Solving WRI's data-augmented wave equation in 3D

Bas Peters
Joint work with Chen Greif & Felix J. Herrmann

SINBAD Consortium Meeting Spring 2015, June 5

SLIM
University of British Columbia

Problem of interest

$$\bar{\mathbf{u}} = \arg\min_{\mathbf{u}} \left\| \begin{pmatrix} \lambda H(\mathbf{m}) \\ P \end{pmatrix} \mathbf{u} - \begin{pmatrix} \lambda \mathbf{q} \\ \mathbf{d} \end{pmatrix} \right\|_2$$

$H(\mathbf{m}) \in \mathbb{C}^{N \times N}$ discrete PDE

$\mathbf{m} \in \mathbb{R}^{N}$ medium parameters

$P \in \mathbb{R}^{m \times N}$ selects field at receivers

$\mathbf{u} \in \mathbb{C}^{N}$ field

$\mathbf{d} \in \mathbb{C}^{m}$ observed data

$\mathbf{q} \in \mathbb{C}^{N}$ source

SLIM

[T. van Leeuwen & F.J. Herrmann, 2013]

WRI

To minimize: $$\min_{\mathbf{m}} \frac{1}{2}\|P\bar{\mathbf{u}} - \mathbf{d}\|_2^2 + \frac{\lambda^2}{2}\|H(\mathbf{m})\bar{\mathbf{u}} - \mathbf{q}\|_2^2$$

at every outer iteration:

- compute $$\bar{\mathbf{u}} = \arg\min_{\mathbf{u}} \left\| \begin{pmatrix} \lambda H(\mathbf{m}) \\ P \end{pmatrix} \mathbf{u} - \begin{pmatrix} \lambda \mathbf{q} \\ \mathbf{d} \end{pmatrix} \right\|_2$$
- evaluate $\bar{\phi}(\mathbf{m}, \bar{\mathbf{u}}, \lambda)$ & $\nabla_{\mathbf{m}}\bar{\phi}(\mathbf{m}, \bar{\mathbf{u}}, \lambda)$
- update $\mathbf{m}$

SLIM

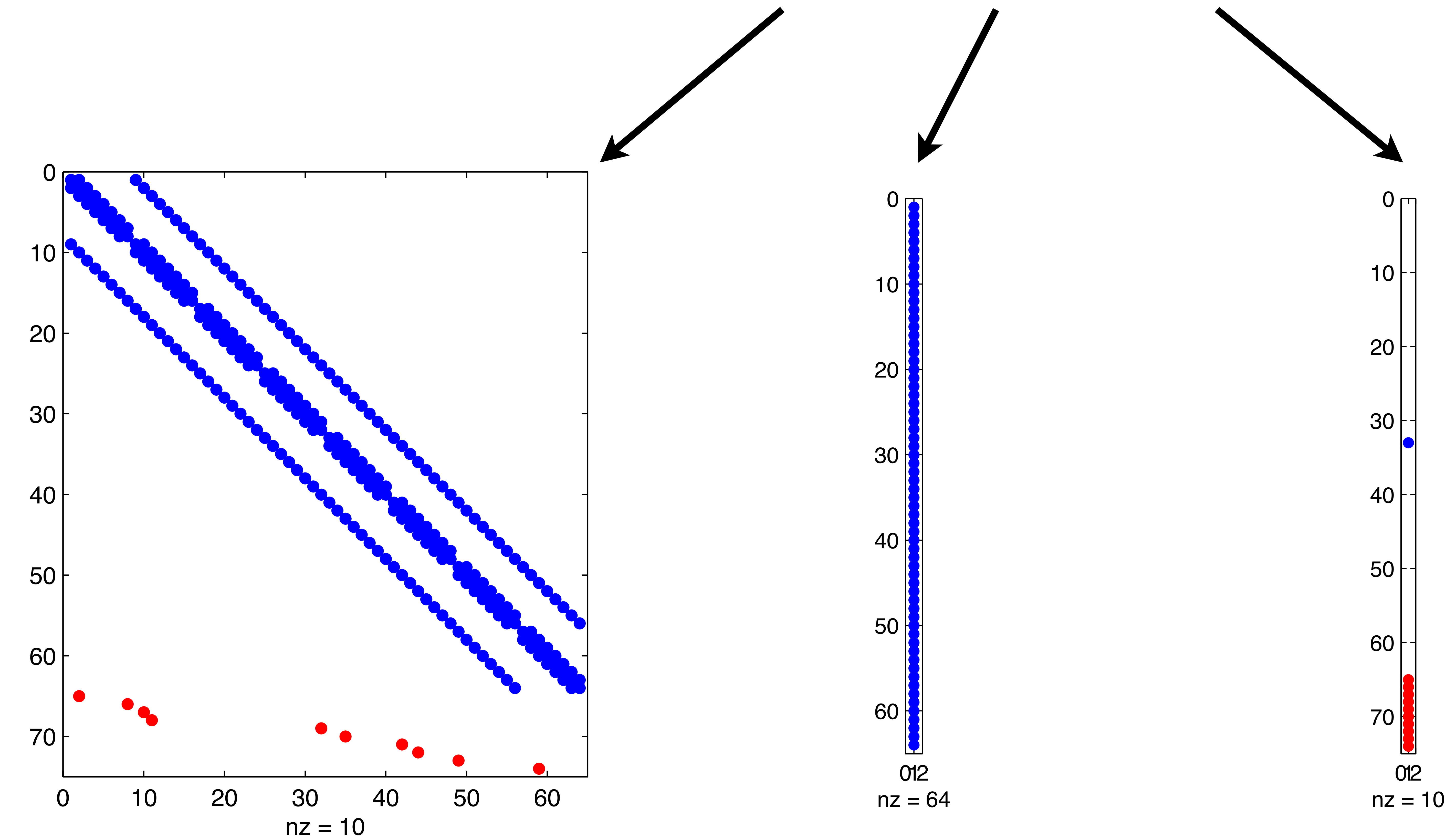
Properties of the problem

$$\bar{\mathbf{u}} = \arg\min_{\mathbf{u}} \left\| \begin{pmatrix} \lambda H(\mathbf{m}) \\ P \end{pmatrix} \mathbf{u} - \begin{pmatrix} \lambda \mathbf{q} \\ \mathbf{d} \end{pmatrix} \right\|_2$$

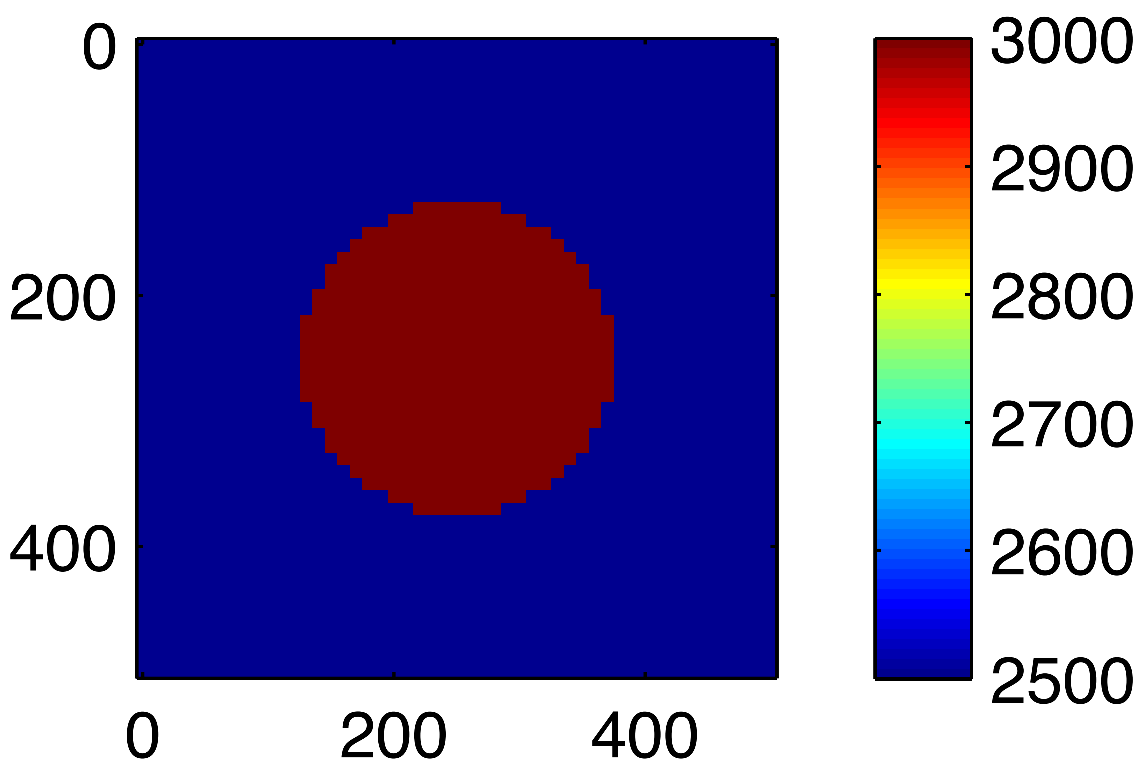
- H is indefinite, non-Hermitian
- inconsistent
- full column rank

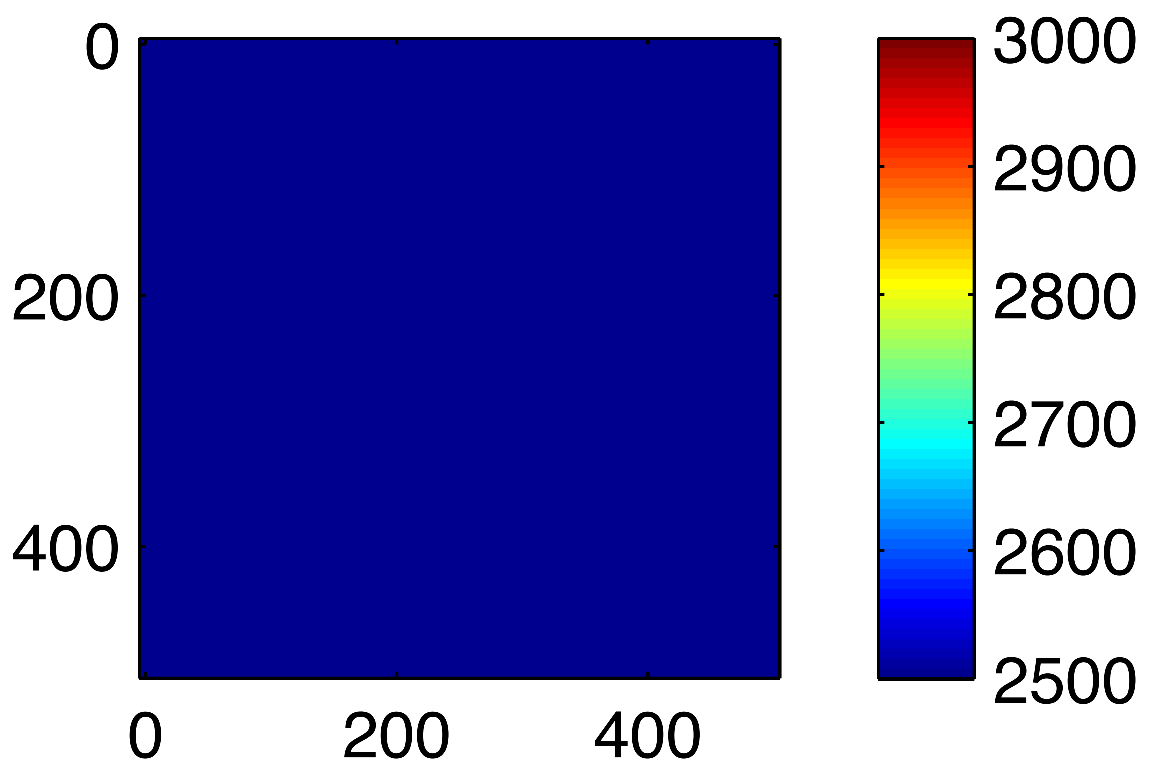
Properties of the problem

$$\bar{\mathbf{u}} = \arg\min_{\mathbf{u}} \left\| \begin{pmatrix} \lambda H(\mathbf{m}) \\ P \end{pmatrix} \mathbf{u} - \begin{pmatrix} \lambda \mathbf{q} \\ \mathbf{d} \end{pmatrix} \right\|_2$$



