SINBAD Consortium 2010

Exploiting Sparsity in FWI: Nonlinear BPDN

Aleksandr Aravkin & Tristan van Leewen

SLIM
University of British Columbia

SLIM

Full Waveform Inversion

- The Full Waveform Inversion (FWI) problem is to find solutions to the Helmholtz PDE that match data from source experiments observed on the surface

- Commonly formulated as nonlinear least-squares problem:

$$\min_{m} \|D - RH^{-1}[m]Q\|_F^2$$

- FWI is ill-posed -- the observed data is not sufficient to recover the solution.
  1) Need to start close
  2) Can't iterate 'too long'

SLIM

Ill-posed Problems

Several techniques are used to deal with ill-posed problems

1) Small fixed iteration count

2) Regularization with respect to initial guess, e.g.:

$$\min_{m} \|D - RH^{-1}[m]Q\|_F^2 + \lambda \|m - m_0\|_2^2$$

3) Our approach: use sparsity by imposing L1 penalty

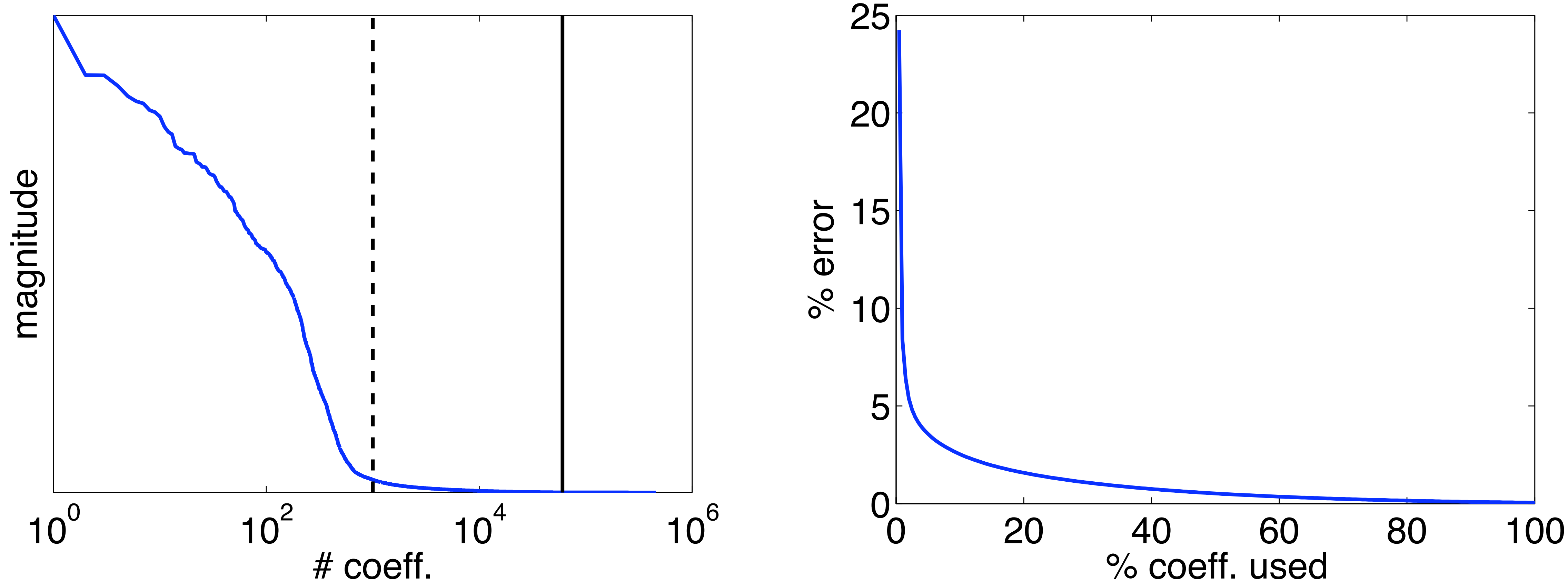
- Images and velocity models are sparse in Curvelets
- Time-lapse difference images are sparse

Sparsity in Curvelets

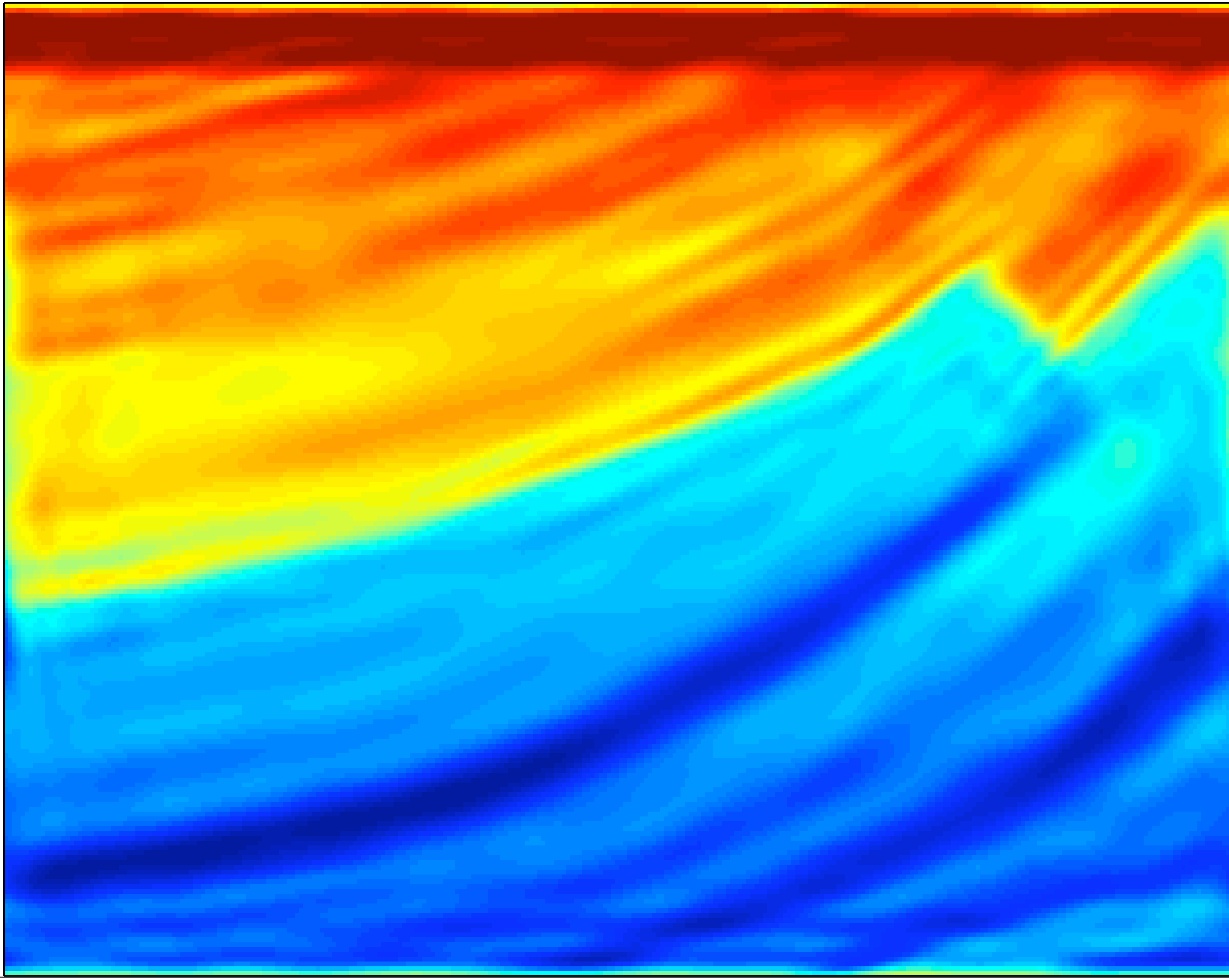
- Typical **velocity model** is sparse (compressible) in Curvelets:

