Wavefield reconstruction via SVD-free low-rank matrix factorization

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

SLIM
University of British Columbia

Motivation

- acquisition challenges
  - missing data
  - noise

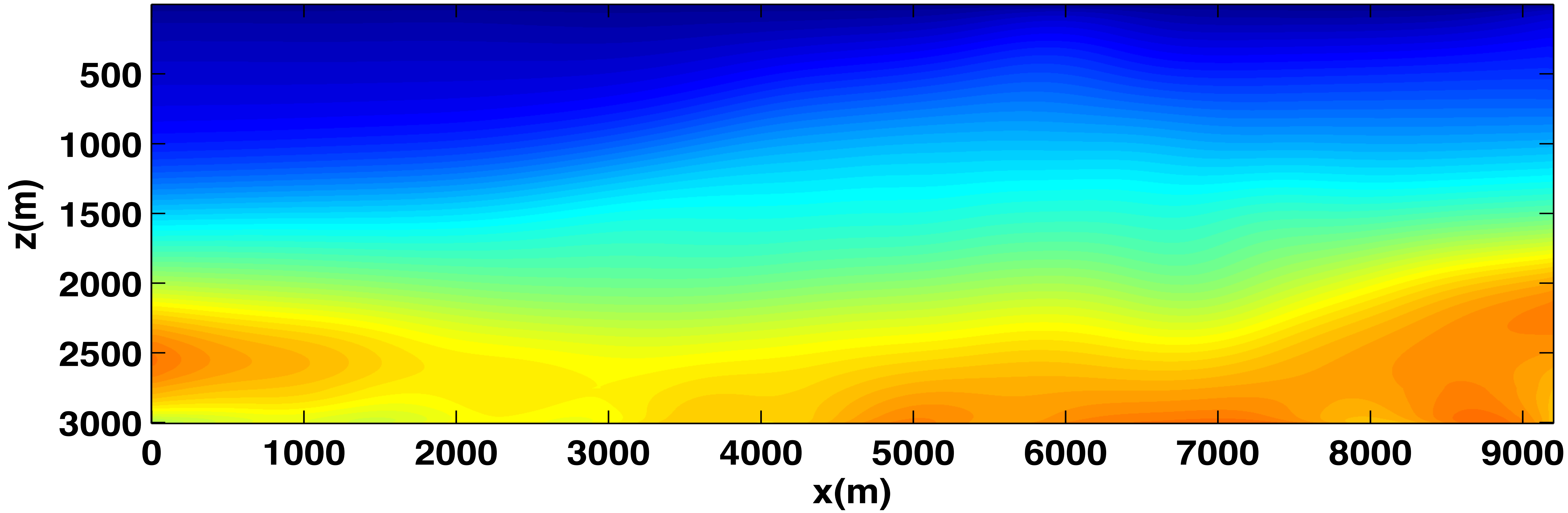
- fully sampled data
  - simultaneous shot based FWI & migration
  - estimation of primaries by sparse inversion & SRME

- exploit *low-rank* structure of seismic data
  - randomized sampling
  - *SVD-free* matrix factorization

[Li et. al. 2012, van Leeuwen and Herrmann 2012]

Full waveform inversion

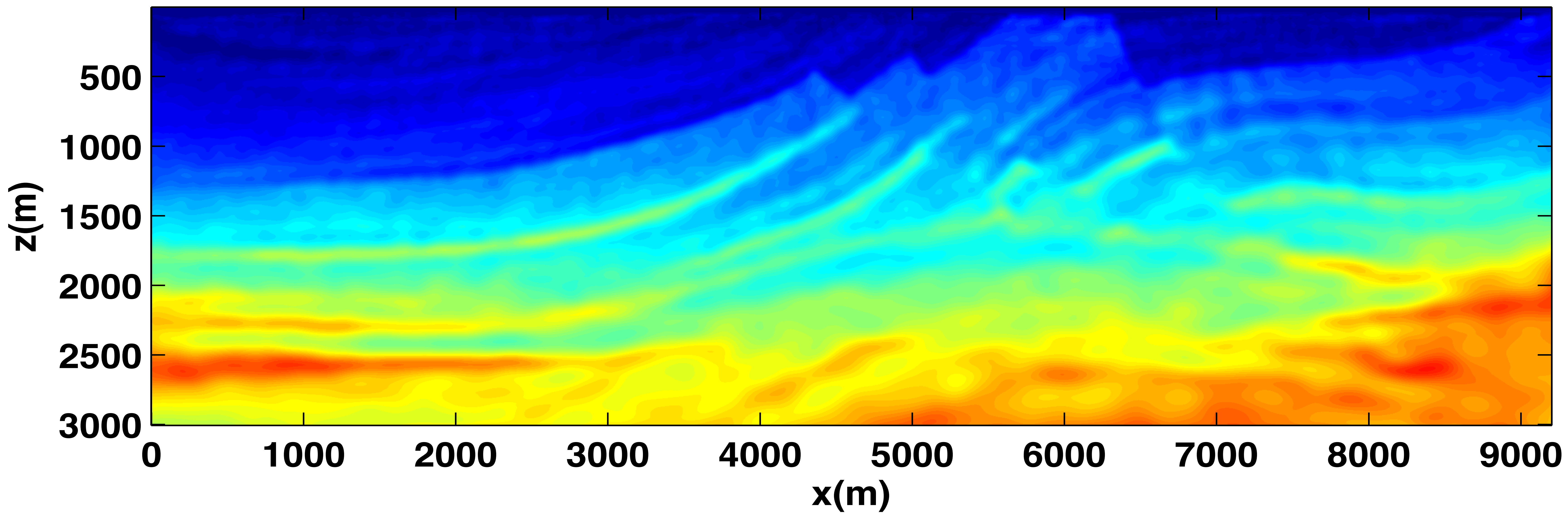
[initial model]



[Li et. al. 2012, van Leeuwen and Herrmann 2012]

Full waveform inversion

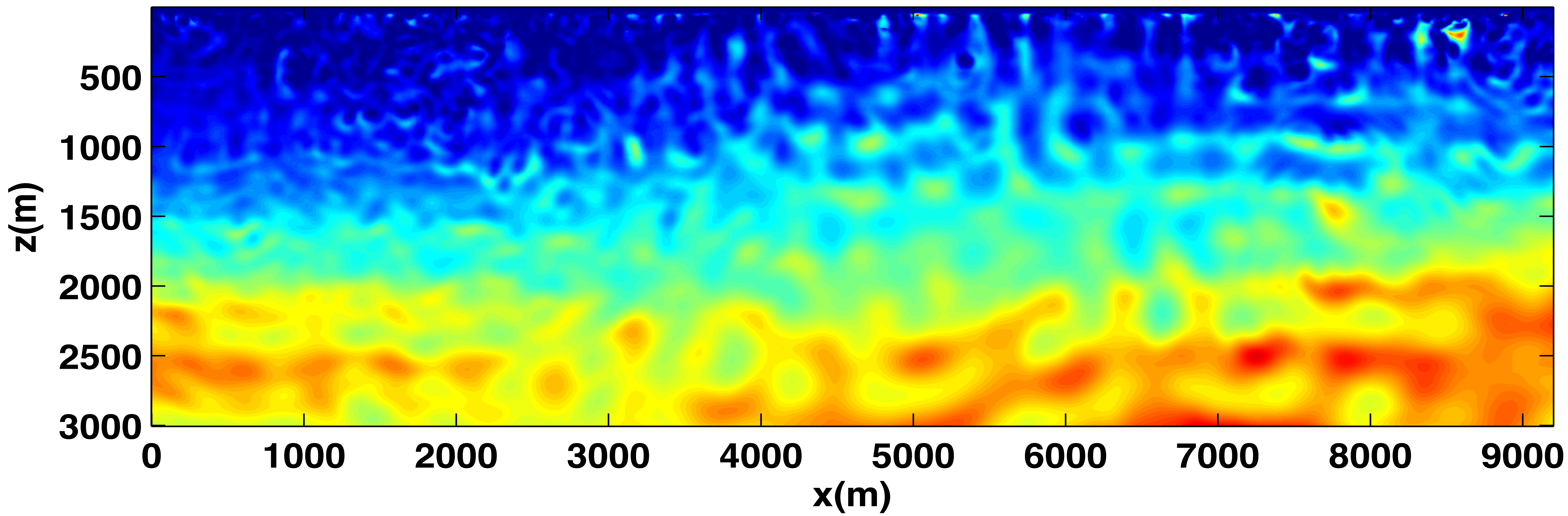
[fully sampled data]



[Li et. al. 2012, van Leeuwen and Herrmann 2012]

Full waveform inversion

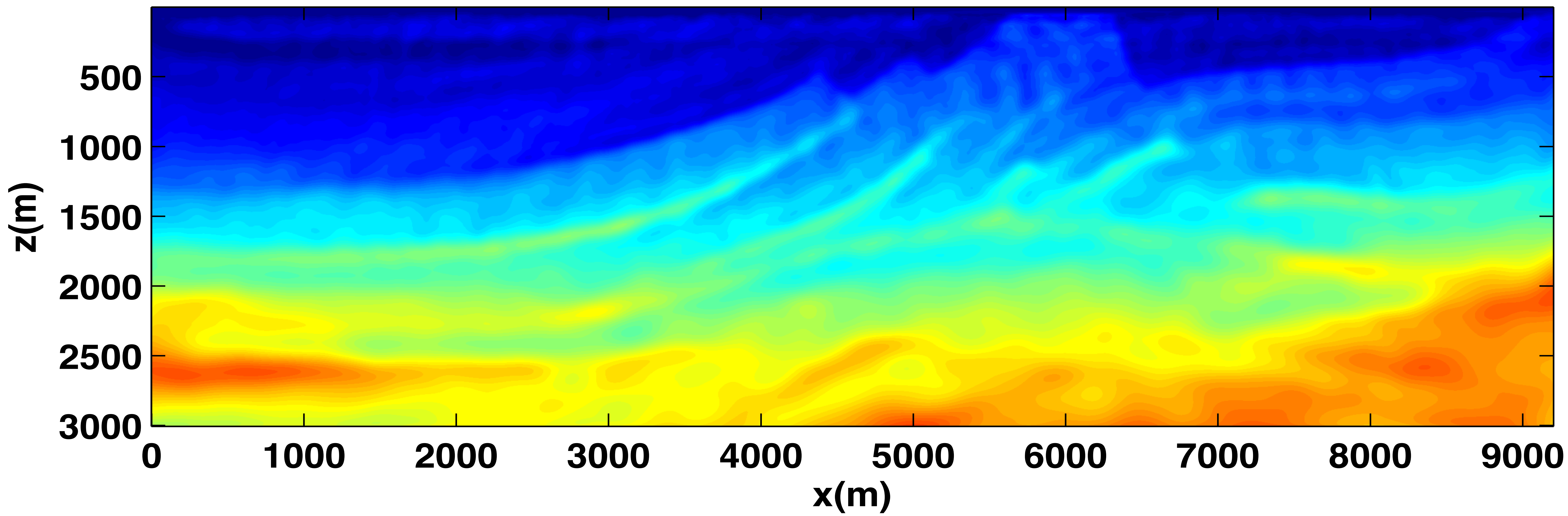
[Subsampled data]



[Li et. al. 2012, van Leeuwen and Herrmann 2012]

Full waveform inversion

[low-rank interpolation]



[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
- uniform random subsampling
  - no control on the maximum gap size

[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??

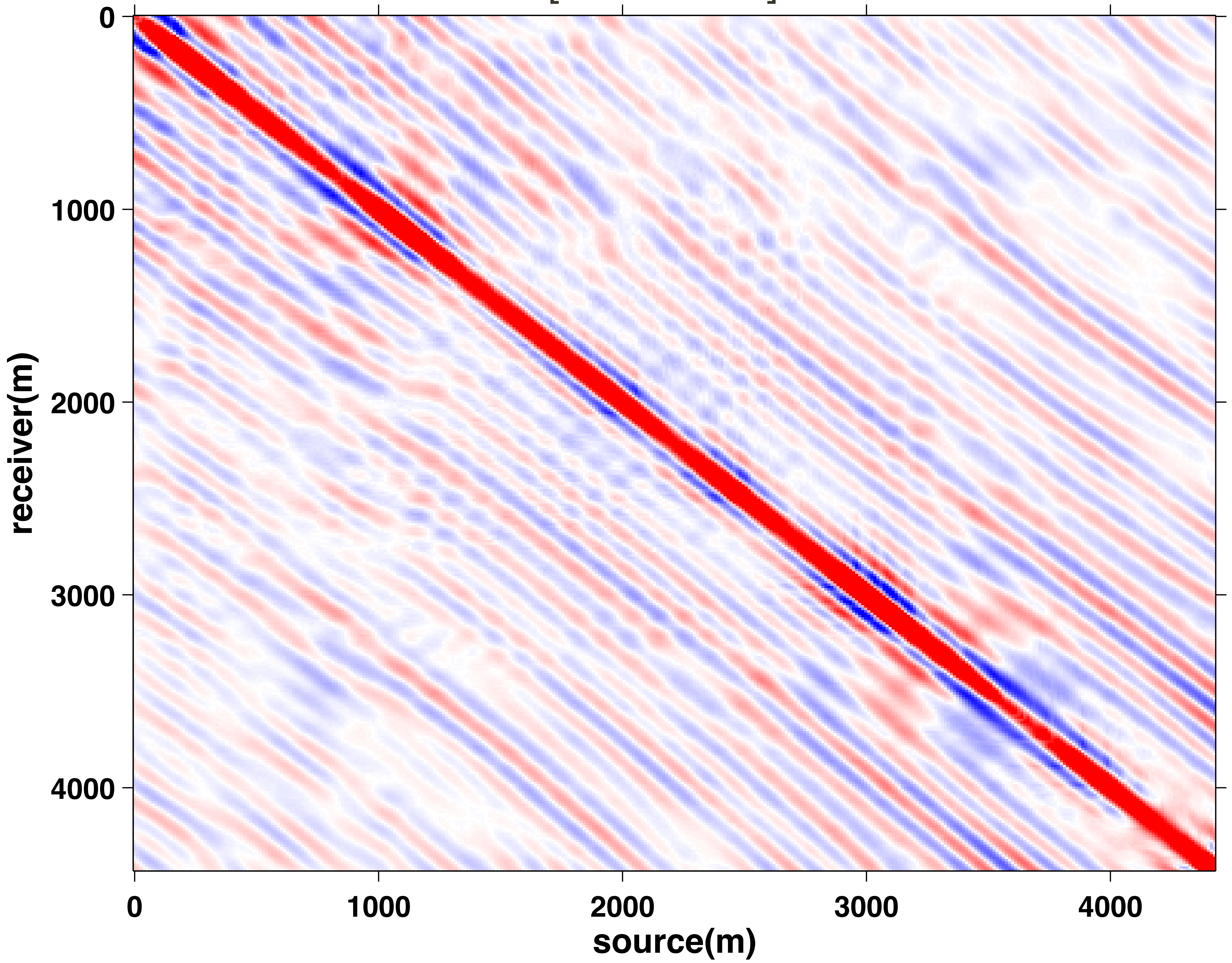
[Candes and Plan 2010, Oropeza and Sacchi 2011]

Matrix completion

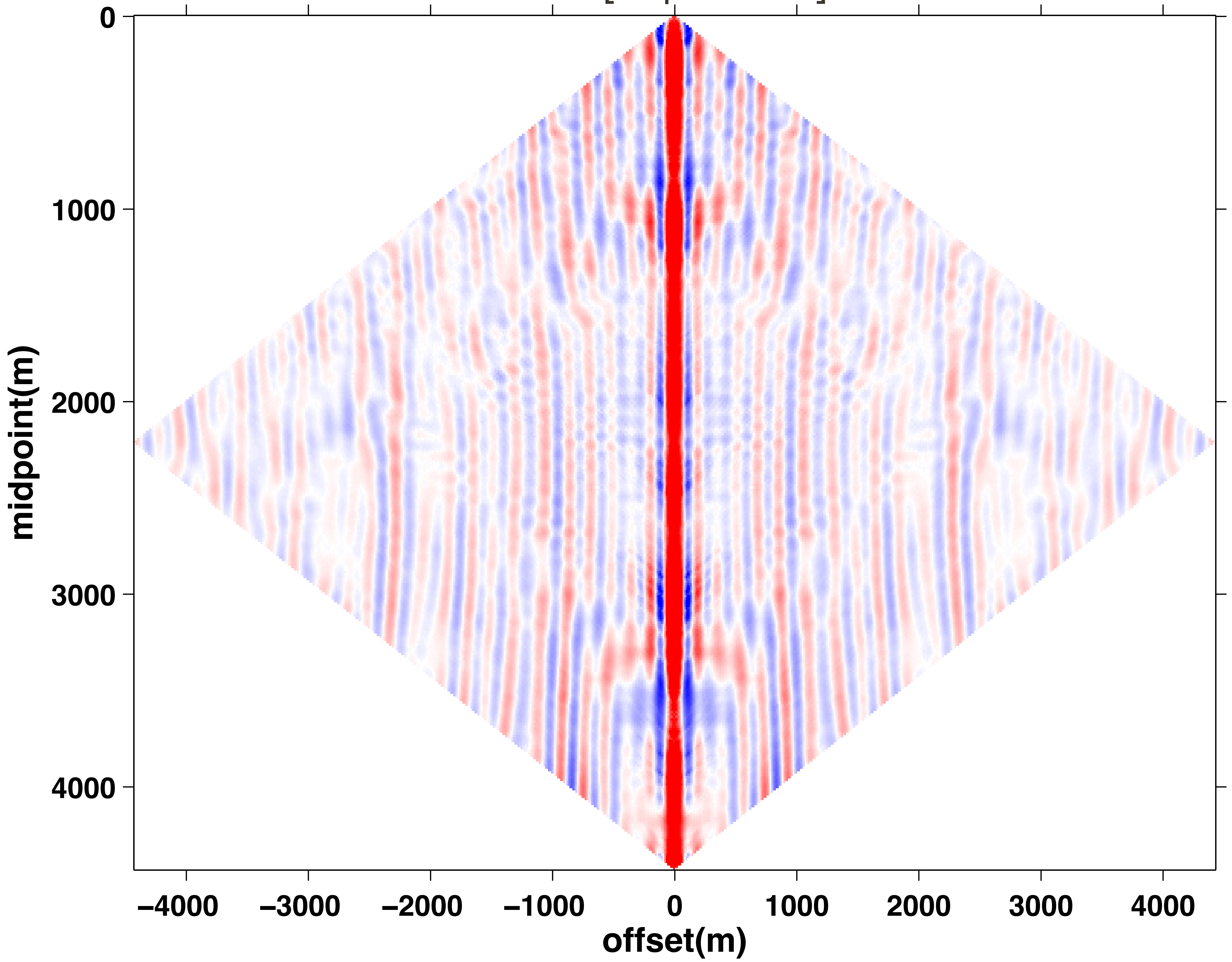
- signal structure
  - *low rank/fast decay* of singular values
- sampling scheme
  - missing data *increase* rank in "transform domain"
- recovery using *rank penalization* scheme