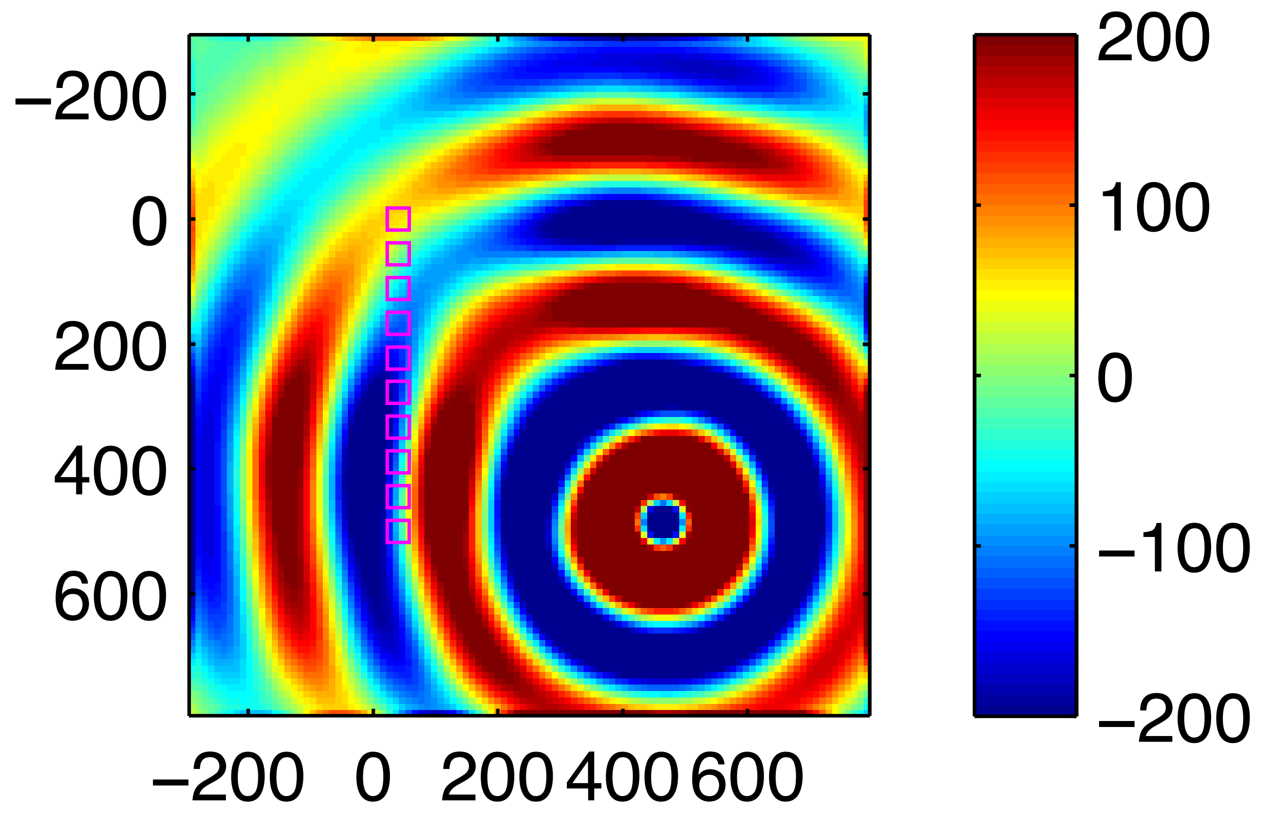
2D example

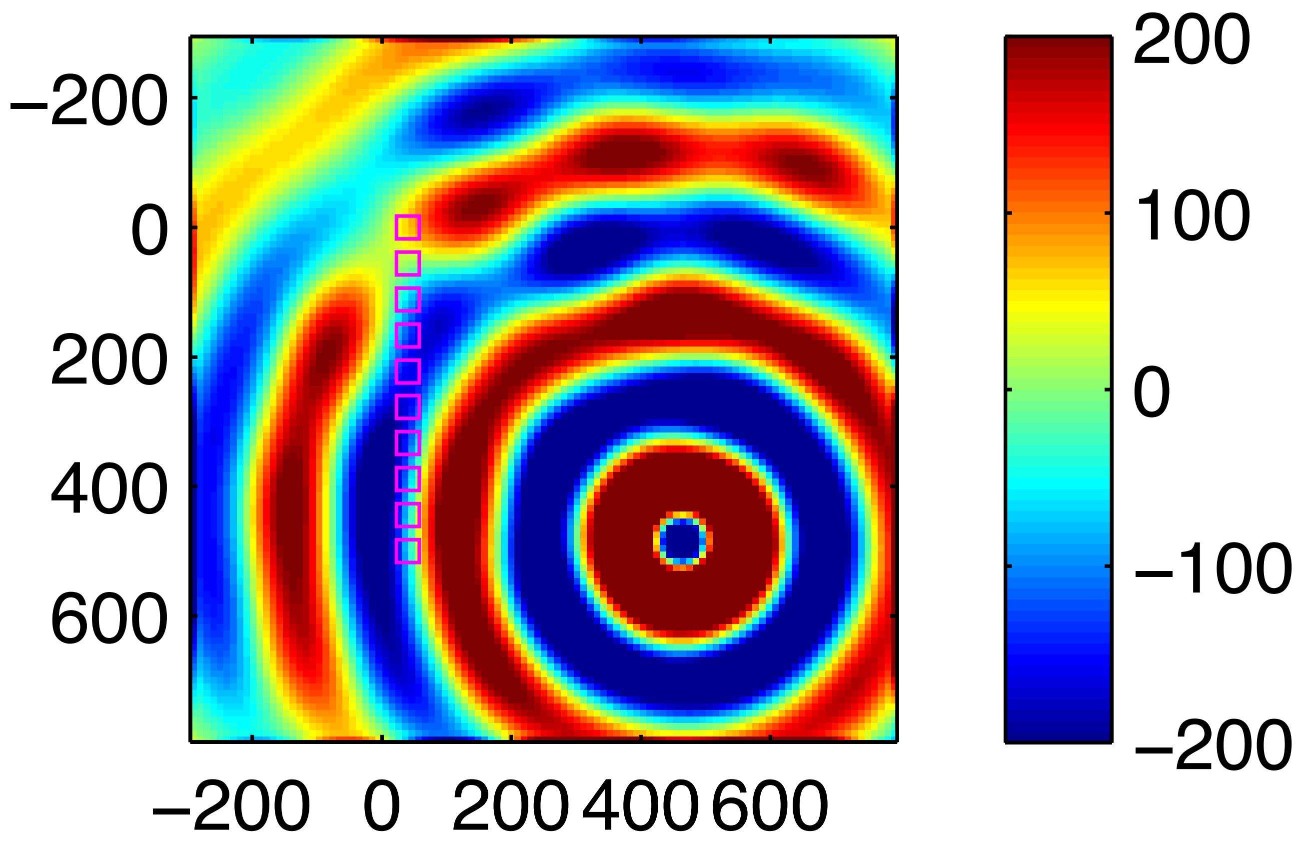


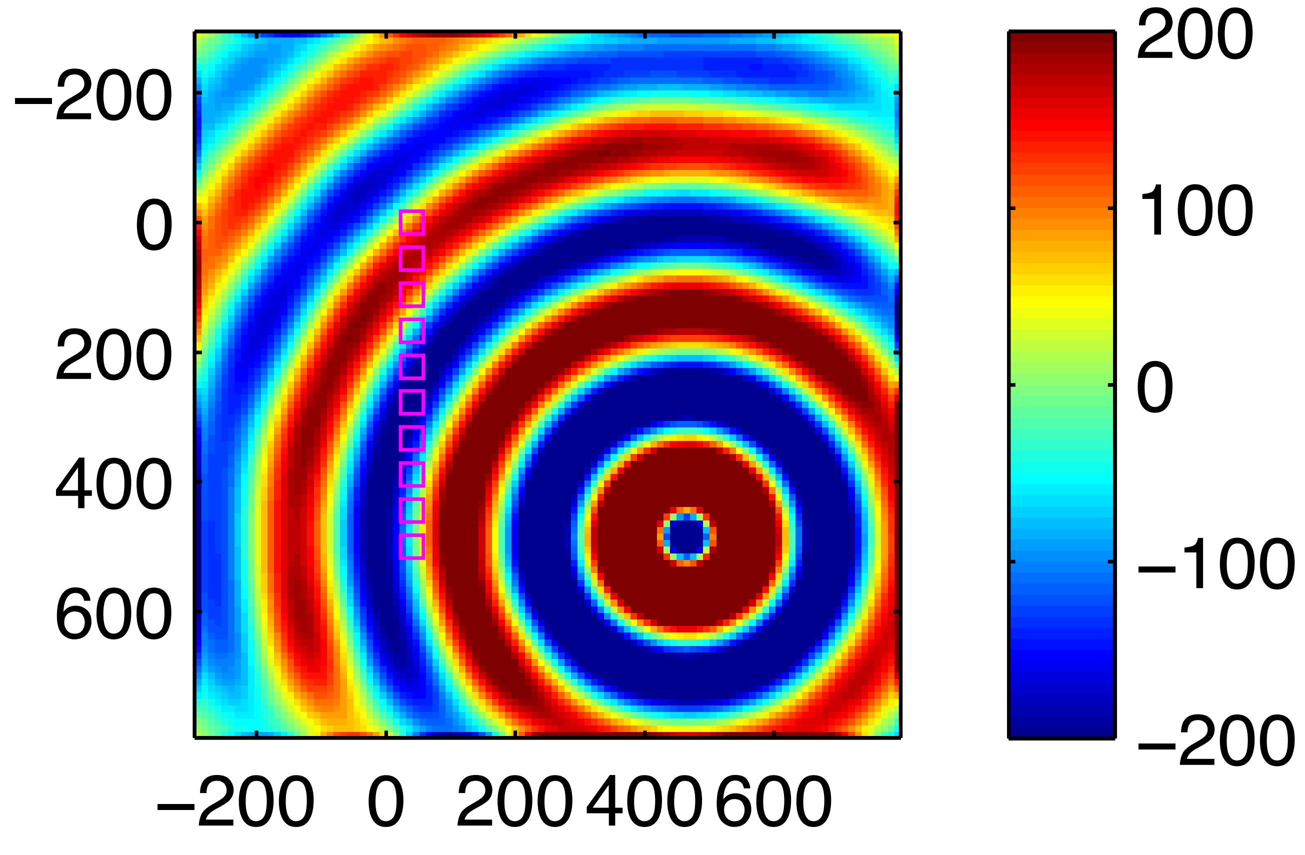


$$\mathbf{u} = H(\mathbf{m}_*)^{-1}\mathbf{q}$$

$$\mathbf{u} = H(\mathbf{m}_0)^{-1}\mathbf{q}$$







$$\bar{\mathbf{u}} = \arg\min_{\mathbf{u}} \left\| \begin{pmatrix} \lambda H(\mathbf{m}_0) \\ P \end{pmatrix} \mathbf{u} - \begin{pmatrix} \lambda \mathbf{q} \\ \mathbf{d} \end{pmatrix} \right\|_2$$

Solution of the sub-problem

Main challenge: solve $$\bar{\mathbf{u}} = \arg\min_{\mathbf{u}} \left\| \begin{pmatrix} \lambda H(\mathbf{m}) \\ P \end{pmatrix} \mathbf{u} - \begin{pmatrix} \lambda \mathbf{q} \\ \mathbf{d} \end{pmatrix} \right\|_2$$

- 2D: direct factorization **[L. M. Delves & I. Barrodale, 1979 ; T. A. Davis, 2011]**

In 3D we want:

- iteratively & matrix-free
- no QR or LU factorizations
- at cost cost of a few PDE solves

Proposed algorithm

LS-problem in normal-equation form:

$$\left(\lambda^2 H(\mathbf{m})^* H(\mathbf{m}) + P^* P\right)\bar{\mathbf{u}} = \lambda^2 H(\mathbf{m})\mathbf{q} + P^*\mathbf{d}$$

Split-preconditioning by λH w/o computations

$$(I + H_\lambda^{-*} P^* P H_\lambda^{-1})\mathbf{y} = \lambda \mathbf{q} + (H_\lambda^*)^{-1} P^* \mathbf{d}, \quad \text{with} \quad H_\lambda \bar{\mathbf{u}} = \mathbf{y}$$

- $m+1$ distinct eigenvalues (identity + low-rank)
- even for inexact Helmholtz

Proposed algorithm

Exploit identity + low-rank structure:

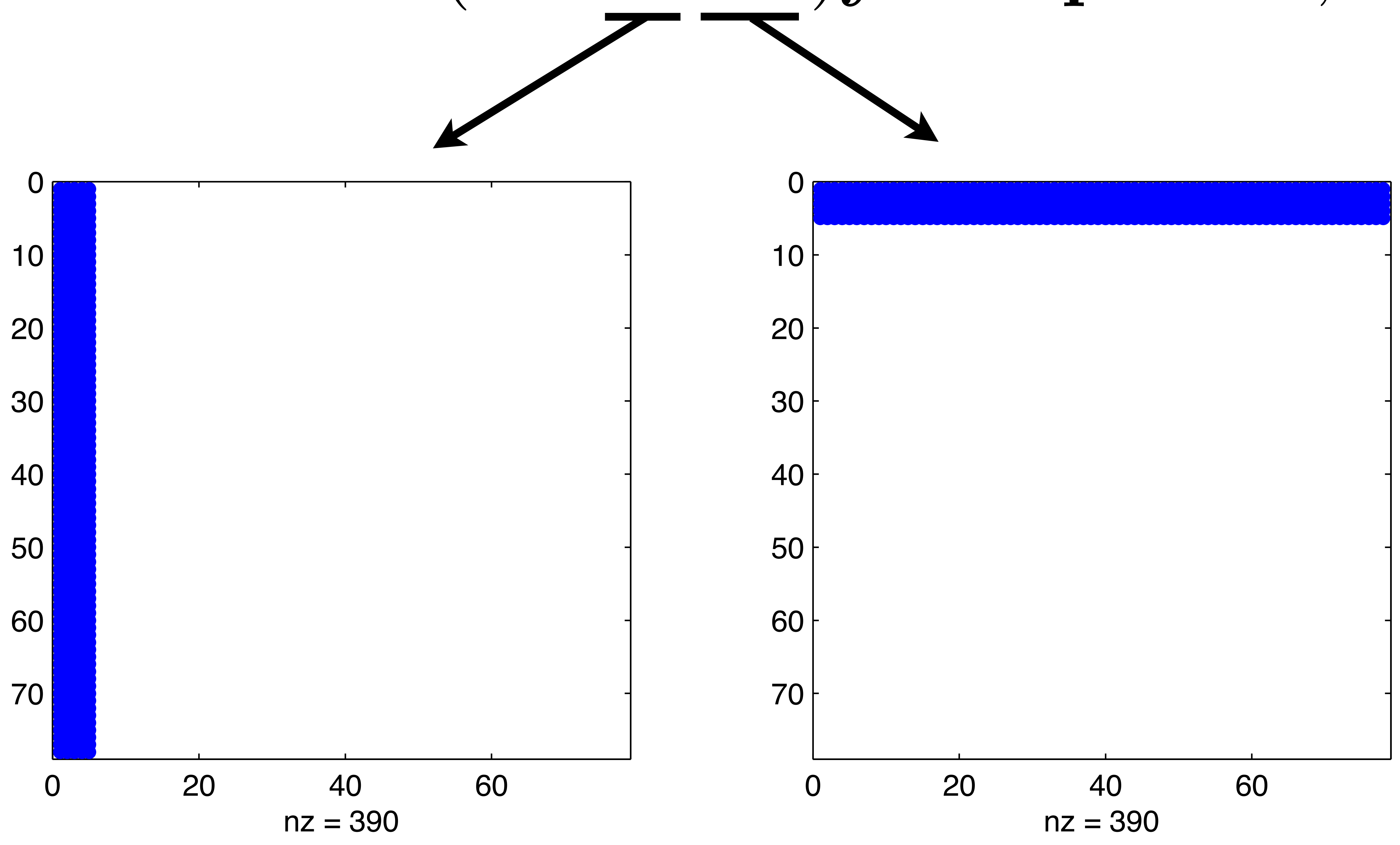
$$(I + H_\lambda^{-*} P^* P H_\lambda^{-1})\mathbf{y} = \lambda \mathbf{q} + (H_\lambda^*)^{-1} P^* \mathbf{d}, \quad \text{with} \quad H_\lambda \bar{\mathbf{u}} = \mathbf{y}$$

by solving $H^{-*} P^* = W$

- n_{rec} Helmholtz problems (inexactly)
- low-rank factorization
- $W \in \mathbb{C}^{N \times m}$ dense but extremely skinny

Proposed algorithm

$$(I + WW^*)\mathbf{y} = \lambda\mathbf{q} + W\mathbf{d}, \quad \text{with} \quad H_\lambda \bar{\mathbf{u}} = \mathbf{y}$$



inverses transformed into identity + low-rank factorization

Proposed algorithm

Leverage low-rank factorization:

$$(I + WW^*)\mathbf{y} = \lambda\mathbf{q} + W\mathbf{d}, \quad \text{with} \quad H_\lambda\bar{\mathbf{u}} = \mathbf{y}$$

and invert system matrix as

$$\mathbf{y} = (I - W(I + W^*W)^{-1}W^*)(\lambda\mathbf{q} + W\mathbf{d}), \quad \text{with} \quad H_\lambda\bar{\mathbf{u}} = \mathbf{y}$$

so we only need to invert $(I + W^*W) \in \mathbb{C}^{m\times m}$

(this is alway small enough to do explicitly)

Proposed algorithm

for angular frequency ω **do**
  // solve m Helmholtz problems
  $H_{\lambda}^{*} W = P^{*}$
  $M = (I + W^{*}W)^{-1}$
  for right hand side i **do**
    $\mathbf{y}_i = \left(I - WMW^{*}\right)\left(\lambda \mathbf{q}_i + W\mathbf{d}_i\right)$
    // solve for $\bar{\mathbf{u}}_i$
    $H_{\lambda}\bar{\mathbf{u}}_i = \mathbf{y}_i$
  end for
end for

Proposed algorithm

Matrix-free algorithm

- no direct solves
- related mildly overdetermined systems **[L. M. Delves & I. Barrodale, 1979]**

Computational cost:

- 1 PDE per receiver
- 1 PDE per source

Memory requirements:

- 1 vector per receiver (W)
- system matrix (H)
- storage for solving systems with H

SLIM

Proposed algorithm

Inexact solutions to the linear systems:

for angular frequency ω **do**
  // solve m Helmholtz problems inexactly
  → $H_\lambda^* \hat{W} = P^* + R_W$
  $\hat{M} = (I + \hat{W}^* \hat{W})^{-1}$
  for right hand side $\mathbf{b}_i$ **do**
    $\hat{\mathbf{y}}_i = \left(I - \hat{W}\hat{M}\hat{W}^*\right)\left(\lambda \mathbf{q}_i + \hat{W}\mathbf{d}_i\right)$
    // solve for $\bar{\mathbf{u}}_i$ inexactly
    → $H_\lambda \hat{\mathbf{u}}_i = \hat{\mathbf{y}}_i + \mathbf{r_u}$
  end for
end for

Proposed algorithm

error propagation (1 right-hand-side, 1 receiver case):

$$H_\lambda^* \hat{\mathbf{w}} = \mathbf{p}^* + \mathbf{r}_\mathbf{w}$$

$$(I + \hat{\mathbf{w}}\hat{\mathbf{w}}^*)\hat{\mathbf{y}} = \lambda \mathbf{q} + \hat{\mathbf{w}} d$$

$$H_\lambda \hat{\mathbf{u}} = \hat{\mathbf{y}} + \mathbf{r}_\mathbf{u}$$

SLIM

Proposed algorithm

error propagation (1 right-hand-side, 1 receiver case):

$$H_\lambda^* \hat{\mathbf{w}} = \mathbf{p}^* + \mathbf{r}_\mathbf{w}$$

$$(I + \hat{\mathbf{w}}\hat{\mathbf{w}}^*)\hat{\mathbf{y}} = \lambda \mathbf{q} + \hat{\mathbf{w}} d$$

$\longrightarrow$ solve as: $\hat{\mathbf{y}} = (I - \hat{m}\hat{\mathbf{w}}\hat{\mathbf{w}}^*)(\lambda \mathbf{q} + \hat{\mathbf{w}} d)$

with $\hat{m} = \dfrac{1}{1 + \hat{\mathbf{w}}^*\hat{\mathbf{w}}}$

$$H_\lambda \hat{\mathbf{u}} = \hat{\mathbf{y}} + \mathbf{r}_\mathbf{u}$$

SLIM

Proposed algorithm

- several bounds derived
- Multi-stage error propagation makes deriving (useful) bounds in terms of observables very challenging.
- Very pessimistic bound may be all what is possible.
- The aim is to obtain (asymptotically) correct dependence on condition numbers / spectral norms of involved matrices and relative residuals.
- work in progress

Suggested PDE-solver

need to store 1 vector per receiver

-> use PDE-solver with low-memory & setup requirements

Helmholtz:

- CGMN (only 4 vectors) / CARP-CG **[A. Bjorck & T. Elfving, 1979; D. Gordon & R. Gordon, 2010; T. van Leeuwen & F.J. Herrmann, 2014]**
- shifted-Laplacian w/ multi-grid **[Y.A. Erlangga, 2008; H. Calandra et al., 2013]**
- combination of the above **[R. Lago & F.J. Herrmann, 2015]**

Simultaneous receivers

Simultaneous sources reduce the number of sources to be modeled.

Can we use similar ideas with the proposed algorithm?

Memory and computational cost now depends on sources + receivers.

Simultaneous receivers

What is the number of receivers is too large, storage wise?

Can we approximate the least-squares problem using randomization & subsampling (simultaneous receivers)?