$$\bar{m} = C^* \bar{x}$$

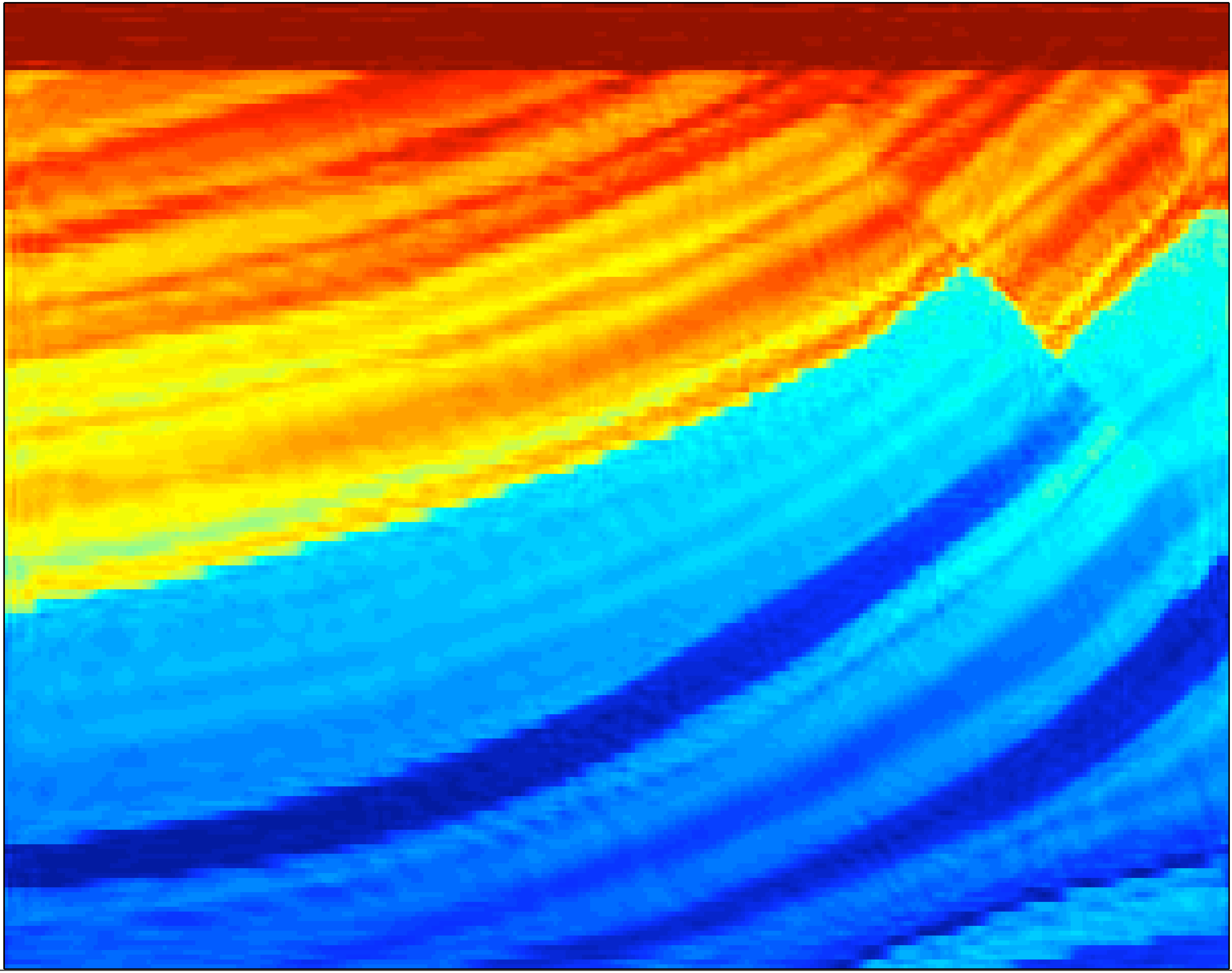
- This means we need very few coefficients to capture the model:



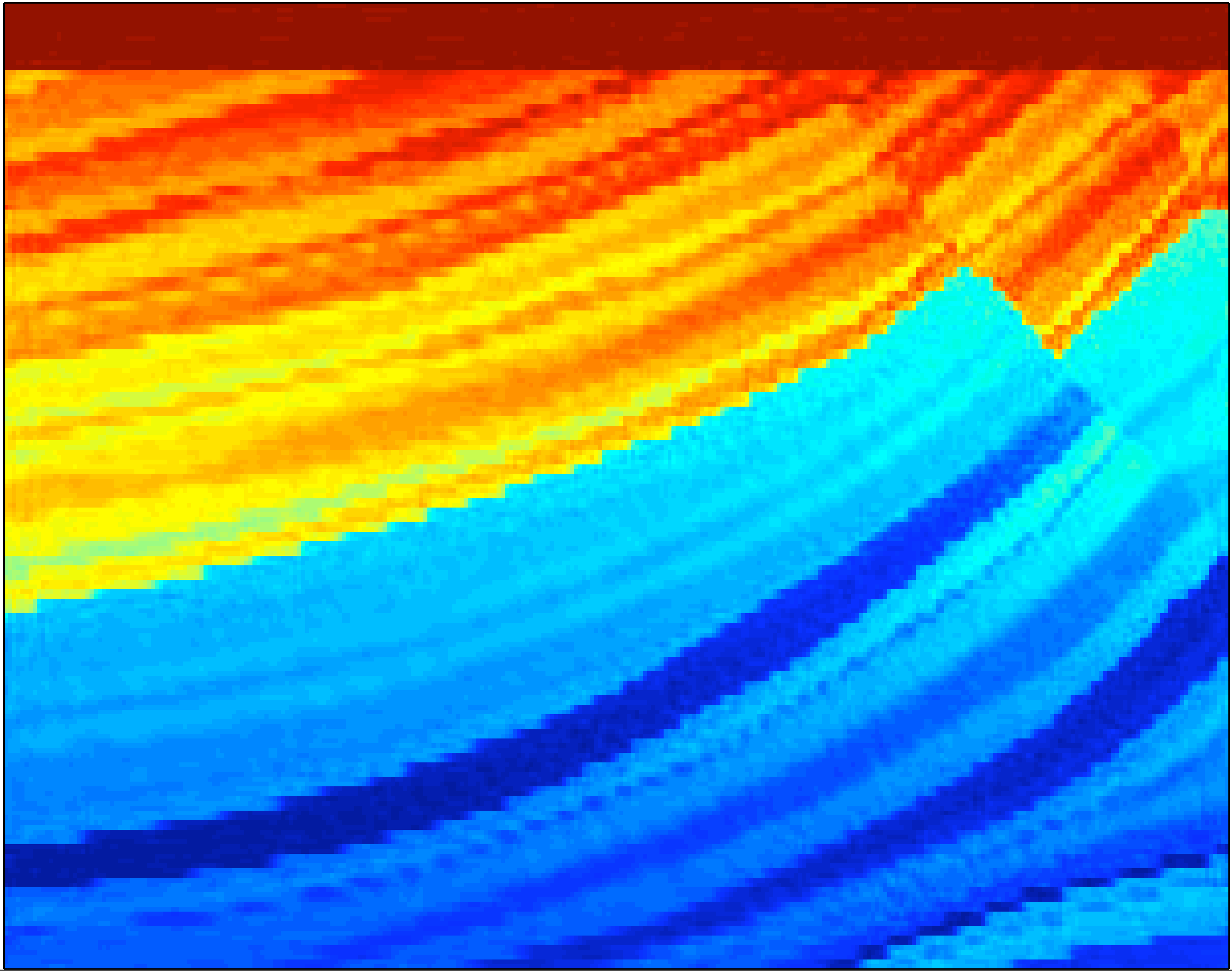
1% of coeff.



