2013 SEG HOUSTON
George R. Brown Convention Center
22–27 September 2013 • Houston, Texas USA

Seismic data interpolation via low-rank matrix factorization in the h(ierarchical) s(emi) s(eparable) representation

Rajiv Kumar, Hassan Mansour, Sasha Aravkin and Felix J. Herrmann

SLIM
University of British Columbia

Motivation

- acquisition challenges
  - missing data
  - noise

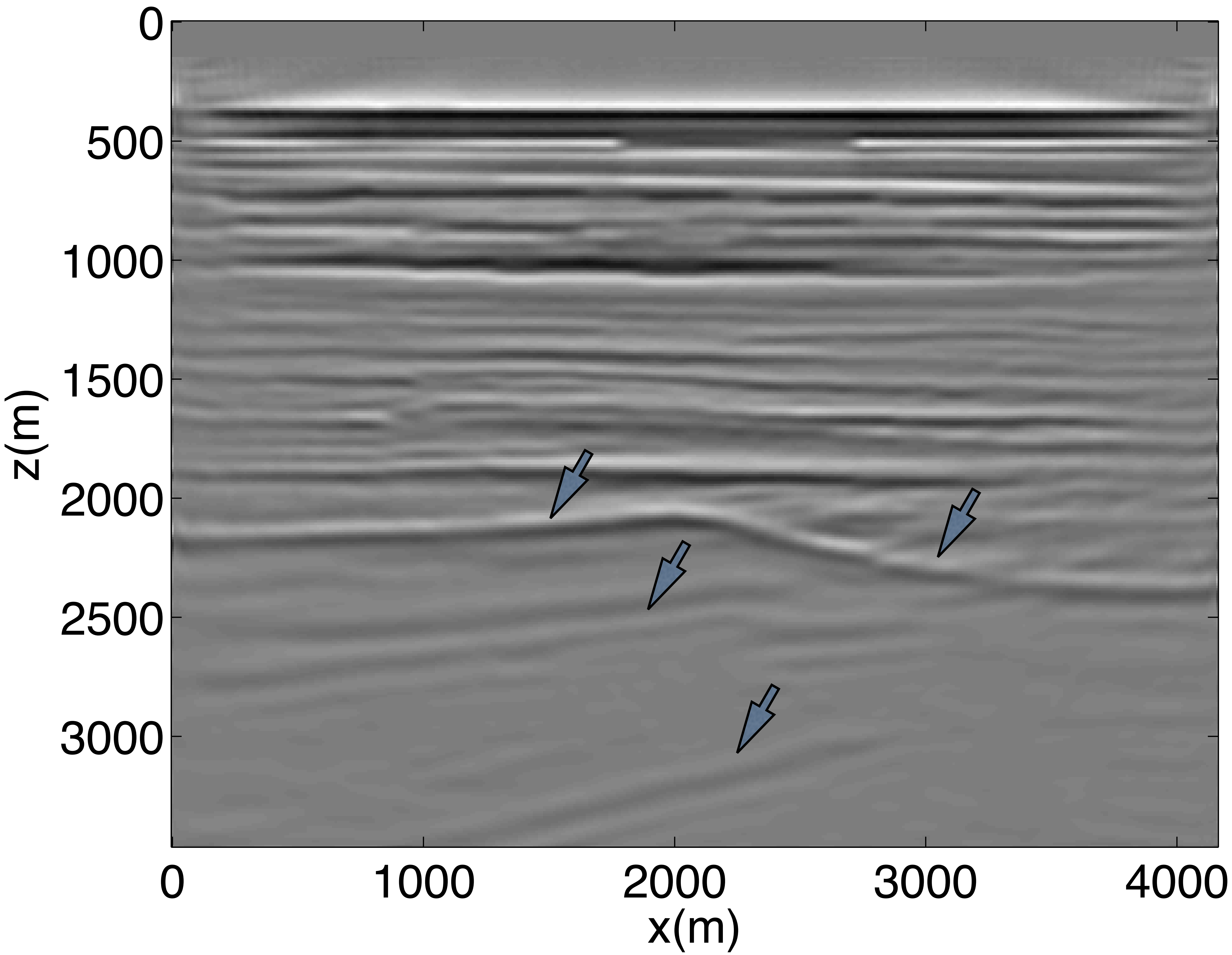
- fully sampled data
  - simultaneous shot based FWI & migration
  - estimation of primaries by sparse inversion & SRME

- exploit *low-rank* structure of seismic data
  - *SVD-free* matrix factorization
  - benefit of HSS

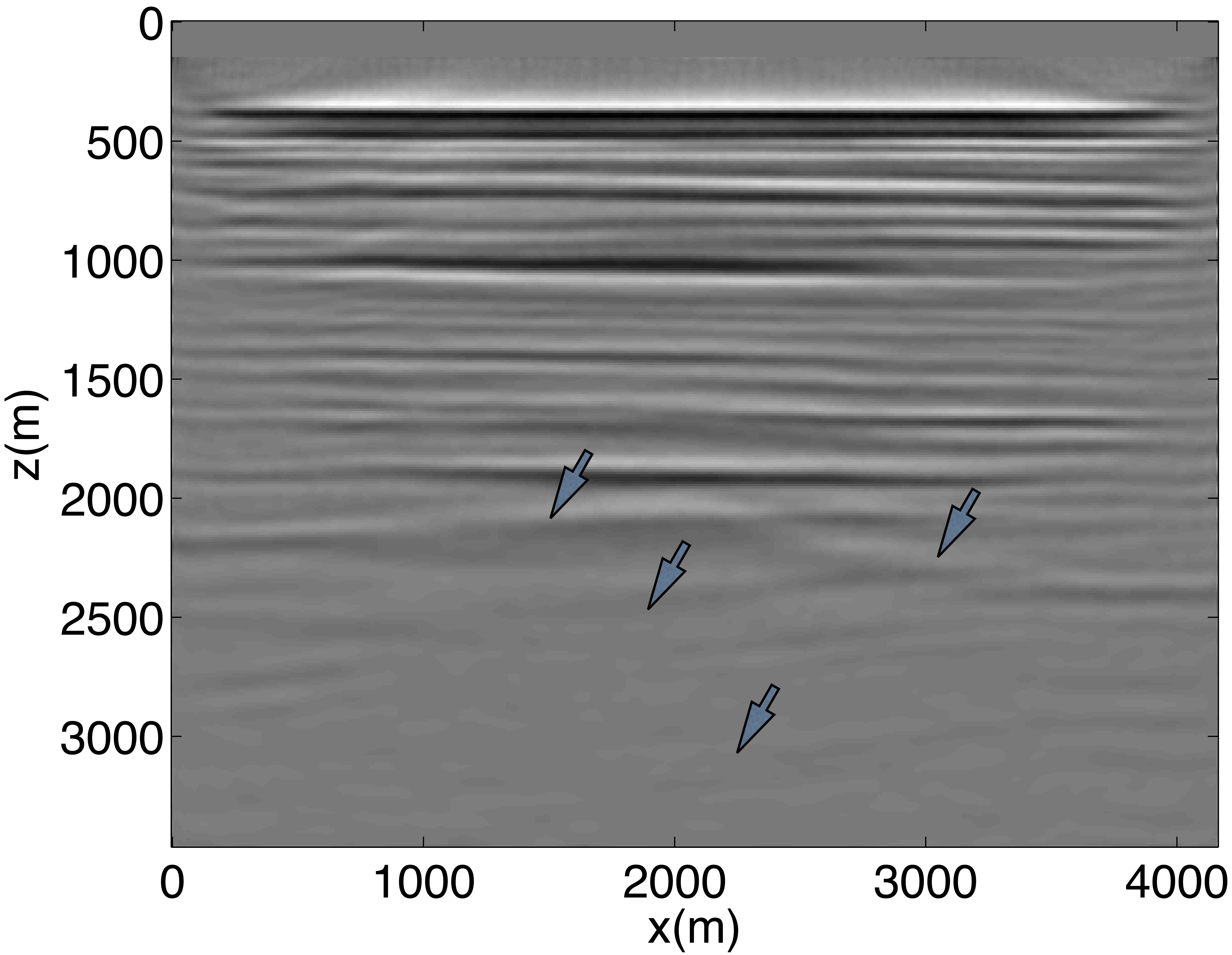
[Tu et. al. 12, li et. al. 12]

Migration

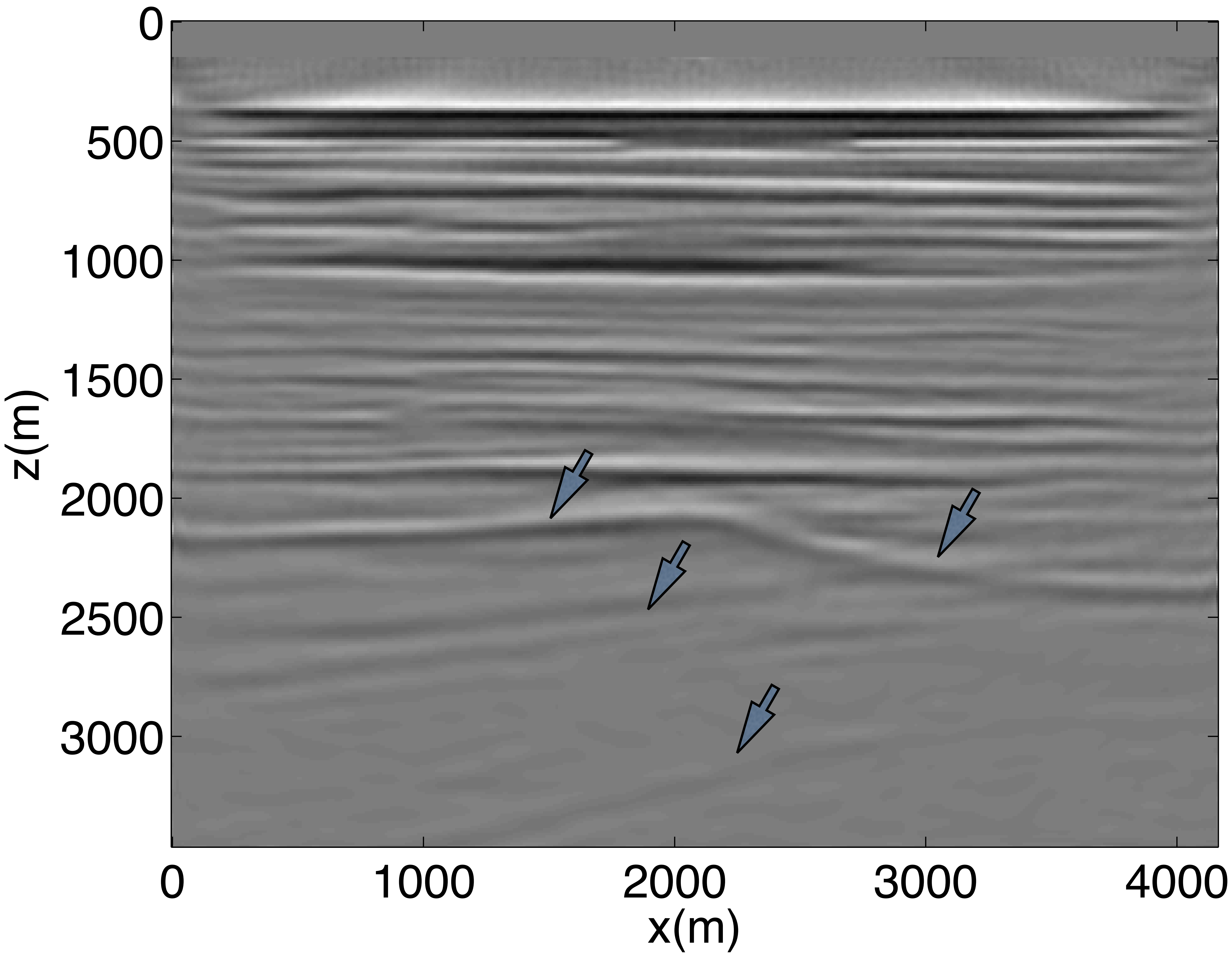
[fully sample data, 30 simultaneous shots, redraw]



Migration
[w/o HSS, 30 simultaneous shots, redraw]



Migration
[w/ HSS, 30 simultaneous shots, redraw]



[1] Oropeza V and M D Sacchi, 2011, Simultaneous seismic data de-noising and reconstruction via Multichannel Singular Spectrum Analysis (MSSA), Geophysics, 76 (3), V25-V32.
[2] Nadia Kreimer, Aaron Stanton and Mauricio D. Sacchi, Tensor completion based on nuclear norm minimization for 5D seismic data reconstruction, Technical report, 2013

Existing work

- SVD computation [1,2]
  - expensive for large scale data
- Overlap windows in space and time [2]
  - heuristic approach

[Candes and Donoho 2000, Herrmann 2008]

Compressive sensing
[transform based]

- signal structure
  - *sparse/compressible*
- sampling scheme
  - random missing traces make signal *less sparse* in transform domain
- recovery using *sparsity promoting* scheme

is *sparsity* the only inherent structure in seismic??

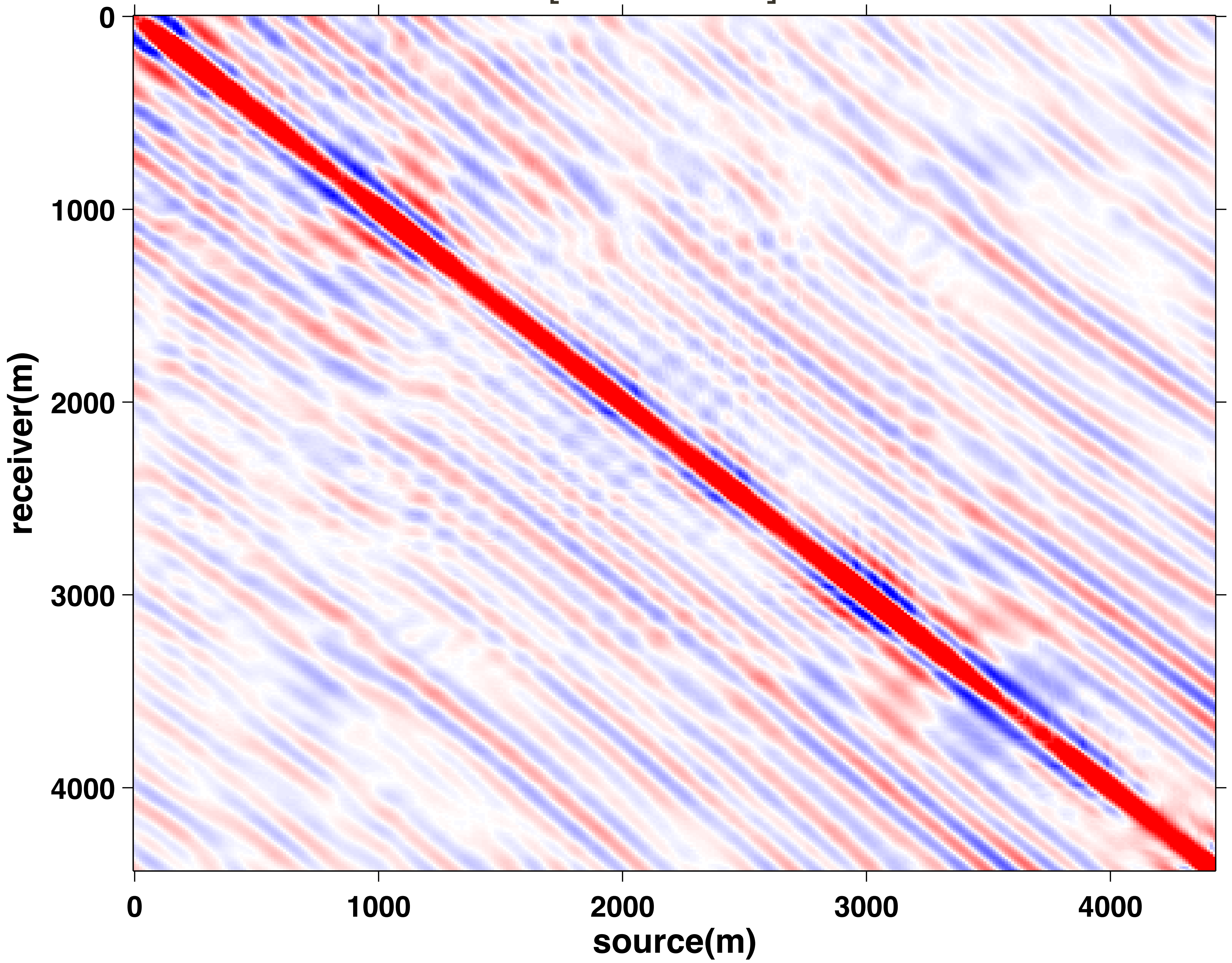
[Candes and Plan 2010, Oropeza and Sacchi 2011]

Matrix completion

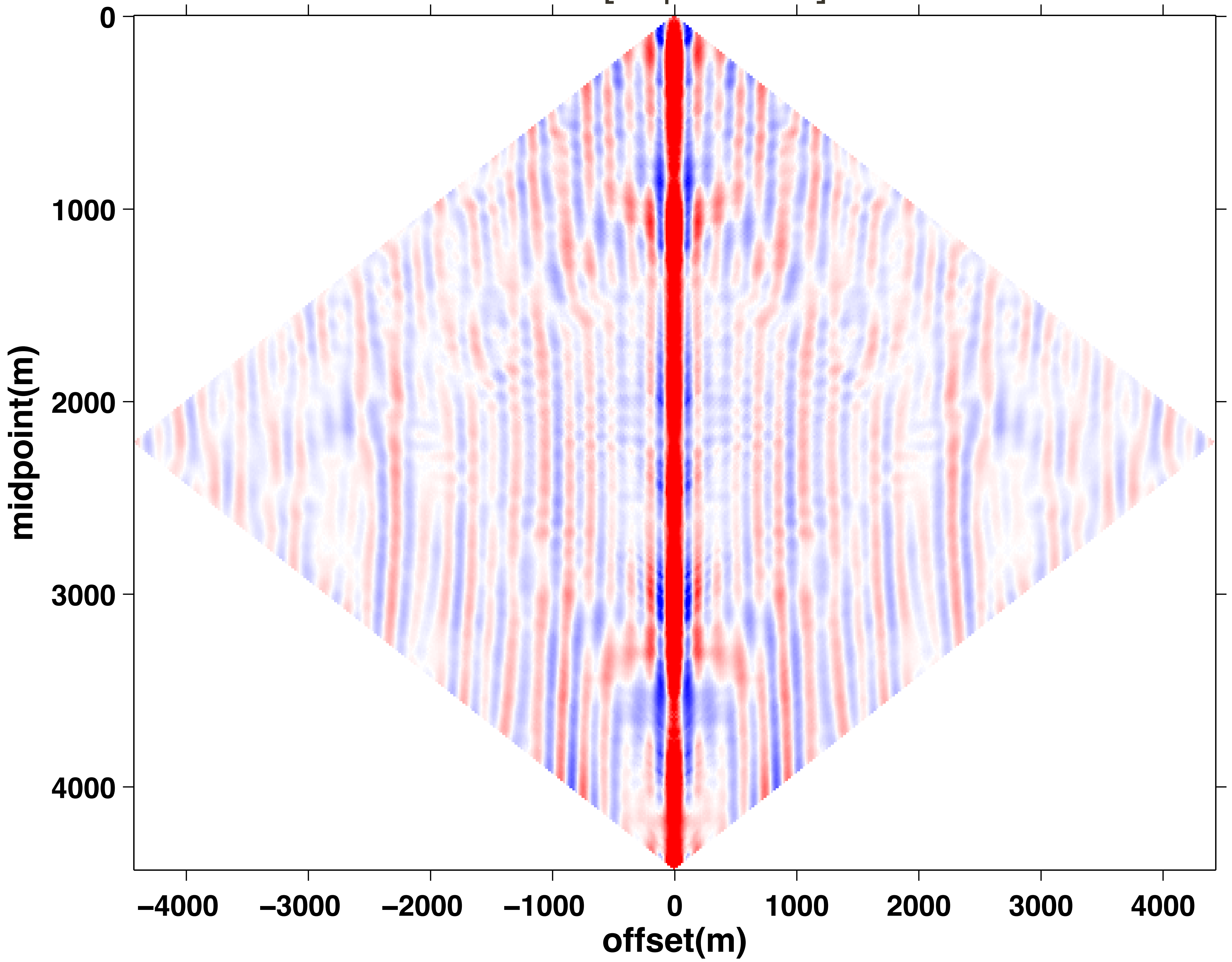
- signal structure
  - *low rank/fast decay* of singular values
- sampling scheme
  - missing data *increase* rank in "transform domain"
- recovery using *rank penalization* scheme

Low-rank structure
[2-D acquisition]

