Off the grid tensor completion for seismic data interpolation

Curt Da Silva and Felix J. Herrmann

June 3, 2015

SLIM
University of British Columbia

Motivation

3D seismic experiments - 5D data

- expensive to acquire, store
- sample at *sub-Nyquist* rates

Data exhibits *low-rank* structure

- exploit structure for interpolation

Fully sampled data

- simultaneous sources in wave-equation based inversion
- mitigating multiples

Motivation

Standard matrix, tensor completion

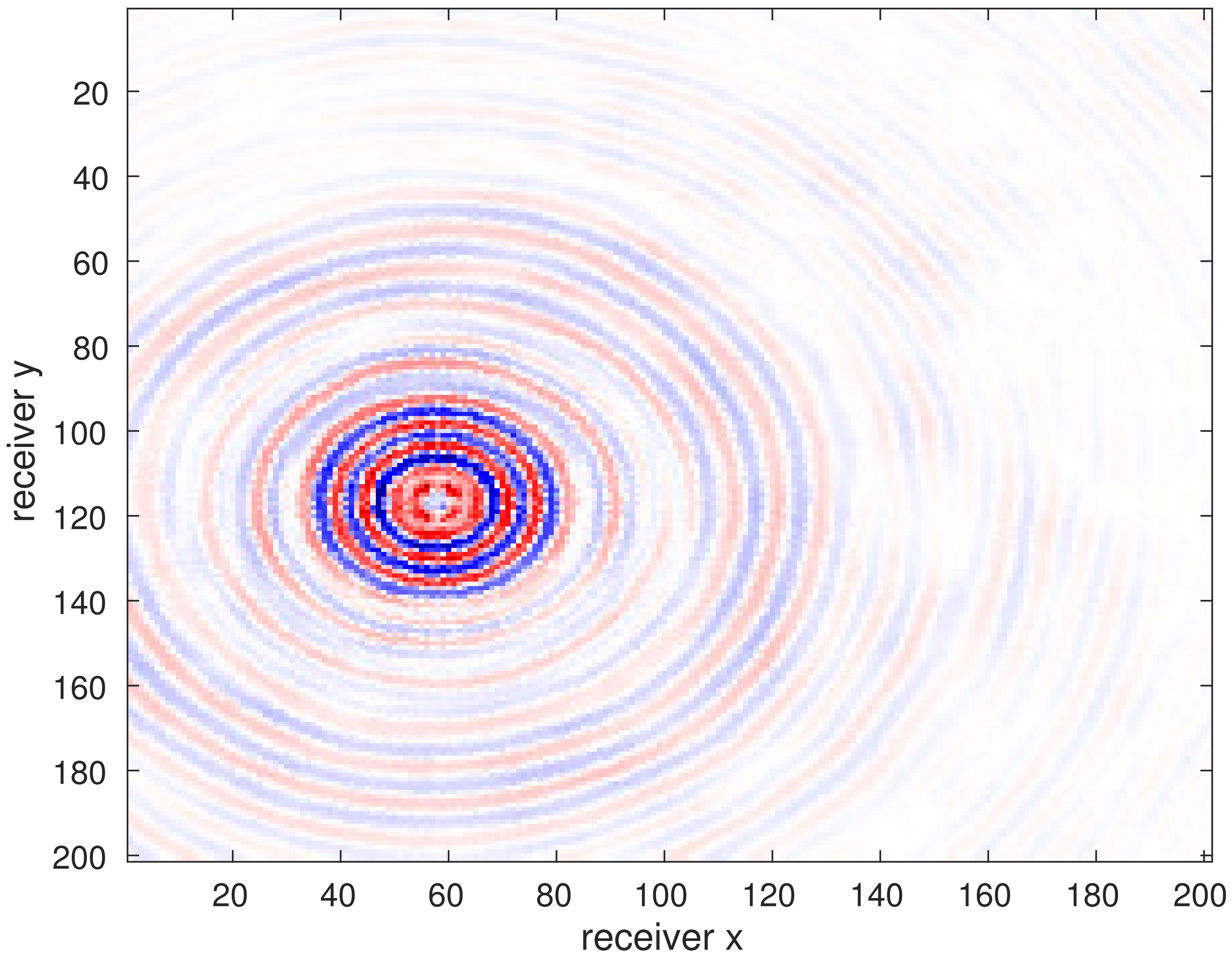
- defined on a regularly spaced grid
- sampling removes points from this regular grid

Degraded results in practice

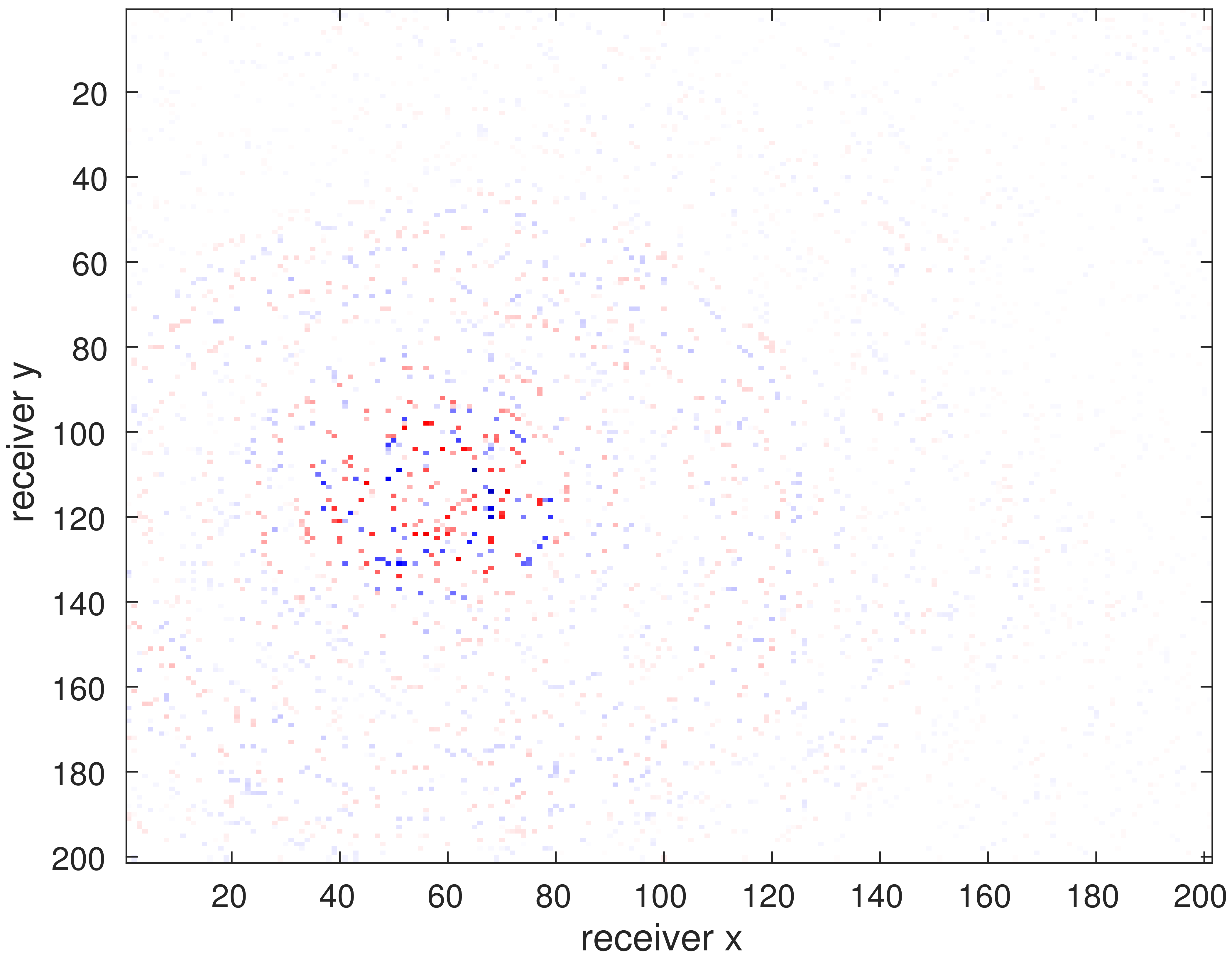
- very difficult to sample regularly in practice
- irregular “off-the-grid” sampling destroys low-rank behaviour

7.34 Hz - 90% missing receivers - irregular sampling

Common source gather



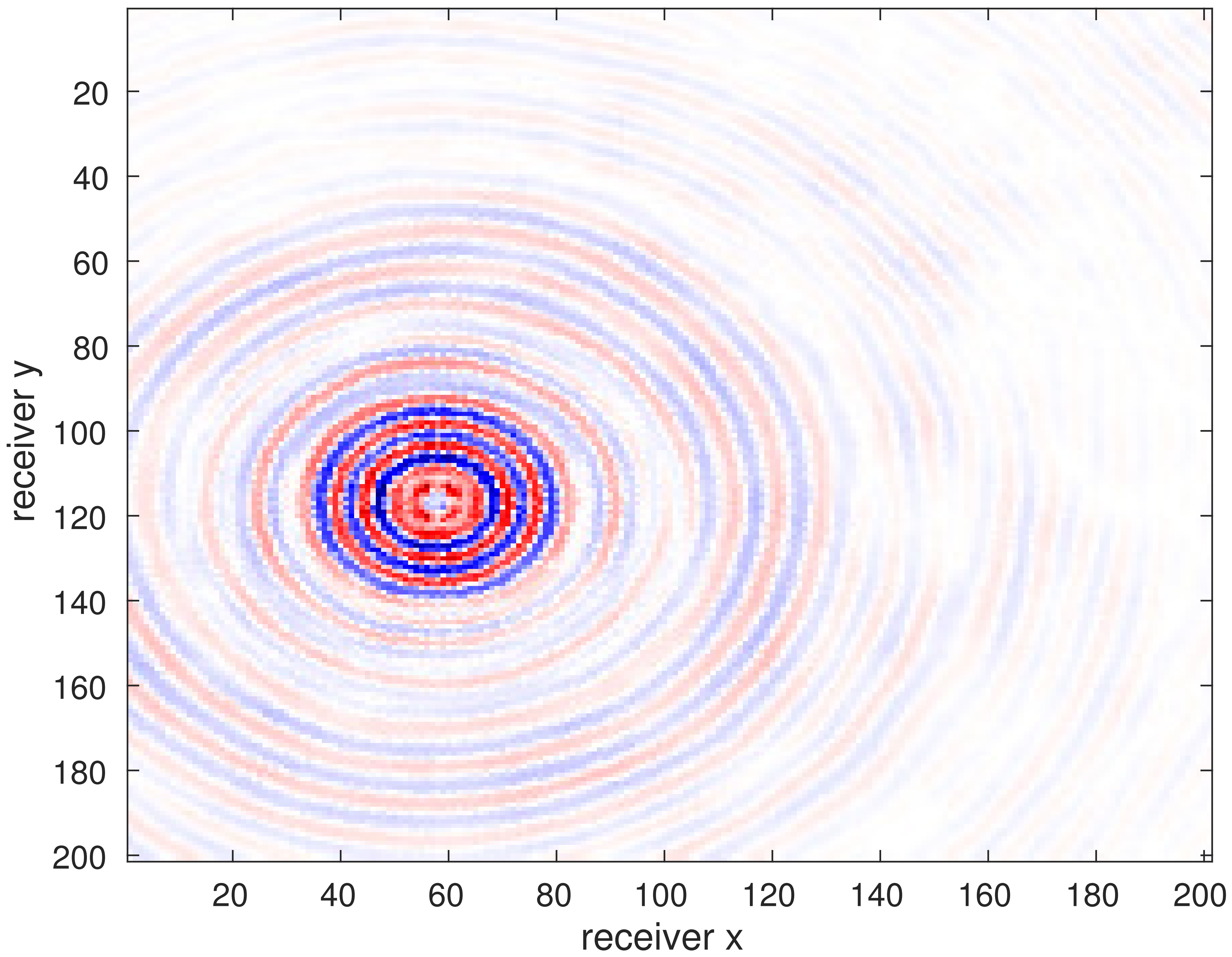
True data



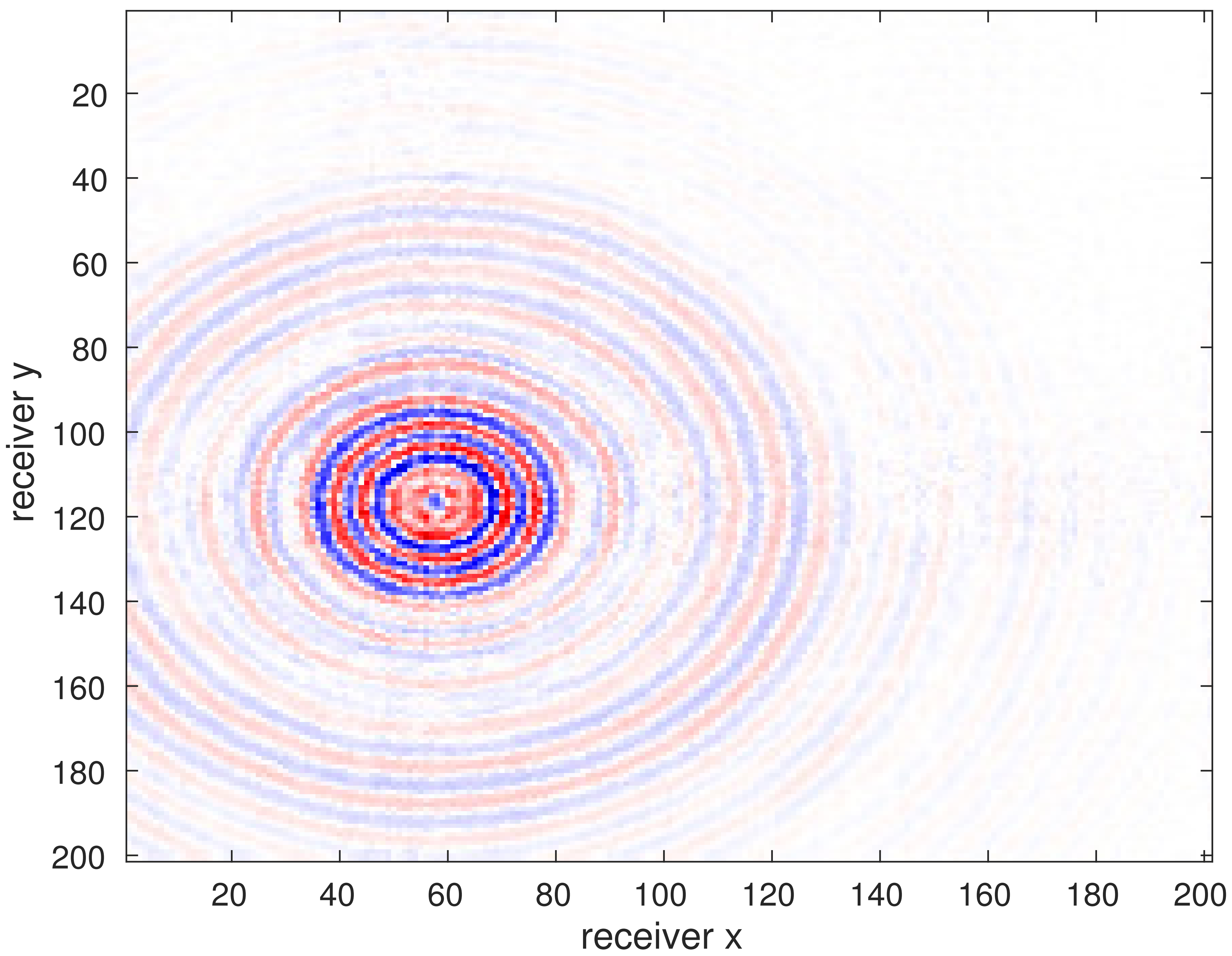
Subsampled data

7.34 Hz - 90% missing receivers - irregular sampling

Common source gather



True data



Regularized tensor completion

SNR 11 dB

[1] Kreimer and Sacchi, "A tensor higher-order singular value decomposition for prestack seismic data noise reduction and interpolation." (2012)
[2] Gao, Vicente, and Sacchi. "Evaluation of a fast algorithm for the eigen-decomposition of large block Toeplitz matrices with application to 5D seismic data interpolation." *(2011)*
[3] Da Silva, Kumar, et al, "Efficient matrix completion for seismic data reconstruction." (To appear)

SLIM

Context

Low-rank matrix/tensor completion via *nuclear norm* projection [1]

- Require SVDs on huge data matrices
- Not scalable to large problem sizes

Data completion via Toeplitz embedding [2]

- Problem size - (# data points)^2
- Ad-hoc windowing - can degrade quality, as demonstrated in [3]

Goals

Review Hierarchical Tucker tensor format, principles of low-rank tensor recovery

Compensate for lack of low-rank behaviour in irregular sampling domain

- Introduce a new domain in which the data is low rank

Compressive sensing
with sparsity promotion

Successful reconstruction scheme

Signal structure

- sparsity

Sampling

- subsampling decreases sparsity

Optimization

- look for sparsest solution

Multidimensional interpolation
with Hierarchical Tucker

Successful reconstruction scheme

Signal structure

- ***Hierarchical Tucker***

Sampling

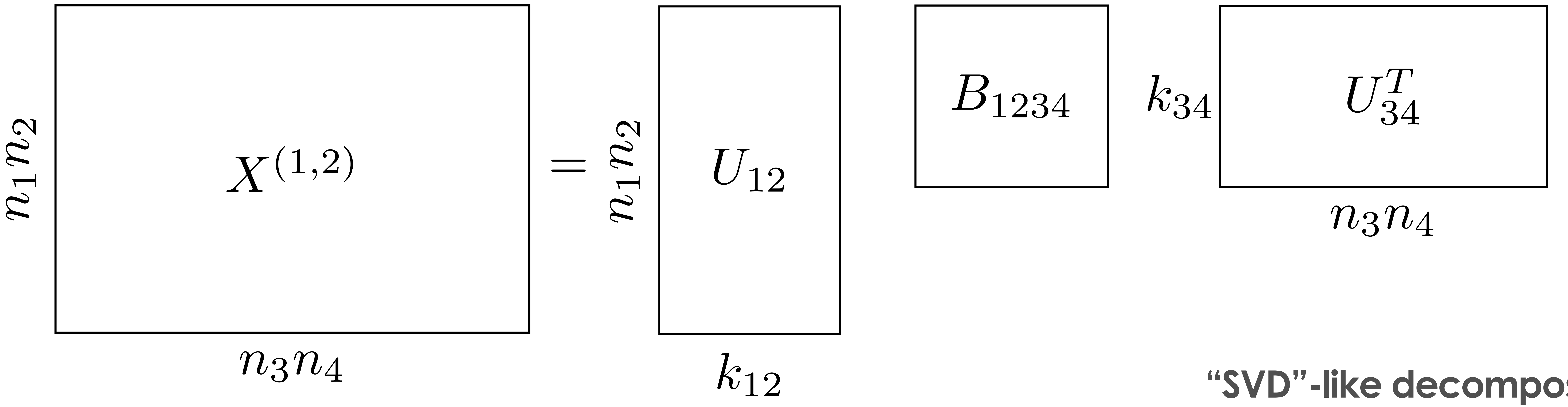
- subsampling, noise increases hierarchical rank

Optimization

- fit data in the Hierarchical Tucker format

Hierarchical Tucker format

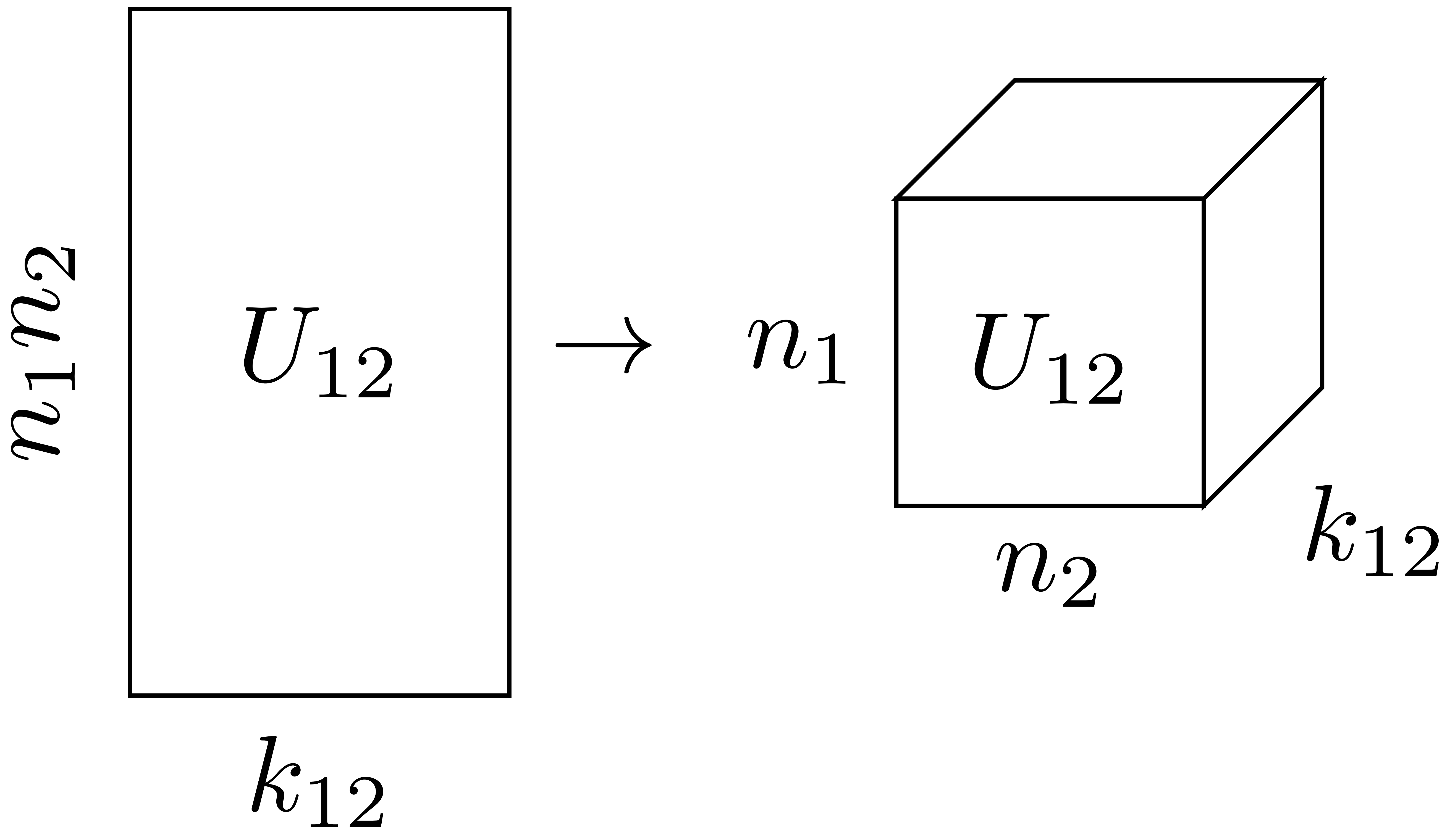
$X - n_1 \times n_2 \times n_3 \times n_4$ tensor