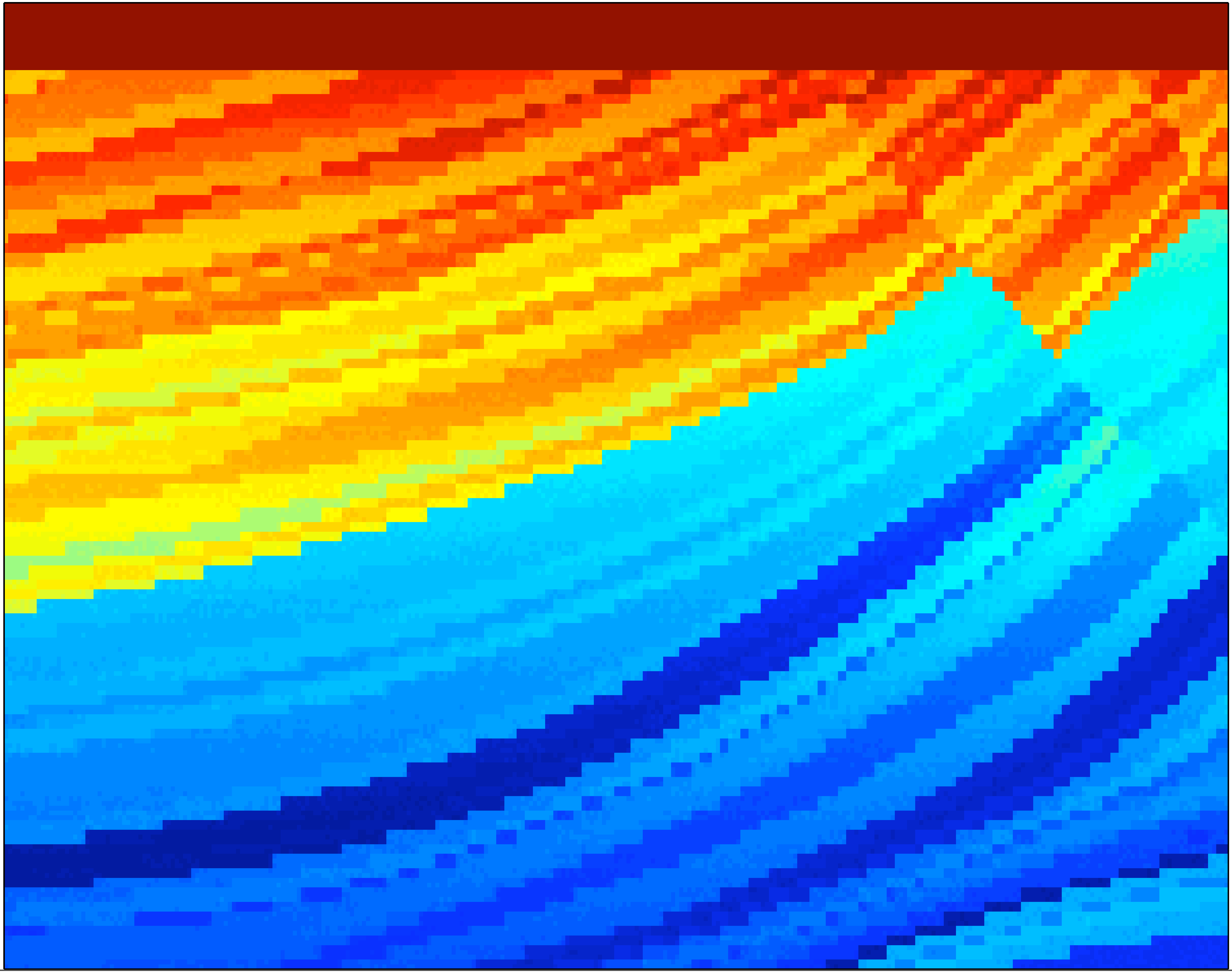
5% of coeff.



10% of coeff.



50% of coeff.



FWI: Sparsity promotion

Sparsity-exploiting formulations

1. $\min_{x} ||D - RH^{-1}[C^* x]Q||_F^2 + \lambda ||x||_1$

2. $\min_{x} ||D - RH^{-1}[C^* x]Q||_F^2 \quad \text{s.t.} \quad ||x||_1 \leq \tau$

3. $\min_{x} ||x||_1 \quad \text{s.t.} \quad ||D - RH^{-1}[C^* x]Q||_F^2 \leq \sigma$

Option 3. has a big advantage: we may be able to derive σ from the data *and* we don't need to guess the sparsity of the solution in Curvelets.

FWI: Sparsity promotion

- Sparsity-exploiting formulation:

$$\begin{array}{ll} \min_x & \|x\|_1 \\ \text{s.t.} & \|D - RH^{-1}[C^*x]Q\|_F^2 \leq \sigma \end{array}$$

$$\bar{m} = C^*\bar{x}$$

- This is a Convex-Composite optimization problem.

Strategy

- Consider a toy model problem:

$$\min_{m} \quad \|m\|_1 \quad \text{s.t.} \quad \|d - g(m)\|_2 \leq \sigma$$

- Implement iterated algorithm:

$$m^{\nu+1} = m^{\nu} + \gamma_{\nu} s^{\nu}$$

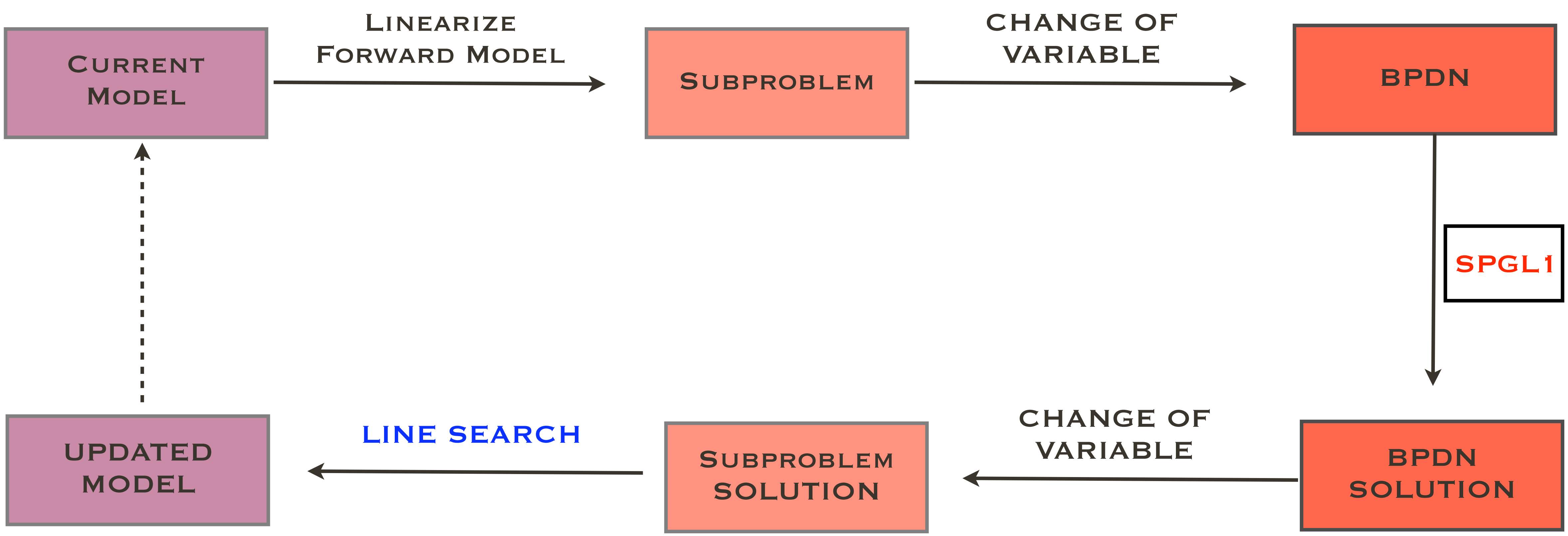
- Direction s^{ν} solves subproblem:

$$\min_{\delta m} \quad \|m^{\nu} + \delta m\|_1 \quad \text{s.t.} \quad \|d - g(m^{\nu}) - \nabla g(m^{\nu})\delta m\|_2 \leq \sigma$$

- Subproblem equivalent to BPDN, which is solved with SPGL1!

