Low-rank Promoting Transformations and Tensor Interpolation - Applications to Seismic Data Denoising

Curt Da Silva and Felix J. Herrmann

June 20, 2014

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

Oropeza, and Sacchi. "Simultaneous seismic data denoising and reconstruction via multichannel singular spectrum analysis." Geophysics 2011.

Why window?

Embedding high dimensional data in an even higher dimensional space

- e.g. 4 dimensions -> 8 dimensions
- can’t handle large volumes

Window with k events -> rank k embedding

- sensitive to choice of k-parameter, no automatic way to choose it

Why window?

SVD-free matrix completion

- applied to the *original* data volume
- *no need* to embed data in higher dimensions
- # parameters << ambient dimension
- less sensitive to rank parameters

Can we still window to reduce computational costs?

Windowing vs signal model

Low rank matrix completion

$$\underset{\mathbf{X}\in\mathbb{R}^{m\times n}}{\text{minimize}} \quad \|\mathbf{X}\|_*$$
$$\text{such that } \|\mathcal{A}(\mathbf{X}) - \mathbf{B}\|_2 \leq \sigma$$

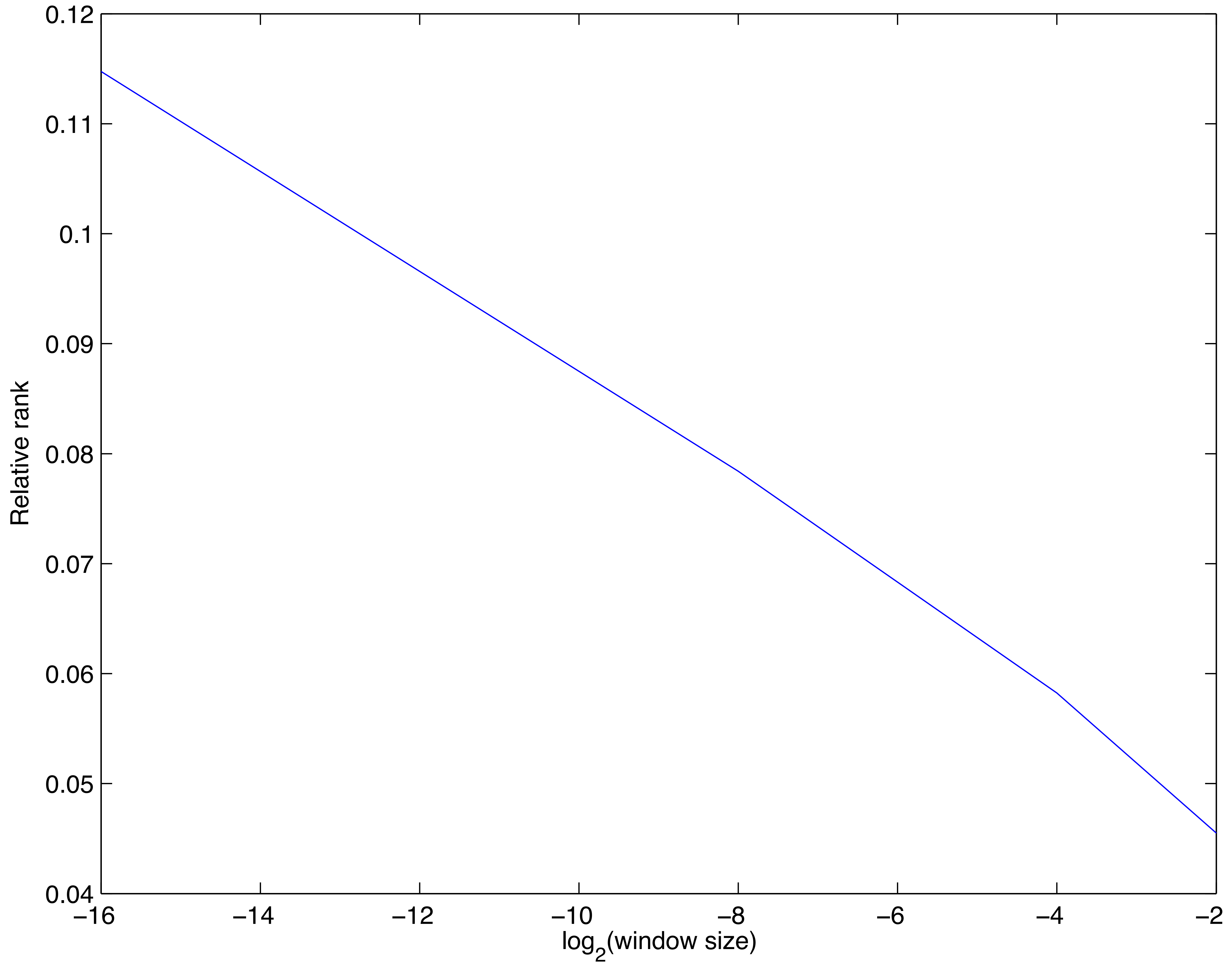
Underlying assumption: low *relative* rank

- not just that $\text{rank}(\mathbf{X})$ is small but $\dfrac{\text{rank}(\mathbf{X})}{\min(m,n)}$ is small

Windowing decreases ambient dimension and rank, but not necessarily *relative* rank

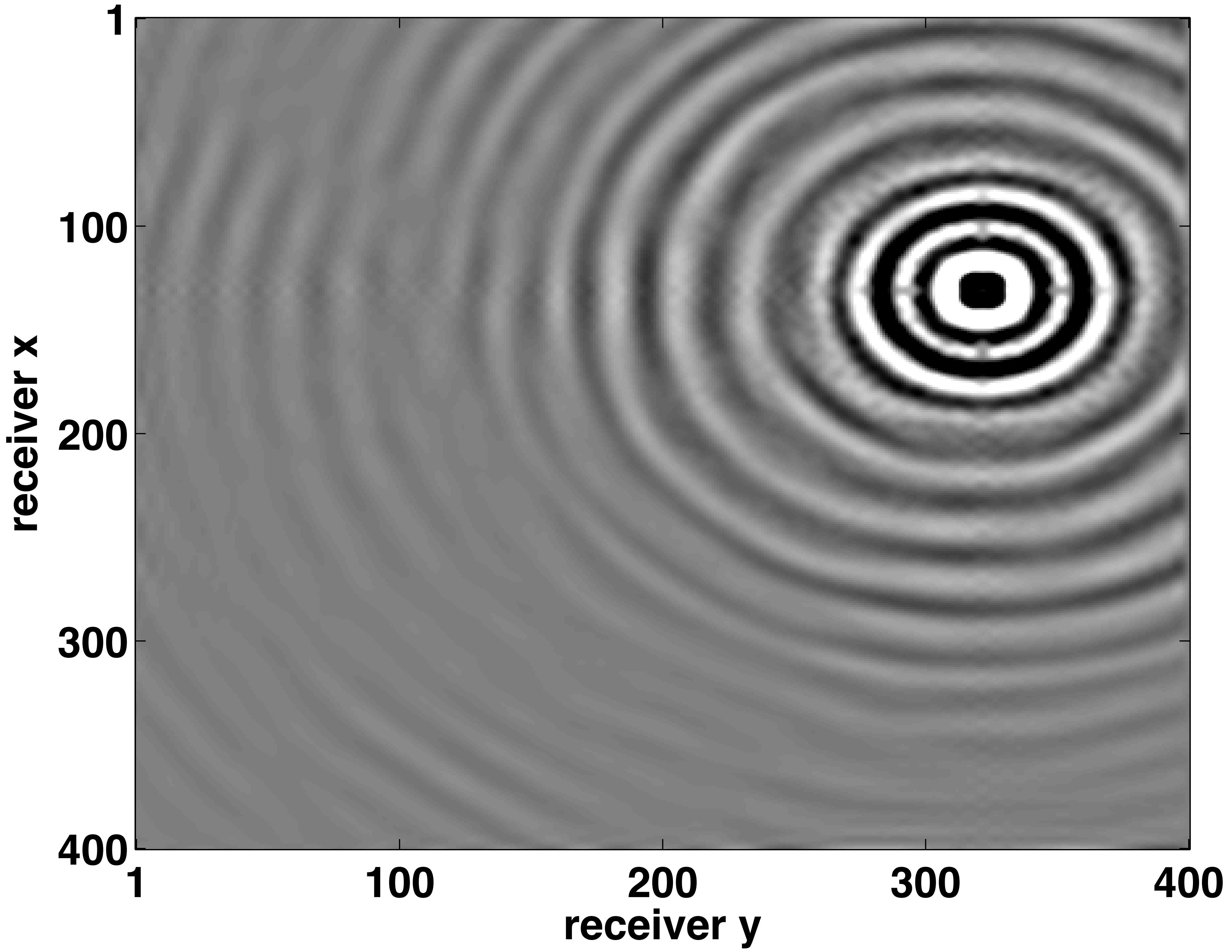
$$\frac{\mathrm{rank}(\mathbf{X})}{\min(m,n)}$$

smaller is better

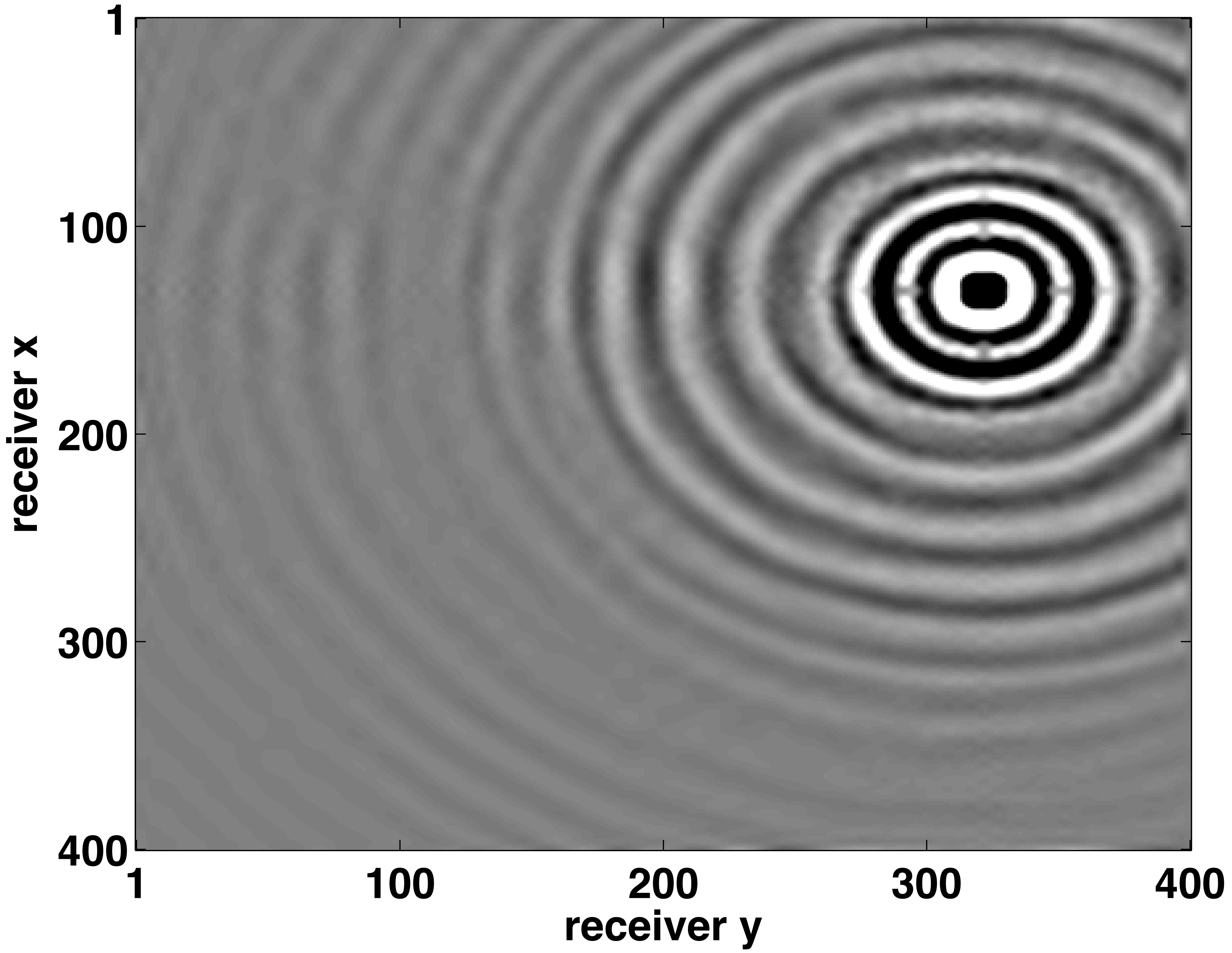


← smaller window bigger window →

68 x 68 sources, 101 x 101 receivers, 4.68 Hz
75% randomly missing sources

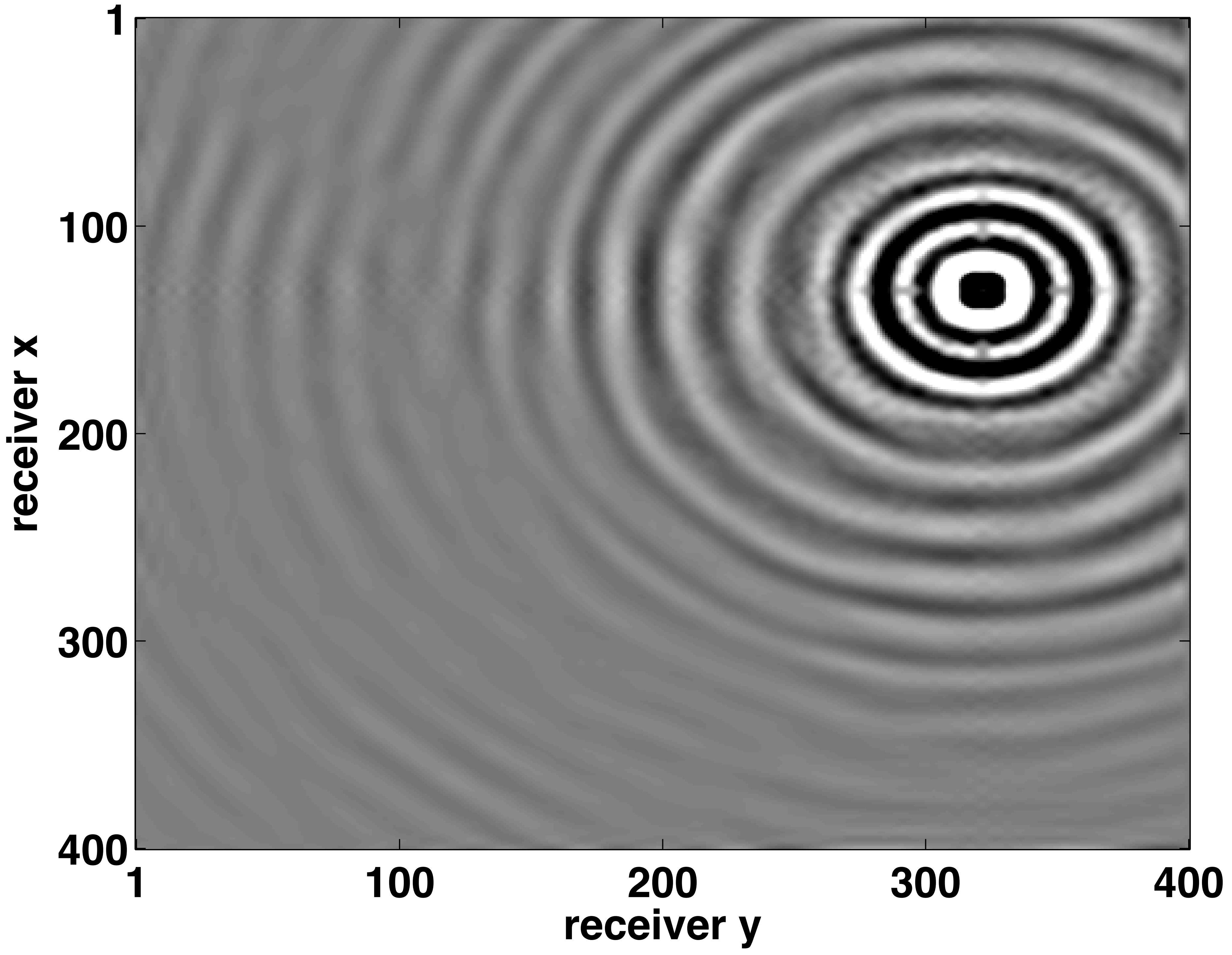


True data

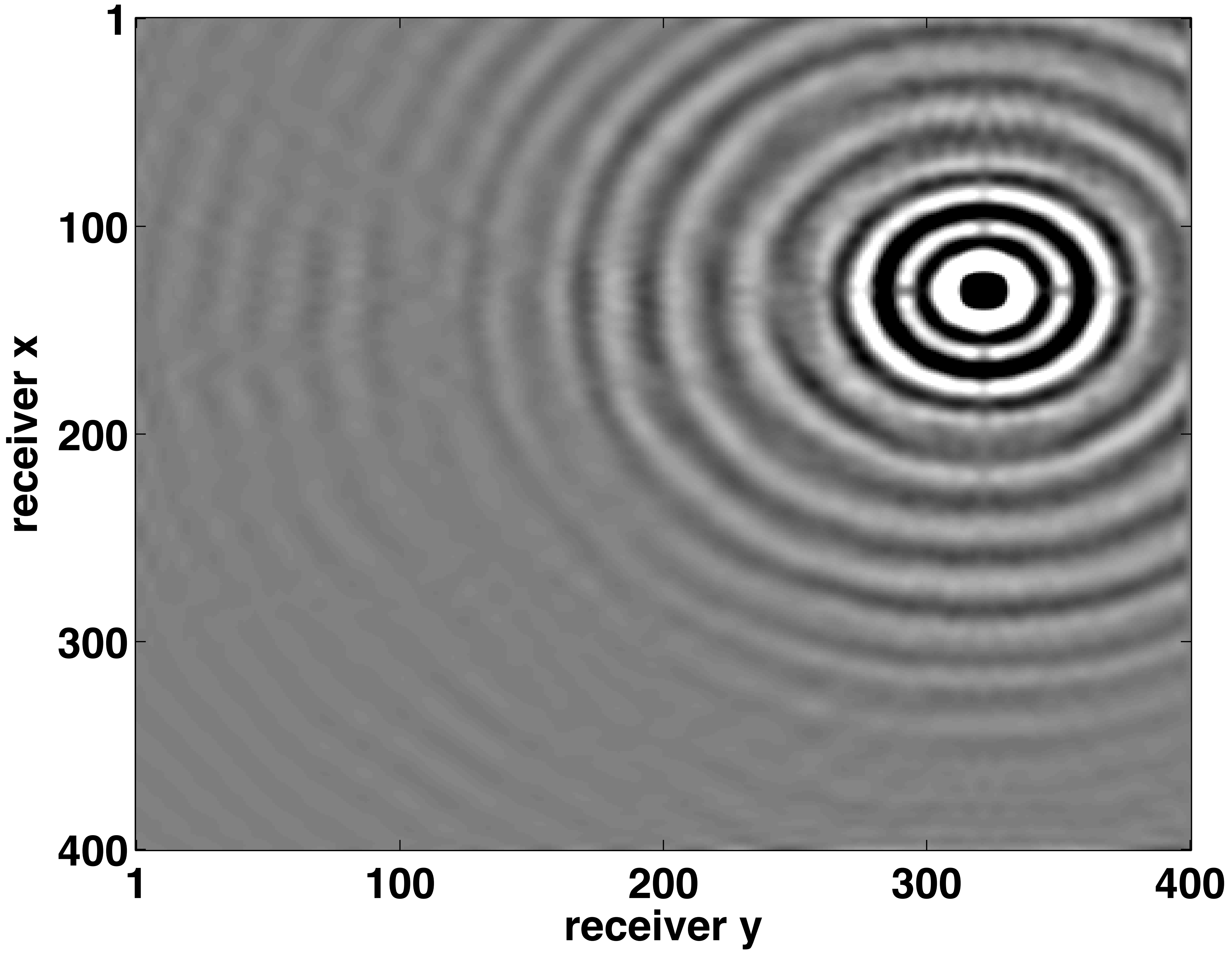


No windowing - SNR **16.7 dB**

68 x 68 sources, 101 x 101 receivers, 4.68 Hz
75% randomly missing sources

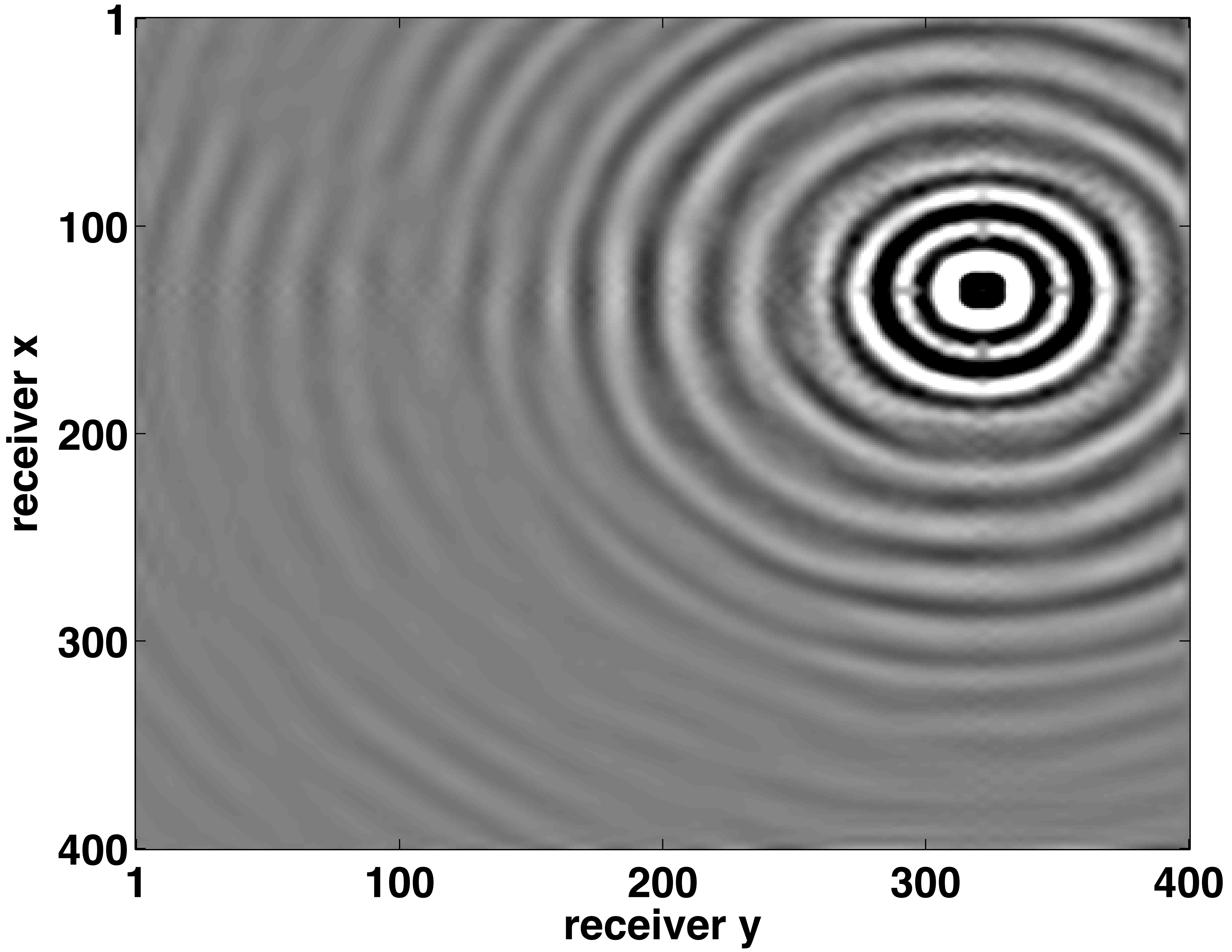


True data

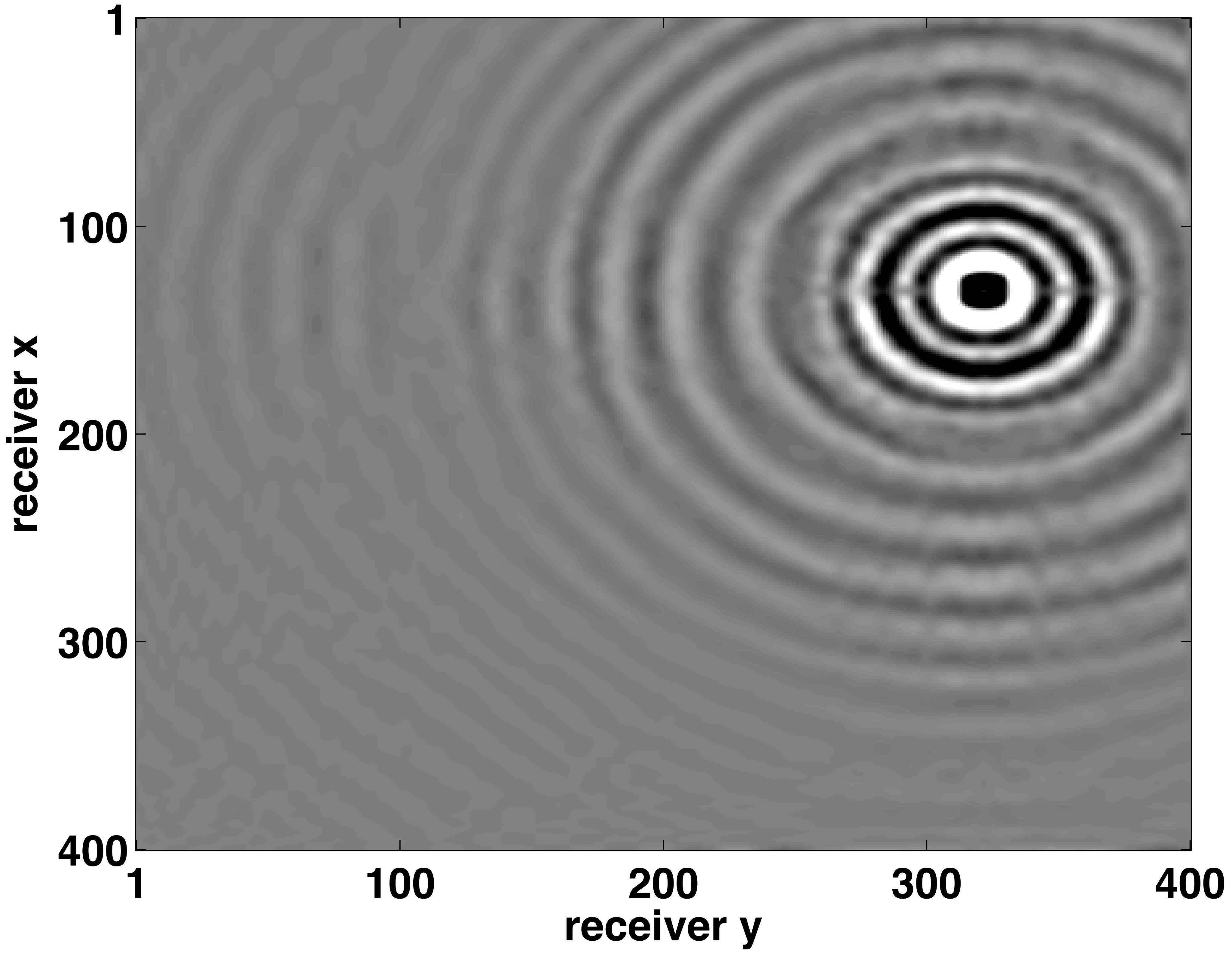


1/4th window - SNR **14.5 dB**

68 x 68 sources, 101 x 101 receivers, 4.68 Hz
75% randomly missing sources



True data



1/16th window - SNR **8.5 dB**

Matrix vs Tensor methods

$\mathbf{X}$ - $n_{\text{src}} \times n_{\text{src}} \times n_{\text{rec}} \times n_{\text{rec}}$ tensor

Assume each *matricization*, $\mathbf{X}^{(i)}, i = 1, \ldots, 4$, is low-rank

$\mathbf{X}^{(i)}$ - ith dimension placed along rows, other dimensions along the columns

Kreimer, Stanton, Sacchi. “Tensor completion based on nuclear norm minimization for 5D seismic data reconstruction.” Geophysics, 2013

SLIM

Matrix vs Tensor methods

Data fit Low-rank in each dimension

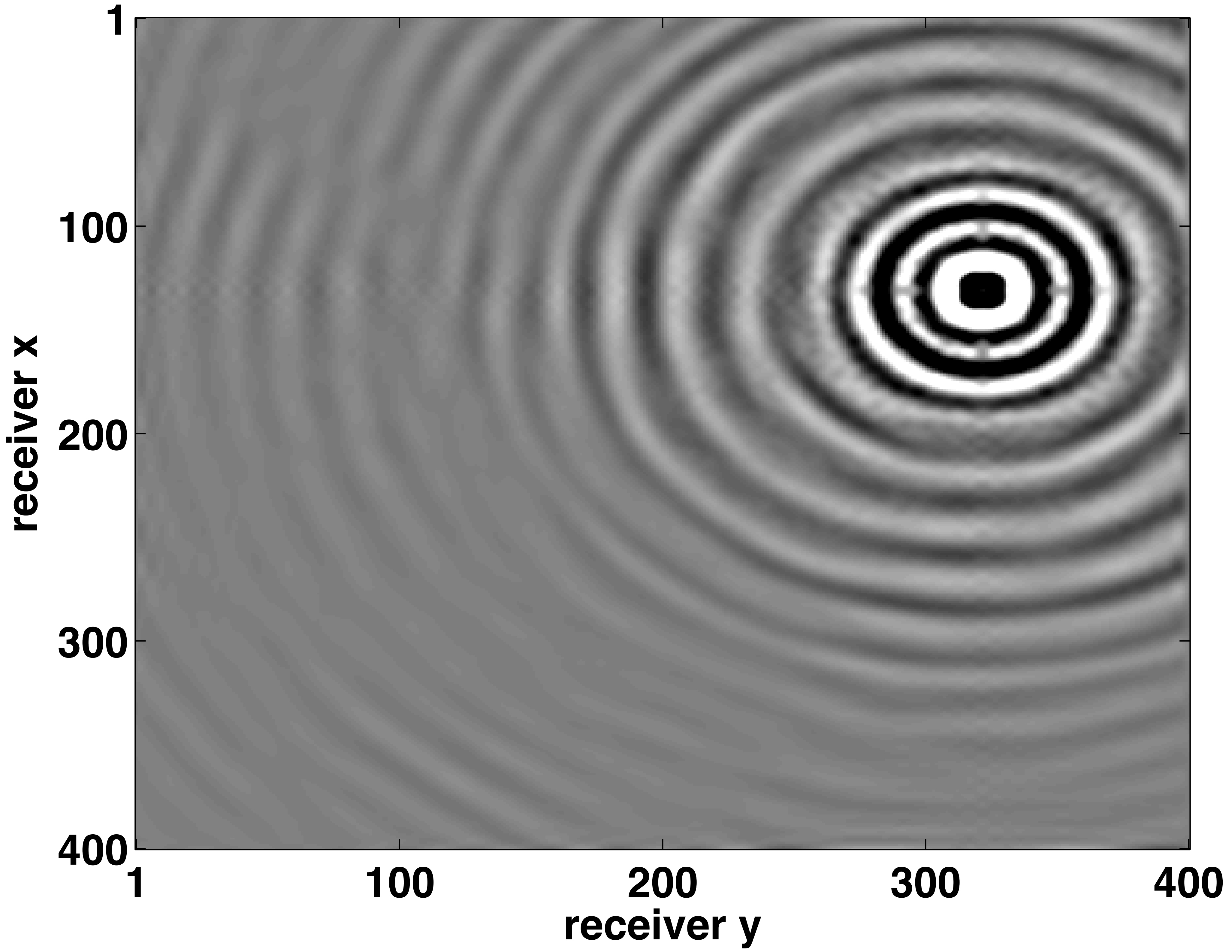
$$\underset{\mathbf{X}\in\mathbb{R}^{n_1\times n_2\times n_3\times n_4}}{\text{minimize}} \ \frac{1}{2}\|\mathcal{A}(\mathbf{X})-\mathbf{B}\|_2^2+\lambda\sum_{i=1}^{4}\|\mathbf{X}^{(i)}\|_*$$

