Released to public domain under Creative Commons license type BY (https://creativecommons.org/licenses/by/4.0).
Copyright (c) 2014 SLIM group @ The University of British Columbia.

Low-rank based matrix/tensor completions for the "real" (seismic) world

Felix J. Herrmann

SLIM
University of British Columbia

People

Haneet Wason

Hassan Mansour

Curt Da Silva

Ben Recht

Okan Akalin

Aleksandr Y. Aravkin

Rajiv Kumar

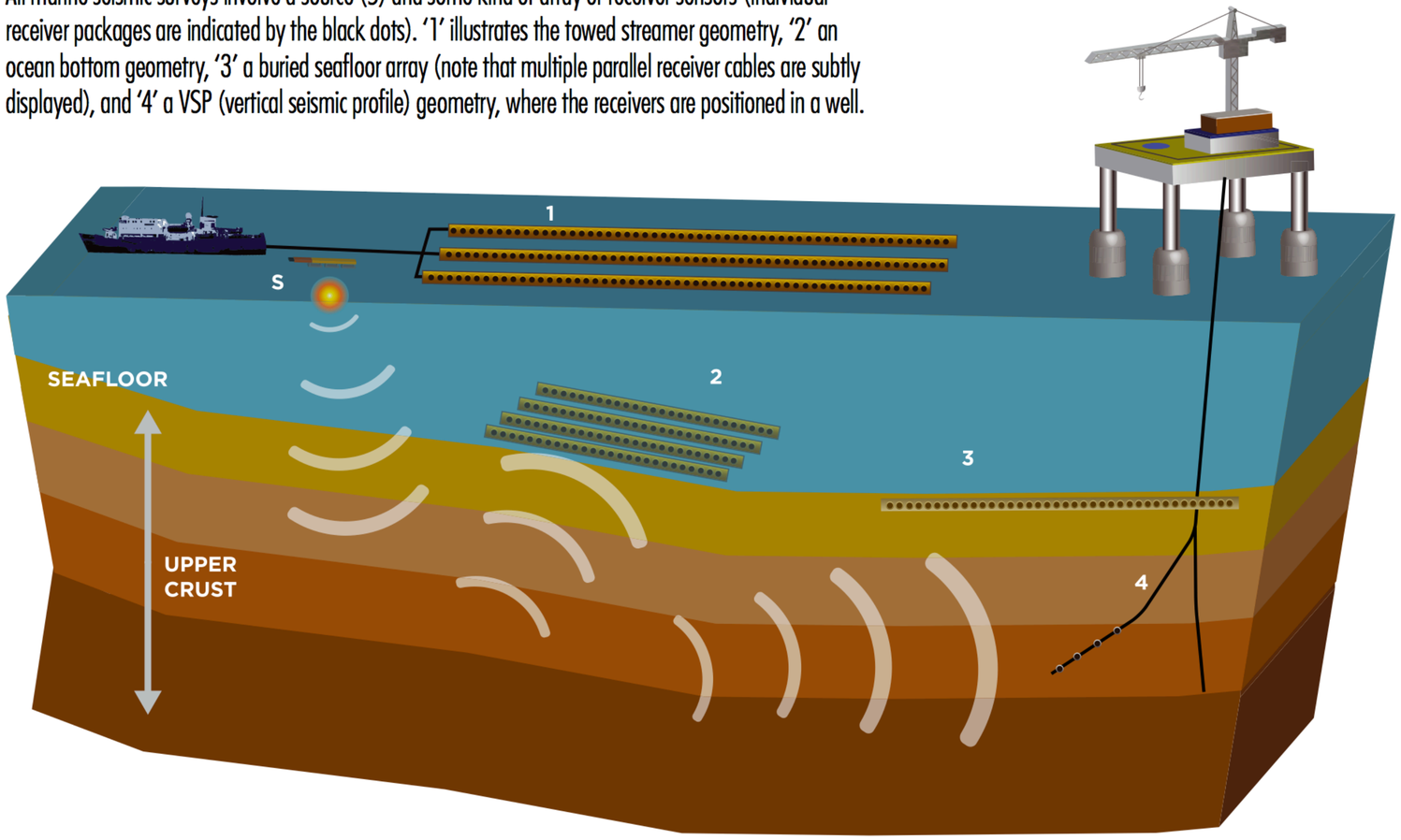
Oscar Lopez

Seismic inversion

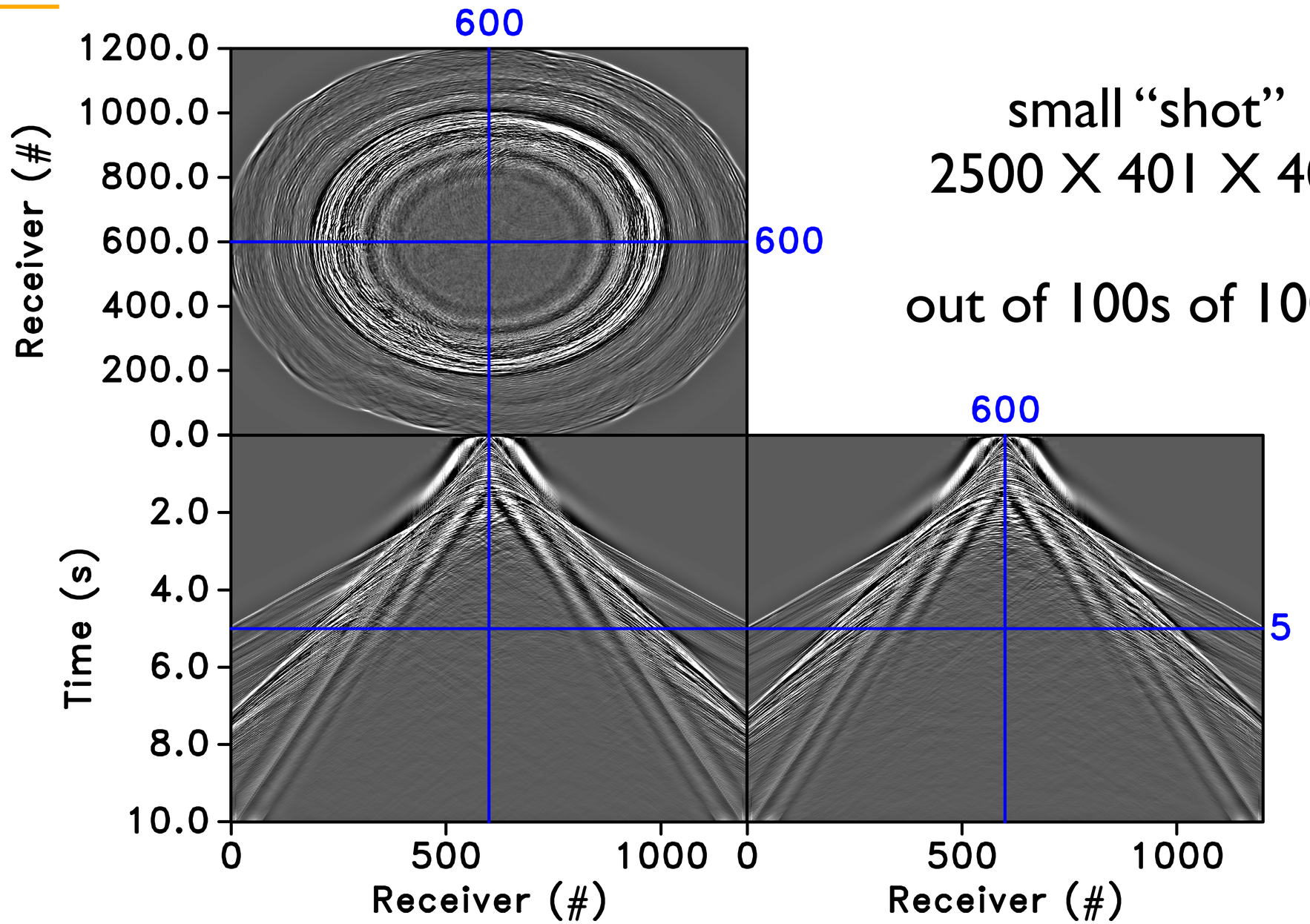
Infer 3D *velocity* model from *multi-experiment* data:

- $\mathcal{O}(10^9)$ unknowns
- $\mathcal{O}(10^{15})$ datapoints
- propagate $\mathcal{O}(10^2)$ wavelengths

All marine seismic surveys involve a source (S) and some kind of array or receiver sensors (individual receiver packages are indicated by the black dots). '1' illustrates the towed streamer geometry, '2' an ocean bottom geometry, '3' a buried seafloor array (note that multiple parallel receiver cables are subtly displayed), and '4' a VSP (vertical seismic profile) geometry, where the receivers are positioned in a well.



Data deluge in 3D seismic



small "shot"
2500 X 401 X 401

out of 100s of 1000s

from: T. Keho & P. Kelamis. the Leading Edge

Data deluge

"Moore's law" for *channel* count:

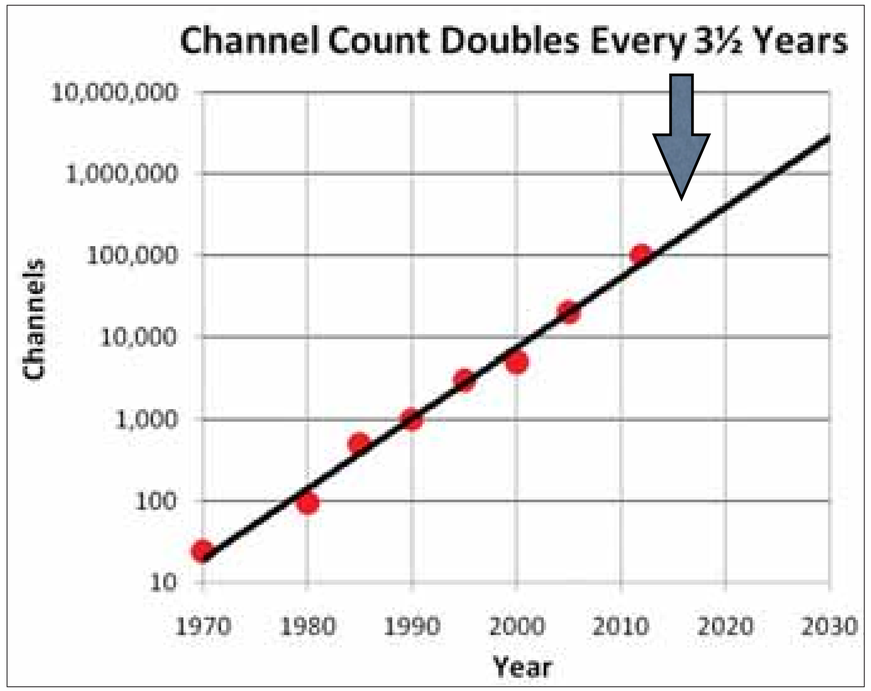


Figure 3. *"Moore's law" for seismic channel count (modified from Monk, 2006).*

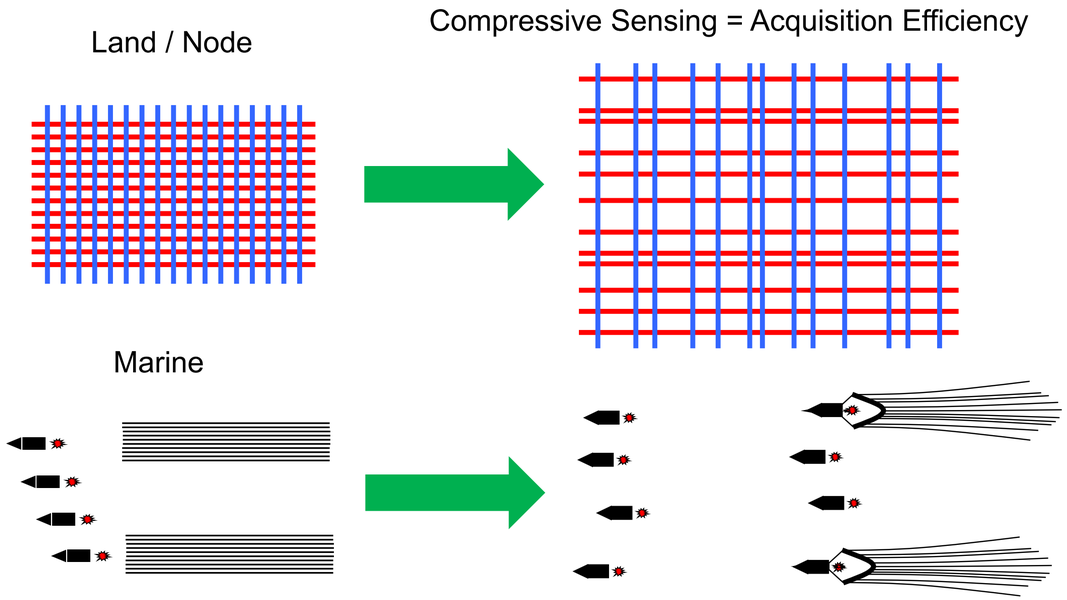
Mosher, C. C., Keskula, E., Kaplan, S. T., Keys, R. G., Li, C., Ata, E. Z., ... & Sood, S. (2012, November). Compressive Seismic Imaging. In *2012 SEG Annual Meeting*. Society of Exploration Geophysicists.

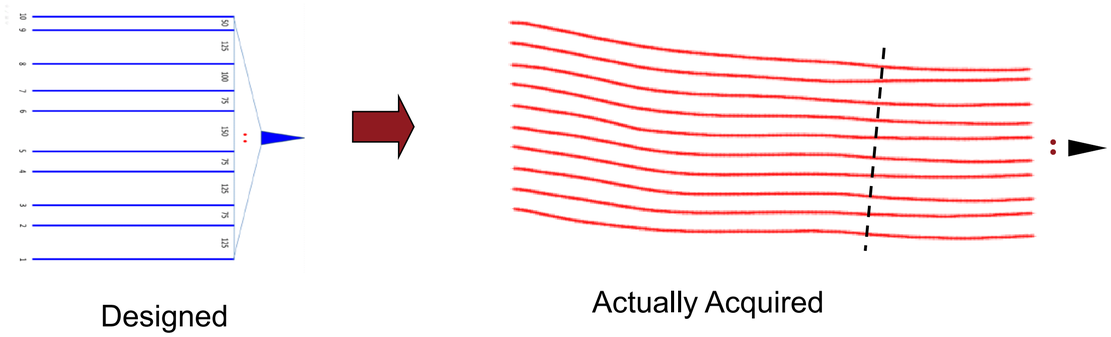
Randomized undersampling

– examples from industry (ConocoPhilips)

Deliberate & natural randomness in acquisition

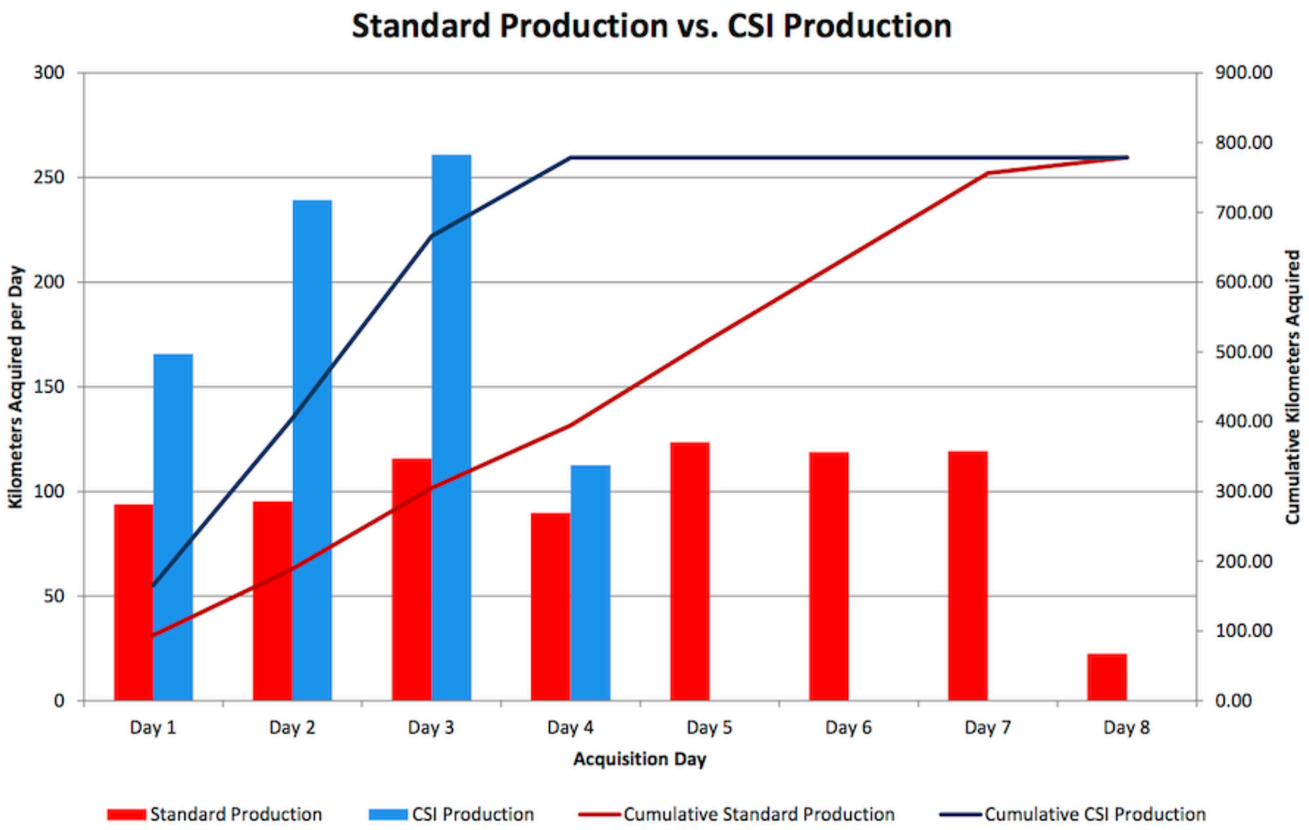
(thanks to Chuck Mosher)





CS in the field

- ConocoPhilips successfully applied CS in land & marine settings
- Lead to 3–5 X speedup
- Big impact towards bottom line & exploration in the Arctic and sensitive areas



Thanks to Chuck Mosher

Matrix completion framework

Successful reconstruction scheme

Low-complexity signal structure

- low-rank matrices in “transformed” factored, HSS,… formats

Sampling increases complexity

- pointwise sampling increases rank

Optimization decreases complexity

- “nuclear” norm-minimization
- manifold optimization

Tensor completion framework

Successful reconstruction scheme

Low-complexity signal structure

- low-rank tensors in “transformed” CP-Parafac, Tucker, Hierarchical Tucker, QTT,… formats

Sampling increases complexity

- pointwise sampling increases rank

Optimization decreases complexity

- alternating least squares, manifold optimization, …

Challenges for seismic data completion

Challenge #1 – Low rank seismic data

Recorded discrete pressure waveforms: $d(s, r, t)$

- assume that sources + receivers are on a grid for now

Fourier transform along time : $d(s, r, \omega)$

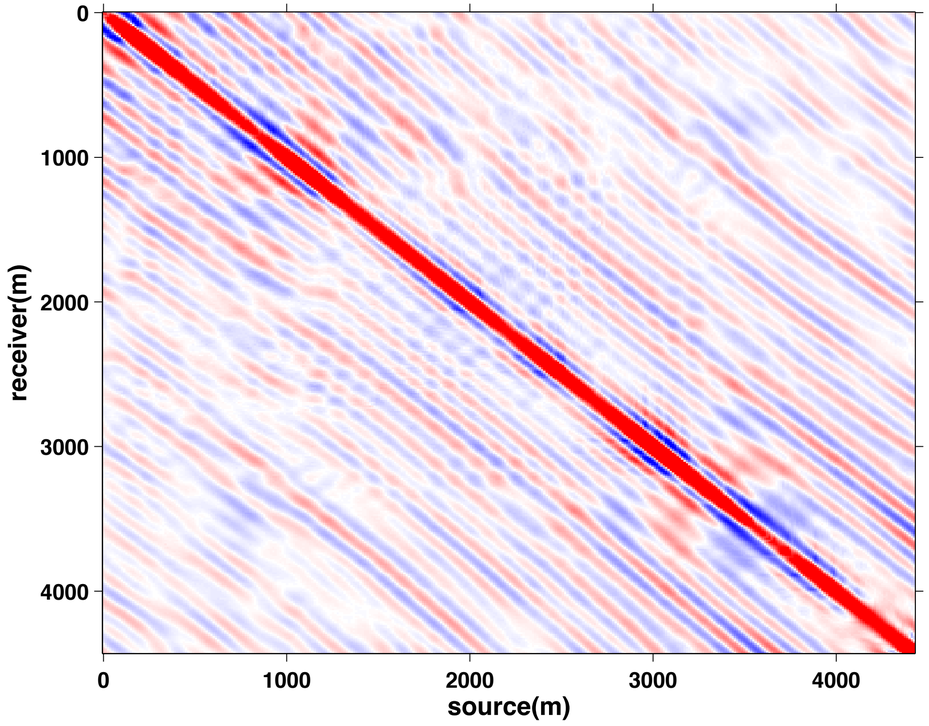
For each fixed ω, consider *frequency slices* $d_\omega(s, r)$

as a matrix

Challenge #1 – Low rank seismic data

Frequency slice of 2D field data as a function of source, receiver positions ("Green's function restricted to surface)

Not low-rank, due to strong diagonal energy, oscillating off-diagonal energy



Challenge #1 – Low-rank seismic data

Midpoint-offset coordinates:

$$\text{midpt} = \frac{s+r}{2}, \text{offset} = \frac{s-r}{2}$$

Non-separable transformation

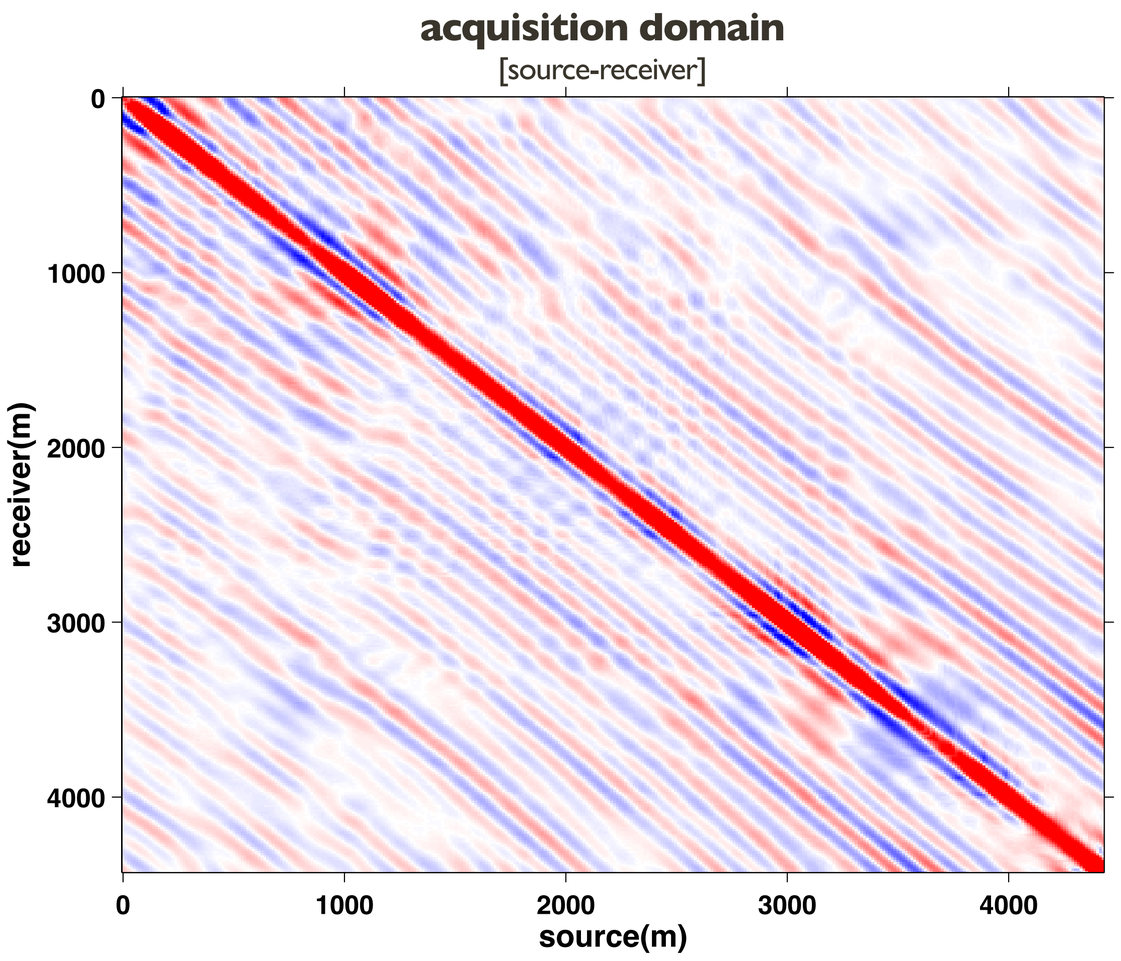
$$\mathcal{M} : \mathbb{R}^{n_s \times n_r} \rightarrow \mathbb{R}^{n_{midpt} \times n_{offset}}$$