Use ideas from algorithms such as

- **[V Rokhlin & M Tygert, 2008]**
- Blendenpik **[H. Avron et. al., 2010]**
- LSRN **[X. Meng, M. A. Saunders, M. W. Mahoney, 2014]**

Simultaneous receivers

Blendenpik:

- Randomize (mix the rows) and subsample a very overdetermined system.
- Use R from QR of the approximated and well conditioned problem as a preconditioner for LSQR to solve the original problem.
- Define randomize & subsample matrix as: $V = SFD$,

$$V \in \mathbb{C}^{l \times m}, \quad l < m$$

$$D \in \mathbb{R}^{m \times m} \quad \text{random } [+1 \text{ , -1}] \text{ on the diagonal}$$

$$F \in \mathbb{C}^{m \times m} \quad \text{DFT matrix}$$

$$S \in \mathbb{R}^{l \times m} \quad \text{subsampling matrix, restriction of the identity}$$

Simultaneous receivers

Initial attempt in this work:

apply randomization and subsampling to the receiver block only for a one-step approximation:

$$\bar{\mathbf{u}} = \arg\min_{\mathbf{u}} \left\| \begin{pmatrix} \lambda H(\mathbf{m}) \\ VP \end{pmatrix} \mathbf{u} - \begin{pmatrix} \lambda \mathbf{q} \\ V\mathbf{d} \end{pmatrix} \right\|_2$$

$$V \in \mathbb{C}^{l \times m}, \quad l < m$$

What should V be ? ongoing research, use $V = SFD$

to illustrate the principle

SLIM

Simultaneous receivers

Initial attempt in this work:

apply randomization and subsampling to the receiver block only for a one-step approximation:

$$\bar{\mathbf{u}} = \arg\min_{\mathbf{u}} \left\| \begin{pmatrix} \lambda H(\mathbf{m}) \\ VP \end{pmatrix} \mathbf{u} - \begin{pmatrix} \lambda \mathbf{q} \\ V\mathbf{d} \end{pmatrix} \right\|_2$$

$$V \in \mathbb{C}^{l \times m}, \quad l < m$$

reduces

- # of PDE solves
- # vectors to be stored

Simultaneous receivers

$$\bar{\mathbf{u}} = \arg\min_{\mathbf{u}} \left\| \begin{pmatrix} \lambda H(\mathbf{m}) \\ VP \end{pmatrix} \mathbf{u} - \begin{pmatrix} \lambda \mathbf{q} \\ V\mathbf{d} \end{pmatrix} \right\|_2$$

Approximates system matrix and right hand side as:

$$\left(\lambda^2 H(\mathbf{m})^* H(\mathbf{m}) + P^* V^* V P\right)\bar{\mathbf{u}} = \lambda^2 H(\mathbf{m})\mathbf{q} + P^* V^* V\mathbf{d}.$$

SLIM

Simultaneous receivers

Example:

investigate error in the fields and in the resulting gradient, introduced by subsampling and randomization.

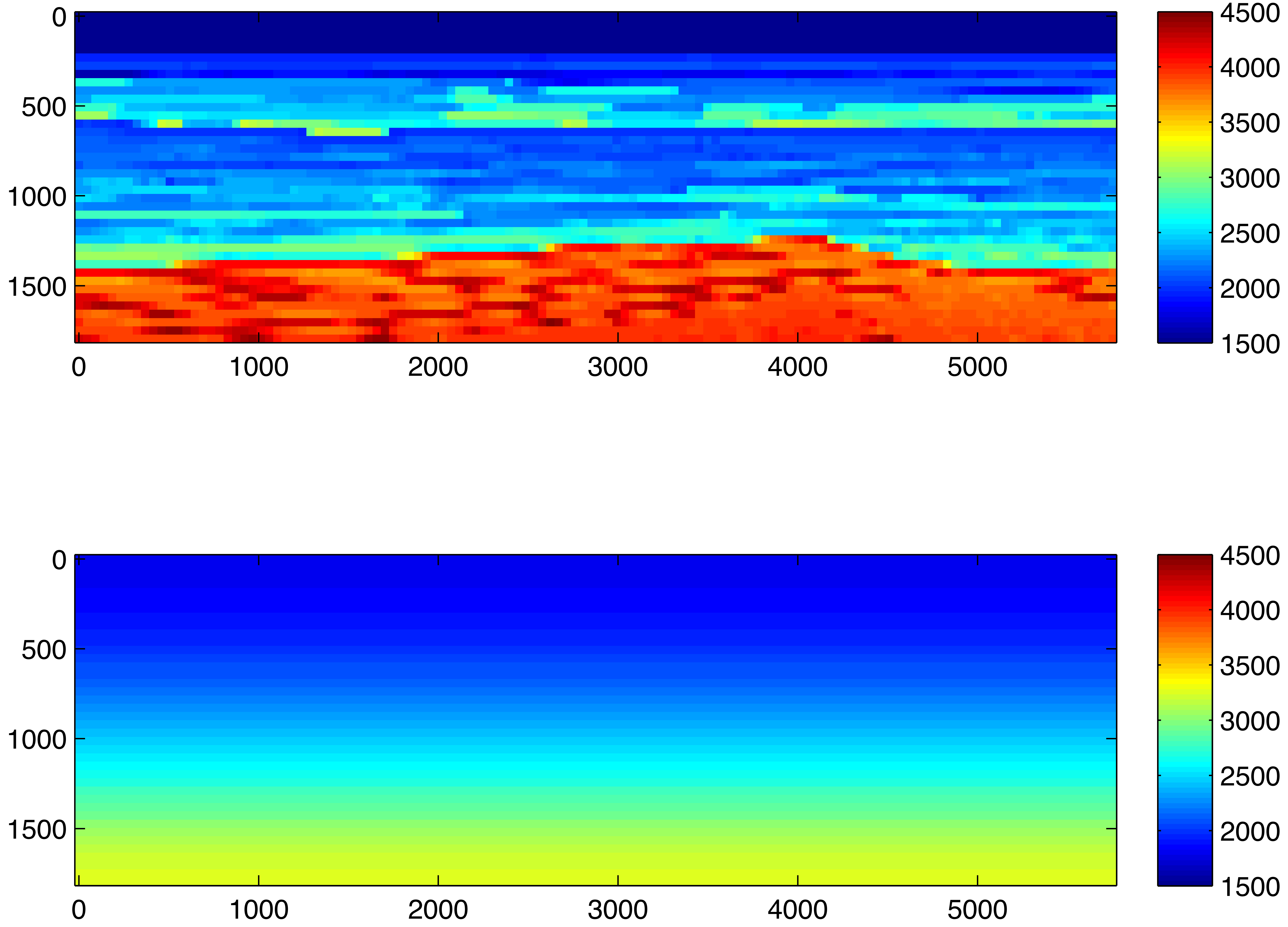
for now, assume:

- solve linear systems exactly
- fixed number of sources (30)
- 1 frequency (4Hz)

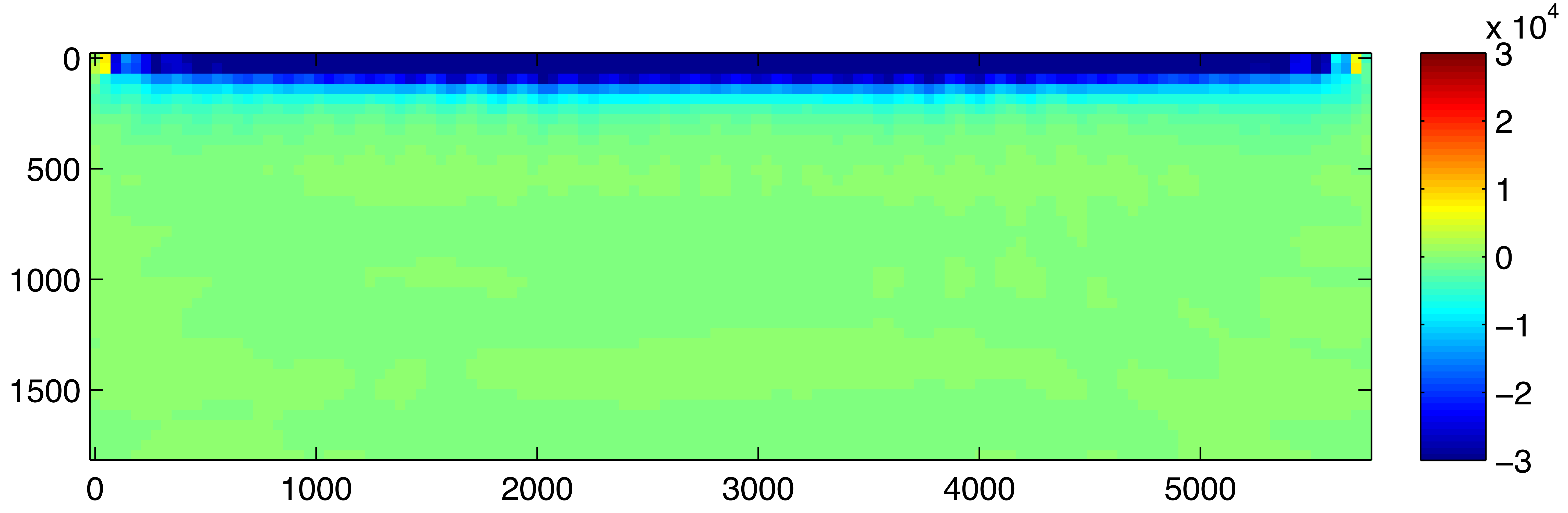
Simultaneous receivers

Example:

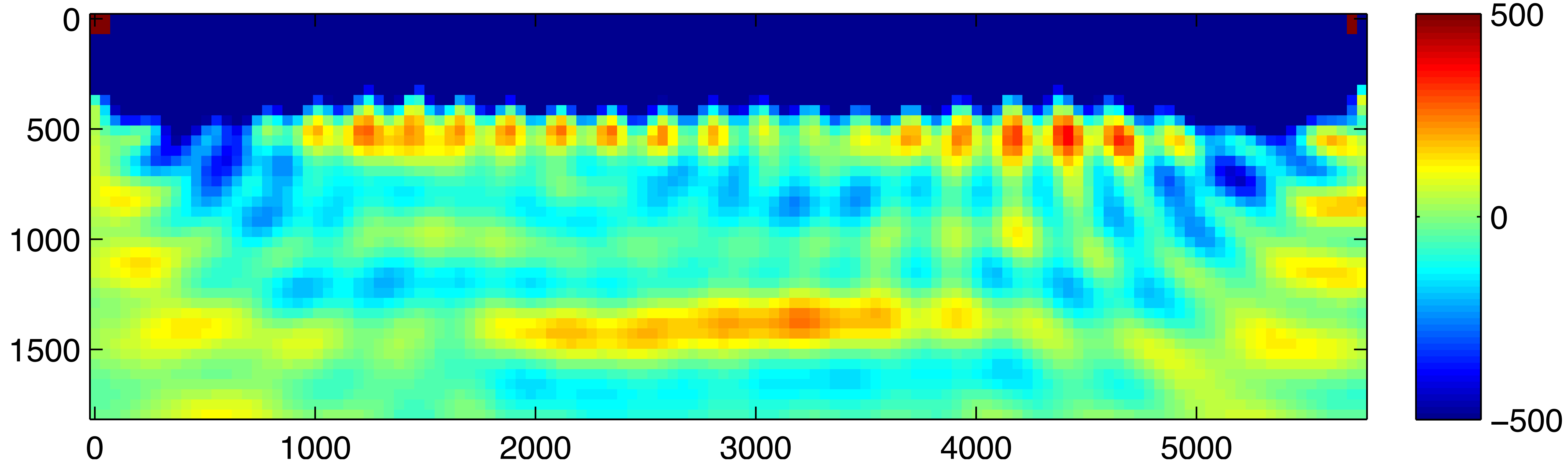
- 100 receivers in total (~55 m interval)
- only work with randomized subsets of varying size (previous slides)
- 30 sources (~200 m interval)
- 1 frequency (4Hz)