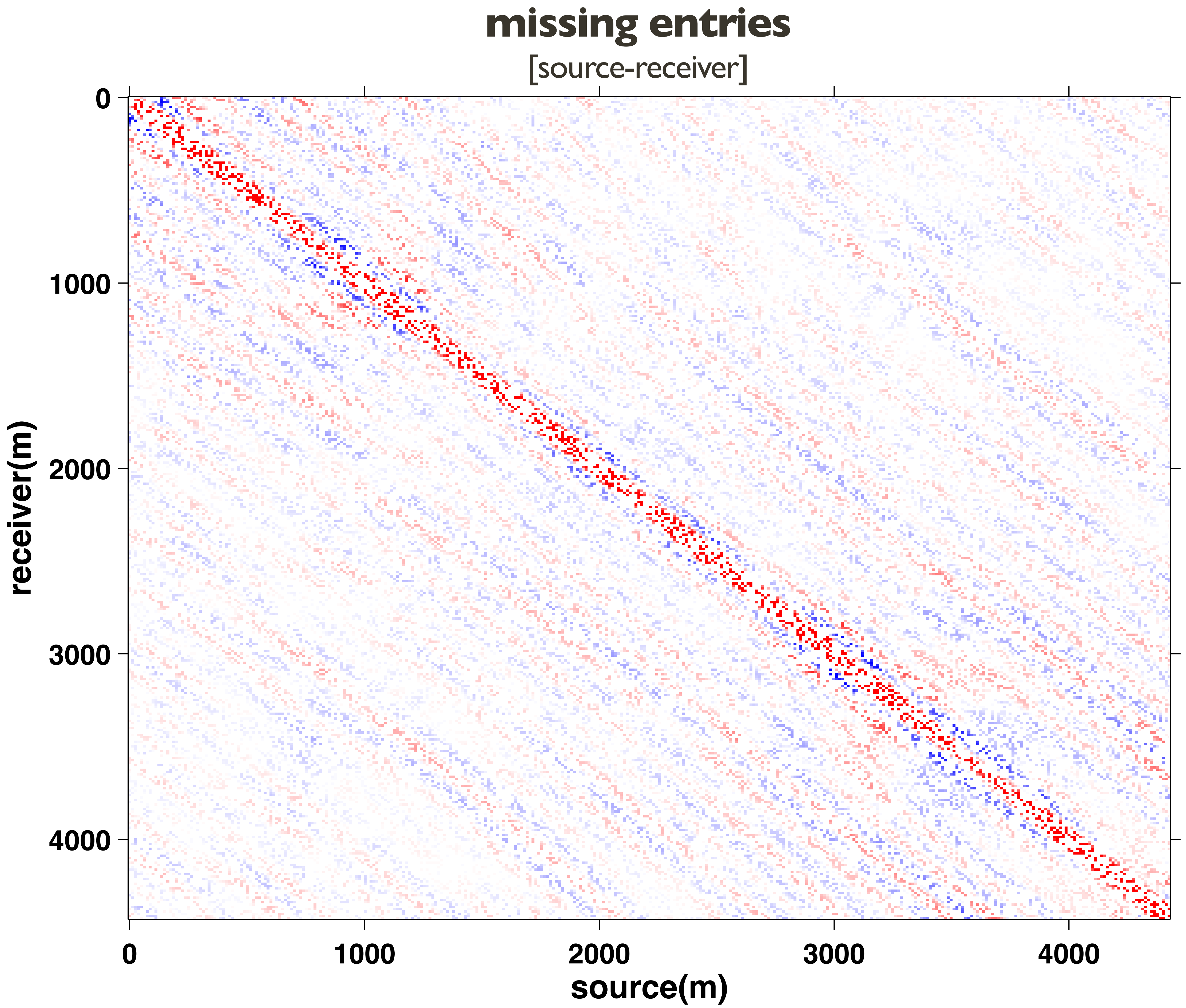
acquisition domain
[source-receiver]

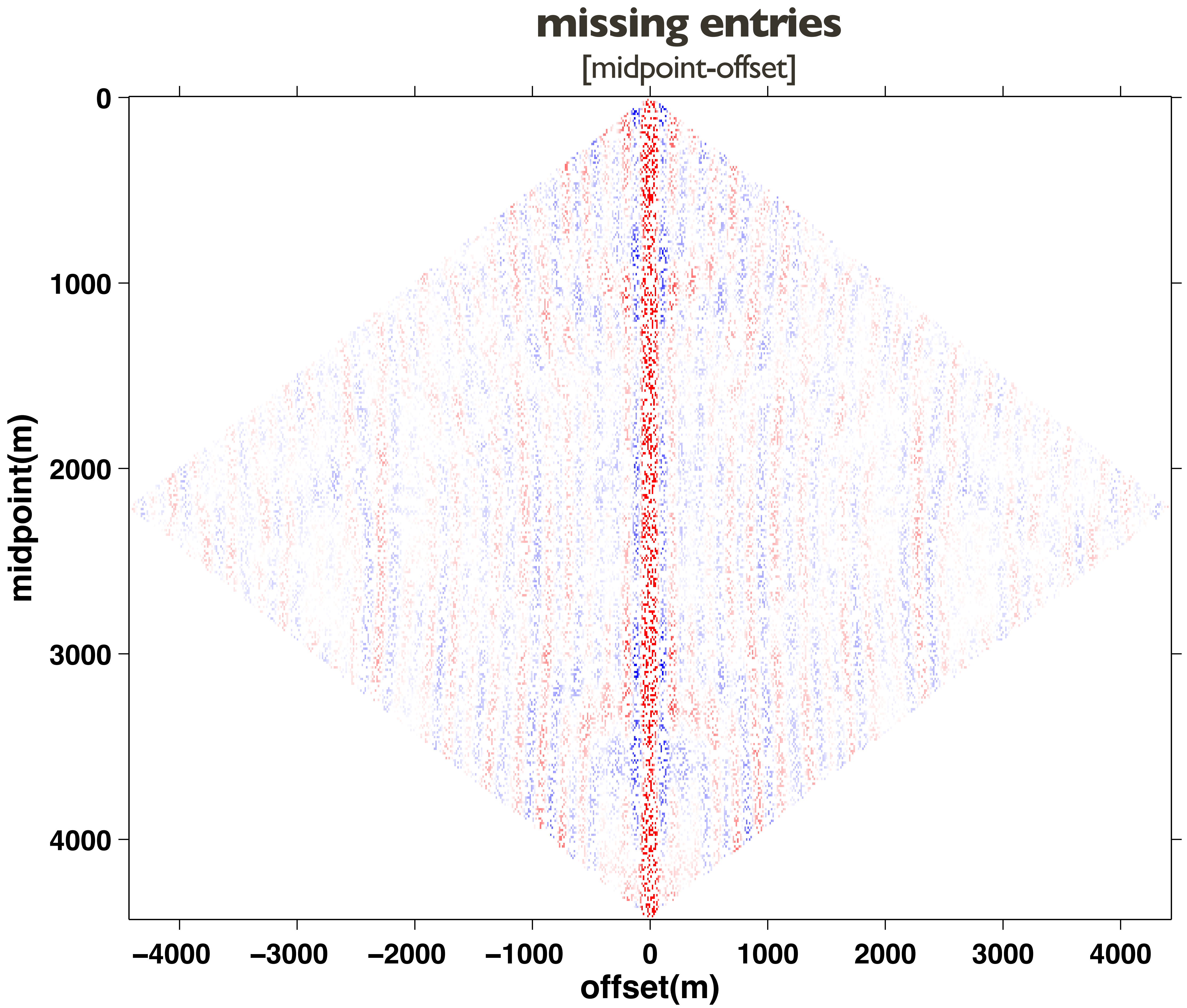


transform domain
[midpoint-offset]

