APPROXIMATE MESSAGE PASSING MEETS EXPLORATION SEISMOLOGY

Felix J. Herrmann*, UBC–Seismic Laboratory for Imaging and Modelling

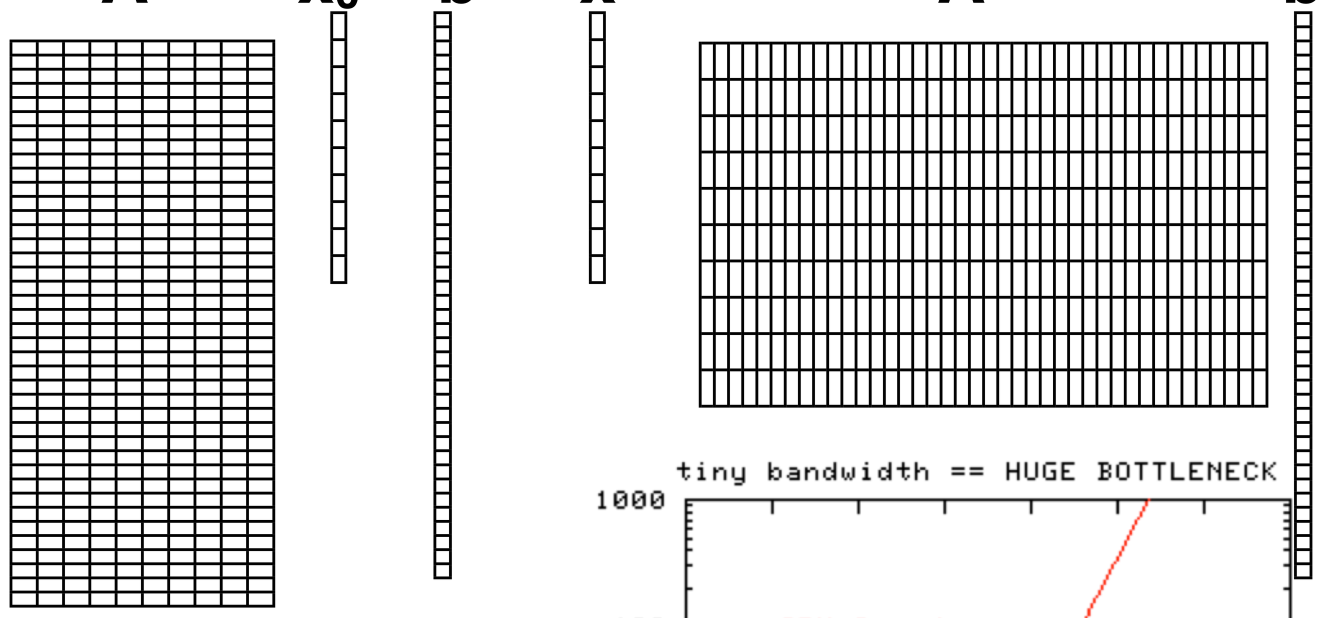
SLIM Seismic Laboratory for Imaging and Modeling

Abstract: *Seismic* imaging *hinges* on *large-scale* sparsity promoting *solvers* to remove *artifacts* caused by *efforts* to *reduce* (computational) costs. Including "Onsager's *message* term" improves *convergence* by *canceling* correlations between model *iterates* and *linearized* forward *models*. *Unfortunately*, the *asymptotic* arguments of *approximate* message passing are oftn *violated* in practice. By using Montanari's *heuristic* argument of "statistical equivalence", we *break correlations* by selecting *independent* experiments via *randomized* subsampling.