Low-rank structure
[2-D acquisition]

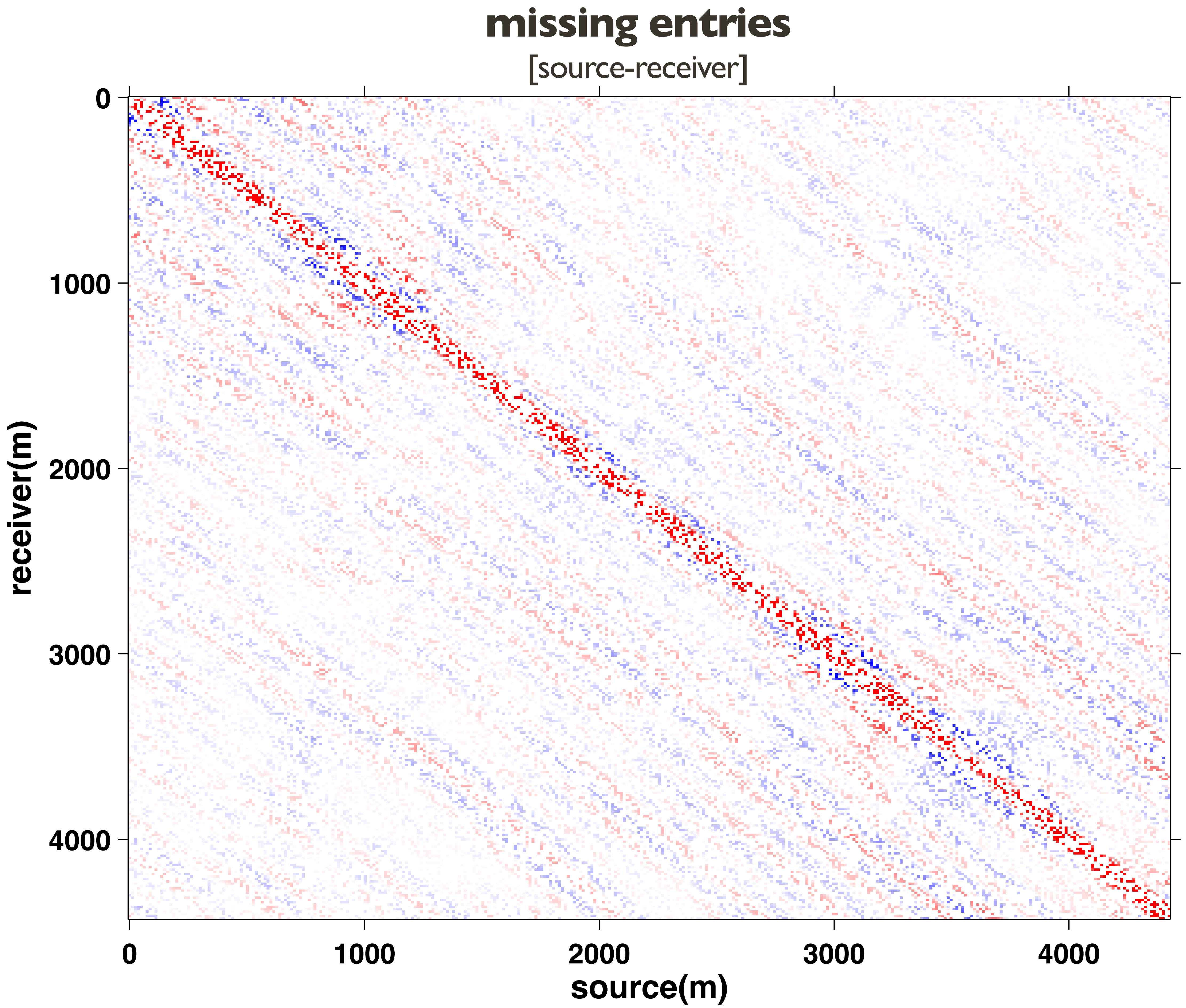
acquisition domain
[source-receiver]

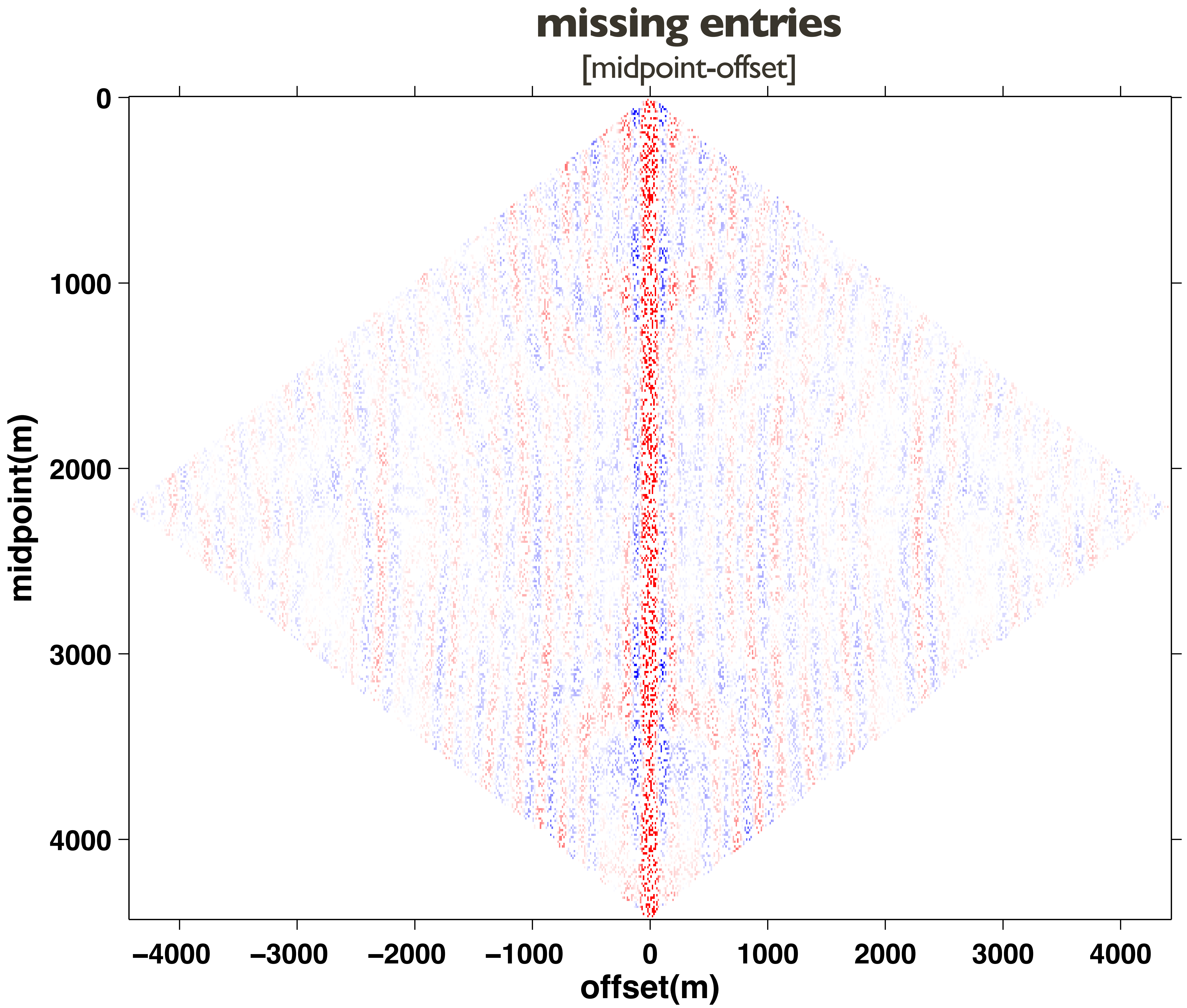


transform domain
[midpoint-offset]

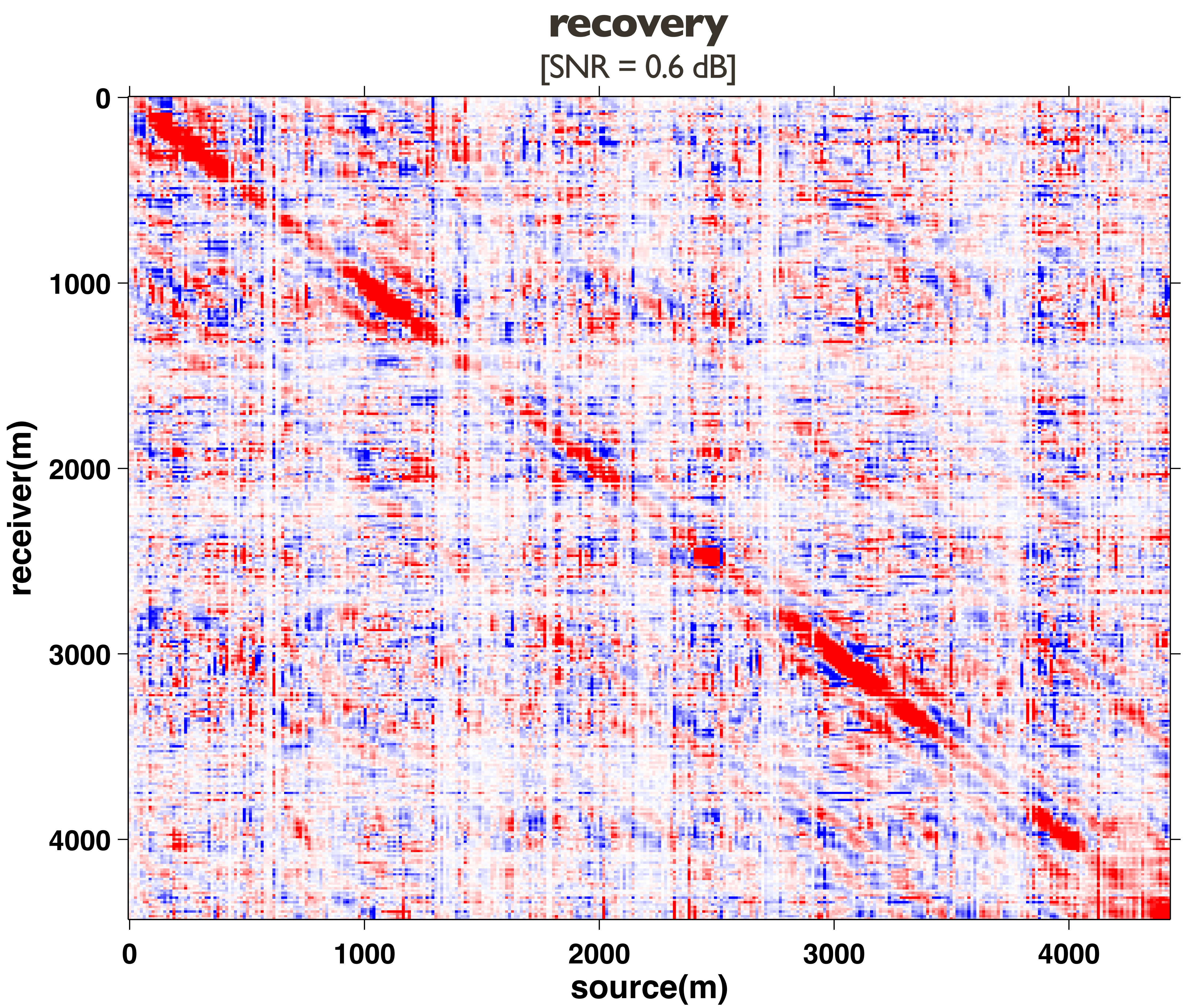


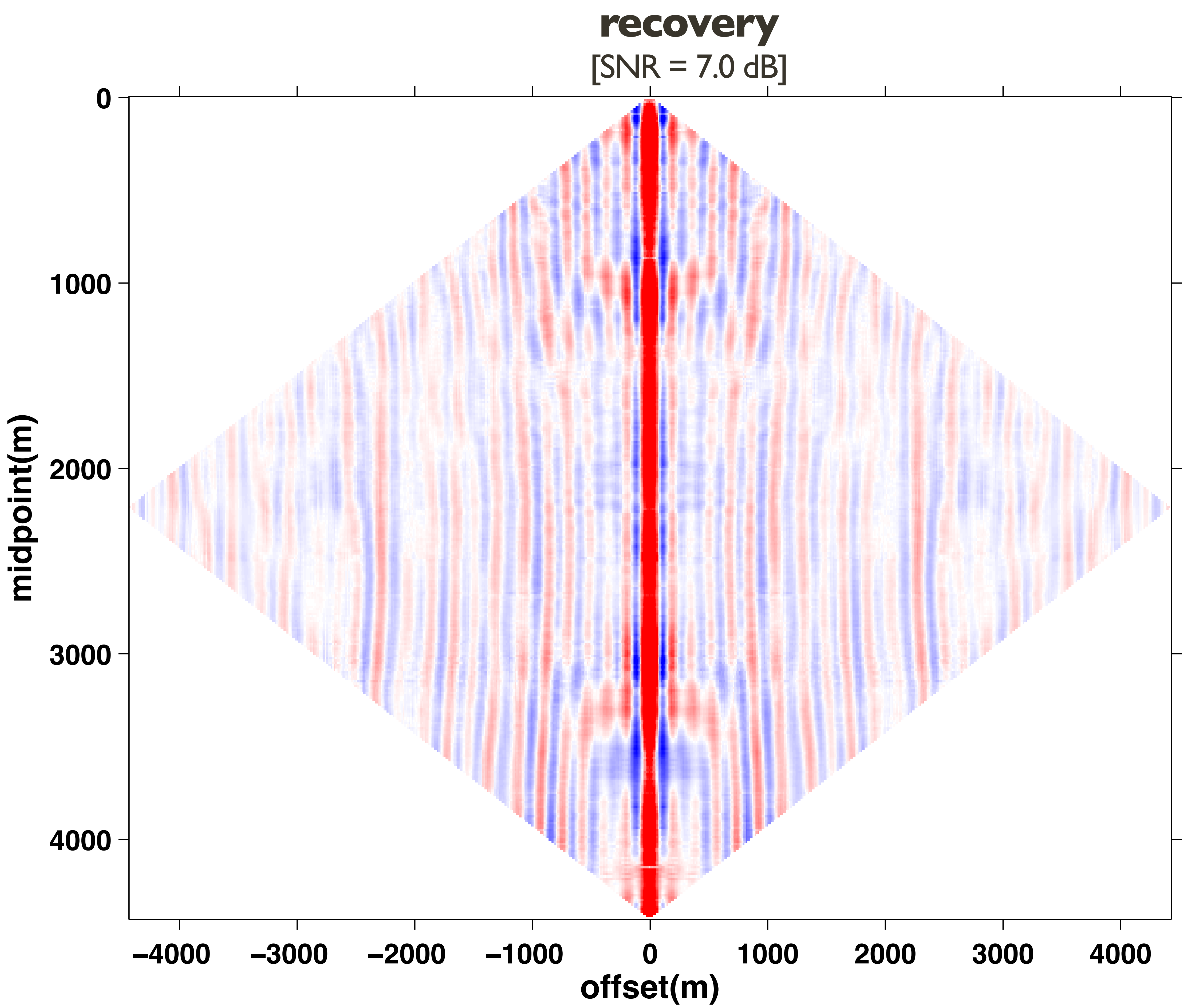
Matrix completion problem



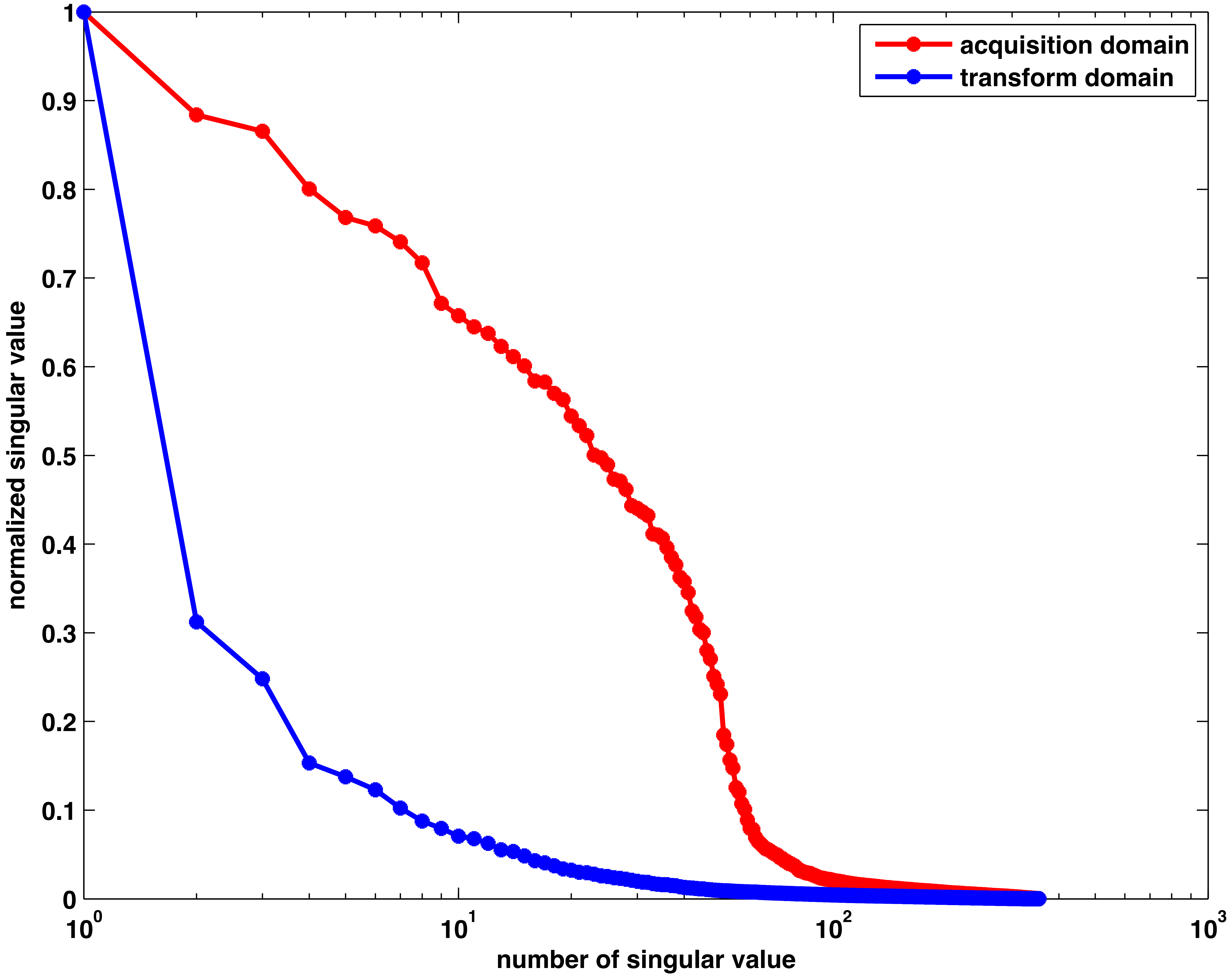


Low-rank interpolation





Singular value decay
[2-D acquisition]