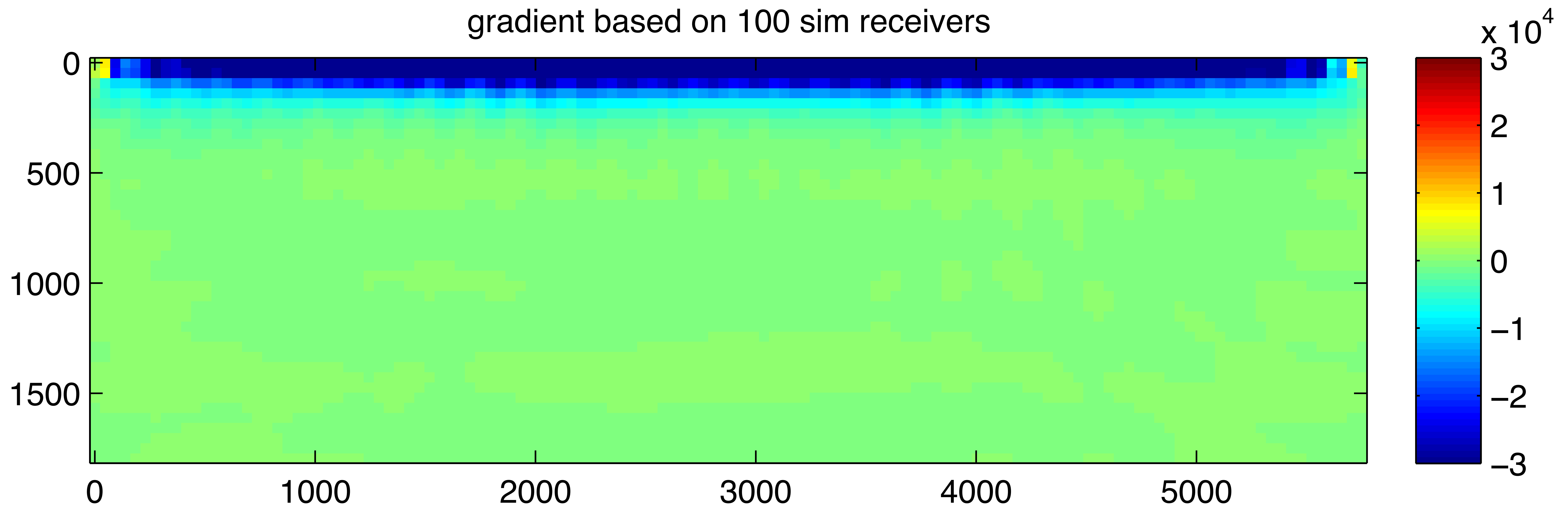
Simultaneous receivers

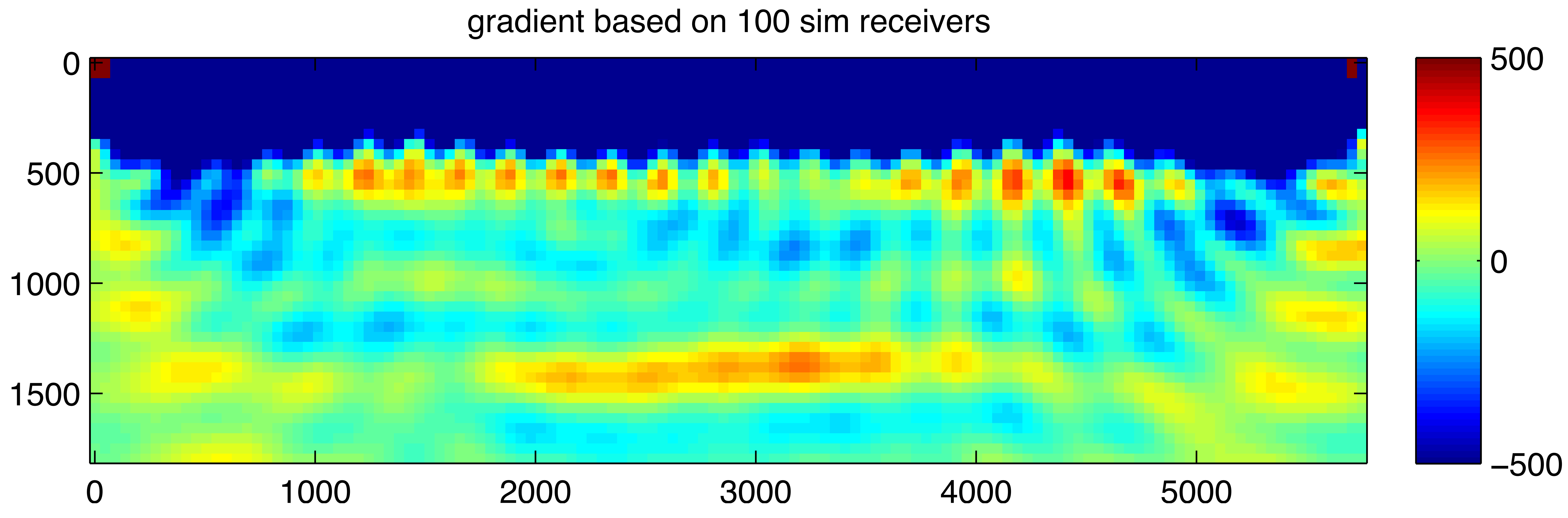


True gradient,
same image twice, different color scale

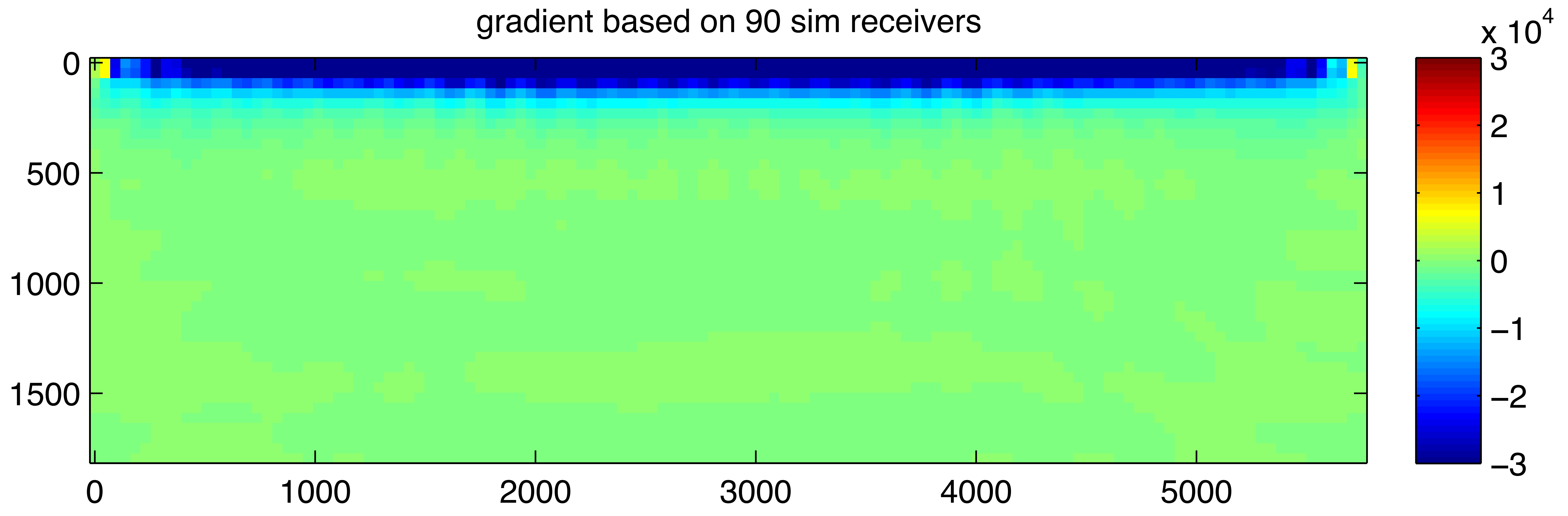


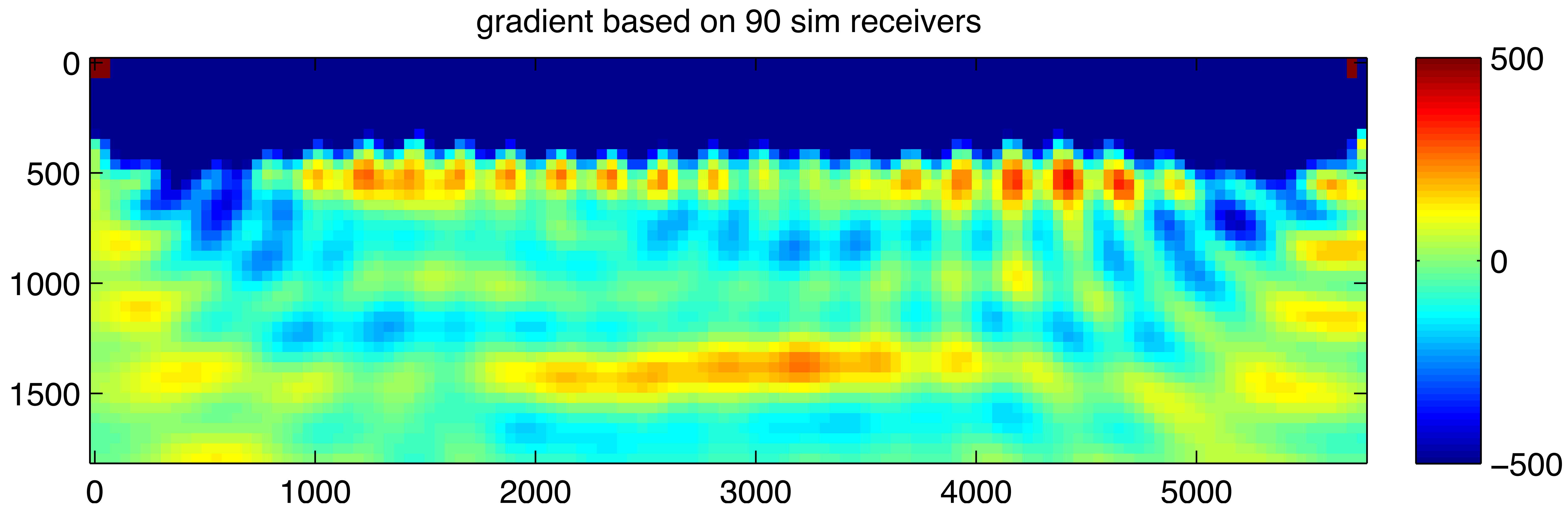
Simultaneous receivers



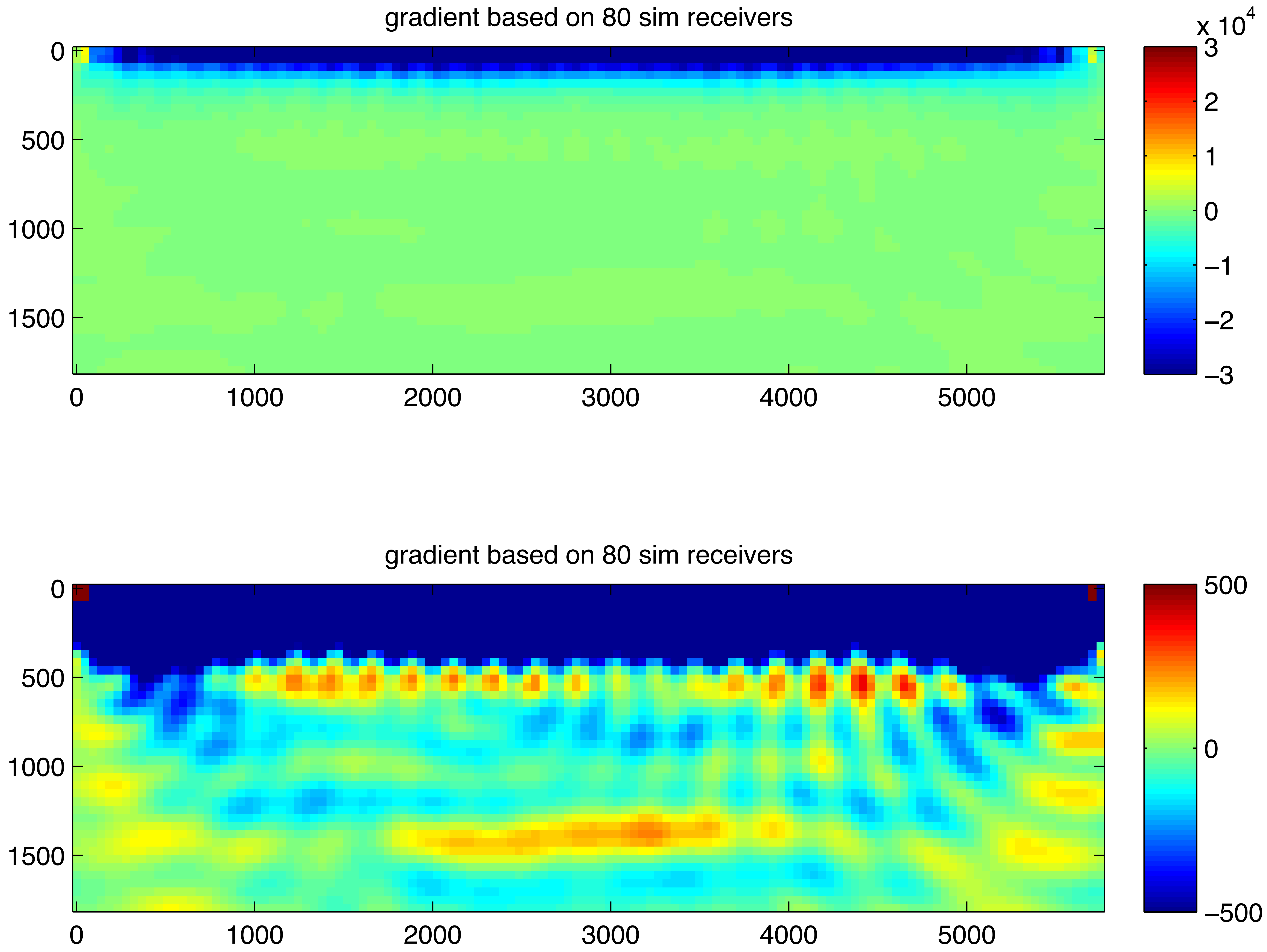


Simultaneous receivers

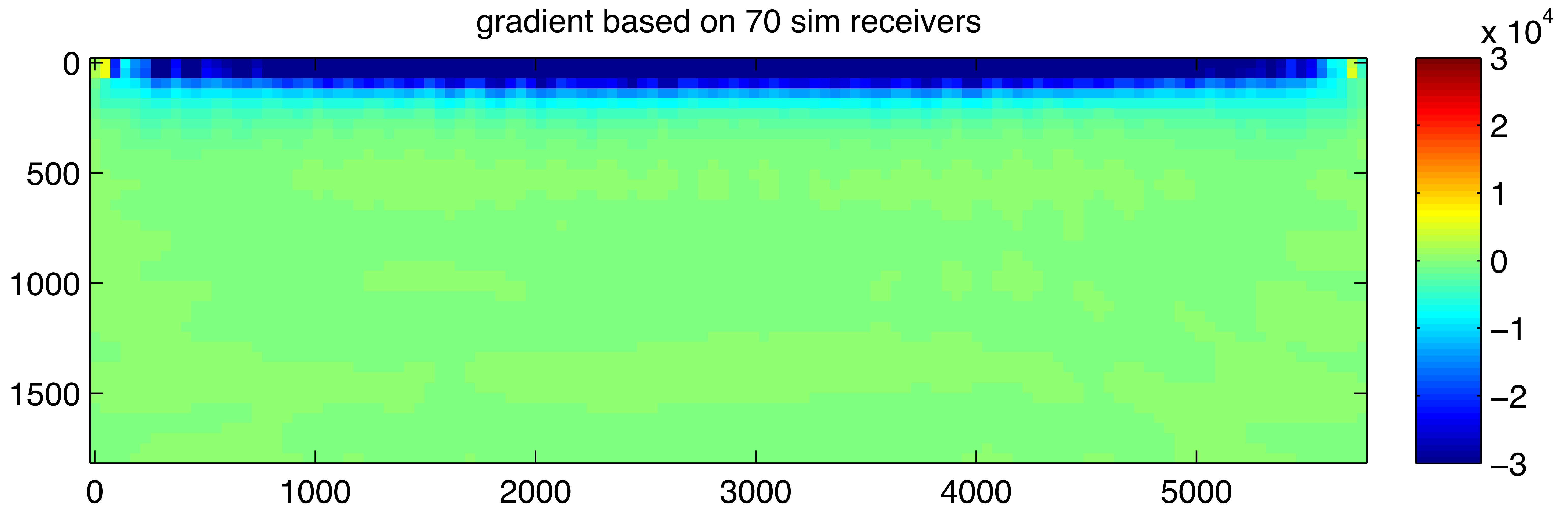


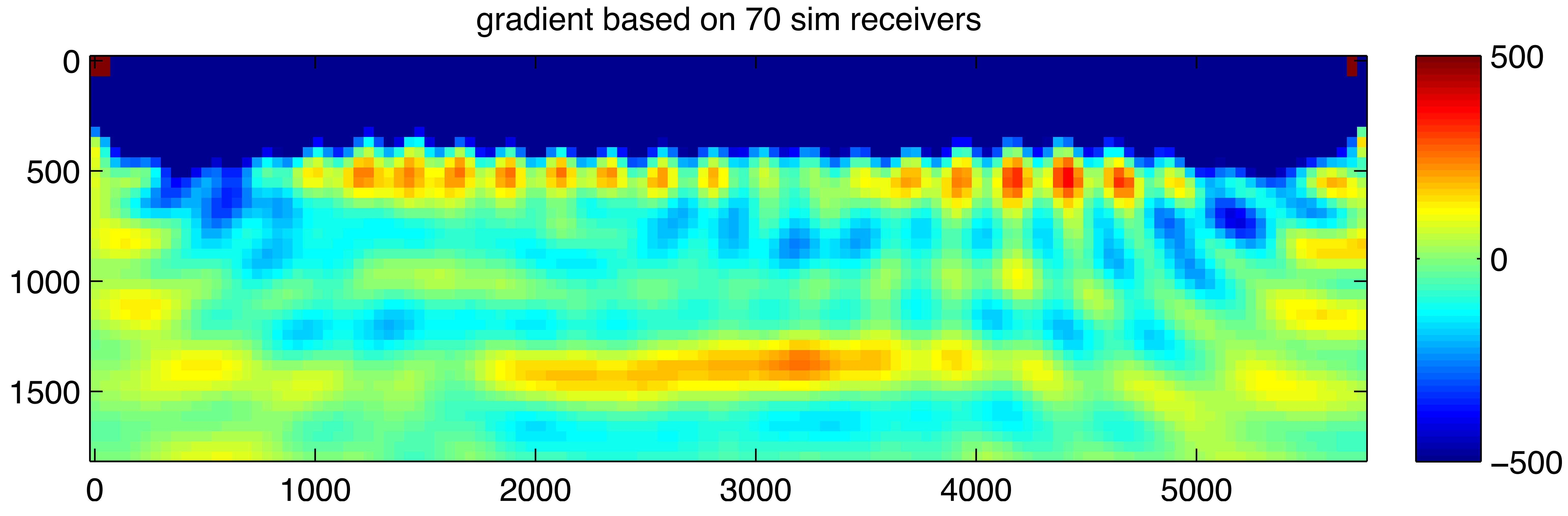