"SVD"-like decomposition

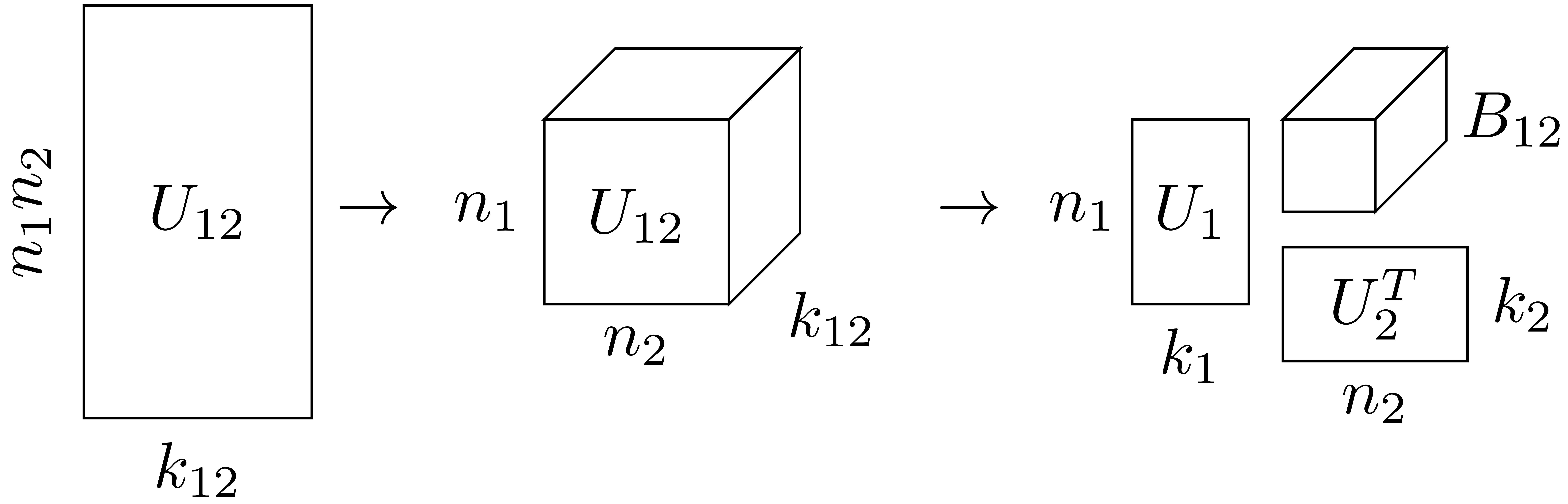
Hierarchical Tucker format

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

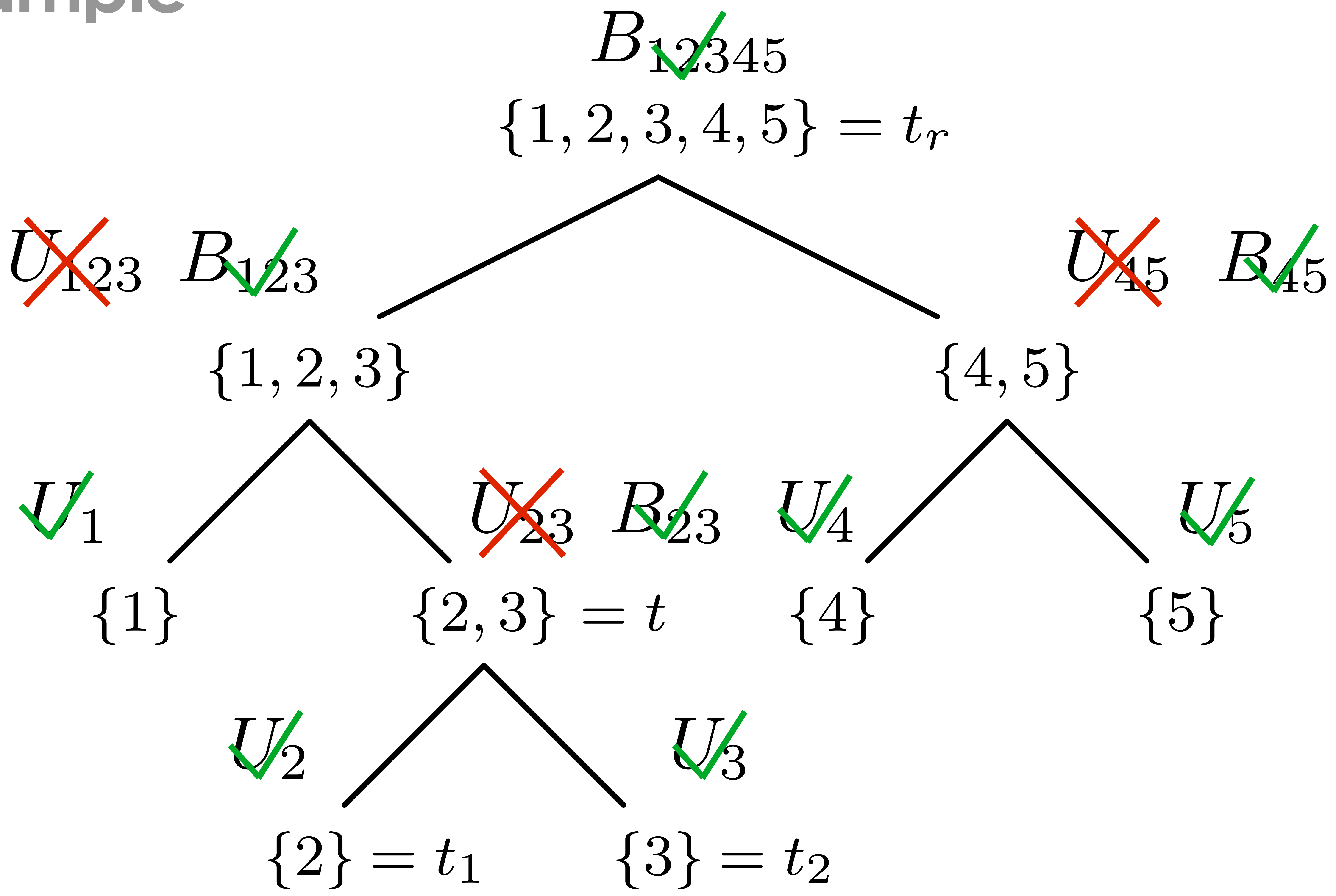
Separating groups of dimensions from each other

- dimension tree

A. Uschmajew, B. Vandereycken. The geometry of algorithms using hierarchical tensors. Linear algebra and its applications, 2013.

SLIM

Example



Hierarchical Tucker format

Storage $\leq dNK + (d-2)K^3 + K^2$

Compare to N^d storage for the full tensor

Effectively breaking the curse of dimensionality when $K \ll N \quad d \geq 4$

Low frequency data compresses in HT

Seismic Hierarchical Tucker

We consider a 3D seismic survey with coordinates
(src x, src y, rec x, rec y, time)

We take a Fourier transform in time and restrict ourselves to a single frequency slice

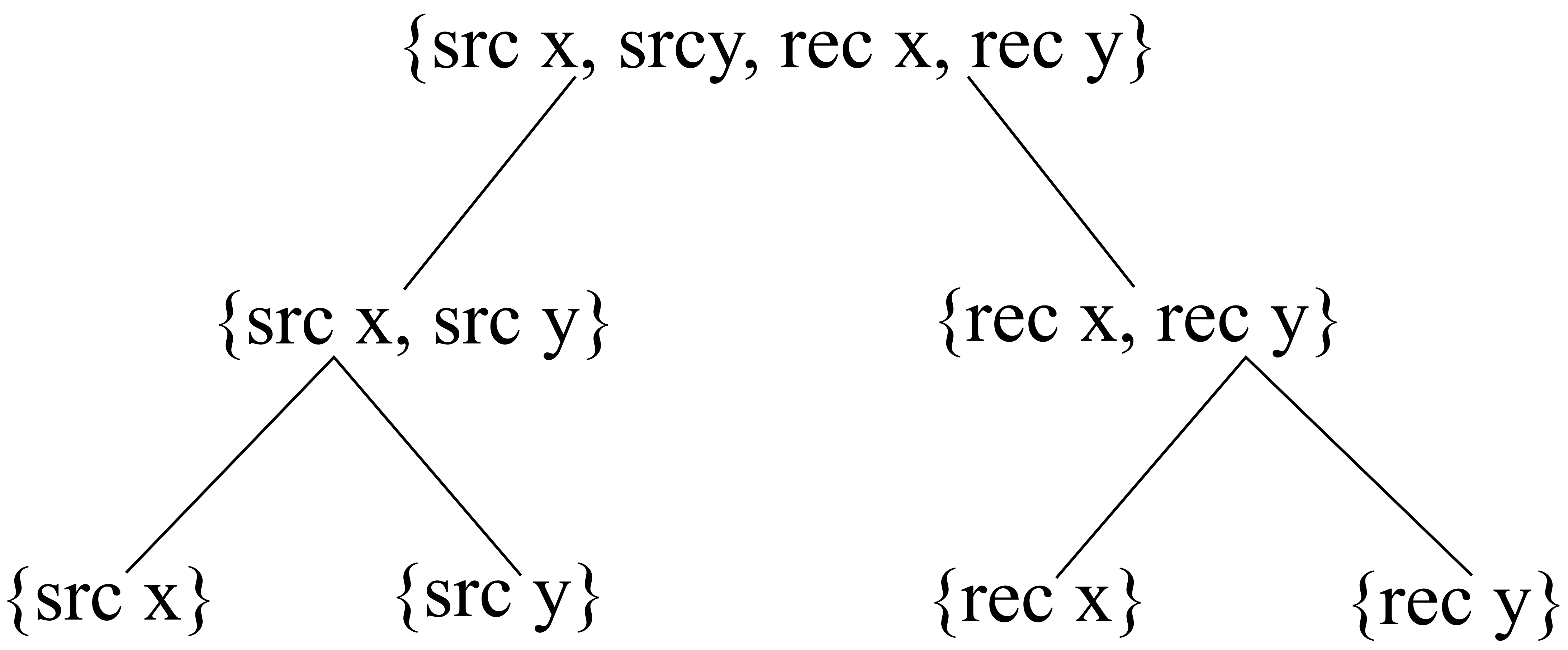
[1] L. Demanet. Curvelets, Wave Atoms, and Wave Equations. PhD Thesis, MIT

Seismic Hierarchical Tucker

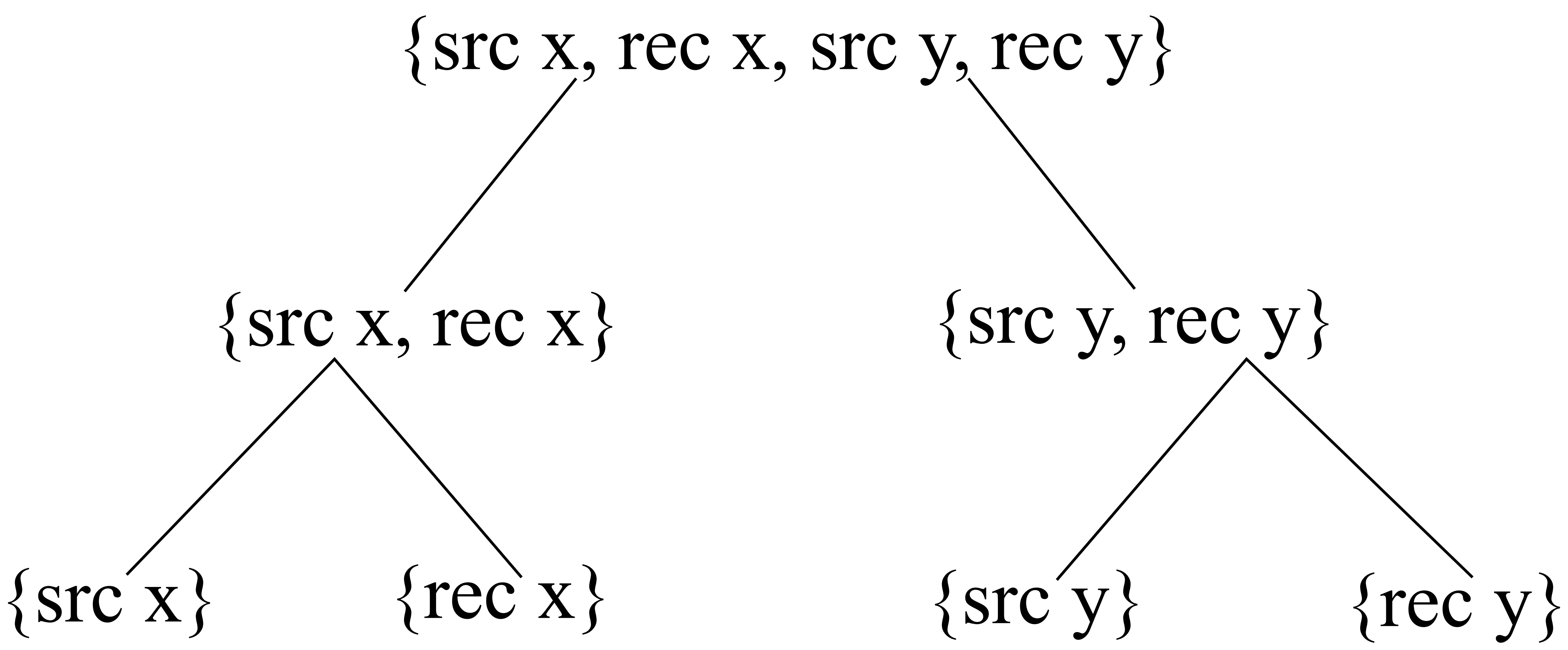
For a frequency slice with coordinates (src x, src y, rec x, rec y),

there are essentially two choices of dimension splitting (by reciprocity)