Introduce $\mathbf{Y}_i = \mathbf{X}^{(i)}, i = 1, \ldots, 4$

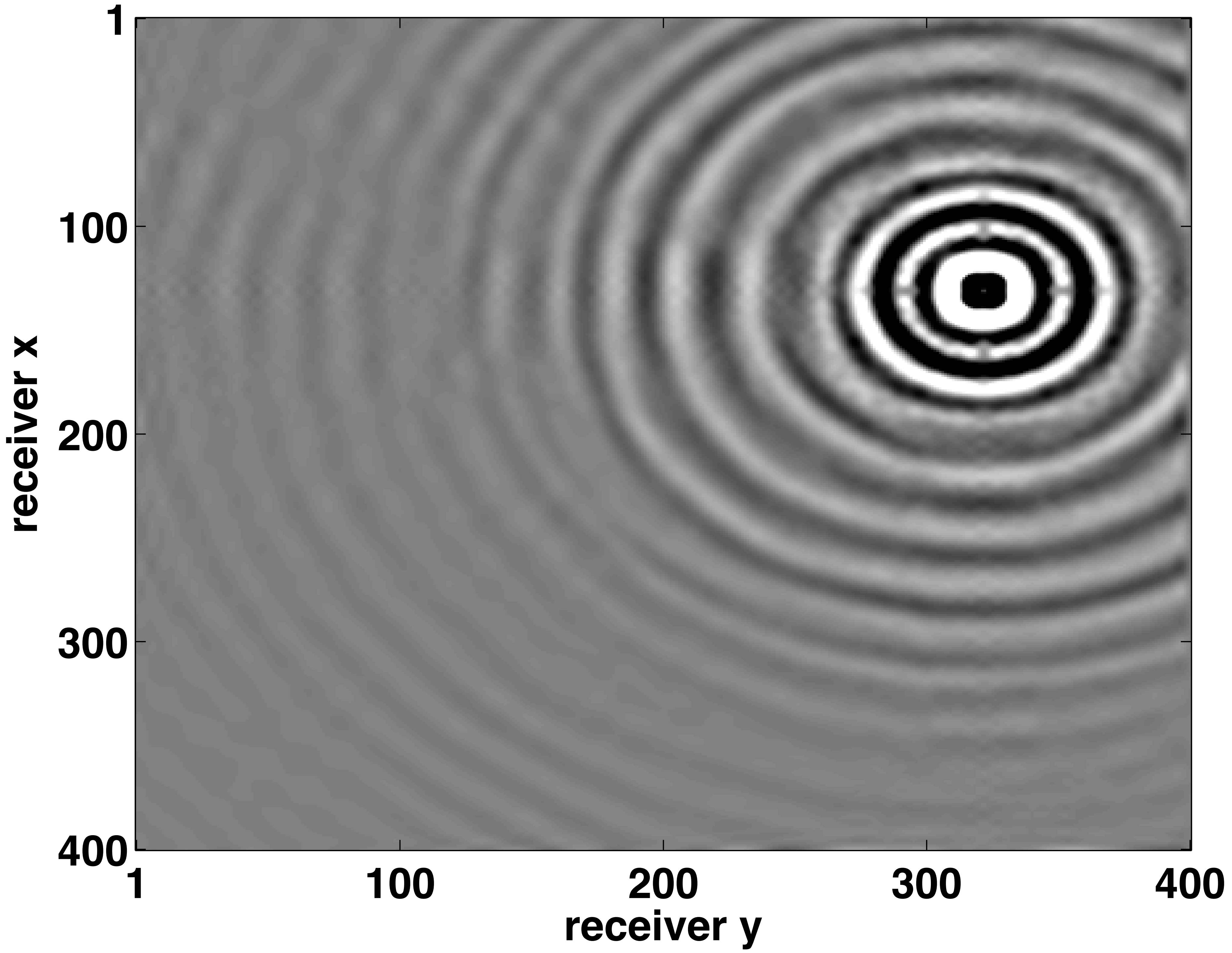
$$\underset{\mathbf{X},\mathbf{Y}_1,\ldots,\mathbf{Y}_4}{\text{minimize}} \ \frac{1}{2}\|\mathcal{A}(\mathbf{X})-\mathbf{B}\|_2^2+\lambda\sum_{i=1}^{4}\|\mathbf{Y}_i\|_*$$

$$\text{such that } \mathbf{Y}_i = \mathbf{X}^{(i)}$$

68 x 68 sources, 101 x 101 receivers, 4.68 Hz
75% randomly missing sources

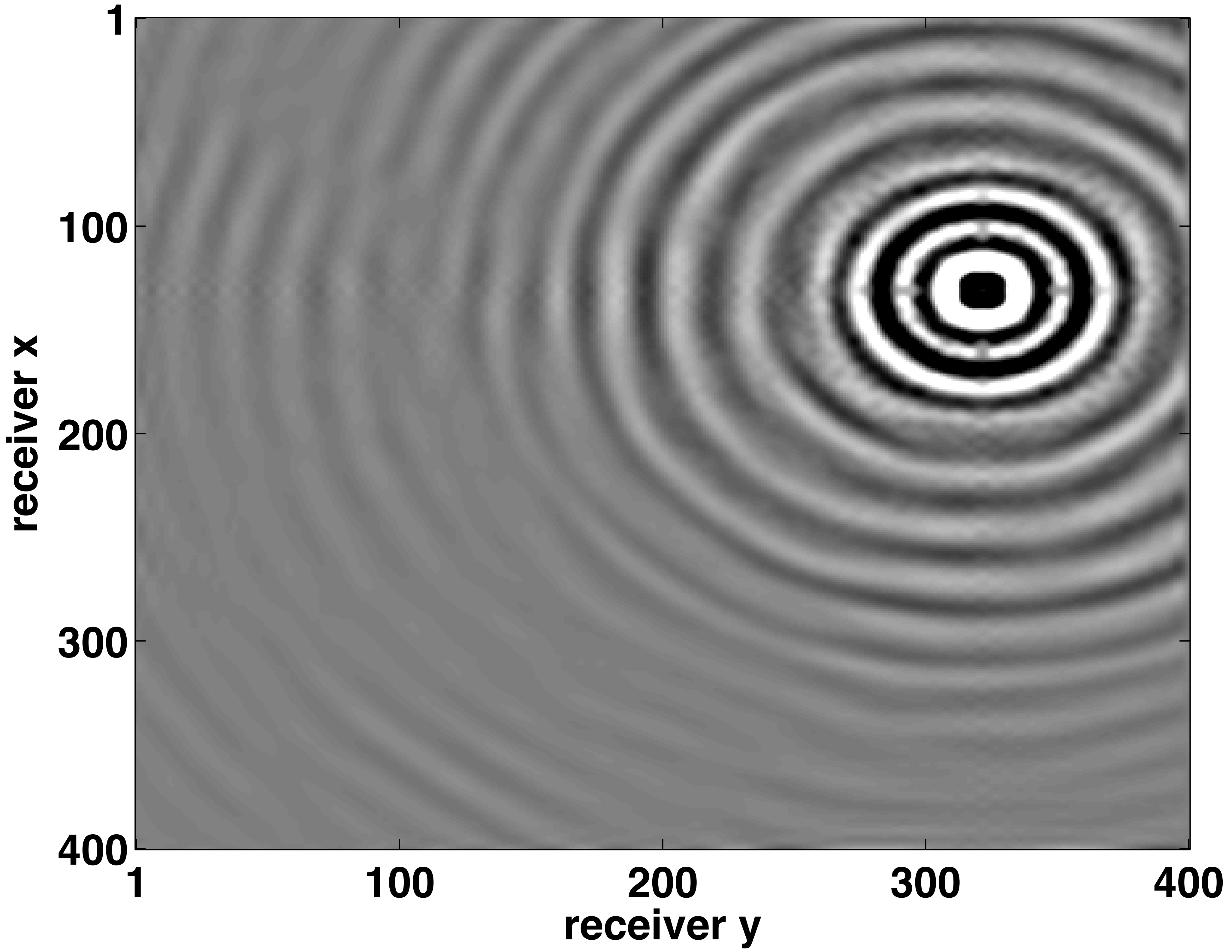


True data

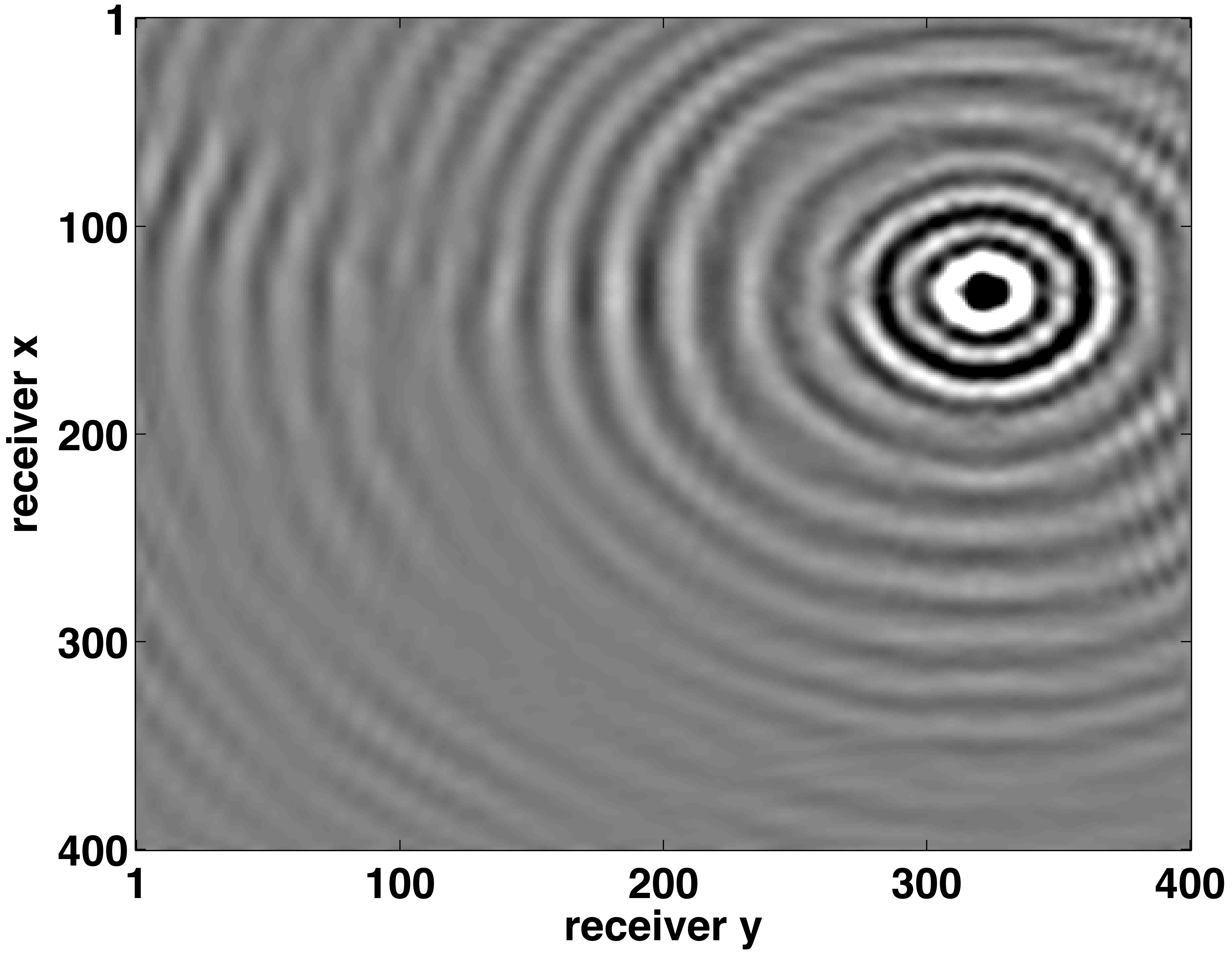


SVD-free MC - SNR **16.7 dB**
Time - **84 minutes**

68 x 68 sources, 101 x 101 receivers, 4.68 Hz
75% randomly missing sources

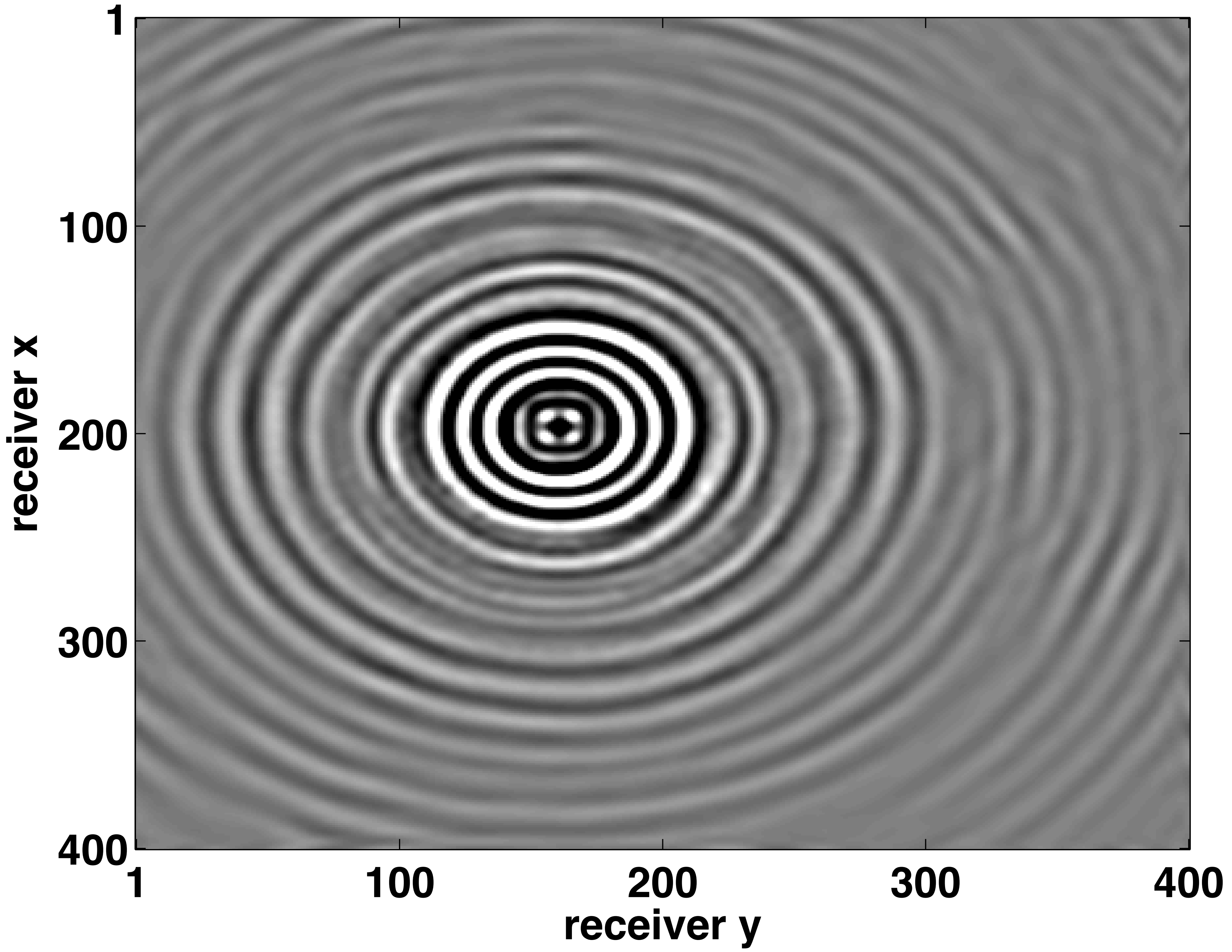


True data

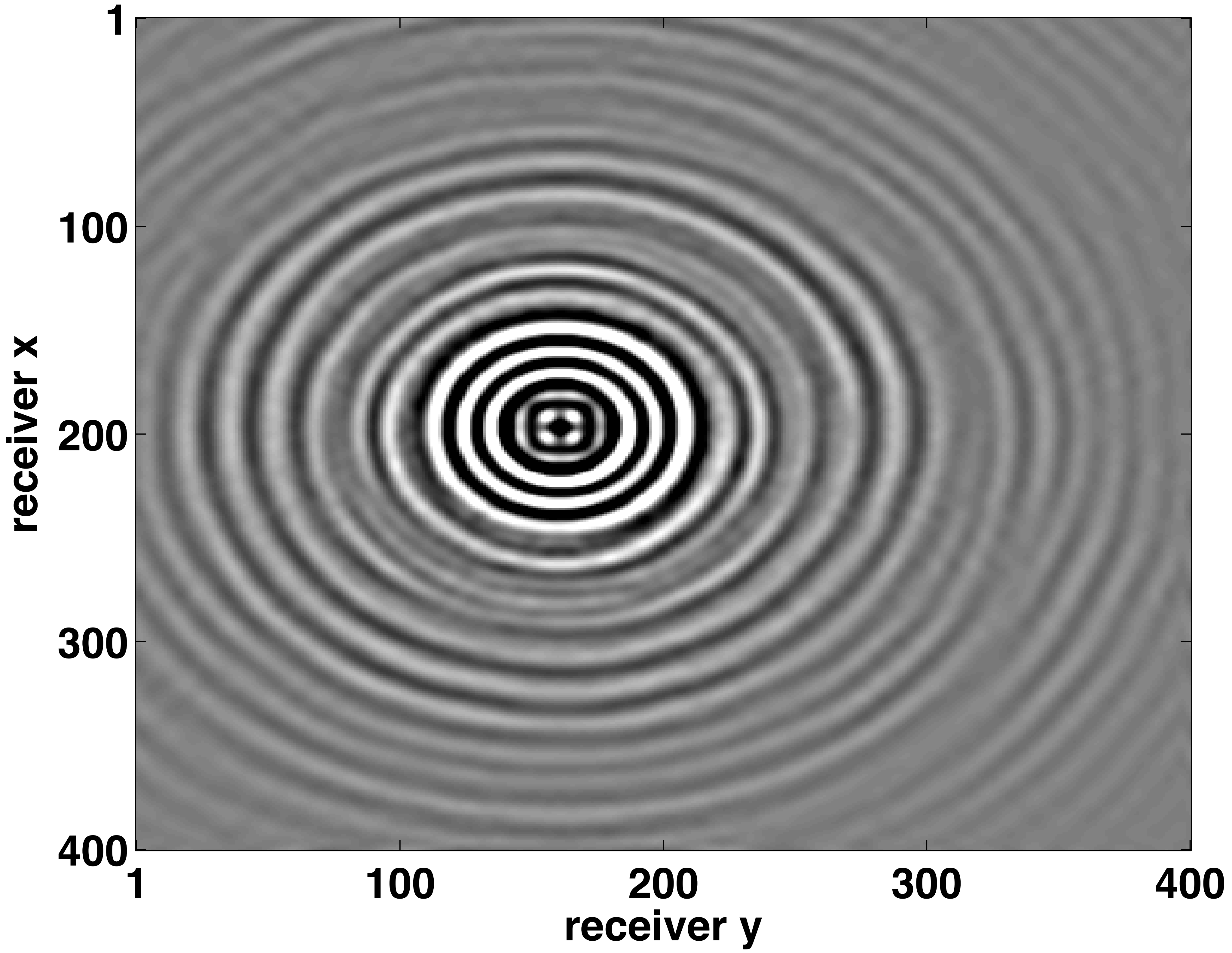


SVD-full Tensor - SNR **6.6 dB**
Time - **1510 minutes**

68 x 68 sources, 101 x 101 receivers, 7.34 Hz
75% randomly missing sources

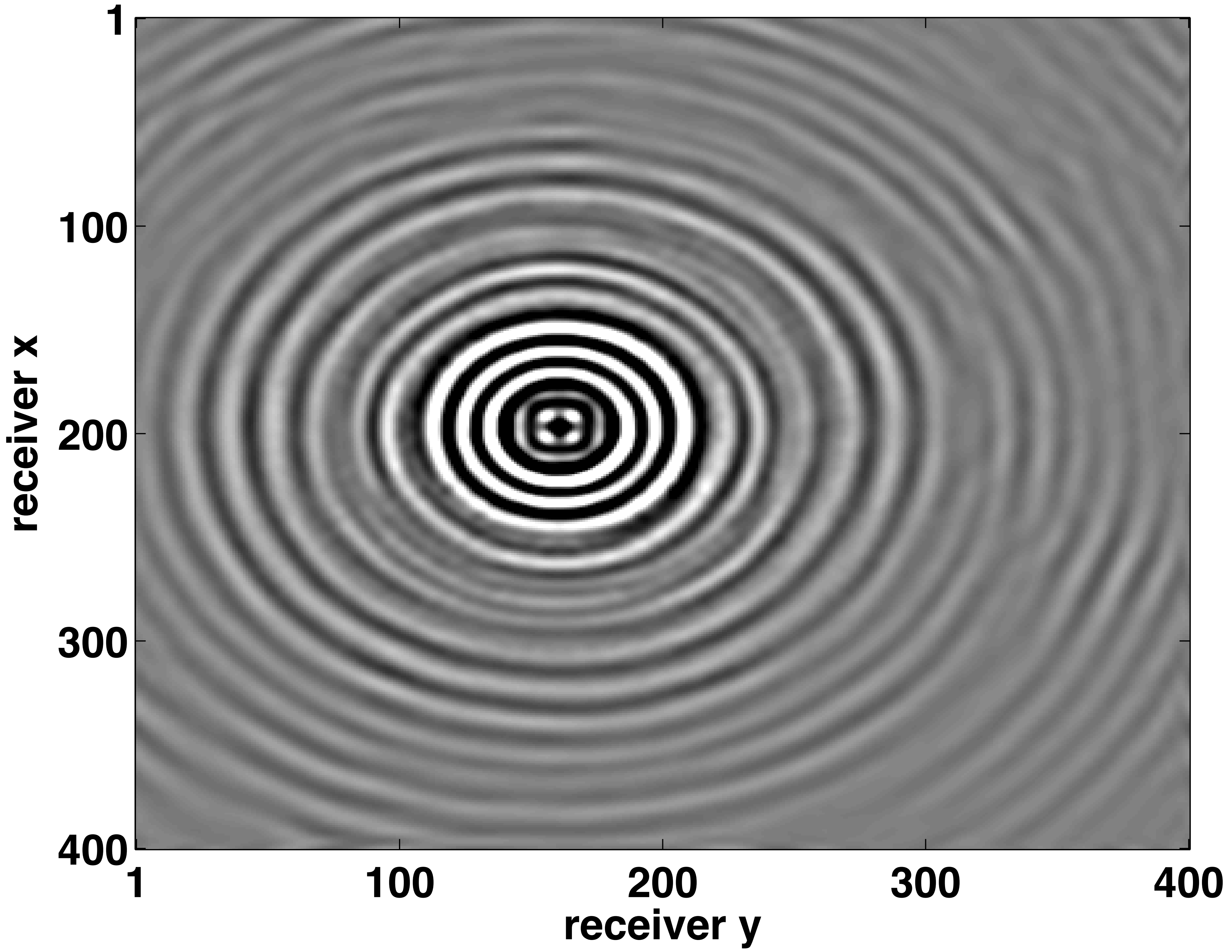


True data

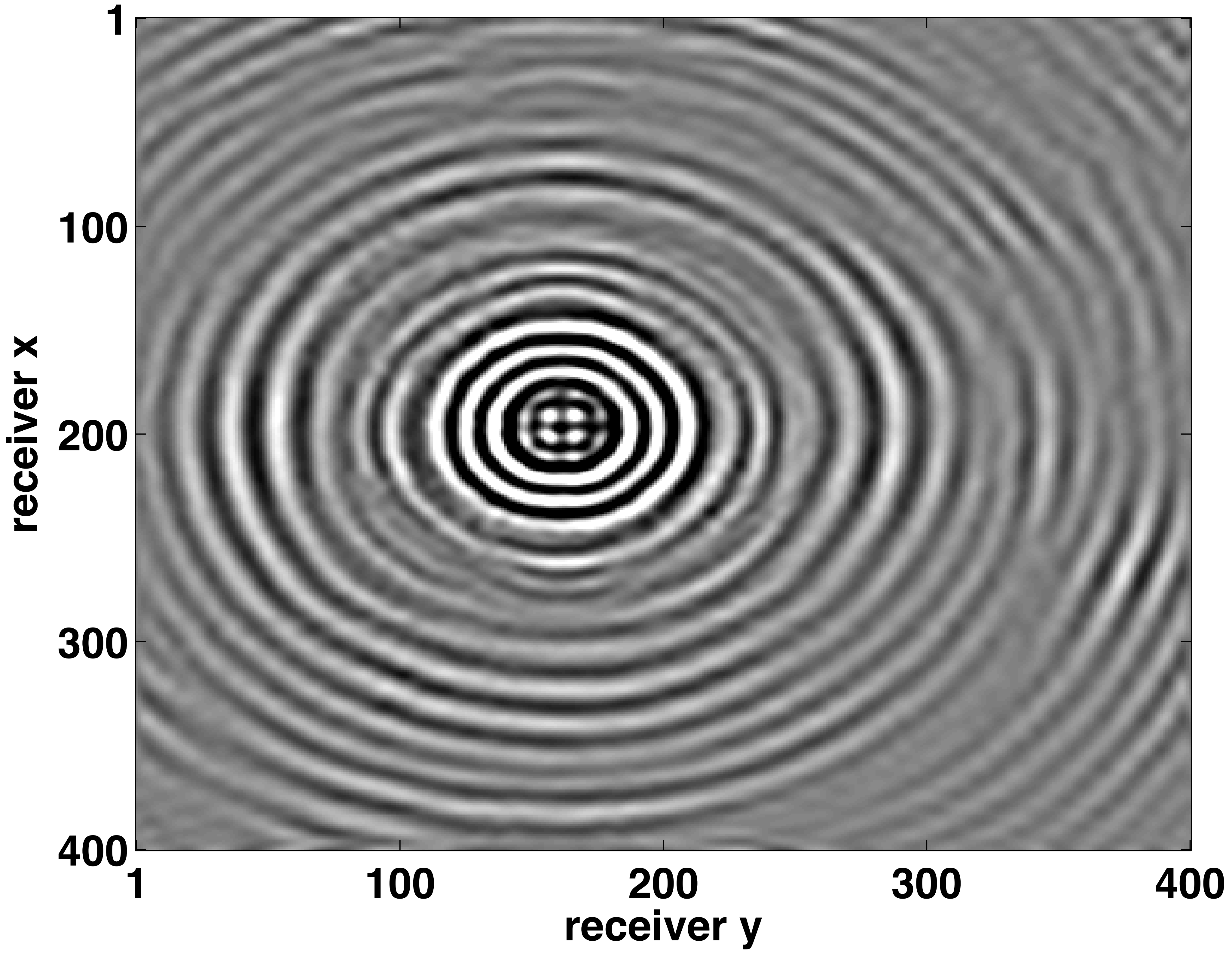


SVD-free MC - SNR **15.6 dB**
Time - **84 minutes**

68 x 68 sources, 101 x 101 receivers, 7.34 Hz
75% randomly missing sources



True data



SVD-ful Tensor - SNR **5.8 dB**
Time - **1512 minutes**

Summary

	SVD-free MC	SVD-ful Tensor
Rank parameter	Explicit, cheap to increase/ decrease	Implicit, expensive to estimate
Optimization variable	Much smaller than data set	At least 5 times the size of the data set
Solver	$SPG\ell1$-based, fast, automatic	Expensive per-iteration, needs parameter tuning

Upcoming paper: SVD-free 4D seismic data reconstruction

Practical principles of compressed sensing/matrix/tensor completion

SVD-free matrix completion vs SVD-ful tensor completion

Windowing?

And more!

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

Motivation

Unattenuated seismic noise can destroy the quality of a seismic image

- garbage in, garbage out

Statistics of the noise can be unknown

- high amplitude, localized
- caused by malfunctioning receivers, wildlife, ambient, unknown sources

[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, "SVD-free 4D seismic data reconstruction." Soon™

SLIM

Context

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

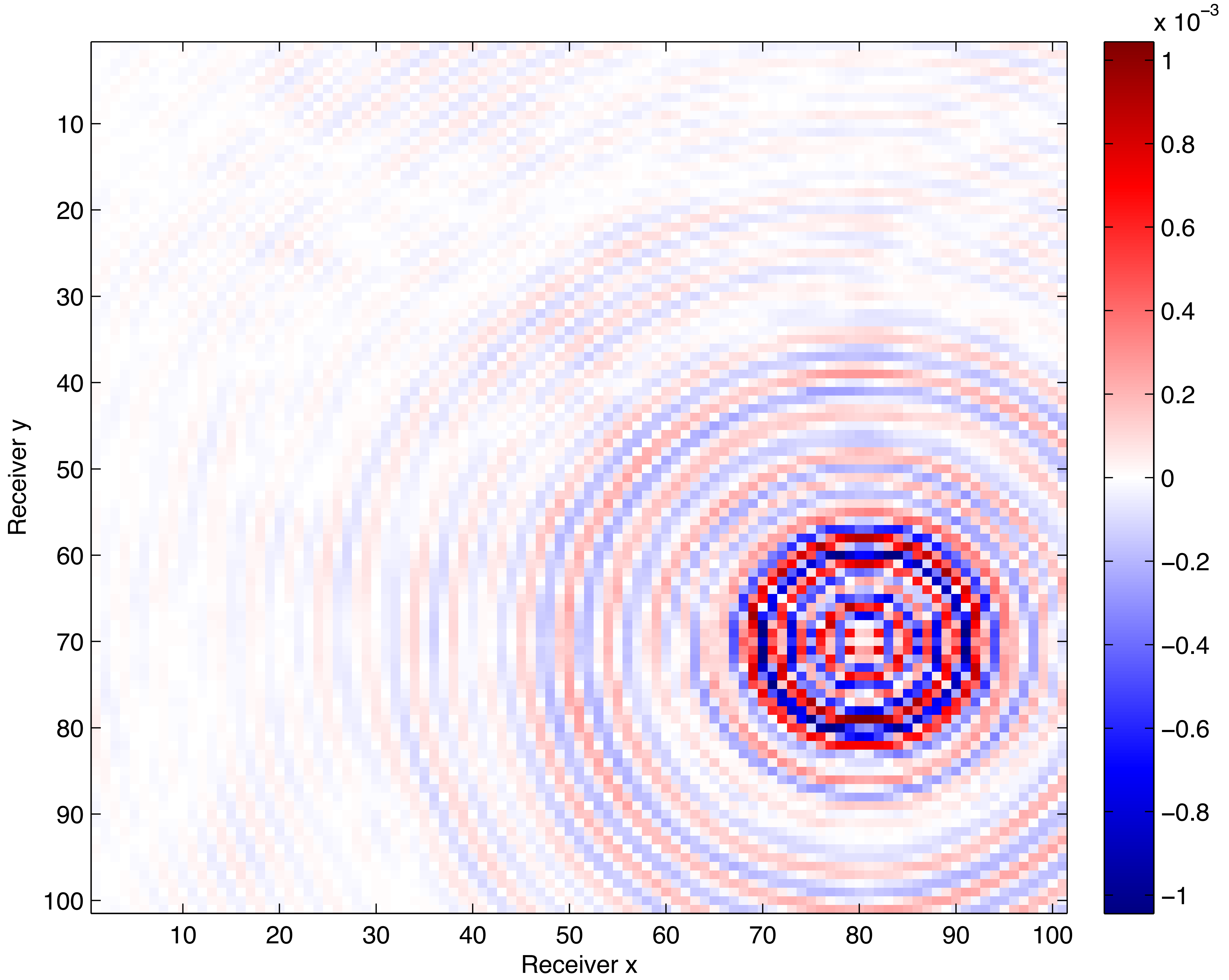
- Require SVDs on huge data matrices
- Not scalable to large problem sizes

Data completion via Toeplitz embedding [2]

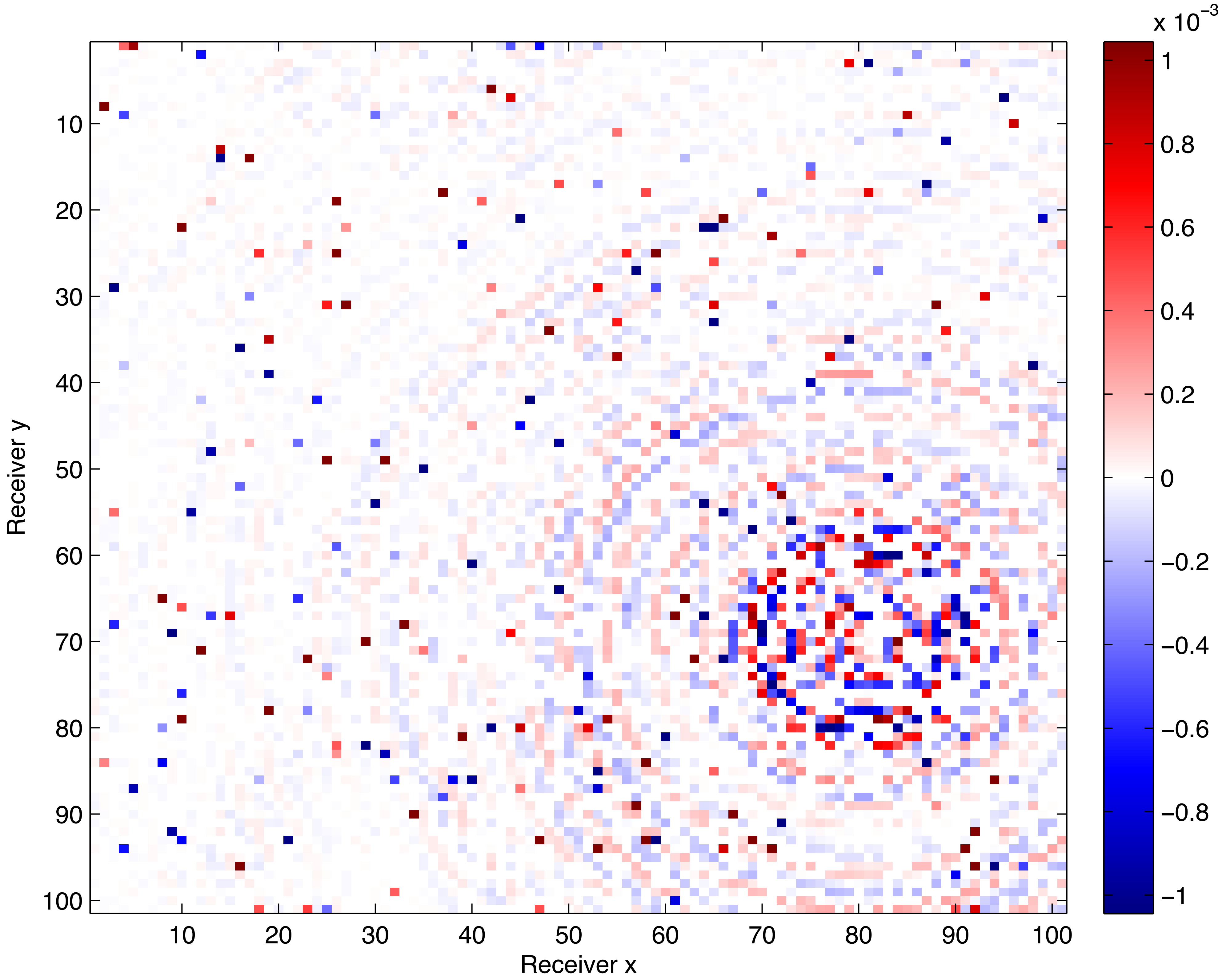
- Problem size - $(\# \text{ data points})^2$
- Ad-hoc windowing - can degrade quality, as demonstrated in [3]

