SINBAD Consortium 2012

Case study: Wave-equation based inversion with joint sparsity promotion

Xiang Li, Rajiv Kumar, Lina Miao & Felix J. Herrmann

SLIM
University of British Columbia

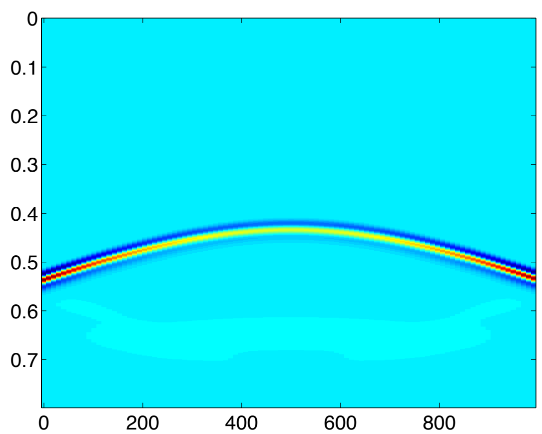
Background

Seismic imaging and inversion assumes

- *linear* relationship between wavefield perturbation and *medium* perturbations (born approximation)
- for *acoustic* wavefields *medium* perturbations are caused by changes in *density* and *velocity*

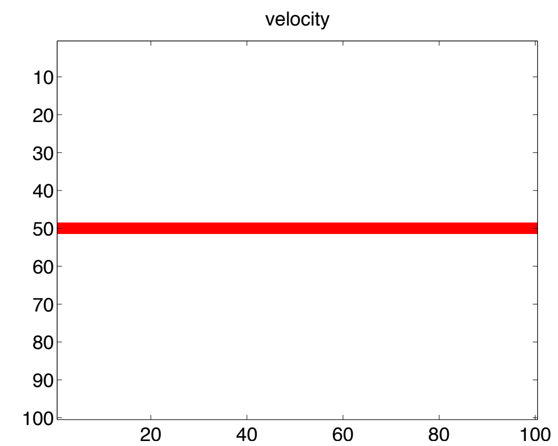
Linearized Born scattering

$$\delta \mathbf{D} = \mathcal{J}_{\mathbf{m}} \delta \mathbf{m}$$



linear wavefield

Jacobian operator (born modeling operator)

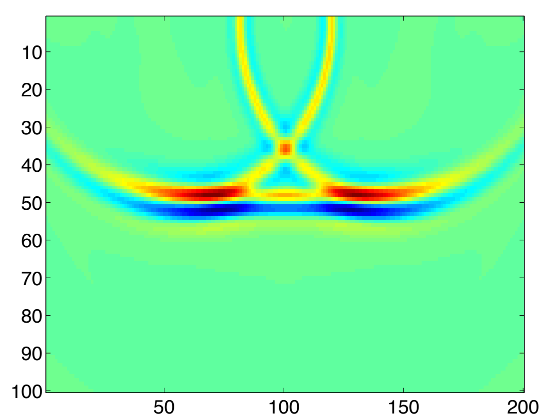


Velocity perturbation

SLIM

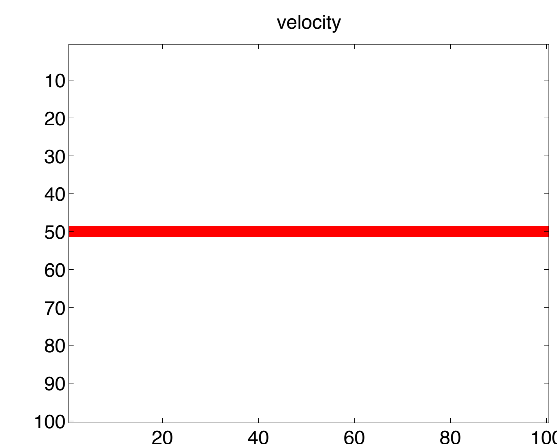
Adjoint imaging (RTM)

$$\color{red}{\mathcal{J}_{\mathbf{m}}^{T}}\,\delta \mathbf{D} = \color{red}{\mathcal{J}_{\mathbf{m}}^{T}}\,\mathcal{J}_{\mathbf{m}}\,\delta \mathbf{m}$$



Jacobian transpose

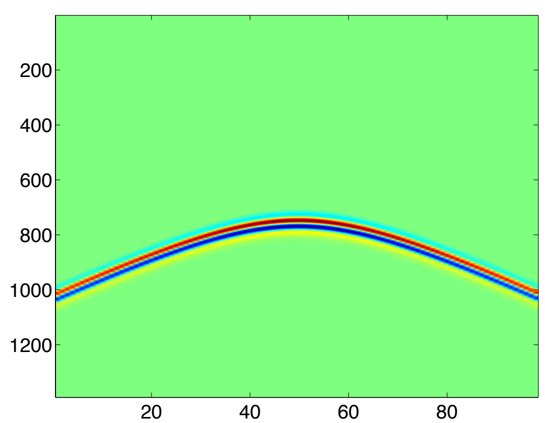
Jacobian operator

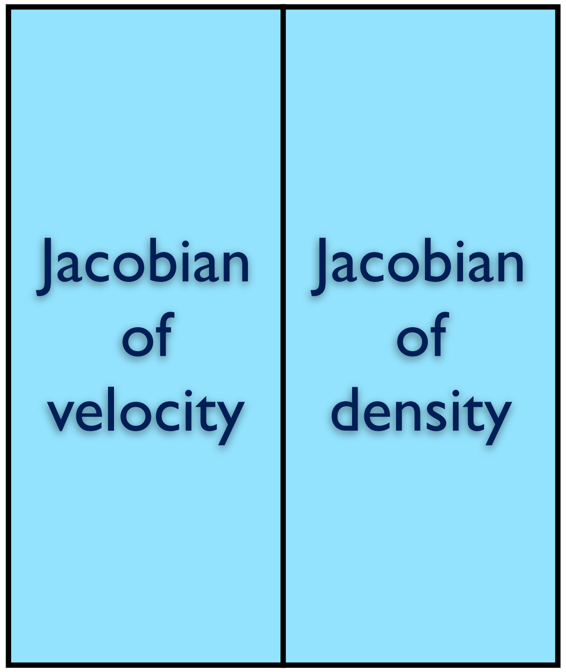


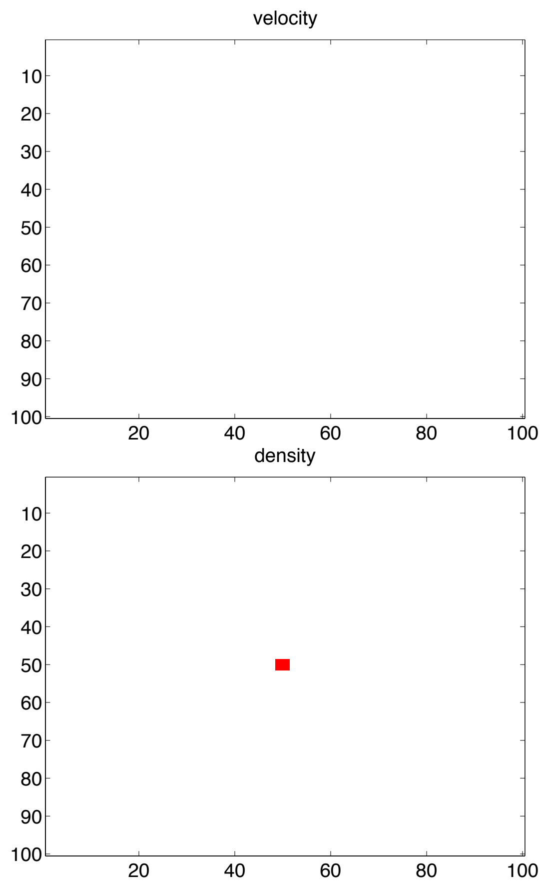
Approximated Hessian: $\mathcal{H} = \mathcal{J}_{\mathbf{m}}^{T}\mathcal{J}_{\mathbf{m}}$

Born modeling
w/ density

$$\delta \mathbf{D} = \begin{bmatrix} \mathcal{J}\,\mathbf{m} & \mathcal{J}\,\boldsymbol{\rho} \end{bmatrix} \begin{bmatrix} \delta \mathbf{m} \\ \delta \boldsymbol{\rho} \end{bmatrix}$$

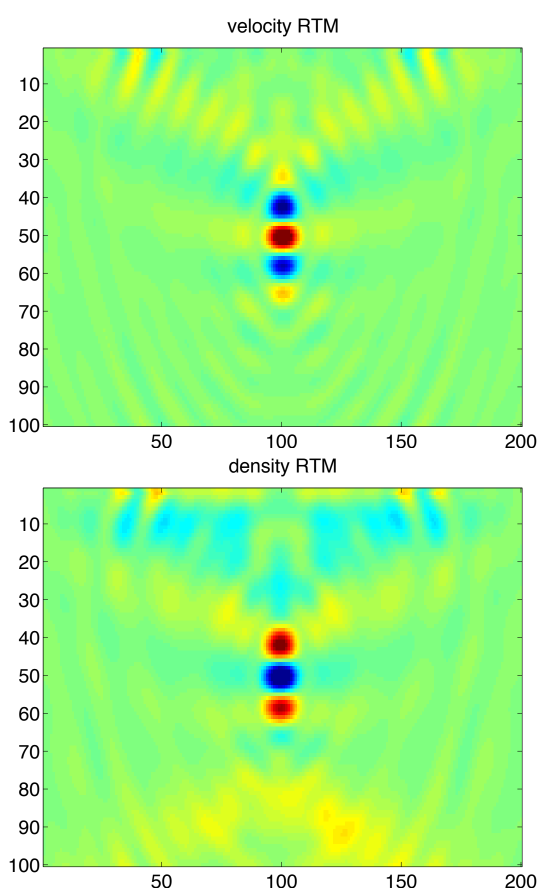




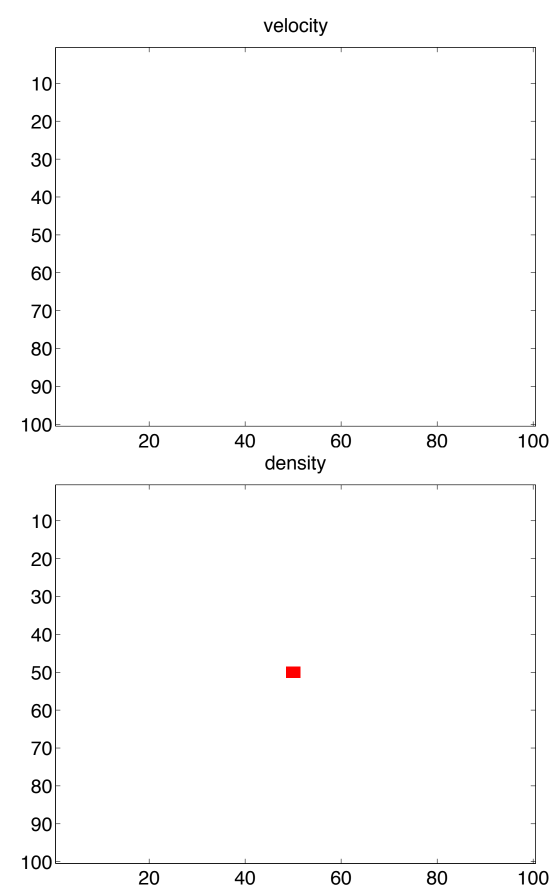


Adjoint imaging
w/ density

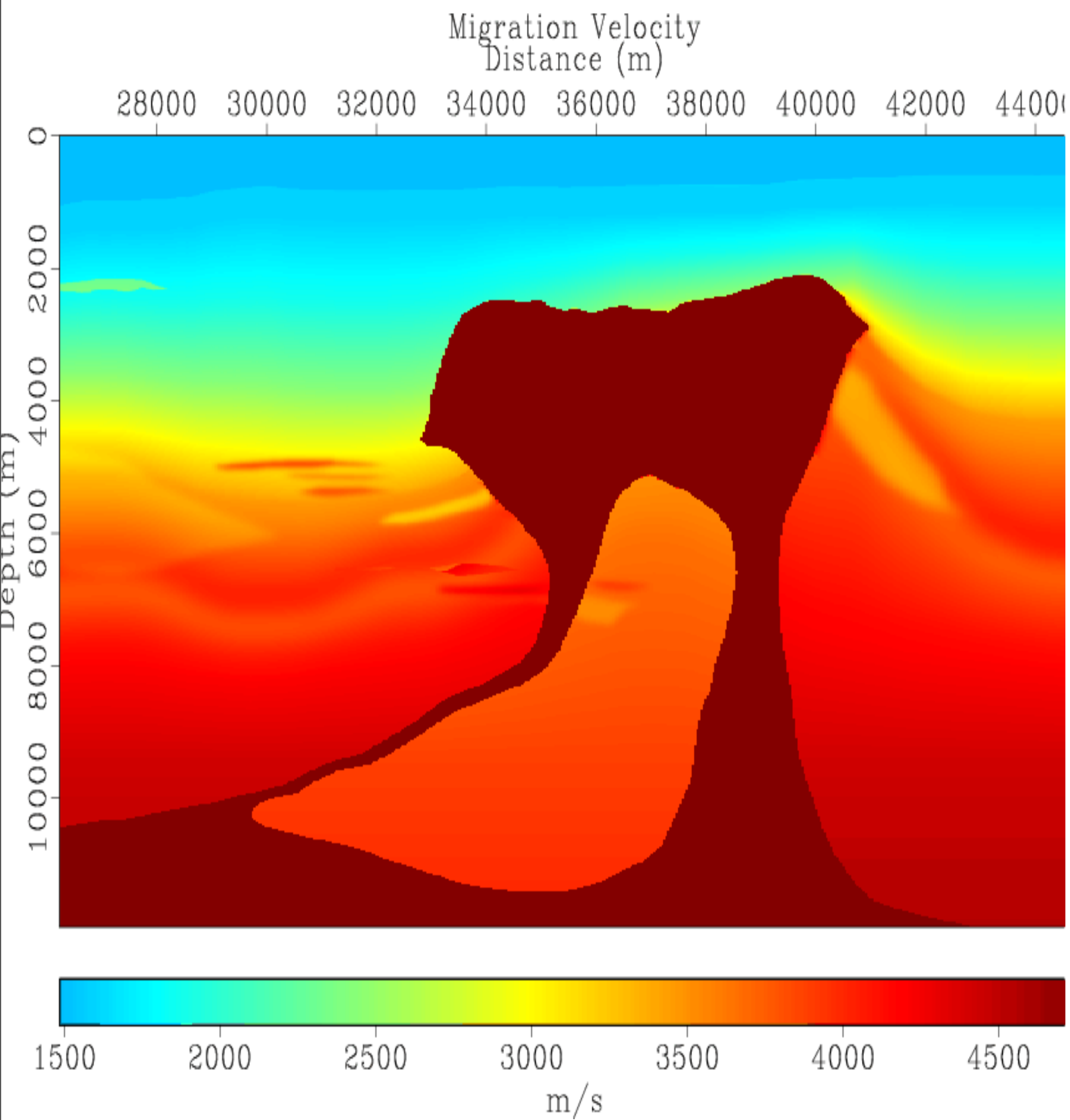
$$\begin{bmatrix}\mathcal{J}_{\mathbf{m}}^T \\ \mathcal{J}_{\boldsymbol{\rho}}^T\end{bmatrix}\delta\mathbf{D} = \mathcal{H}\begin{bmatrix}\delta\mathbf{m} \\ \delta\boldsymbol{\rho}\end{bmatrix}$$

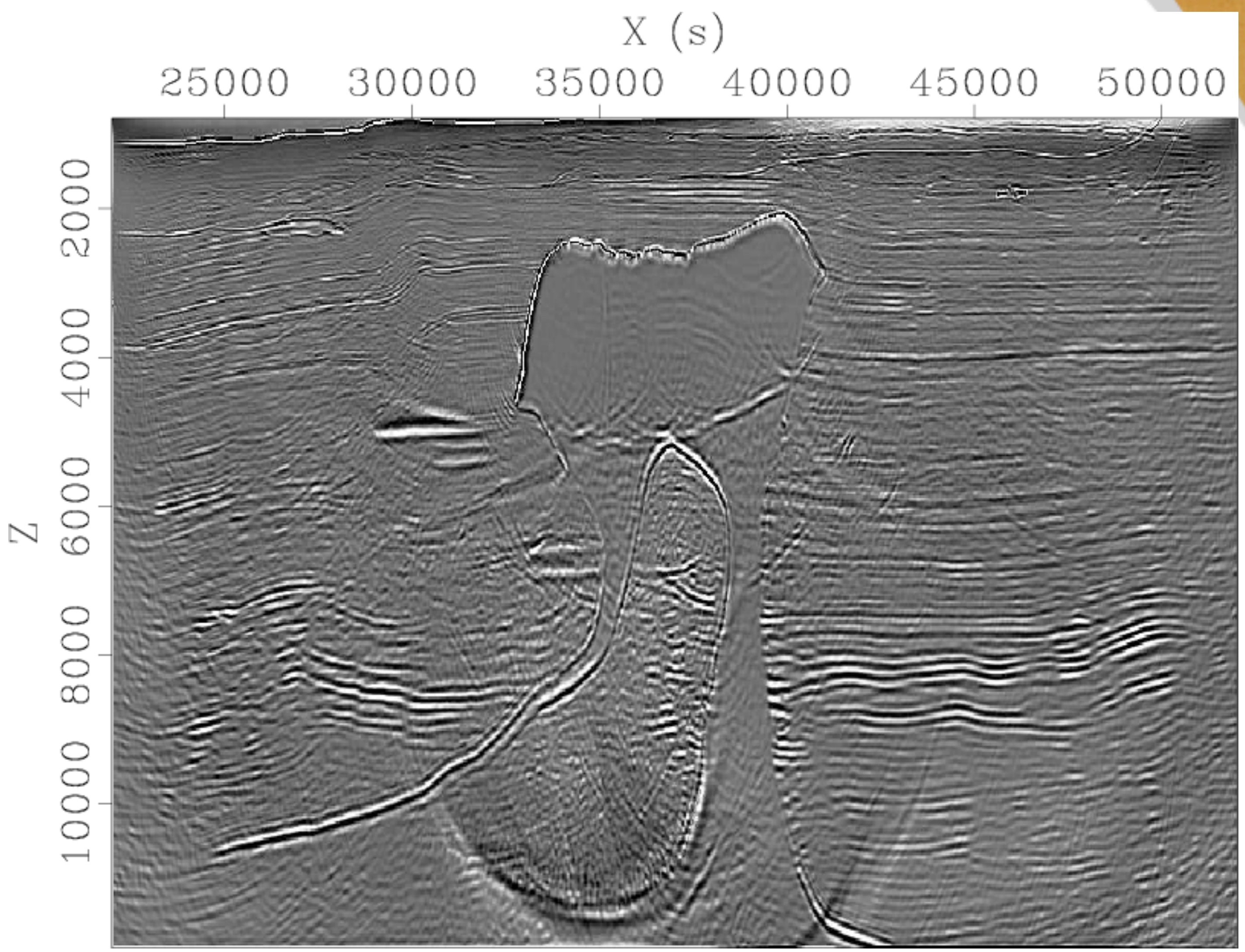


Hessian



Motivation





http://www.reproducibility.org/RSF/book/jsg/ffd/paper_html/paper.html
'Fourier finite-difference wave propagation', Xiaolei Song and Sergey Fomel

Motivation

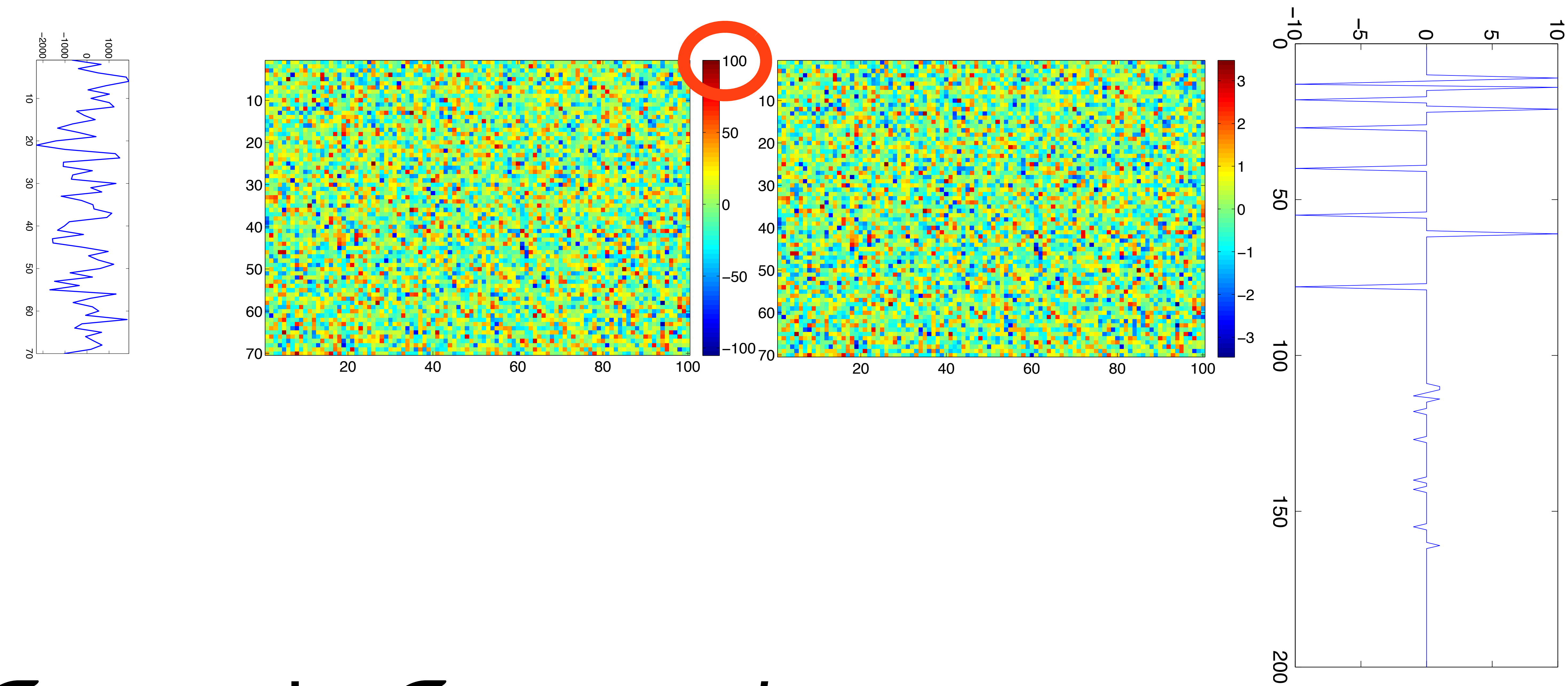
To invert density & velocity *simultaneously*, we should aware that

- *density* also contributes to the *wavefield*
- conventional imaging methods will cause *leakage* between *velocity* and *density*
- In geological settings, *locations* of *unconformities* in *velocity* & *density* often, but not always, overlap
  - exploit this property by *joint* sparsity promotion