First explored in [1] - low rank solution operators for wave equations

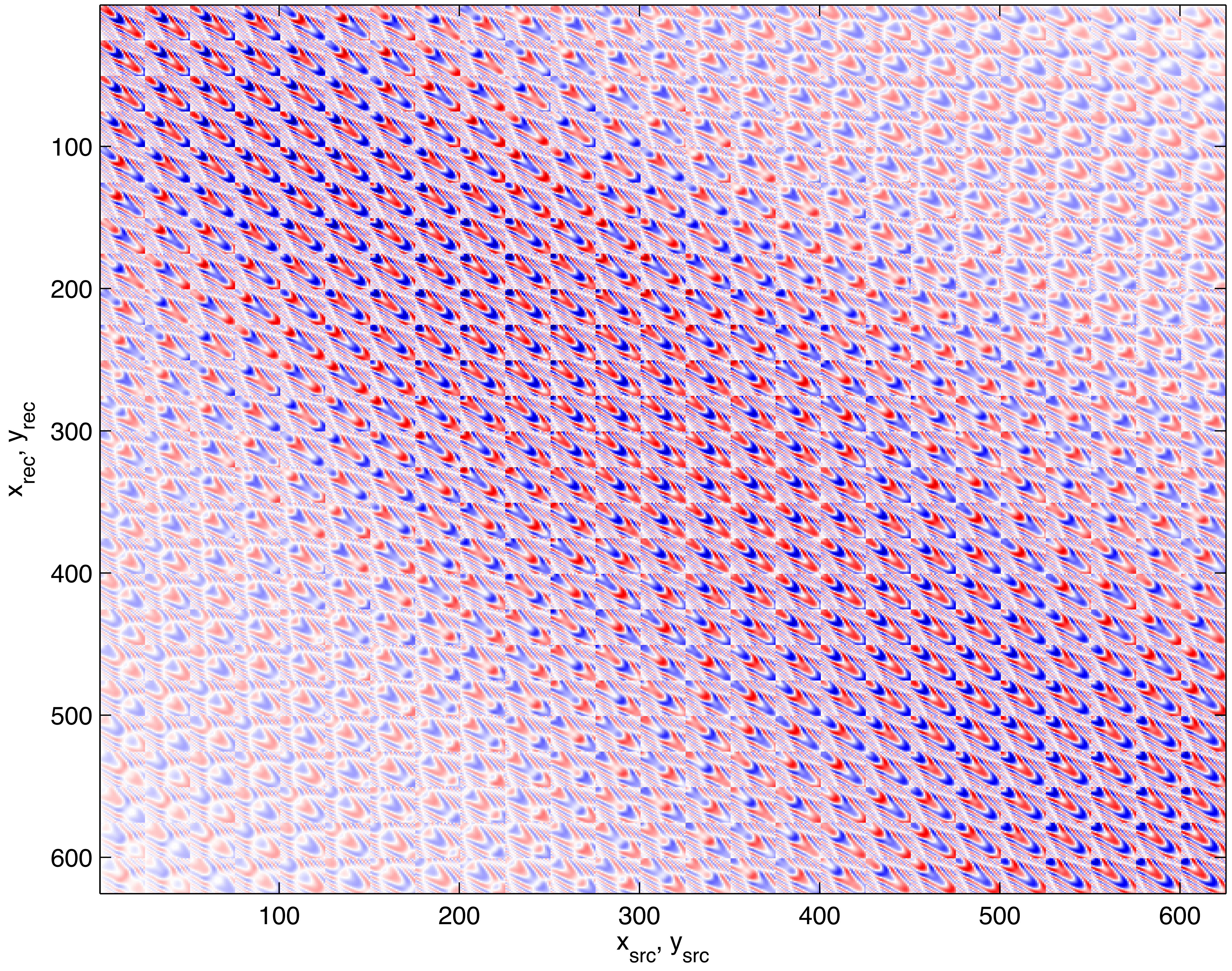


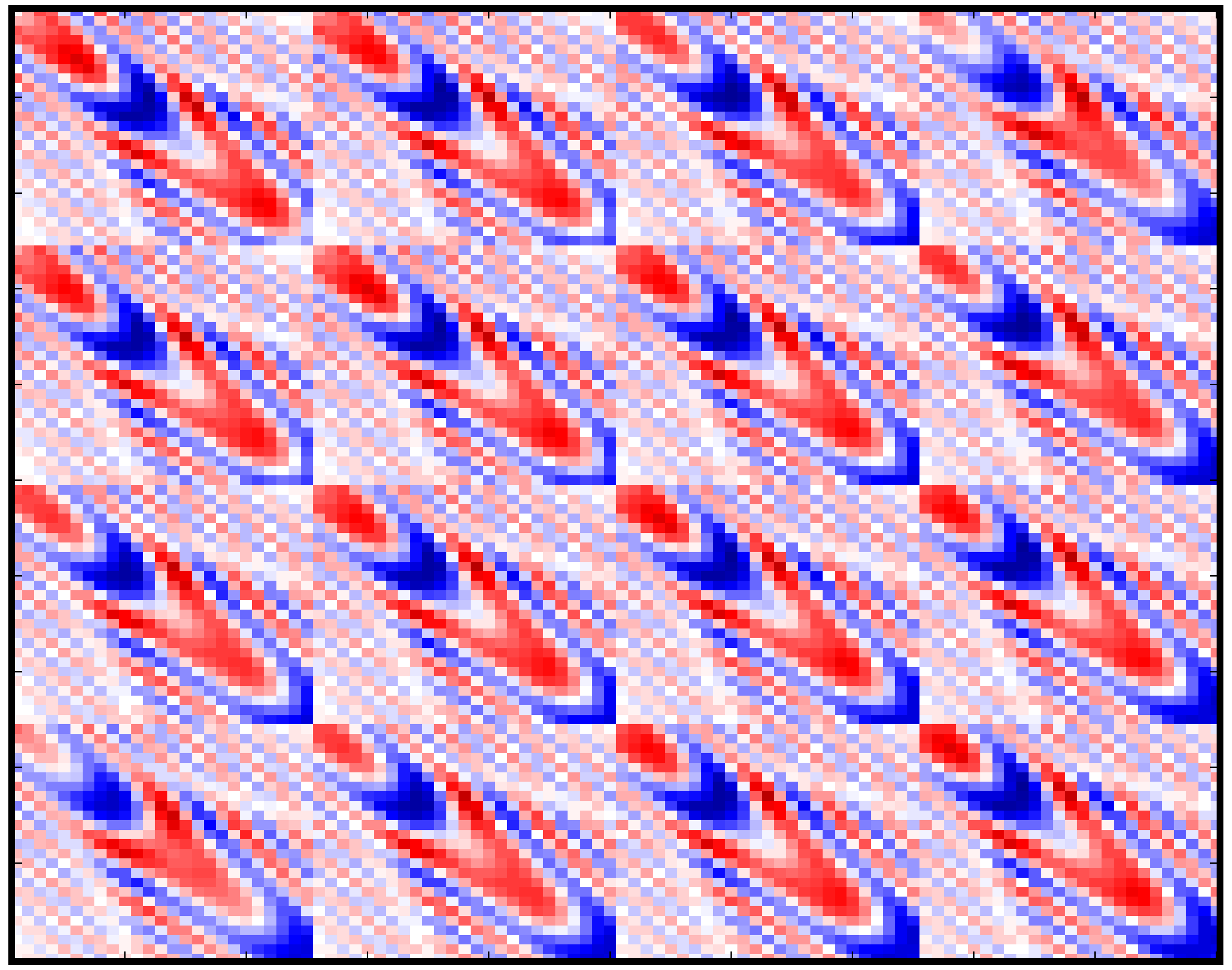
Canonical Decomposition

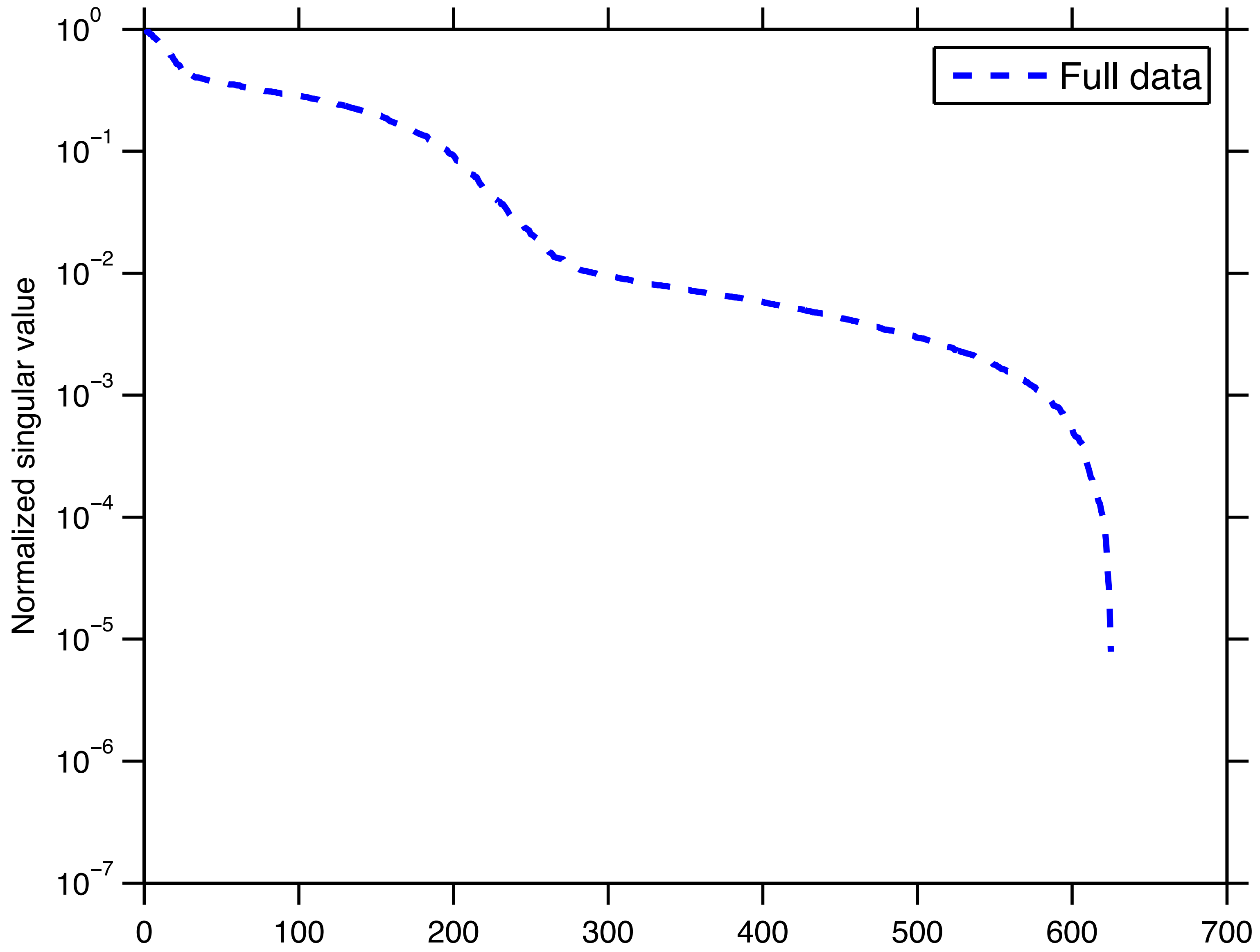


Non-canonical Decomposition

Matricizations

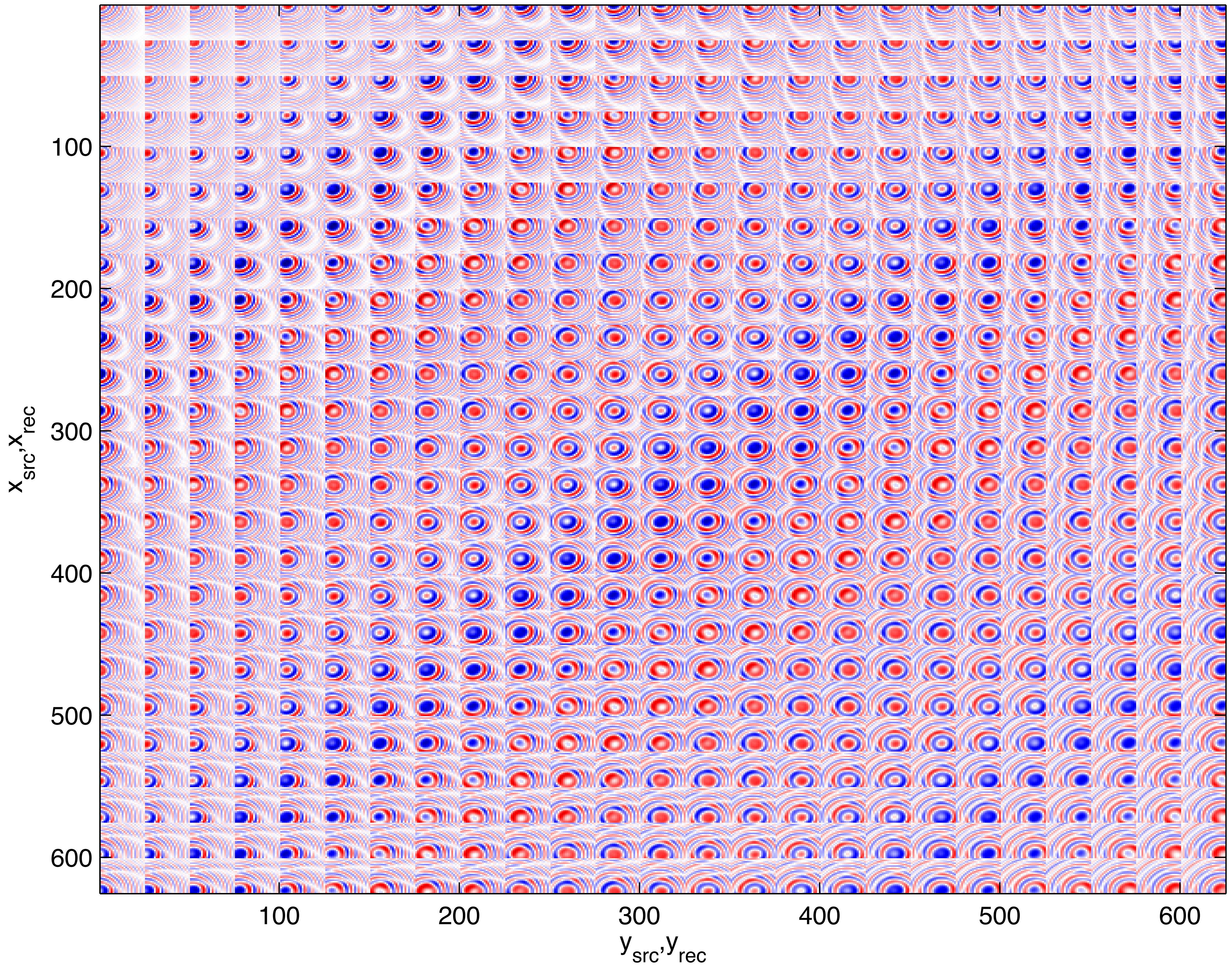


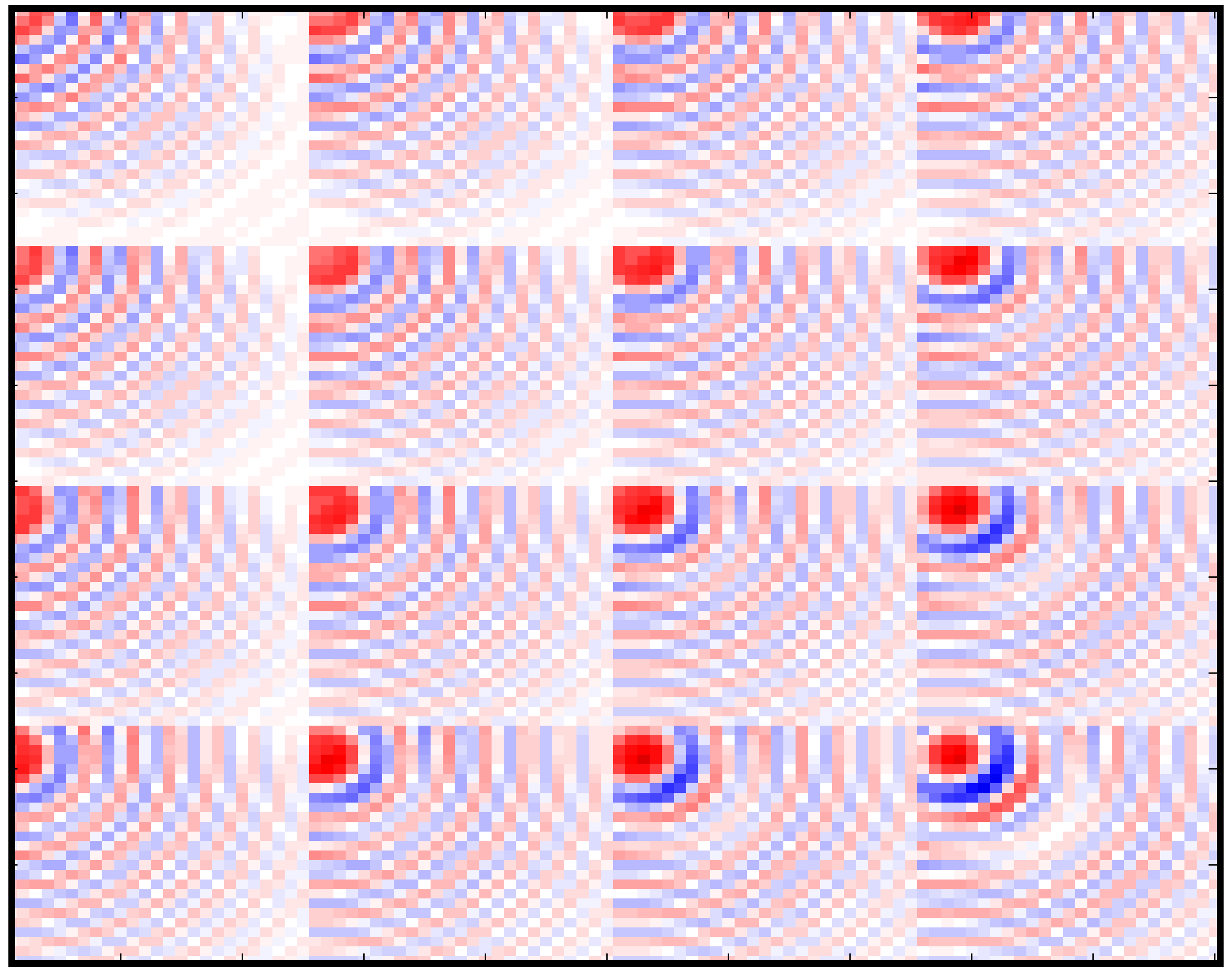


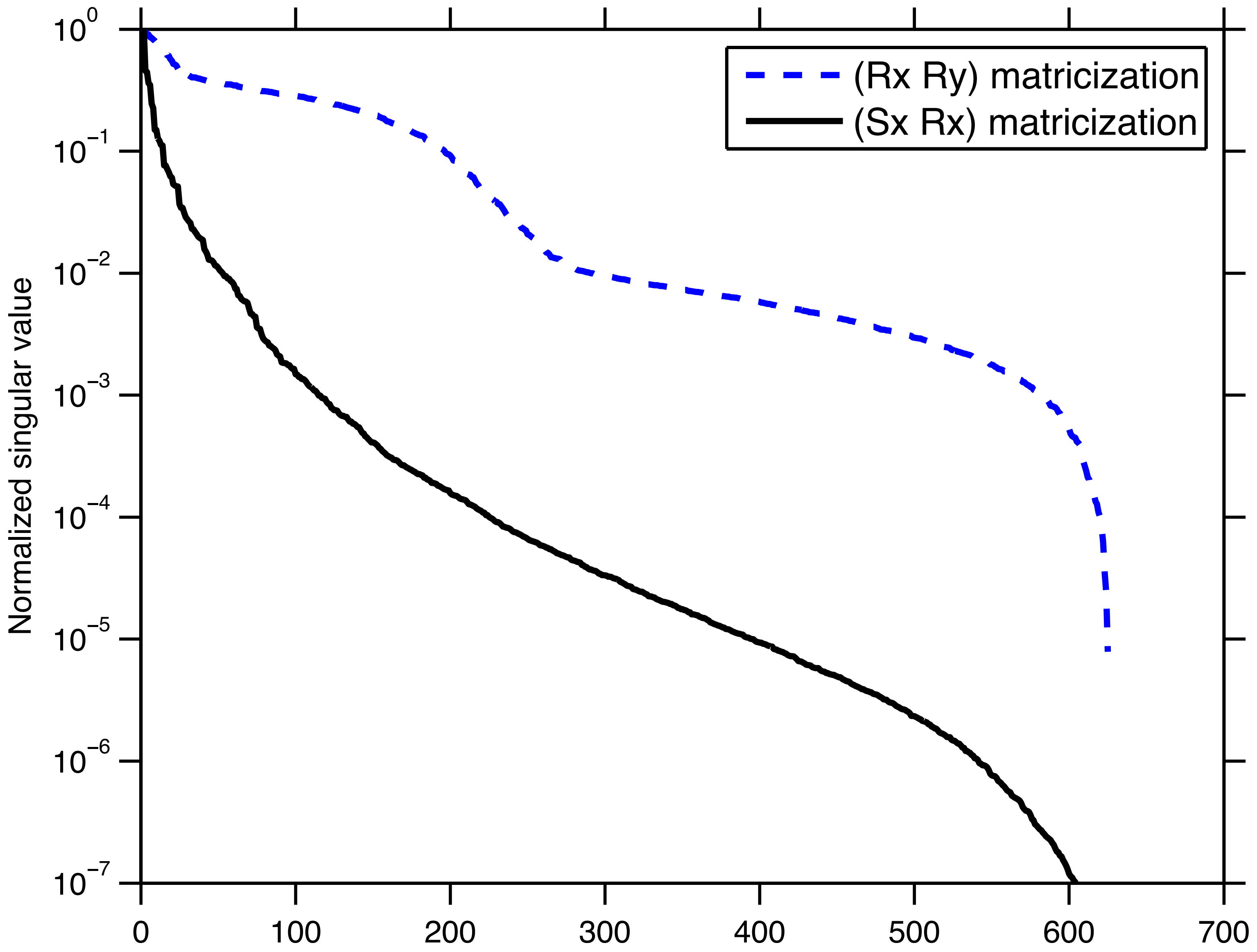


(Rec x, Rec y) matricization - Canonical ordering

Matricizations







(Src x, Rec x) matricization - Noncanonical ordering

Multidimensional interpolation

with Hierarchical Tucker

Successful reconstruction scheme

Signal structure

- Hierarchical Tucker

Sampling

- ***subsampling, noise increases hierarchical rank***

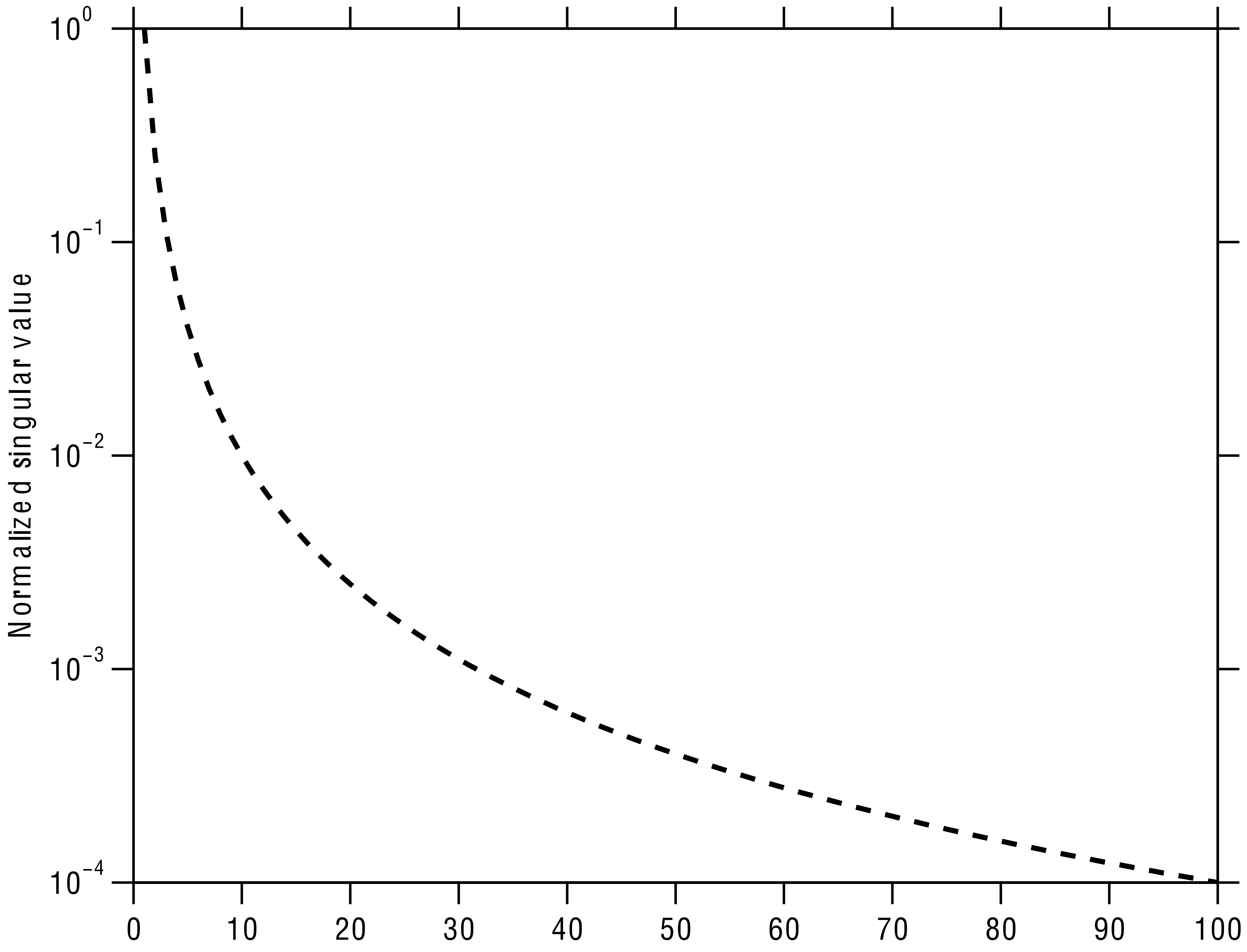
Optimization

- fit data in the Hierarchical Tucker format

Matrix Completion

$$\mathbf{X}$$

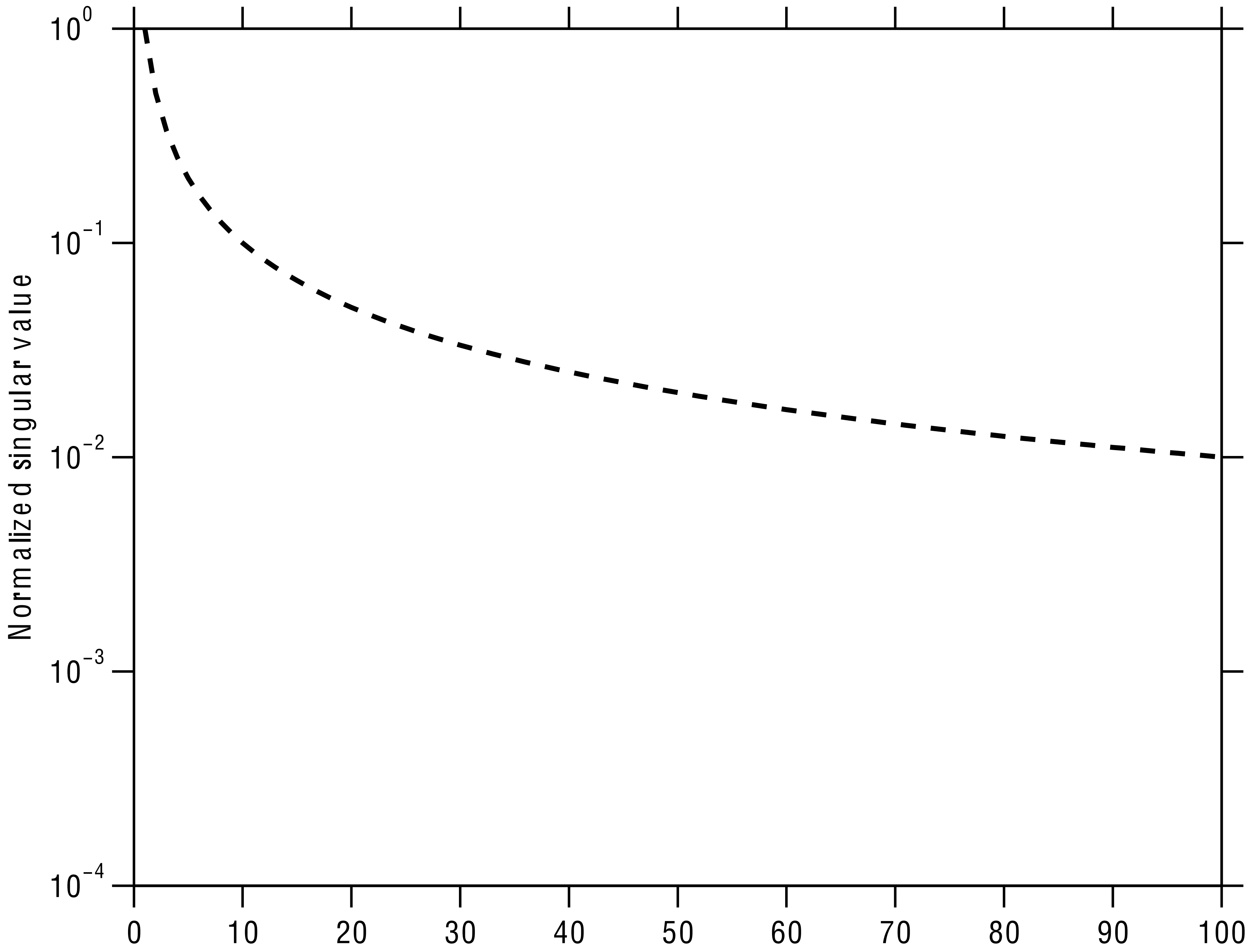
$$\begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}$$



Matrix Completion

$$\mathcal{A}(\mathbf{X})$$

$$\begin{bmatrix} * & * & * & 0 & * \\ * & 0 & 0 & * & 0 \\ * & * & * & * & * \\ * & * & 0 & * & * \\ 0 & * & * & * & 0 \end{bmatrix}$$



Sampling

Sampling $(x_{\text{src}}, y_{\text{src}}, x_{\text{rec}}, y_{\text{rec}})$ points from the data

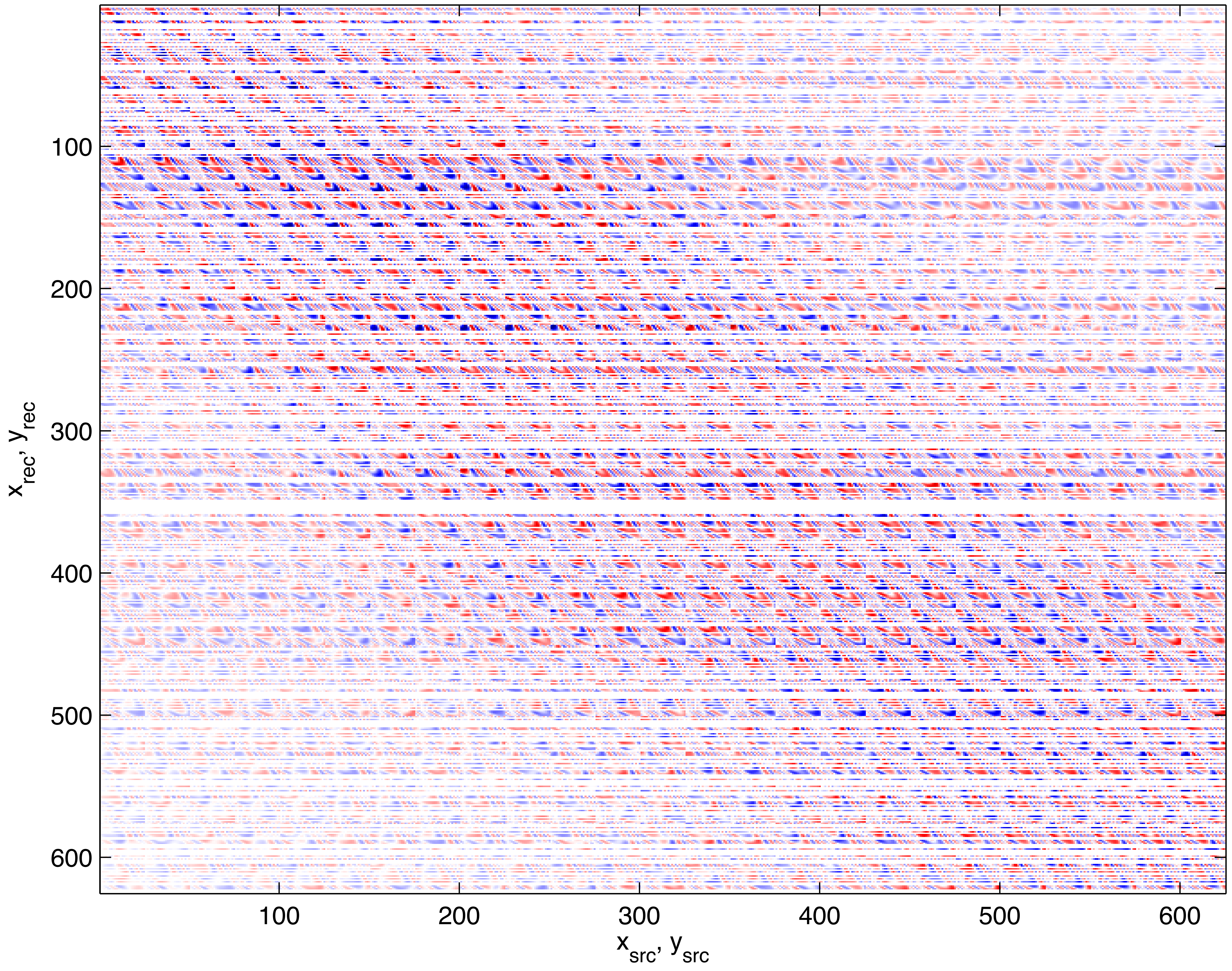
- idealized recovery
- impossible to physically implement

Sampling $(x_{\text{rec}}, y_{\text{rec}})$ points from the data

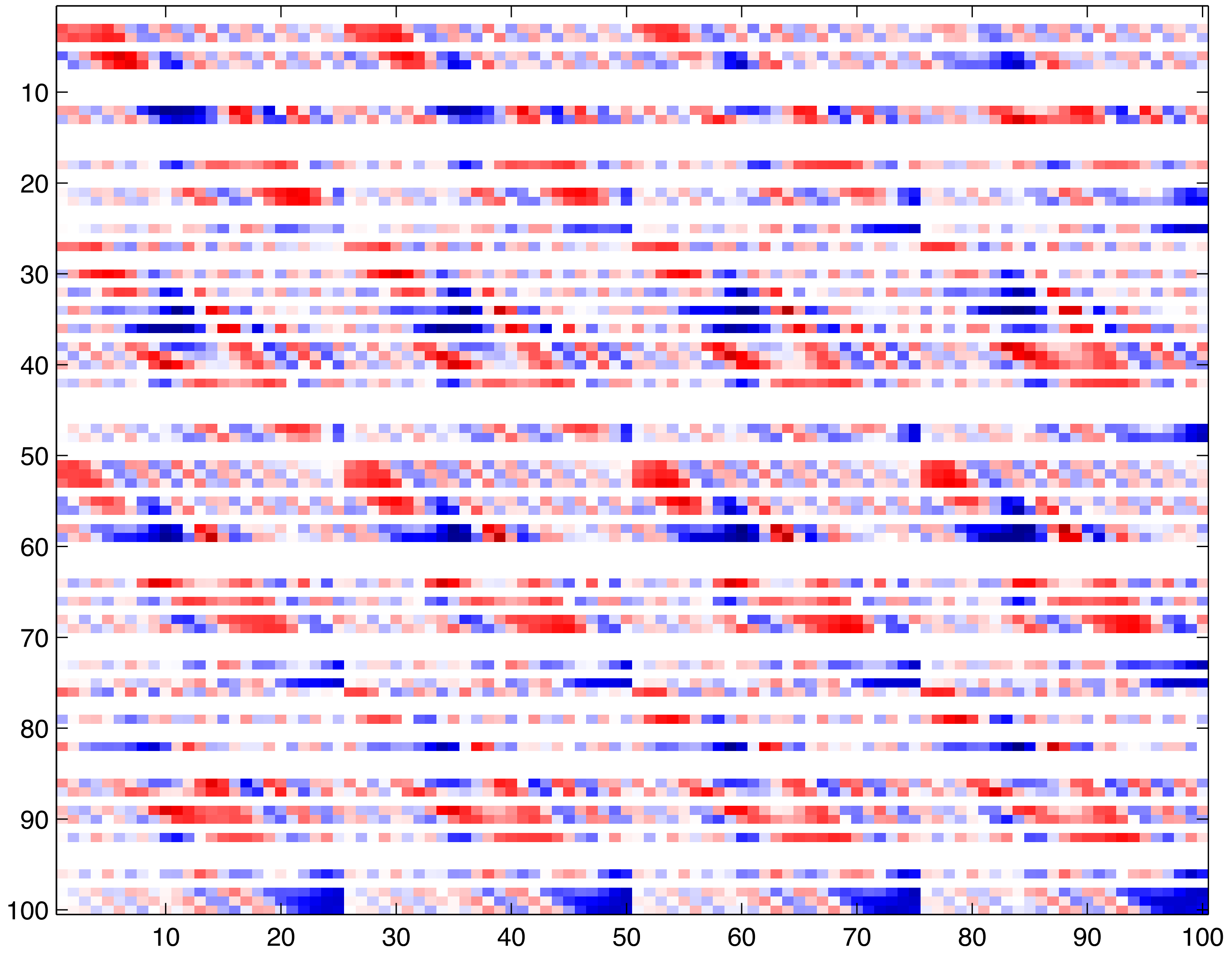
- less idealized
- possible to acquire data - e.g. ocean bottom nodes