Simultaneous receivers

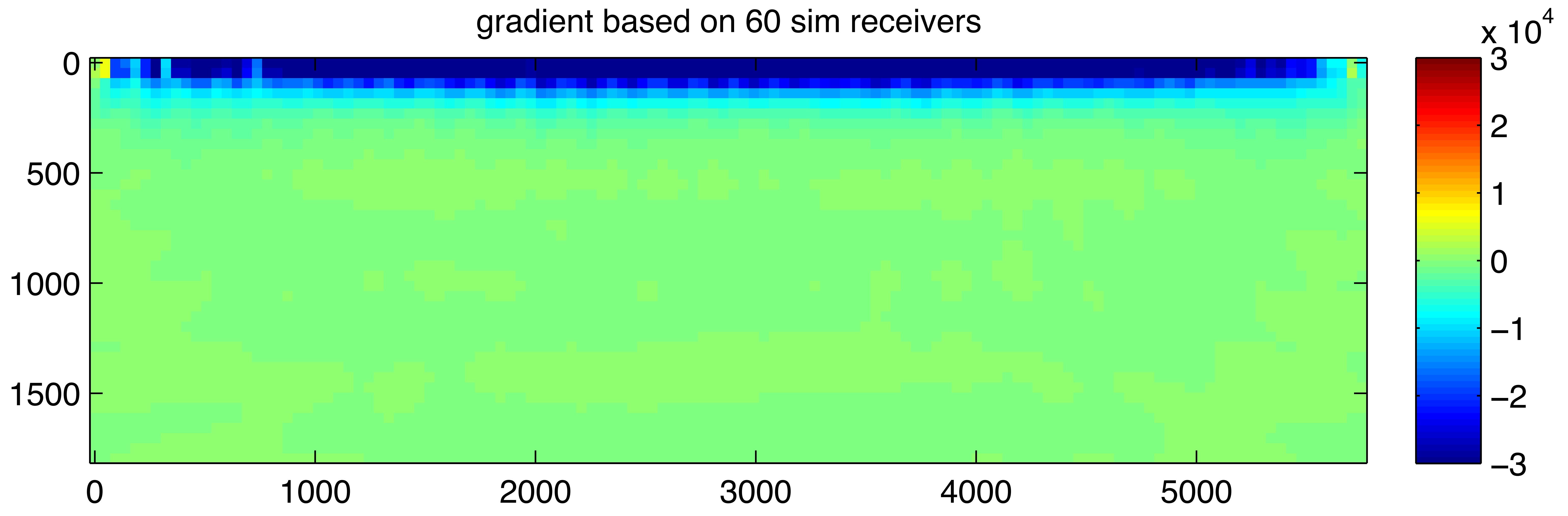


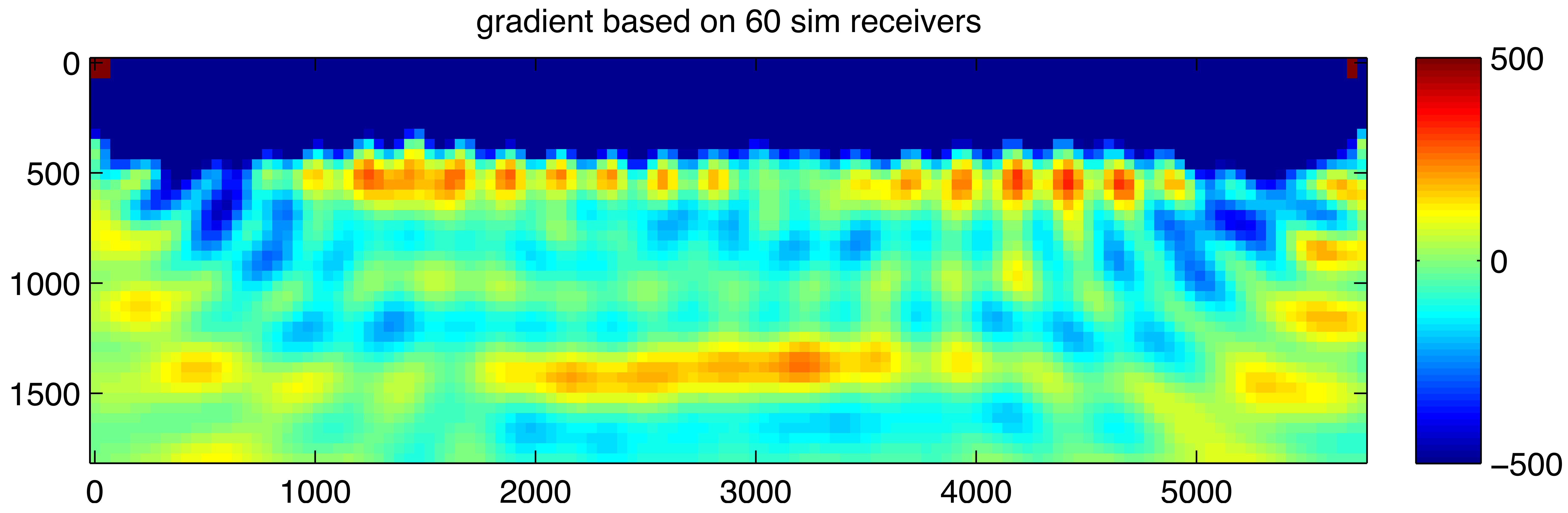
Simultaneous receivers



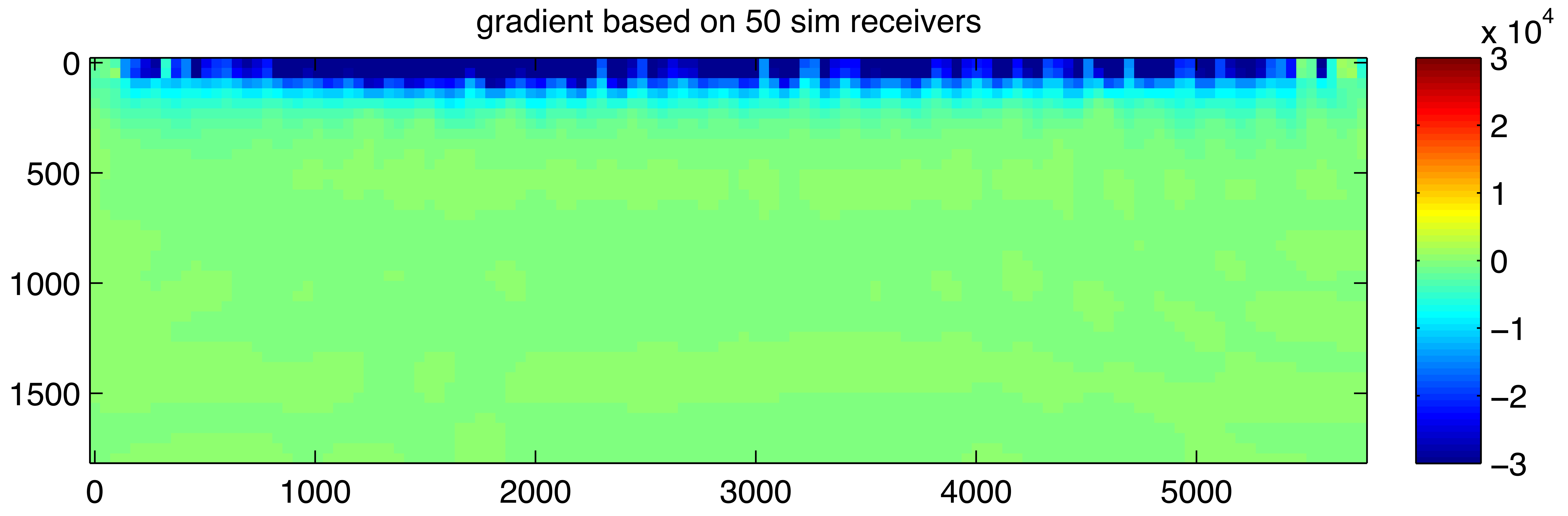


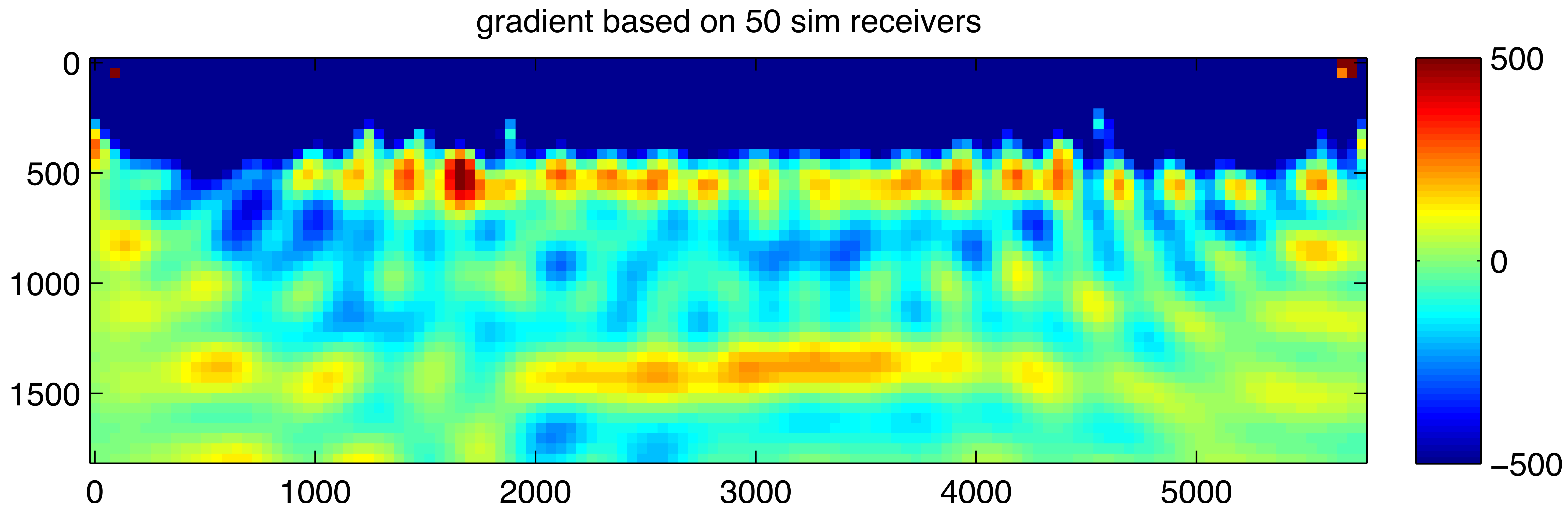
Simultaneous receivers



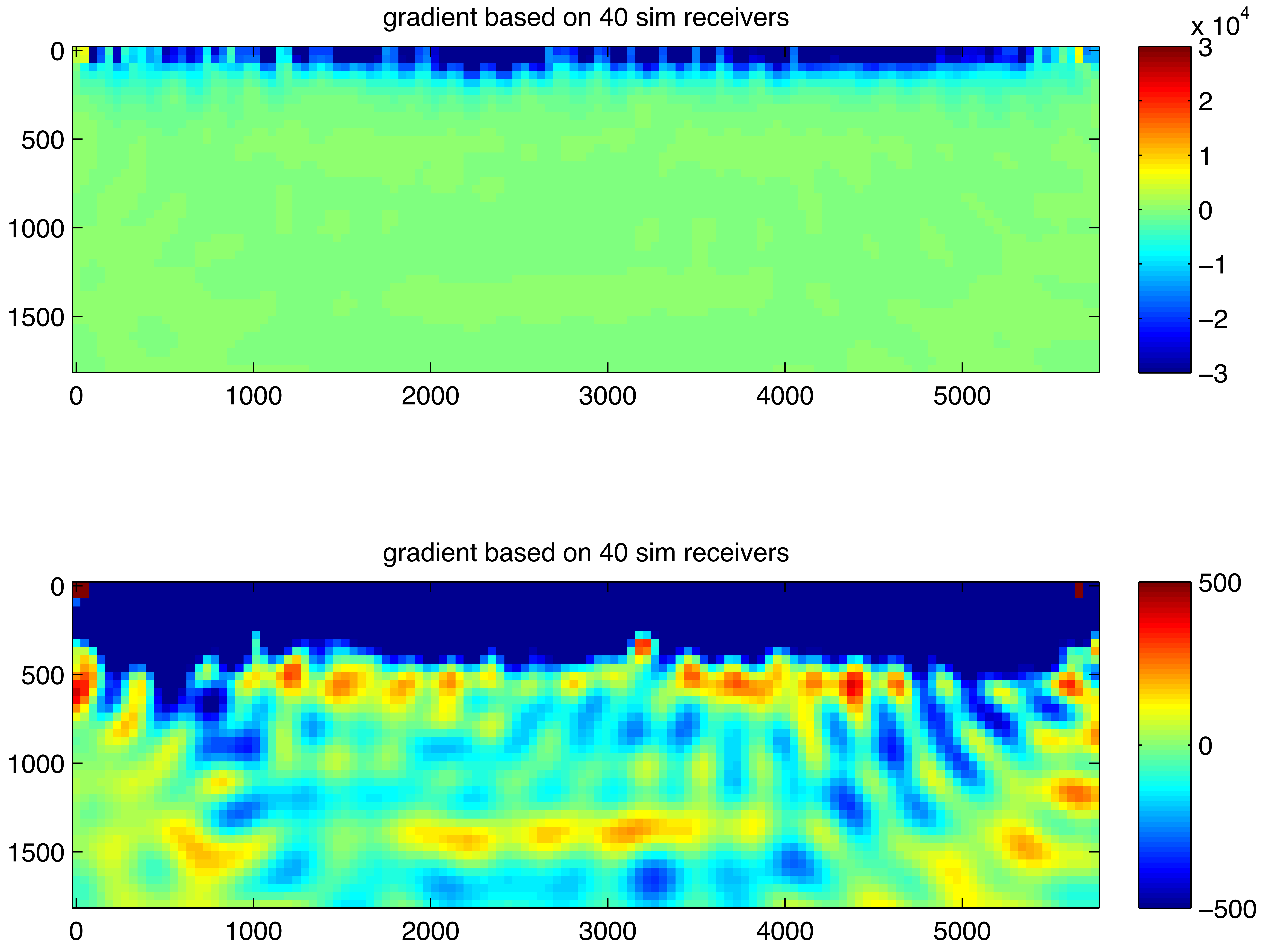


Simultaneous receivers

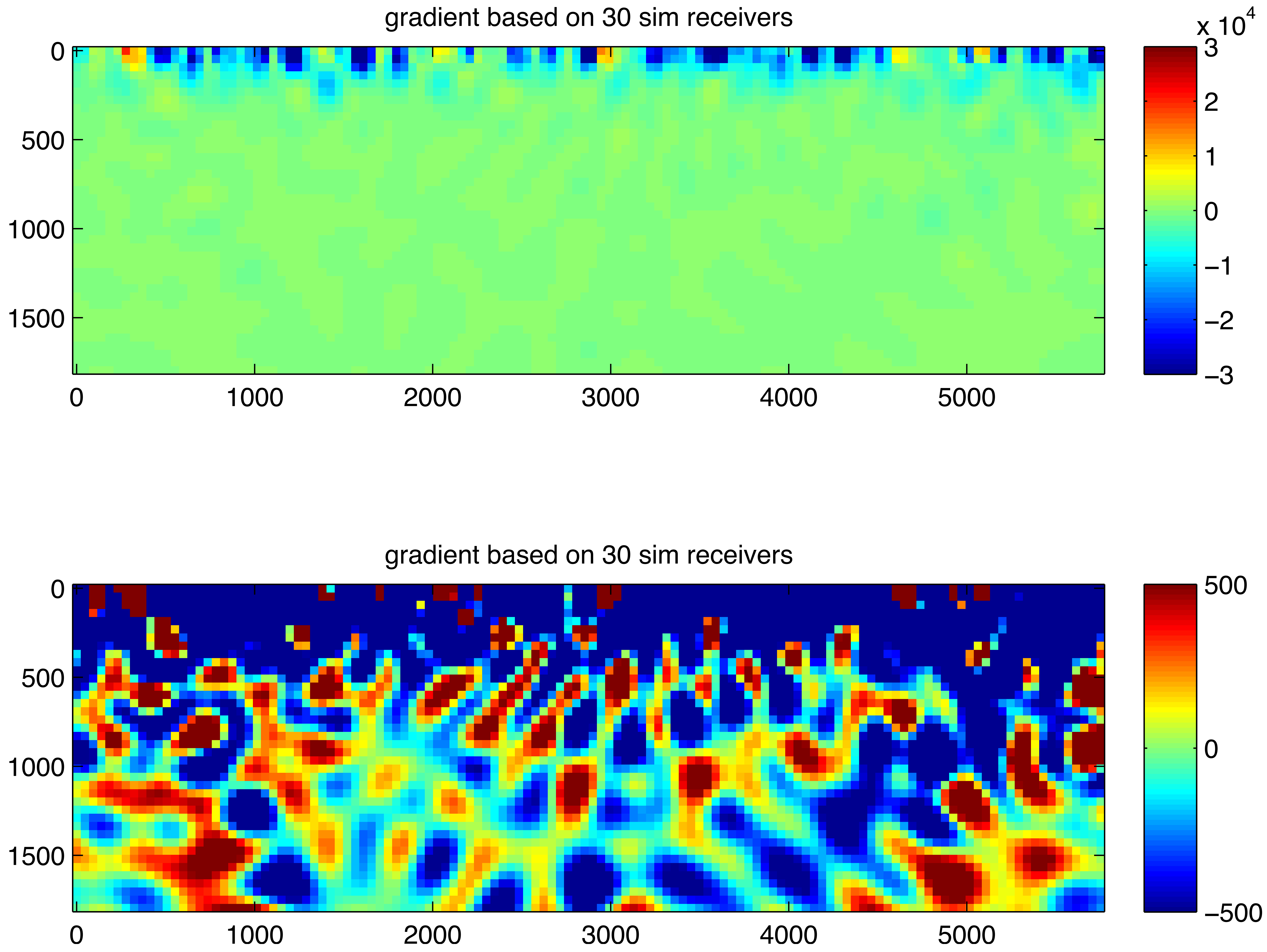




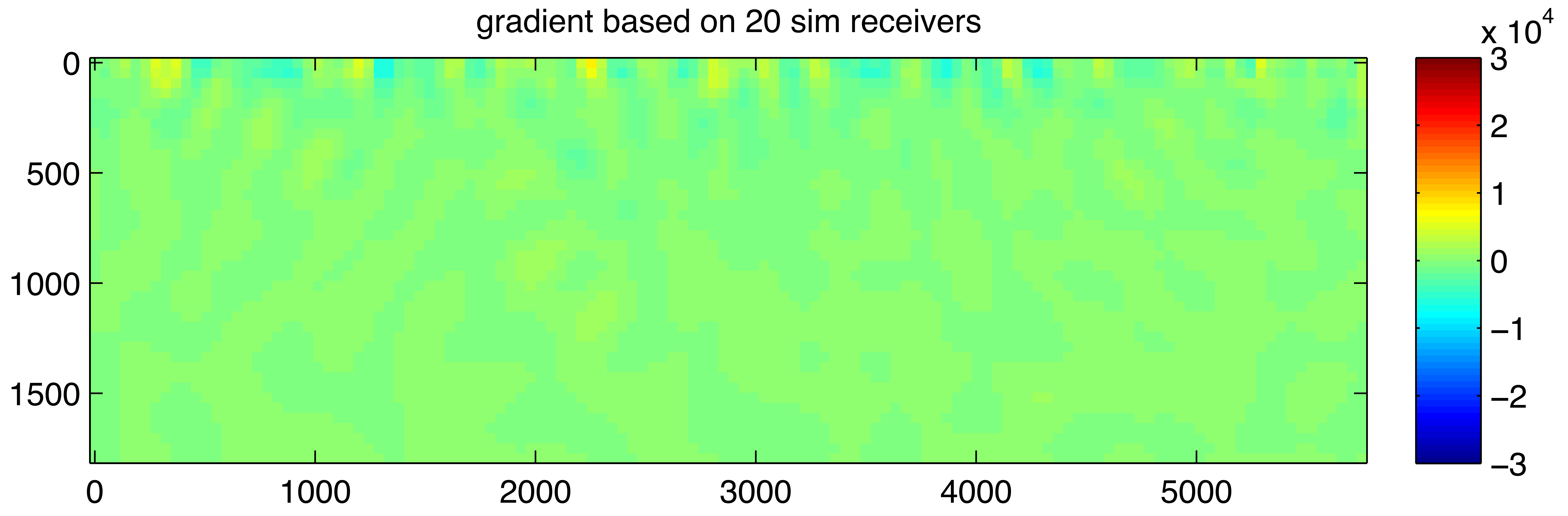
Simultaneous receivers

