Latest developments in seismic data recovery

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

SLIM

Seismic Laboratory for Imaging and Modeling
the University of British Columbia

Premise

Signals in nature including seismic wavefields & sedimentary basins *exhibit* some sort of *structure*

- transform-domain sparsity
- low-rank property

Come up with *new* cost-*effective randomized* sampling *strategies*

- for *land* with *randomized* arrays or *simultaneous* sweeps
- for *marine* with *randomized* time-*dithered simultaneous* sources

Compressive sensing

Compressive sensing *delivers* on this *premise* by coming up

- a *rigorous* theory with *recovery* guarantees
- *constructive* recovery algorithms by *convex* optimization

SINBAD is a world-leader in *adapting* compressive sensing

- seismic-data acquisition (land & marine)
- seismic-data processing (RTM & FWI)

This talk

Weighted one-norm minimization (Felix):

- theoretical *recovery* results
- extension to 3D seismic
- recovery based on curvelet-domain *sparsity* promotion

Nuclear-norm minimization (Sasha):

- *new* solver using factor approximation of *nuclear* norm
- recovery based on *low-rank* promotion

Trace interpolation

Involves an *underdetermined* inverse problem.

Given a seismic line of N_s sources, N_r receivers and N_t time samples, arranged into a vector $\mathbf{f}$ of length $N = N_s N_r N_t$.

Recover a sparse approximation $\tilde{\mathbf{f}}$ from *irregularly* random sampled measurements

$$\mathbf{b} = \mathbf{RMf}$$

where $\mathbf{RM}$ is the sampling operator.

Today's focus

Move to missing trace interpolation for 3D seismic

Work on *frequency* slices with 2D curvelet transform

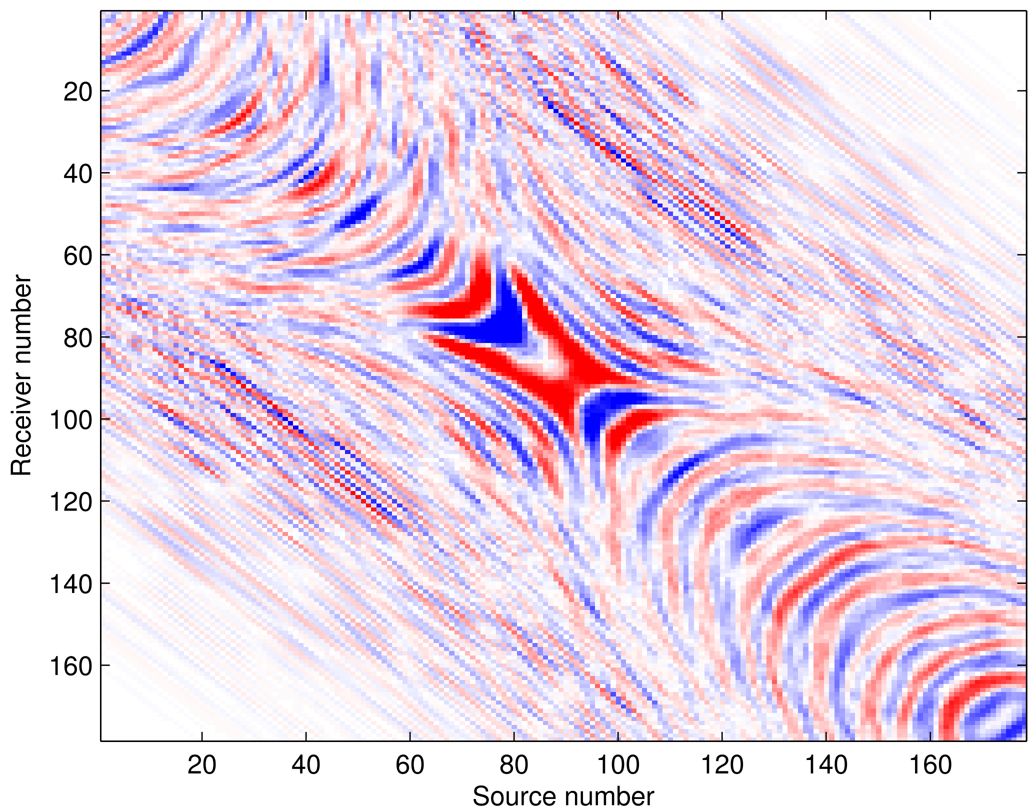
- exploit *correlations* in the support
- use *weighted* one-norm minimization

Opens the way to do 4D interpolation

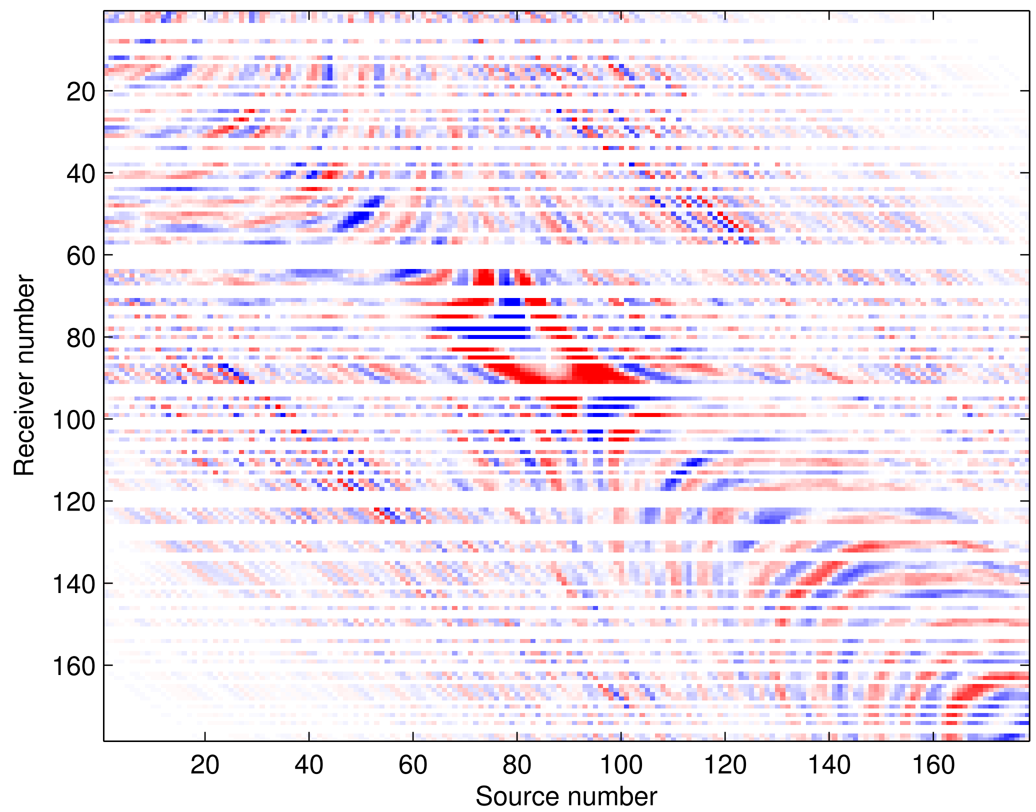
Subsampled traces

Example of a time slice with missing receivers.

$$\mathbf{f}$$



$$\mathbf{b} = \mathbf{RMf}$$



CS and random trace interpolation

When data are *randomly* subsampled & admit a *sparse* representation, the trace-interpolation problem falls under the CS paradigm.

Corresponds to finding the sparsest representation $\tilde{\mathbf{x}}$ of the data in some domain $\mathbf{S}$, by solving

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{u}\in\mathbb{C}^P} \|\mathbf{u}\|_1 \text{ subject to } \mathbf{b} = \mathbf{A}\mathbf{u}$$

where $\mathbf{A} := \mathbf{RMS^H}$.

Compressed Sensing

Candes, Romberg, and Tao; and Donoho proved that CS recovery is stable to model mismatch and robust to noise.

Theorem (CRT'06):

Suppose there exists an $a > 1$, and that A has the RIP with $\delta_{(a+1)k} < \frac{a-1}{a+1}$. Then the sparse approximation $\tilde{x}$ of x can be obtained from the solution to the ℓ_1 minimization problem and obeys

$$\|\tilde{x} - x\|_2 \leq C_0 \epsilon + C_1 \frac{\|x - x_k\|_1}{\sqrt{k}}.$$

* Remark: If x is k-sparse, then the recovery is exact.

CS implications

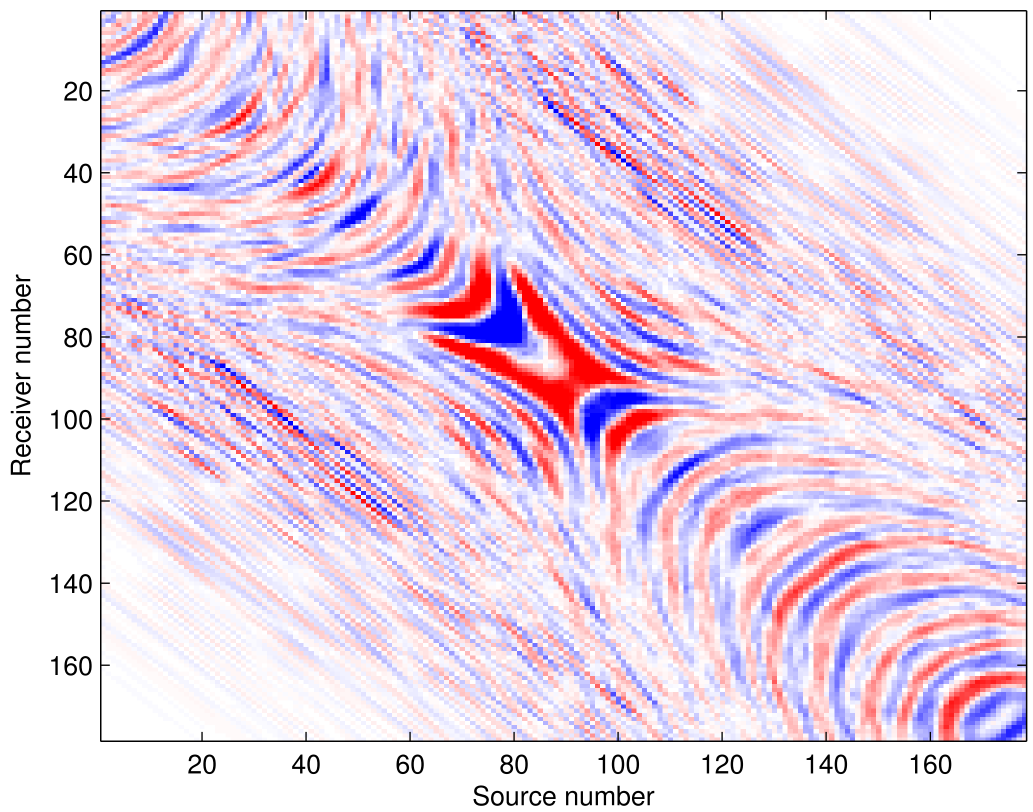
For nice measurement matrices $\mathbf{A}$, CS guarantees exact recovery for signals $\mathbf{f}$ that are *strictly* sparse