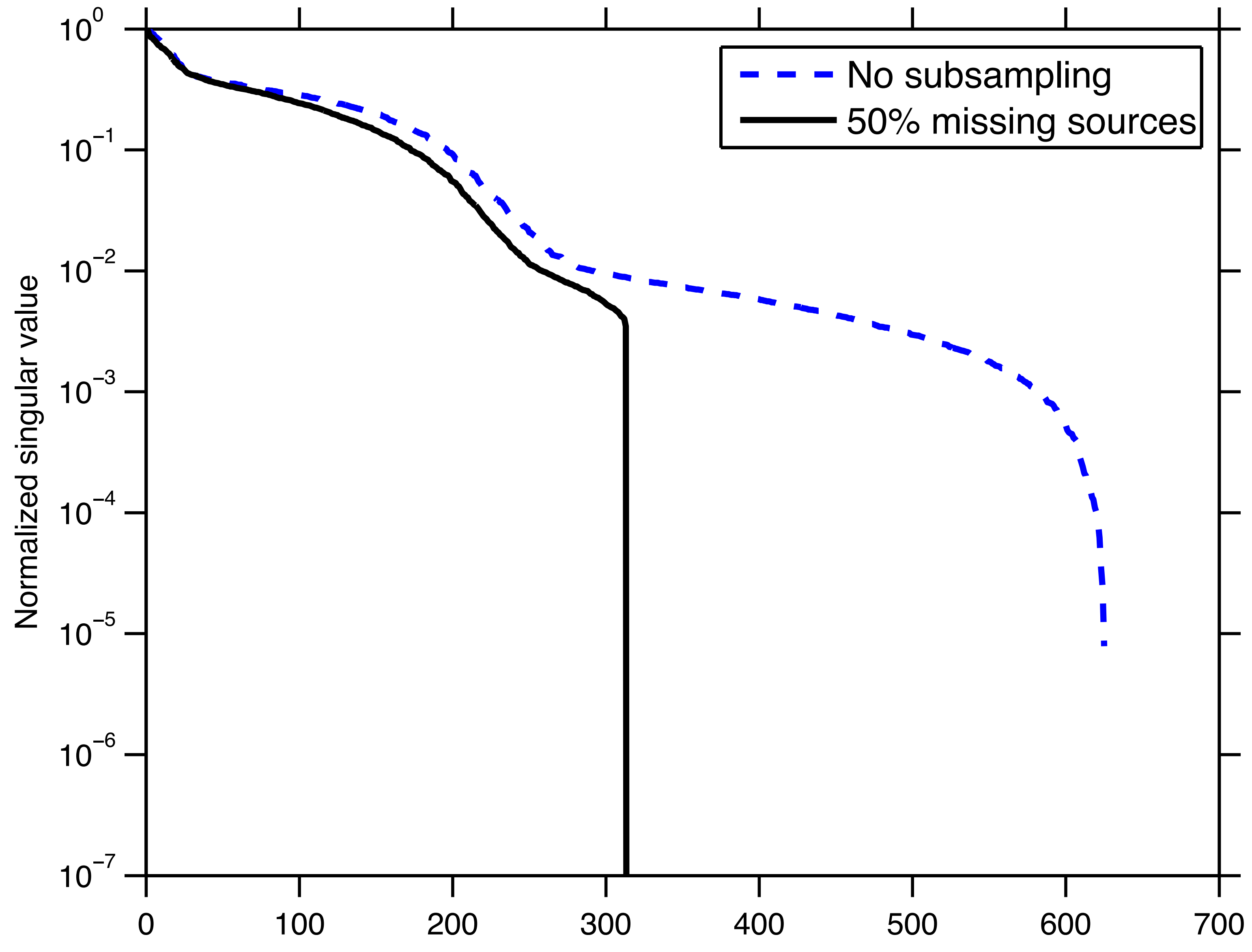
Realistic recovery

50% random receivers removed



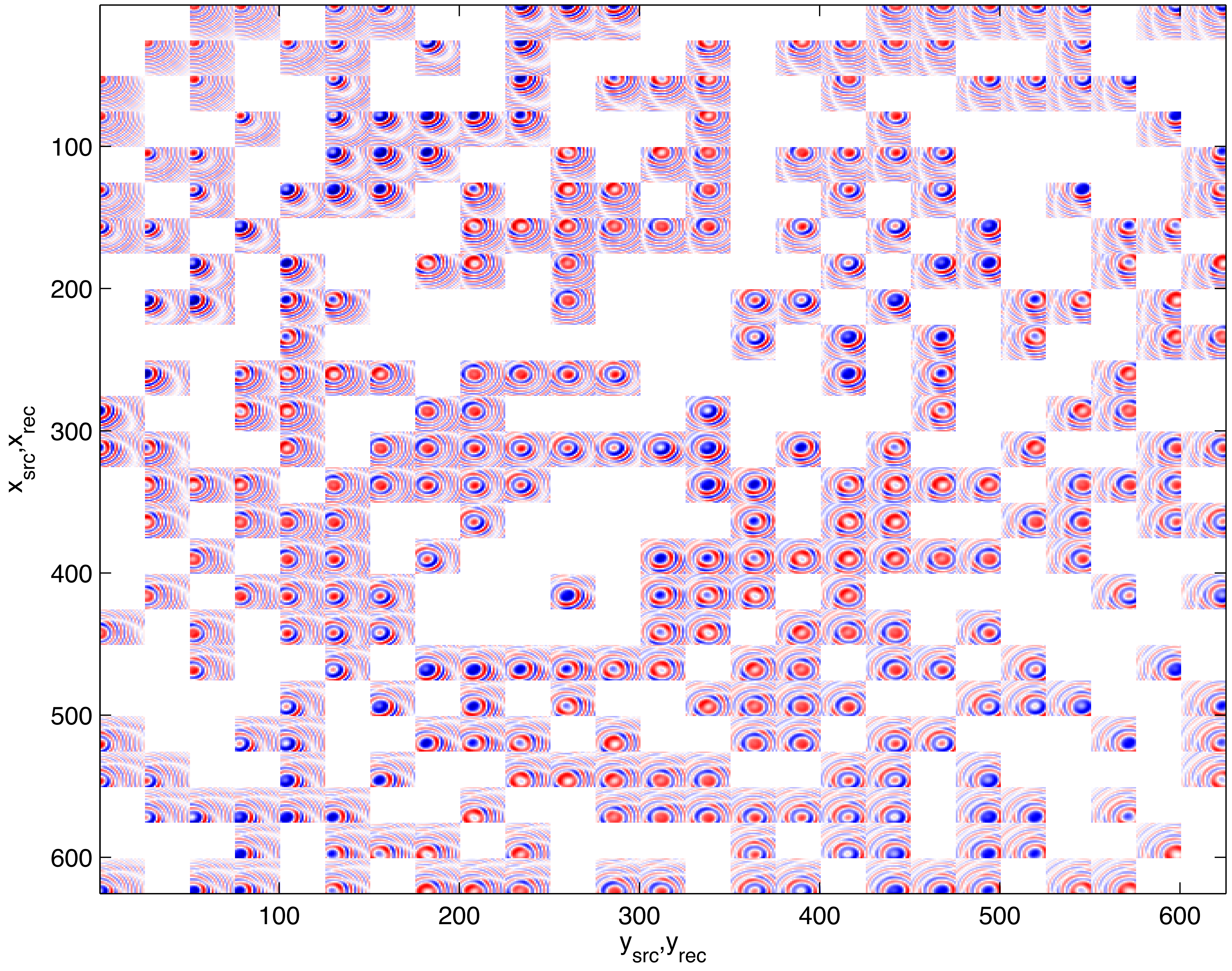
(Rec x, Rec y) matricization - Canonical ordering

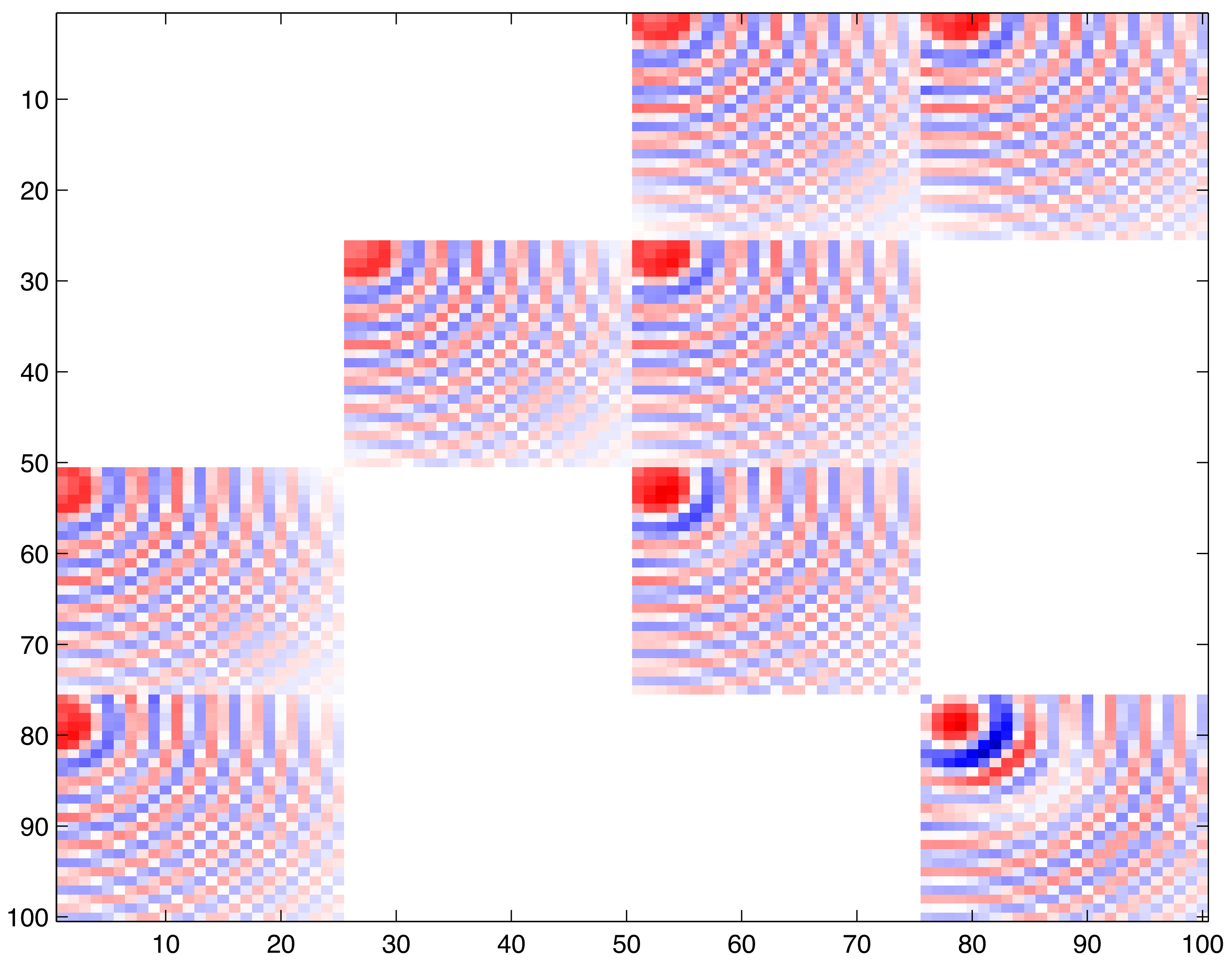


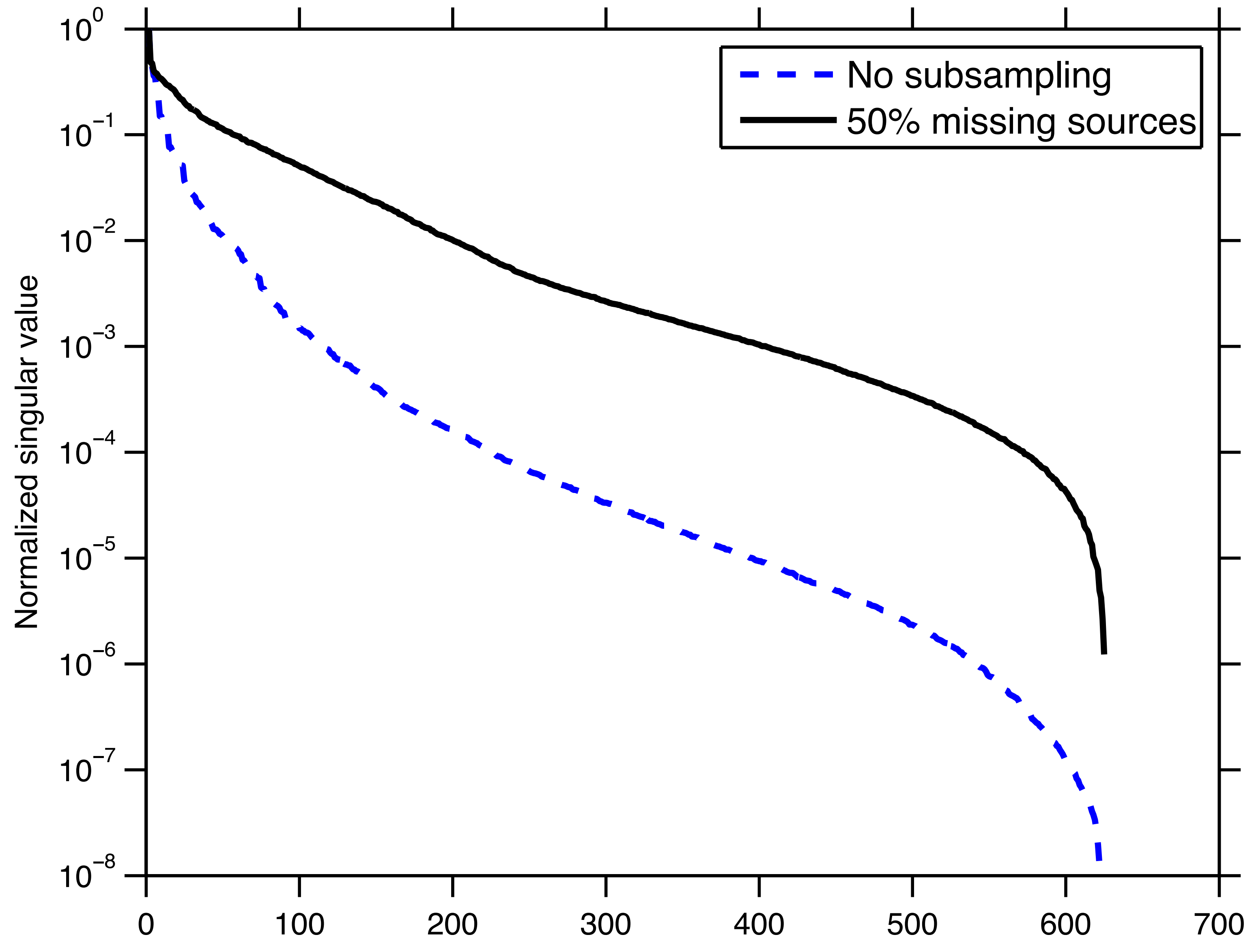


Realistic recovery

50% random receivers removed







(Src x, Rec x) matricization - Noncanonical ordering

Data organization

(rec x, rec y) organization

- High rank
- Missing receivers operator - removes rows
- Poor recovery scenario

(src x, rec x) organization

- Low rank
- Missing receivers operator - removes blocks
- Closer to ideal recovery scenario

Nonuniform sampling

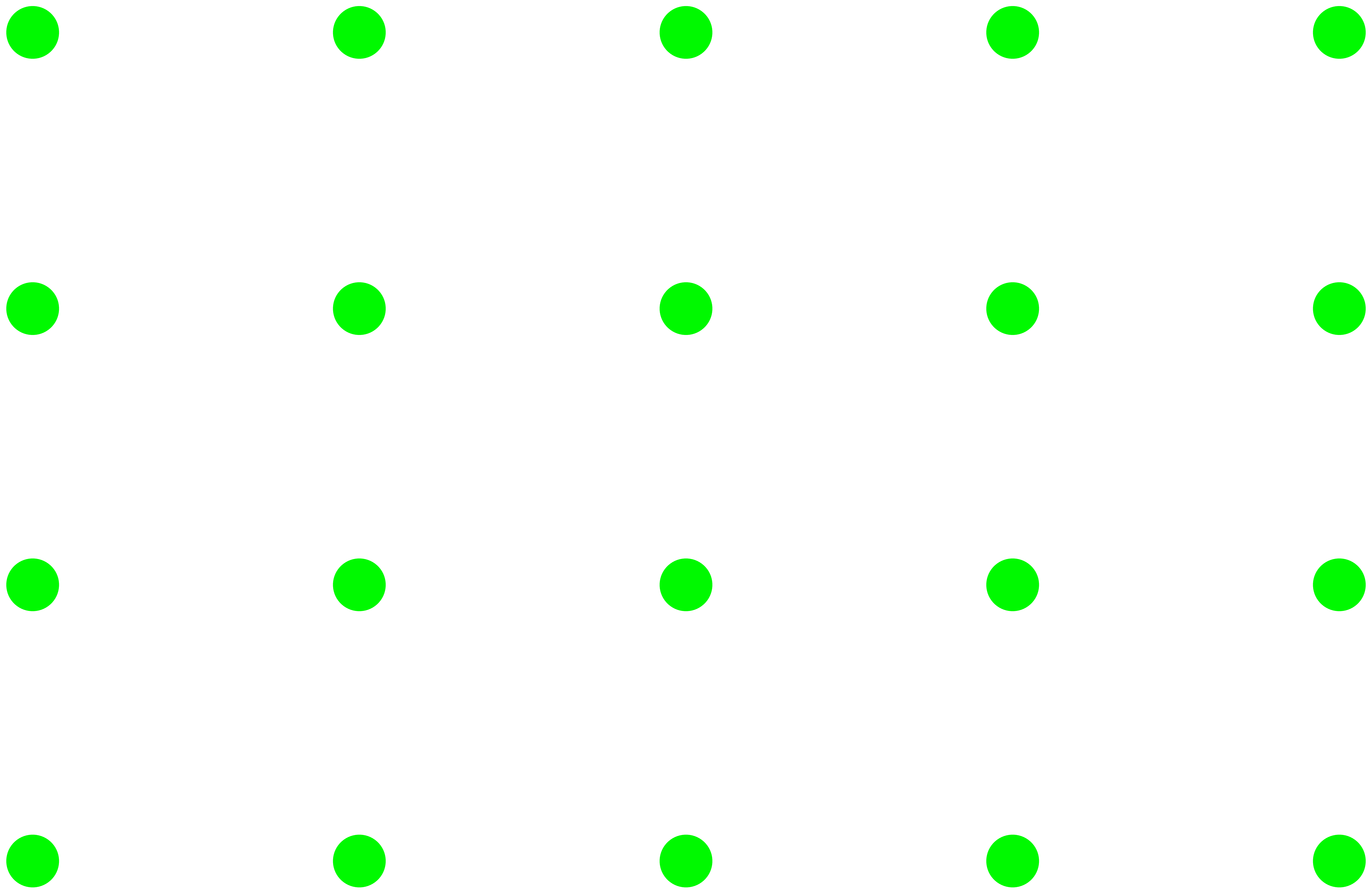
BG data set

- 68 x 68 sources, 150m spacing
- 401 x 401 receivers, 25m spacing

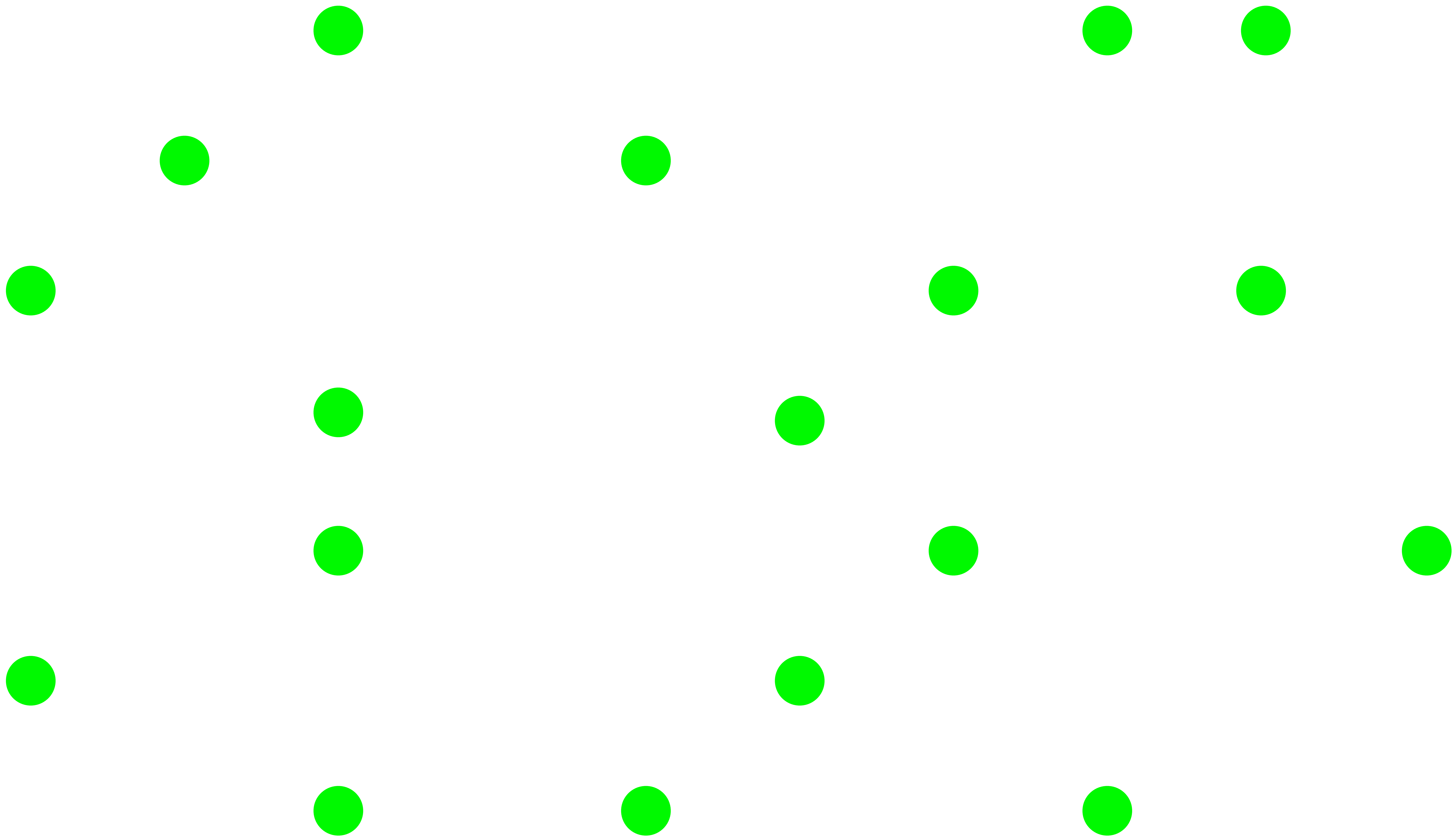
Consider two sampling situations

- regular 201 x 201 grid, 50m spacing
- irregular 201 x 201 grid
  - random 25m perturbations from the regular 201 x 201 grid
  - load true data from the irregular grid, avoid generating inverse crime data