Current *imaging* paradigm

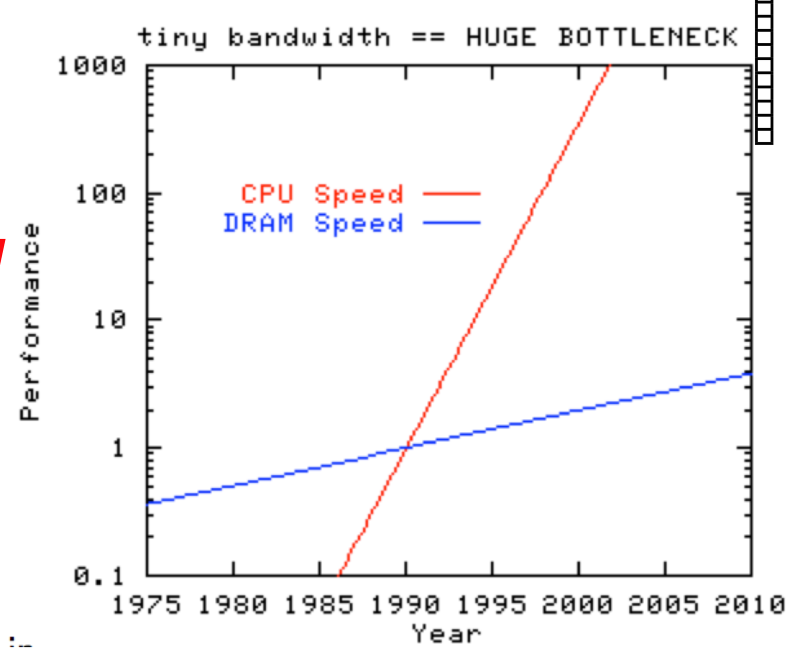
Linear forward model: **Imaging:**

A $\mathbf{x}_0 = \mathbf{b}$ $\mathbf{x} =$ $\mathbf{A}^H$ **b**



Challenge of *big data...*

Credit to John McCaplin, University of Texas, HPC.



IO bottle neck of seismic imaging caused by insisting on touching all data.

Data deluge

Seismic *imaging* calls for:

- *large* data volumes (peta bytes)
- *inversion* instead of "matched" filtering
- more *sophisticated* physics
- *multiple* passes through *all* data

Exposes:

- IO *bottlenecks*
- lack of growth in *sequential* processing
- pressures on *parallelization*

Our solution:

- *work* on small *randomized* subsets of *data*
- go from *data*– to *model*–space *parallelism*

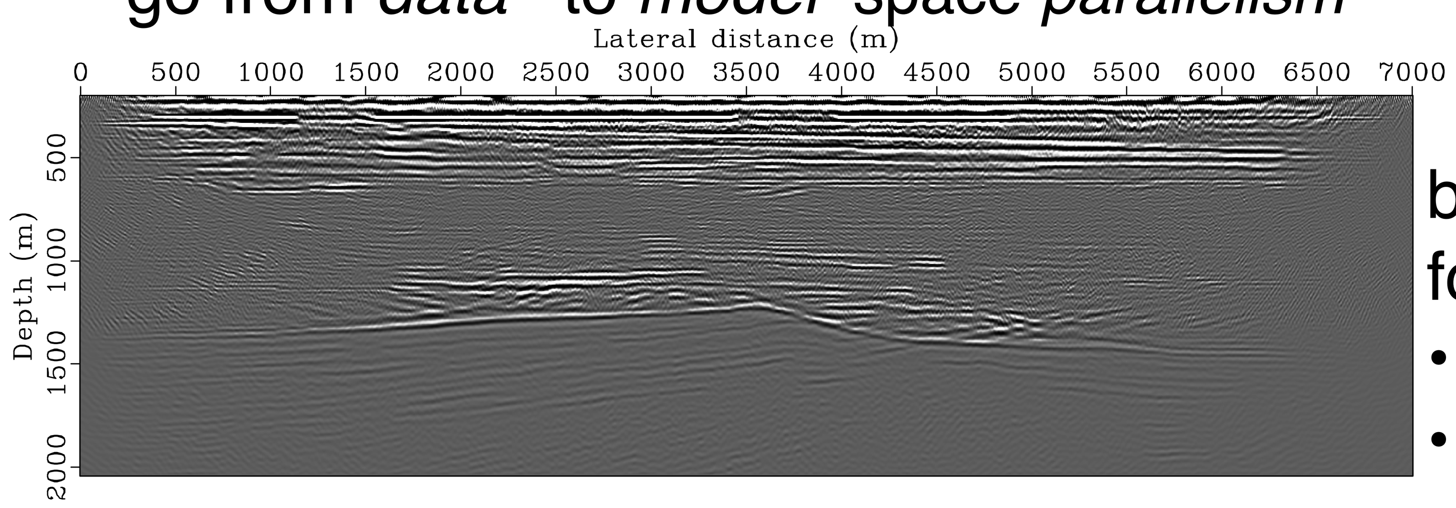
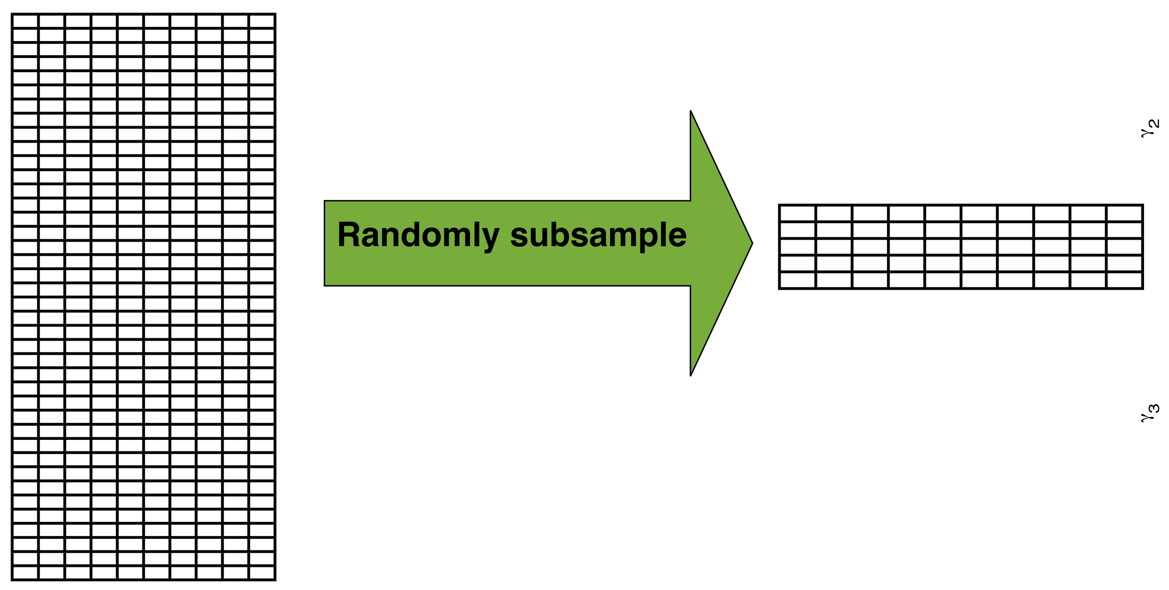