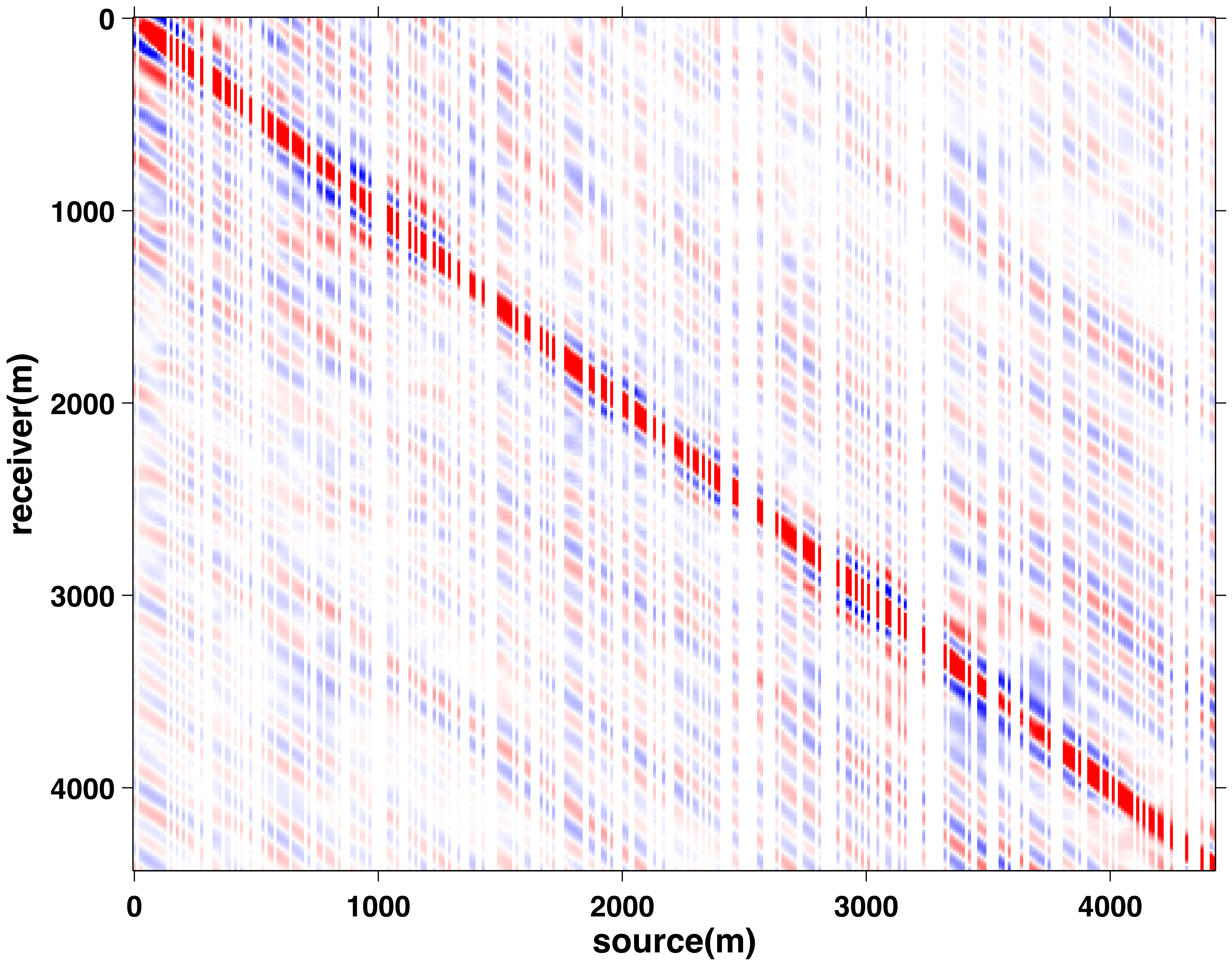
Matrix completion

- 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]