Regular grid



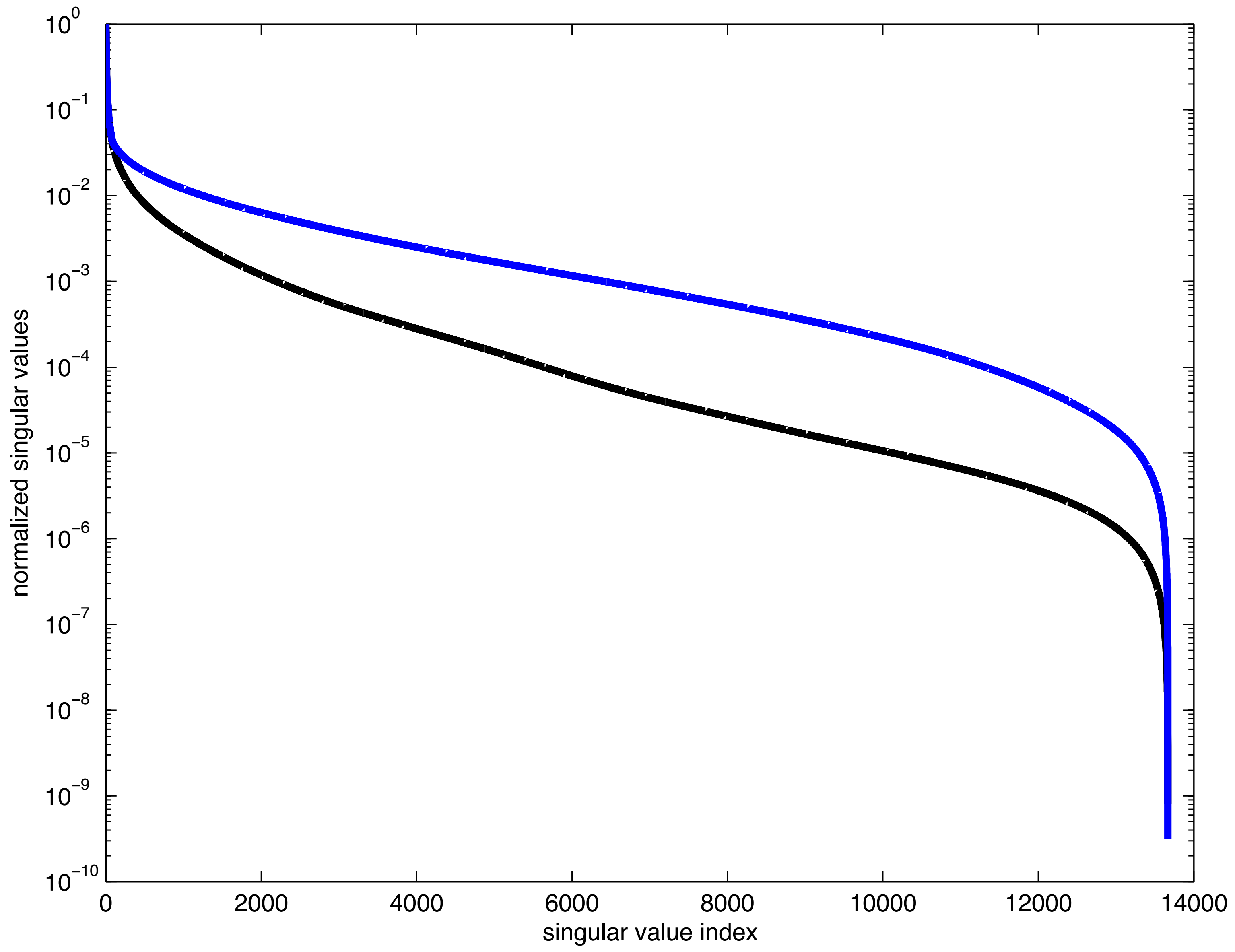
Unstructured grid



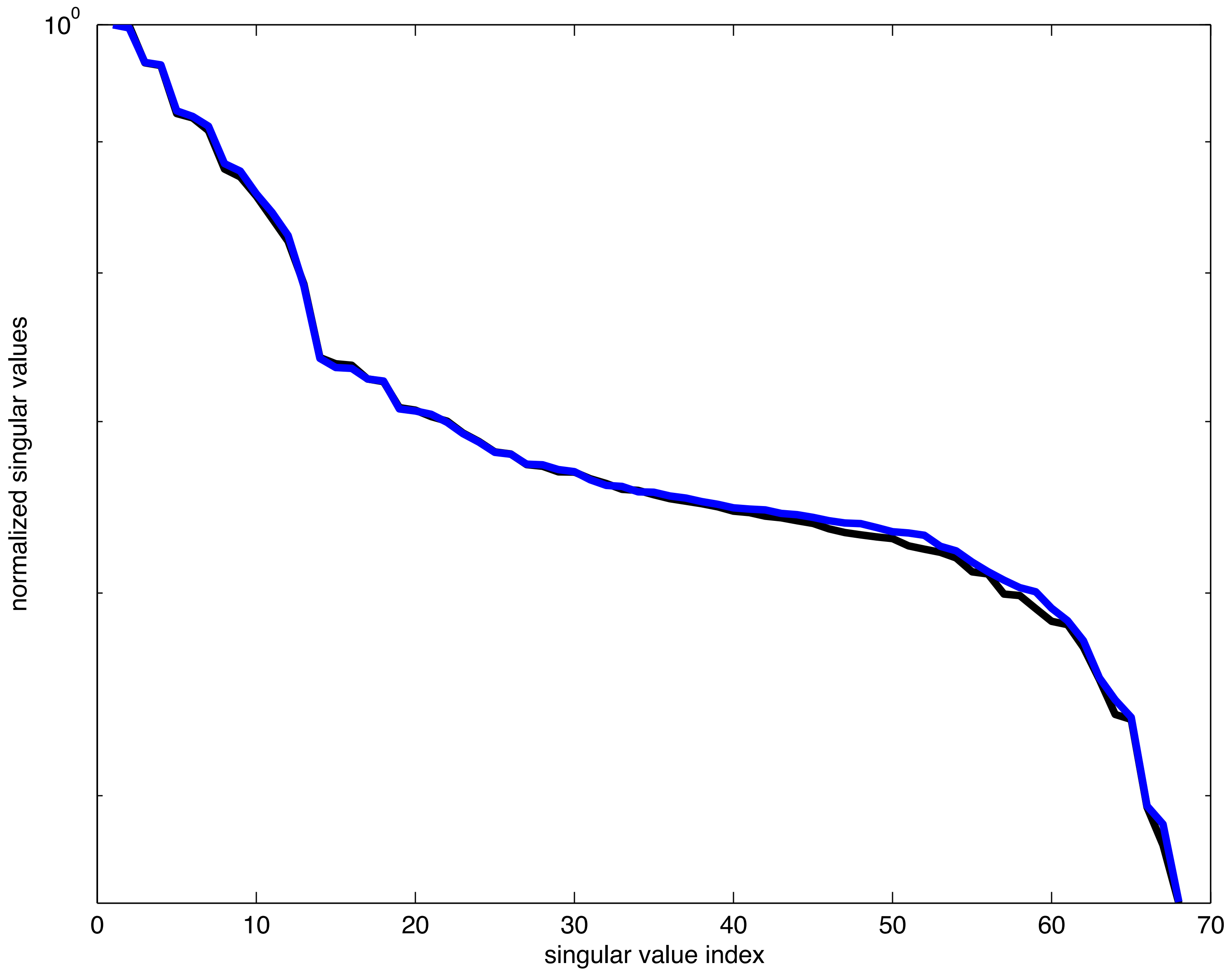
Regular vs irregular grid - singular values

Black - regular grid

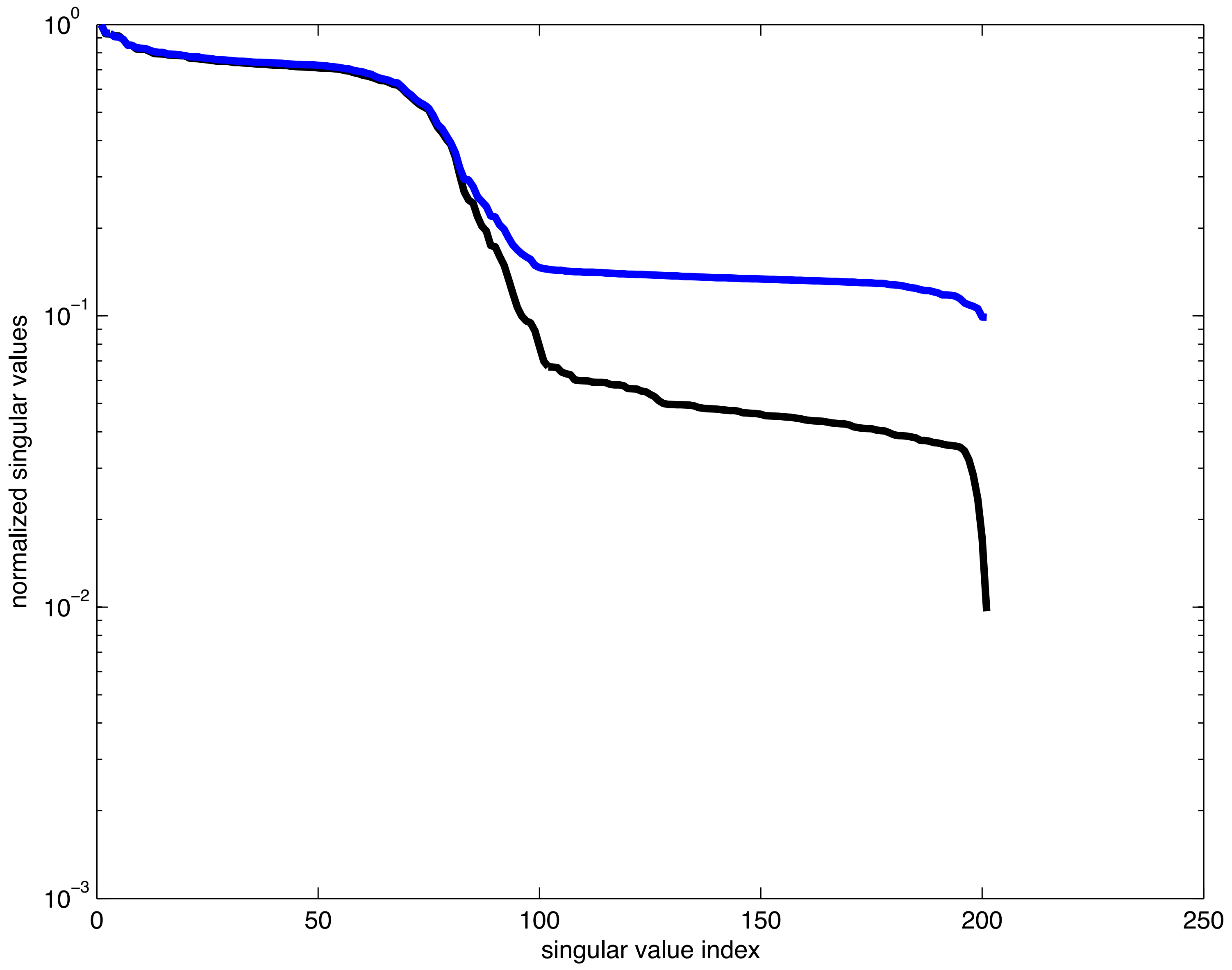
Blue - irregular grid



source x, receiver x



source x



receiver x

Off the grid tensor interpolation

The data volume is **no longer** low rank when irregularly sampled

- standard tensor completion framework won't work well

Solution

- construct a domain where the data **is** low rank
- choose an appropriate transform : low rank domain -> sampling domain
- incorporate transform in to the optimization problem

Multidimensional interpolation
with Hierarchical Tucker

Successful reconstruction scheme

Signal structure

- Hierarchical Tucker

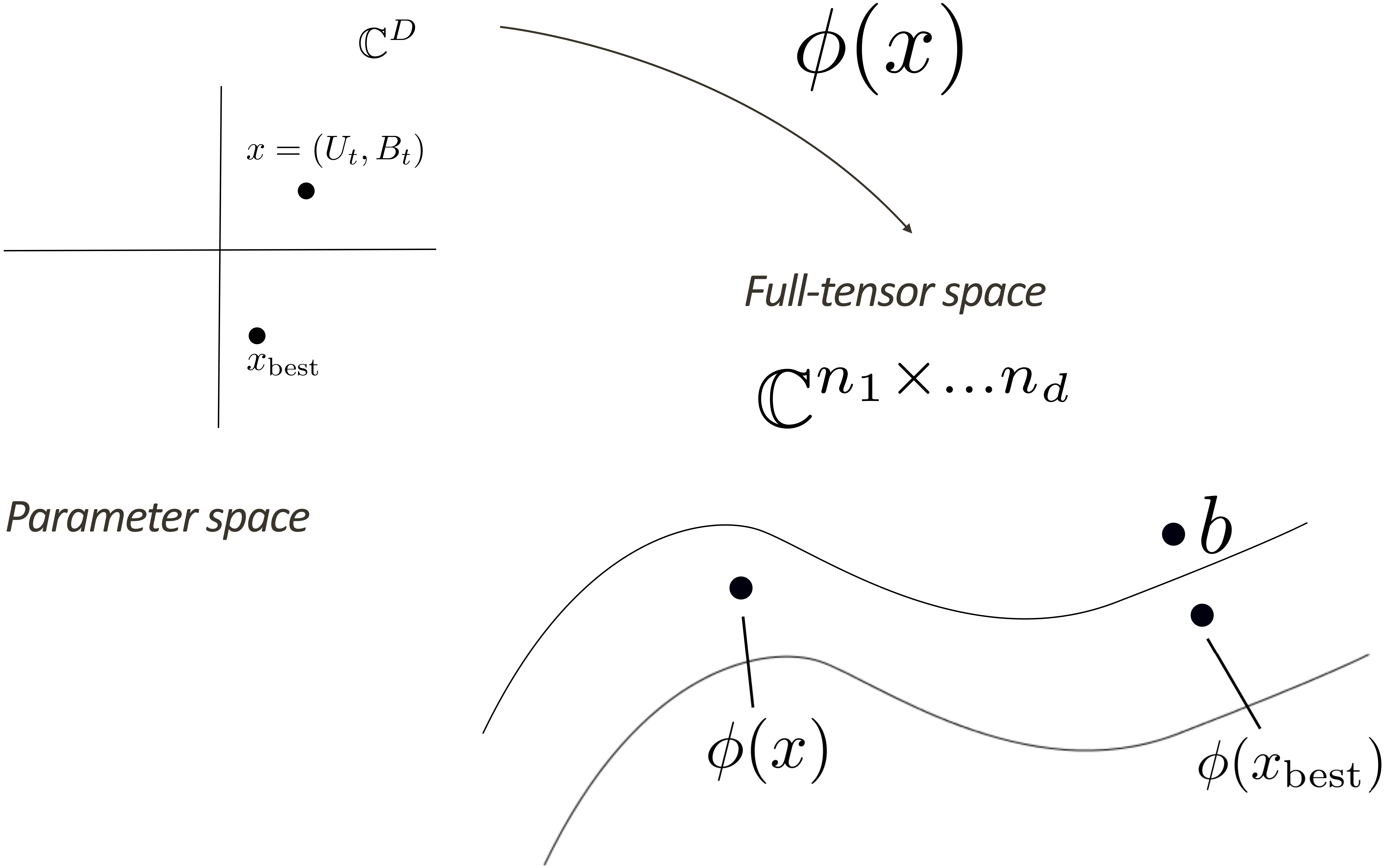
Sampling

- subsampling, noise increases hierarchical rank

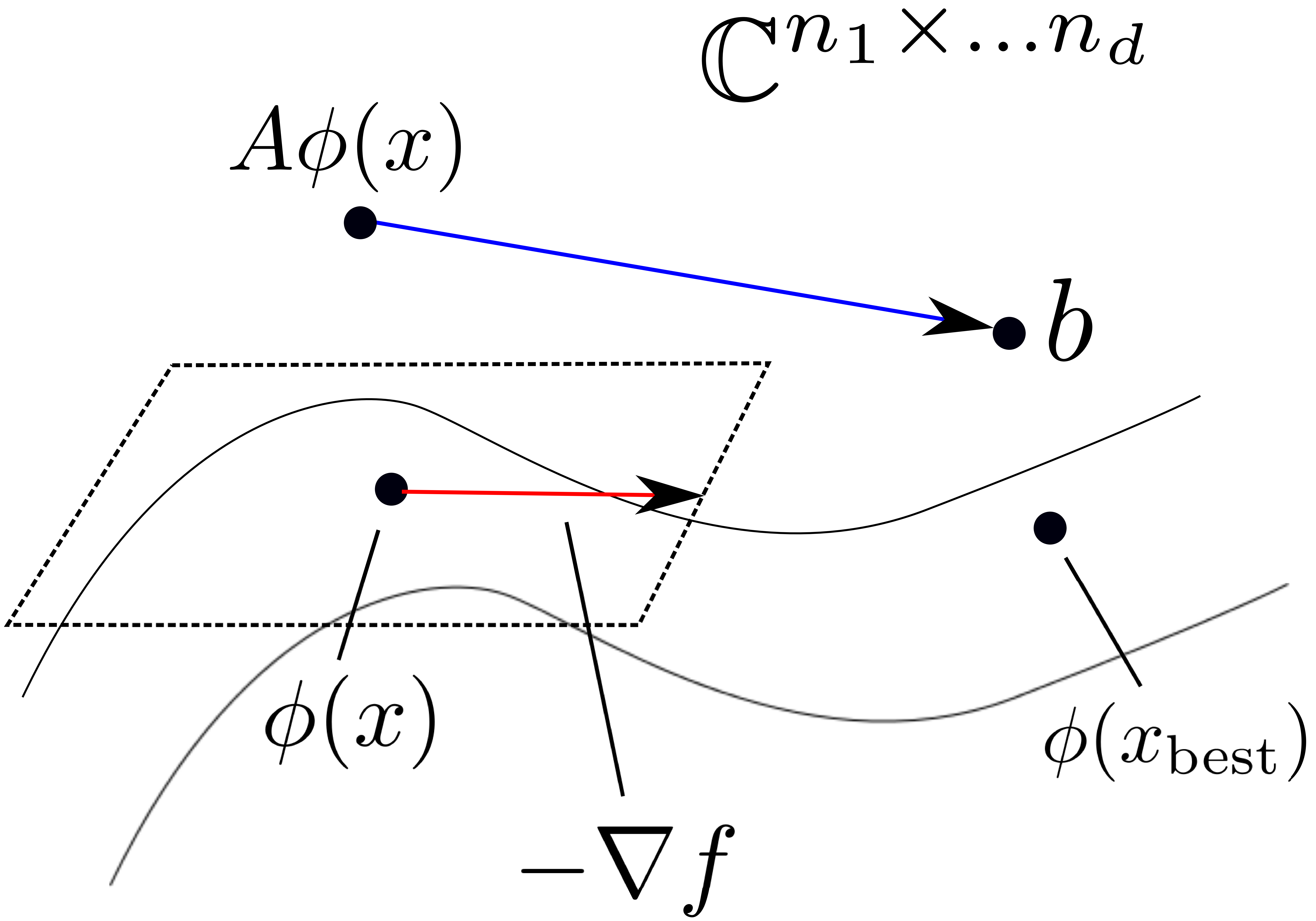
Optimization

- ***fit data in the Hierarchical Tucker format***

Optimization program



Optimization program



Optimization problem

The standard problem we solve is

$$\min_{x} \|\mathcal{A}\phi(x) - b\|_2^2$$

Our sampling operator is typically

$$\mathcal{A} = \mathcal{R}\mathcal{P}$$

where

$$\mathcal{R} : \text{regular full grid} \rightarrow \text{subsampled grid}$$

$$\mathcal{P} : (\text{src x}, \text{rec x}, \text{src y}, \text{rec y}) \rightarrow (\text{src x}, \text{src y}, \text{rec x}, \text{rec y})$$

[1] R. Kumar, O. Lopez, E. Esser, and F. J. Herrmann. Matrix completion on unstructured grids : 2-D seismic data regularization and interpolation. EAGE 2015.

SLIM

Optimization problem

In the irregular grid case, the subsampling operator is in fact

$$\mathcal{R} : \text{irregular full grid} \rightarrow \text{subsampled grid}$$

In order to take this discrepancy in to account, we introduce an operator

$$\mathcal{F} : \text{regular full grid} \rightarrow \text{irregular full grid}$$

This is an extension of [1] to the tensor case

[1] C. Da Silva and F. J. Herrmann. Optimization on the hierarchical tucker manifold – applications to tensor completion. Linear Algebra and its Applications
[2] L Greengard and J.Y. Lee. Accelerating the nonuniform fast fourier transform. SIAM Review.

SLIM

Optimization problem

The sequence of operators is then

$$\mathcal{R} : \text{irregular full grid} \rightarrow \text{subsampled grid}$$
$$\mathcal{F} : \text{regular full grid} \rightarrow \text{irregular full grid}$$
$$\mathcal{P} : (\text{src x}, \text{rec x}, \text{src y}, \text{rec y}) \rightarrow (\text{src x}, \text{src y}, \text{rec x}, \text{rec y})$$

We set $\mathcal{A} = \mathcal{R}\mathcal{F}\mathcal{P}$ and use the same optimization code as previously in [1]

In our examples, we use the non-uniform Fourier transform [2]

Optimization problem

Main computational costs

- Gauss-Newton method -> convergence in ~15 iterations
  - ~ 4 objective evaluations per iteration

- ~60 applications of the interpolation operator
  - main source of computational costs

Results

Synthetic BG Group data

Unknown model

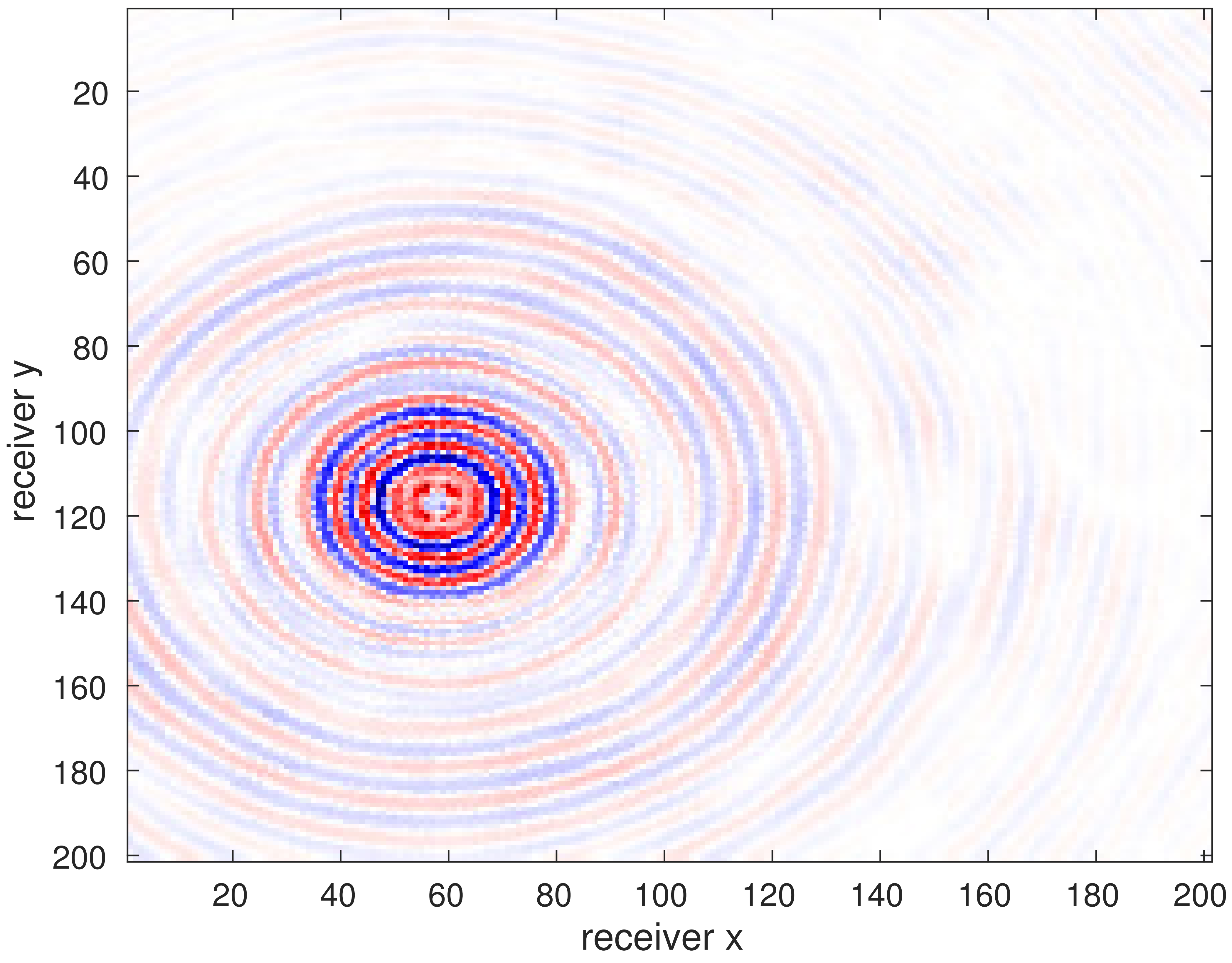
- 68 x 68 sources with 401 x 401 receivers, data at 7.34 Hz, 12.3 Hz

Receivers sampled on a randomly perturbed 201 x 201 grid

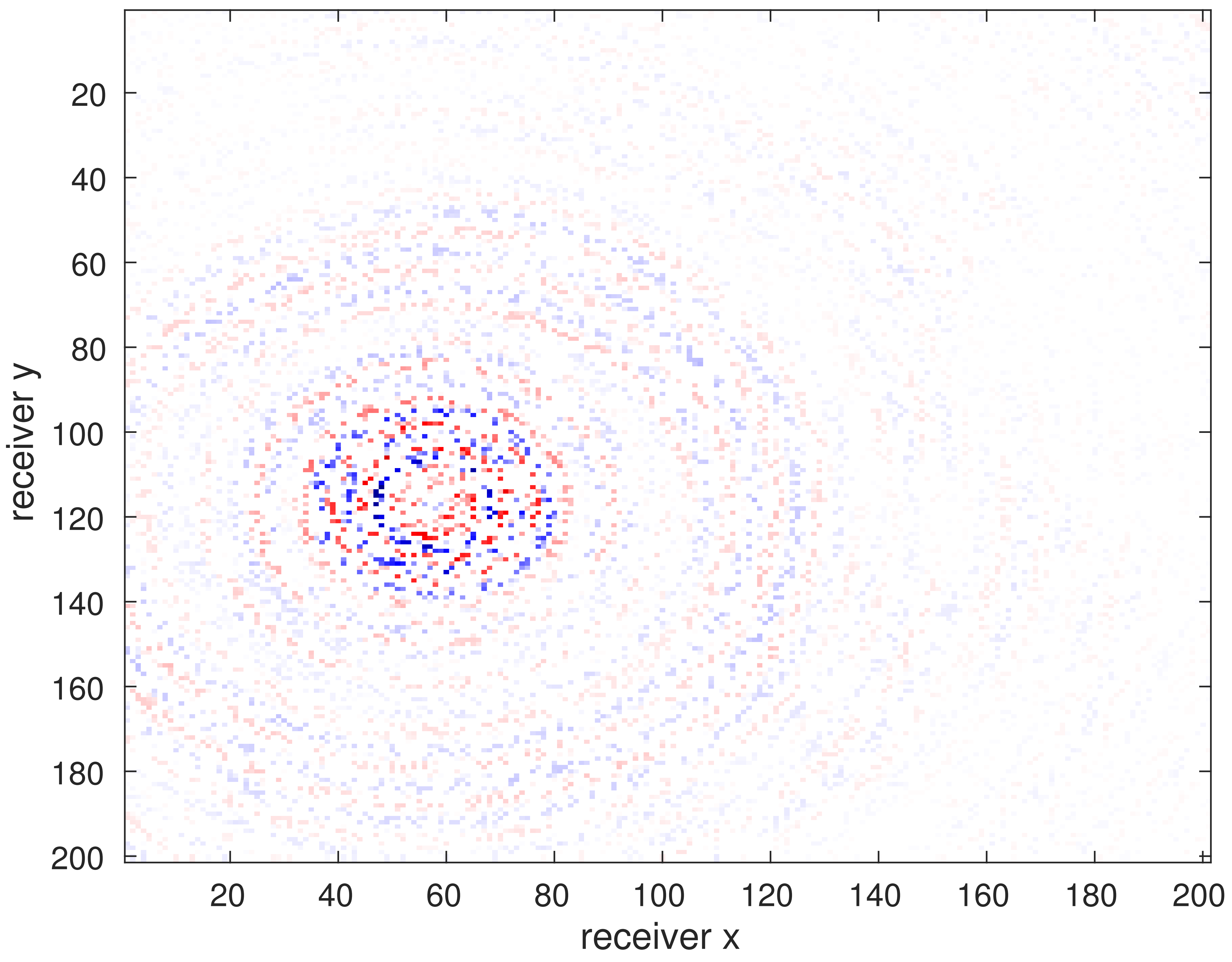
We compare to ‘vanilla’ tensor completion, where we just bin the data to a regular grid