Image by touching all 350 source experiments once with "matched" filter.

***Thanks to the sponsors of the SINBAD Consortium & NSERC CRD Grant DNOISE (375142-08)**

New paradigm

***Tall* "overdetermined" system** ***Wide* underdetermined system**



Dimensionality reduction with *CS*

- ability to *exploit* model-space *structure*
- reduces *size* data volumes
- reduces *wave* simulation costs
- at the *price* of introducing *interference*

Convex optimization with *cooling*

Solve

$$\underset{\mathbf{x}}{\text{minimize}}\ \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{b}$$

via a *series* of LASSO *subproblems*

$$\underset{\mathbf{x}}{\text{minimize}}\ \frac{1}{2}\|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 \quad \text{subject to} \quad \|\mathbf{x}\|_1 \leq \tau$$

for *appropriately* chosen *increasing* τ's.

- uses *Pareto* trade-off curve [v.d. Berg & Friedlander, '08]
- *suffers* from *harmful* correlation *buildup*
- solvers need *too* many *matvecs*

Approximate Message passing

Add "*Onsager*"-term to *iterative* thresholding

$$\mathbf{x}^{t+1} = \eta_t\left(\mathbf{A}^*\mathbf{r}^t + \mathbf{x}^t\right)$$
$$\mathbf{r}^t = \mathbf{b} - \mathbf{A}\mathbf{x}^t + \frac{\|\mathbf{x}^{t+1}\|_0}{n}\mathbf{r}^{t-1}$$