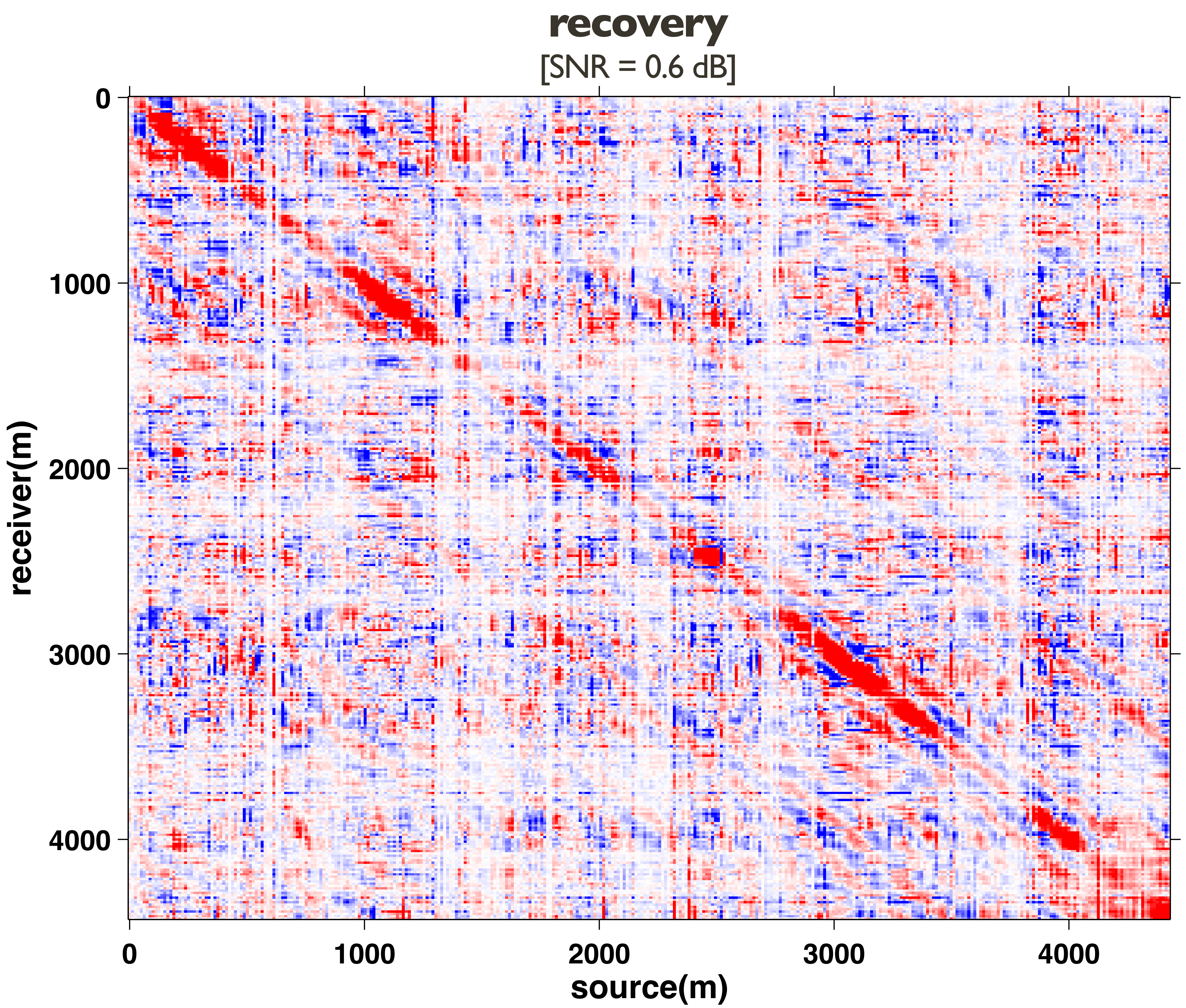


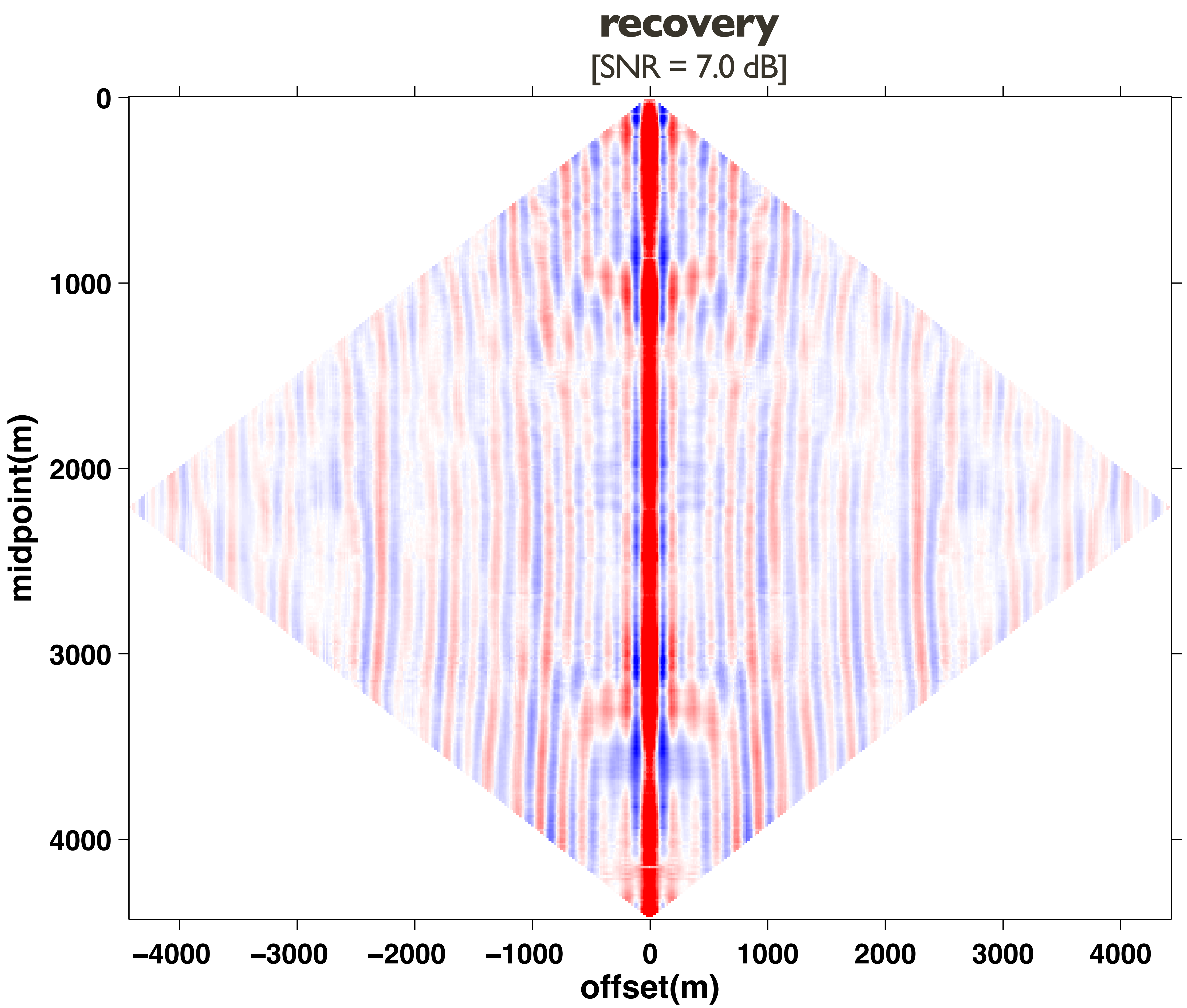
Matrix completion problem



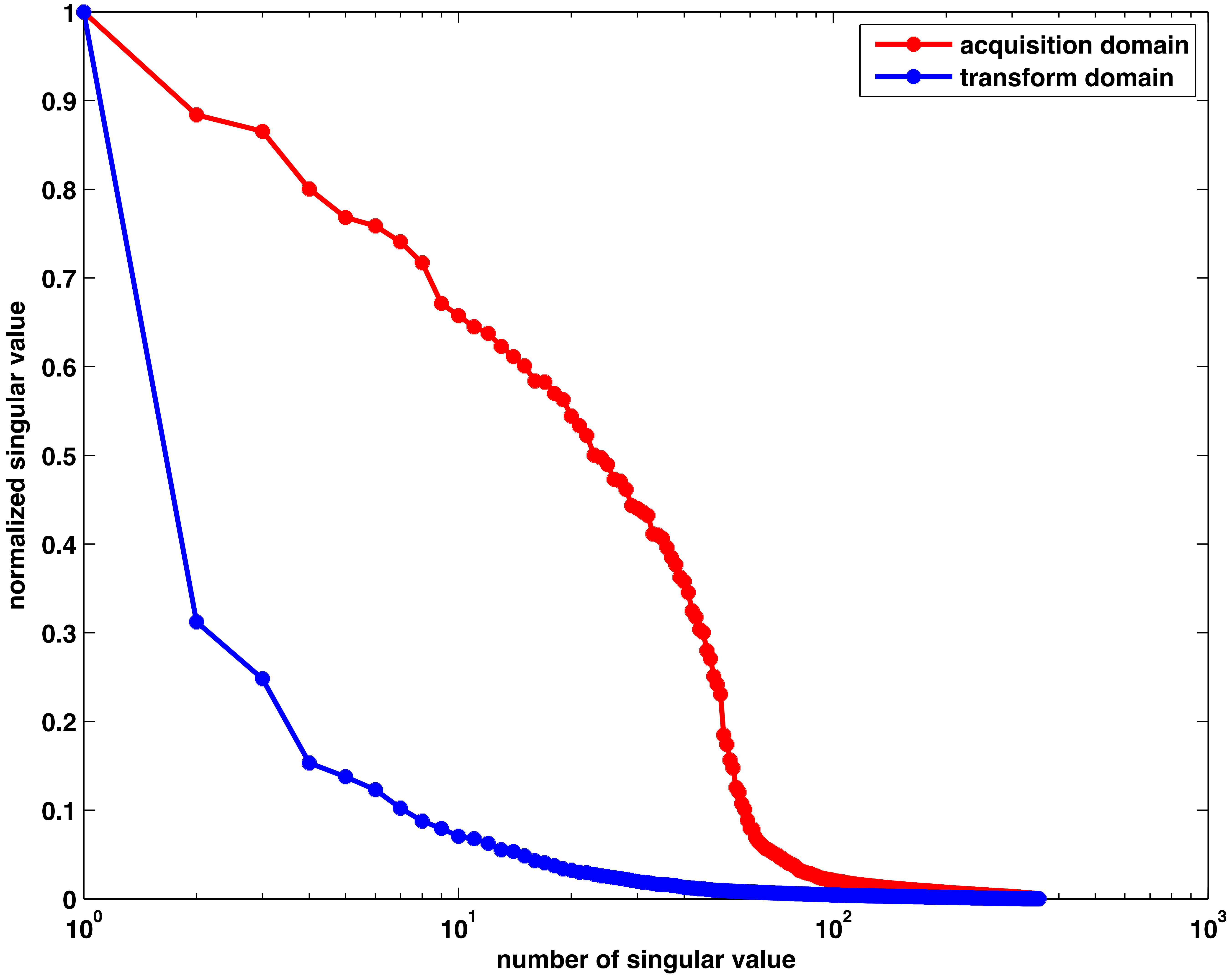


Low-rank interpolation

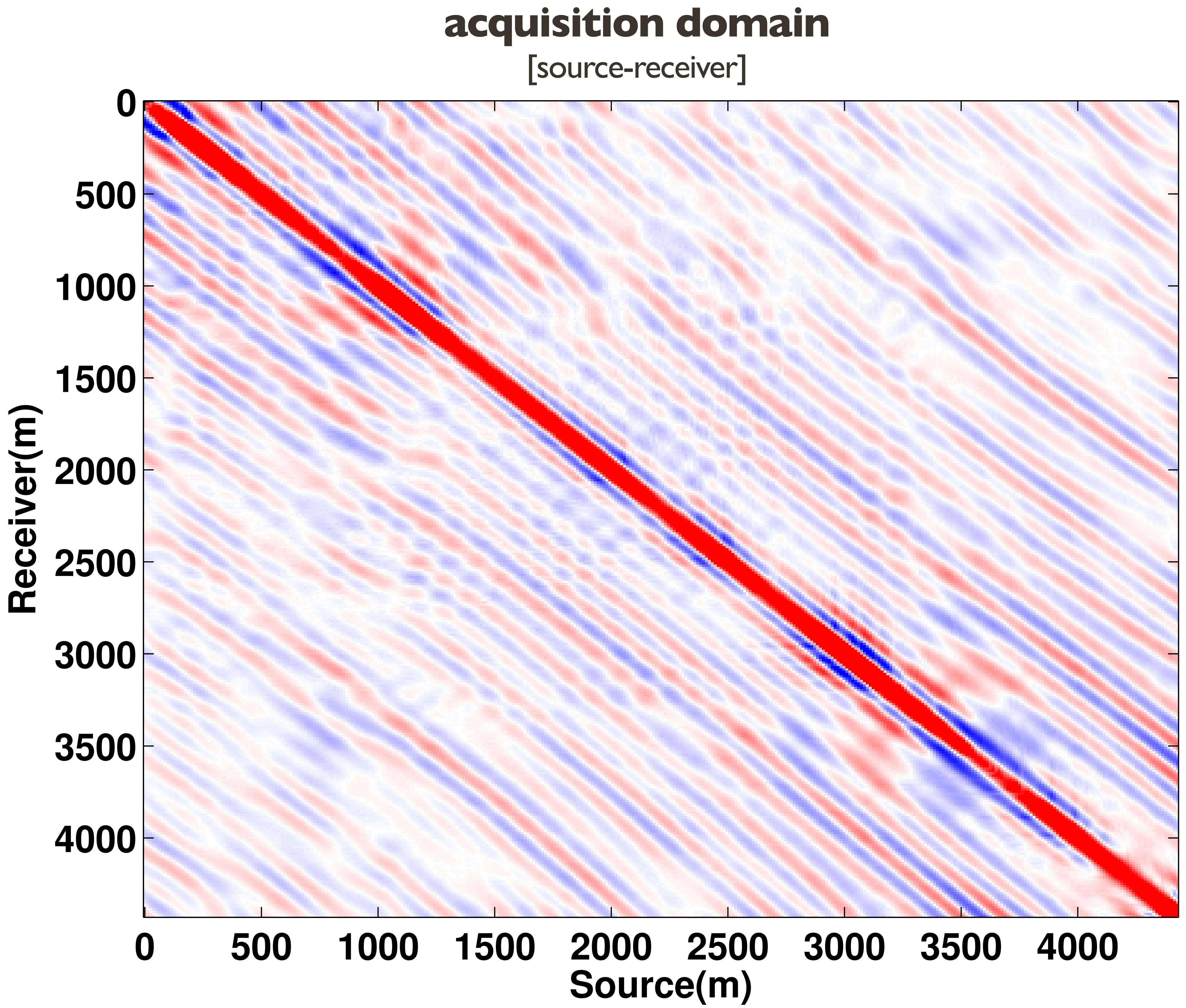


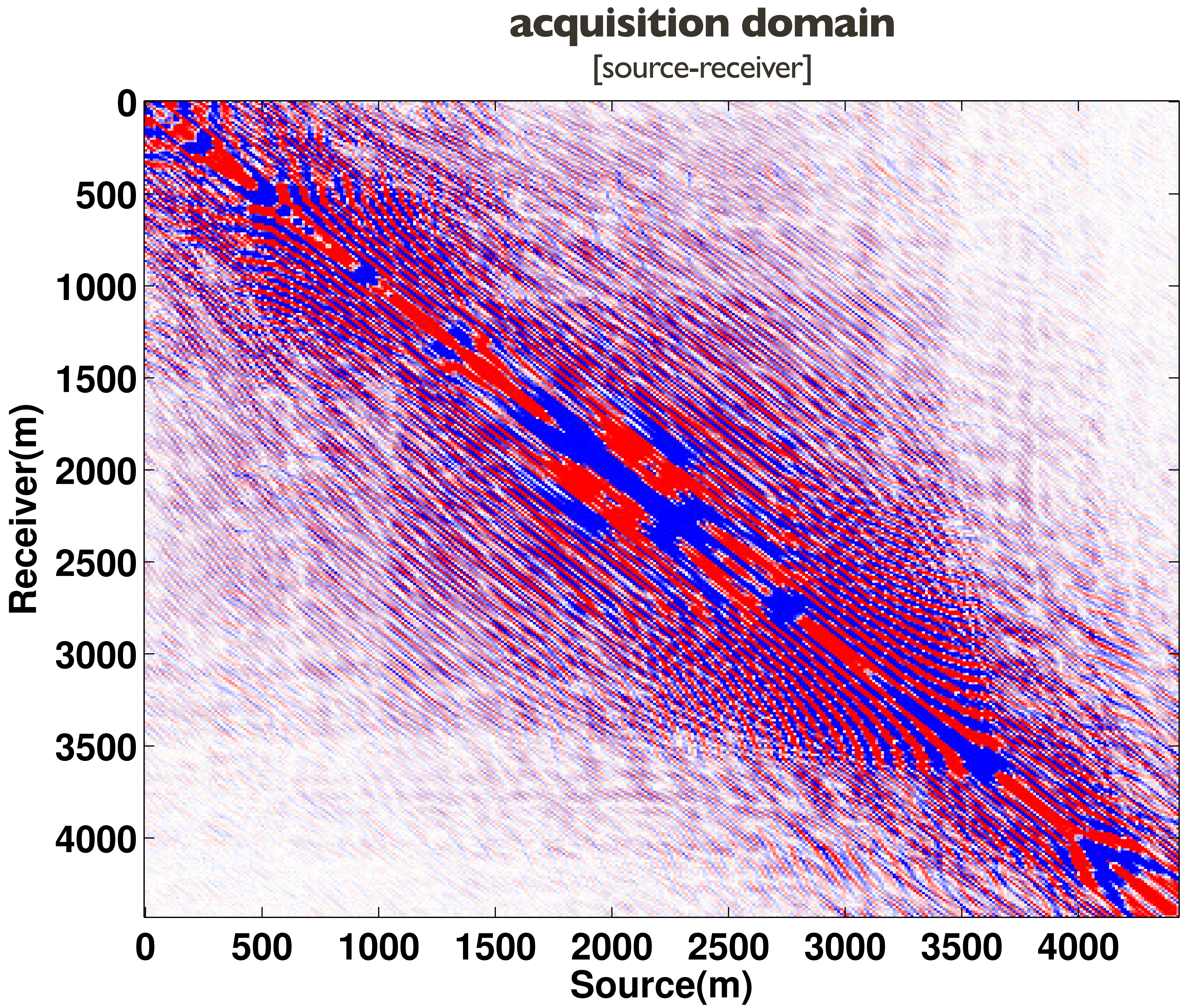


Singular value decay
[2-D acquisition]

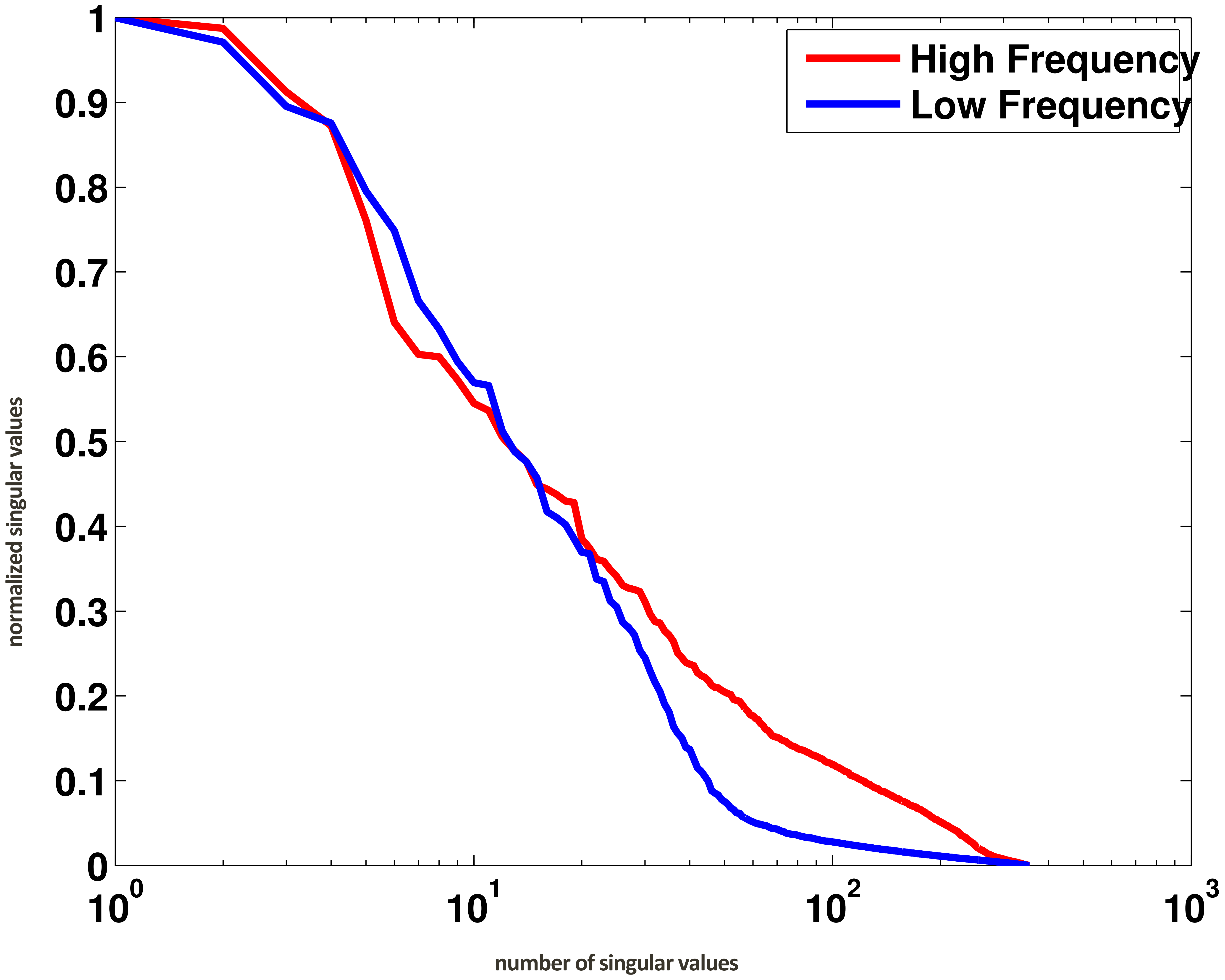


Is high frequency low-rank ?



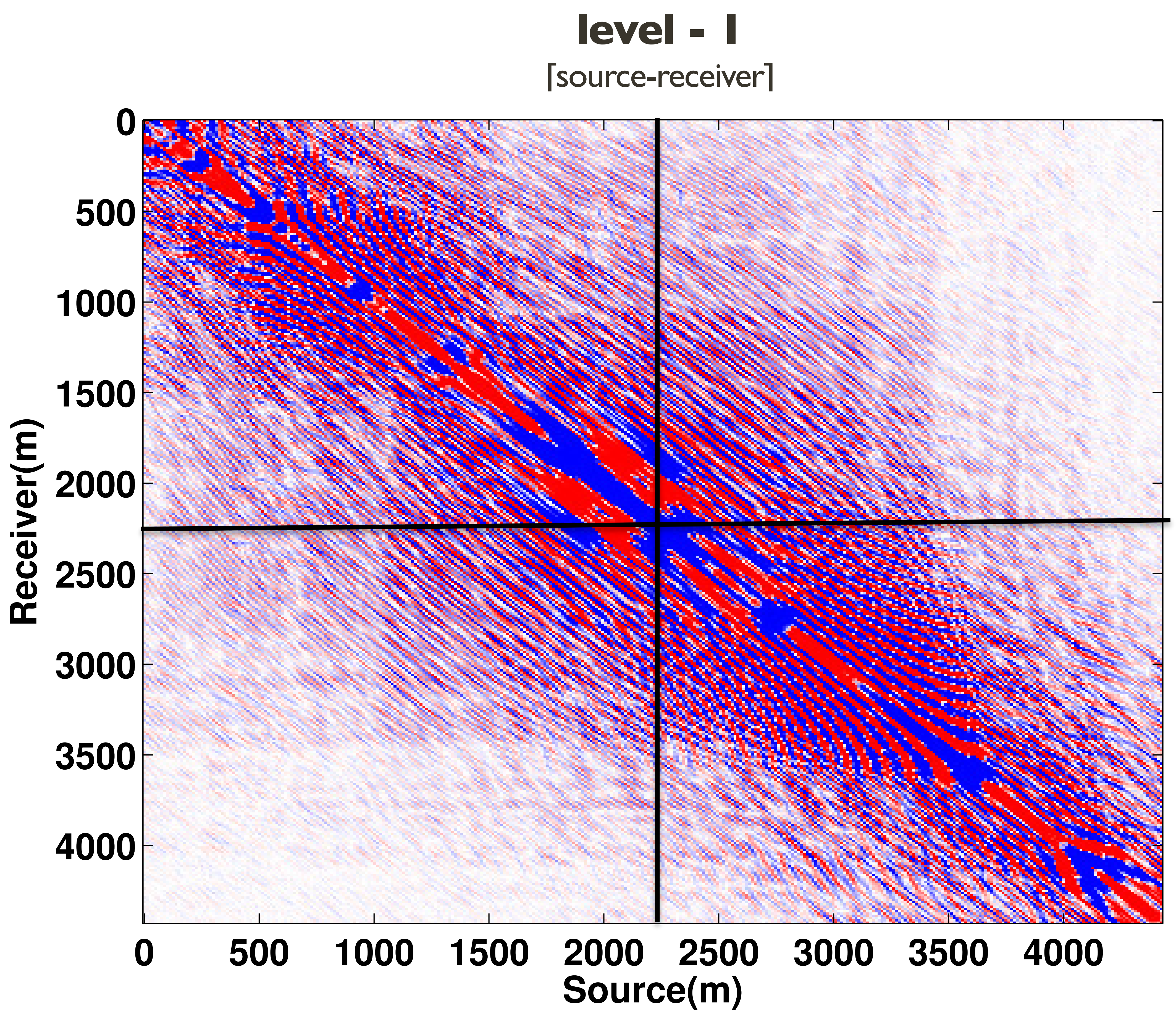


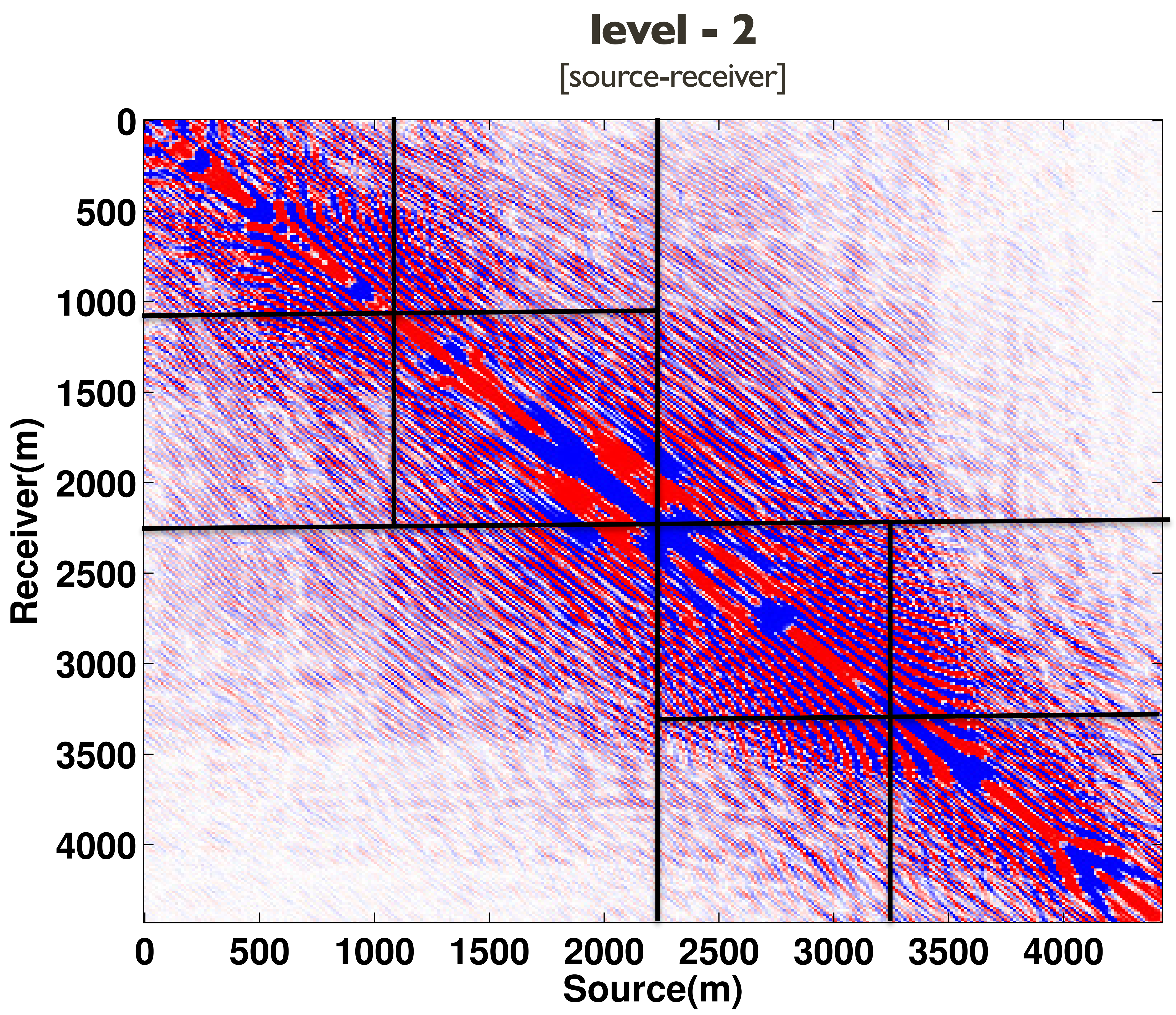
Singular value decay



[Chandrasekaran et al. (2006)]

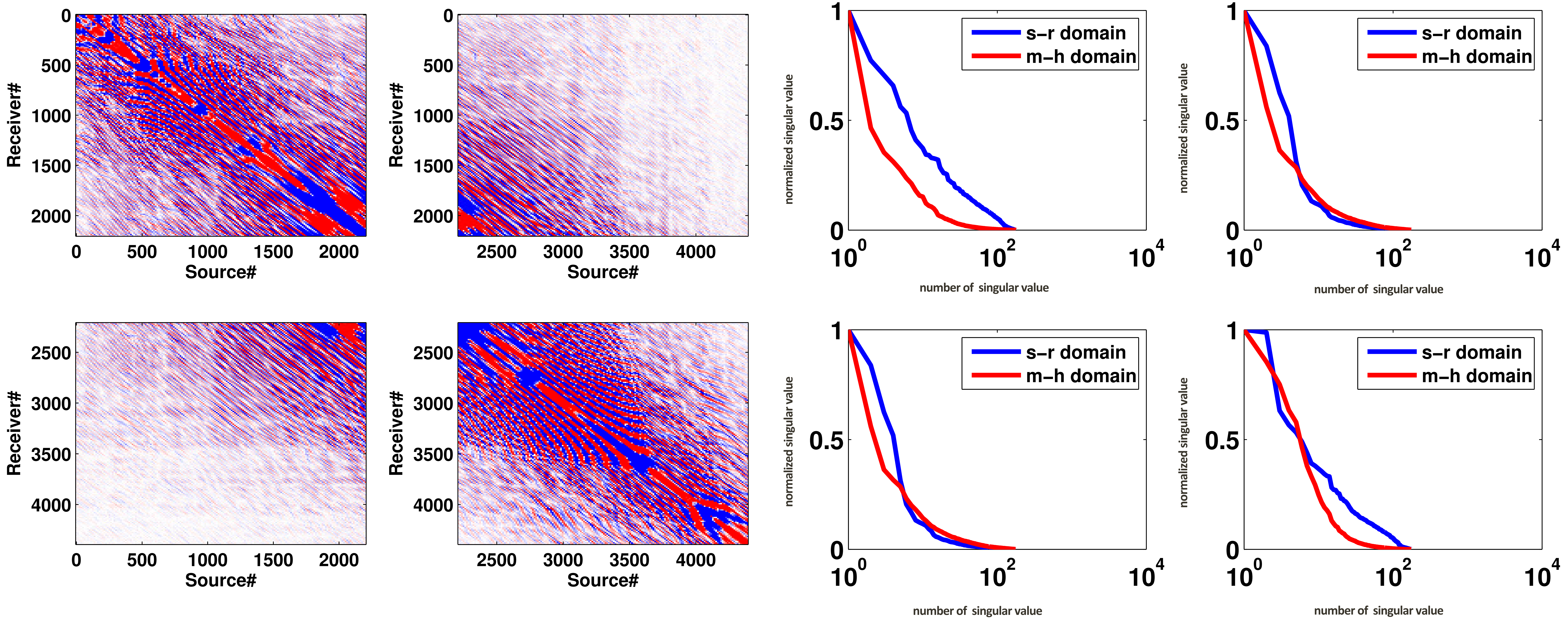
HSS representation





HSS representation

[level-1]



Matrix completion

- signal structure
  - *low rank/fast decay* of singular values

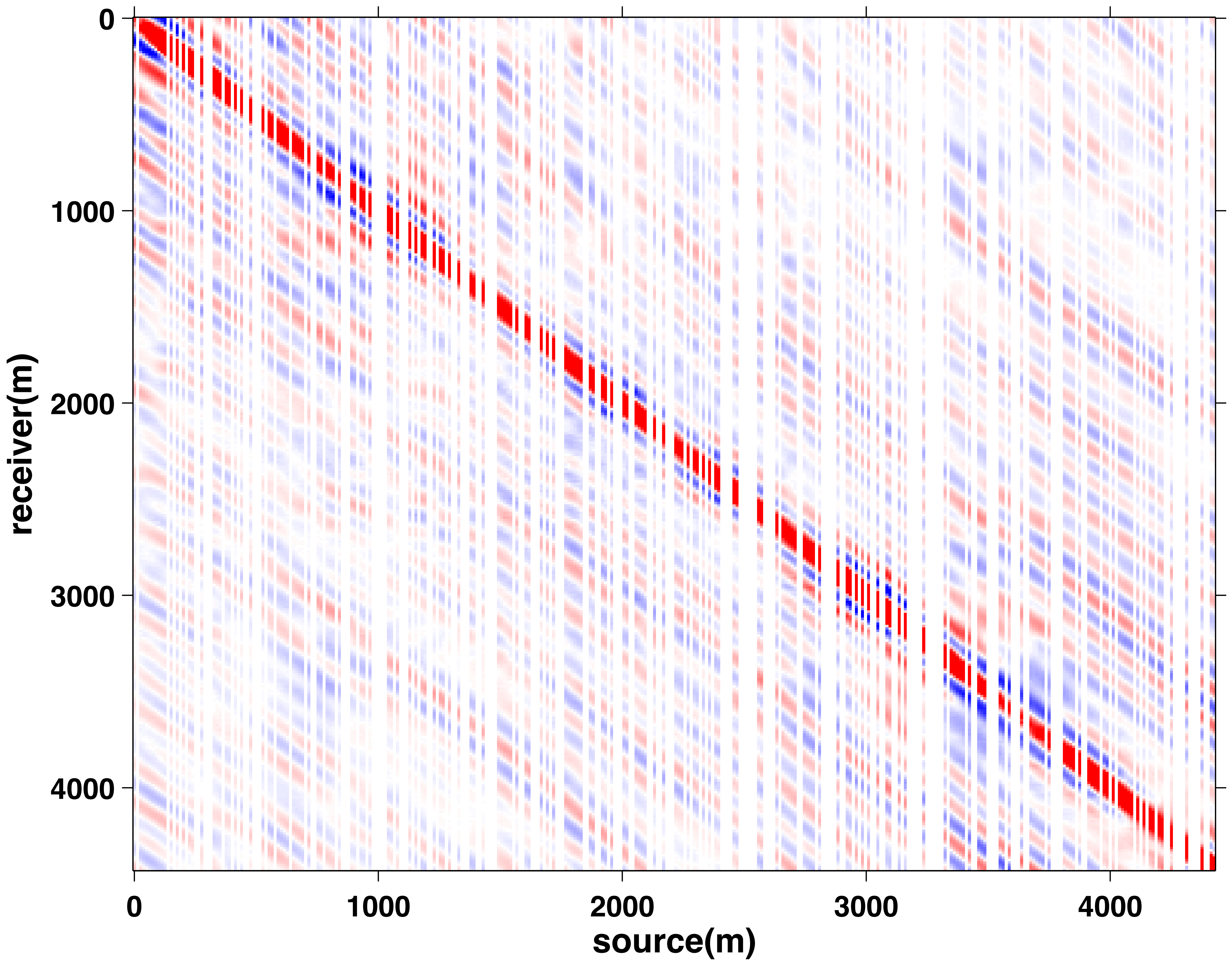
- sampling scheme
  - missing data *increase* rank in "transform domain"