7.34 Hz - 75% missing receivers

Common source gather



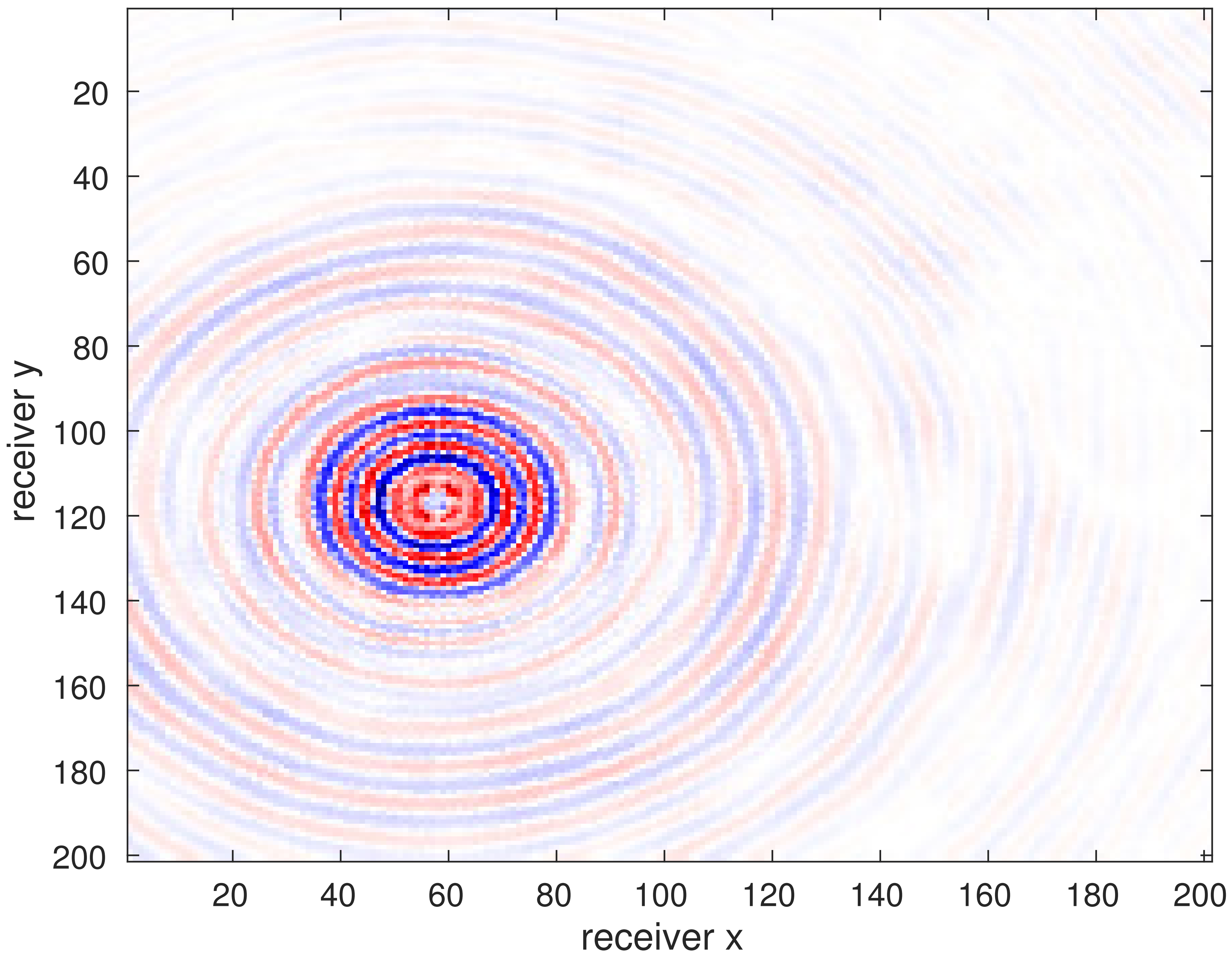
True data



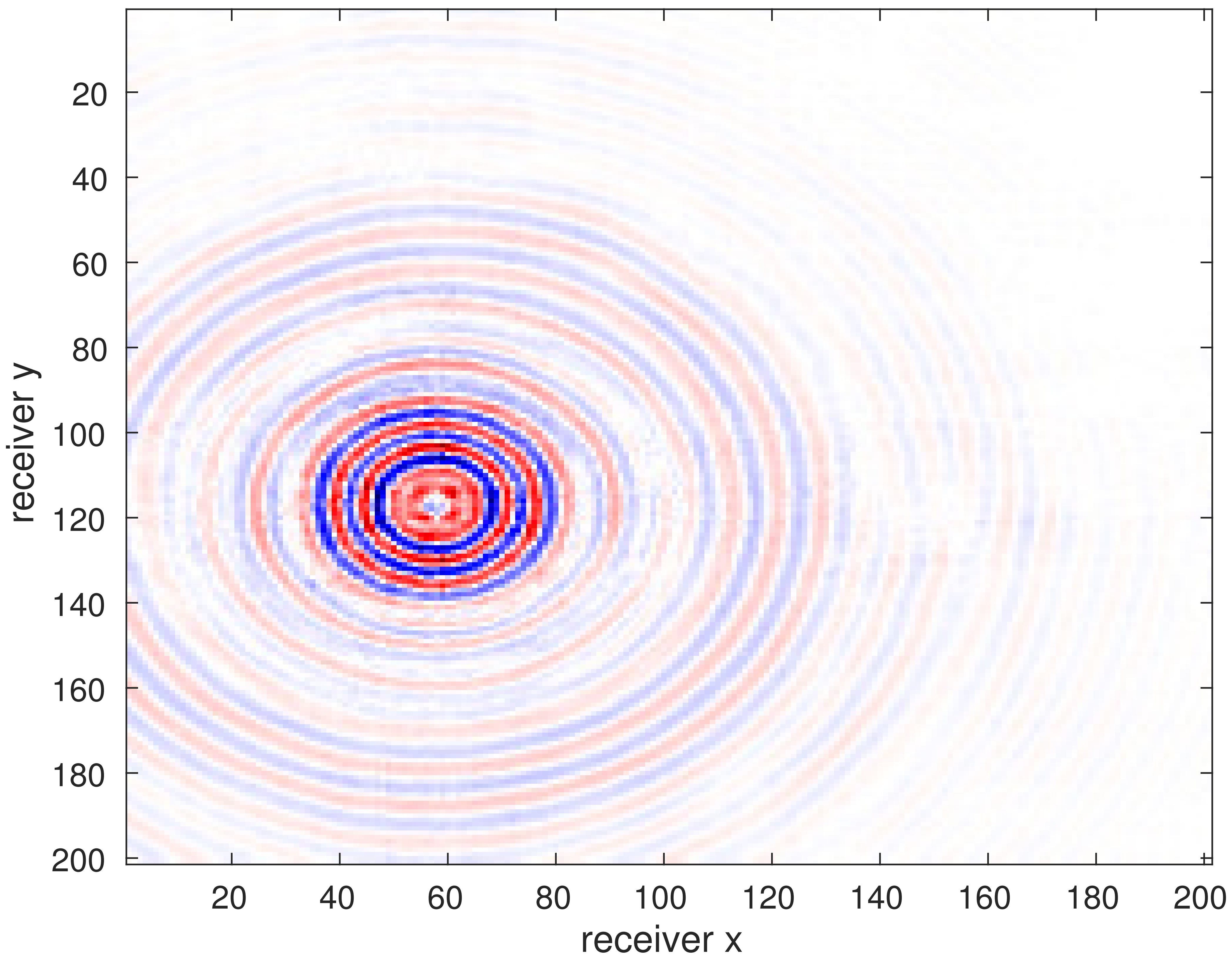
Subsampled data

7.34 Hz - 75% missing receivers

Common source gather



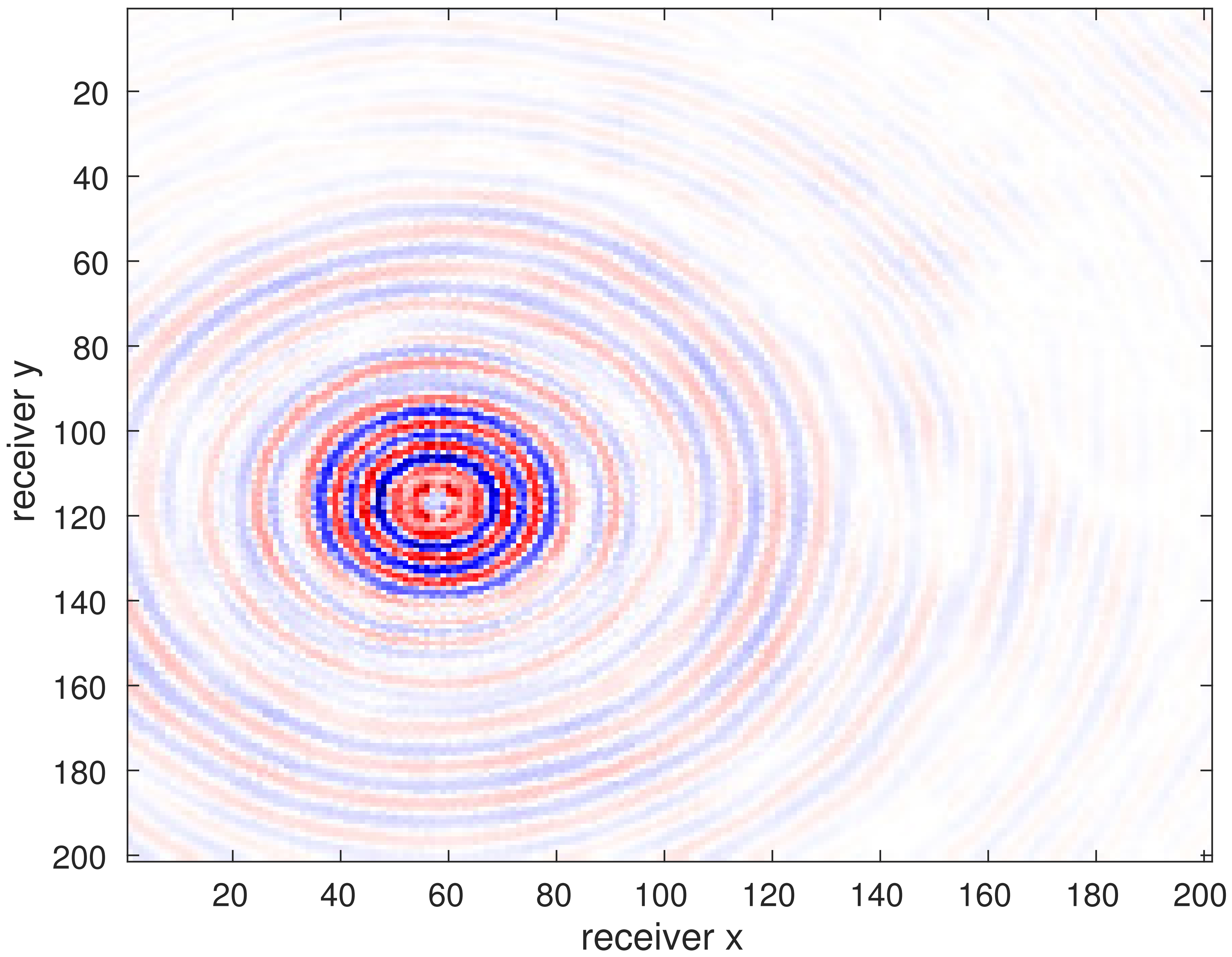
True data



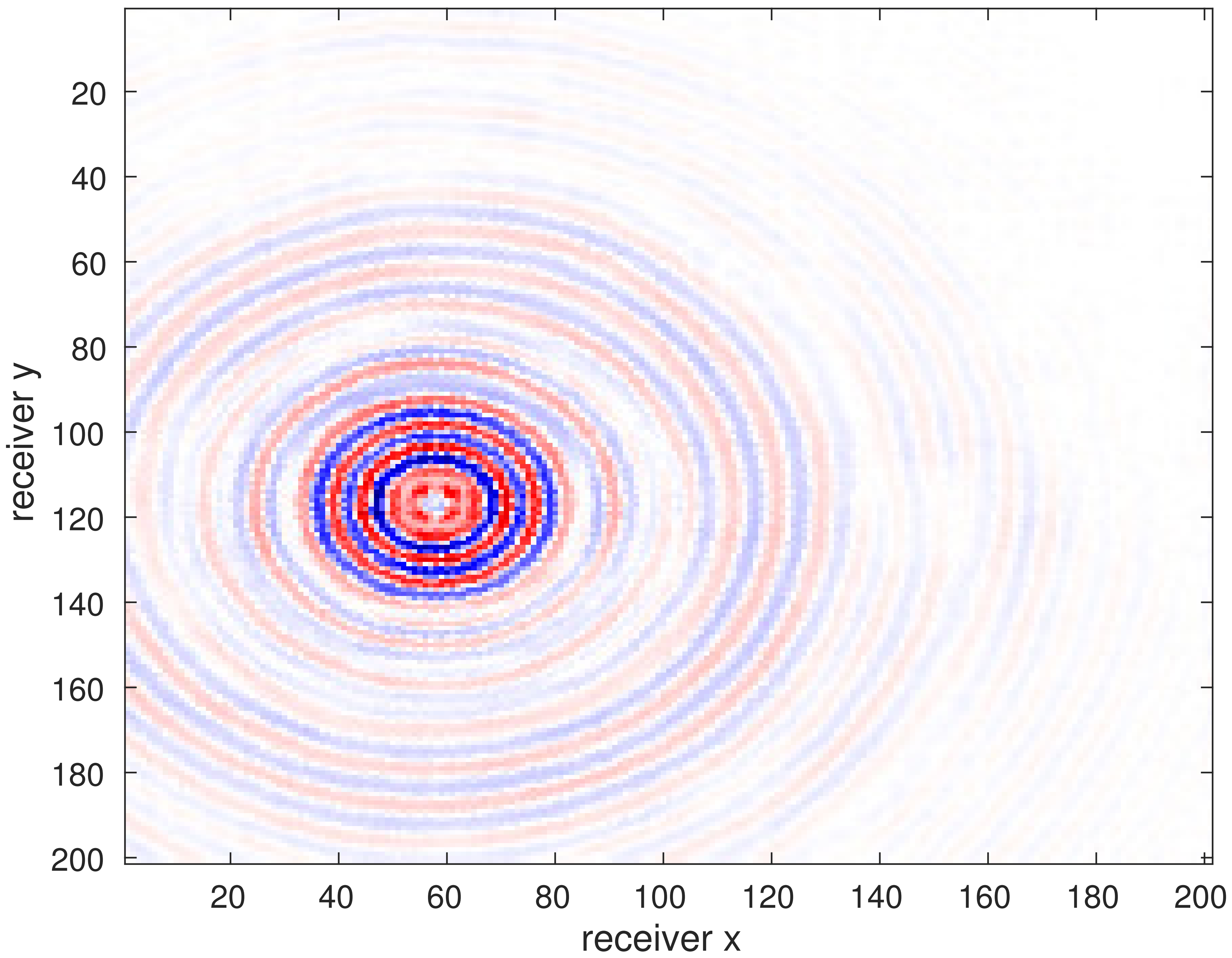
Vanilla tensor completion
SNR 9.46 dB

7.34 Hz - 75% missing receivers

Common source gather



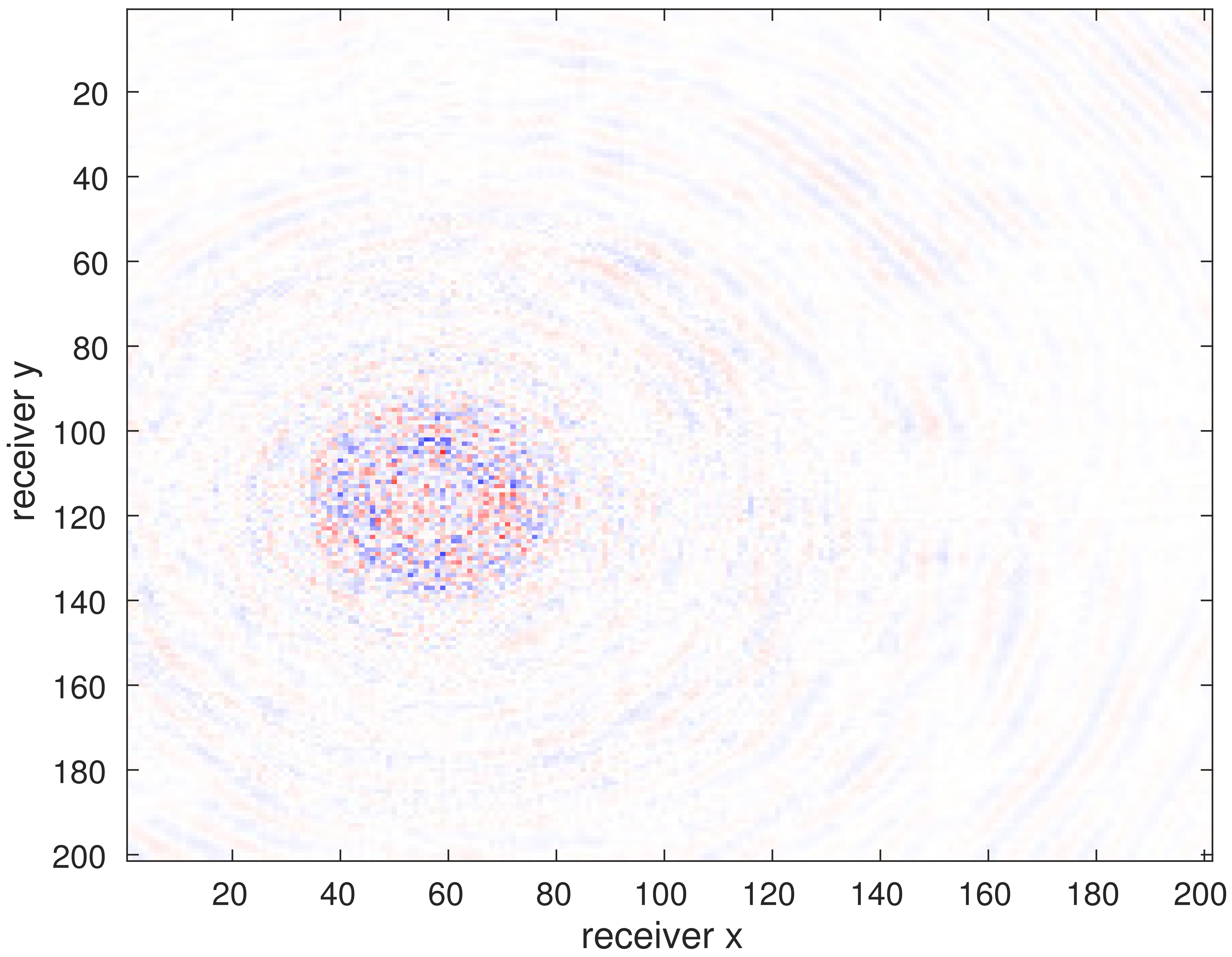
True data



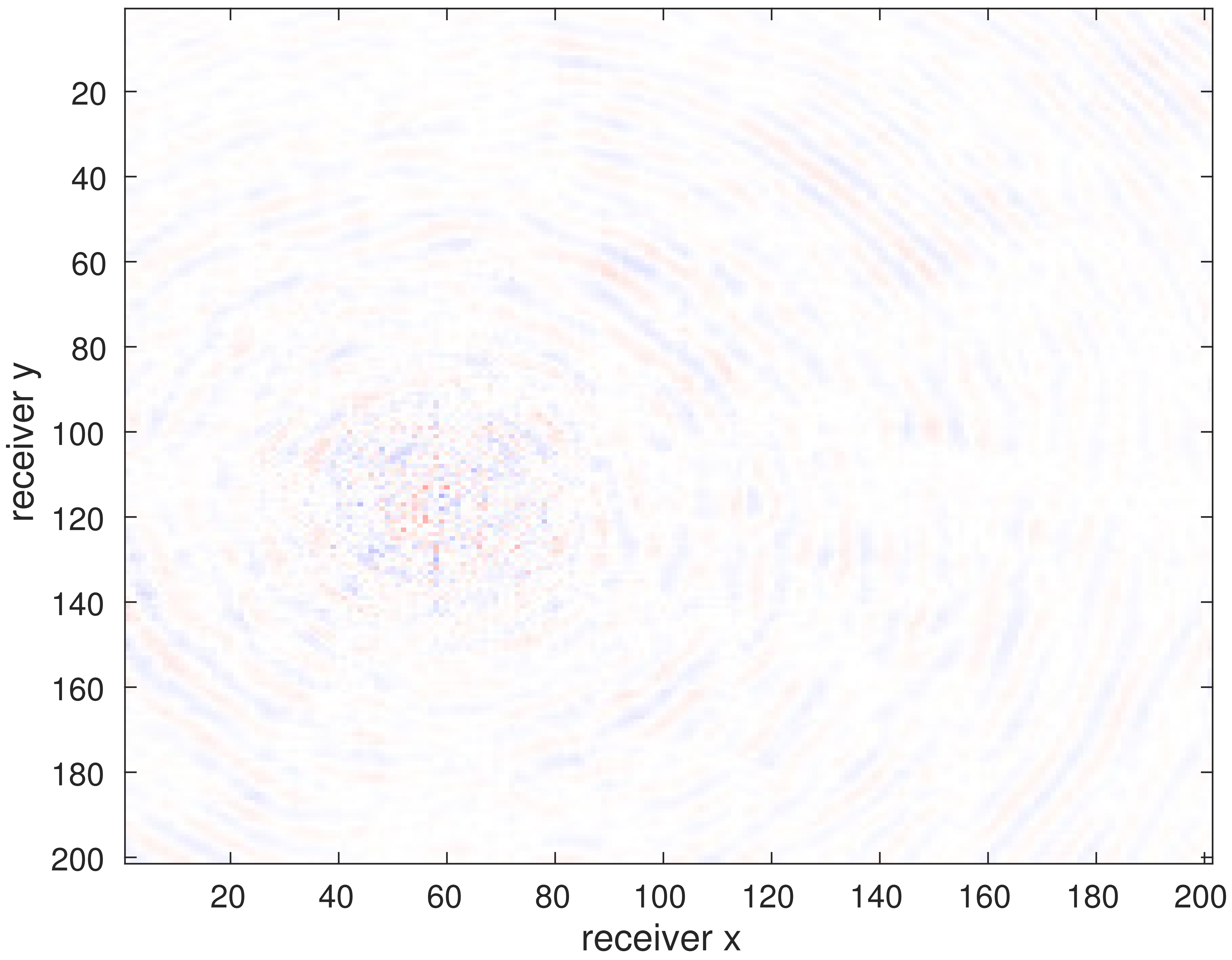
Regularized tensor completion
SNR 15.7 dB

