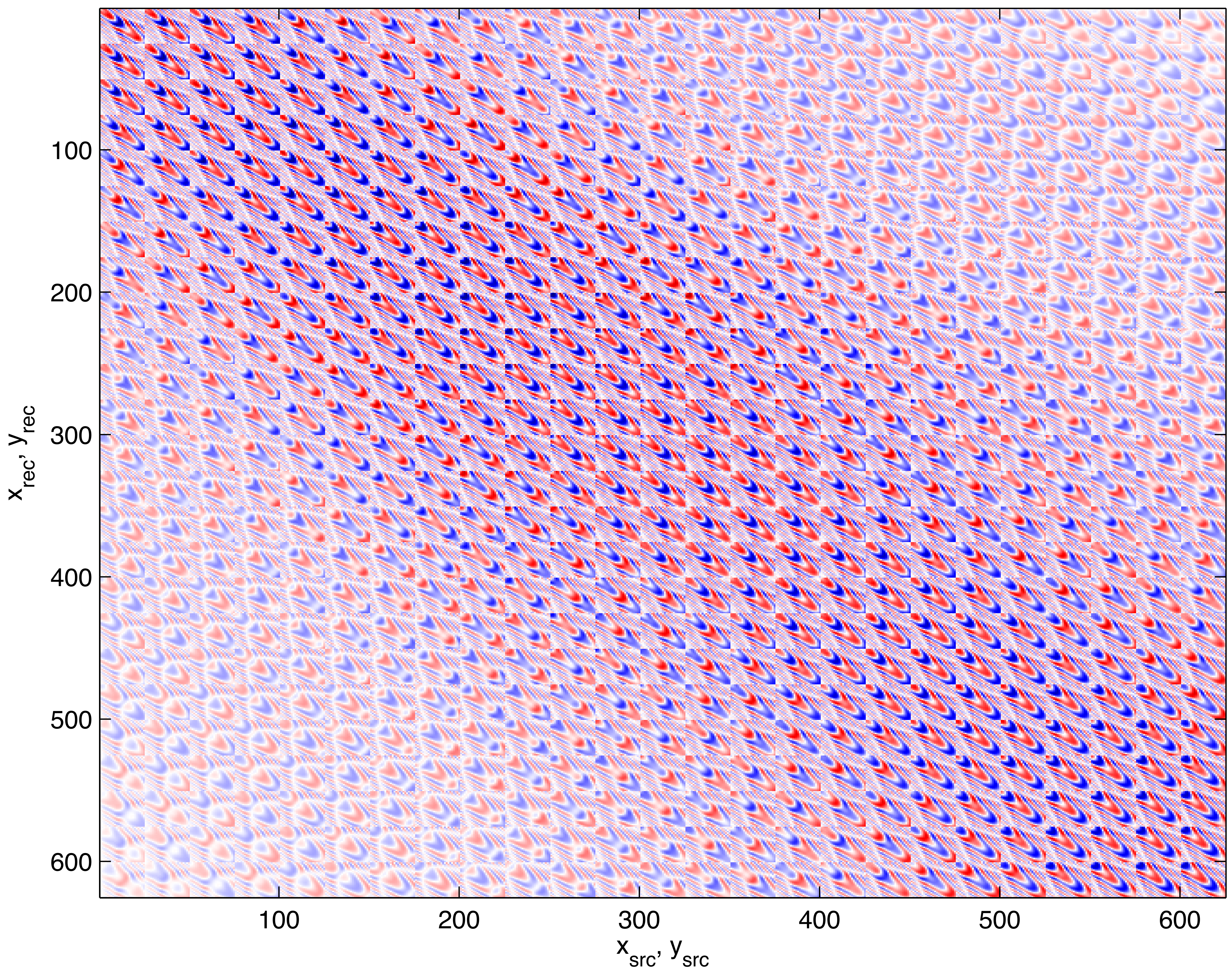
Questions:

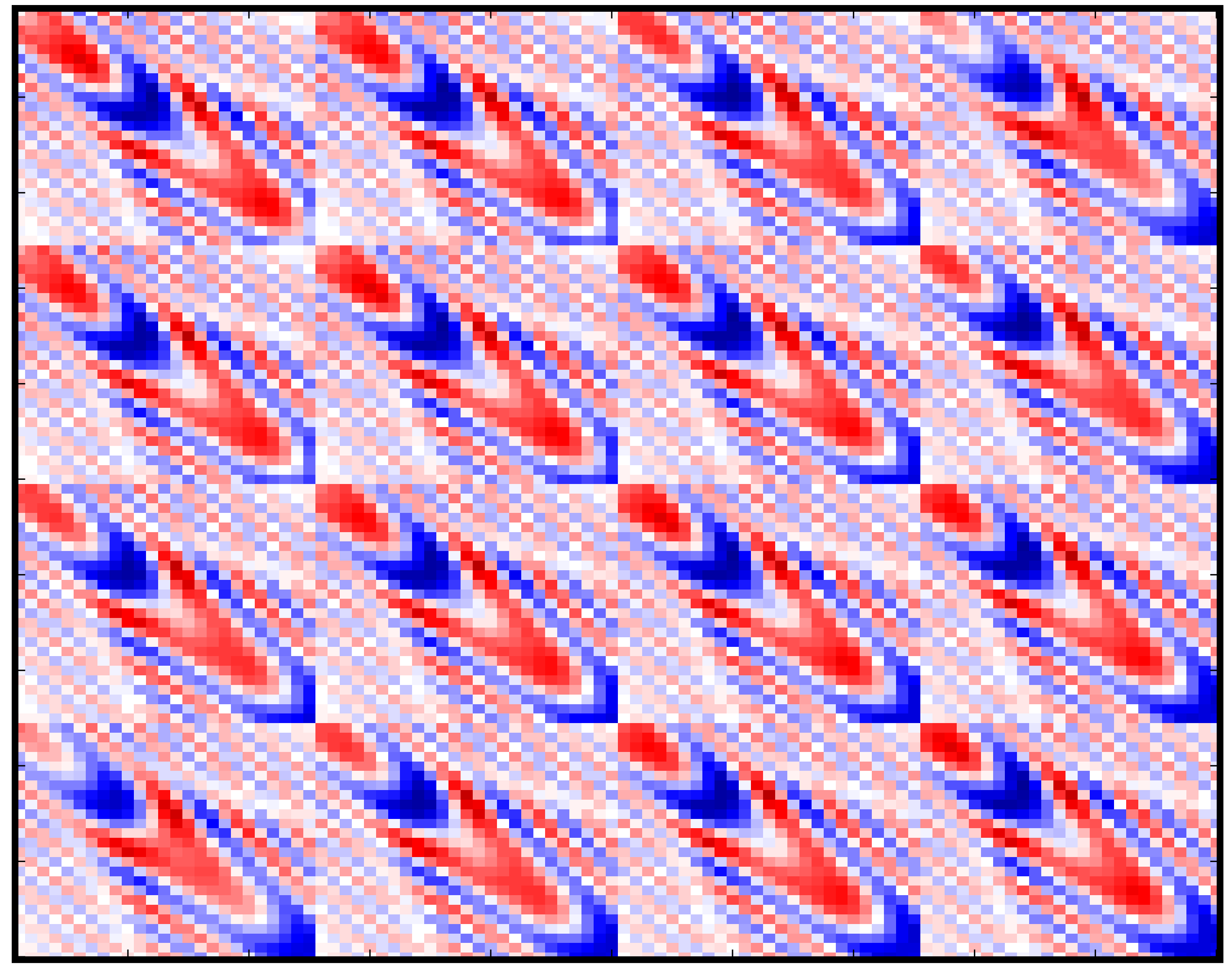
What *structure* can we *exploit* & how can we *break* it?

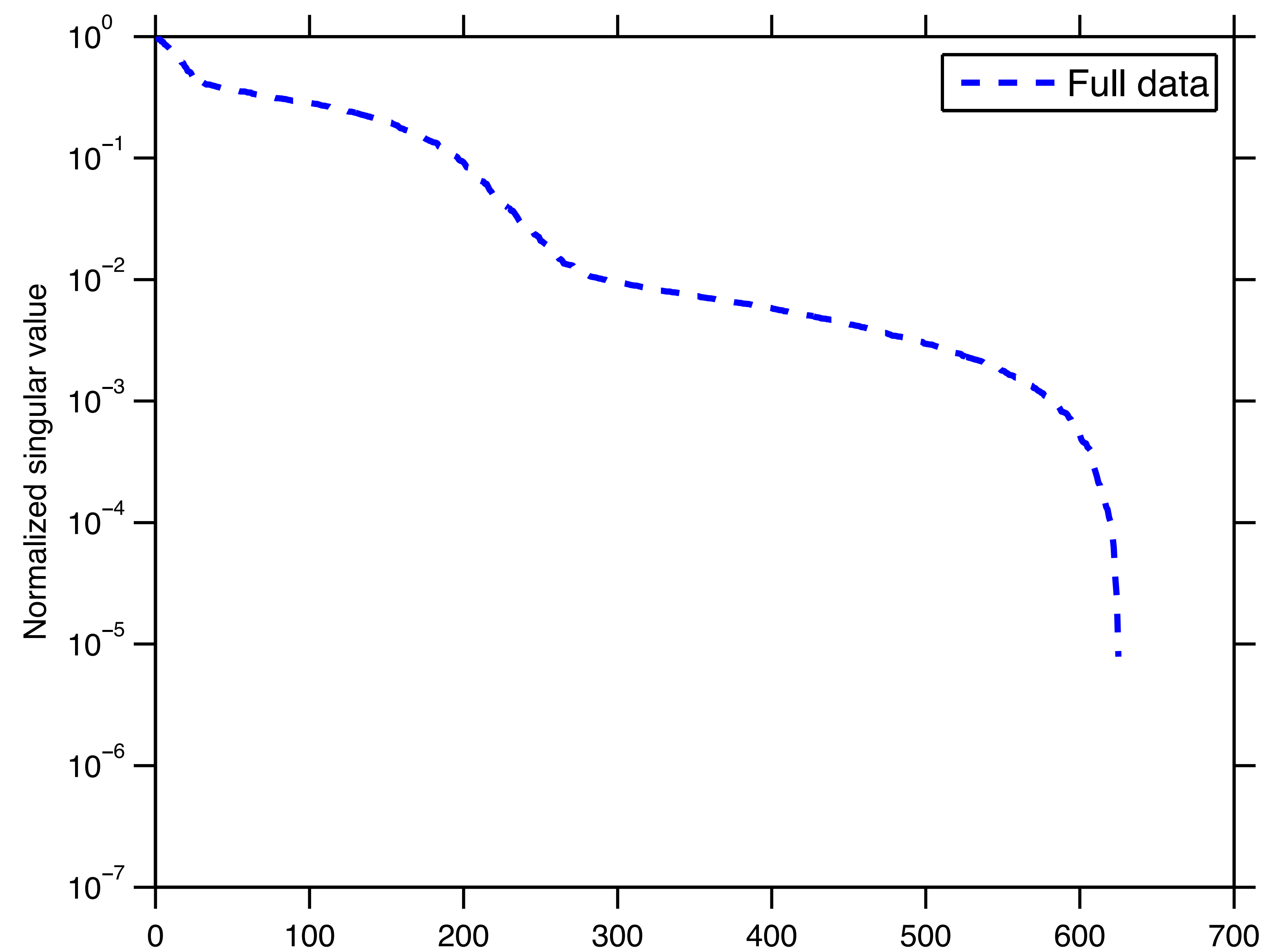
What is the best "*domain*" to exploit *low*-rank *properties*?

What is the *most* appropriate recovery technique?

Matricizations

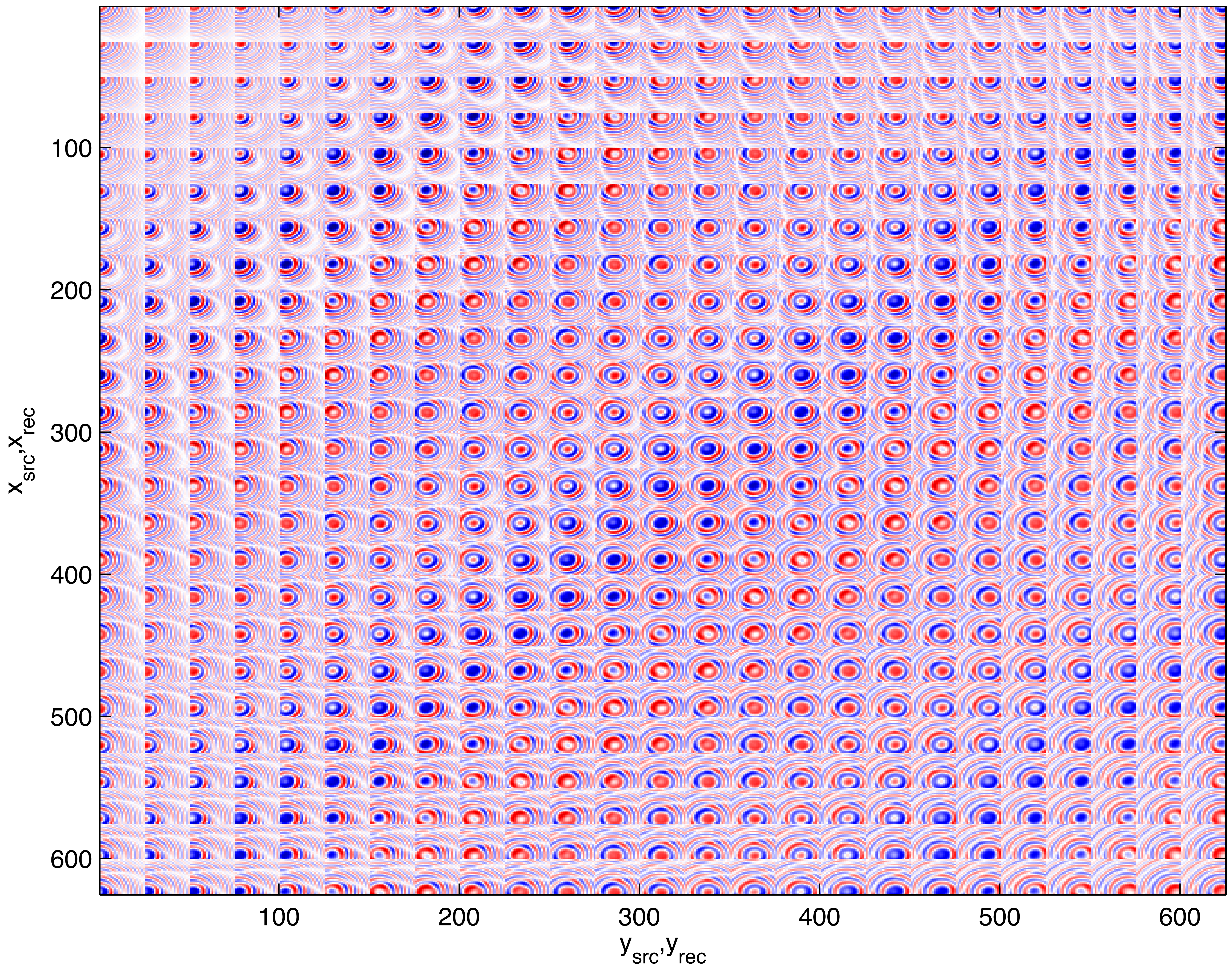


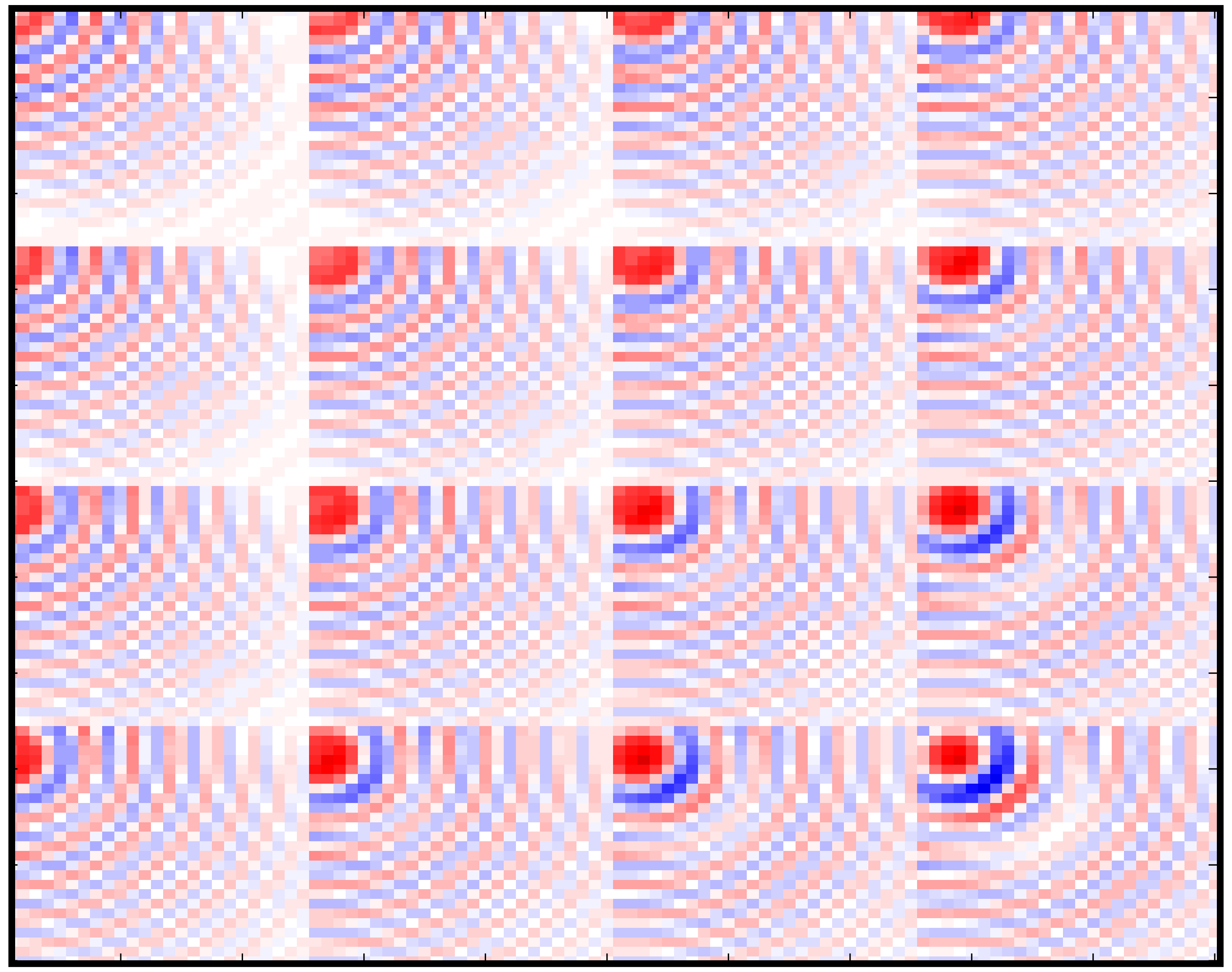


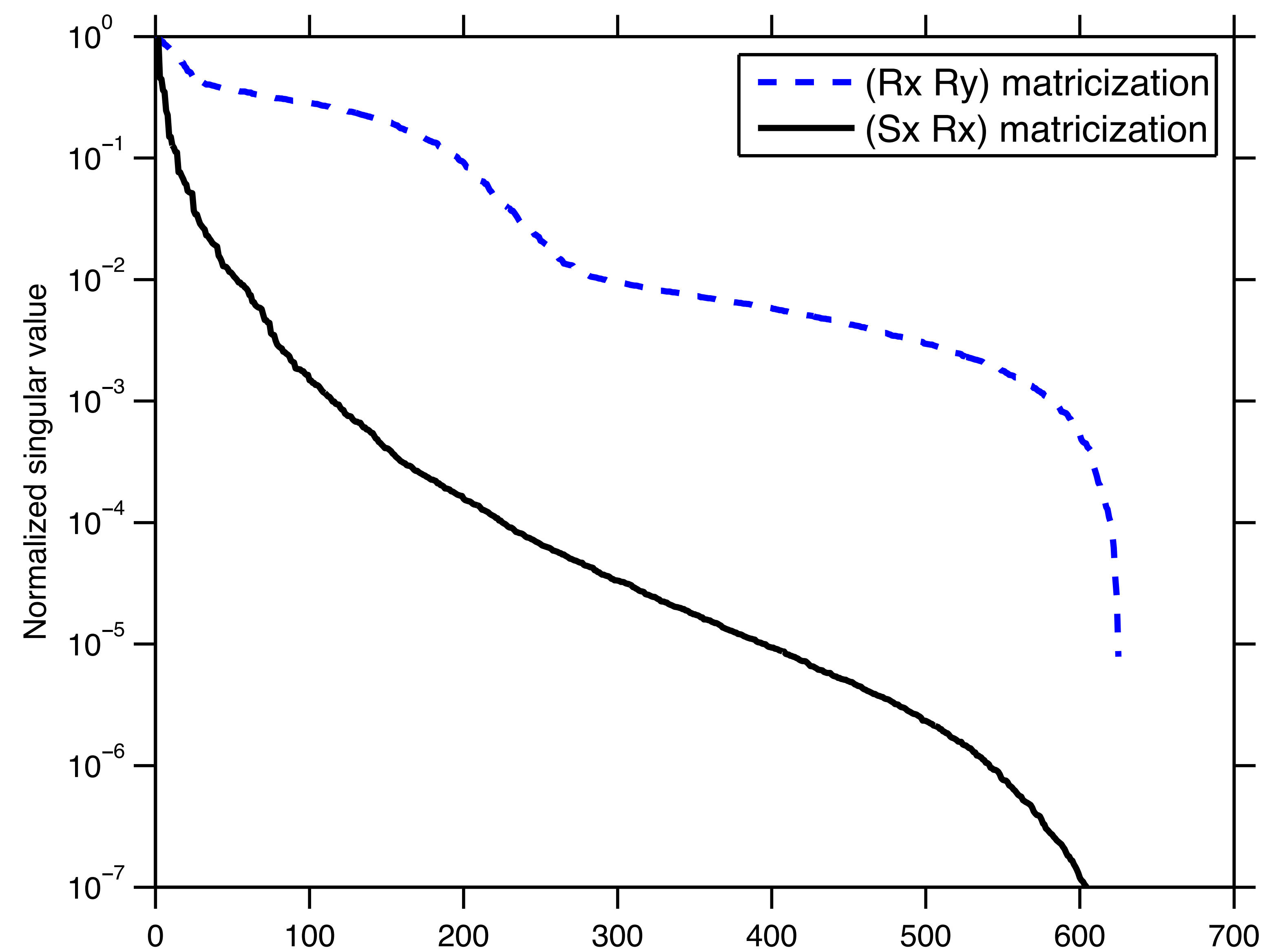


(Rec x, Rec y) matricization - Canonical ordering

Matricizations



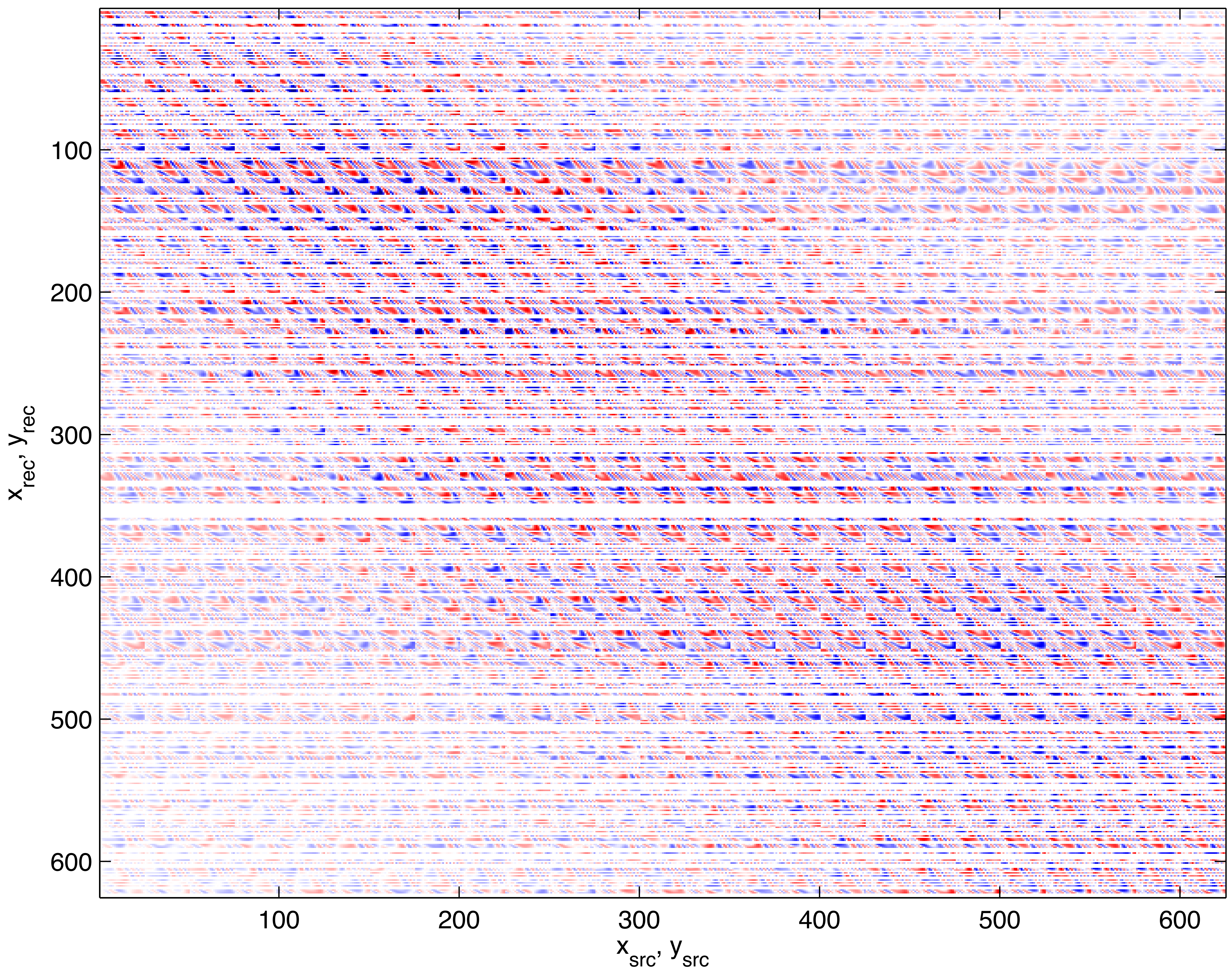


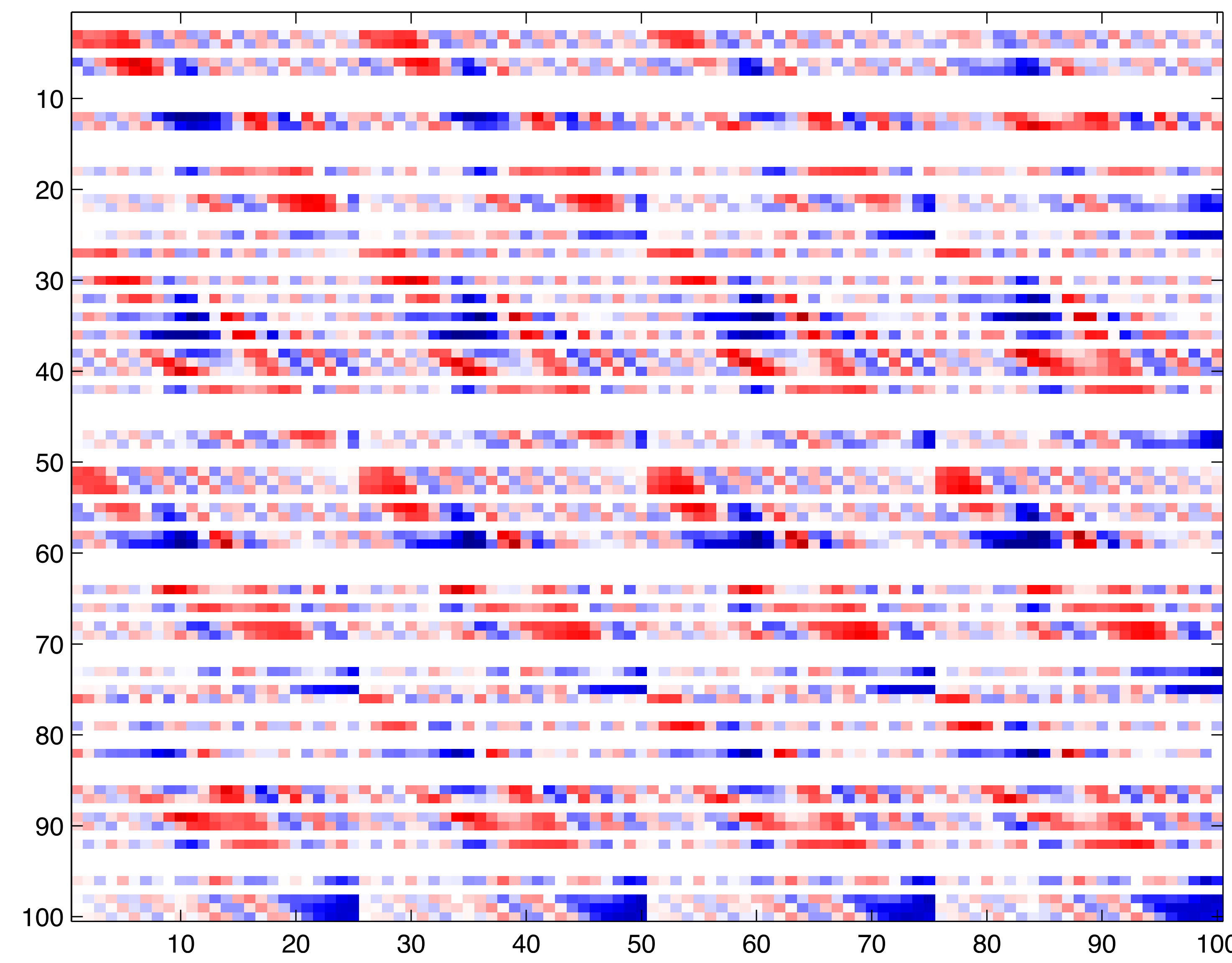


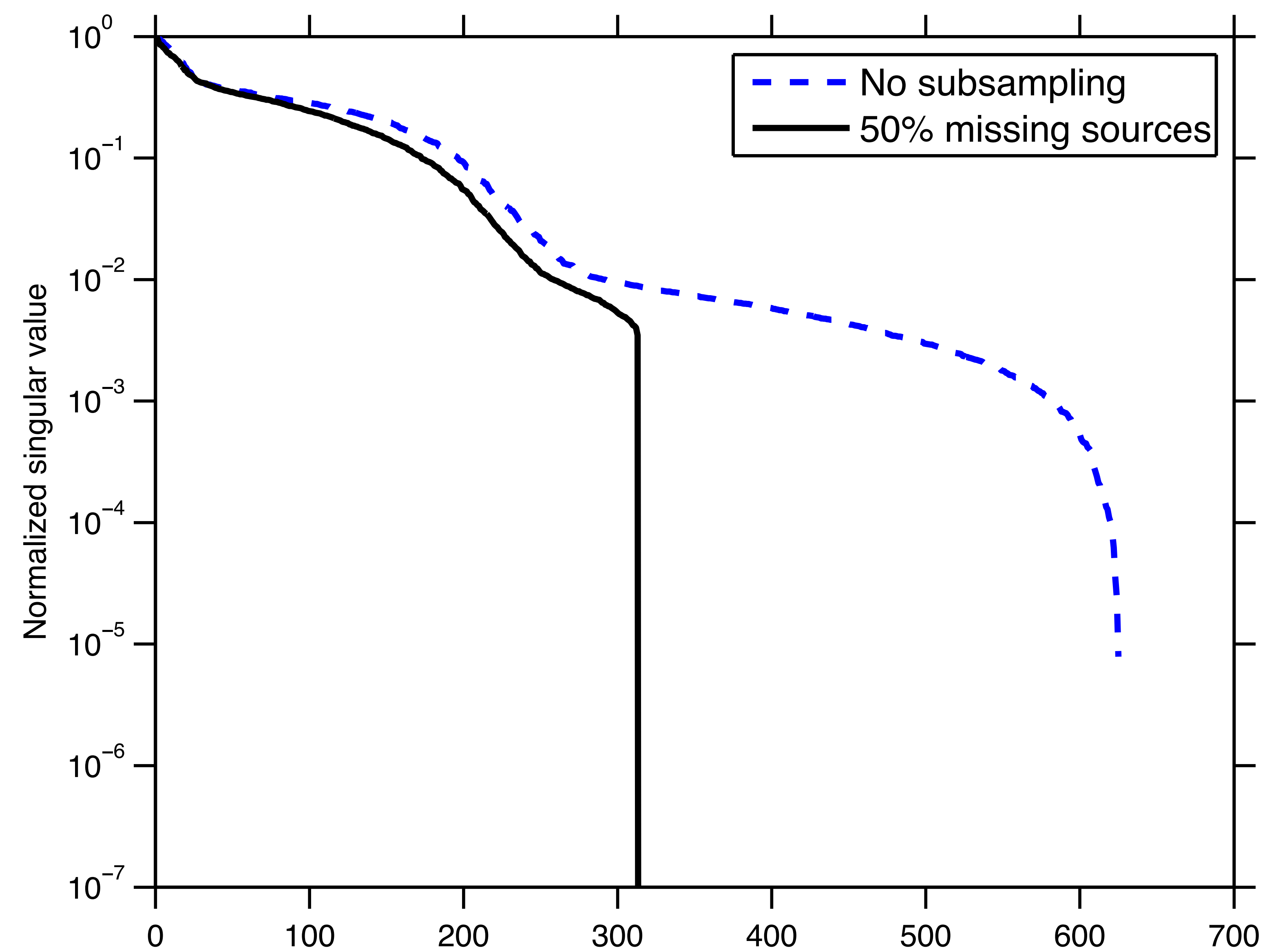
(Src x, Rec x) matricization - Noncanonical ordering