$$\min_{y} \quad \|y\|_1 \quad \text{s.t.} \quad \|b - Ay\|_2 \leq \sigma$$

Algorithm



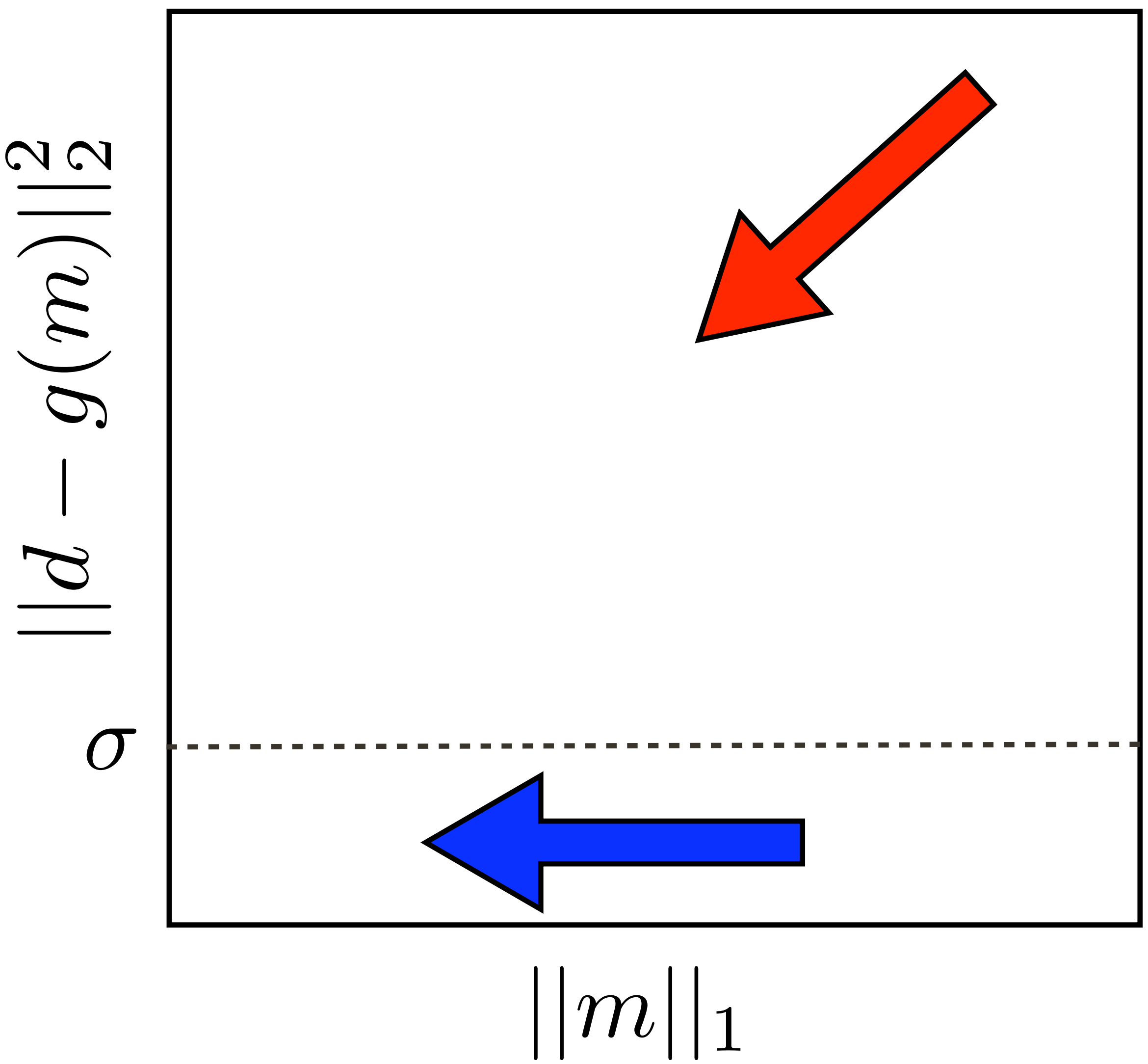
Current Model
Linearize Forward Model
Subproblem
Change of Variable
BPDN
SPGL1
BPDN Solution
Change of Variable
Subproblem Solution
Line Search
Updated Model

Updating the model

$$\min_m \quad \|m\|_1$$
$$\text{s.t.} \quad \|d - g(m)\|_2 \leq \sigma$$

- Competing interests: minimize the 1-norm (sparsity) and decrease misfit.
- Main idea: weigh interests differently as the algorithm proceeds.
- Define merit function $P_\alpha(m) = \|m\|_1 + \alpha(\|d - g(m)\|_2 - \sigma)_+$
- Line search ensures $P_\alpha(m^\nu + \gamma_\nu s^\nu) < P_\alpha(m^\nu)$
- Increase α as the algorithm proceeds.

Updating the model



SLIM

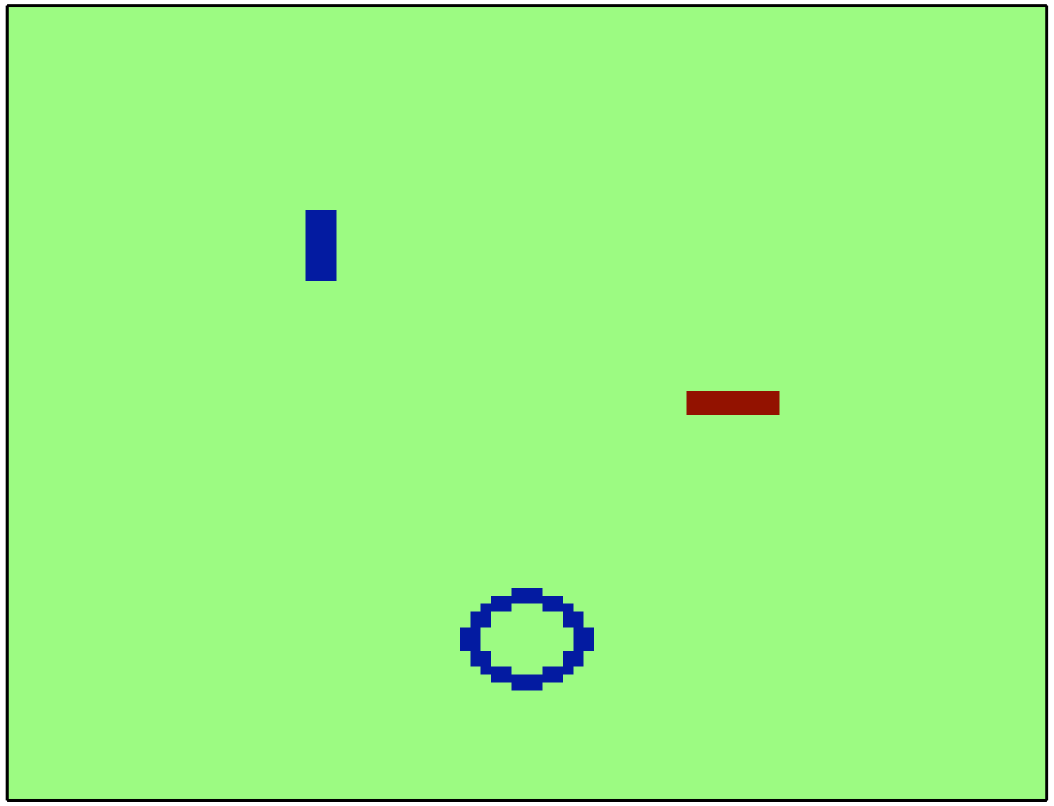
Numerical results

- preliminary tests on model that is sparse in pixels
- cross-well setting, 101 sources and receivers.
- solve

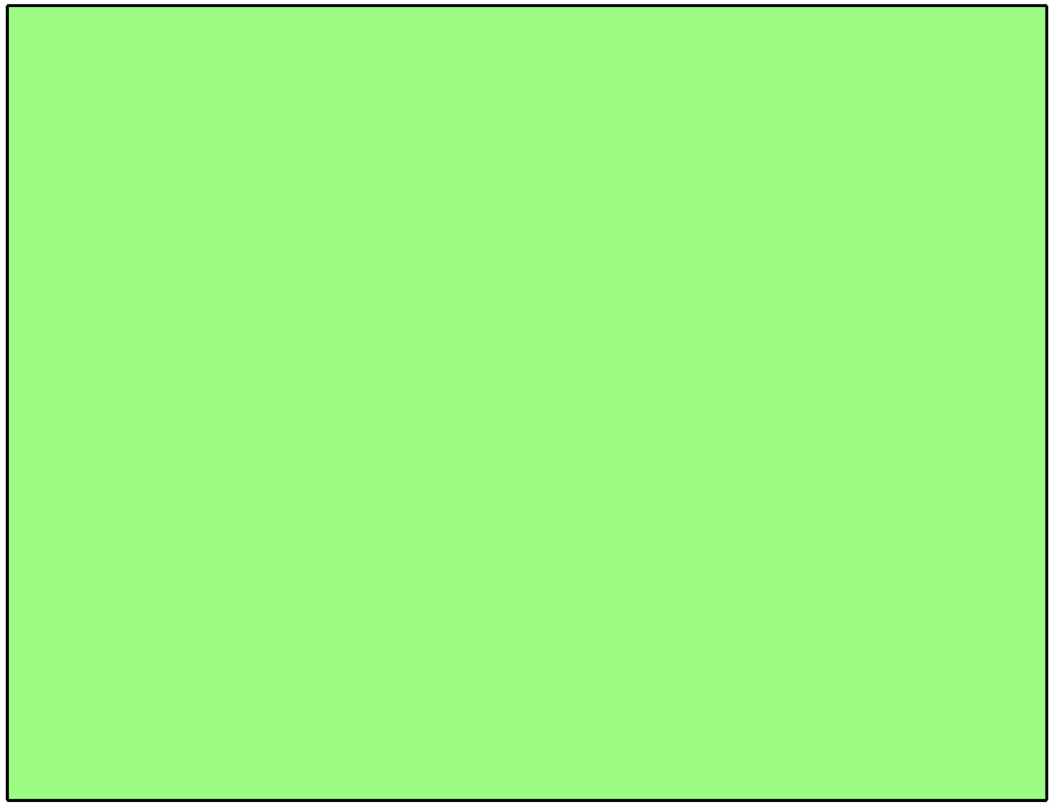
$$\min_{x} \quad \|m\|_1 \qquad \text{s.t.} \quad \|D - RH^{-1}[m_0 + m]Q\|_F^2 \leq \sigma$$