for $\mathbf{A}^{n\times N} \in n^{-1/2}N(0,1)$ and $N \rightarrow \infty$

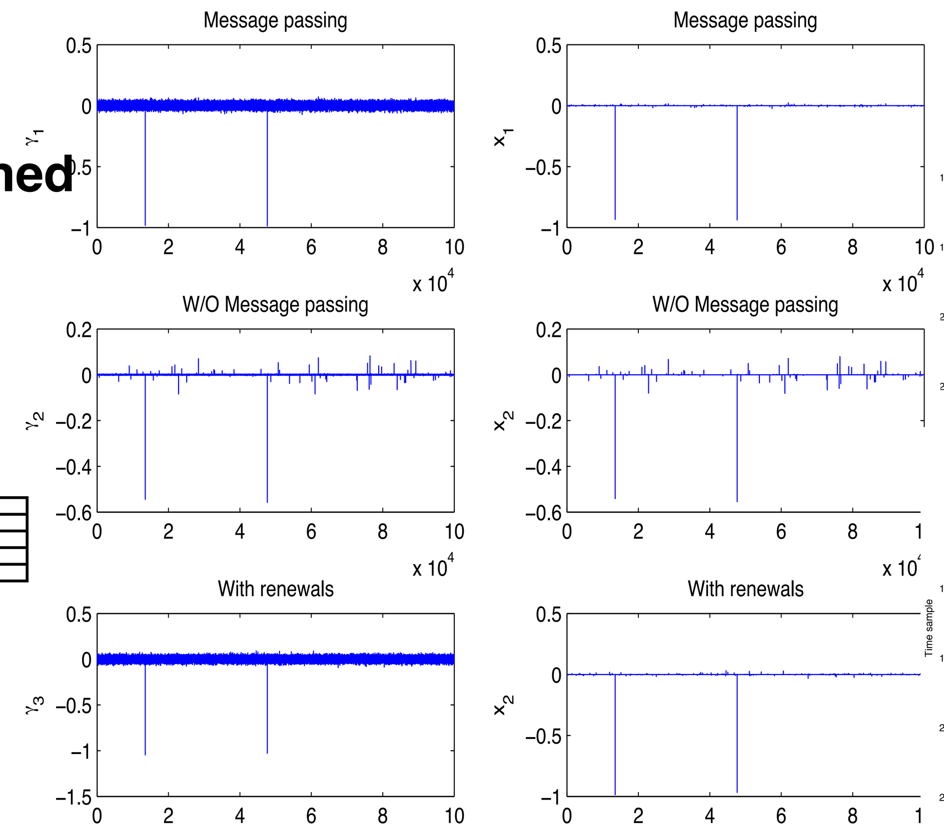
Statistically equivalent to [Montanari, '12]

$$\mathbf{x}^{t+1} = \eta_t\left(\mathbf{A}_t^*\mathbf{r}^t + \mathbf{x}^t\right)$$
$$\mathbf{r}^t = \mathbf{b}_t - \mathbf{A}_t\mathbf{x}^t$$

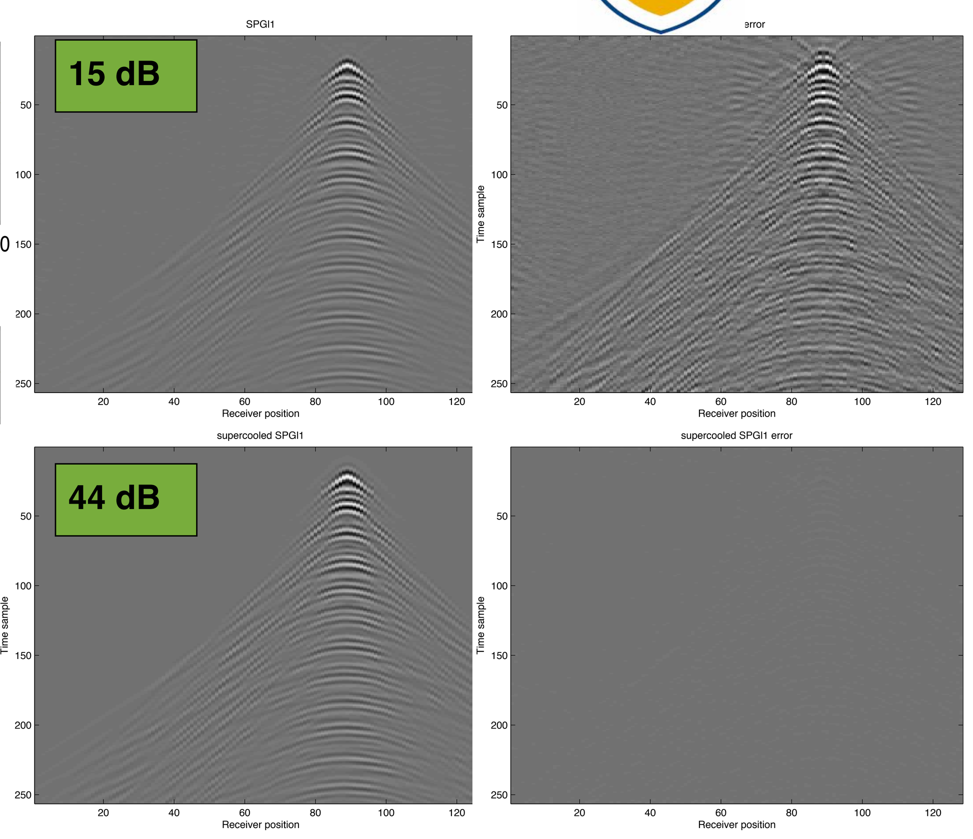
by drawing *new* independent $\{\mathbf{b}_t, \mathbf{A}_t\}$ for *each* iteration.

