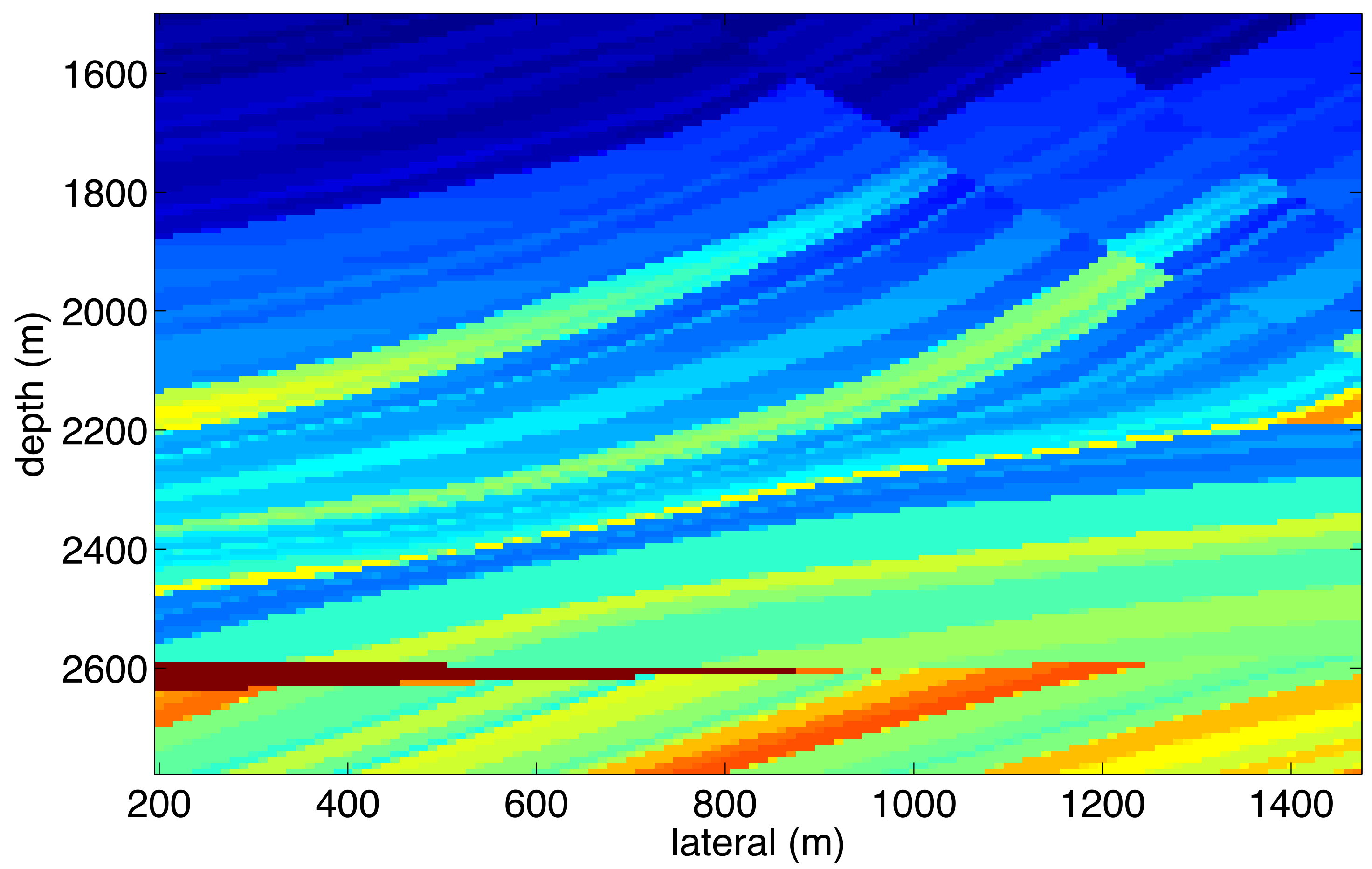
Why the modified Gauss-Newton method?