7.34 Hz - 50% missing receivers - high noise

Common source gather



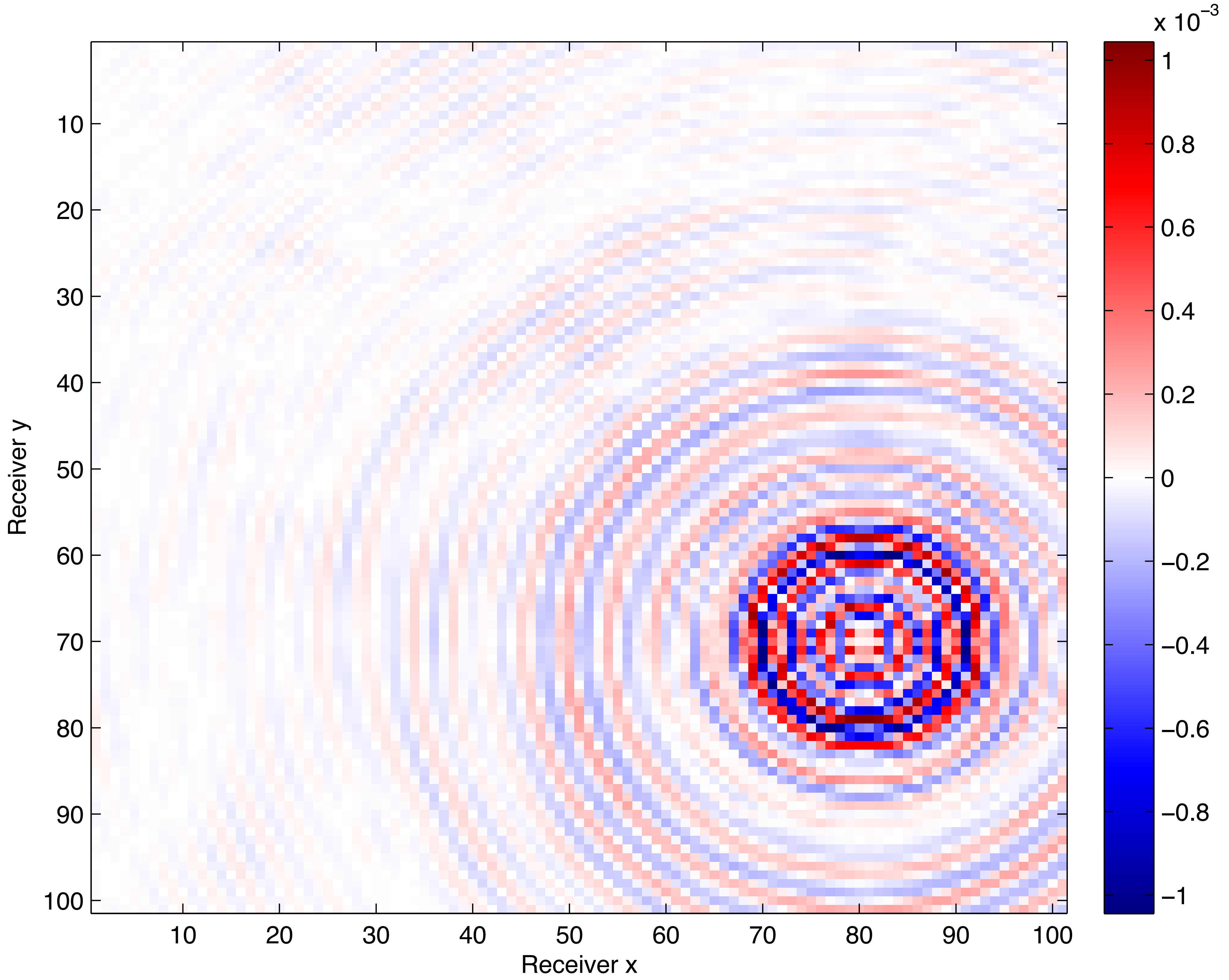
True data



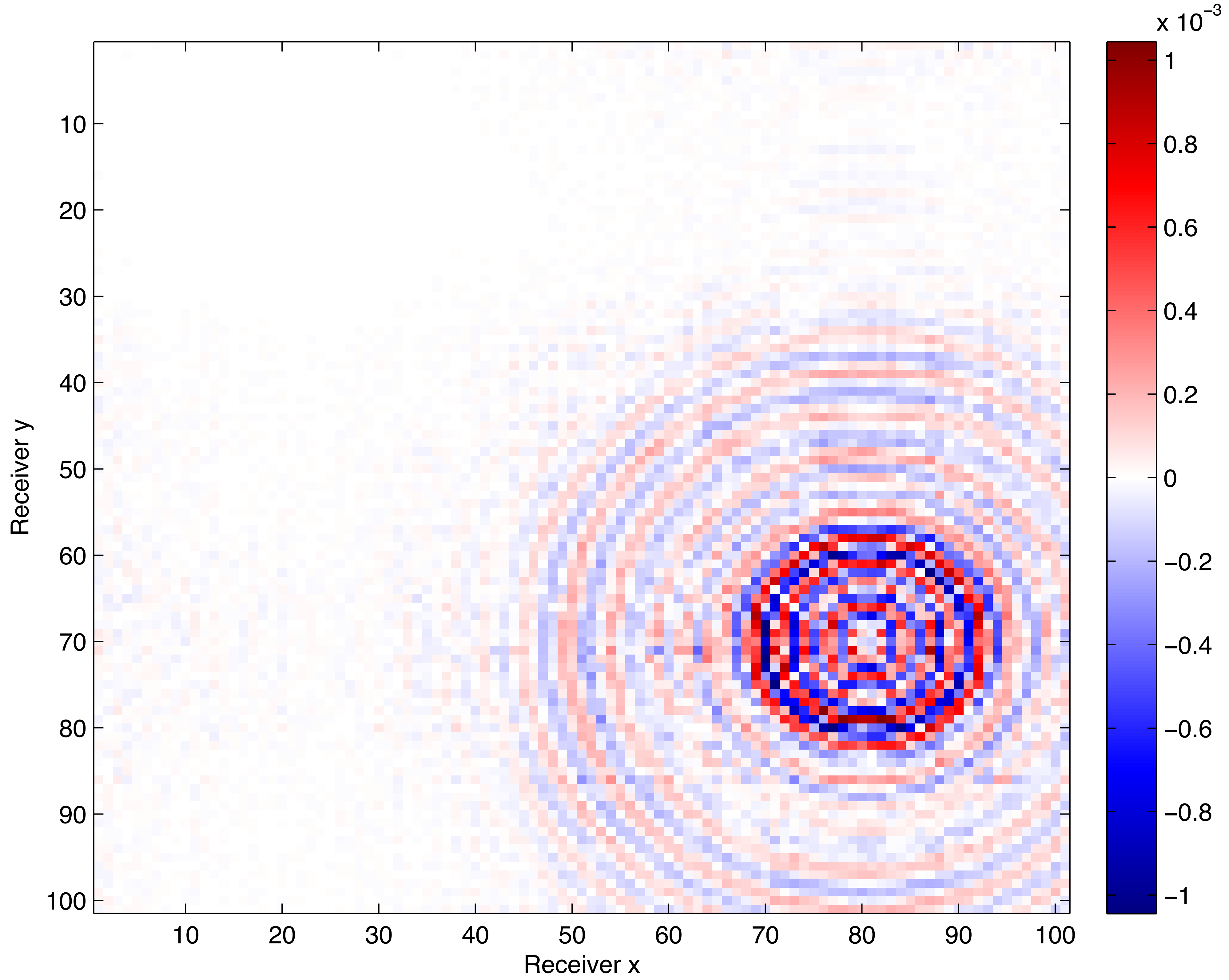
Subsampled data

7.34 Hz - 50% missing receivers - high noise

Common source gather



True data



MH Recovery - SNR 8.95 dB

Goals

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

Explore effect of transform domain on noise

- determine favourable recovery scenario

Multidimensional interpolation

with Hierarchical Tucker

Successful reconstruction scheme

Signal structure

- ***Hierarchical Tucker***

Sampling

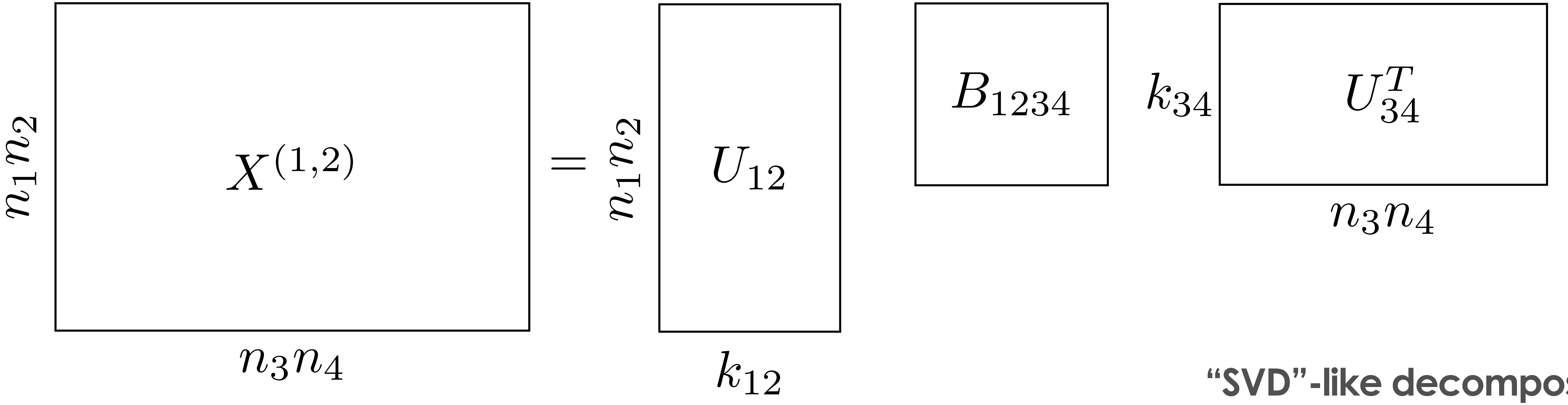
- subsampling, noise increases hierarchical rank

Optimization

- fit data in the Hierarchical Tucker format

Hierarchical Tucker format

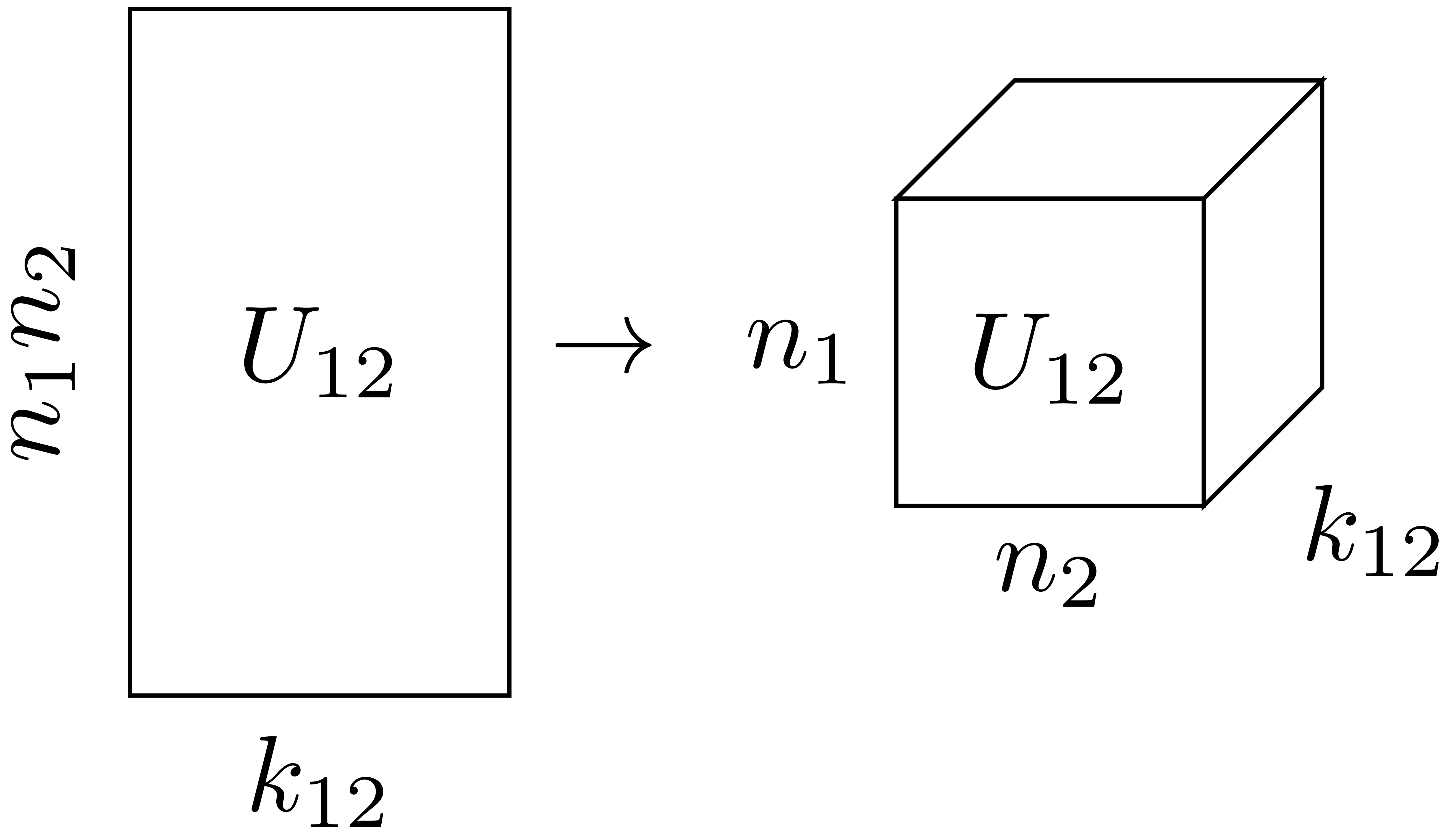
$X - n_1 \times n_2 \times n_3 \times n_4$ tensor



"SVD"-like decomposition

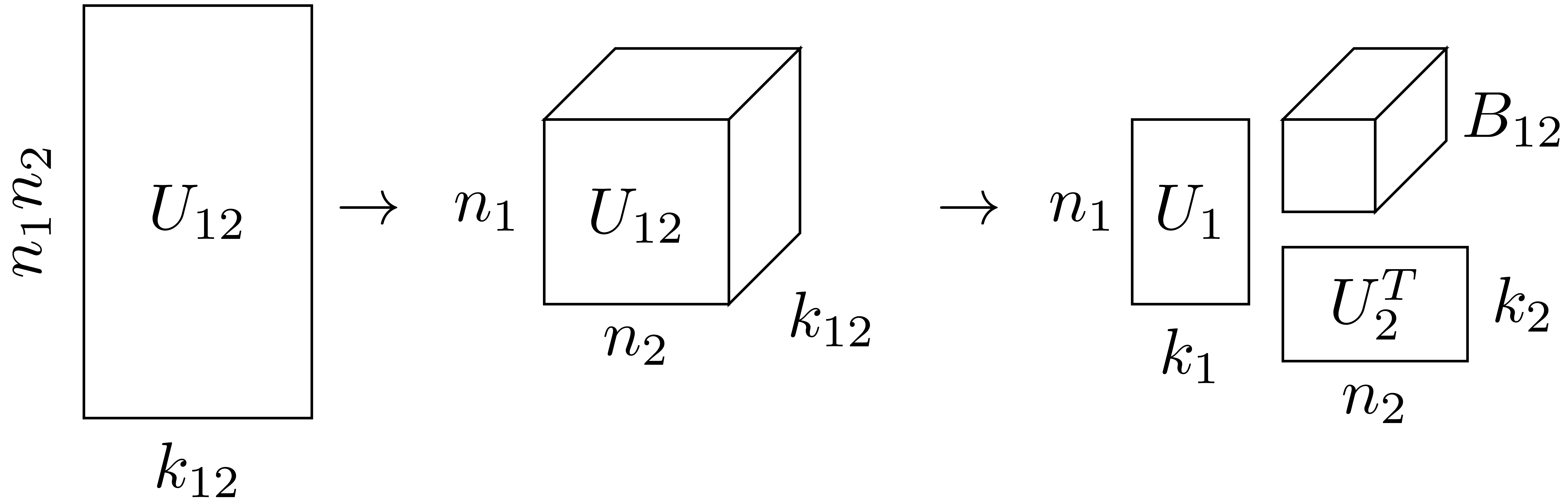
Hierarchical Tucker format

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

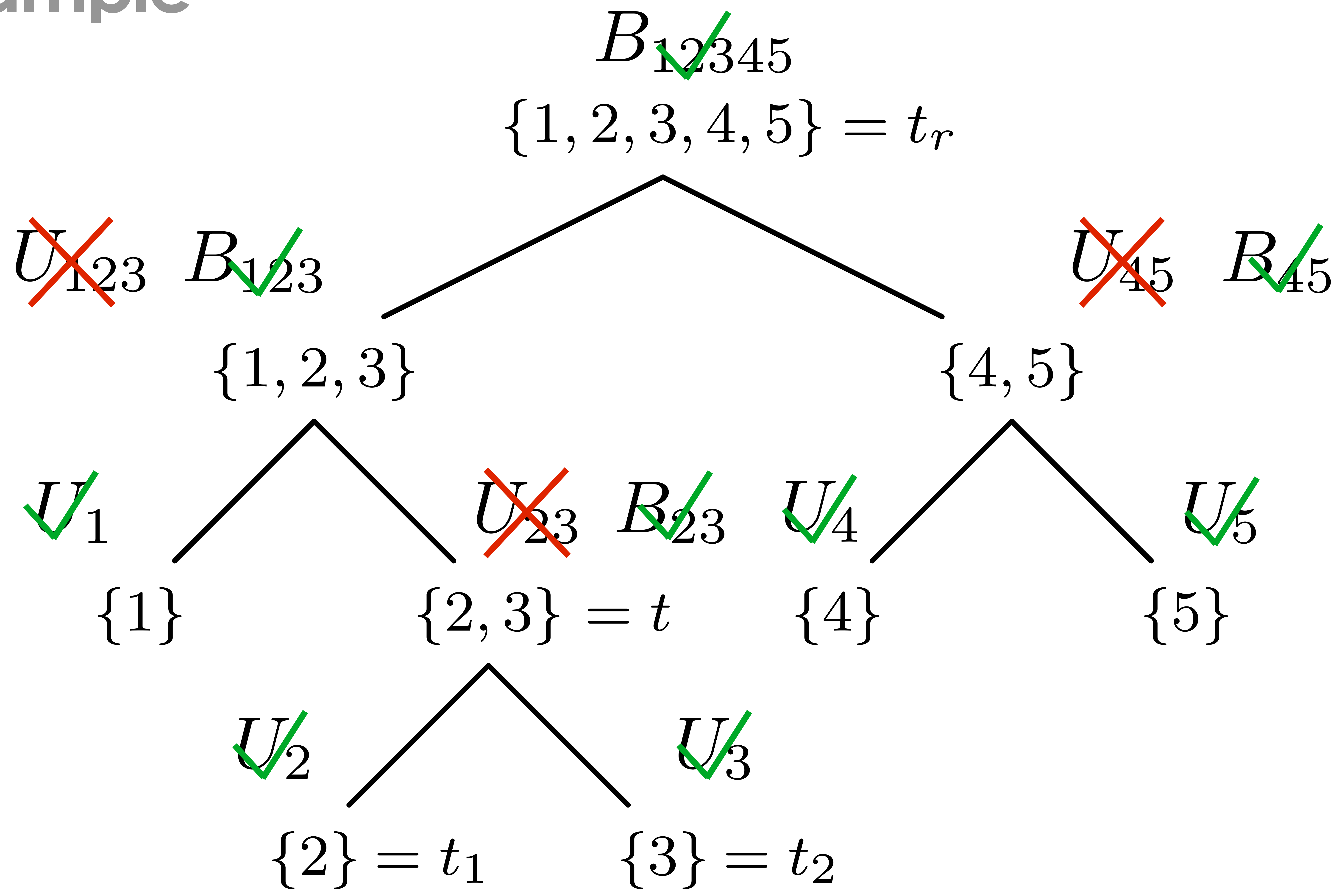
- specify the tensor completely

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.

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

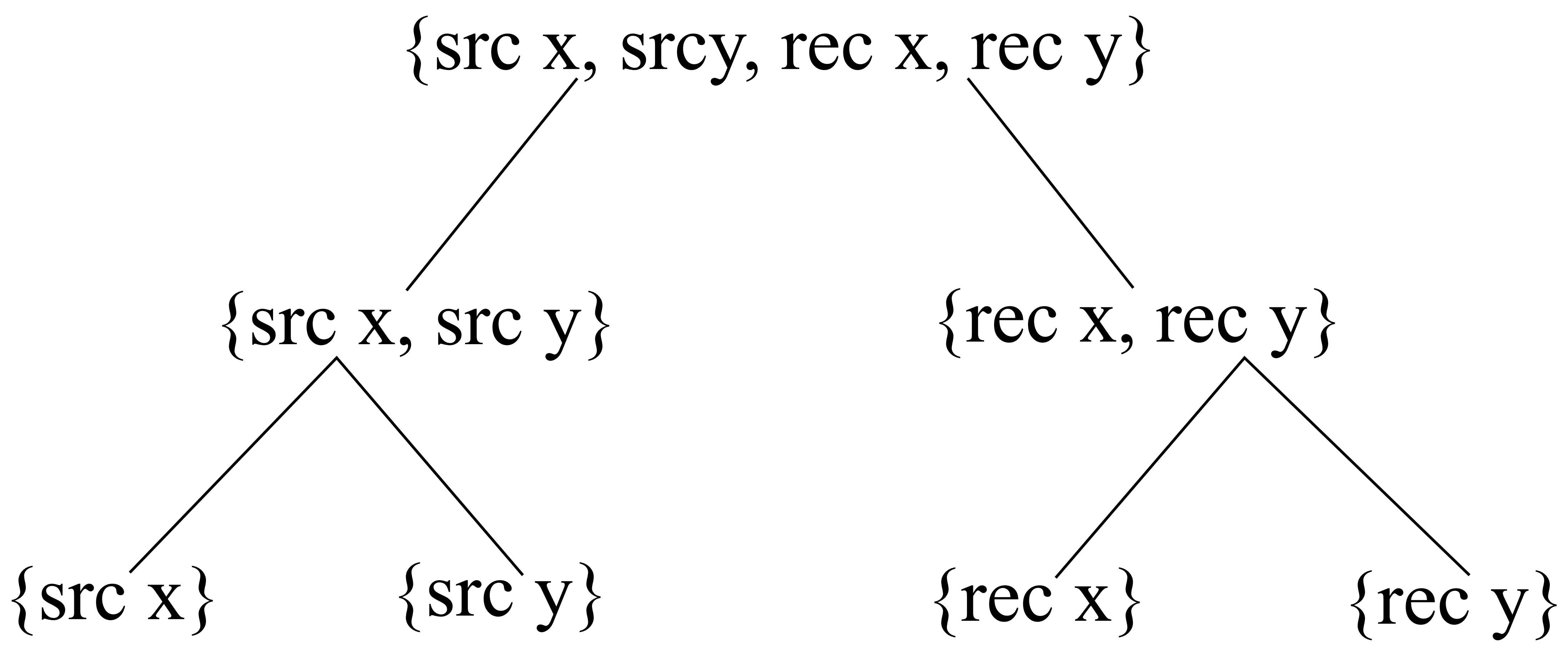
We consider a 3D seismic survey with coordinates
(src x, src y, rec x, rec y, time)