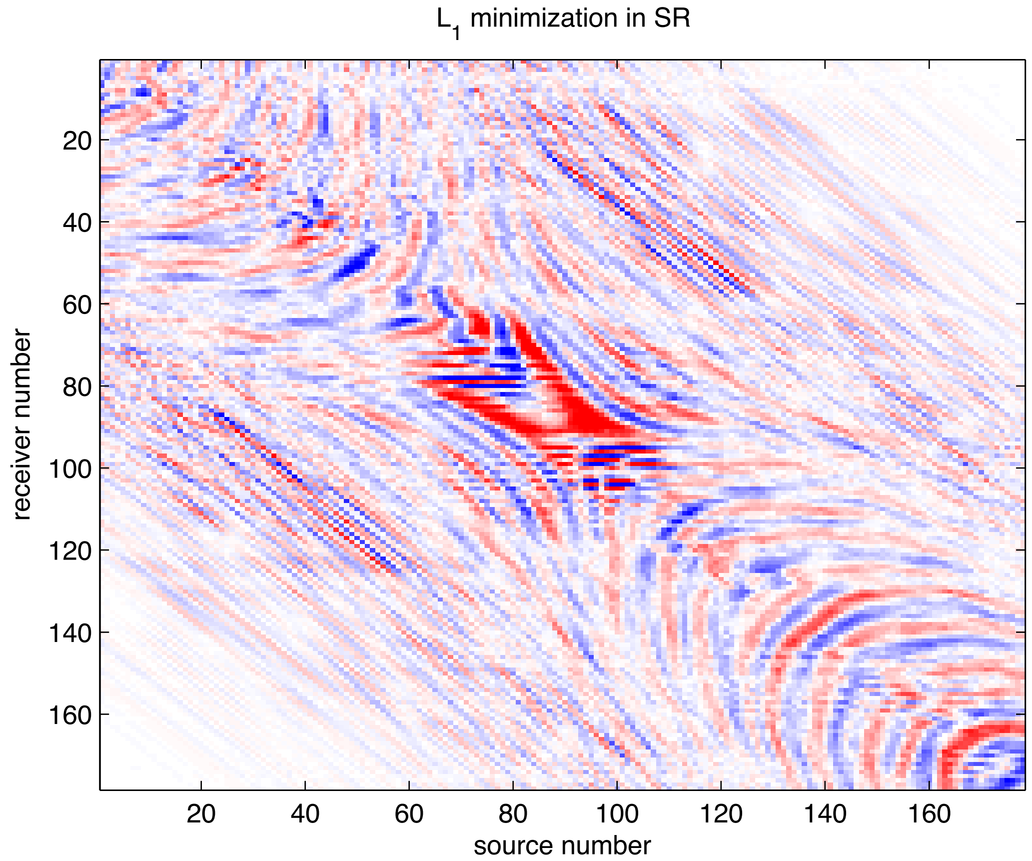
- more sparse (fewer nonzeros) => better recovery
- more randomization (better RIP) => better recovery

Extends to *compressible* signals

- more compressible => better recovery
- *recovery* akin *nonlinear* approximation

CS interpolation (time-slice)





What more can be done?

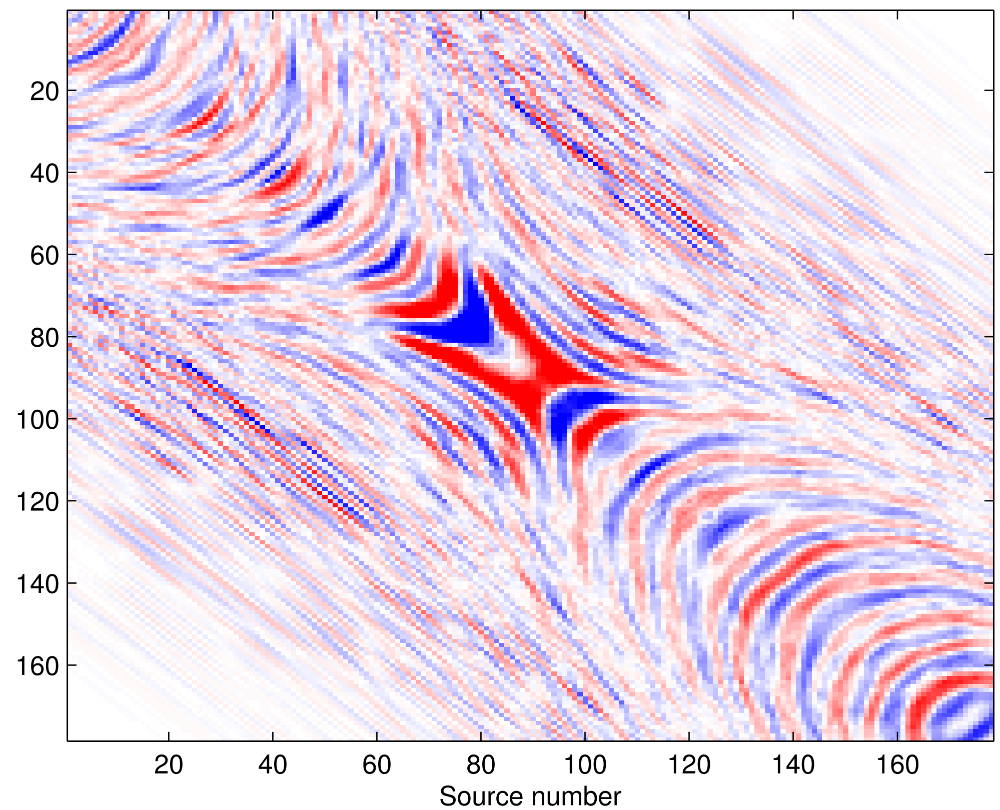
Improve the RIP of the measurement matrix $\mathbf{A}$

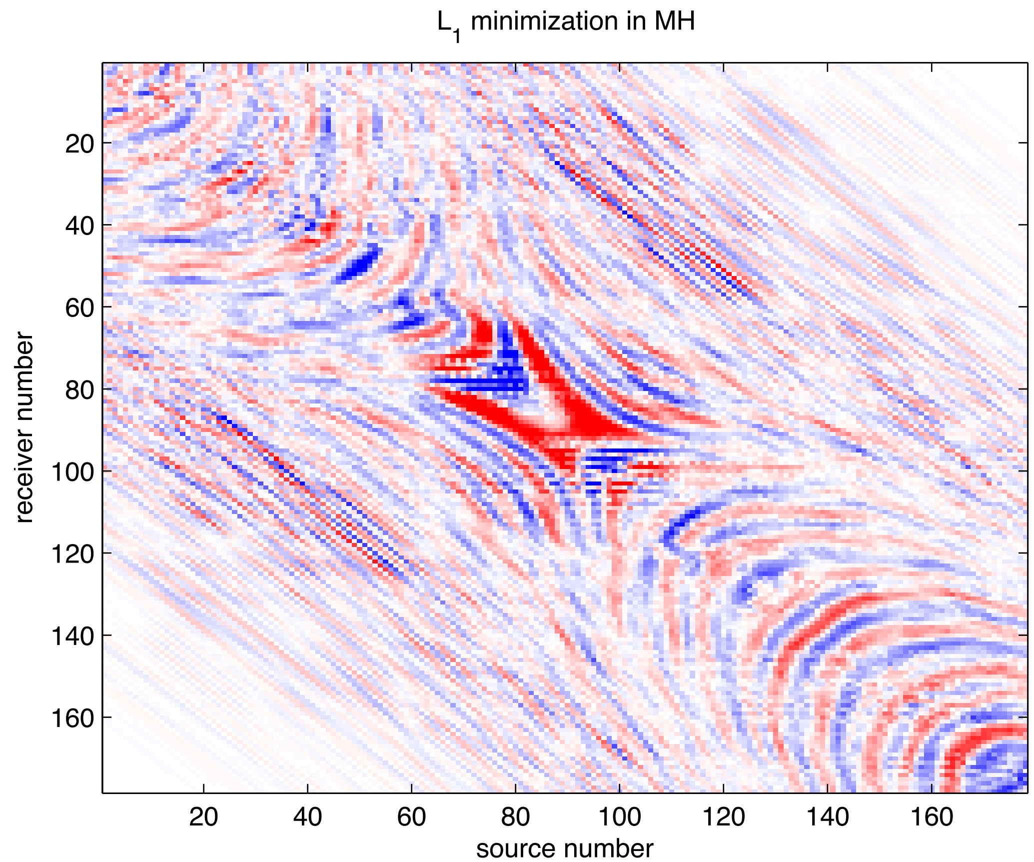
- perform recovery in *other* domains
  - e.g. midpoint-offset, time-midpoint,...

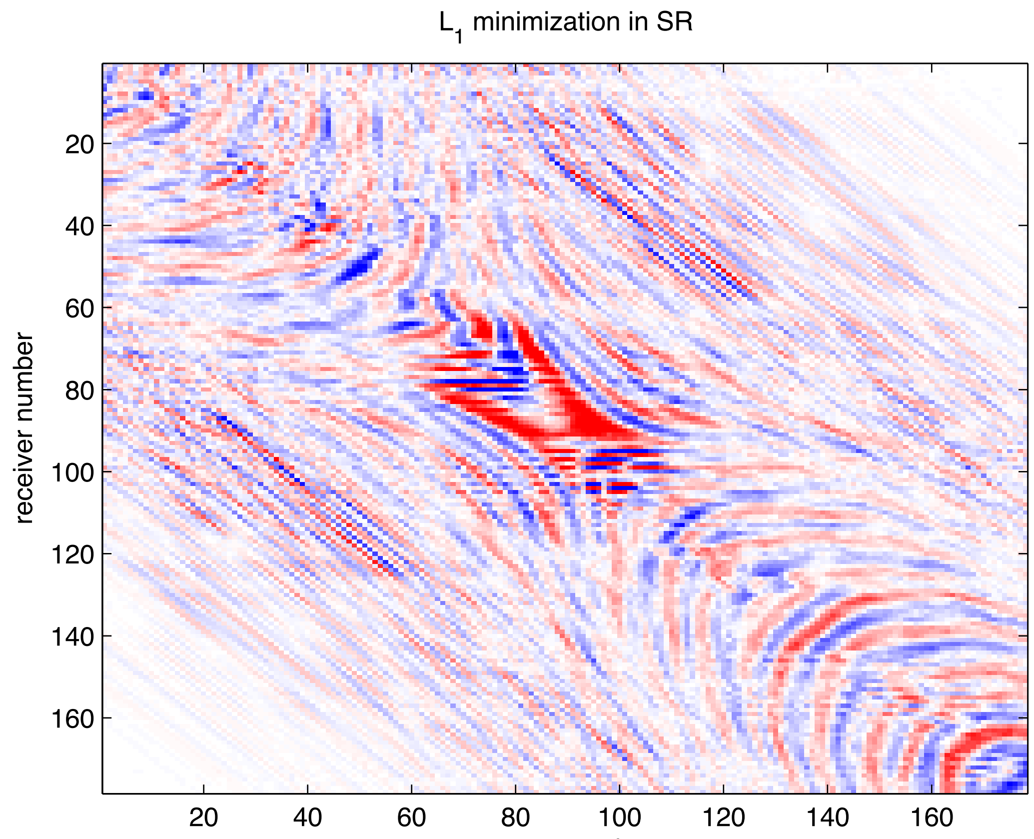
Incorporate *additional* information in the *recovery*

- *frequency* slices of seismic lines are *highly* correlated
  - incorporate *prior support* information in the *recovery*

CS in MH domain







Correlations in seismic data

Frequency slices exhibit considerable *structure* that is shared between *adjacent* frequencies.