SLIM

Stylized numerical example

$$\delta \mathbf{D} = \begin{bmatrix} \mathcal{J}_{\mathbf{m}} & \mathcal{J}_{\boldsymbol{\rho}} \end{bmatrix} \begin{bmatrix} \delta \mathbf{m} \\ \delta \boldsymbol{\rho} \end{bmatrix}$$



$\mathcal{J}_{\mathbf{m}}$ and $\mathcal{J}_{\rho}$ are *coherent*

Conventional method

$$\min \| \begin{bmatrix} \delta \mathbf{m} \\ \delta \boldsymbol{\rho} \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta \mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta \mathbf{m} \\ \delta \boldsymbol{\rho} \end{bmatrix} \|_2 \leq \sigma$$

Conventional method

$$\min \| \begin{bmatrix} \delta \mathbf{m} \\ \delta \boldsymbol{\rho} \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta \mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta \mathbf{m} \\ \delta \boldsymbol{\rho} \end{bmatrix} \|_2 \leq \sigma$$

Hessian:

Conventional method

$$\min \| \begin{bmatrix} \delta\mathbf{m} \\ \delta\boldsymbol{\rho} \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta\mathbf{D} - [\boldsymbol{\mathcal{J}}\mathbf{m} \quad \boldsymbol{\mathcal{J}}\boldsymbol{\rho}] \begin{bmatrix} \delta\mathbf{m} \\ \delta\boldsymbol{\rho} \end{bmatrix} \|_2 \leq \sigma$$

Hessian:

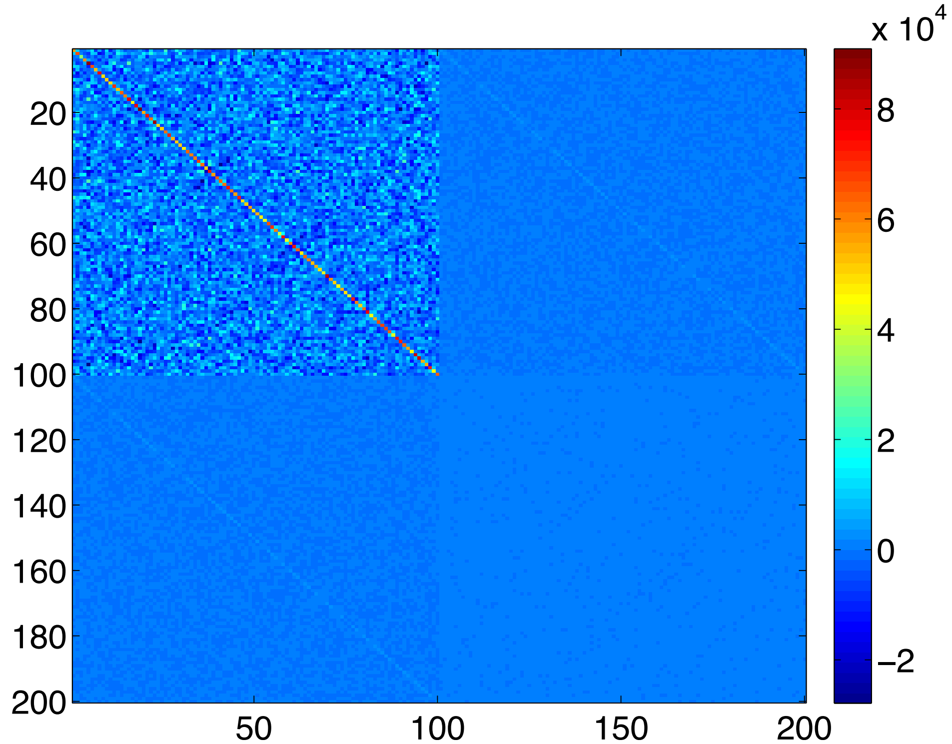
$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \\ \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \end{bmatrix} [\boldsymbol{\mathcal{J}}\mathbf{m} \quad \boldsymbol{\mathcal{J}}\boldsymbol{\rho}] = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T\boldsymbol{\mathcal{J}}\mathbf{m} & \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T\boldsymbol{\mathcal{J}}\boldsymbol{\rho} \\ \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T\boldsymbol{\mathcal{J}}\mathbf{m} & \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T\boldsymbol{\mathcal{J}}\boldsymbol{\rho} \end{bmatrix}$$

Conventional method

$$\min \| \begin{bmatrix} \delta \mathbf{m} \\ \delta \boldsymbol{\rho} \end{bmatrix} \|_p \quad \text{s.t.} \quad \| \delta \mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta \mathbf{m} \\ \delta \boldsymbol{\rho} \end{bmatrix} \|_2 \leq \sigma$$

Hessian:

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \\ \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \end{bmatrix} [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \boldsymbol{\mathcal{J}}_{\mathbf{m}} & \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \\ \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \boldsymbol{\mathcal{J}}_{\mathbf{m}} & \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \end{bmatrix}$$

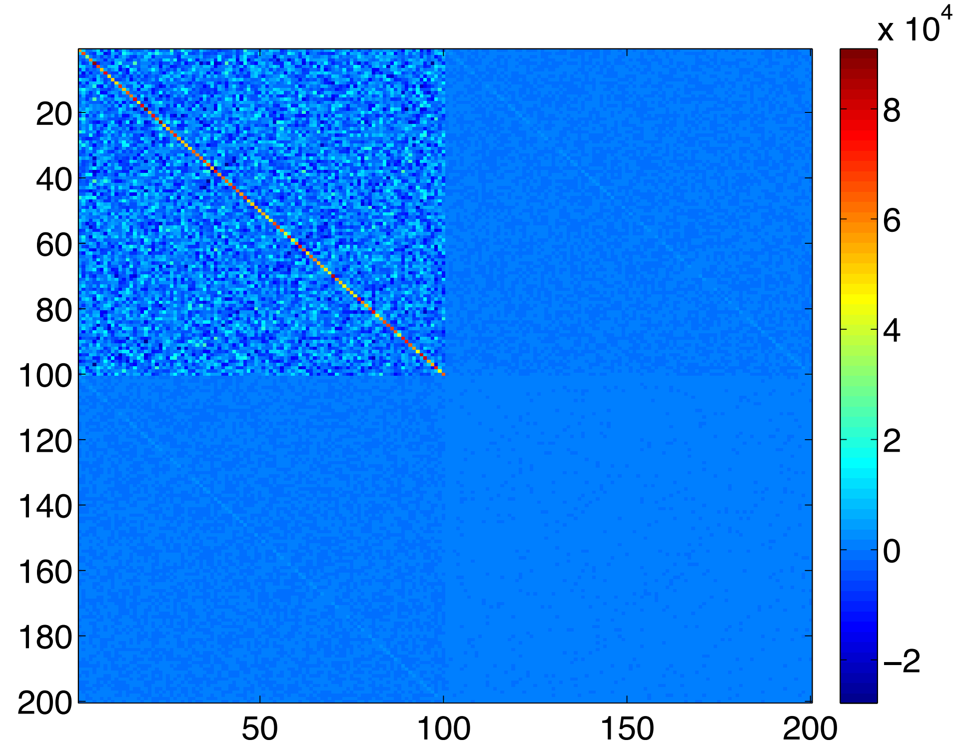


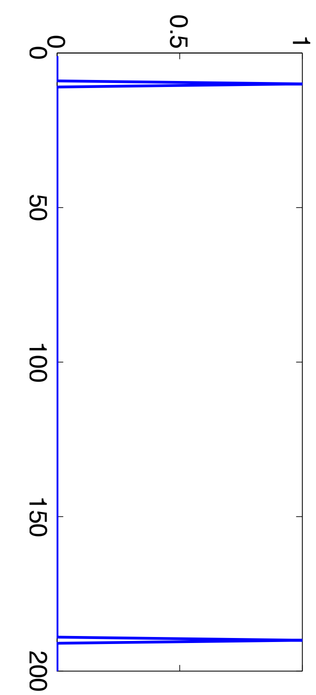
Conventional method

$$\min \left\| \begin{bmatrix} \delta \mathbf{m} \\ \delta \boldsymbol{\rho} \end{bmatrix} \right\|_p \quad \text{s.t.} \quad \left\| \delta \mathbf{D} - \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}} & \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \end{bmatrix} \begin{bmatrix} \delta \mathbf{m} \\ \delta \boldsymbol{\rho} \end{bmatrix} \right\|_2 \leq \sigma$$

Hessian:

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \\ \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}} & \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \boldsymbol{\mathcal{J}}_{\mathbf{m}} & \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \\ \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \boldsymbol{\mathcal{J}}_{\mathbf{m}} & \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \end{bmatrix}$$



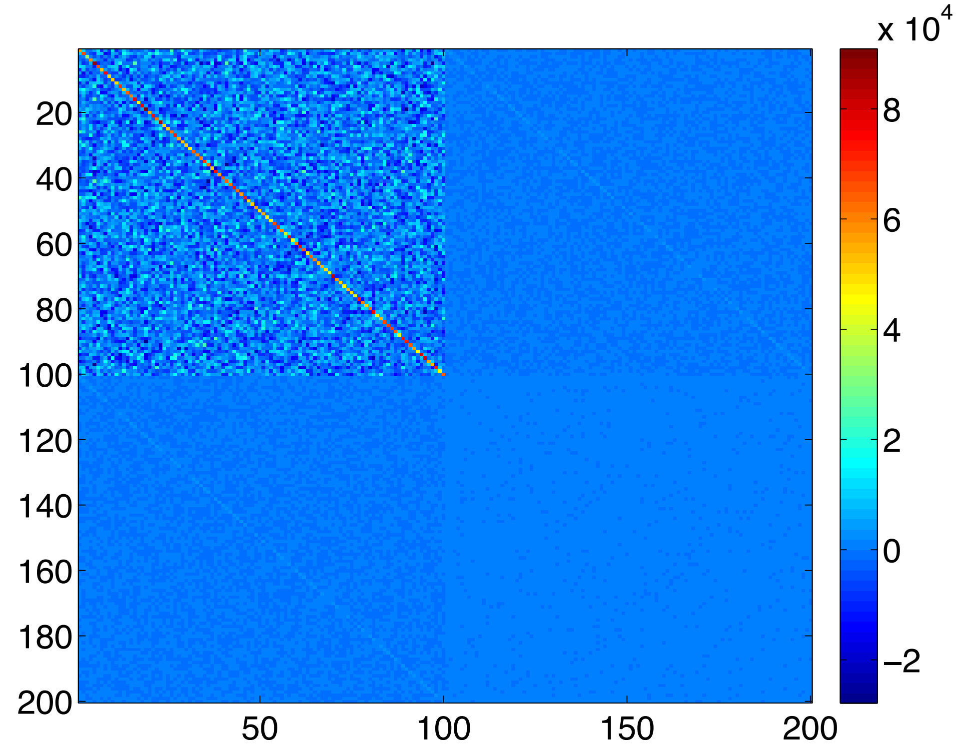


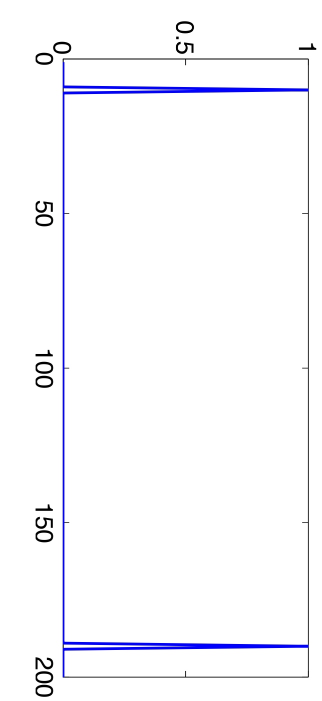
Conventional method

$$\min \left\| \begin{bmatrix} \delta \mathbf{m} \\ \delta \boldsymbol{\rho} \end{bmatrix} \right\|_p \quad \text{s.t.} \quad \left\| \delta \mathbf{D} - \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}} & \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \end{bmatrix} \begin{bmatrix} \delta \mathbf{m} \\ \delta \boldsymbol{\rho} \end{bmatrix} \right\|_2 \leq \sigma$$

Hessian:

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \\ \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}} & \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \boldsymbol{\mathcal{J}}_{\mathbf{m}} & \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \\ \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \boldsymbol{\mathcal{J}}_{\mathbf{m}} & \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \end{bmatrix}$$



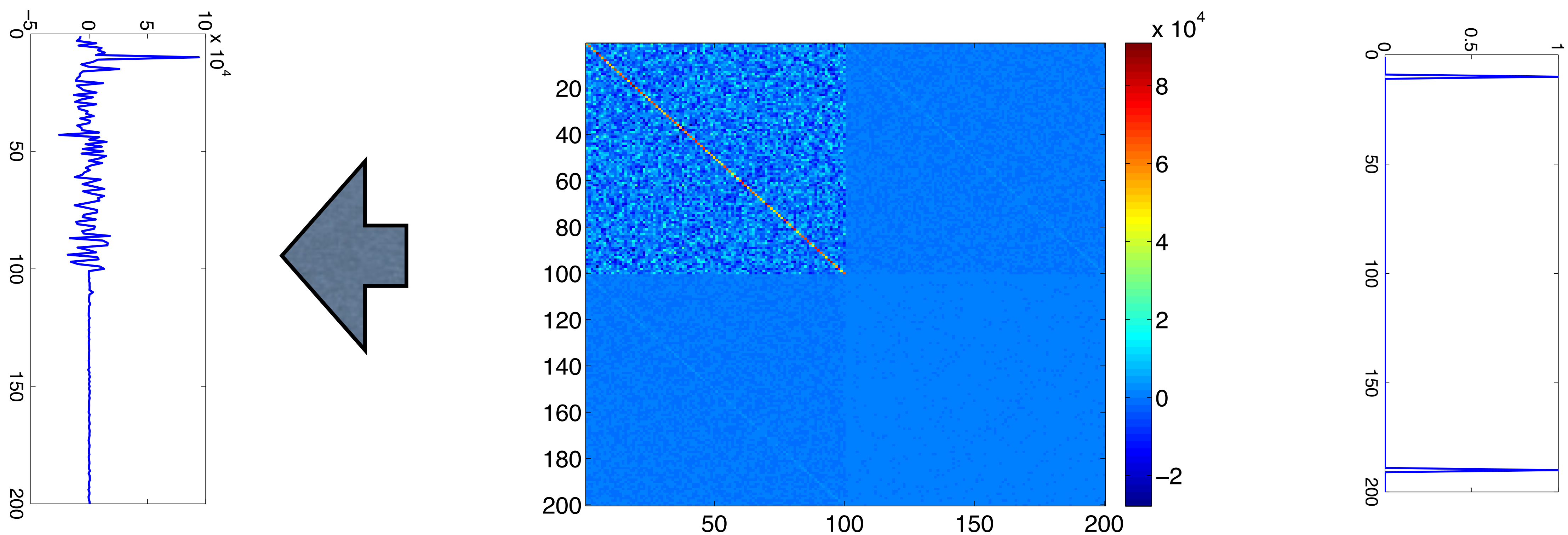


Conventional method

$$\min \left\| \begin{bmatrix}\delta\mathbf{m}\\ \delta\boldsymbol{\rho}\end{bmatrix} \right\|_p \quad \text{s.t.} \quad \left\| \delta\mathbf{D} - \begin{bmatrix}\boldsymbol{\mathcal{J}}_{\mathbf{m}} & \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}\end{bmatrix} \begin{bmatrix}\delta\mathbf{m}\\ \delta\boldsymbol{\rho}\end{bmatrix} \right\|_2 \leq \sigma$$

Hessian:

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix}\boldsymbol{\mathcal{J}}_{\mathbf{m}}^T\\ \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T\end{bmatrix} \begin{bmatrix}\boldsymbol{\mathcal{J}}_{\mathbf{m}} & \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}\end{bmatrix} = \begin{bmatrix}\boldsymbol{\mathcal{J}}_{\mathbf{m}}^T\boldsymbol{\mathcal{J}}_{\mathbf{m}} & \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}\\ \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T\boldsymbol{\mathcal{J}}_{\mathbf{m}} & \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}\end{bmatrix}$$

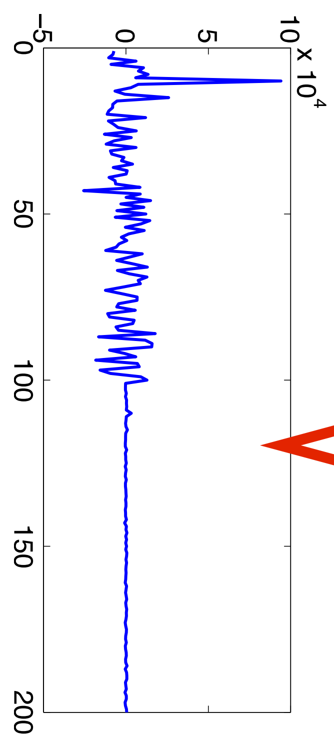


Conventional method

$$\min \| \begin{bmatrix} \delta \mathbf{r} \\ \delta \end{bmatrix}$$

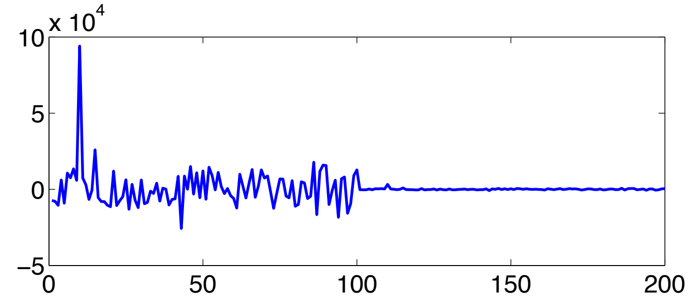
Hessian:

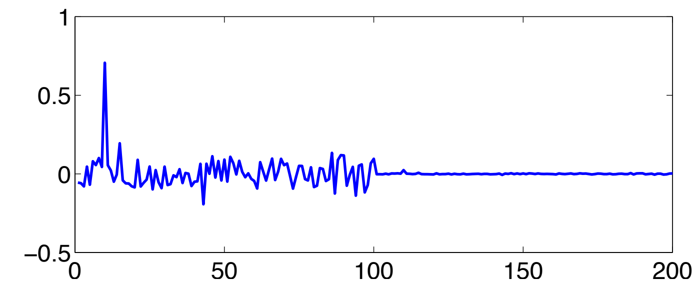
$\mathcal{H}$

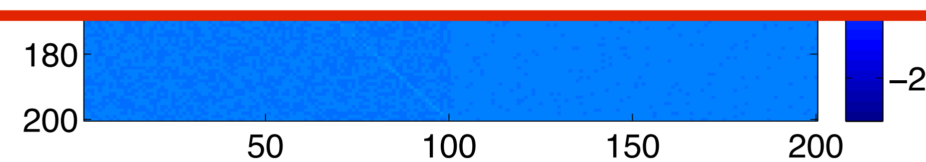


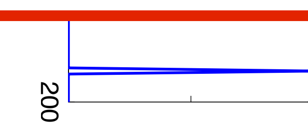
Regularization methods

$p = 2$







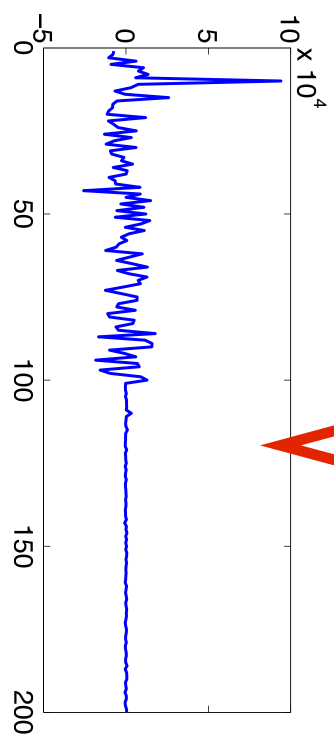


Conventional method

$\min \|$ $\Big[\begin{matrix}\delta \mathbf{r} \\ \delta\end{matrix}$

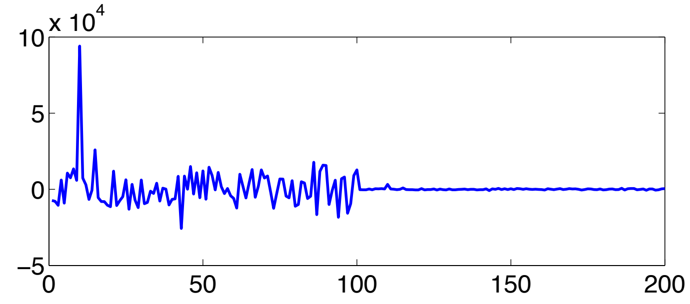
Hessian:

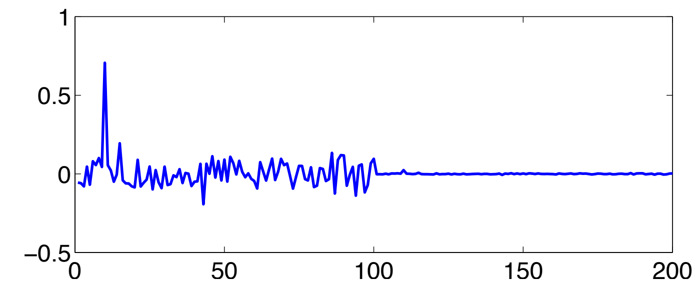
$\mathcal{H}$



Regularization methods

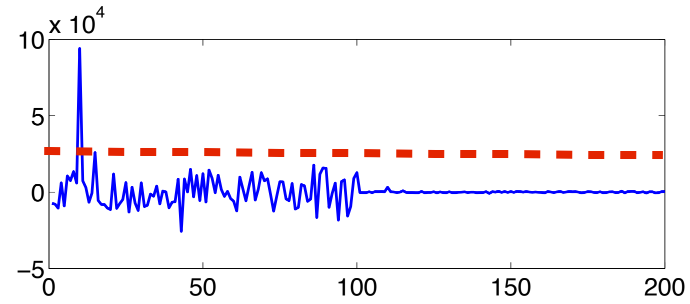
$p = 2$

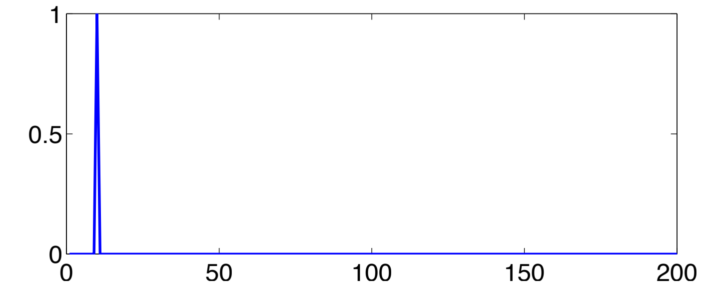


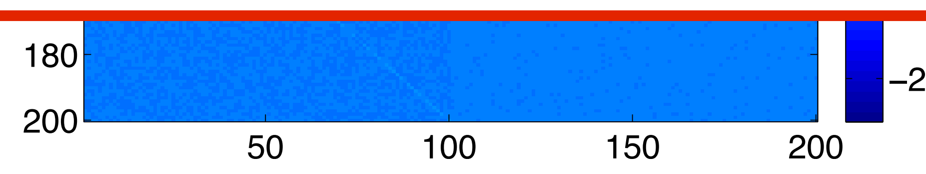


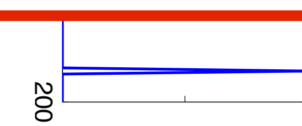
$p = 1$

Sparsity





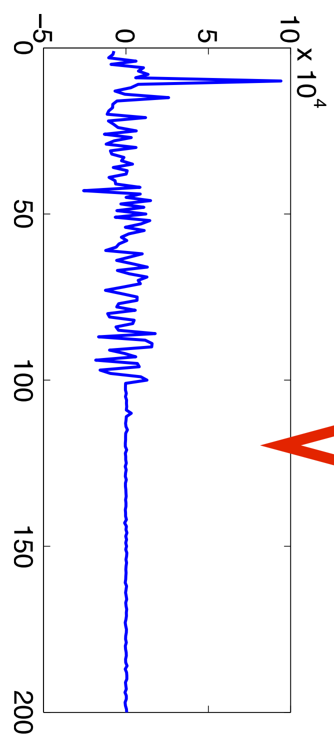




Conventional method

$\min \|$

Hessian:



Regularization methods

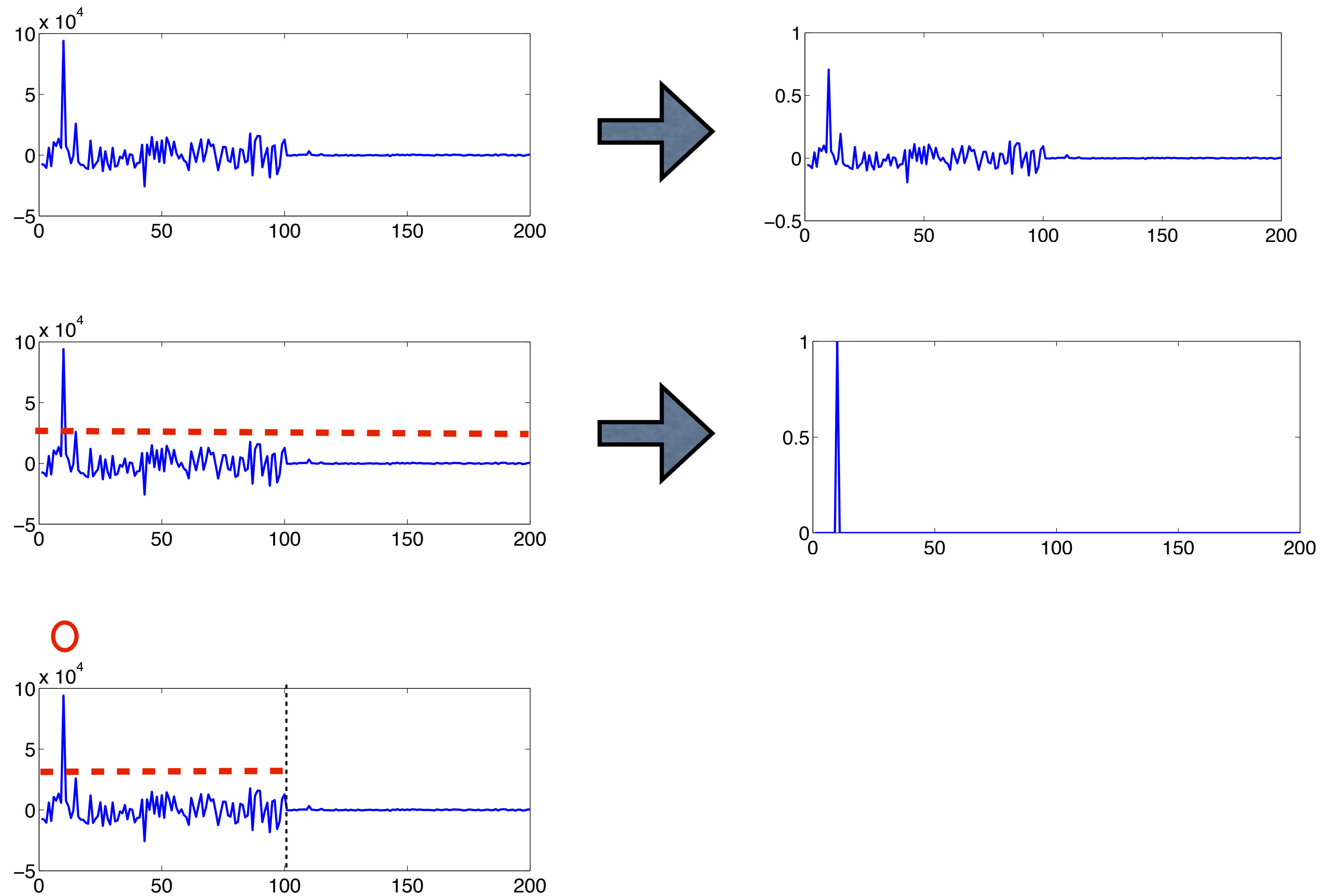
$p = 2$

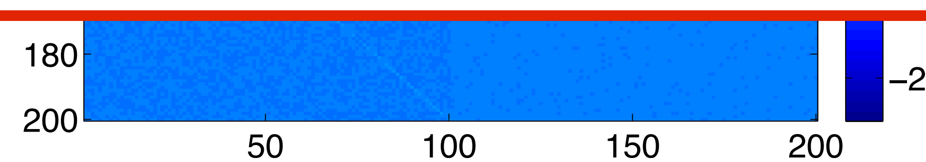
$p = 1$

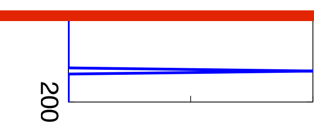
Sparsity

$p = 1, 2$

Joint-sparsity







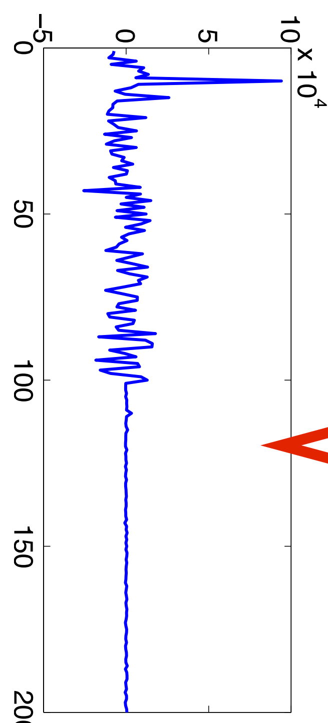
Conventional method

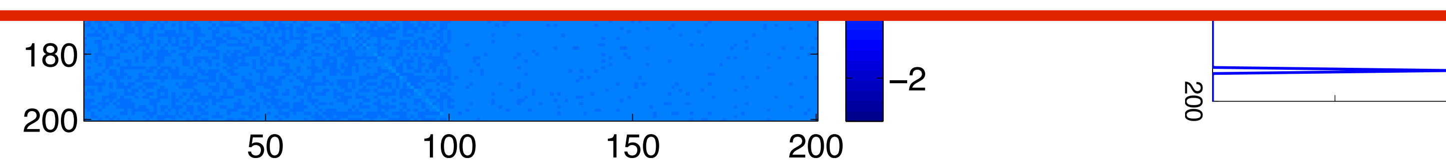
SLIM

$$\min \| \begin{bmatrix} \delta \mathbf{r} \\ \delta \end{bmatrix}$$

Hessian:

$\mathcal{H}$





Regularization methods

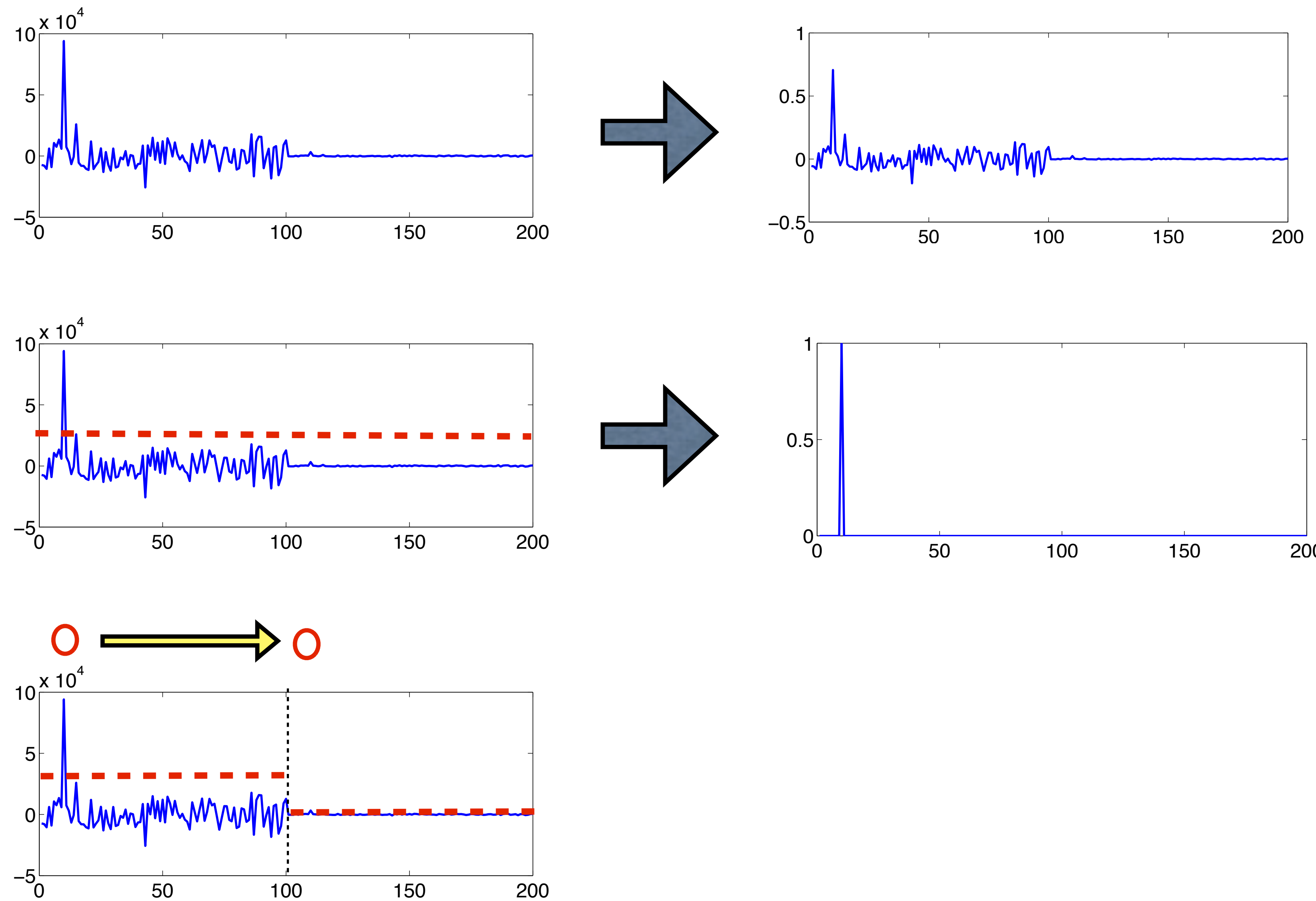
$p = 2$

$p = 1$

Sparsity

$p = 1, 2$

Joint-sparsity



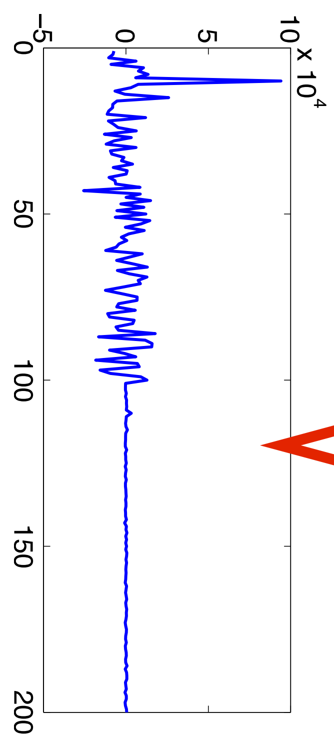
Conventional method

Regularization methods

$$\min \left\| \begin{bmatrix} \delta \mathbf{r} \\ \delta \end{bmatrix} \right.$$

Hessian:

$\mathcal{H}$



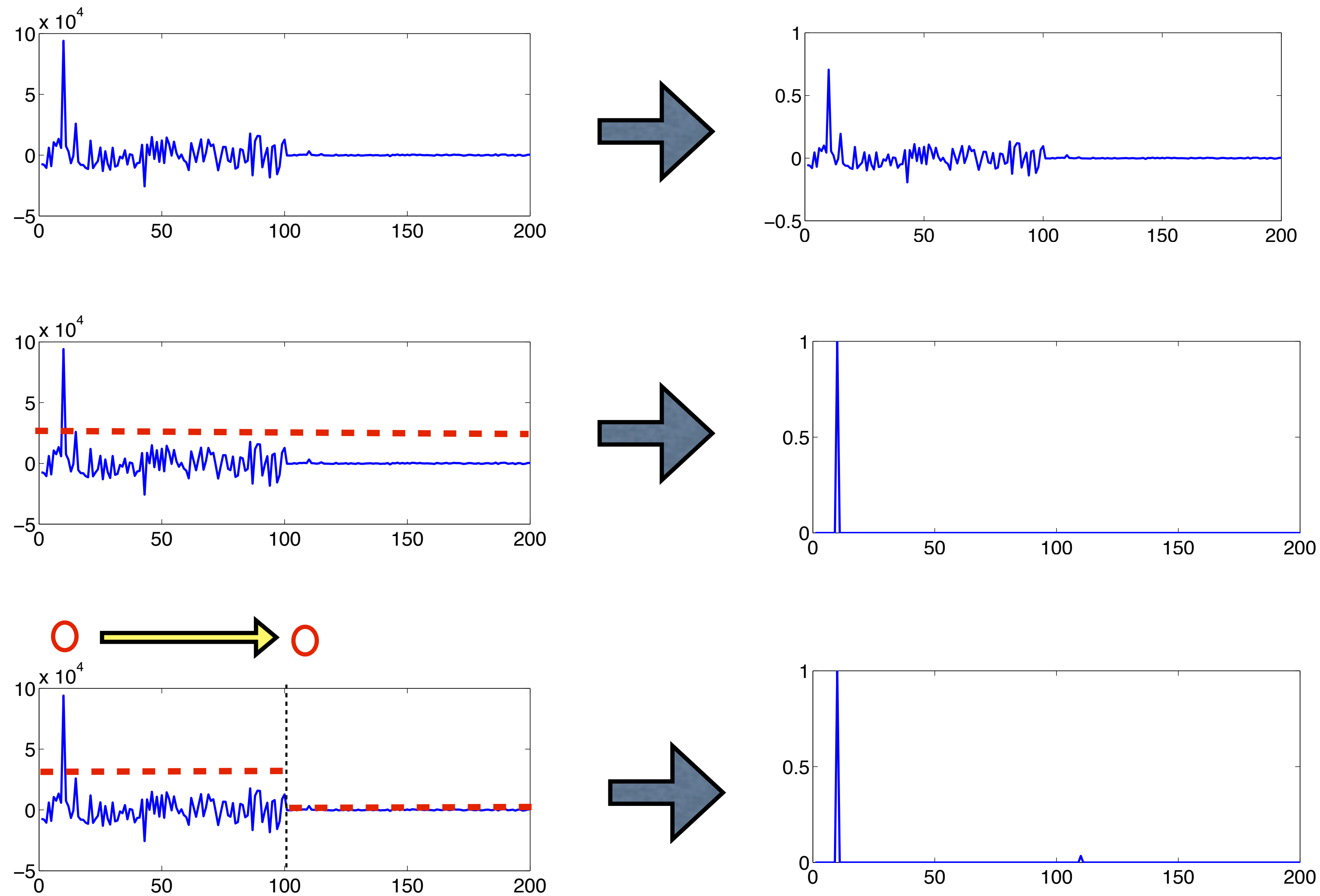
$p = 2$

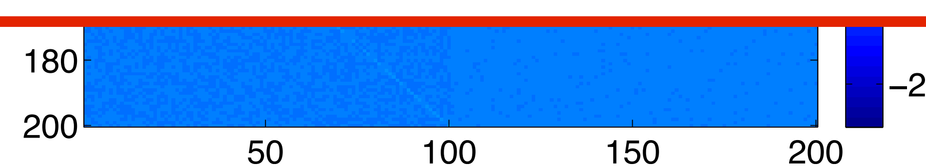
$p = 1$

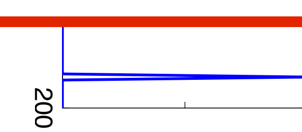
Sparsity

$p = 1, 2$

Joint-sparsity

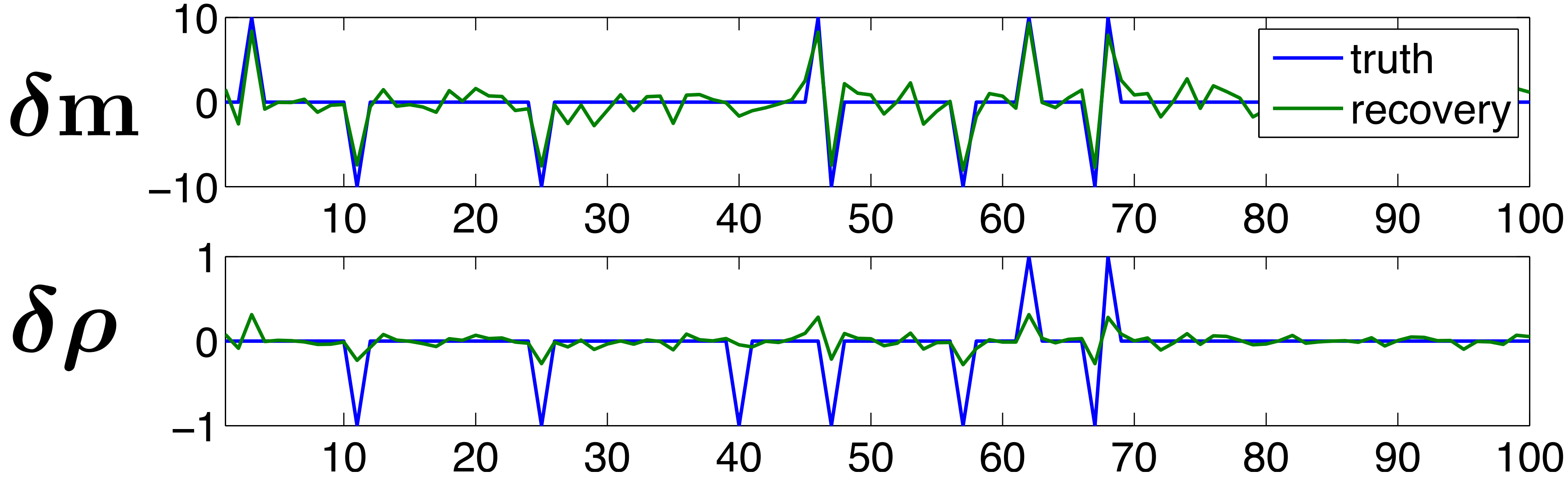




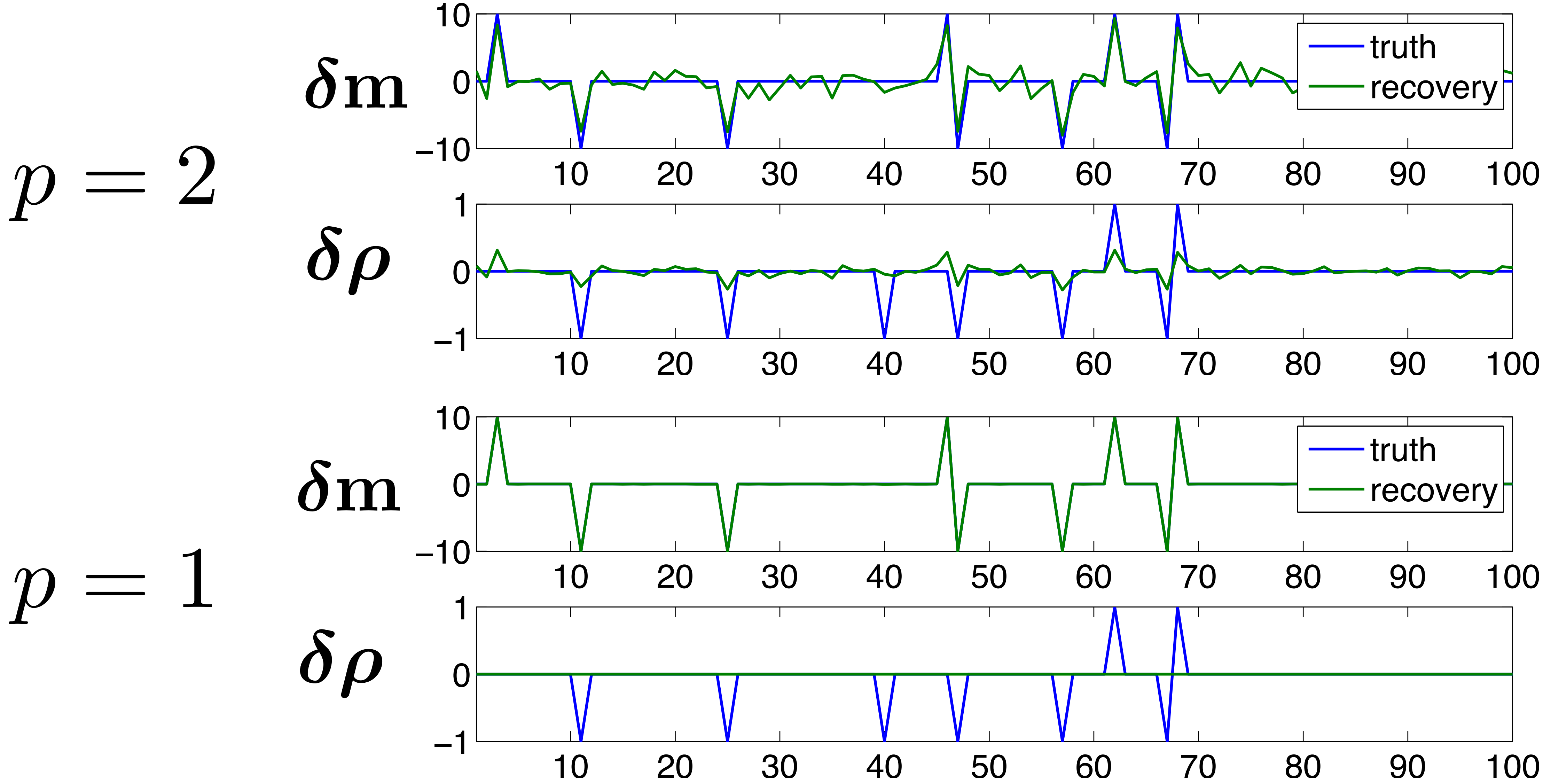


Conventional method

$p = 2$



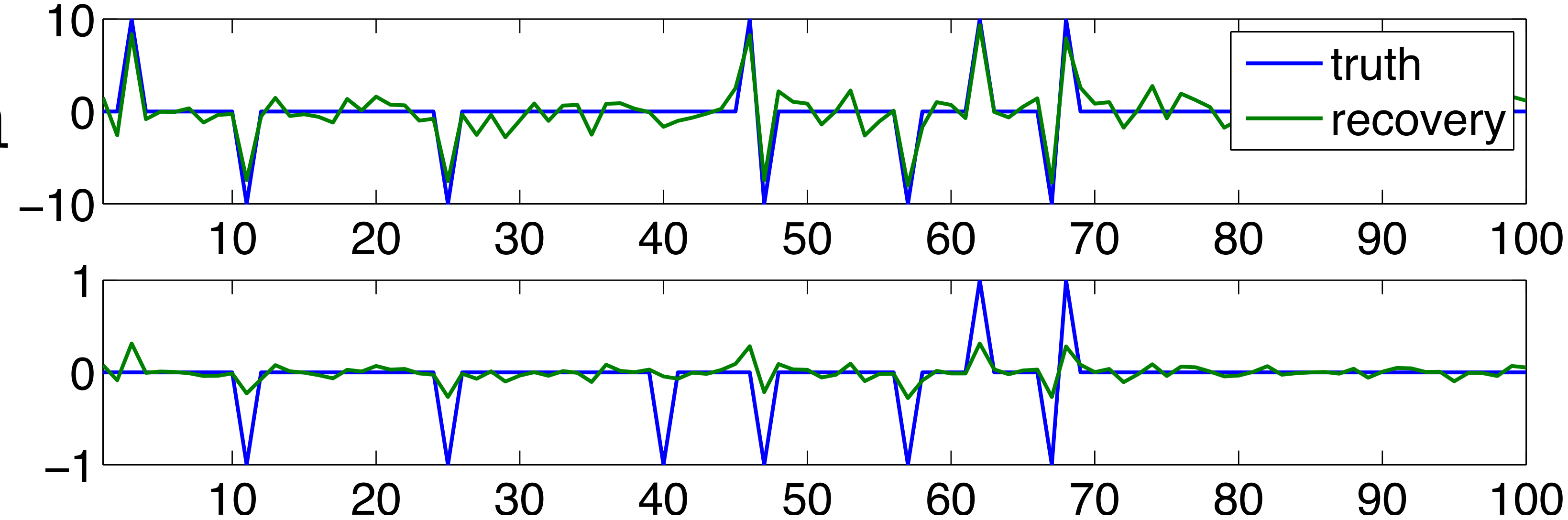
Conventional method



Conventional method

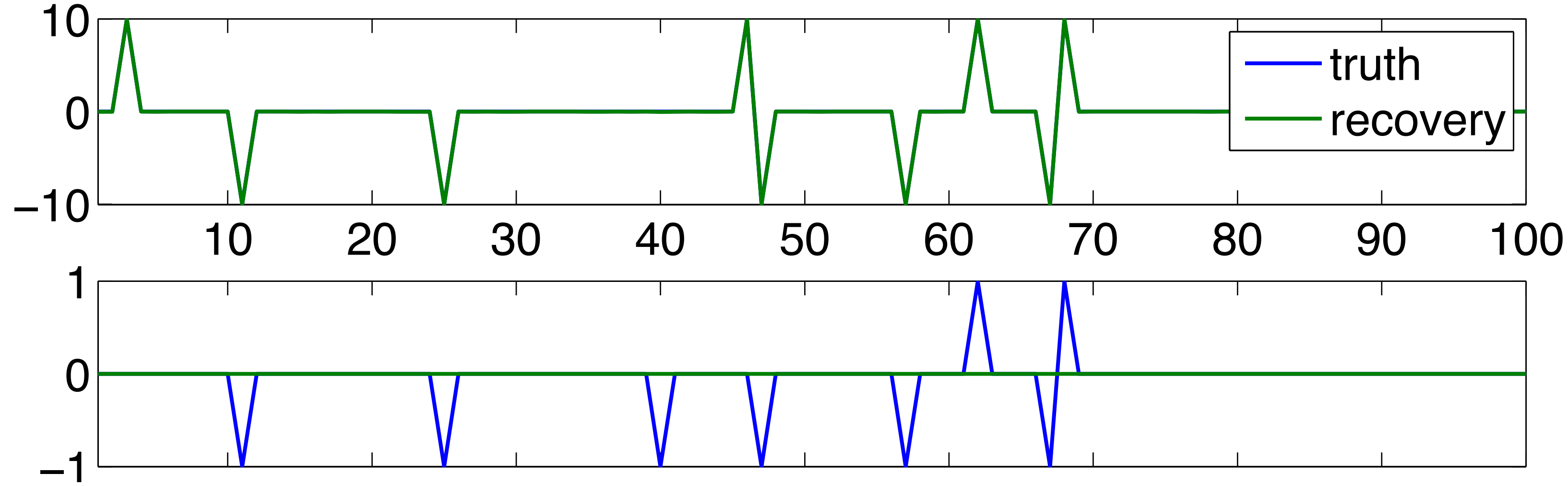
$p = 2$

$\boldsymbol{\delta}\mathbf{m}$, $\boldsymbol{\delta\rho}$



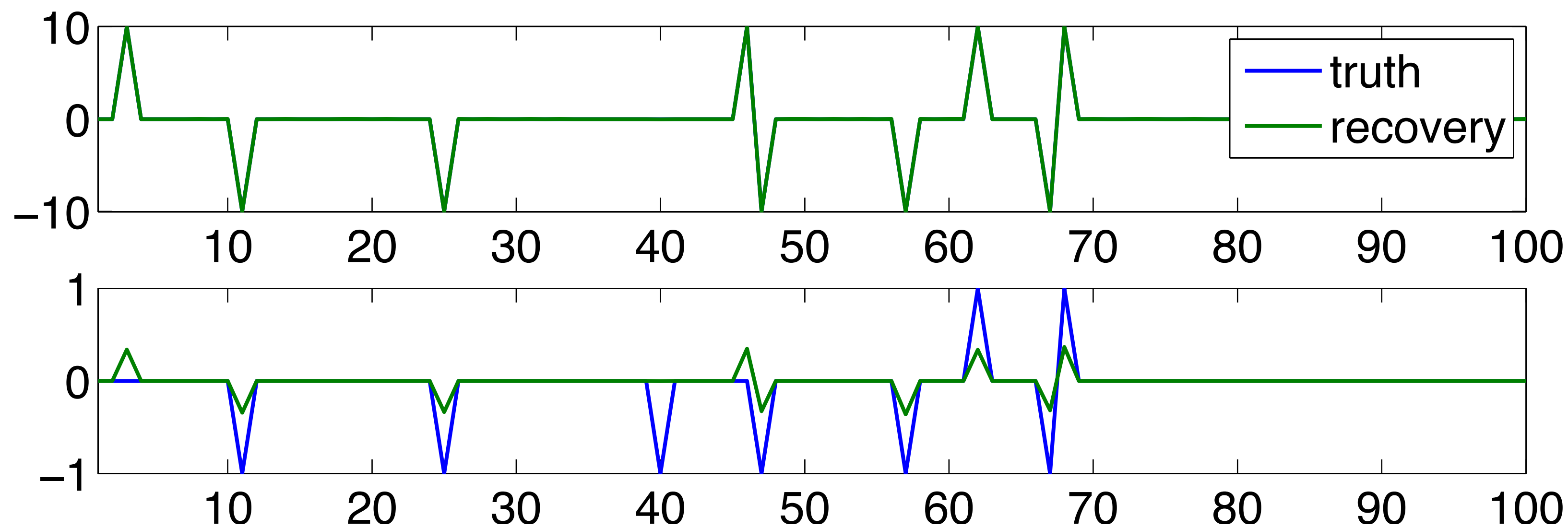
$p = 1$

$\boldsymbol{\delta}\mathbf{m}$, $\boldsymbol{\delta\rho}$



$p = 1, 2$

$\boldsymbol{\delta}\mathbf{m}$, $\boldsymbol{\delta\rho}$



Scaled method

$$\min \| \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\alpha}\delta \boldsymbol{\rho} \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta \mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\alpha}\delta \boldsymbol{\rho} \end{bmatrix} \|_2 \leq \sigma \qquad \alpha = \frac{\|\boldsymbol{\mathcal{J}}_{\mathbf{m}}\mathbf{x}\|_2}{\|\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}\mathbf{x}\|_2}$$

$\mathbf{x}$ is a random gaussian vector

Scaled method

$$\min \| \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\alpha}\delta\boldsymbol{\rho} \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta\mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\alpha}\delta\boldsymbol{\rho} \end{bmatrix} \|_2 \leq \sigma \qquad \alpha = \frac{\|\boldsymbol{\mathcal{J}}_{\mathbf{m}}\mathbf{x}\|_2}{\|\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}\mathbf{x}\|_2}$$

$\mathbf{x}$ is a random gaussian vector

Scaled Hessian:

Scaled method

$$\min \| \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\alpha}\delta\boldsymbol{\rho} \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta\mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\alpha}\delta\boldsymbol{\rho} \end{bmatrix} \|_2 \leq \sigma \qquad \alpha = \frac{\|\boldsymbol{\mathcal{J}}_{\mathbf{m}}\mathbf{x}\|_2}{\|\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}\mathbf{x}\|_2}$$

$\mathbf{x}$ is a random gaussian vector

Scaled Hessian:

$$\boldsymbol{\mathcal{H}}_{\boldsymbol{\alpha}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \\ \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \end{bmatrix} [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T\boldsymbol{\mathcal{J}}_{\mathbf{m}} & \alpha\boldsymbol{\mathcal{J}}_{\mathbf{m}}^T\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \\ \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T\boldsymbol{\mathcal{J}}_{\mathbf{m}} & \alpha^2\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \end{bmatrix}$$

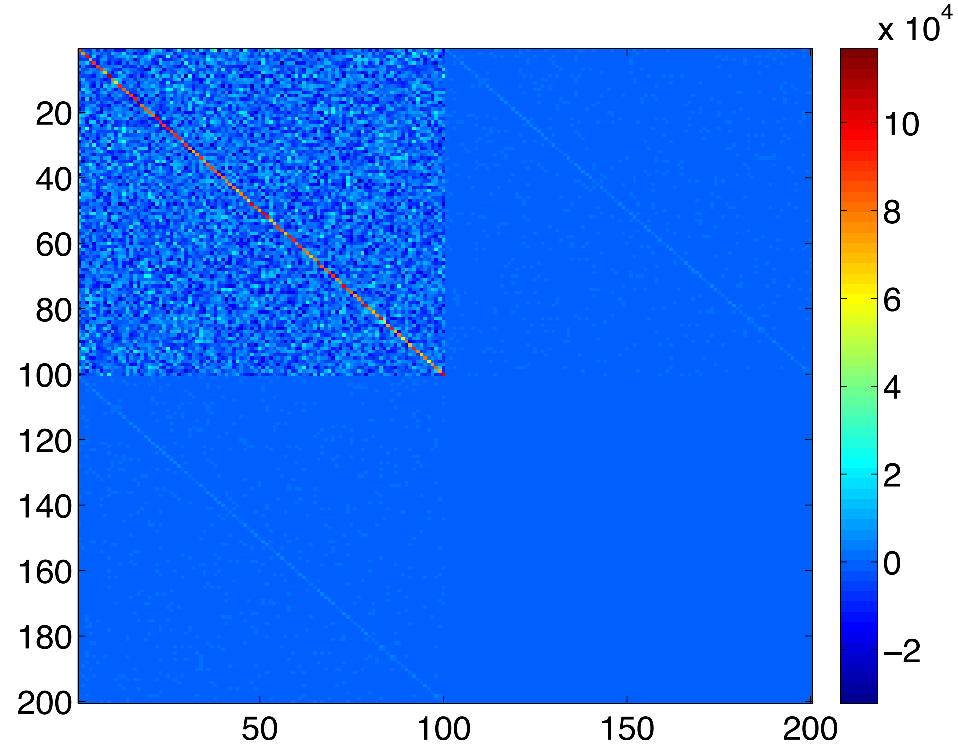
Scaled method

$$\min \left\| \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\alpha}\delta \boldsymbol{\rho} \end{bmatrix} \right\|_p \quad \text{s.t.} \quad \left\| \delta \mathbf{D} - [\mathcal{J}_{\mathbf{m}} \quad \alpha \mathcal{J}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\alpha}\delta \boldsymbol{\rho} \end{bmatrix} \right\|_2 \leq \sigma \qquad \alpha = \frac{\|\mathcal{J}_{\mathbf{m}}\mathbf{x}\|_2}{\|\mathcal{J}_{\boldsymbol{\rho}}\mathbf{x}\|_2}$$

$\mathbf{x}$ is a random gaussian vector

Scaled Hessian:

$$\mathcal{H}_{\boldsymbol{\alpha}} = \begin{bmatrix} \mathcal{J}_{\mathbf{m}}^T \\ \alpha \mathcal{J}_{\boldsymbol{\rho}}^T \end{bmatrix} [\mathcal{J}_{\mathbf{m}} \quad \alpha \mathcal{J}_{\boldsymbol{\rho}}] = \begin{bmatrix} \mathcal{J}_{\mathbf{m}}^T \mathcal{J}_{\mathbf{m}} & \alpha \mathcal{J}_{\mathbf{m}}^T \mathcal{J}_{\boldsymbol{\rho}} \\ \alpha \mathcal{J}_{\boldsymbol{\rho}}^T \mathcal{J}_{\mathbf{m}} & \alpha^2 \mathcal{J}_{\boldsymbol{\rho}}^T \mathcal{J}_{\boldsymbol{\rho}} \end{bmatrix}$$



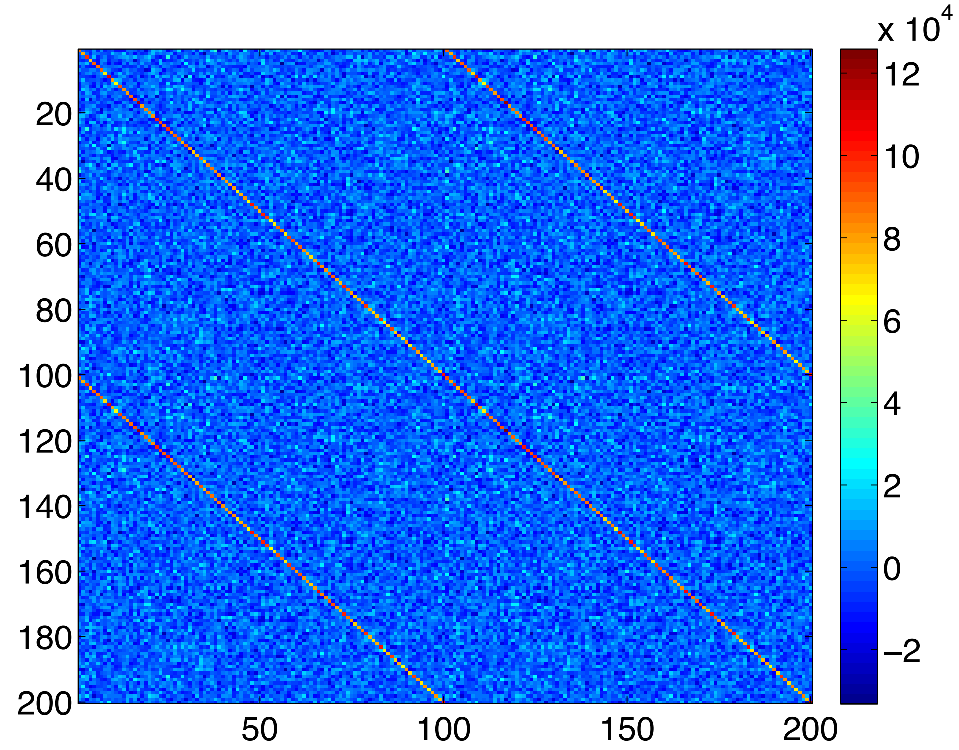
Scaled method

$$\min \| \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\alpha}\delta \boldsymbol{\rho} \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta \mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\alpha}\delta \boldsymbol{\rho} \end{bmatrix} \|_2 \leq \sigma \qquad \alpha = \frac{\|\boldsymbol{\mathcal{J}}_{\mathbf{m}}\mathbf{x}\|_2}{\|\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}\mathbf{x}\|_2}$$

$\mathbf{x}$ is a random gaussian vector

Scaled Hessian:

$$\boldsymbol{\mathcal{H}}_{\boldsymbol{\alpha}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \\ \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \end{bmatrix} [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \boldsymbol{\mathcal{J}}_{\mathbf{m}} & \alpha \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \\ \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \boldsymbol{\mathcal{J}}_{\mathbf{m}} & \alpha^2 \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \end{bmatrix}$$



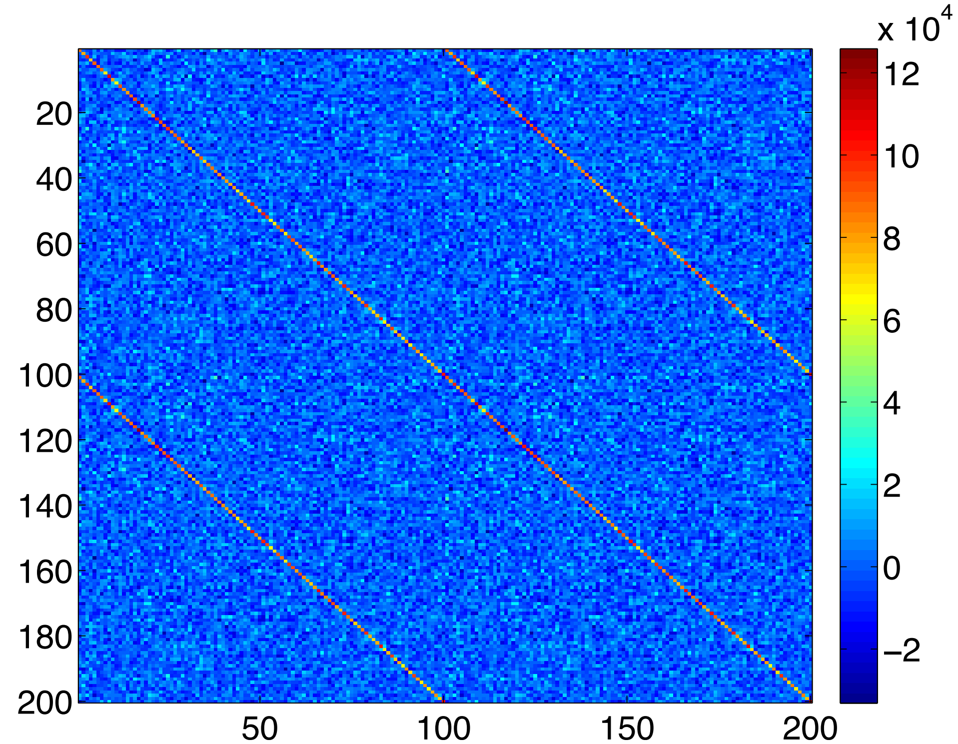
Scaled method

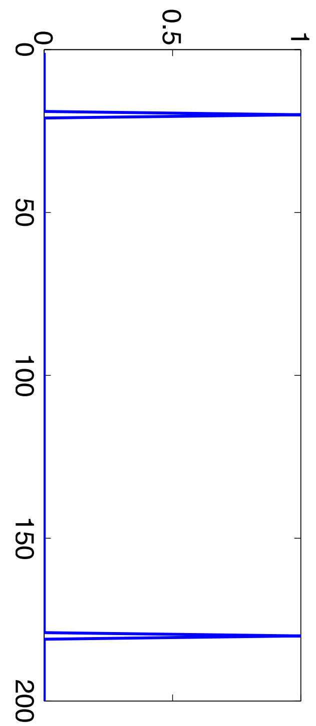
$$\min \left\| \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\alpha}\delta\boldsymbol{\rho} \end{bmatrix} \right\|_p \quad \text{s.t.} \quad \left\| \delta\mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\alpha}\delta\boldsymbol{\rho} \end{bmatrix} \right\|_2 \leq \sigma \qquad \alpha = \frac{\|\boldsymbol{\mathcal{J}}_{\mathbf{m}}\mathbf{x}\|_2}{\|\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}\mathbf{x}\|_2}$$

$\mathbf{x}$ is a random gaussian vector

Scaled Hessian:

$$\boldsymbol{\mathcal{H}}_{\boldsymbol{\alpha}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \\ \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \end{bmatrix} [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T\boldsymbol{\mathcal{J}}_{\mathbf{m}} & \alpha\boldsymbol{\mathcal{J}}_{\mathbf{m}}^T\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \\ \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T\boldsymbol{\mathcal{J}}_{\mathbf{m}} & \alpha^2\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \end{bmatrix}$$





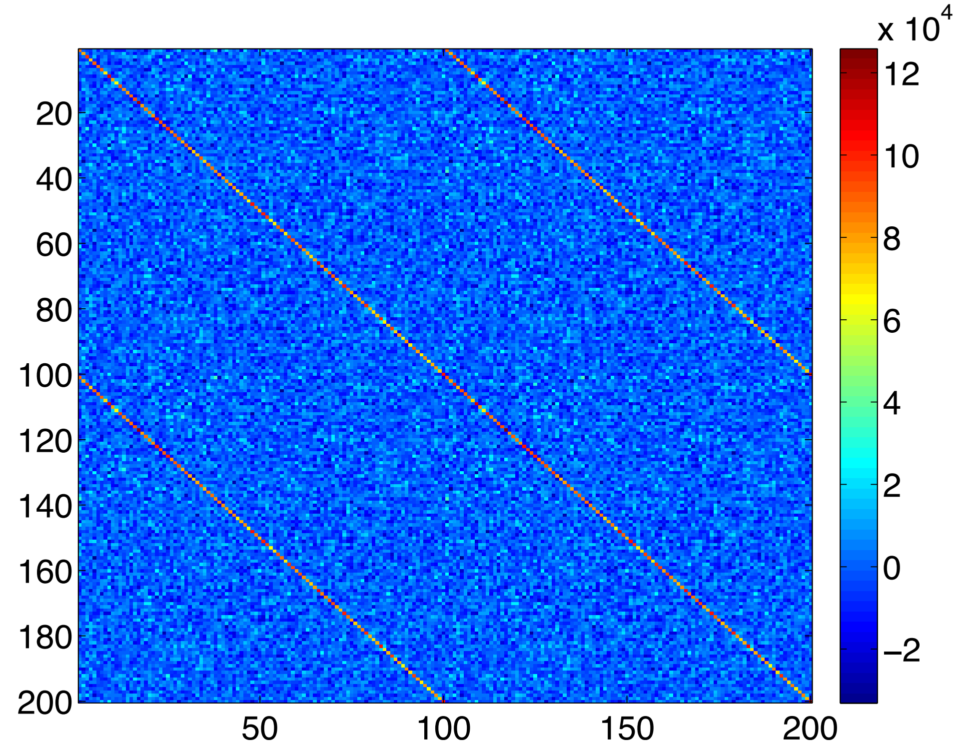
Scaled method

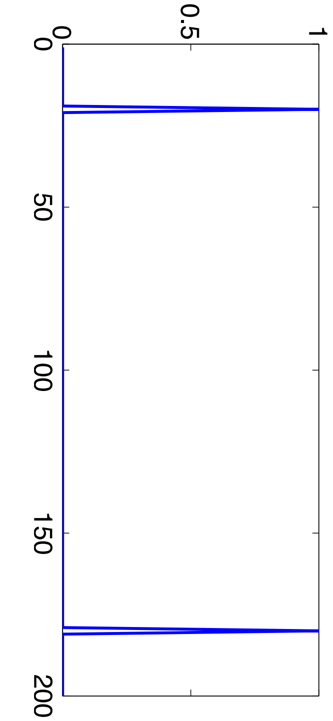
$$\min \left\| \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\alpha}\delta \boldsymbol{\rho} \end{bmatrix} \right\|_p \quad \text{s.t.} \quad \left\| \delta \mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\alpha}\delta \boldsymbol{\rho} \end{bmatrix} \right\|_2 \leq \sigma \qquad \alpha = \frac{\|\boldsymbol{\mathcal{J}}_{\mathbf{m}}\mathbf{x}\|_2}{\|\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}\mathbf{x}\|_2}$$