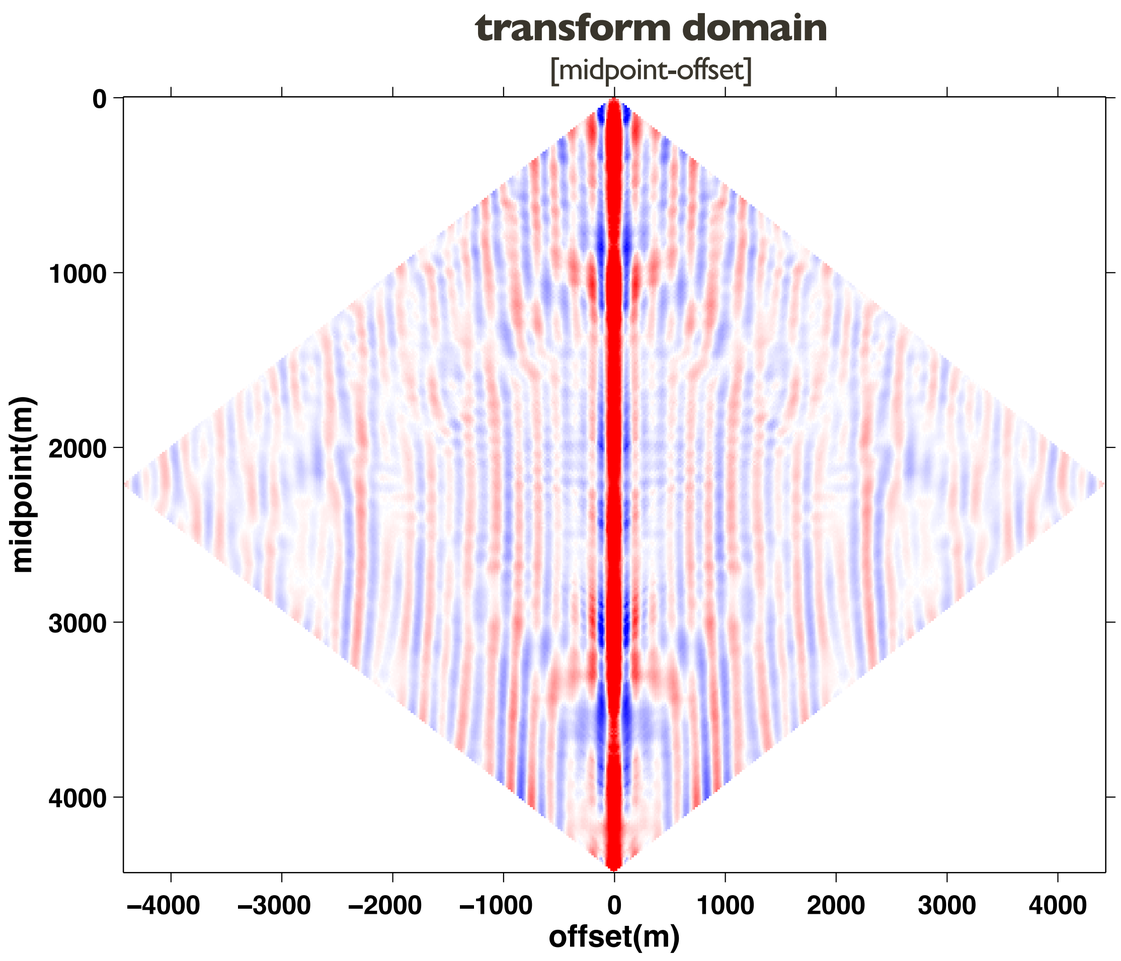
Tight frame

- linear operator, preserves 2-norm

Low-rank structure

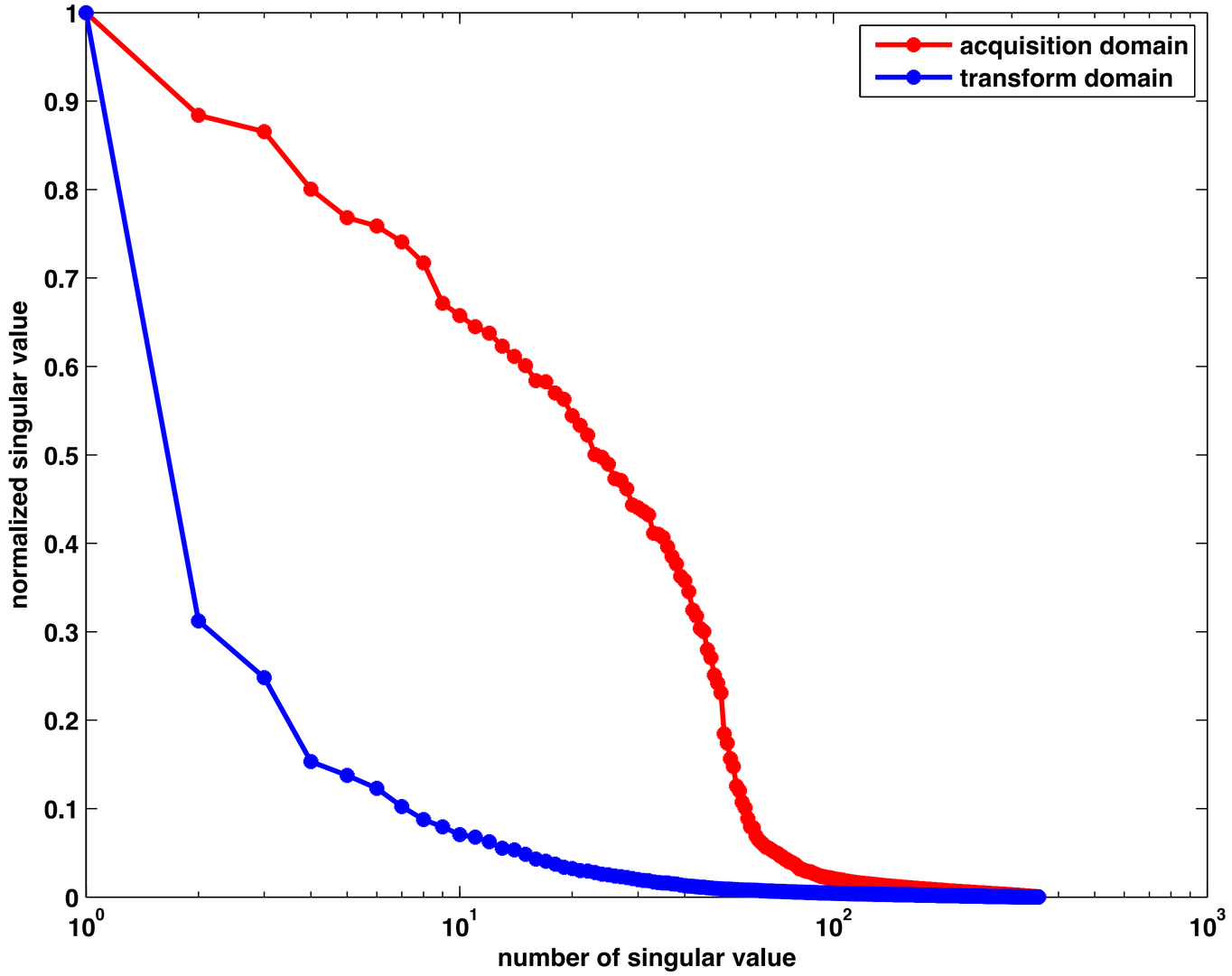
2-D acquisition, 10Hz



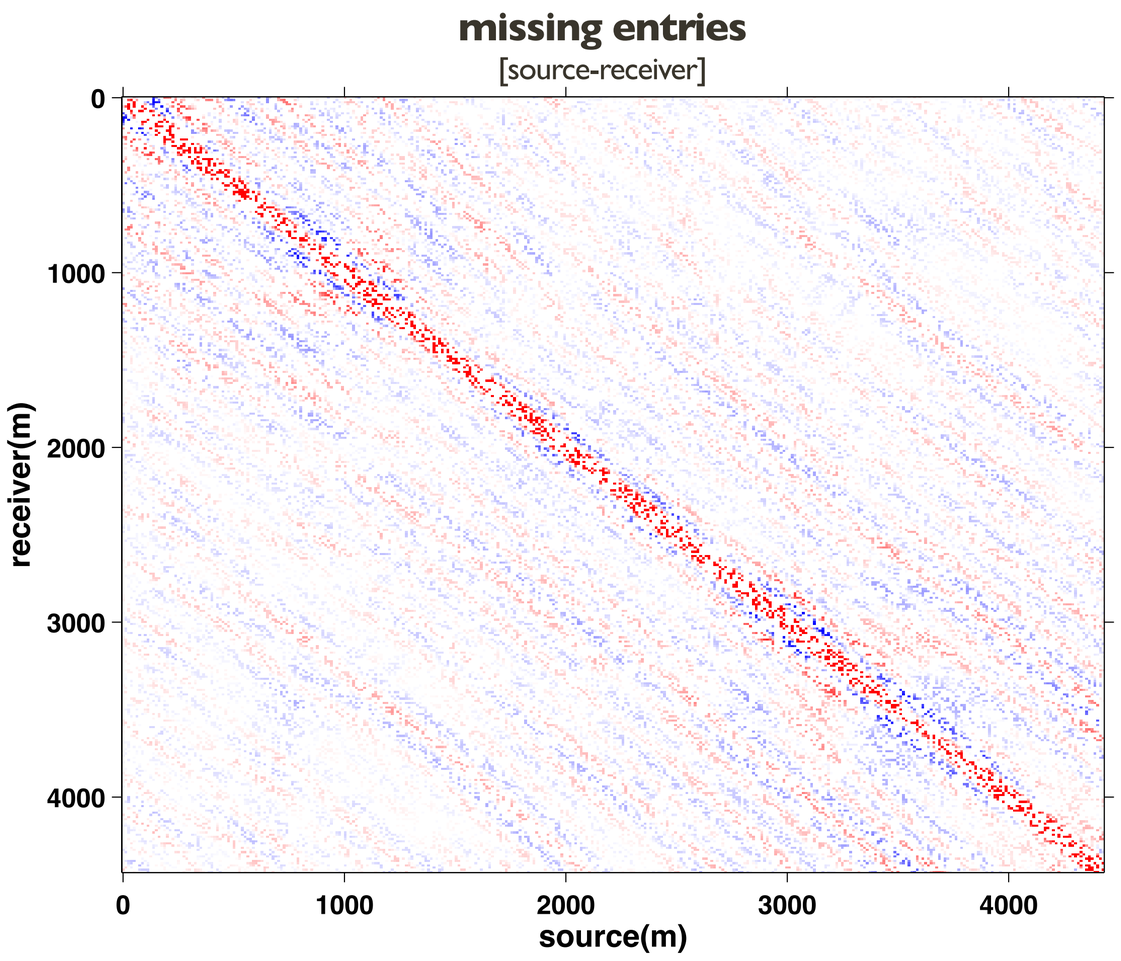


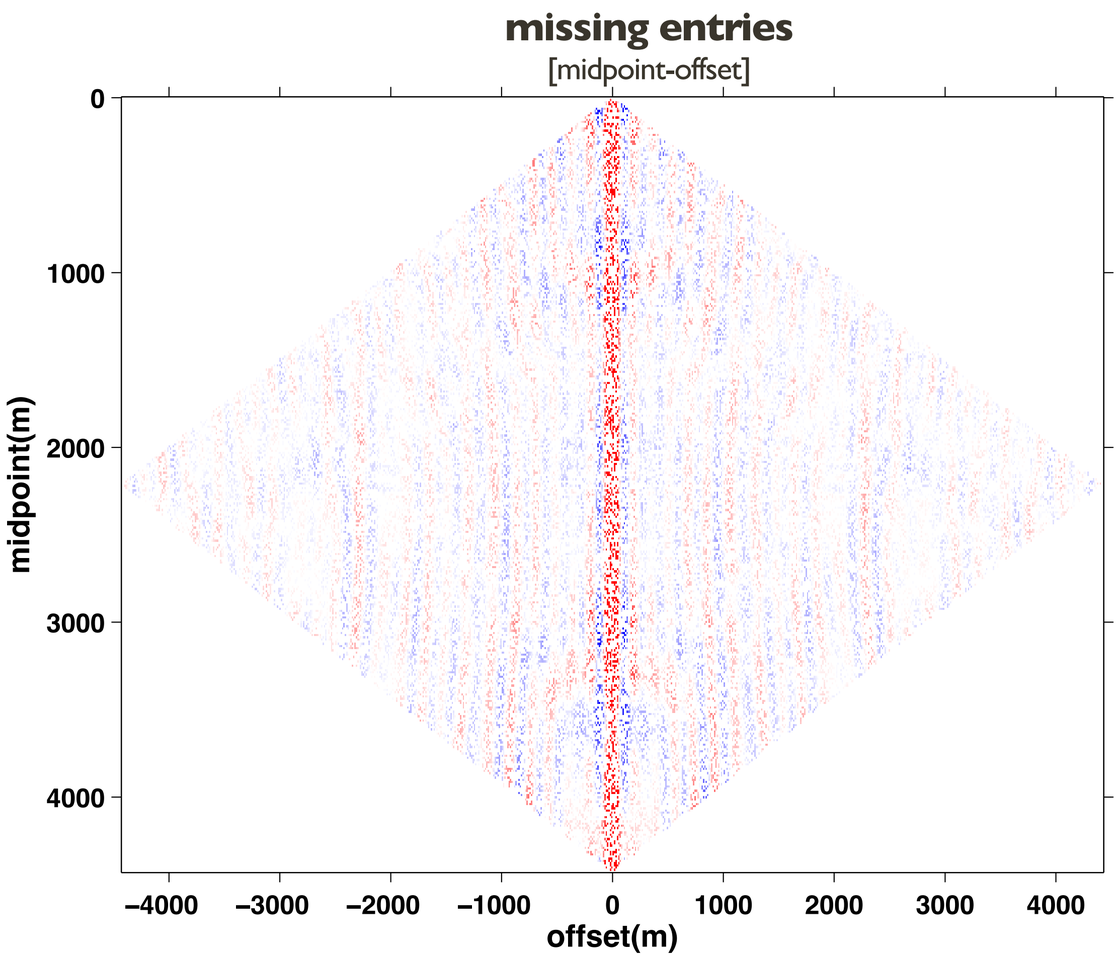
Singular value decay

2-D acquisition, 10Hz

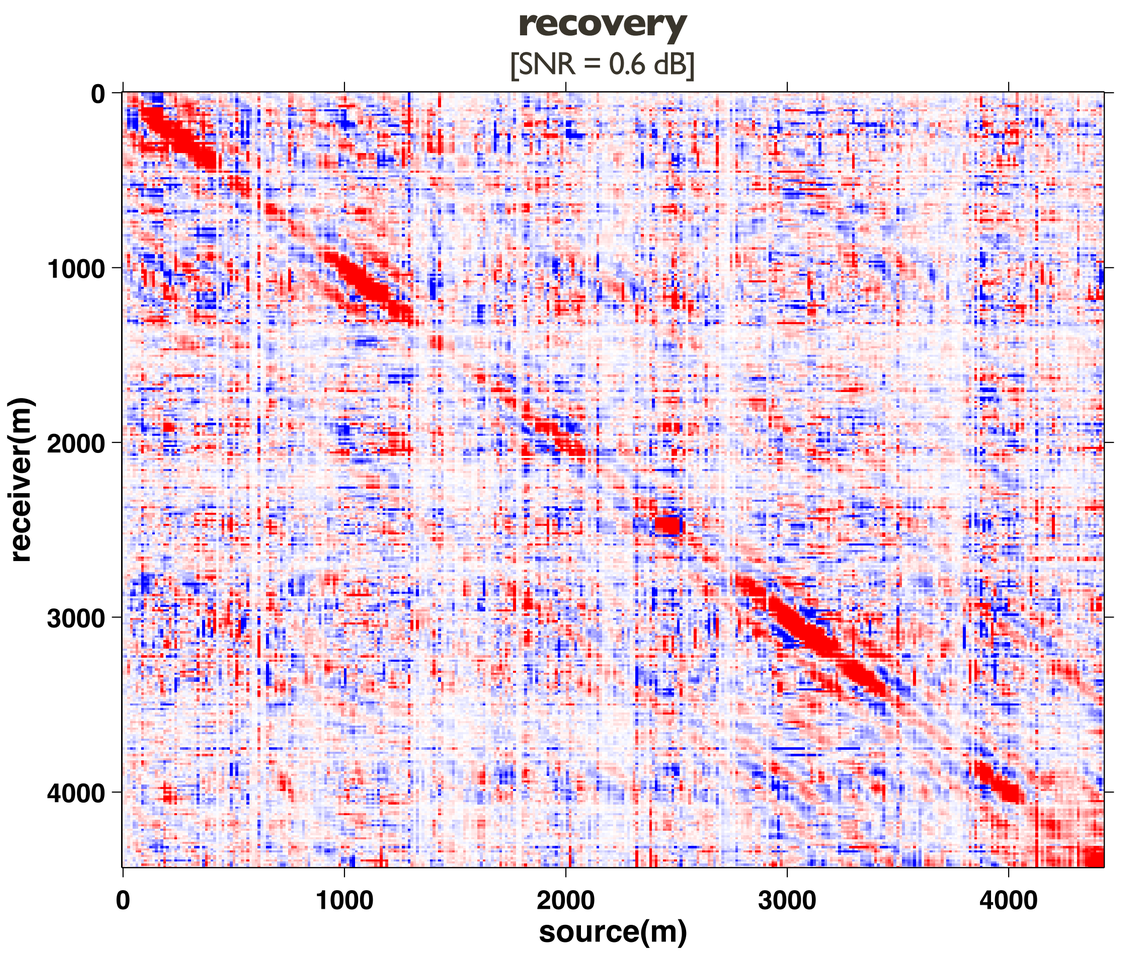


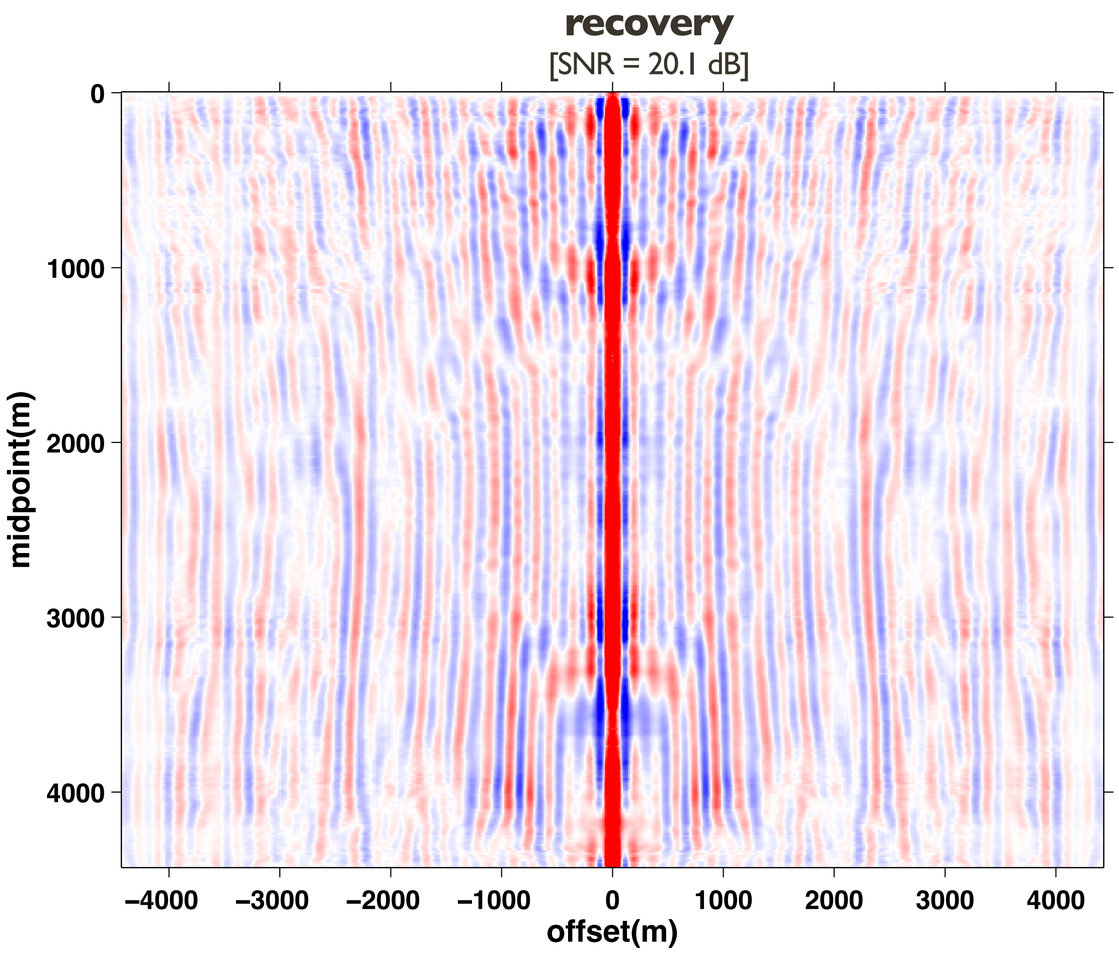
Matrix completion – missing entries



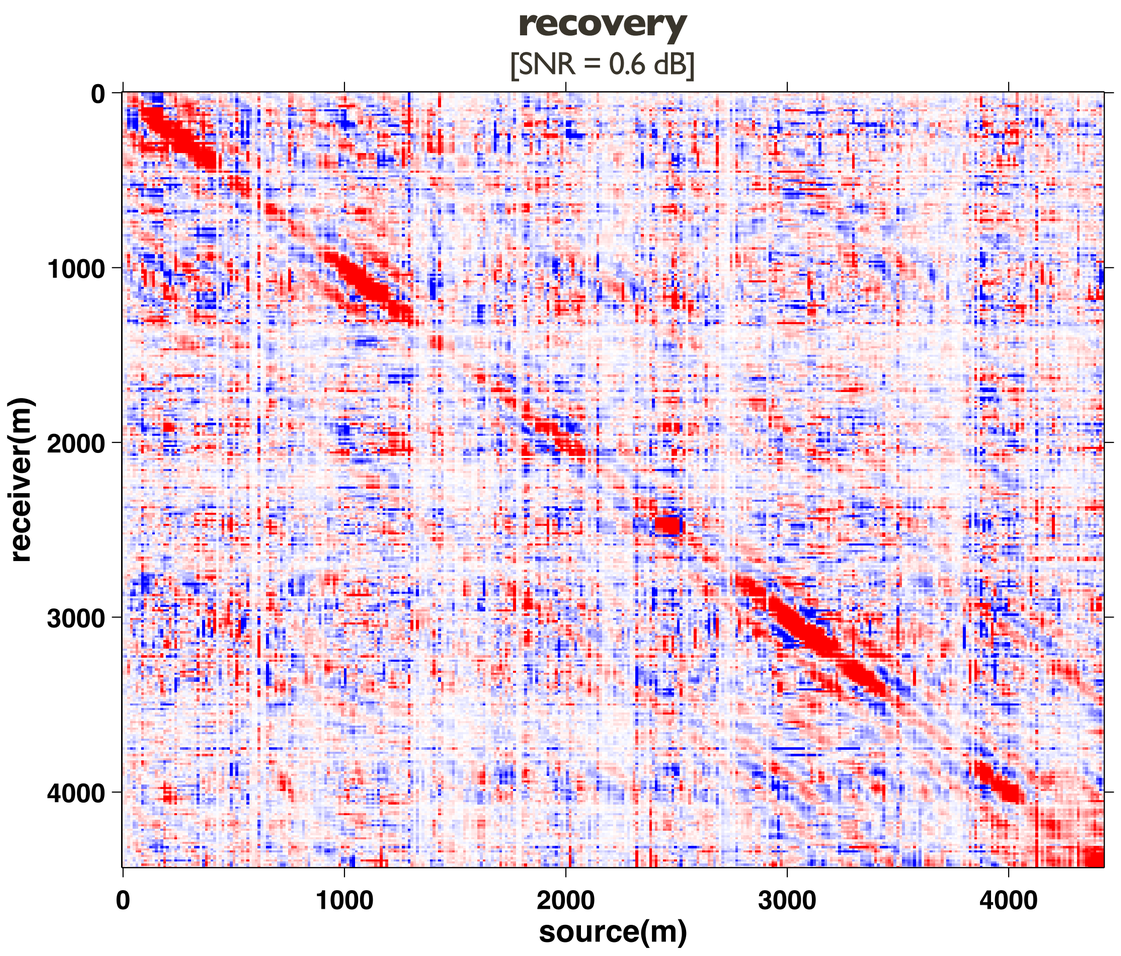


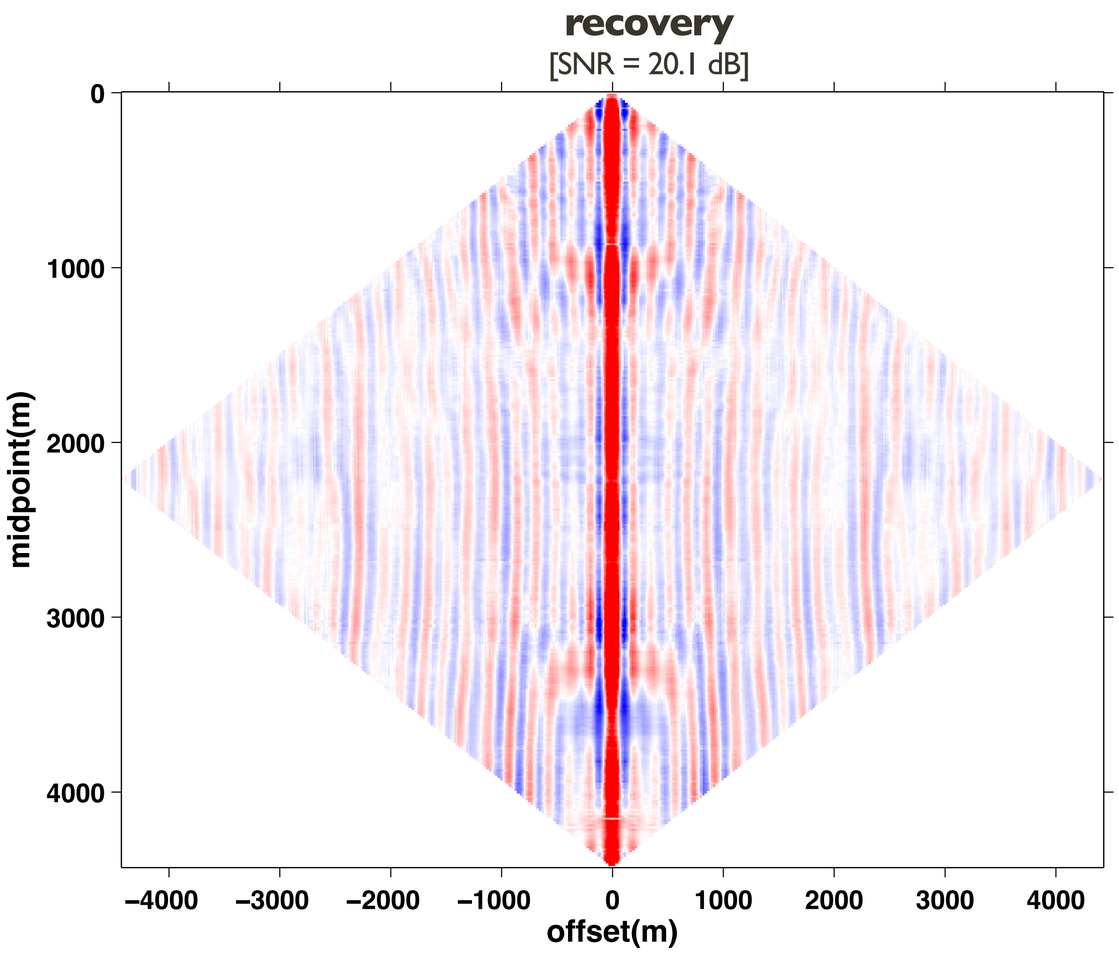
Low-rank interpolation



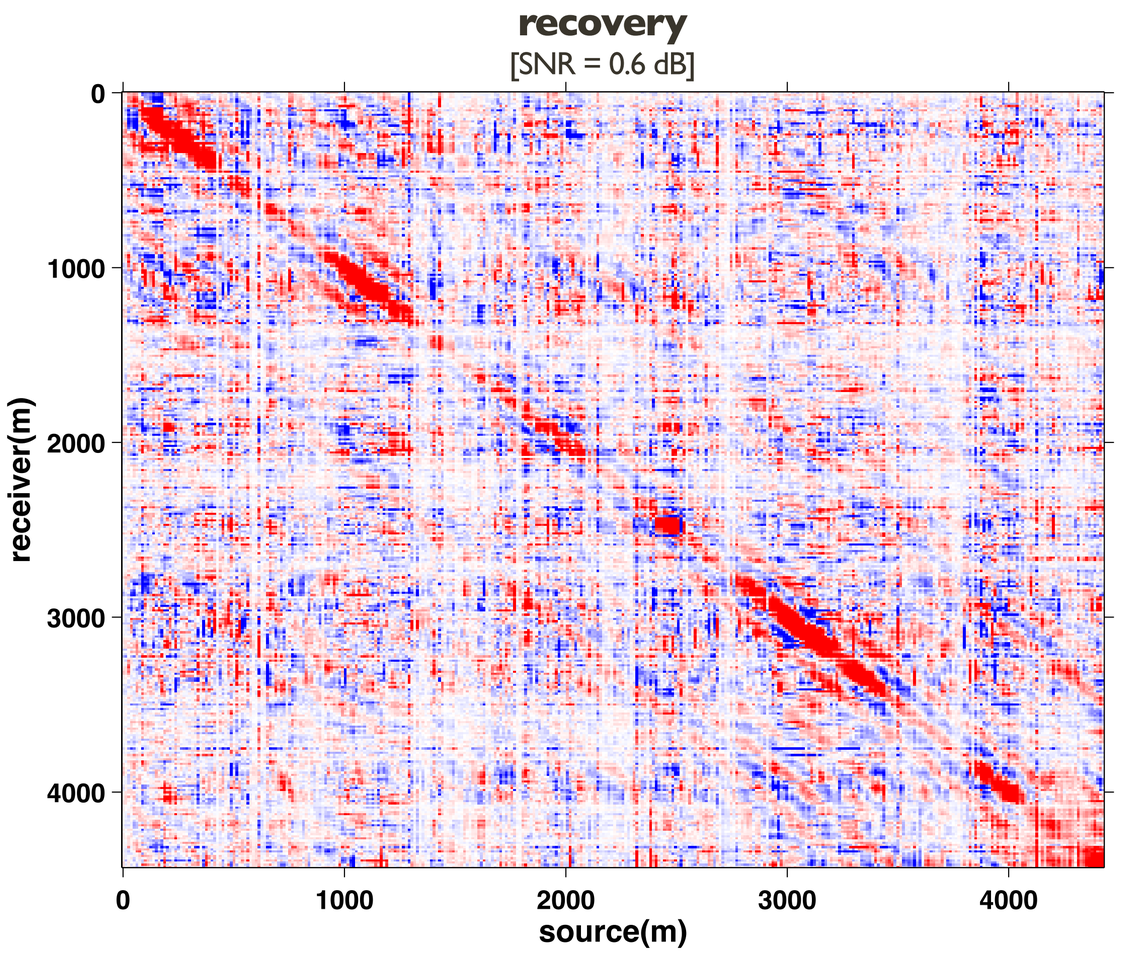


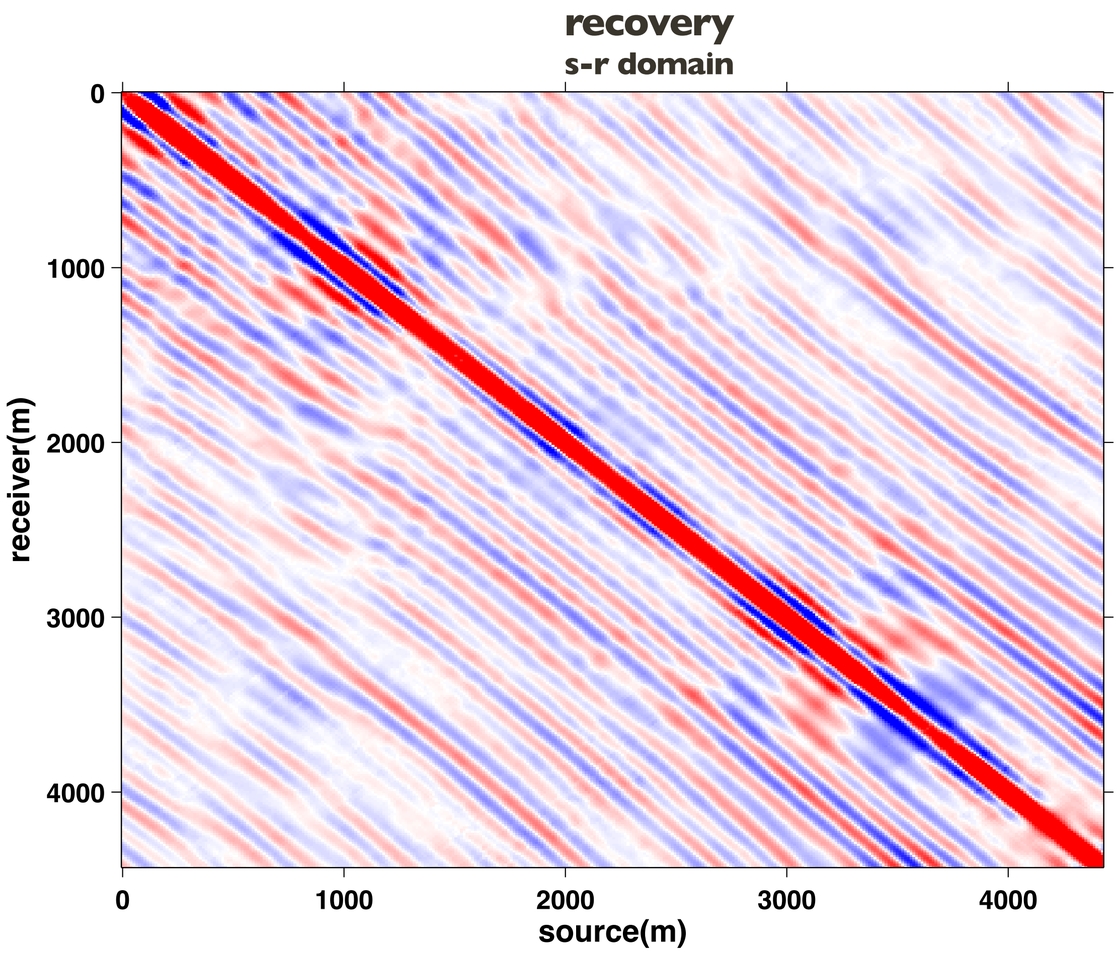
Low-rank interpolation





Low-rank interpolation





Challenge #1 – Low-rank seismic data

Ideal pointwise sampling corresponds to removing randomized (source, receiver) pairs from each frequency slice

- impractical to implement in a real survey

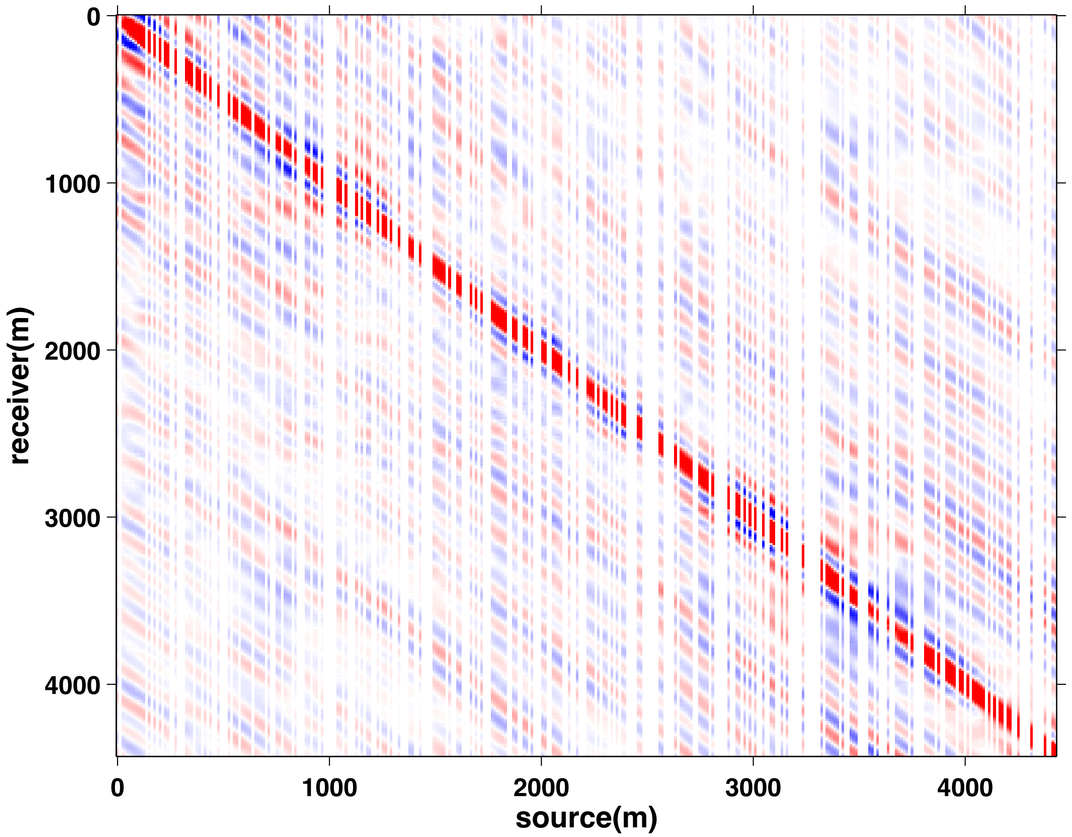
Possible to remove random sources, receivers