This *shared* structure translates into a *high* degree of *overlap* in the *support* sets of curvelet coefficients.

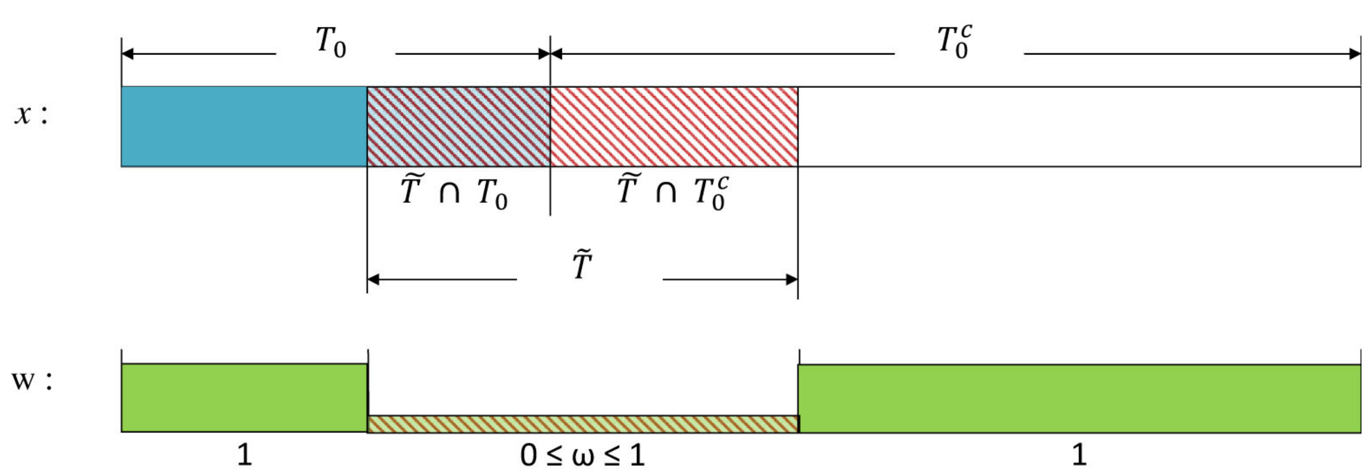
CS
with prior *support* information

Mansour et al. proposed *weighted* one-norm minimization to incorporate *prior support* information.

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{u}} \ \|\mathbf{u}\|_{1,\mathbf{w}} \text{ subject to } \|\mathbf{A}\mathbf{u} - \mathbf{b}\|_2 \leq \epsilon$$

where $\|\mathbf{u}\|_{1,\mathbf{w}} := \sum_i \mathrm{w}_i |u_i|$ is the weighted ℓ_1 norm, and the weights are assigned such that

$$\mathrm{w}_i = \begin{cases} 1, & i \in \widetilde{T}^c, \\ \omega, & i \in \widetilde{T}. \end{cases}$$



CS
with prior *support* information

Let T_0 be the support of x_k and given a *support* estimate $\widetilde{T}$ of size k and *accuracy*

$$\alpha = \frac{|\widetilde{T} \cap T_0|}{|T_0|}$$

Theorem (FMSY'12)

Suppose there exists an $a > 1$, and that A has the RIP with $\delta_{(a+1)k} < \frac{a-\gamma^2}{a+\gamma^2}$. Then the sparse approximation $\tilde{x}$ of x can be obtained from the solution to the weighted ℓ_1 minimization problem and obeys

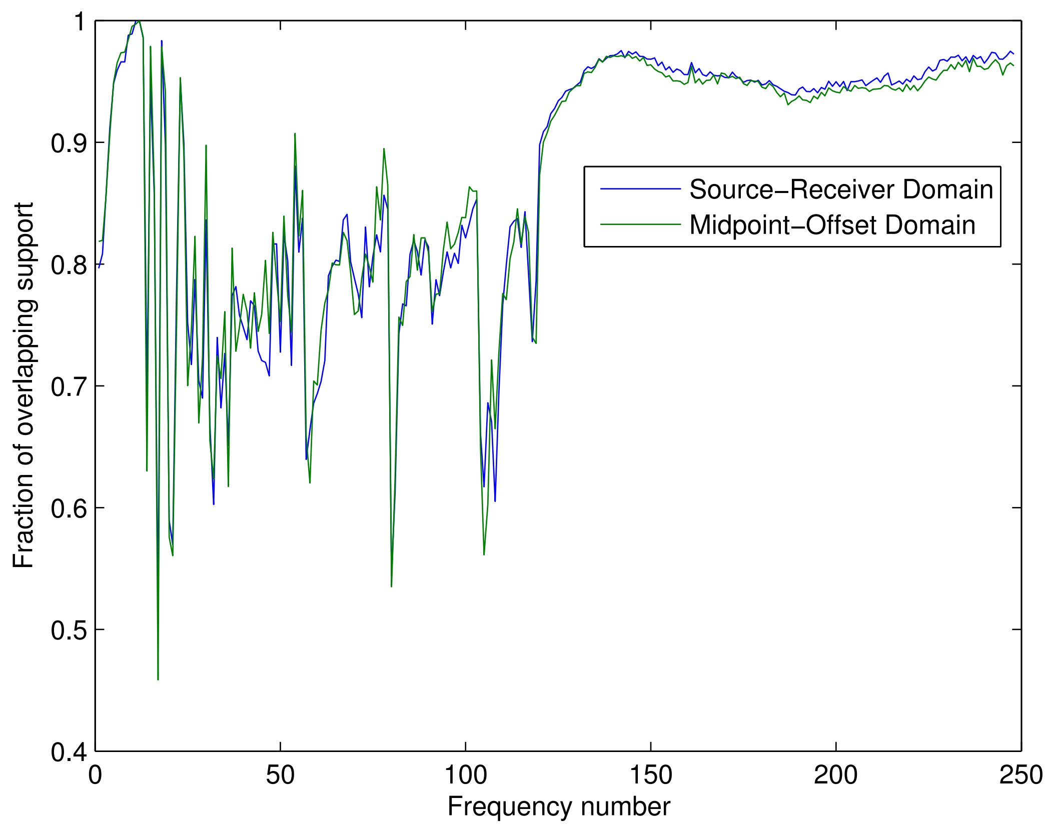
$$\|\tilde{x} - x\|_2 \leq C_0(\gamma)\epsilon + C_1(\gamma)\frac{\omega\|x_{T_0^c}\|_1 + (1-\omega)\|x_{\widetilde{T}^c \cap T_0^c}\|_1}{\sqrt{k}}.$$

Note: $\gamma = \omega + (1-\omega)\sqrt{2-2\alpha}$, where $\alpha = \frac{|\widetilde{T} \cap T_0|}{|T_0|}$.