$\mathbf{x}$ is a random gaussian vector

Scaled Hessian:

$$\mathcal{H}_{\boldsymbol{\alpha}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \\ \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \end{bmatrix} [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \boldsymbol{\mathcal{J}}_{\mathbf{m}} & \alpha \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \\ \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \boldsymbol{\mathcal{J}}_{\mathbf{m}} & \alpha^2 \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \end{bmatrix}$$





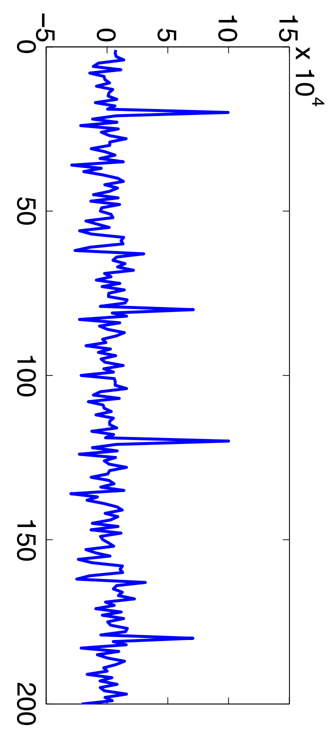
Scaled method

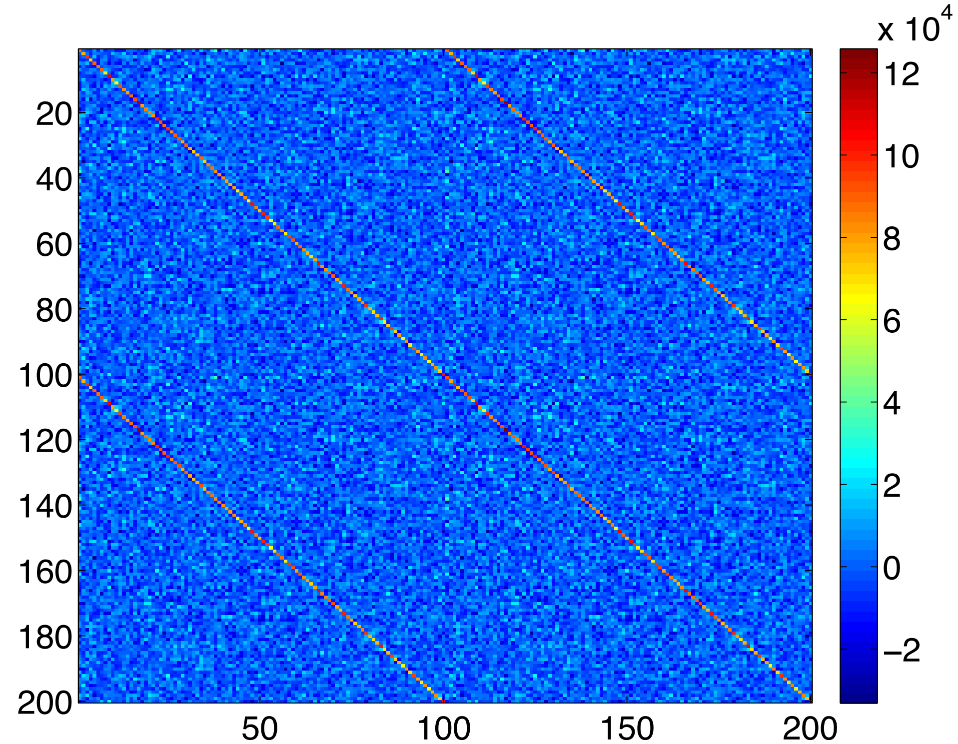
$$\min \left\| \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\alpha}\delta \boldsymbol{\rho} \end{bmatrix} \right\|_p \quad \text{s.t.} \quad \left\| \delta \mathbf{D} - [\mathcal{J}_{\mathbf{m}} \quad \alpha \mathcal{J}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\alpha}\delta \boldsymbol{\rho} \end{bmatrix} \right\|_2 \leq \sigma \qquad \alpha = \frac{\|\mathcal{J}_{\mathbf{m}} \mathbf{x}\|_2}{\|\mathcal{J}_{\boldsymbol{\rho}} \mathbf{x}\|_2}$$

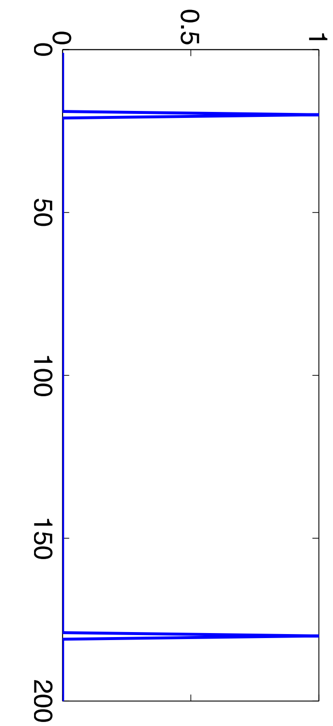
$\mathbf{x}$ is a random gaussian vector

Scaled Hessian:

$$\mathcal{H}_{\boldsymbol{\alpha}} = \begin{bmatrix} \mathcal{J}_{\mathbf{m}}^T \\ \alpha \mathcal{J}_{\boldsymbol{\rho}}^T \end{bmatrix} [\mathcal{J}_{\mathbf{m}} \quad \alpha \mathcal{J}_{\boldsymbol{\rho}}] = \begin{bmatrix} \mathcal{J}_{\mathbf{m}}^T \mathcal{J}_{\mathbf{m}} & \alpha \mathcal{J}_{\mathbf{m}}^T \mathcal{J}_{\boldsymbol{\rho}} \\ \alpha \mathcal{J}_{\boldsymbol{\rho}}^T \mathcal{J}_{\mathbf{m}} & \alpha^2 \mathcal{J}_{\boldsymbol{\rho}}^T \mathcal{J}_{\boldsymbol{\rho}} \end{bmatrix}$$



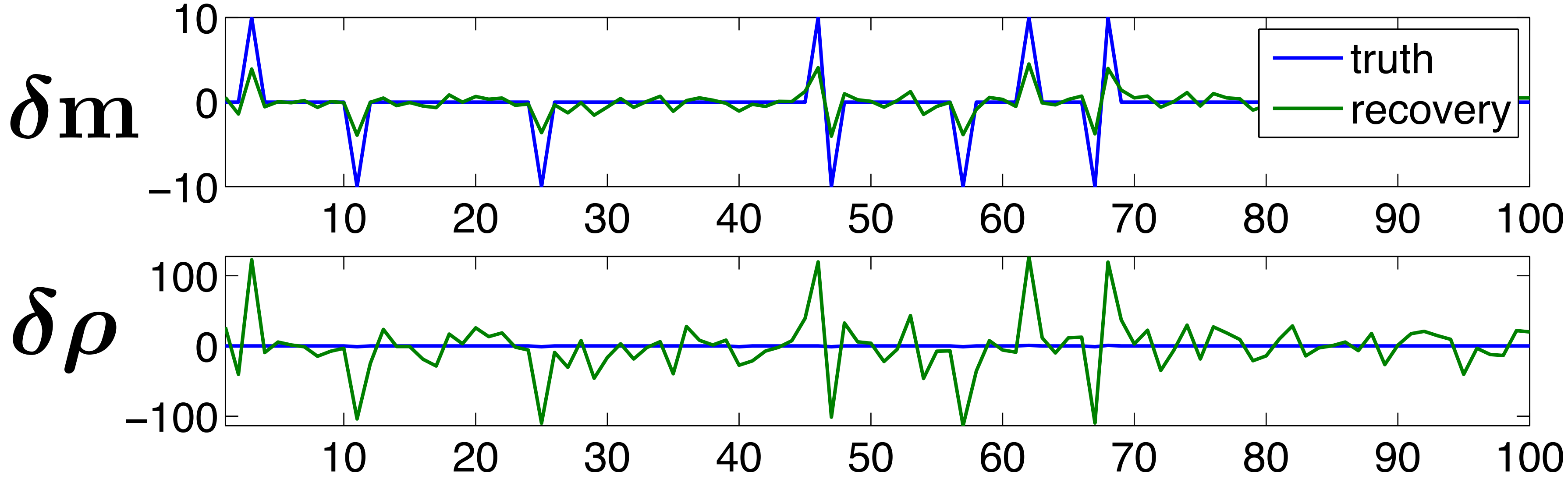




Scaled method

Scaled method

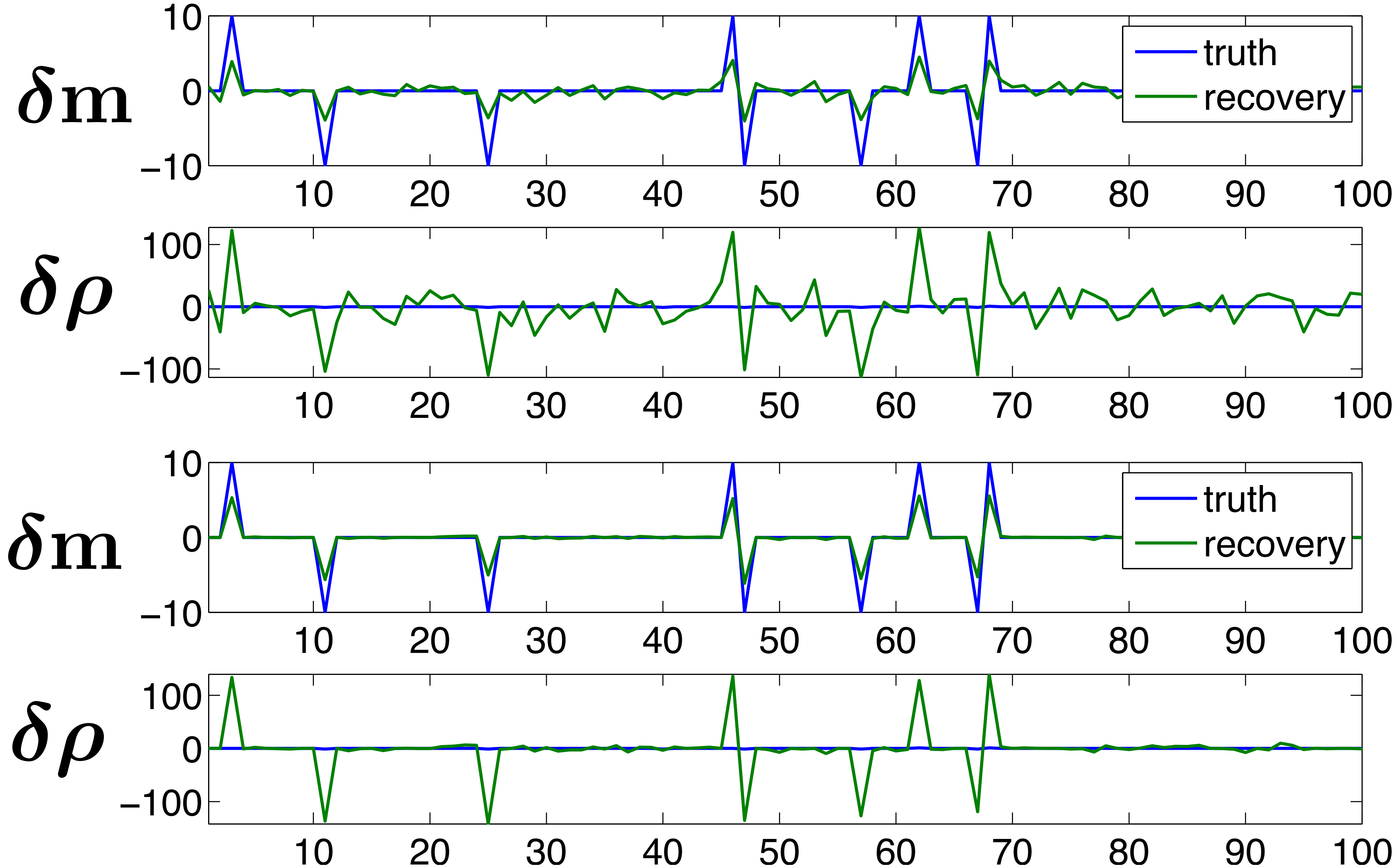
$$p = 2$$



Scaled method

$p = 2$

$p = 1$



Scaled method

$p = 2$

$\boldsymbol{\delta}\mathbf{m}$

$\boldsymbol{\delta\rho}$

$p = 1$

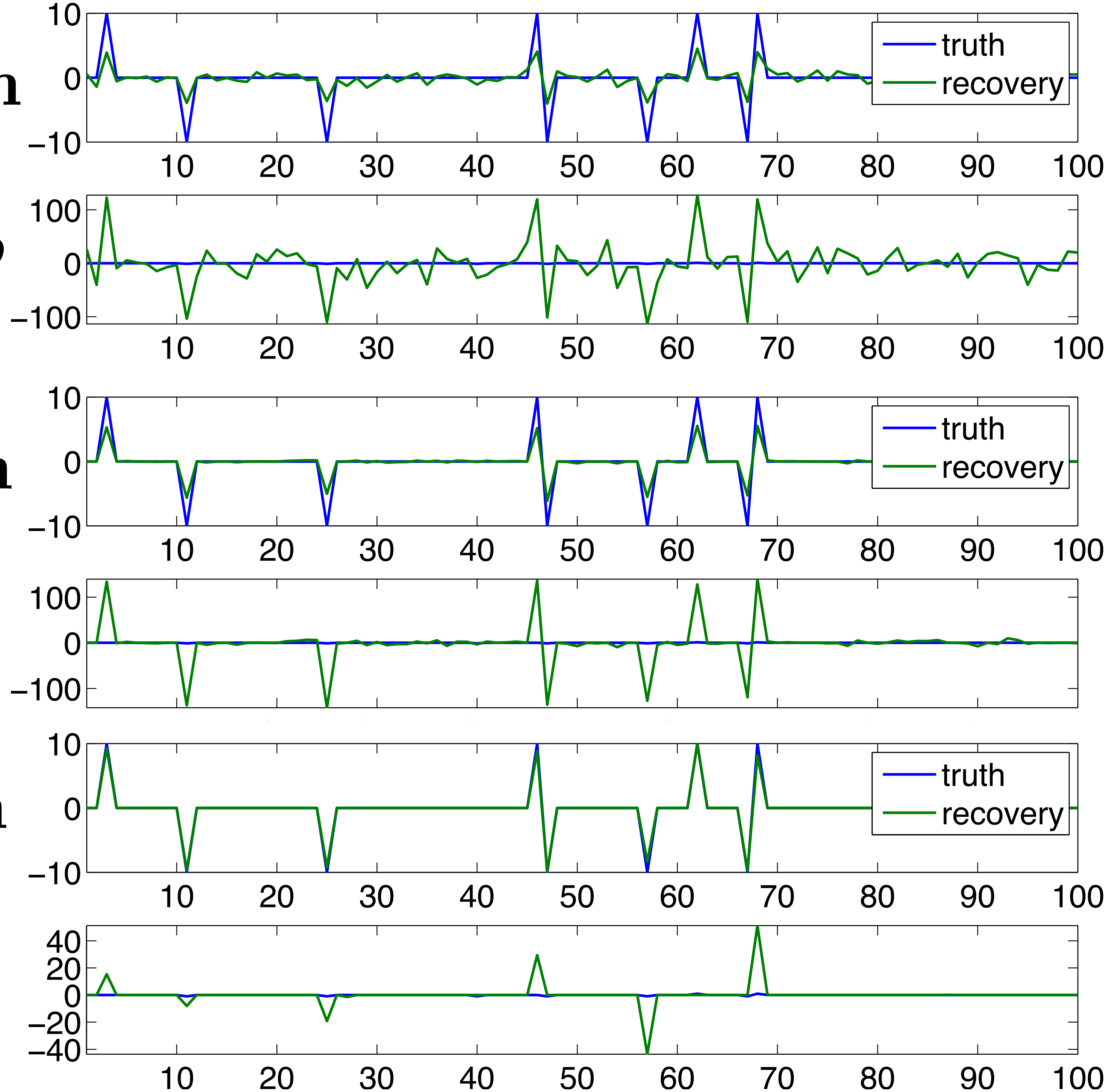
$\boldsymbol{\delta}\mathbf{m}$

$\boldsymbol{\delta\rho}$

$p = 1, 2$

$\boldsymbol{\delta}\mathbf{m}$

$\boldsymbol{\delta\rho}$



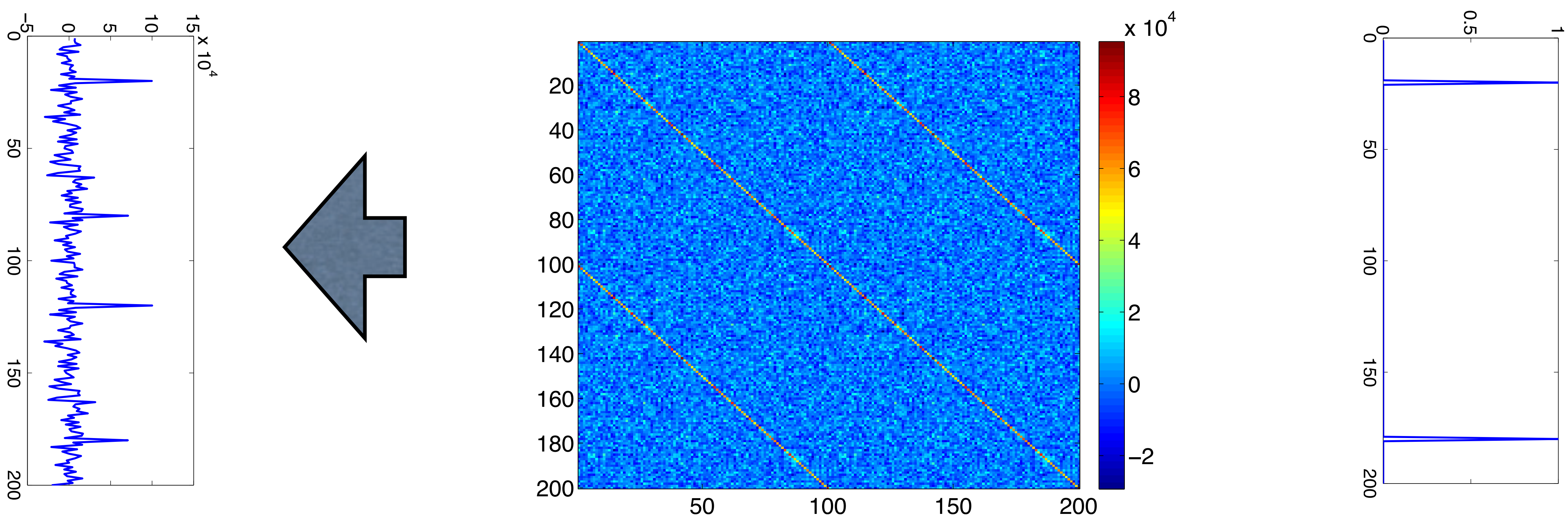
Scaled method

$$\min \left\| \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\alpha}\delta \boldsymbol{\rho} \end{bmatrix} \right\|_p \quad \text{s.t.} \quad \left\| \delta \mathbf{D} - [\mathcal{J}_{\mathbf{m}} \quad \alpha \mathcal{J}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\alpha}\delta \boldsymbol{\rho} \end{bmatrix} \right\|_2 \leq \sigma \qquad \alpha = \frac{\|\mathcal{J}_{\mathbf{m}}\mathbf{x}\|_2}{\|\mathcal{J}_{\boldsymbol{\rho}}\mathbf{x}\|_2}$$

$\mathbf{x}$ is a random gaussian vector

Scaled Hessian:

$$\mathcal{H}_{\boldsymbol{\alpha}} = \begin{bmatrix} \mathcal{J}_{\mathbf{m}}^T \\ \alpha \mathcal{J}_{\boldsymbol{\rho}}^T \end{bmatrix} [\mathcal{J}_{\mathbf{m}} \quad \alpha \mathcal{J}_{\boldsymbol{\rho}}] = \begin{bmatrix} \mathcal{J}_{\mathbf{m}}^T \mathcal{J}_{\mathbf{m}} & \alpha \mathcal{J}_{\mathbf{m}}^T \mathcal{J}_{\boldsymbol{\rho}} \\ \alpha \mathcal{J}_{\boldsymbol{\rho}}^T \mathcal{J}_{\mathbf{m}} & \alpha^2 \mathcal{J}_{\boldsymbol{\rho}}^T \mathcal{J}_{\boldsymbol{\rho}} \end{bmatrix}$$



Preconditioned method

$$\min \| \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\alpha}\delta\boldsymbol{\rho} \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta\mathbf{D} - [\boldsymbol{\mathcal{J}}_\mathbf{m} \quad \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\alpha}\delta\boldsymbol{\rho} \end{bmatrix} \|_2 \leq \sigma \qquad \alpha = \frac{\|\boldsymbol{\mathcal{J}}_\mathbf{m}\mathbf{x}\|_2}{\|\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}\mathbf{x}\|_2}$$

$\mathbf{x}$ is a random gaussian vector

Preconditioned method

$$\min \| \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\alpha}\delta \boldsymbol{\rho} \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta \mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\alpha}\delta \boldsymbol{\rho} \end{bmatrix} \|_2 \leq \sigma \qquad \alpha = \frac{\|\boldsymbol{\mathcal{J}}_{\mathbf{m}}\mathbf{x}\|_2}{\|\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}\mathbf{x}\|_2}$$

$\mathbf{x}$ is a random gaussian vector

break the "*coherence*" of the two *Jacobians* by

$$\min \| \begin{bmatrix} \delta \mathbf{m} \\ (\frac{1}{\alpha}\delta \boldsymbol{\rho} + \delta \mathbf{m}) \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta \mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \quad \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta \mathbf{m} \\ (\frac{1}{\alpha}\delta \boldsymbol{\rho} + \delta \mathbf{m}) \end{bmatrix} \|_2 \leq \sigma$$