Realistic acquisition

50% random receivers removed

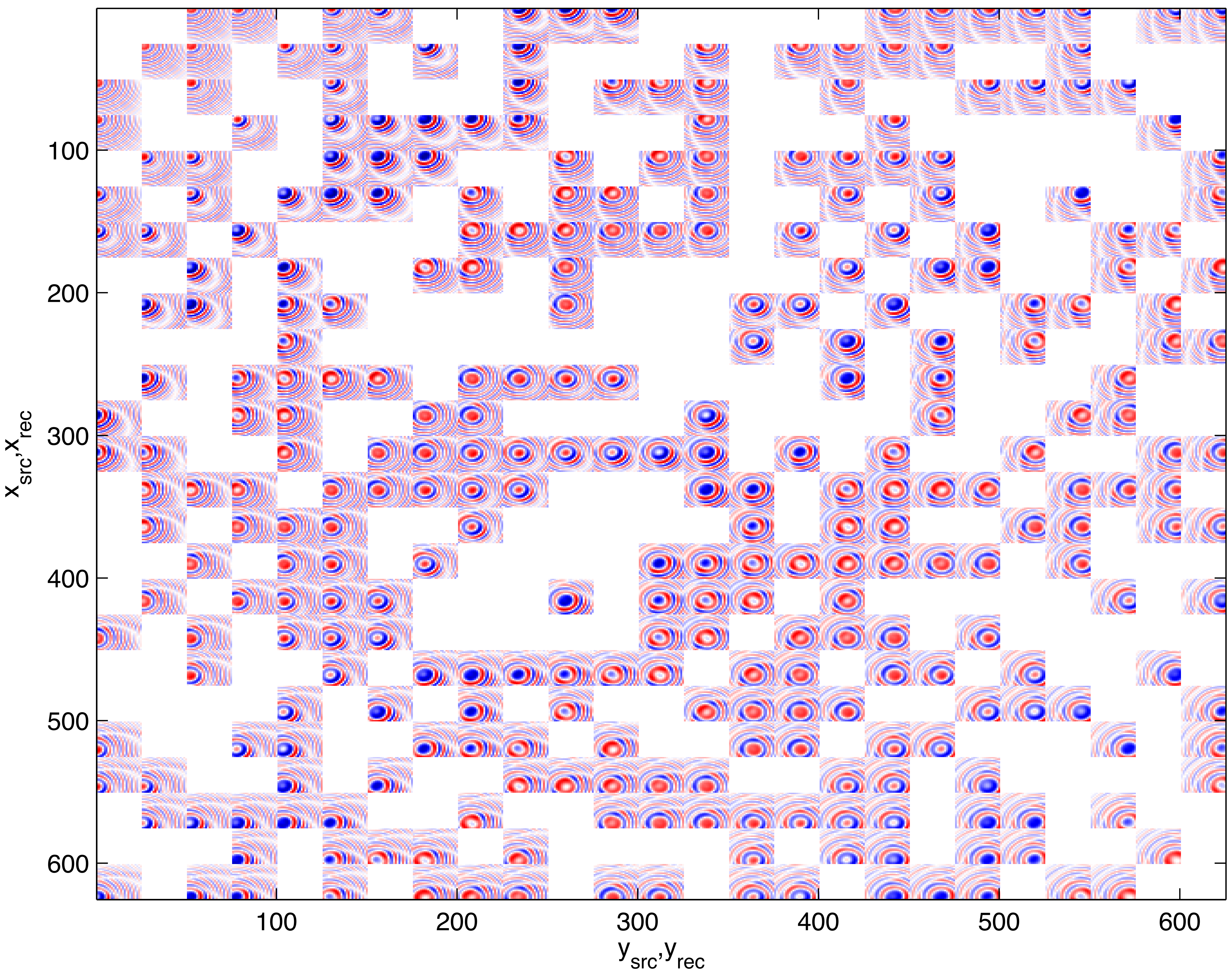


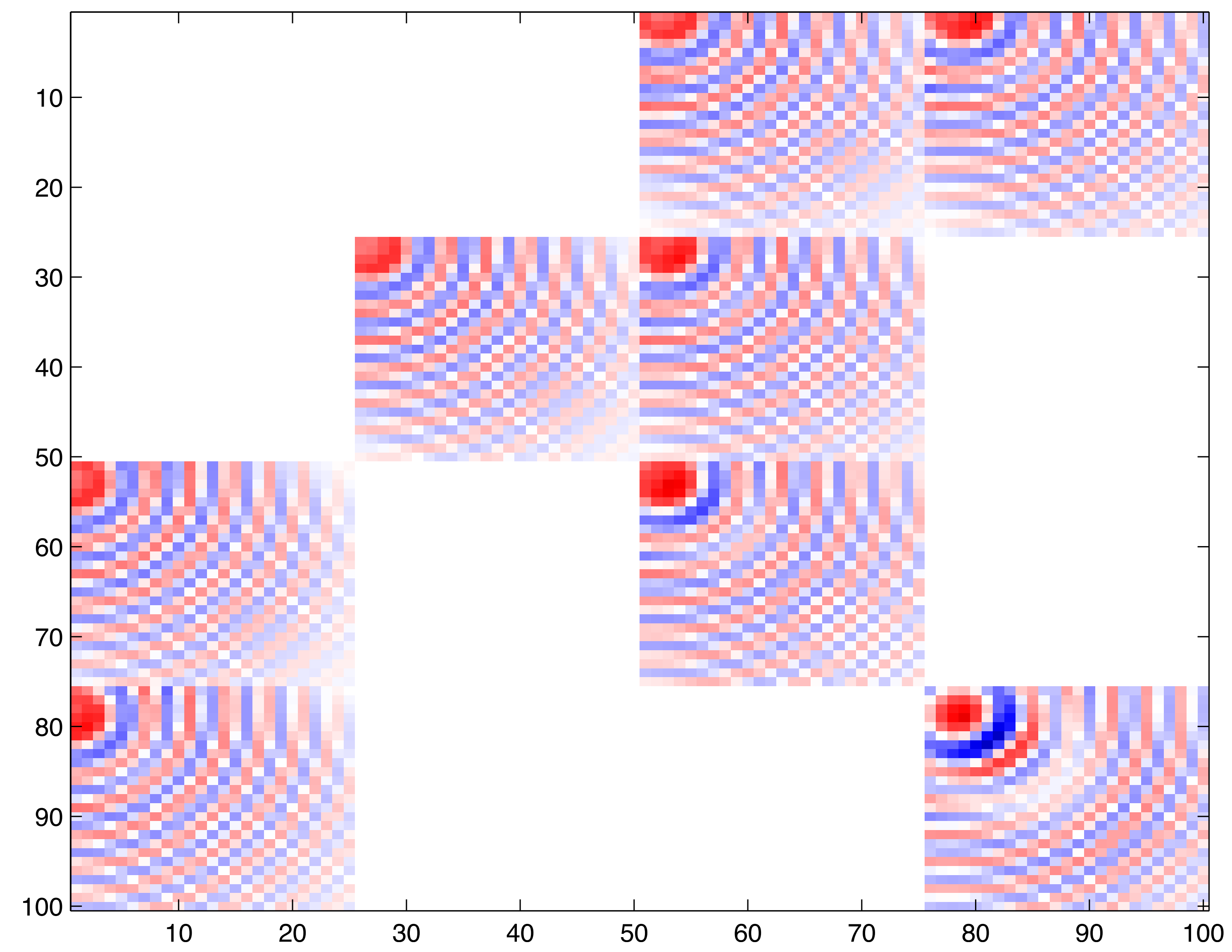


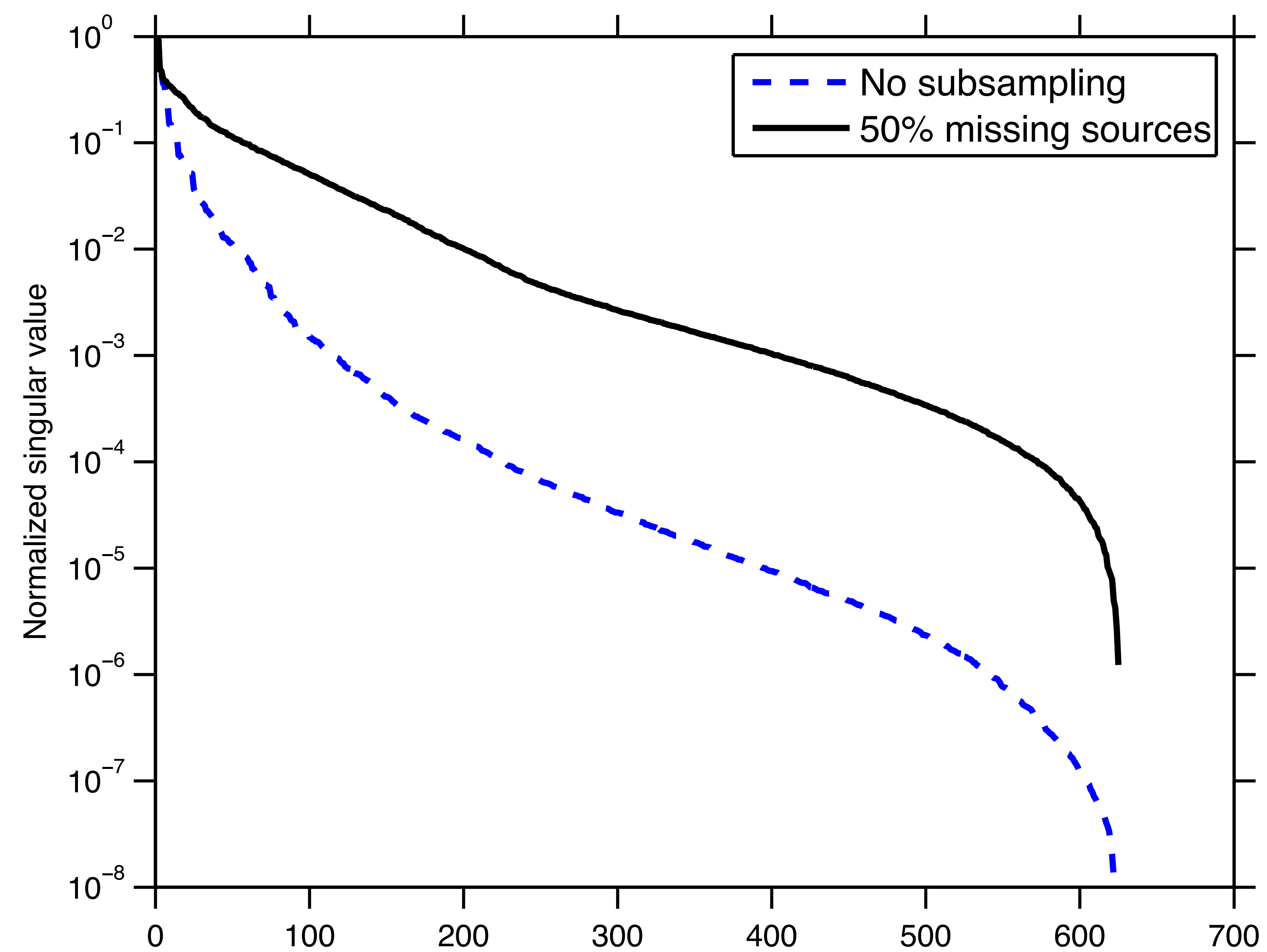


(Rec x, Rec y) matricization - Canonical ordering

Realistic acquisition
50% random receivers removed



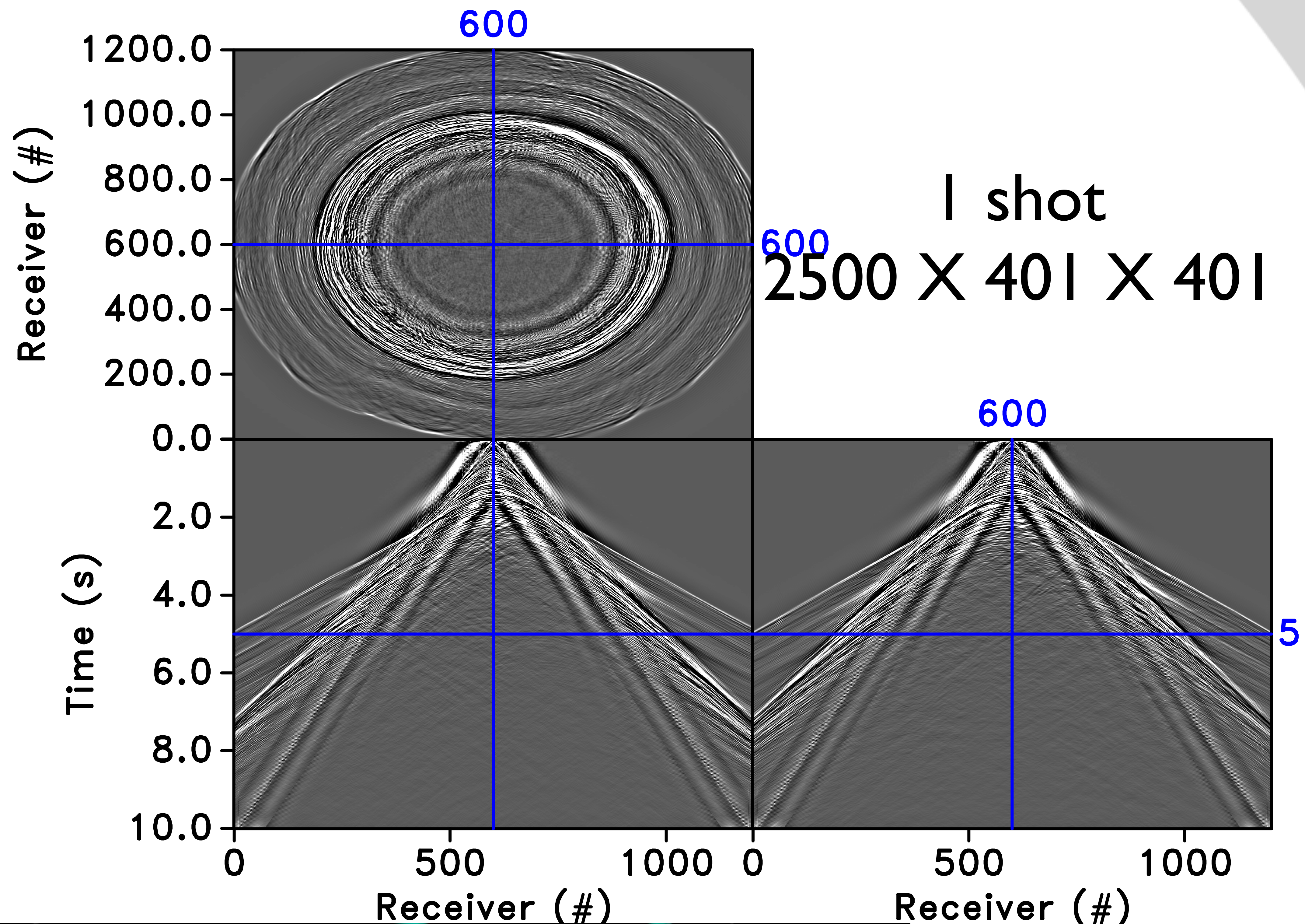


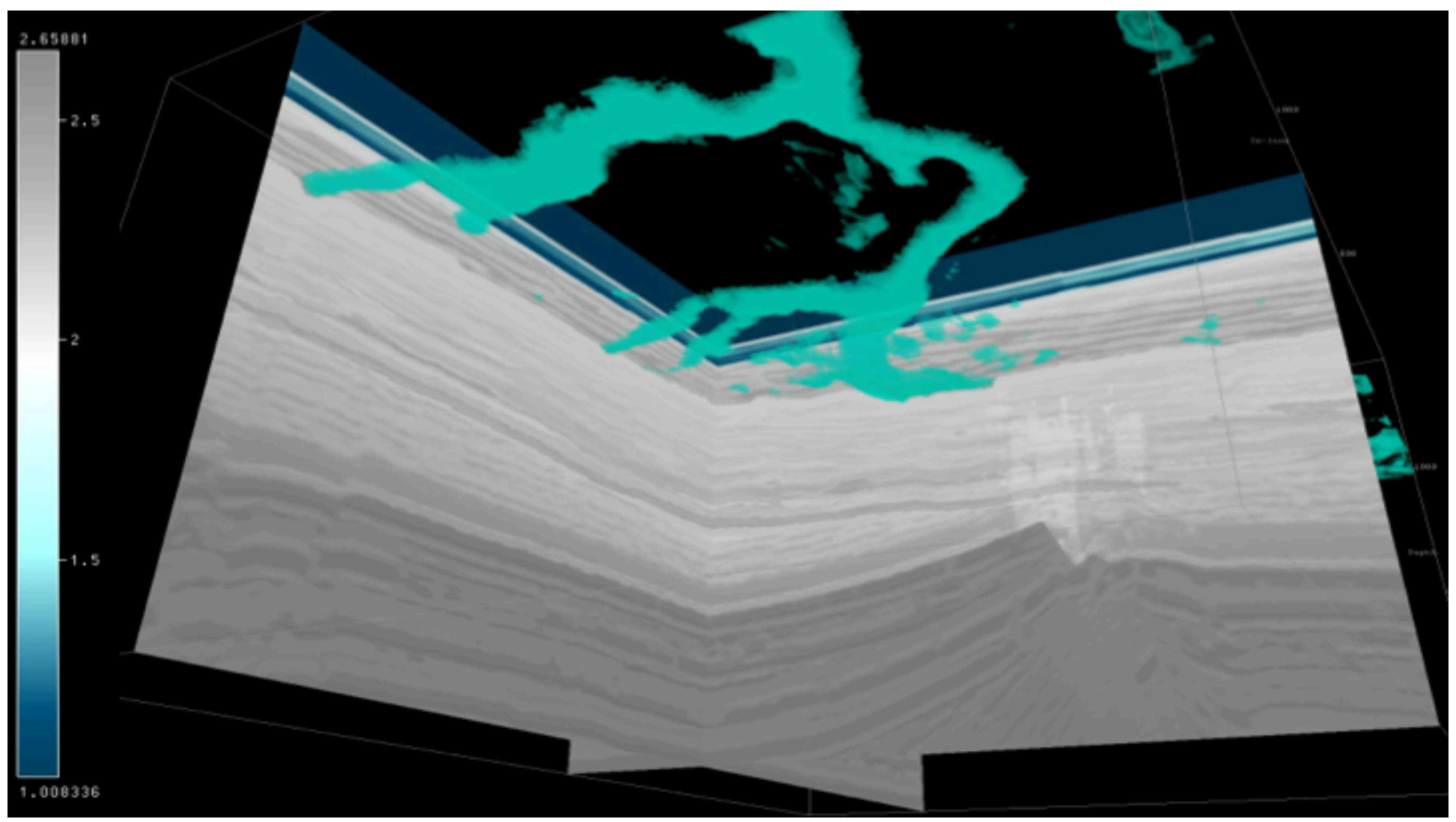


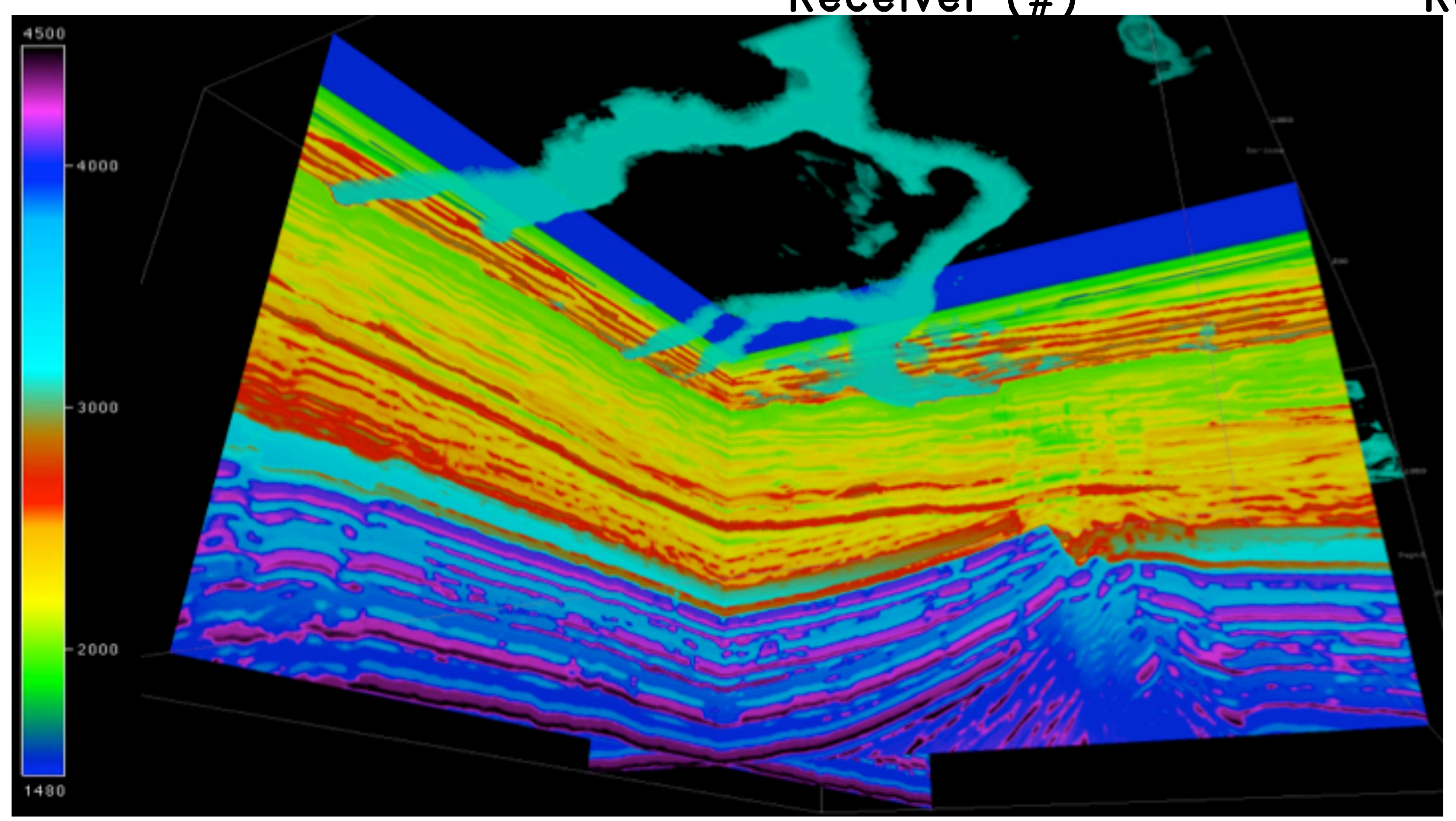
(Src x, Rec x) matricization - Noncanonical ordering