- corresponds to removing rows, columns

Source-receiver versus midpoint-offset recovery

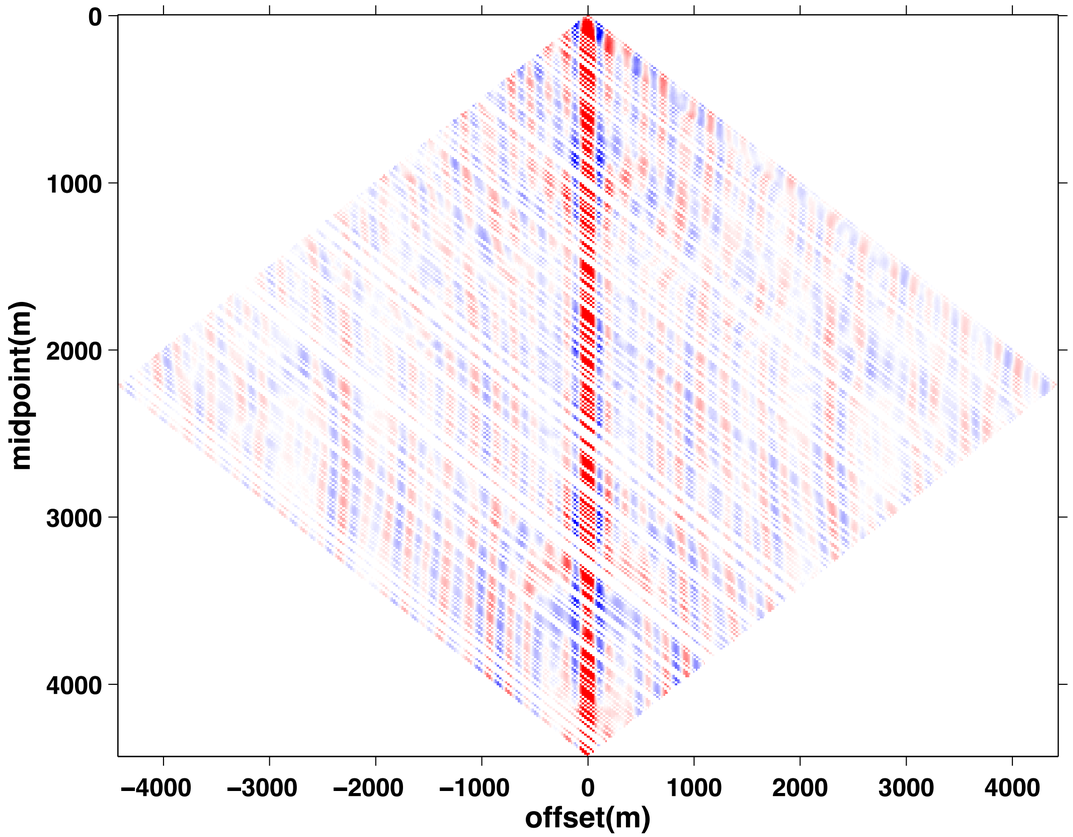
acquisition domain

missing columns *do not* increase rank



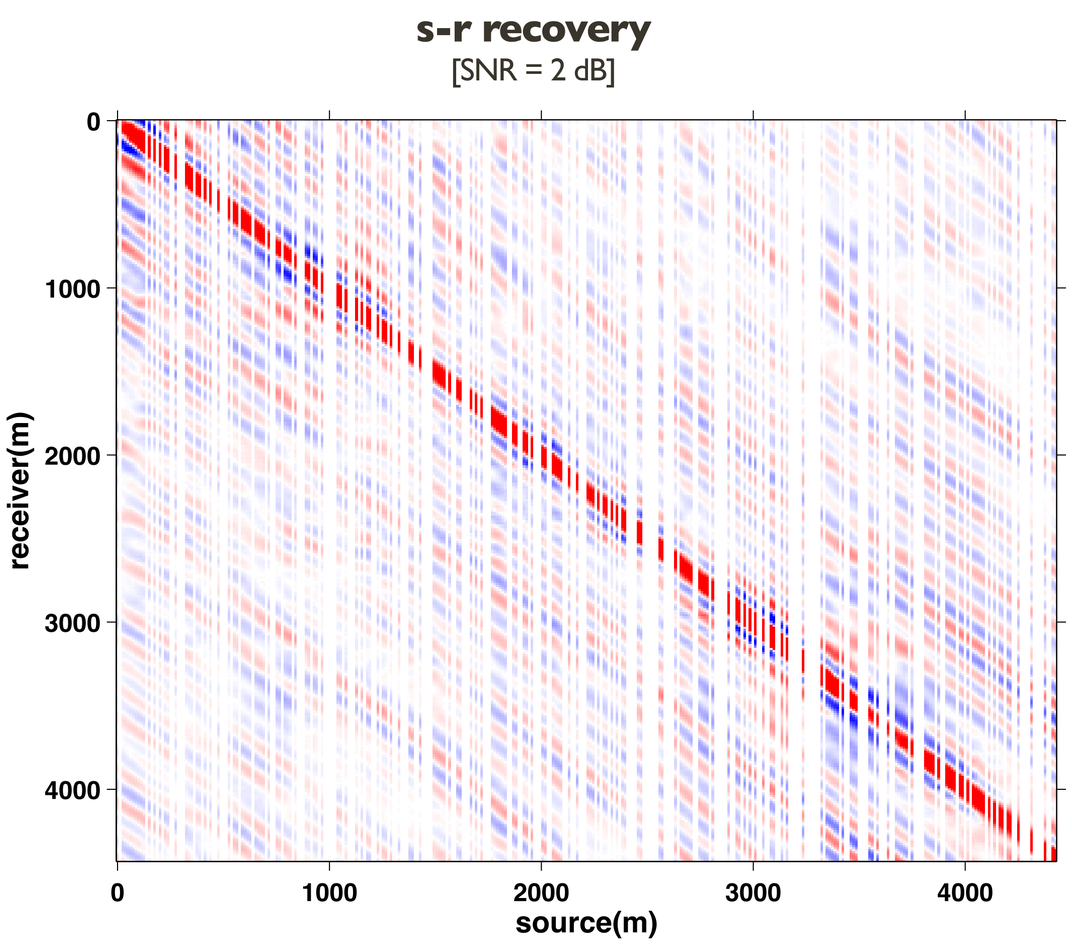
transform domain

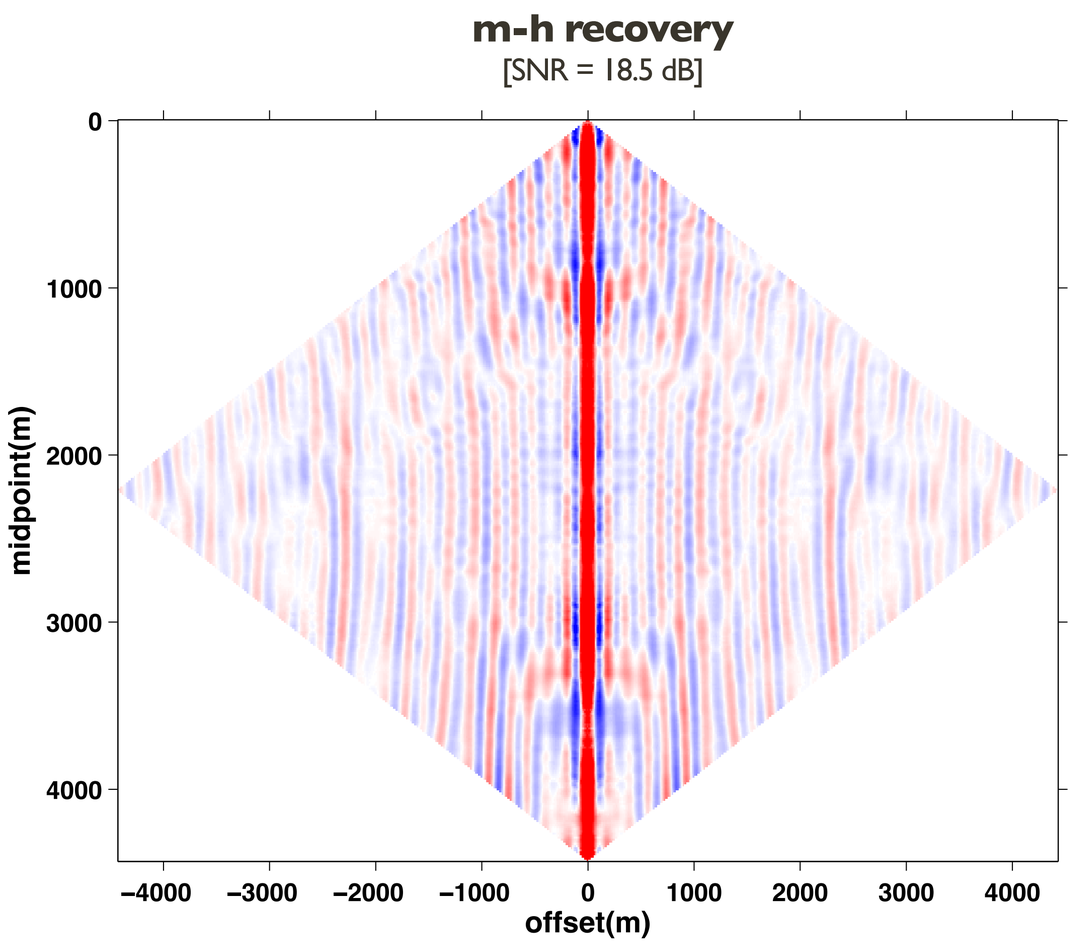
missing columns *increase* rank



Source-receiver versus midpoint-offset recovery

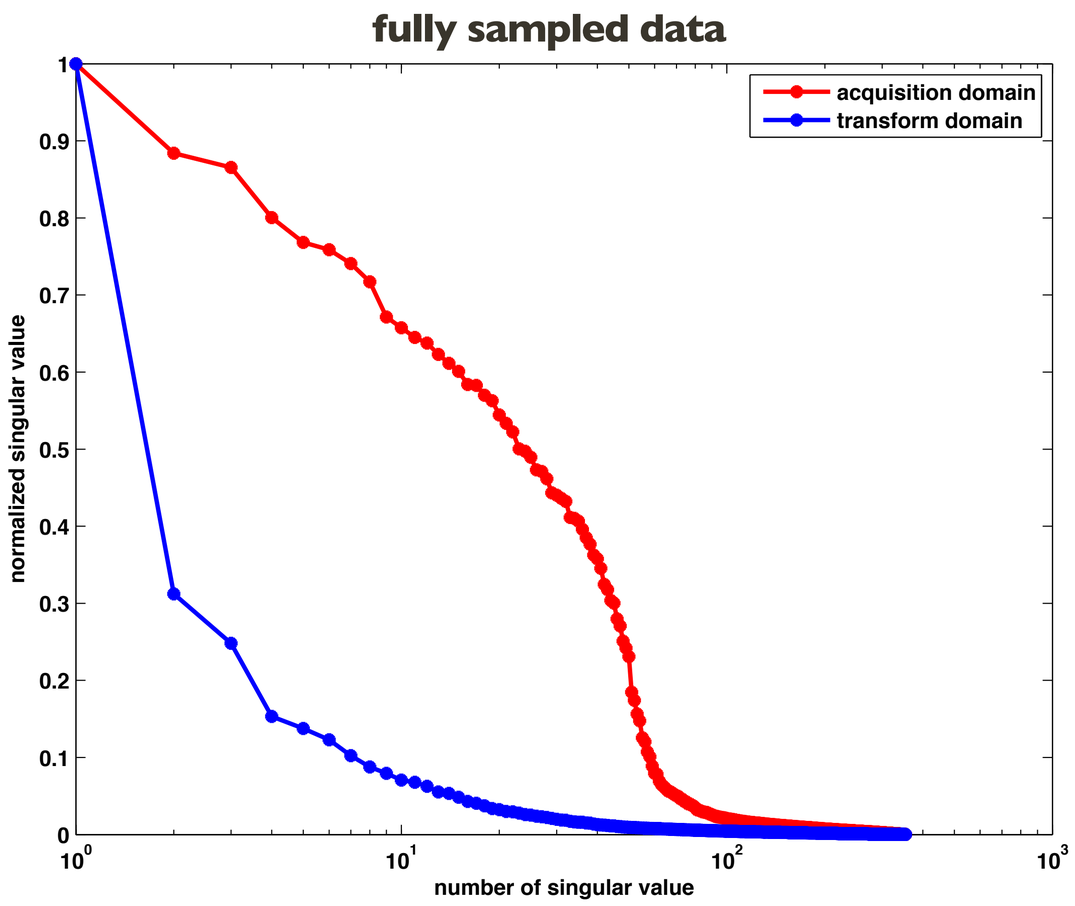
w/ mask

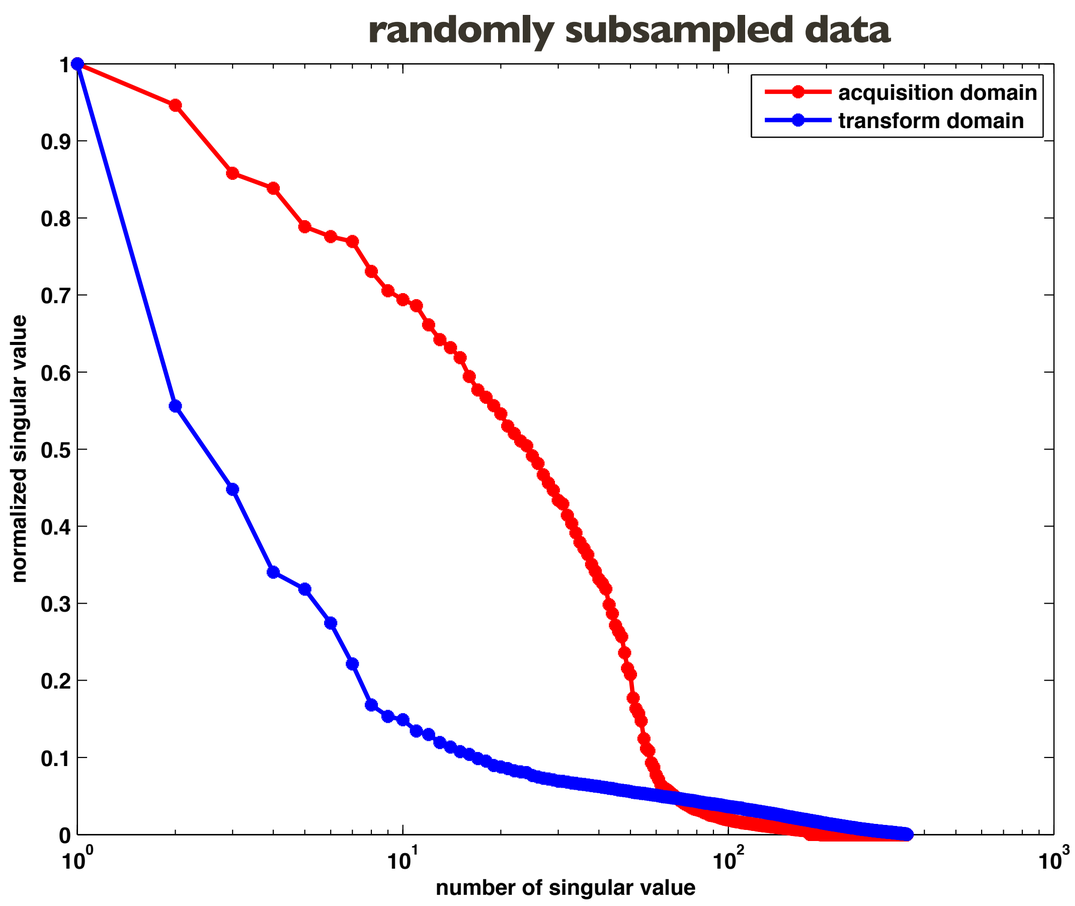




Randomized sampling

singular value decay





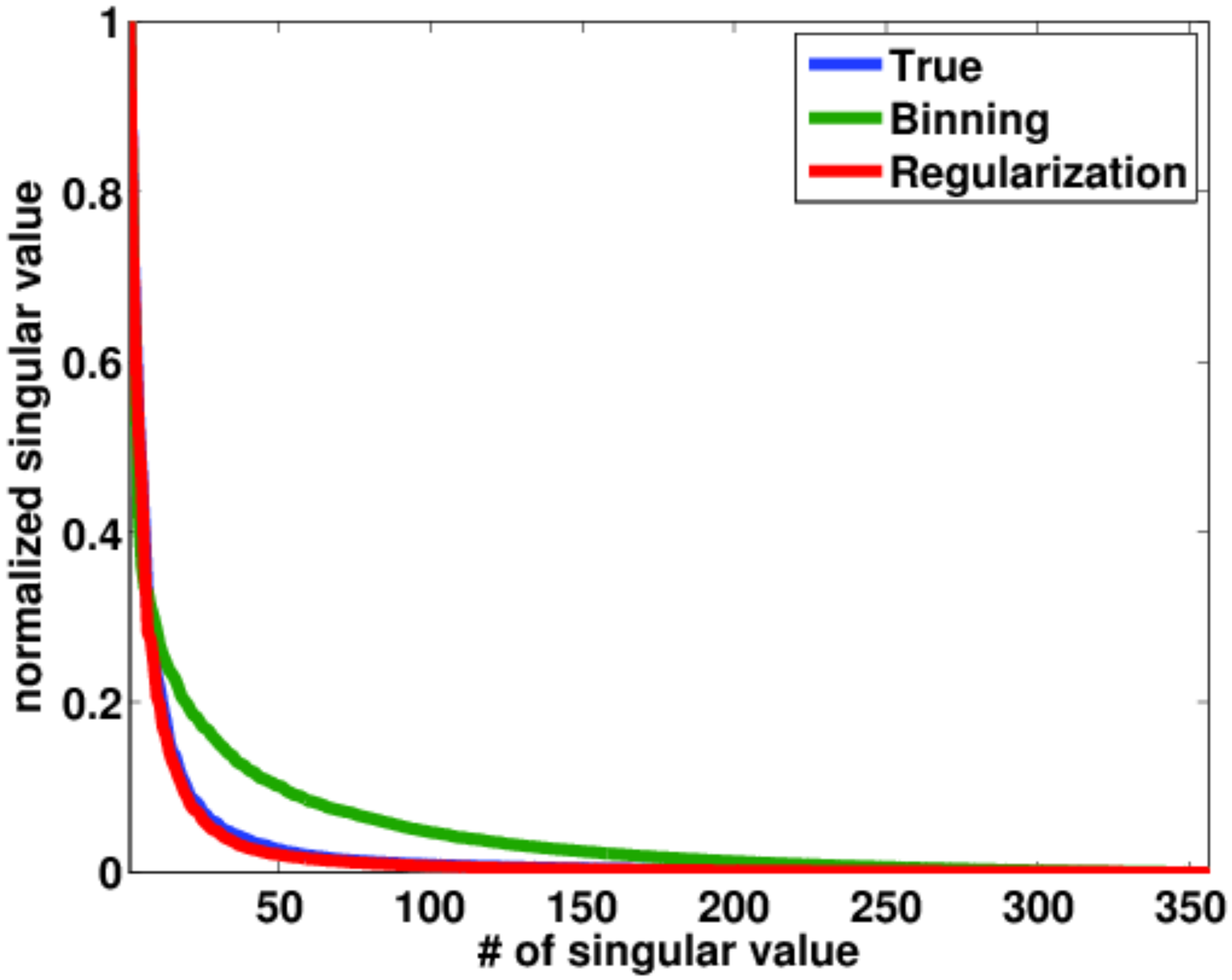
Challenge #2 – Sampling off the grid

Nearest neighbour interpolation (= binning) is inadequate

Introduce irregular sampling operator in optimization framework

Singular-value decay

regularization vs binning



Challenge #3 – Big data

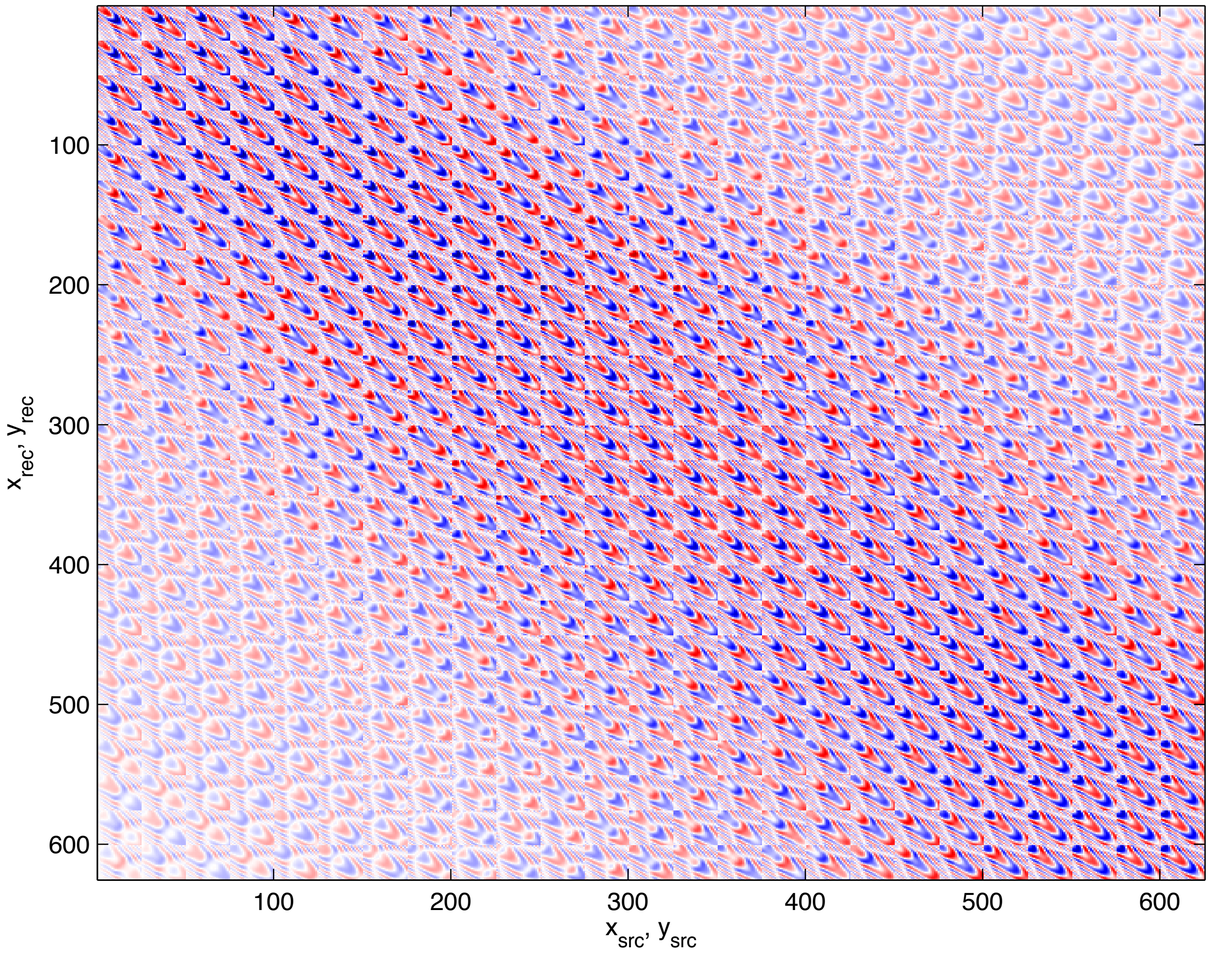
3D seismic data – 5D

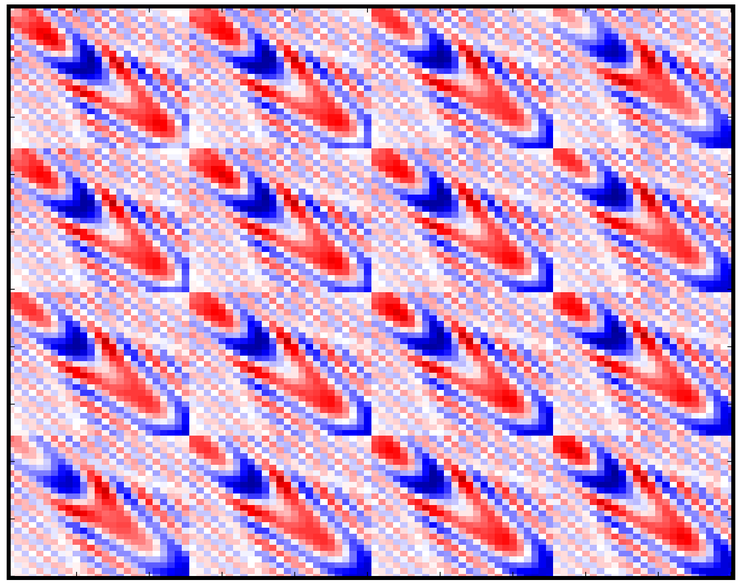
- source x, source y, receiver x, receiver y, time

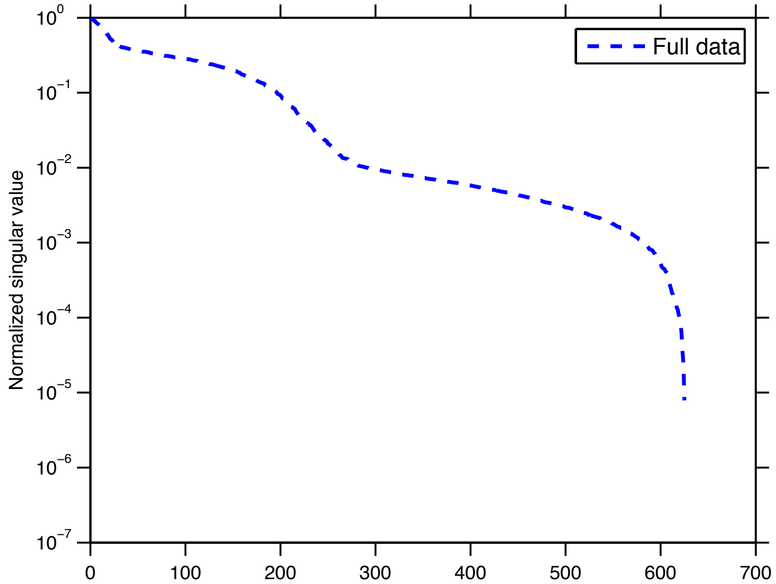
Consider frequency slices $d_\omega(x_{src}, y_{src}, x_{rec}, y_{rec})$

Different *groupings* of indices promote low-rank behaviour

Matricizations

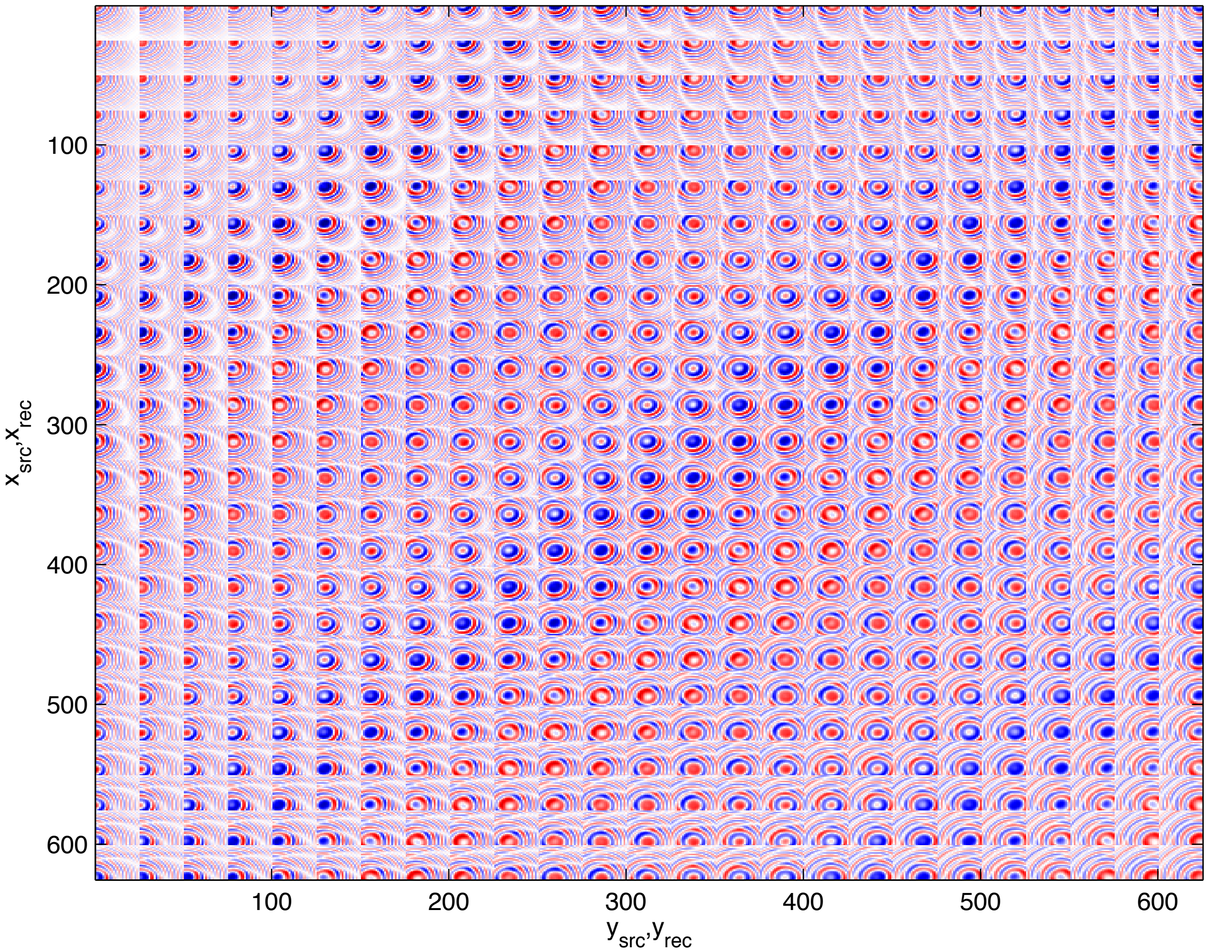




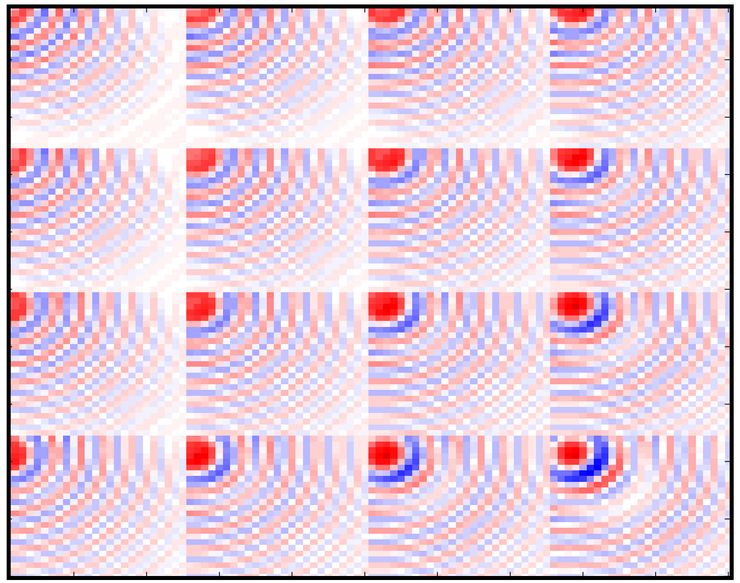


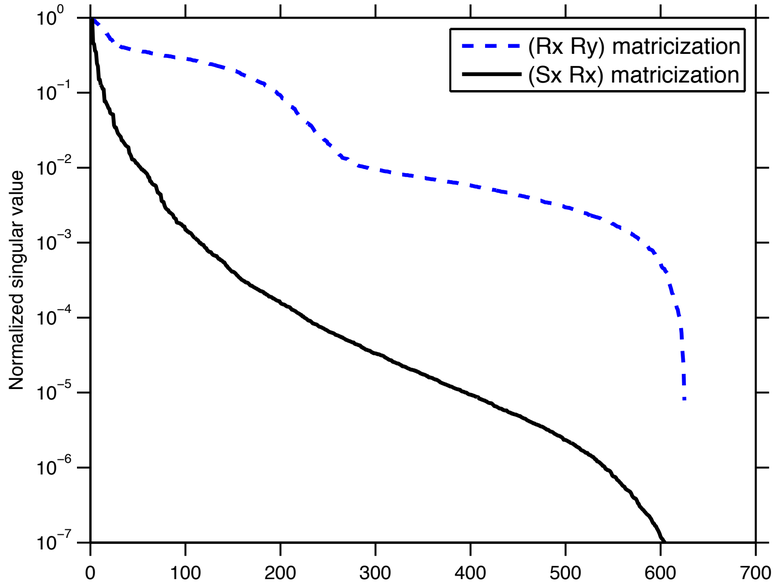
(Rec x, Rec y) matricization - Canonical ordering