Preconditioned method

$$\min \| \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\alpha}\delta\boldsymbol{\rho} \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta\mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\alpha}\delta\boldsymbol{\rho} \end{bmatrix} \|_2 \leq \sigma \qquad \alpha = \frac{\|\boldsymbol{\mathcal{J}}_{\mathbf{m}}\mathbf{x}\|_2}{\|\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}\mathbf{x}\|_2}$$

$\mathbf{x}$ is a random gaussian vector

break the "*coherence*" of the two *Jacobians* by

$$\min \| \begin{bmatrix} \delta\mathbf{m} \\ (\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta\mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \quad \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta\mathbf{m} \\ (\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \|_2 \leq \sigma$$

then, we *scale* the Hessian

$$\min \| \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta\mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \quad \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \|_2 \leq \sigma$$

Preconditioned method

SLIM

$$\min \| \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\alpha}\delta \boldsymbol{\rho} \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta \mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} \quad \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\alpha}\delta \boldsymbol{\rho} \end{bmatrix} \|_2 \leq \sigma \qquad \alpha = \frac{\|\boldsymbol{\mathcal{J}}_{\mathbf{m}}\mathbf{x}\|_2}{\|\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}\mathbf{x}\|_2}$$

$\mathbf{x}$ is a random gaussian vector

break the "*coherence*" of the two *Jacobians* by

$$\min \| \begin{bmatrix} \delta \mathbf{m} \\ (\frac{1}{\alpha}\delta \boldsymbol{\rho} + \delta \mathbf{m}) \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta \mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \quad \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta \mathbf{m} \\ (\frac{1}{\alpha}\delta \boldsymbol{\rho} + \delta \mathbf{m}) \end{bmatrix} \|_2 \leq \sigma$$

then, we *scale* the Hessian

$$\min \| \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta \boldsymbol{\rho} + \delta \mathbf{m}) \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta \mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \quad \beta\alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta \mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta \boldsymbol{\rho} + \delta \mathbf{m}) \end{bmatrix} \|_2 \leq \sigma$$

$$\beta = \frac{\|(\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}})\mathbf{x}\|_2}{\|\alpha \boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}\mathbf{x}\|_2}$$

Improved "*incoherence*"

$$\min \| \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta\mathbf{D} - [\boldsymbol{\mathcal{J}}\mathbf{m} - \alpha\boldsymbol{\mathcal{J}}\boldsymbol{\rho} \quad \beta\alpha\boldsymbol{\mathcal{J}}\boldsymbol{\rho}] \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \|_2 \leq \sigma$$

Improved "*incoherence*"

$$\min \| \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta\mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \quad \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \|_2 \leq \sigma$$

Hessian:

Improved "*incoherence*"

$$\min \| \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \|_p \quad \text{s.t.} \quad \|\delta\mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \quad \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \|_2 \leq \sigma$$

Hessian:

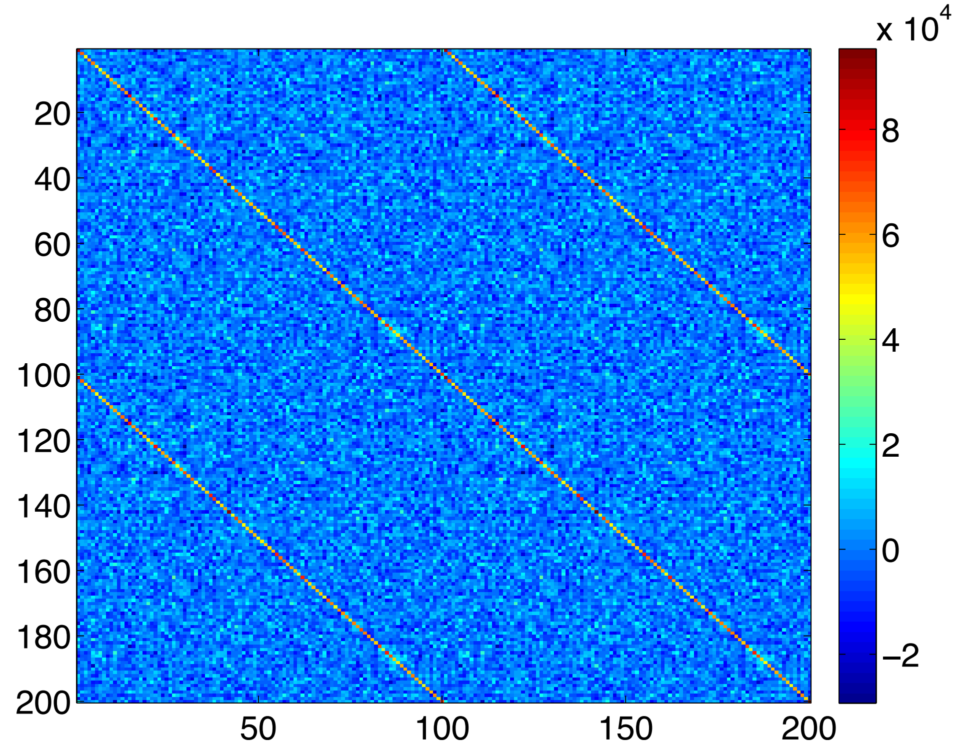
$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \\ \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \end{bmatrix} [\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \quad \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}]$$

Improved "*incoherence*"

$$\min \left\| \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \right\|_p \quad \text{s.t.} \quad \left\| \delta\mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \quad \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \right\|_2 \leq \sigma$$

Hessian:

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \\ \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \end{bmatrix} [\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \quad \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}]$$

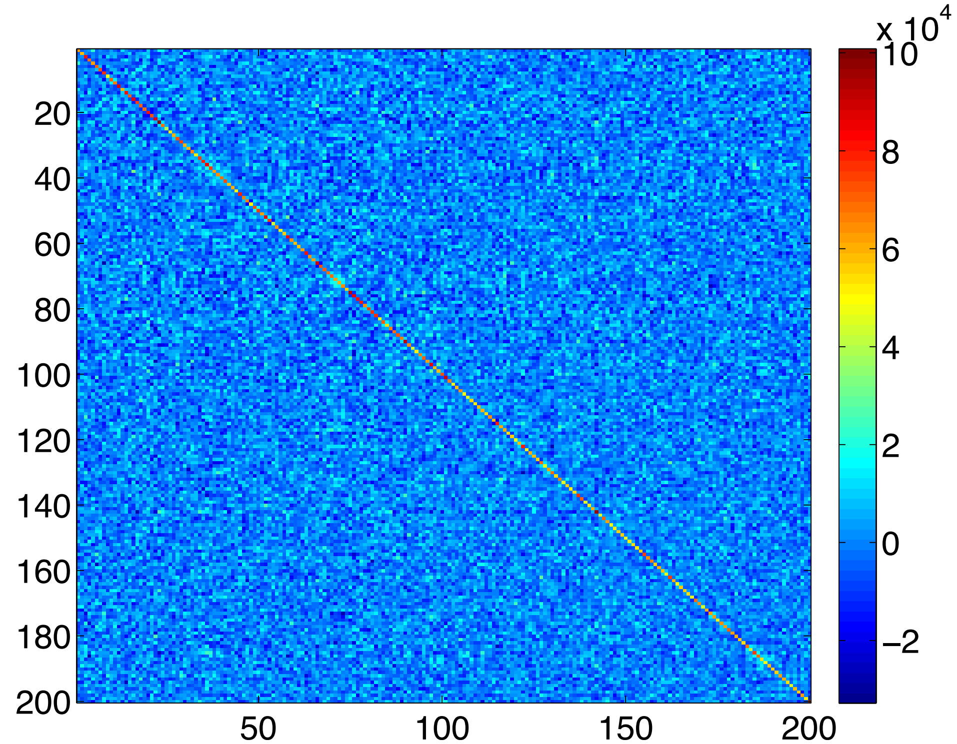


Improved "*incoherence*"

$$\min \left\| \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \right\|_p \quad \text{s.t.} \quad \left\| \delta\mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \quad \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \right\|_2 \leq \sigma$$

Hessian:

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \\ \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \end{bmatrix} [\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \quad \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}]$$

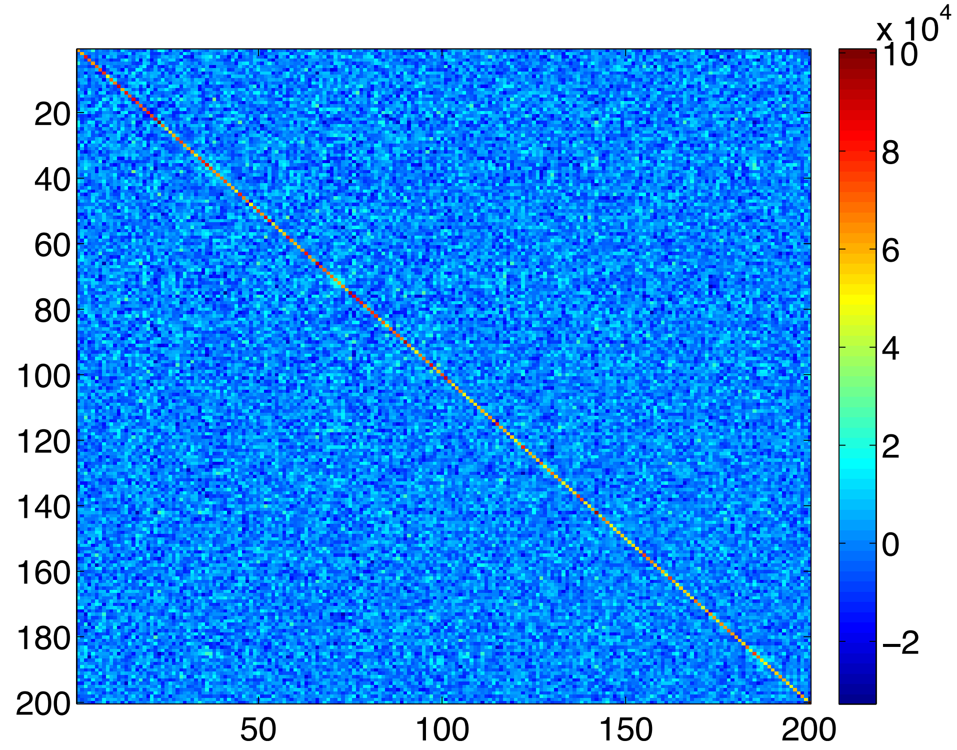


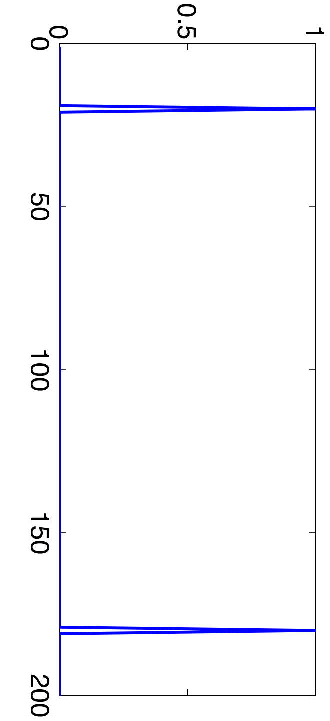
Improved "*incoherence*"

$$\min \left\| \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \right\|_p \quad \text{s.t.} \quad \left\| \delta\mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \quad \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \right\|_2 \leq \sigma$$

Hessian:

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \\ \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \end{bmatrix} [\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \quad \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}]$$



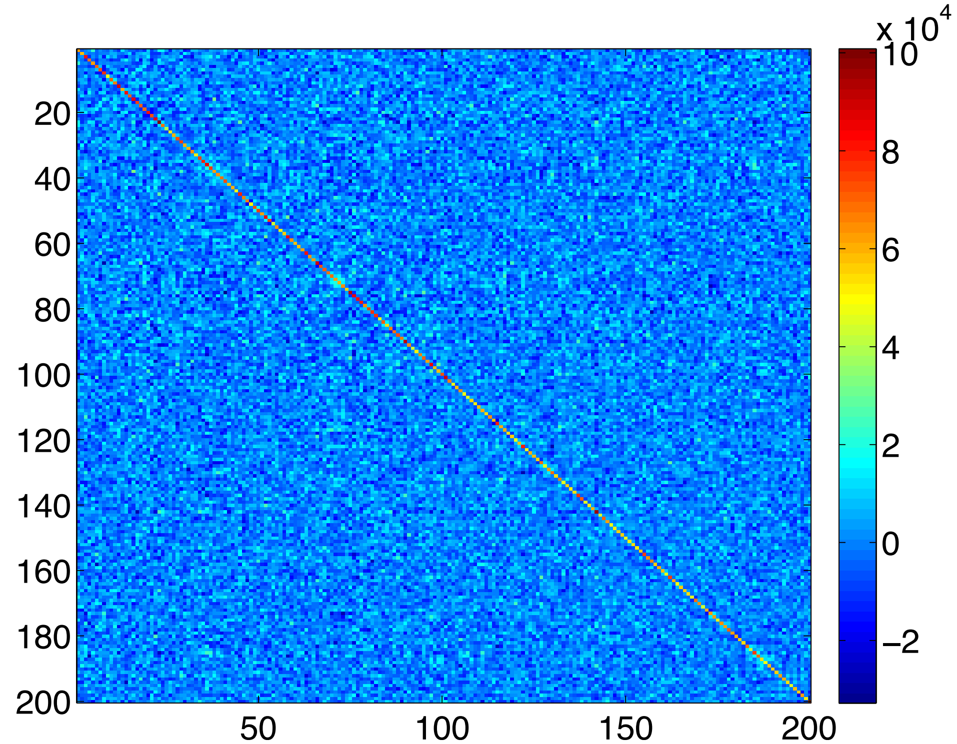


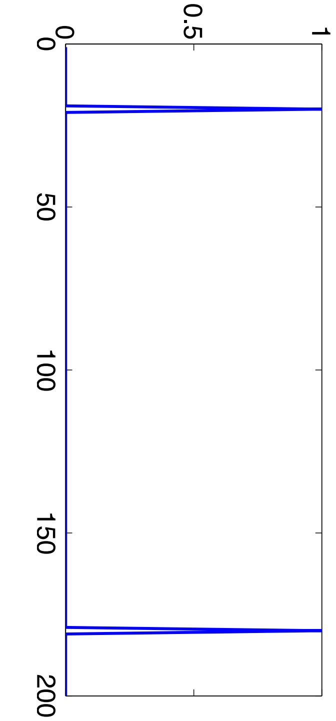
Improved "*incoherence*"

$$\min \left\| \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \right\|_p \quad \text{s.t.} \quad \left\| \delta\mathbf{D} - [\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \quad \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}] \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \right\|_2 \leq \sigma$$

Hessian:

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \\ \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \end{bmatrix} [\boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \quad \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}]$$



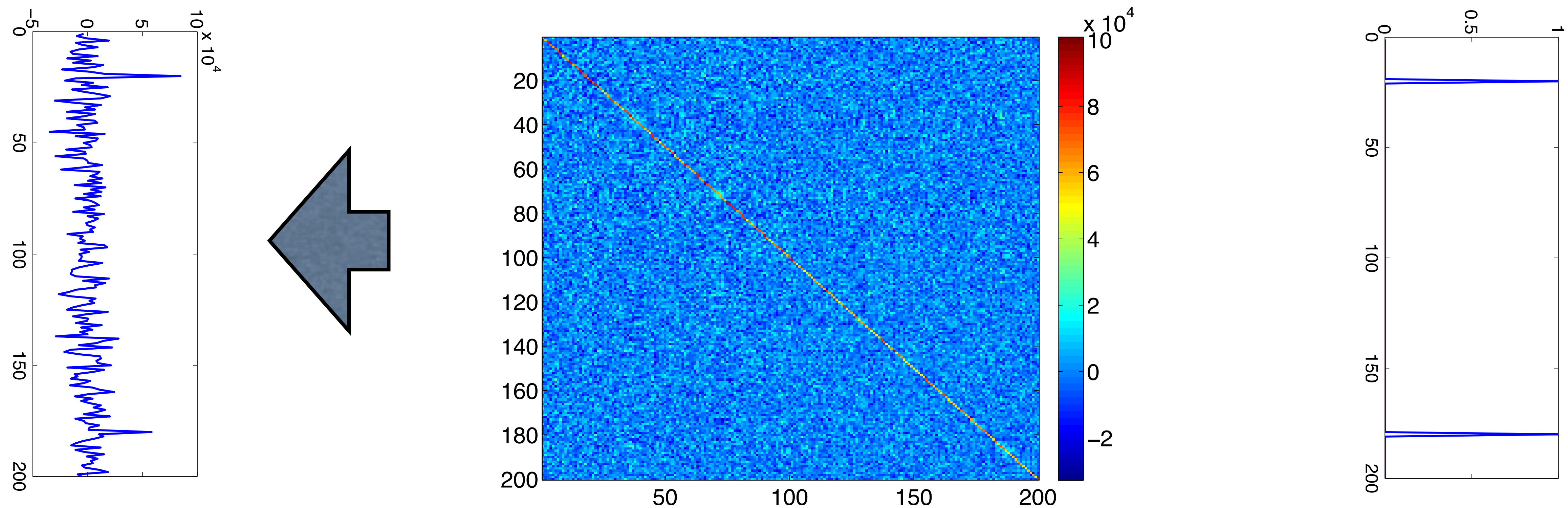


Improved "*incoherence*"

$$\min \left\| \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \right\|_p \quad \text{s.t.} \quad \left\| \delta\mathbf{D} - \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} & \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \end{bmatrix} \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \right\|_2 \leq \sigma$$

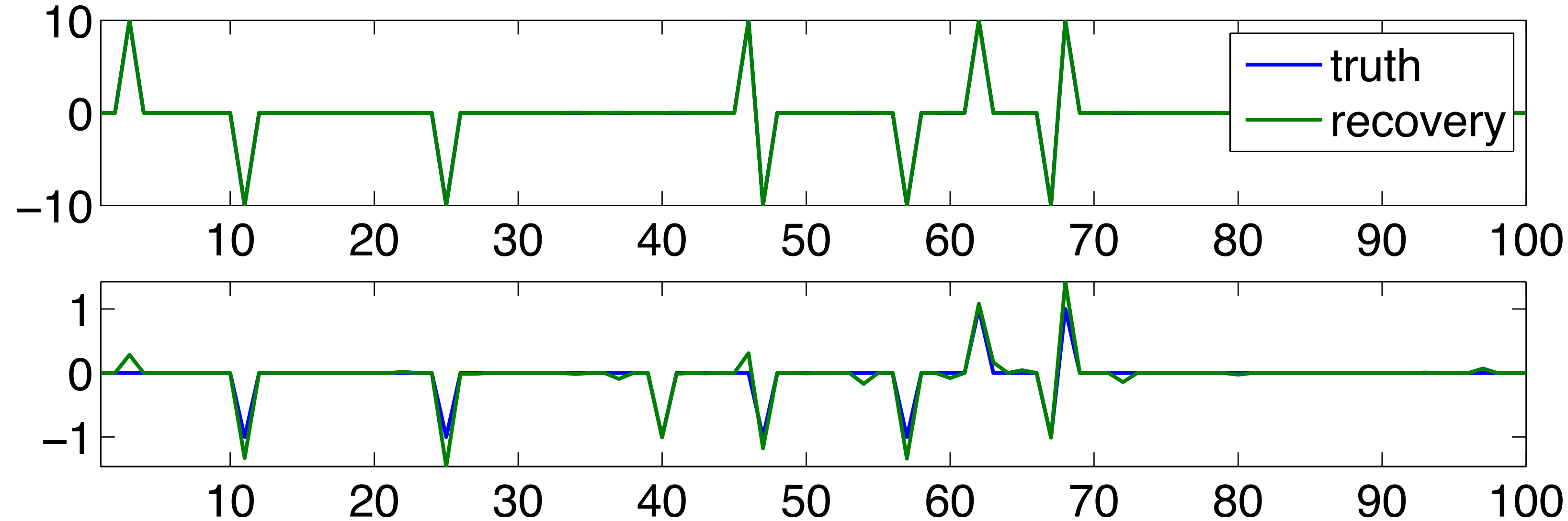
Hessian:

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \\ \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} & \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \end{bmatrix}$$

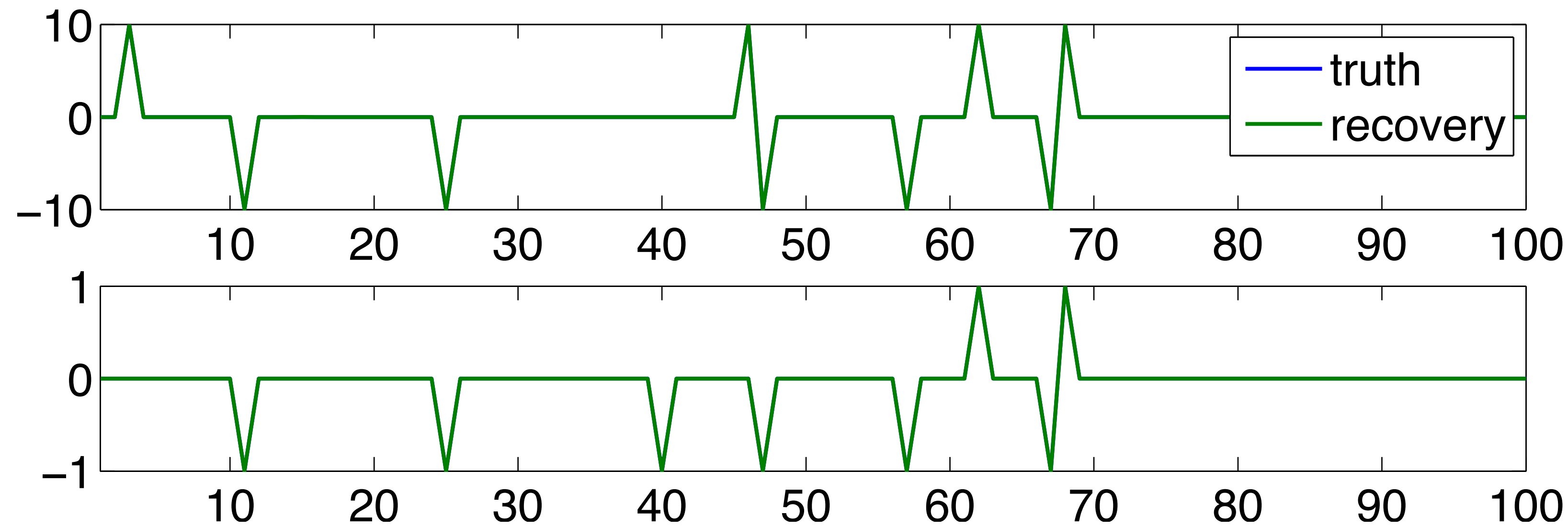


Improved "*incoherence*"

$$p = 1$$



$$p = 1, 2$$

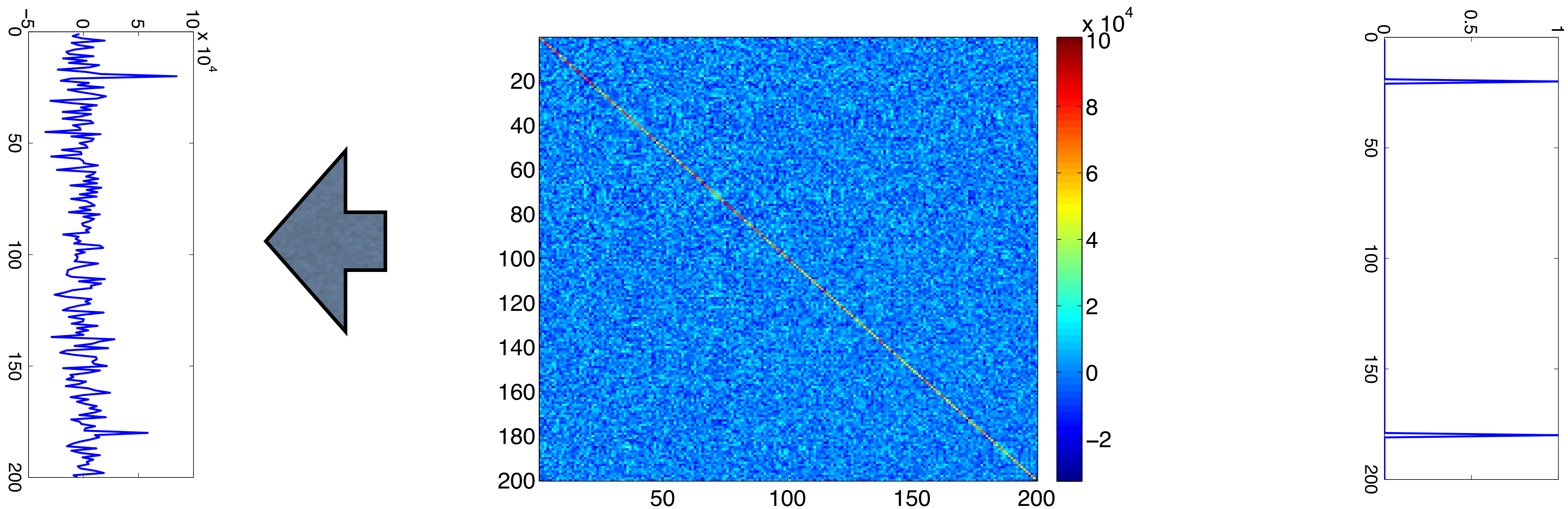


Improved "*incoherence*"

$$\min \left\| \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \right\|_p \quad \text{s.t.} \quad \left\| \delta\mathbf{D} - \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} & \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \end{bmatrix} \begin{bmatrix} \delta\mathbf{m} \\ \frac{1}{\beta}(\frac{1}{\alpha}\delta\boldsymbol{\rho} + \delta\mathbf{m}) \end{bmatrix} \right\|_2 \leq \sigma$$