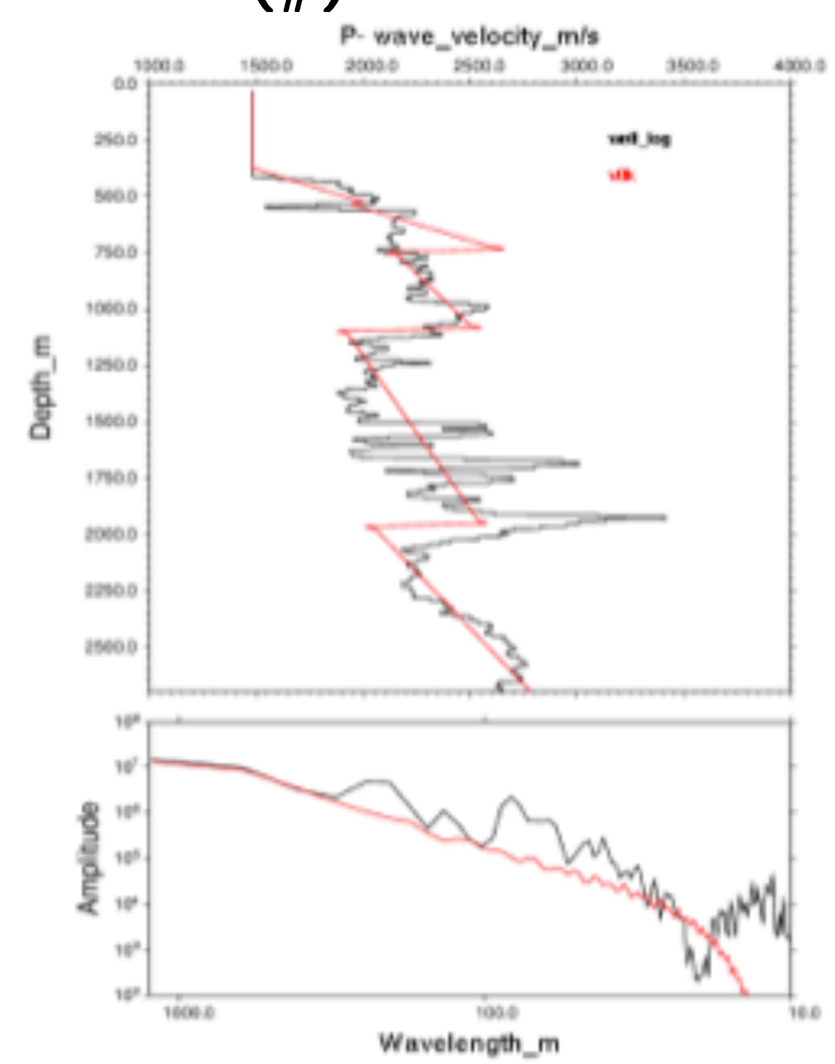
4-D Marine synthetic

BG Compass model



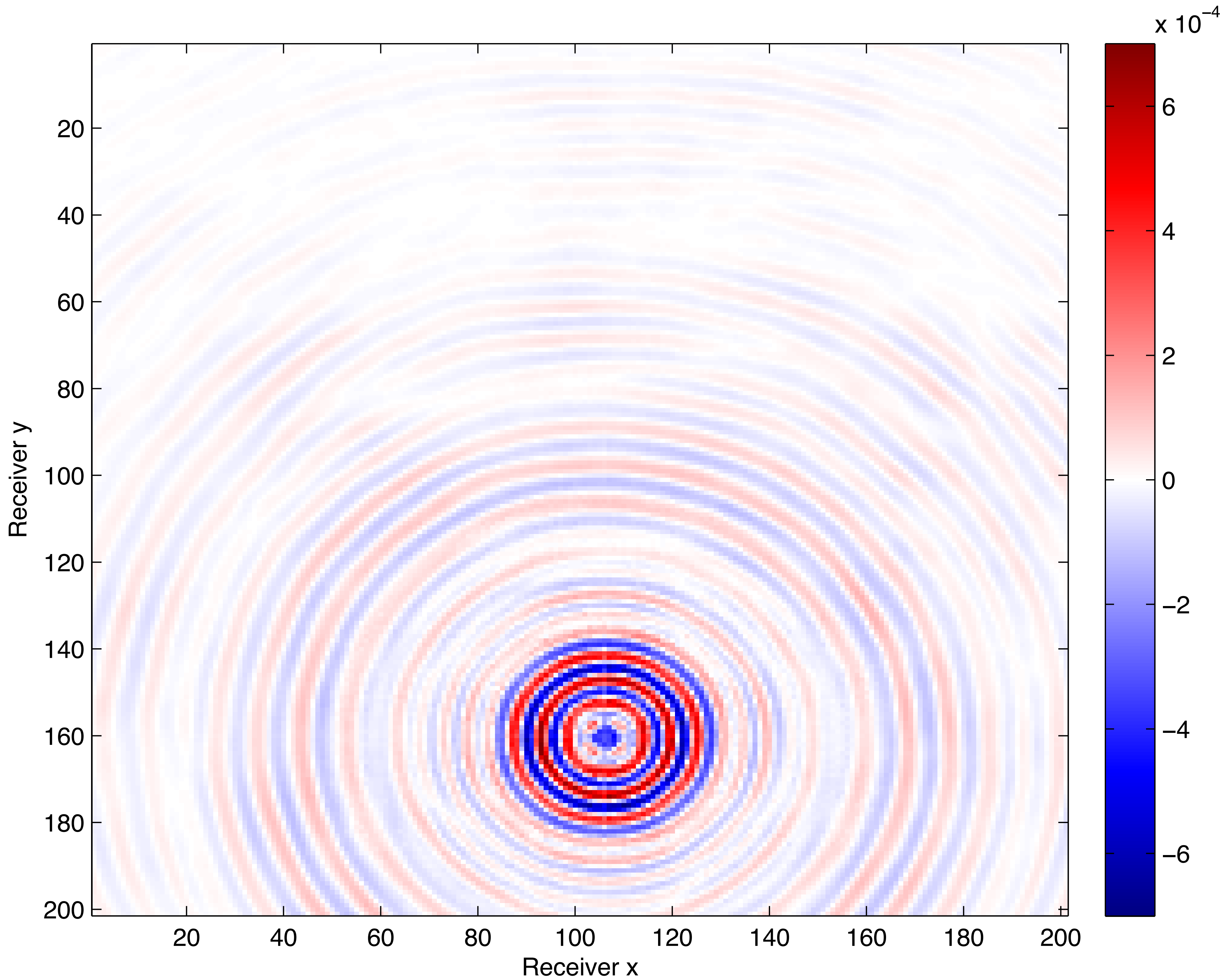




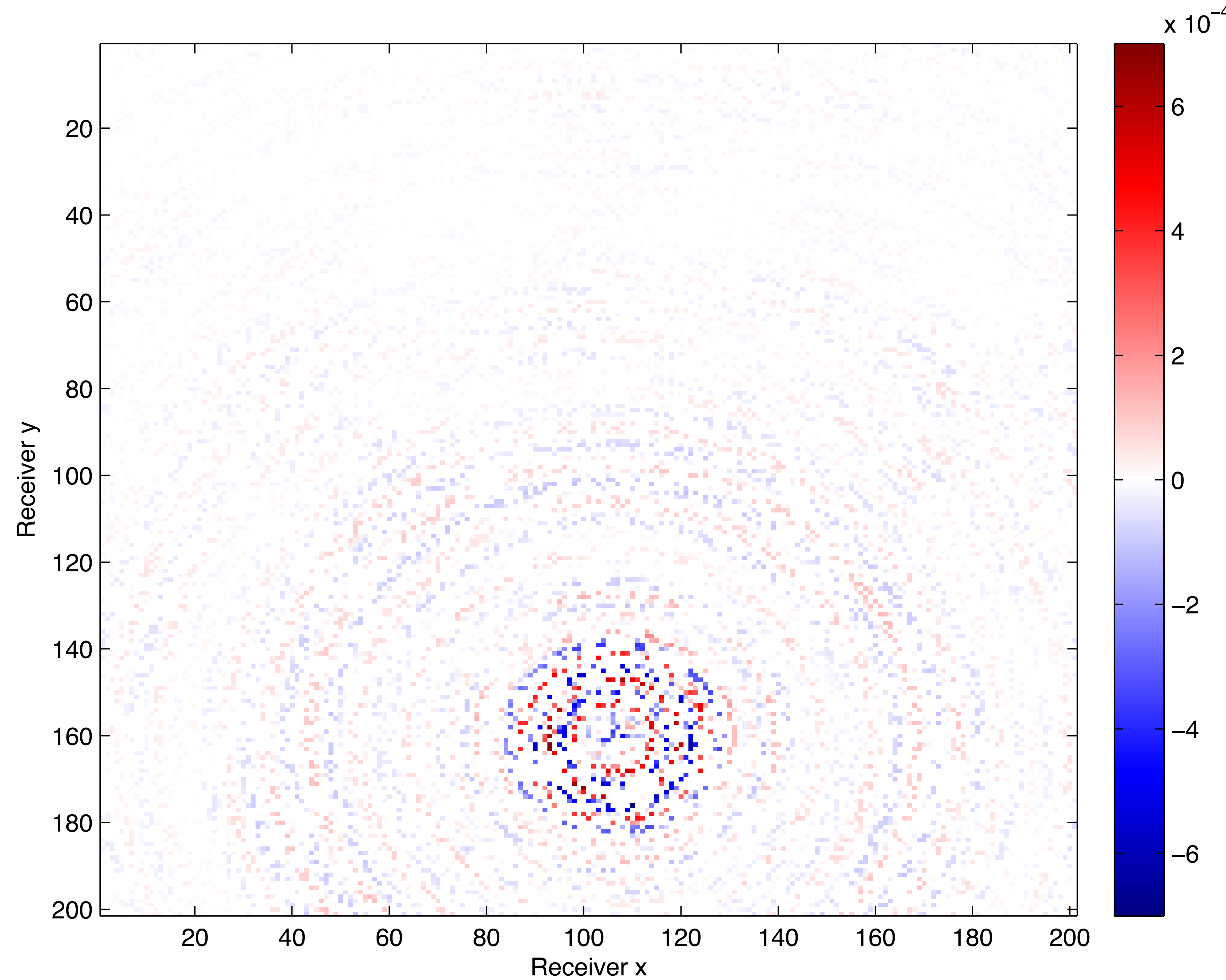


7.34 Hz - 75% missing receivers

Common source gather - noncanonical permutation



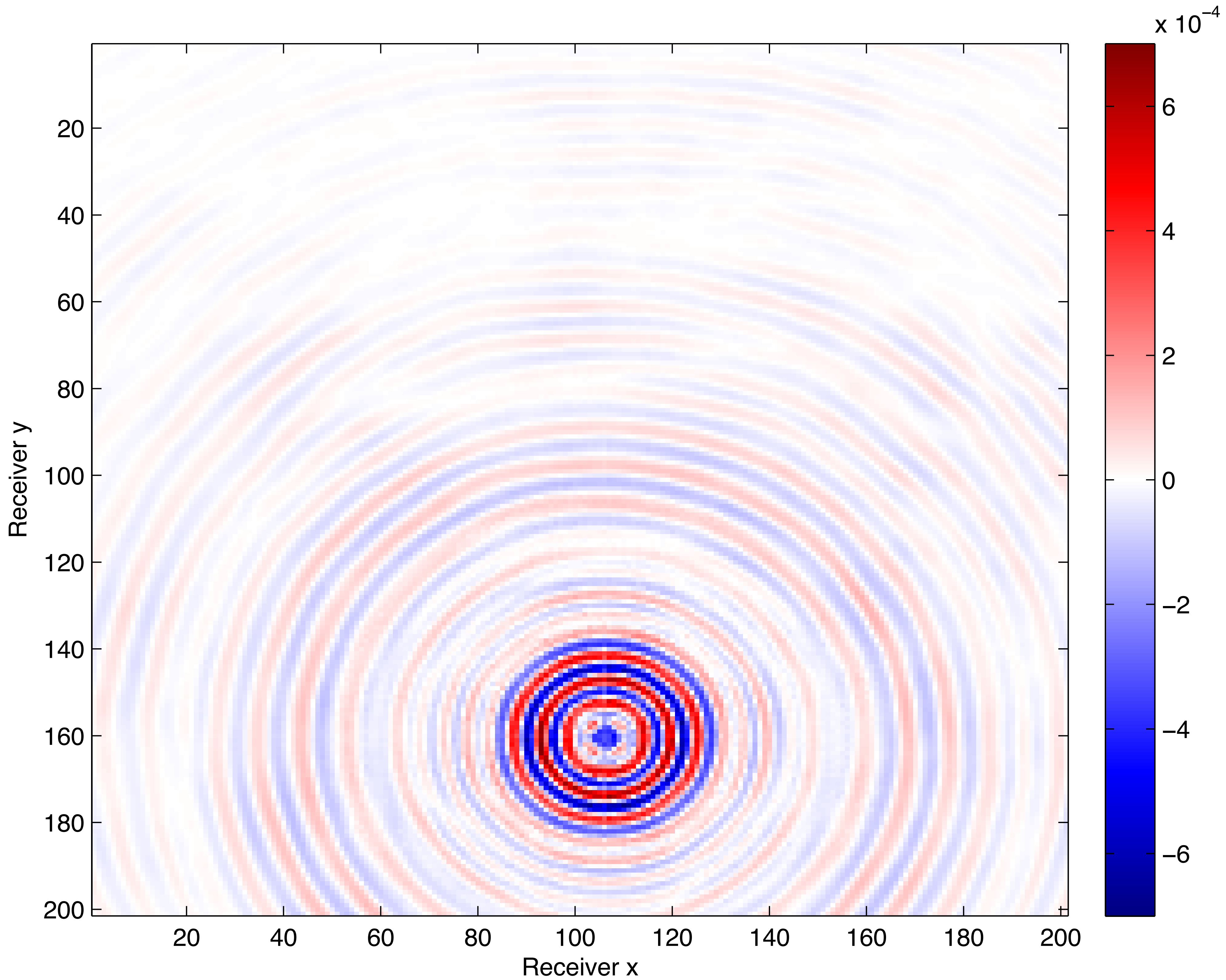
True data



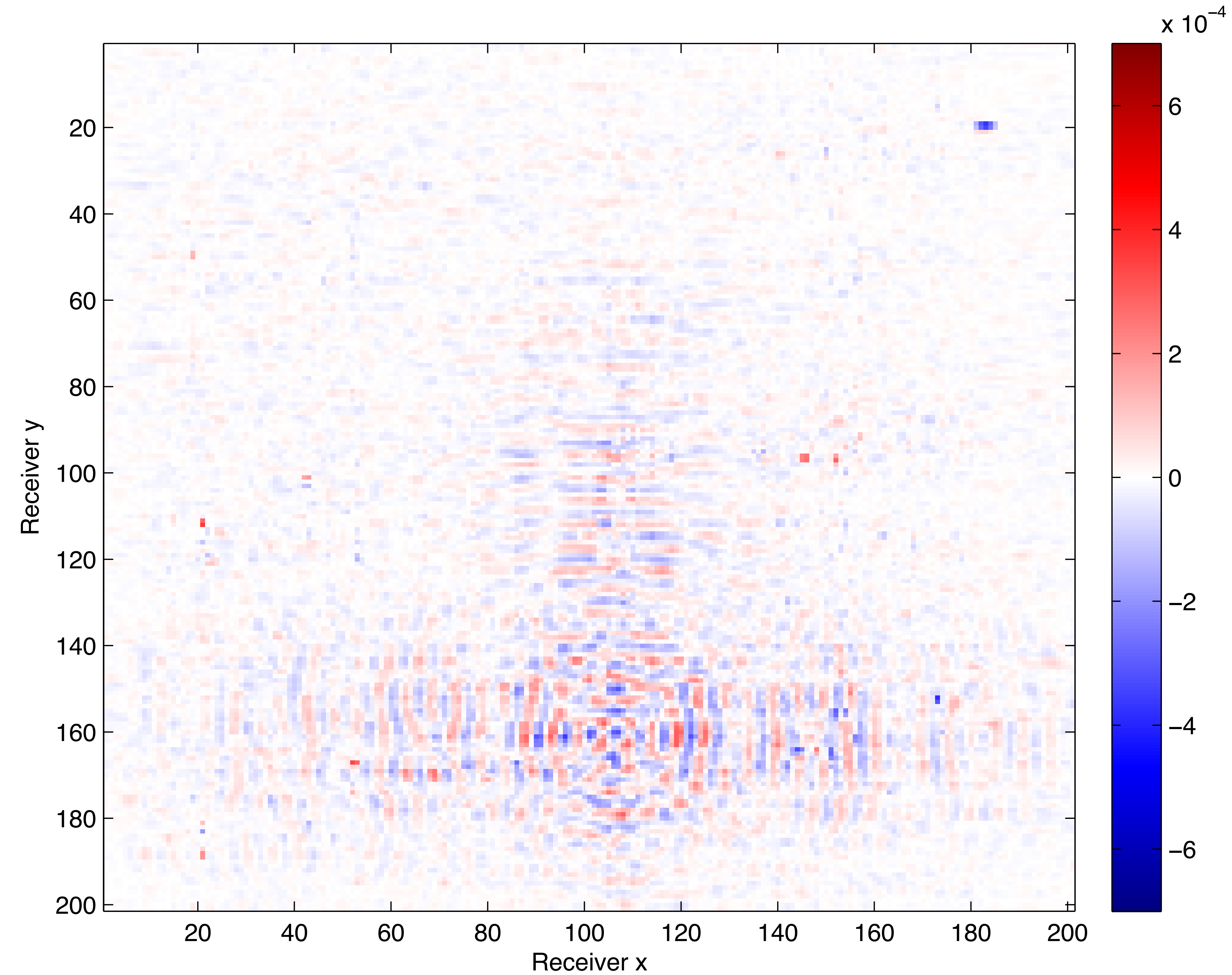
Subsampled data

7.34 Hz - 75% missing receivers

Common source gather - canonical permutation



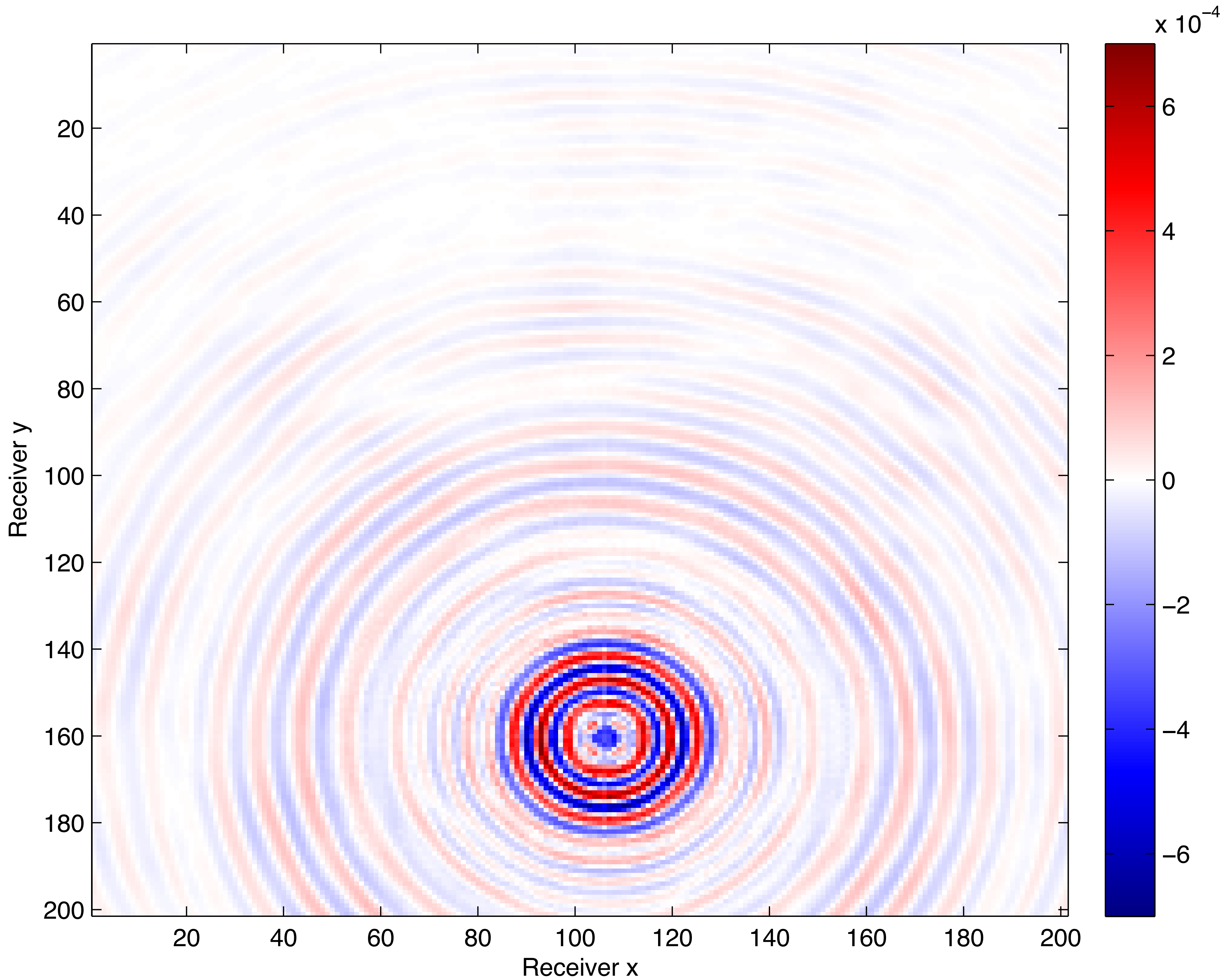
True data



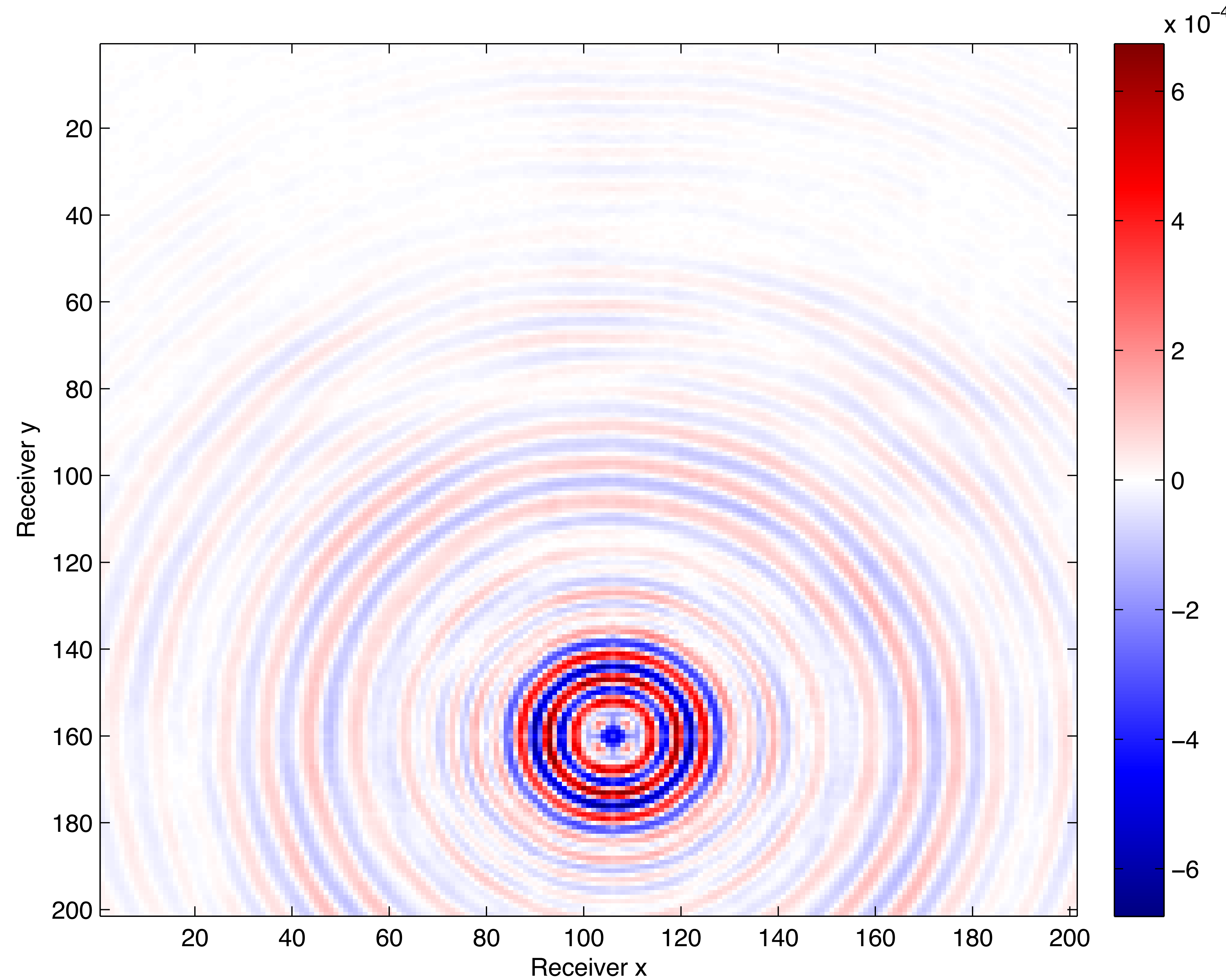
Recovered data - SNR 0.49 dB

7.34 Hz - 75% missing receivers

Common source gather - noncanonical permutation



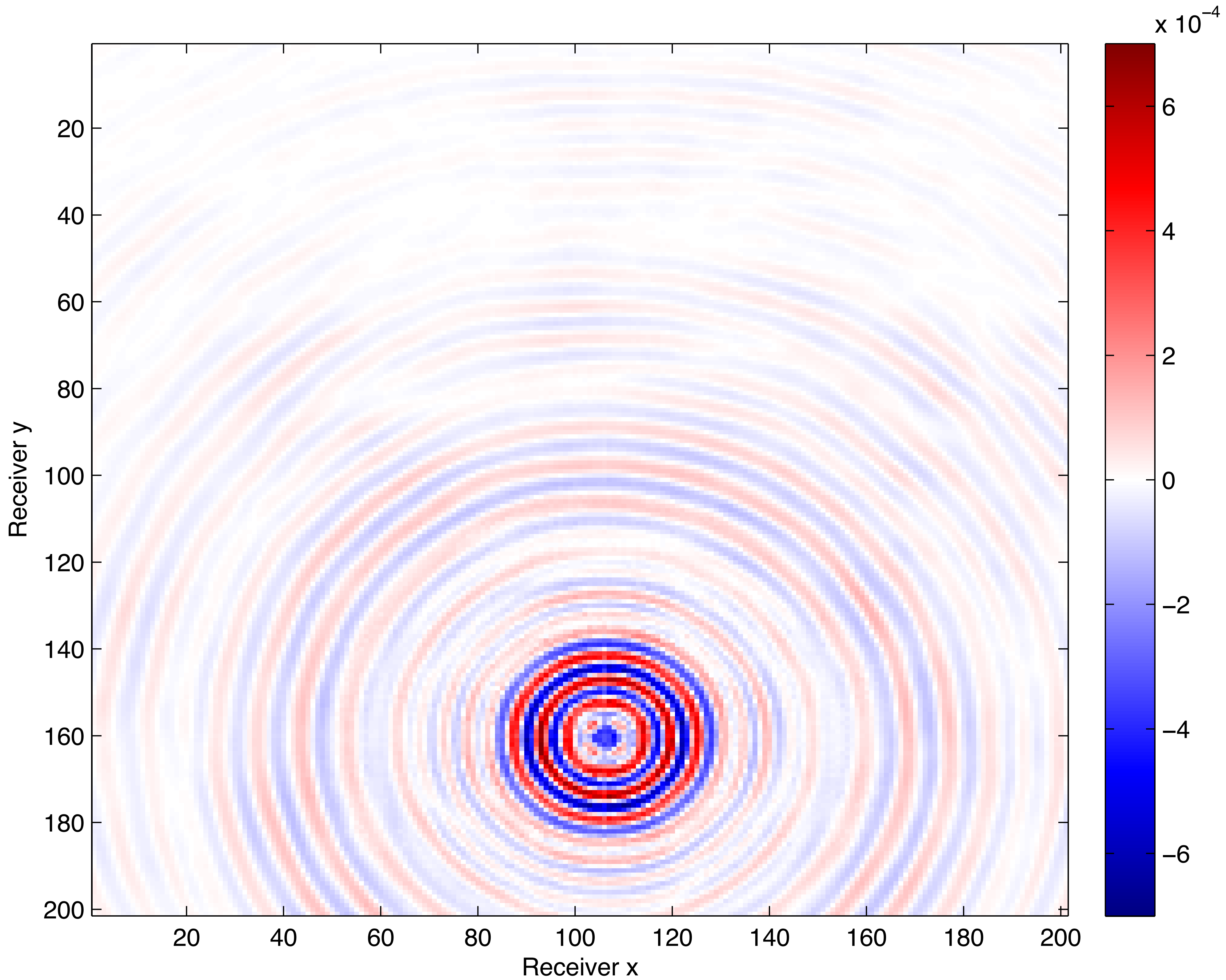
True data



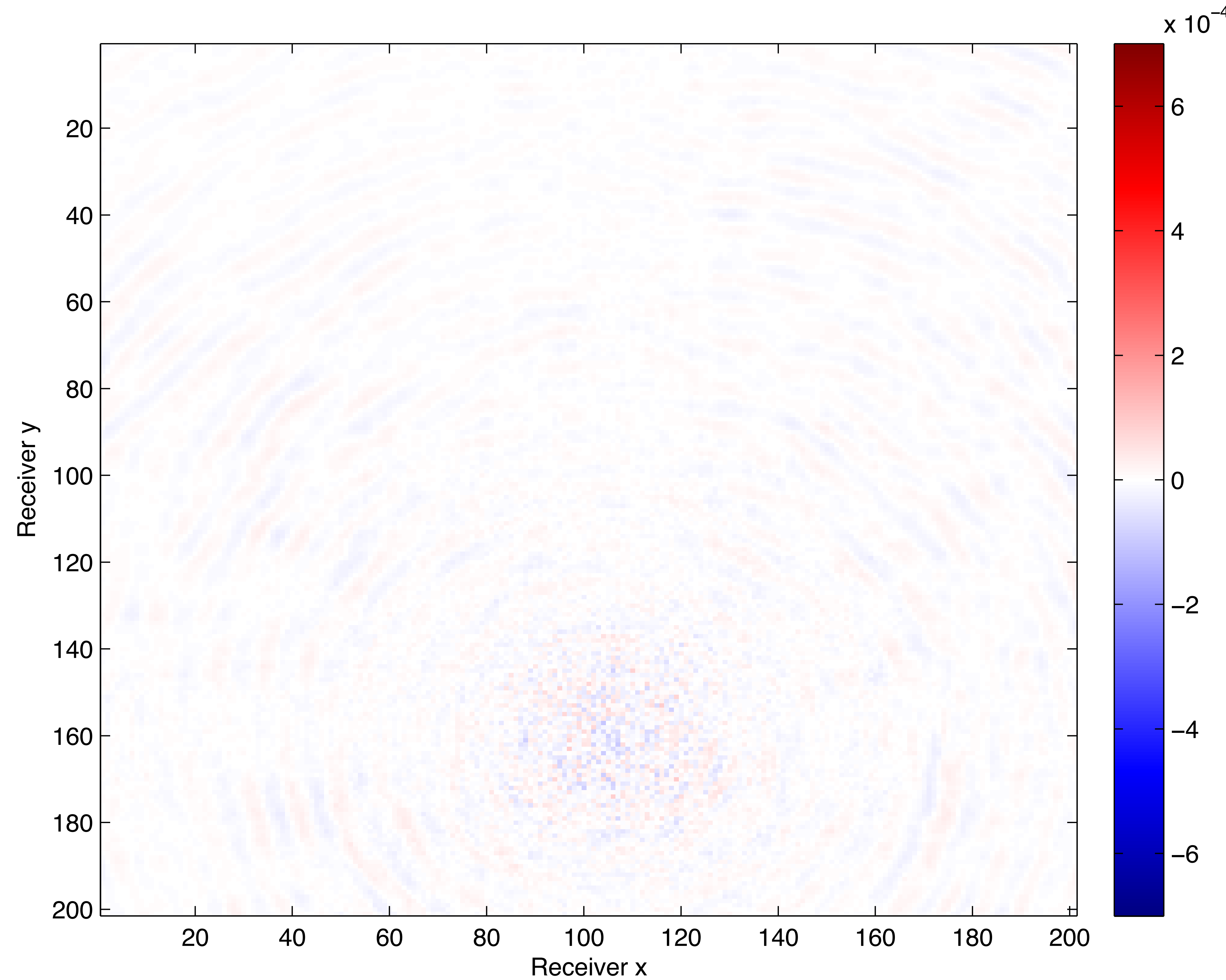
Recovered data - SNR 17.6 dB

7.34 Hz - 75% missing receivers

Common source gather - noncanonical permutation



True data



Difference

Curt Da Silva and Felix J. Herrmann, “**Optimization on the Hierarchical Tucker manifold - applications to tensor completion**”. 2013
Curt Da Silva and Felix J. Herrmann, “**Structured tensor missing-trace interpolation in the Hierarchical Tucker format**”, *SEG*, 2013
Rajiv Kumar, Hassan Mansour, Aleksandr Y. Aravkin, and Felix J. Herrmann, “**Reconstruction of seismic wavefields via low-rank matrix factorization in the hierarchical-separable matrix representation**”, *SEG*, 2013

SLIM

Observations

What *structure* can we *exploit* & how can we *break* it?

✓ *non*-canonical* matrization leads to *fast* decay of *singular* values & to *slow* decay after *randomized* sampling

What is the best “*domain*” to exploit *low*-rank *properties*?

✓ “SVD-free” hierarchical Tucker format but there are other candidates such as HSS matrices

What is the *most* appropriate recovery technique?

✓ manifold optimization for *tensors* or LR factorizations for *matrices*

*Same holds for midpoint/offset/azimuth.

Time-*jittered* marine acquisition

SLIM

University of British Columbia

Objective

***Shorten* marine acquisition *times* & *increase* source sample *density*.**

Questions:

Does increased *variability* of firing times improve recovery?

Is *regularization* (vs *binning*) necessary?

Are current industry practices in *jittered* towed-array surveys optimal?