7.34 Hz - 75% missing receivers

Common source gather

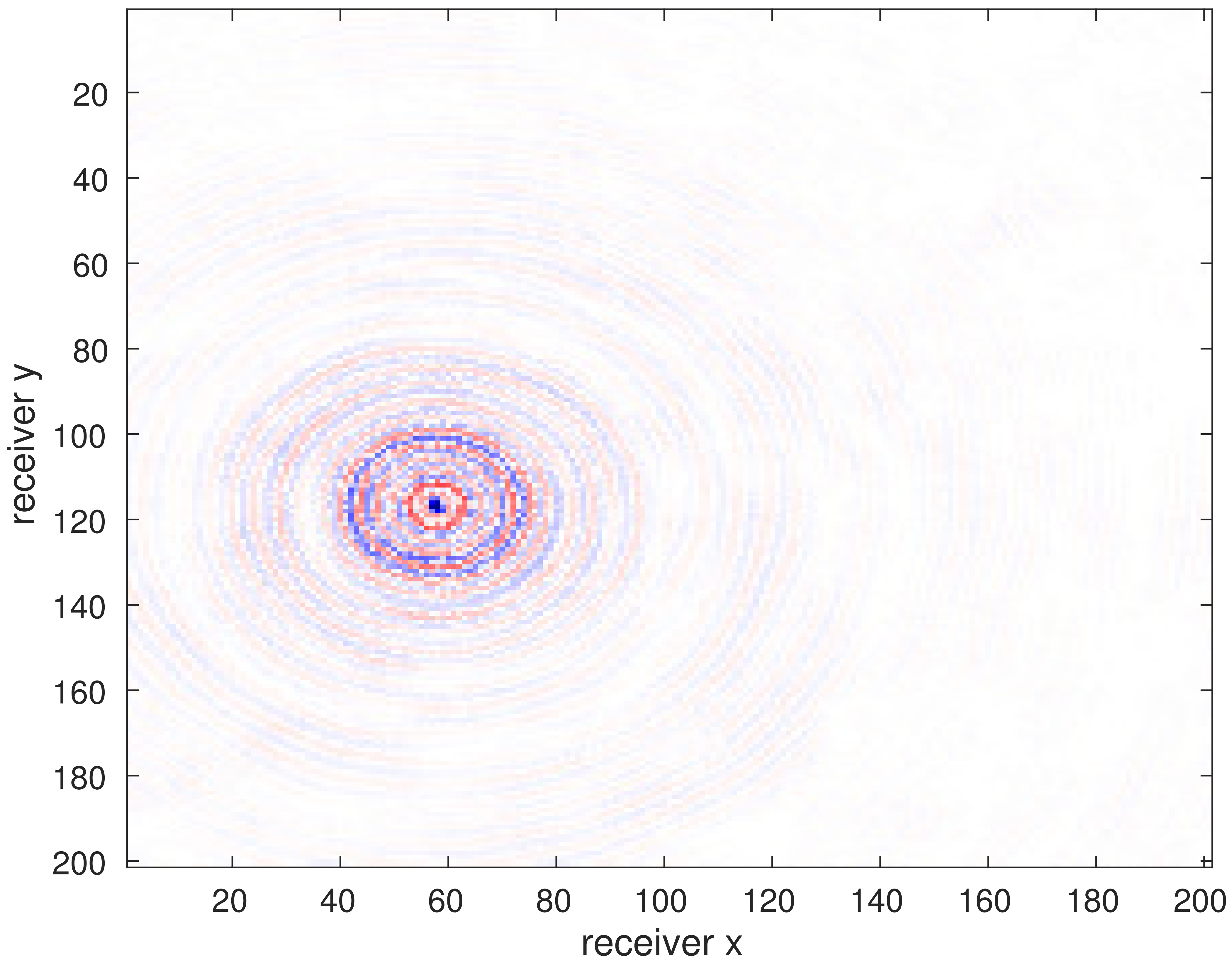


Vanilla tensor completion difference

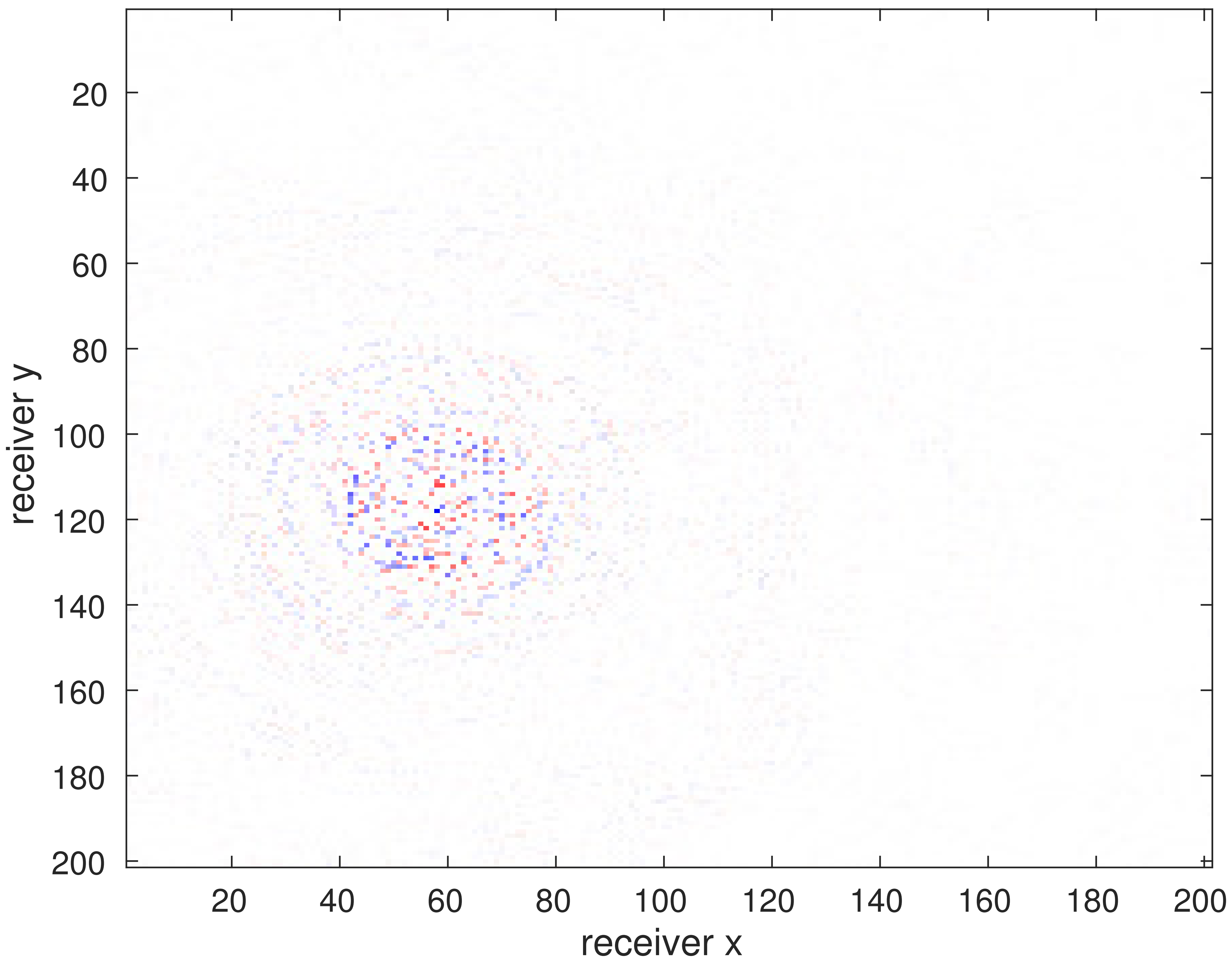


Regularized tensor completion difference

12.3 Hz - 75% missing receivers
Common source gather



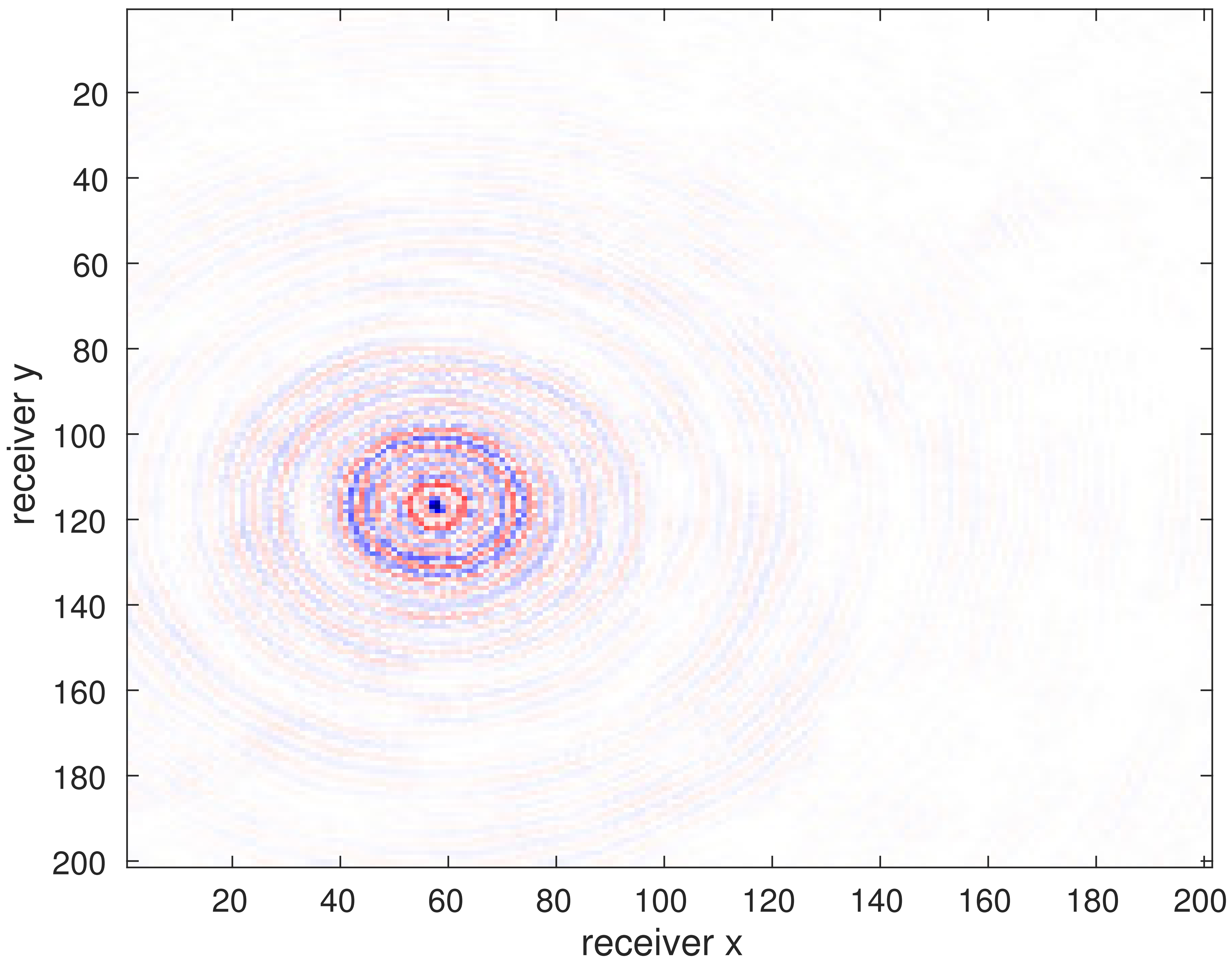
True data



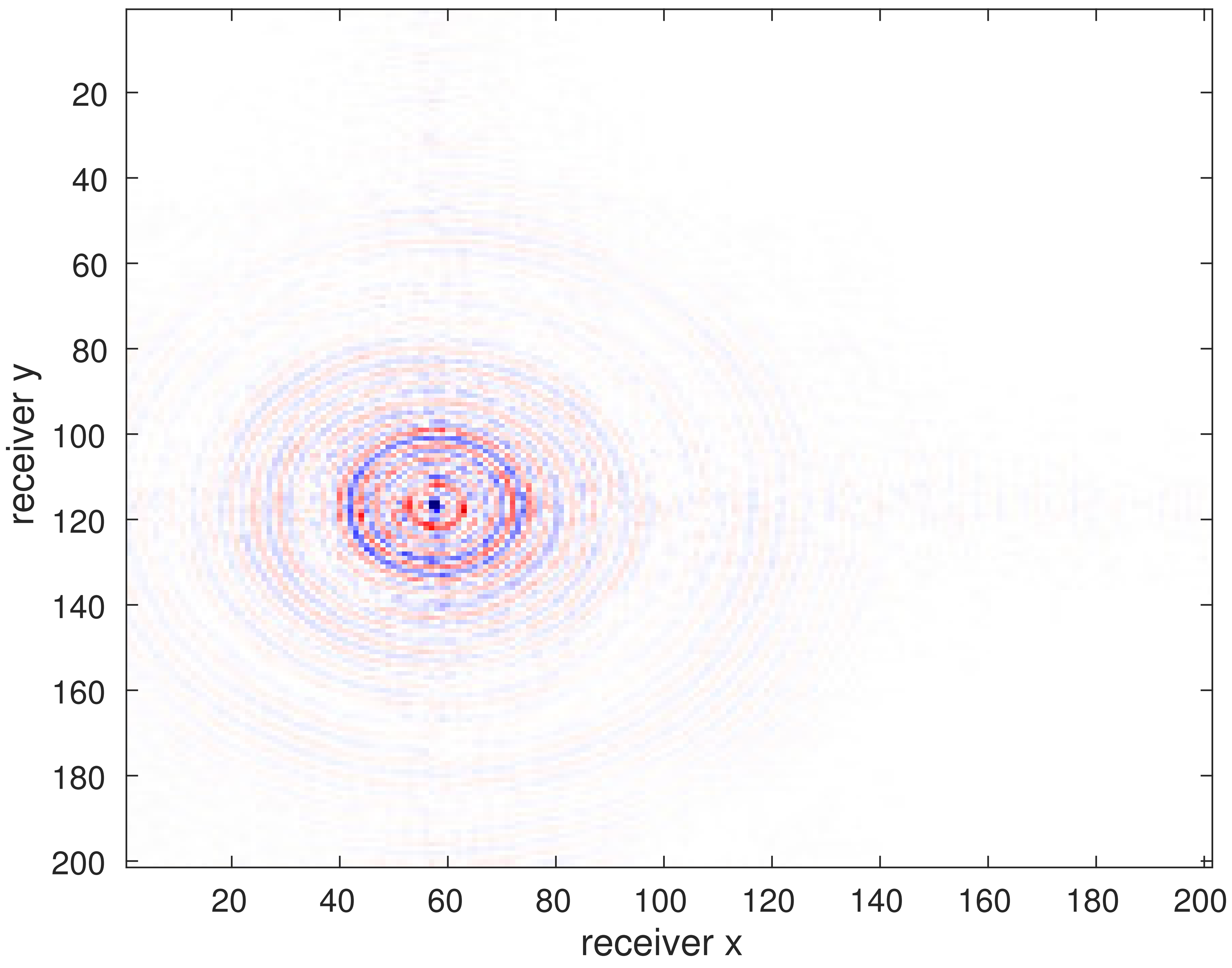
Subsampled data

12.3 Hz - 75% missing receivers

Common source gather



True data

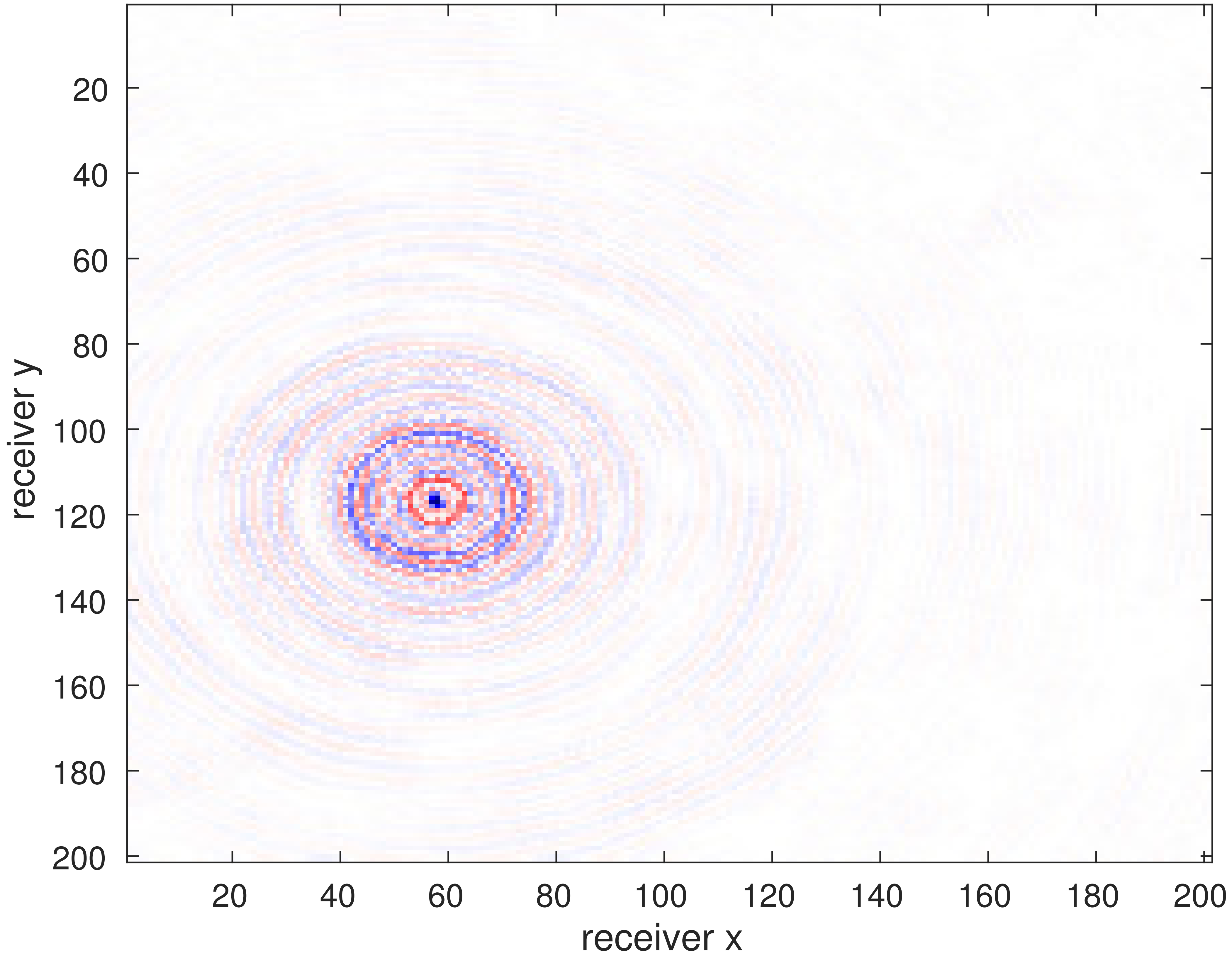


Vanilla tensor completion

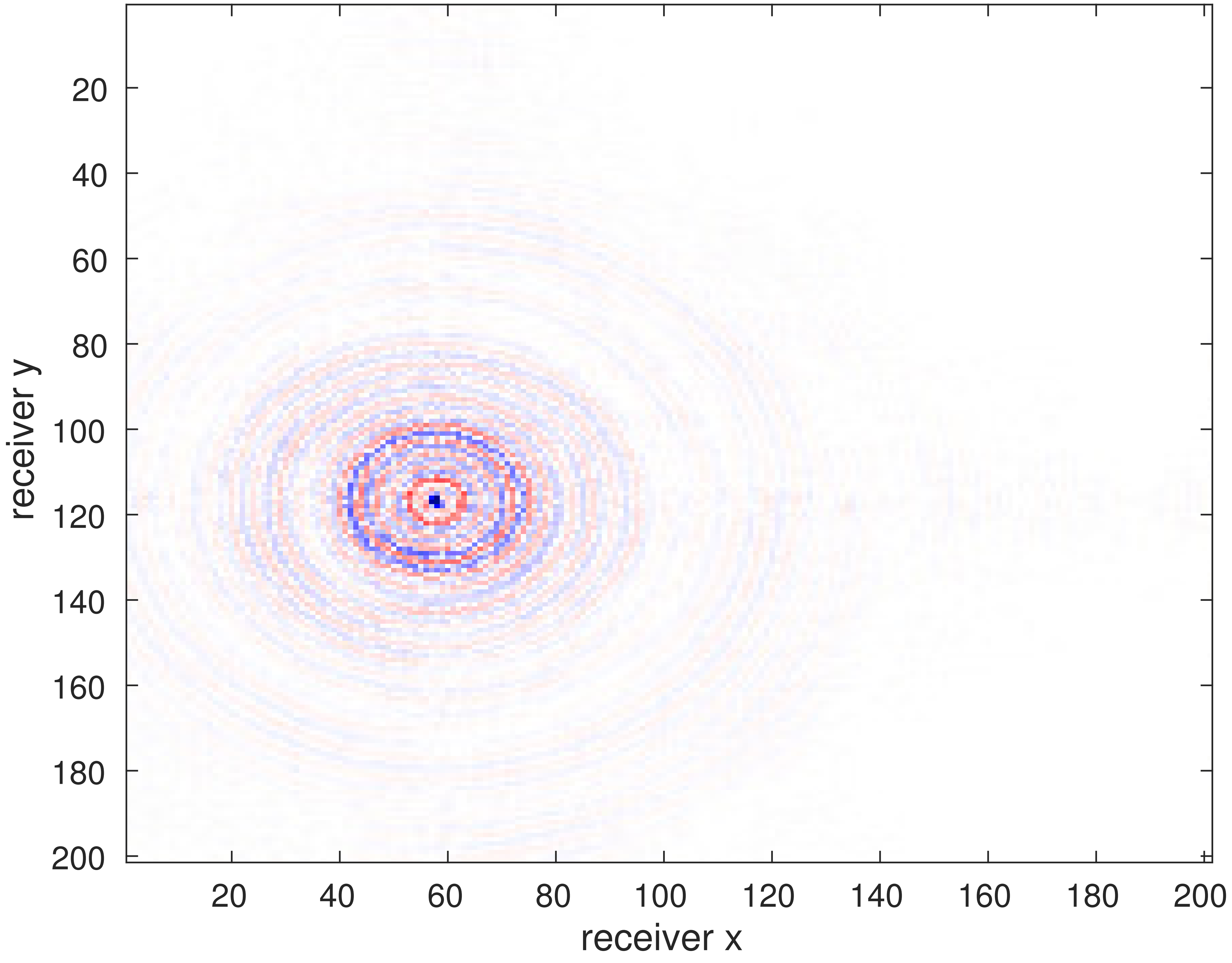
SNR 4.45 dB

12.3 Hz - 75% missing receivers

Common source gather



True data

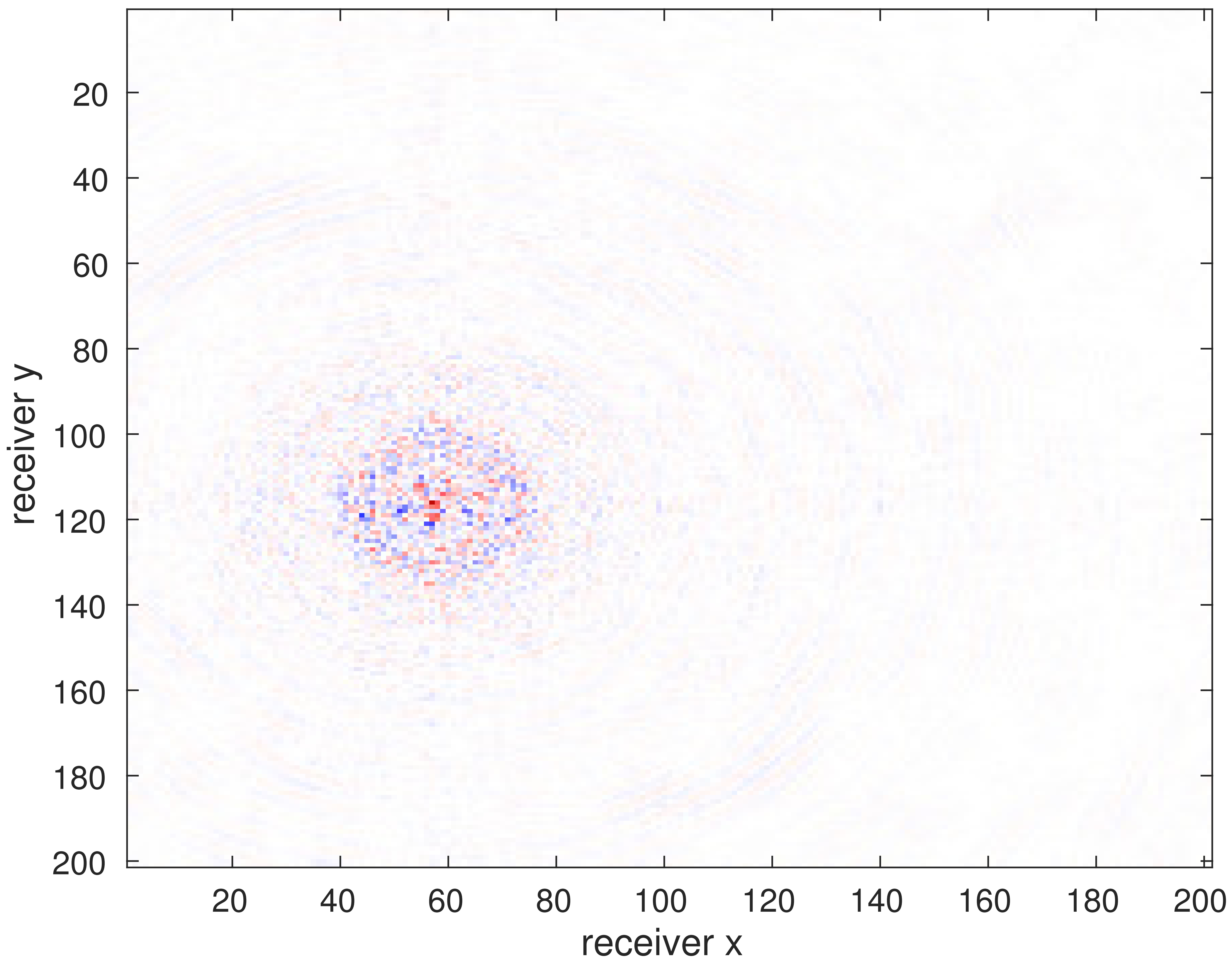


Regularized tensor completion

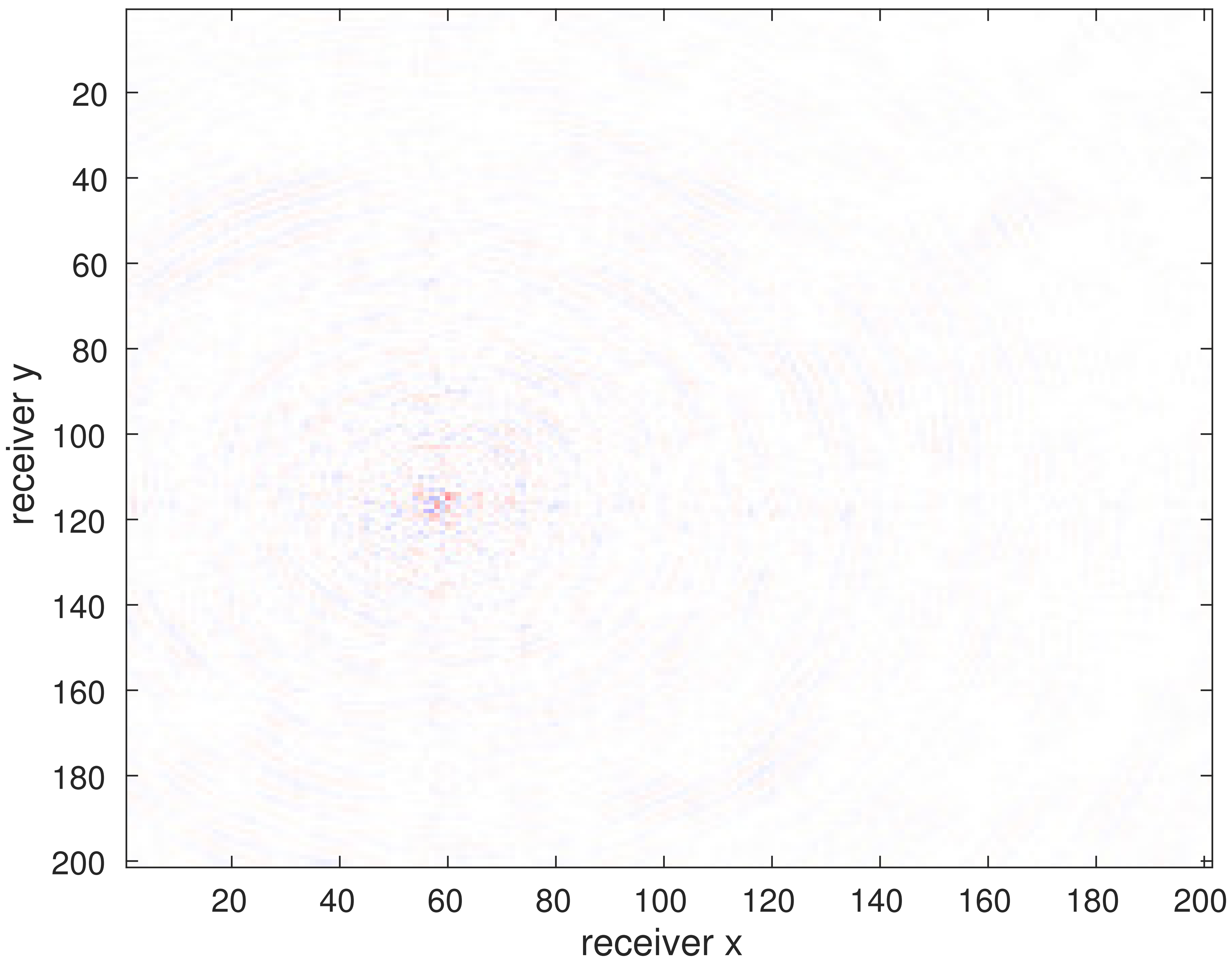
SNR 11.4 dB

12.3 Hz - 75% missing receivers

Common source gather



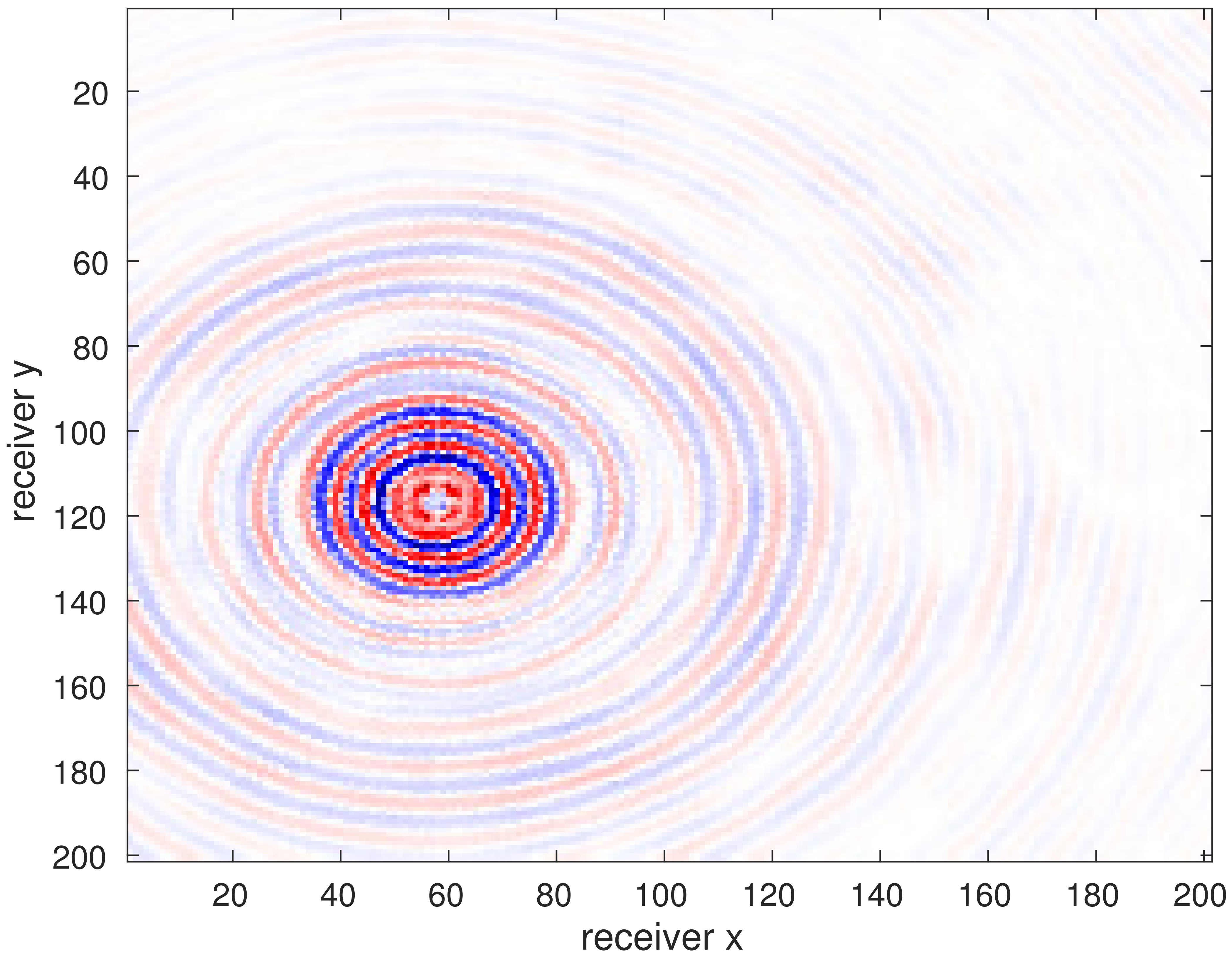
Vanilla tensor completion difference



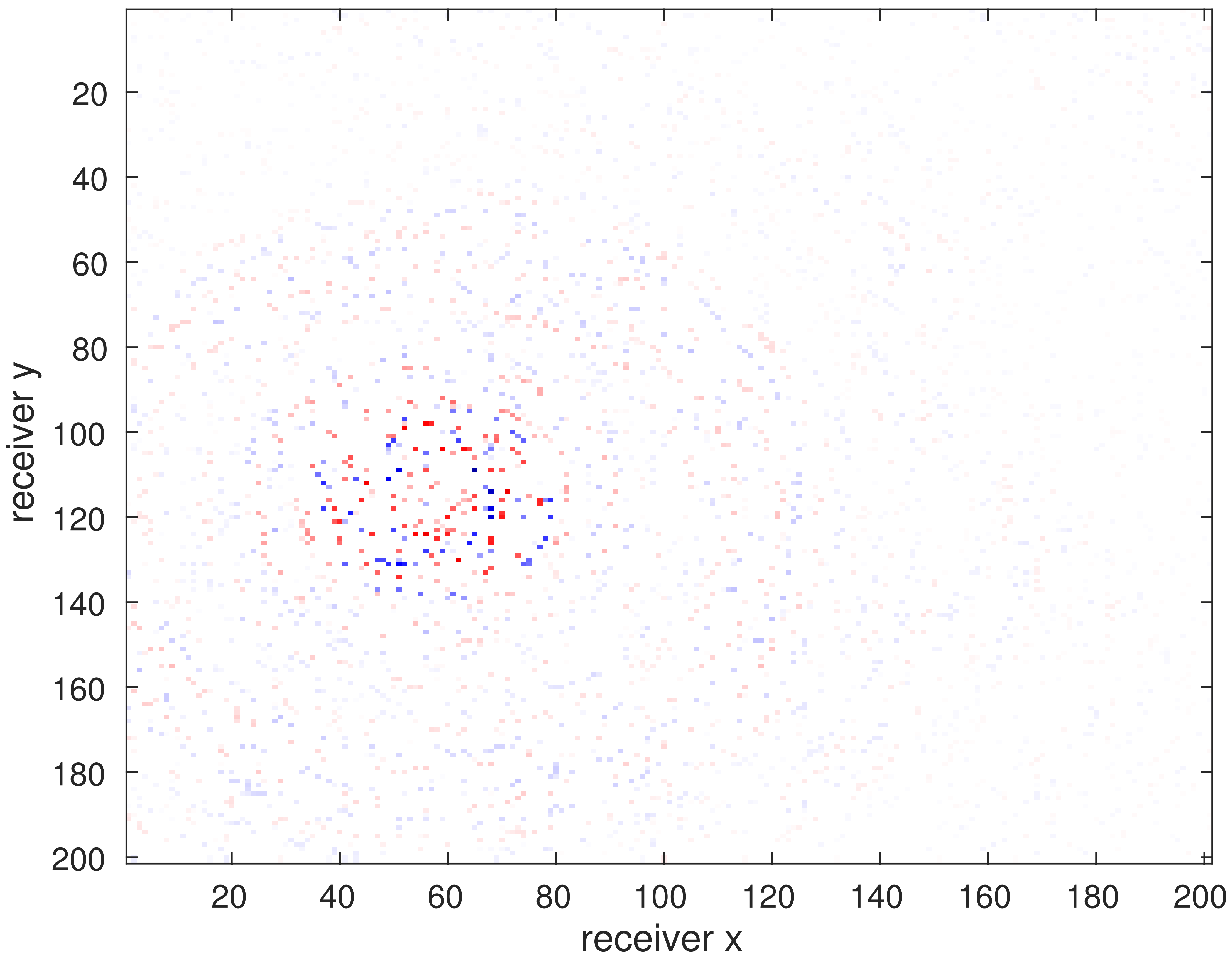
Regularized tensor completion difference