Matricizations



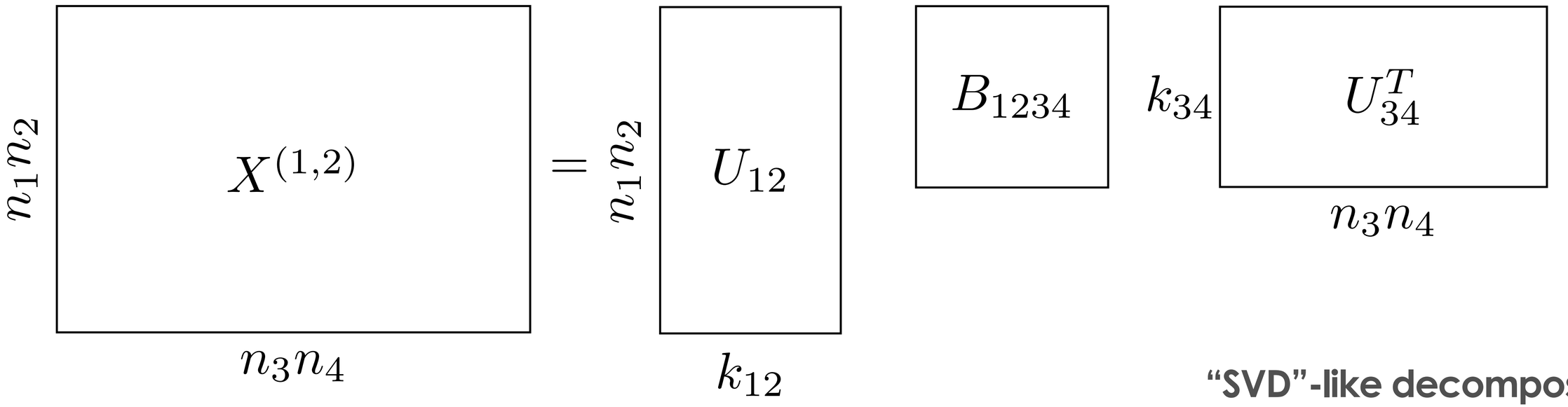
(Src x, Rec x) matricization - Noncanonical ordering





Hierarchical Tucker format

$X - n_1 \times n_2 \times n_3 \times n_4$ tensor

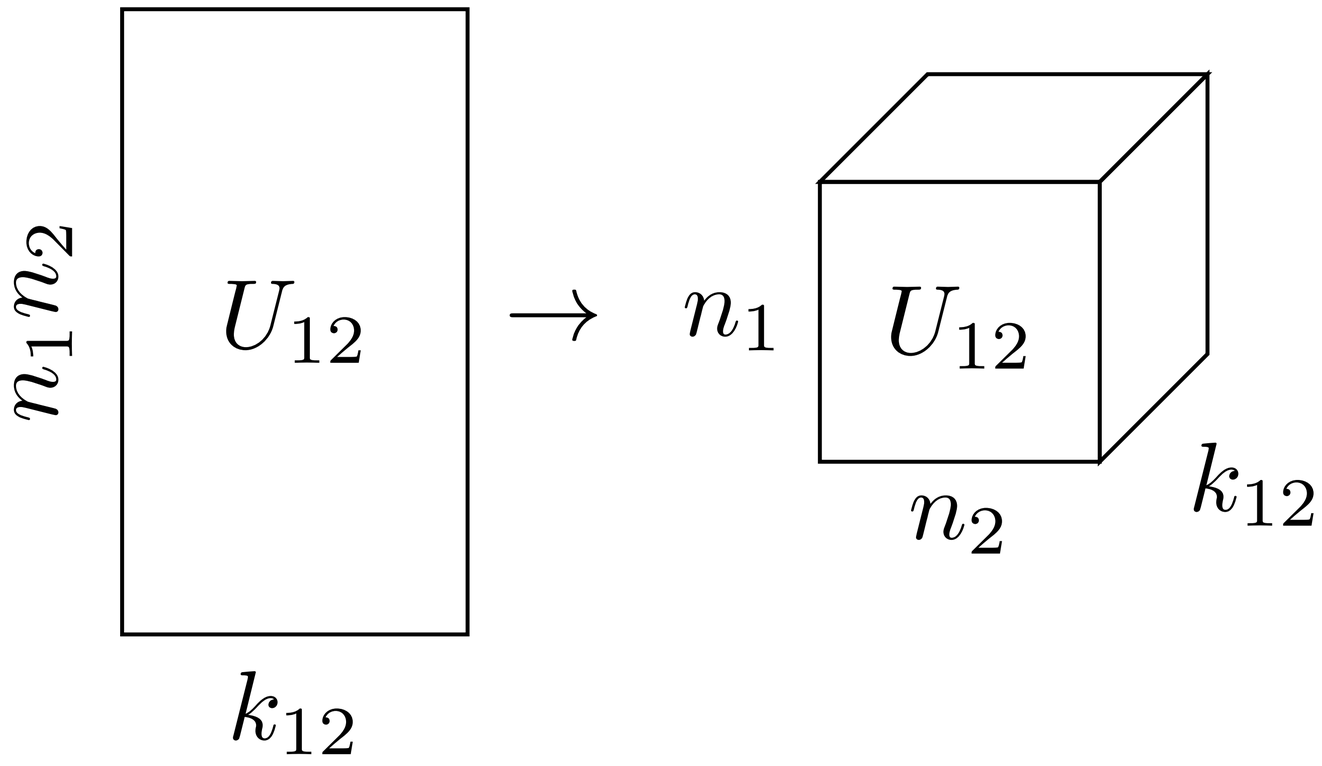


"SVD"-like decomposition

"Hierarchical singular value decomposition of tensors" L. Grasedyck

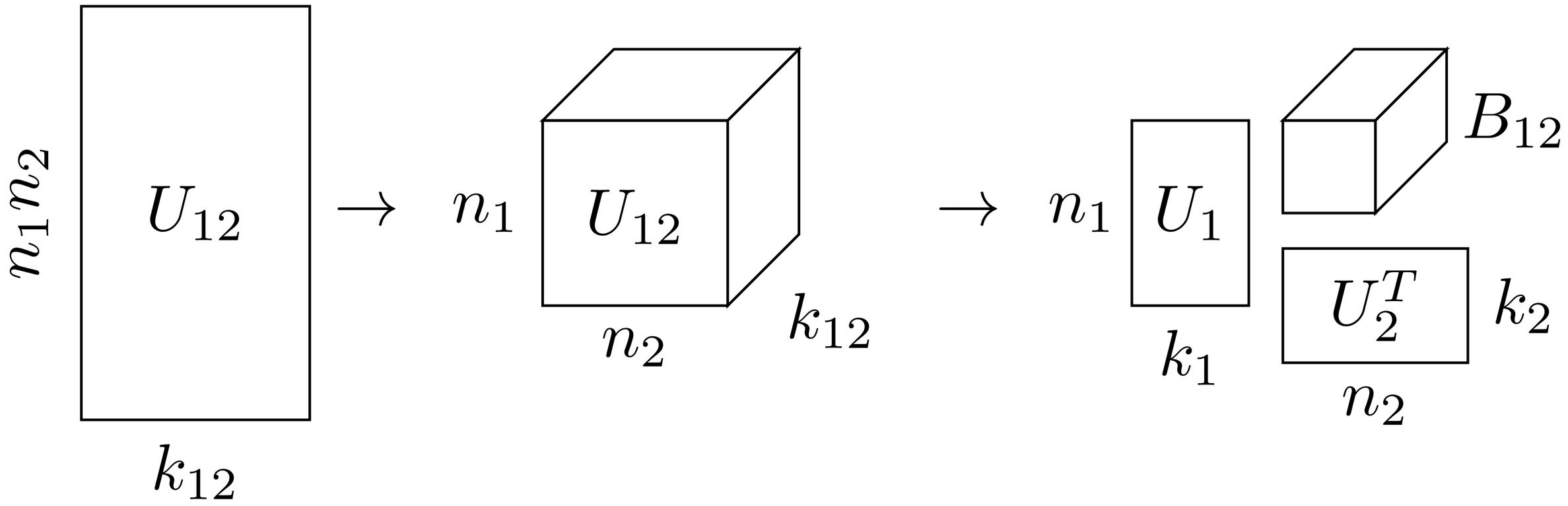
Hierarchical Tucker format

$X - n_1 \times n_2 \times n_3 \times n_4$ tensor



Hierarchical Tucker format

$X - n_1 \times n_2 \times n_3 \times n_4$ tensor



Hierarchical Tucker format

Intermediate matrices don't need to be stored

U_t, B_t - small parameter matrices

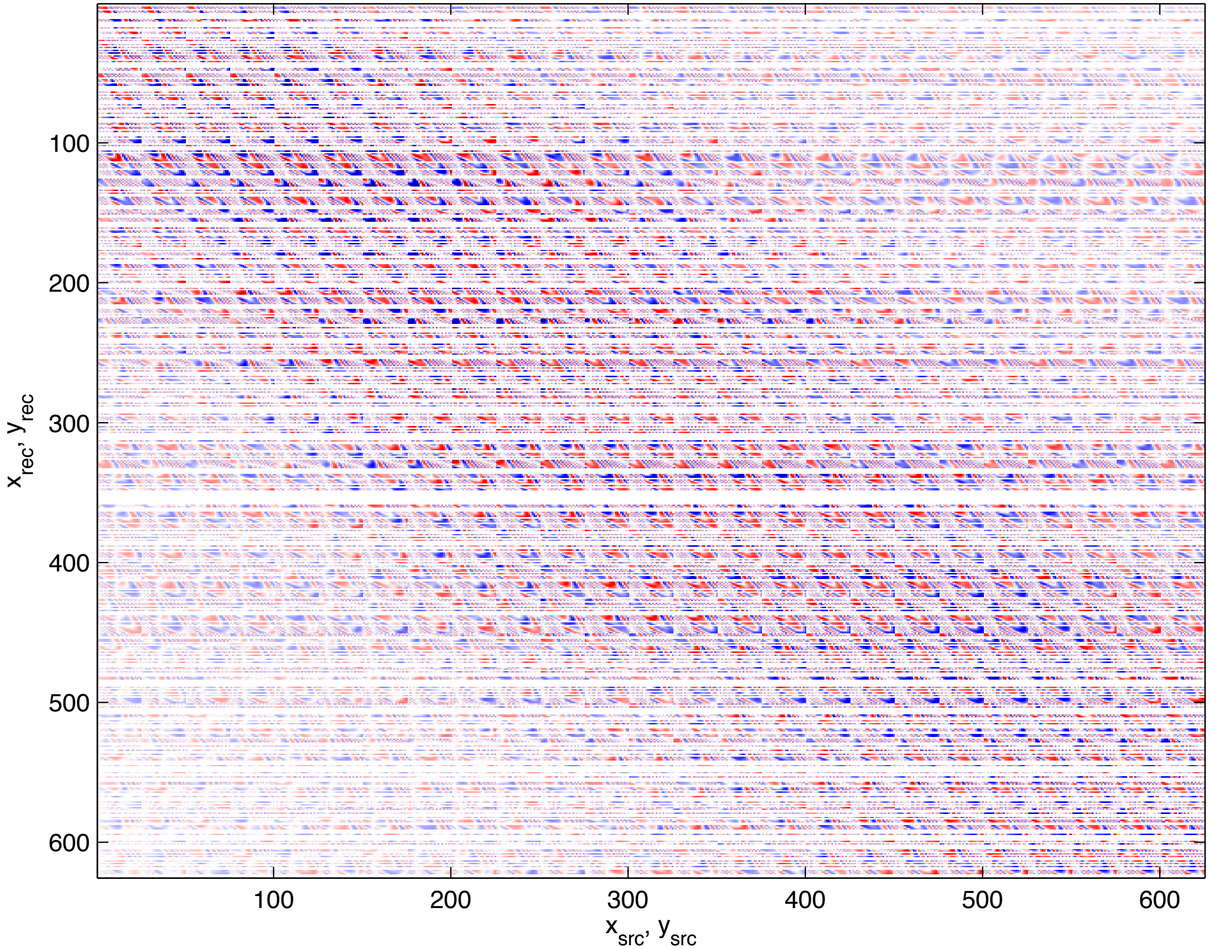
- specify the tensor completely

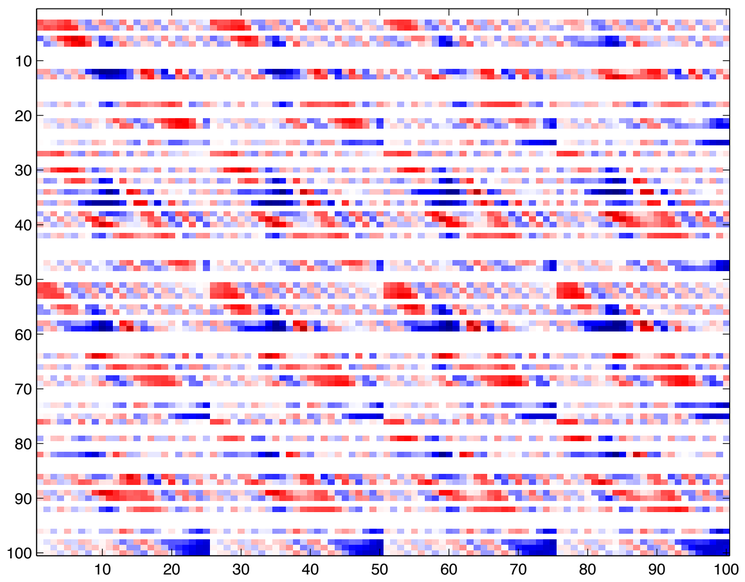
Separating groups of dimensions from each other

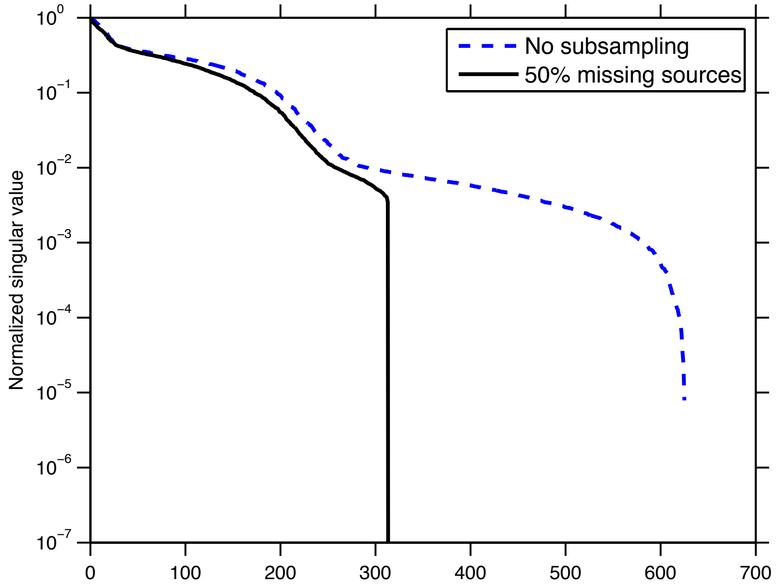
- dimension tree

Realistic sampling

50% random receivers removed



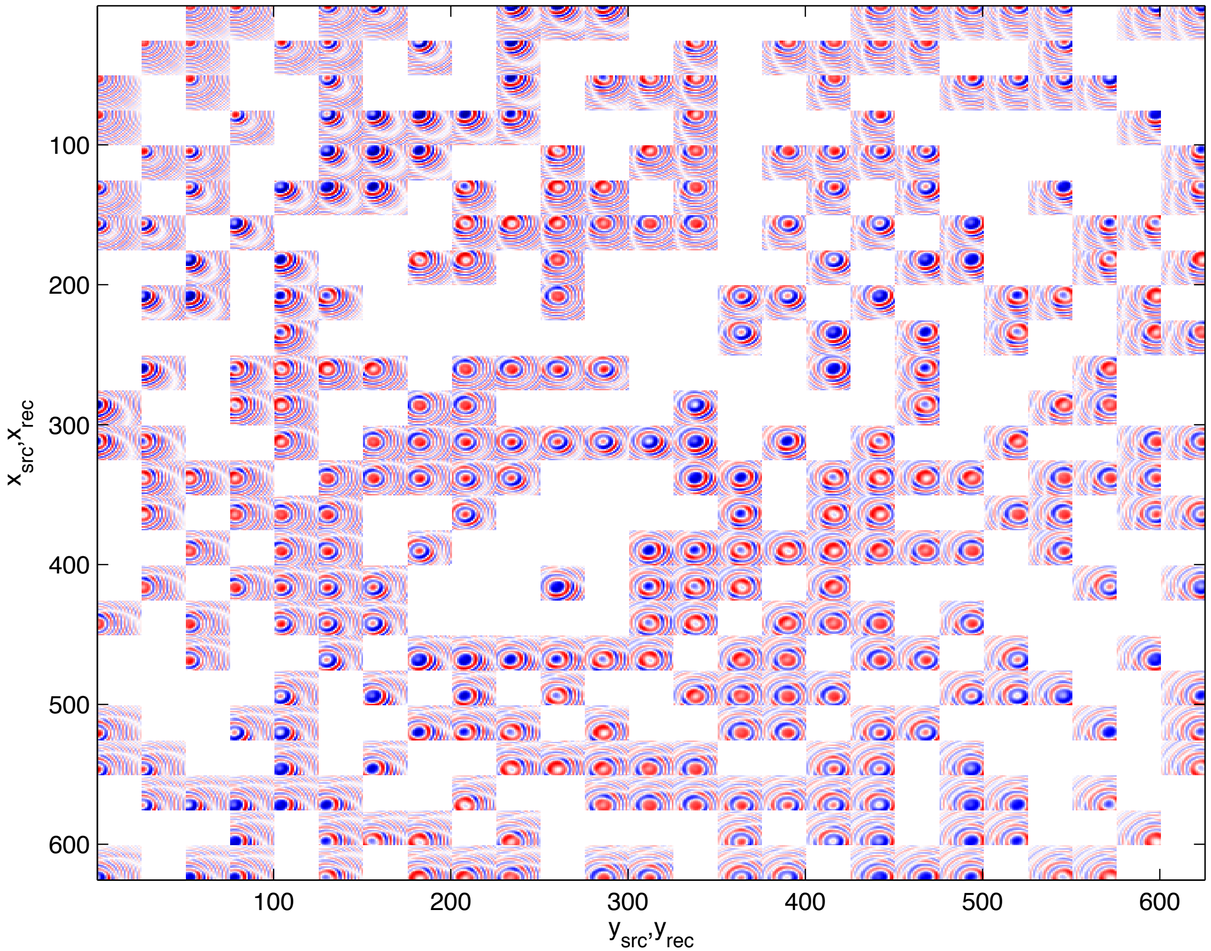


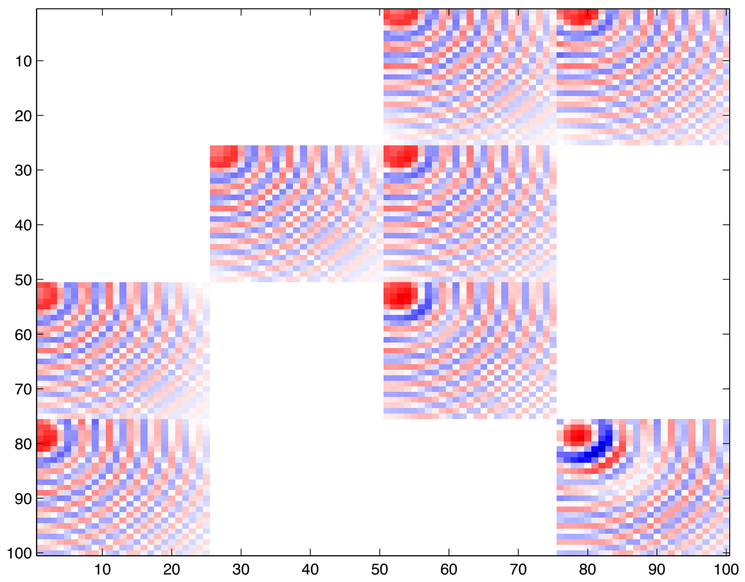


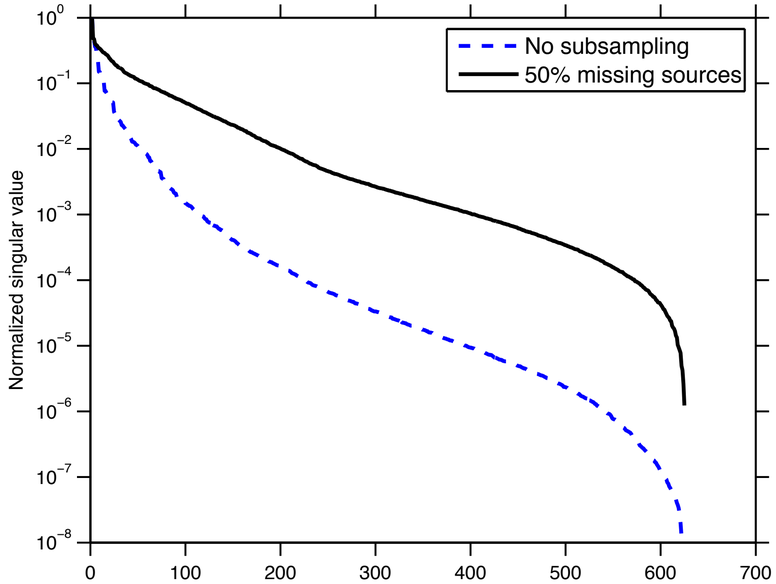
(Rec x, Rec y) matricization - Canonical ordering

Realistic sampling

50% random receivers removed







(Src x, Rec x) matricization - Noncanonical ordering

Challenge #3 – Big data

3D seismic data in Hierarchical Tucker format

- Parametrize a low-dimensional manifold in high-dimensional space
- Exploit this manifold structure for completion
- Minimize computations in high-dimensional space

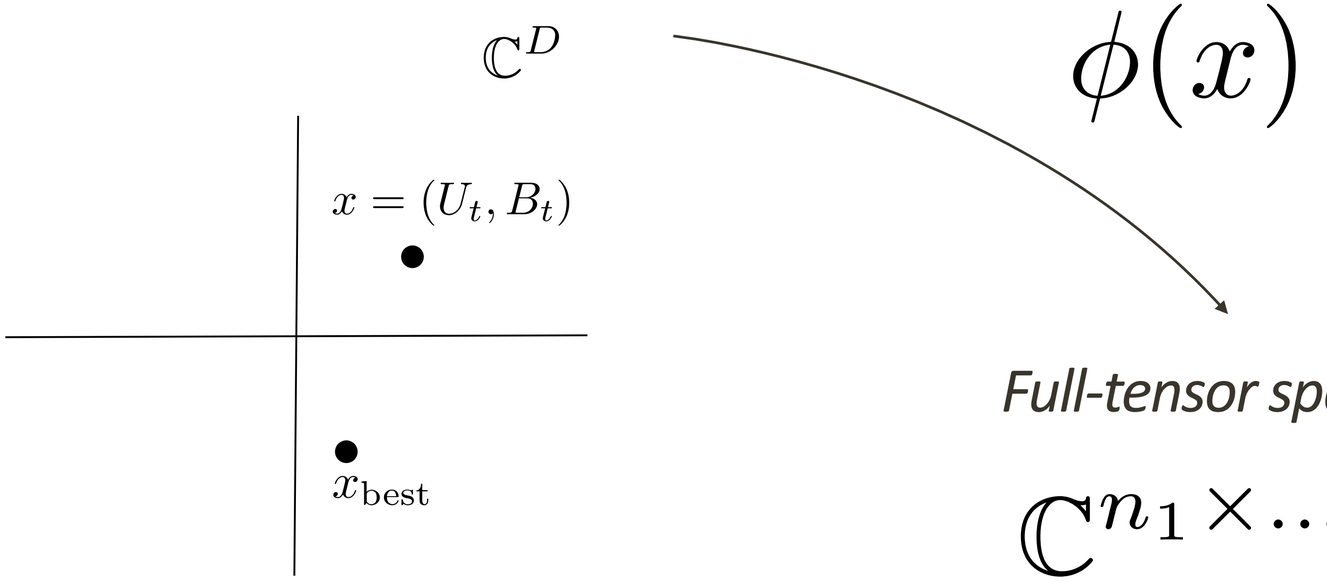
"The geometry of algorithms using hierarchical tensors", A Uschmajew, B Vandereycken
"Optimization on the Hierarchical Tucker manifold - applications to tensor completion", C. Da Silva and F. J. Herrmann

Challenge #3 – Big data

Hierarchical Tucker tensors parametrize a *smooth submanifold* of d-dimensional tensors

- equip with a Riemannian metric to implement smooth optimization algorithms
- avoid ambiguities in the parameters via respecting the quotient geometry of the format

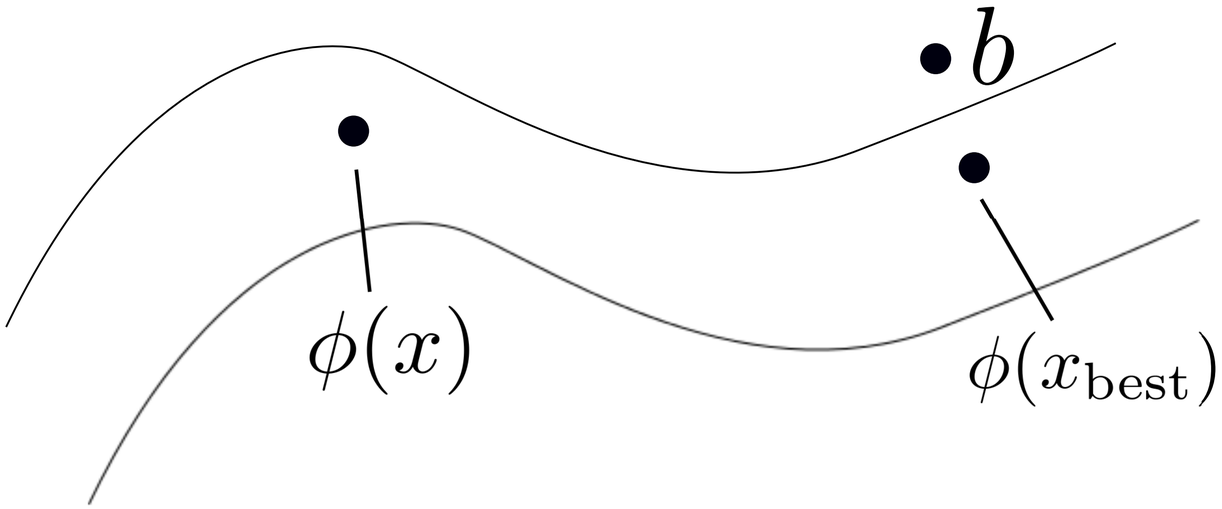
Optimization program



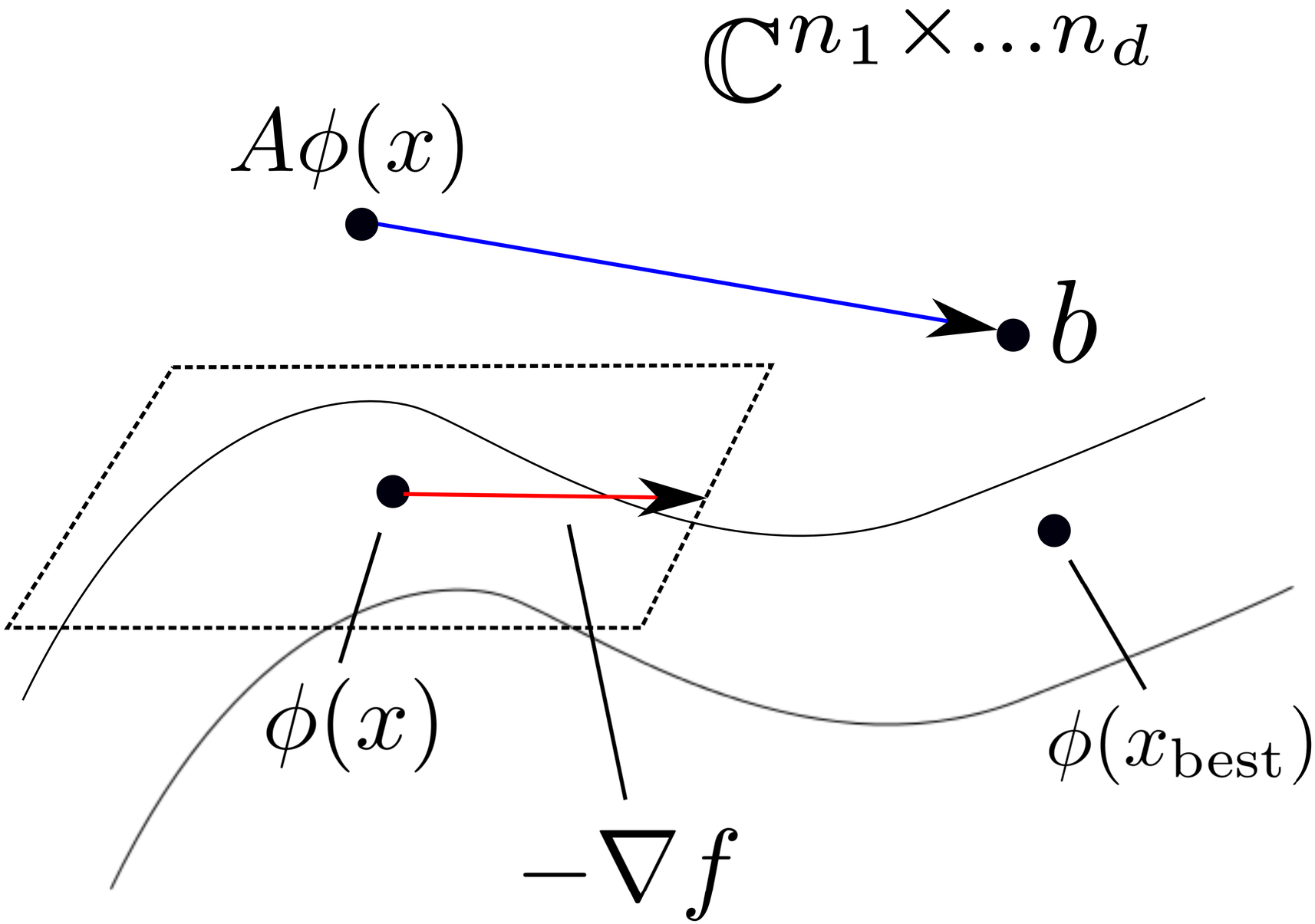
Full-tensor space

$$\mathbb{C}^{n_1 \times \ldots n_d}$$

Parameter space



Optimization program



Derivatives

Only involves matrix-matrix multiplications of small matrices compared to the full-tensor space