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

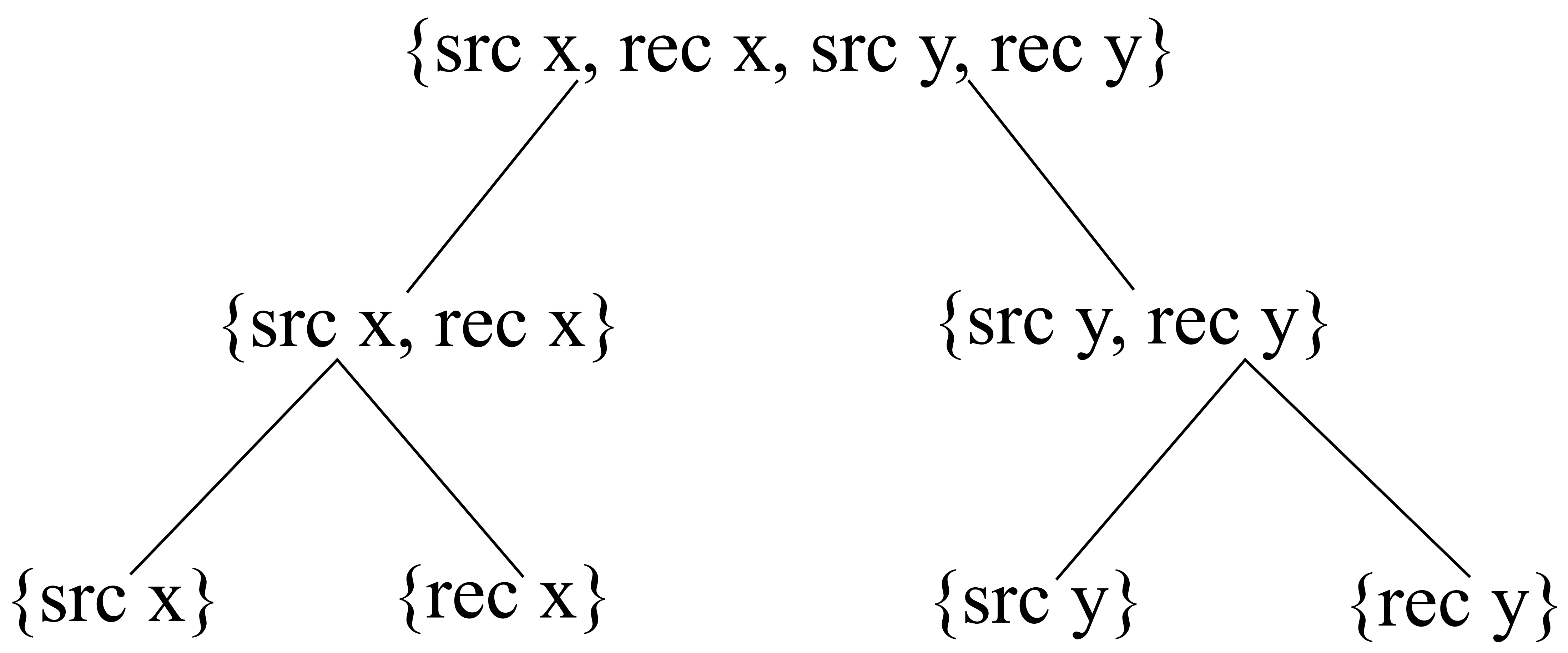
Seismic Hierarchical Tucker

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

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

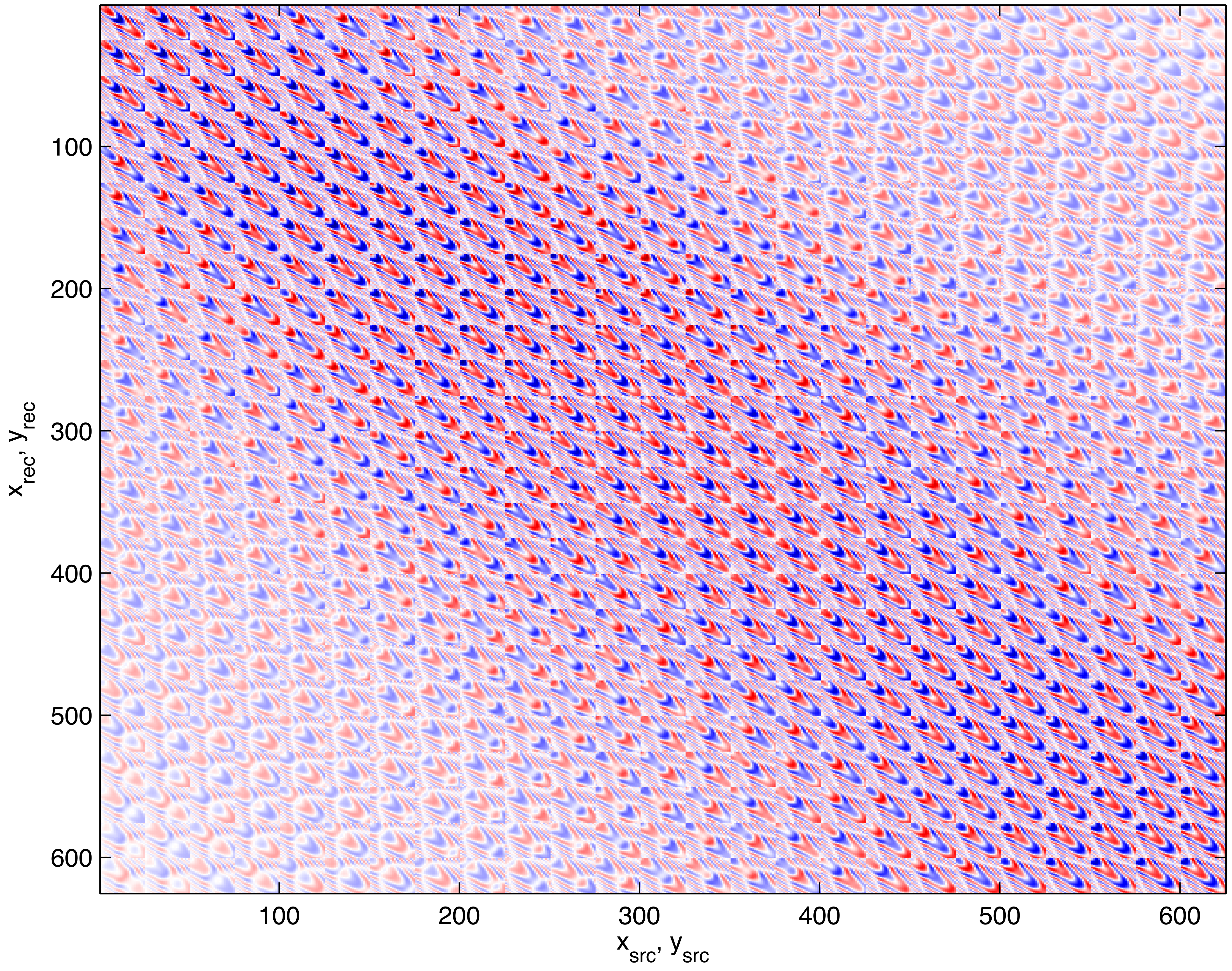


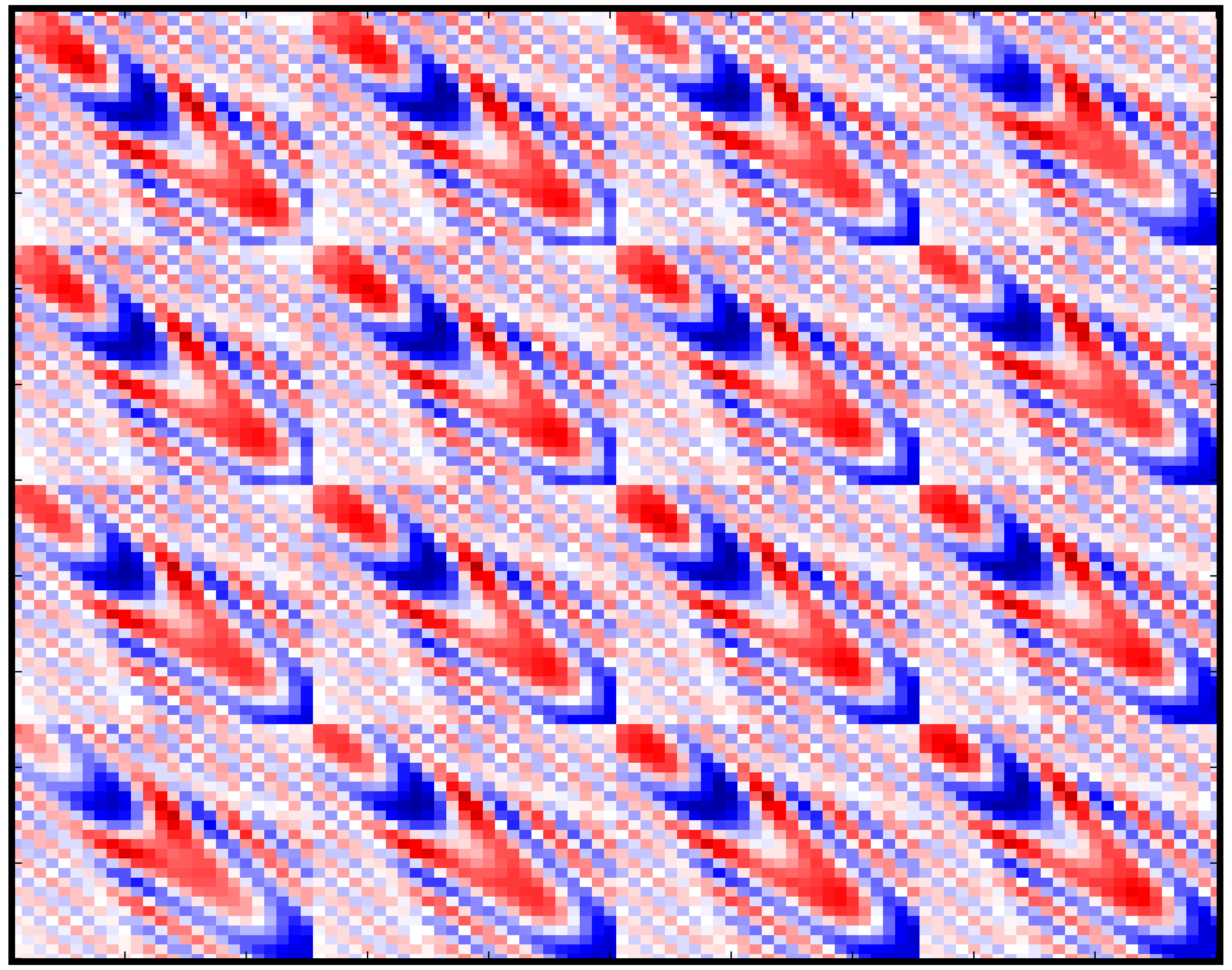
Canonical Decomposition

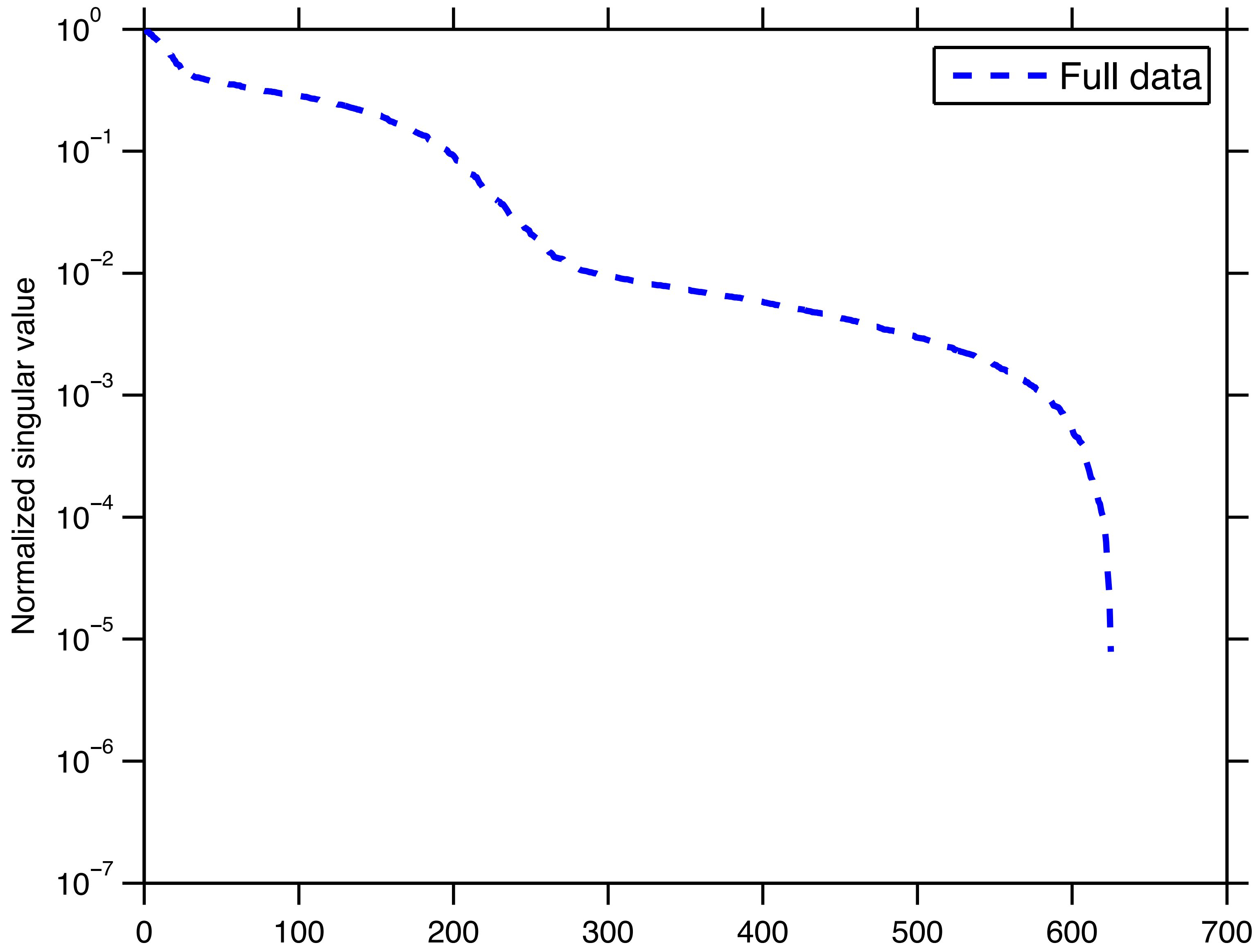


Non-canonical Decomposition

Matricizations

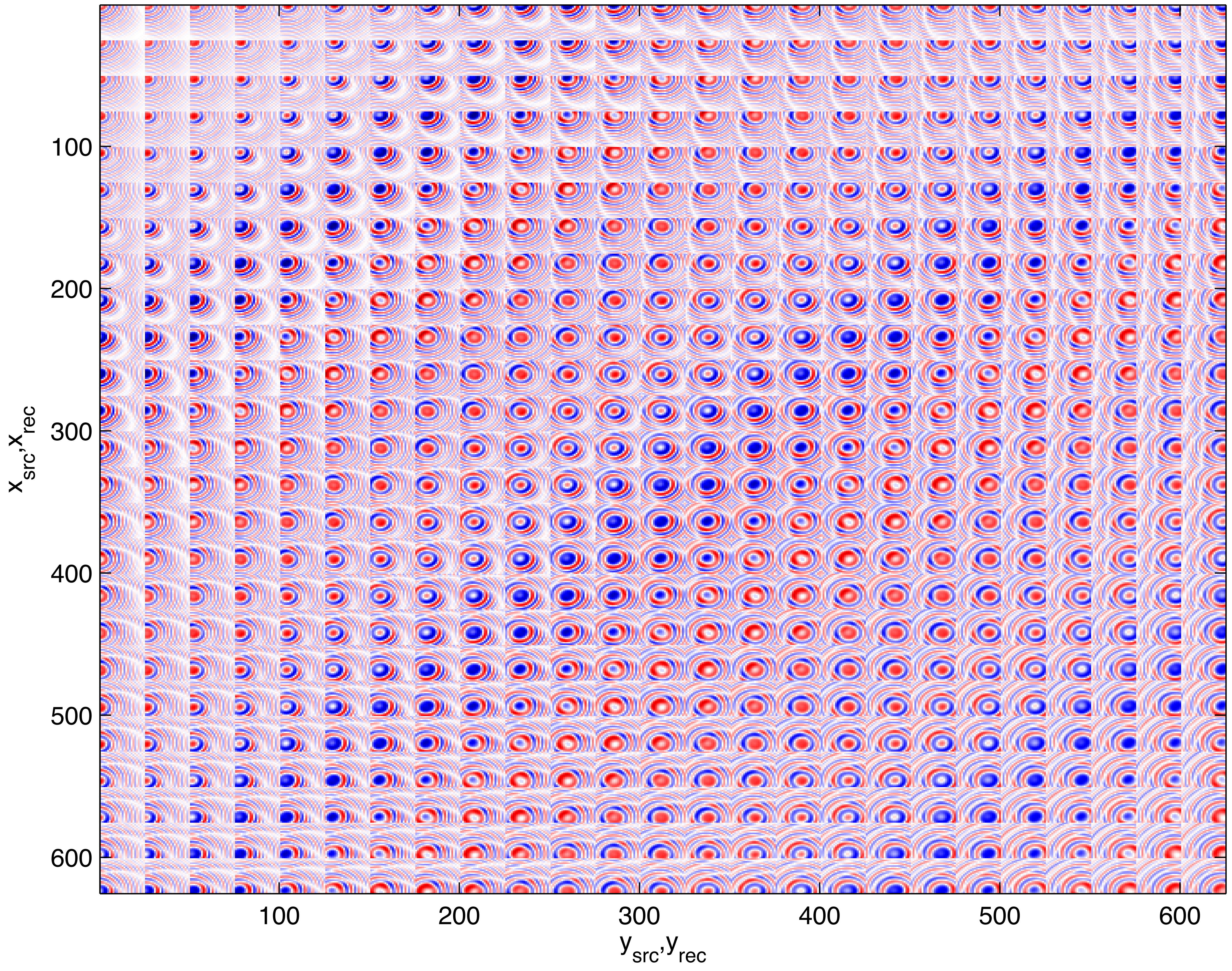


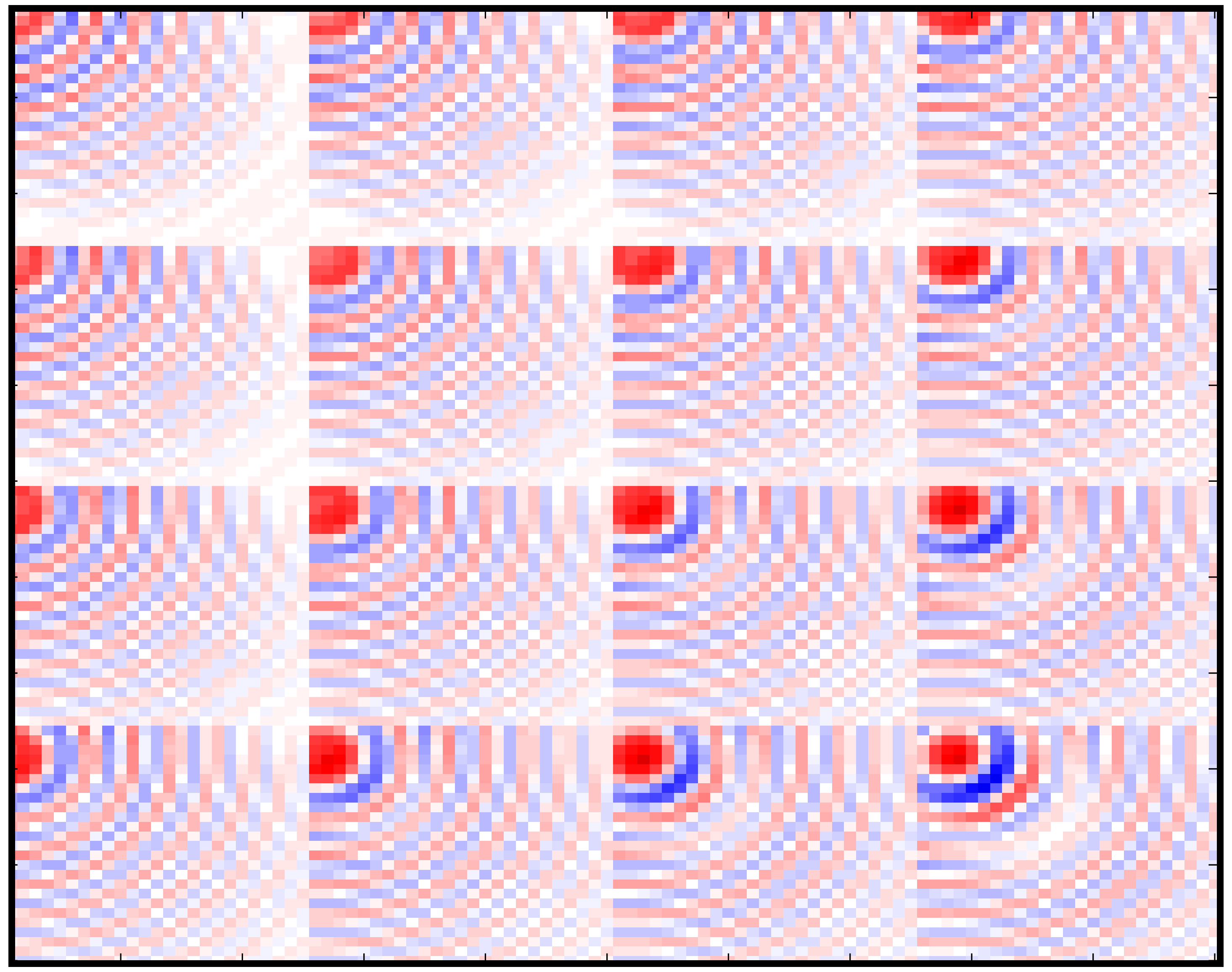


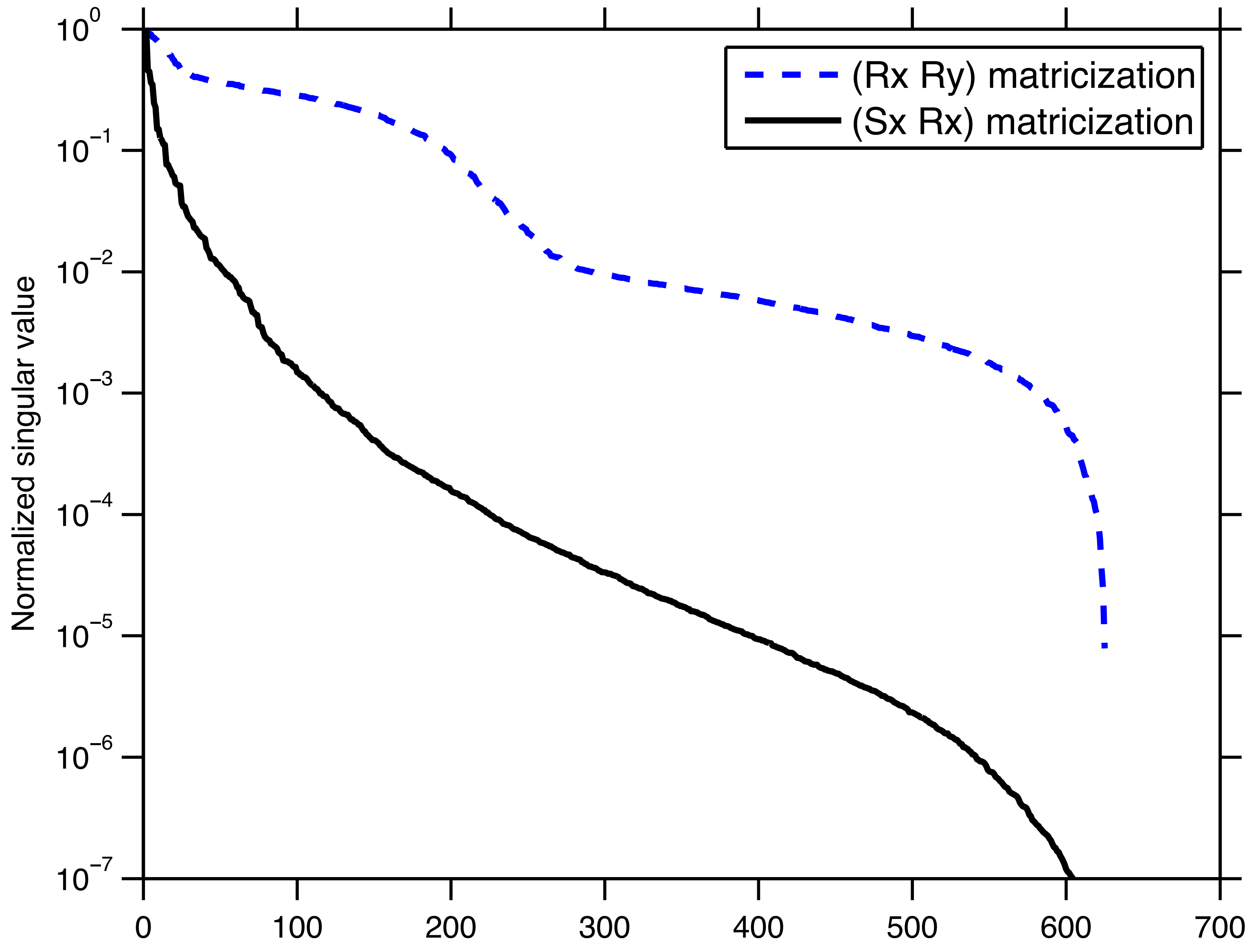


(Rec x, Rec y) matricization - Canonical ordering

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***

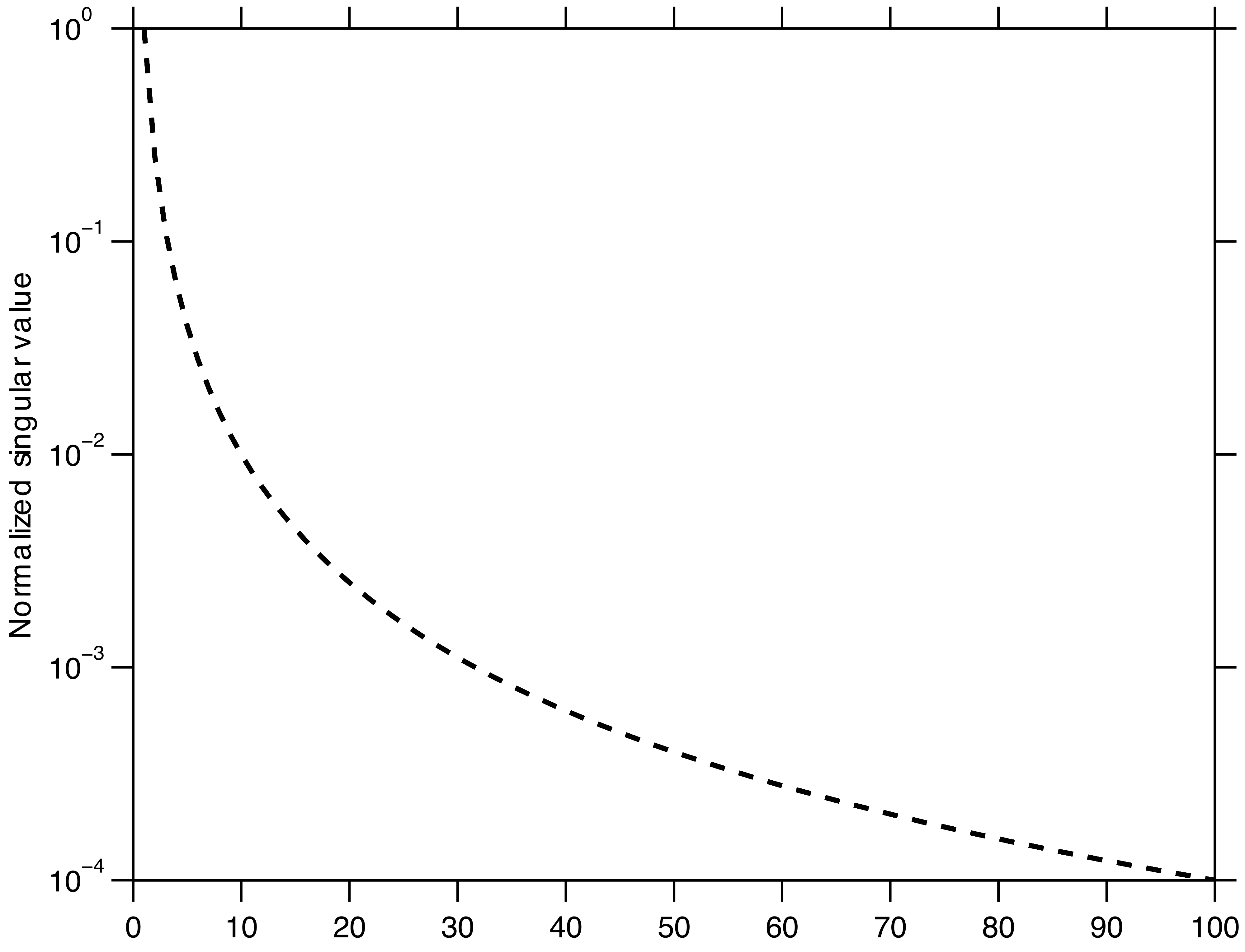
Optimization

- fit data in the Hierarchical Tucker format

Matrix Completion

$$\mathbf{X}$$

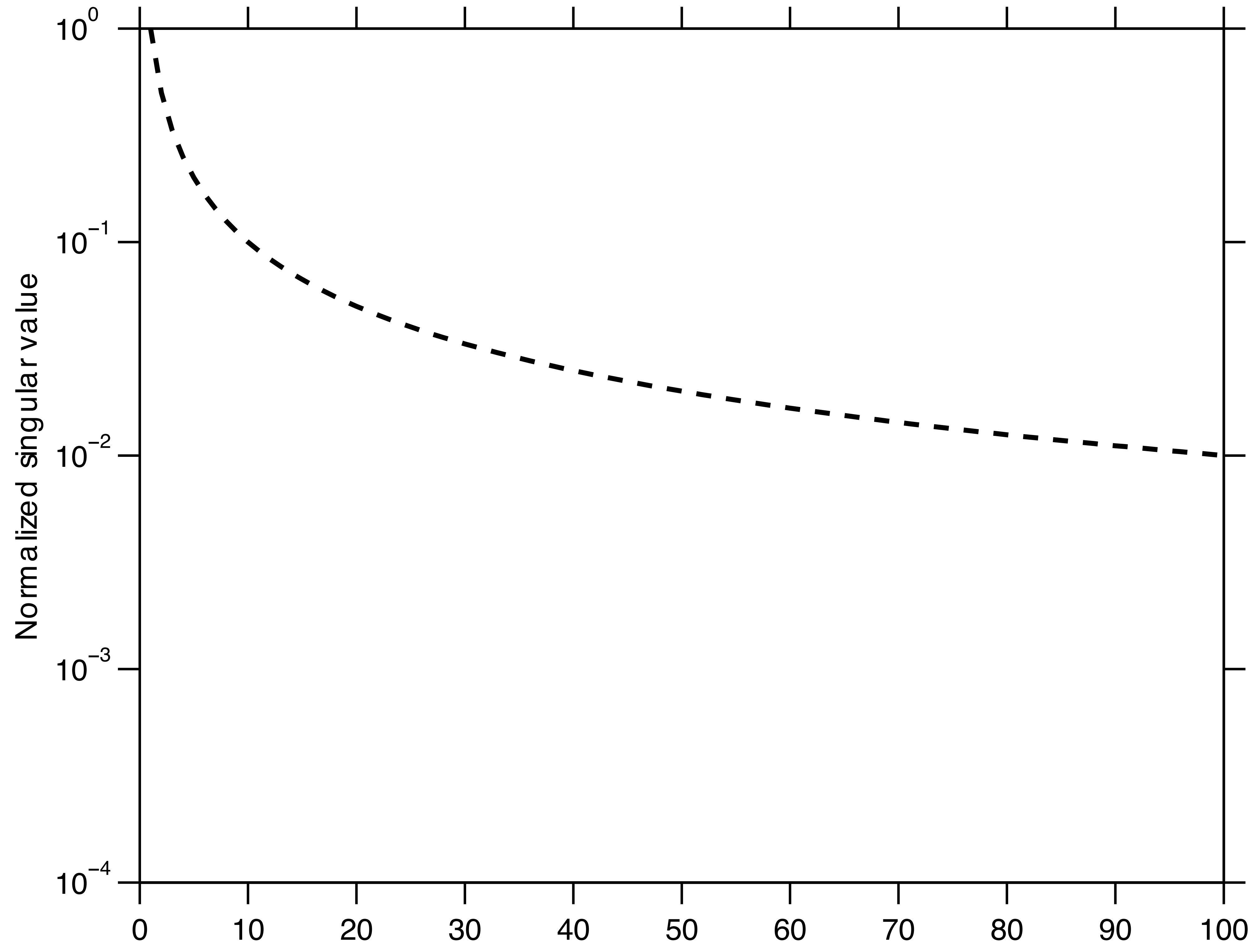
$$\begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix}$$



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

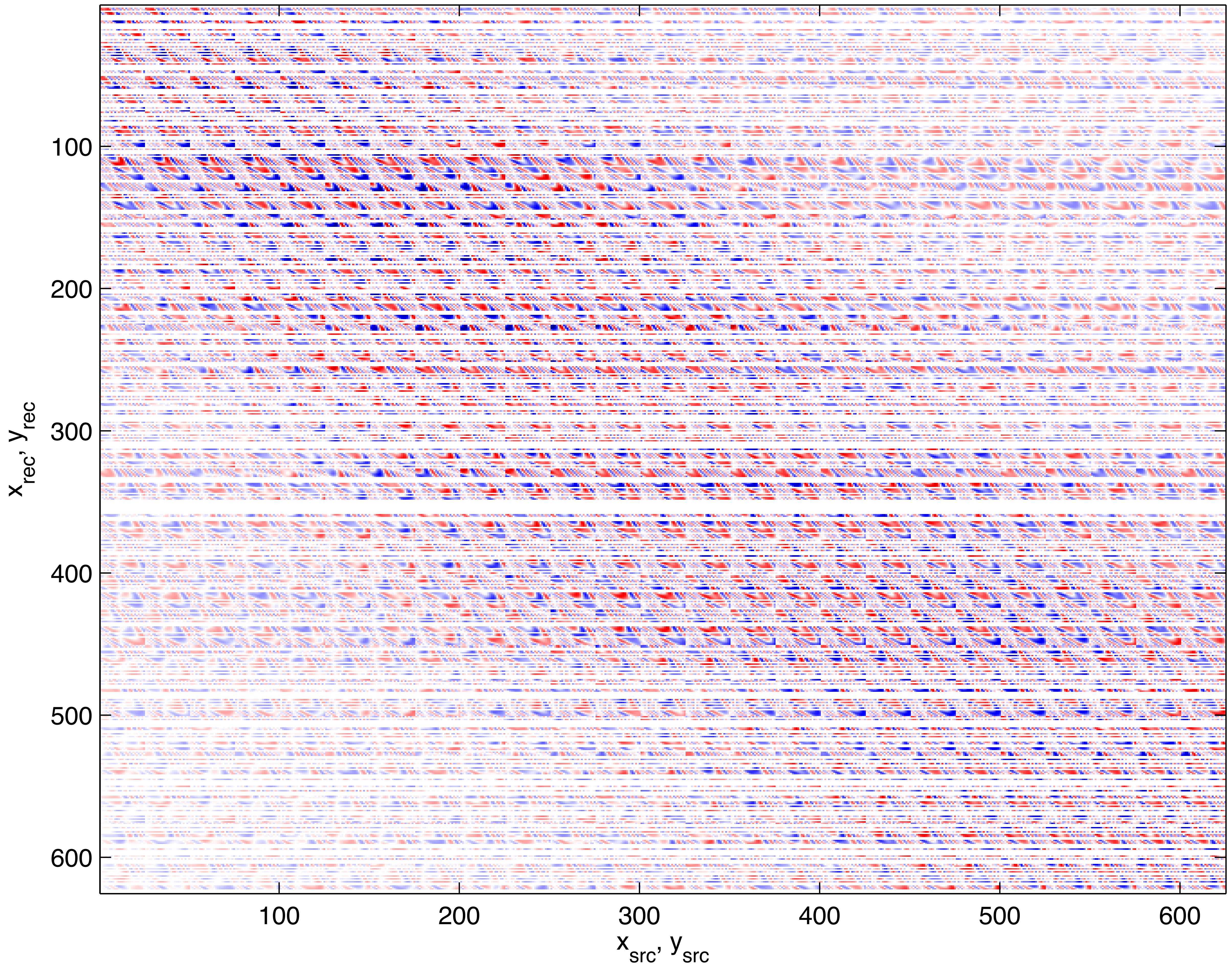
- idealized recovery
- impractical to physically implement

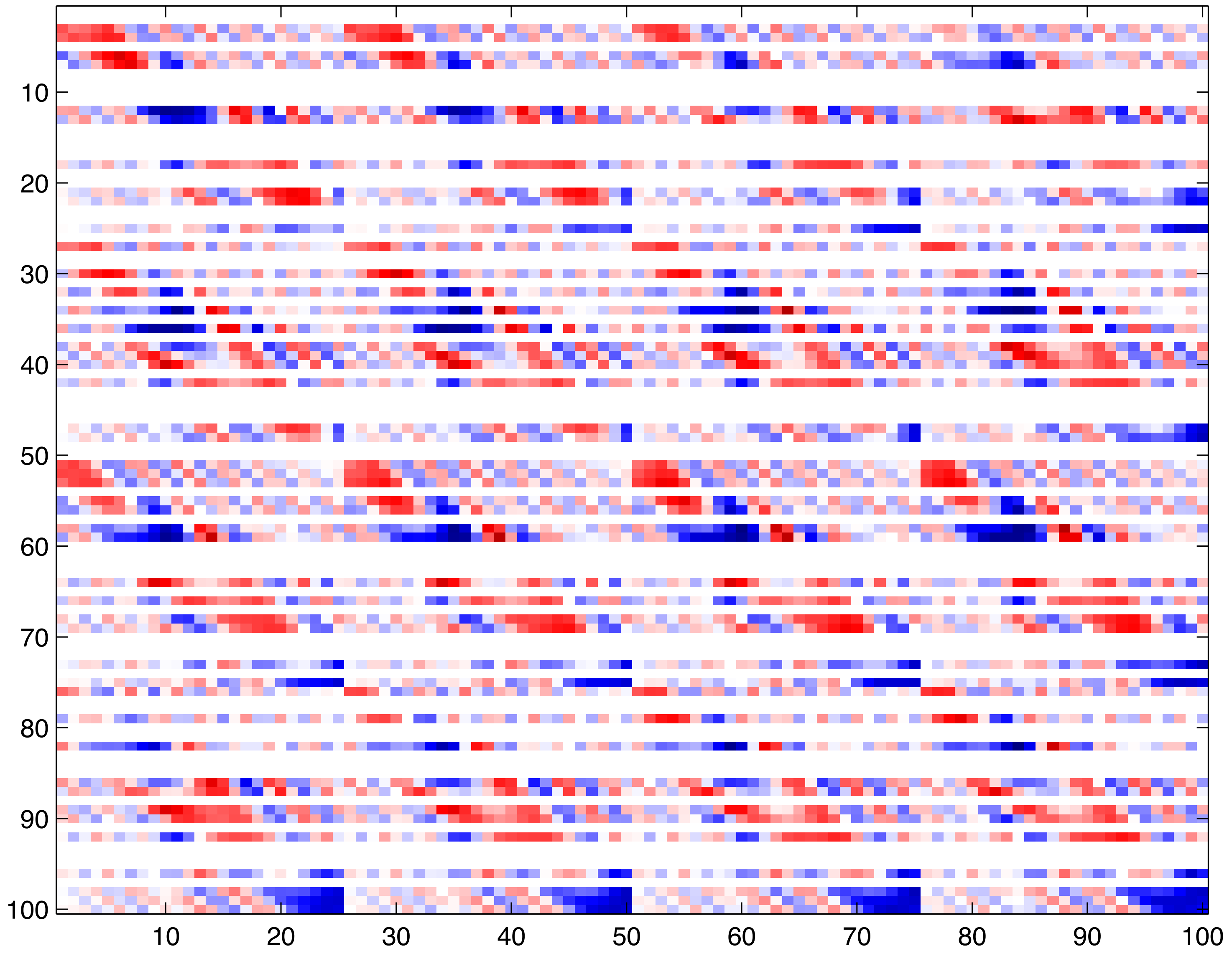
Sampling $(x_{\text{rec}}, y_{\text{rec}})$ points from the data

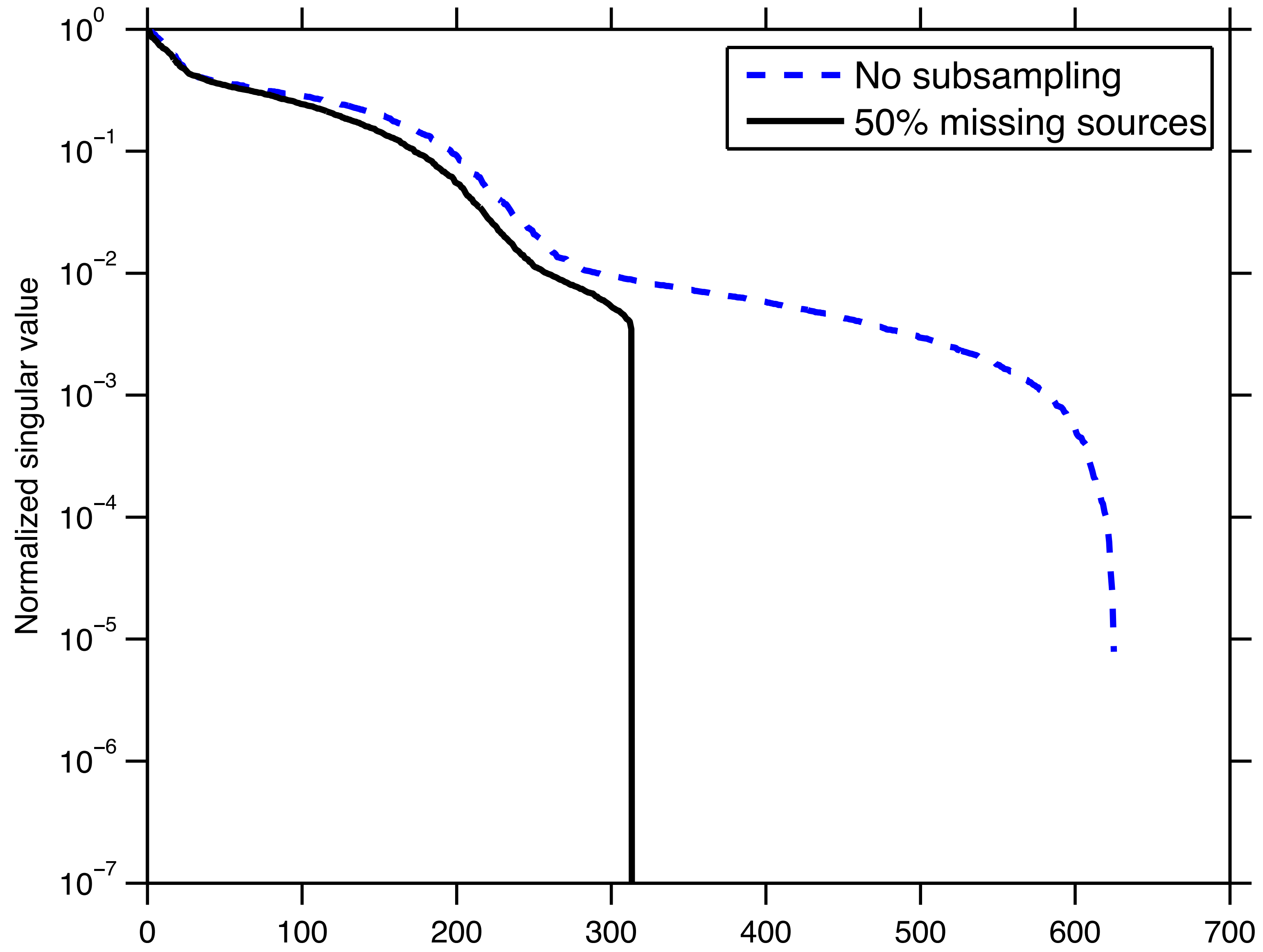
- less idealized
- possible to acquire data - e.g. ocean bottom nodes