Regular vs. *jittered* locations

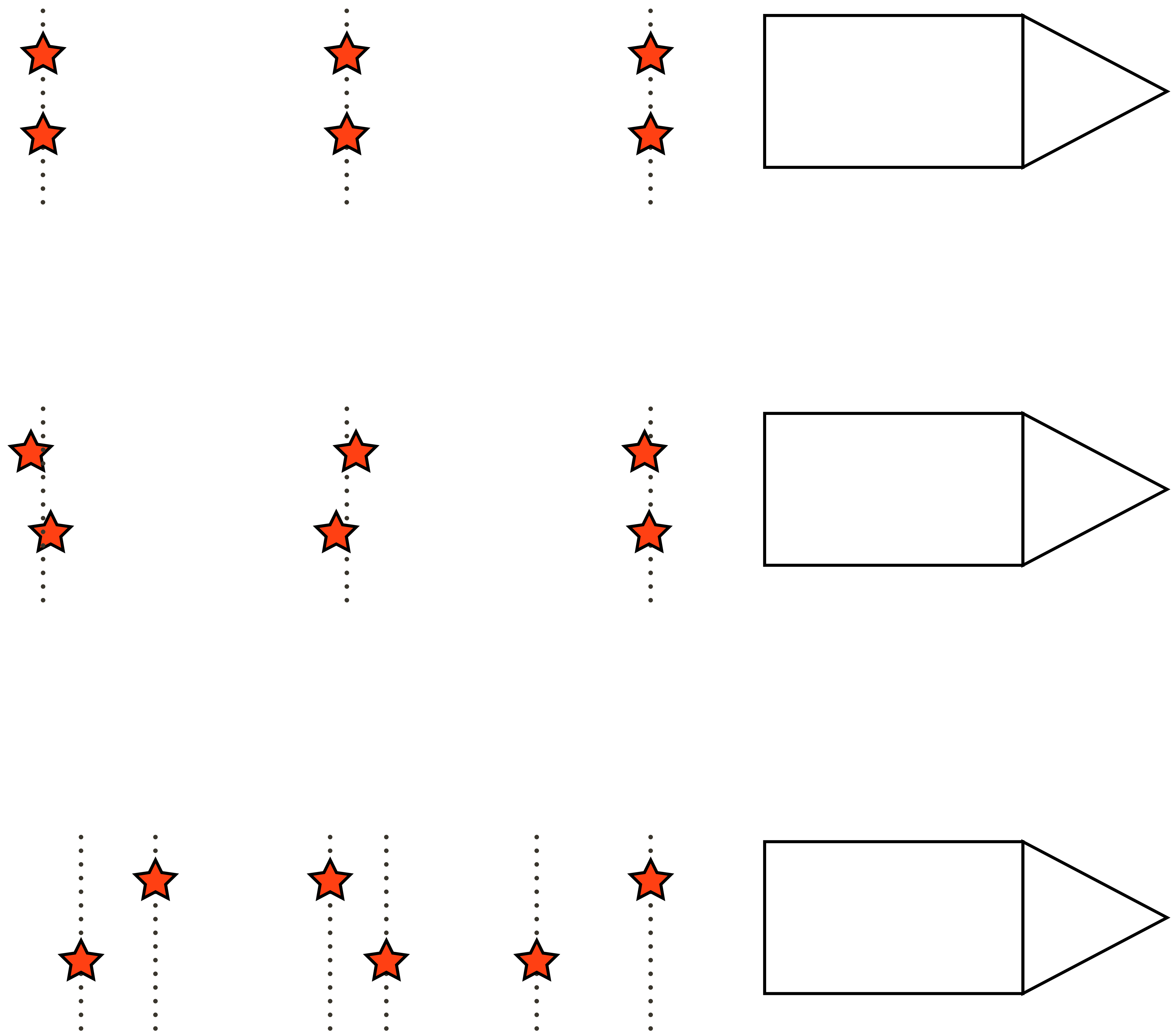
regularly sampled spatial grid

almost regularly sampled spatial grid

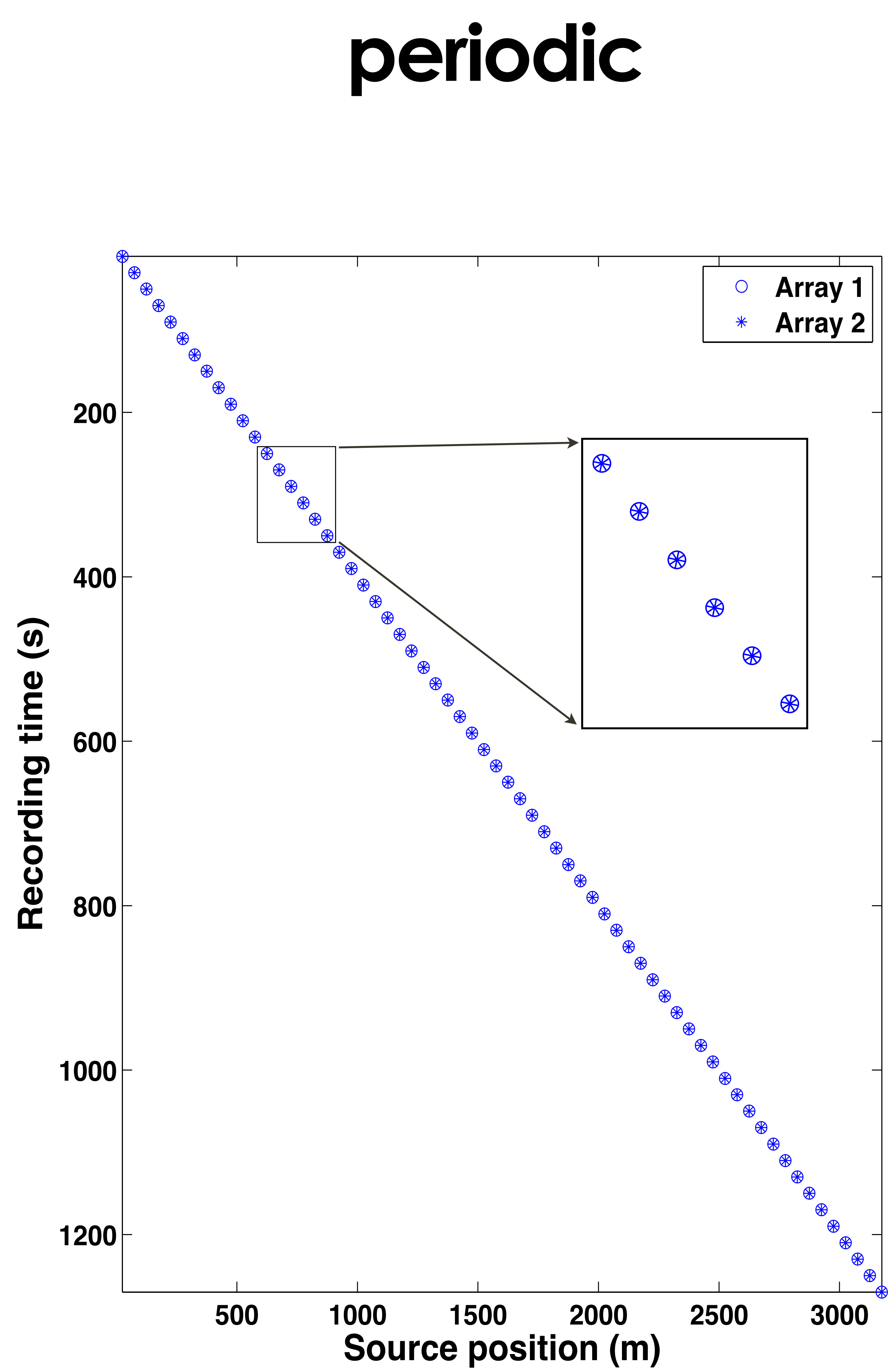
(low variability)

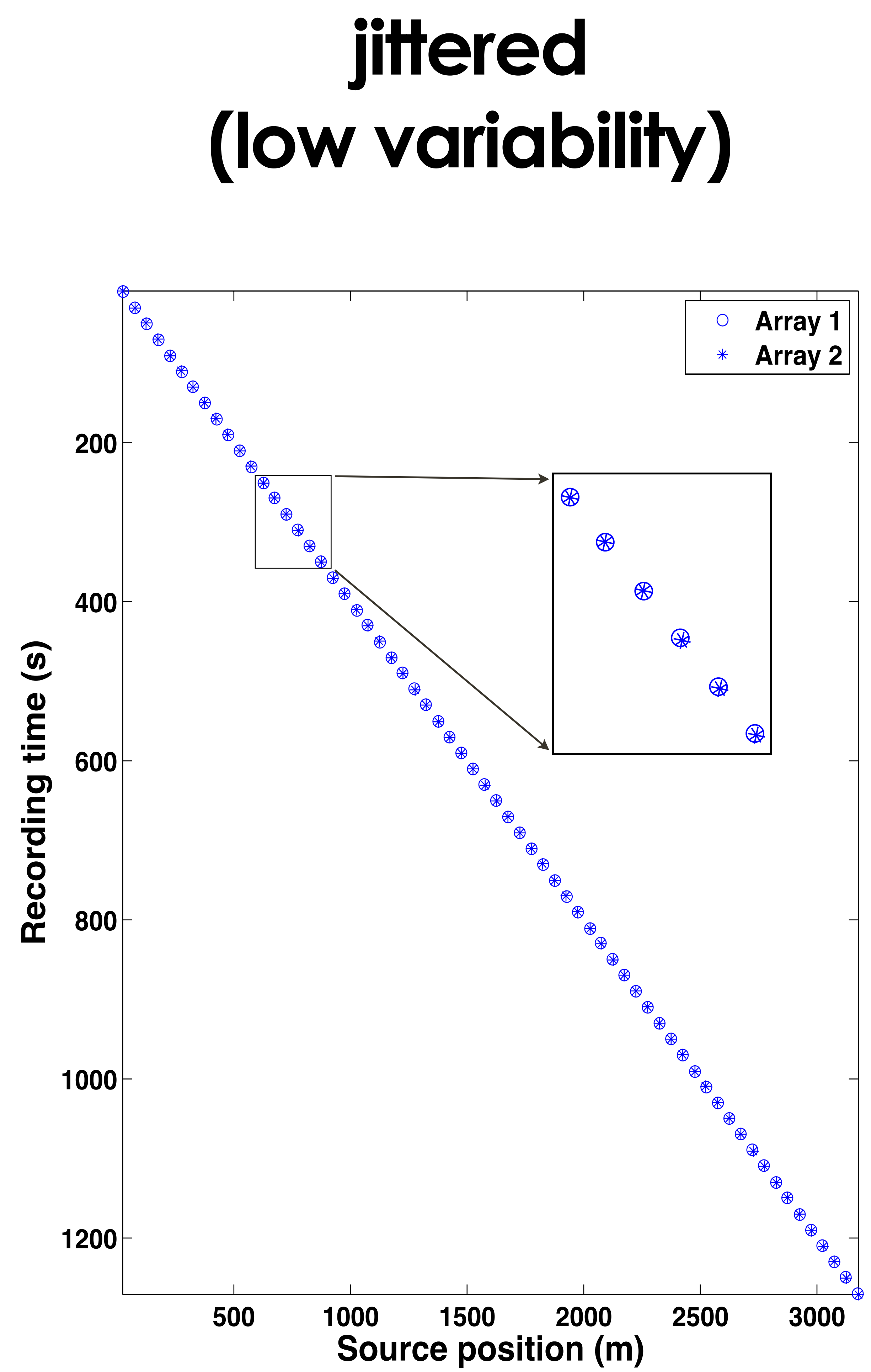
irregularly sampled spatial grid

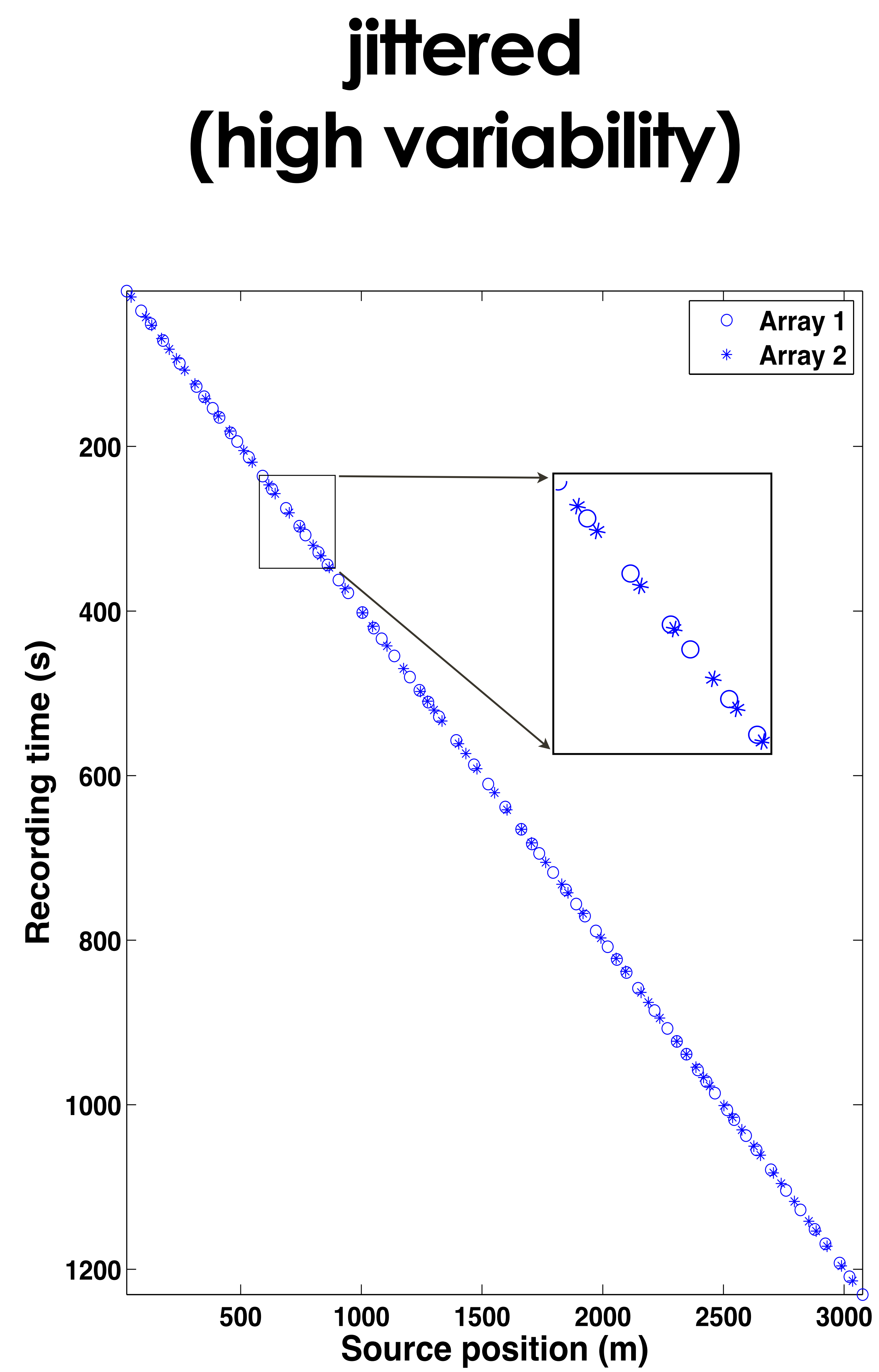
(high variability)



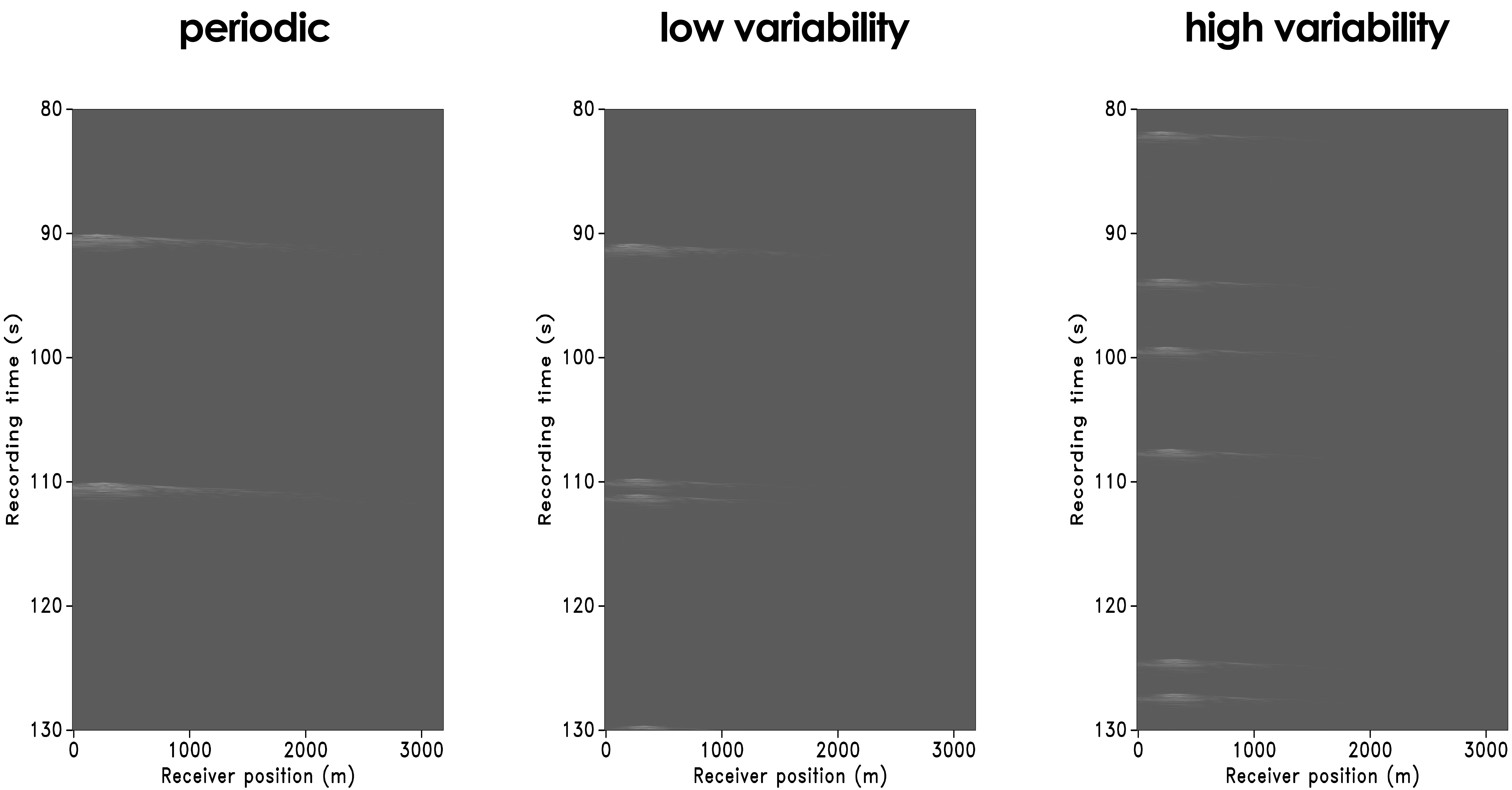
Regular vs. *jittered* times/locations





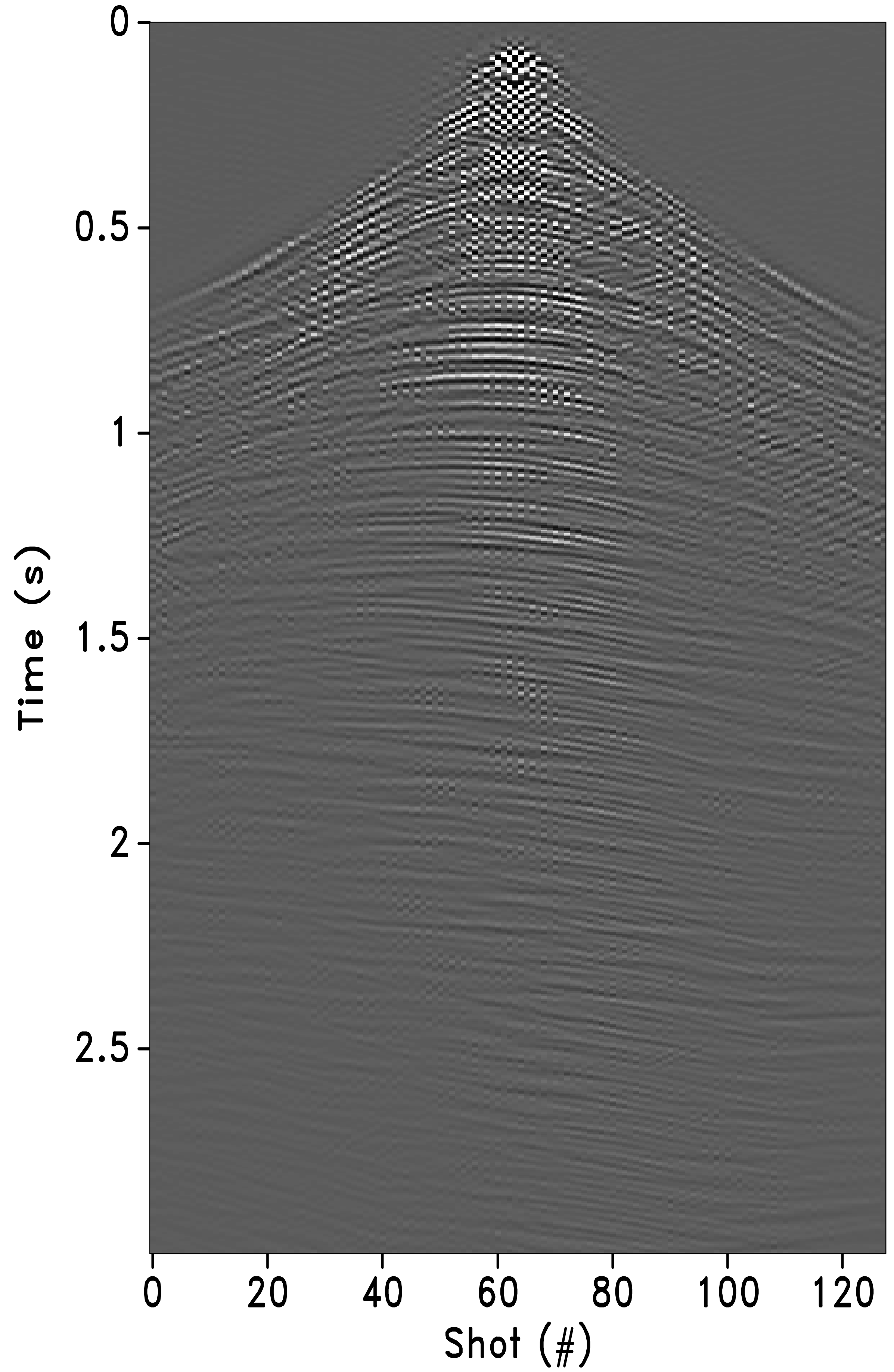


Measurements

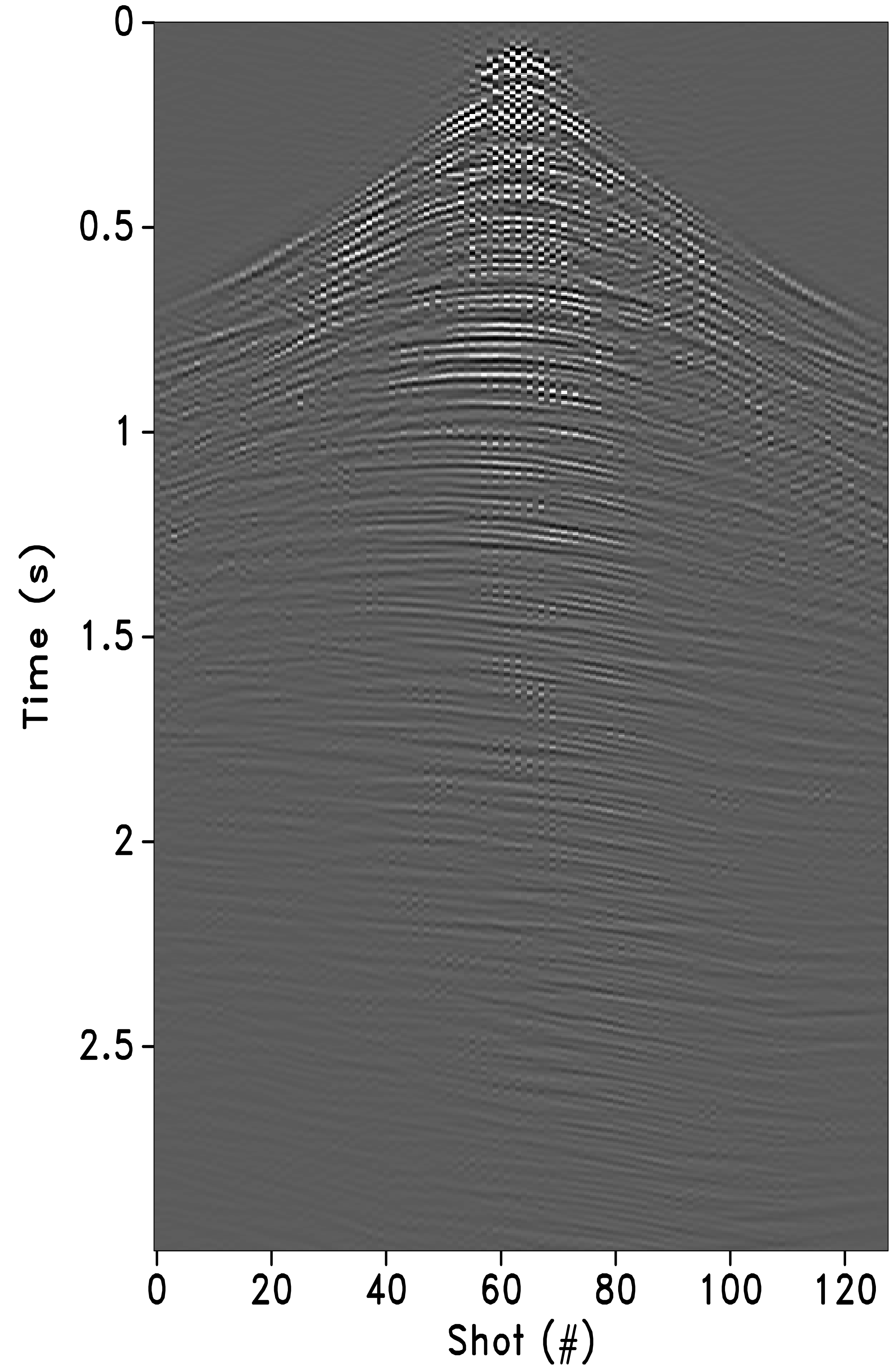


Recovery [“deblending” from 50m grid to 25m grid]

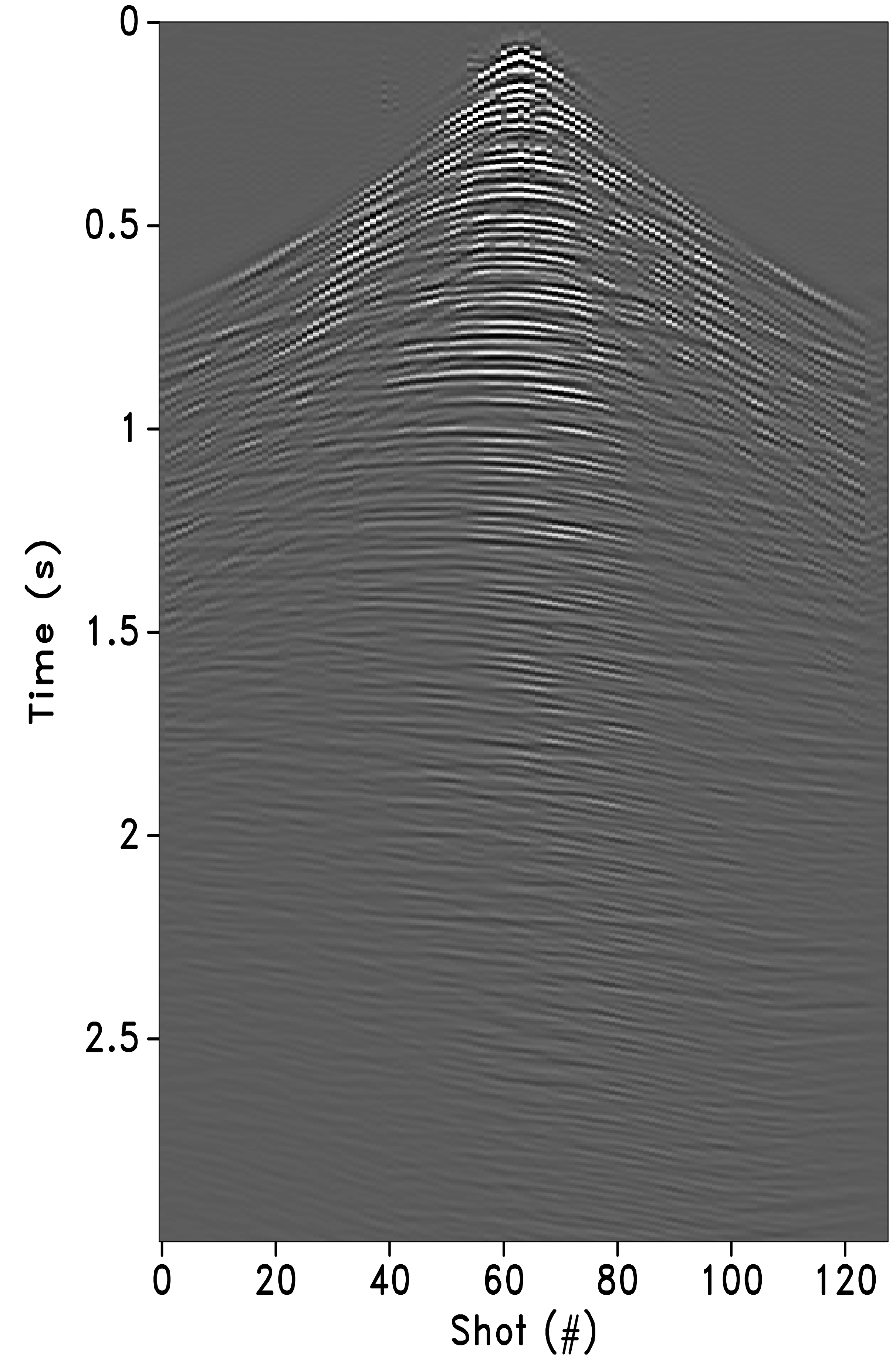
periodic
(1.6 dB)