- *Onsager* term *cancels* harmful *correlations*
- *renewals* yield similar *decorrelation*
- improves convergence *data-rich problems*

***Possibility* to *better* & *speedup* imaging...**



Model iterates (t=4) before (left) and after (right) soft thresholding. Top: with Onsager; Middle: without Onsager; Bottom: with renewals.



Recovery (t=500) of curvelet vector from CS with Gaussian matrix with N=248759 and $\delta = n/N = 0.13$ Top: recovery & error without renewals. Bottom with.

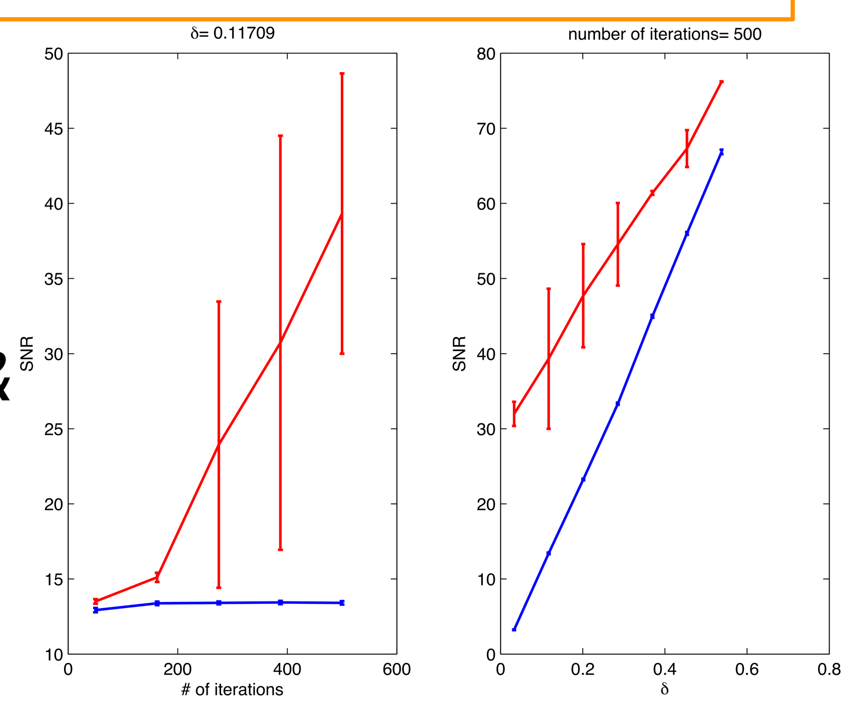
Supercooled convex optimization

LASSO's for $\{\mathbf{b}_t, \mathbf{A}_t\}$ with τ_t from *previous* solution

$$\underset{\mathbf{x}}{\text{minimize}}\ \frac{1}{2}\|\mathbf{b_t} - \mathbf{A_t}\mathbf{x}\|_2 \quad \text{subject to} \quad \|\mathbf{x}\|_1 \leq \tau_t$$

MCC recovery experiments

- 25 *runs* with & without *renewals*
- *different* subsampling *ratios* & *number* of *iterations (t)*
- renewals *improve* results *dramatically*