Realistic recovery

50% random receivers removed



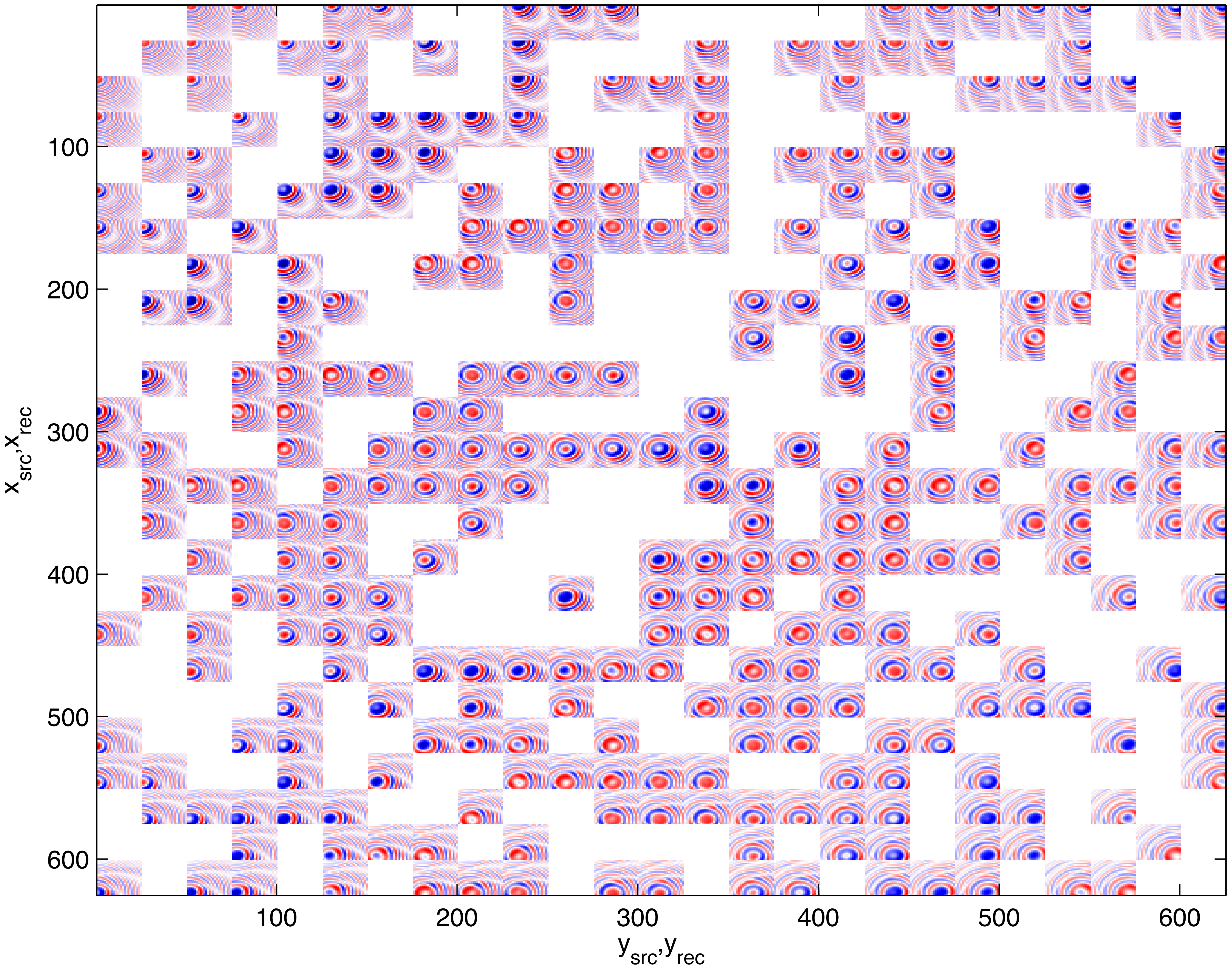


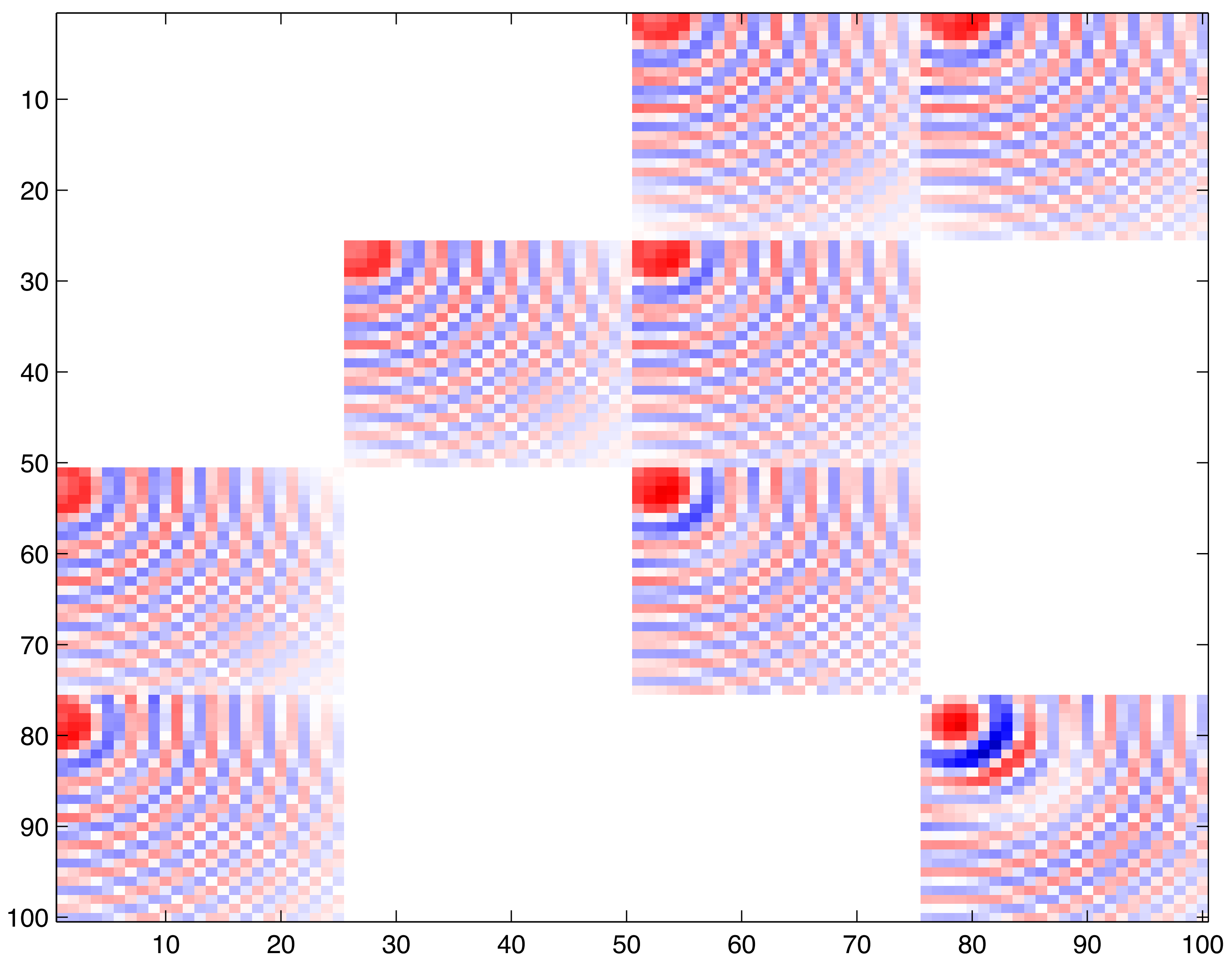


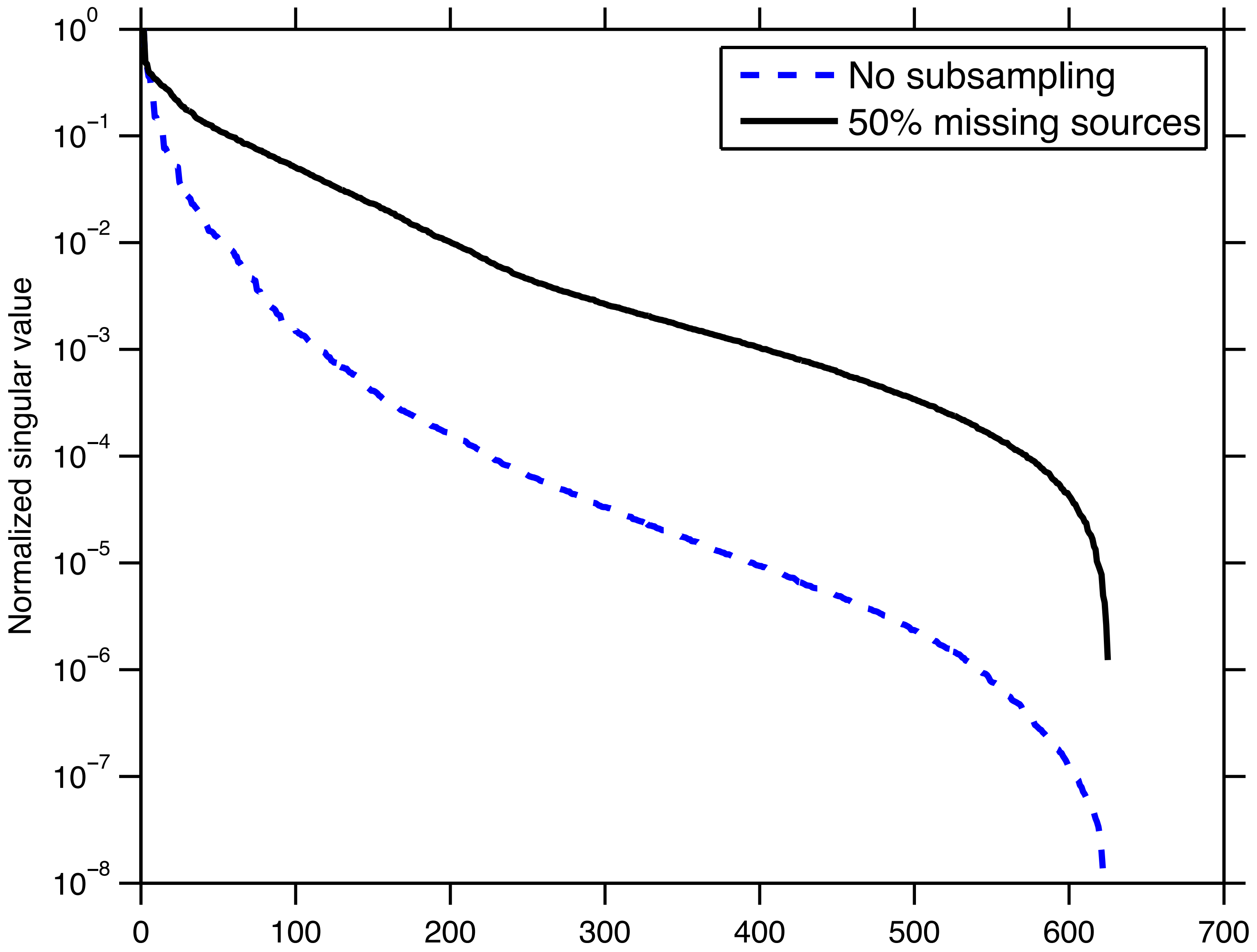
(Rec x, Rec y) matricization - Canonical ordering

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

Interpolation with noise

Malfunctioning receivers

- unknown, malfunctioning receivers generating Gaussian noise

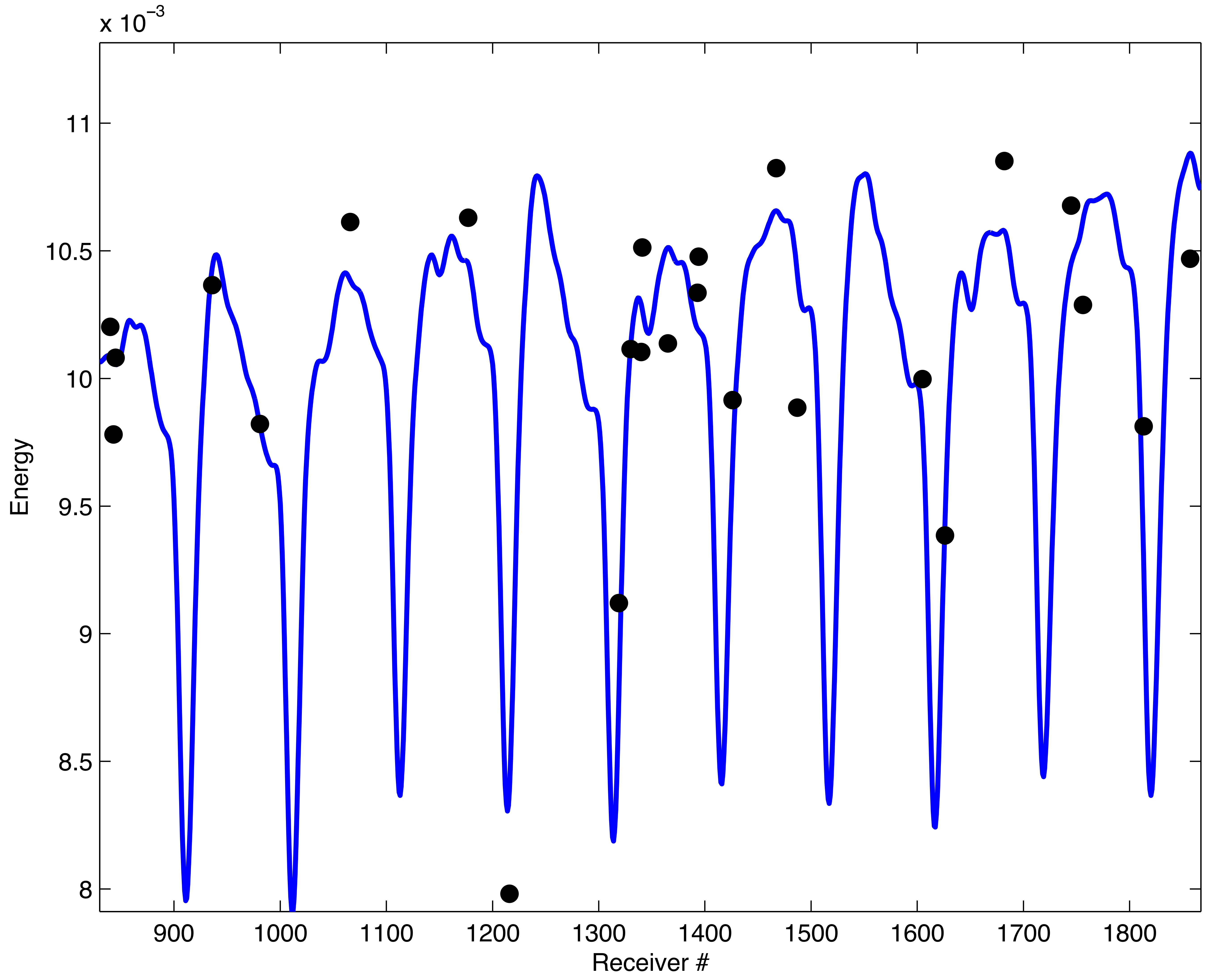
Low noise

- energy scaled to energy of removed receivers

High noise

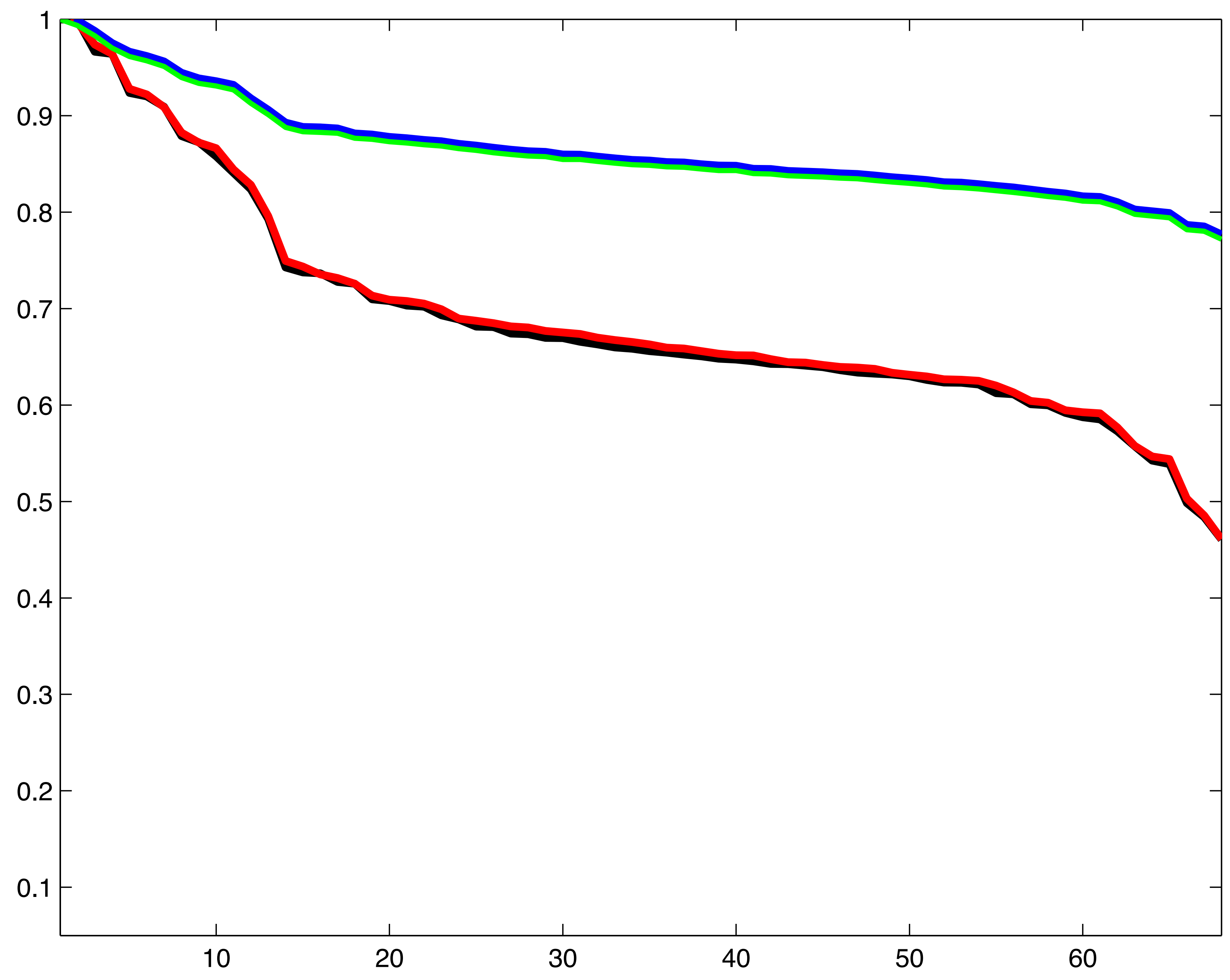
- total noise energy scaled to entire data energy

Receiver energy - low noise

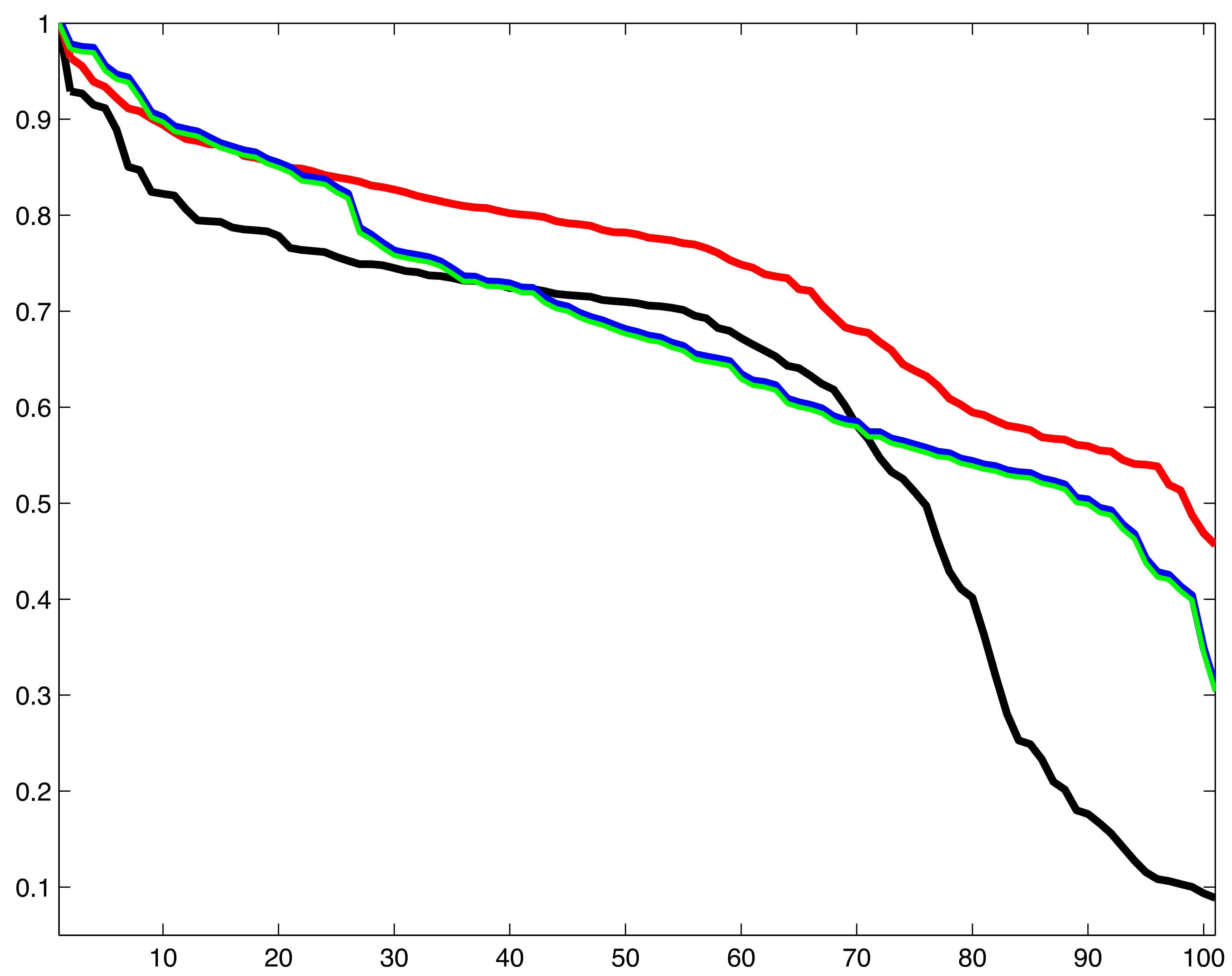


Interpolation with noise

Black - original
Red - subsampled
Blue - low noise
Green - high noise



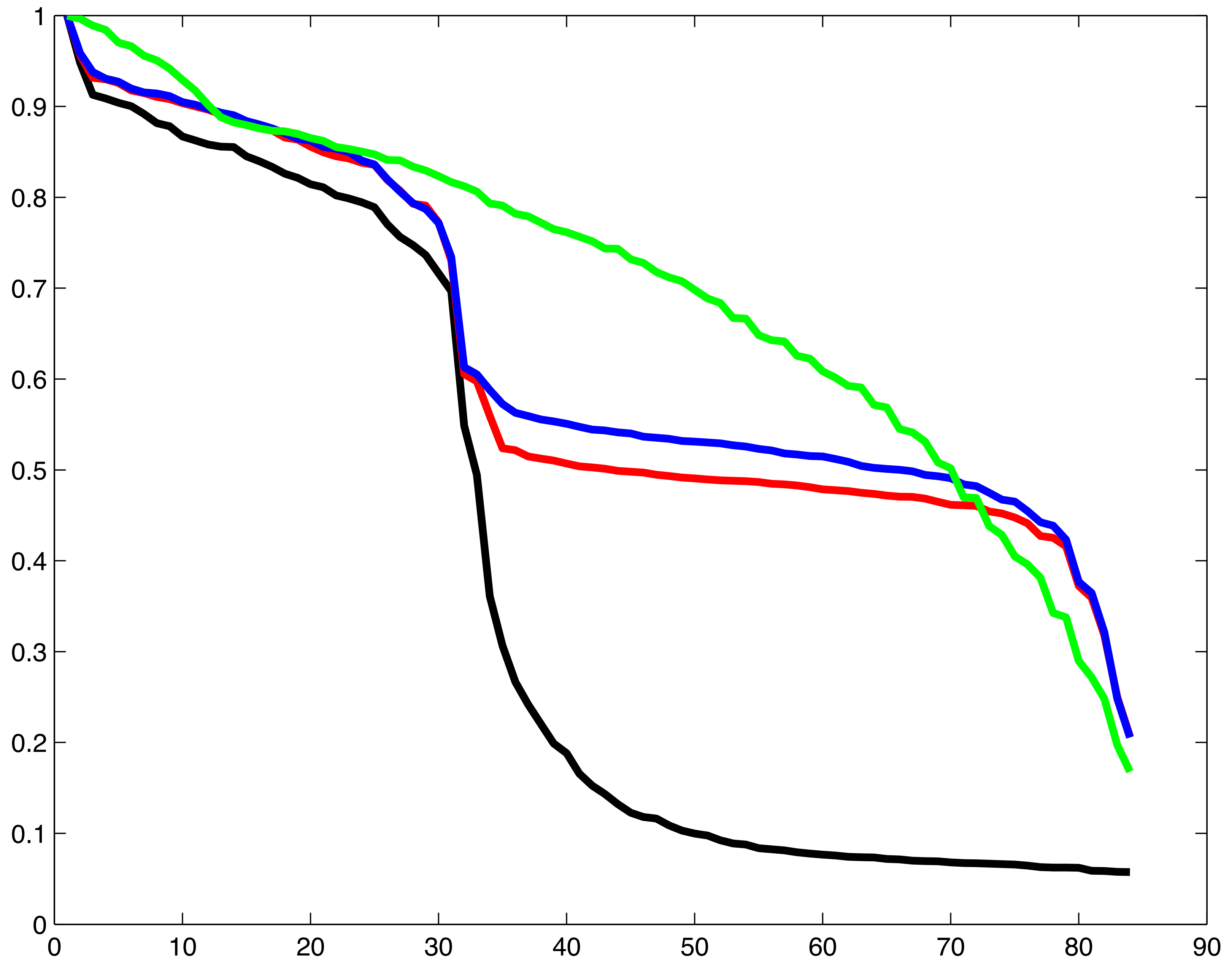
x_{src} singular values



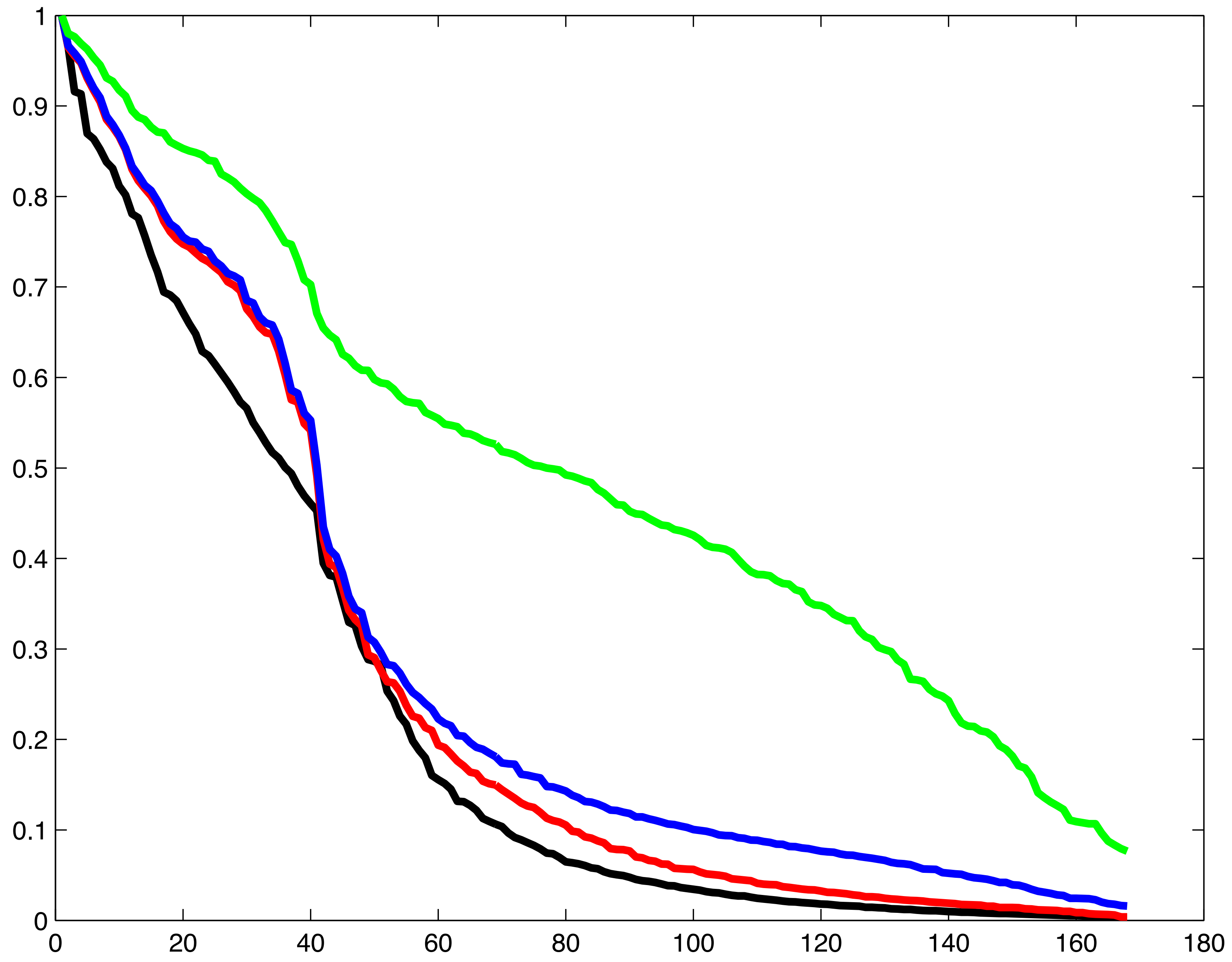
x_{rec} singular values

Interpolation with noise

Black - original
Red - subsampled
Blue - low noise
Green - high noise



$x_{midpoint}$ singular values



x_{offset} singular values

Interpolation with noise

Source-receiver domain

- subsampling ***increases*** the singular values in all dimensions
- noise ***does not change*** source-side singular values, ***decreases*** receiver-side singular values

Conclusion, *in this domain*

- Low-rank HT optimization ***will*** interpolate values in noiseless case
- Low-rank HT optimization ***cannot*** distinguish between noise & signal in noisy case

Interpolation with noise

Midpoint-offset domain

- subsampling ***increases*** the singular values in all dimensions
- noise ***increases*** the singular values in both the midpoint and offset dimensions

Conclusion, in this domain

- Low-rank HT optimization ***will*** interpolate values in noiseless case
- Low-rank HT optimization ***will*** interpolate and distinguish between noise and signal