- recovery using *rank penalization* scheme

2-D acquisition
[randomized sampling]

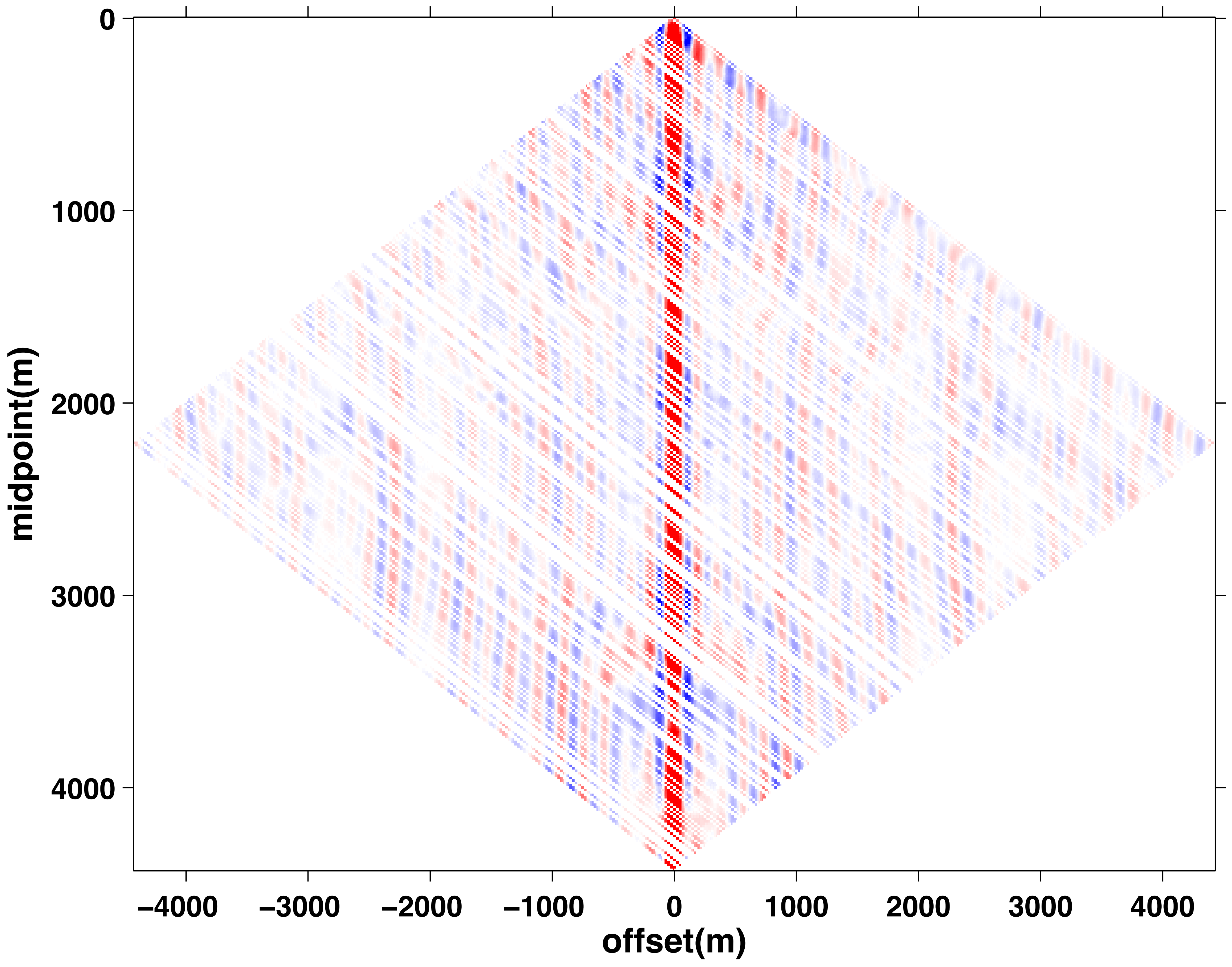
acquisition domain

missing columns *do not* increase rank



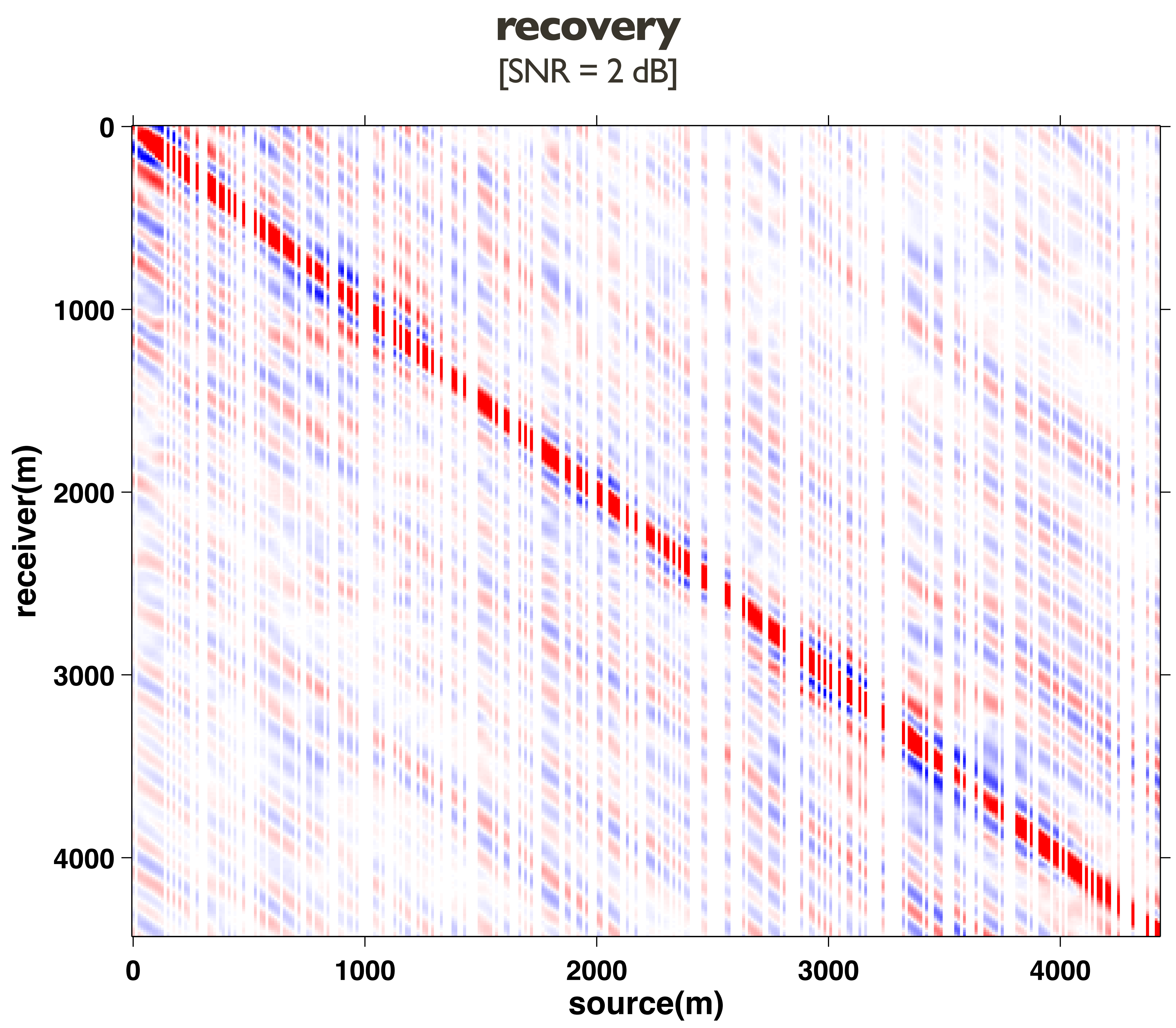
transform domain

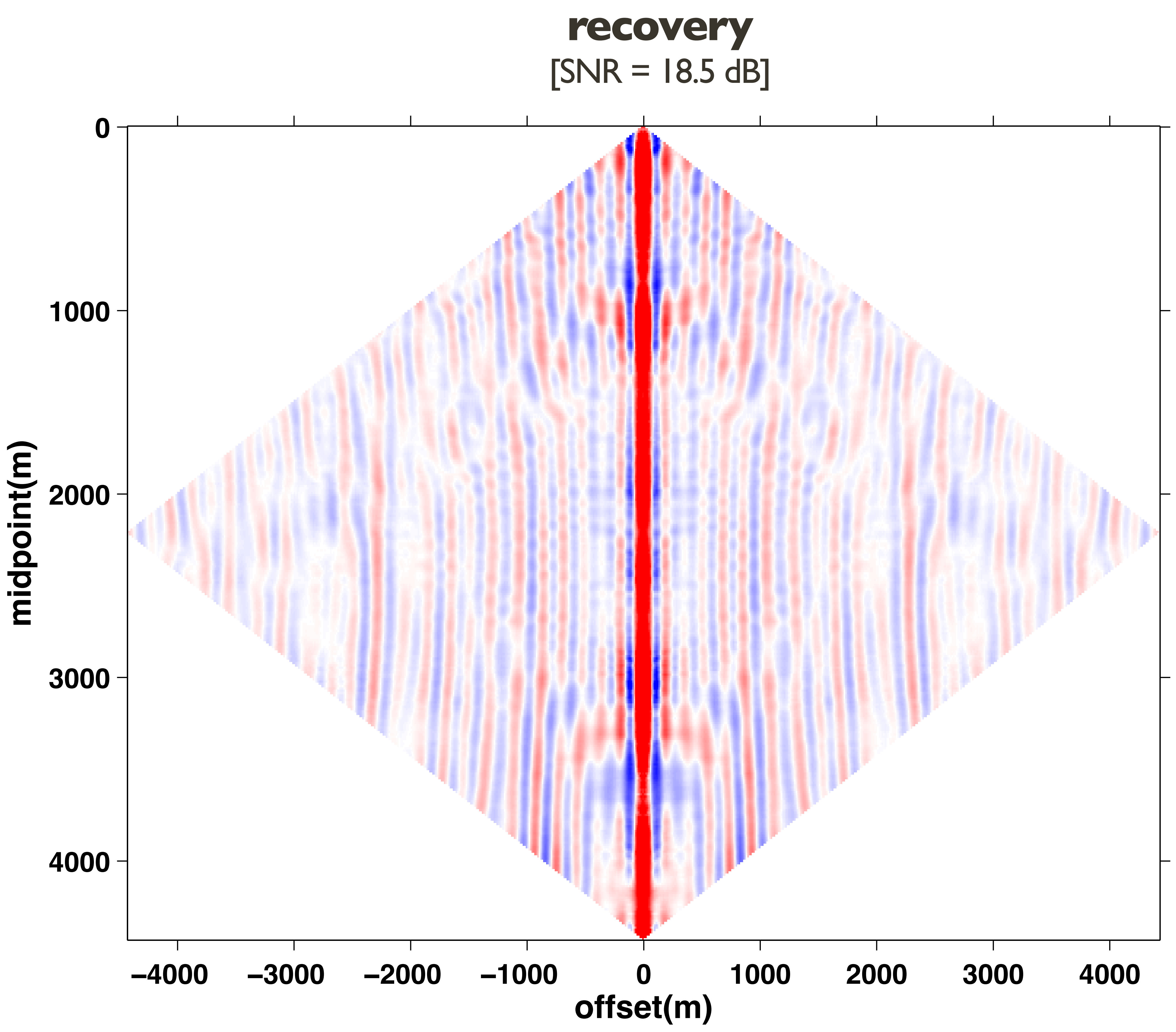
missing columns *do* increase rank



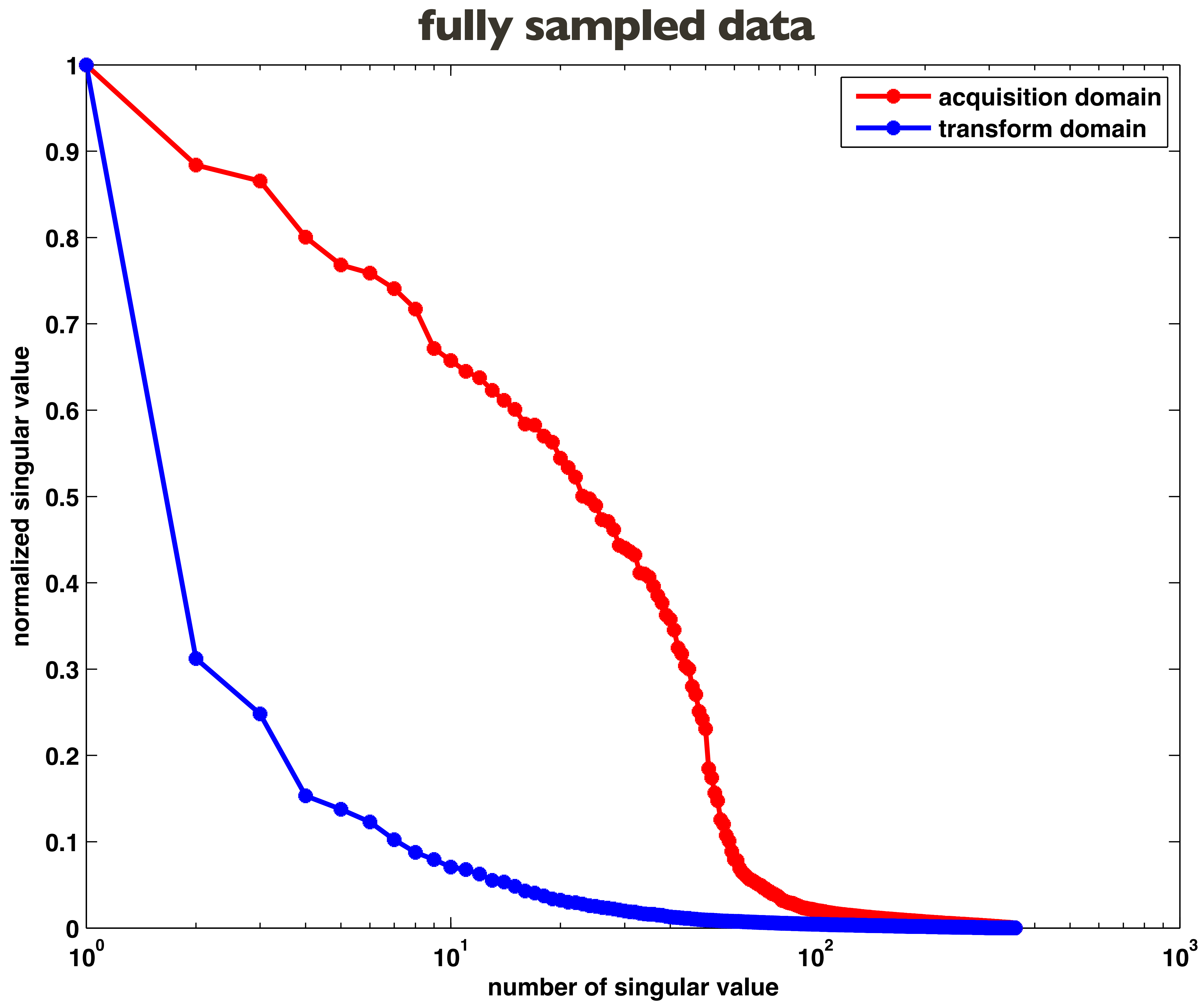
SLIM

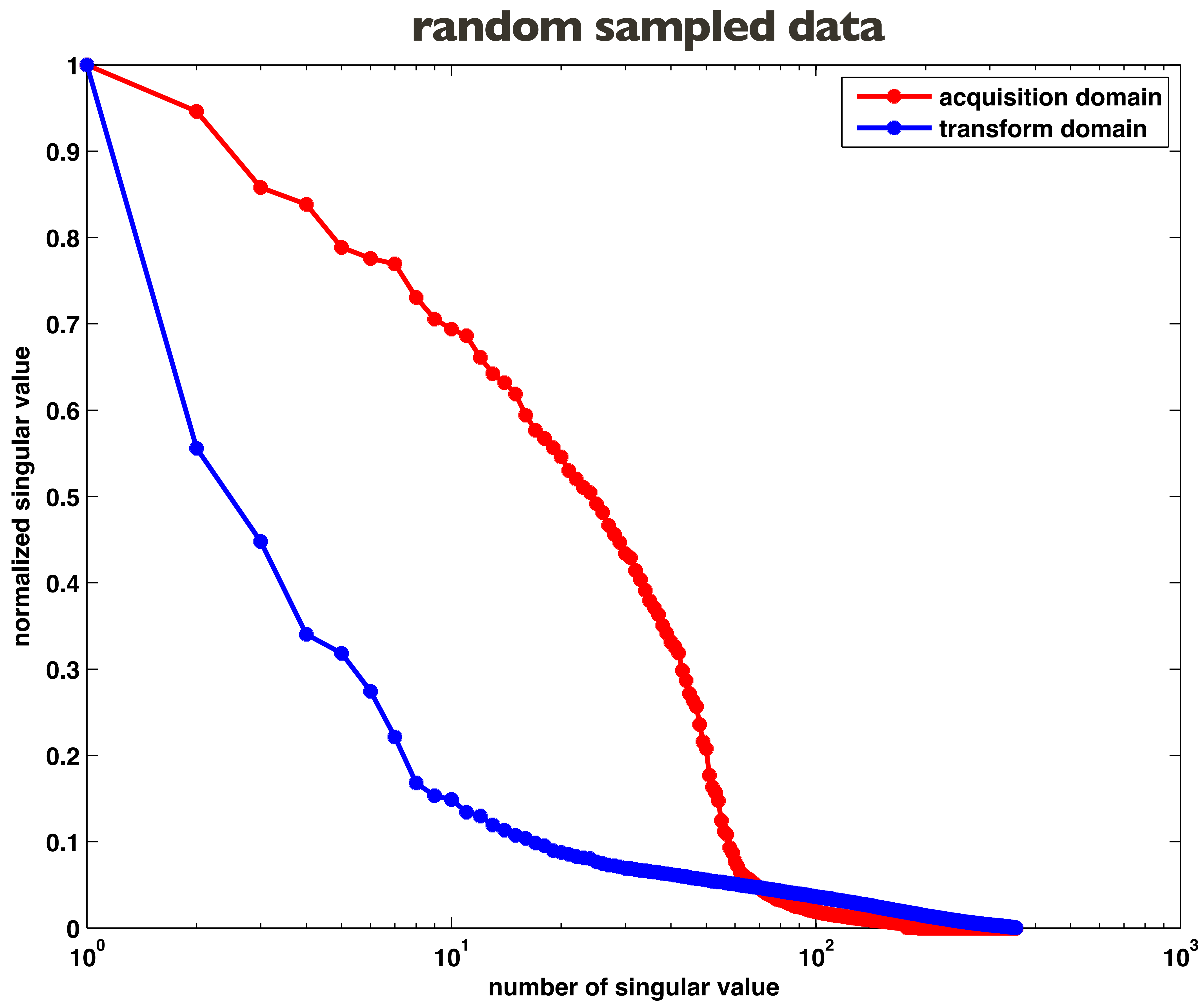
Low-rank interpolation





Randomized sampling
[singular value decay]