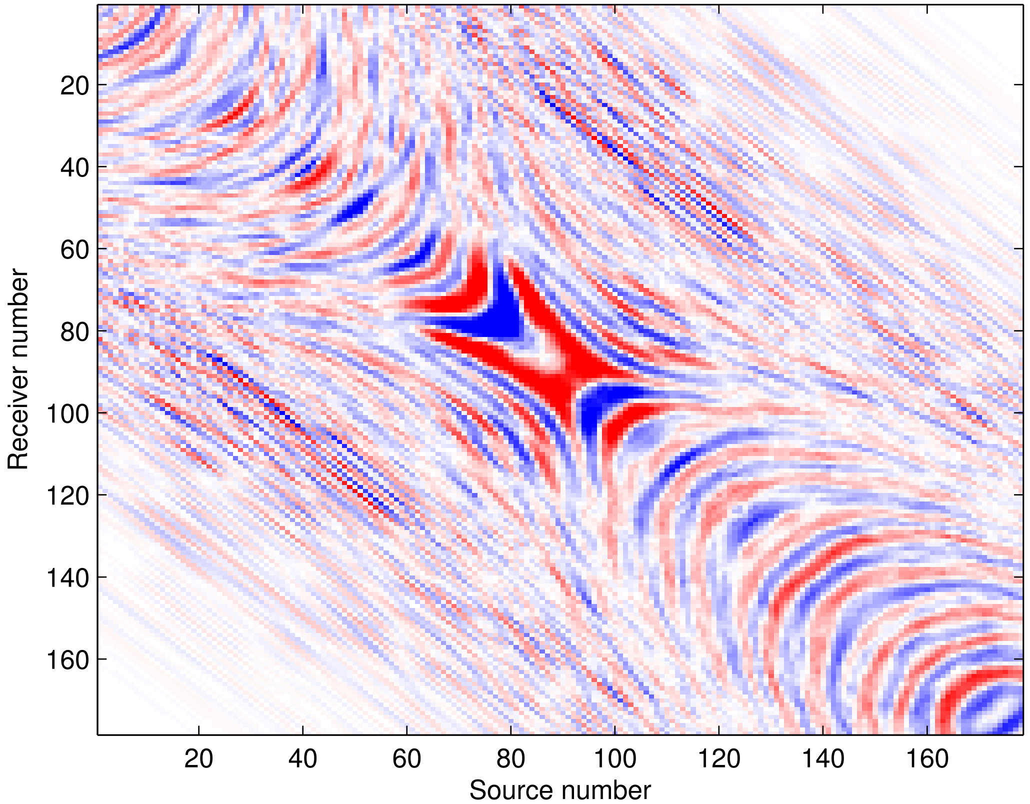
Weighted one-norm implications

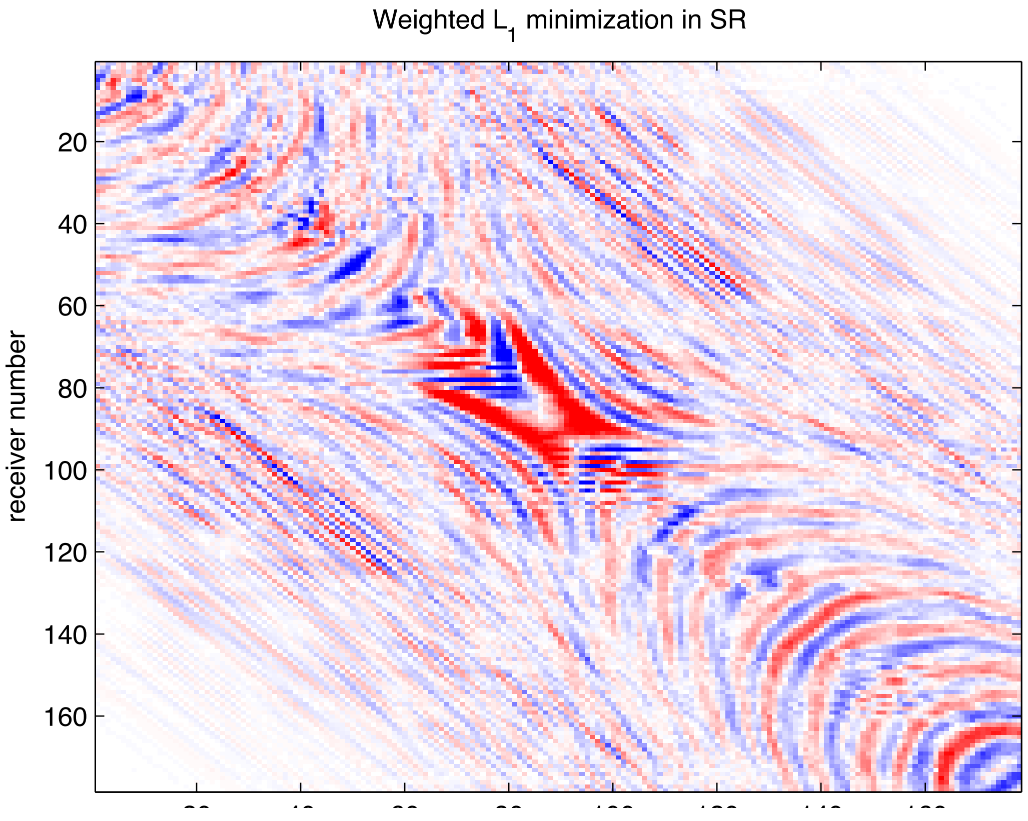
When $\alpha > \frac{1}{2}$, weighted one-norm minimization has better recovery guarantees than standard CS.

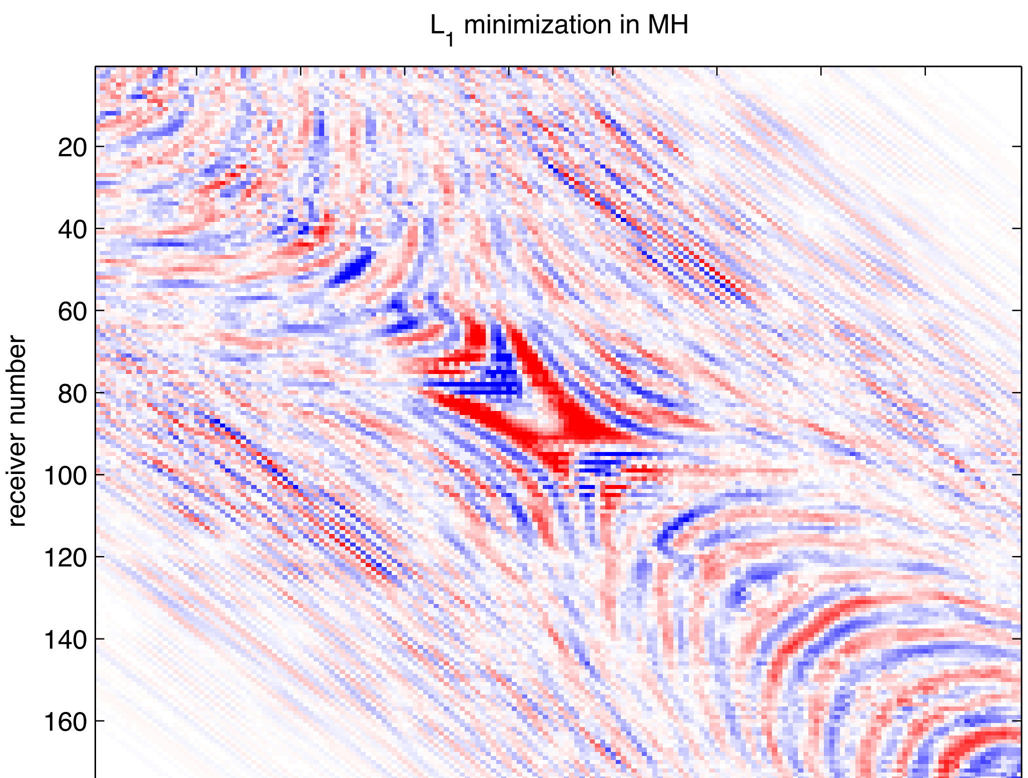
The support sets of the curvelet coefficients of seismic lines make good estimates for adjacent frequency slices.



Weighted L1 recovery
[source-receiver]