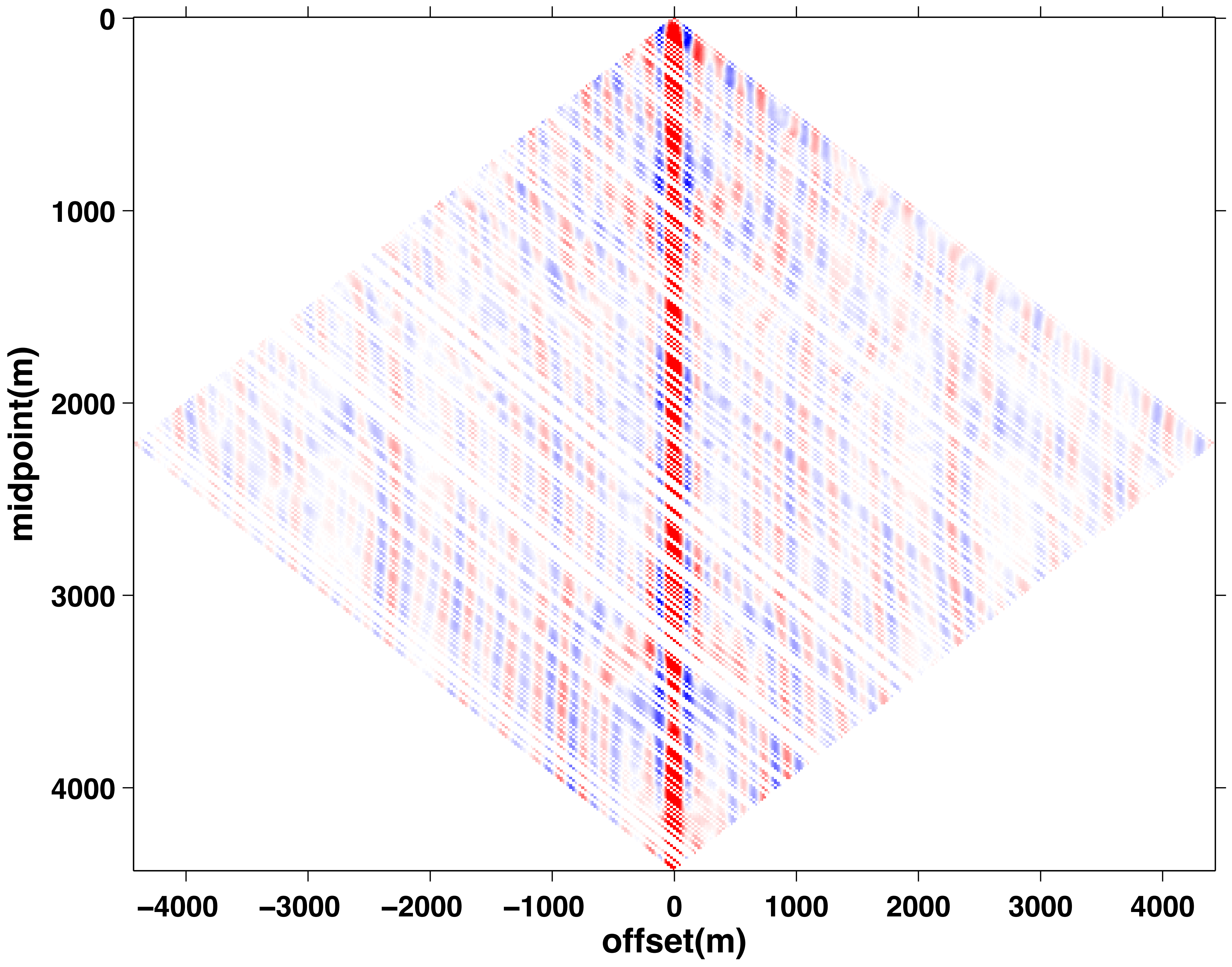
acquisition domain

missing columns *do not* increase rank

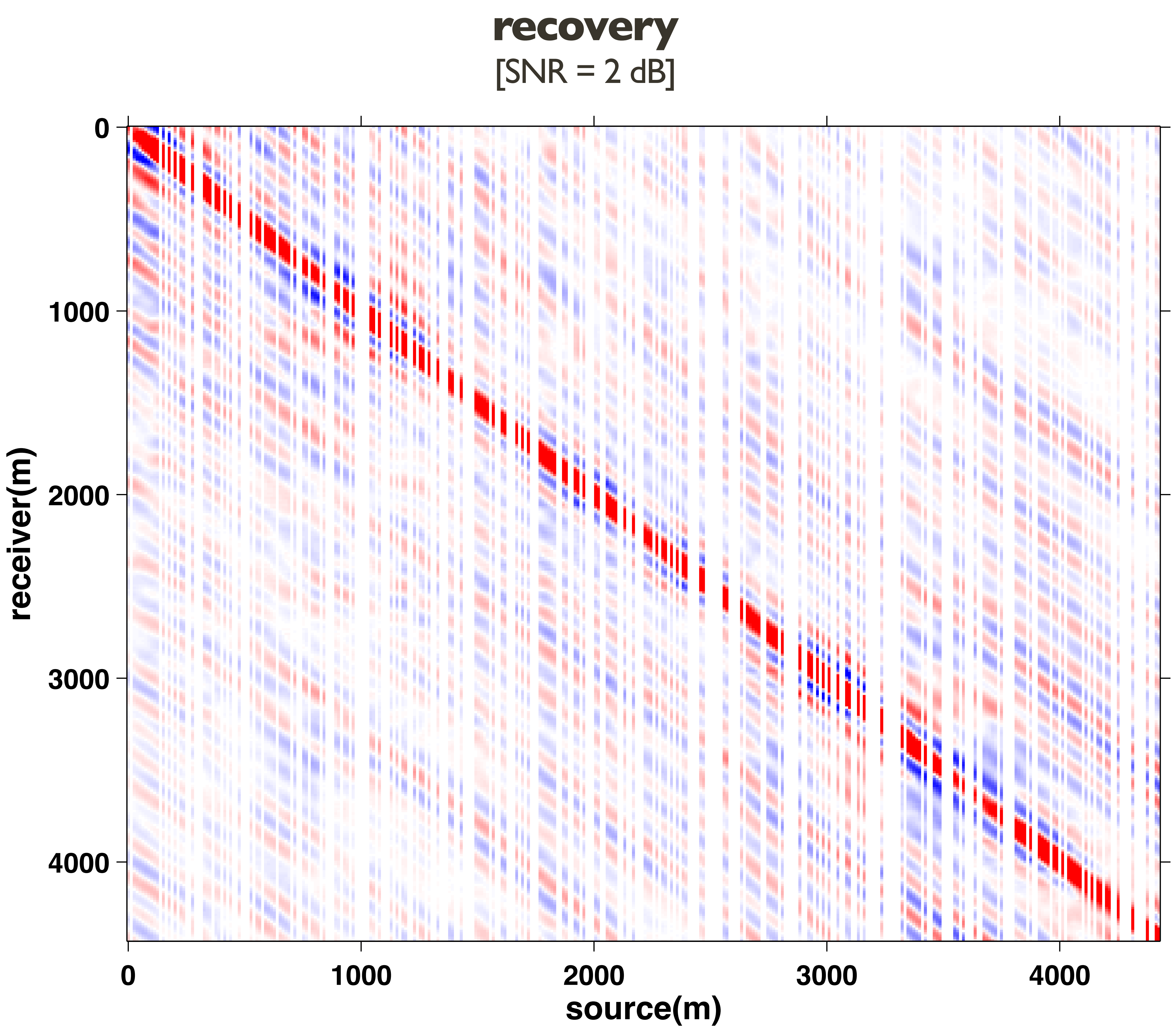


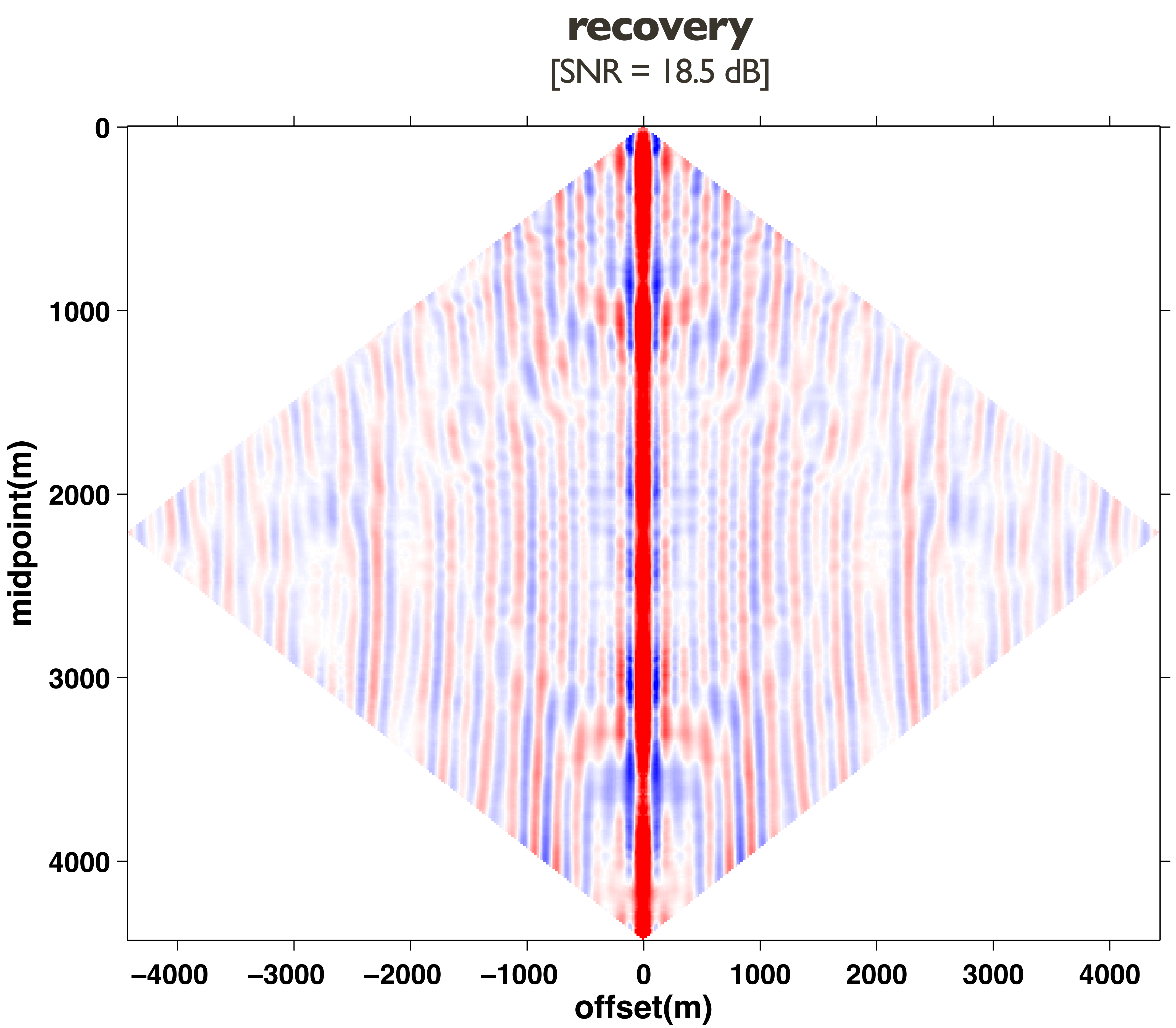
transform domain

missing columns *do* increase rank

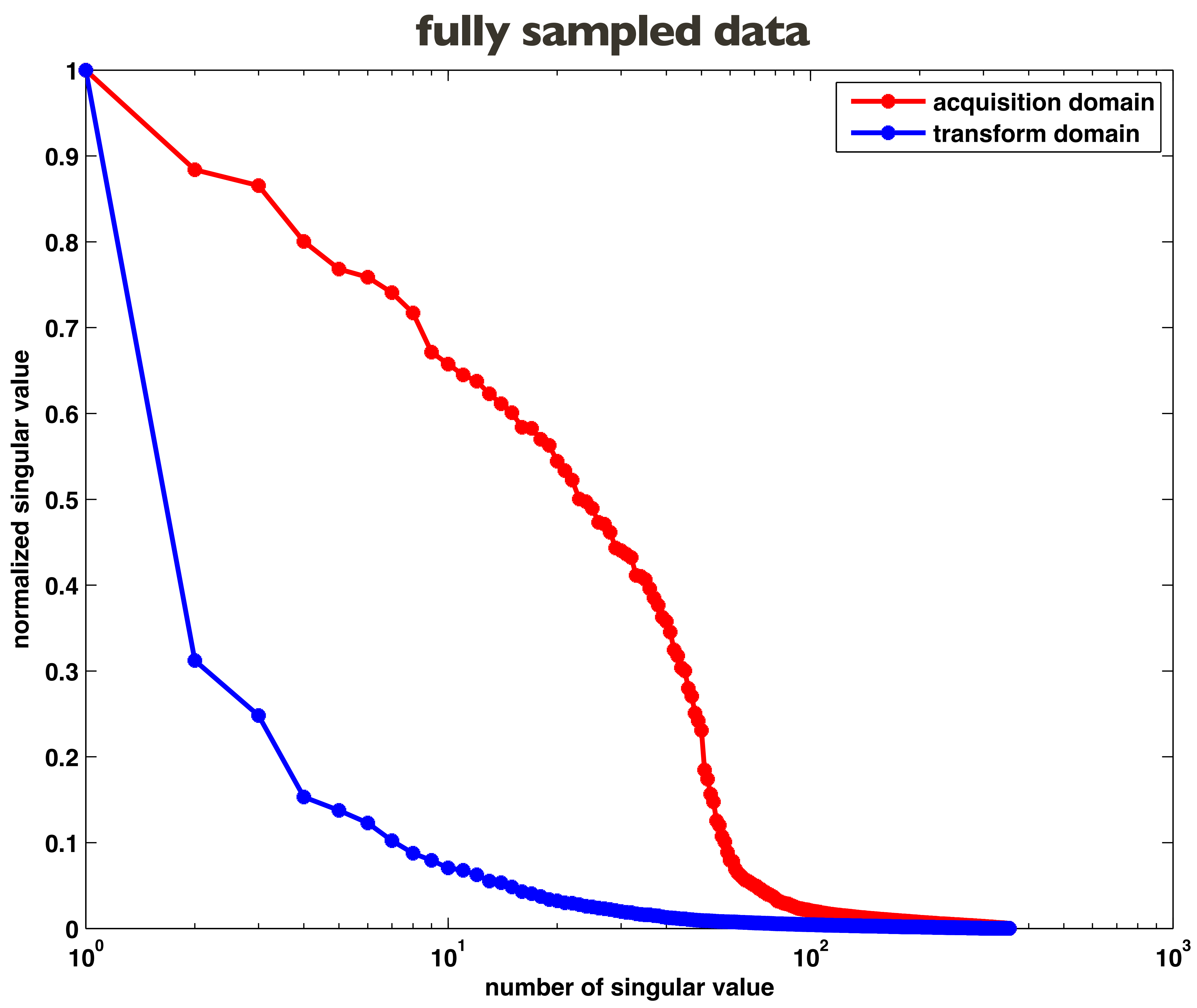


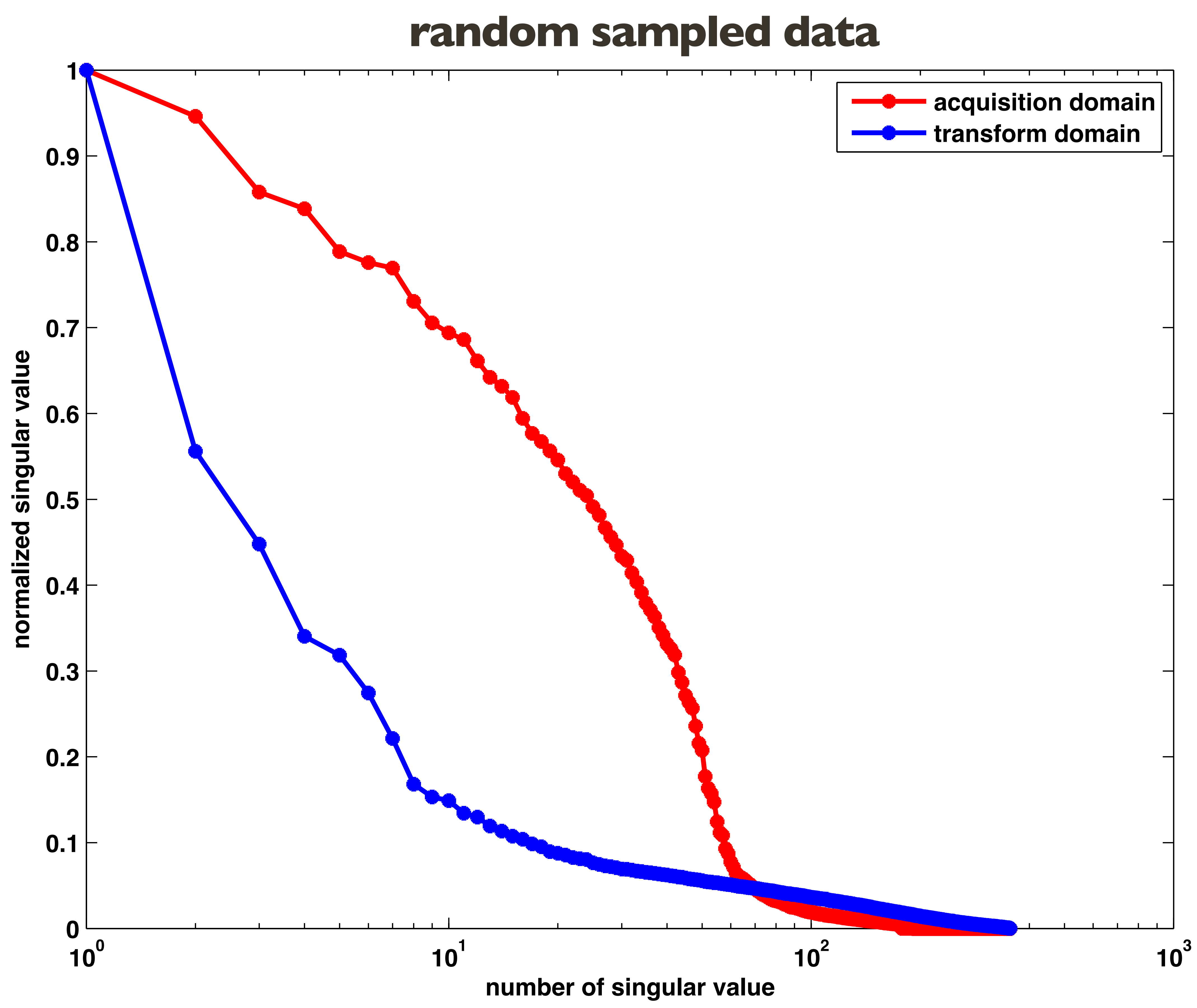
Low-rank interpolation





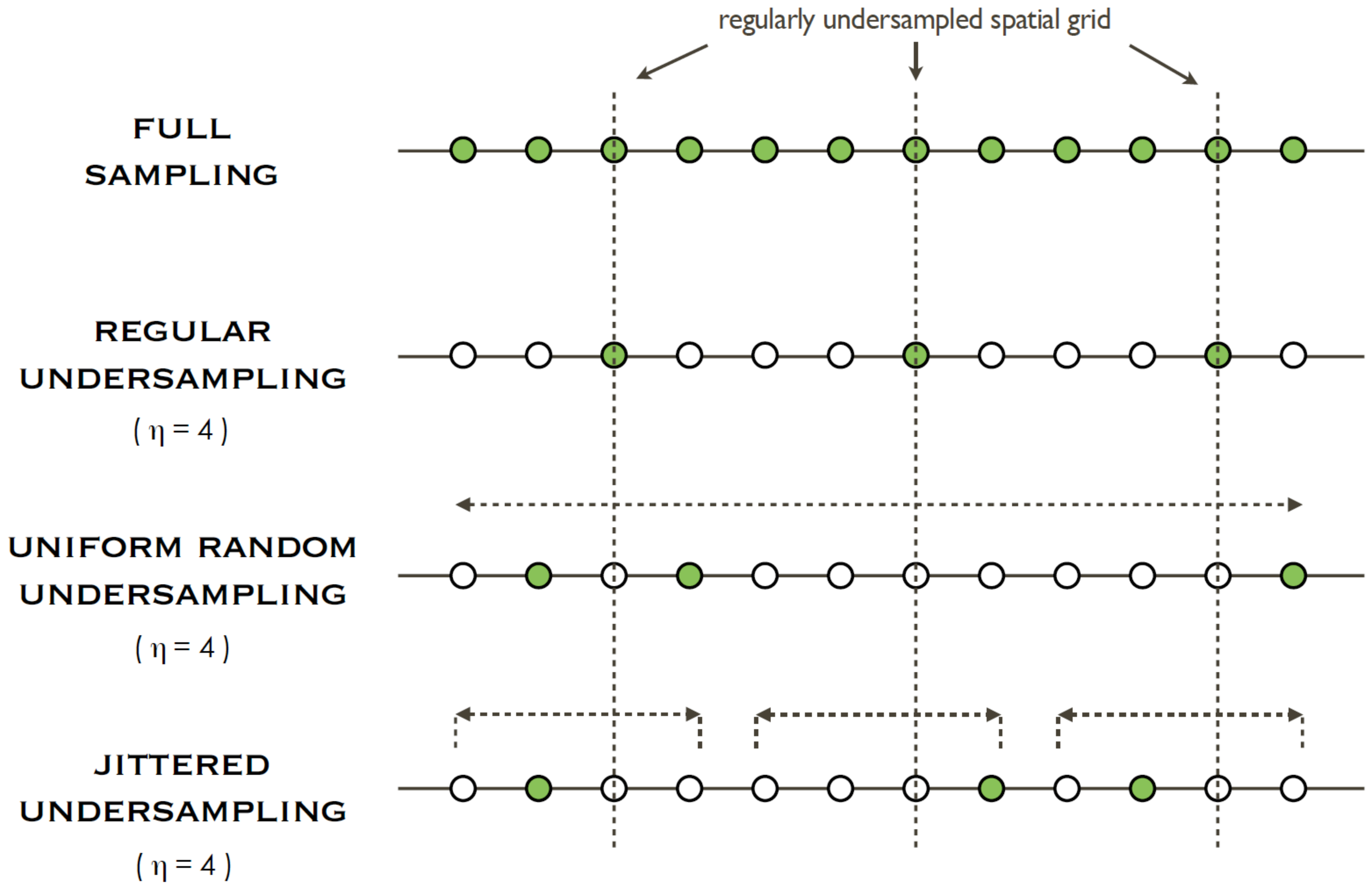
Randomized sampling
[singular value decay]





[Hennenfent et. al. 2008]

Sampling schemes



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

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

where $\text{rank}(\mathbf{X}) =$ 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]

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

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

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

[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

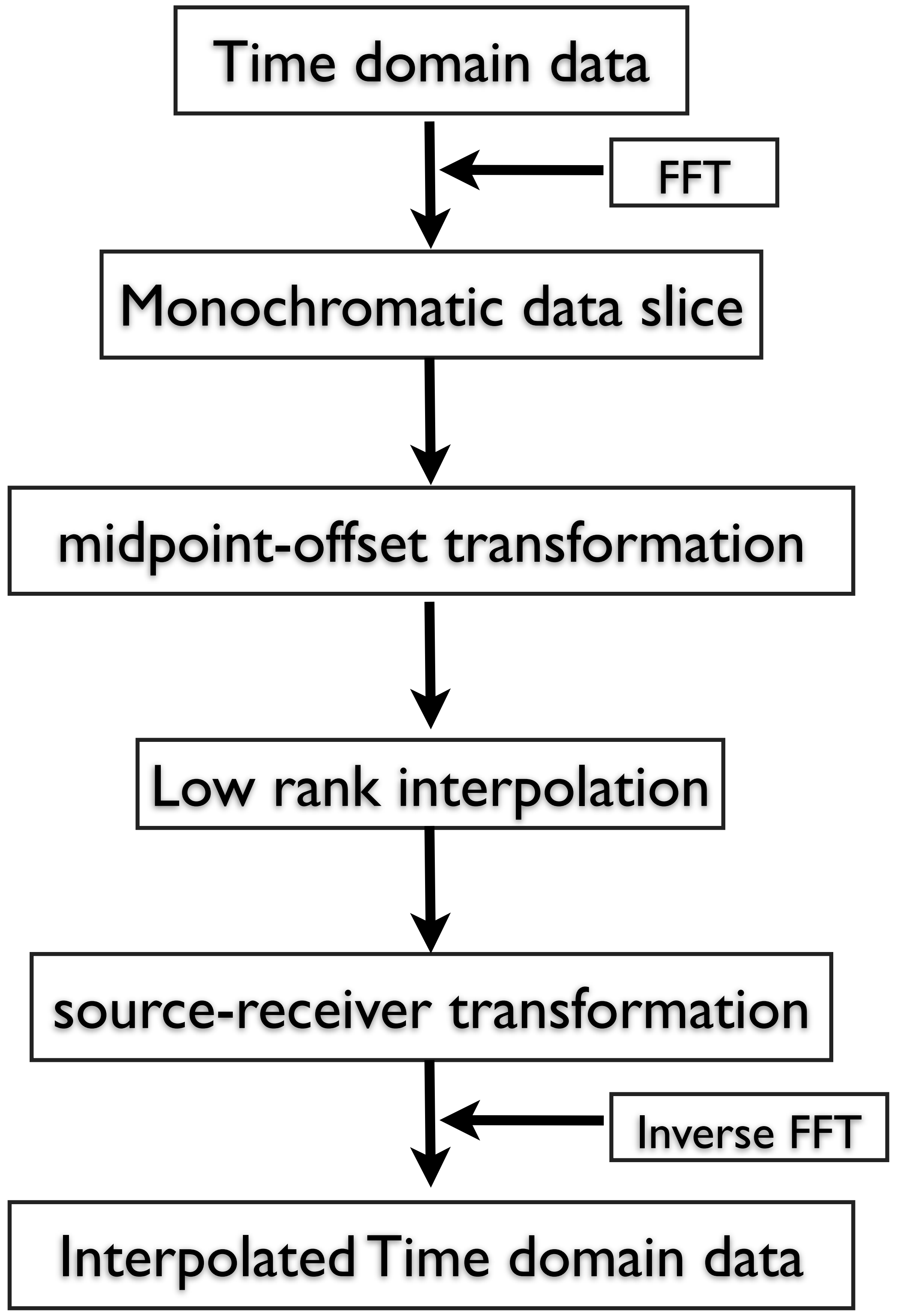
- Upper-bound on nuclear norm is defined as

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

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

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

Interpolation flow chart



Experiments and Results

Case 1 : Uniform random subsampling

Case 2 : Jittered subsampling

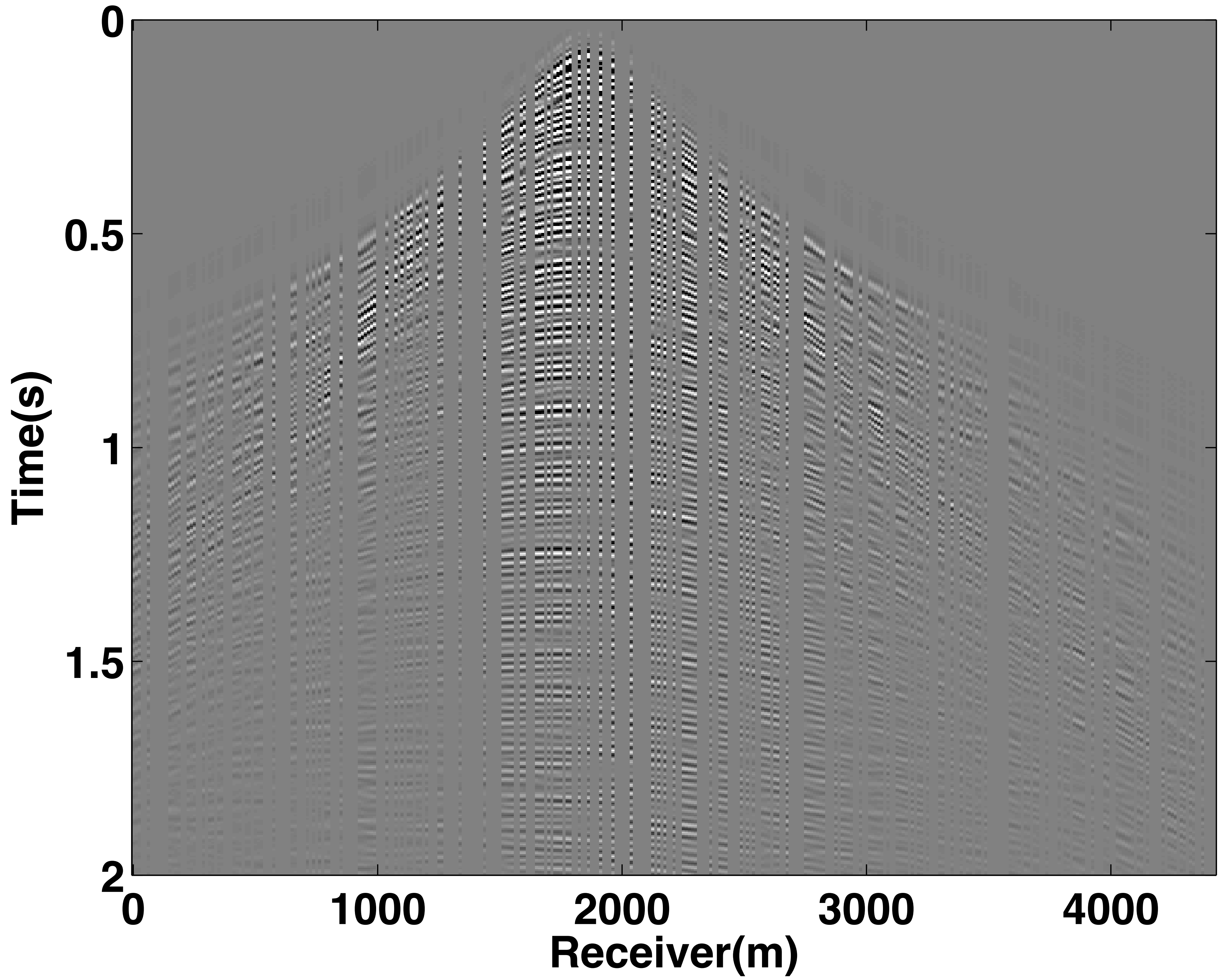
Case 3 : Jittered + Reciprocity

Case 4 : Simultaneous acquisition (Land)