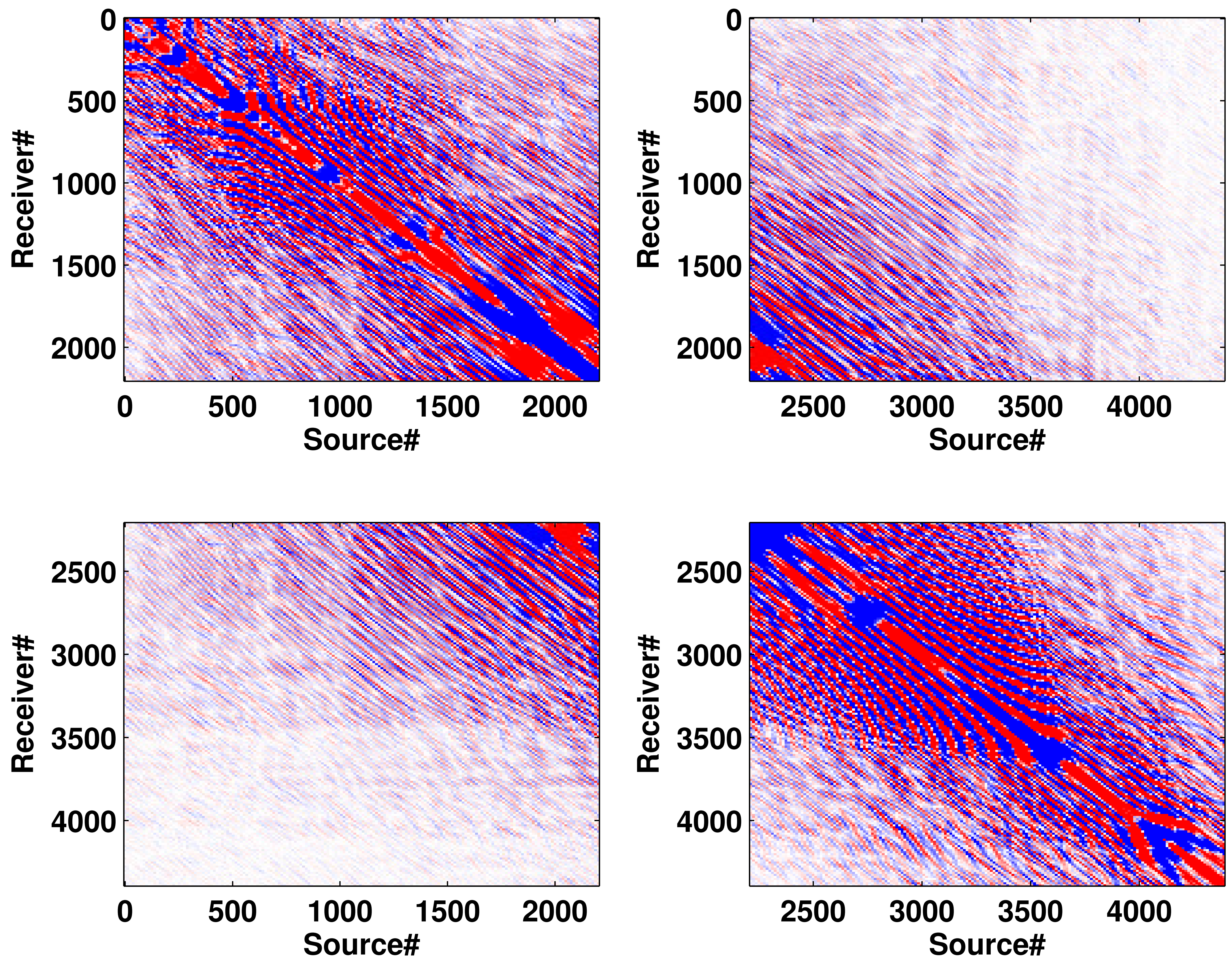




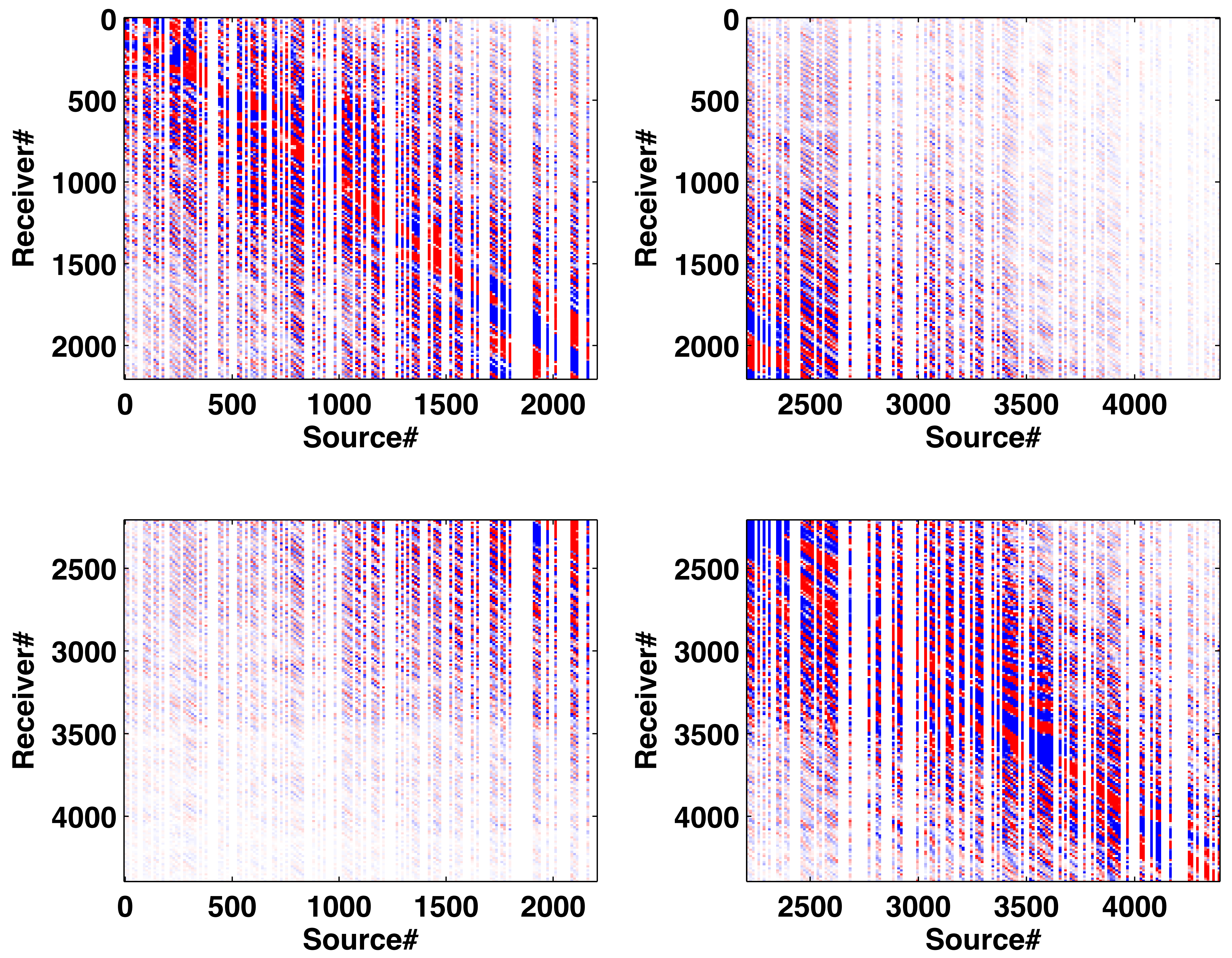
HSS representation

[level-1]

fully sampled data

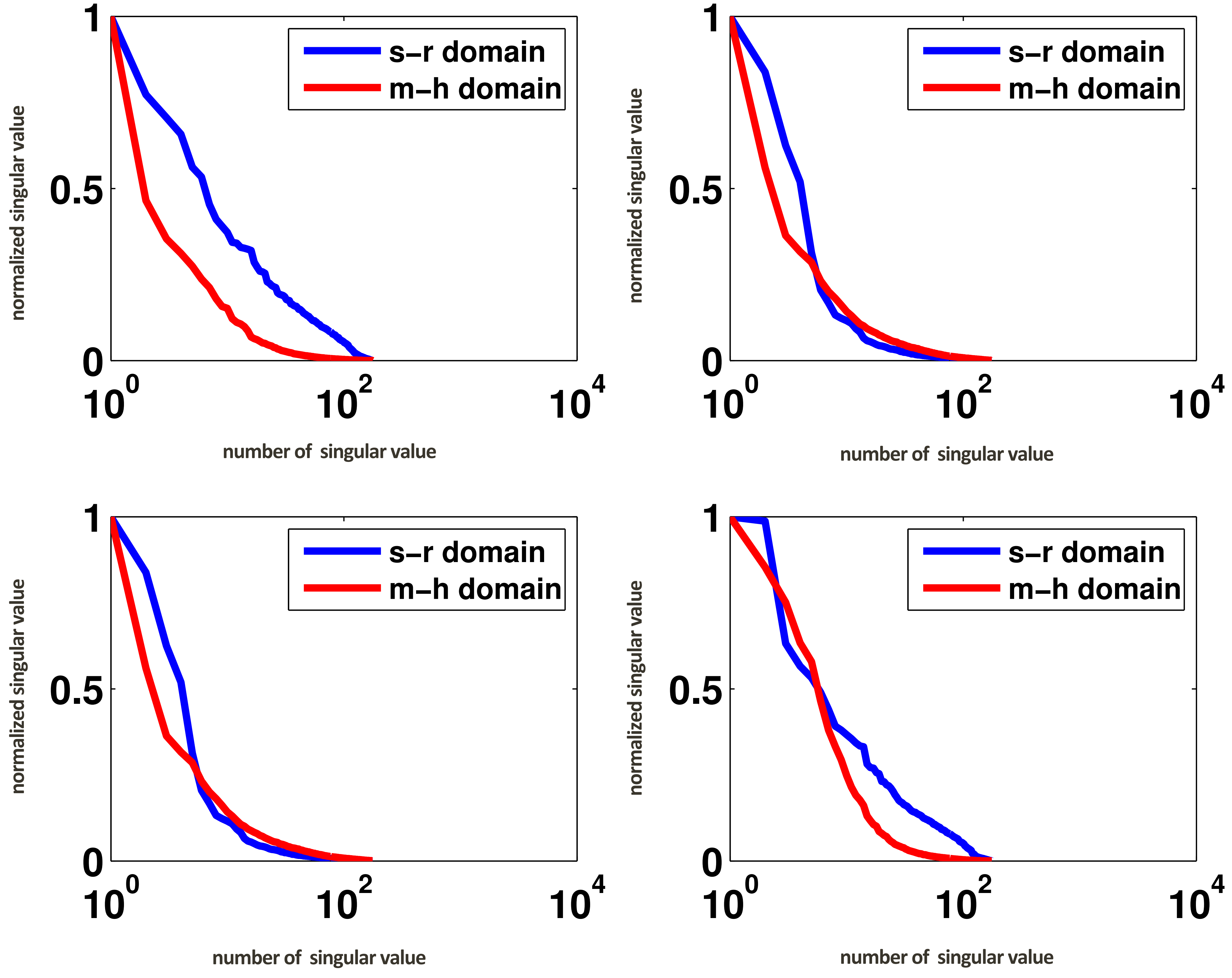


random sampled data

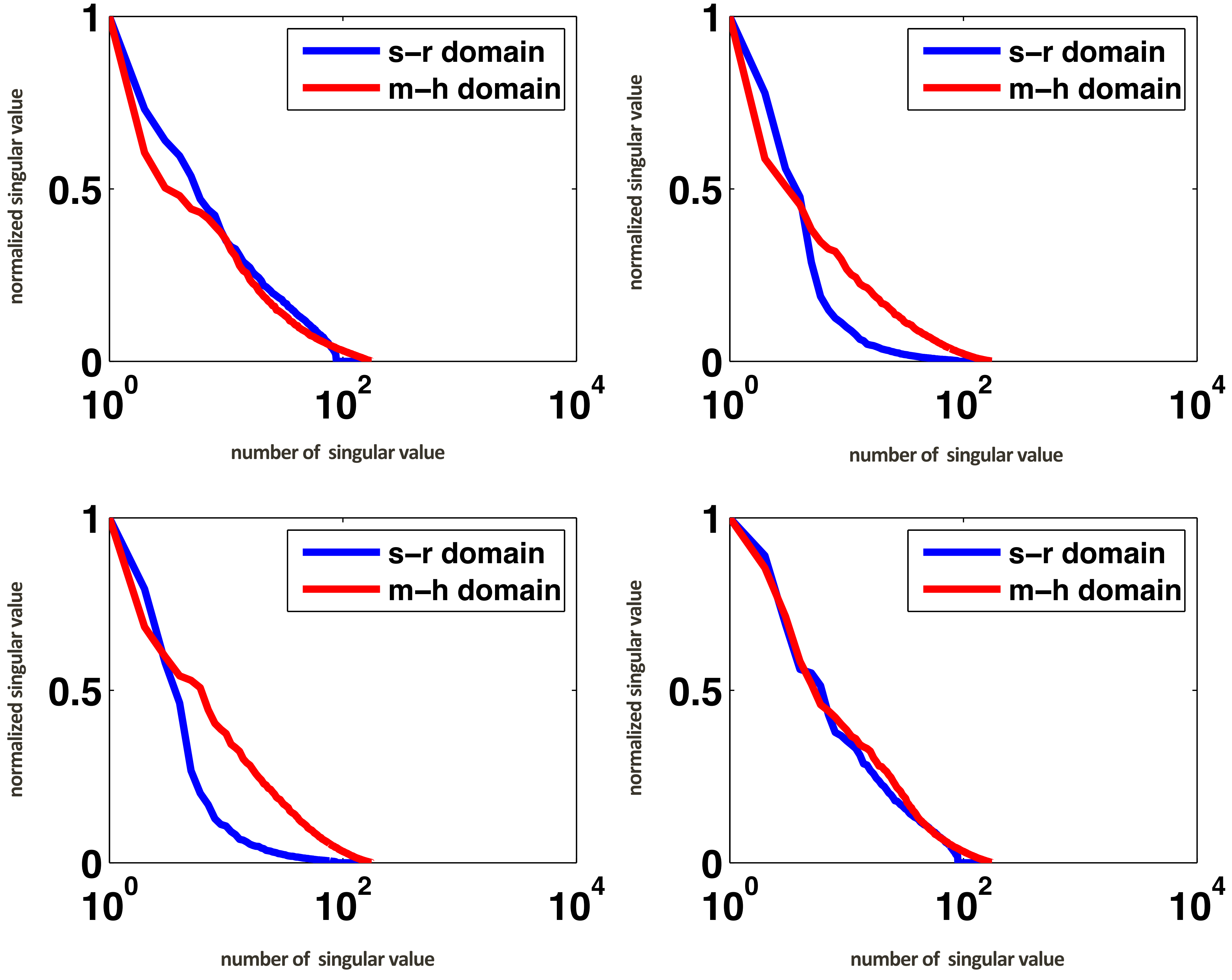


Singular value decay
[HSS, level-1]

fully sampled data



random sampled data



Matrix completion

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

Rank minimization

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

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

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

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

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

where

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

Rank minimization

- prohibitively *expensive*
  - do not know rank value in advance
  - search over all possible values of rank
- instead solve nuclear-norm minimization
  - convex relaxation of rank-minimization [Recht et. al. 2010]

[Recht et. al. 2010]

SLIM

Nuclear-norm minimization

- we want to solve

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

where

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

where λ_i are the *singular* values

Challenges

- requires repeated application of *SVD* for projections

- expensive to compute for large system
  - curse of dimensionality

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

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

Factorized formulation

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

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

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

Factorized formulation

- reformulate $(BPDN_\sigma)$ formulation

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

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

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

where τ is a rank regularization parameter

[Rennie and Srebro 2005]

Factorized formulation

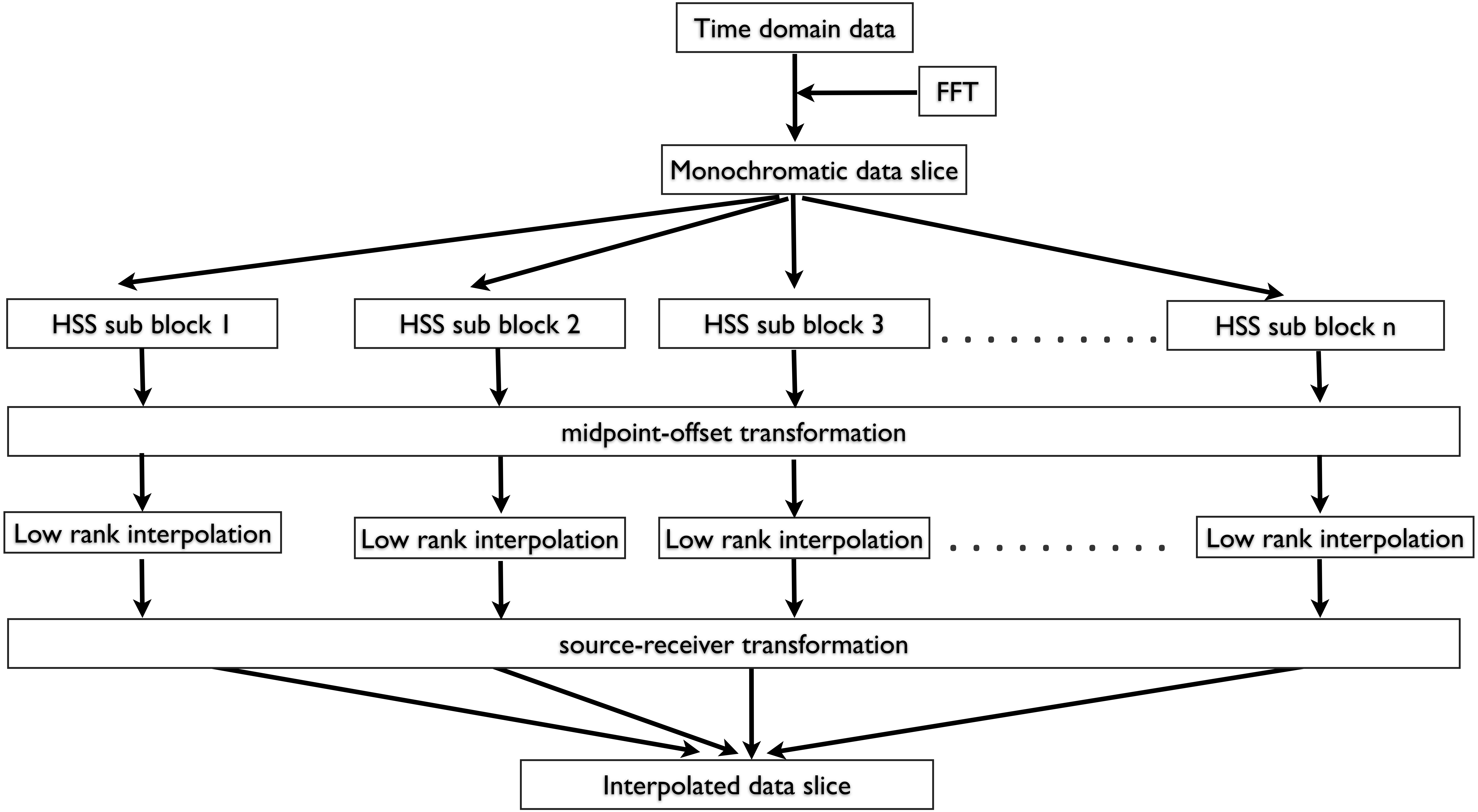
- Upper-bound on nuclear norm is defined as