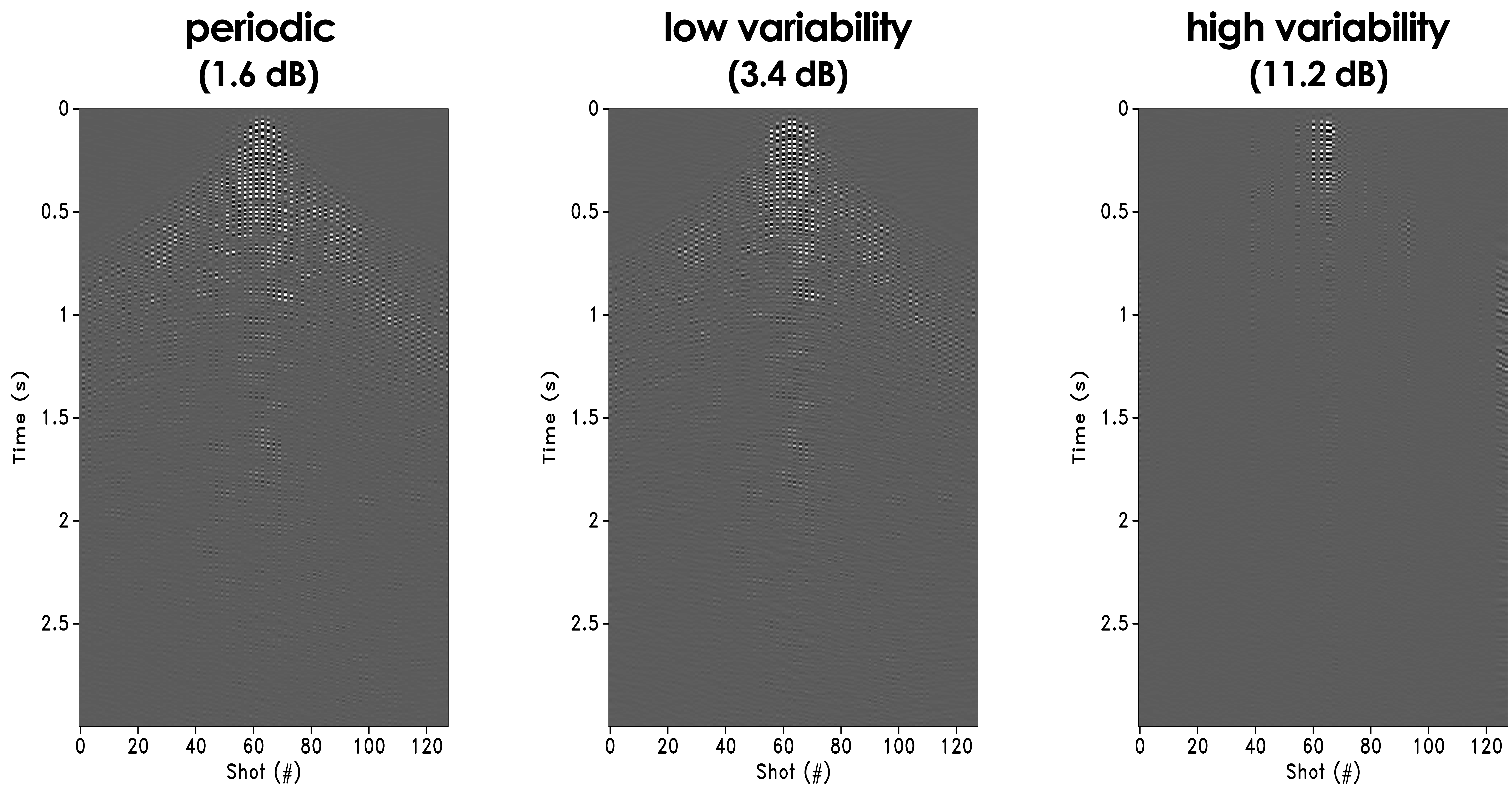
low variability
(3.4 dB)



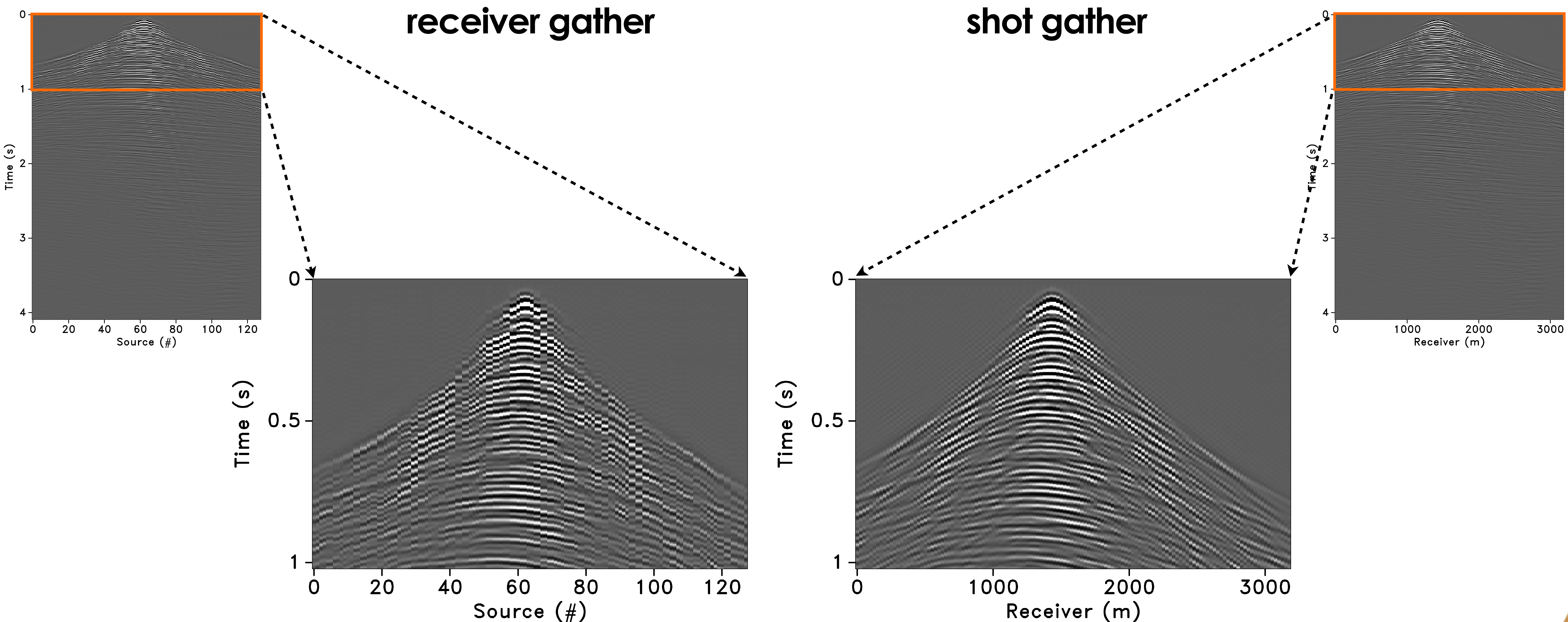
high variability
(11.2 dB)



Residual [“deblending” from 50m grid to 25m grid]

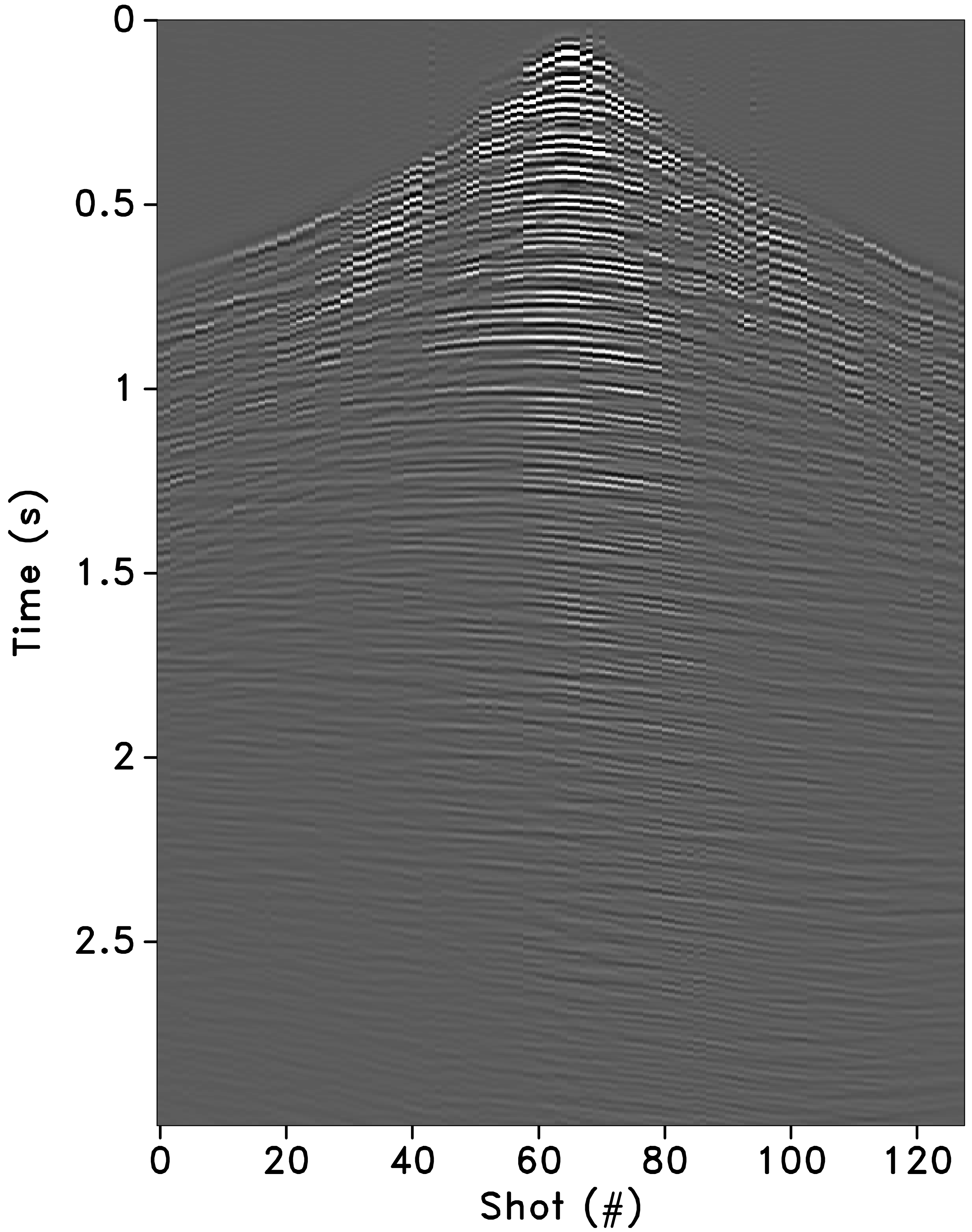


Shot-time *variability* ⇒ *irregular* traces

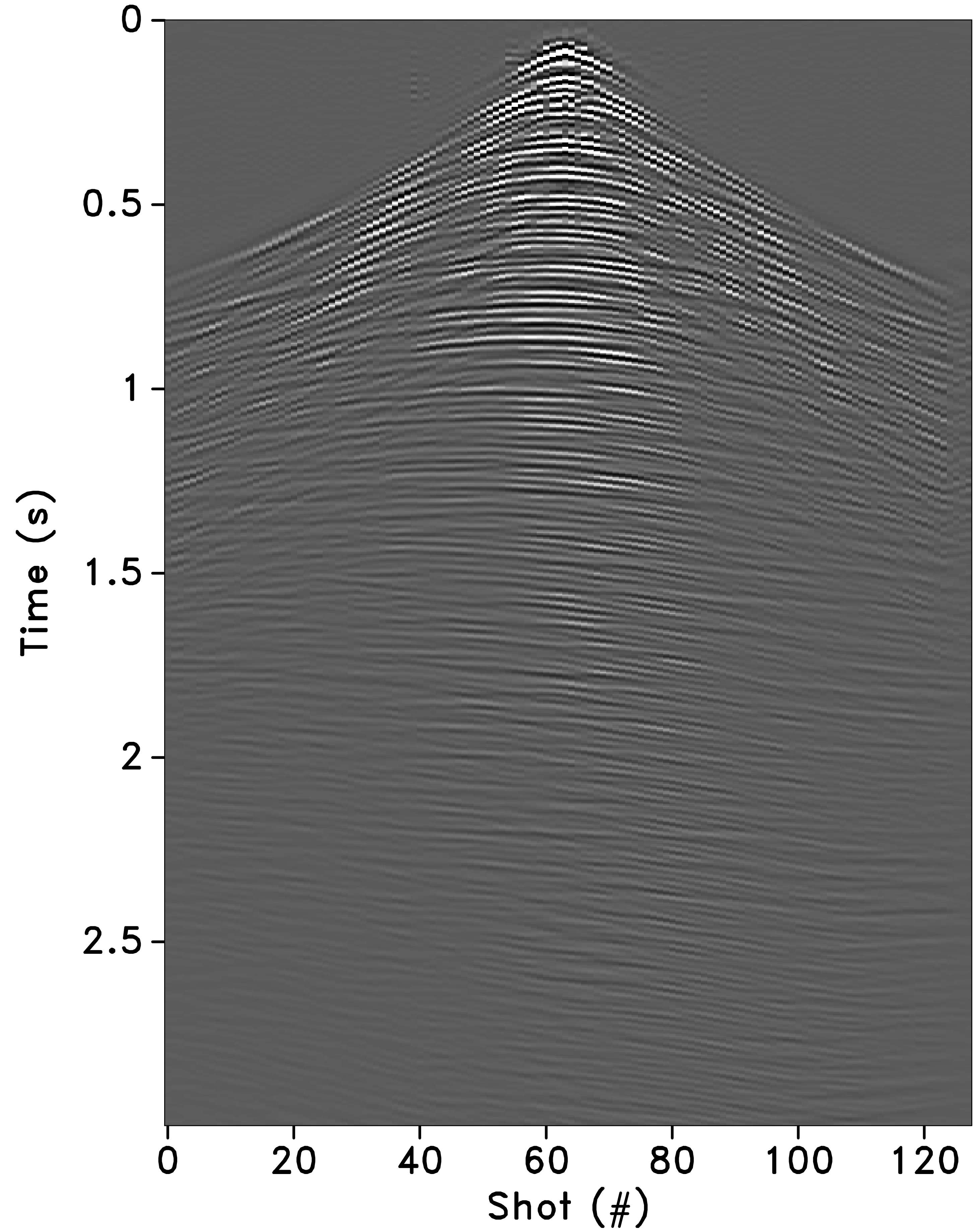


Recovery (do *not* bin!)

regular **curvelet transform (*'binning'*)**
(-1.47 dB)

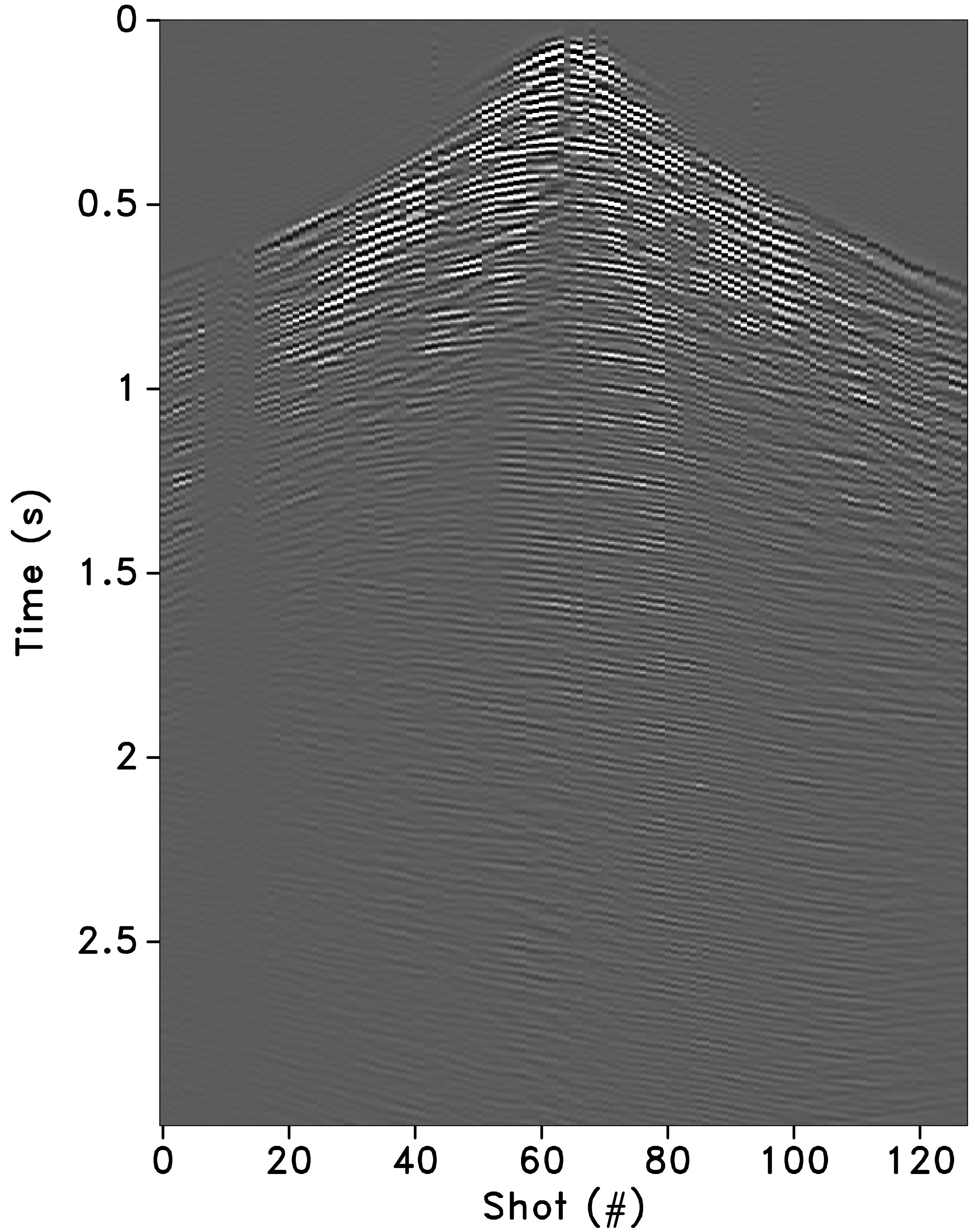


non-equispaced **curvelet transform**
(11.2 dB)

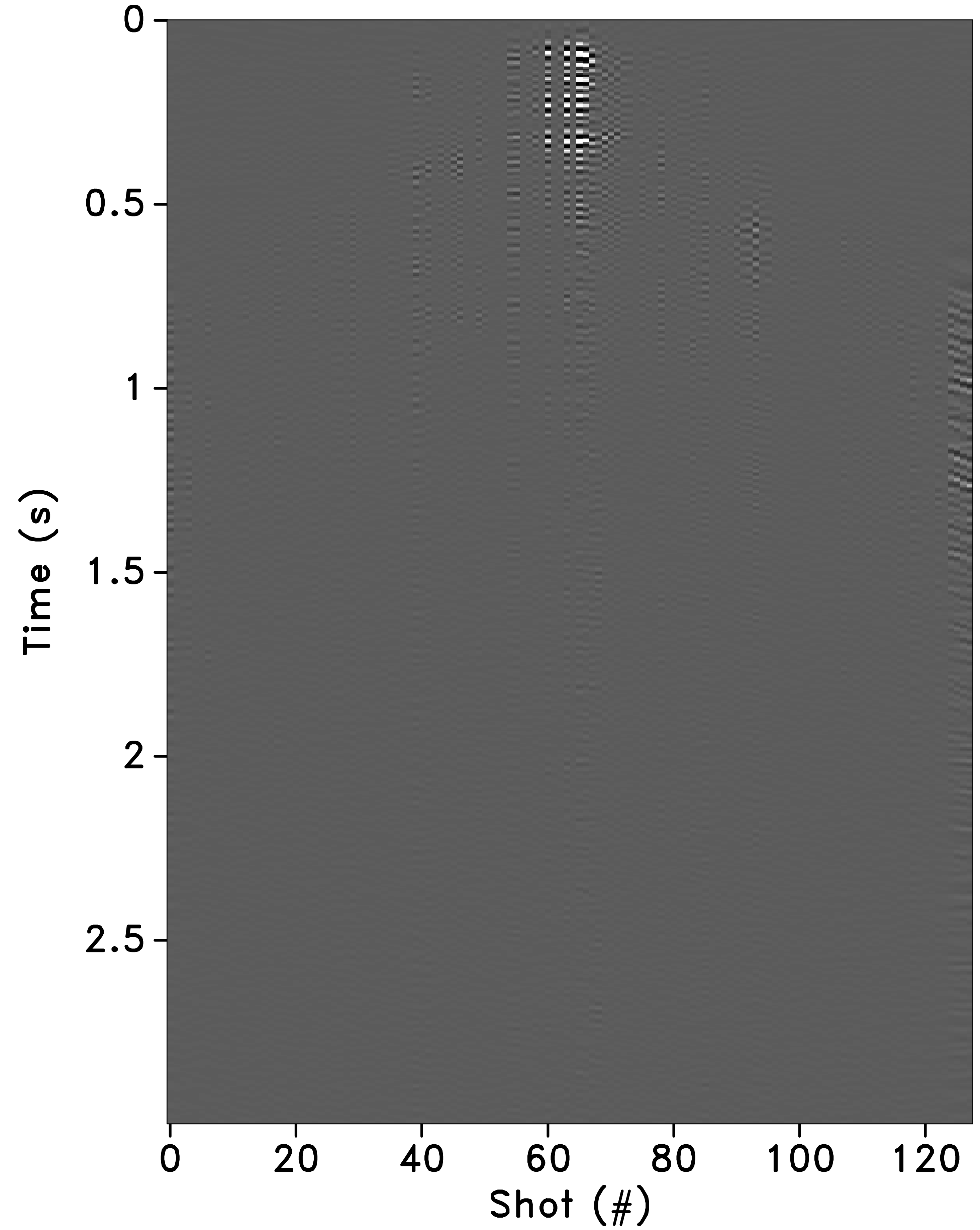


Residual (do *not* bin!)

regular **curvelet transform (*'binning'*)**
(-1.47 dB)



non-equispaced **curvelet transform**
(11.2 dB)



Haneet Wason and Felix J. Herrmann, "Time-jittered ocean bottom seismic acquisition", *SEG*, 2013
Hassan Mansour, Haneet Wason, Tim T.Y. Lin, and Felix J. Herrmann, "Randomized marine acquisition with compressive sampling matrices", *Geophysical Prospecting*, vol. 60, p. 648-662, 2012

Observations

Questions:

Does increased *variability* of firing times improve curvelet recovery?

- ✓ yes, but only for node acquisition since it is challenging for towed arrays

Is *regularization* (vs *binning*) necessary?

- ✓ yes, binning leads to significant errors especially for *high* variable case

Are current industry practices in *jittered* towed-array surveys optimal?

- no, but the challenge is that sources do not see same receivers

Time-*lapse* w/o repetition

SLIM
University of British Columbia

Objective

Make *time-lapse* acquisition more *feasible* & *affordable.*

Questions:

Can we avoid mandatory *repeatability* of *conventional* 4-D acquisition?

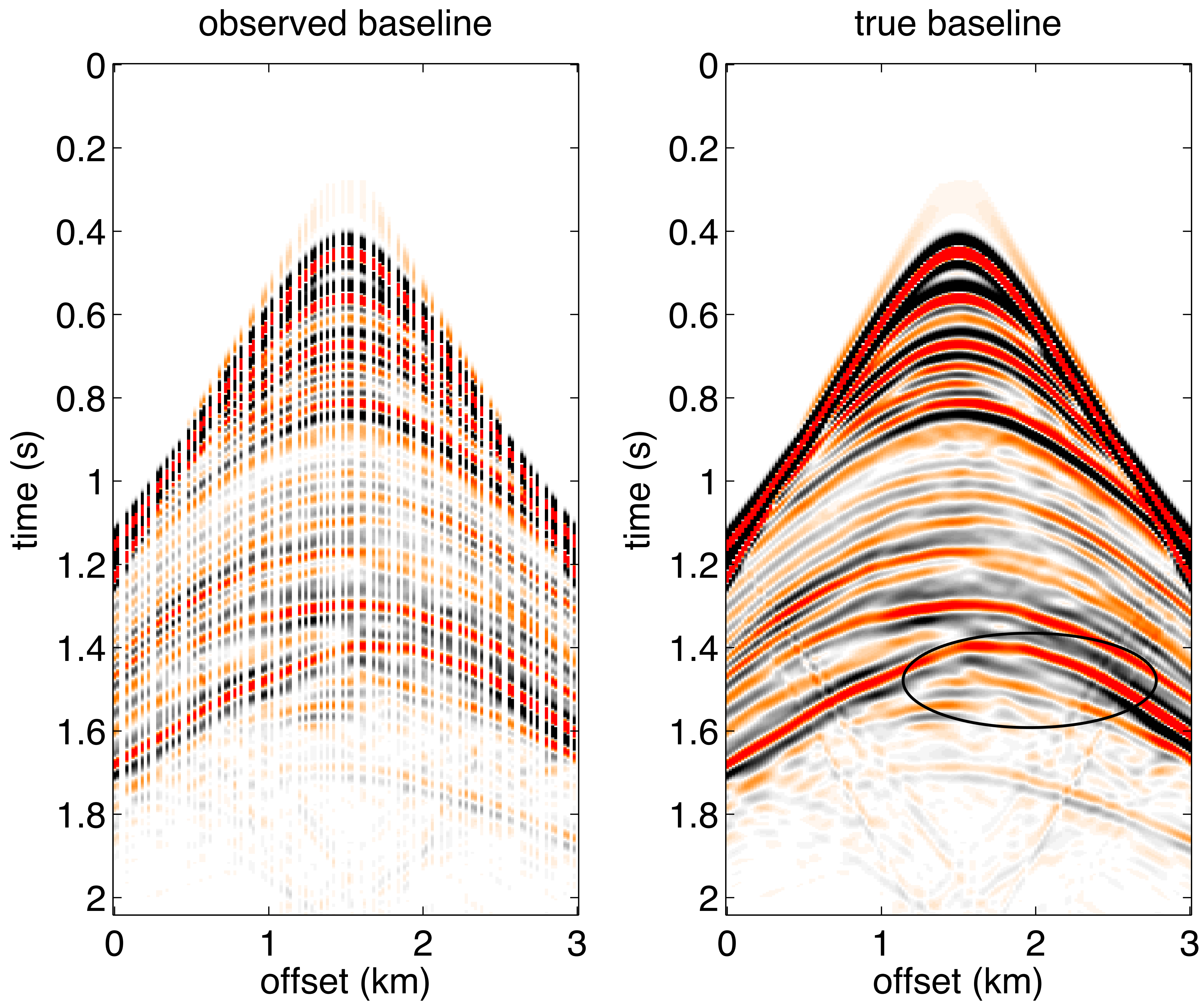
Can we *improve* recovery by exploiting *shared* structure?

Can we make time-lapse acquisition more affordable?