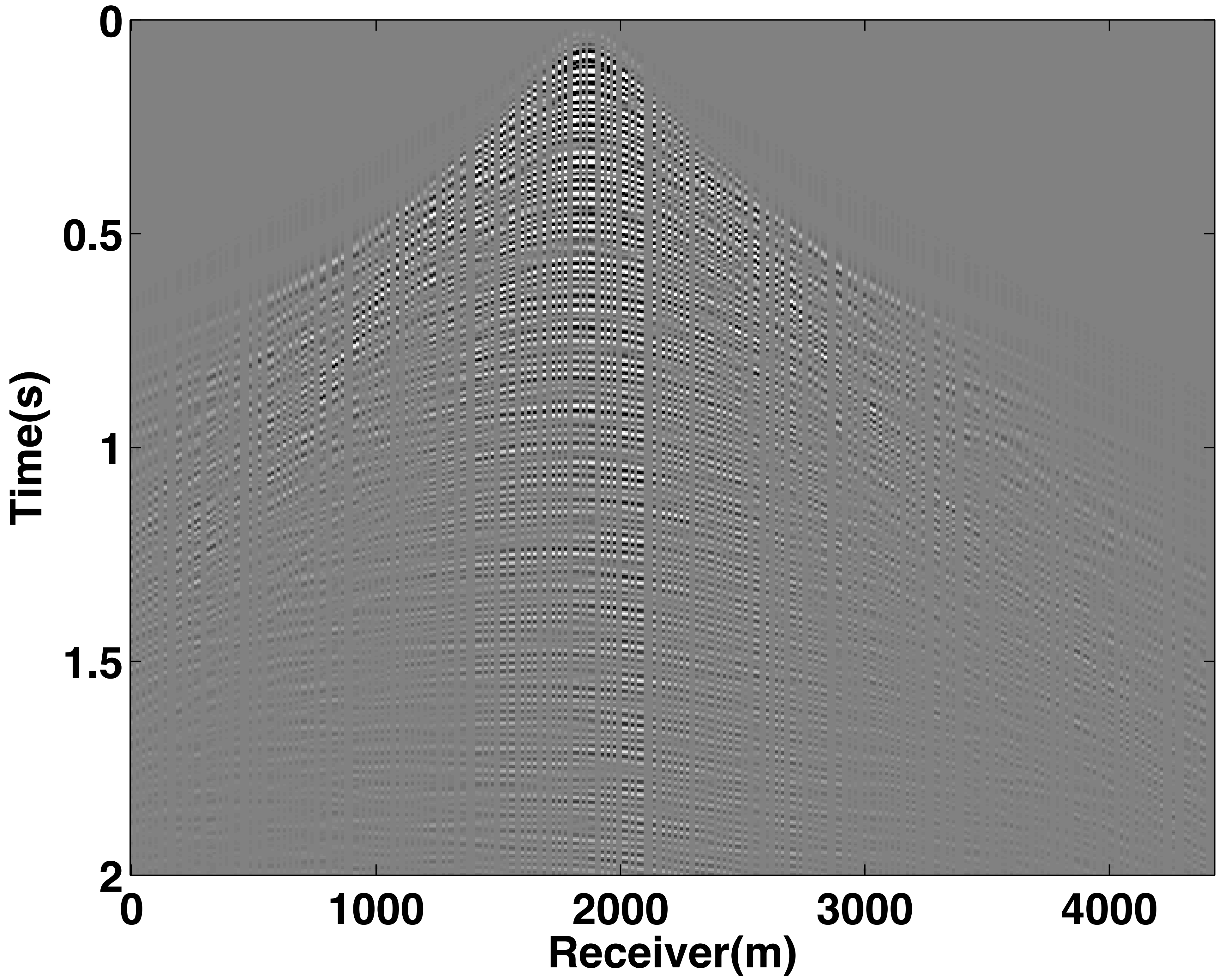
Experiments and Results

- Gulf of Suez
  - 2-D seismic line
  - 50 % missing traces
  - rank adjusted from low to high frequency
  - 150 iterations

Uniform random v/s Jittered



Discrete random

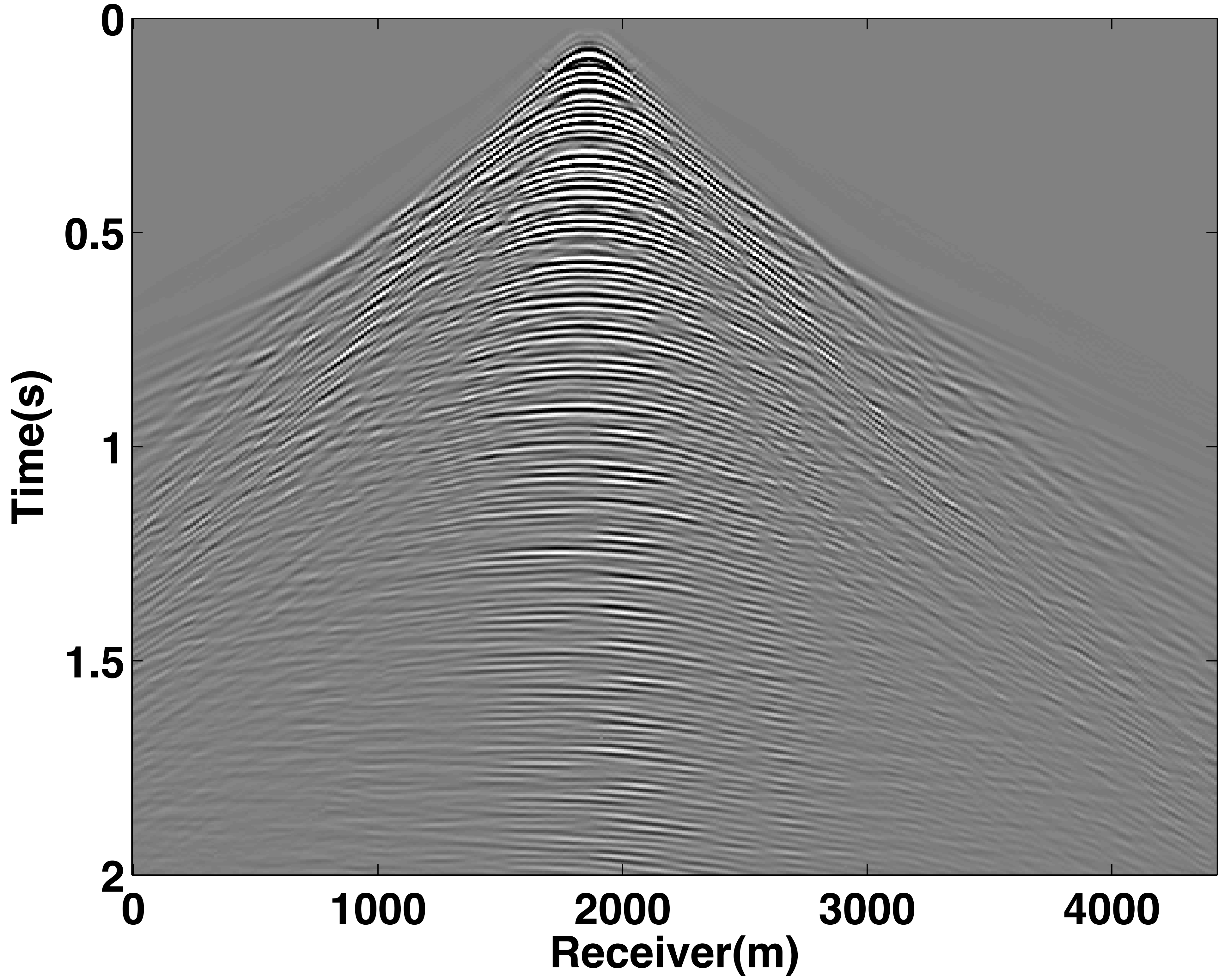


Jittered

Case I

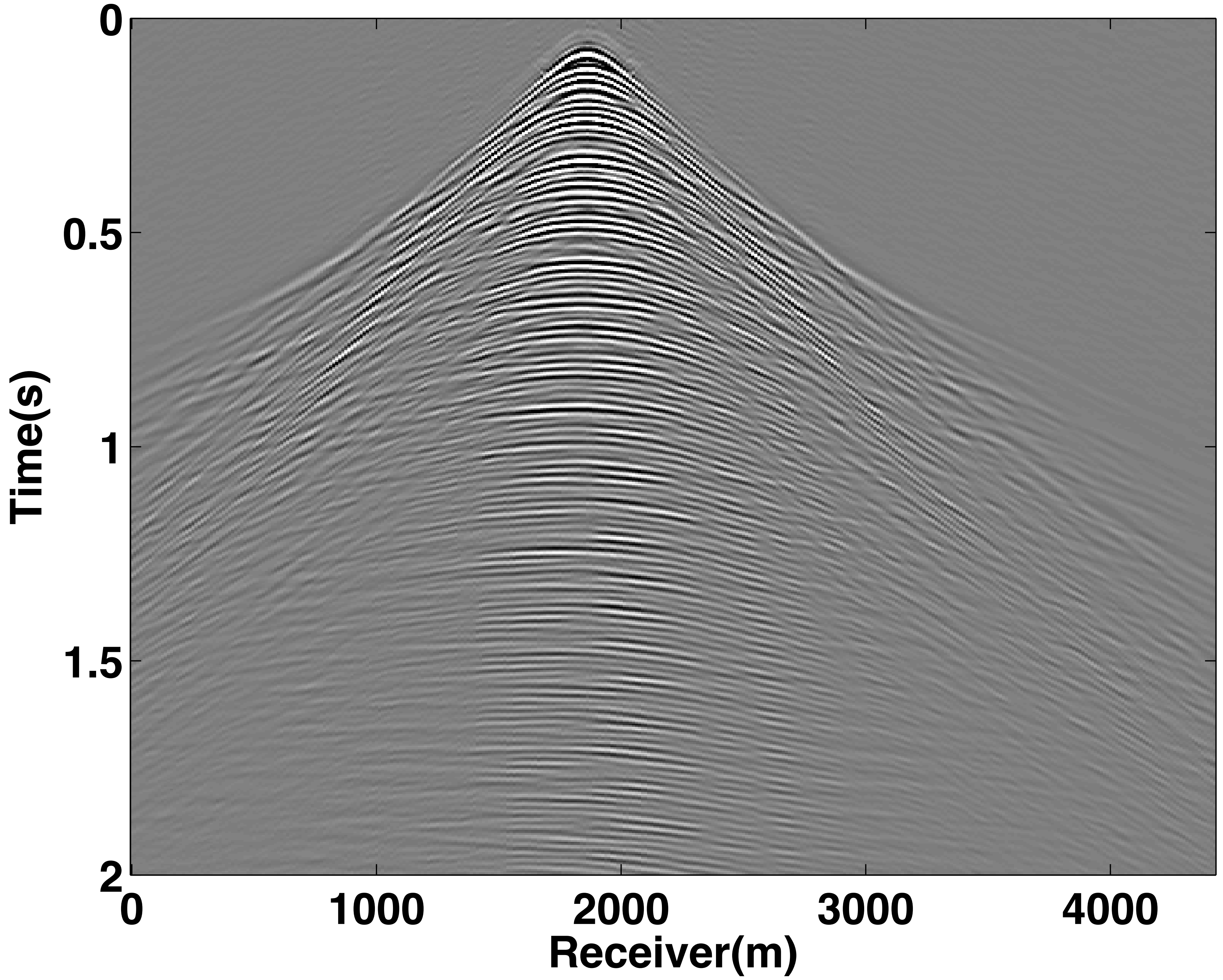
[uniform random subsampling]

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



Ground Truth

SNR = 16.2 dB

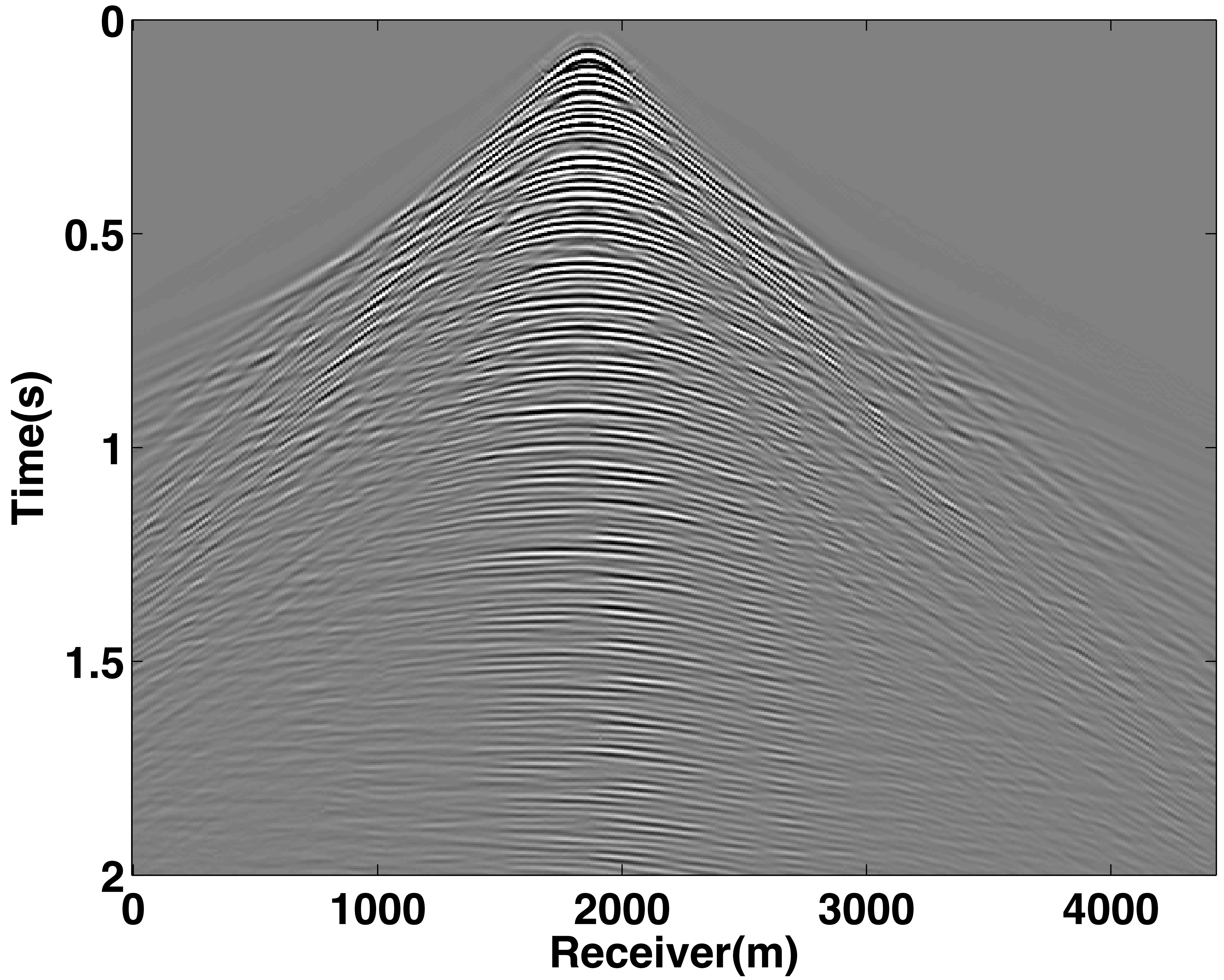


Recovery

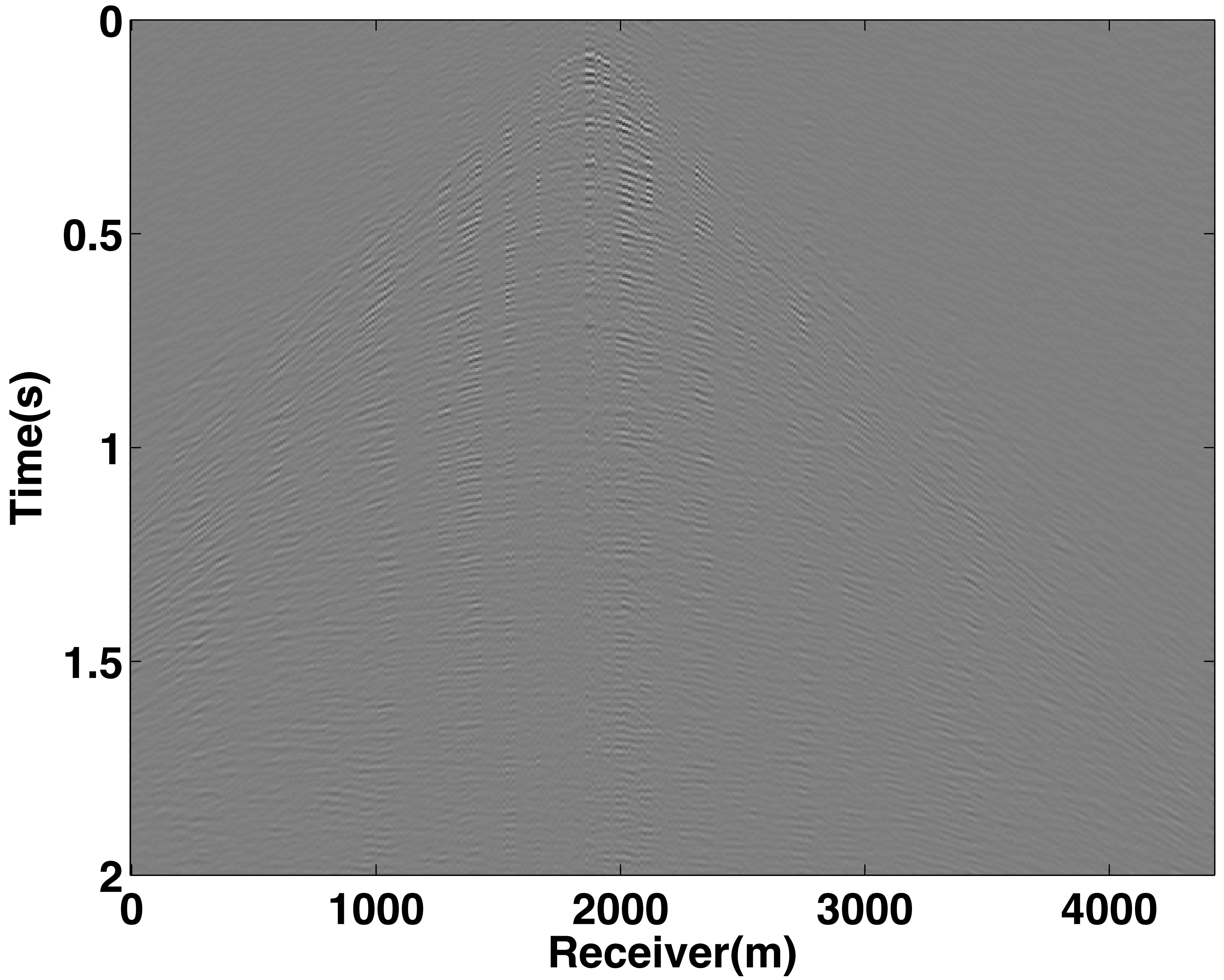
Case I

[uniform random subsampling]