Hessian:

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \\ \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}} - \alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} & \beta\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} \end{bmatrix}$$

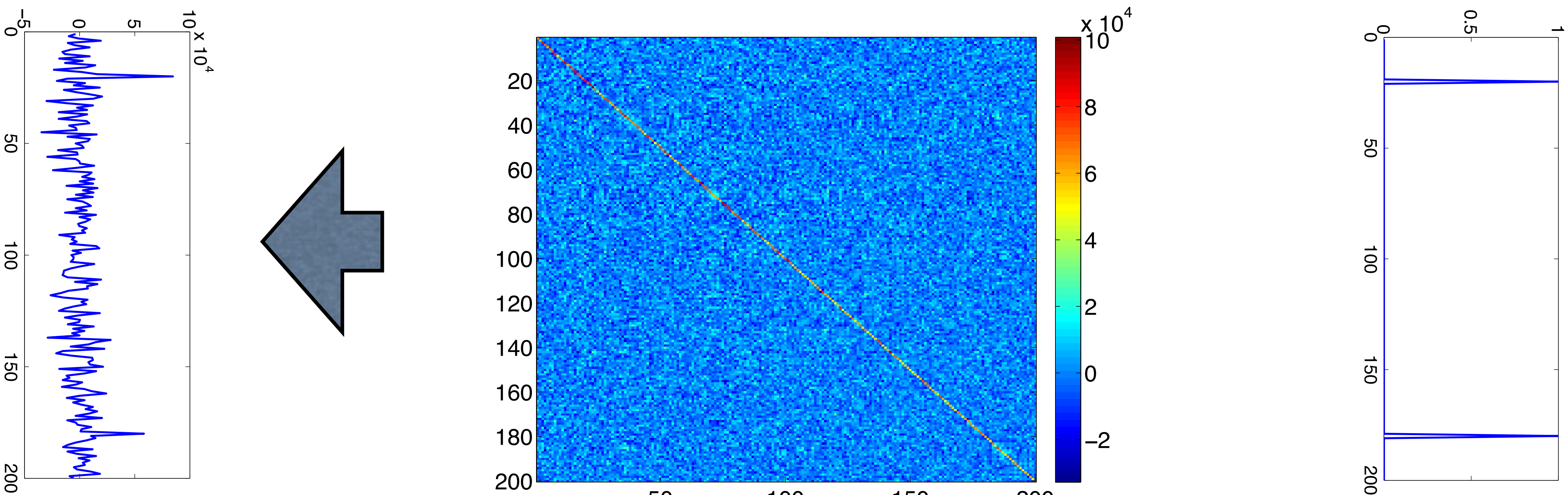


Improved "*incoherence*"

$$\min \left\| \begin{bmatrix} \delta\mathbf{m} + \frac{1}{\alpha}\delta\boldsymbol{\rho} \\ \frac{1}{\beta}\frac{1}{\alpha}\delta\boldsymbol{\rho} \end{bmatrix} \right\|_p \quad \text{s.t.} \quad \left\| \delta\mathbf{D} - \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}} & \beta(\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} - \boldsymbol{\mathcal{J}}_{\mathbf{m}}) \end{bmatrix} \begin{bmatrix} \delta\mathbf{m} + \frac{1}{\alpha}\delta\boldsymbol{\rho} \\ \frac{1}{\beta}\frac{1}{\alpha}\delta\boldsymbol{\rho} \end{bmatrix} \right\|_2 \leq \sigma$$

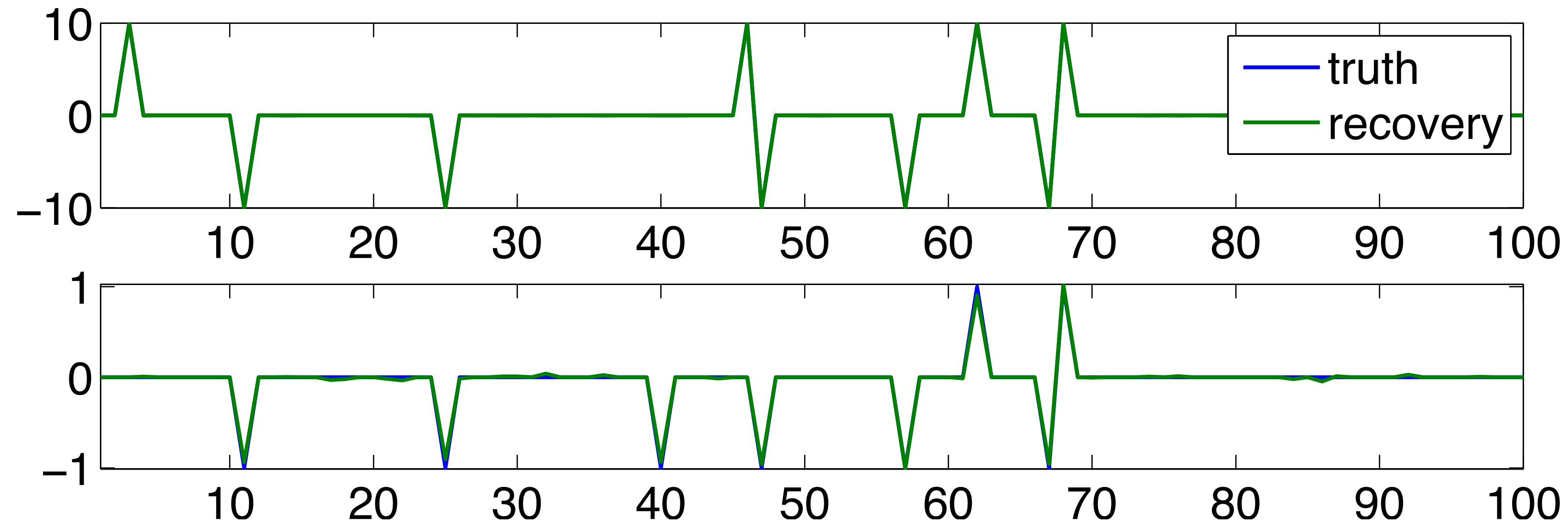
Hessian:

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T \\ \beta(\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}}^T - \boldsymbol{\mathcal{J}}_{\mathbf{m}}^T) \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{J}}_{\mathbf{m}} & \beta(\alpha\boldsymbol{\mathcal{J}}_{\boldsymbol{\rho}} - \boldsymbol{\mathcal{J}}_{\mathbf{m}}) \end{bmatrix}$$

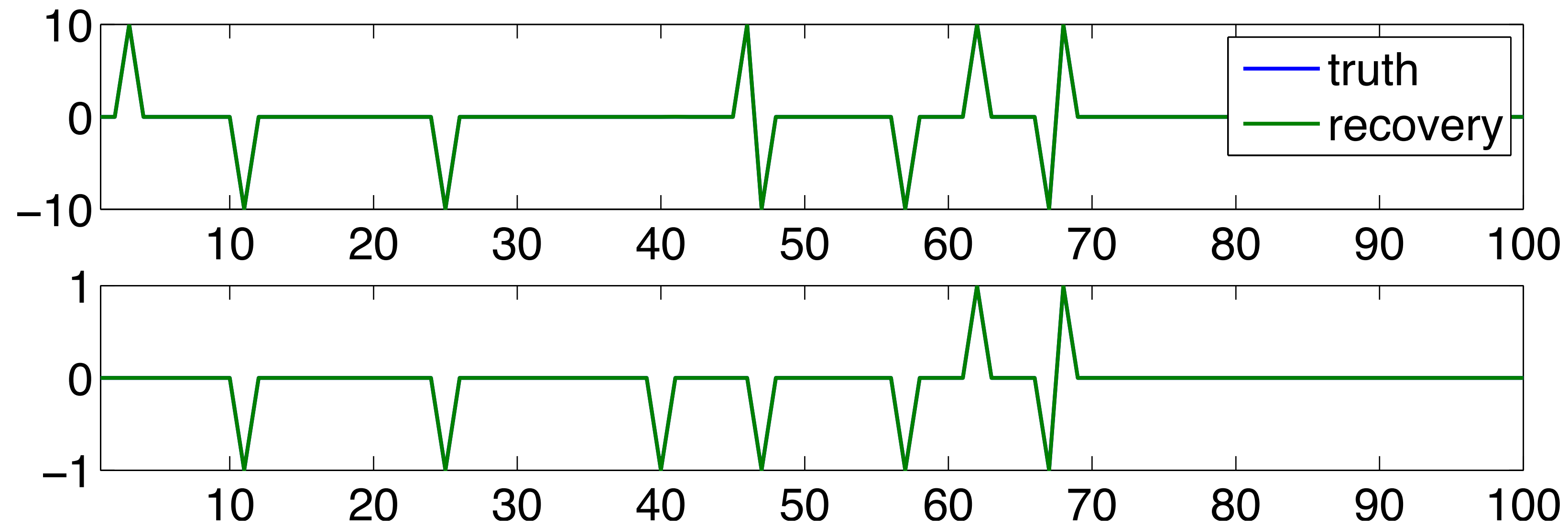


Improved "*incoherence*"

$$p = 1$$



$$p = 1, 2$$

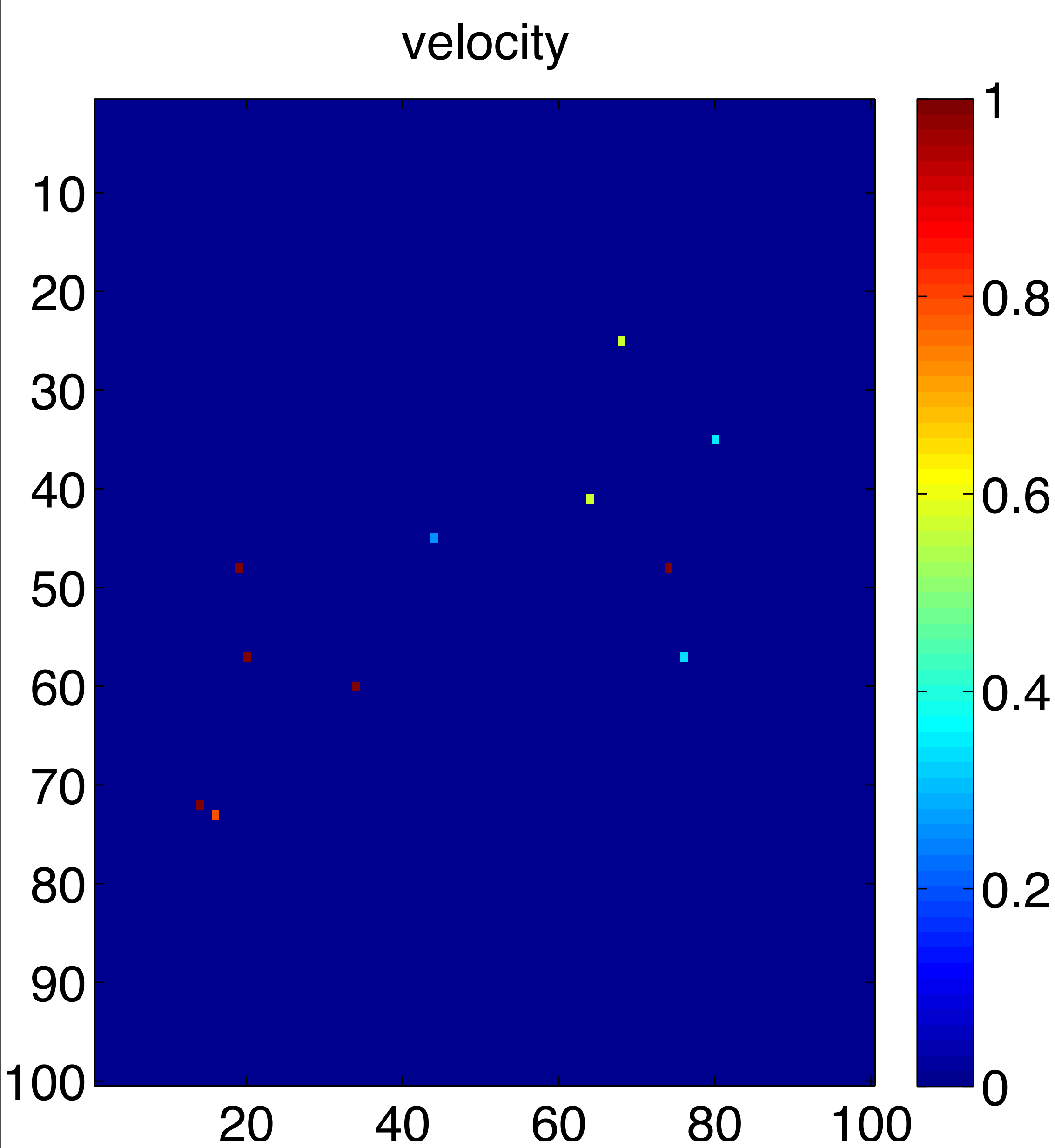


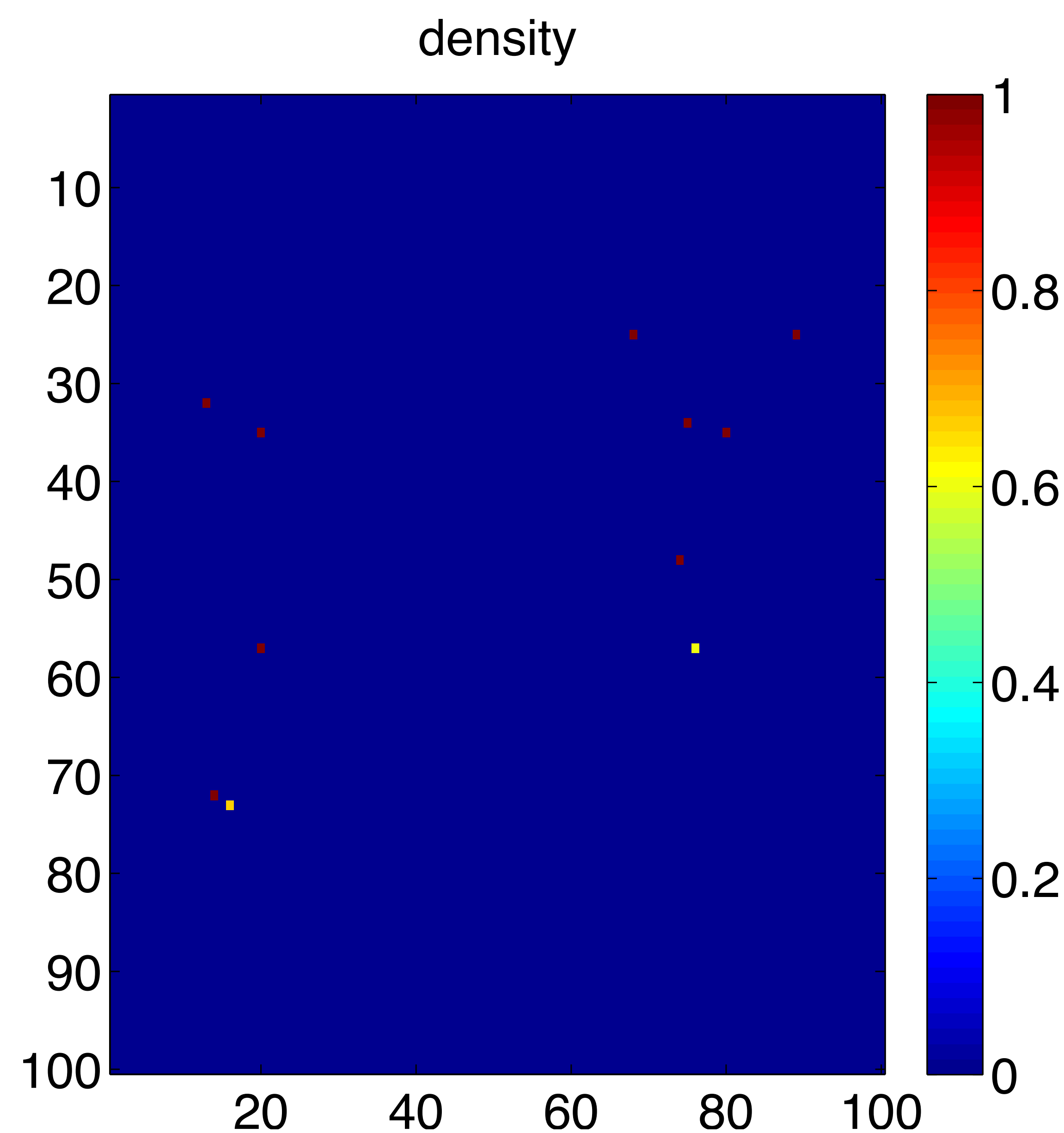
Examples

In this experiment, we use

- wave-equation based Jacobians
- constant background model (100x100)
- linear observed data
- 90 shots and 90 receivers with 10m interval