- use simultaneous sources to reduce computational load

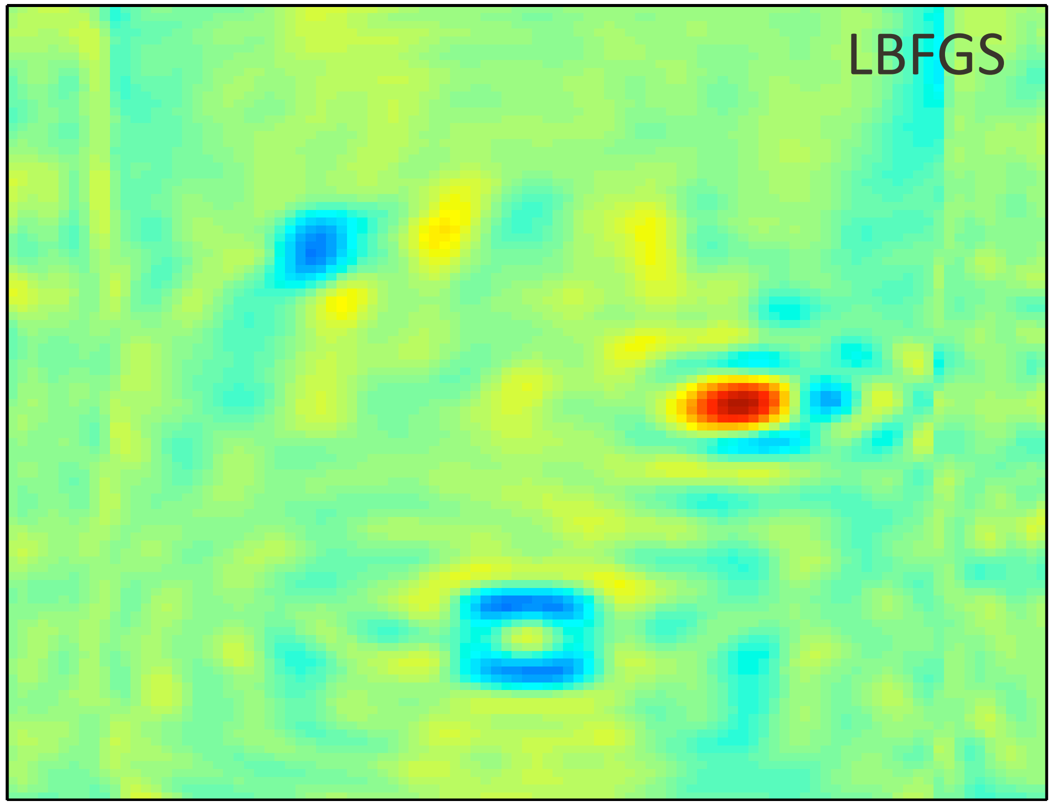
True Model



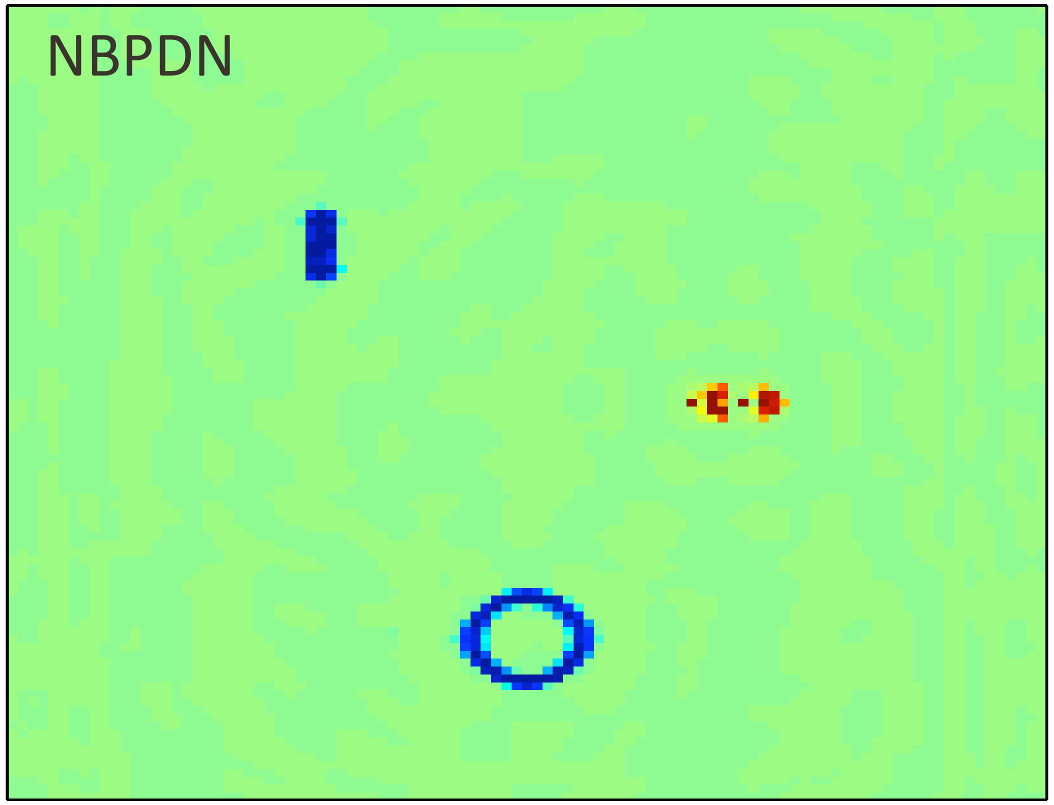
Initial Model

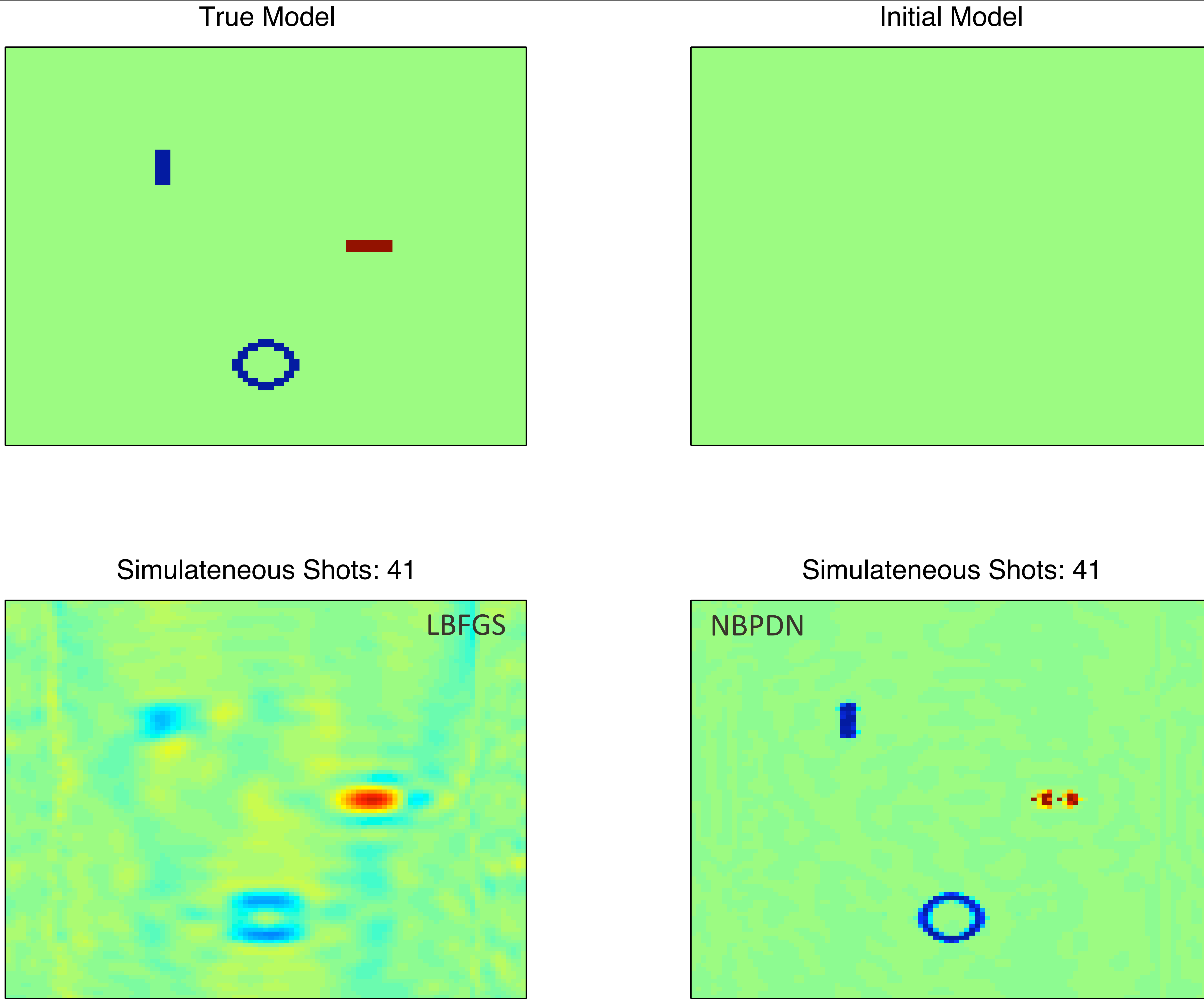