Parallelizable – multilinear product can be done in parallel

SVD-free – no large-scale SVDs, unlike nuclear norm-based methods

Seismic matrix/tensor completion examples

Synthetic, Gulf of Mexico

LR factorization

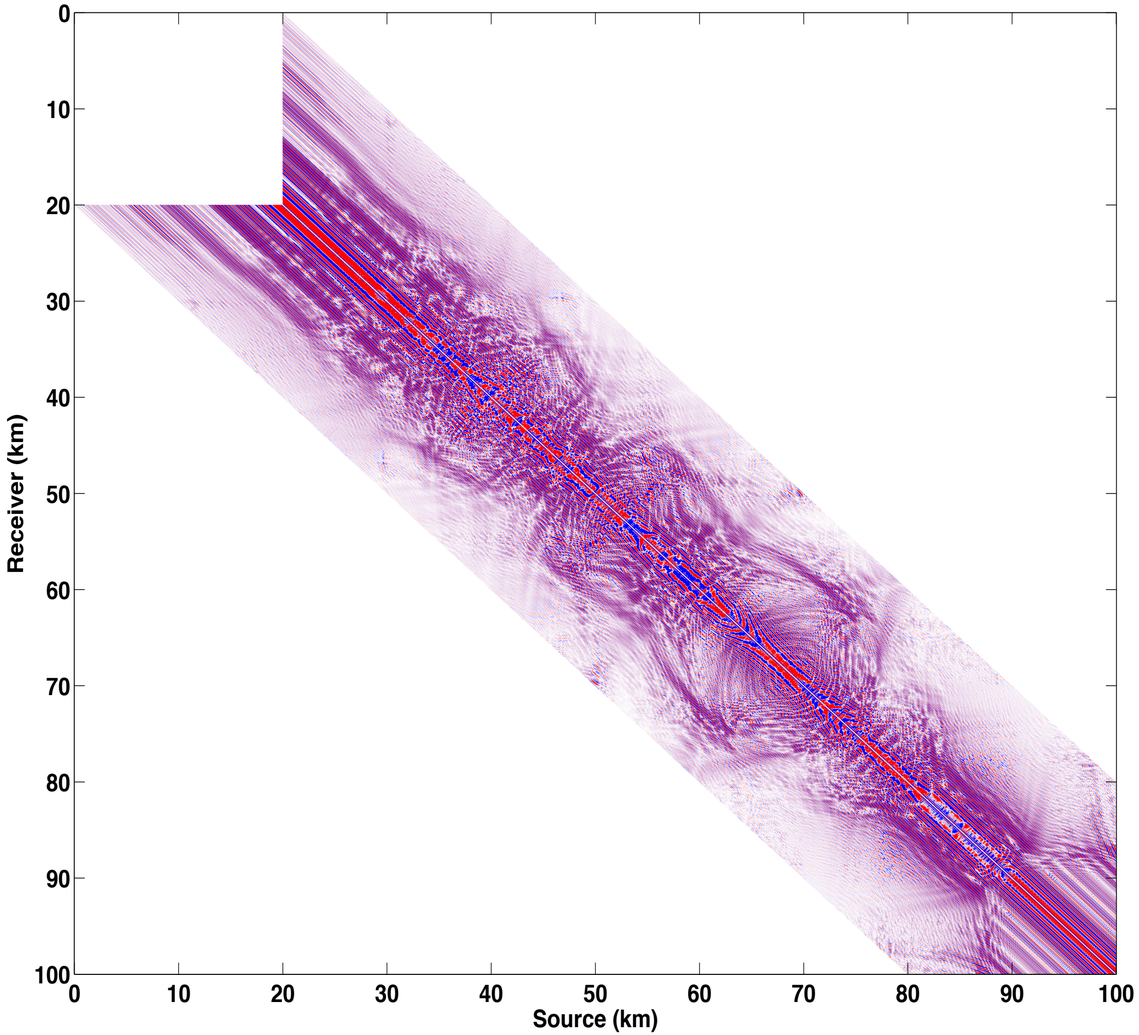
High-rank model

- 3201 sources with 801 receivers, data at 7 Hz

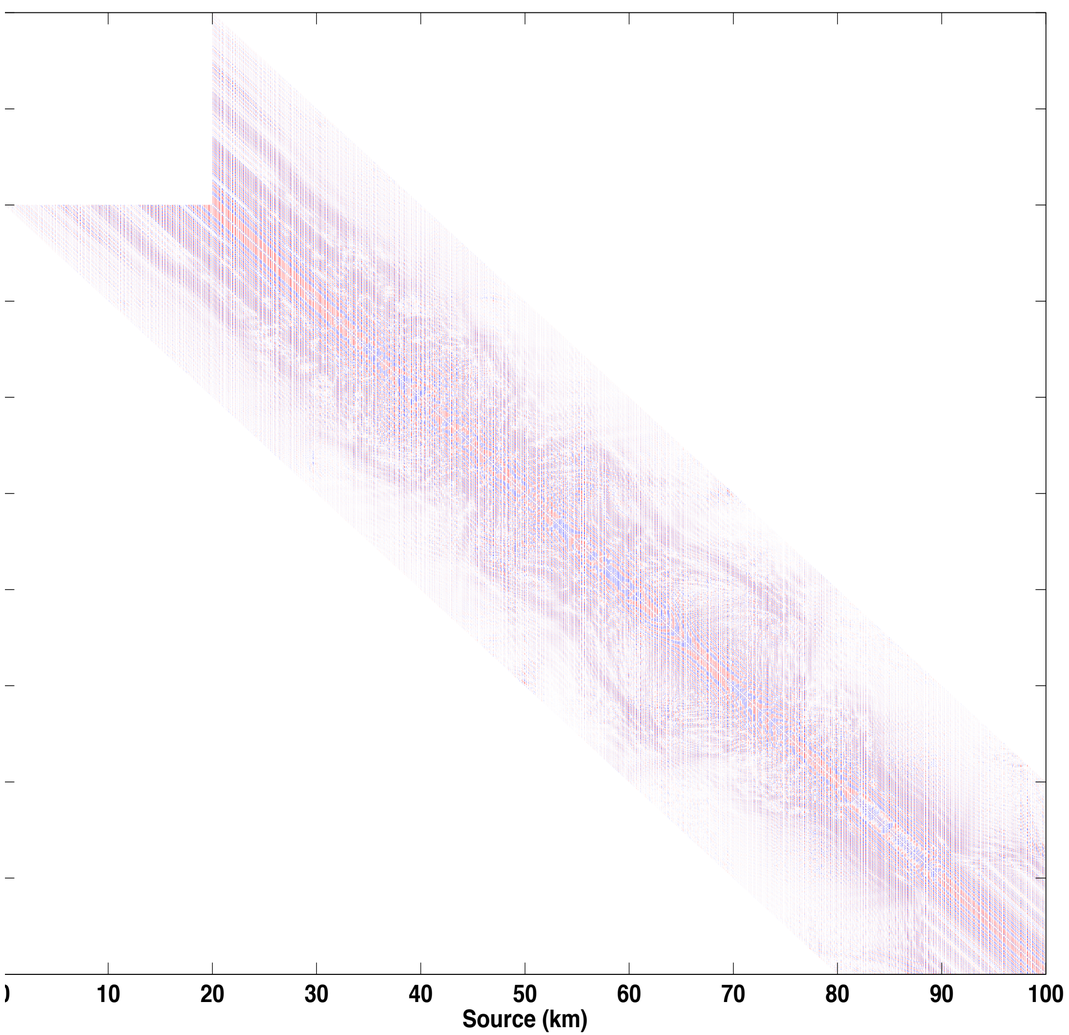
80 % missing sources

Comparison: LR-factorization versus Curvelet transform

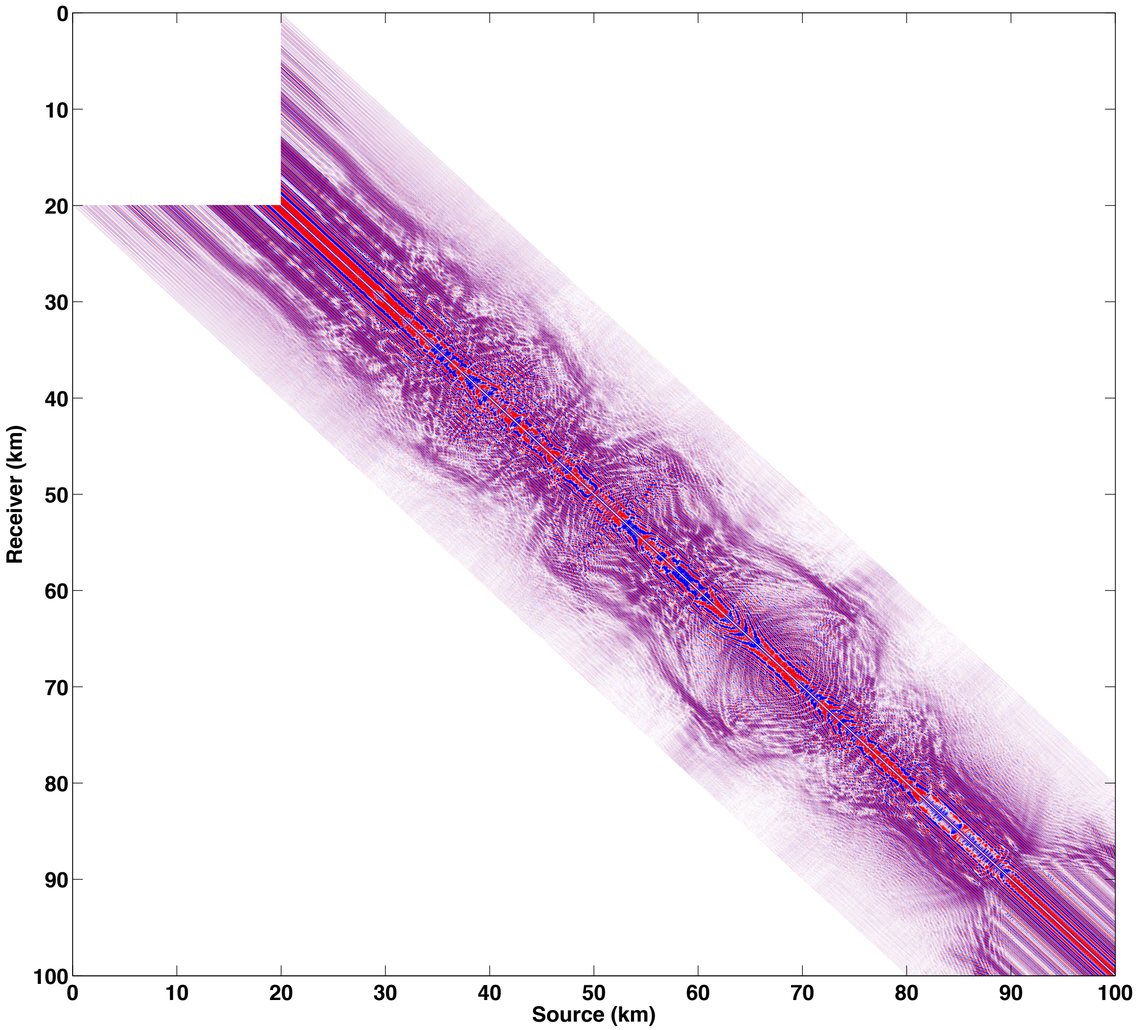
4001 X 4001



True data

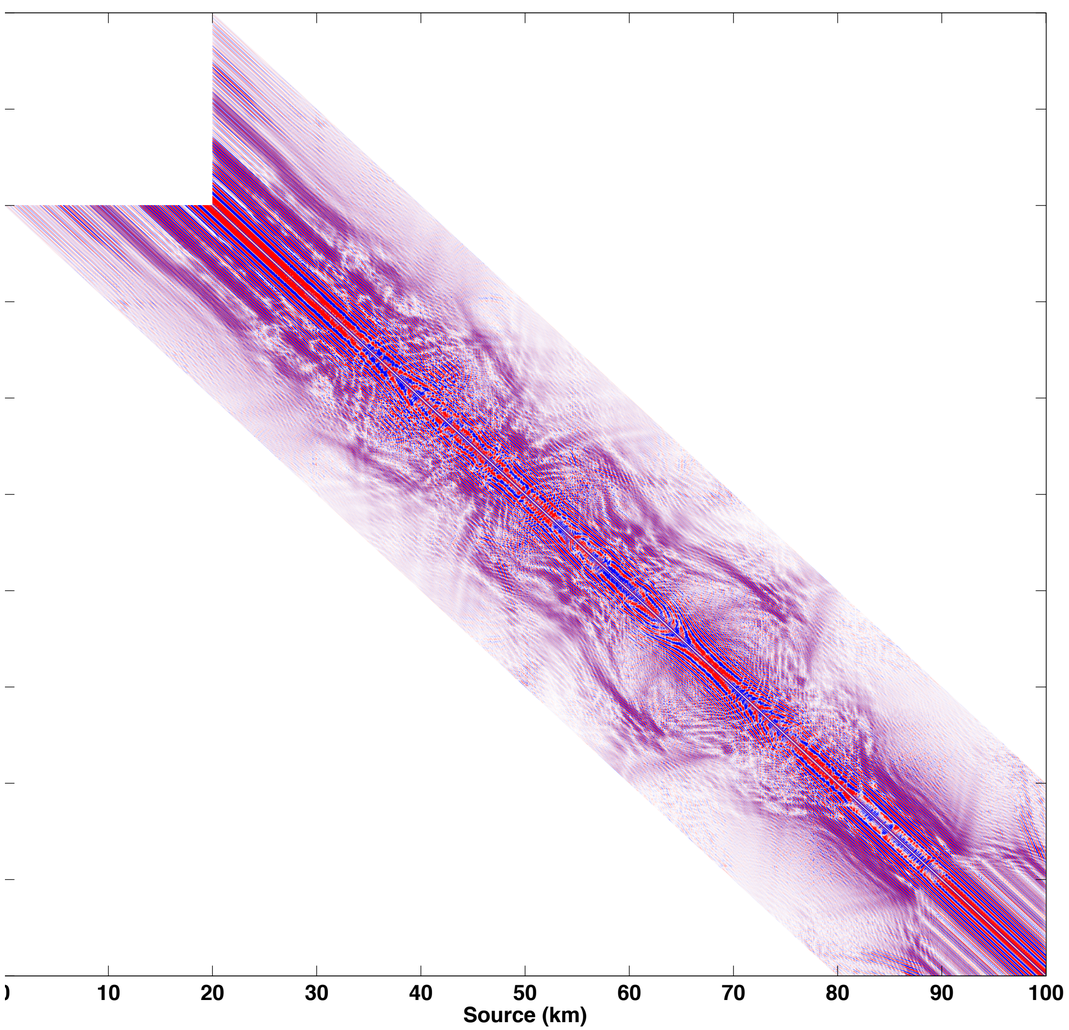


Subsampled data



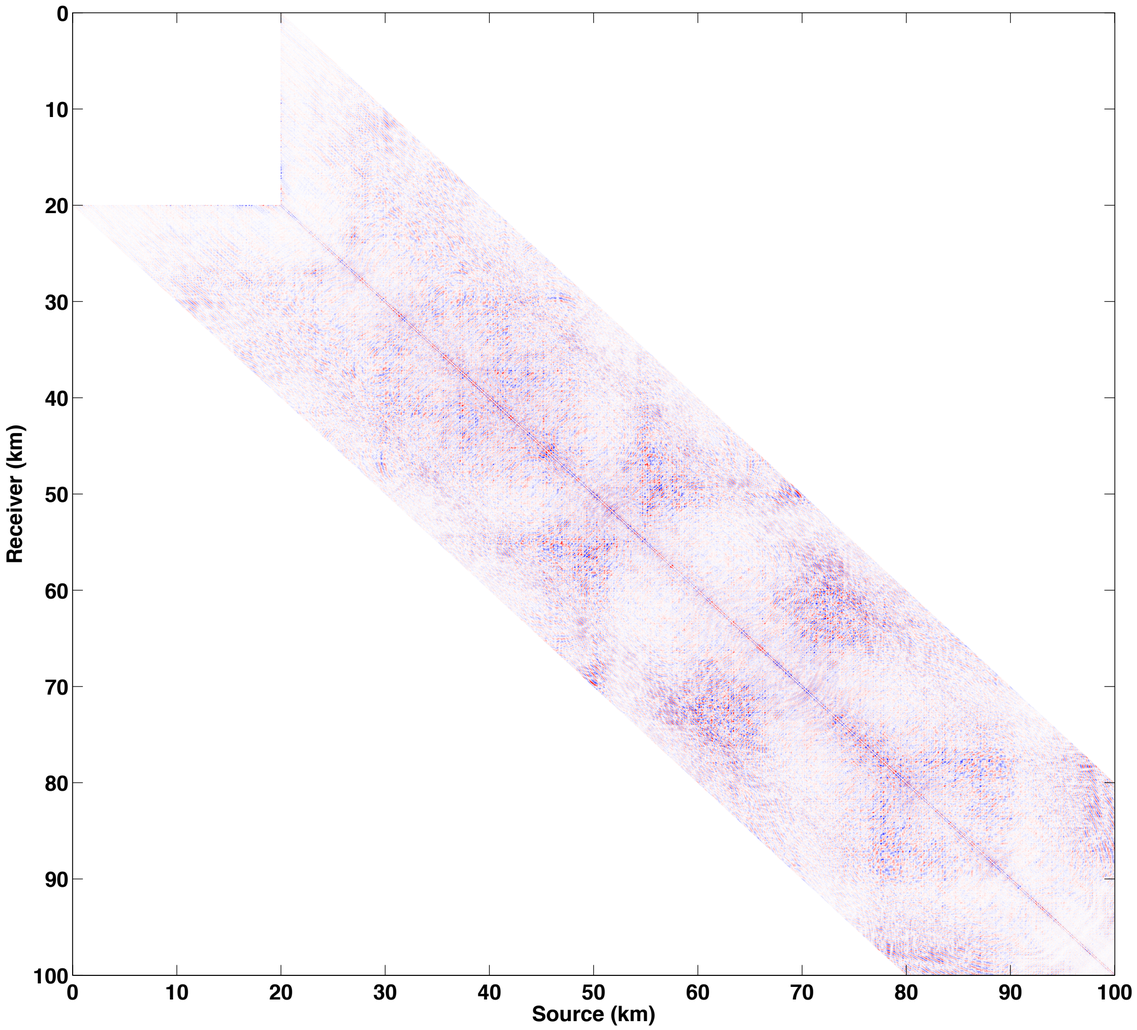
Recovery - LR

Time = 20 minutes, SNR = 14.2 dB

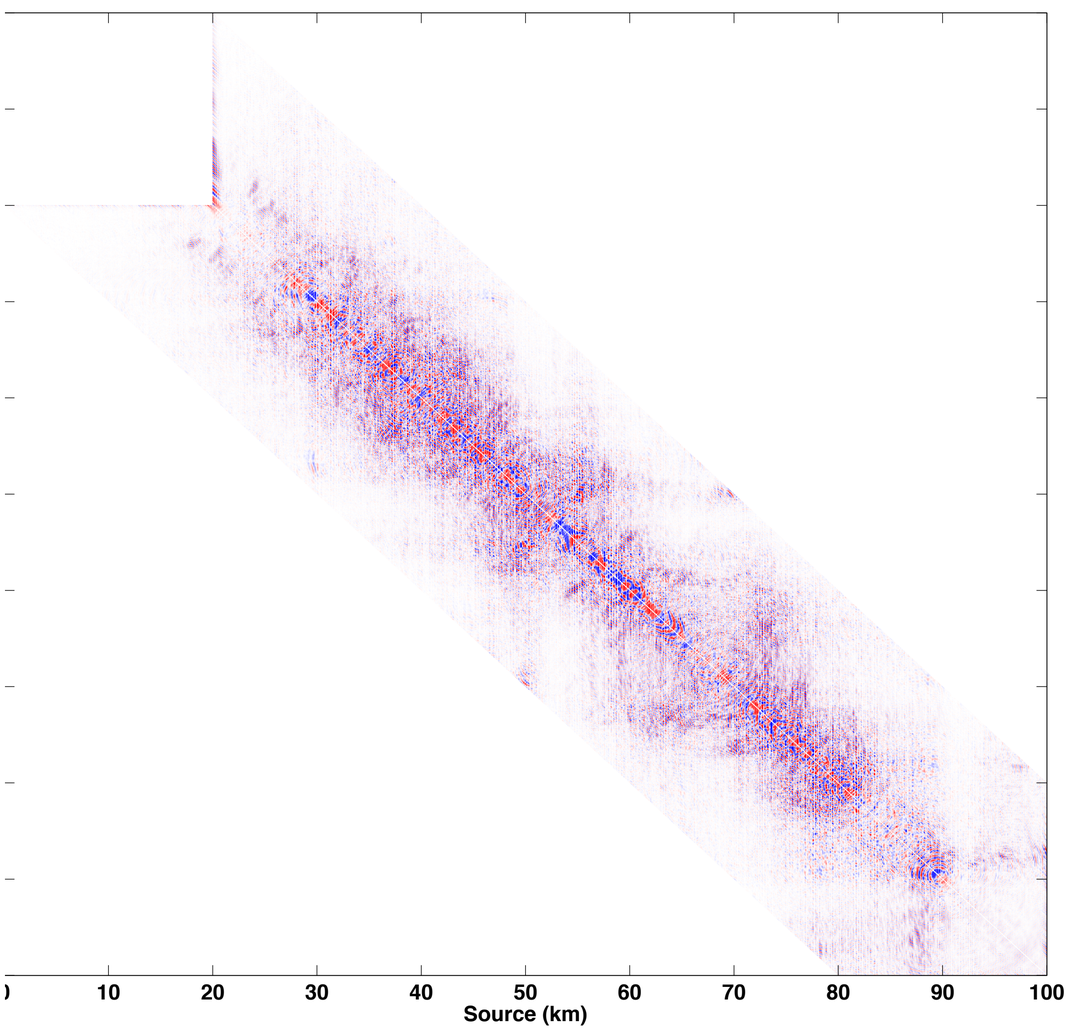


Recovery - curvelet

Time = 22 hours, SNR = 9.4 dB



difference - LR



difference - curvelet

Synthetic BG Group data

Hierarchical Tucker

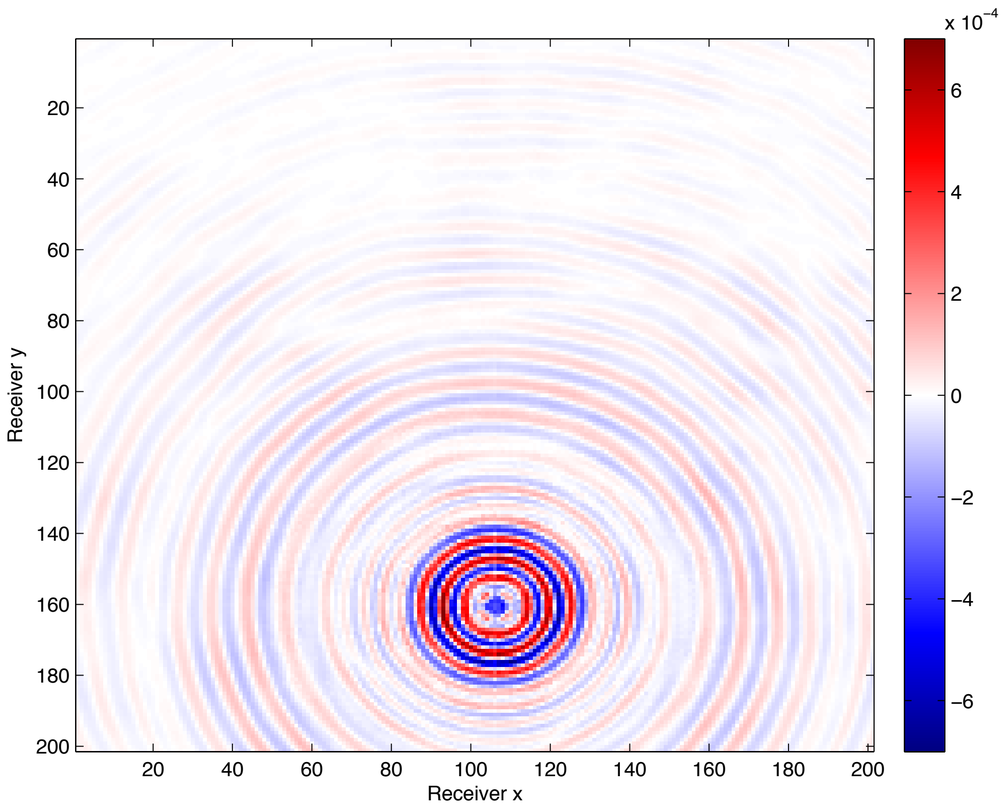
Unknown model

- 68 x 68 sources with 401 x 401 receivers, data at 7.34 Hz

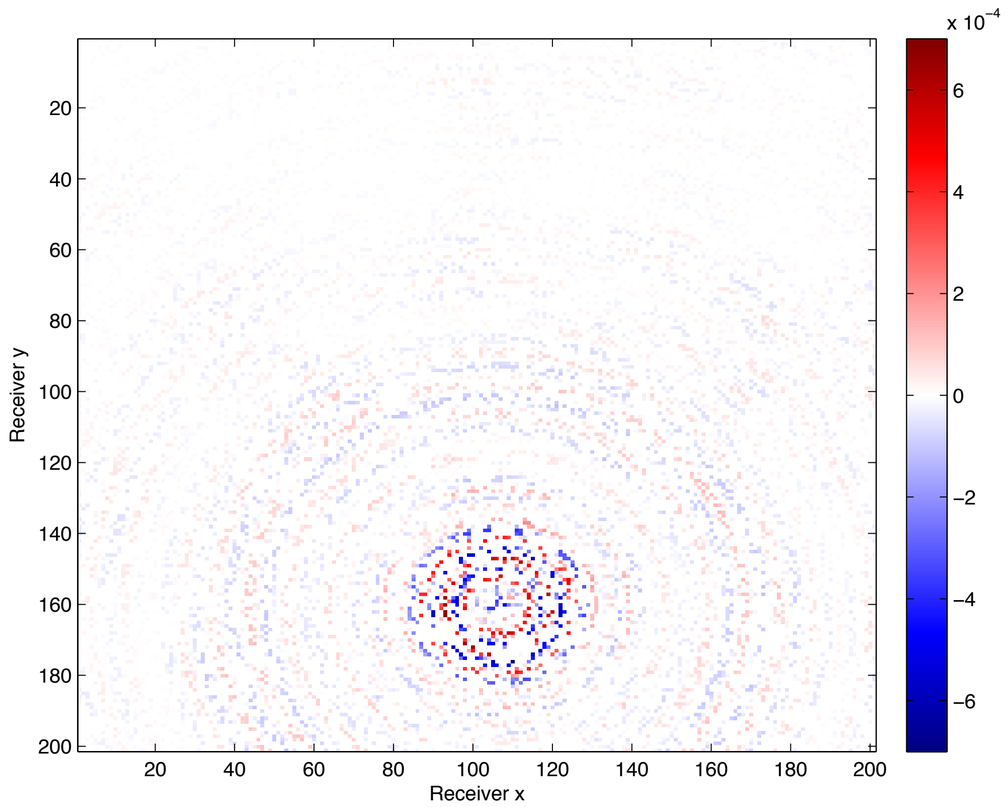
Receivers subsampled to 101 x 101, Fourier interpolated to 401 x 401