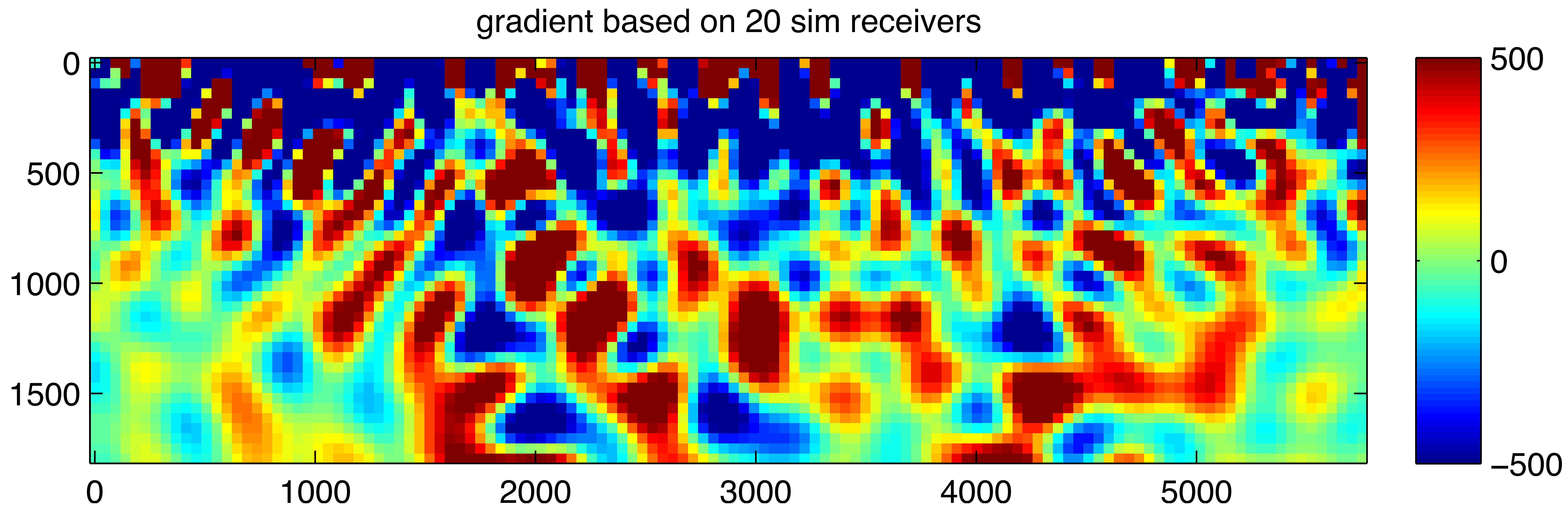


Simultaneous receivers

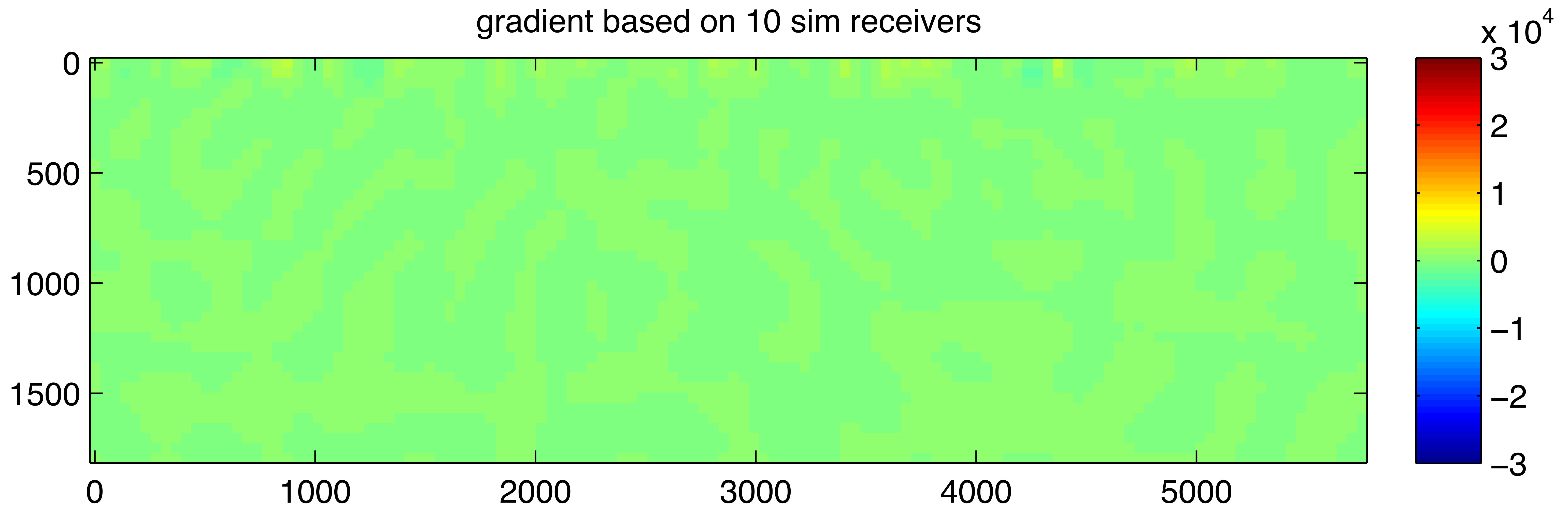


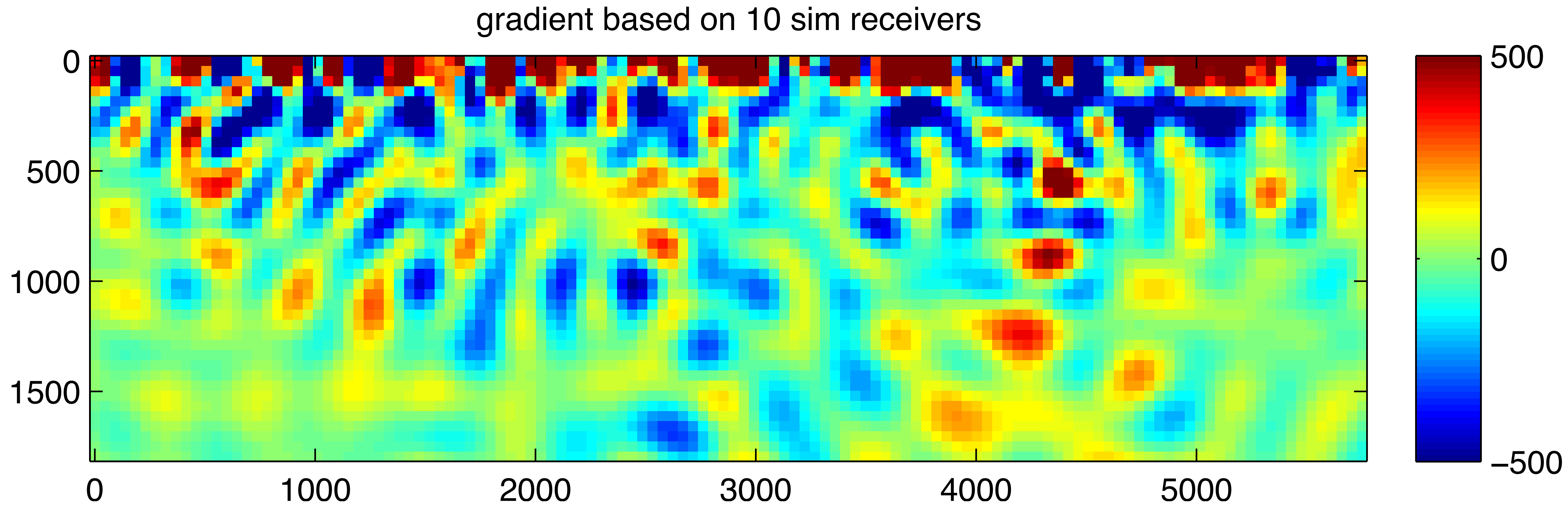
Simultaneous receivers



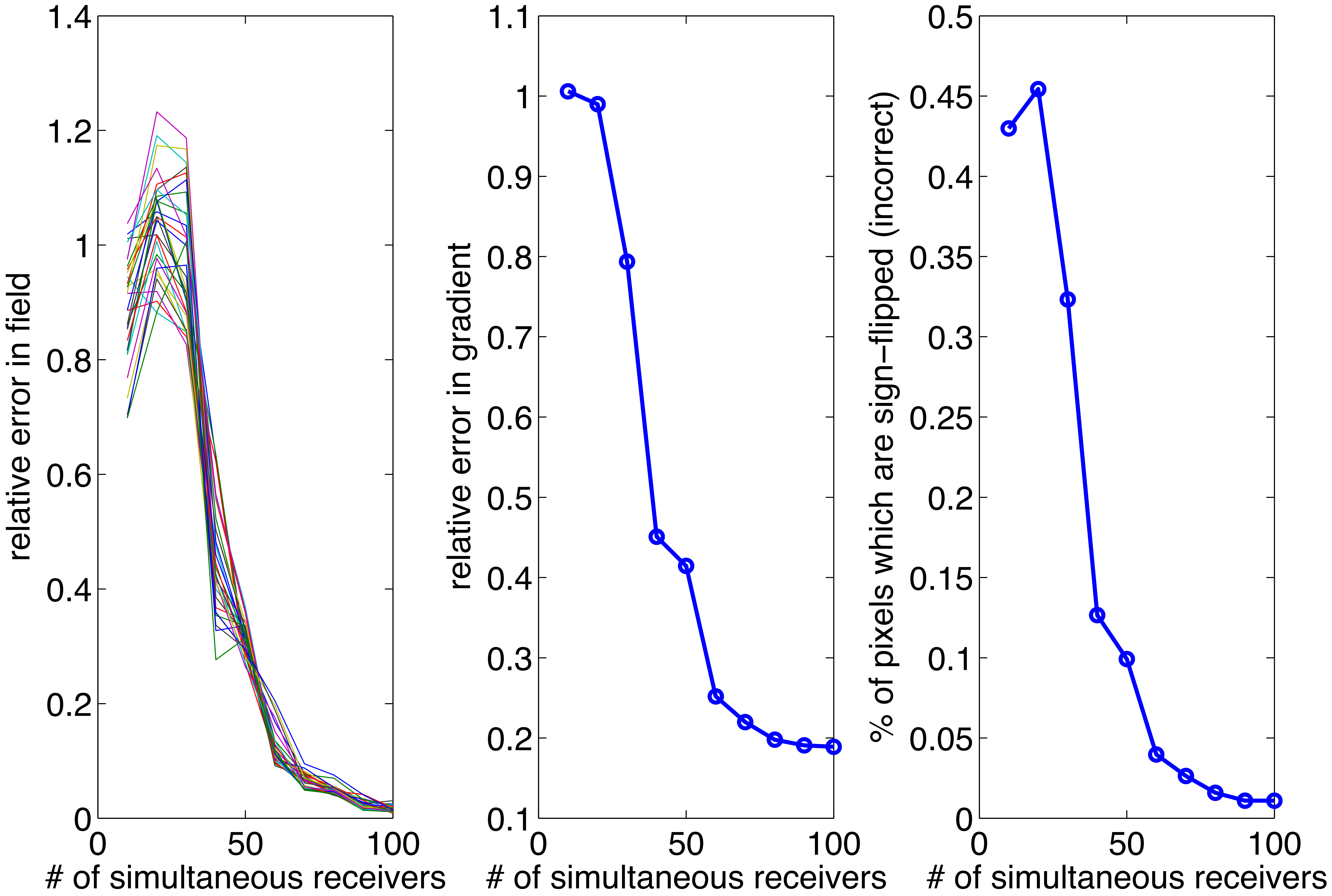


Simultaneous receivers



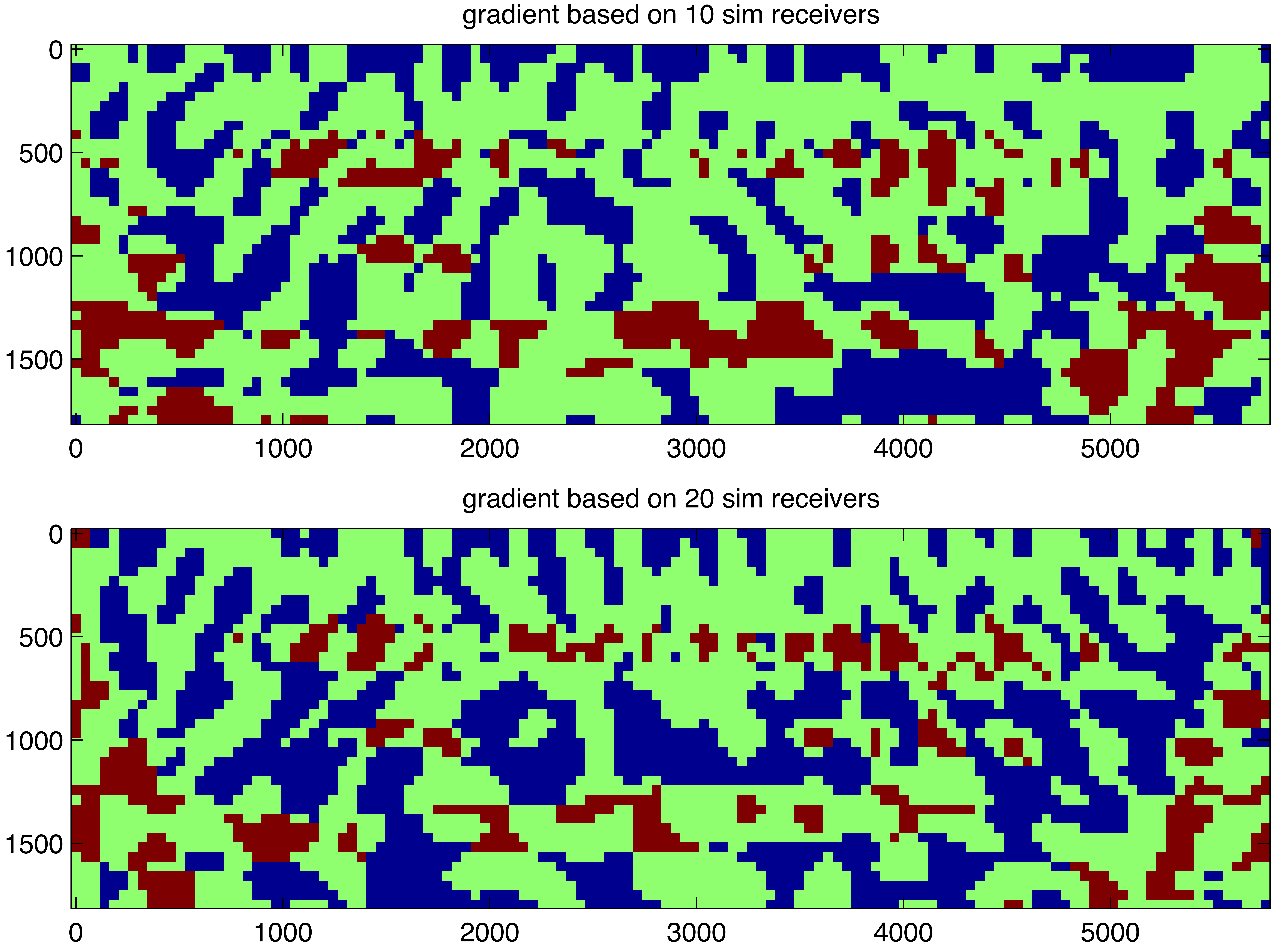


Simultaneous receivers



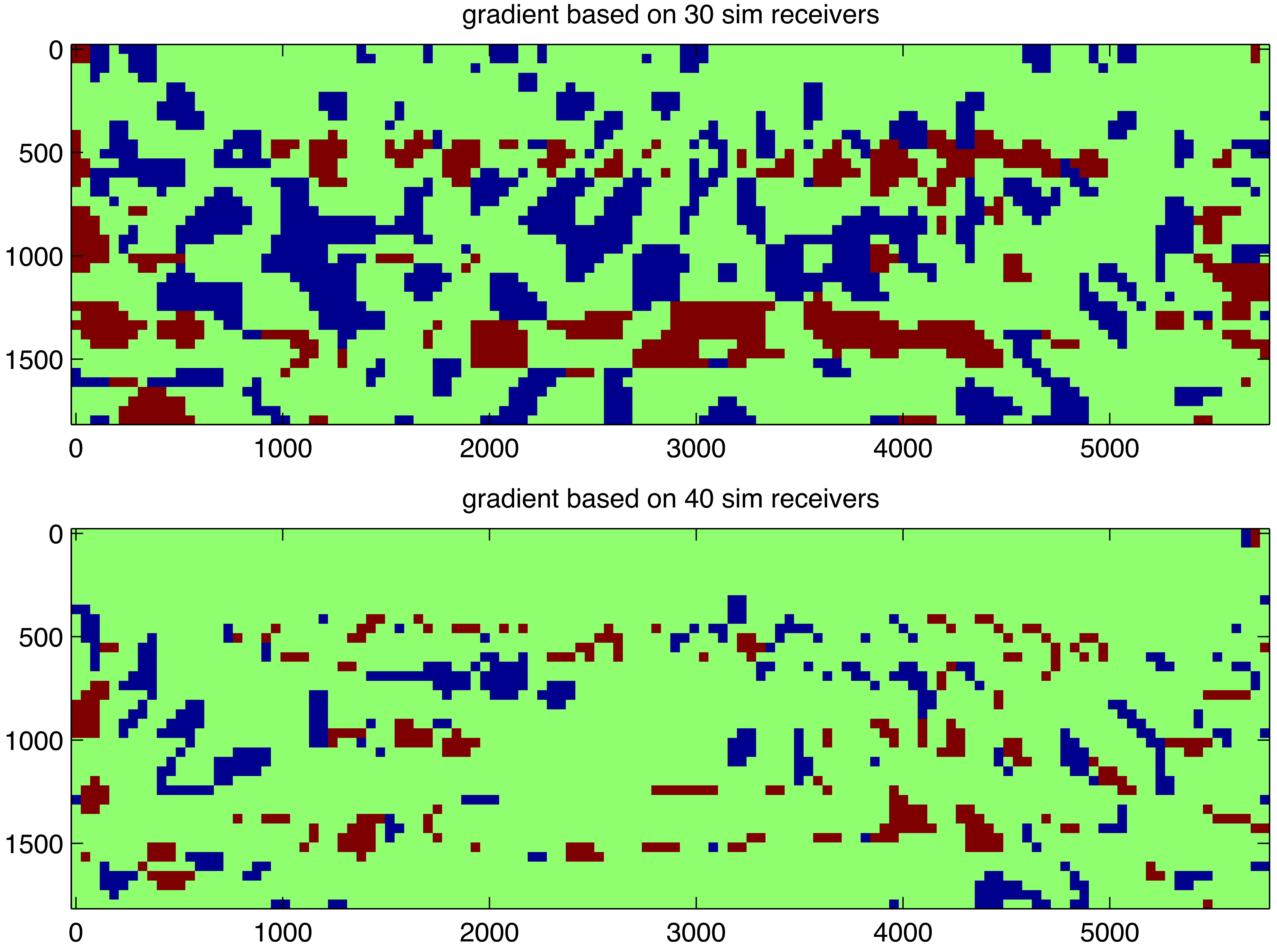
Simultaneous receivers

$$\text{sign}(\mathbf{g}_*) - \text{sign}(\hat{\mathbf{g}})$$