True perturbation

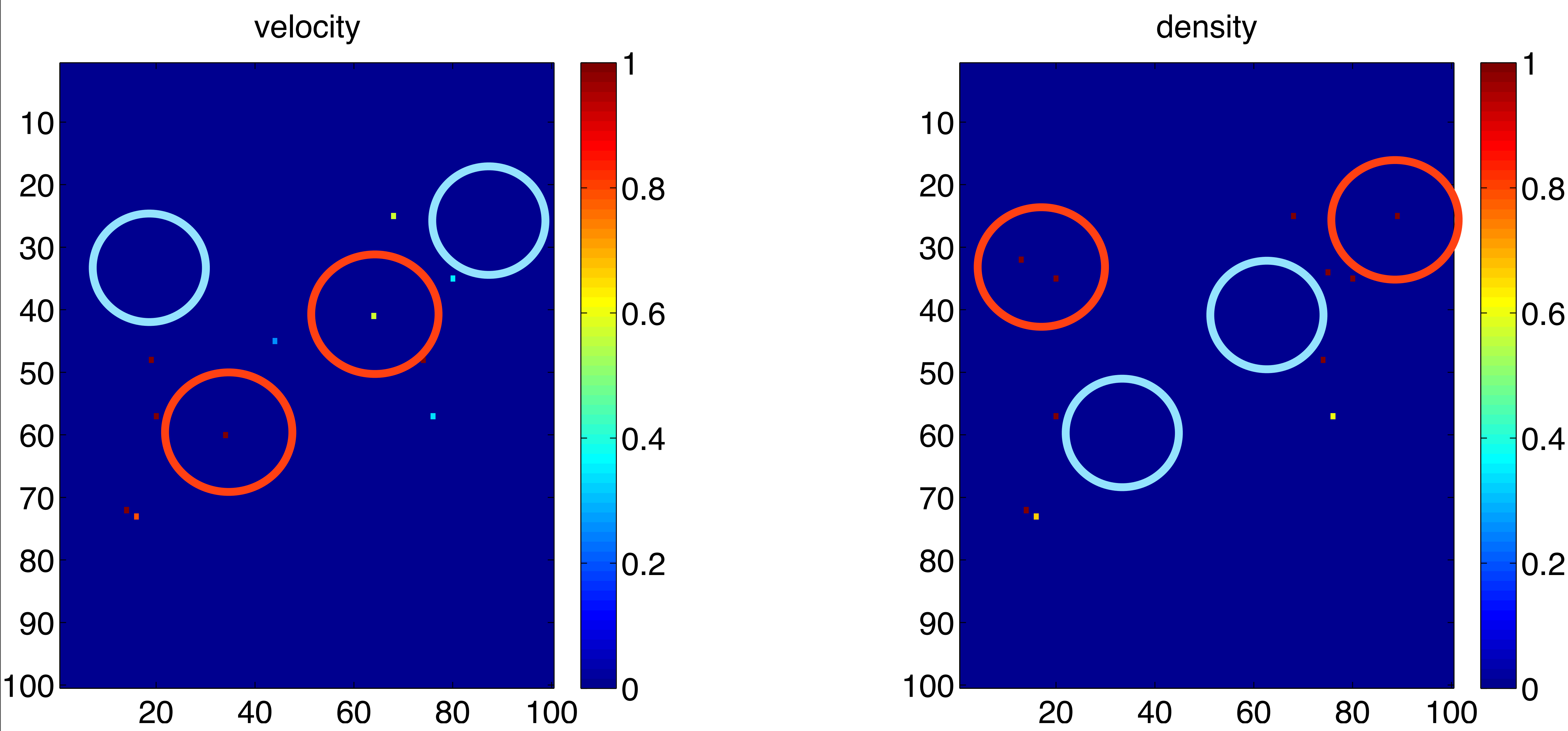




True perturbation

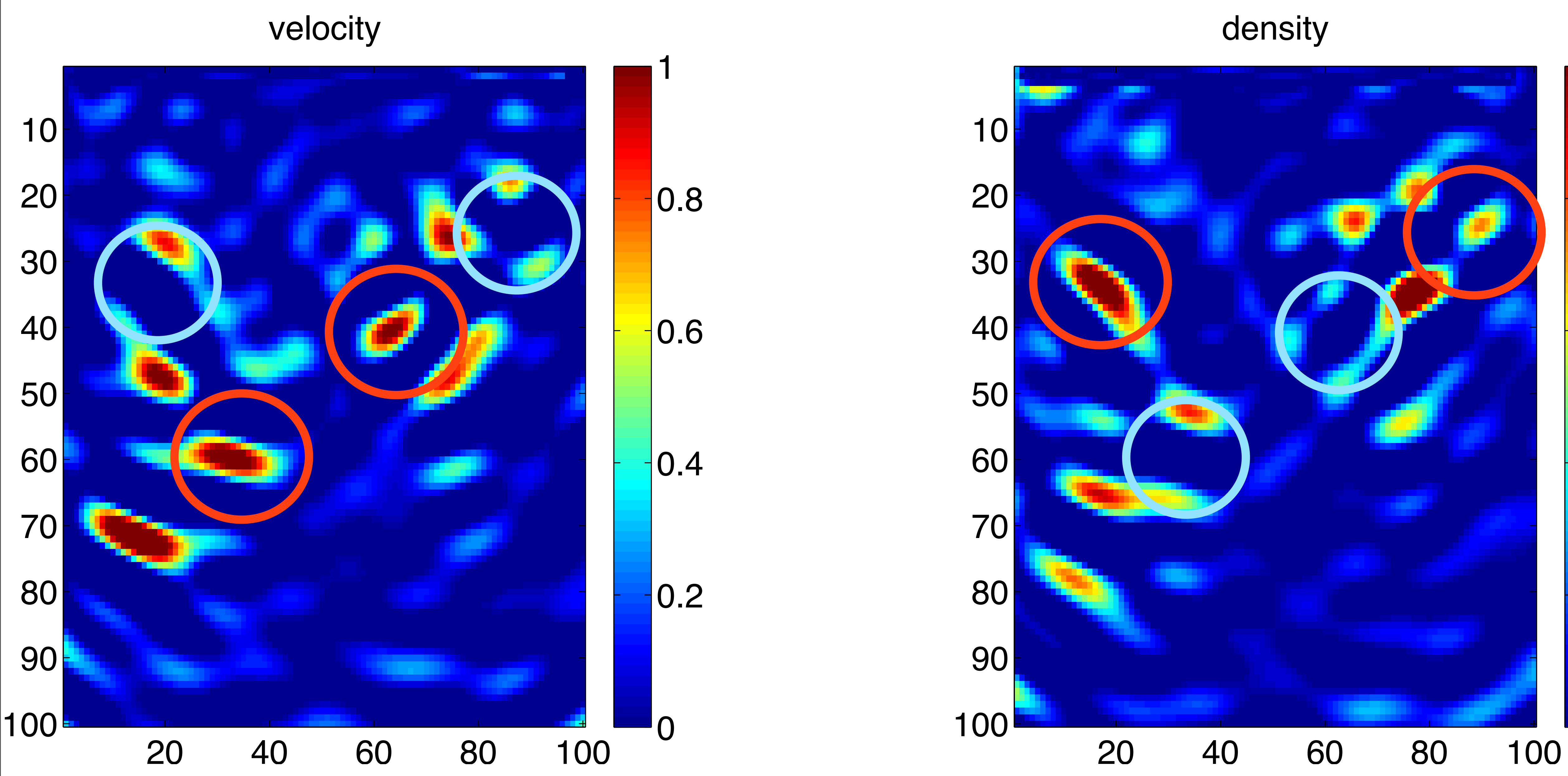
: with spot in it

: empty



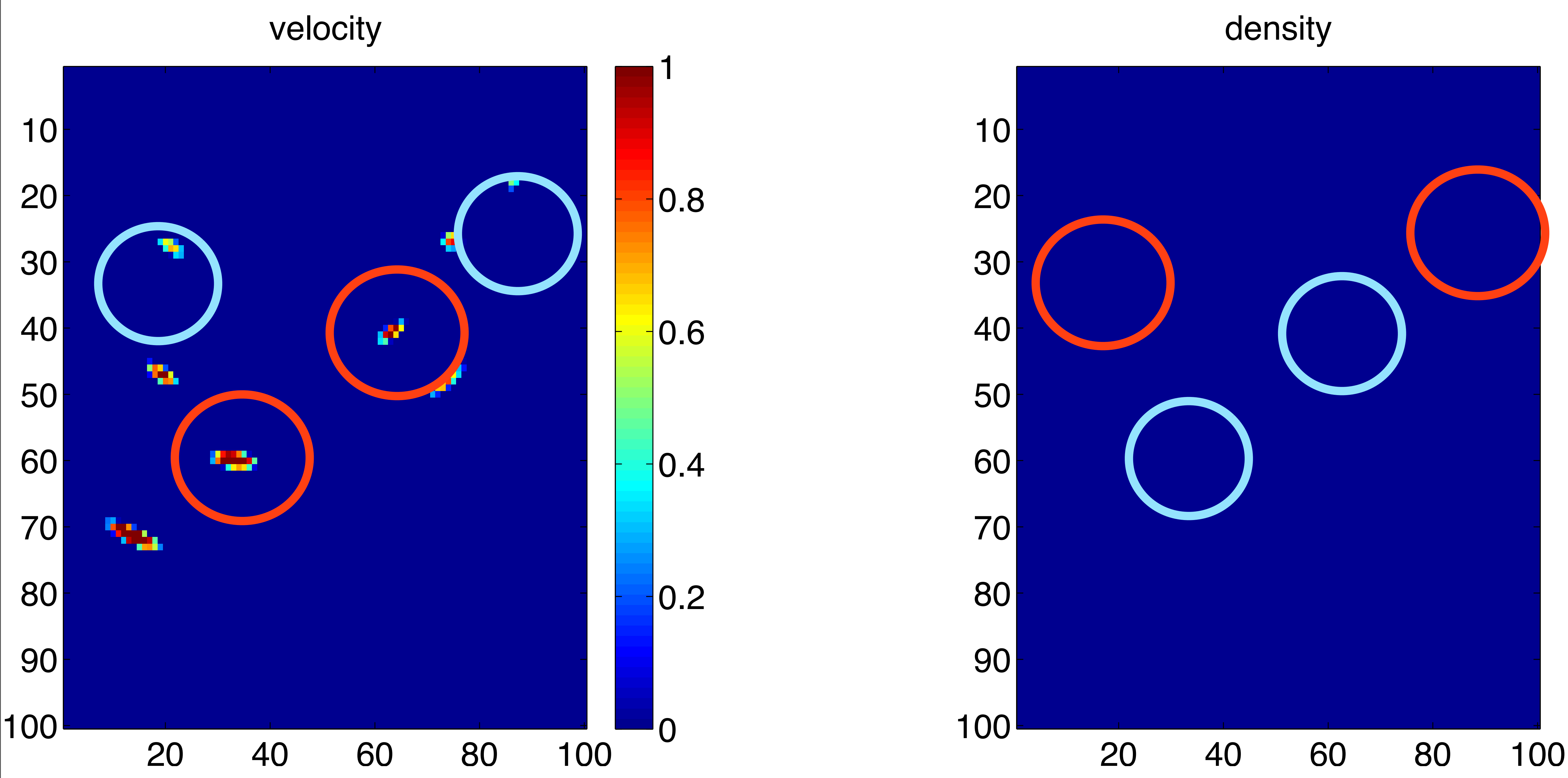
Inverted results

Conventional method without regularization (L2)



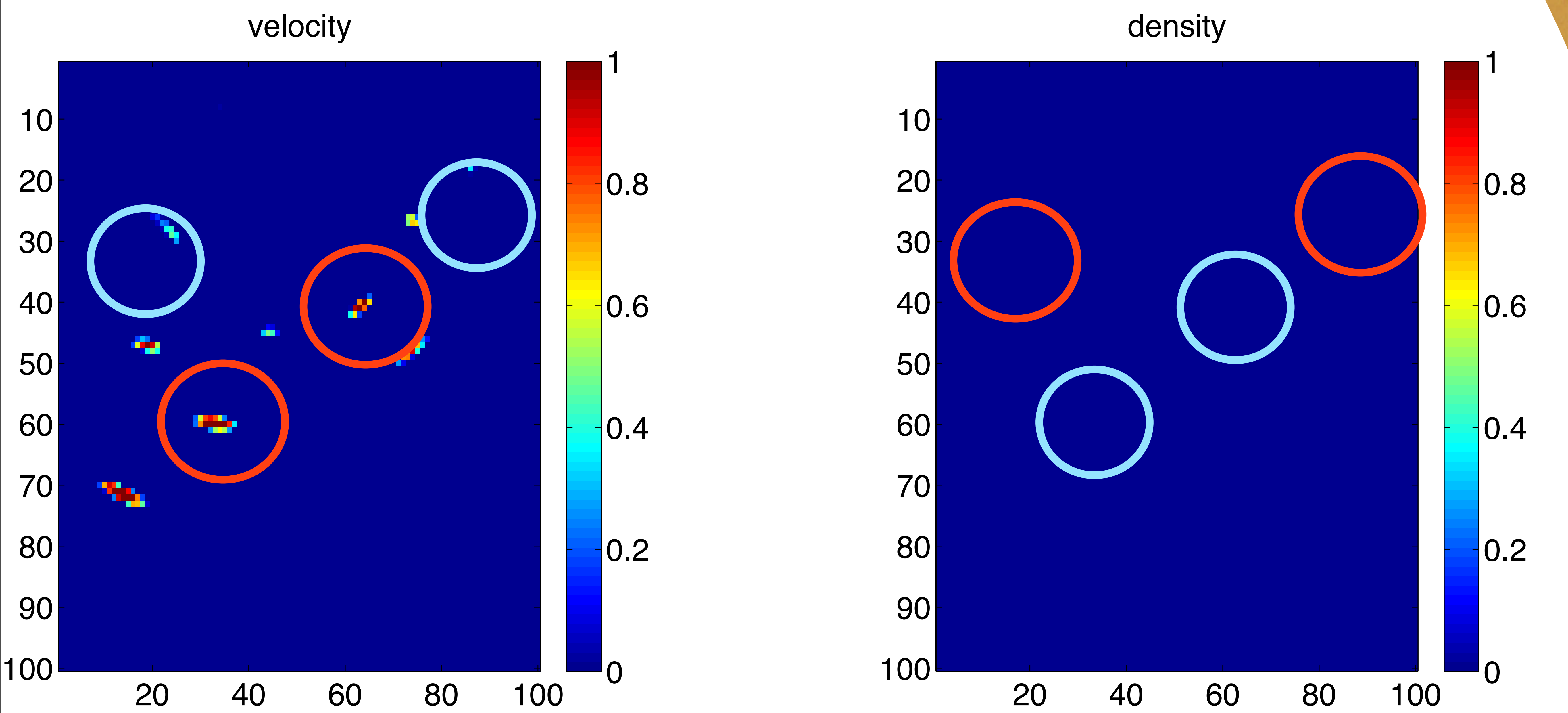
Inverted results

Conventional method with *sparsity* regularization (L1)



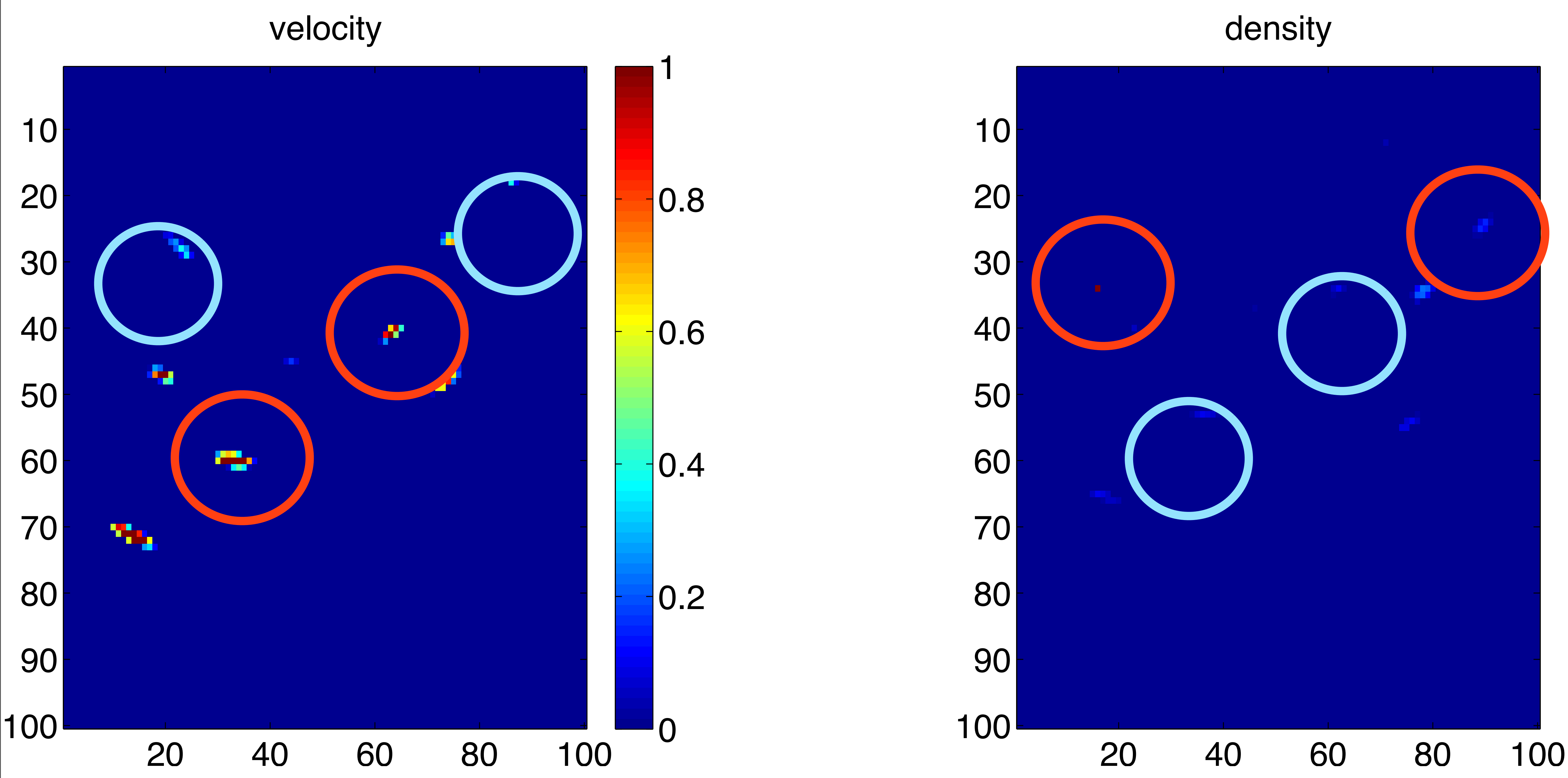
Inverted results

Scaled method with *sparsity* regularization



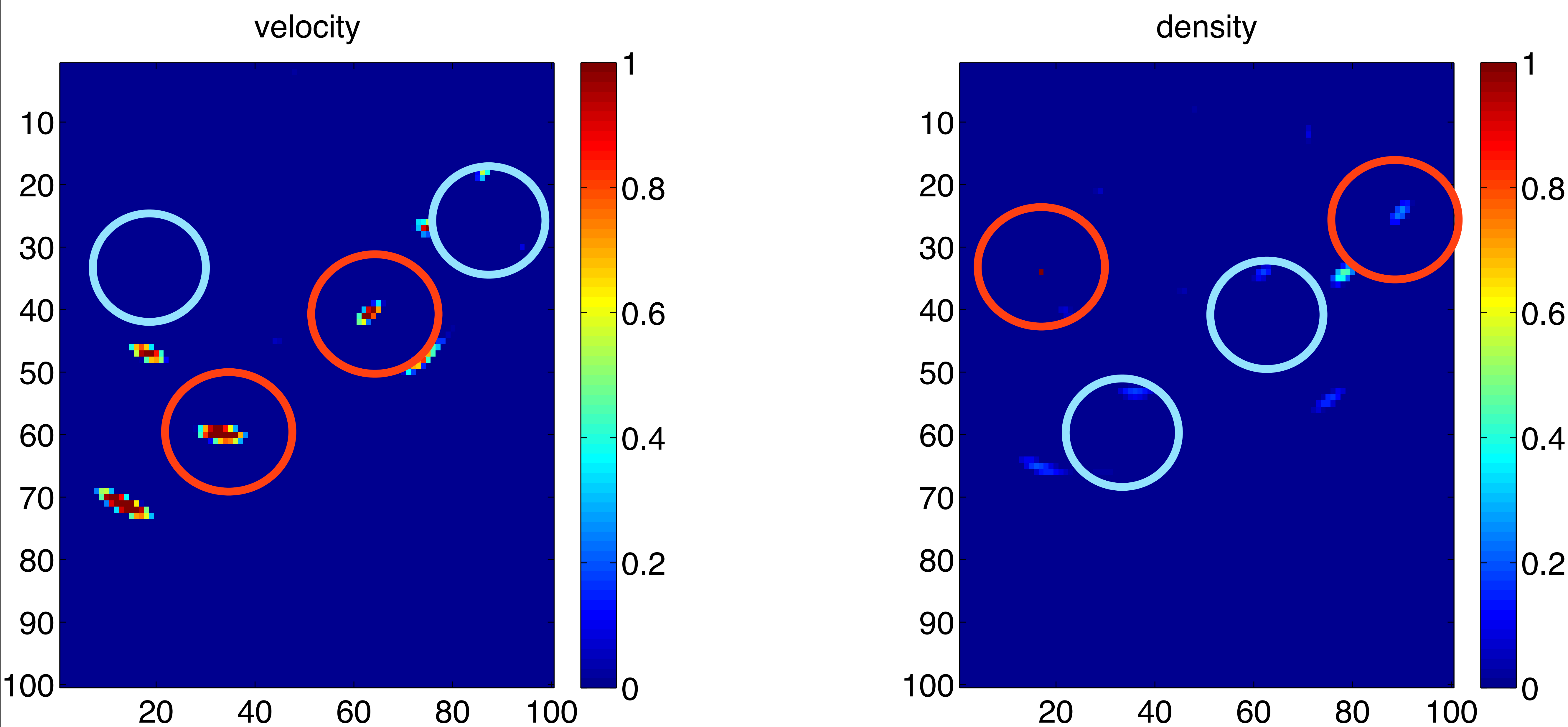
Inverted results

Scaled method with *joint-sparsity* regularization



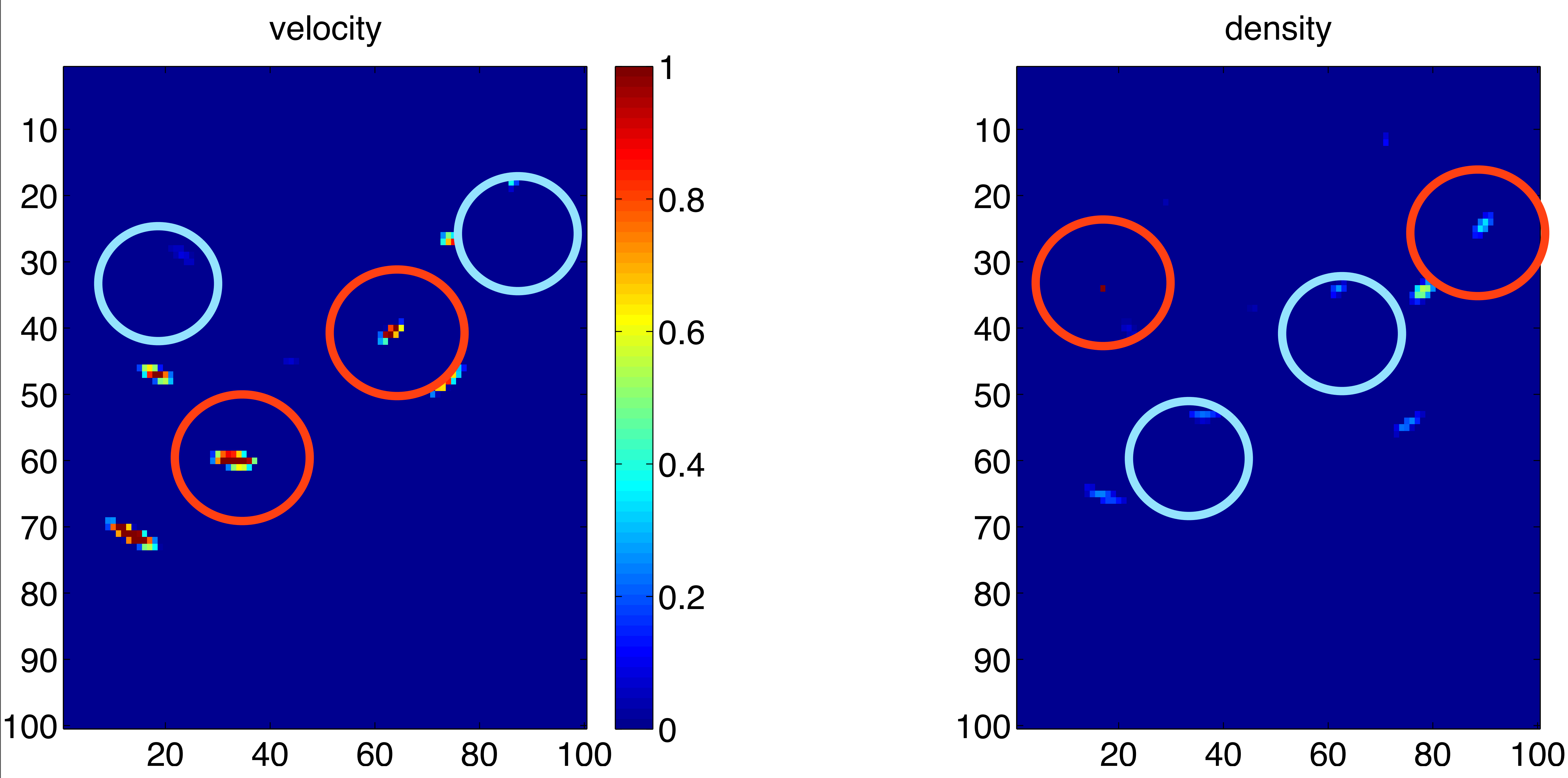
Inverted results

"*Incoherence*" enhanced method with *joint-sparsity* regularization



Inverted results

"*Incoherence*" enhanced method with *sparsity* regularization



Conclusion

- sparsity regularization can efficiently suppress artifacts
- simultaneously invert density and velocity is possible by diagonalizing Hessian
- joint sparsity can help improve the resolution when density and velocity share similar sparsity support

Acknowledgements

- All my colleagues

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

Thank you

https://www.slim.eos.ubc.ca