7.34 Hz - 75% missing receivers

Fixed source coordinate, varying receiver coordinates



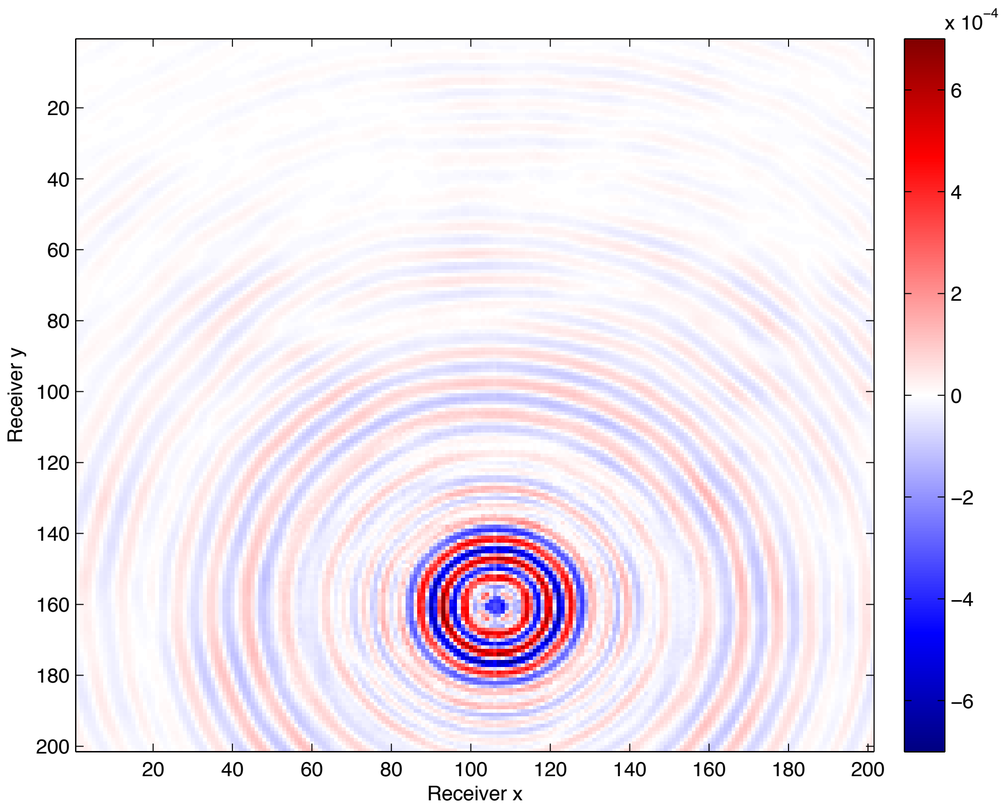
True data



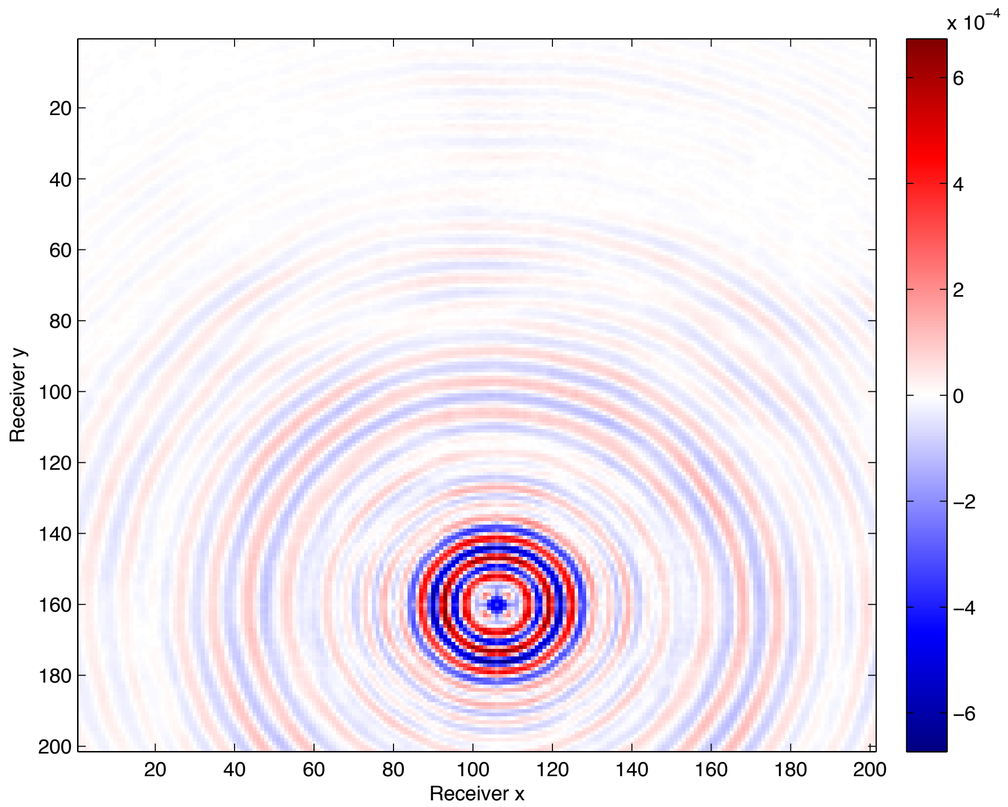
Subsampled data

7.34 Hz - 75% missing receivers

Fixed source coordinate, varying receiver coordinates



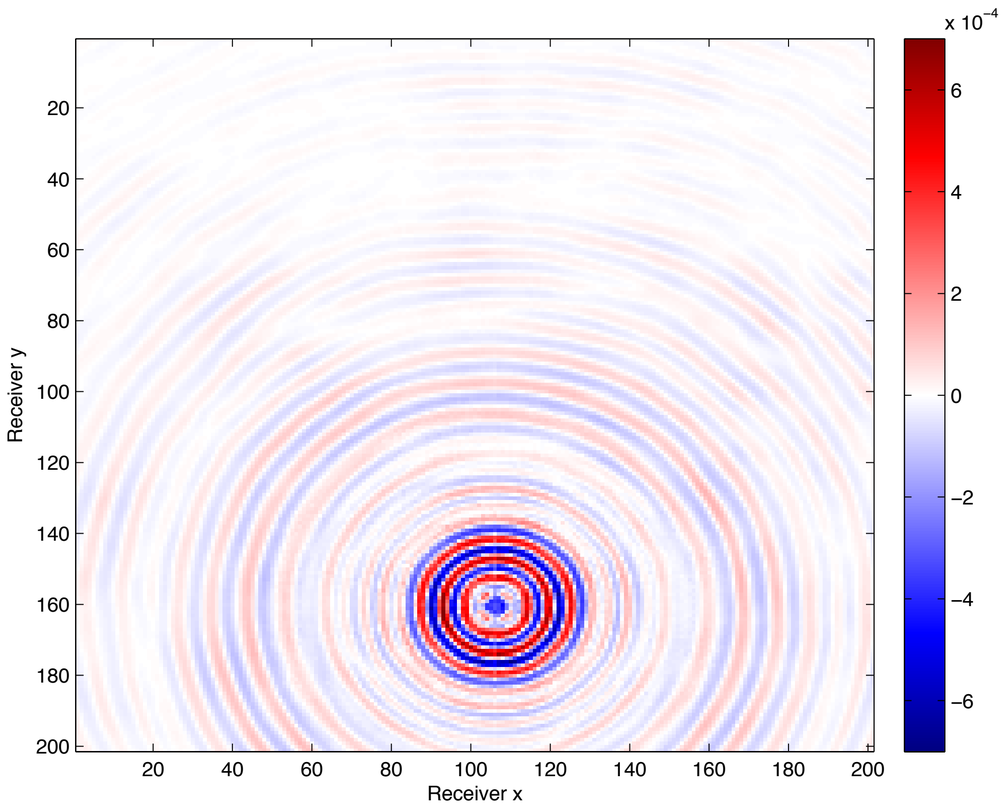
True data



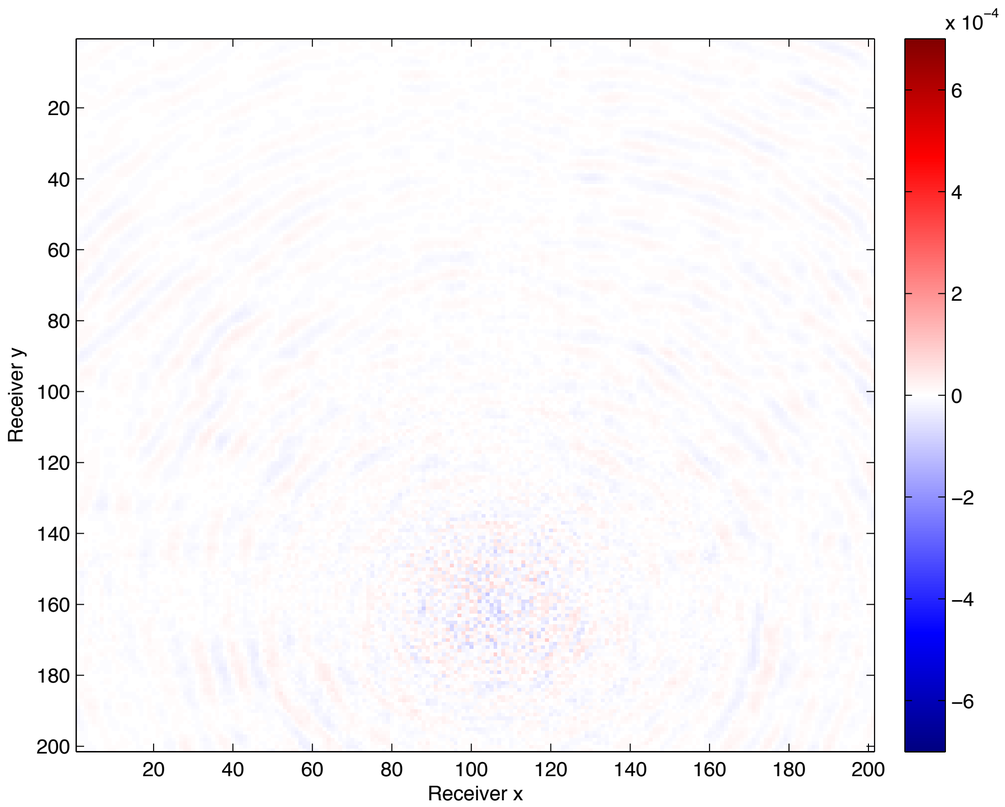
Recovered data - SNR 17.6 dB

7.34 Hz - 75% missing receivers

Fixed source coordinate, varying receiver coordinates



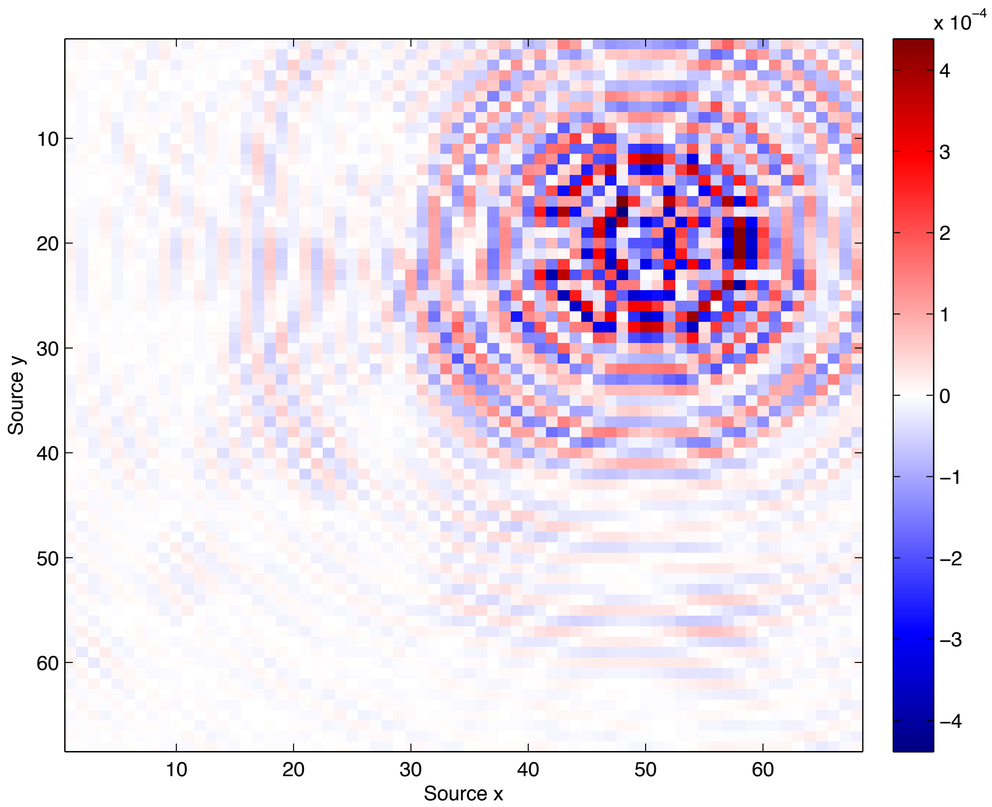
True data



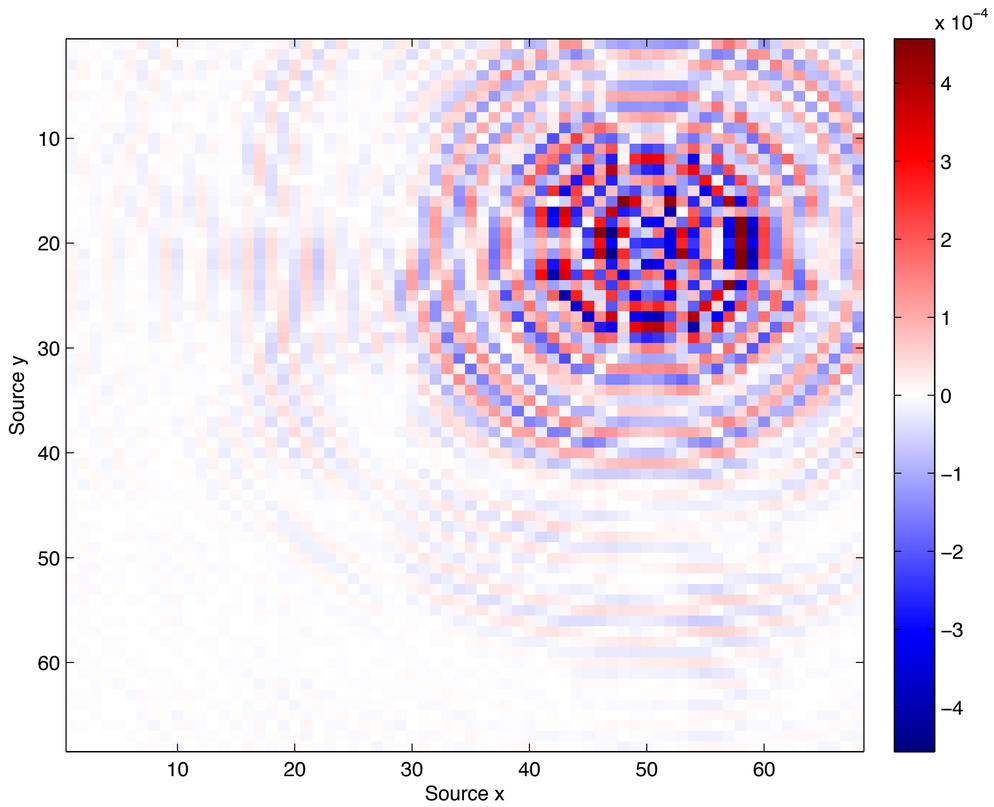
Difference

7.34 Hz - 75% missing receivers

Fixed receiver coordinates (not in training data), varying source coordinates



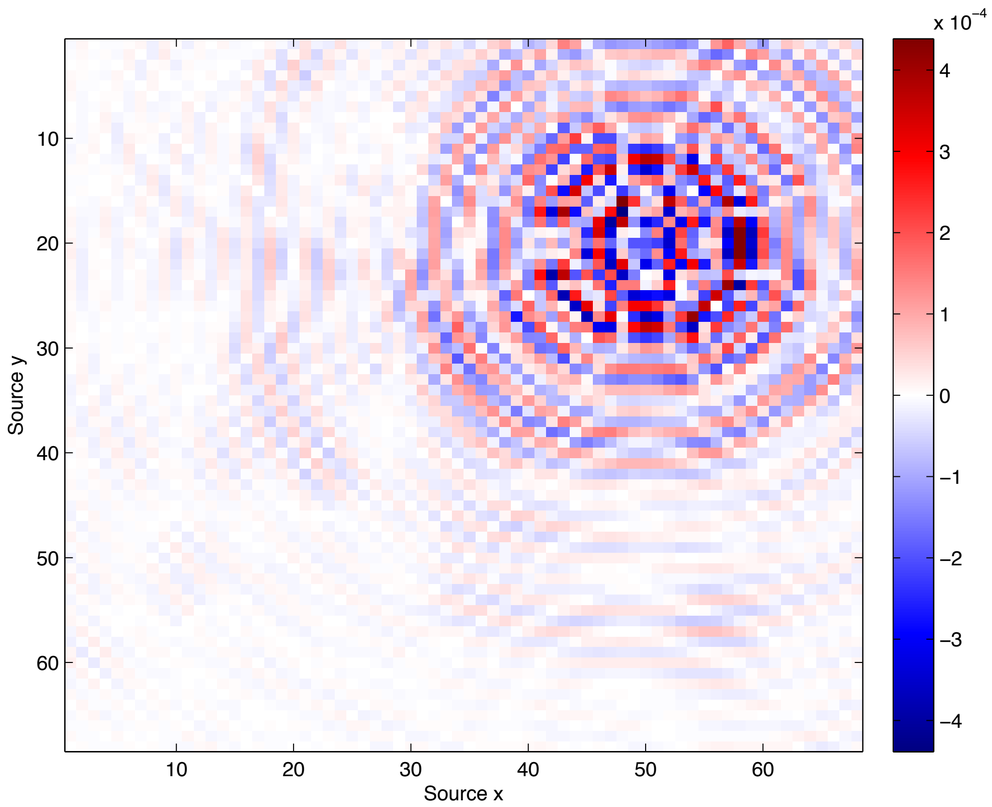
True data



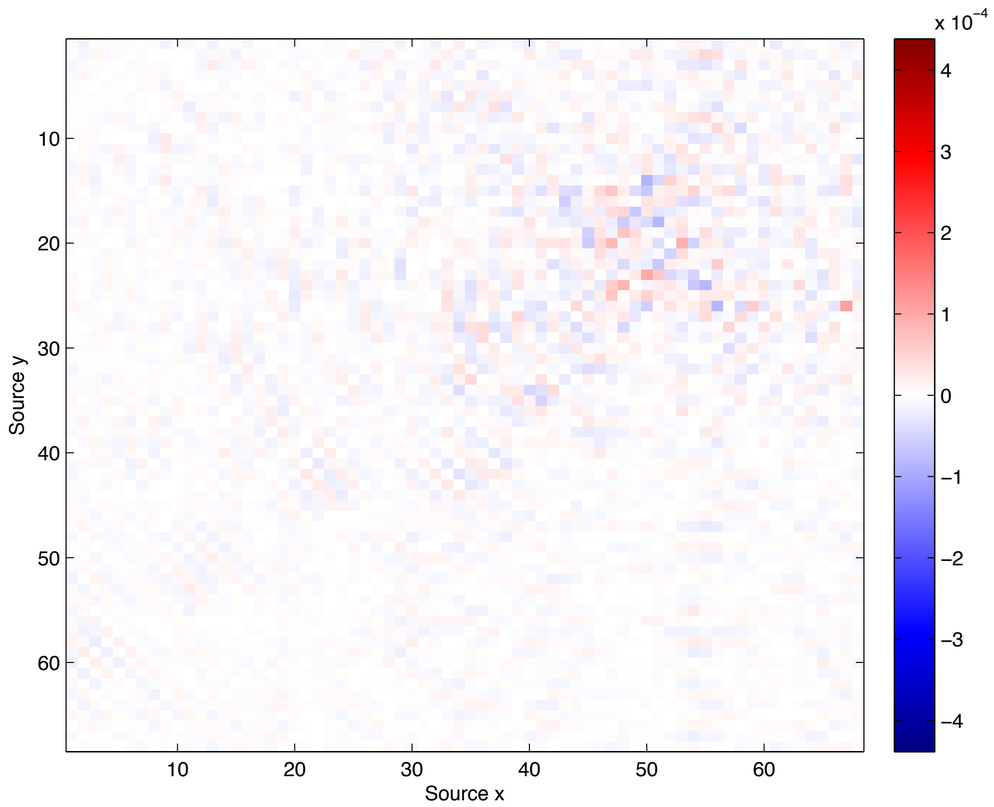
Recovered data - SNR 16.2 dB

7.34 Hz - 75% missing receivers

Fixed receiver coordinates (not in training data), varying source coordinates



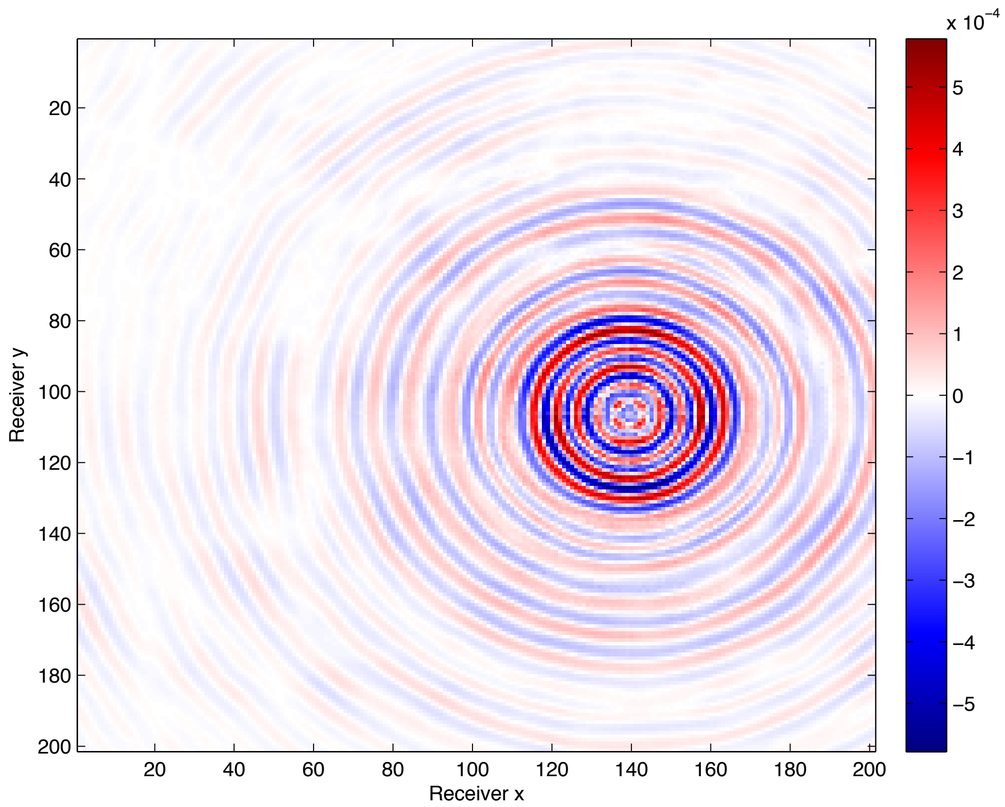
True data



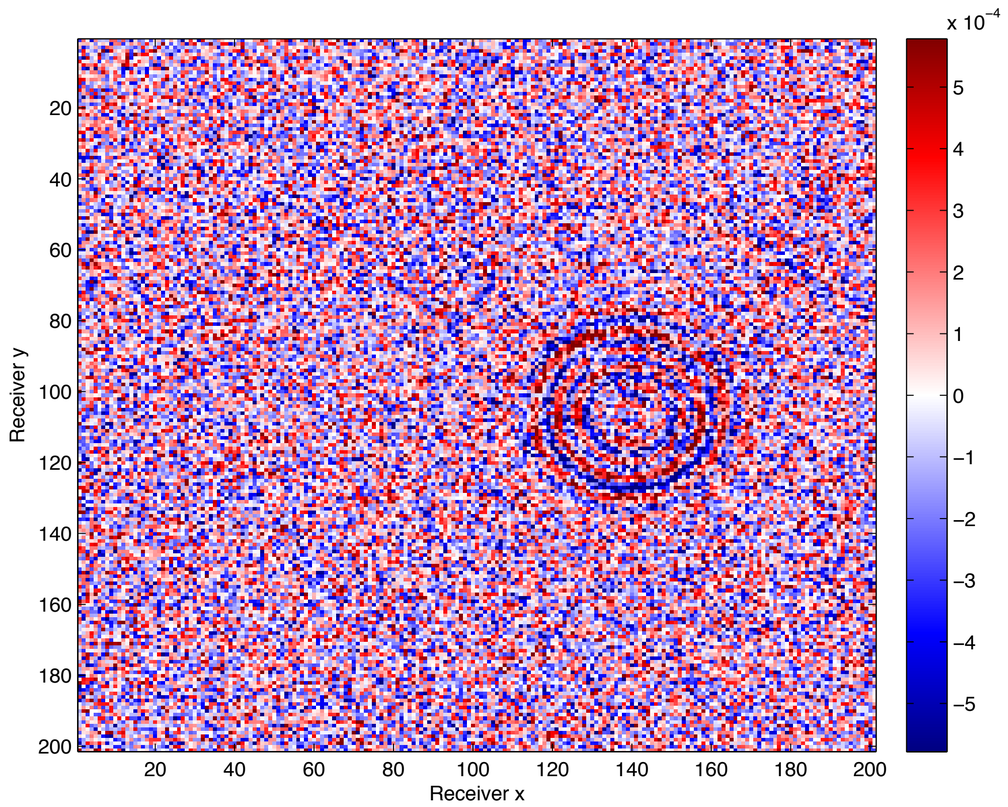
Difference

7.34 Hz - simultaneous receivers - 90% data reduction

Fixed source coordinate, varying receiver coordinates



True data

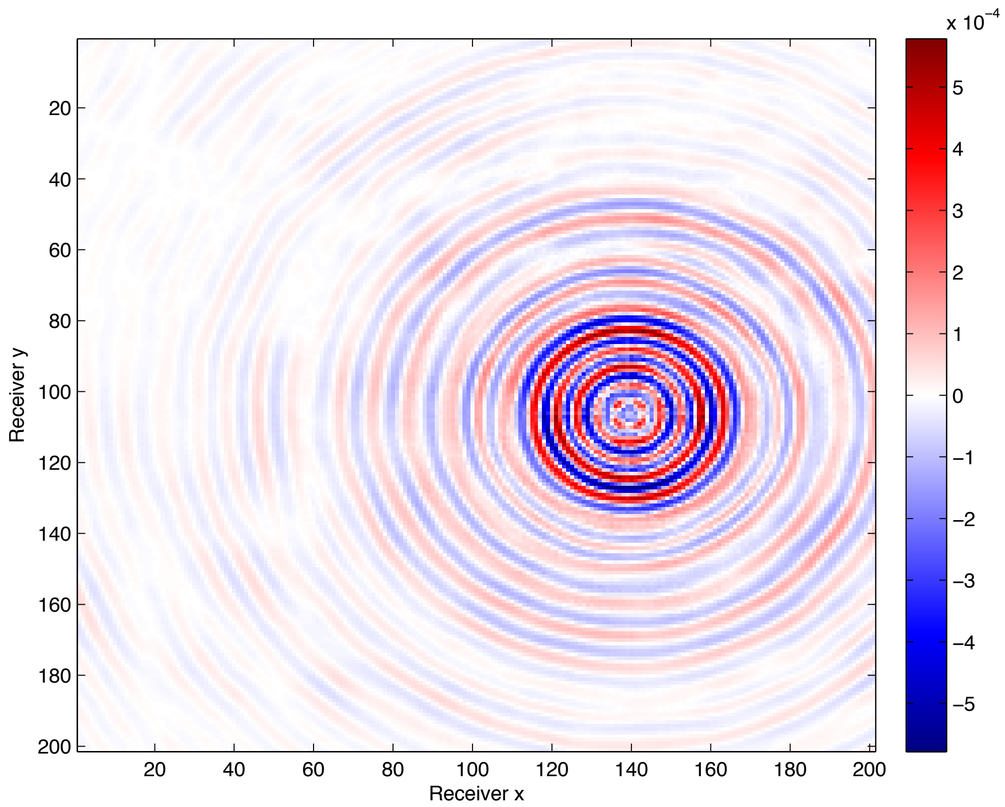


Input data - $A^T Ab$

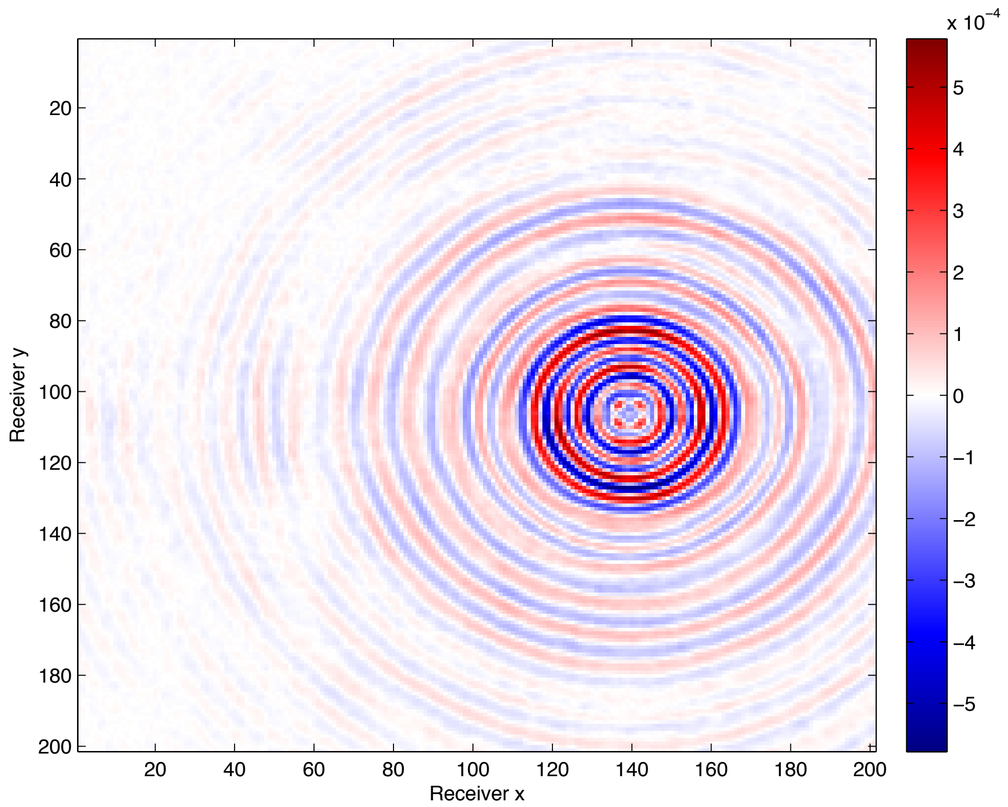
A - subsampling operator b - full data

7.34 Hz - simultaneous receivers - 90% data reduction

Fixed source coordinate, varying receiver coordinates



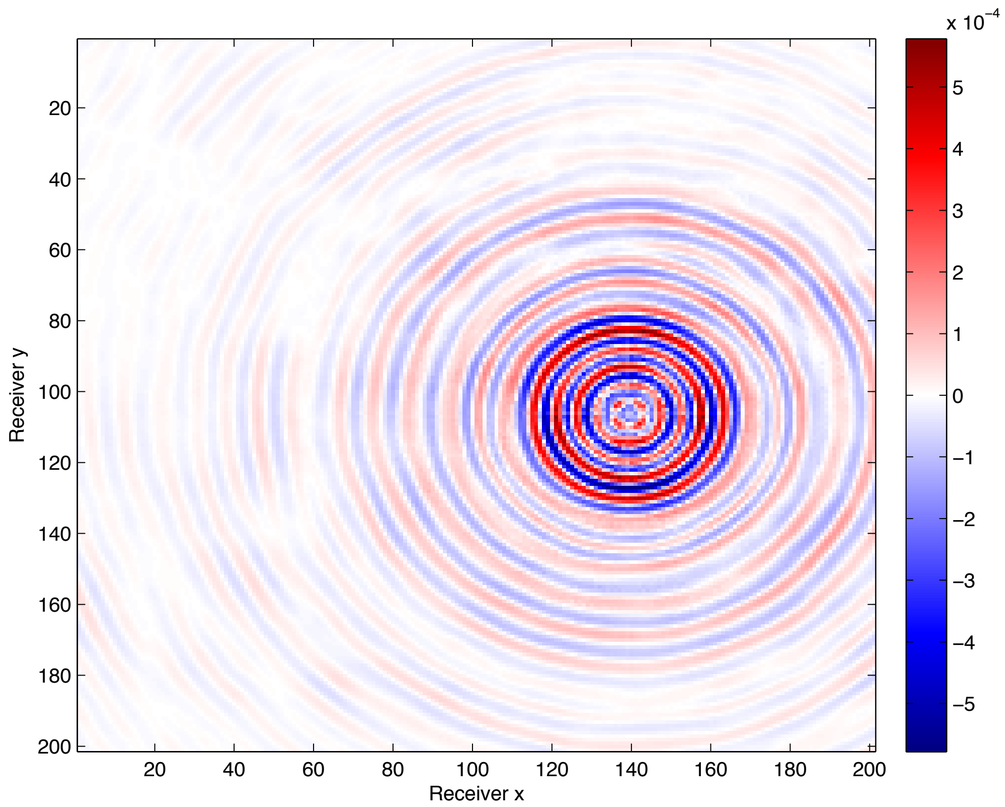
True data



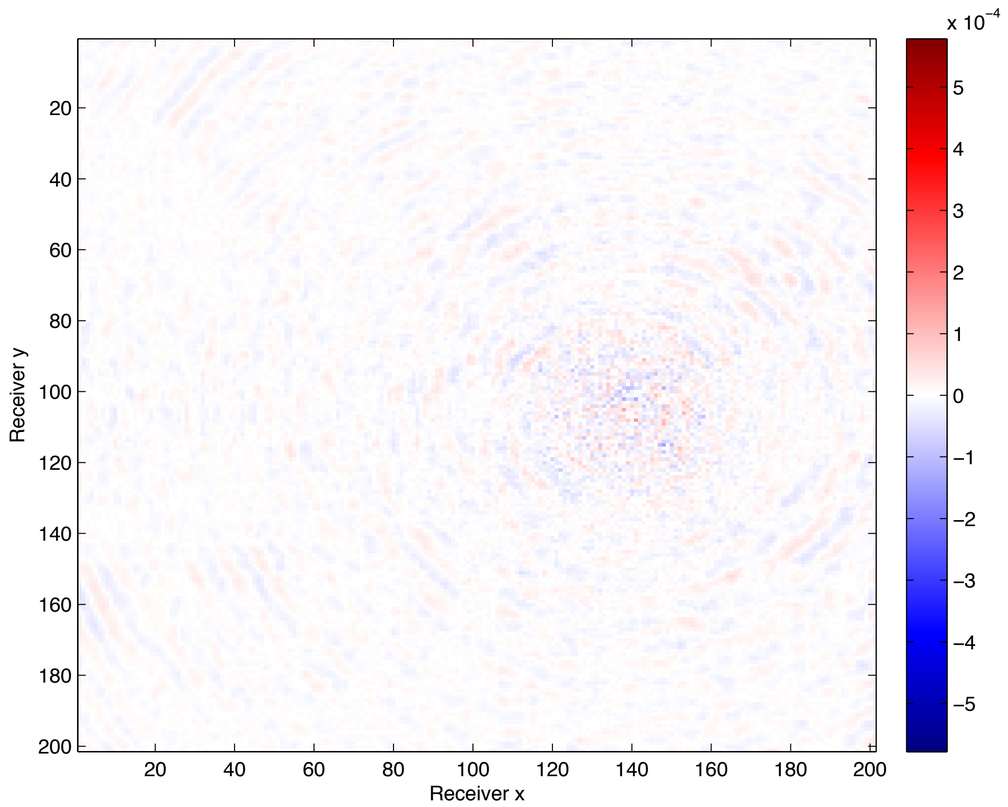
Recovered data - SNR 16.4 dB

7.34 Hz - simultaneous receivers - 90% data reduction

Fixed source coordinate, varying receiver coordinates



True data



Difference

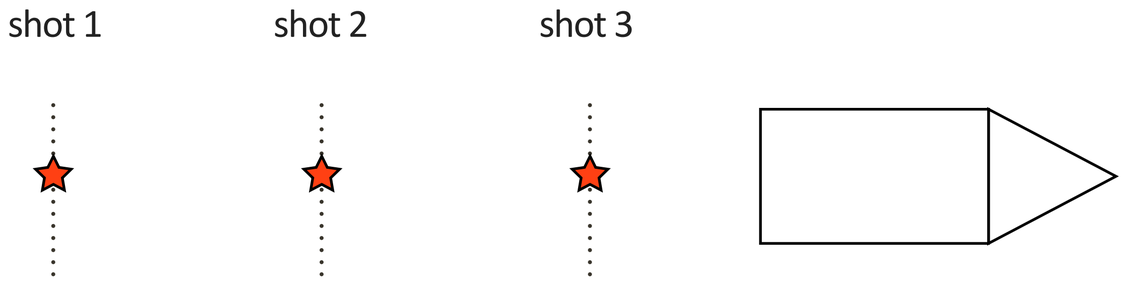
Source separation for *short*-time jitter in marine