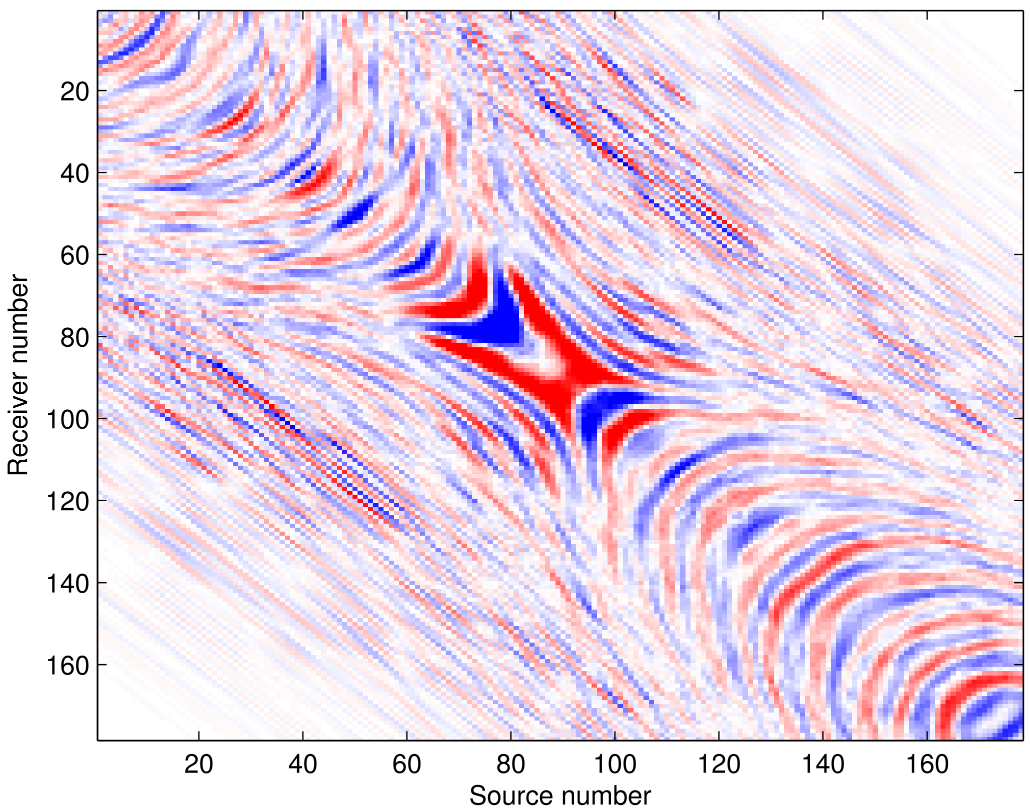


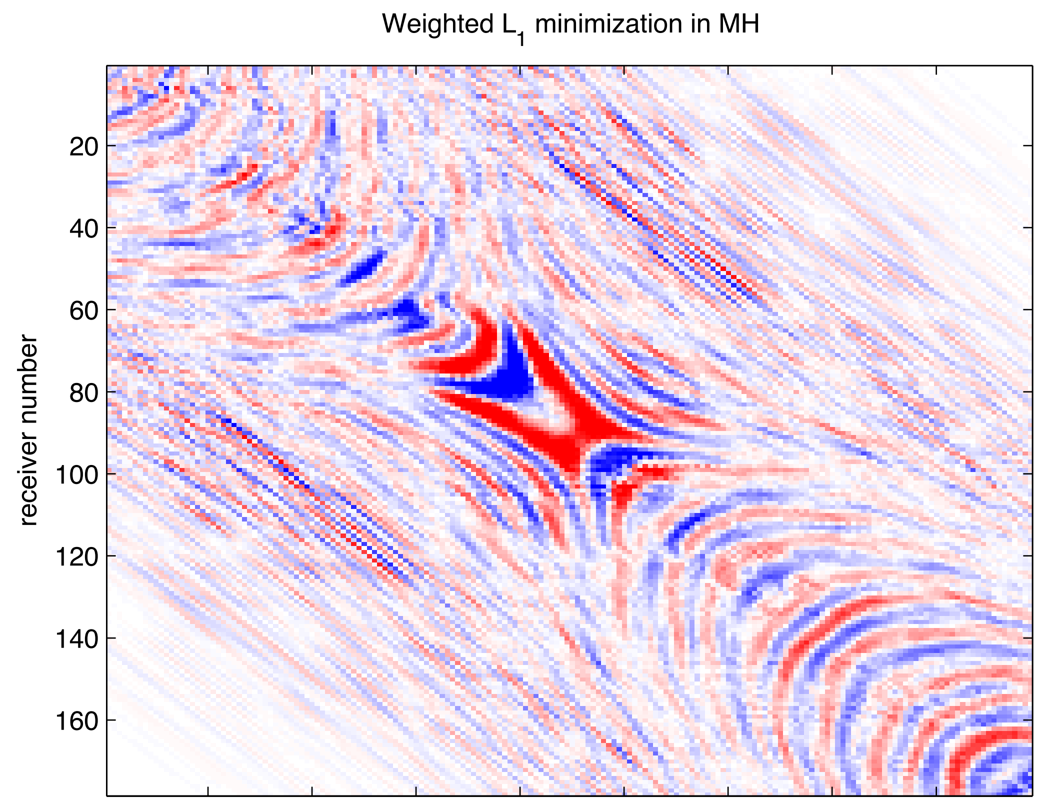


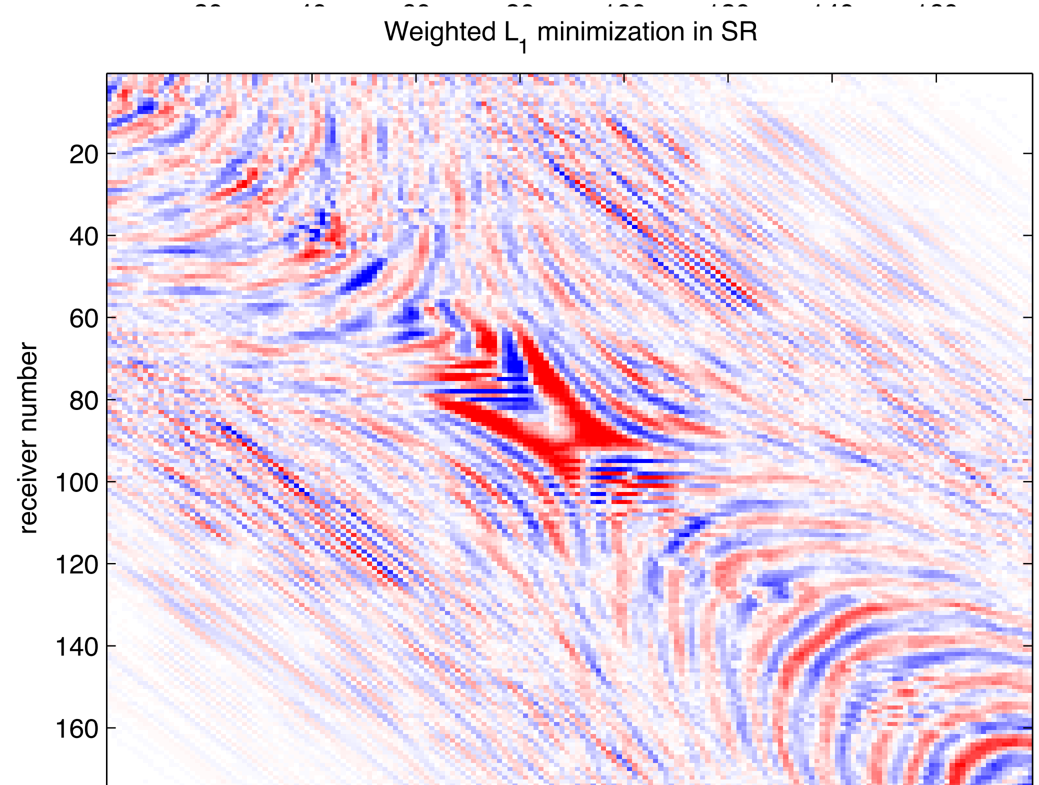


Weighted L1 recovery

[midpoint-offset]



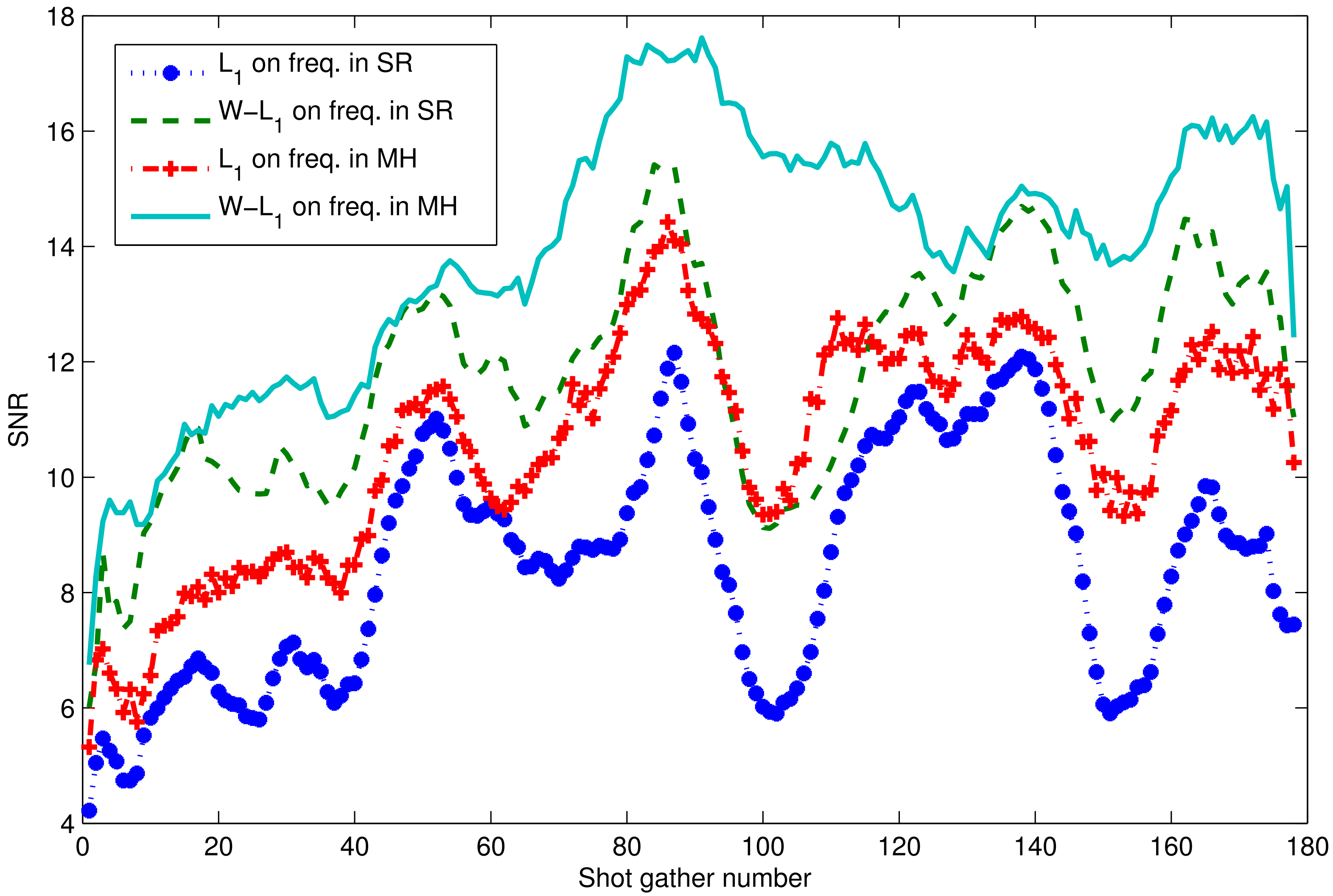




SLIM

Weighted L1 recovery

Signal-to-noise-ratio (SNR) comparison



Promoting Low rank

- Seismic data is *low rank* in some frequencies/ transformed domains (e.g. midpoint-offset), motivating *low-rank* optimization to denoise/recover missing data.
- Low-rank optimization classically relies on *nuclear* norm and SVD, and hence is prohibitively *expensive*.
- We can work with a new *factorized* formulation of *nuclear* (and other) norms, bringing down the cost
- We formulate and solve the problem in an SPGL1-type setting.

Nuclear Norm

- Given any matrix $X = USV^T$,

  the nuclear norm is $\|X\|_* = \sum(\text{diag}(S))$.
- Just like the 1-norm approximates the 0-norm, so the *nuclear* norm approximates the rank.
- Therefore, to find a low rank solution, solve:

$$\min_X \|X\|_*$$

$$\text{such that } \|b - \mathcal{F}(X)\|_2 \leq \sigma \ .$$

Bring on the Pareto!

$$\min_{X} \|X\|_* \quad \text{such that } \|b - \mathcal{F}(X)\|_2 \le \sigma \ .$$

- We can use SPGL1 to solve such problems if
  - It is easy to project onto $\mathbb{B}_*^\tau := \{X : \|X\|_* \le \tau\}$
  - It is easy to evaluate the *dual* norm.
- *Dual* norm is simply *maximum* singular value (op norm)
- But just computing the nuclear norm requires SVDs. Fortunately, we can use a clever trick...

Factorization Approach

- The *Nuclear* norm has a convenient property:

$$\|X\|_* = \inf_{X=LR^*} \frac{1}{2}\left(\|L\|_F^2 + \|R\|_F^2\right)$$

- We can work with L, R rather than X:

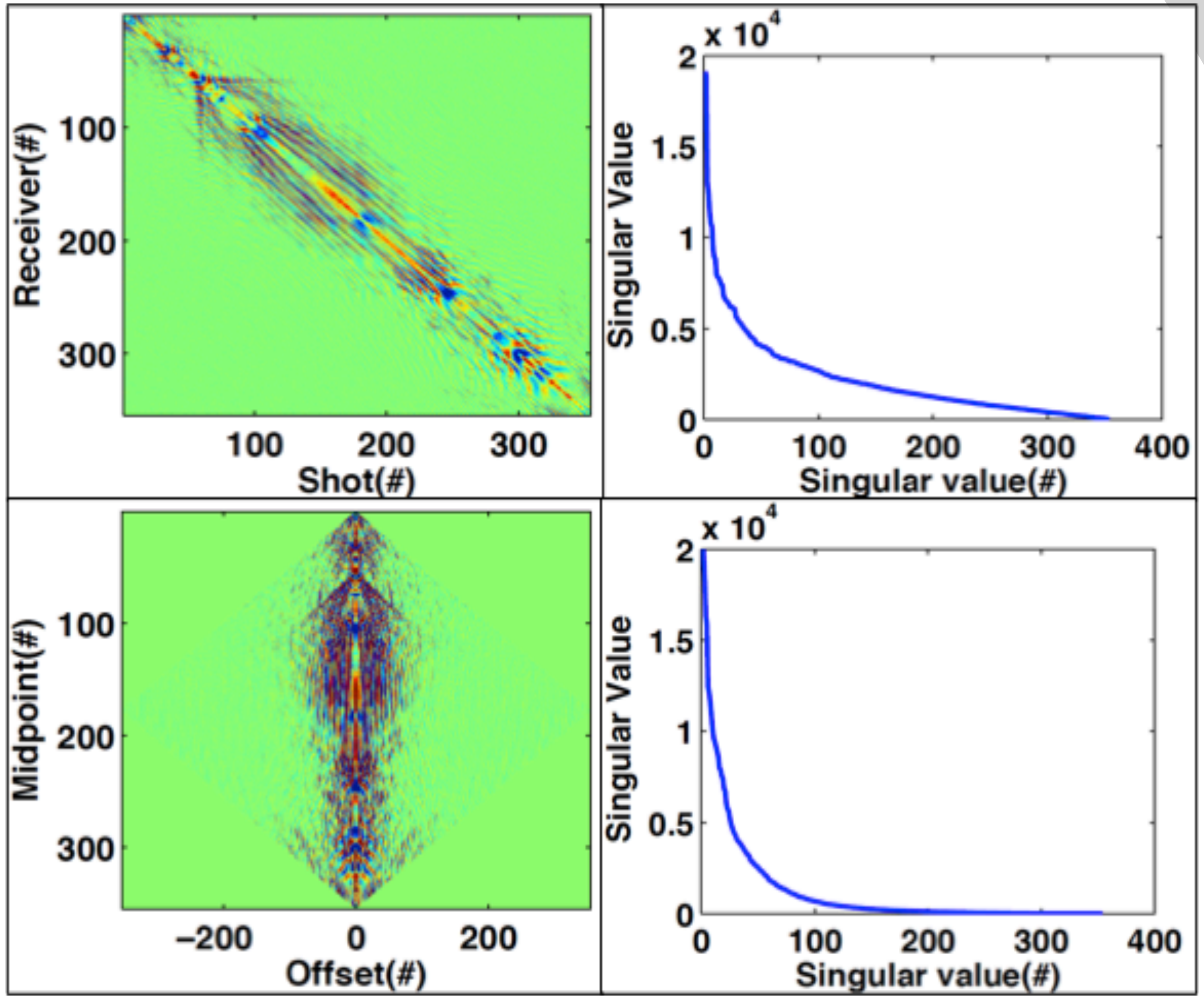
$$\min_{L,R} \frac{1}{2}\left(\|L\|_F^2 + \|R\|_F^2\right)$$

$$\text{such that } \|b - \mathcal{F}(LR^*)\|_2 \leq \sigma \ .$$

- Advantages: no SVD required; trivial projection; potential to use factors L, R downstream.