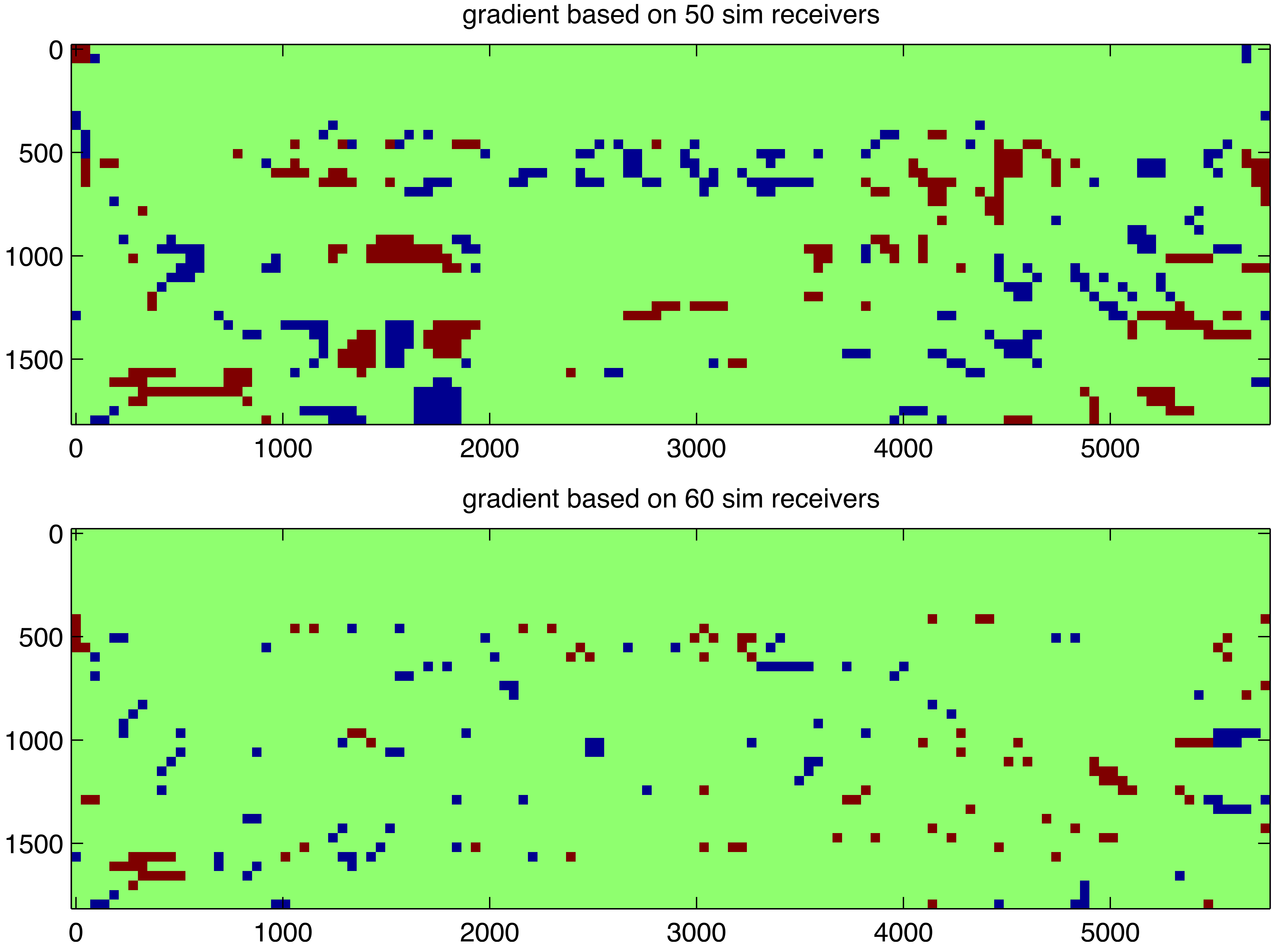
Simultaneous receivers

$\text{sign}(\mathbf{g}_*) - \text{sign}(\hat{\mathbf{g}})$



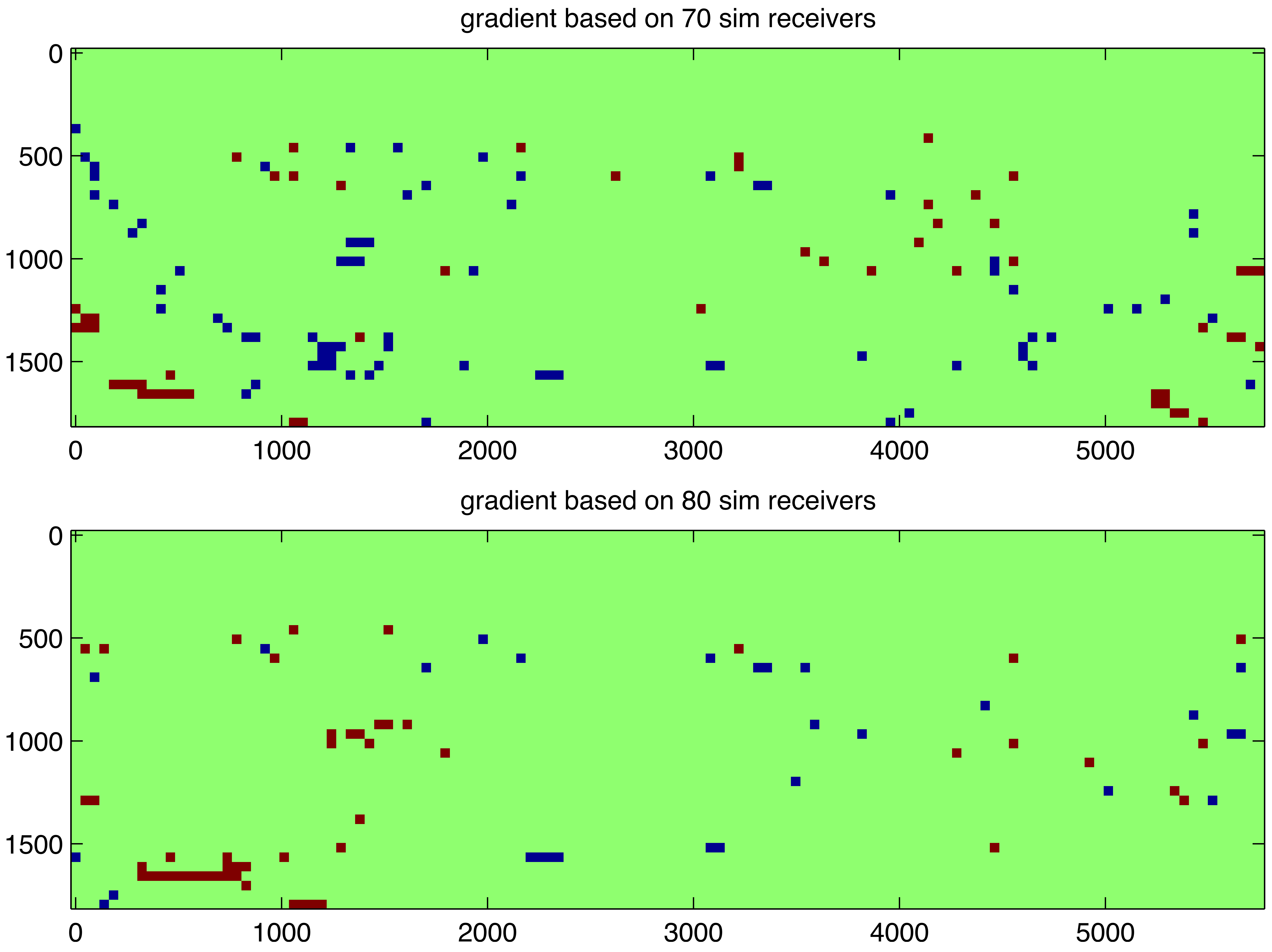
Simultaneous receivers

$$\text{sign}(\mathbf{g}_*) - \text{sign}(\hat{\mathbf{g}})$$



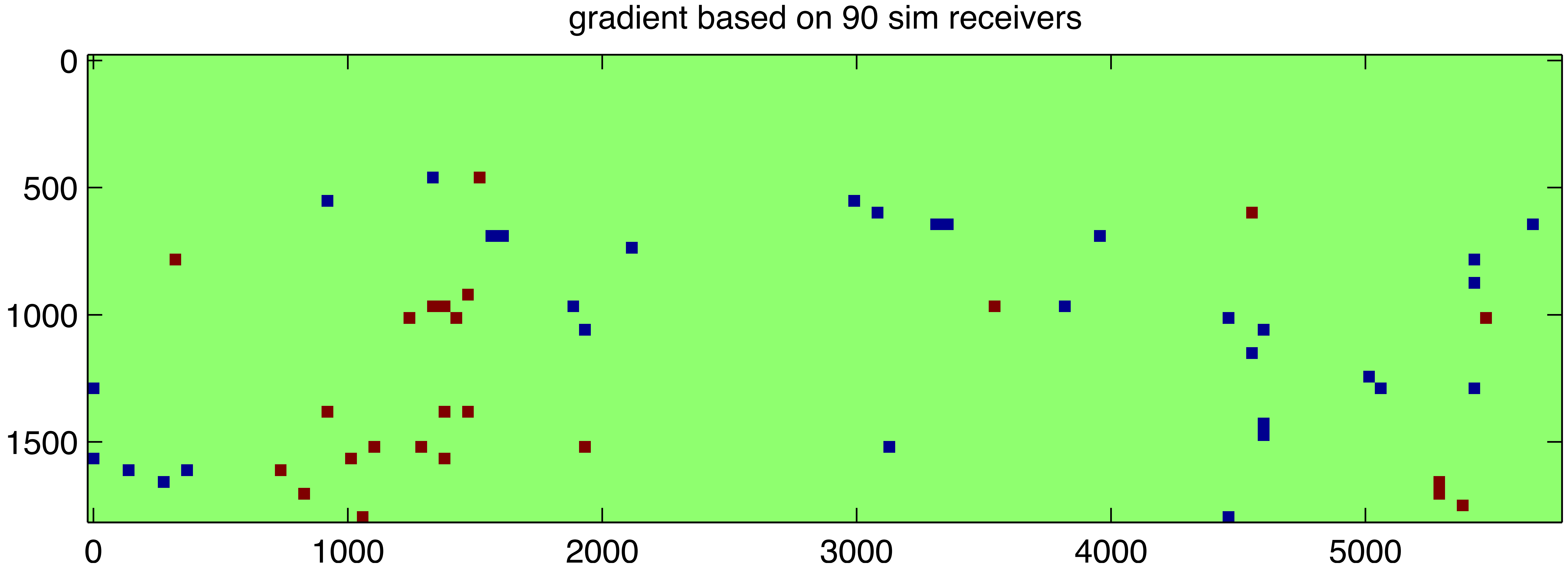
Simultaneous receivers

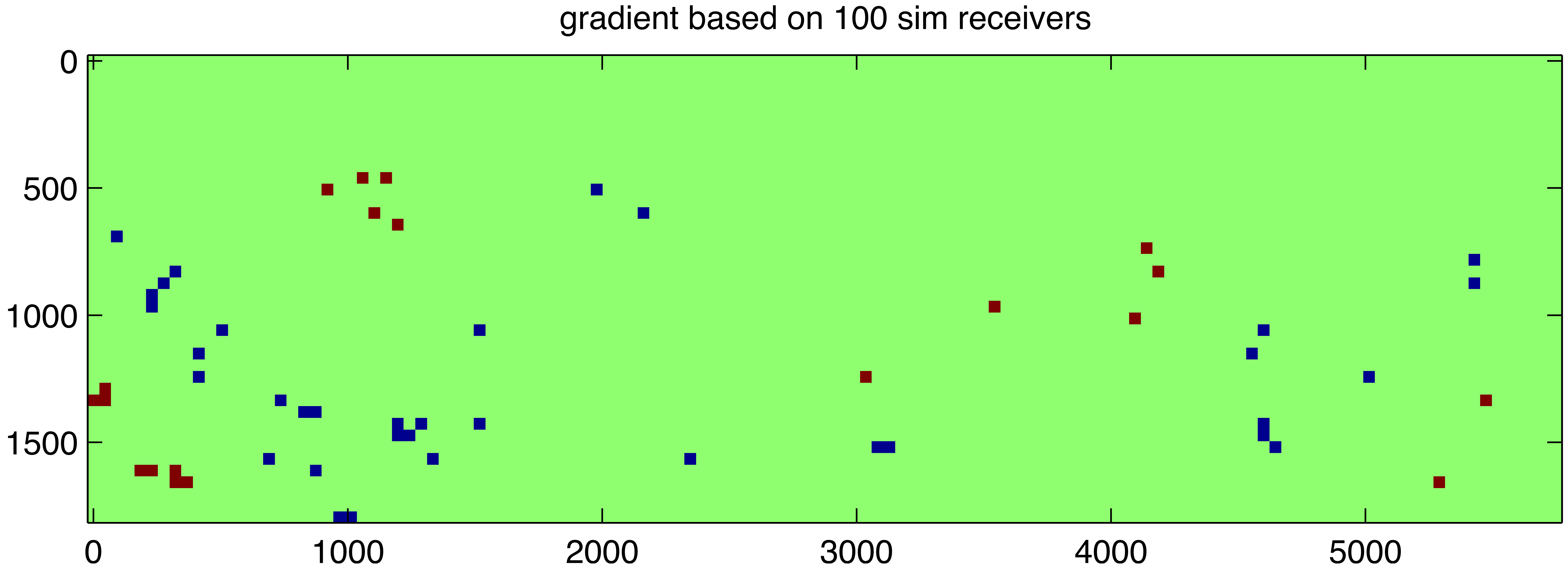
$$\text{sign}(\mathbf{g}_*) - \text{sign}(\hat{\mathbf{g}})$$