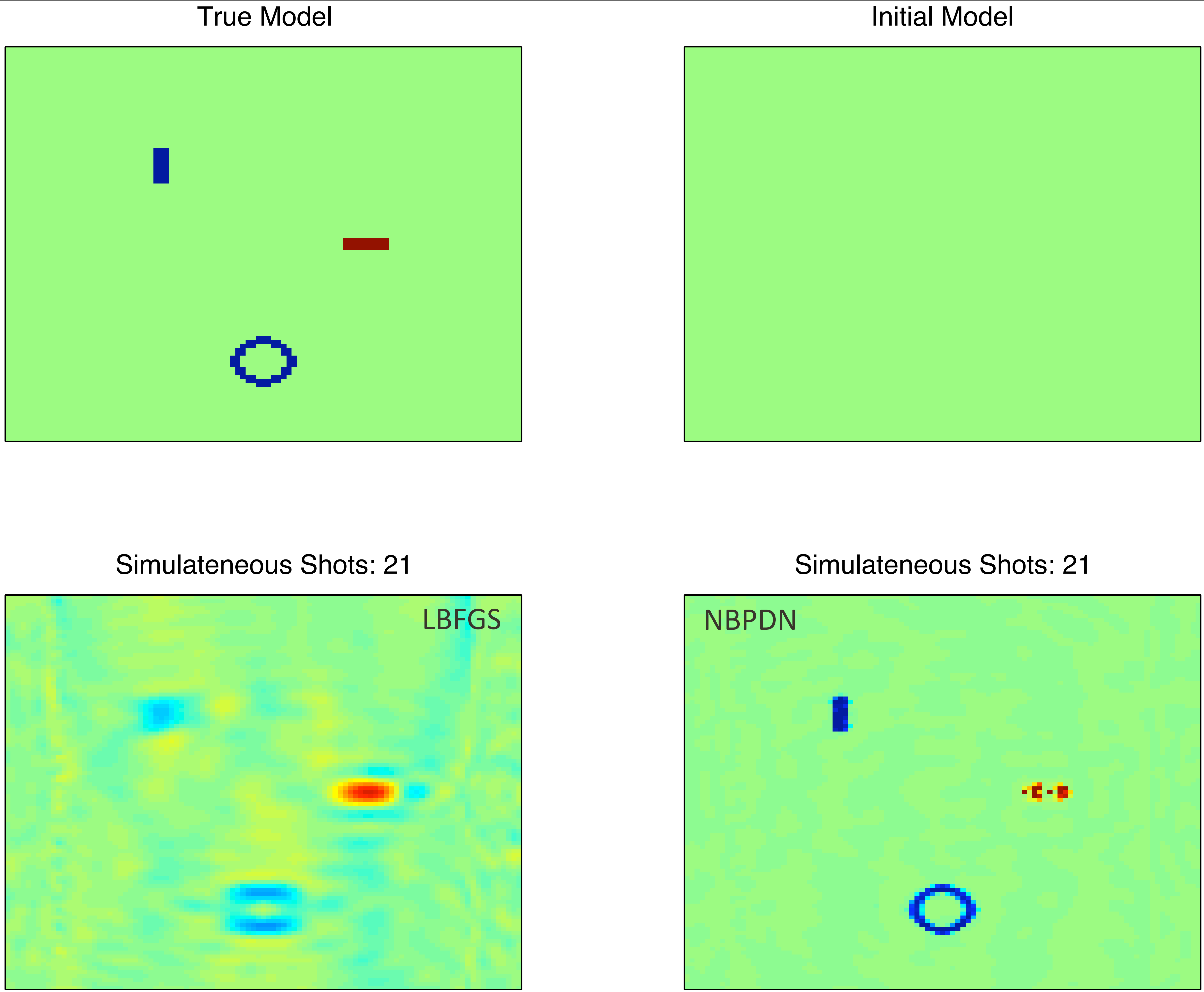
Simulateneous Shots: Full

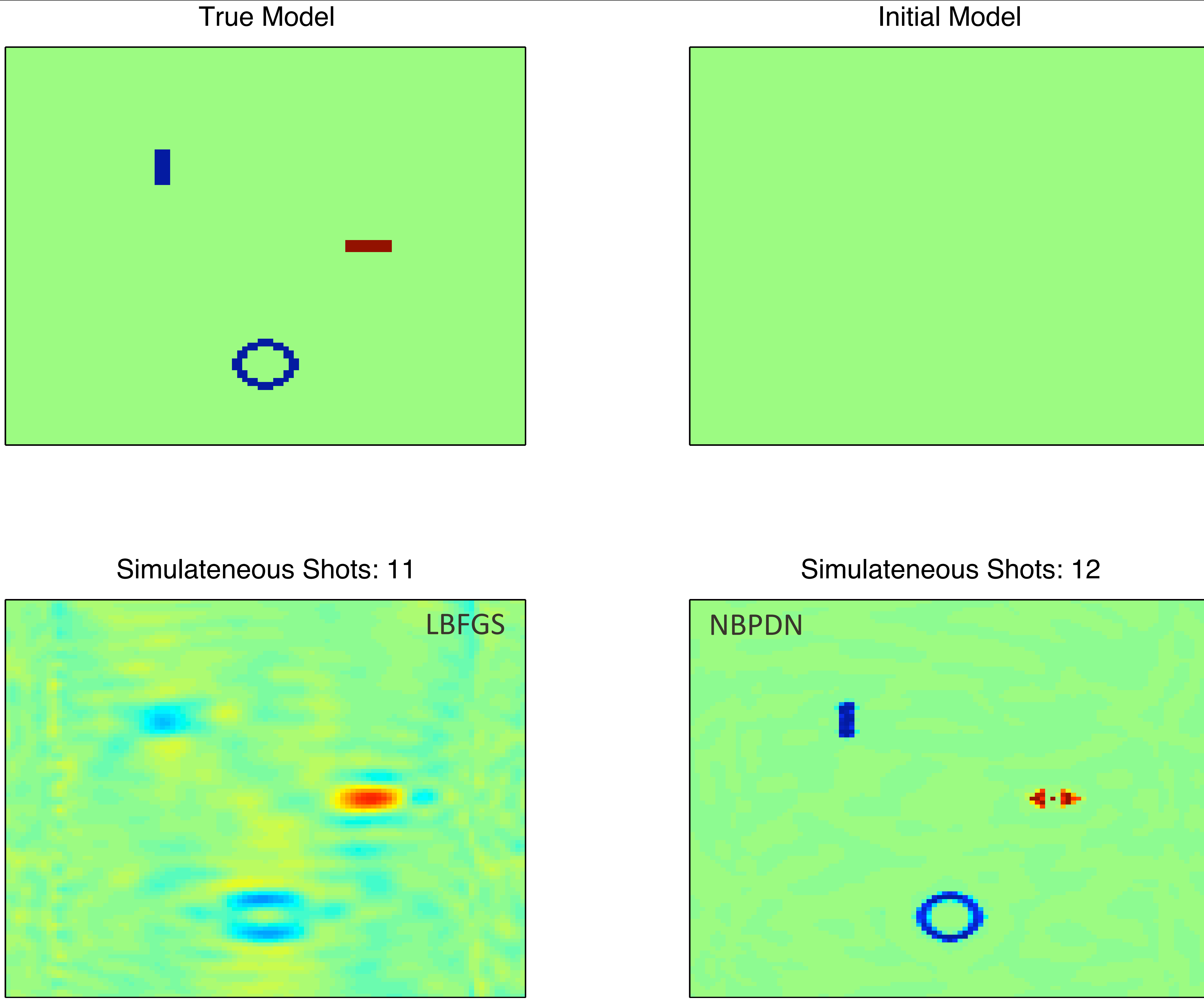


Simulateneous Shots: Full

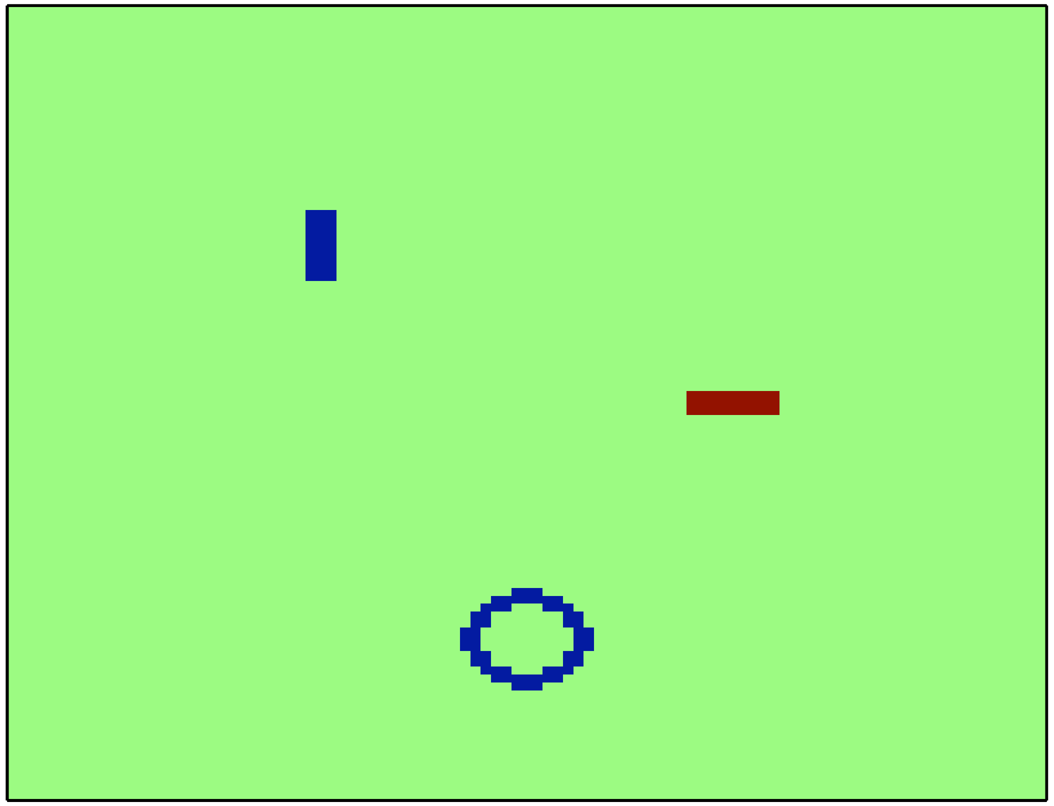



True Model
Initial Model
Simulateneous Shots: 41
LBFGS
Simulateneous Shots: 41