Rank Optimization in Midpoint-Offset

- Seismic data have faster singular value decay in midpoint-offset domain
- We recover 50% missing data by solving the rank optimization problem for high (70) and low (20) frequencies.
- nr = ns = 354.



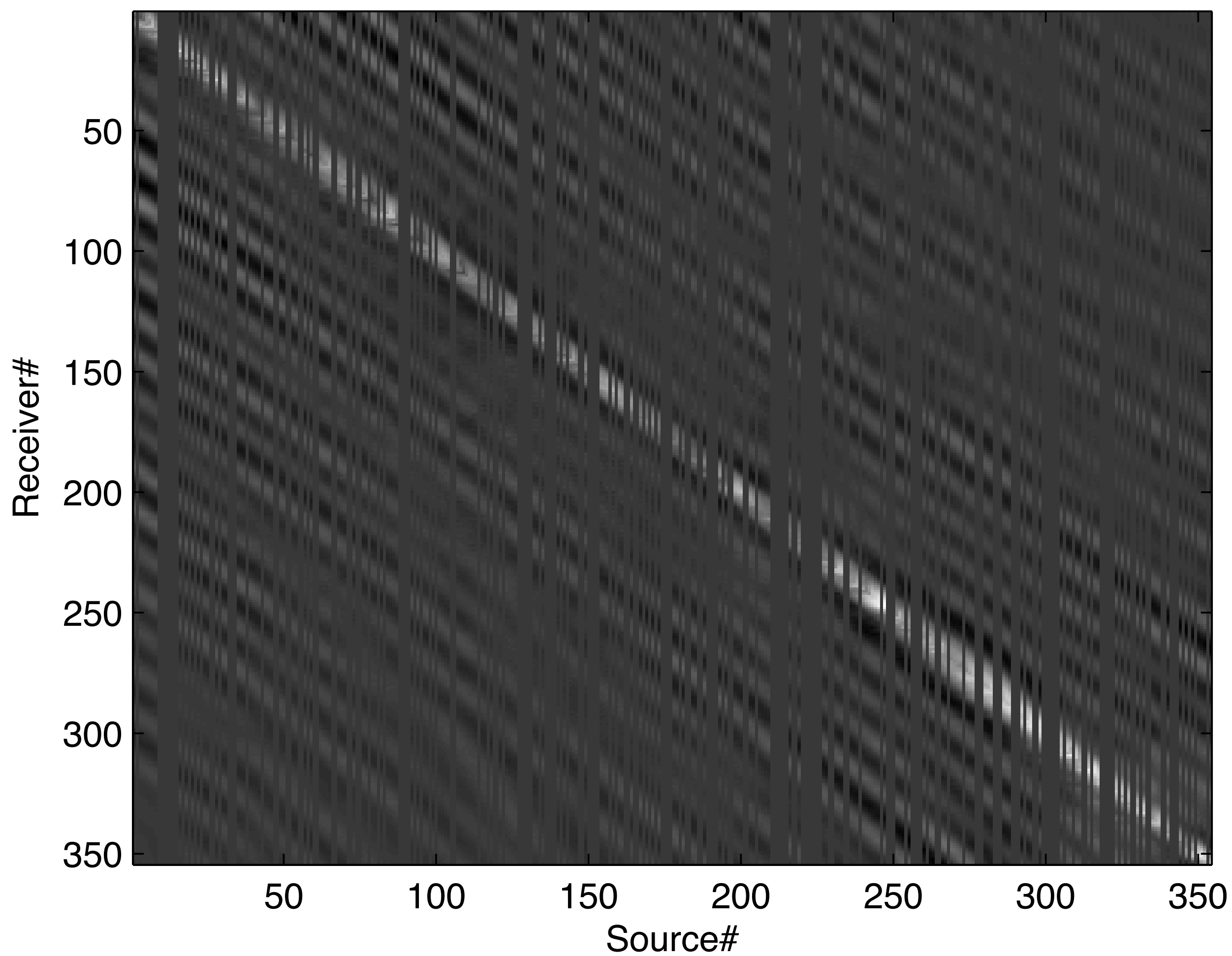
Complete data before and after transformation

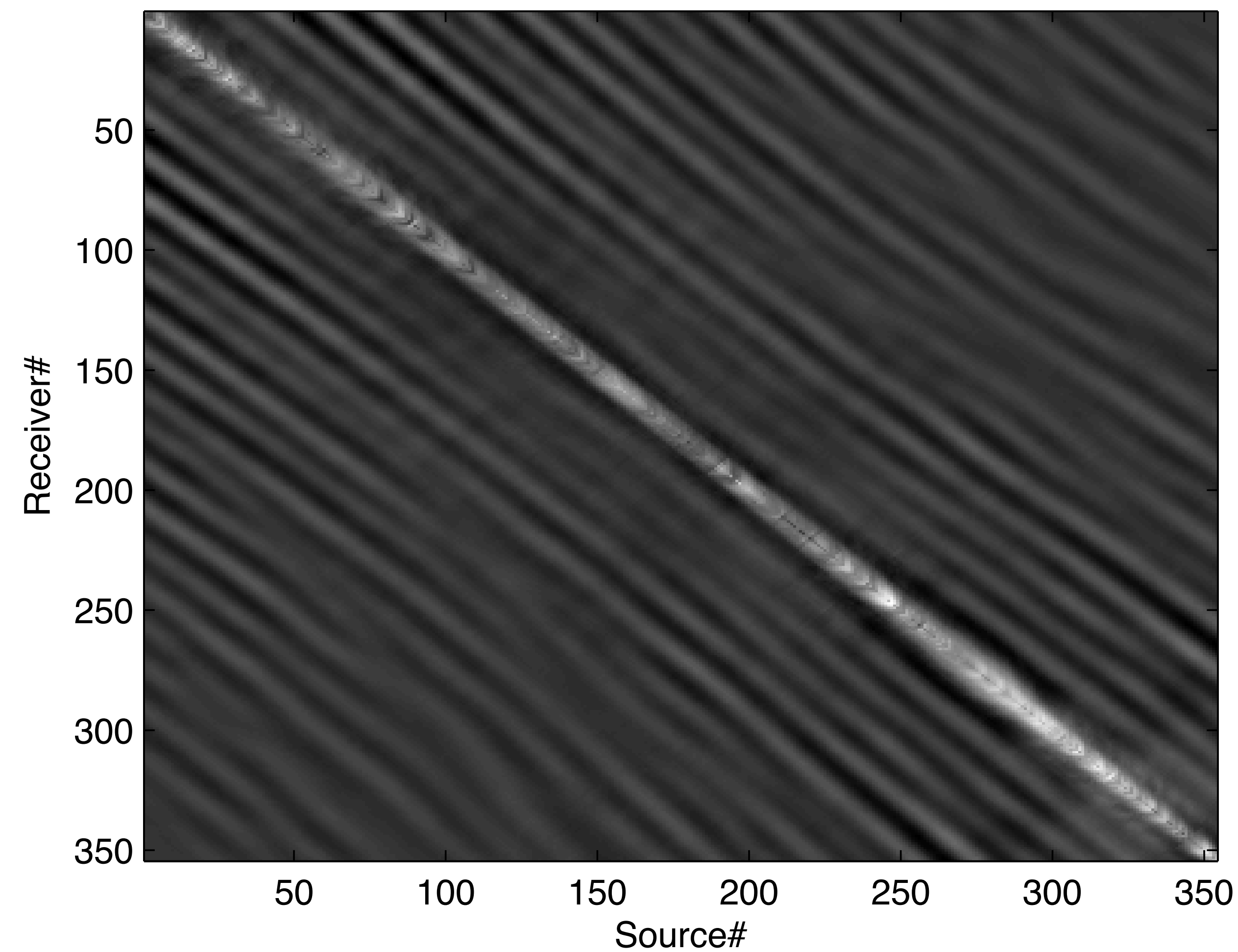
Work flow:

- Convert data with missing traces to M-O domain.
- Initialize L, R factors of pre-selected rank.
- Run rank optimization algorithm (SPGL1+).
- Form dense solution X = LR*
- Convert solution back to source-receiver domain.