Numerical experiment

For Baseline model

- (**Proposed Setup**) randomized acquisition with missing shots



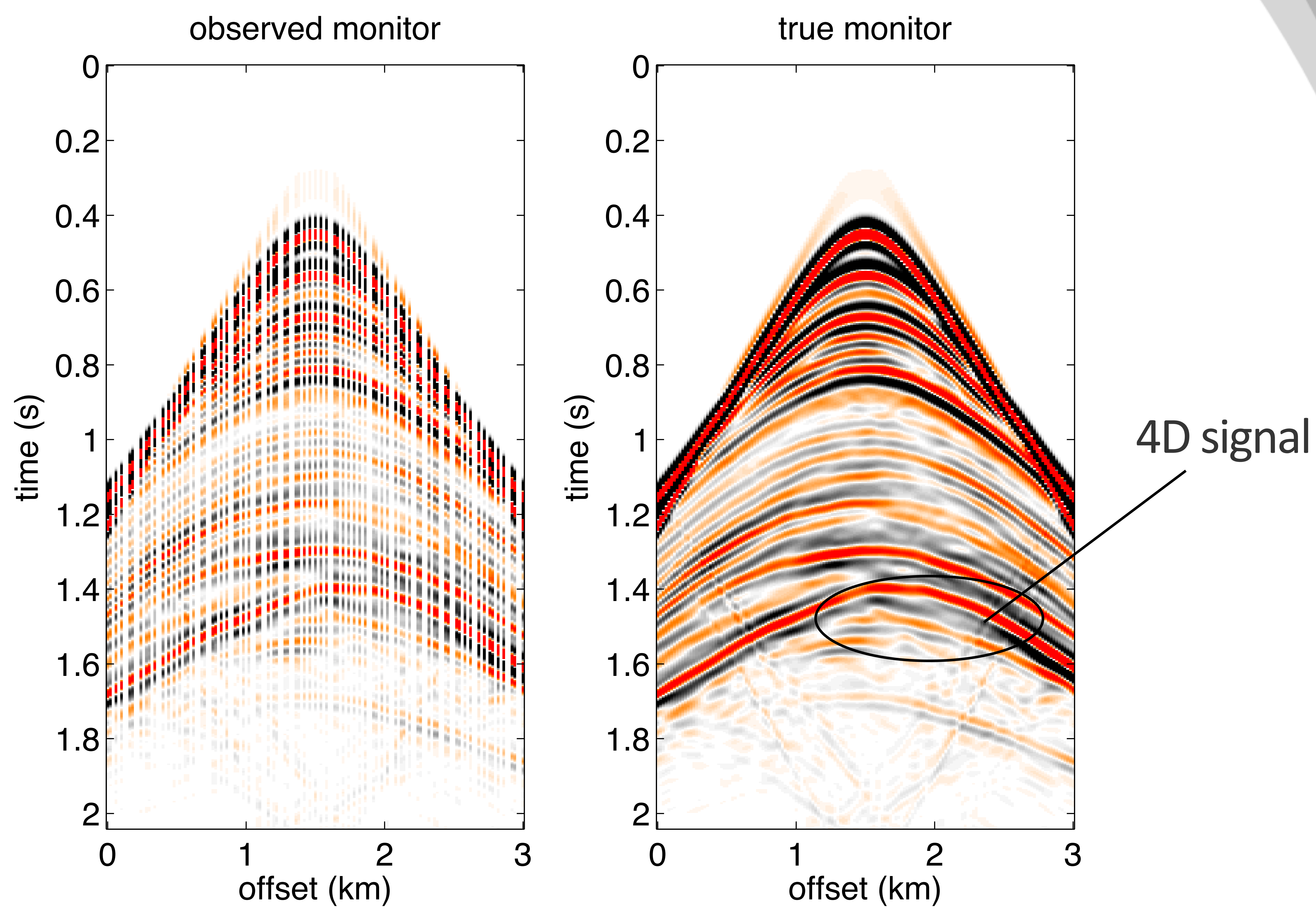
Numerical Experiment

For Baseline model

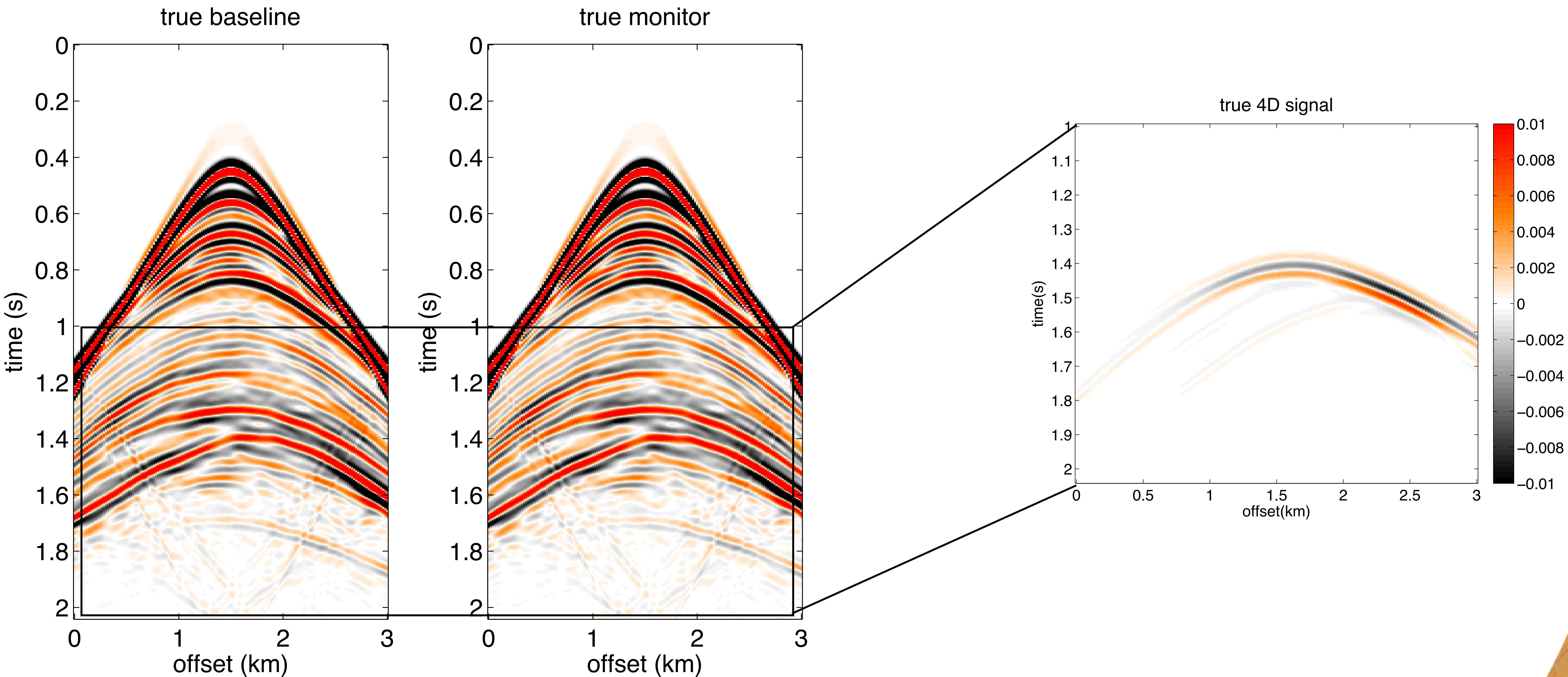
- (**Proposed Setup**) randomized acquisition with missing shots

For Monitor

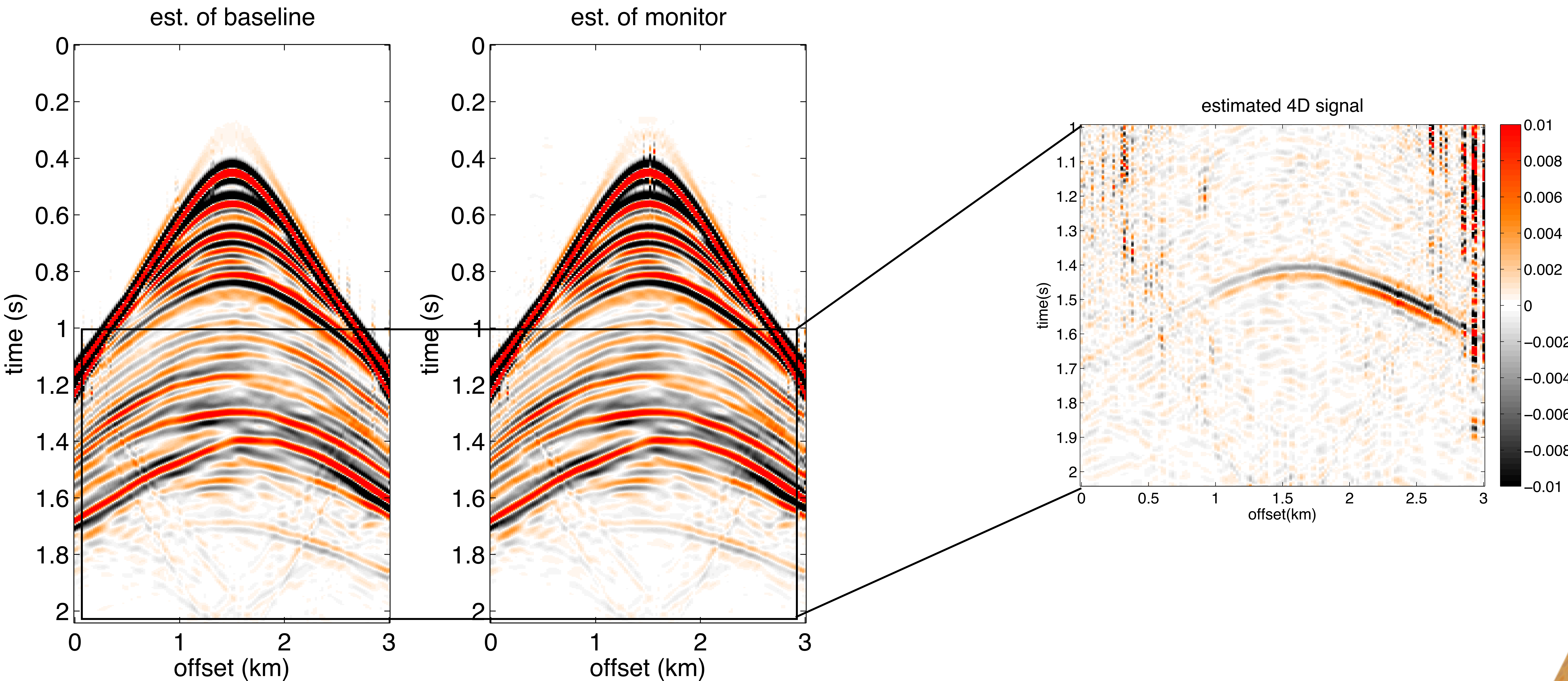
- use a ***different*** randomized acquisition geometry ?



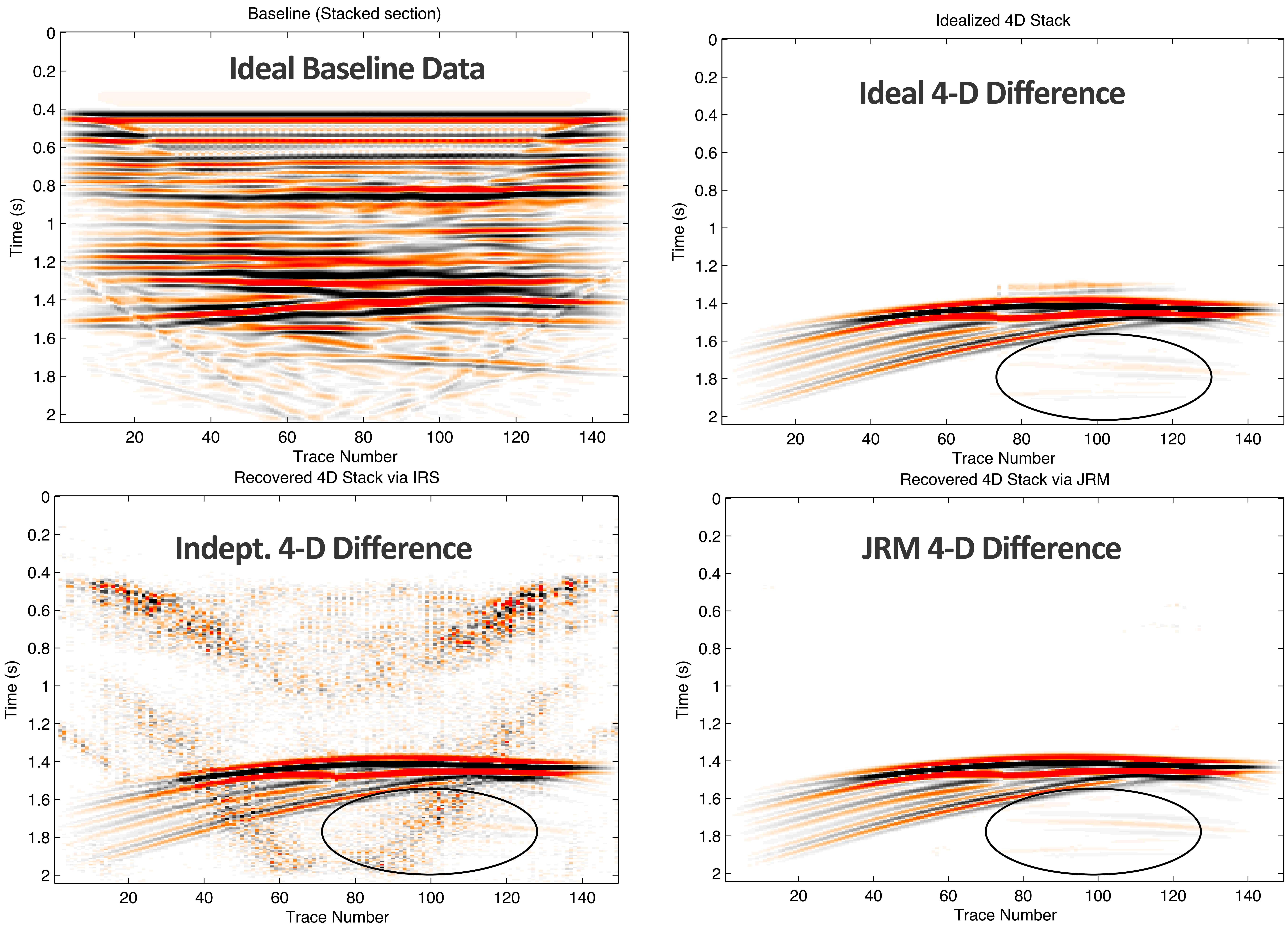
Perfectly repeated geometry (ideal "unrealistic" case)



Independent recovery for each vintage



Stacked sections: *independent* vs *joint* recovery



Felix Oghenekohwo, Ernie Esser, and Felix J. Herrmann, "Time-lapse seismic without repetition: reaping the benefits from randomized sampling and joint recovery". 2014

SLIM

Objective

Questions:

Can we avoid mandatory *repeatability* of *conventional* 4-D acquisition?

- ✓ yes, *via* sparse recovery from *independent* randomized acquisitions

Can we *improve* recovery by exploiting *shared* structure?

- ✓ yes, *joint* recovery explores information *shared* amongst *different* vintages

Can we make time-lapse acquisition more affordable?

- ✓ yes, since the formulations permits *randomized* subsampling

Randomized inversion–a new *paradigm*

Approximate* data *misfits* & model *updates

- exploit *separable* structure of *map-reduce* problems (machine learning/big data)
- work with *randomized* subsets of data only to limit # *passes* through *data*
- approximate simulations (e.g. PDE solves)

Control* the *errors

- by *increasing* the *randomized* subsets
- by *increasing* the *accuracy* of *simulations*
- by exploring sparsity of Gauss-Newton updates

Improve* the *convergence

- by drawing *independent* randomized *subsets* to remove *biases*
- by *repeated* drawings of *subsets* to touch *all* information

Wave-equation based *inversion*

Xiang Li, Tu Ning, Tristan van Leeuwen

SLIM
University of British Columbia

Challenges

Wave-equation based *inversion* is *prohibitively* expensive

Fundamentally rethink how we *invert* seismic data

Challenges

Wave-equation based *inversion* is *prohibitively* expensive

Fundamentally rethink how we *invert* seismic data

- **by *simulating* waves w/ *dynamic* accuracy**

Challenges

Wave-equation based *inversion* is *prohibitively* expensive

Fundamentally rethink how we *invert* seismic data

- **by *simulating* waves w/ dynamic accuracy**
- **by working w/ *subsets* of *shots* only**

Challenges

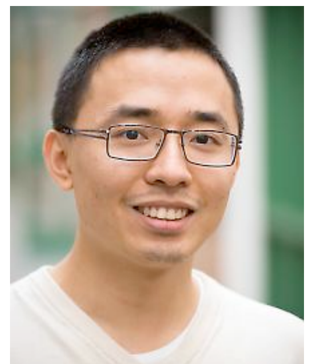
Wave-equation based *inversion* is *prohibitively* expensive

Fundamentally rethink how we *invert* seismic data

- **by simulating waves w/ dynamic accuracy**
- **by working w/ *subsets* of *shots* only**

Examples in wave-equation based inversion:

- *frugal* FWI w/ batching & dynamic accuracy
- FWI w/ *compressive* sensing & sparse recovery
- *Fast* RTM w/ multiples



Frugal FWI

SLIM

University of British Columbia

Objective

Make 3-D FWI computationally affordable.

Questions:

Can we *approximate* the “physics” when we are *far* from the *solution*?

Is it ok to work with *subsets* of *shots*?

Do we *arrive* at the *same* result?

Frugal misfit

Algorithm 1 $\{f, \mathbf{g}\} = \mathsf{misfit}(\mathbf{m}, \mathcal{I}, \eta)$

$\epsilon = 10^{-2}$, $\alpha = 0.5$// Initialization
for $i \in \mathcal{I}$ **do**
  for $k = 0 \to 10$ **do**
    solve $A(\mathbf{m})\mathbf{u} = \mathbf{s}_i$ up to ϵ // solve forward equation
    $r_k = \rho(P_i\mathbf{u} - \mathbf{d}_i)$ // compute residual
    if $|r_k - r_{k-1}| \leq \eta r_k$ **then**
      break
    else
      $\epsilon = \alpha\epsilon$
    end if
  end for
  solve $A(\mathbf{m})^*\mathbf{v} = {P_i}^*\nabla\rho(P_i\mathbf{u} - \mathbf{d}_i)$ up to ϵ
  $f = f + |\mathcal{I}|^{-1}\rho(P_i\mathbf{u} - \mathbf{d}_i)$ // misfit
  $\mathbf{g} = \mathbf{g} + |\mathcal{I}|^{-1}G(\mathbf{m}, \mathbf{u})^*\mathbf{v}$ // gradient
end for

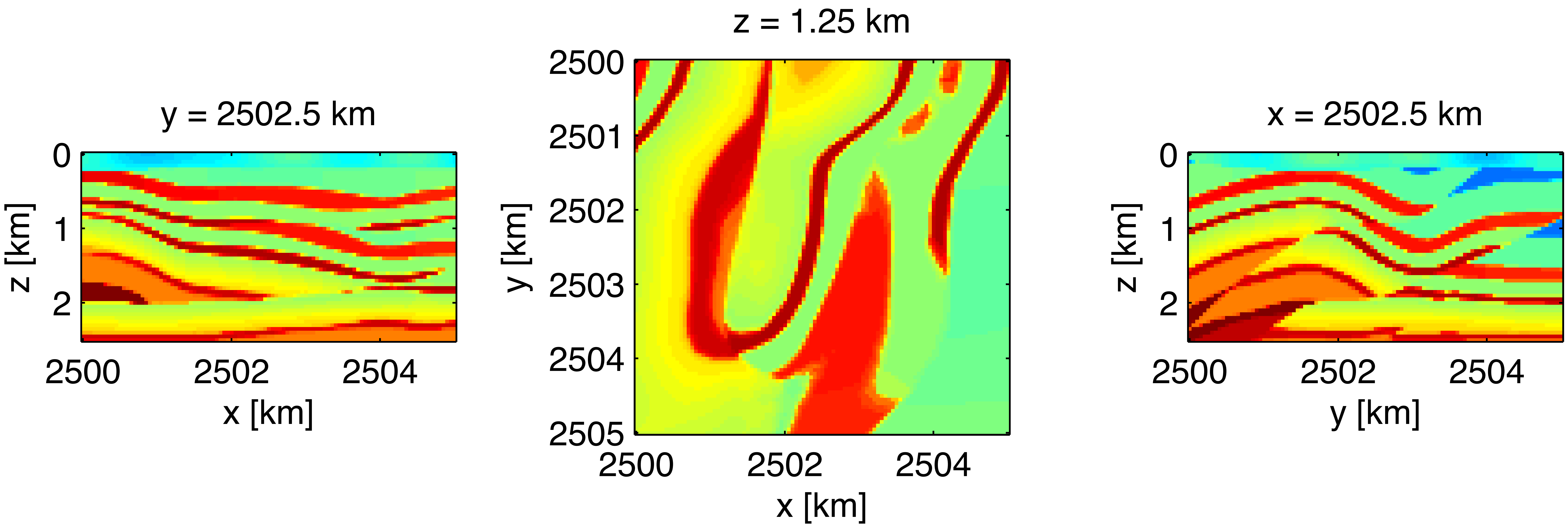
Stochastic
Quasi-Newton

Algorithm 1 Stochastic L-BFGS method

$\eta = 0.1,\ b = 1,\ \beta = 1,\ b_{\max} = M$ // Initialize
choose $\mathcal{I}_0 \subseteq \{1, 2, \dots, M\}$ s.t. $|\mathcal{I}_0| = b$
$\{f_0, \mathbf{g}_0\} = \mathsf{misfit}(\mathbf{m}_0, \mathcal{I}_0, \eta)$ // *frugal* misfit & gradient at initial guess
while not converged **do**
  $\delta\mathbf{m}_k = \mathsf{lbfgs}(-\mathbf{g}_k, \{\mathbf{t}_l\}_{l=k-m}^{k}, \{\mathbf{y}_l\}_{l=k-m}^{k})$ // low-rank inverse Hessian
  $\{\mathbf{m}_{k+1}, f_{k+1}, \mathbf{g}_{k+1}\} = \mathsf{linesearch}(f_k, \mathbf{g}_k, \delta\mathbf{m}_k)$
  if linesearch successfull **then**
    $\mathbf{t}_{k+1} = \mathbf{m}_{k+1} - \mathbf{m}_k,\ \mathbf{y}_{k+1} = \mathbf{g}_{k+1} - \mathbf{g}_k$ // update L-BFGS vectors
    choose $\mathcal{I}_{k+1} \subseteq \{1, 2, \dots, M\}$ s.t. $|\mathcal{I}_{k+1}| = b$ // draw new sample
    $\{f'_{k+1}, \mathbf{g}'_{k+1}\} = \mathsf{misfit}(\mathbf{m}_{k+1}, \mathcal{I}_{k+1}, \eta)$ // misfit & gradient new sample
    if $(f_{k+1} + f'_{k+1}) \geq (f_k + f'_k)$ **then**
      $b = \min(b + \beta, b_{\max})$ // increase batch
    end if
    $f_{k+1} = f'_{k+1},\ \mathbf{g}_{k+1} = \mathbf{g}'_{k+1},\ k = k + 1$ // Use new misfit & gradient
  else
    $\eta = \eta/2$ // narrow tolerenance
  end if
end while

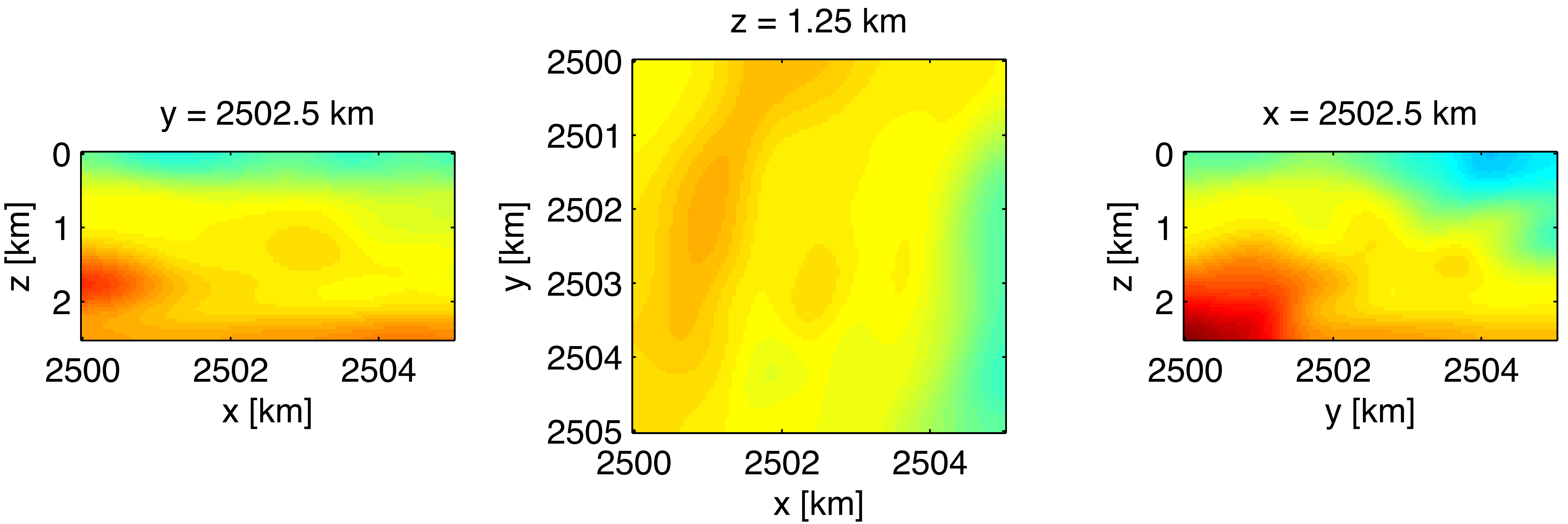
Overthrust model

true model
5km X 5km X 2.5Km
121 sources & 2601 receivers



Overthrust model

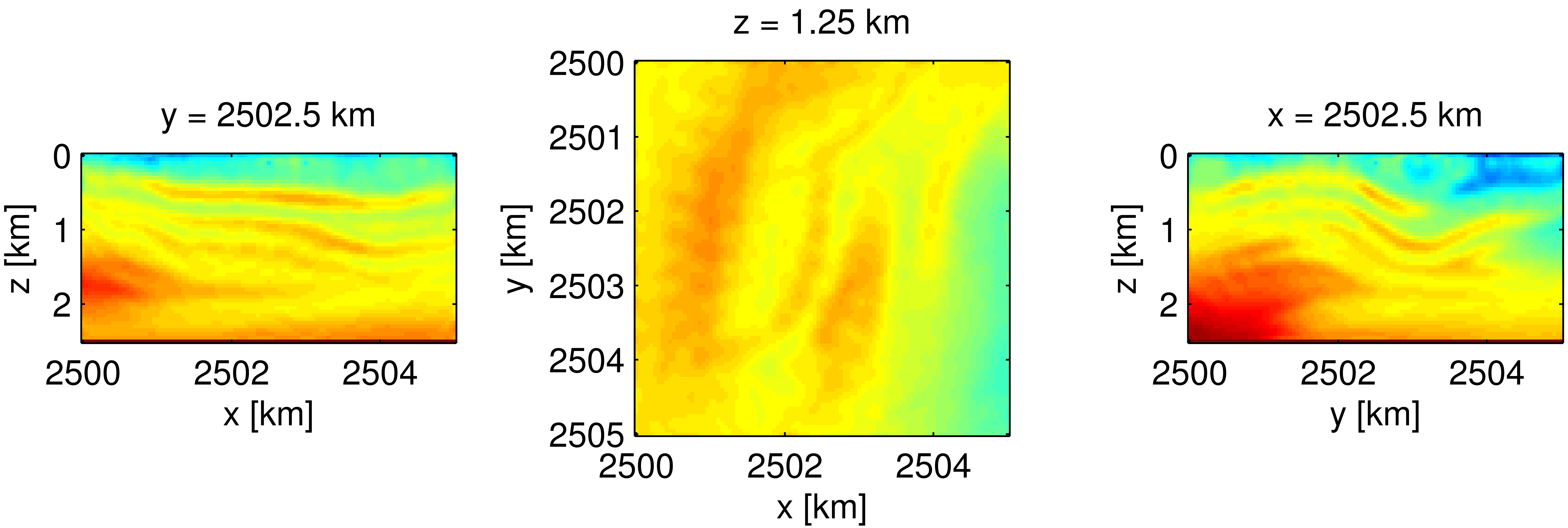
initial model



Overthrust model

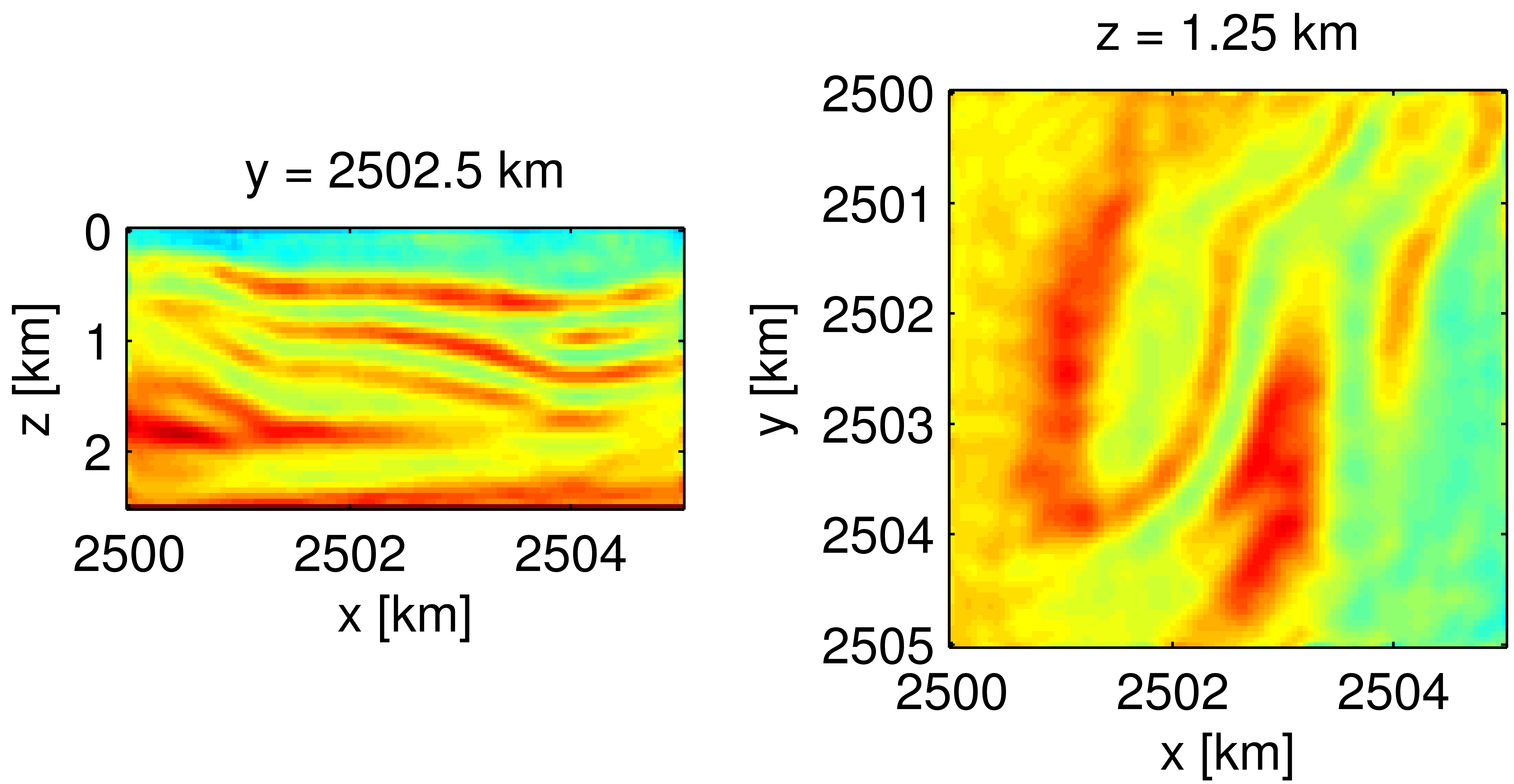
recovered** model **w/ b=121

2 passes** through **data** for each **(4,6,8) Hz



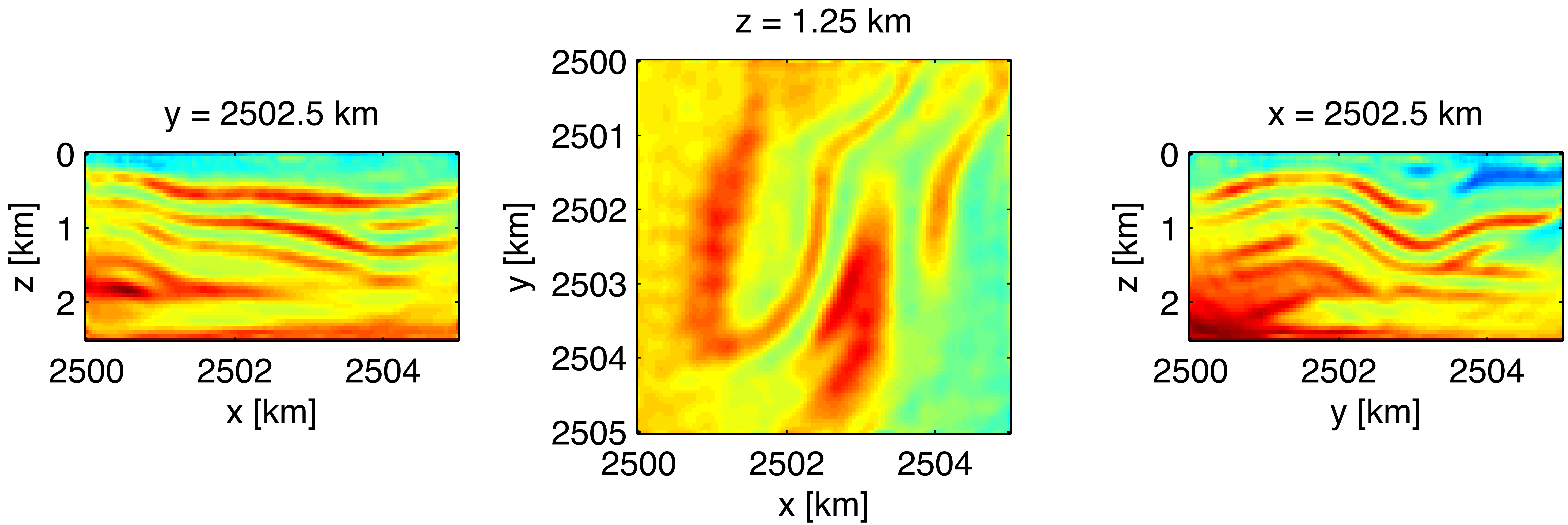
Overthrust model

growing sample size
2 passes through *data* for each *(4,6,8) Hz*



Overthrust model

growing sample size
10 passes through *data* for each *(4,6,8) Hz*



Aleksandr Y. Aravkin, Michael P. Friedlander, Felix J. Herrmann, and Tristan van Leeuwen, "**Robust inversion, dimensionality reduction, and randomized sampling**", *Mathematical Programming*, vol. 134, p. 101-125, 2012
Felix J. Herrmann et. al., "**Frugal full-waveform inversion: from theory to a practical algorithm**", *The Leading Edge*, vol. 32, p. 1082-1092, 2013

SLIM

Observations

Questions:

Can we *approximate* the "physics" when we are *far* from the *solution*?