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



Ground Truth

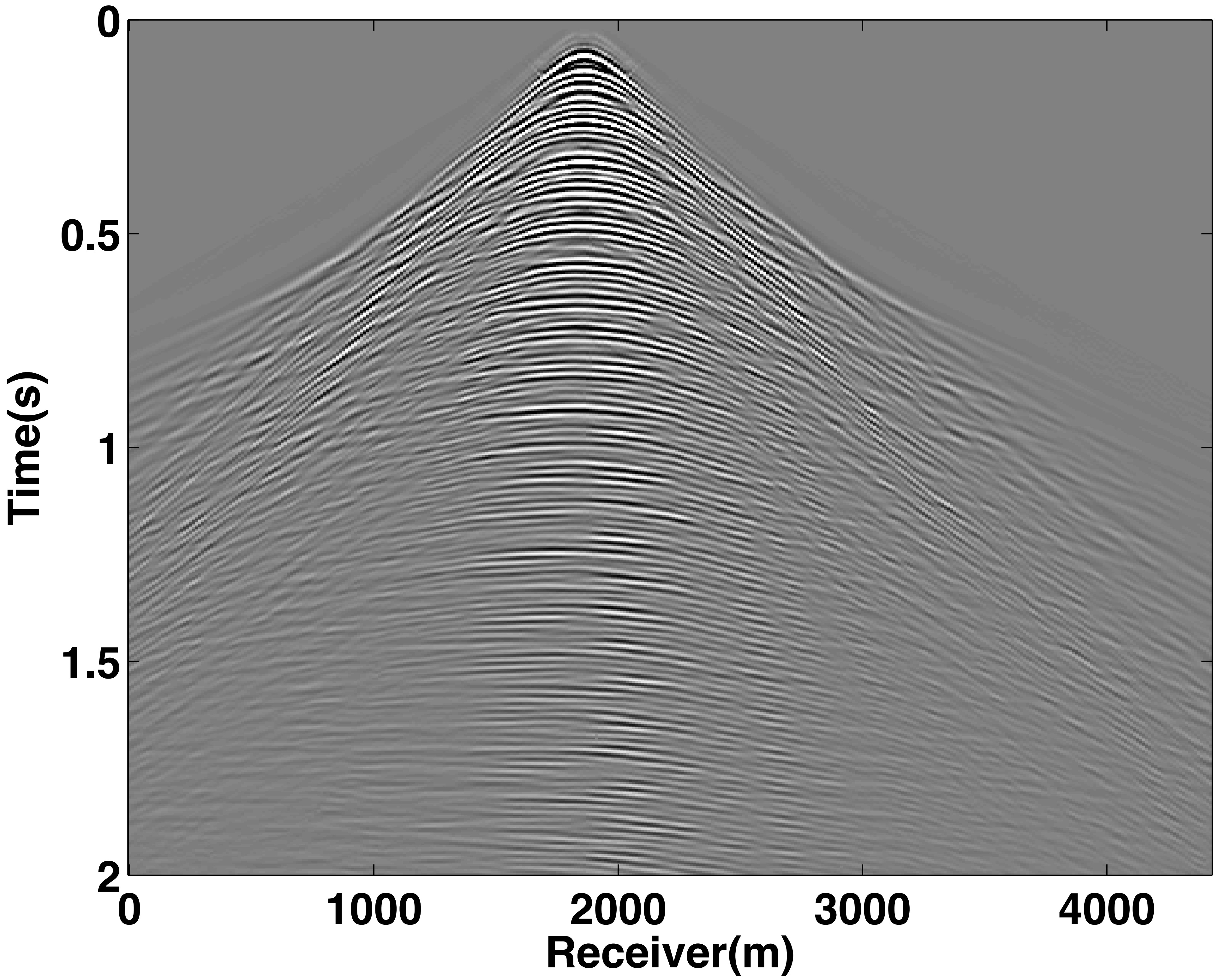


Difference

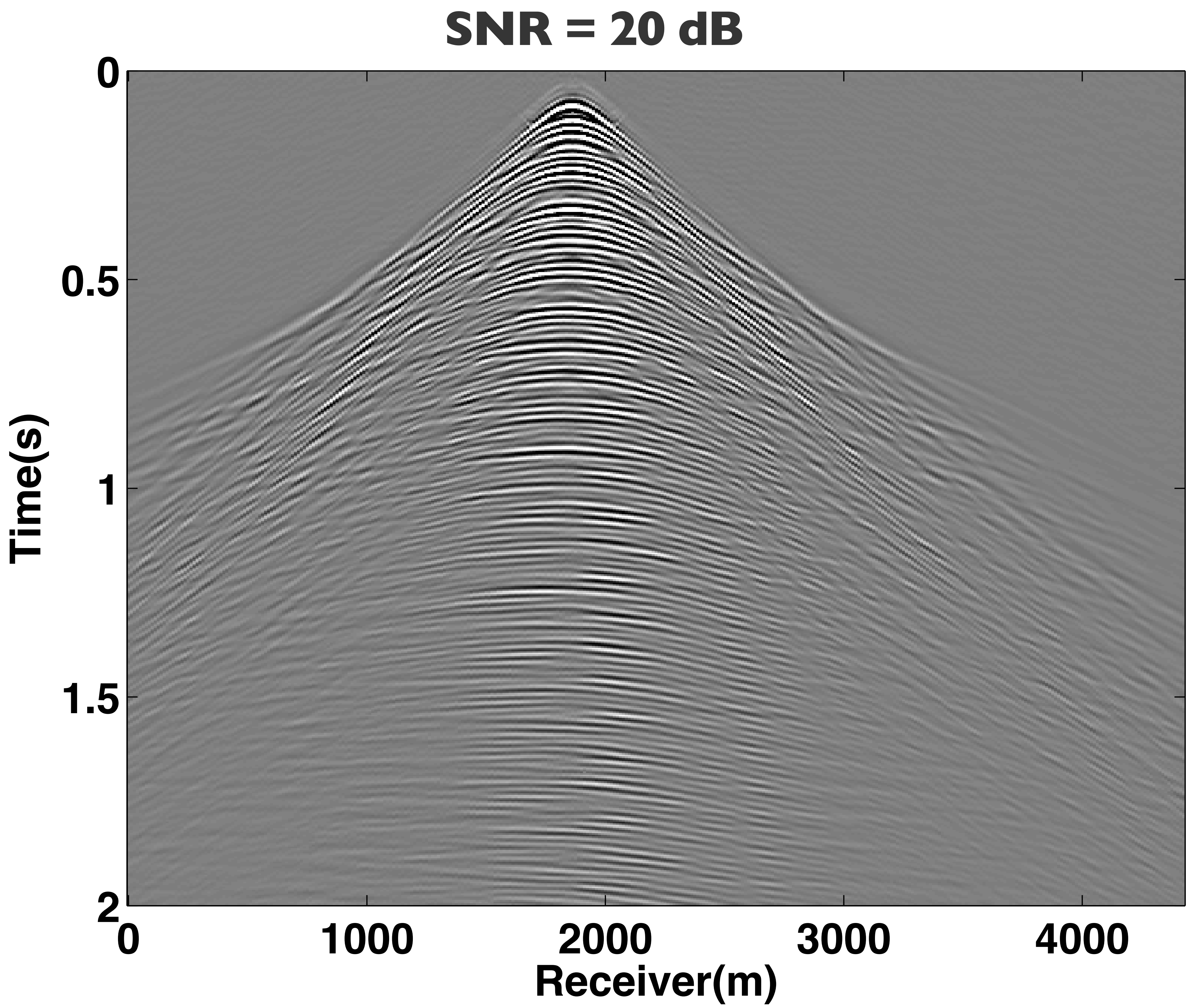
Case 2

[jittered subsampling]

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



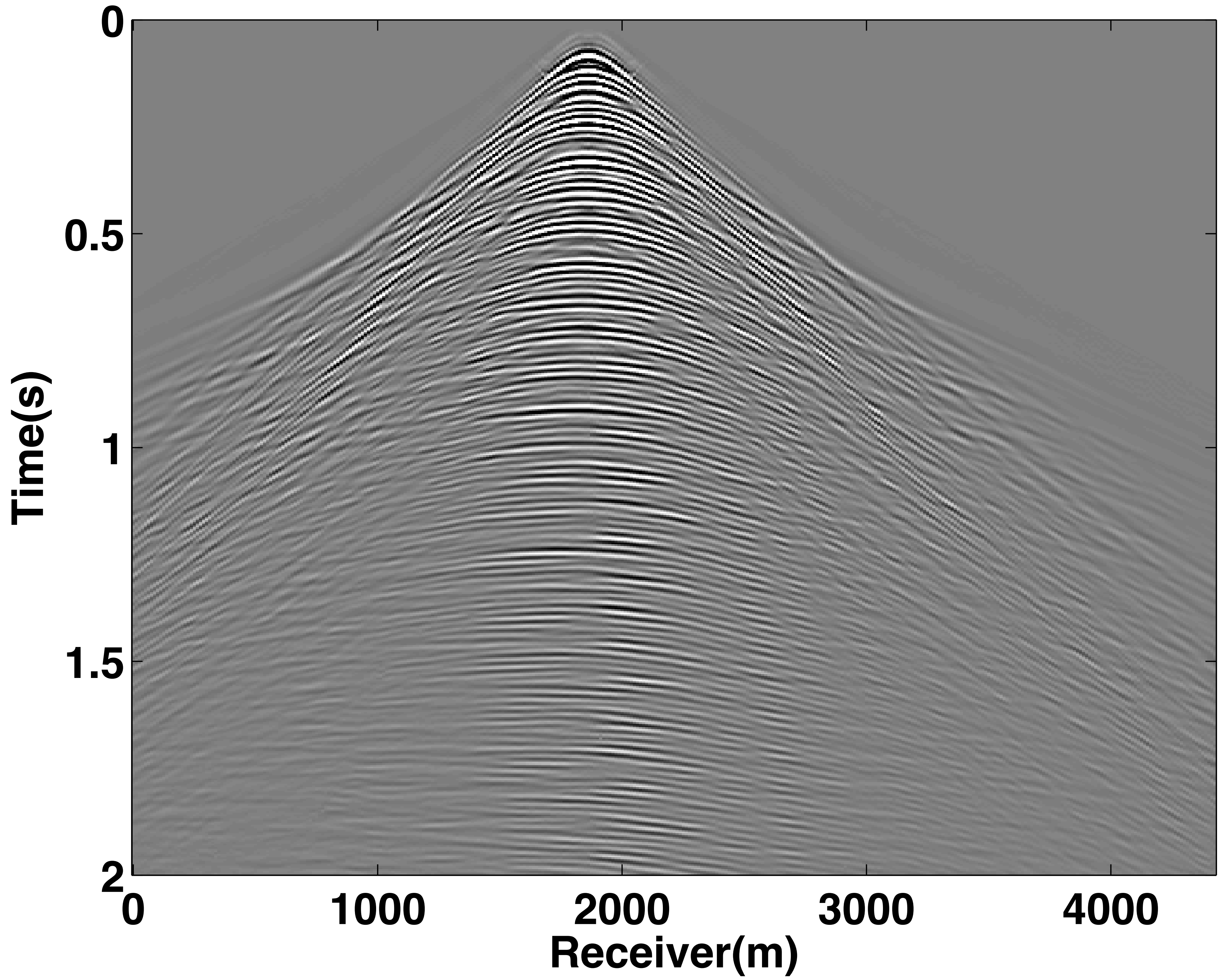
Ground Truth



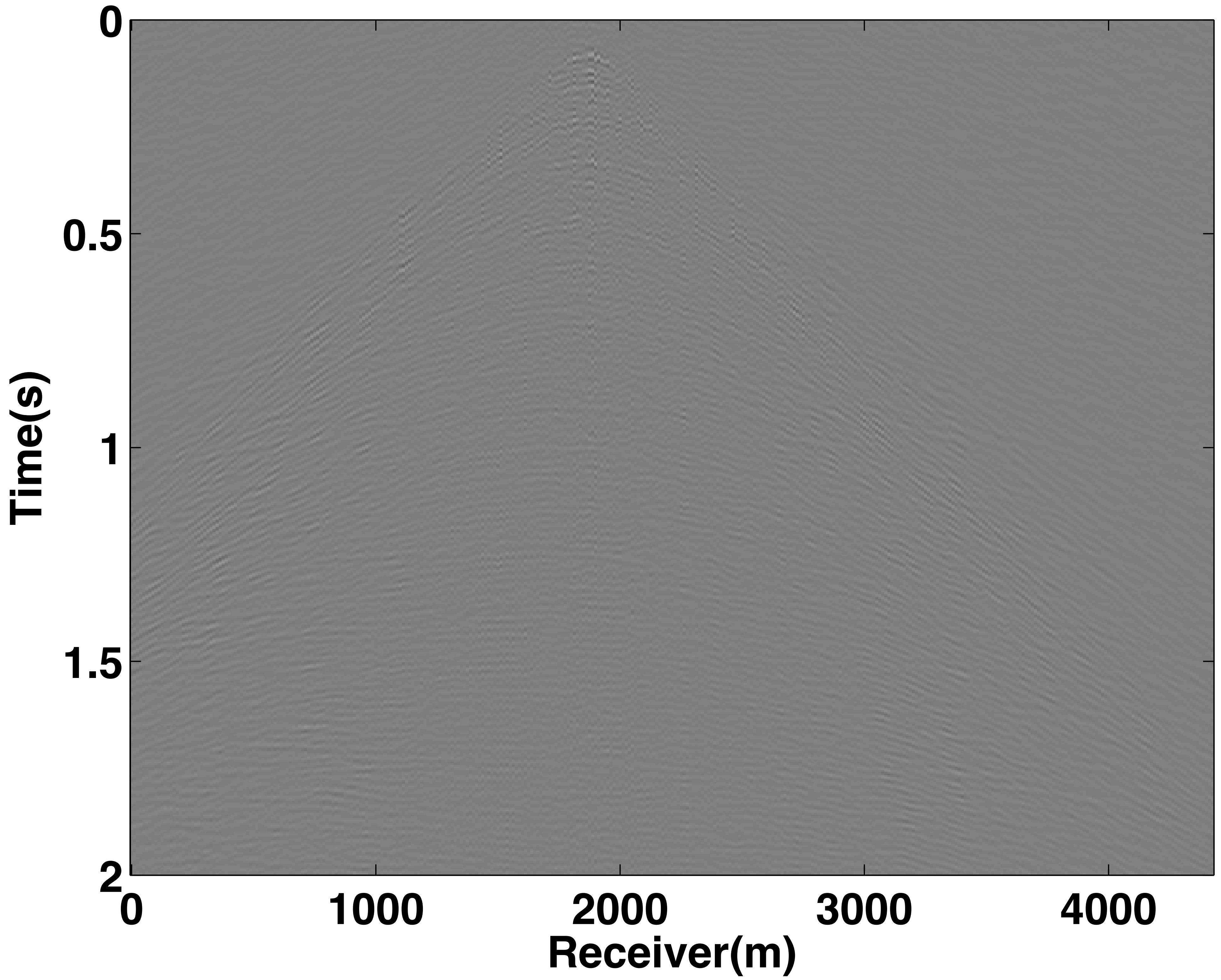
Recovery

Case 2
[jittered subsampling]