Gulf of Suez Data: Least Squares+Low Rank

Frequency Slice : 20 Hz, Rank : 20





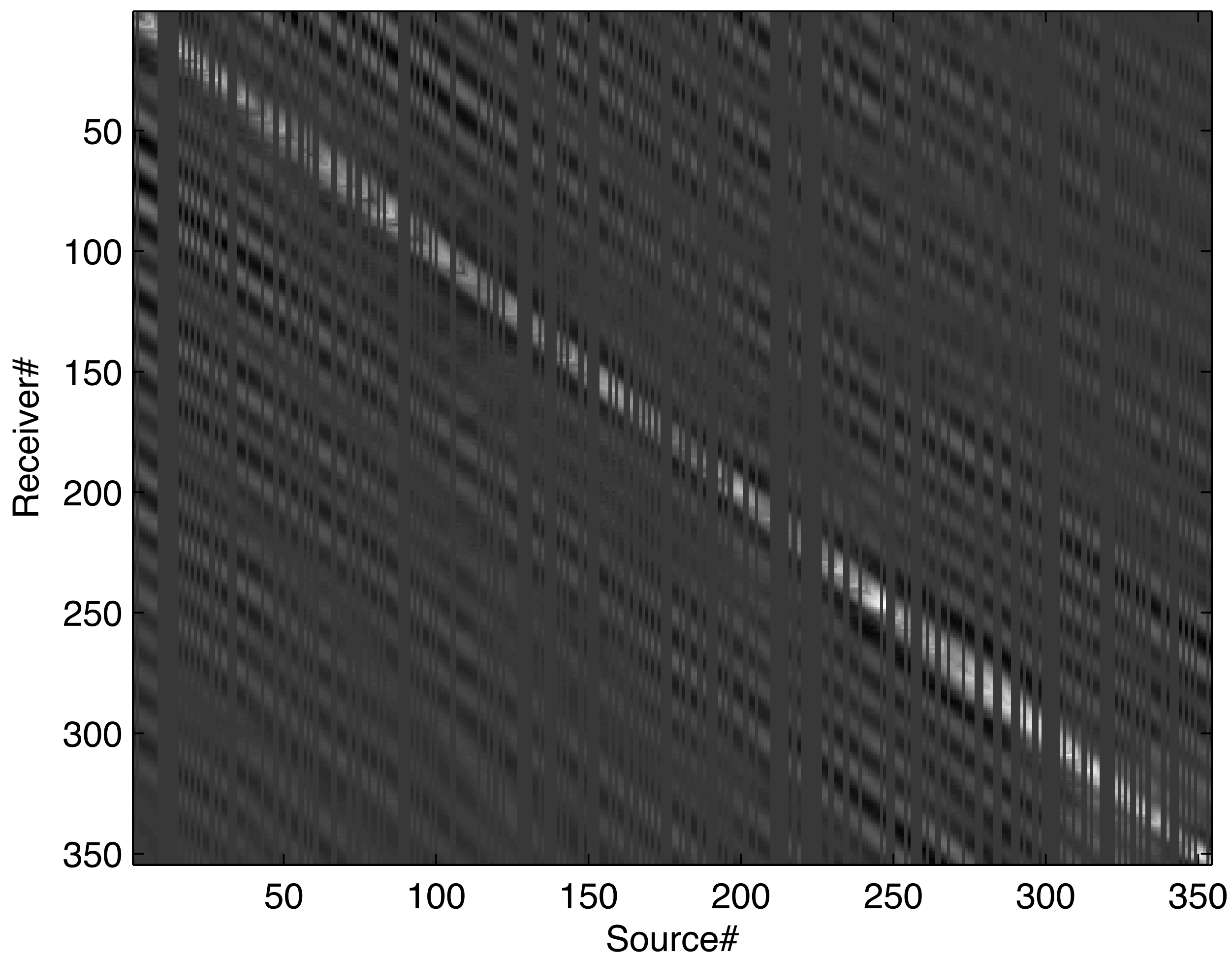
50% Missing data before interpolation

Data after interpolation, **SNR = 41.7 db**

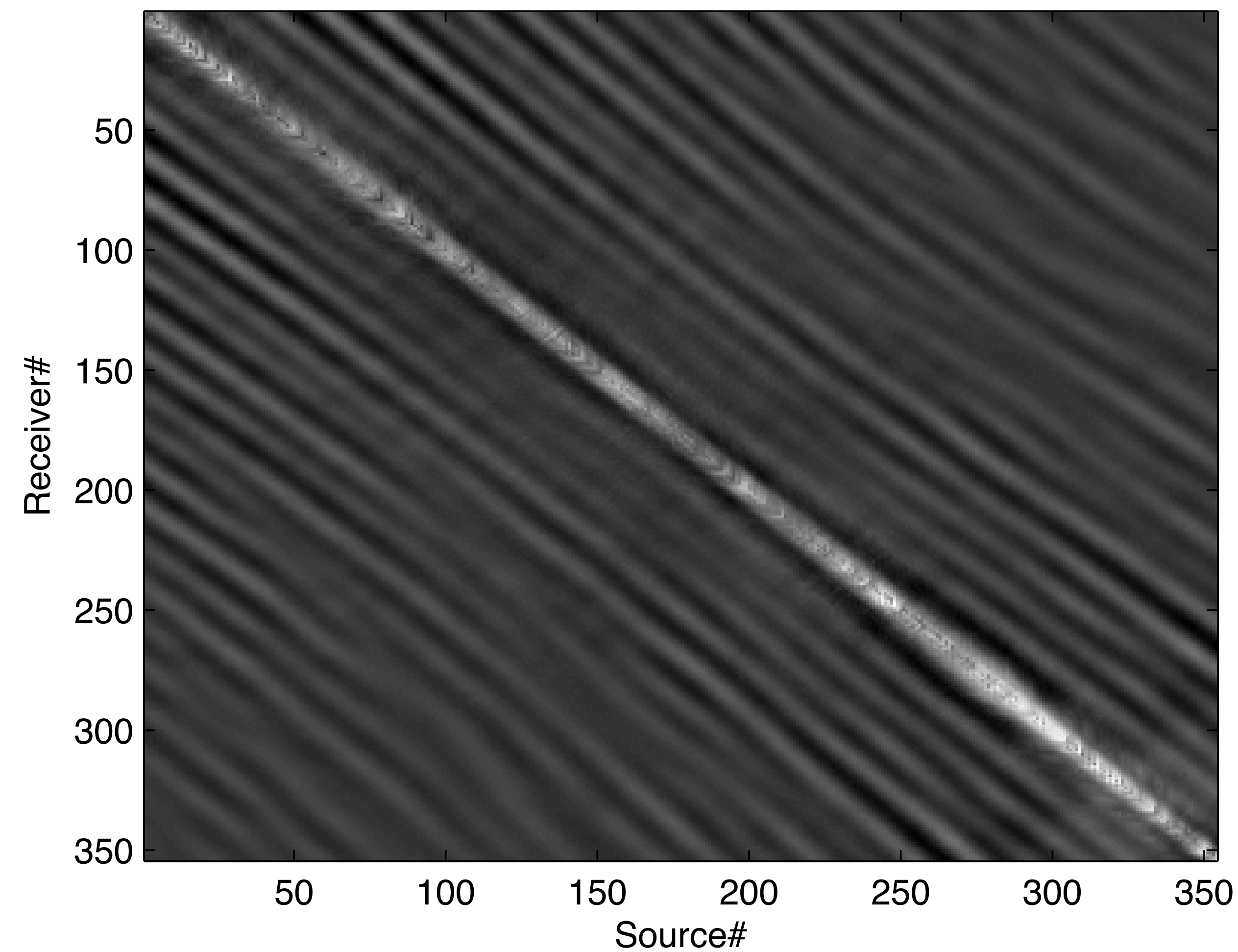
- 150 SPGL1 iterations; sigma = 1e-6, nr = ns = 354.
- Curvelet Sparsity (200 SPGL1 iterations): SNR of 18-20dB