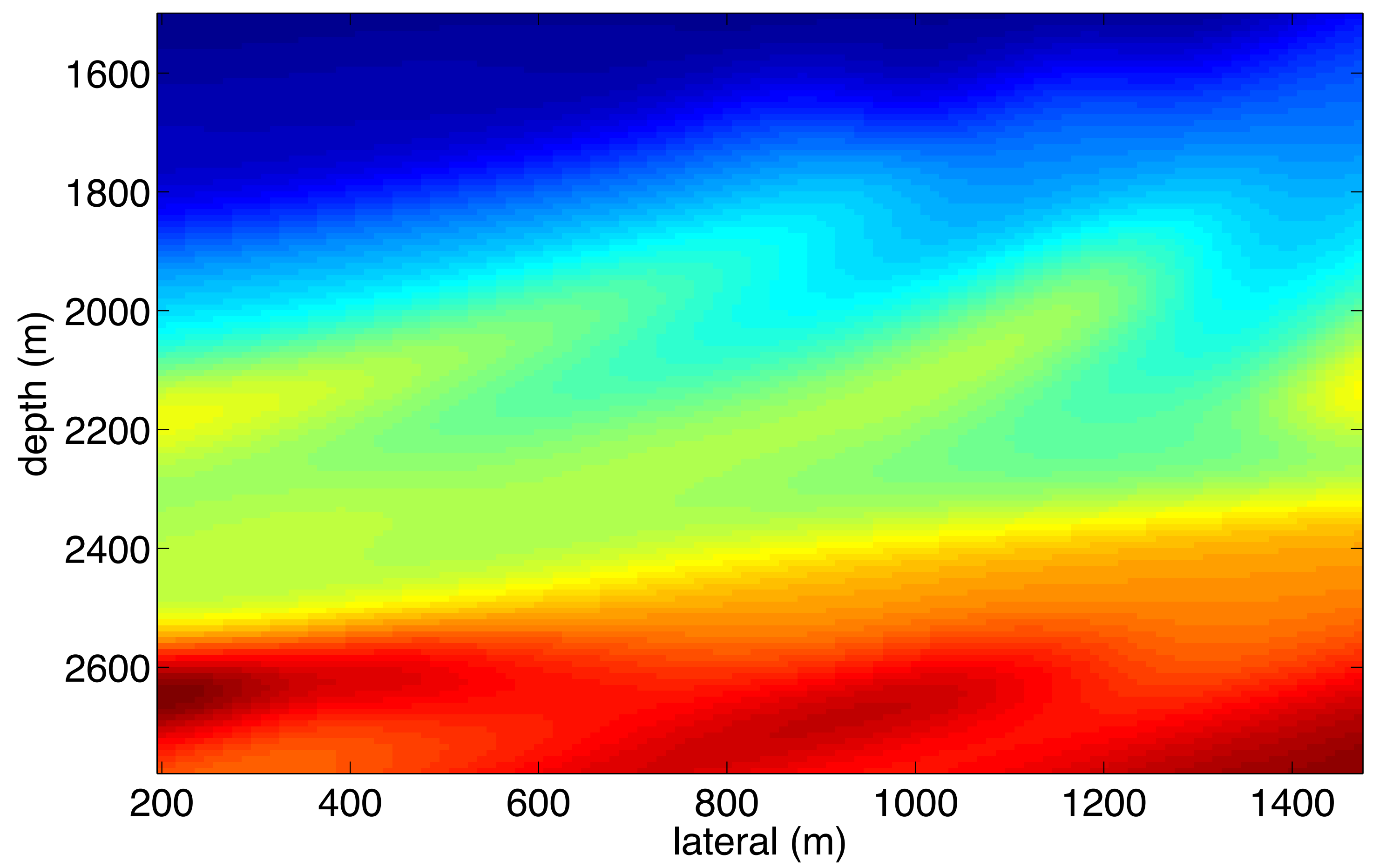
Xiang Li, Ernie Esser and Felix Herrmann