NBPDN

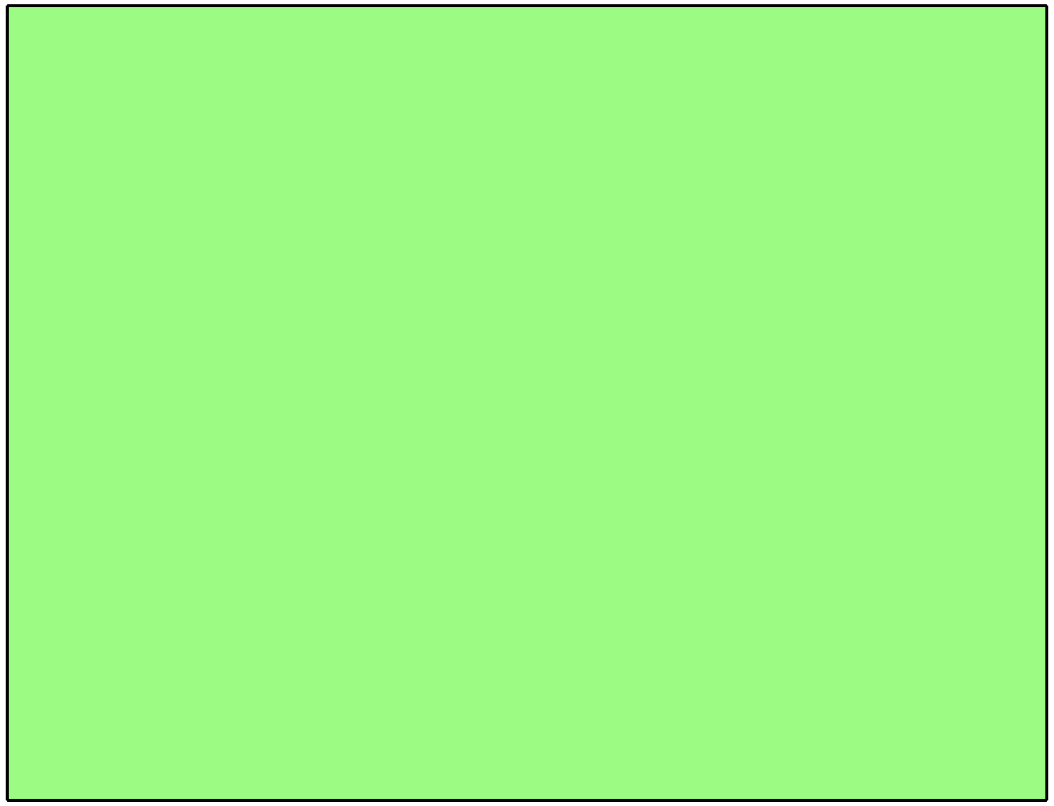

True Model
Initial Model
Simulateneous Shots: 21
LBFGS
Simulateneous Shots: 21
NBPDN


True Model
Initial Model
Simulateneous Shots: 11
LBFGS
Simulateneous Shots: 12
NBPDN