Gulf of Suez Data: Least Squares+Low Rank

Frequency Slice : 20 Hz, Rank : 20



50% Missing data before interpolation

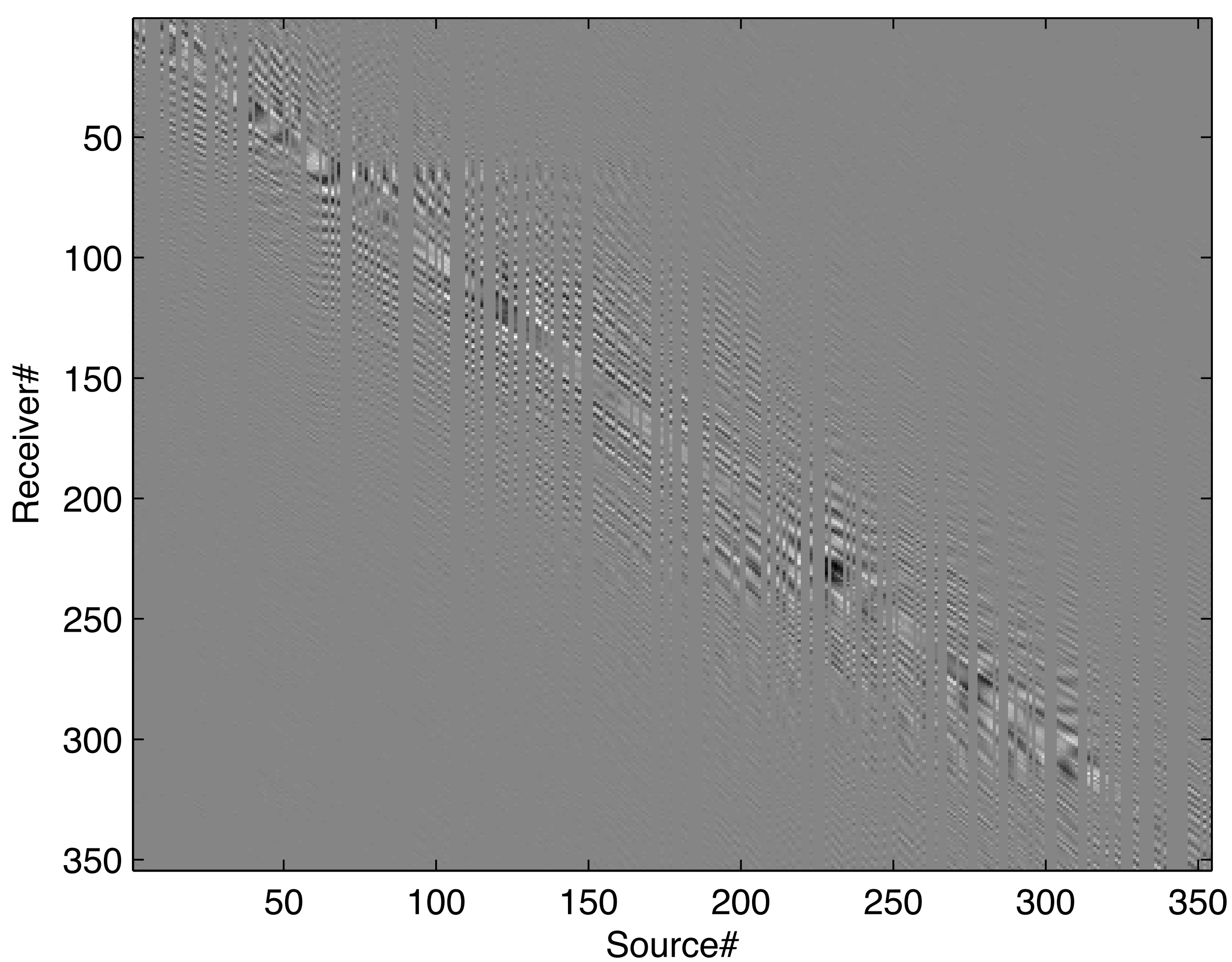


Data after interpolation, **SNR = 42.2 db**

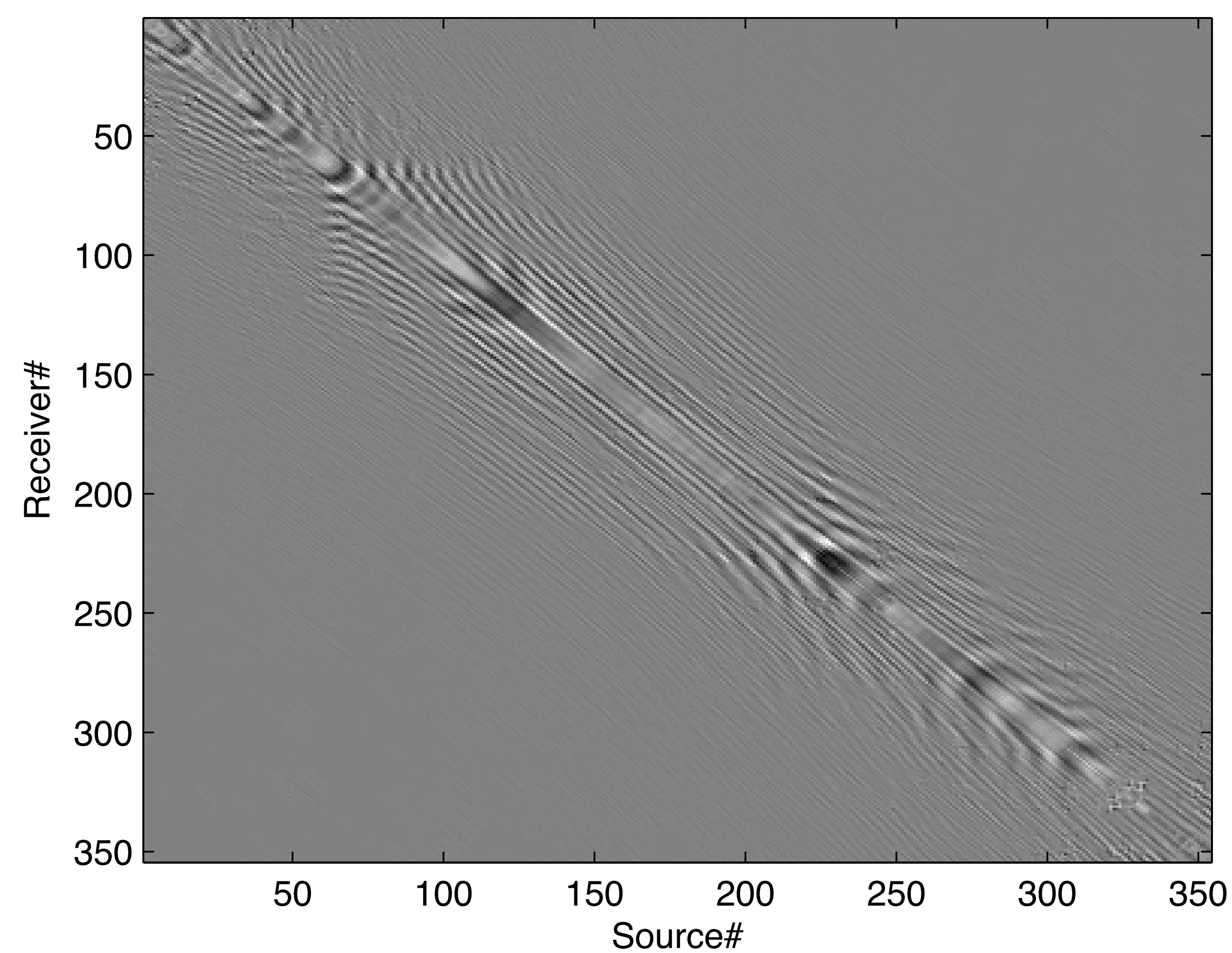
- 150 SPGL1 iterations; sigma = 1e-6, nr = ns = 354.
- Curvelet Sparsity (200 SPGL1 iterations): SNR of 18-20dB

Gulf of Suez: Least Squares + Low Rank

Frequency Slice : 70 Hz, Rank : 20



50% Missing data, before interpolation

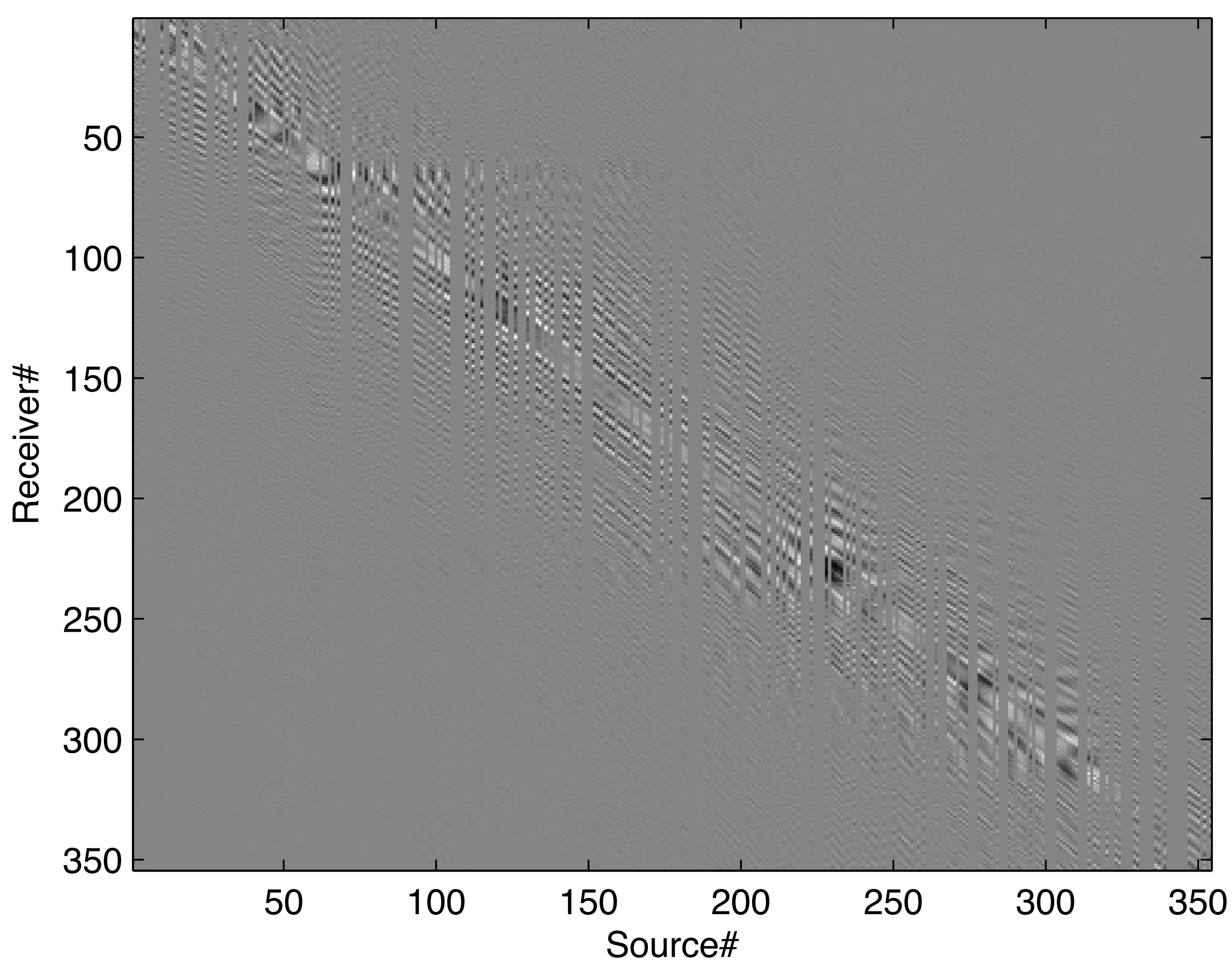


Data after interpolation, **SNR = 22.7 db**

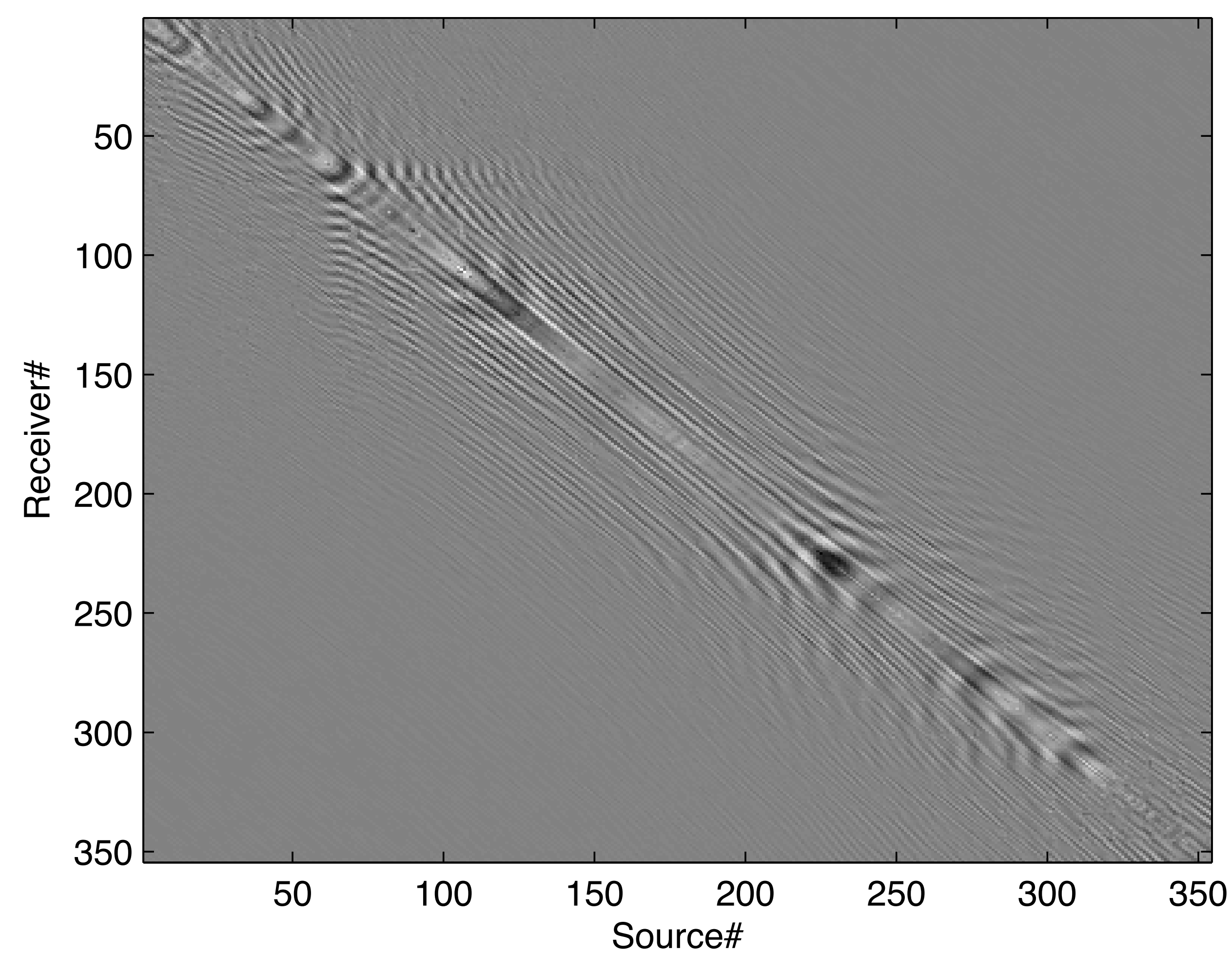
- 150 SPGL1 iterations; sigma = 1e-6, nr = ns = 354.
- Curvelet Sparsity (200 SPGL1 iterations): SNR of 9-10dB

Gulf of Suez: Least Squares + Low Rank

Frequency Slice : 70 Hz, Rank : 40



50% Missing data, before interpolation



Data after interpolation, **SNR = 29.3 db**

- 150 SPGL1 iterations; sigma = 1e-6, nr = ns = 354.
- Curvelet Sparsity (200 SPGL1 iterations): SNR of 9-10dB

SLIM

Current/Future work:

- Find transforms and representations that make seismic data *low-rank*.
- Can factorizations $X = LR^*$ in these representations be used directly?
- Combine *low rank* with *robust* penalties (later talk)
- Use other norms that have computational advantages, such as the Max norm.