Multidimensional interpolation

with Hierarchical Tucker

Successful reconstruction scheme

Signal structure

- Hierarchical Tucker

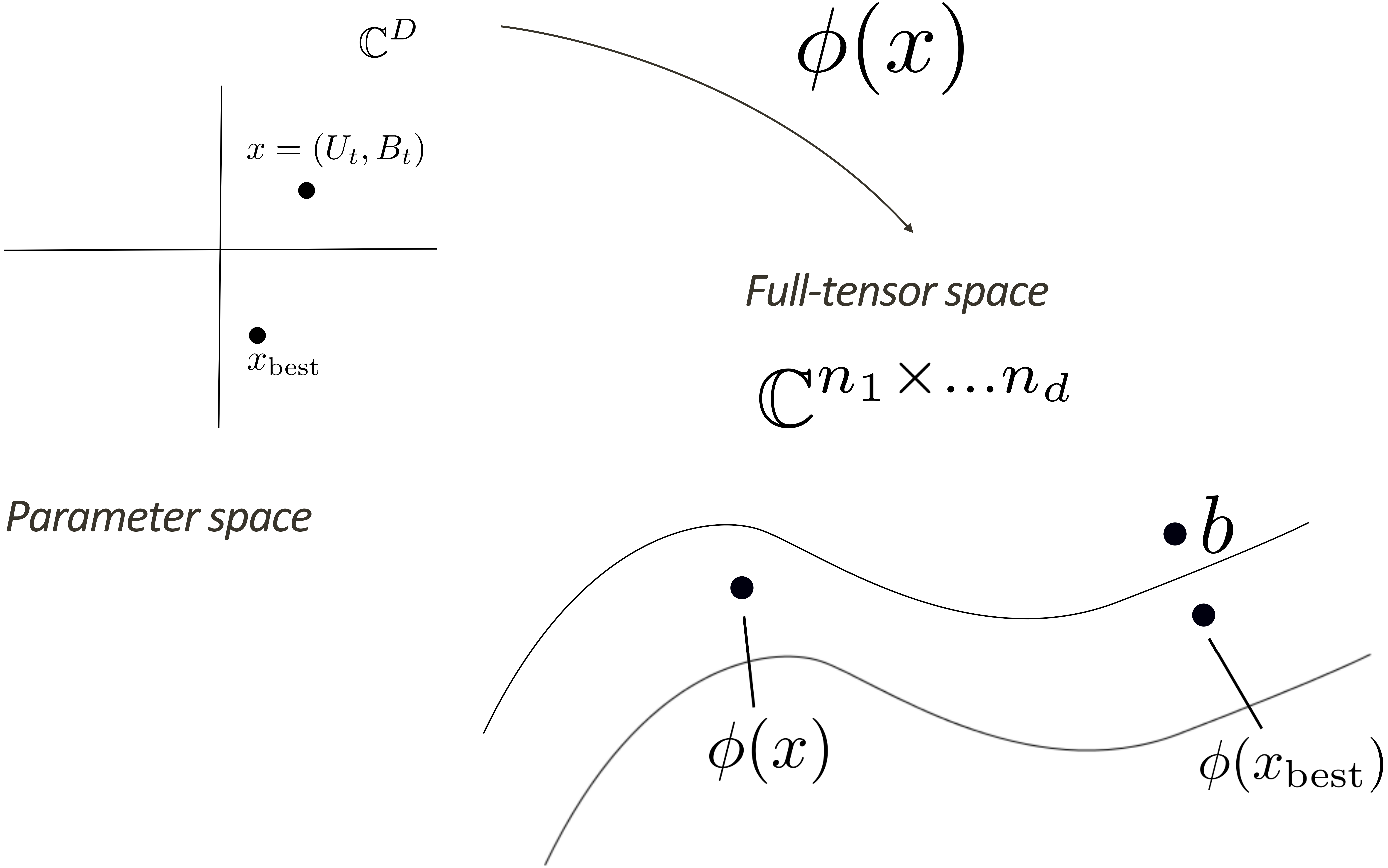
Sampling

- subsampling, noise increases hierarchical rank

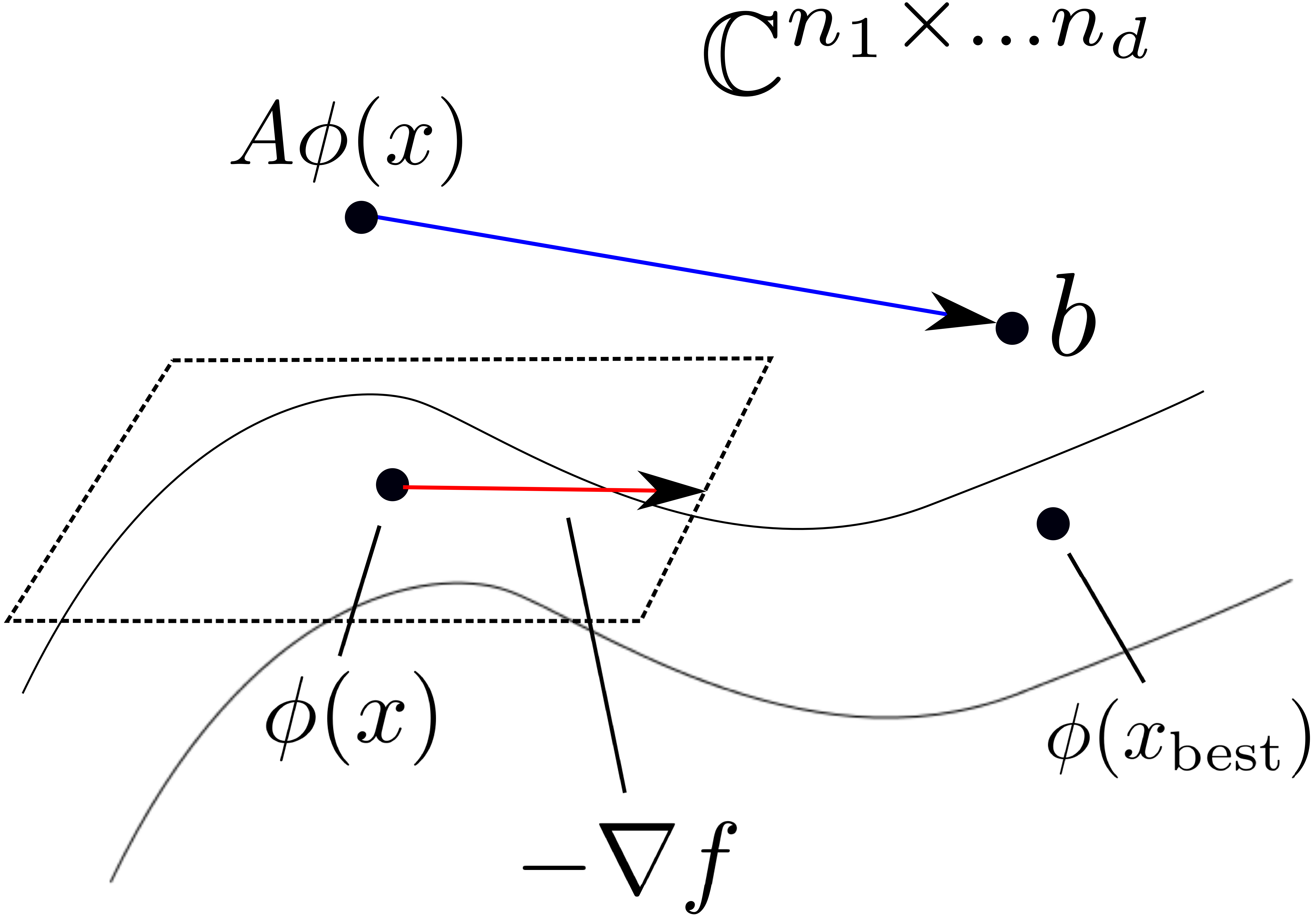
Optimization

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

Optimization program



Optimization program



C. Da Silva and F. J. Herrmann, *Optimization on the Hierarchical Tucker manifold - applications to tensor completion*, 2014

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

Results

Synthetic BG Group data

Unknown model

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

Receivers subsampled to 101 x 101

Recovered with Gauss-Newton

Noise

Removed 50% of receivers randomly

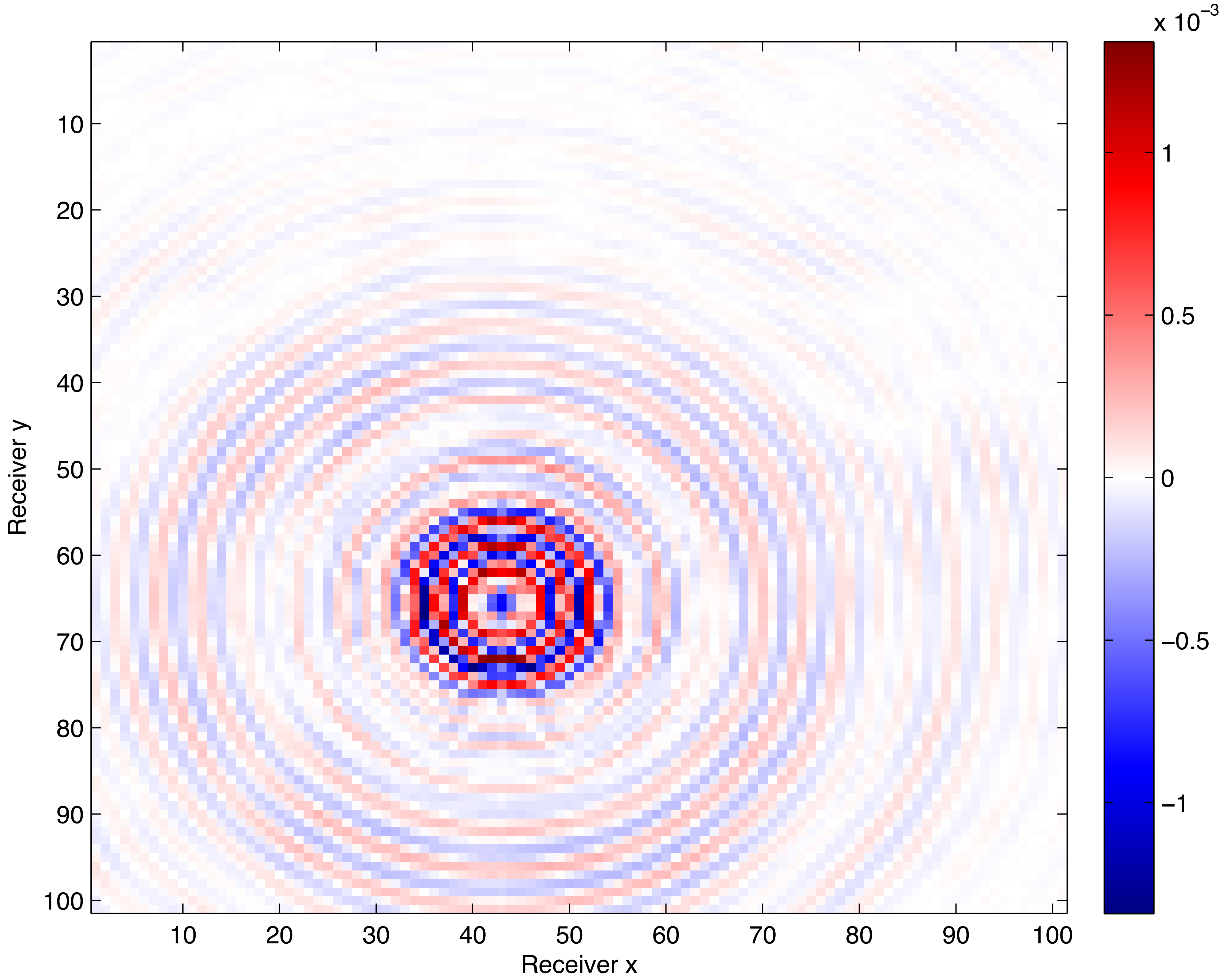
5% of remaining receivers replaced with random Gaussian noise

Low noise - energy scaled to energy of removed receivers

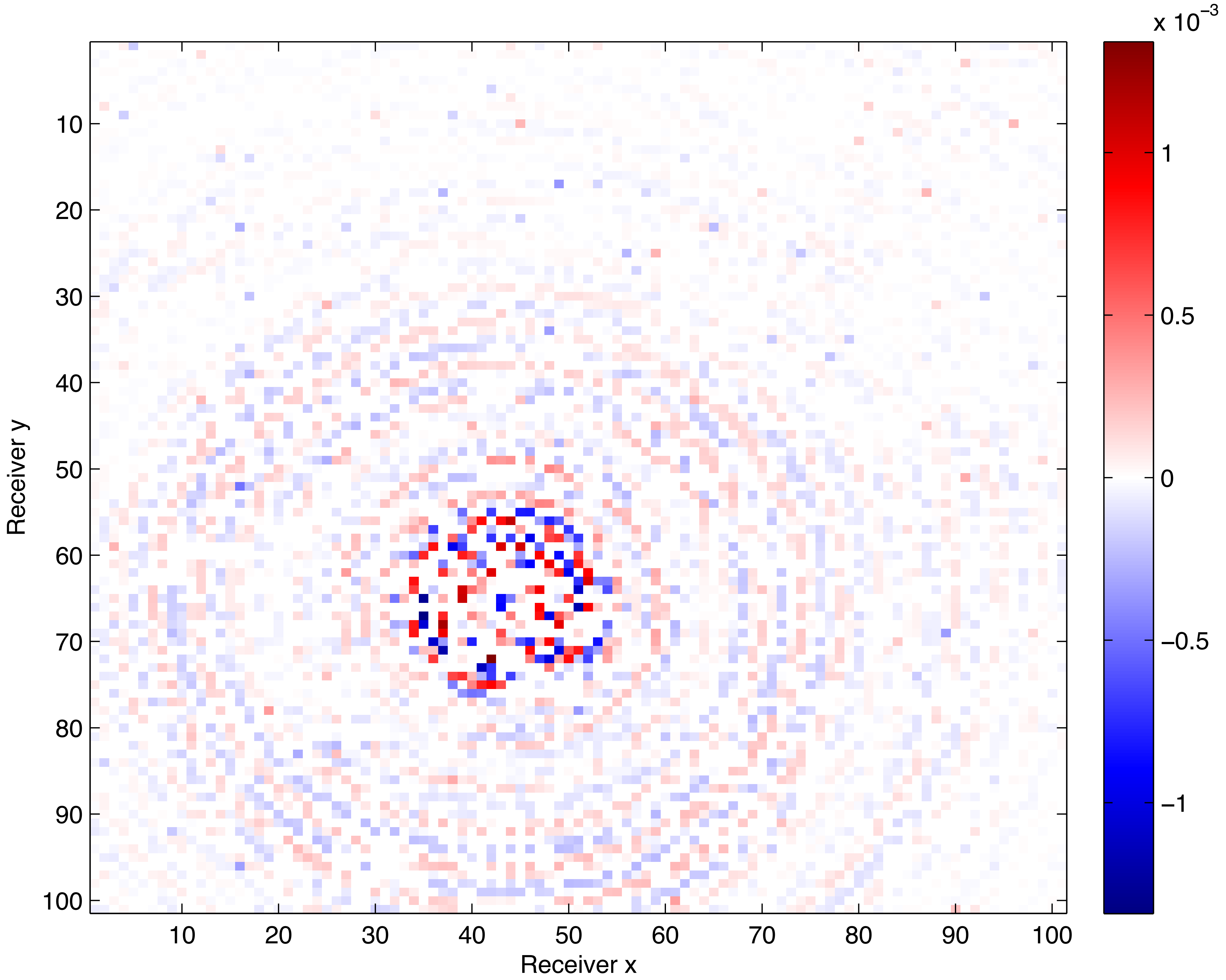
High noise - total noise energy scaled to entire data energy