7.34 Hz - 90% missing receivers

Common source gather



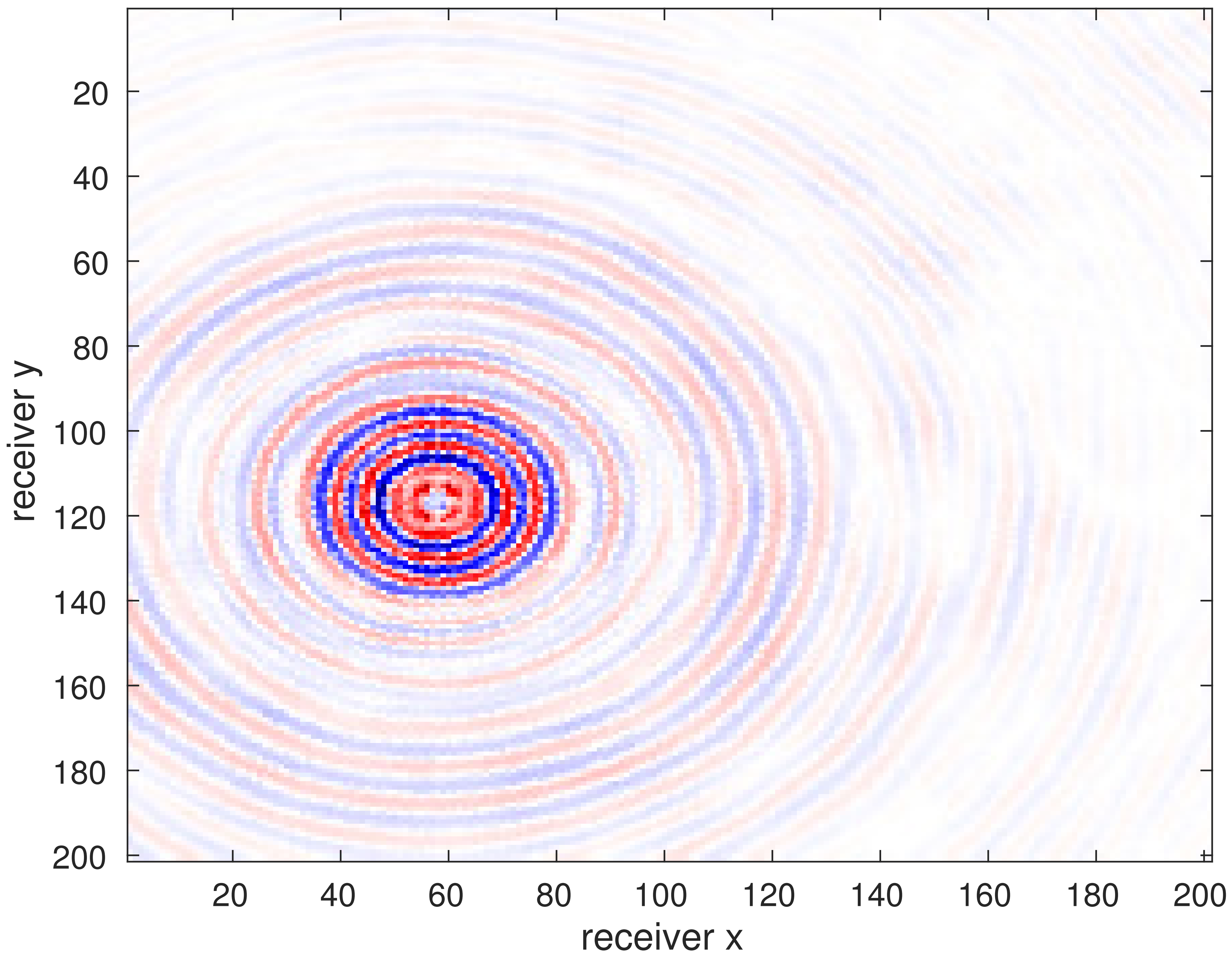
True data



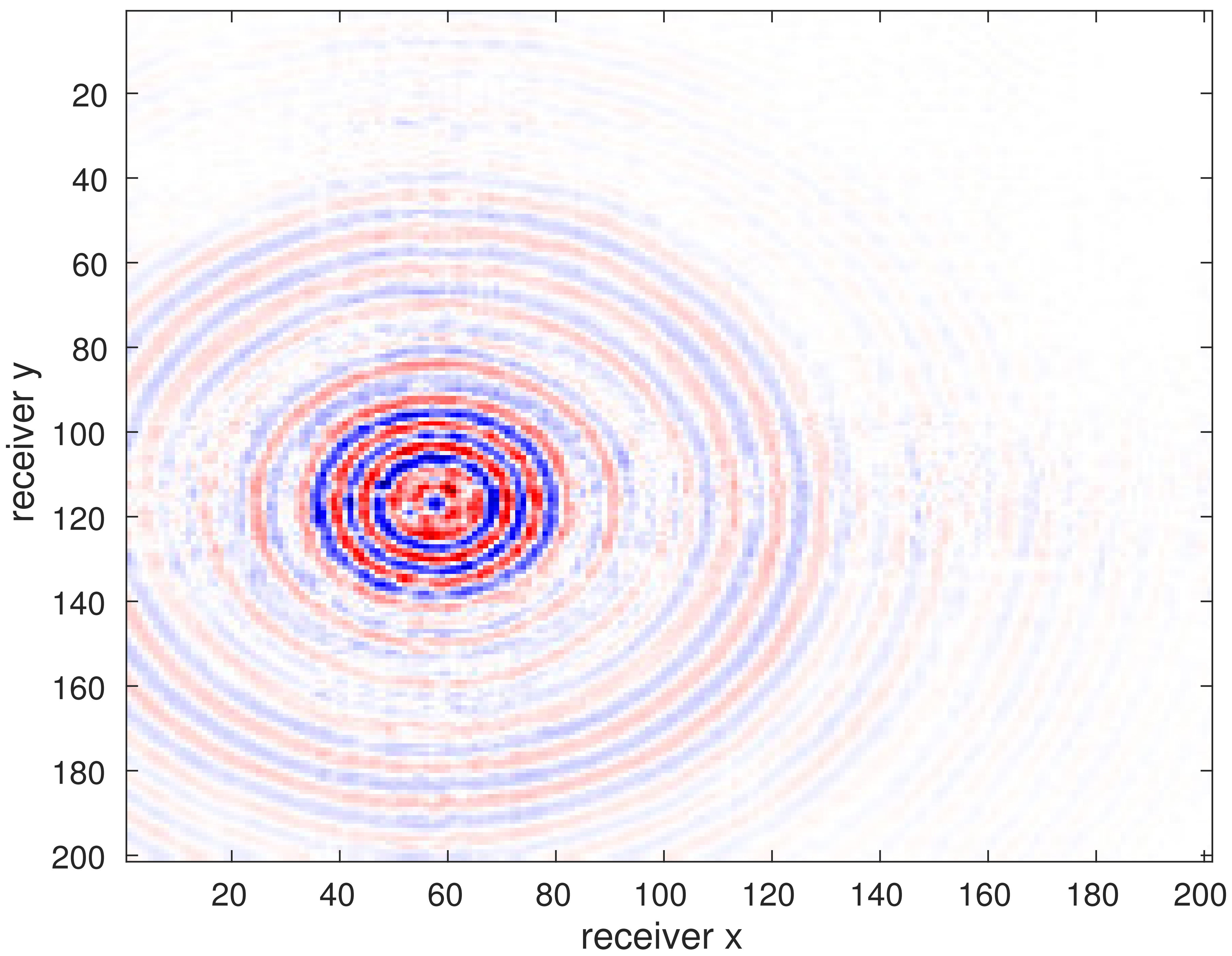
Subsampled data

7.34 Hz - 90% missing receivers

Common source gather

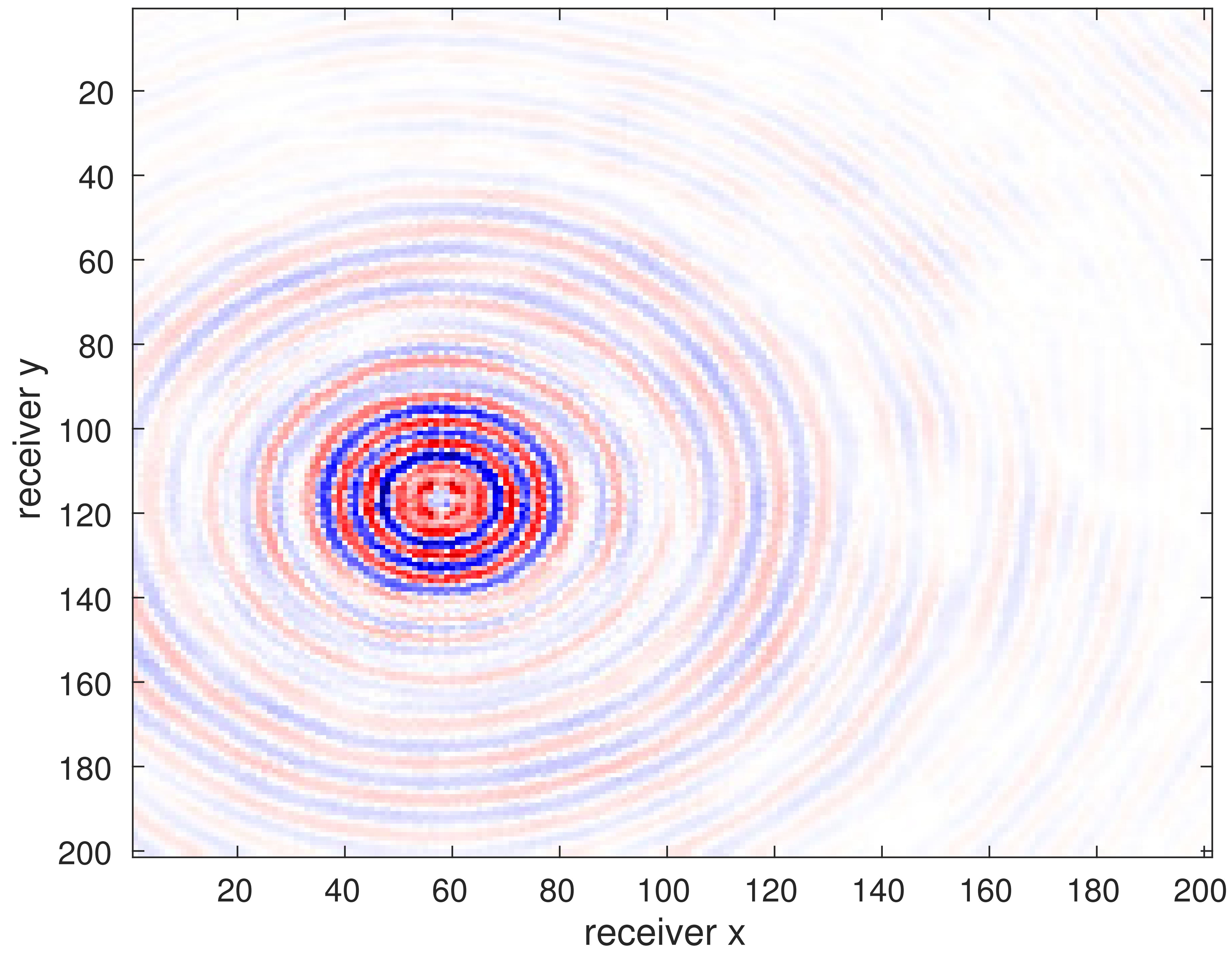


True data

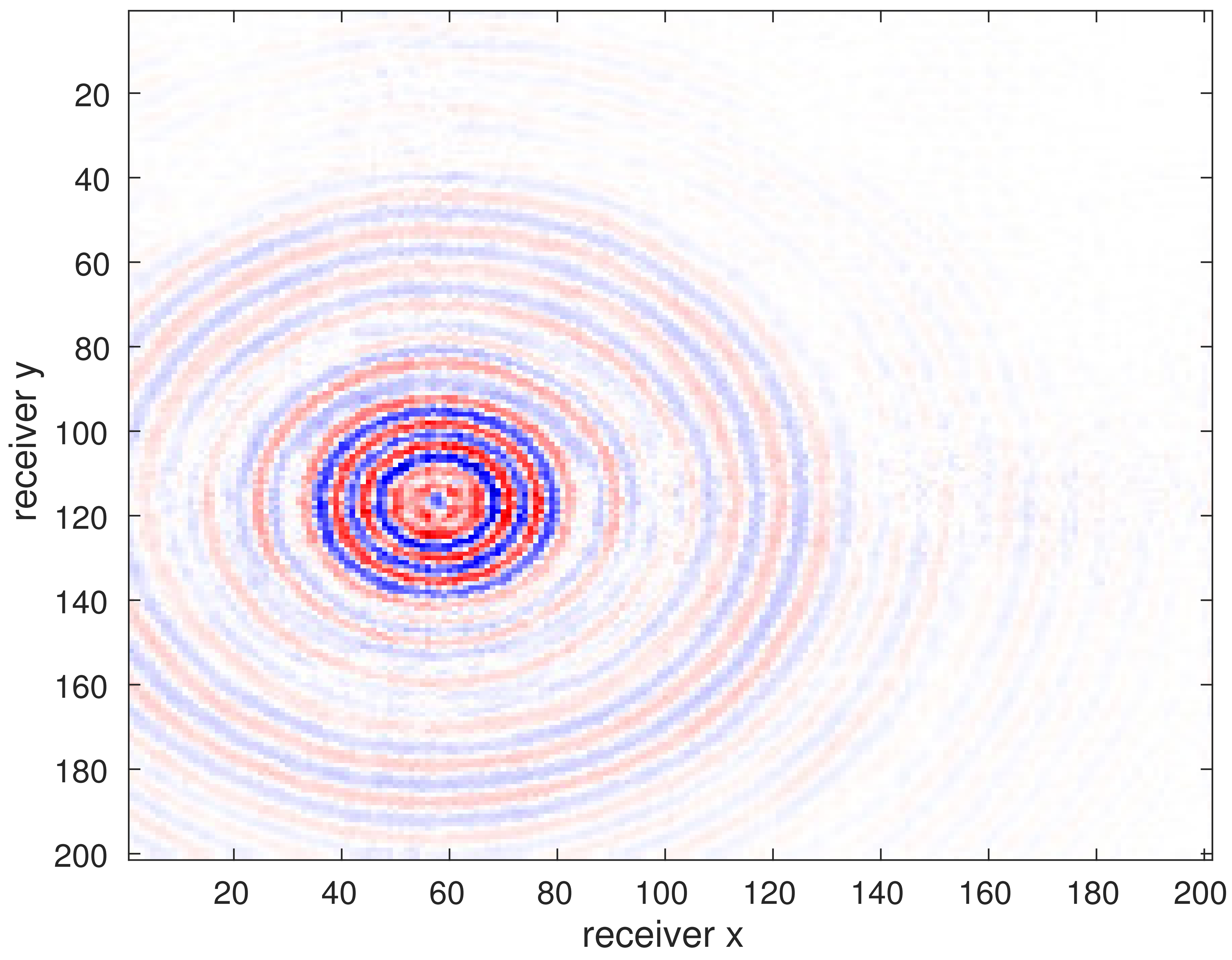


Vanilla tensor completion
SNR 6.95 dB

7.34 Hz - 90% missing receivers
Common source gather



True data



Regularized tensor completion
SNR 11 dB

Conclusion

3D seismic data has an underlying structure that we can exploit for interpolation (Hierarchical Tucker format)

Different schemes for organizing data - important for recovery

Conclusion

We can interpolate HT tensors with missing entries using the Riemannian manifold structure of the HT format

It is important and worthwhile to account for the off-grid nature of sampling when interpolating data

- binning just doesn't cut it

SLIM

Acknowledgements

Thank you for your attention

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