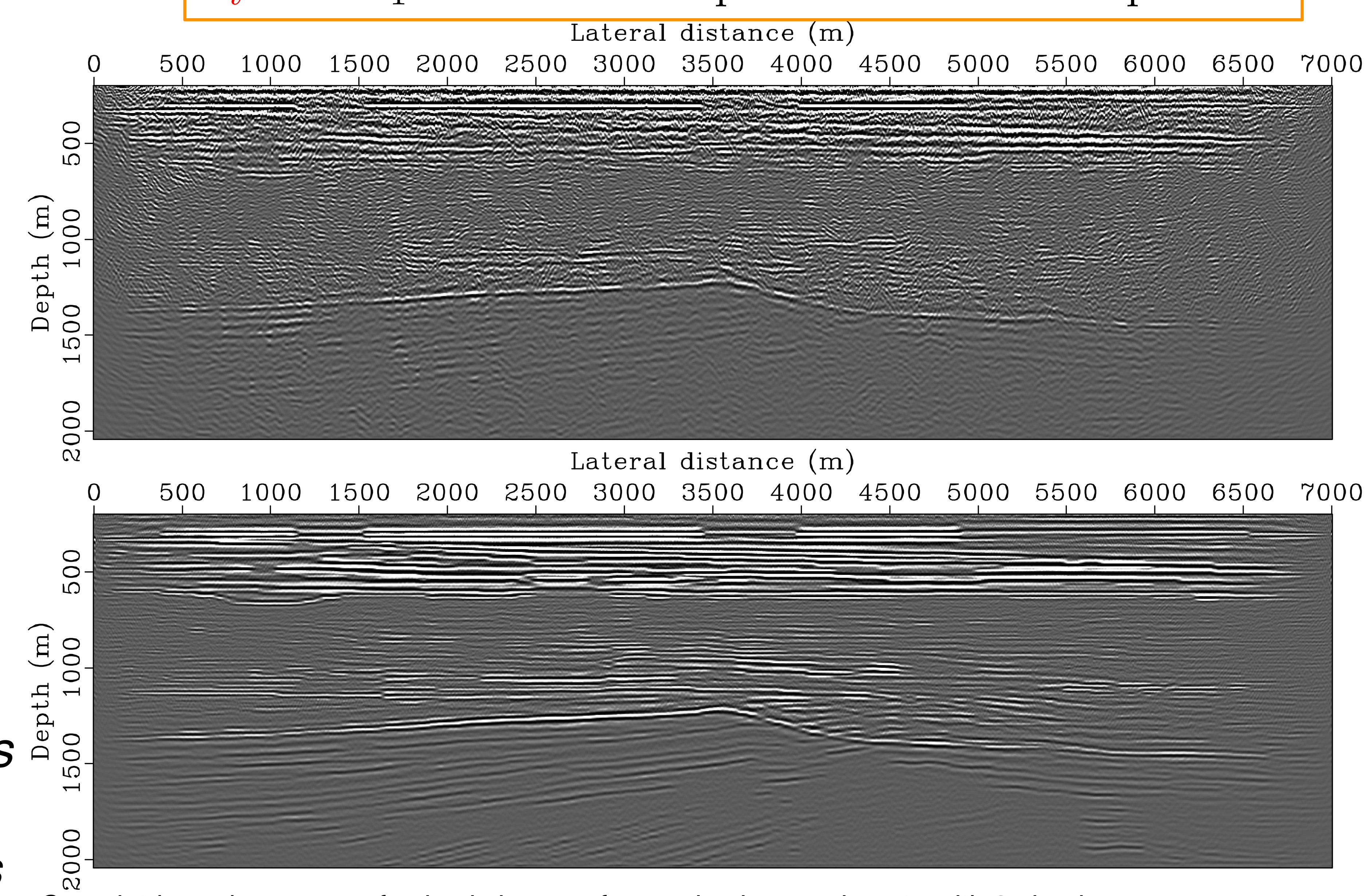
Recovery with (red) and without renewals.

New *imaging* paradigm

- *reduces* number of PDE solves.
- improves *convergence*
- *possible* because seismic *imaging* is "*data rich*"...

Set

$$\mathbf{b}_t = \text{Independent dimensionality reduced data}$$
$$\mathbf{A}_t = \text{Lin. Born scattering operator}$$
$$\tau_t = \ell_1\text{-norm solution previous LASSO subproblem}$$



Curvelet-based recovery of seismic images from seismic experiments with 3 simultaneous sources opposed to 350 sequential sources. Top: inverted image by sparsity promotion without renewals. Bottom: the same but with renewals. Computational cost is the same for both.