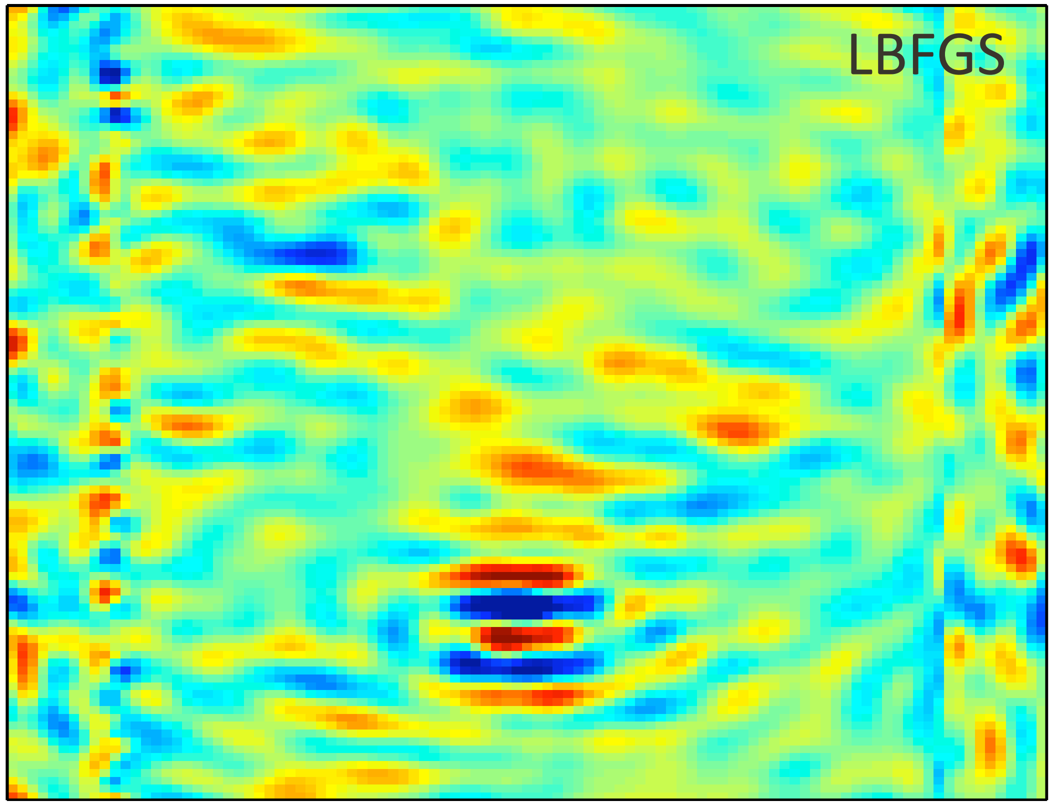
True Model



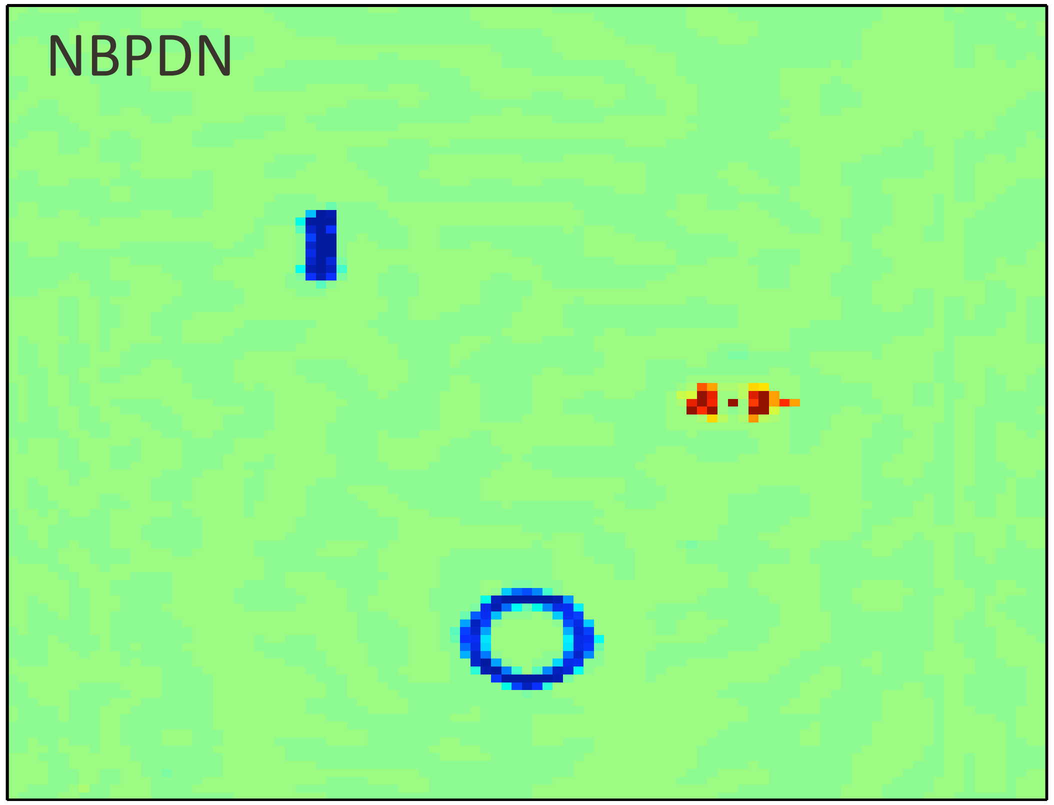
Initial Model

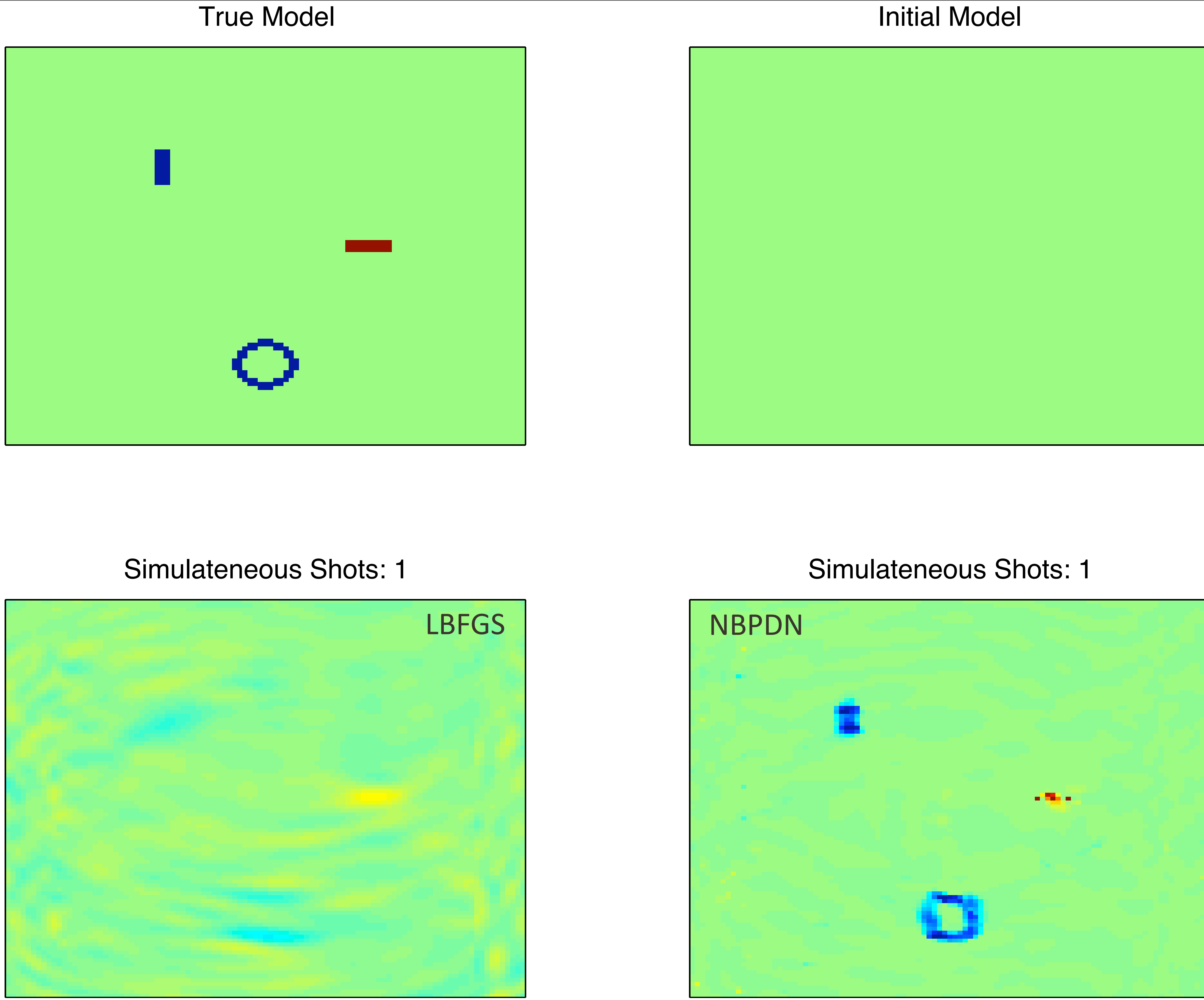


Simulateneous Shots: 5



Simulateneous Shots: 5




True Model
Initial Model
Simulateneous Shots: 1
LBFGS
Simulateneous Shots: 1
NBPDN

SLIM

Conclusions

- Way to introduce sparsity constraints into FWI
- Non-linear formulation of BDPN does not require us to guess sparsity level of the solution
- Preliminary results are promising: we can recover a sparse solution from undersampled data.

The Road Ahead

- Test on realistic models with Curvelet sparsity and noise on data
- Apply to time-lapse seismic
- Implement renewal strategy for simultaneous shots
- Extend the entire framework to be robust against outliers (i.e, different norms on the data residual)
- Investigate relation to compressed sensing L1 recovery

SLIM

Acknowledgements

- Thanks to James Burke (UW Mathematics)

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, BP, Chevron, ConocoPhillips, Petrobras, Total SA, and WesternGeco.

References

Burke, J.V., 1989, A sequential quadratic programming method for potentially infeasible mathematical programs, *Journal of Mathematical Analysis and Applications, 139,2:319-351*

Burke, J.V., 1992, A robust trust region method for constrained nonlinear programming problems, *Siam J. Optimization, 2,2:325-347, 1992*

Burke, J.V., and Ferris, M.C, A Gauss-Newton method for convex composite optimization, *Technical Report, University of Wisconsin-Madison, 1993*

R. Fletcher, Second order correction for nondifferentiable optimization, in *G. A. Watson, editor, Numerical Analysis, 85-114, Springer-Verlag, Berlin, 1982*

R. S. Womerseley, Local properties of algorithms for minimizing composite functions, *Mathematical Programming, 32:69-89, 1985*