7.34 Hz - 50% missing receivers - low noise

Common source gather



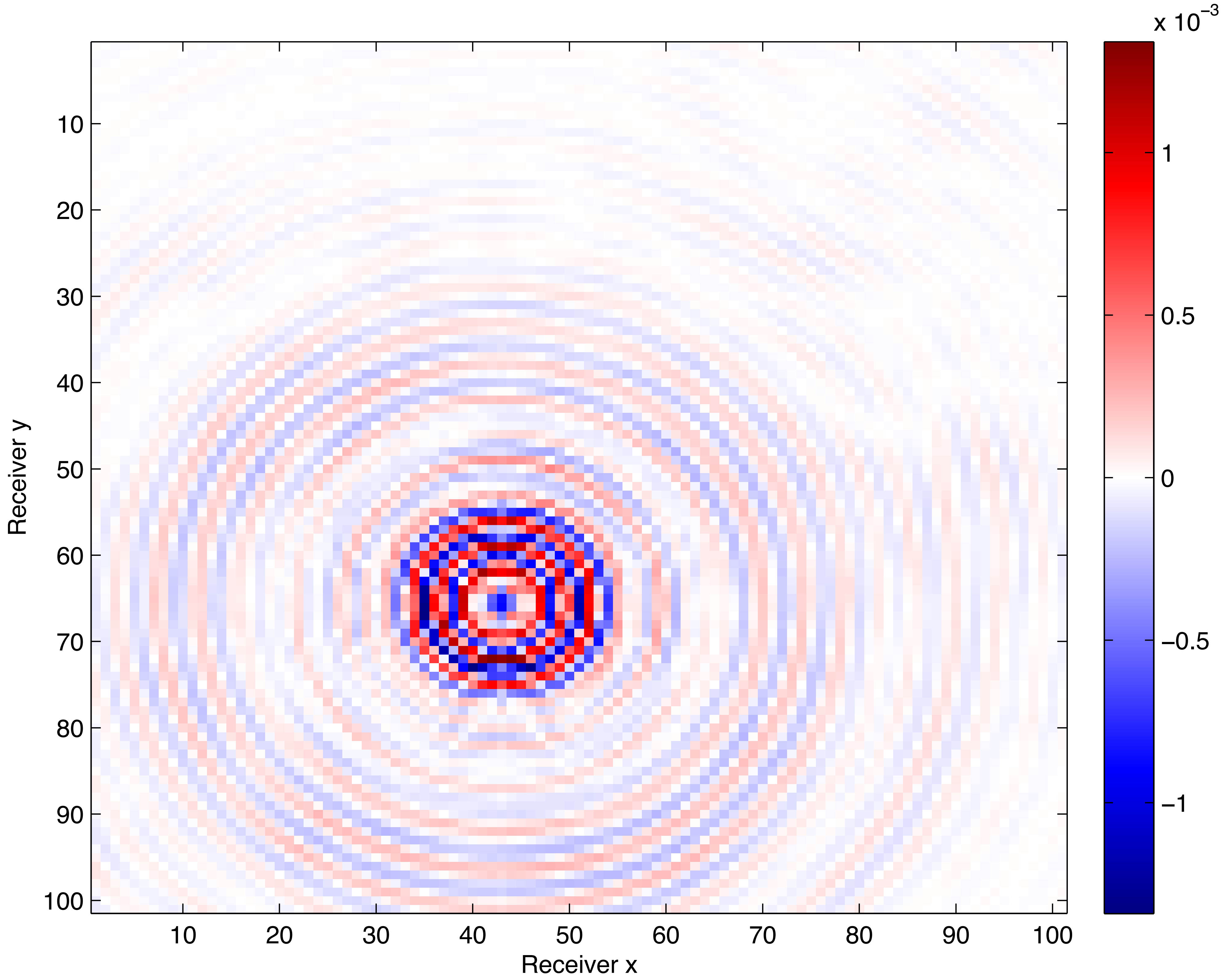
True data



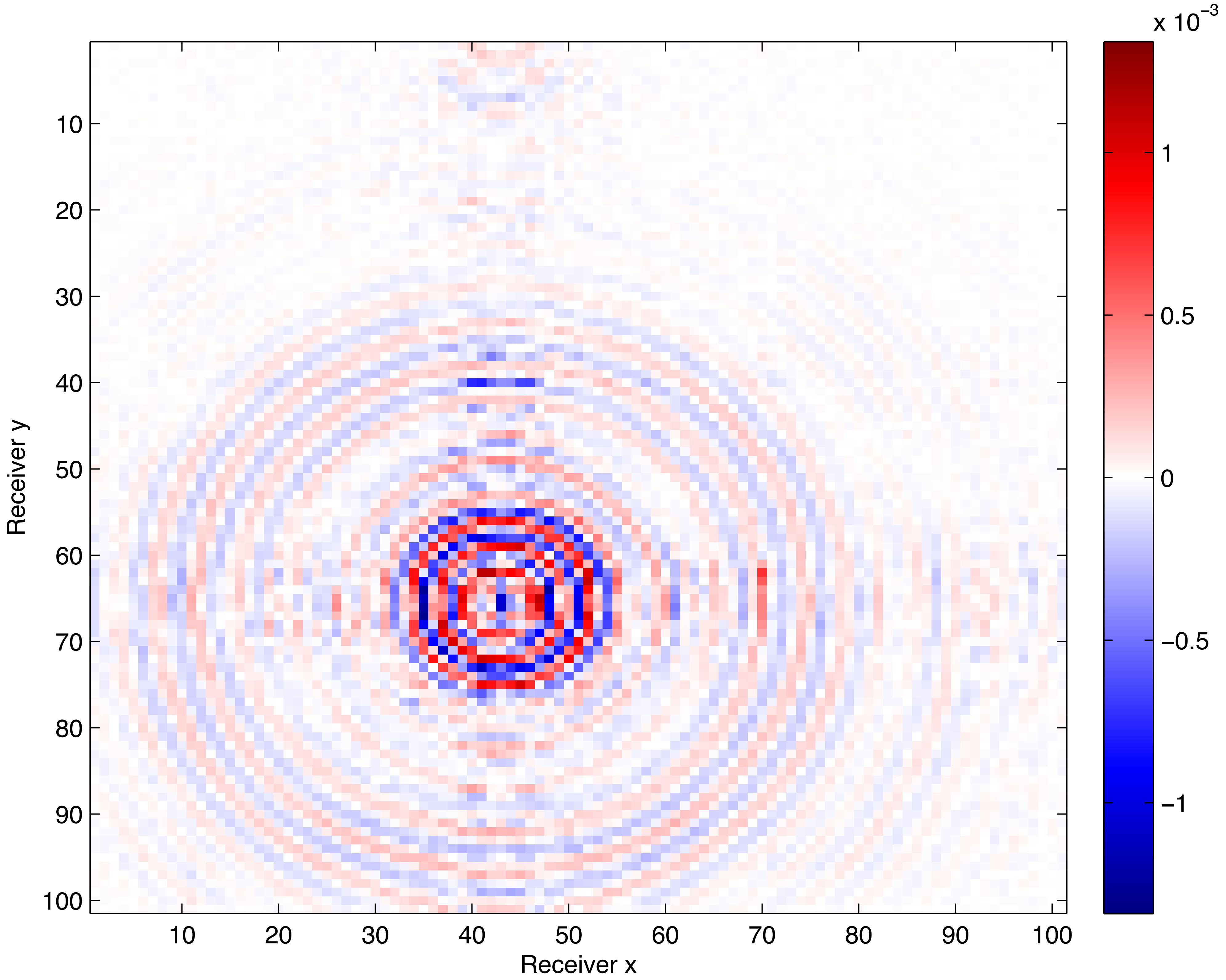
Subsampled data

7.34 Hz - 50% missing receivers - low noise

Common source gather



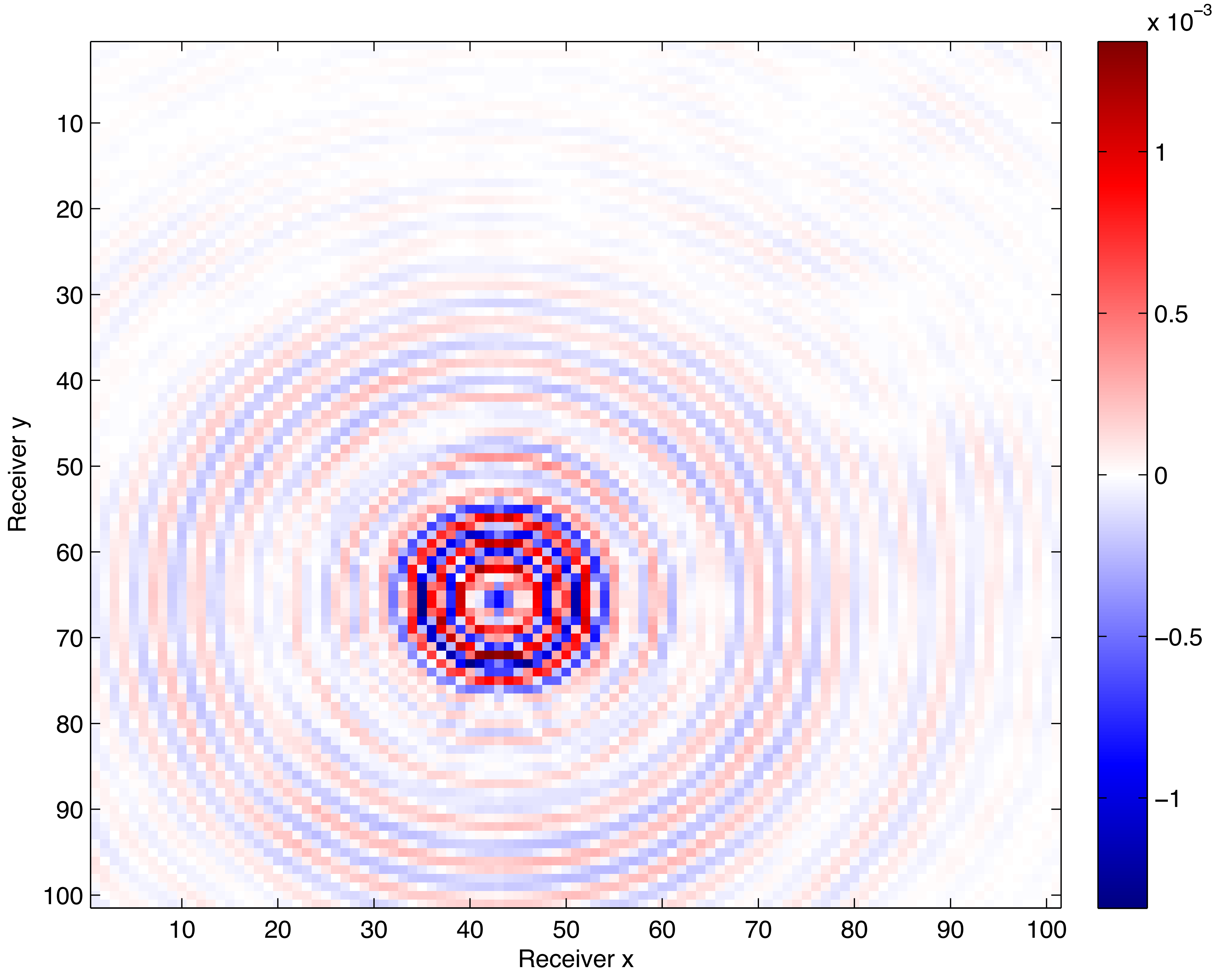
True data



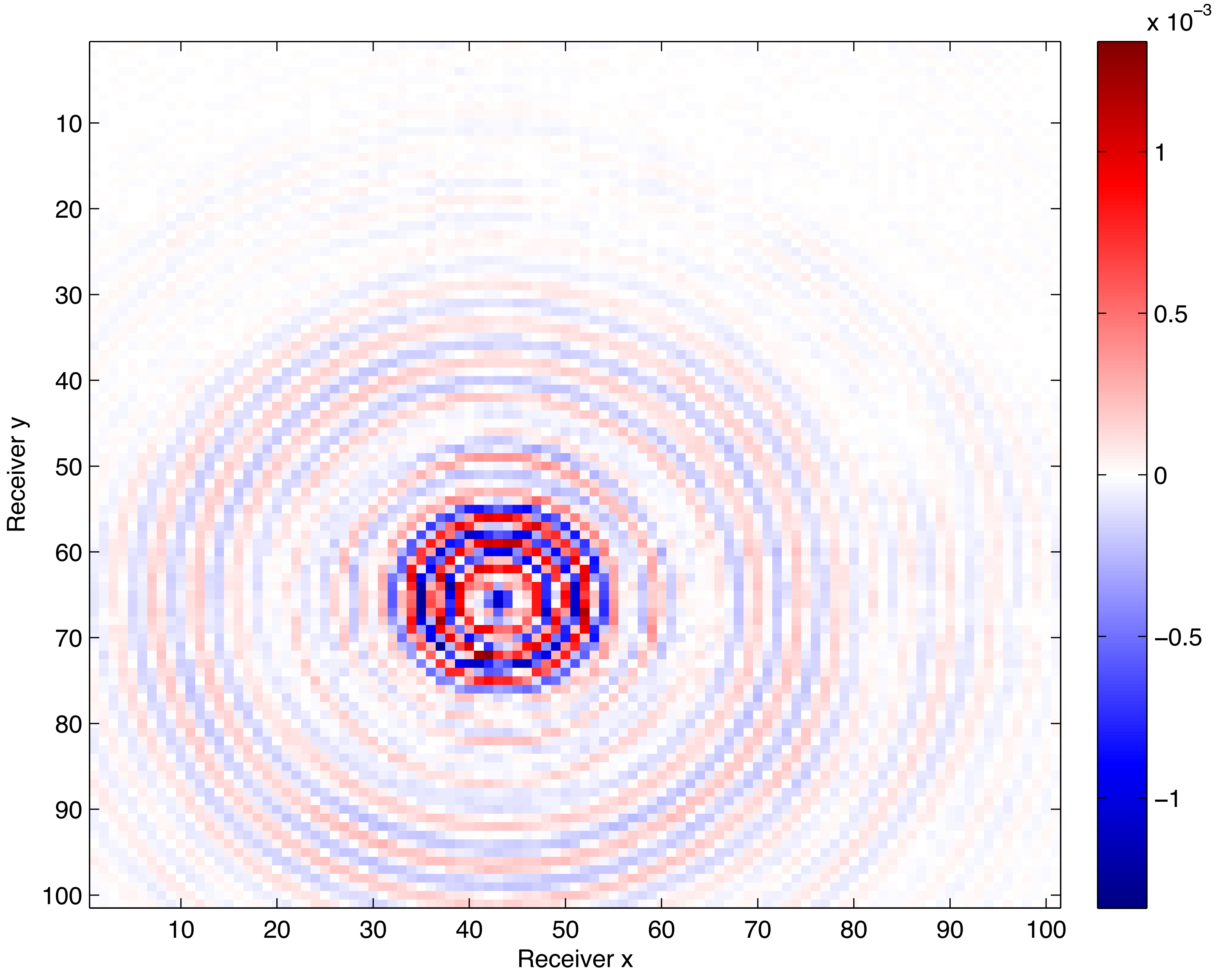
SR Recovery - SNR 7.8 dB

7.34 Hz - 50% missing receivers - low noise

Common source gather



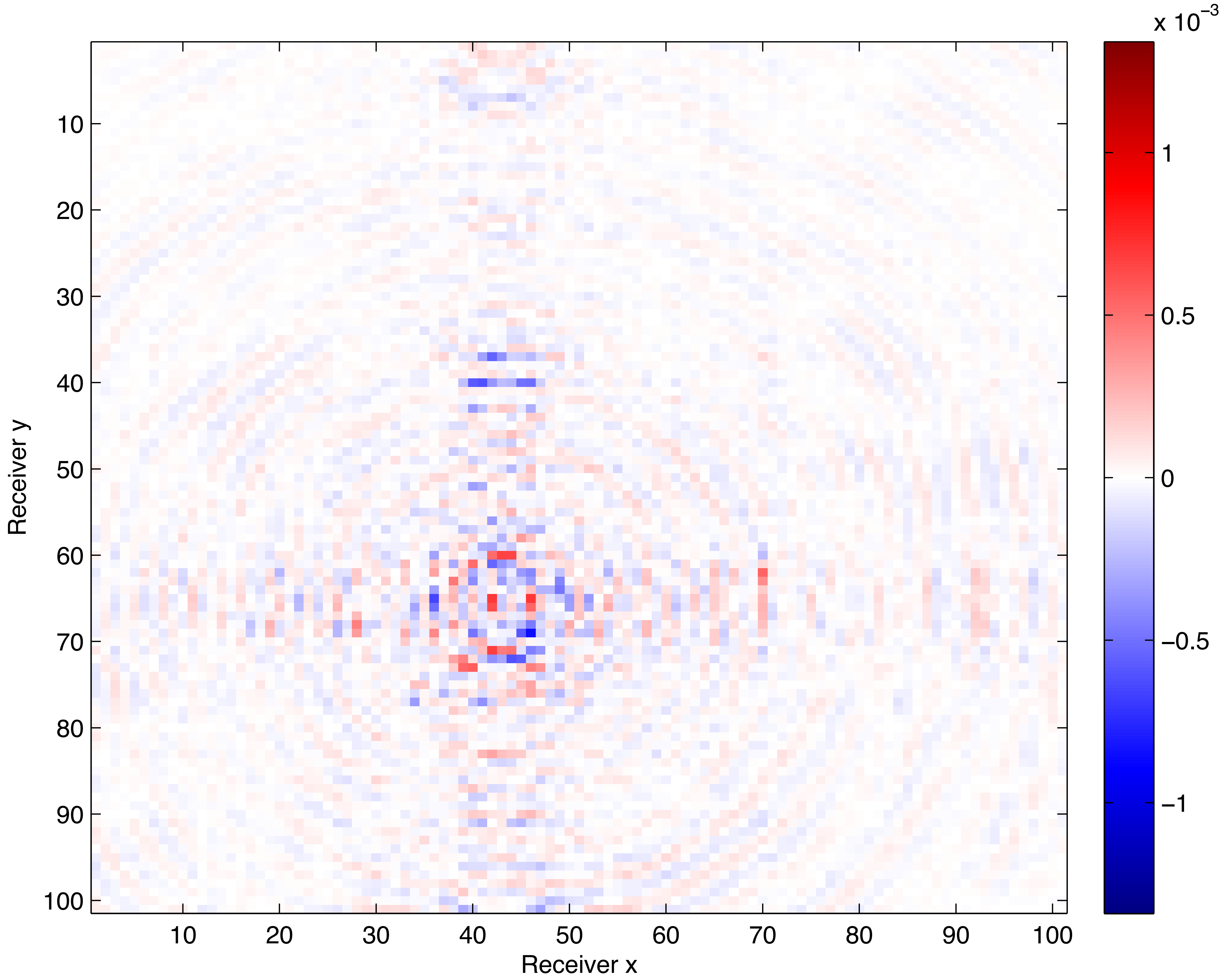
True data



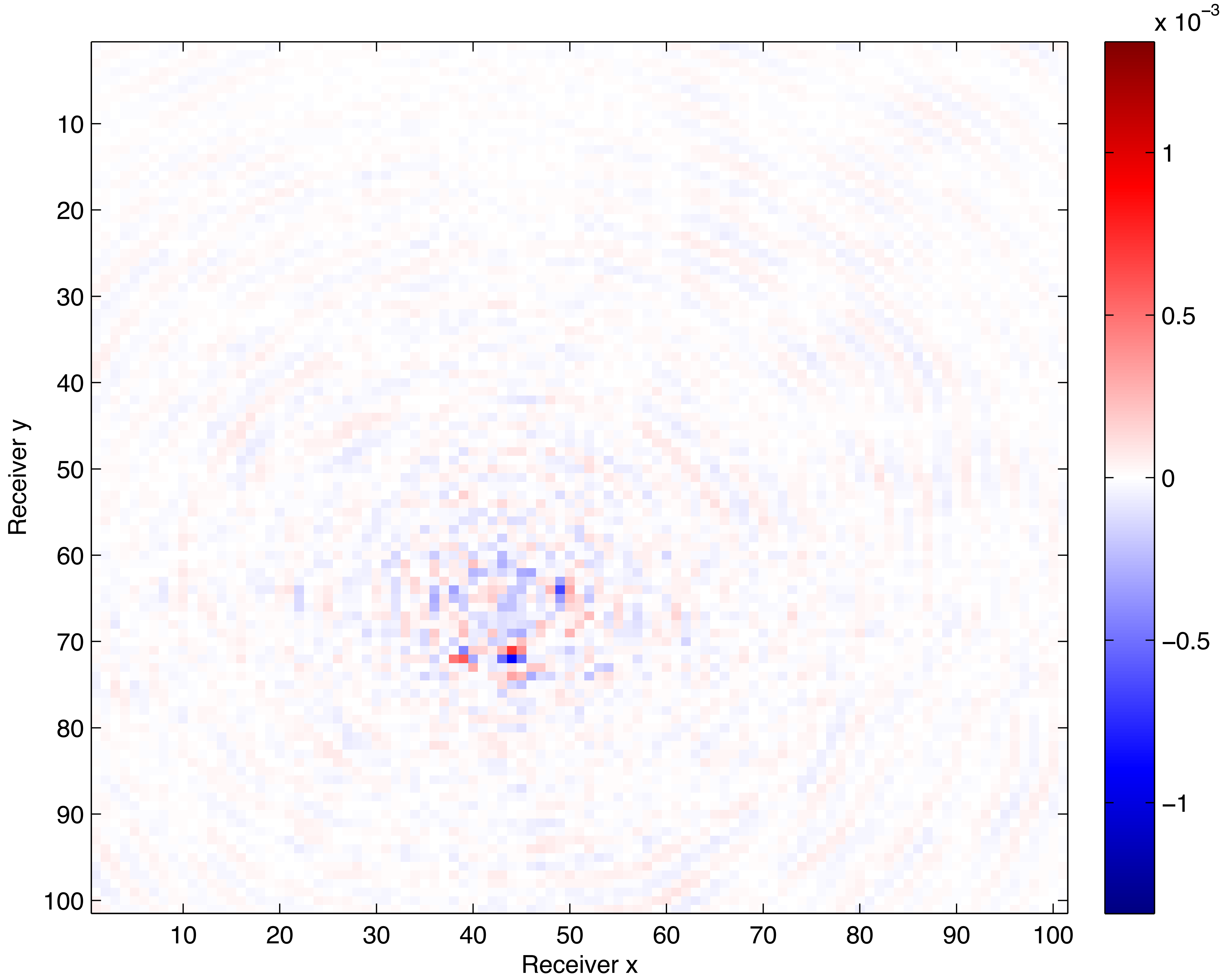
MH Recovery - SNR 12.6 dB

7.34 Hz - 50% missing receivers - low noise

Common source gather



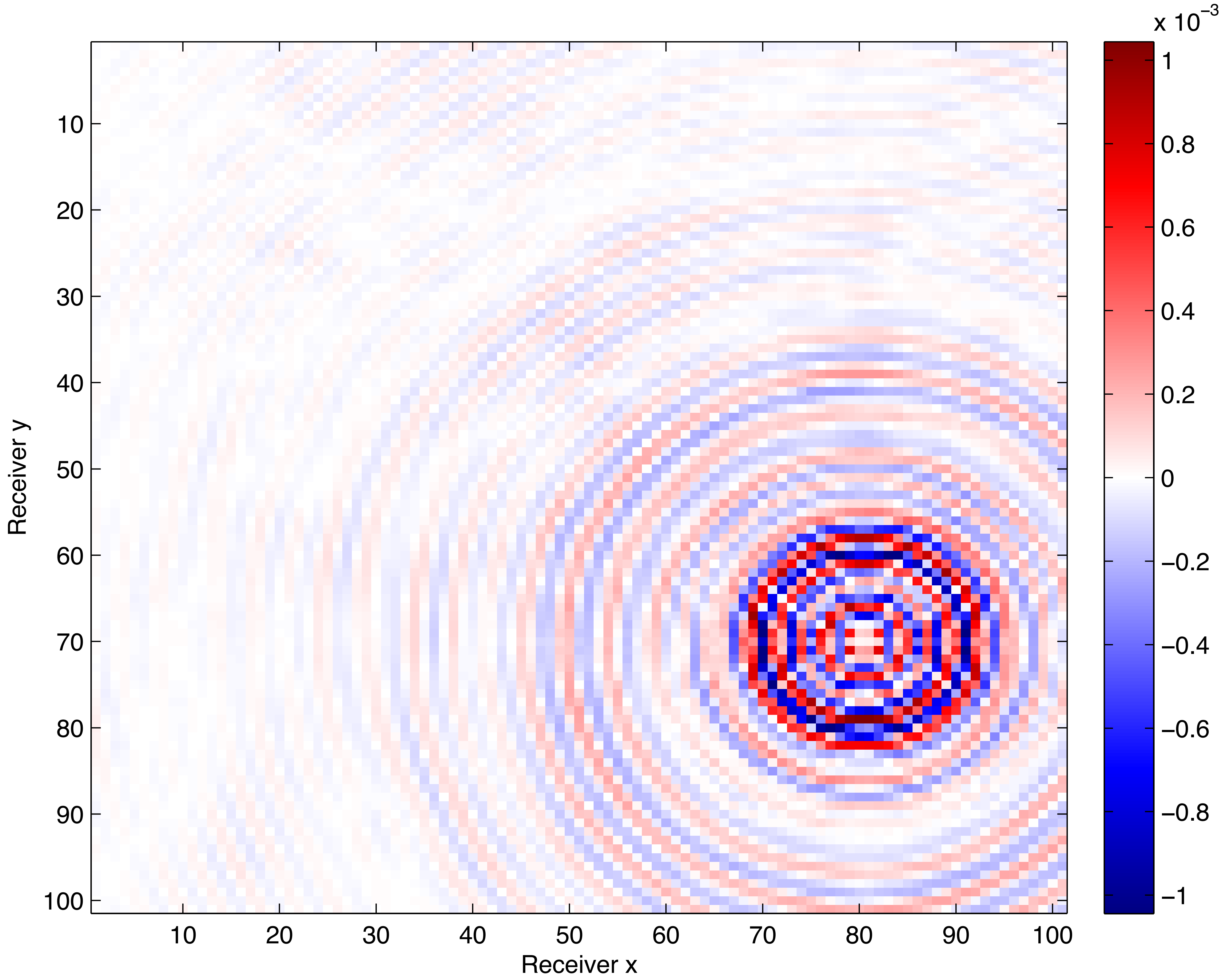
SR Difference



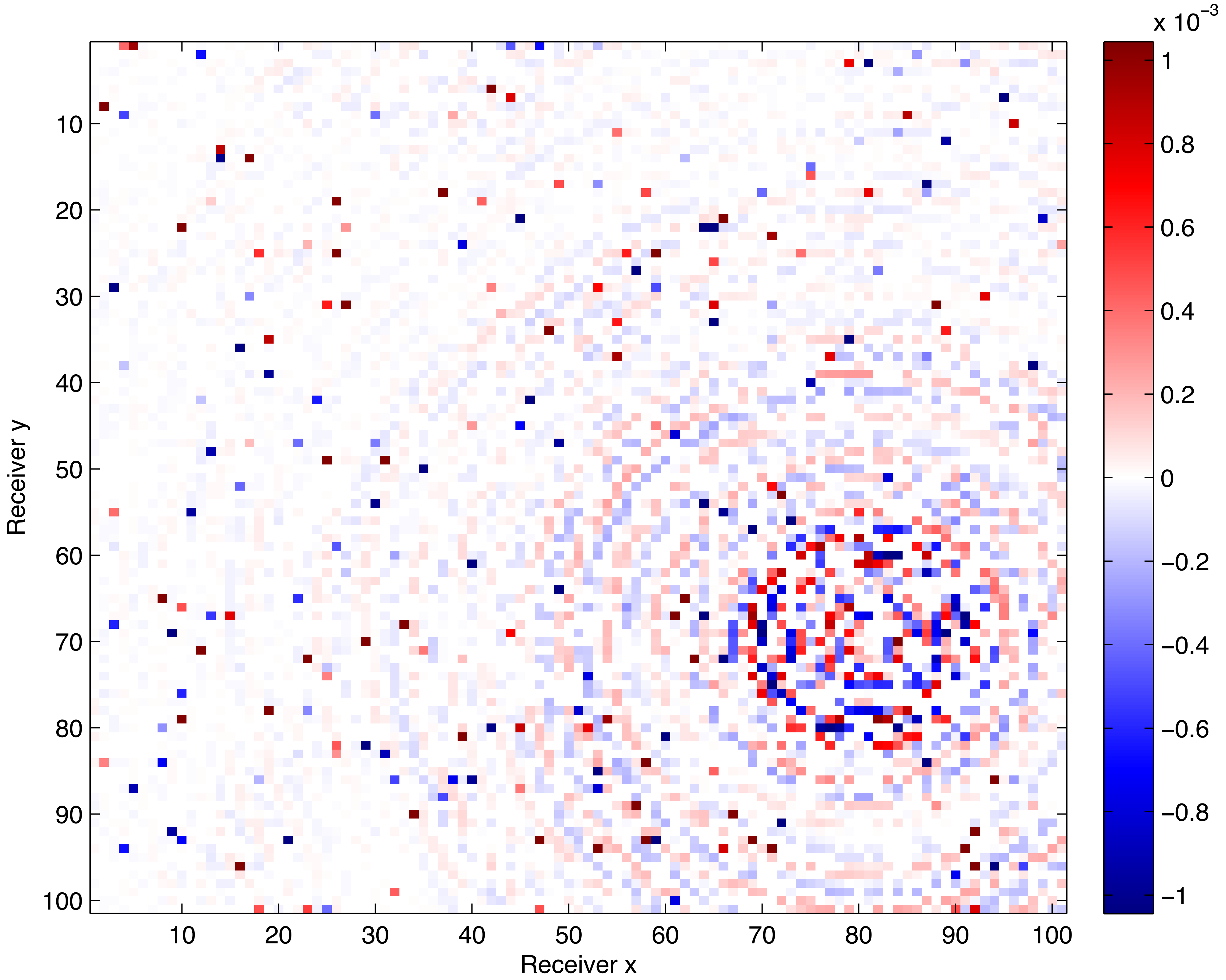
MH Difference

7.34 Hz - 50% missing receivers - high noise

Common source gather



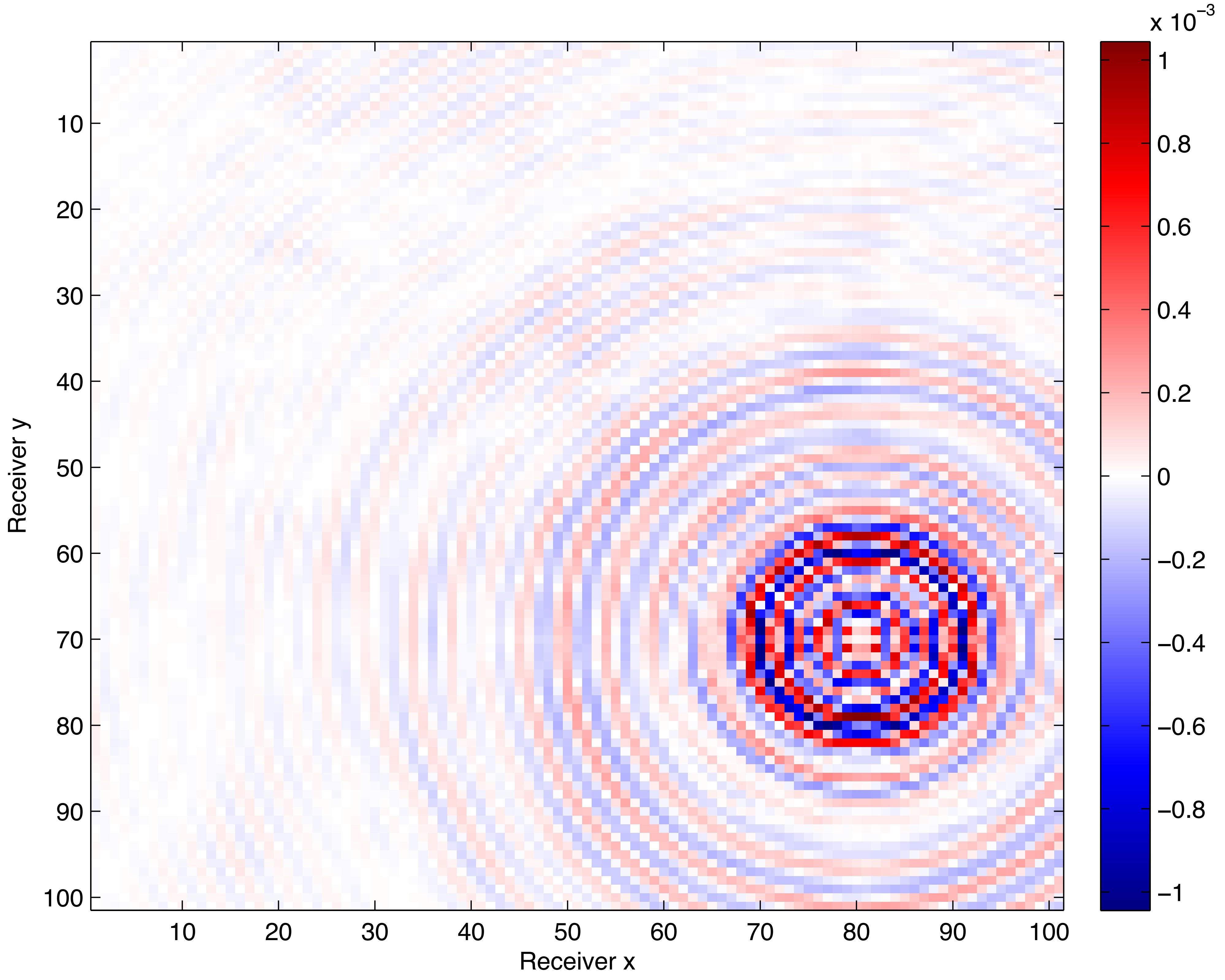
True data



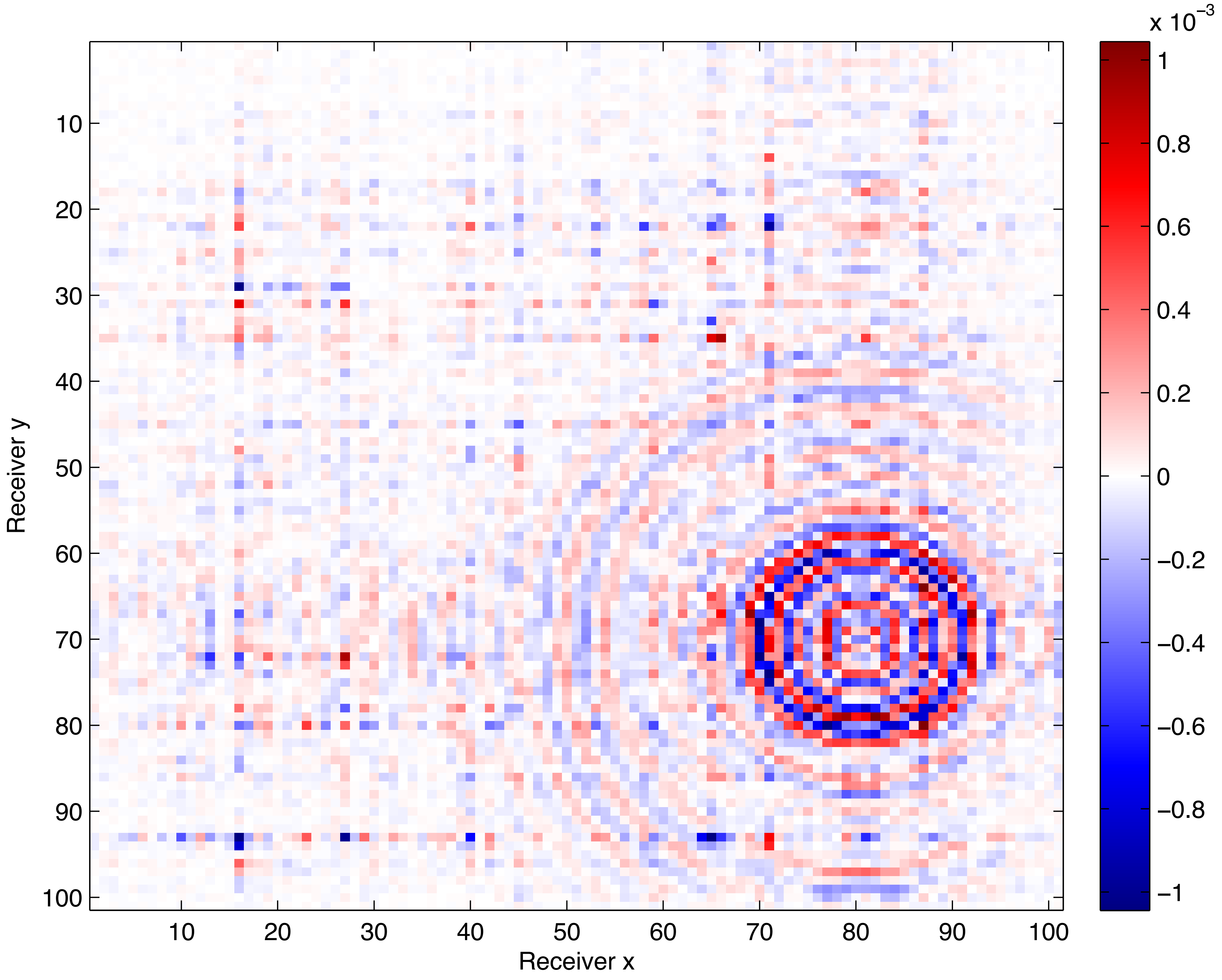
Subsampled data

7.34 Hz - 50% missing receivers - high noise

Common source gather



True data

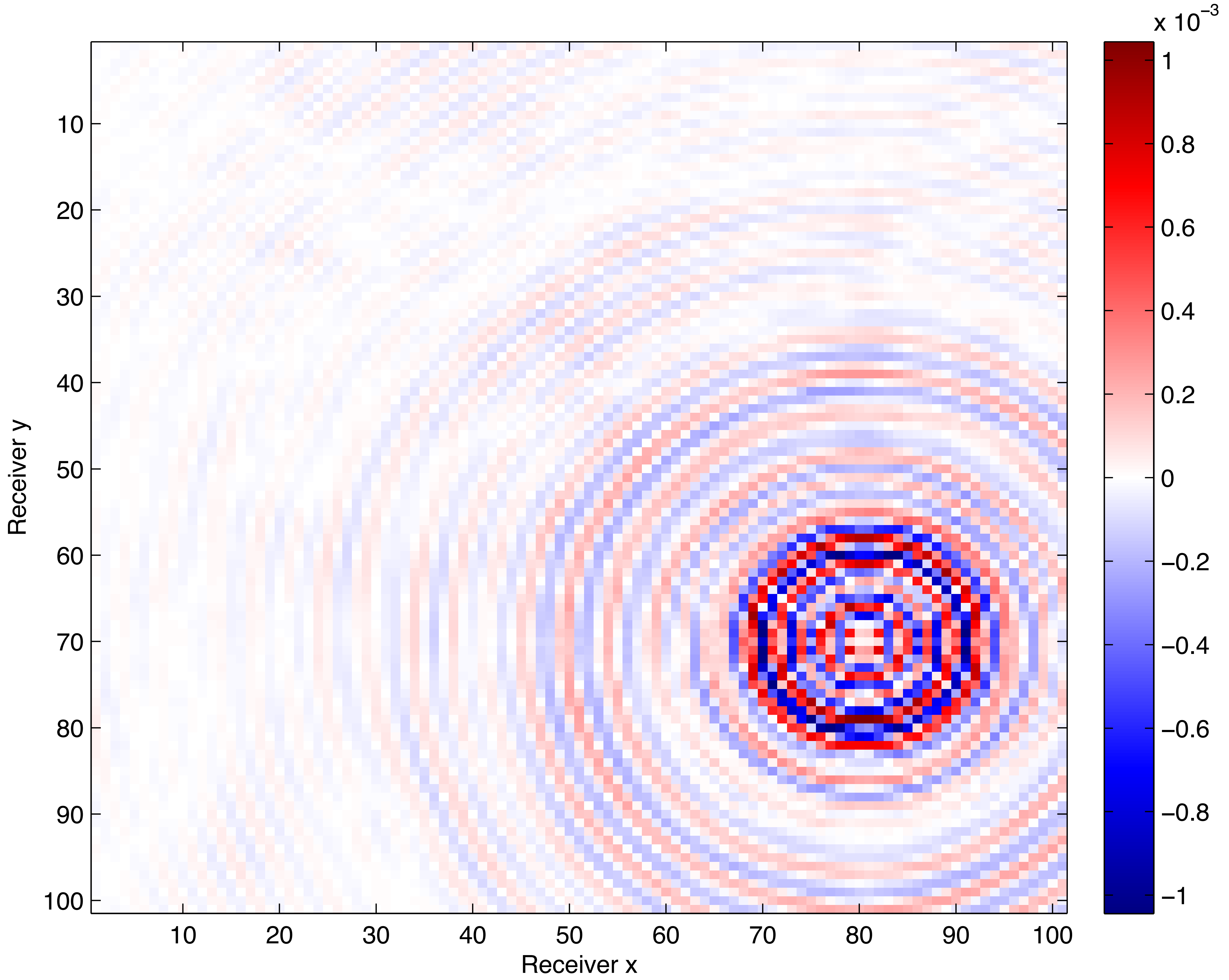


SR Recovery - SNR 3.05 dB

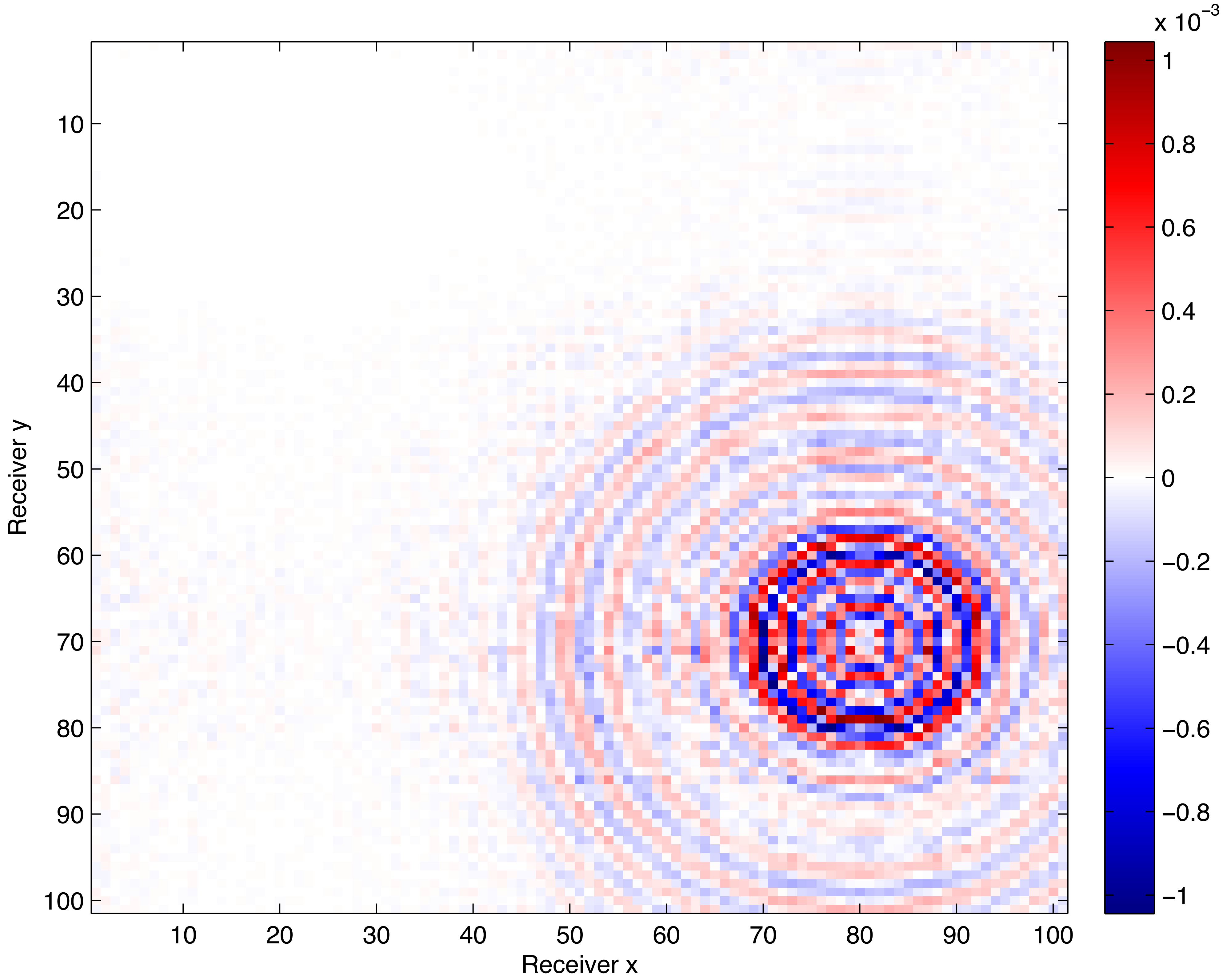
SLIM

7.34 Hz - 50% missing receivers - high noise

Common source gather