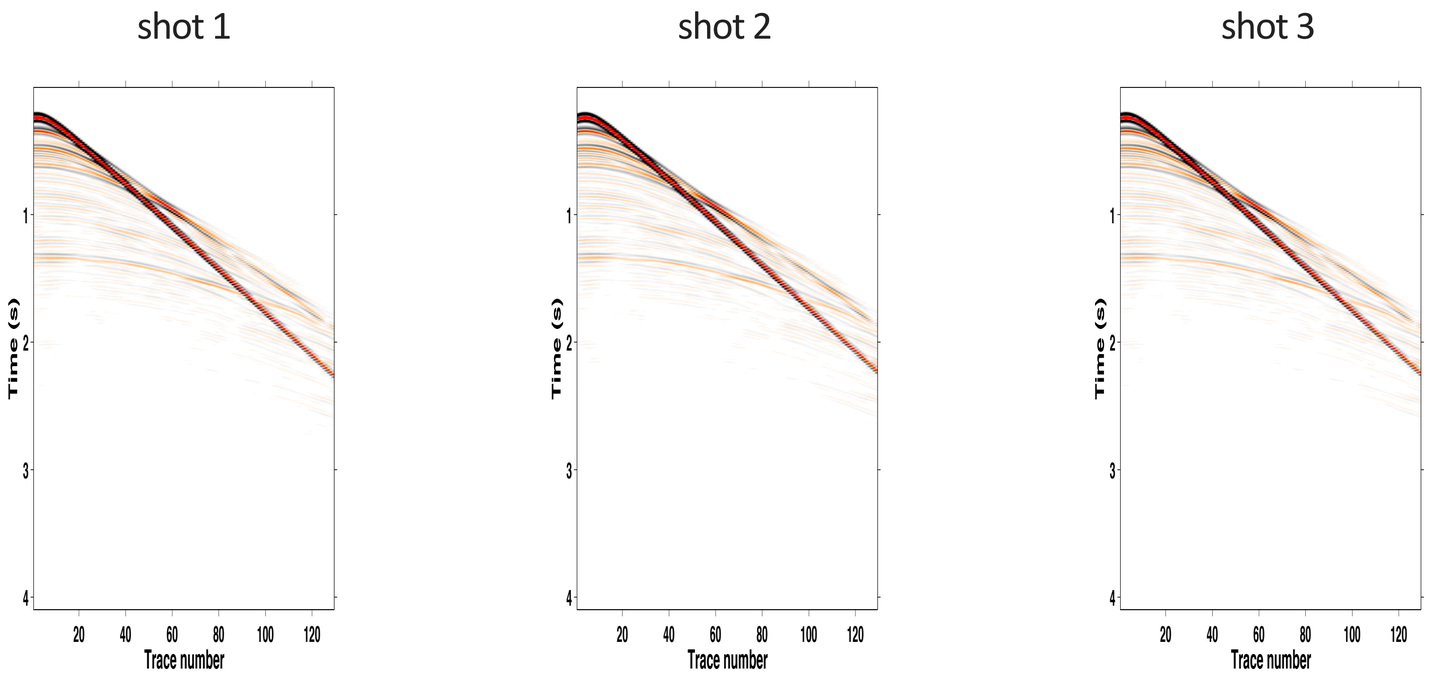
Conventional marine acquisition

source depth = 10 m

(below sea level)

regularly sampled spatial grid





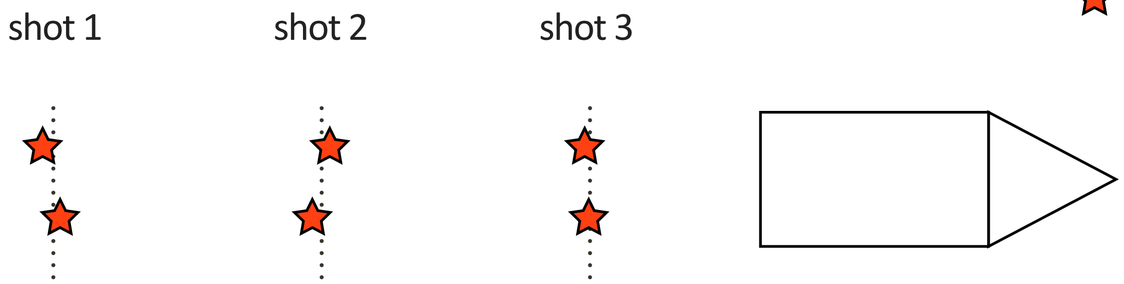
Blended/Simultaneous marine acquisition

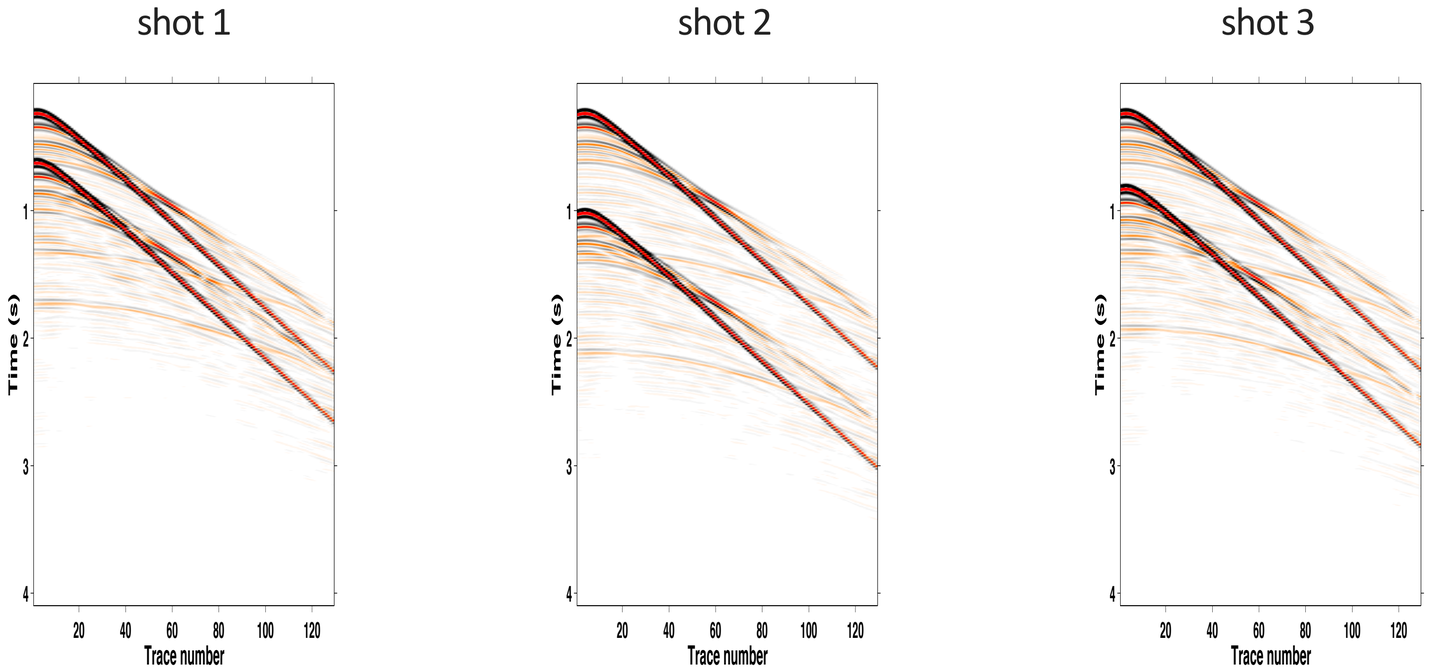
source1 depth = 10 m

source2 depth = 15 m

regularly sampled spatial grid (almost)

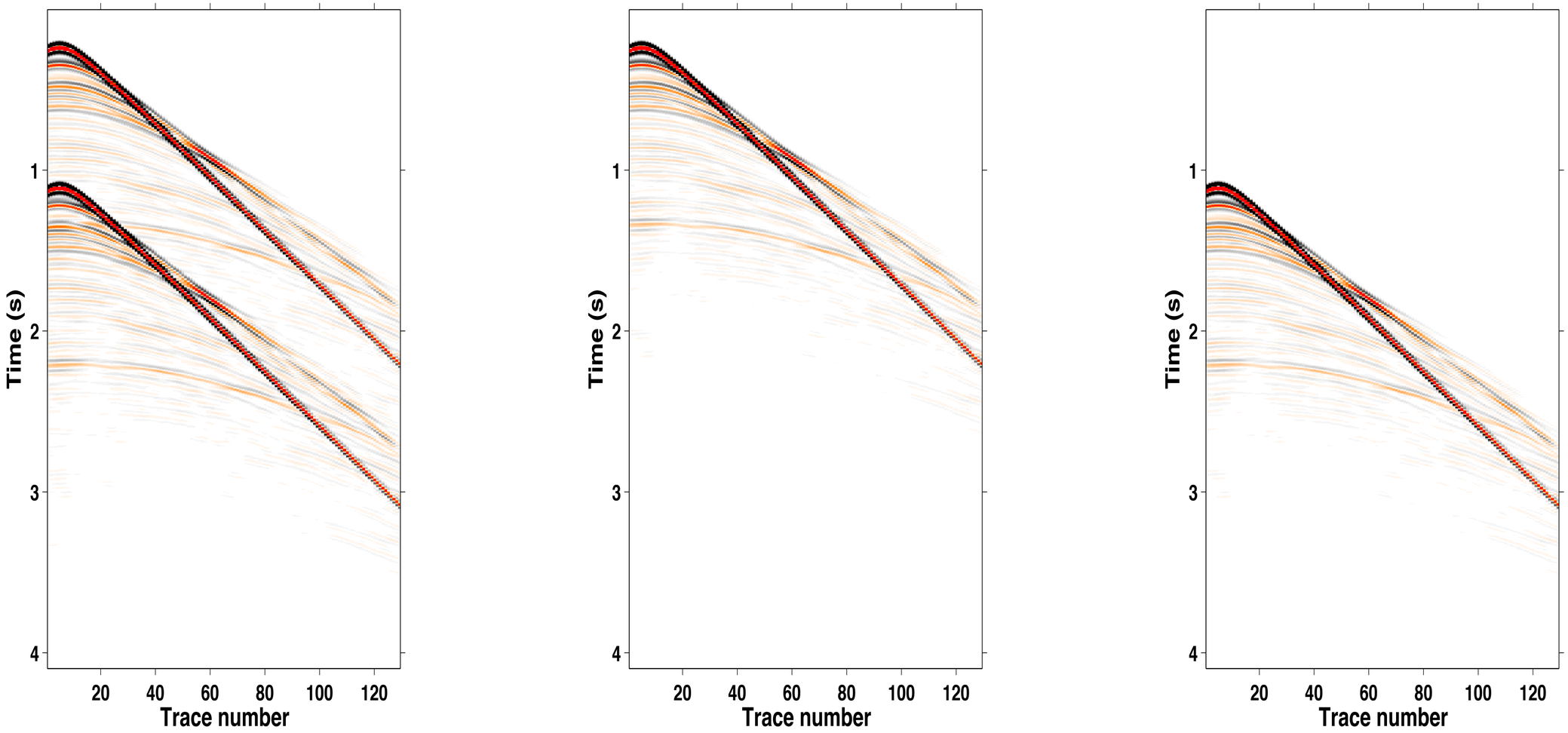
shot-time randomness - ***LOW***





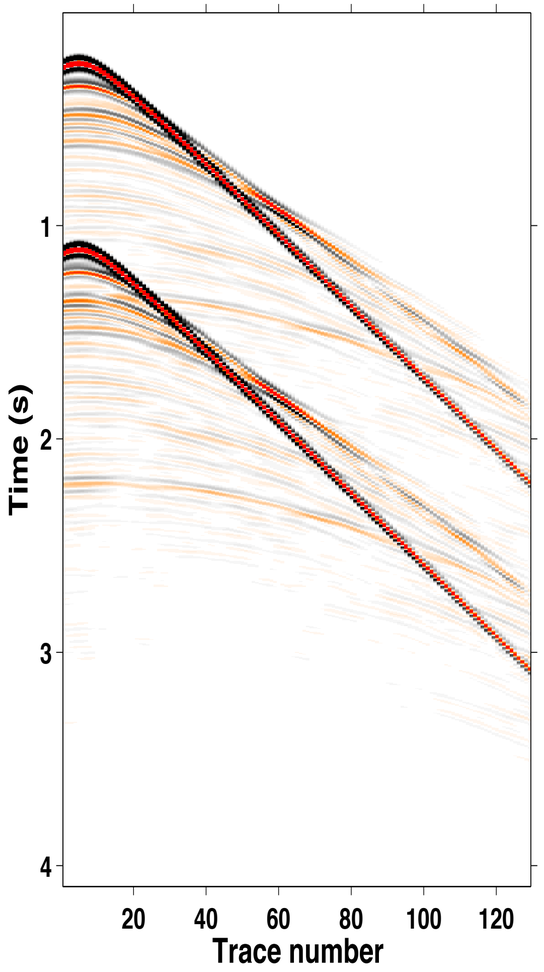
Blended data

blended shot = source 1 + source 2 (time-delayed)

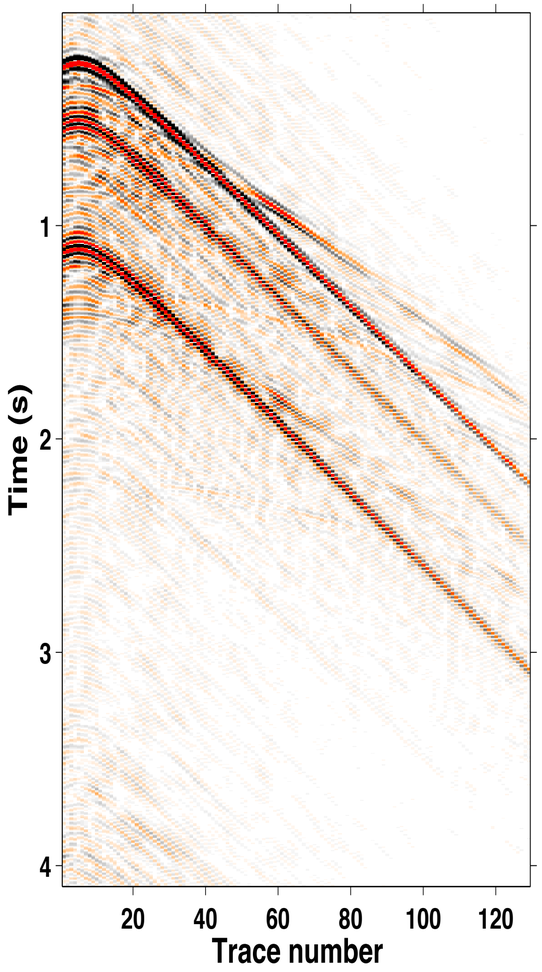


Source separation via *sparsity*-promotion

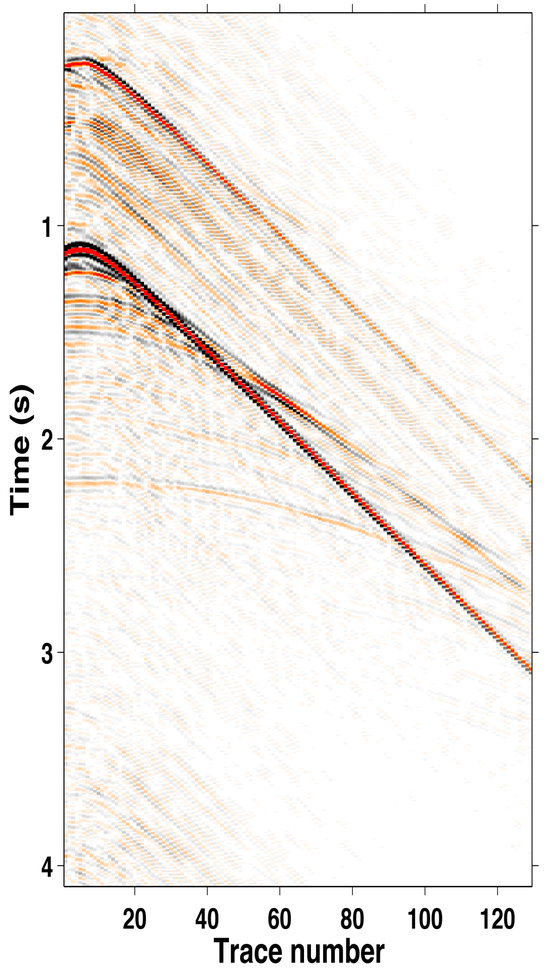
blended shot



source 1



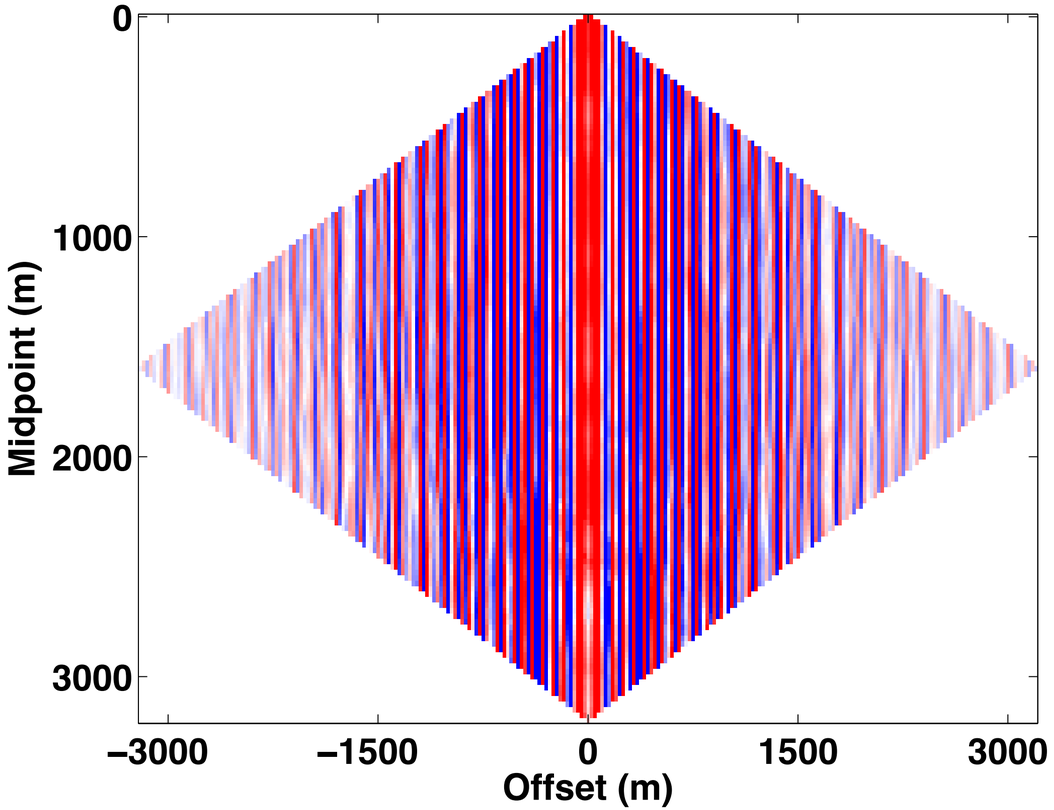
source 2
(time-delayed)



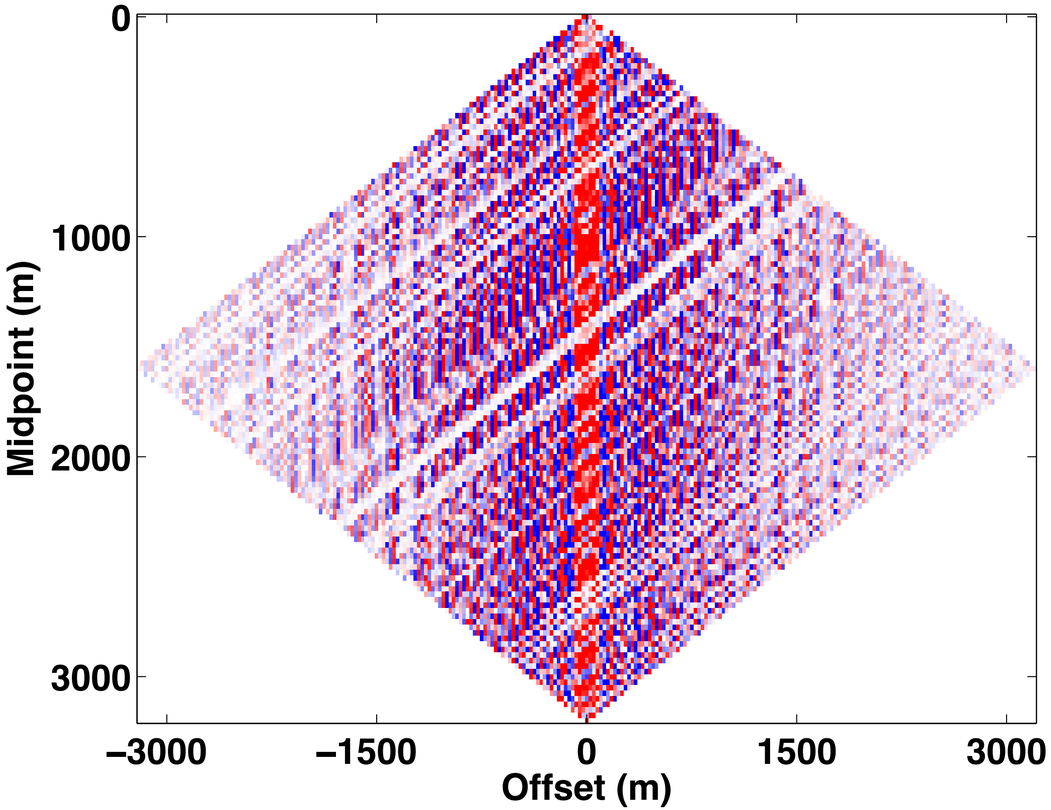
Low-rank structure – frequency slice at 25 Hz

in midpoint-offset domain?

without delay

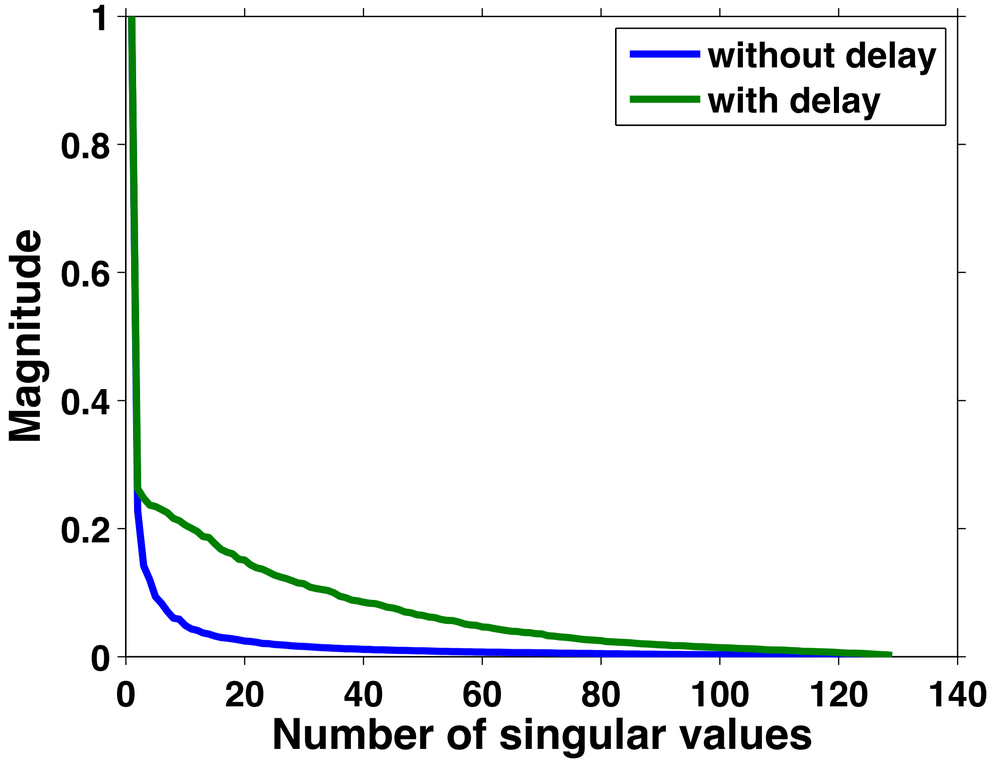


with delay

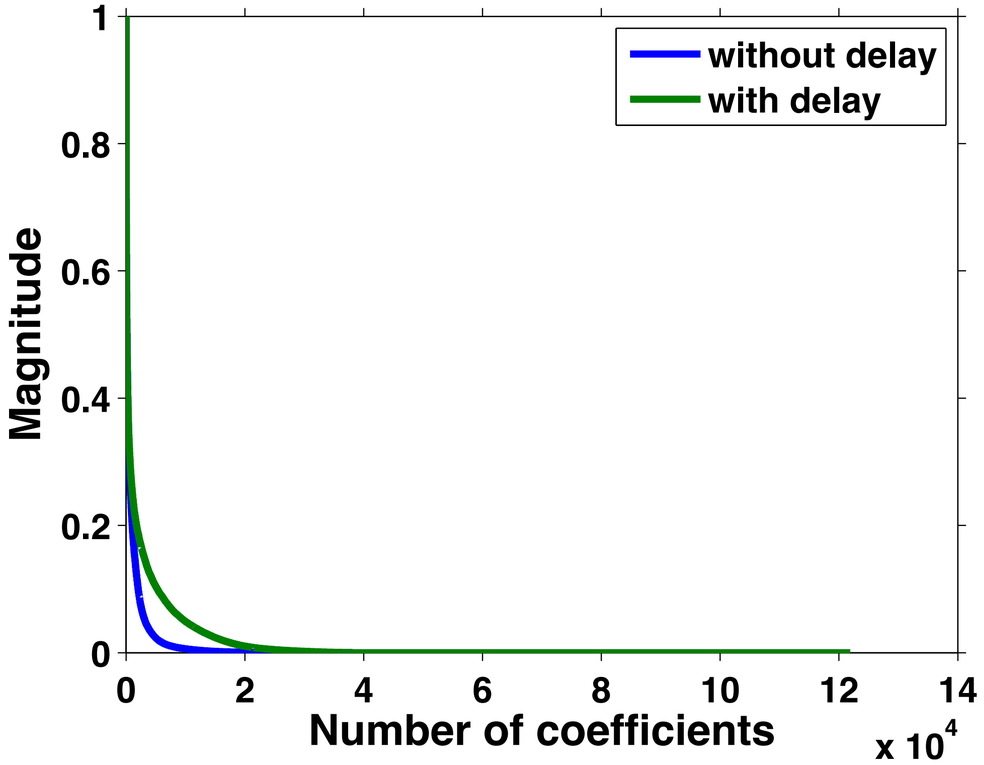


Rank vs. sparsity

rank-minimization
(midpoint-offset domain)

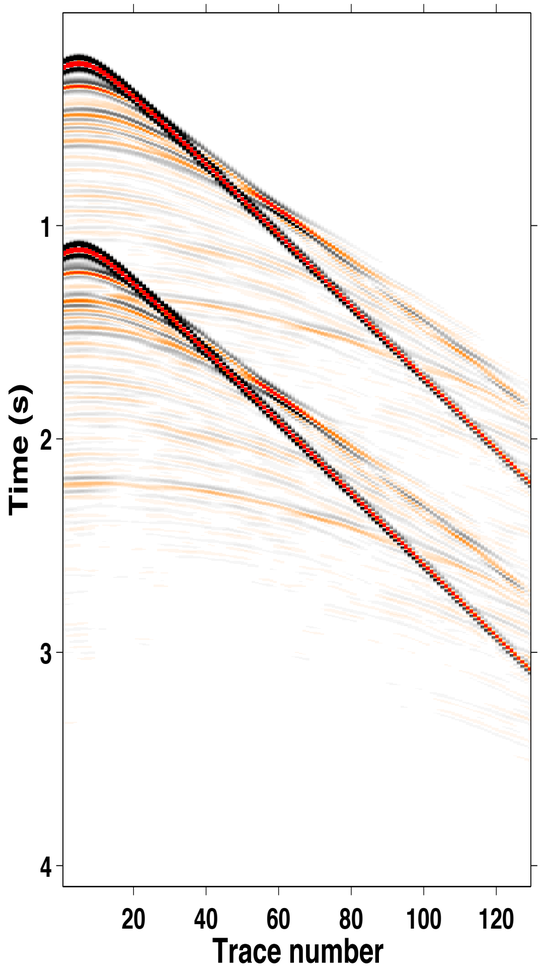


sparsity-promotion
(source-receiver domain)