- ✓ yes, as long as we *dial in* the *accuracy* to *guarantee* convergence

Is it ok to work with *subsets* of shots?

- ✓ yes, as long as we increase the size of the subsets to *guarantee* convergence

Do we *arrive* at the *same* result?

- ✓ yes, under certain conditions

FWI w/ compressive sensing & sparse recovery

SLIM

University of British Columbia

Objective

Make FWI work w/ incomplete physics (elastic data / acoustic modeling kernel)

Questions:

Can we use *sparsity* promotion on model updates to improve FWI?

Can we do this efficiently?

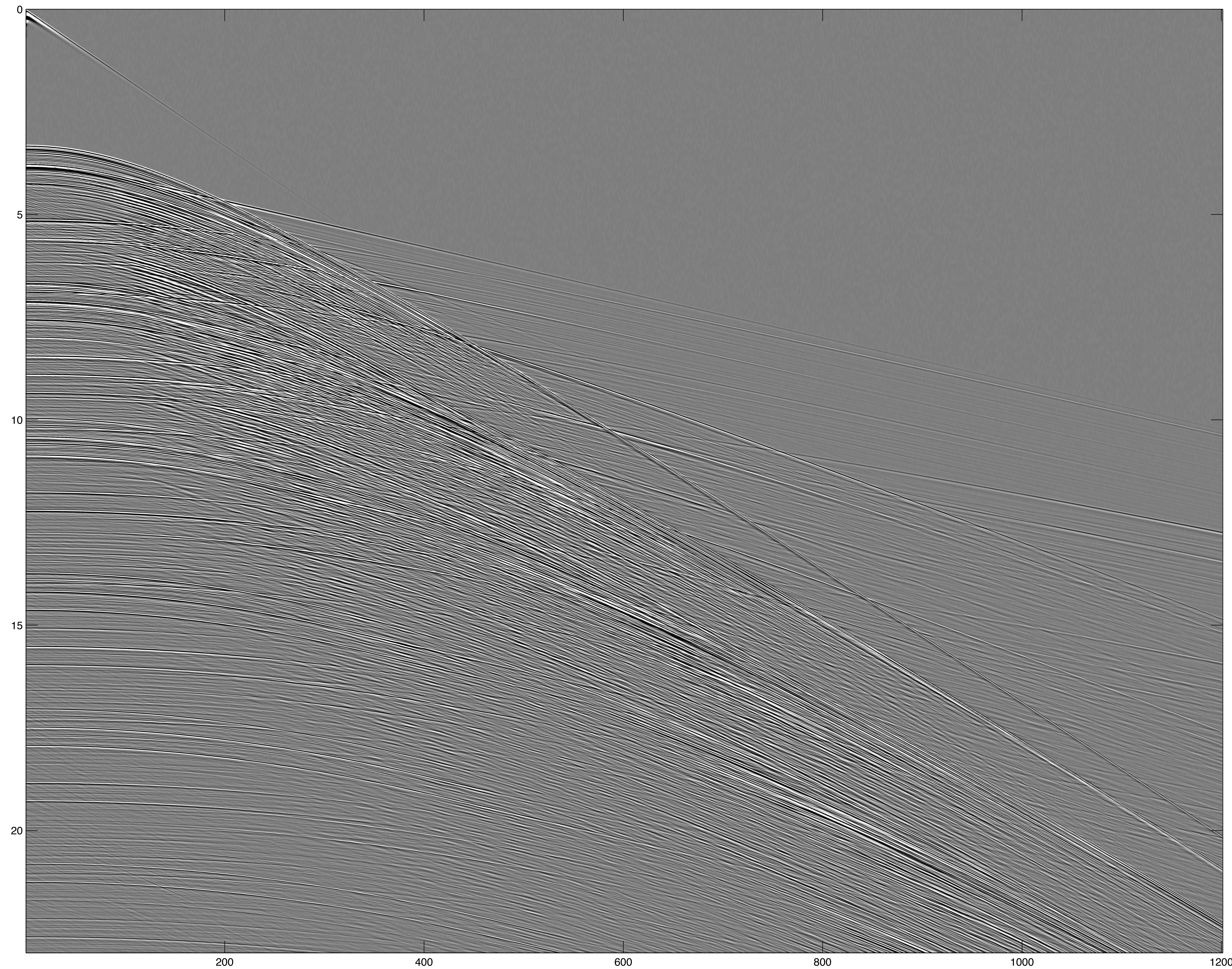
Is FWI a *mature* technology?

Chevron GoM FWI dataset

Large 2D elastic but isotropic dataset

30 km offset

22 s of data



Xiang Li, Aleksandr Y. Aravkin, Tristan van Leeuwen, and Felix J. Herrmann, "Fast randomized full-waveform inversion with compressive sensing", *Geophysics*, vol. 77, p. A13-A17, 2012

Algorithm

modified Gauss Newton

Result: Output estimate for the model $\mathbf{m}$

$k \longleftarrow 0; \mathbf{m}^k \longleftarrow \mathbf{m}_0$

while not converged **do**

$\quad \{\underline{\mathbf{D}}^k, \underline{\mathbf{Q}}^k\} \longleftarrow \{\mathbf{D}\mathbf{W}^k, \mathbf{Q}\mathbf{W}^k\}$ with $\mathbf{W}^k \subset [\mathbf{e}_1, \cdots, \mathbf{e}_{n_s}]$

$\quad \underline{\delta\mathbf{D}}^k \longleftarrow \underline{\mathbf{D}}^k - \boldsymbol{\mathcal{F}}[\mathbf{m}^k; \underline{\mathbf{Q}}^k]$

$\quad \tau^k \longleftarrow \|\underline{\delta\mathbf{D}}^k\|_F \,/\, \|\mathbf{C}_2 \nabla\boldsymbol{\mathcal{F}}^*[\mathbf{m}^k; \underline{\mathbf{Q}}^k]\underline{\delta\mathbf{D}}^k\|_\infty$

$\quad \delta\mathbf{x} \leftarrow \arg\min_{||\mathbf{x}||_1 \leq \tau_k} \|\underline{\delta\mathbf{D}}^k - \nabla\boldsymbol{\mathcal{F}}[\mathbf{m}^k; \underline{\mathbf{Q}}^k]\mathbf{C}_2^H\mathbf{x}\|_F^2$

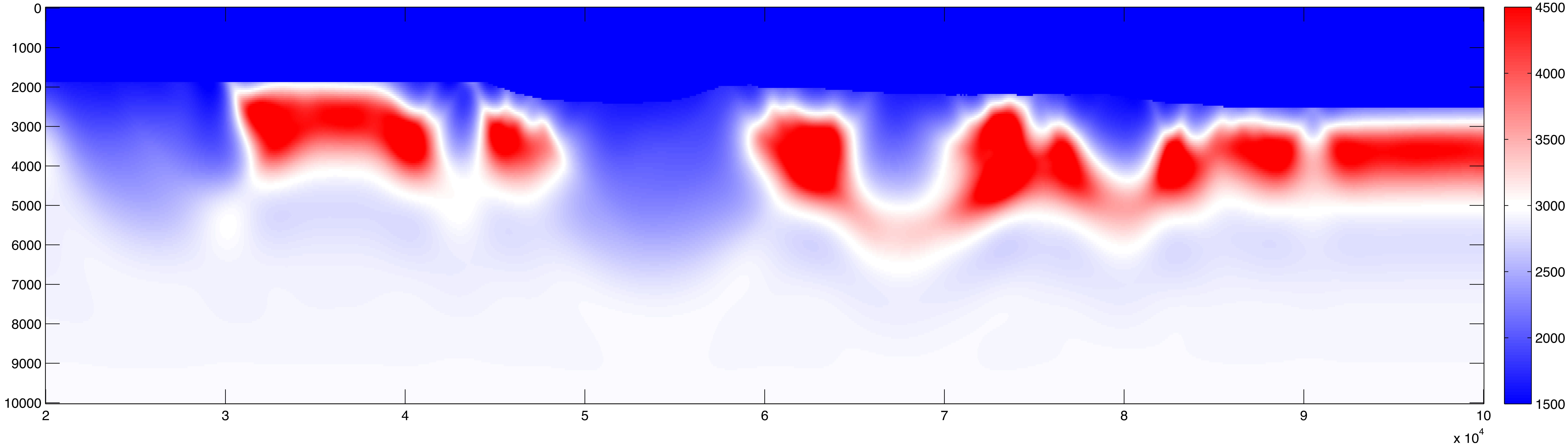
$\quad \mathbf{m}^{k+1} \longleftarrow \mathbf{m}^k + \gamma^k \mathbf{C}_2^H \delta\mathbf{x}$

$\quad k \longleftarrow k + 1;$

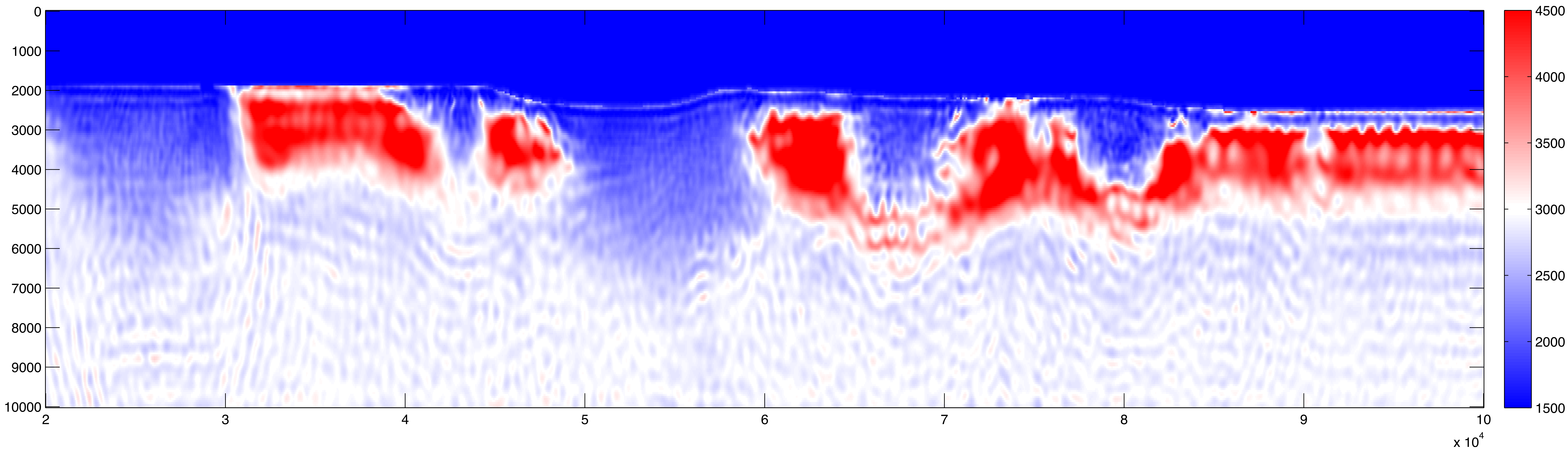
end

Algorithm 1: modified Gauss Newton with sparsity promotion

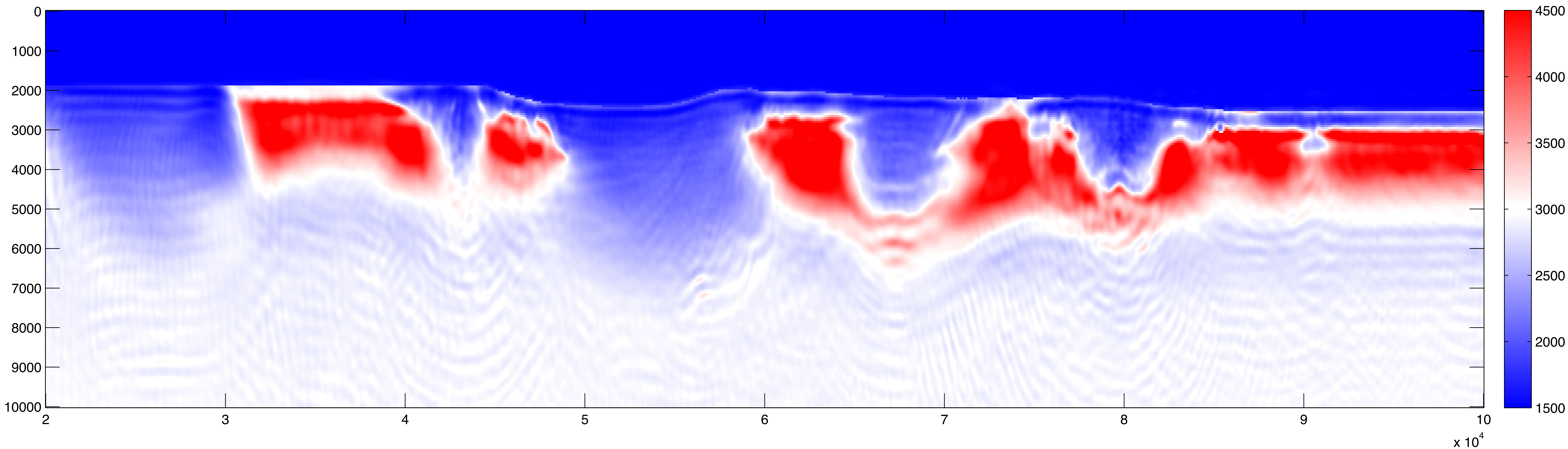
Initial model
ray-based tomography



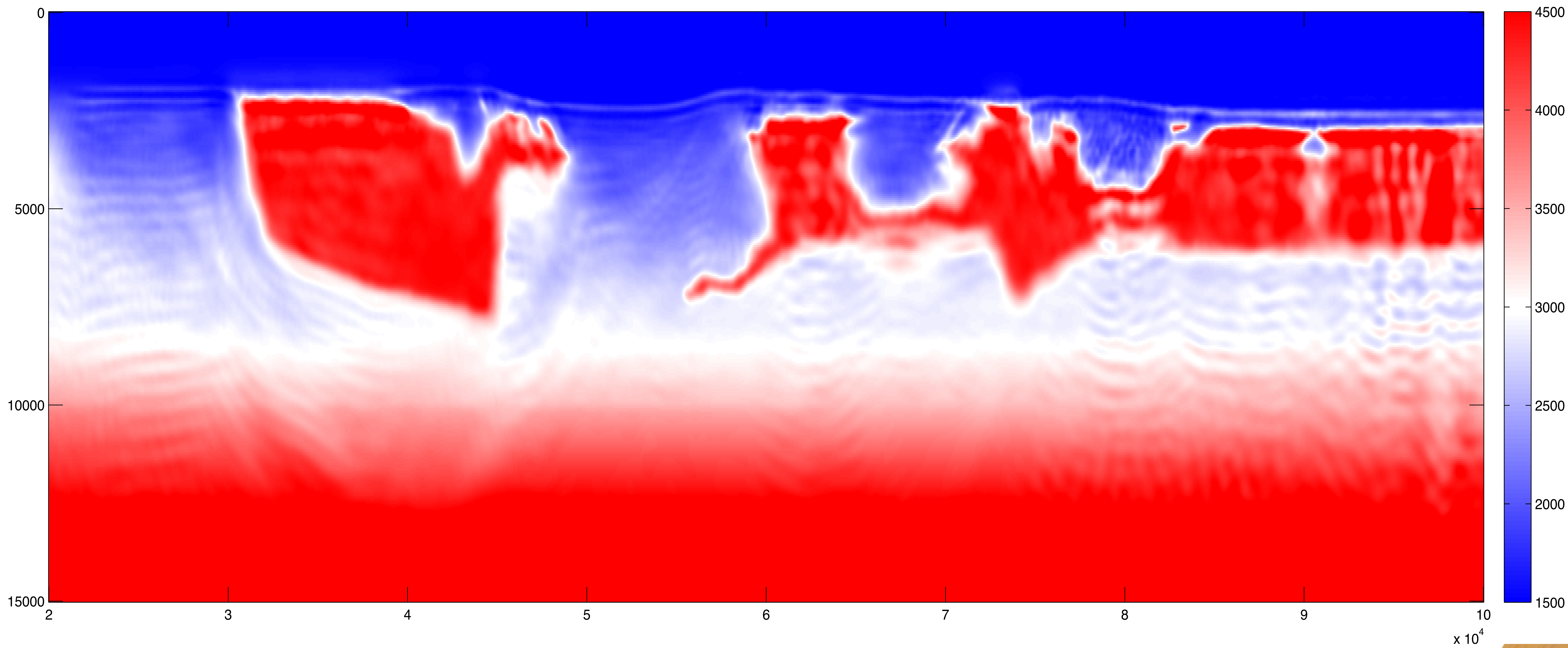
FWI result
w/ Gauss Newton



FWI result
w/ modified Gauss Newton (sparsity promotion)



“Latest” result



Xiang Li, Aleksandr Y. Aravkin, Tristan van Leeuwen, and Felix J. Herrmann, "Fast randomized full-waveform inversion with compressive sensing", *Geophysics*, vol. 77, p. A13-A17, 2012
Felix J. Herrmann et. al., "Frugal full-waveform inversion: from theory to a practical algorithm", *The Leading Edge*, vol. 32, p. 1082-1092, 2013

SLIM

Observations

Questions:

Can we use *sparsity* promotion on model updates to improve FWI?

- ✓ yes/no, inversion improves but FWI for *salt* remains *extremely* challenging

Can we do this efficiently?

- ✓ yes, by using rerandomizations–drawing *independent* subsets of shots

Is FWI a *mature* technology?

- no, major issues with local *minima* & *quality* starting models

Fast RTM w/ multiples

SLIM

University of British Columbia

Objective

***Sparsity*-promoting RTM w/ multiples at cost of a ‘single’ RTM.**

Questions:

Do *multiples* provide *additional* information?

Can we map *multiples* to *primaries* instead of *eliminating* them?

Can we do this *cheaply*?

Model & acquisition geometry

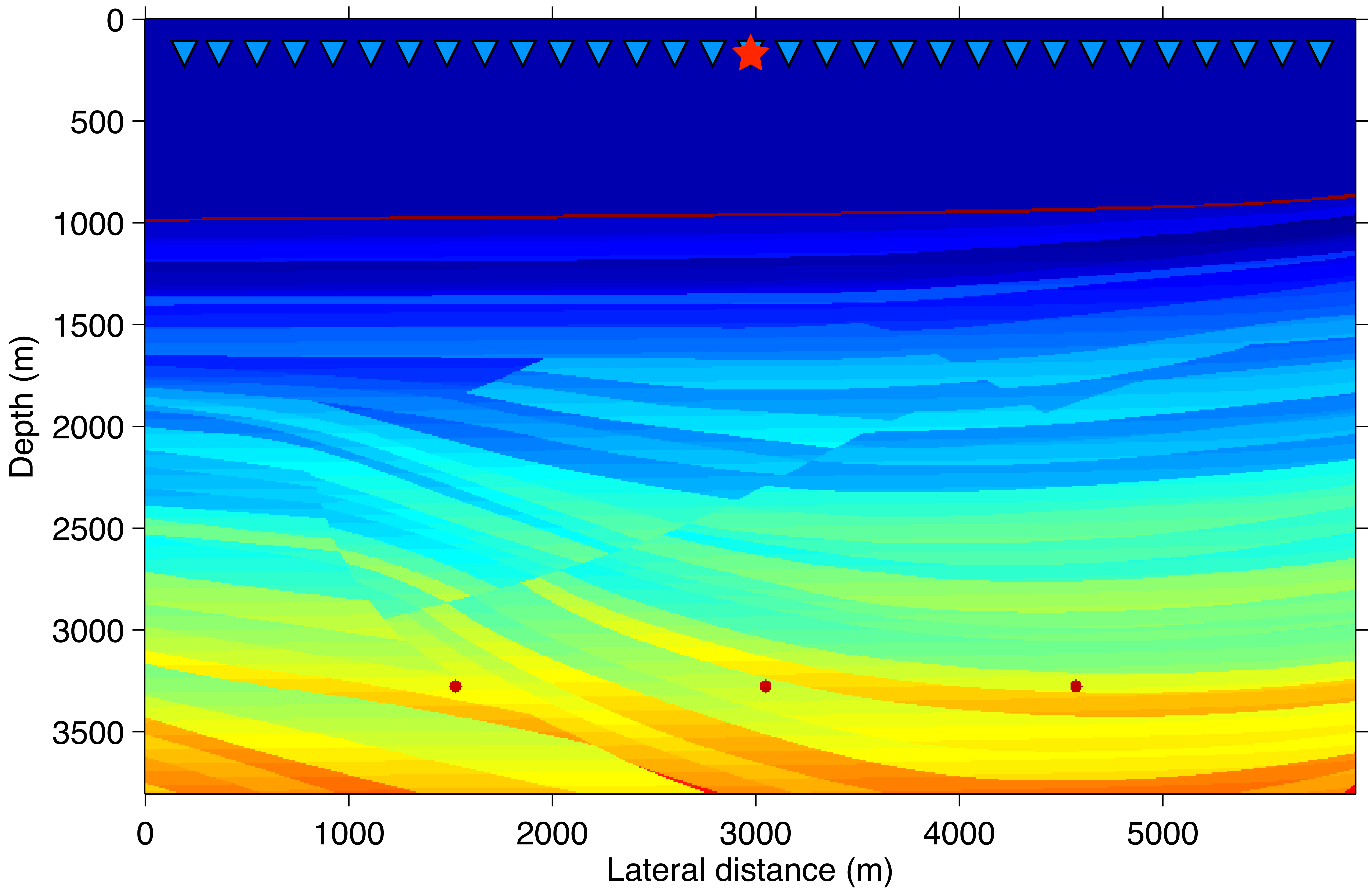


Illustration: primary wave propagation

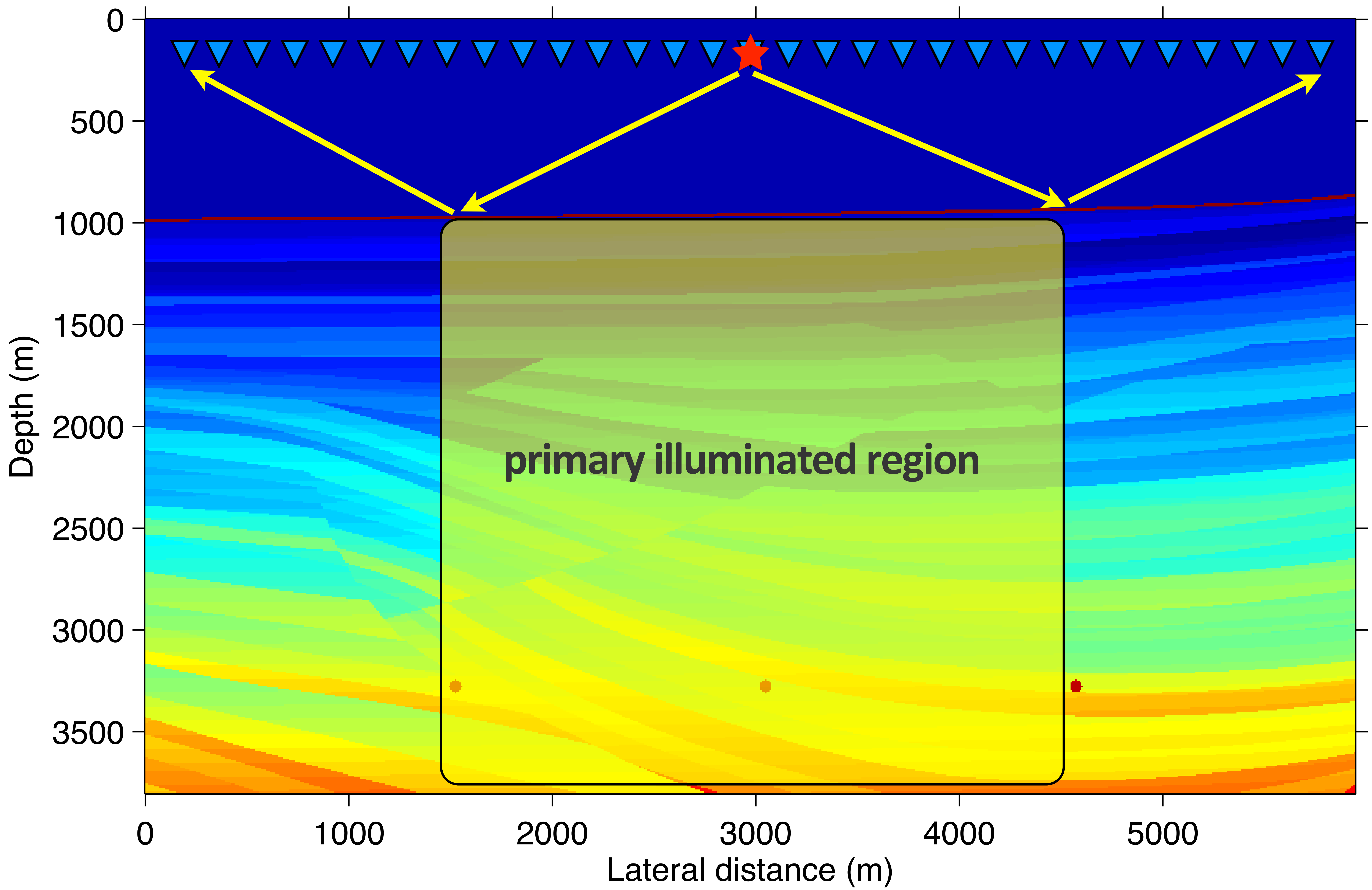
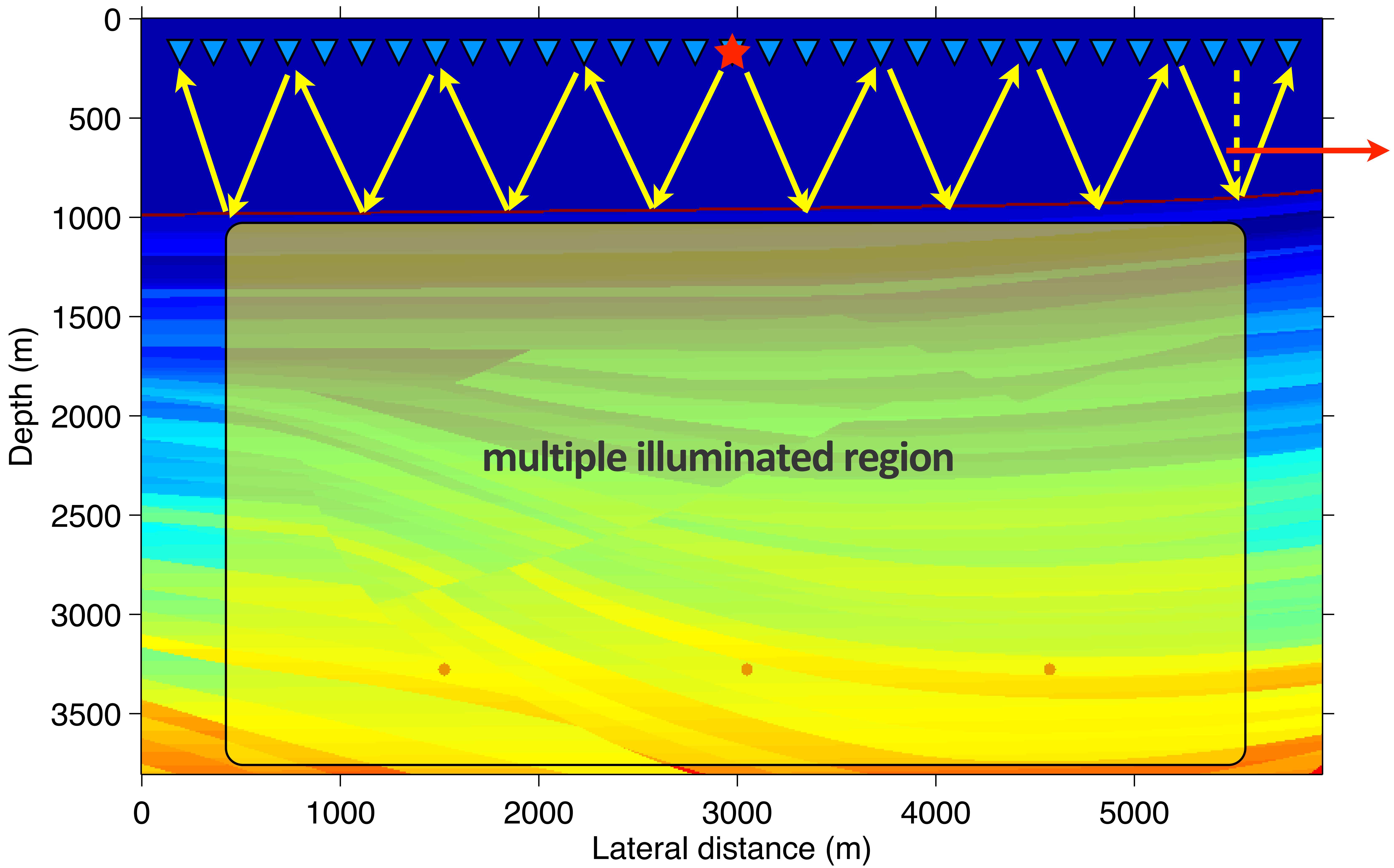