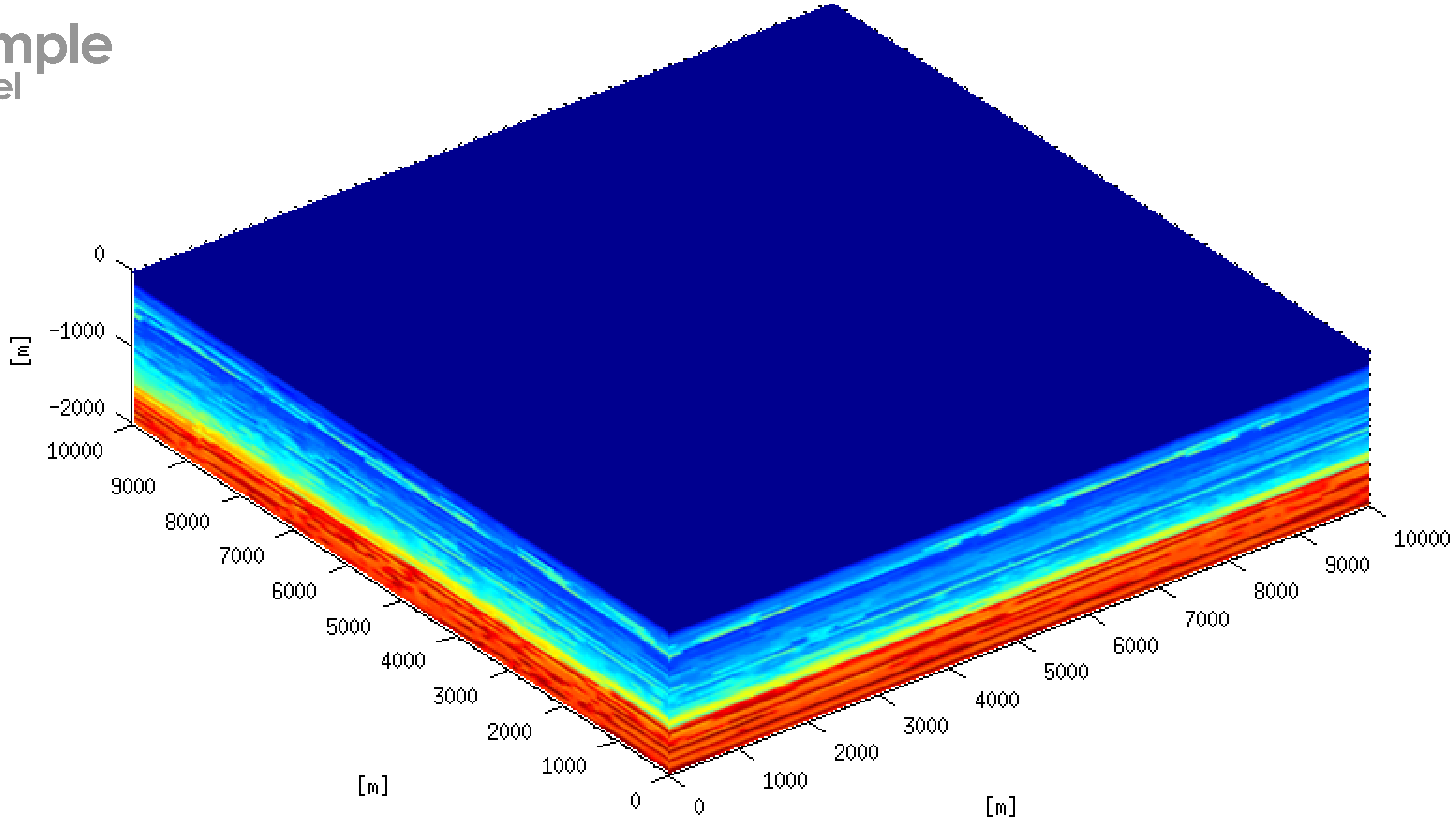
Simultaneous receivers

$$\text{sign}(\mathbf{g}_*) - \text{sign}(\hat{\mathbf{g}})$$



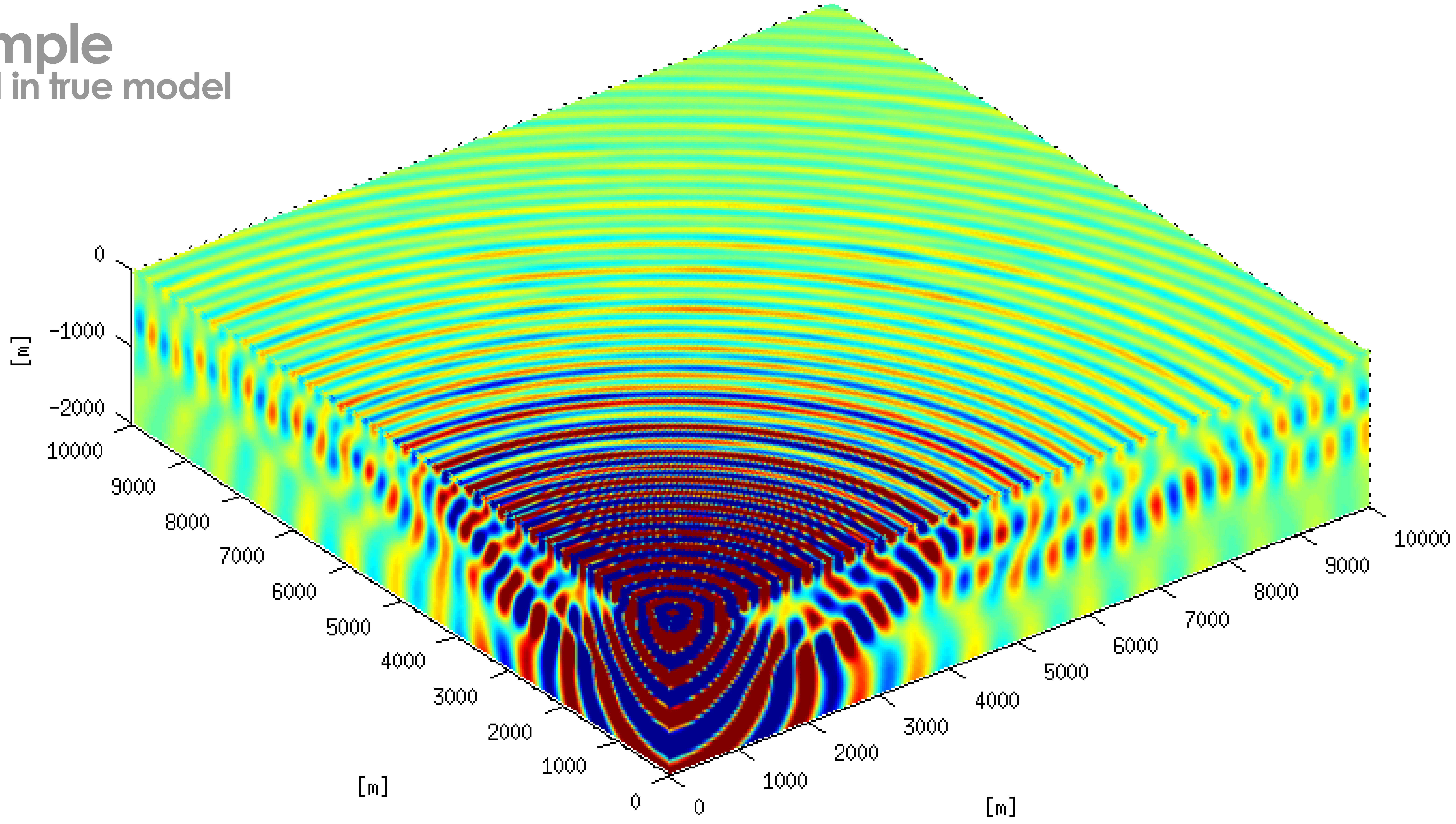


3D Example
– true model

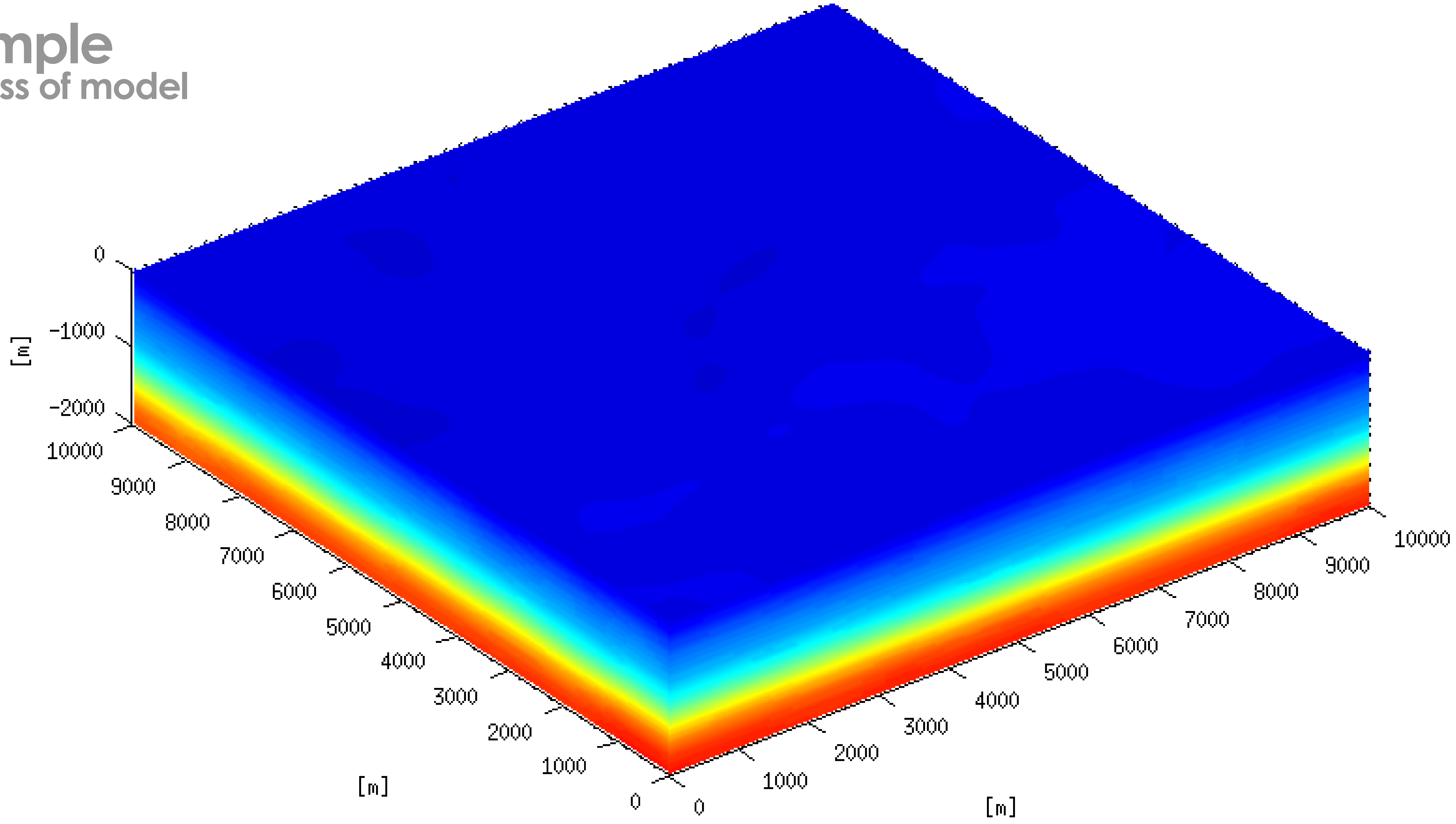


10 x 10 x 2 km, 5 Hz, 27-point discretization, ~1e7 grid points, source at [0,0,0]

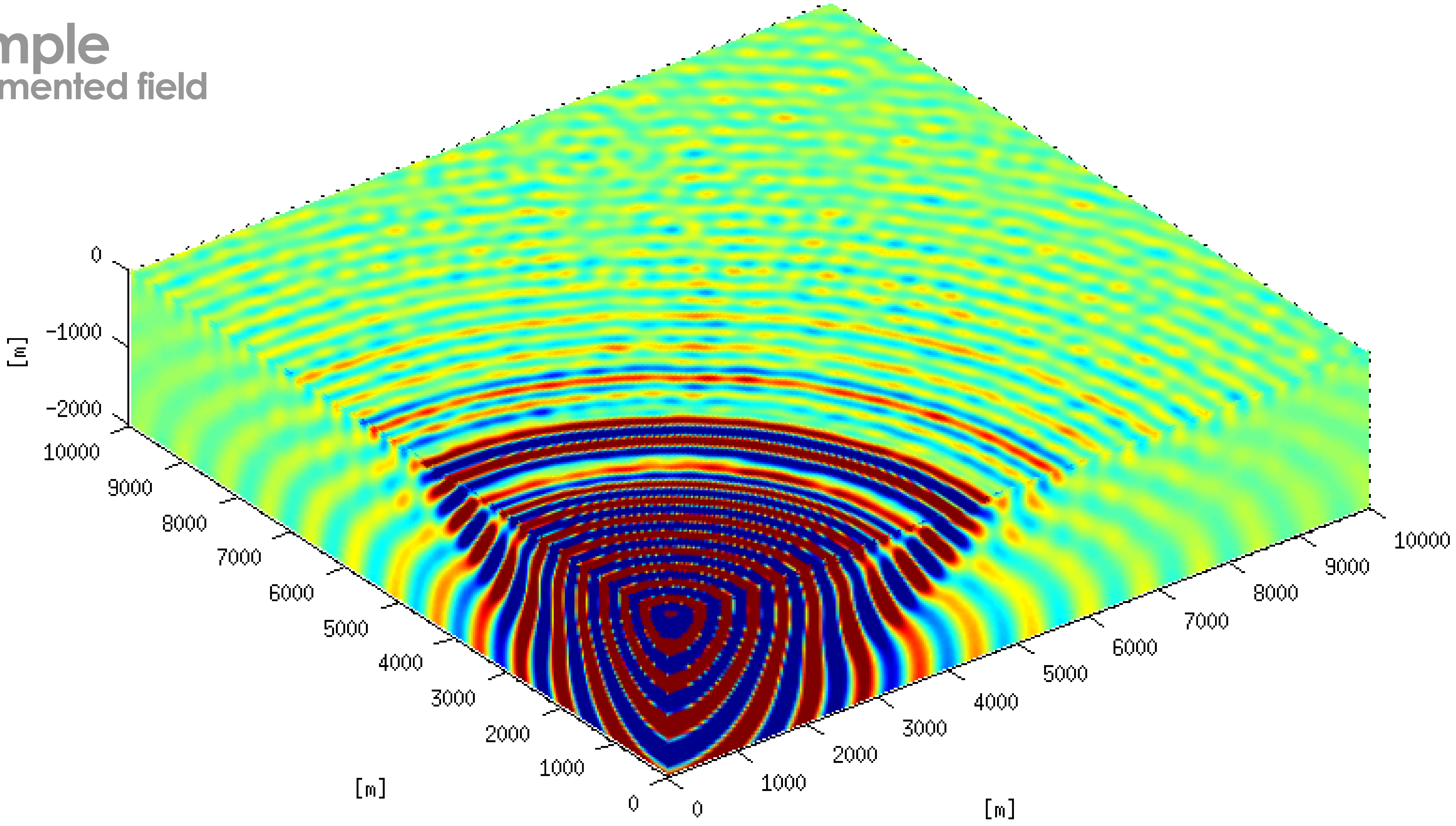
3D Example
– wavefield in true model



3D Example
– initial guess of model



3D Example
– data-augmented field



Conclusions

- Enables WRI in 3D.
- Accepts any Helmholtz solver for the sub-problems.
- Compute 1 Helmholtz problem per source and 1 per receiver.
- Store 1 vector per receiver.
- Can use simultaneous receivers to reduce computational cost and memory use.
- Proposed algorithm might be used for other large-scale mildly overdetermined problems w/ many variables & few constraints.

SLIM

Future work

- Answer some open questions on how to combine simultaneous receiver and simultaneous source principles.
- Adapt optimization algorithms to be able to work with simultaneous sources & receivers. (partially done)
- Error bound derivations are work in progress.

Acknowledgements

Tristan van Leeuwen, Art Petrenko & Rafael Lago for the CGMN & CARP-CG implementation

SINBAD

NSERC CRSNG

This work was in part financially supported by the Natural Sciences and Engineering Research Council of Canada via the Collaborative Research and Development Grant DNOISEII (375142--08). This research was carried out as part of the SINBAD II project which is supported by the following organizations: BG Group, BGP, CGG, Chevron, ConocoPhillips, DownUnder GeoSolutions, Hess, Petrobras, PGS, Schlumberger, Sub Salt Solutions and Woodside

References

1. A. Bjorck and T. Elfving, Accelerated projection methods for computing pseudoinverse solutions of systems of linear equations, BIT, 19 (1979), pp. 145–163.
2. D. Gordon and R. Gordon, CARP-CG: A robust and efficient parallel solver for linear systems, applied to strongly convection dominated pdes, Parallel Computing, 36 (2010), pp. 495– 515.
3. Calandra, H., Gratton, S., Pinel, X. and Vasseur, X. [2013] An improved two-grid preconditioner for the solution of three-dimensional Helmholtz problems in heterogeneous media. Numerical Linear Algebra with Applications.
4. Erlangga, Y.A. [2008] Advances in iterative methods and preconditioners for the Helmholtz equation. Archives of Computational Methods in Engineering, 15, 37–66.
5. Delves, L. M., and I. Barrodale. "A fast direct method for the least squares solution of slightly overdetermined sets of linear equations." IMA Journal of Applied Mathematics 24.2 (1979): 149-156.
6. Rafael Lago, Felix J. Herrmann. Towards a robust geometric multigrid scheme for {Helmholtz} equation, Tech Report, UBC, 2015.
7. Avron, Haim, Petar Maymounkov, and Sivan Toledo. "Blendenpik: Supercharging LAPACK's least-squares solver." SIAM Journal on Scientific Computing 32.3 (2010): 1217-1236.
8. Meng, Xiangrui, Michael A. Saunders, and Michael W. Mahoney. "LSRN: A parallel iterative solver for strongly over-or underdetermined systems." SIAM Journal on Scientific Computing 36.2 (2014): C95-C118.