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



Ground Truth

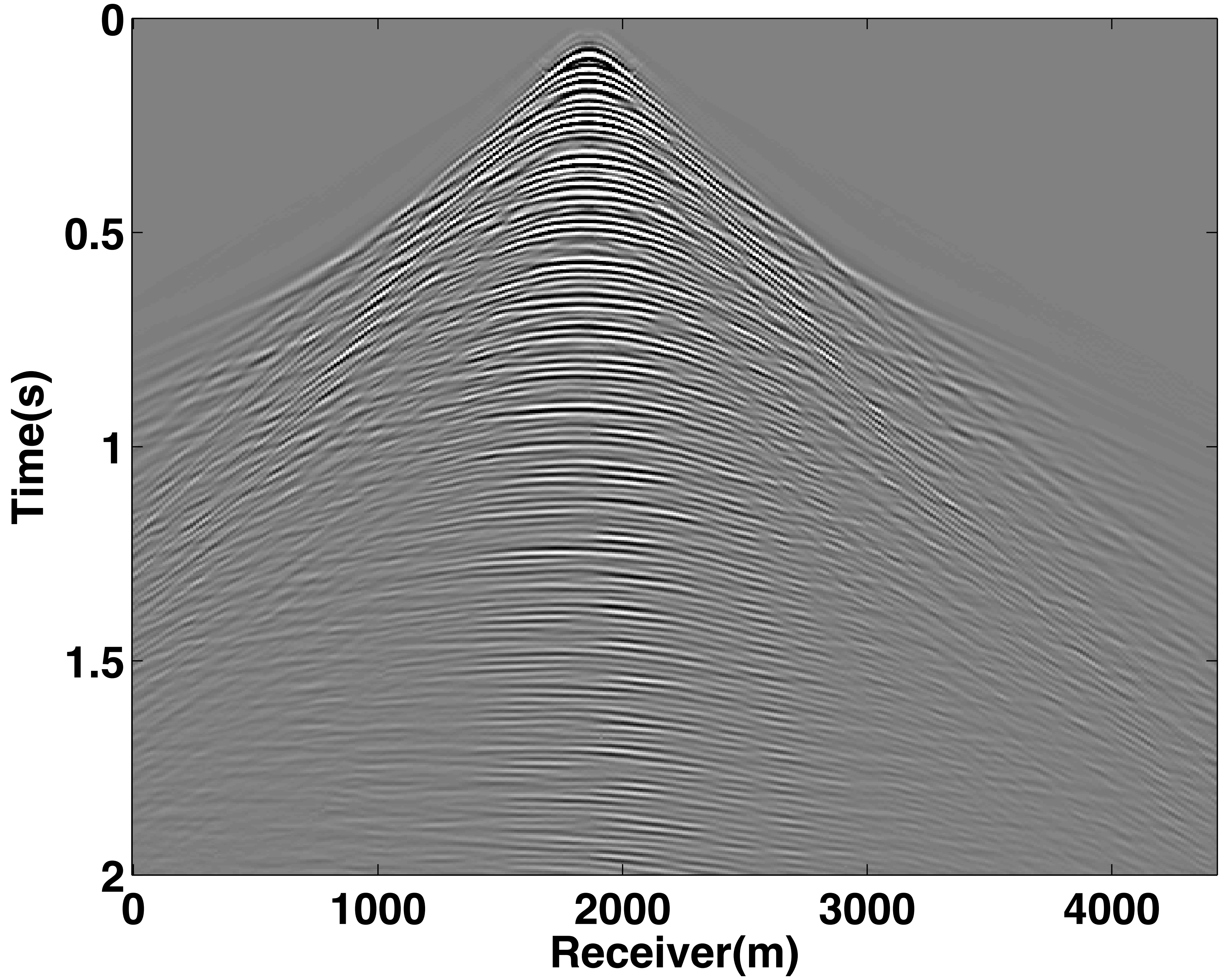


Difference

Case 3

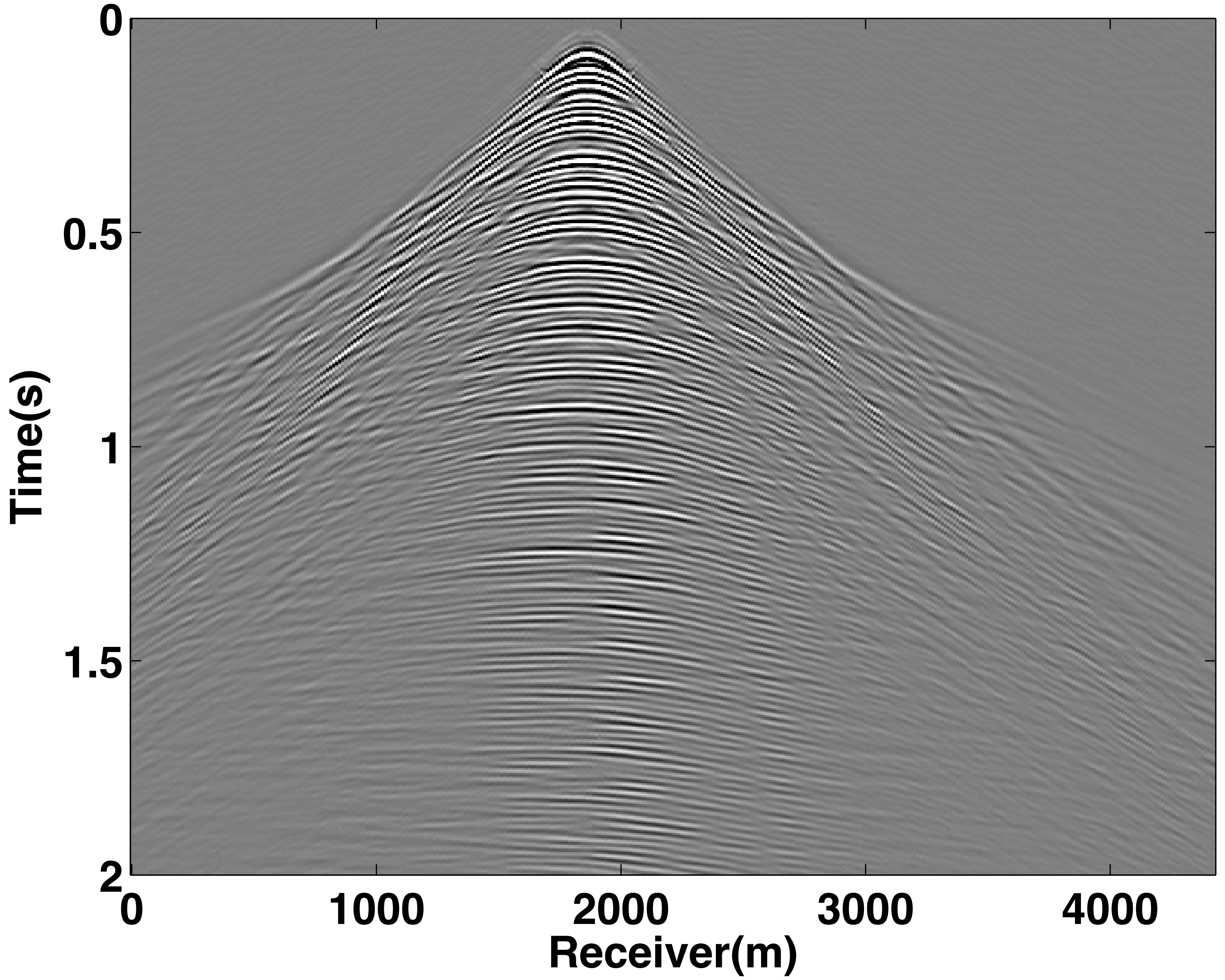
[jittered subsampling + reciprocity]

$$\mathcal{A} = \mathbf{R}\mathbf{M}\frac{(\mathbf{I}+\mathbf{T})}{\mathbf{2}}\mathcal{S}^H$$
$$\mathbf{M} = \mathbf{I}$$



Ground Truth

SNR = 20.6 dB

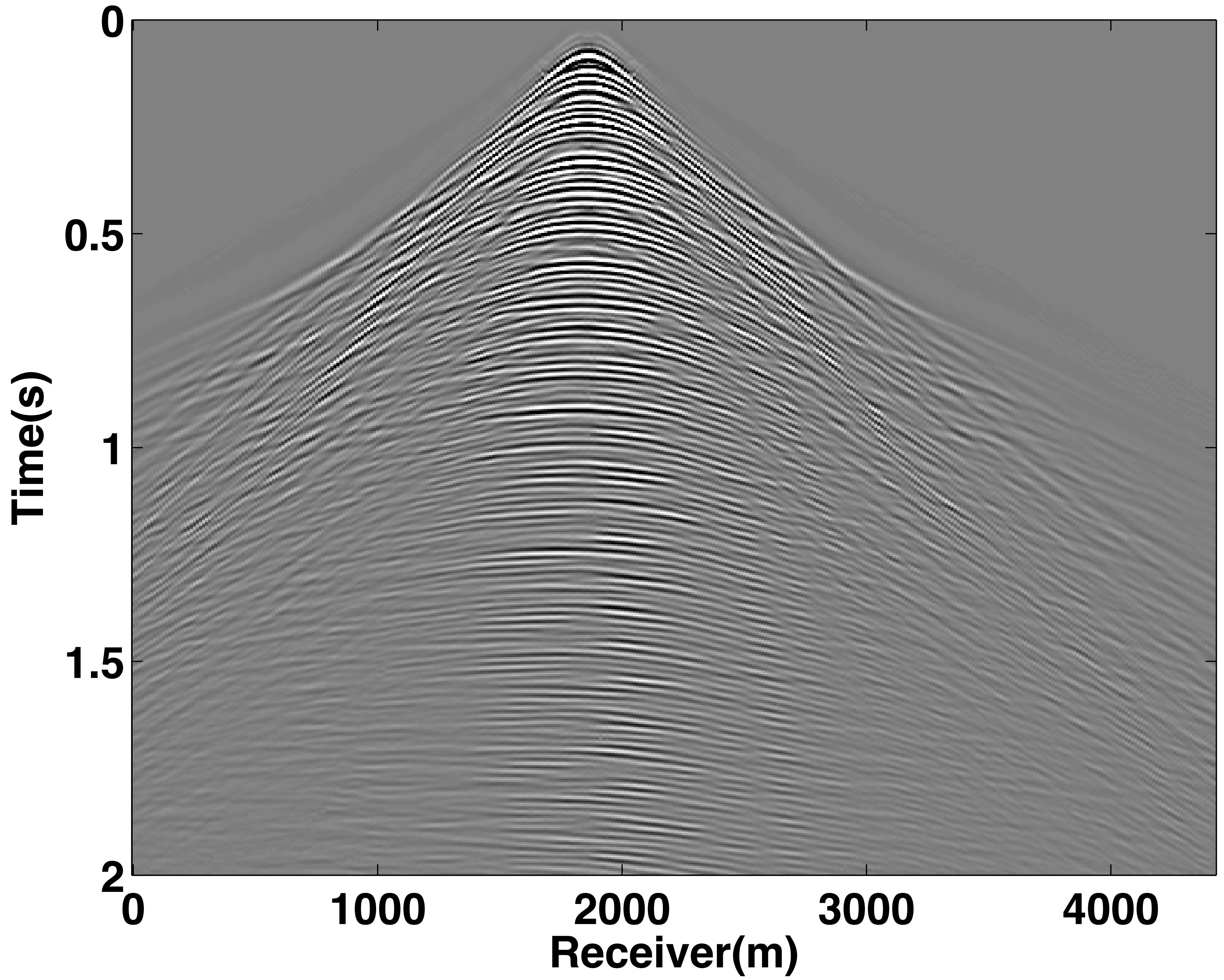


Recovery

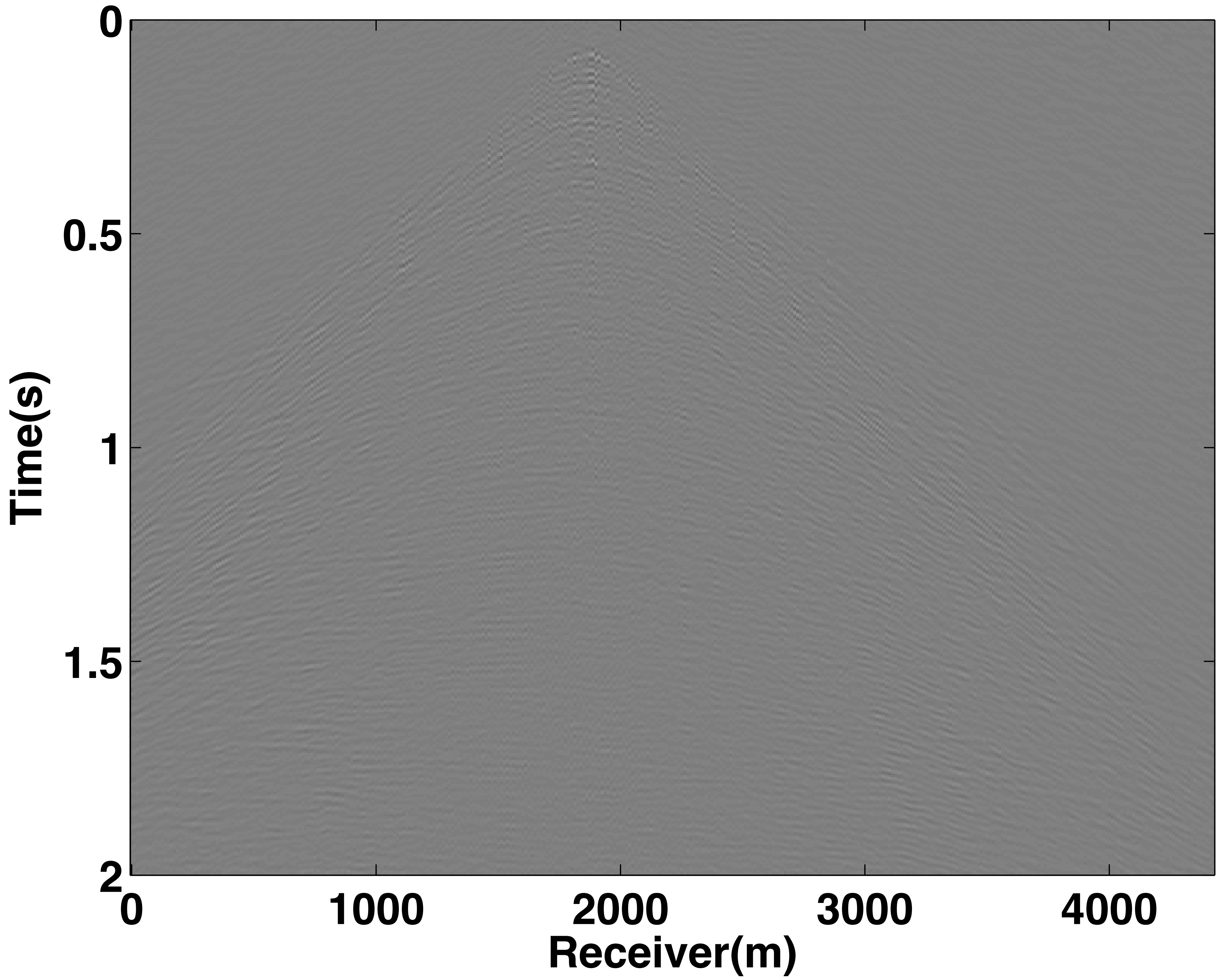
Case 3

[jittered subsampling + reciprocity]

$$\mathcal{A} = \mathbf{R}\mathbf{M}\frac{(\mathbf{I}+\mathbf{T})}{\mathbf{2}}\mathcal{S}^H$$
$$\mathbf{M} = \mathbf{I}$$



Ground Truth



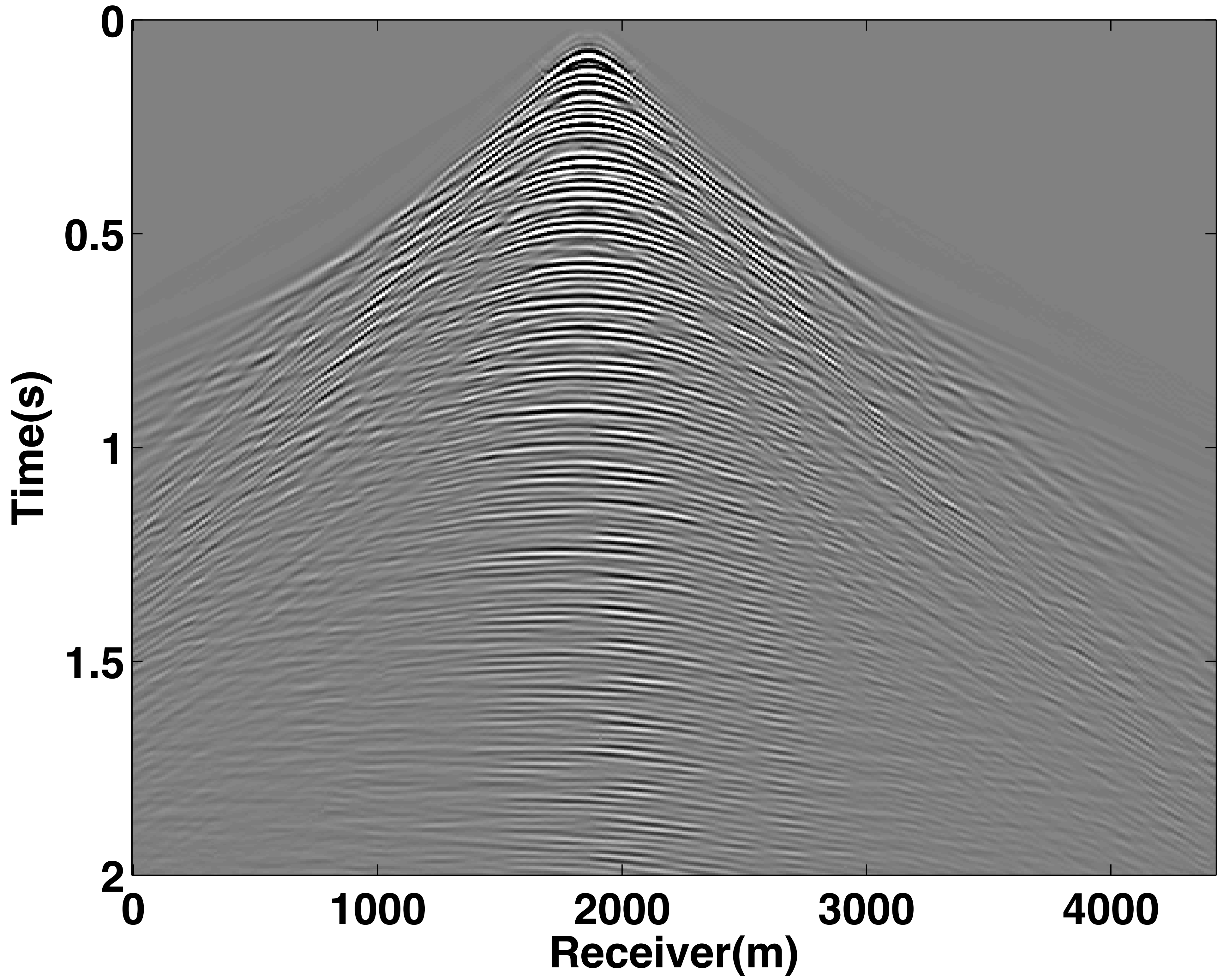
Difference

Case 4

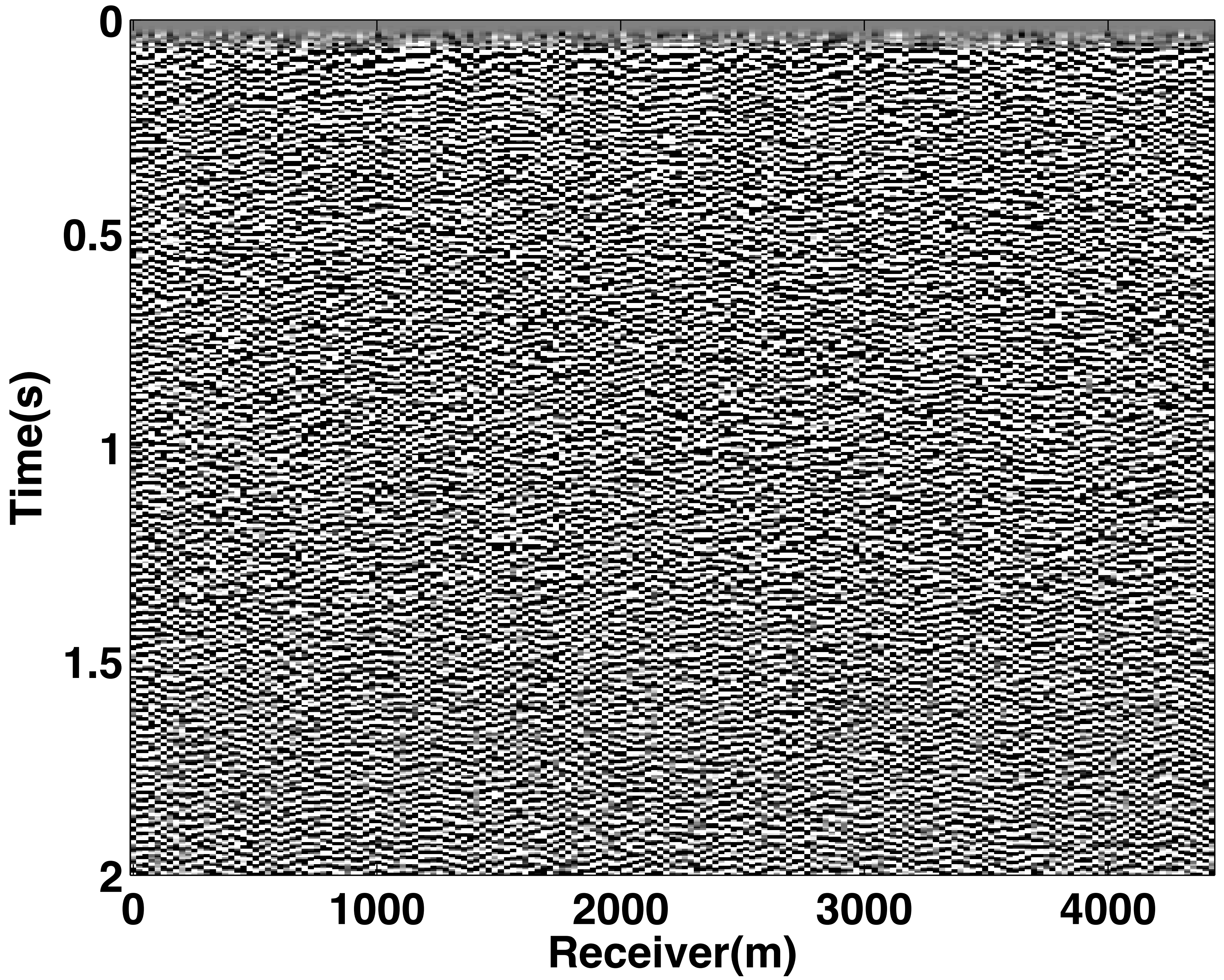
[Simultaneous Source : Land]