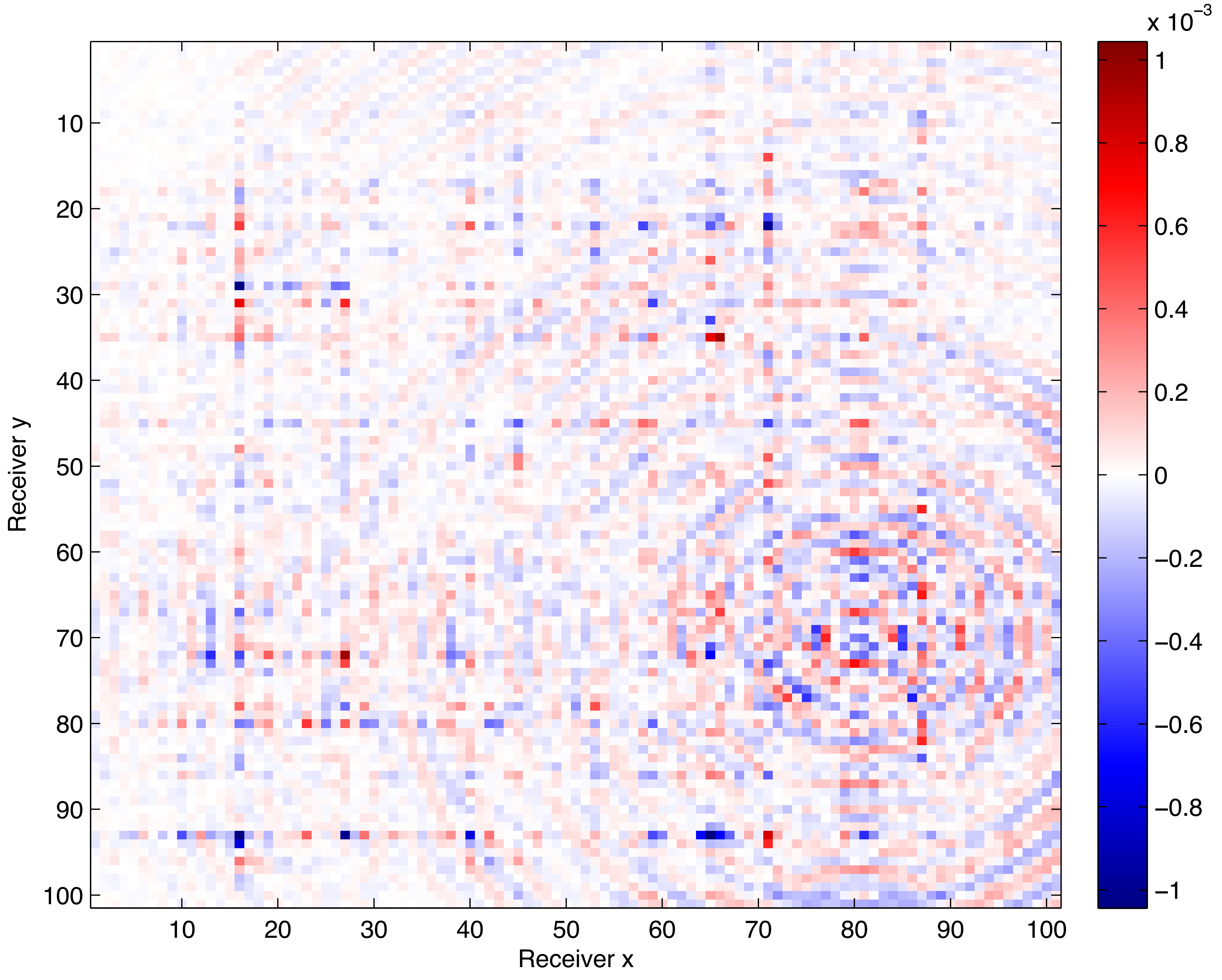
True data



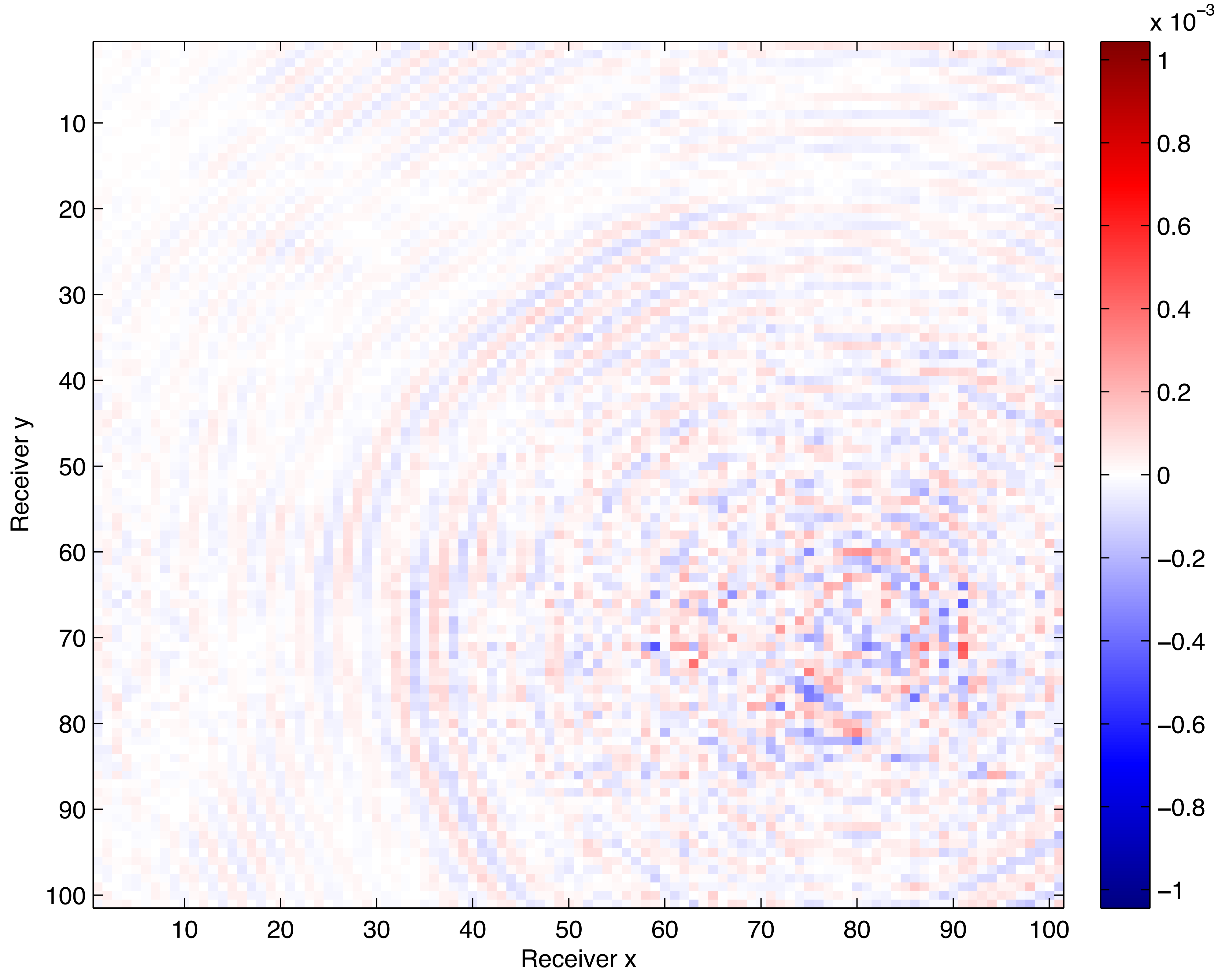
MH Recovery - SNR 8.95 dB

7.34 Hz - 50% missing receivers - high noise

Common source gather

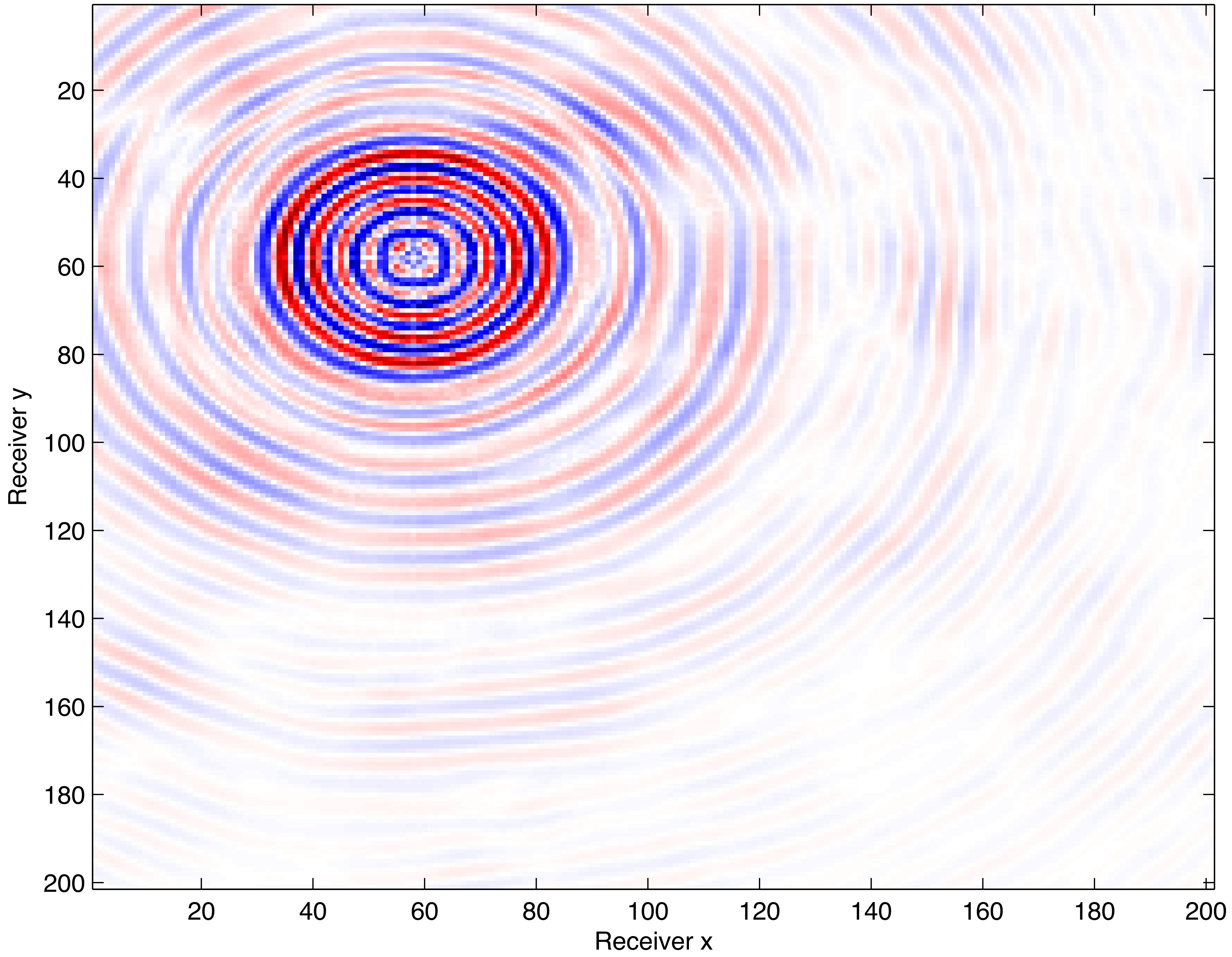


SR Difference

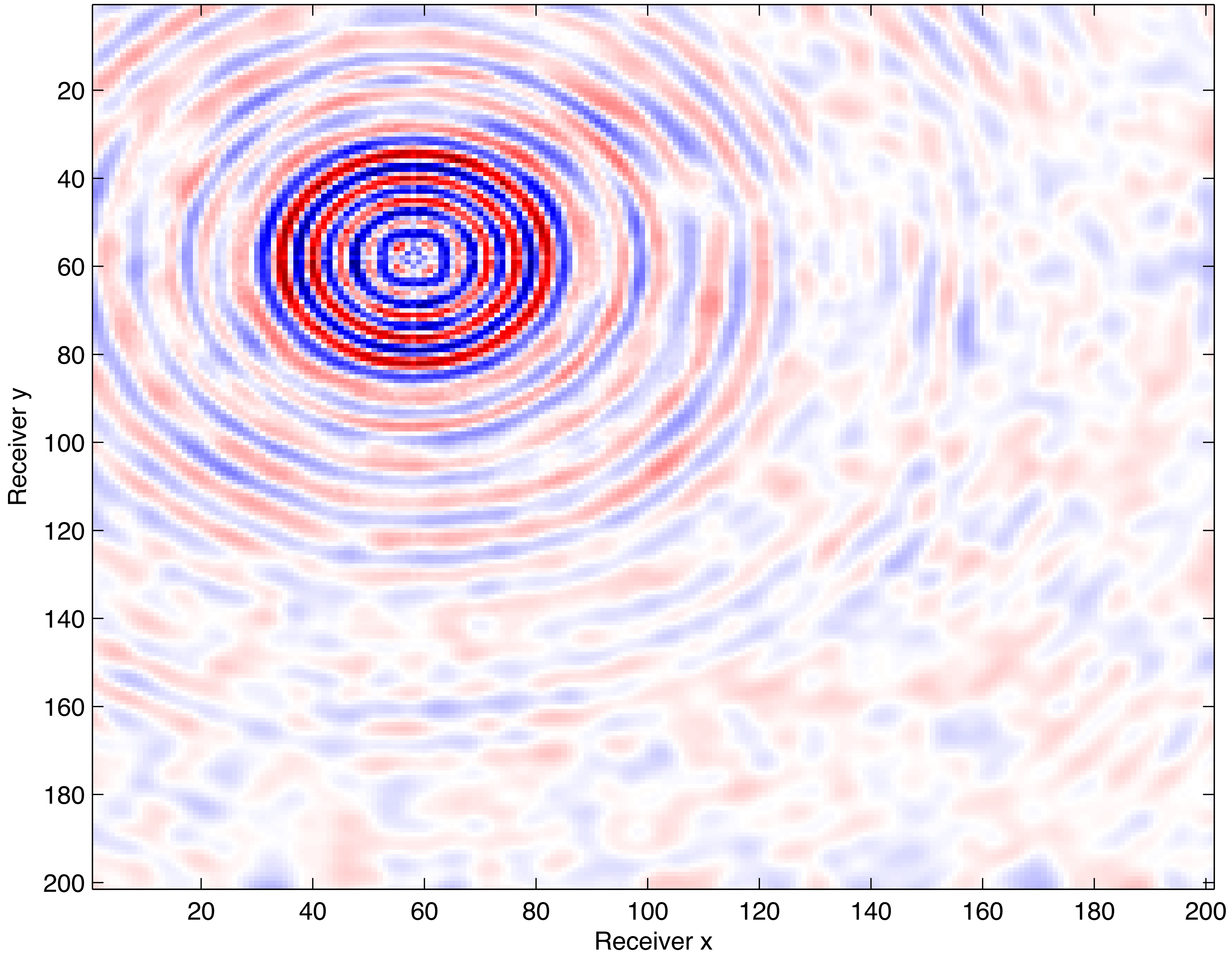


MH Difference

7.34 Hz - Denoising

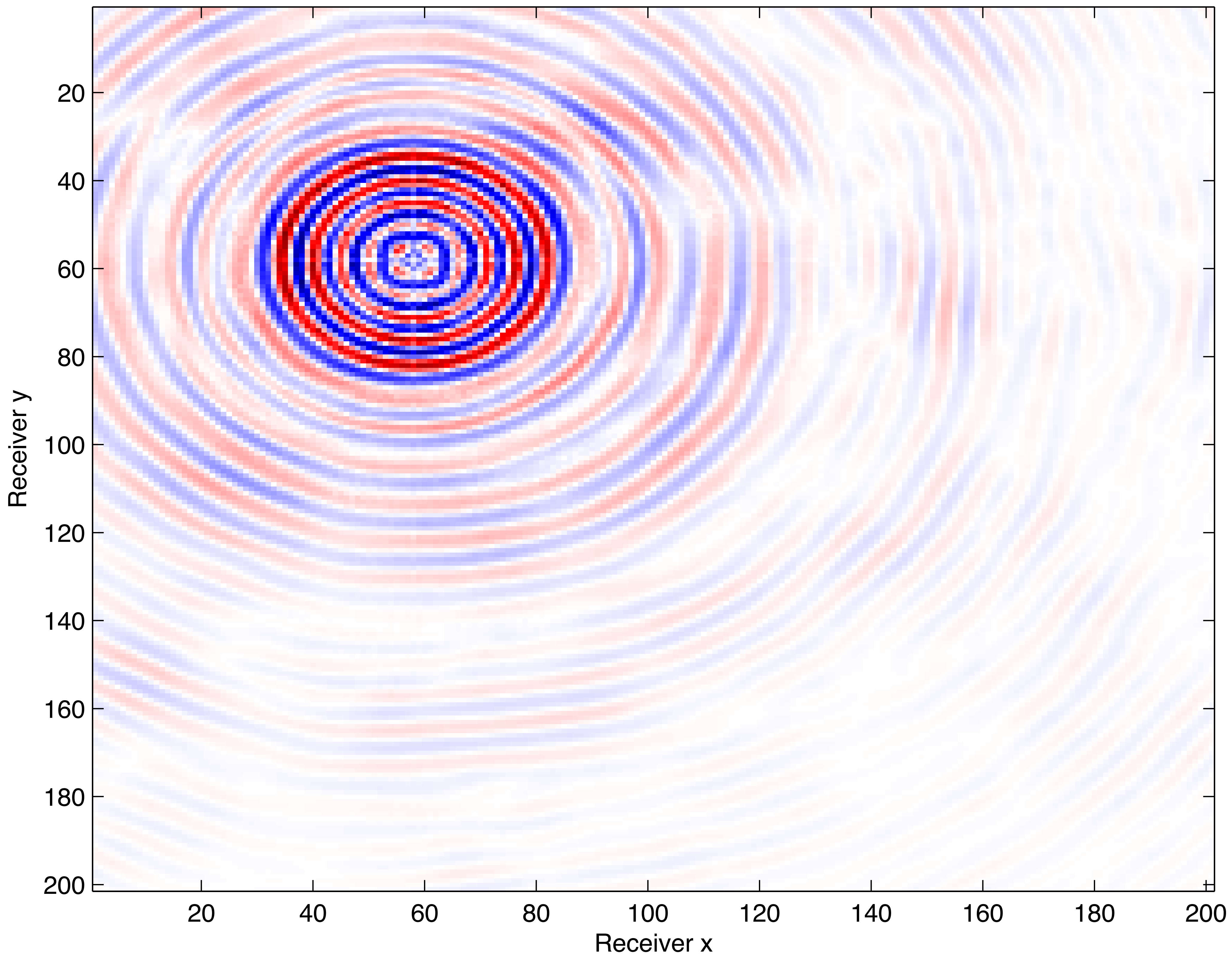


True data

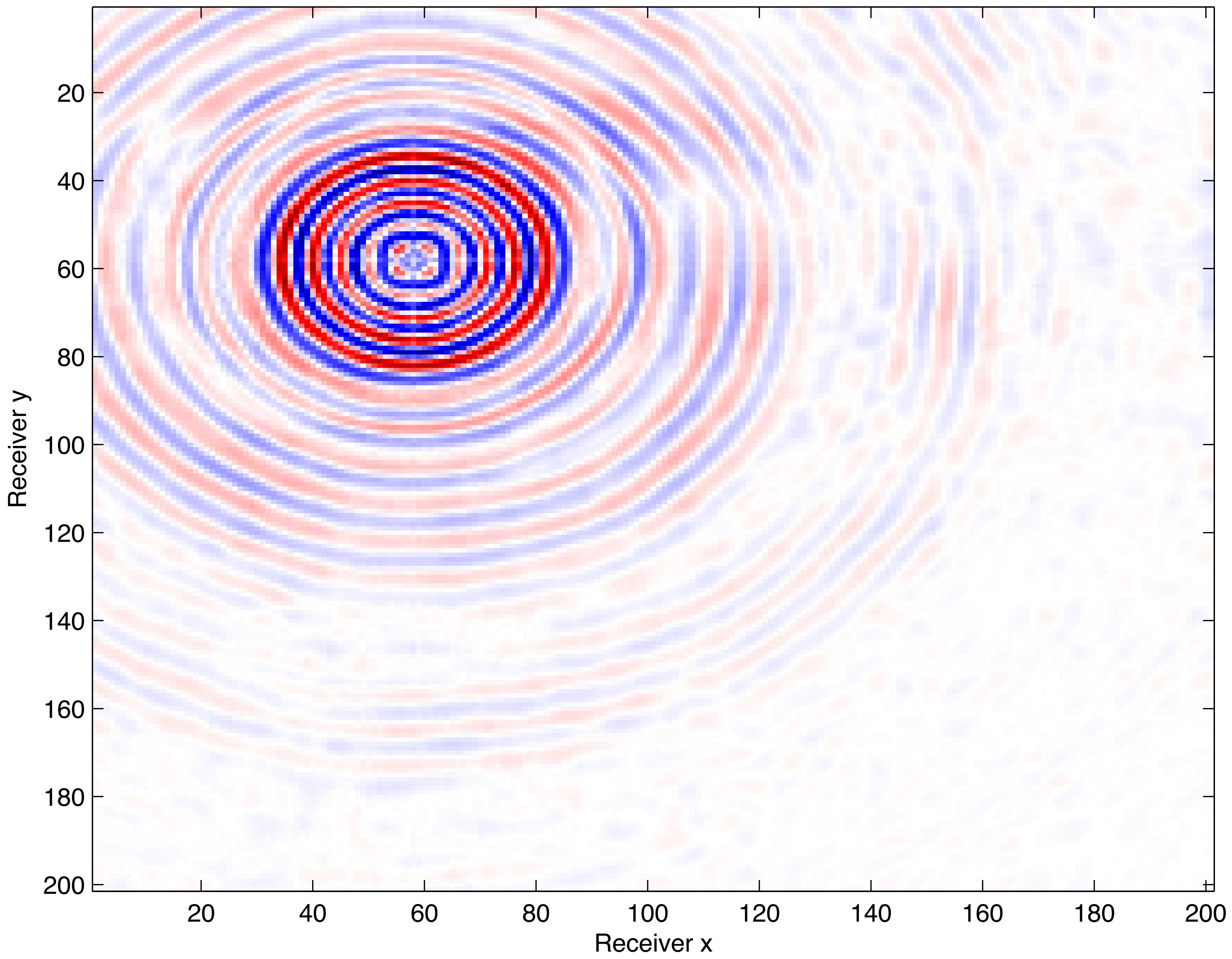


Noisy input - SNR 14.9 dB

7.34 Hz - Denoising

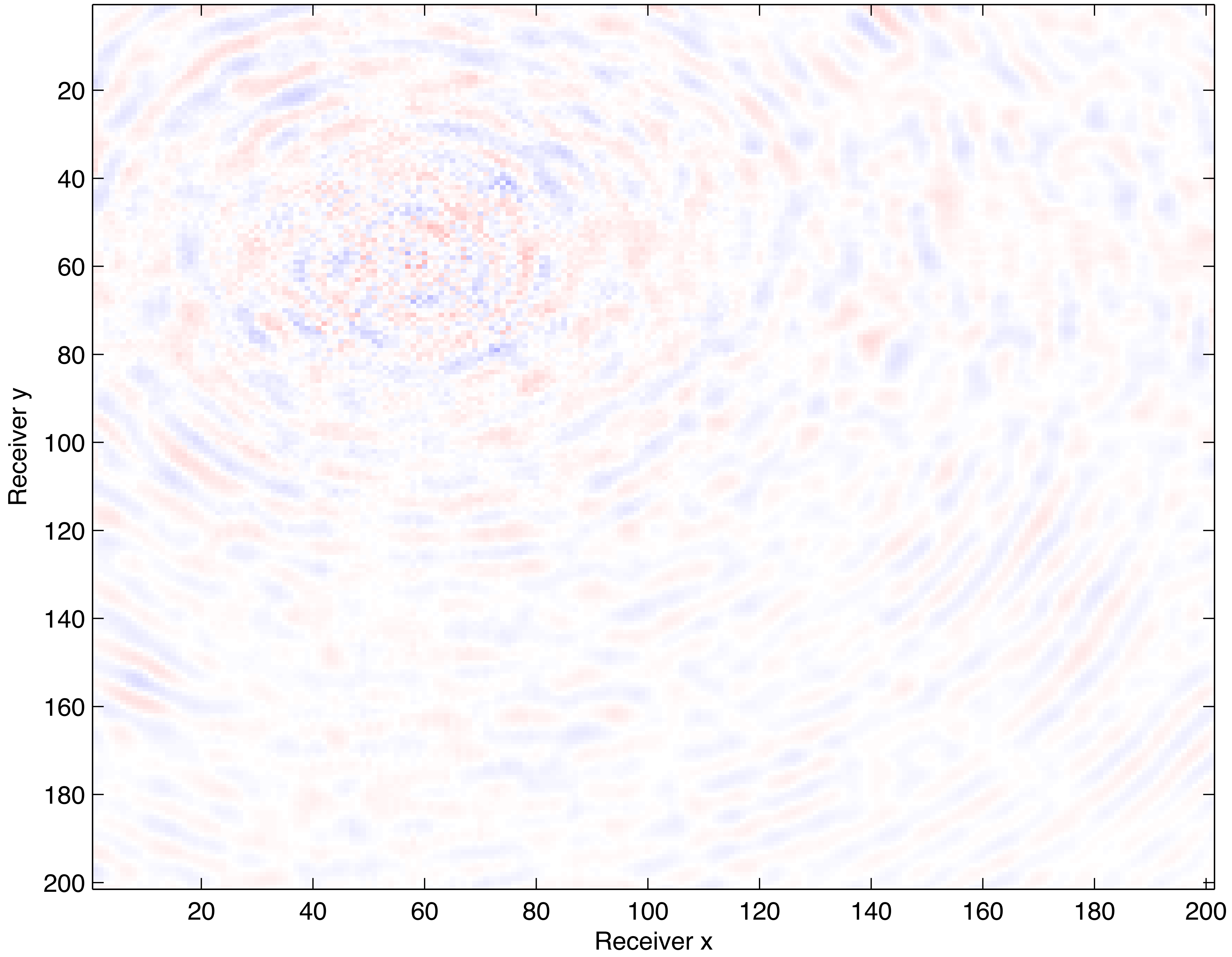


True data

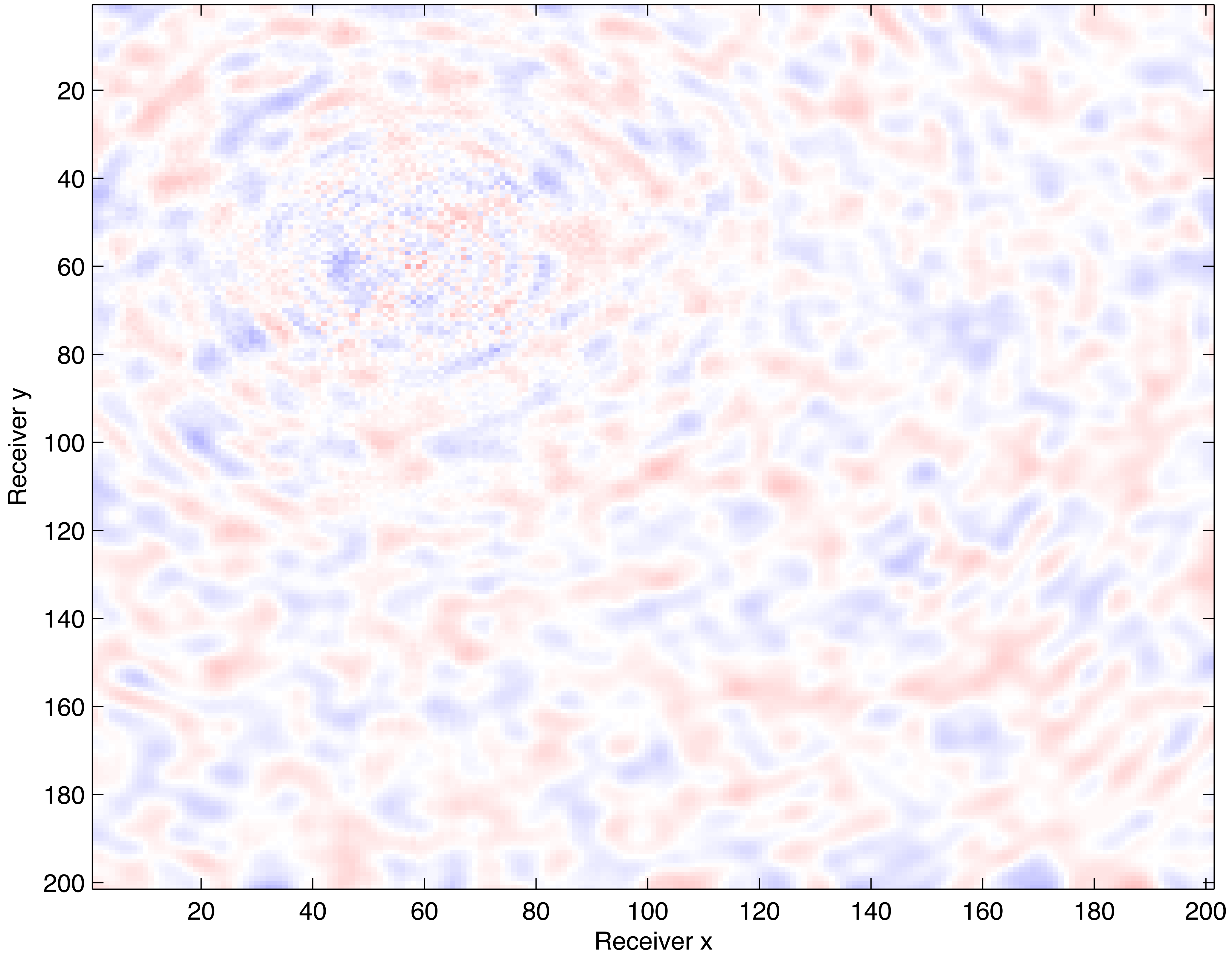


Estimated - SNR 20.1 dB

7.34 Hz - Denoising

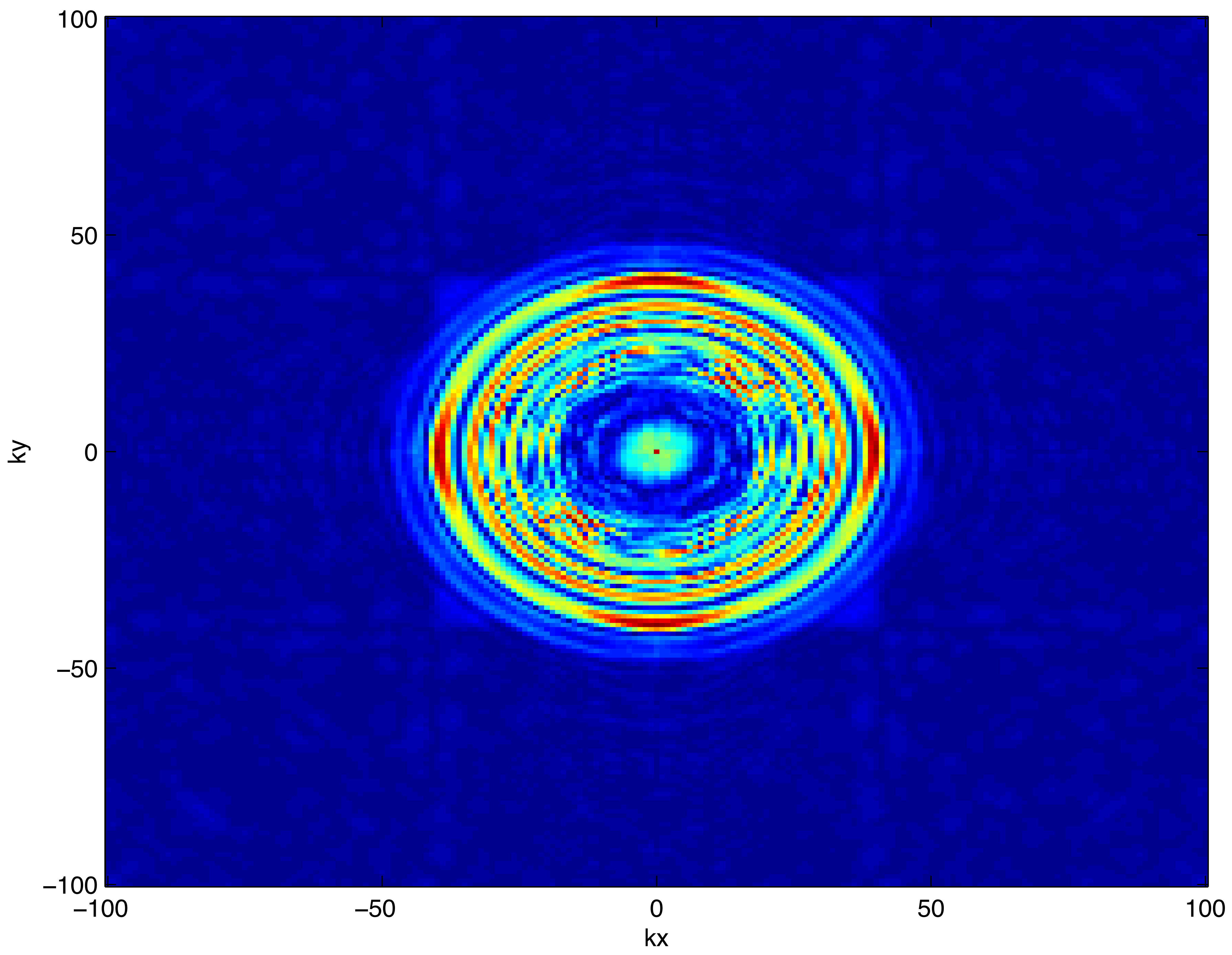


Difference

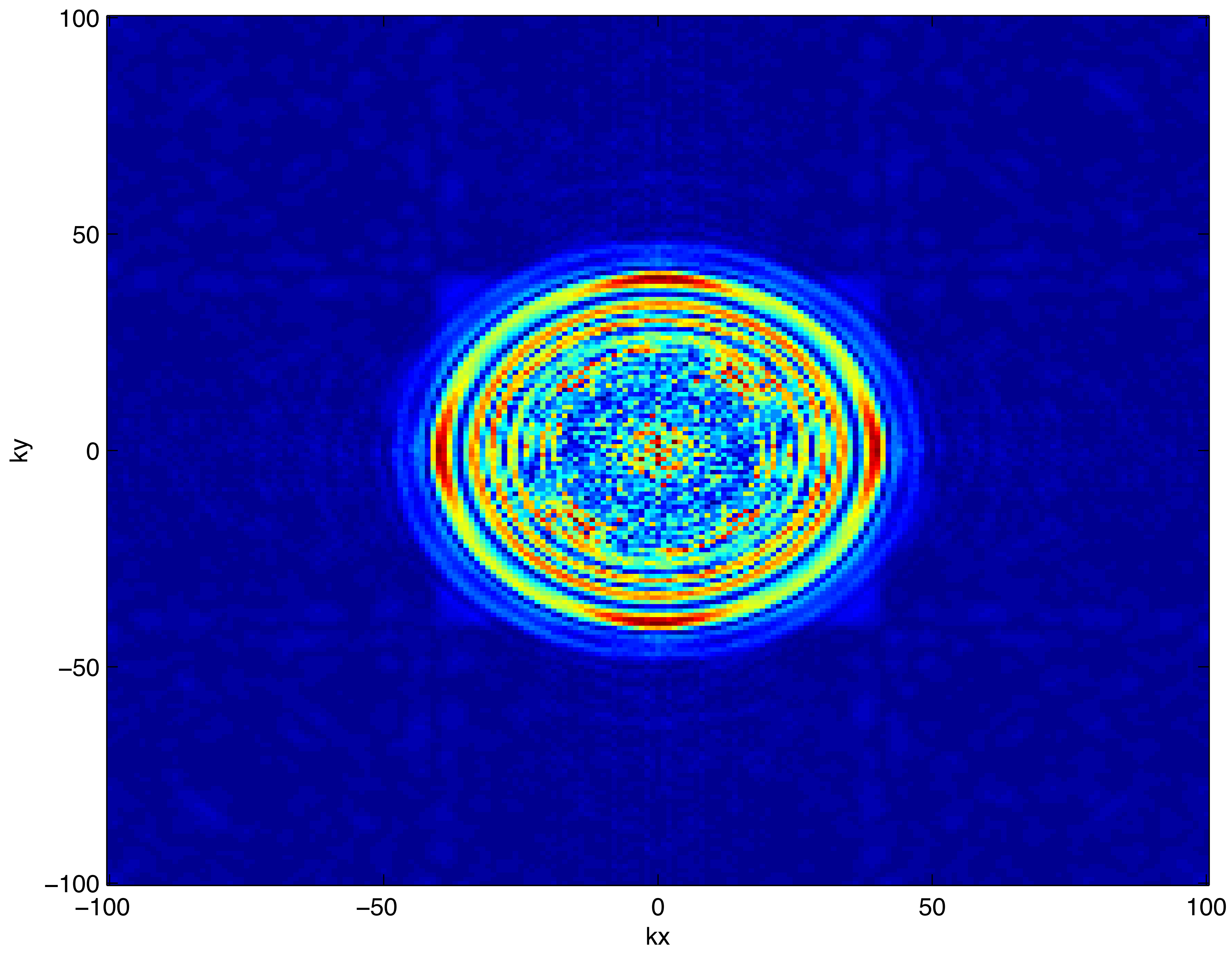


Input + output difference

7.34 Hz - Denoising

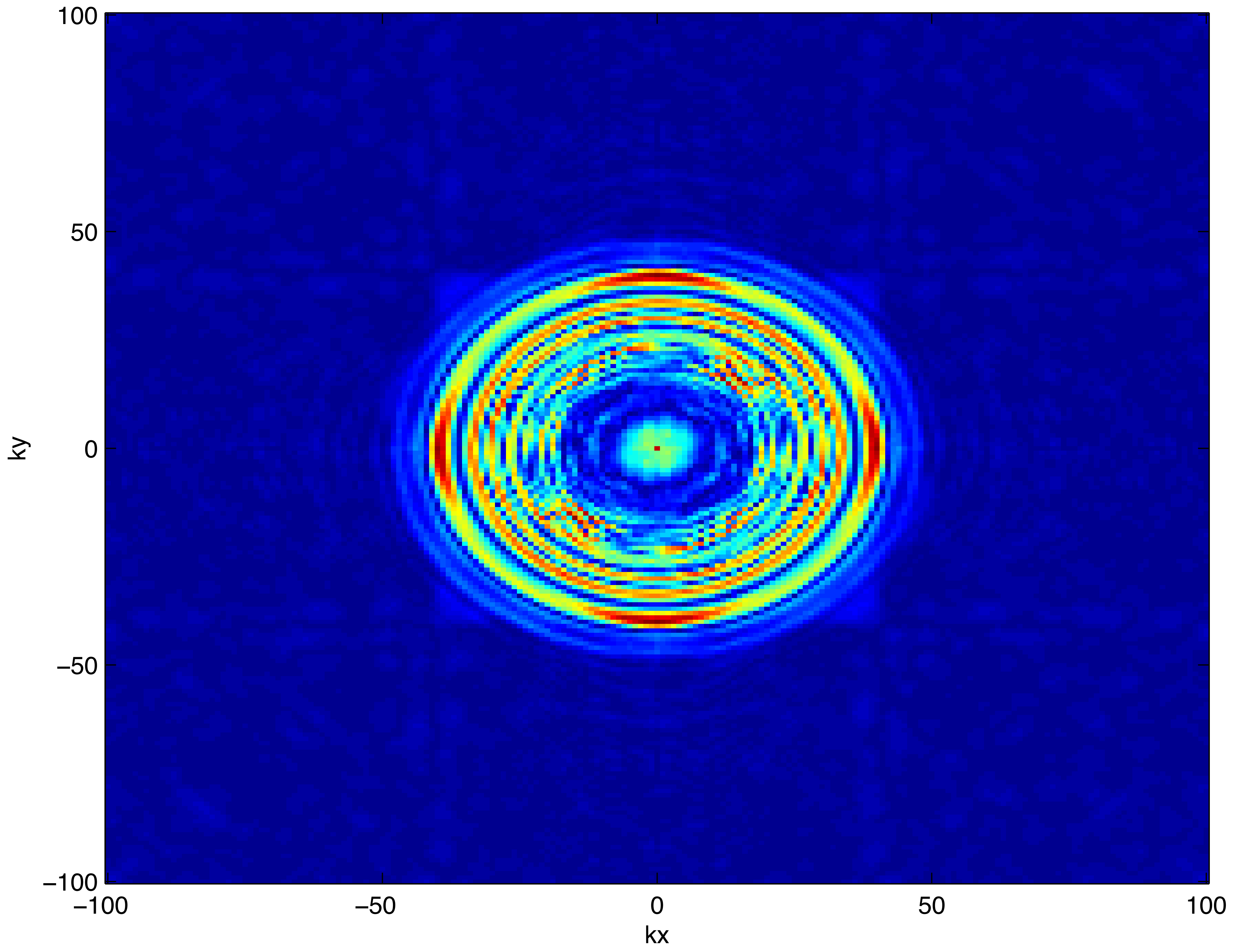


True spectrum

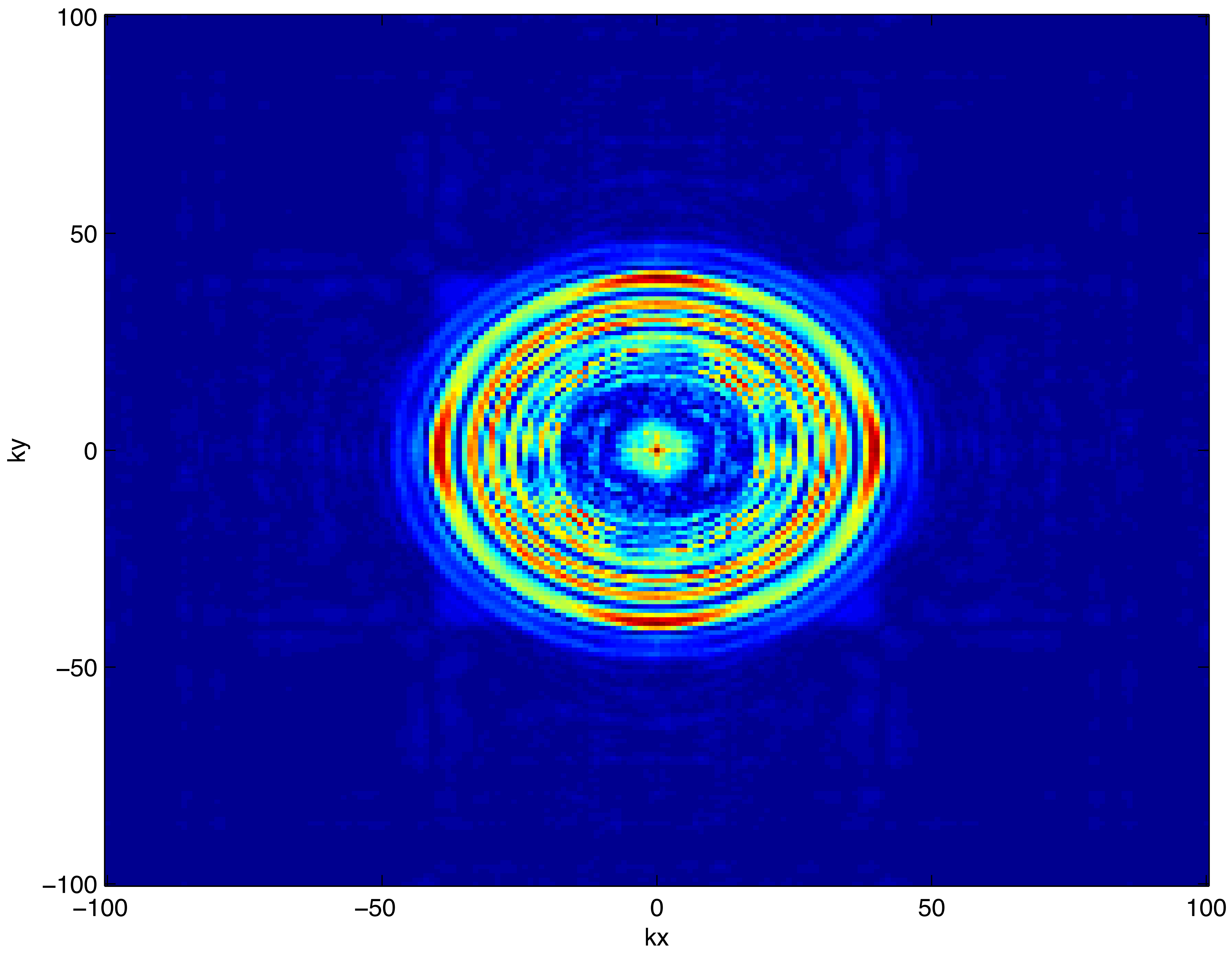


Noisy spectrum

7.34 Hz - Denoising



True spectrum



Estimated spectrum

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

Important to use an appropriate *transform domain* (e.g. midpoint offset) so that sampling + noise *increase* the singular values in that domain

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