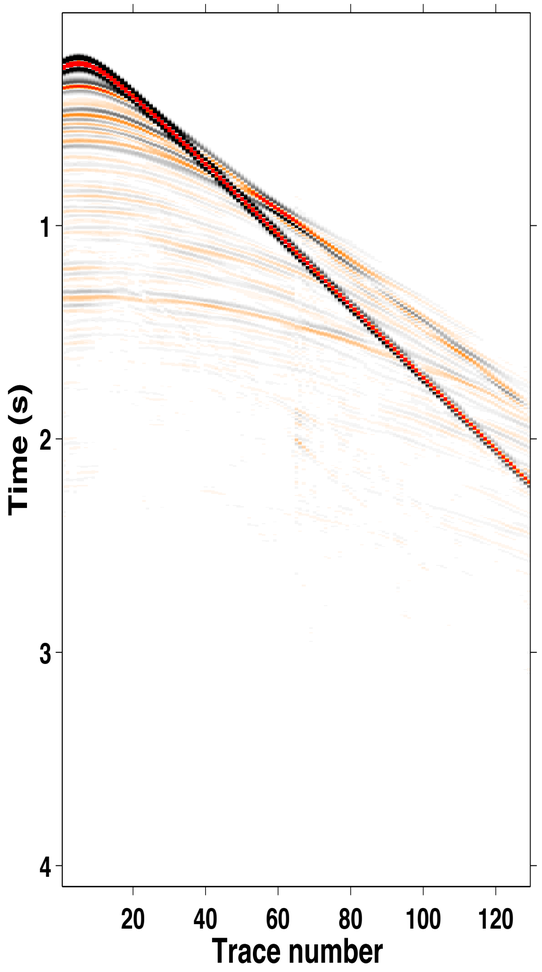


Source separation via *rank*-minimization

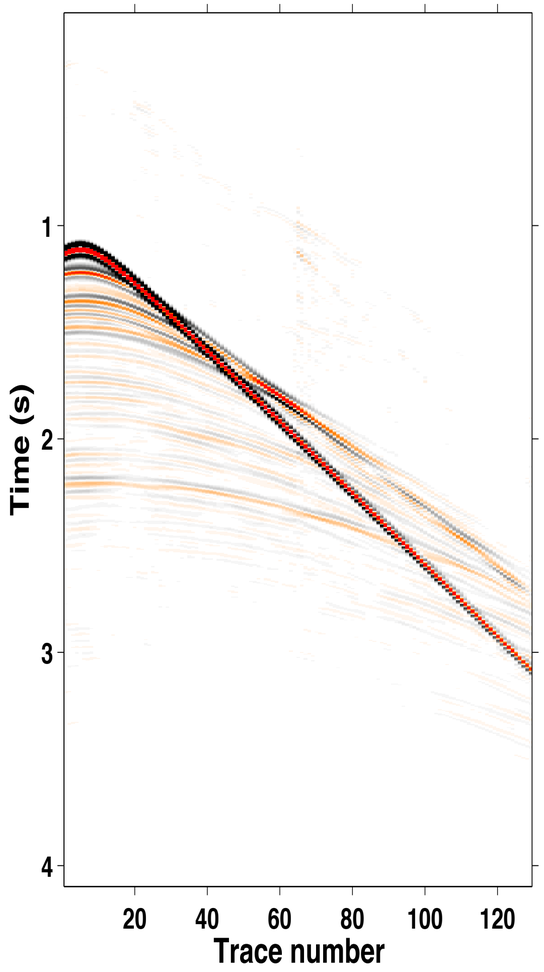
blended shot



source 1

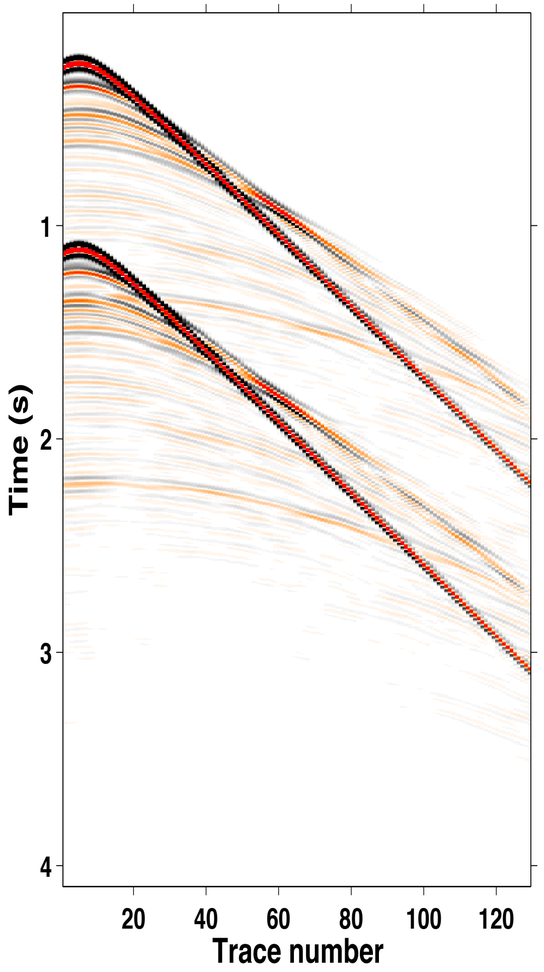


source 2
(time-delayed)

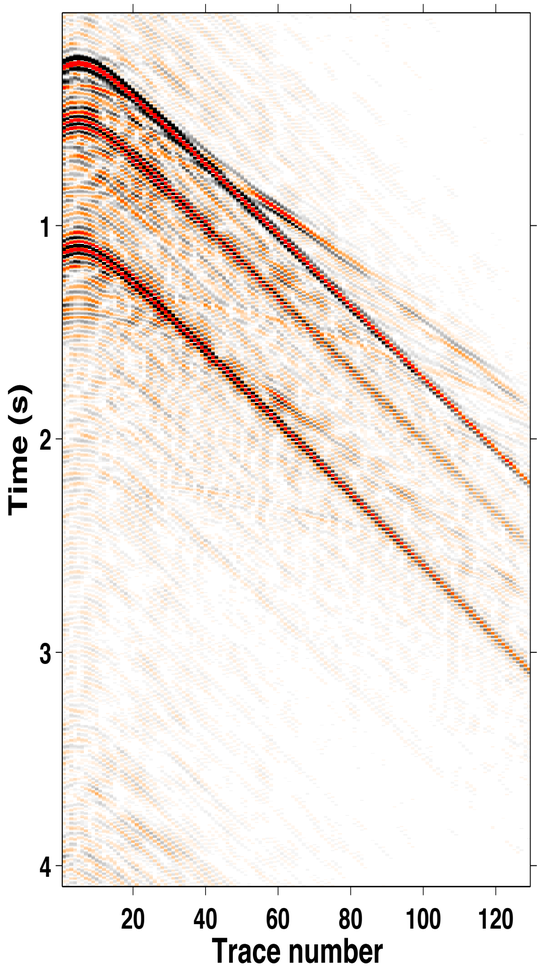


Source separation via *sparsity*-promotion

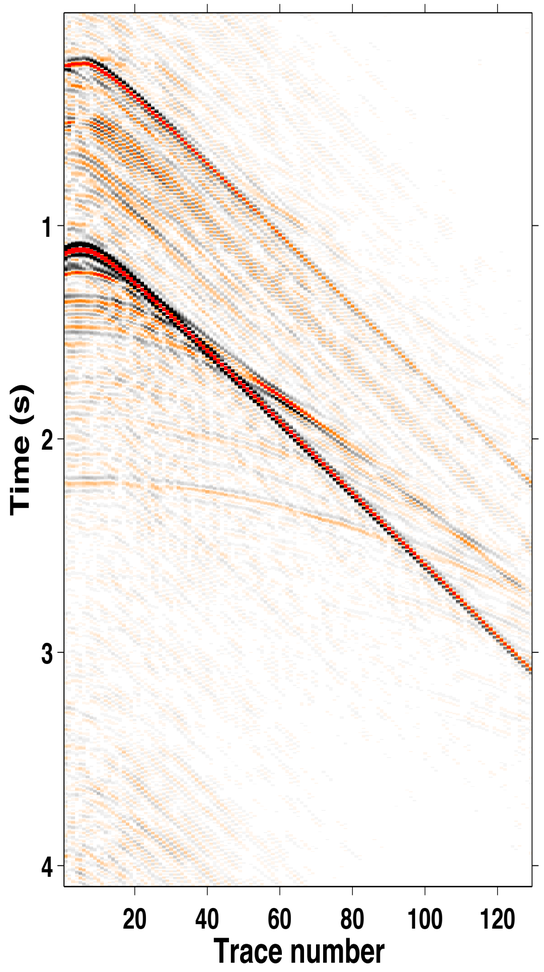
blended shot



source 1



source 2
(time-delayed)



Observations & open problems

Seismic data exhibits low-rank structure that can be exploited

Non-convex (manifold) optimizations seem to do the job

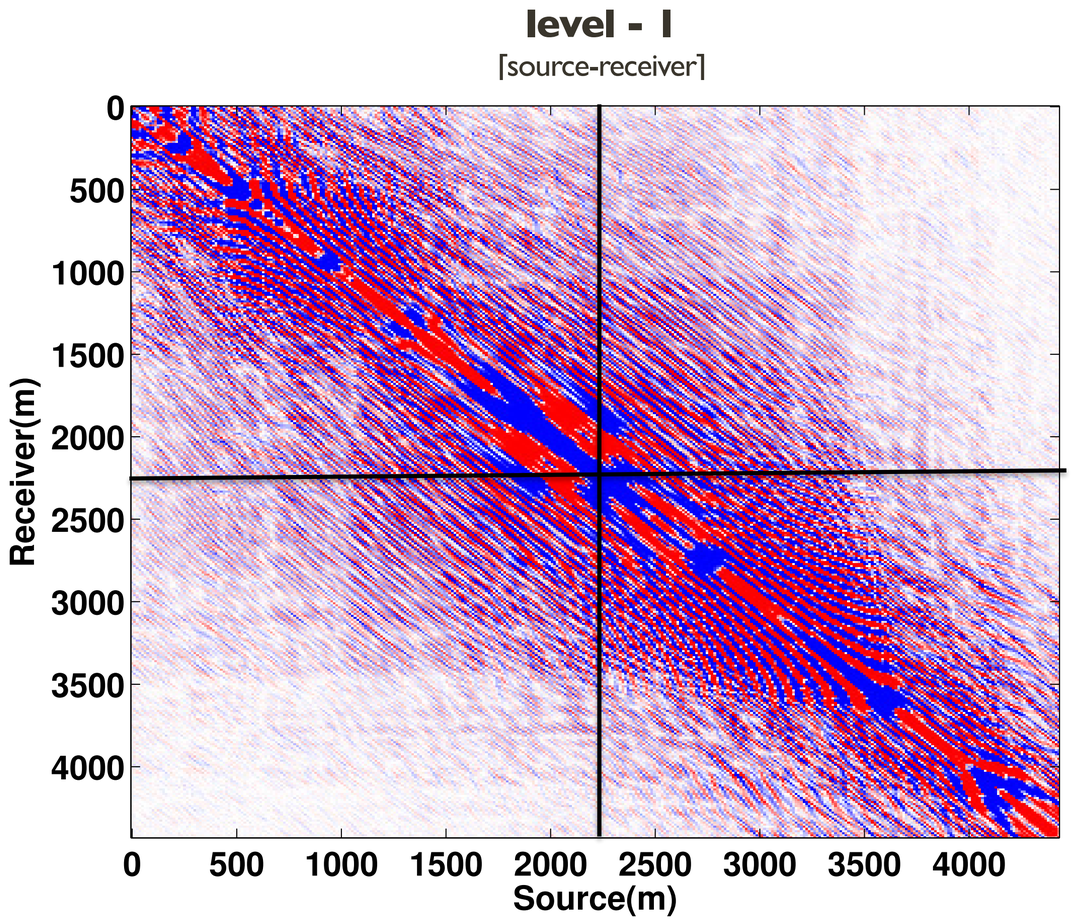
However,

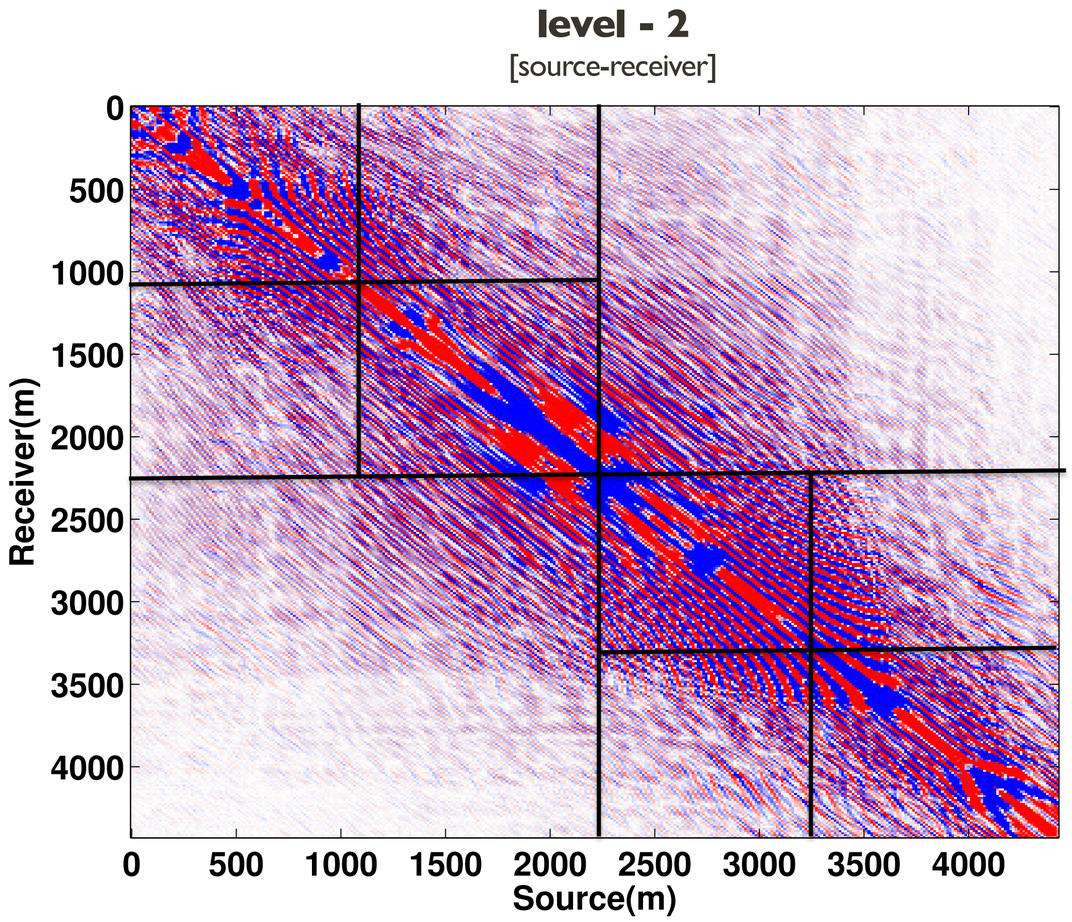
- no practical quantitative design principles, recovery guarantees and UQ
  - how to parameterize manifolds (ranks of LR & HT matrices) is an open question
  - what are “optimal” possibly semi-adaptive sampling schemes, e.g. jitter sampling improves the completion
- approach supposedly only works for low-frequencies …
  - success with HSS type decompositions
- matrix or tensor approaches have not yet been compared
- scale up to large-scale problems pretty open question

How to exploit low-rank structure in the high-frequency regime?

* “A fast solver for HSS representations via sparse matrices”, Chandrasekaran, S., et. al. , SIAM, 2006

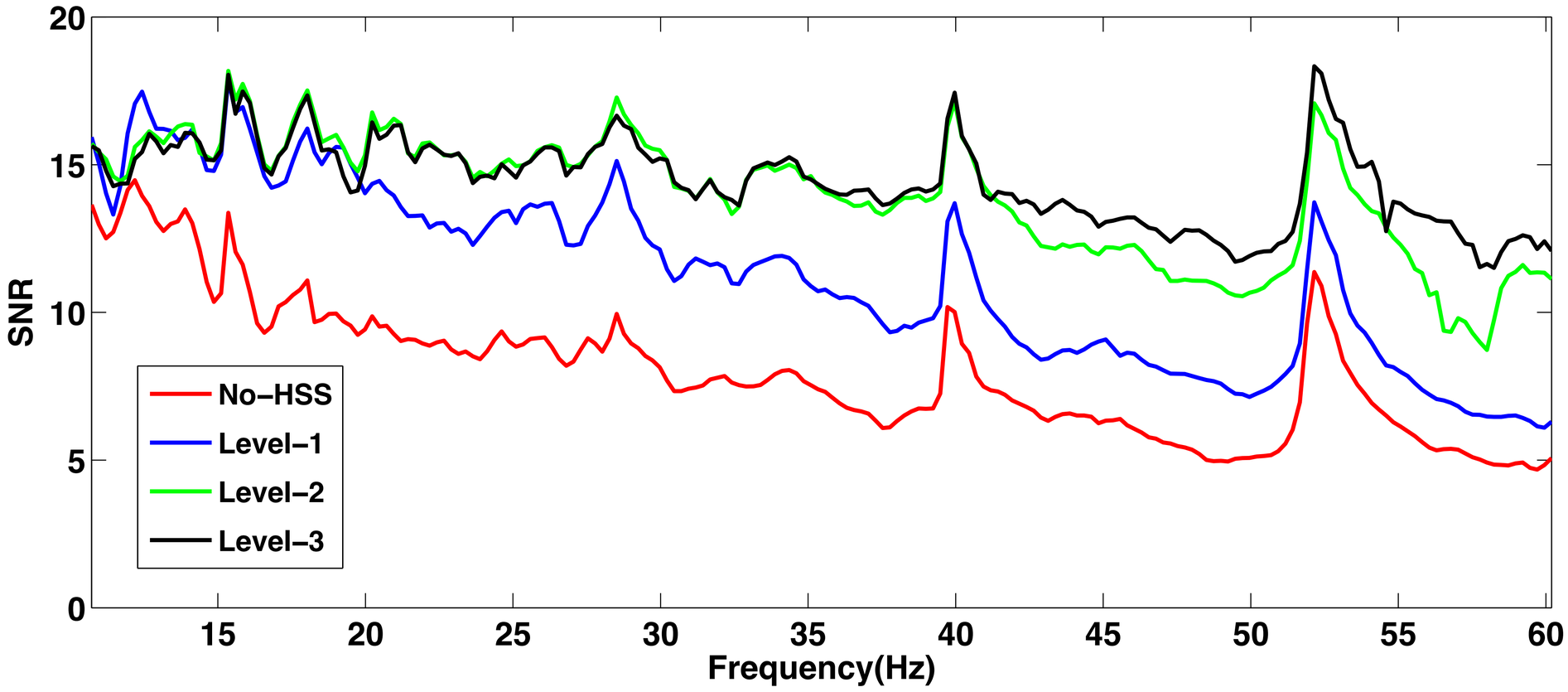
HSS representation





Low-rank interpolation

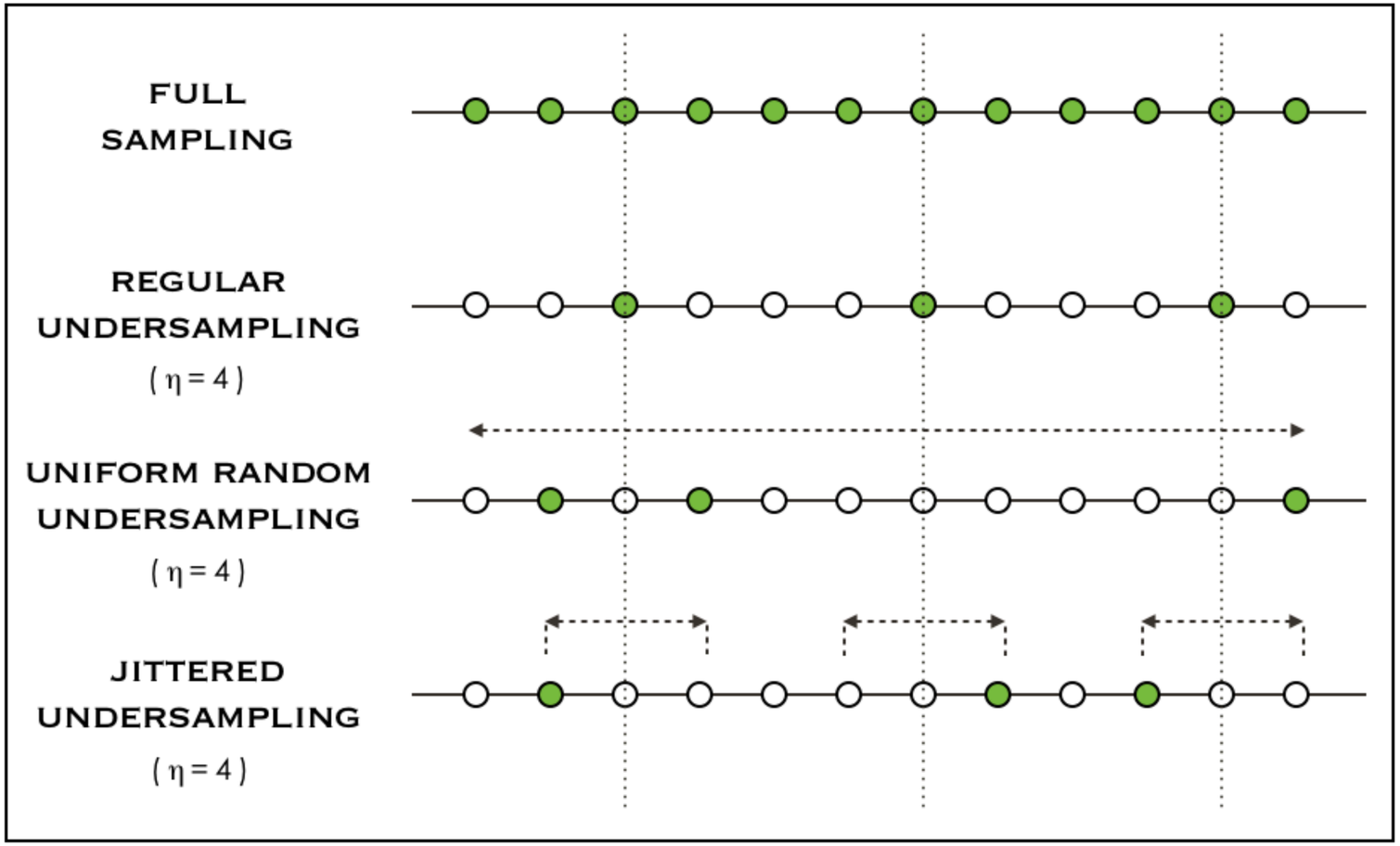
w/o and w/ HSS



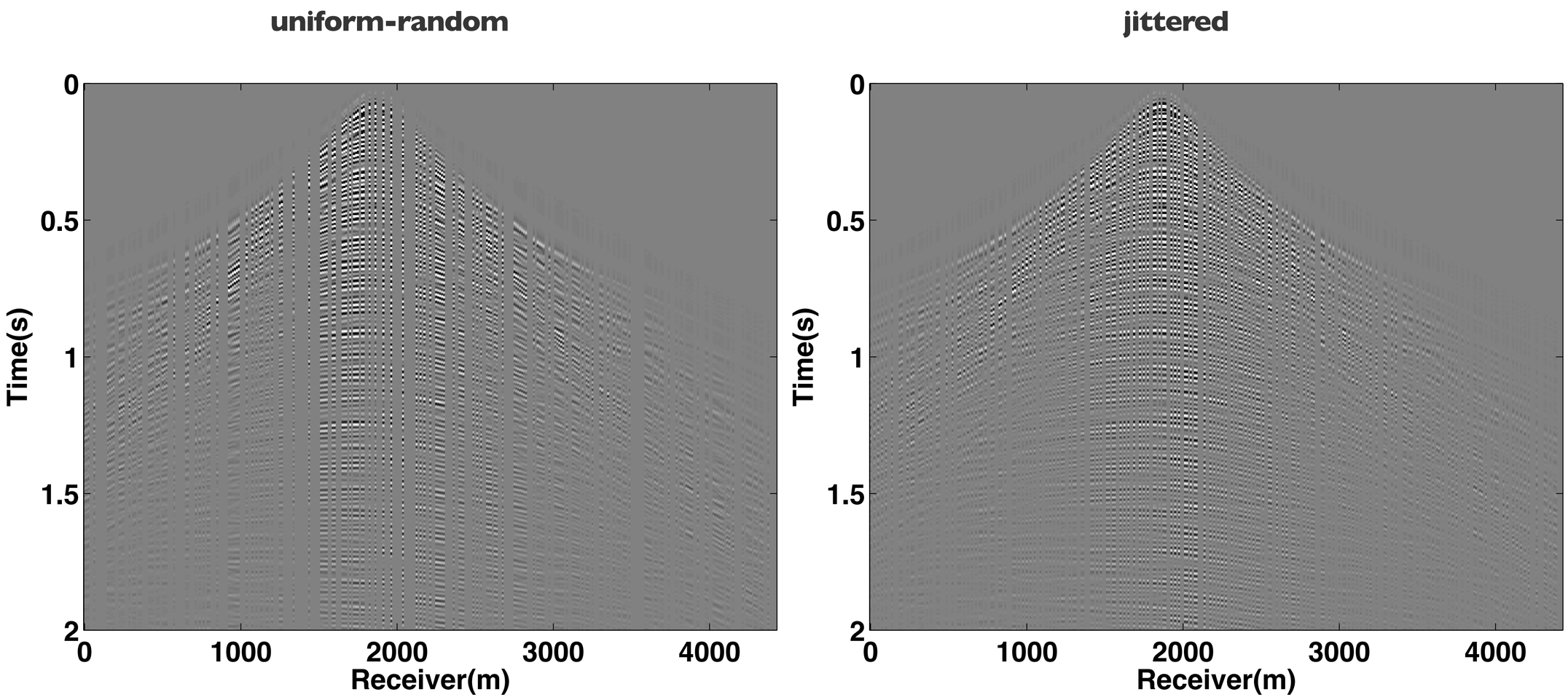
What should be the “optimal” sampling?

* “Simply denoise: wavefield reconstruction via jittered undersampling”, Hennenfent, G., Herrmann, F., Geophysics, 2008

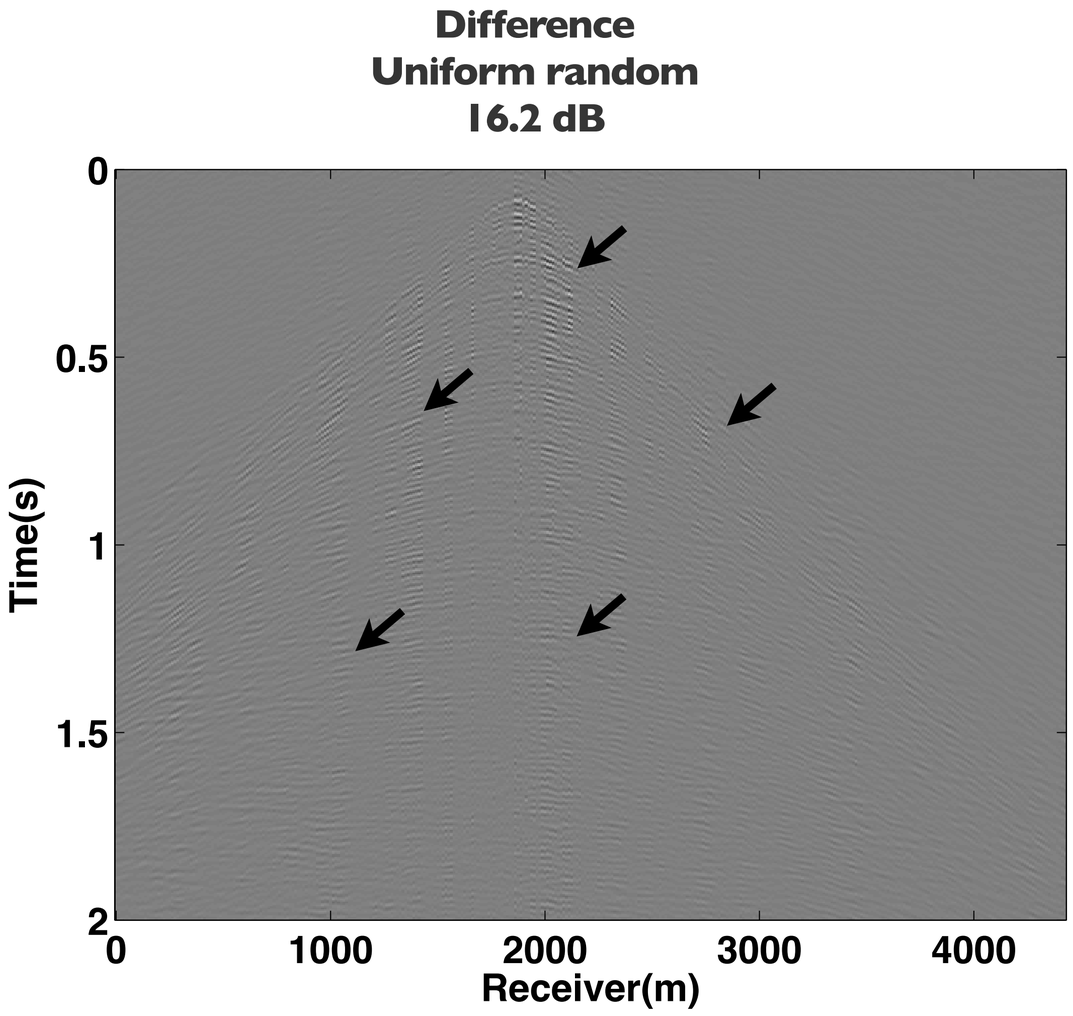
Sampling scheme

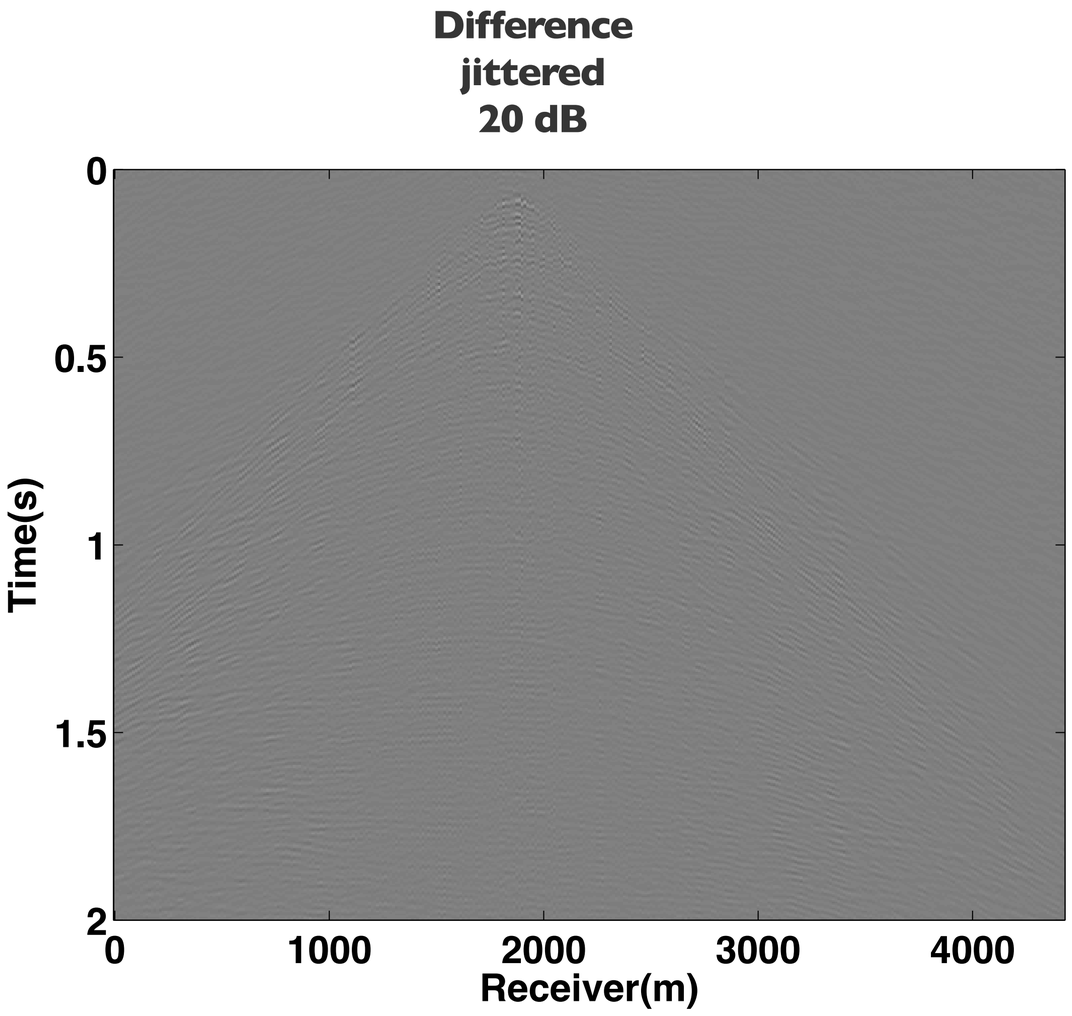


Uniform random v/s Jittered



Uniform-random v/s Jittered





→ areas with large gap

Papers + software

Available at www.slim.eos.ubc.ca under 'Publications'

"Fast methods for denoising matrix completion formulations, with application to robust seismic data interpolation," Aravkin et al.

"Efficient matrix completion for seismic data reconstruction," Kumar et al.

"Optimization on the Hierarchical Tucker manifold - applications to tensor completion," Da Silva and Herrmann.

"Source separation via SVD-free rank minimization in the hierarchical semi-separable representation," Wason, Kumar, et al.

Software: www.slim.eos.ubc.ca/software-documentation under the 'Processing' heading for LR matrix completion, HT tensor completion

Acknowledgements

Thank you for your attention

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, Subsalt Solutions, Total SA, WesternGeco, and Woodside.