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

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

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

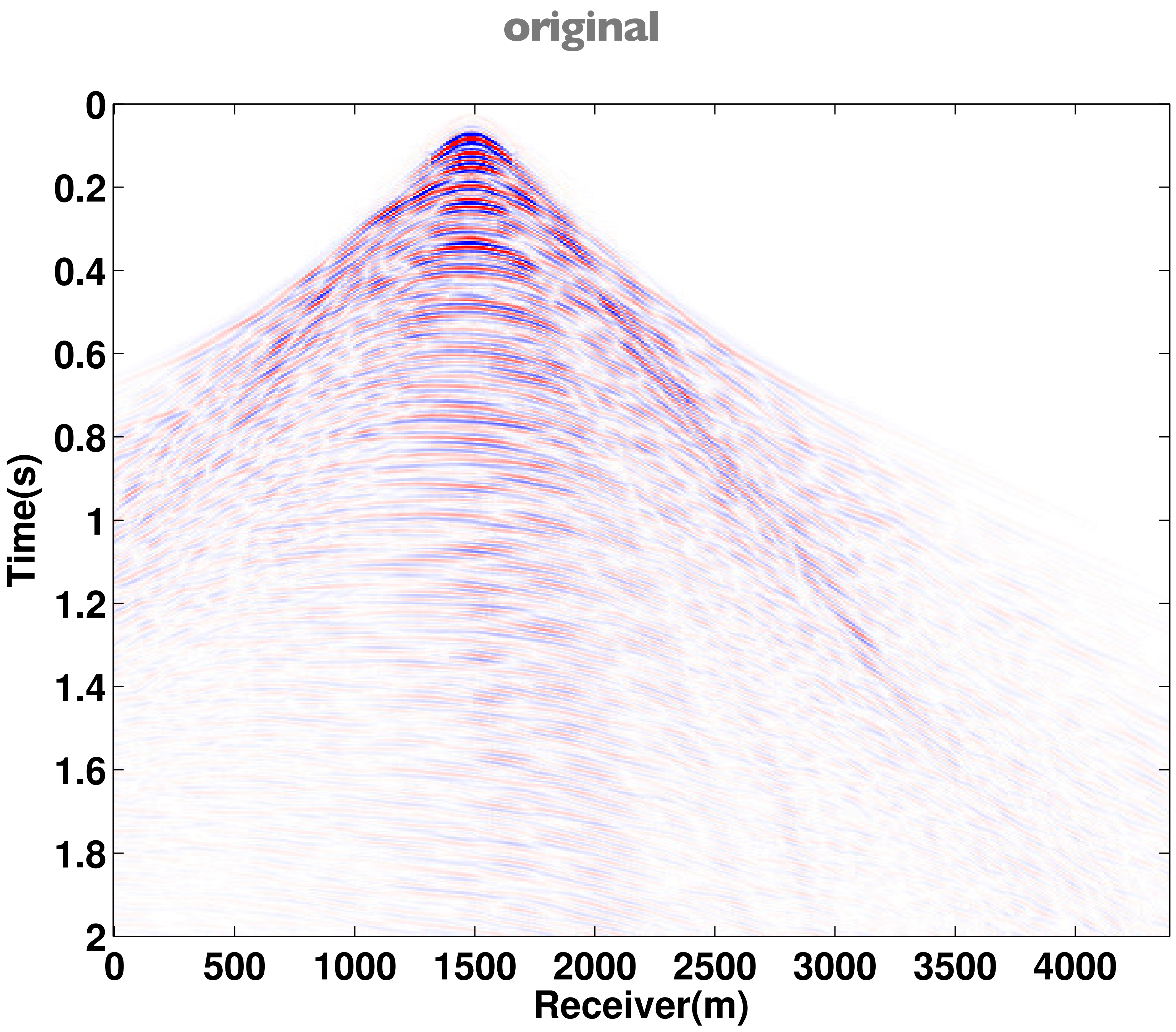
Interpolation flow chart

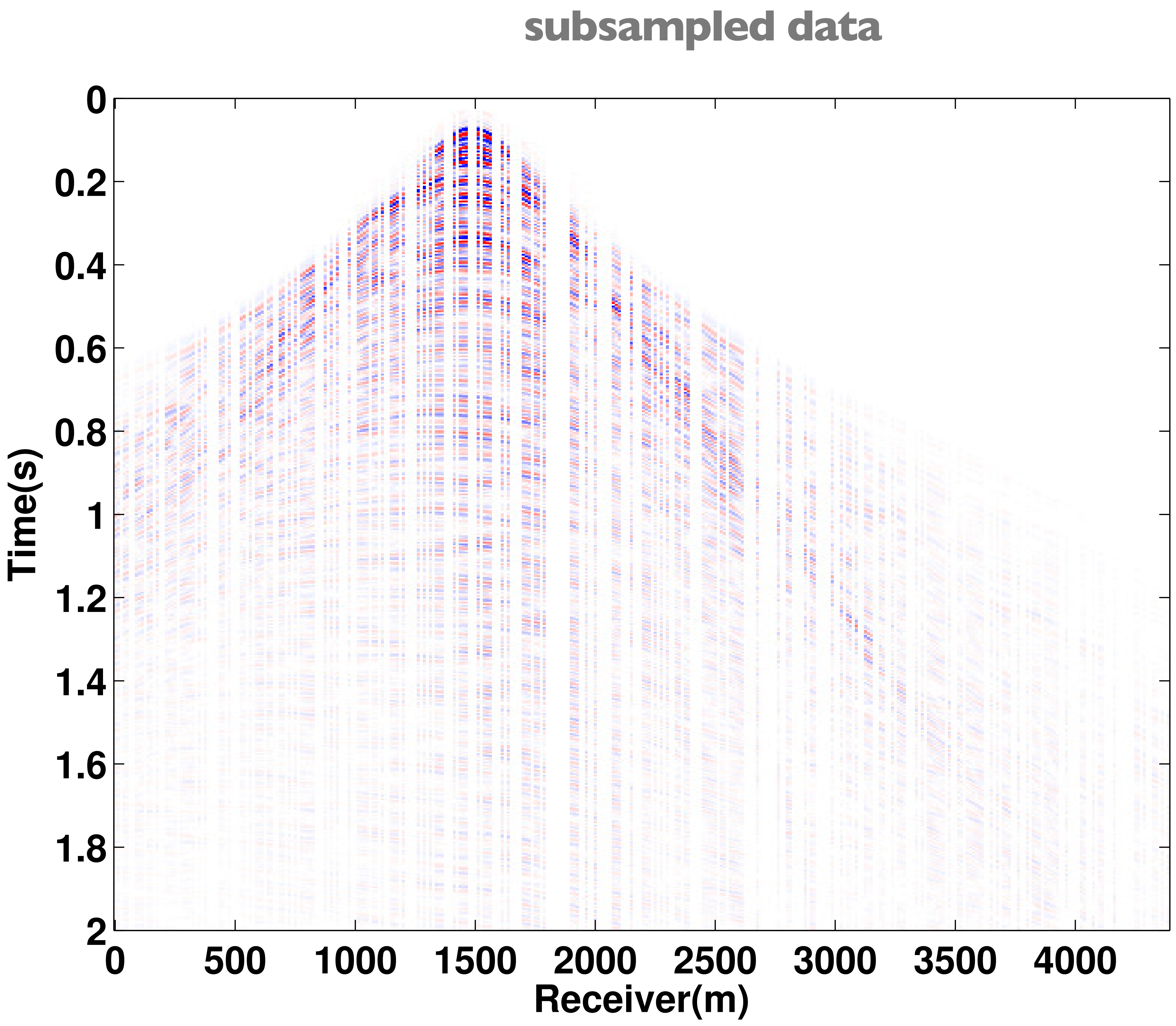


Experiments and Results

- Gulf of Suez
  - 2-D seismic line
  - 50 % missing traces
  - interpolation with HSS level 1,2,3
  - rank = 5
  - 300 iterations

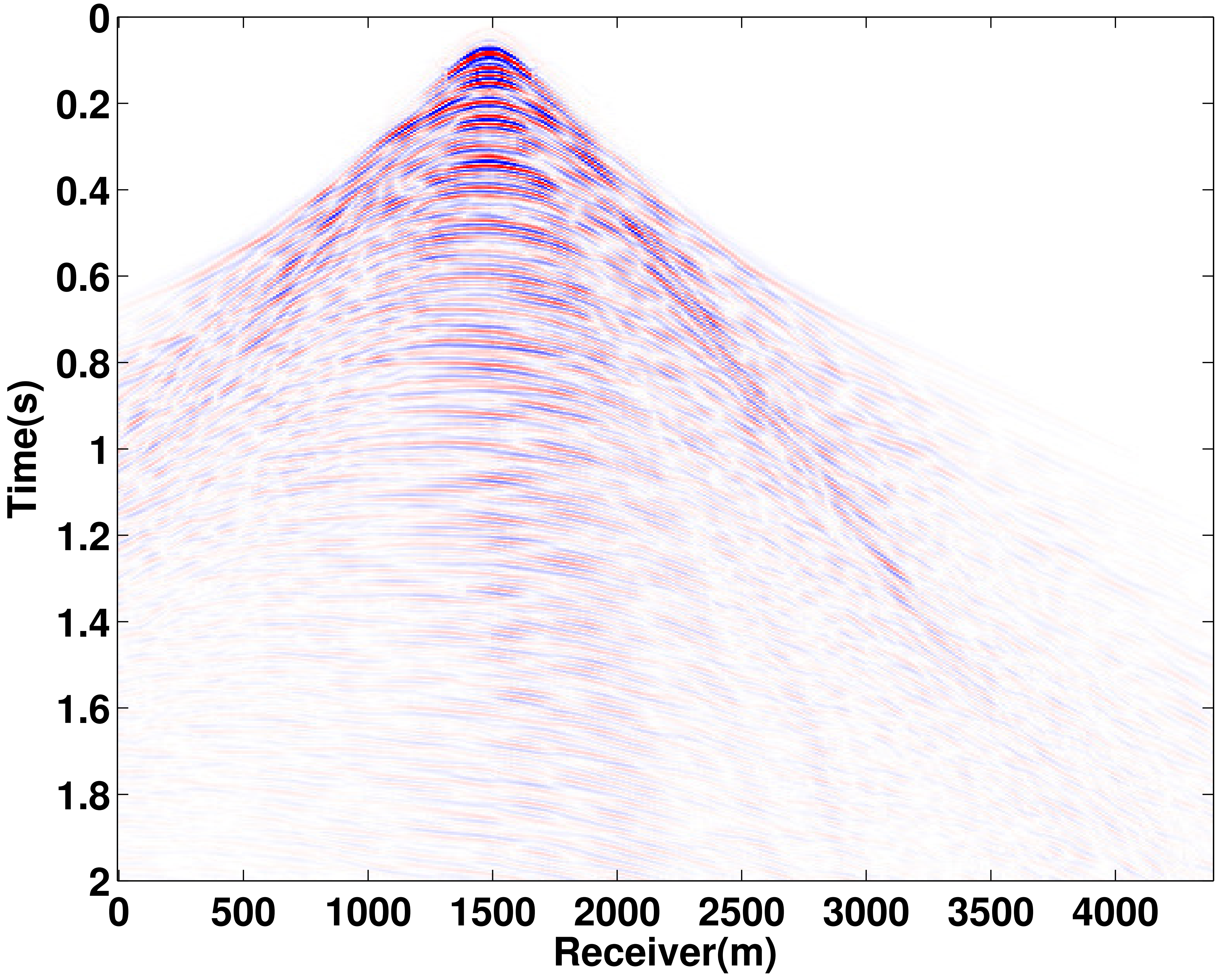
Gulf of Suez


original
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
Time(s)
0
500
1000
1500
2000
2500
3000
3500
4000
Receiver(m)

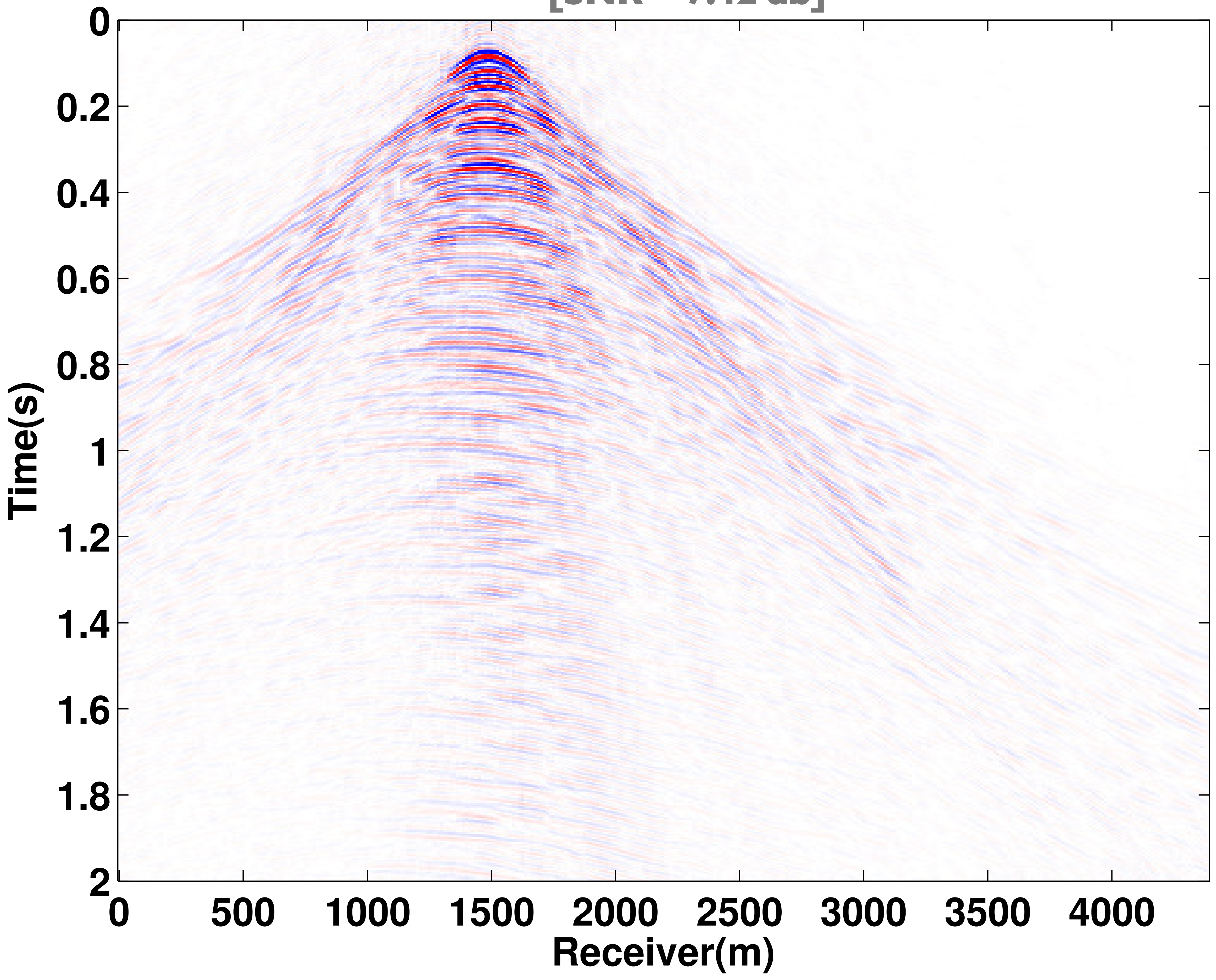

subsampled data
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
Time(s)
0
500
1000
1500
2000
2500
3000
3500
4000
Receiver(m)

Gulf of Suez
[w/o HSS]

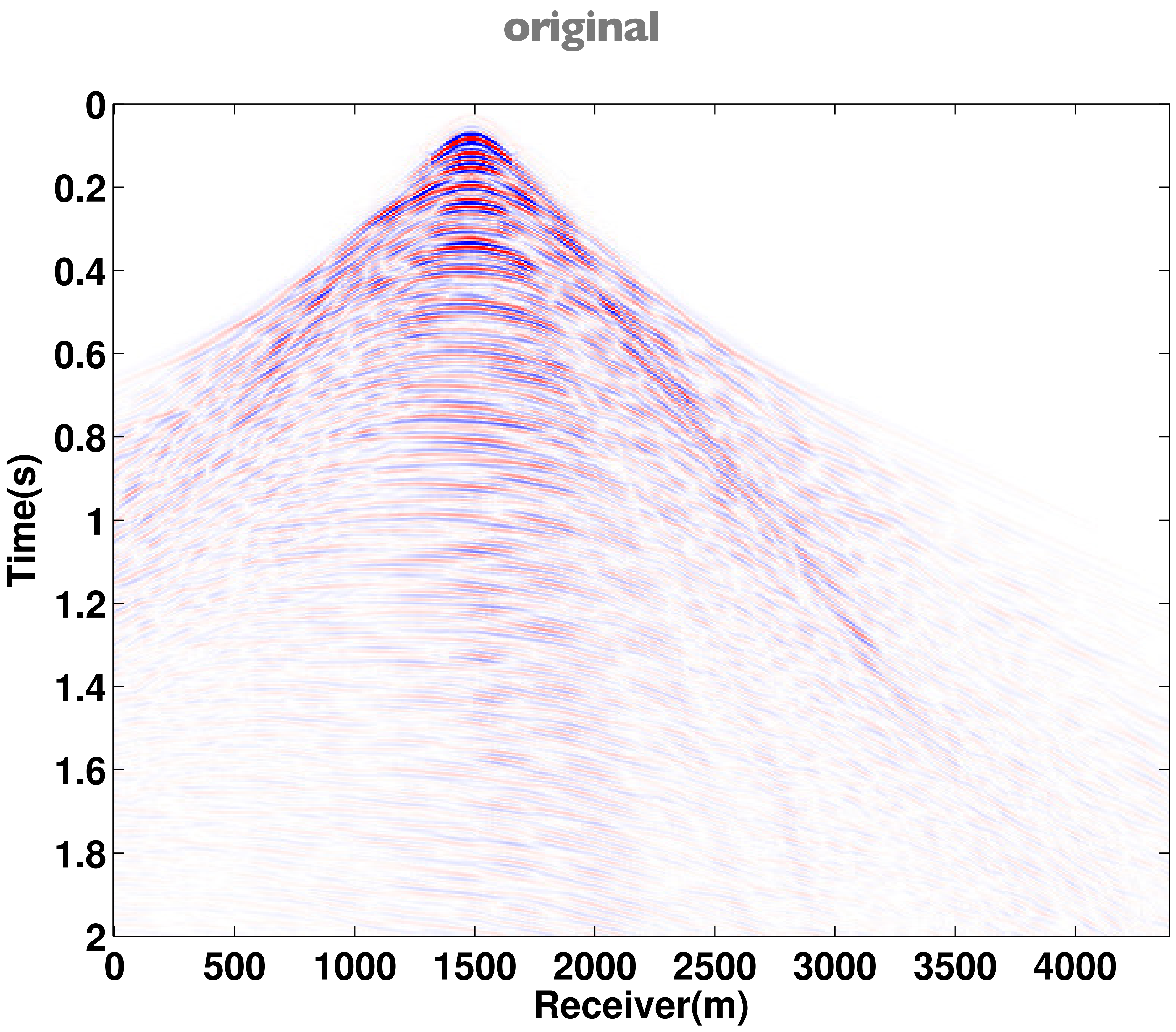
original

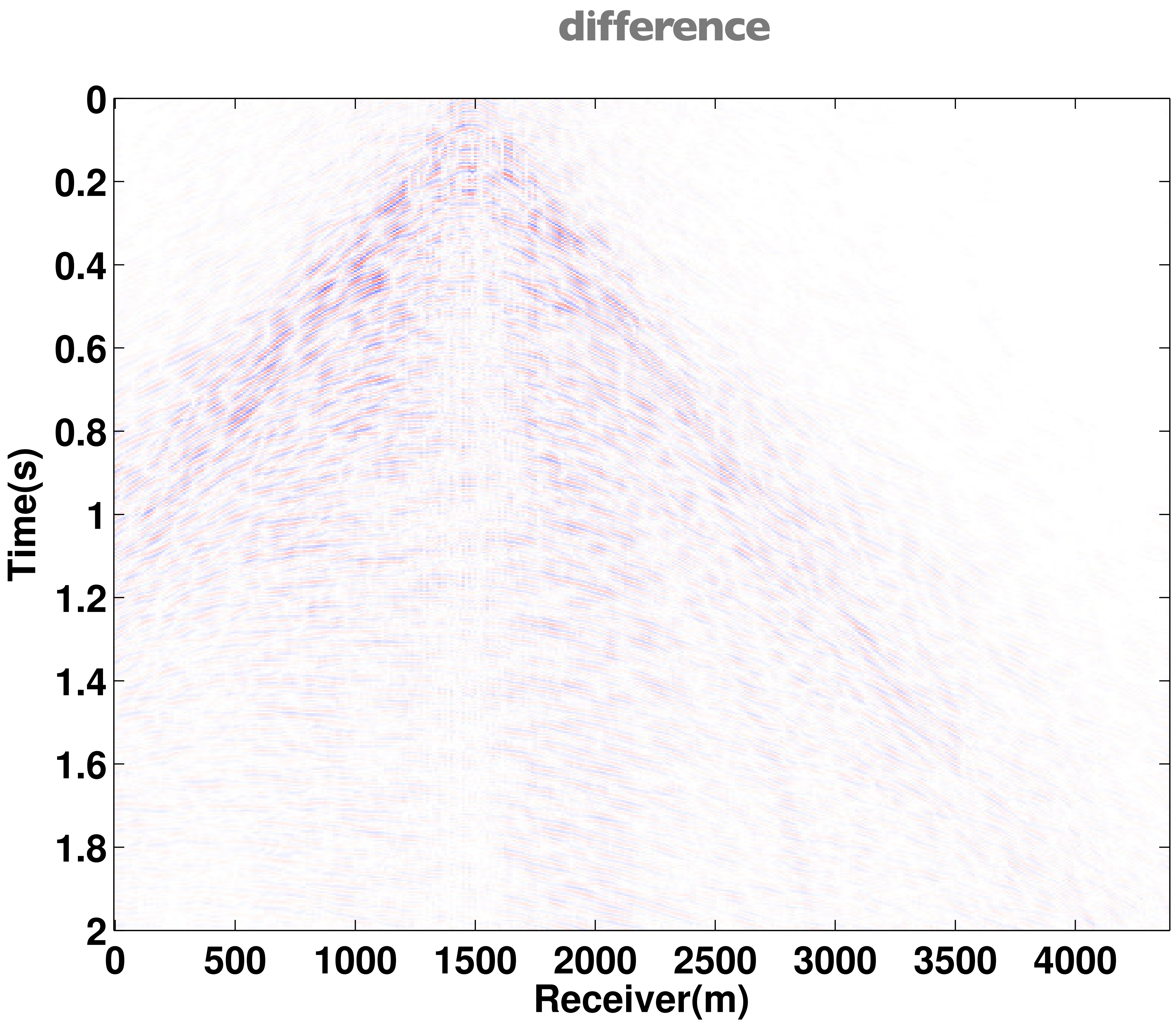


interpolation
[SNR = 7.42 db]



Gulf of Suez
[w/o HSS]