Illustration: surface-multiples propagation

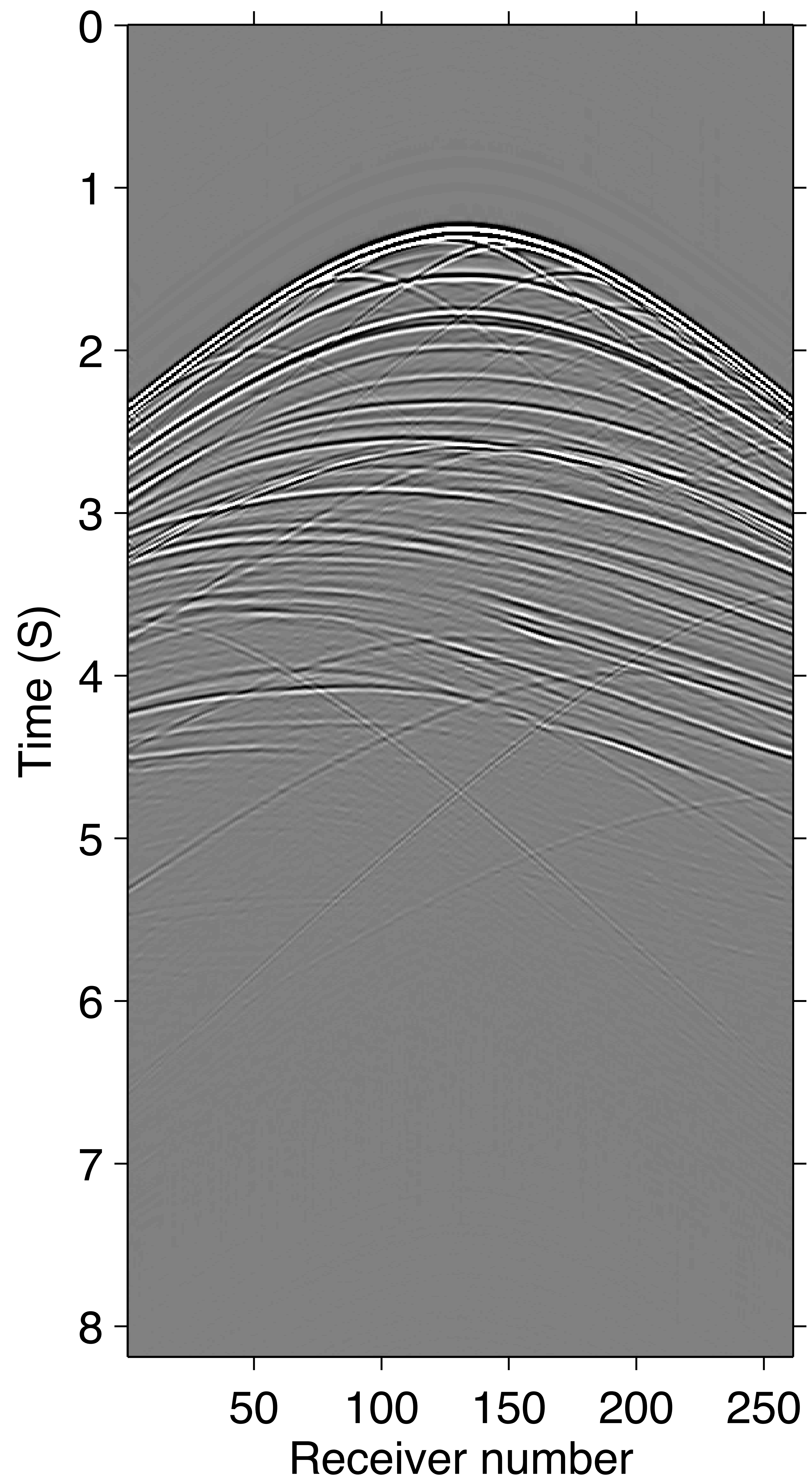
[Each receiver serves as a virtual secondary source!]



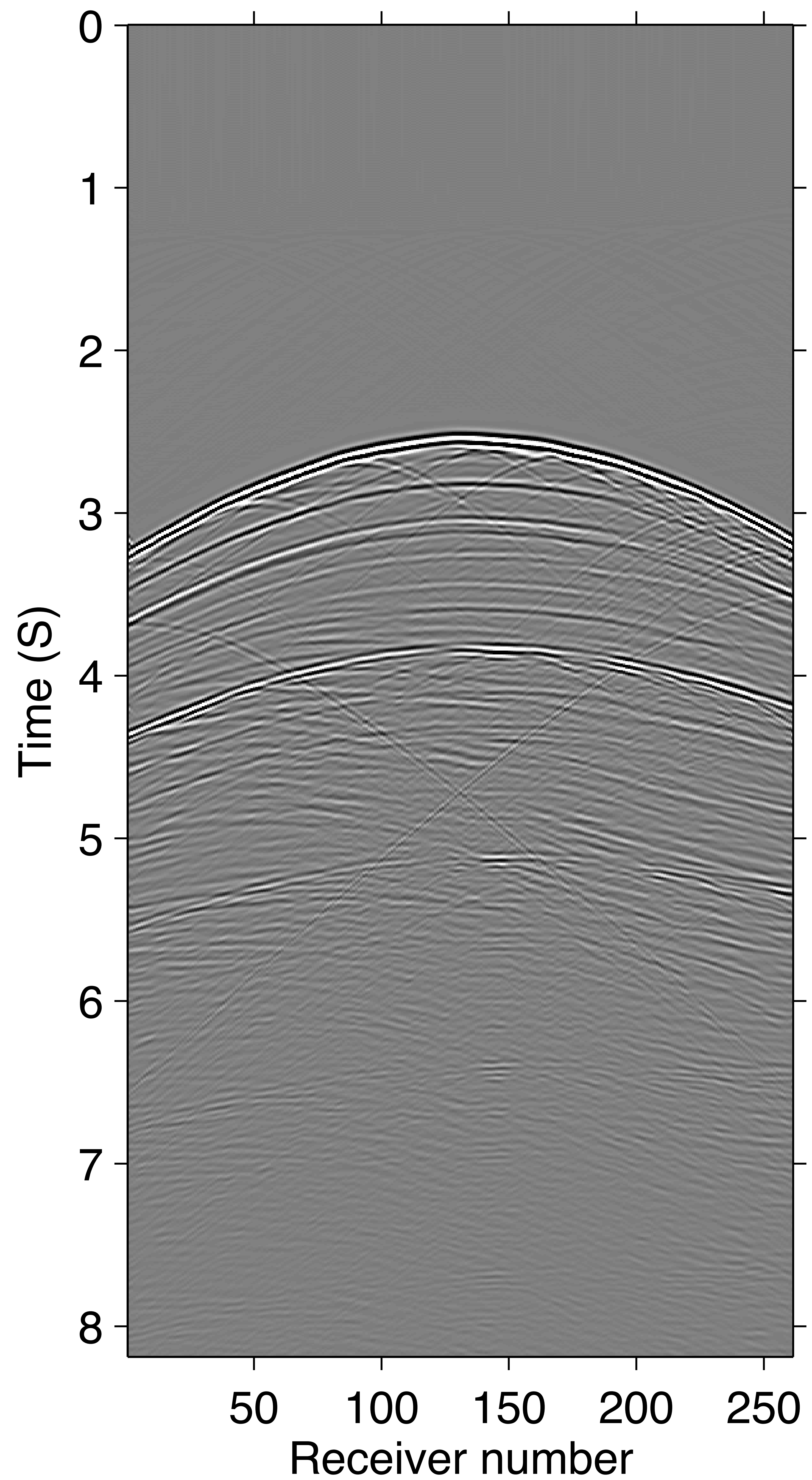
much smaller incident angles compared to primaries

Shot record: primaries

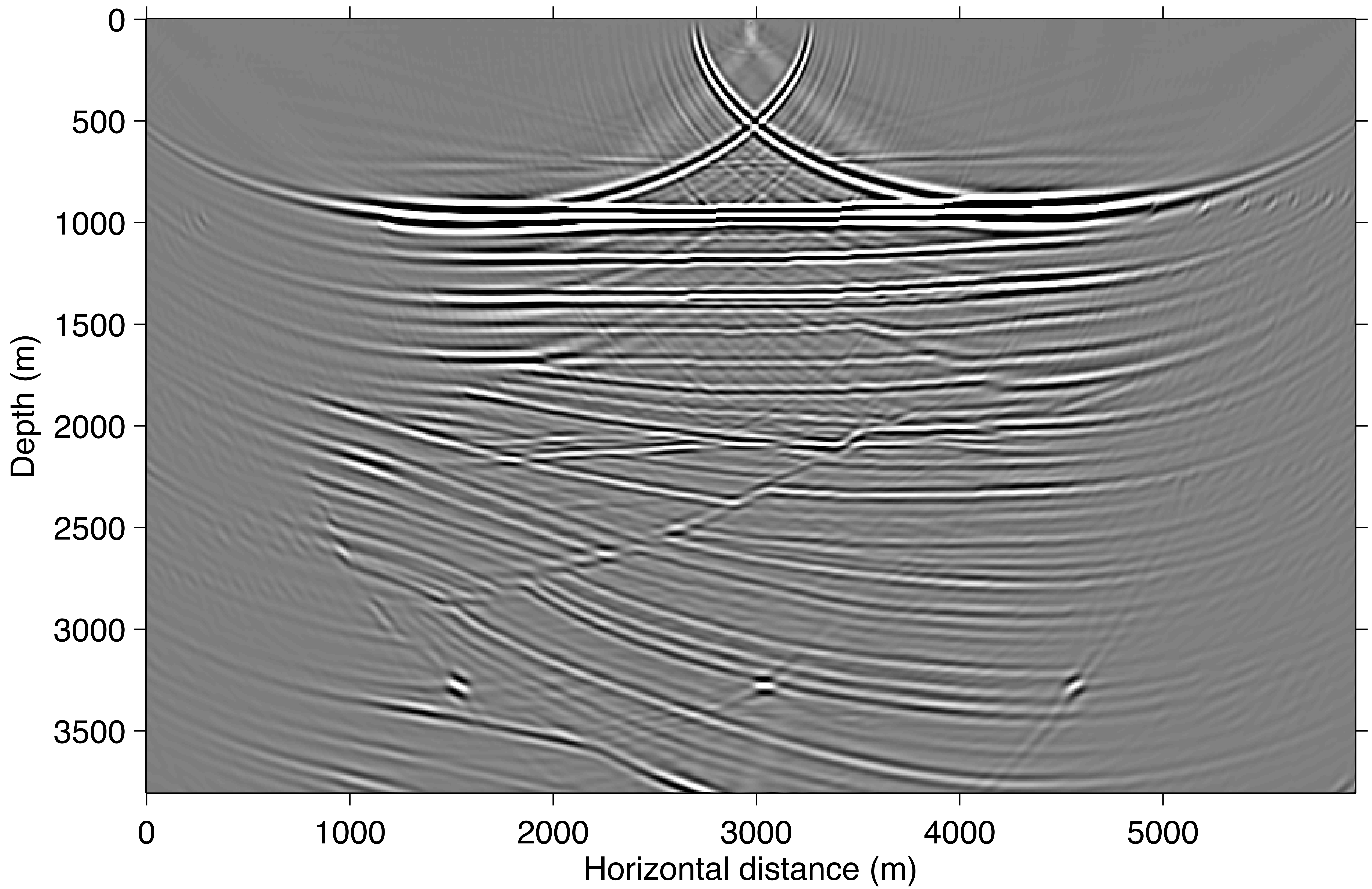
- 15Hz Ricker wavelet
- data generated w/ different code
- inversions w/ 311 frequencies



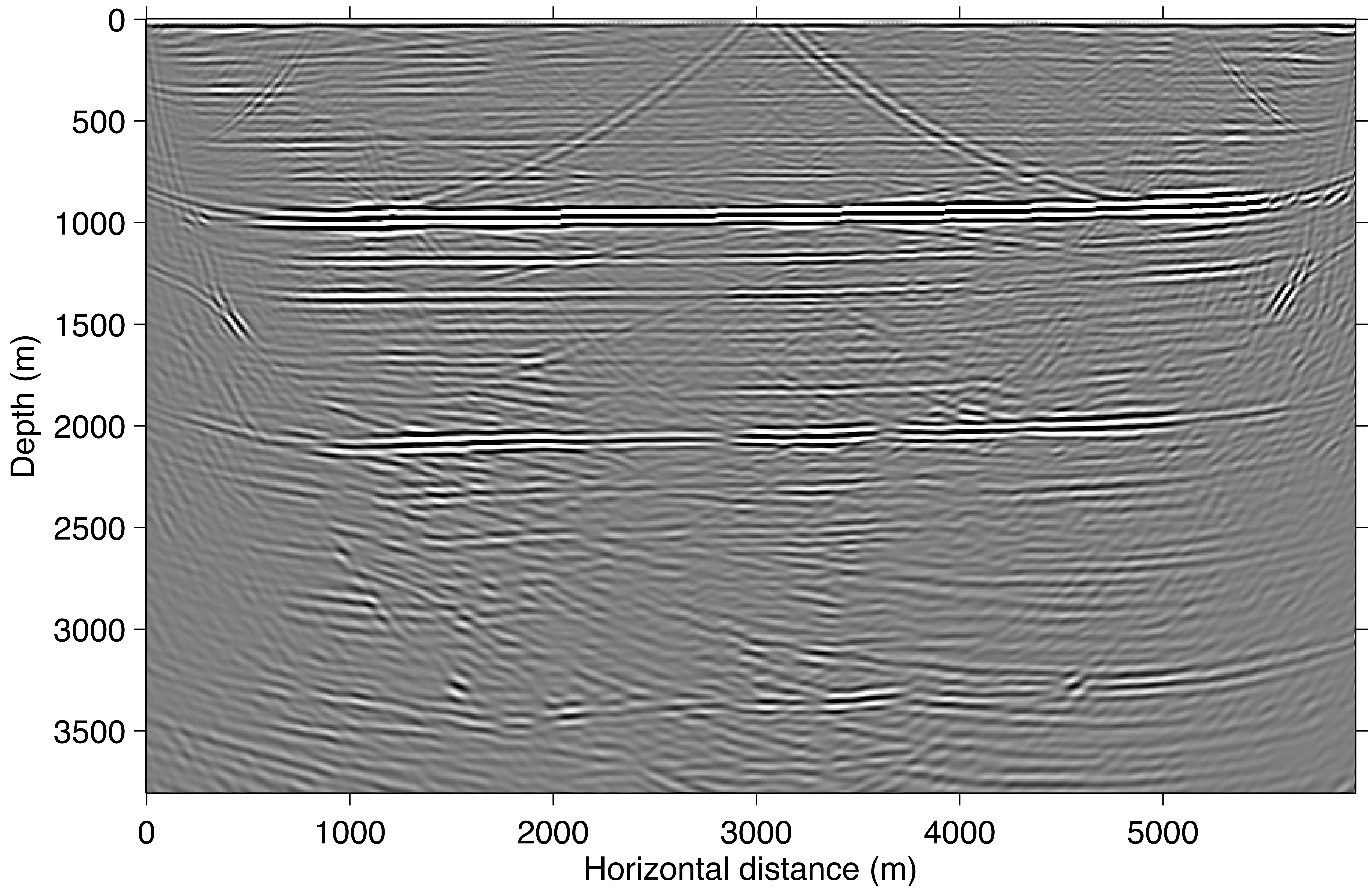
Shot record: multiples



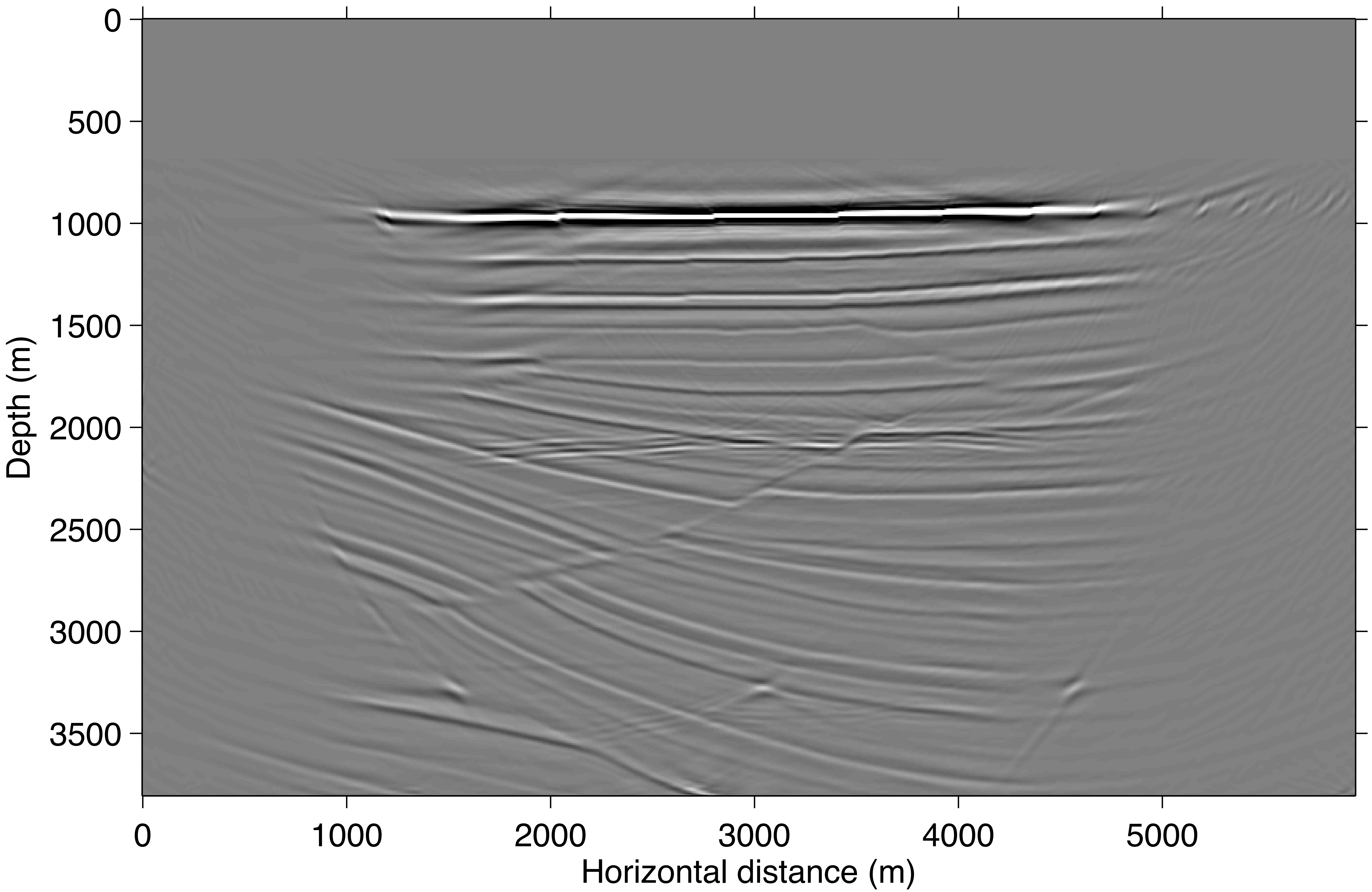
Illuminated region by primaries (RTM image)
single shot



Illuminated region by multiples (RTM image)
single shot

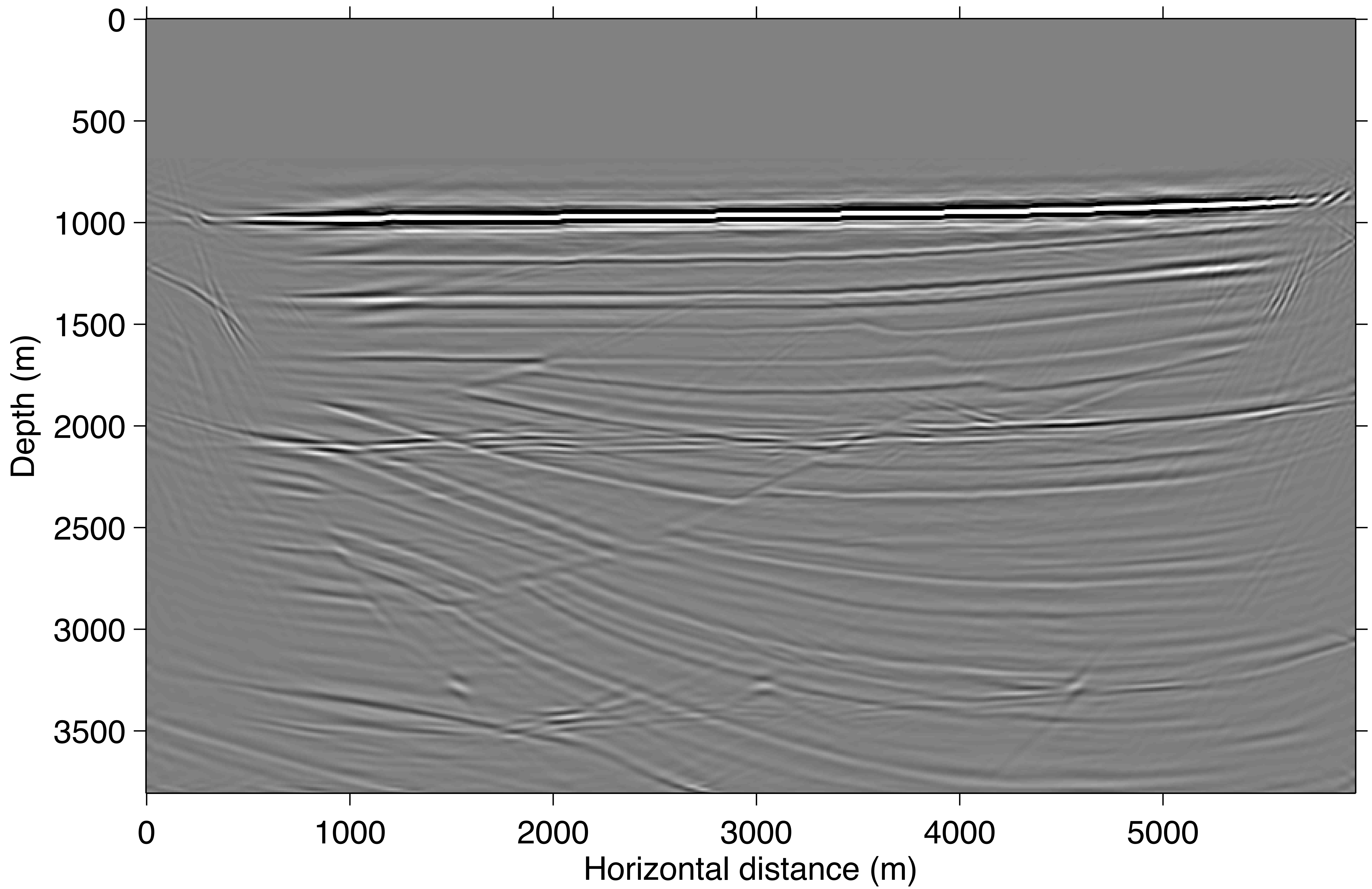


Sparsity-promoting inversion w/ primaries only
single shot



Sparsity-promoting inversion w/ multiples only

single shot



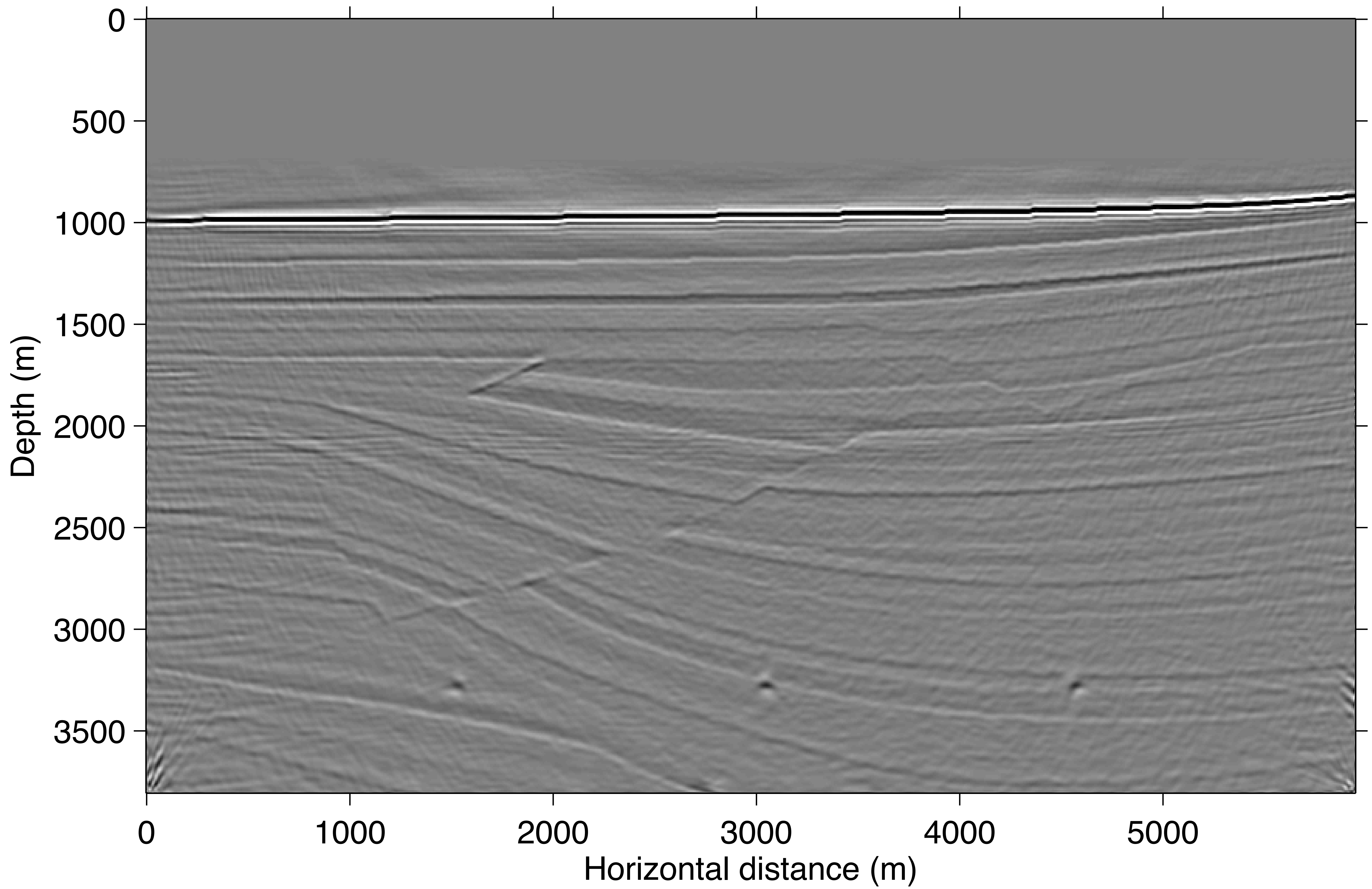
Formulation that embraces multiples

$$\min_{\mathbf{x}} \sum_{i\in\mathbb{F}} \|\underline{\mathbf{d}}_i - \nabla\mathbf{F}_i[\mathbf{m}_0, \underline{\mathbf{Q}}(q_i)]\mathbf{S}^*\mathbf{x}\|_2^2$$

$$\text{subject to } \|\mathbf{x}\|_1 \leq \tau$$

- $\mathbf{d}_i$: vectorized total *up*-going wavefield, primaries and surface multiples
- $\mathbf{Q}(q_i) = q_i\mathbf{I} - \mathbf{D}_i$: *areal* source wavefield w/ total *down*-going wavefield

Fast inversion w/ multiples [assuming we have the true source]
cost of ~1 RTM



Felix J. Herrmann and Xiang Li, "**Efficient least-squares imaging with sparsity promotion and compressive sensing**", *Geophysical Prospecting*, vol. 60, p. 696-712, 2012
Ning Tu and Felix J. Herrmann, "**Imaging with multiples accelerated by message passing**", *SEG*, 2012.

Observations

Questions:

Do *multiples* provide *additional* information?

✓ yes, they provide *complementary* broader *aperture* information

Can we map *multiples* to *primaries* instead of *eliminating* them?

✓ yes, by *sparse* inversion & inclusion of *total* downgoing wavefields as *areal* sources

Can we do this *cheaply*?

✓ yes, if we use *rerandomization* techniques it costs 1–2 RTMs w/ *all* data

Carry home message

Randomization of field-data *acquisition*

- can lead to *improved* wavefield reconstruction
- *new* insights how to *acquire* & *process* data
- **possible pathway towards WAZ & long-offset acquisition**

Randomization of wave-equation based *inversion*

- can lead to *fast* RTM & FWI
- *new* insights how to *invert* large data volumes
- **possible pathway towards more involved physics & uncertainty quantification**

Incorporation of more *wave* physics

- can lead to *hifi* & *fast* inversions from sparsely *sampled* marine data
- **but relies on accurate initial models and that's where the challenge lies...**

Acknowledgements

Thank you for your attention !

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

BG, BP & Chevron are thanked for providing us with data

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, BP, CGG, Chevron, ConocoPhillips, ION, Petrobras, PGS, Statoil, Total SA, WesternGeco, and Woodside.