SLIM
University of British Columbia

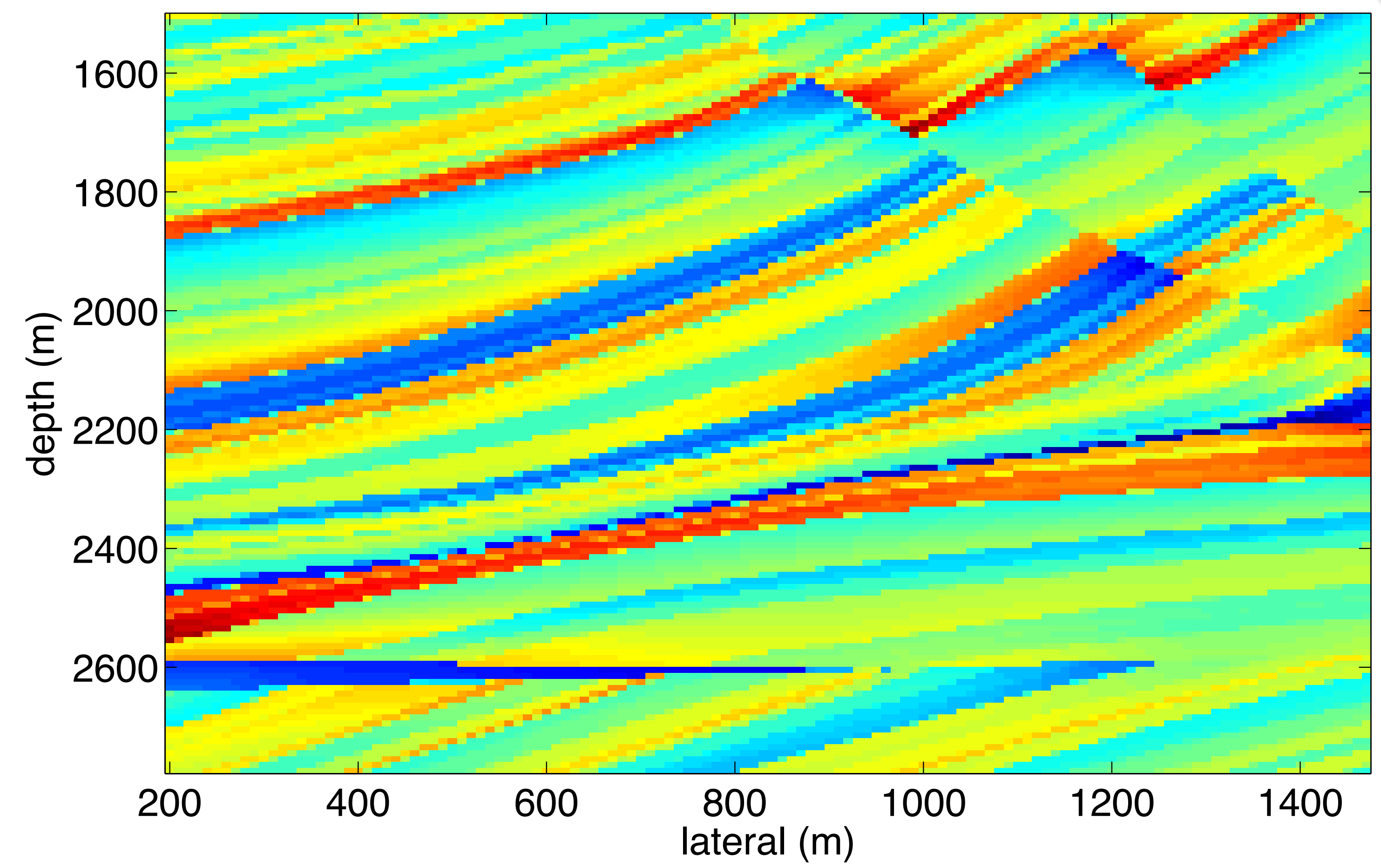
Motivation



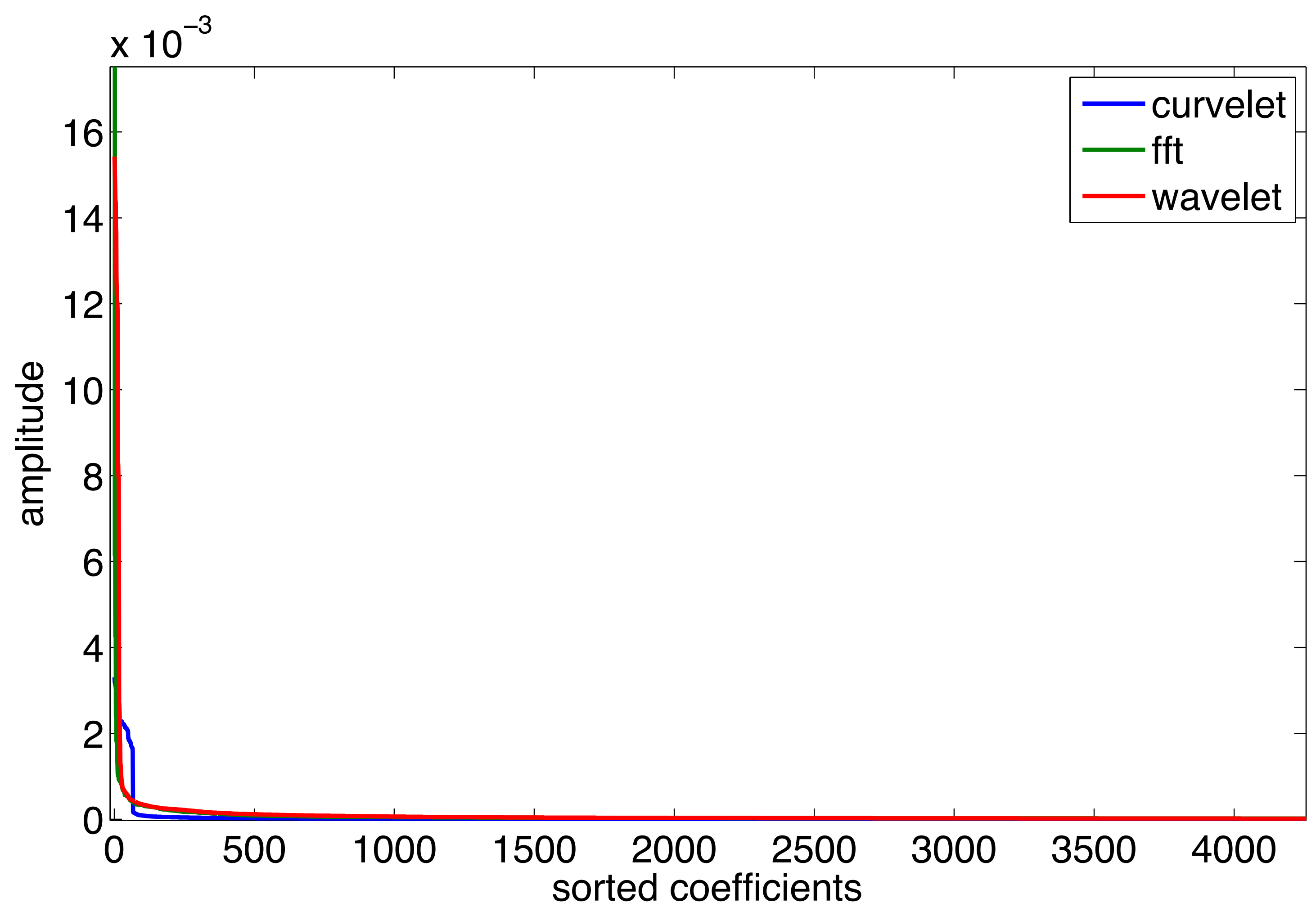
true model



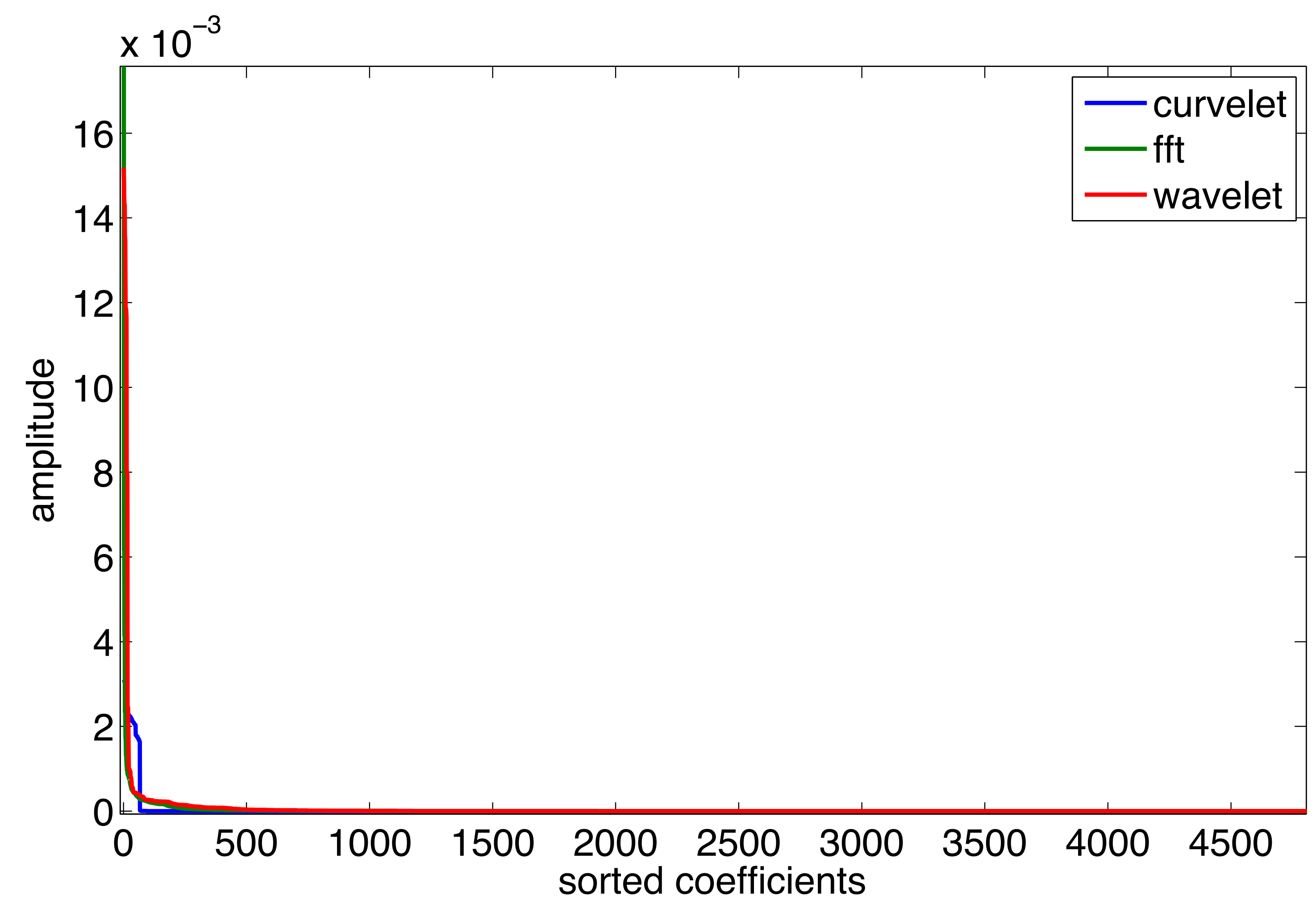
initial model



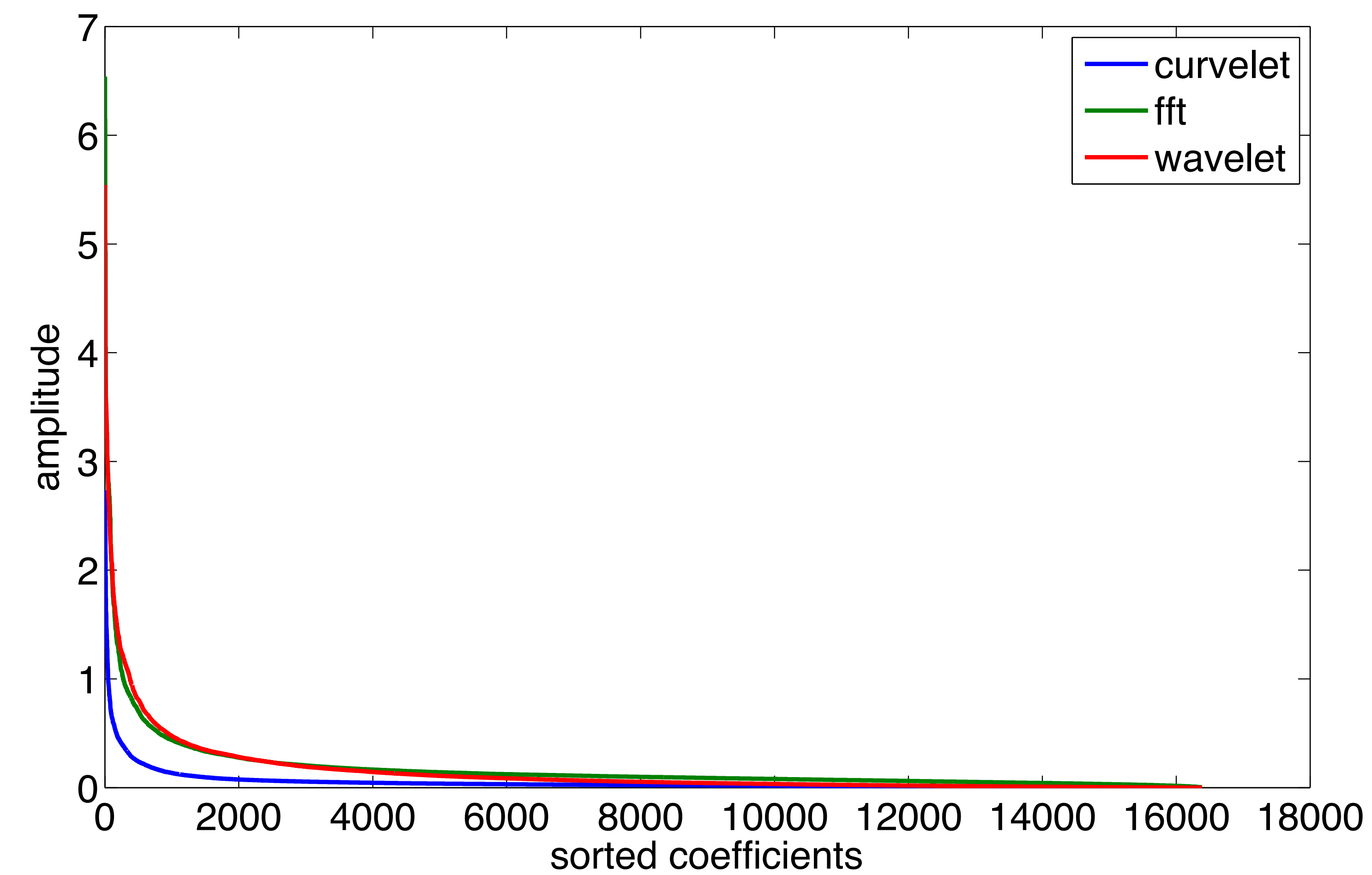
perturbation



transform domain coefficients



transform domain coefficients



transform domain coefficients

What is the modified Gauss-Newton?

FWI objective