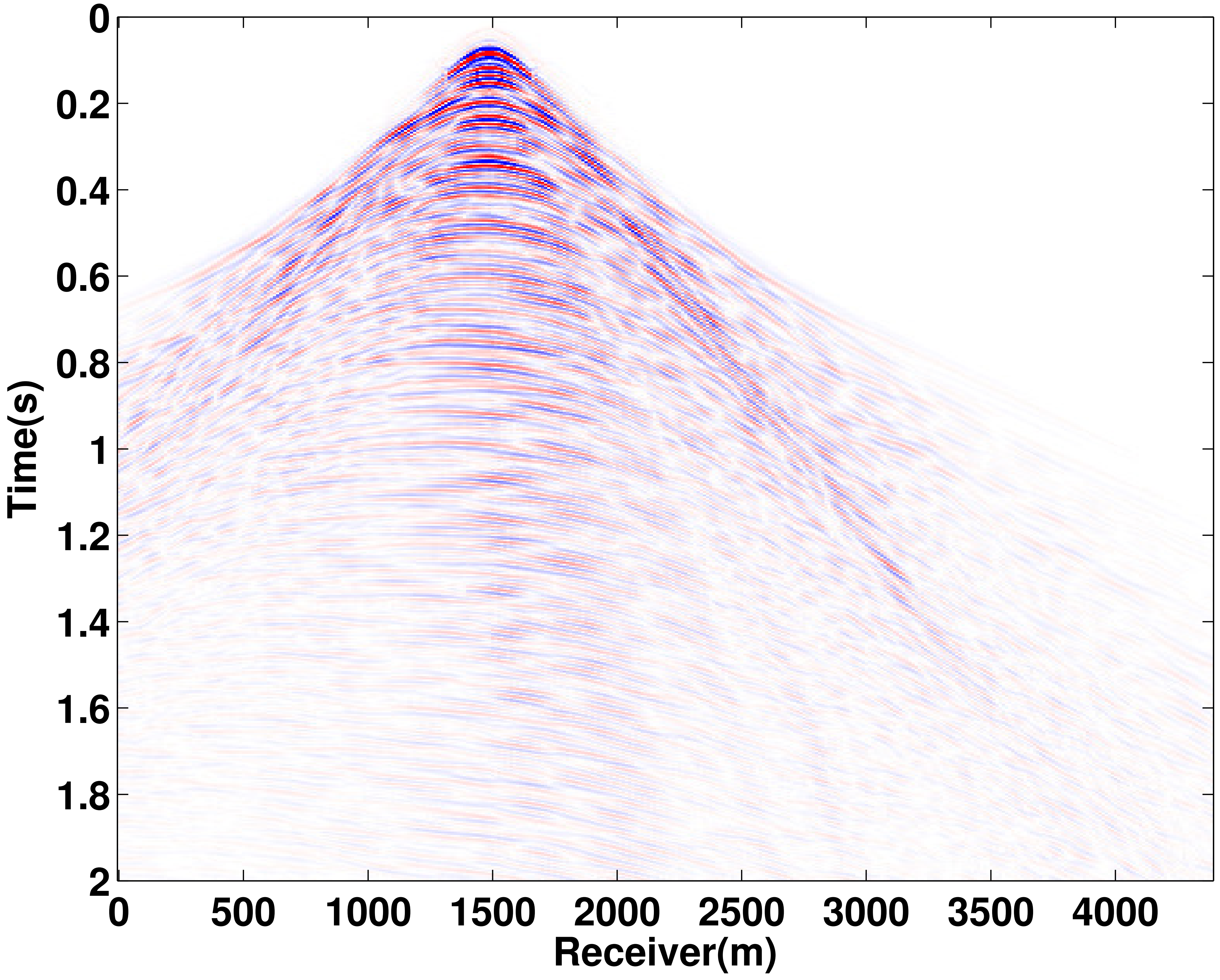




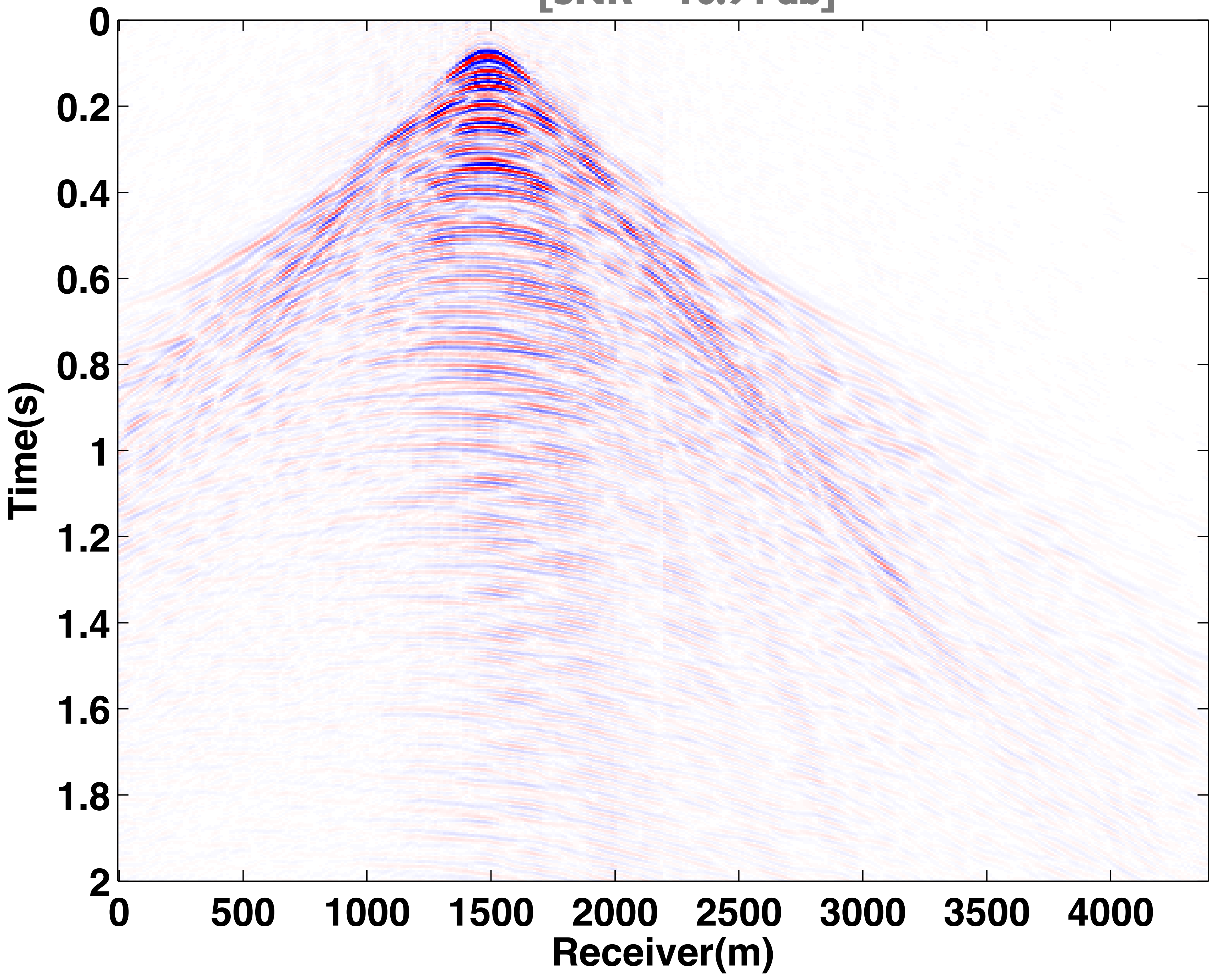
Gulf of Suez

[w/ HSS, level-1]

original

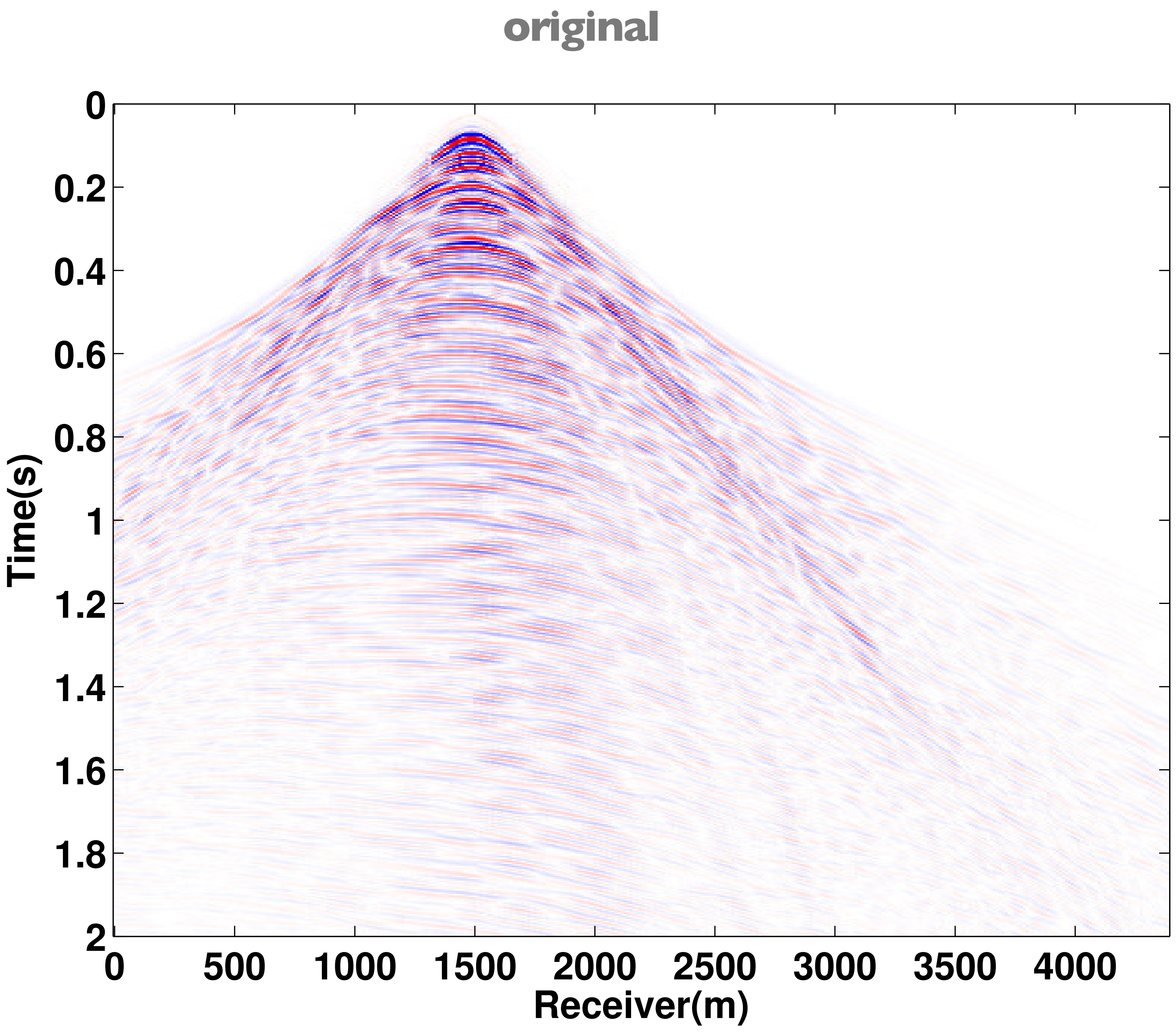


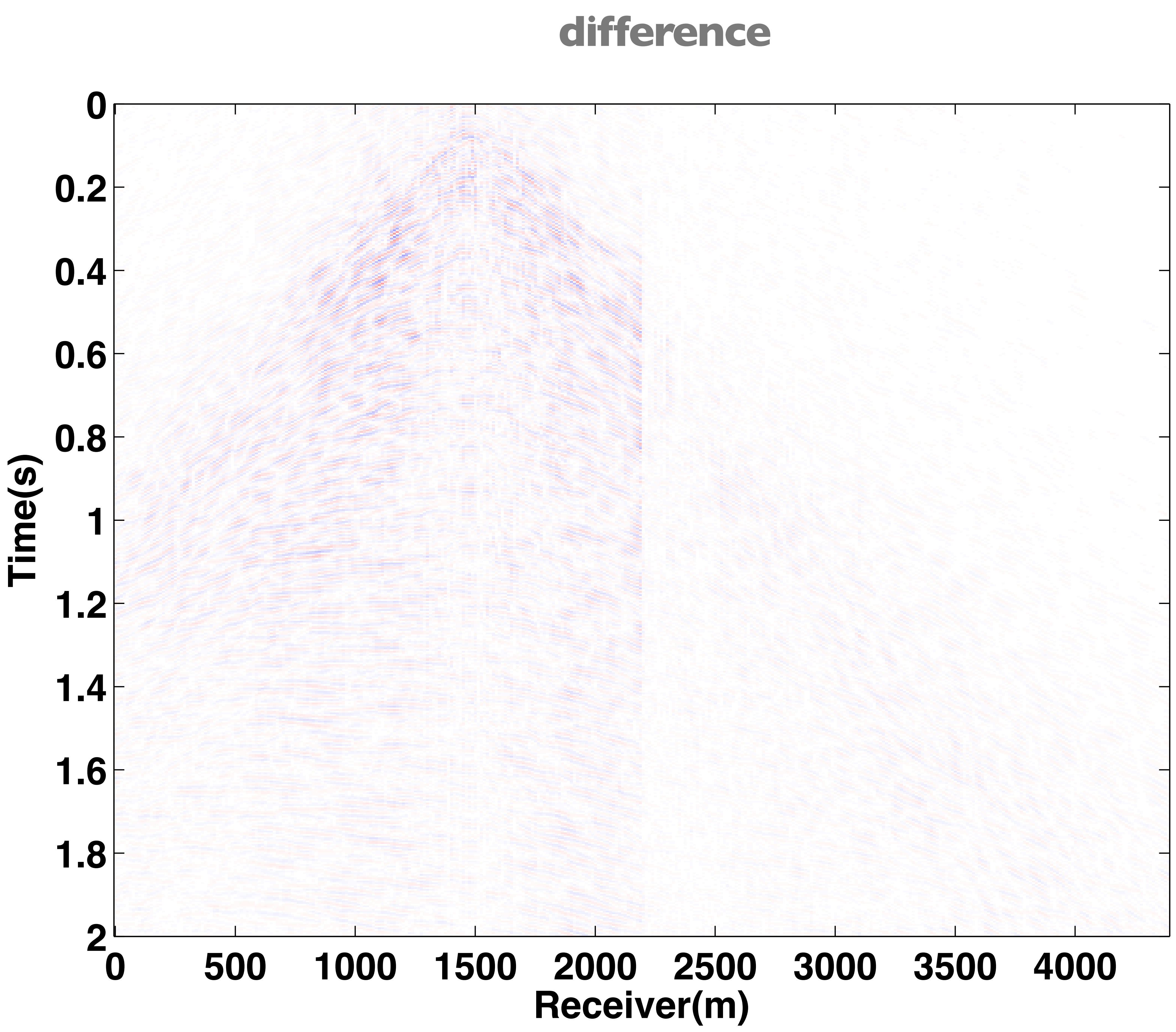
interpolation
[SNR = 10.94 db]



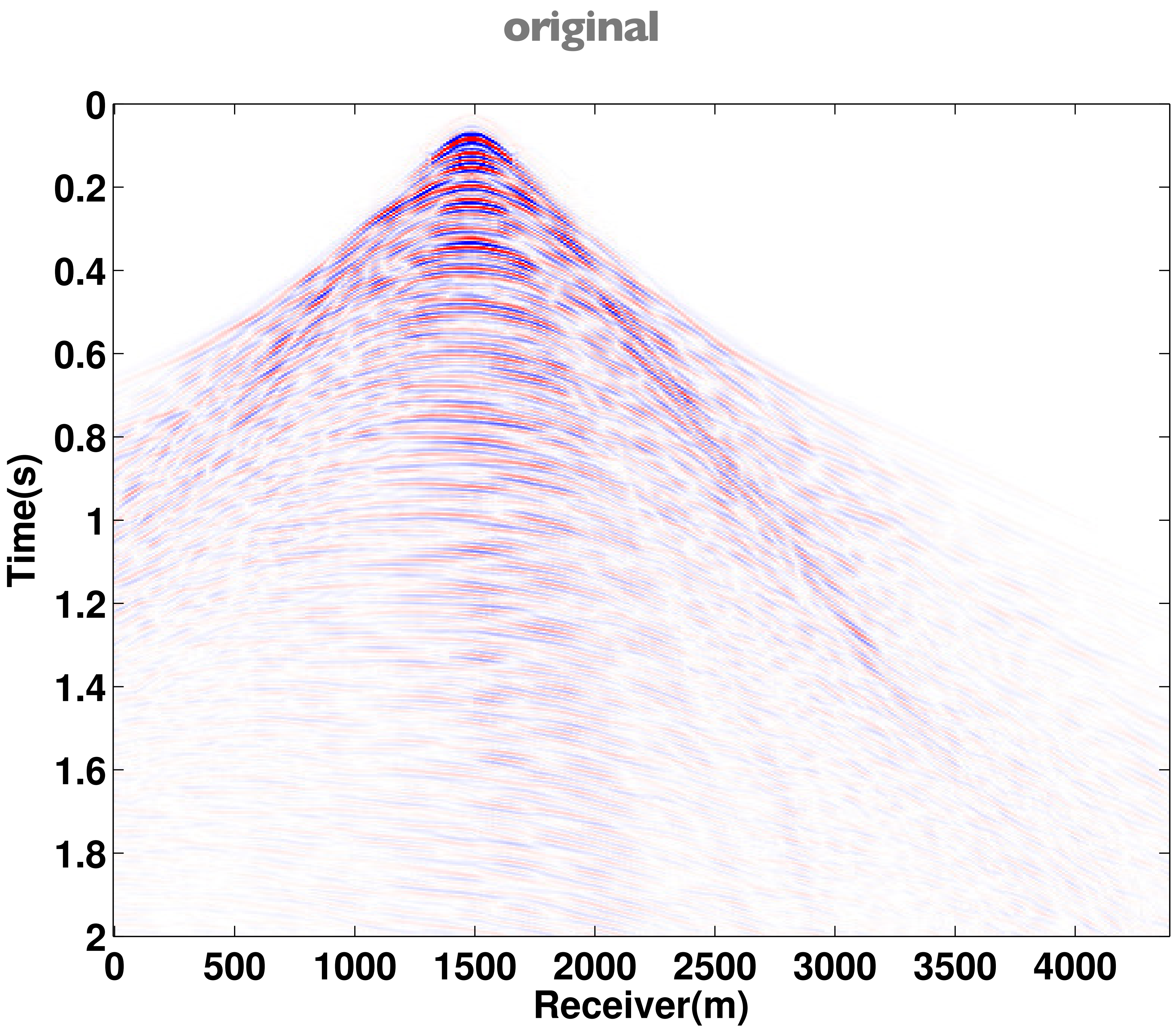
Gulf of Suez

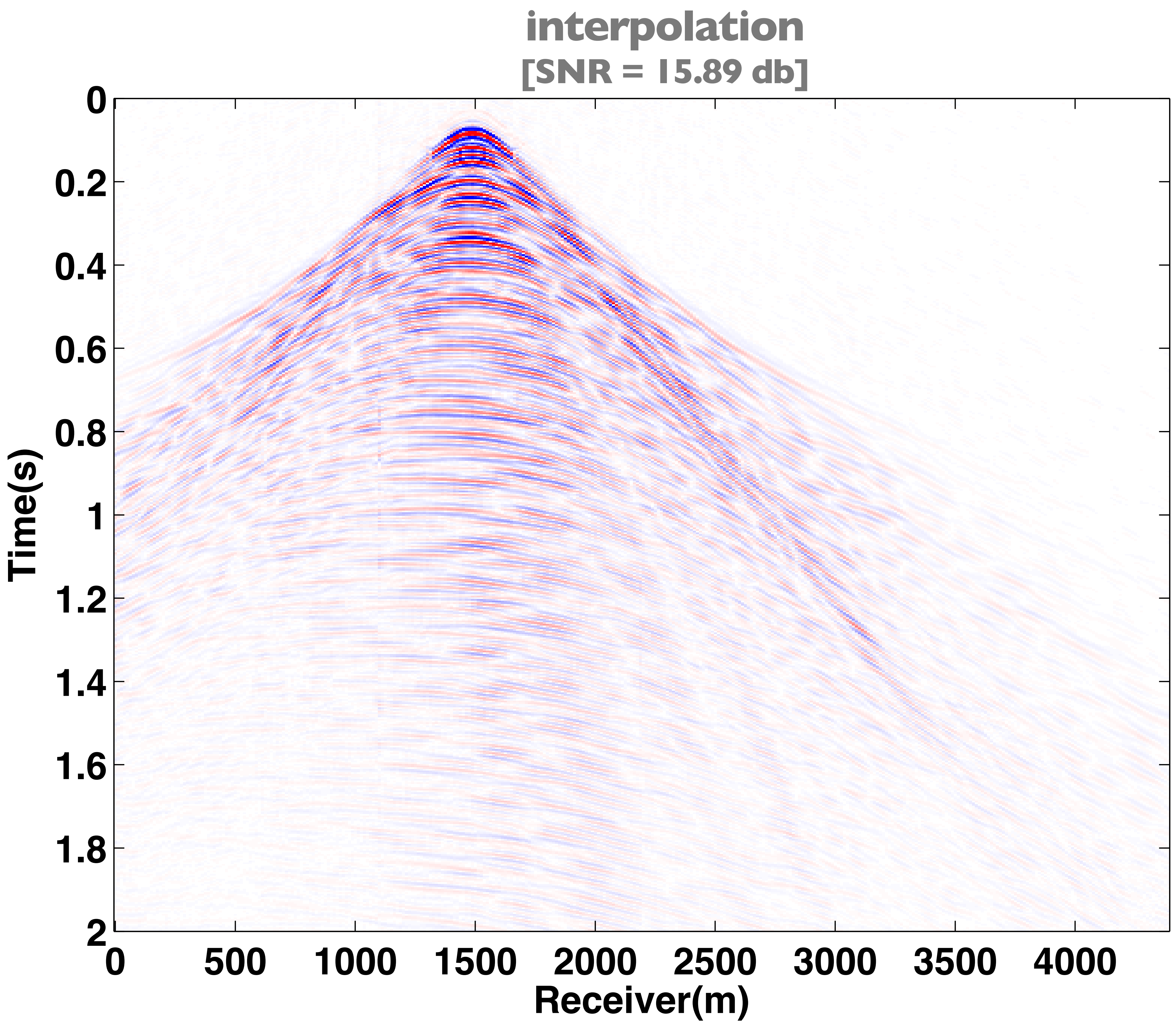
[w/ HSS, level-1]