$$\mathcal{A} = \mathbf{R}\mathbf{M}\mathcal{S}^H$$

$$\mathbf{M} \stackrel{\text{def}}{=} [\mathbf{I} \otimes \mathbf{diag}(\eta)\mathcal{F}_{\mathbf{s}}{}^{*}\mathbf{diag}(\mathbf{e}^{\mathbf{i}\theta})\mathcal{F}_{\mathbf{s}} \otimes \mathbf{I}]$$ [Herrmann et. al. 2009]

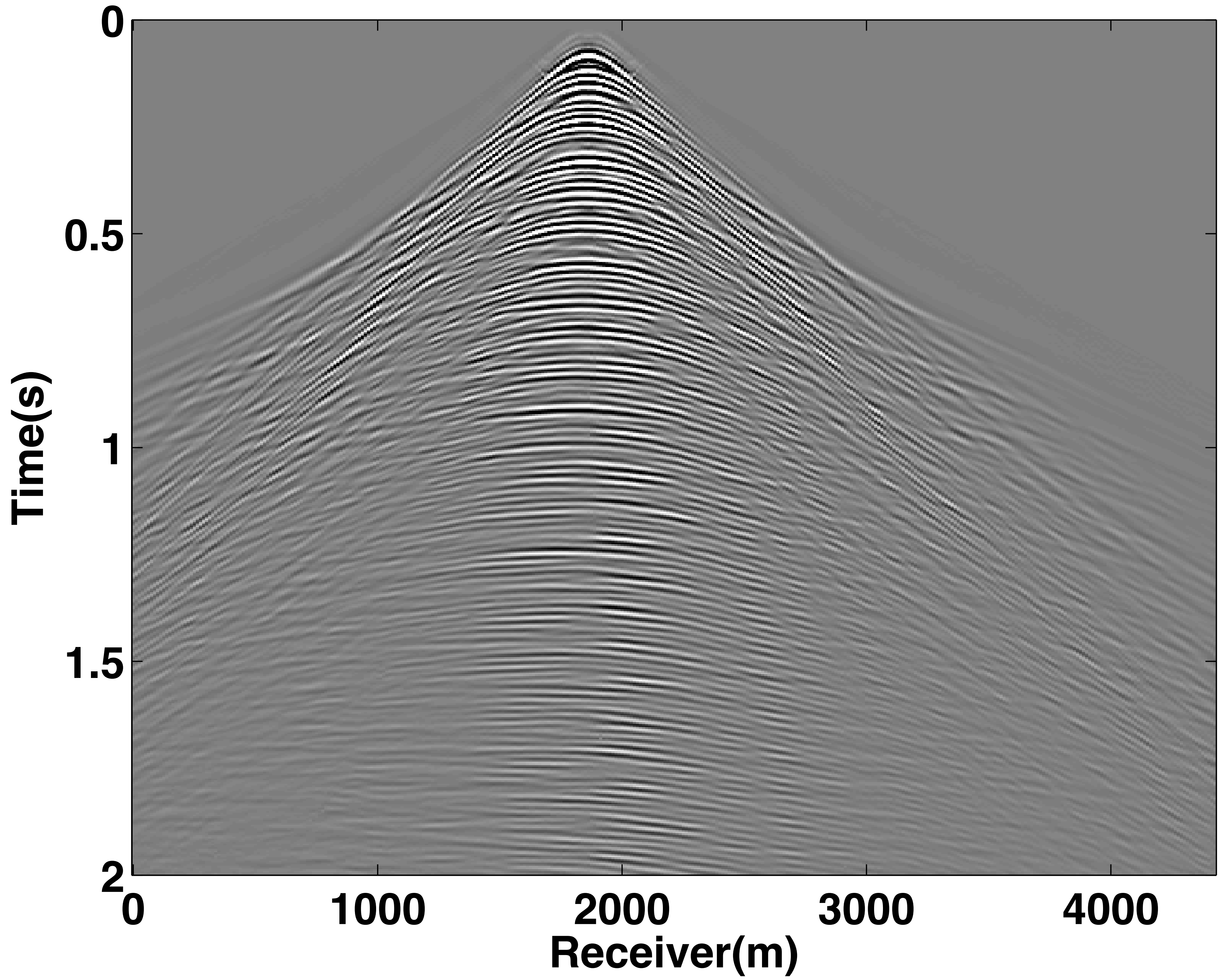


Ground Truth



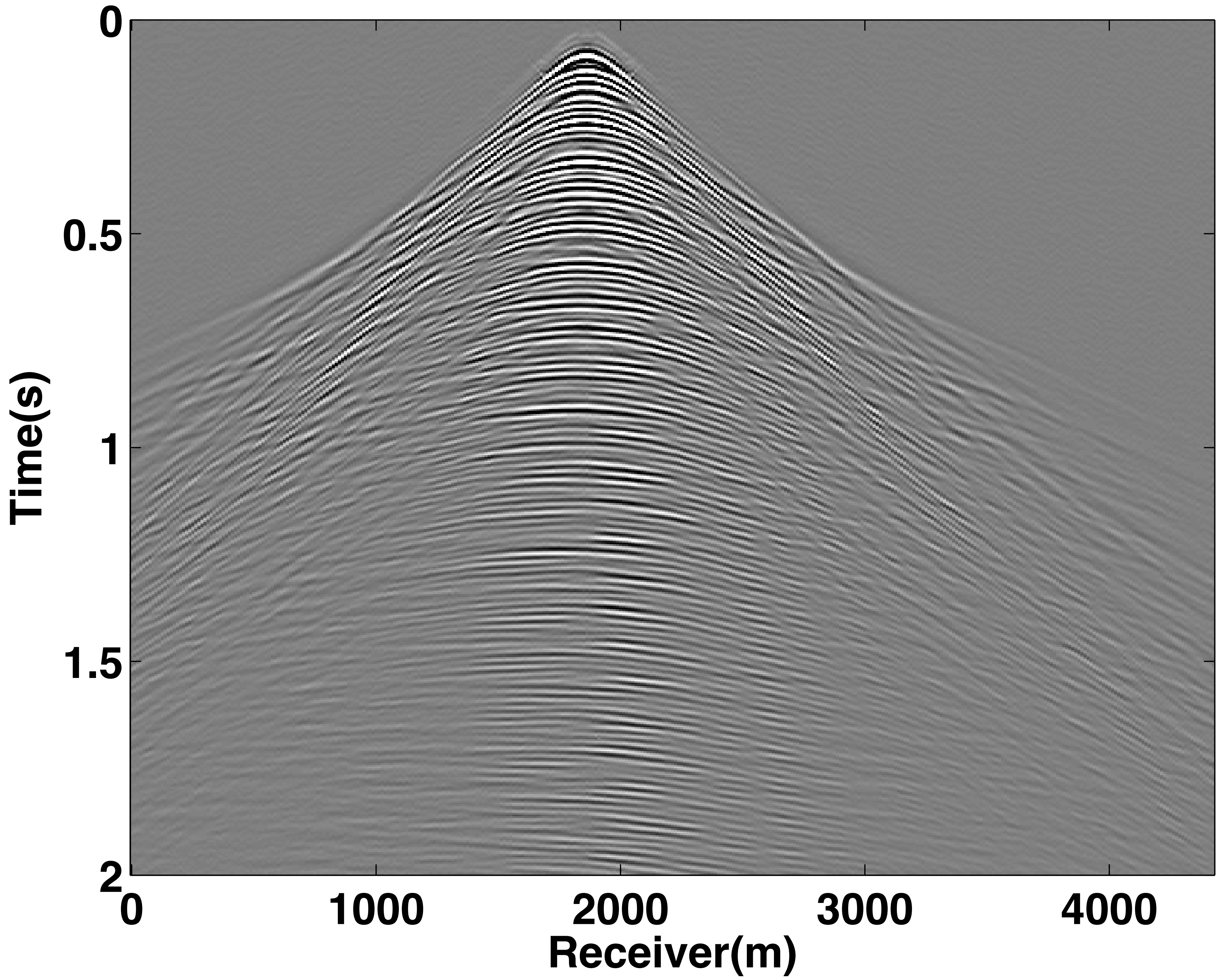
Case 4
[Simultaneous Source : Land]

$$\mathcal{A} = \mathbf{R}\mathbf{M}\mathcal{S}^H$$
$$\mathbf{M} \overset{\text{def}}{=} [\mathbf{I} \otimes \mathbf{diag}(\eta)\mathcal{F}_\mathbf{s}{}^*\mathbf{diag}(\mathbf{e}^{\mathrm{i}\theta})\mathcal{F}_\mathbf{s} \otimes \mathbf{I}]$$ [Herrmann et. al. 2009]



Ground Truth

SNR = 21.2 dB

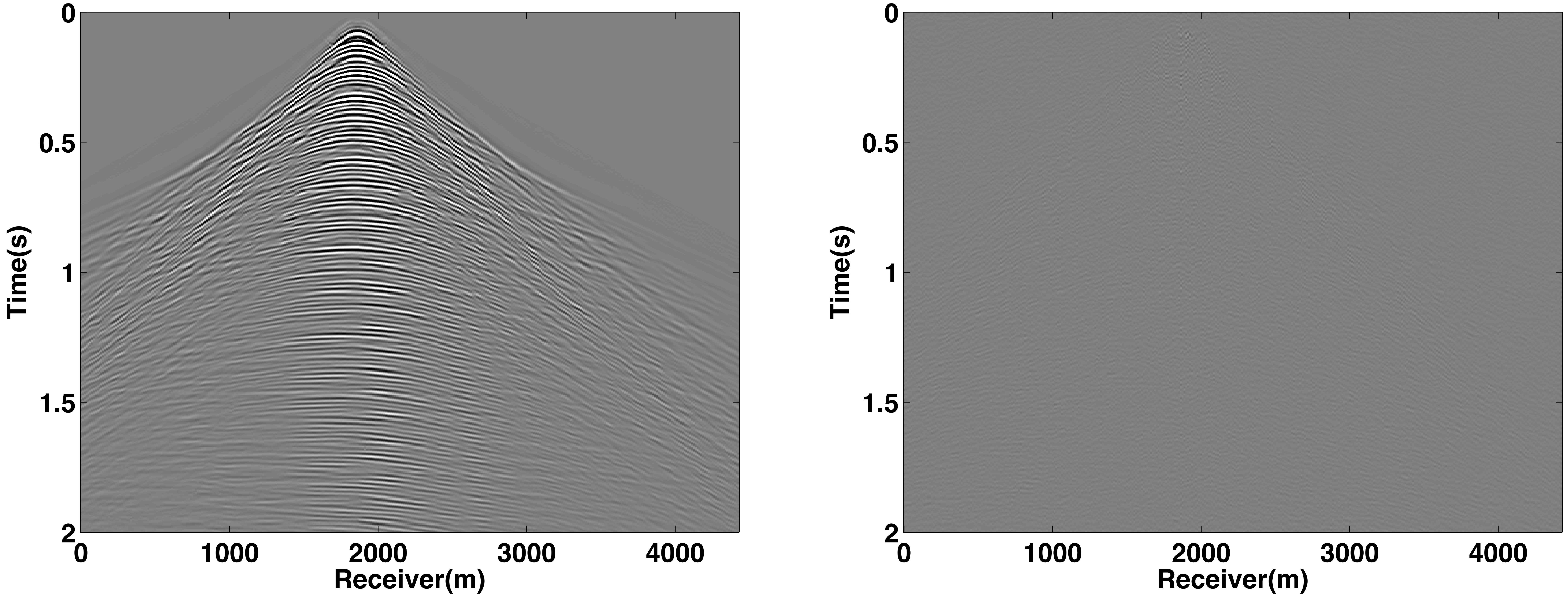


Recovery

Case 4
[Simultaneous Source : Land]

$$\mathcal{A} = \mathbf{R}\mathbf{M}\mathcal{S}^H$$

$$\mathbf{M} \stackrel{\text{def}}{=} [\mathbf{I} \otimes \mathbf{diag}(\eta)\mathcal{F}_{\mathbf{s}}{}^{*}\mathbf{diag}(\mathbf{e}^{\mathbf{i}\theta})\mathcal{F}_{\mathbf{s}} \otimes \mathbf{I}]$$ [Herrmann et. al. 2009]

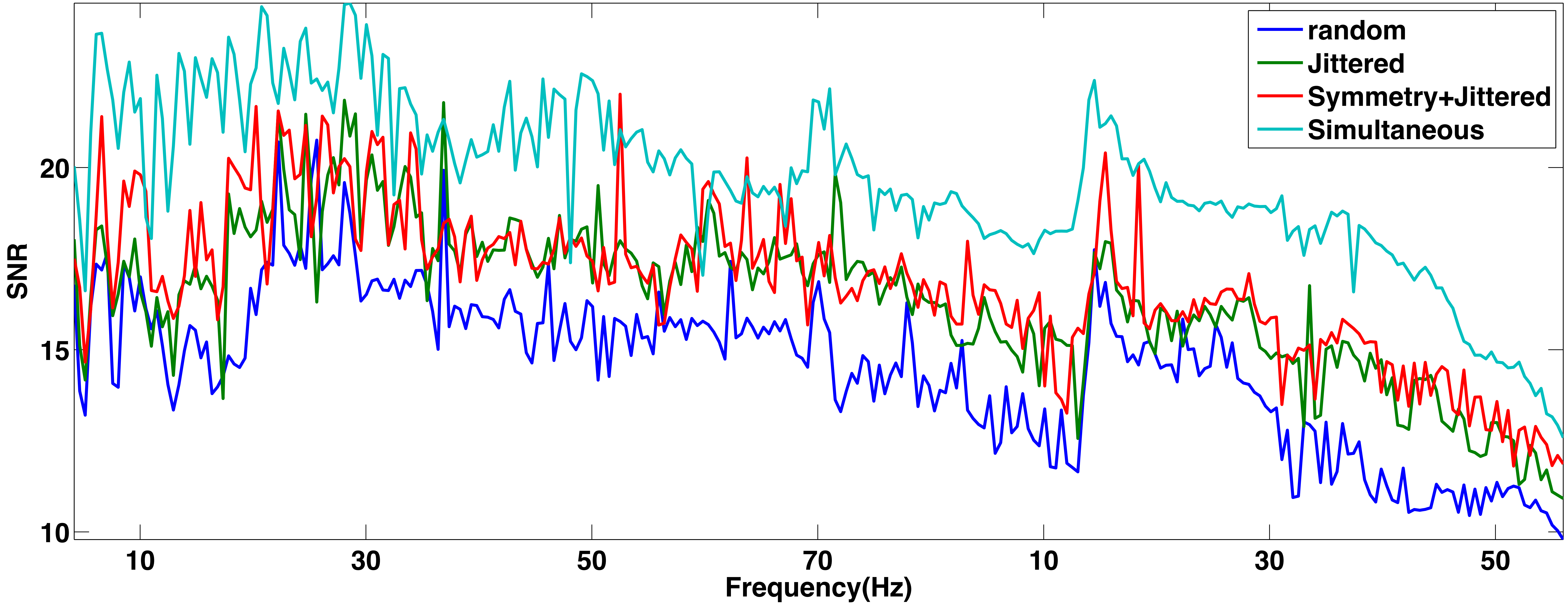


Ground Truth

Difference

Comparison

[uniform random vs jittered vs jittered + reciprocity]



Conclusion

- jittered sampling gives advantage of controlling the gap size
- 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

Future Work

- Incorporate jittered sampling in 3D
- High frequency are *not low-rank* in nature, explore *HSS* (Hierarchical semi-separable representation) in 5D
- weighted low-rank interpolation

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, Statoil, Total SA, WesternGeco, and Woodside.