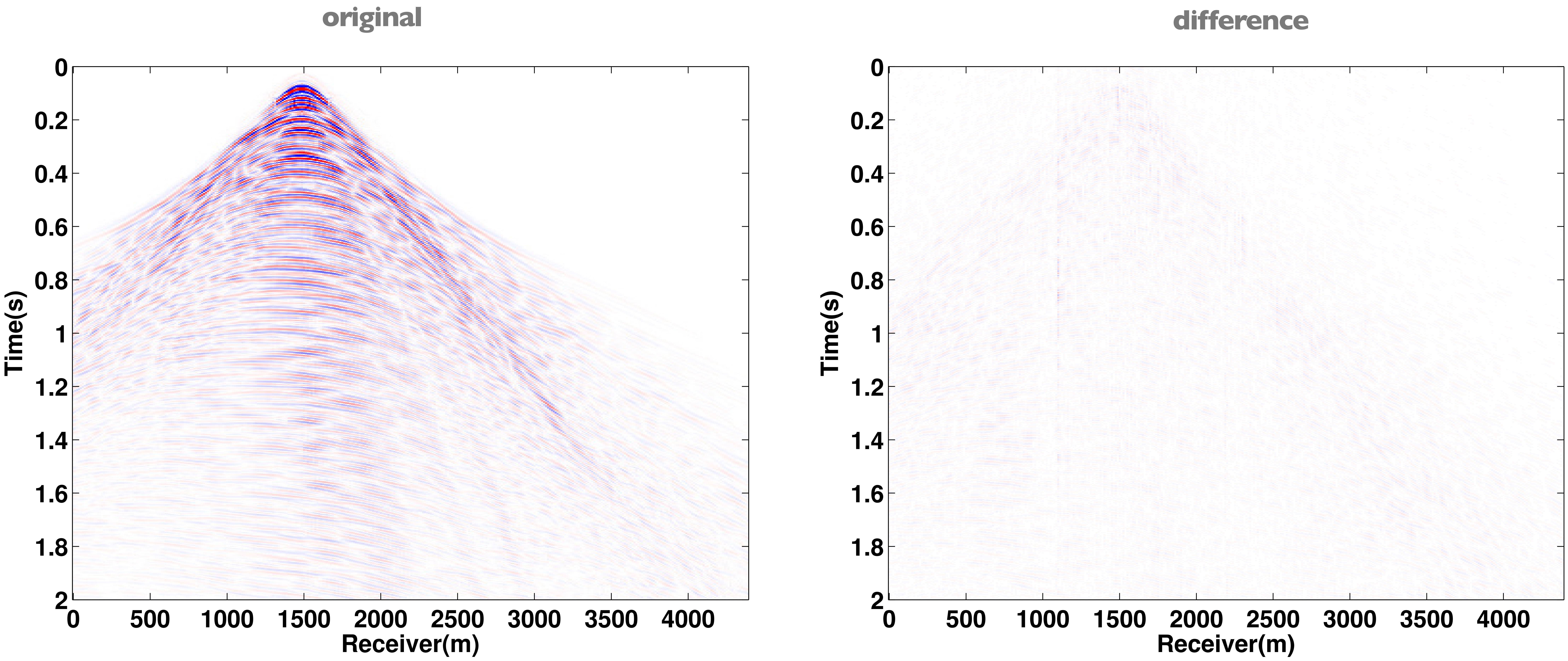


Gulf of Suez
[w/ HSS, level-2]



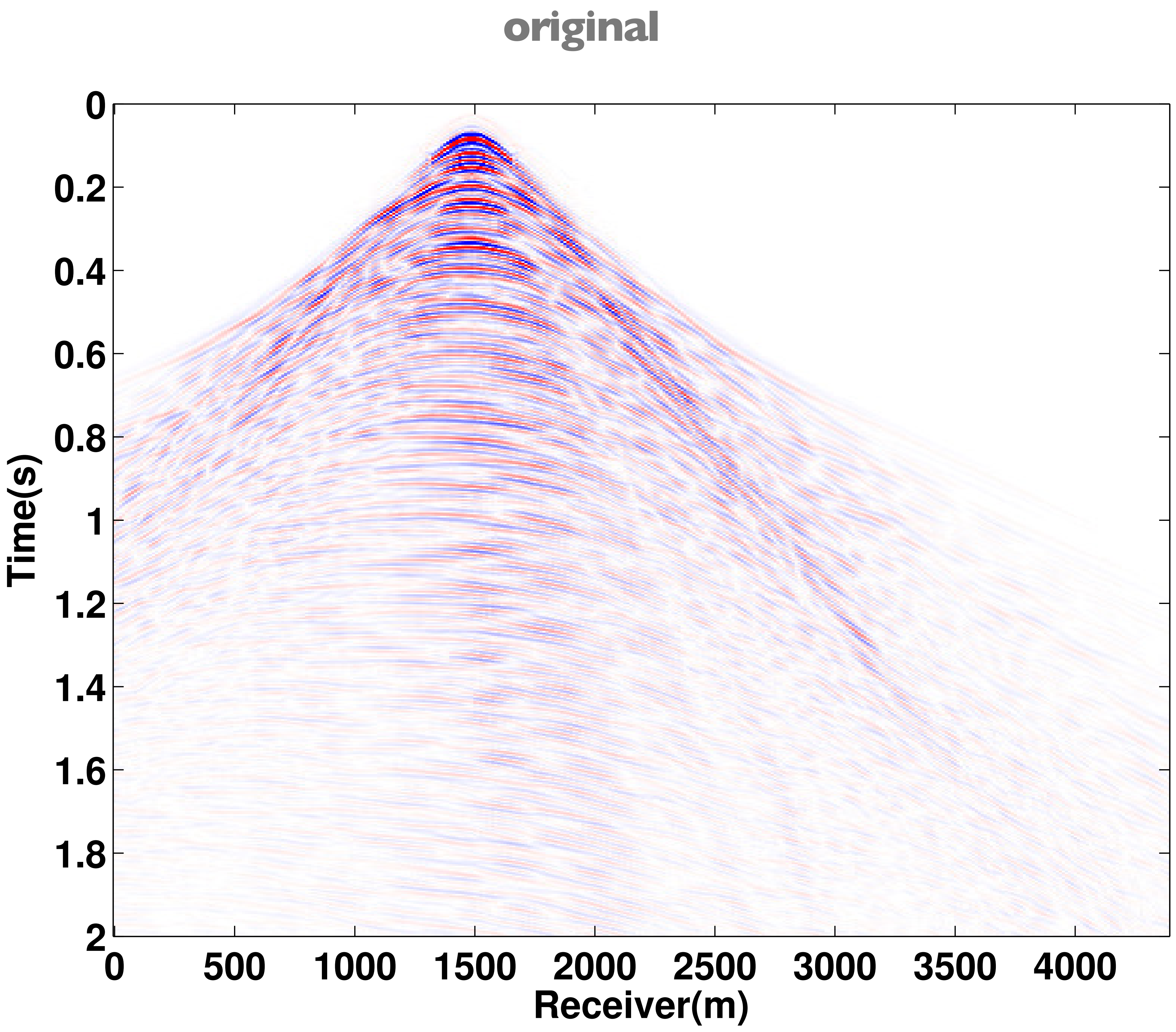


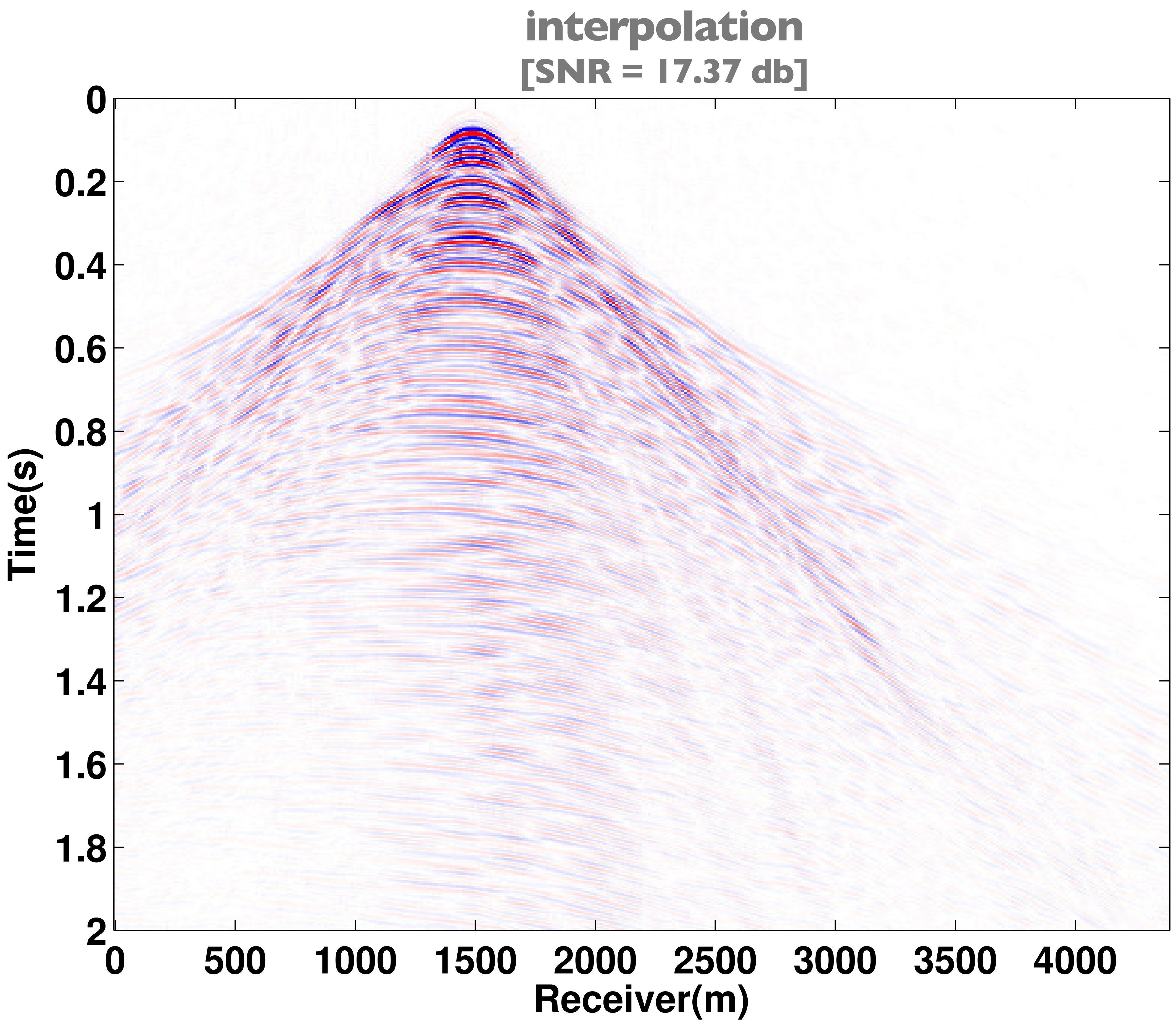
Gulf of Suez
[w/ HSS, level-2]



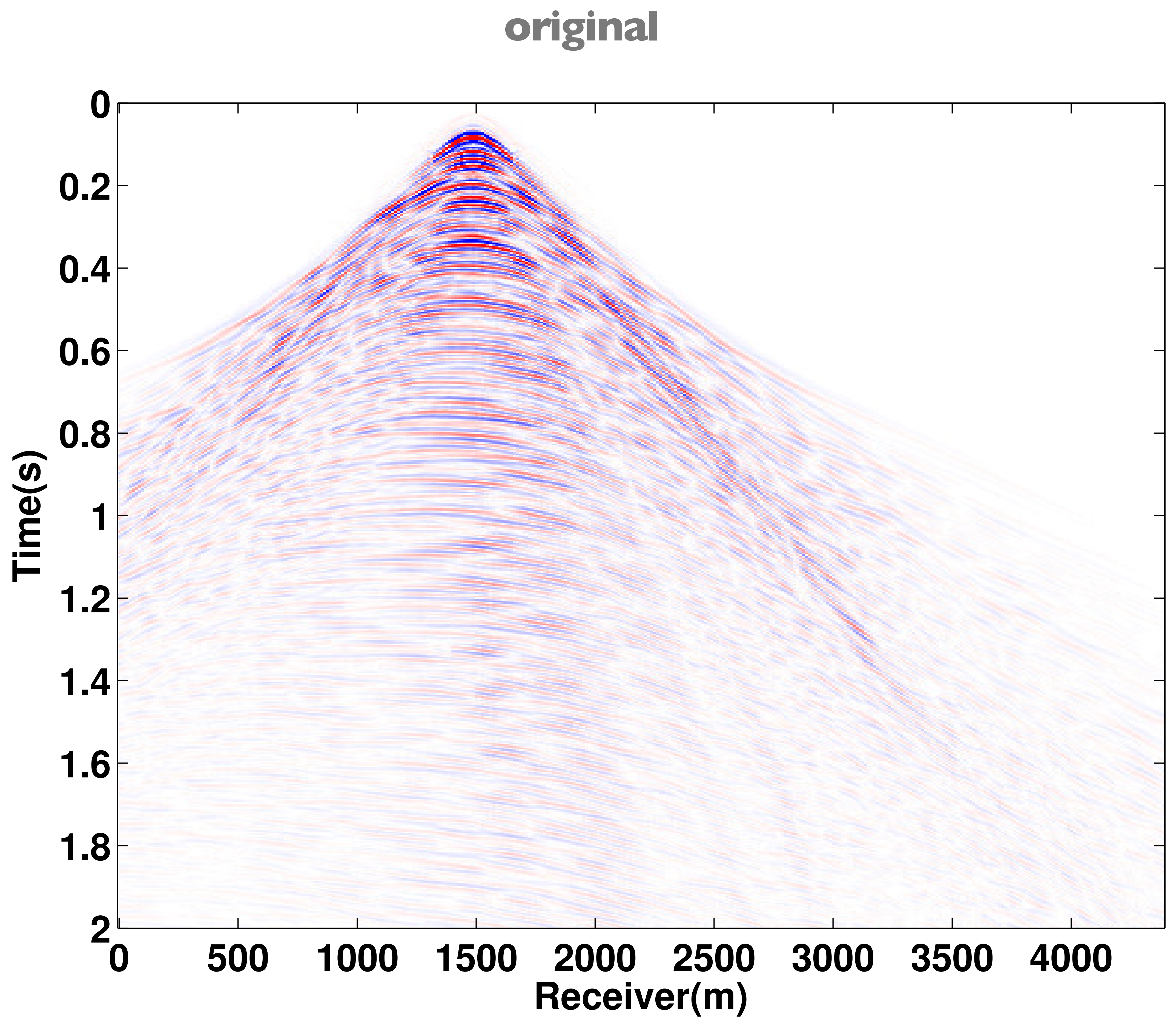
Gulf of Suez

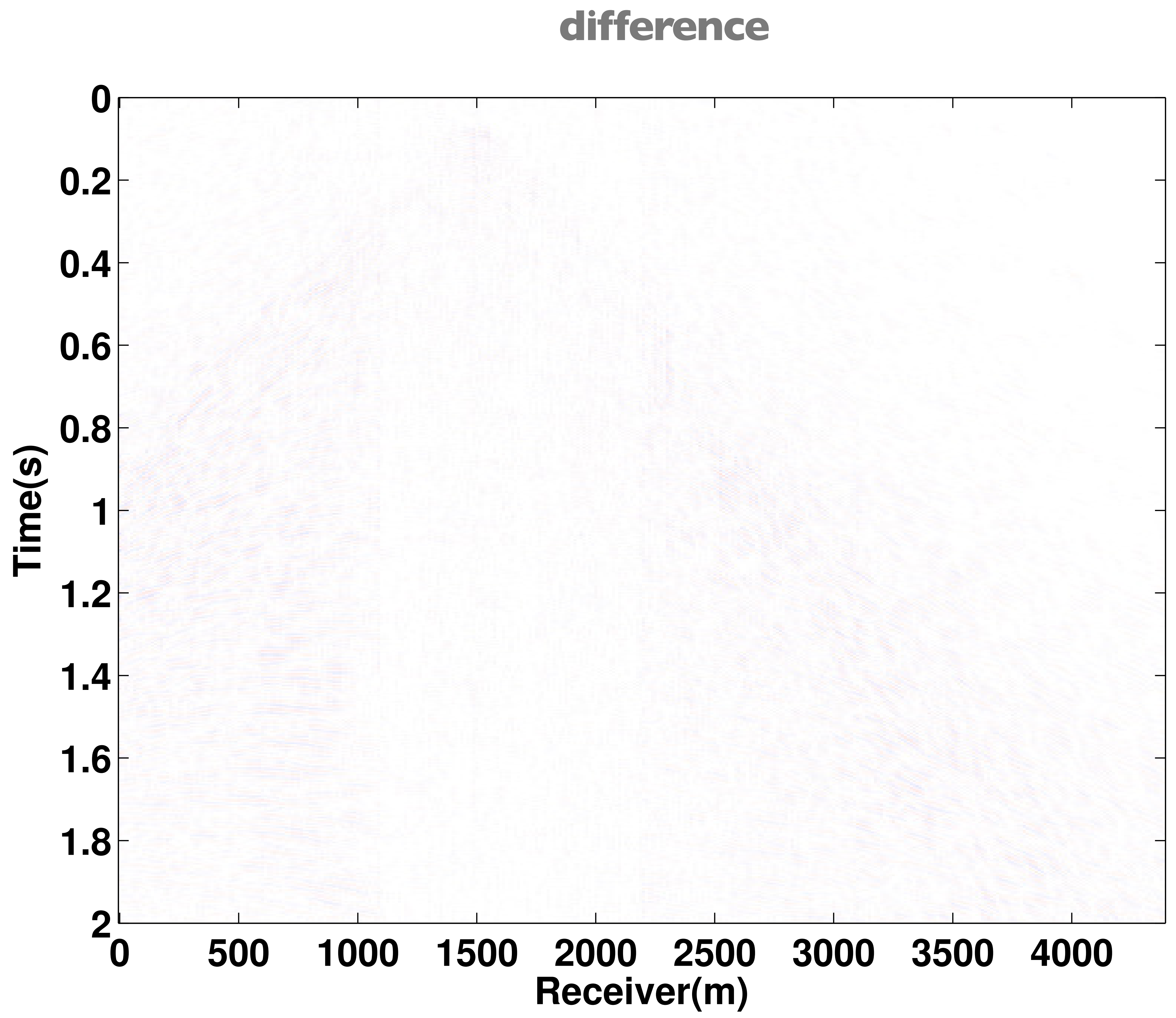
[w/ HSS, level-3]



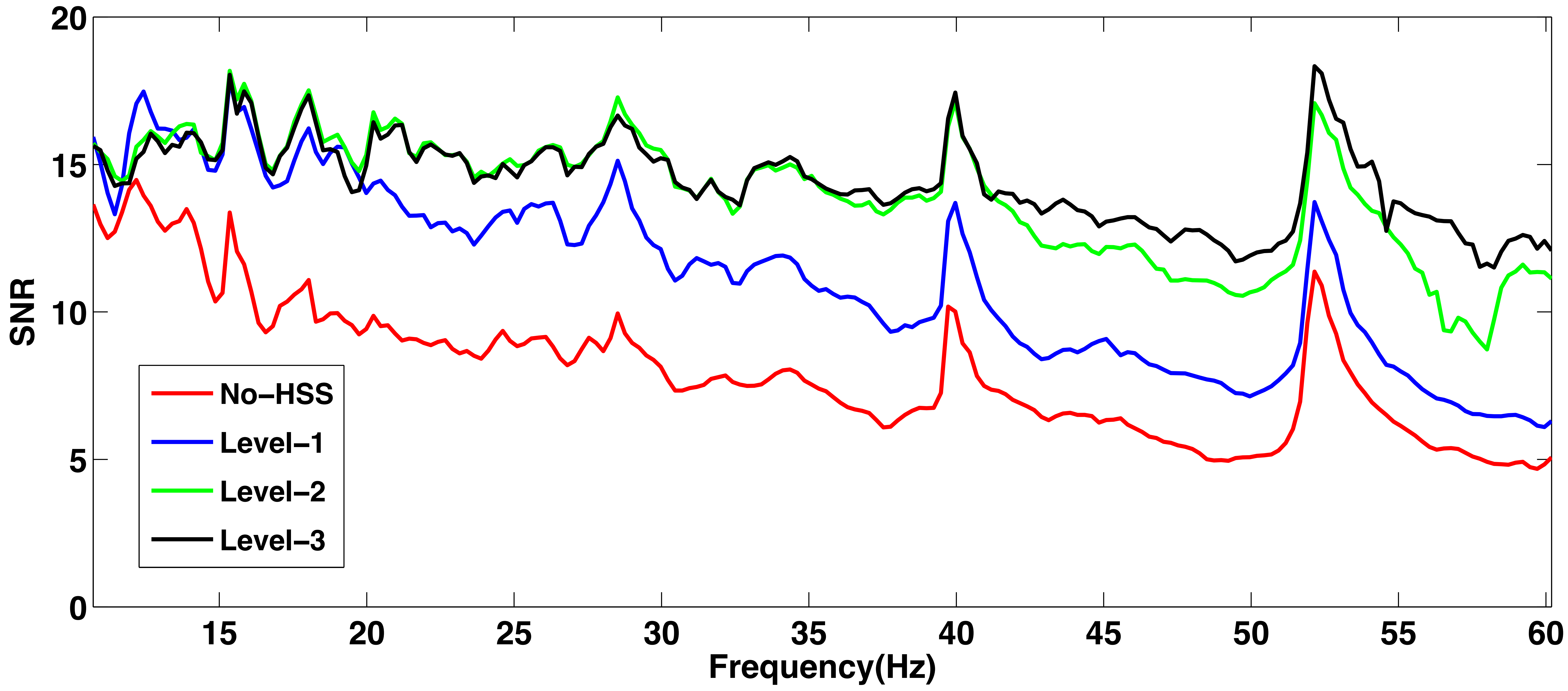


Gulf of Suez
[w/ HSS, level-3]





Comparison
[w/o and w/ HSS]



Conclusion

- firmly established a connection to compressive sensing
  - low-rank structure
  - rank revealing transform domain
  - proper optimization framework

- HSS gives advantage to attack high-rank structures using the partitioning

Conclusion

- matrix factorization allows *SVD-free* low-rank methods that work fast on large data
- low-rank structure holds promise for data recovery and more compact representation

SLIM

Acknowledgements

Thank you for your attention !

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

SINBAD

NSERC CRSNG

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada Discovery Grant (22R81254) and the Collaborative Research and Development Grant DNOISE II (375142-08). This research was carried out as part of the SINBAD II project with support from the following organizations: BG Group, BGP, BP, CGG, Chevron, ConocoPhillips, ION, Petrobras, PGS, Total SA, WesternGeco, and Woodside.