$$\Phi(\mathbf{m}) := \frac{1}{2}\|\underbrace{\mathbf{D} - \mathbf{P}_r\mathbf{H}(\mathbf{m})^{-1}\mathbf{Q}}_{\delta\mathbf{D}}\|_F^2$$

gradient (action of Jacobian adjoint or RTM operator)

$$\nabla\Phi(\mathbf{m}) := \mathbf{J}^{\mathrm{T}}\delta\mathbf{D} = \mathrm{conj}\left(\frac{\partial\mathbf{H}}{\partial\mathbf{m}}(\mathbf{H}^{-1}\mathbf{Q})\right) \odot \left((\mathbf{H}^{\mathrm{T}})^{-1}(\mathbf{P}_r^{\mathrm{T}}\delta\mathbf{D})\right)$$

Gauss-Newton update

$$\delta\mathbf{m} := \arg\min_{\delta\mathbf{m}} \|\delta\mathbf{D} - \mathbf{J}\delta\mathbf{m}\|_F^2$$

Gauss-Newton analytic solution: $\delta\mathbf{m} = (\underbrace{\mathbf{J}^{\mathrm{T}}\mathbf{J}}_{\text{Gauss-Newton Hessian of FWI objective}})^{-1}\underbrace{\mathbf{J}^{\mathrm{T}}\delta\mathbf{D}}_{\text{RTM}}$

the modified Gauss-Newton update

$$\delta\mathbf{m} := \mathbf{S}^{\mathrm{T}}\arg\min_{\mathbf{x}} \|\delta\mathbf{D} - \mathbf{J}\mathbf{S}^{\mathrm{T}}\mathbf{x}\|_F^2 \quad \text{s.t.} \quad \|\mathbf{x}\|_1 \leq \tau$$

The modified Gauss-Newton algorithm

Algorithm 1: the modified Gauss-Newton method for FWI

Result: Output estimate for the model $\mathbf{m}$

$\mathbf{m} \longleftarrow \mathbf{m}_0;\ k \longleftarrow 0; \epsilon$; `// initial model`

while $\frac{1}{2}\|\mathbf{D} - \mathbf{P}_r\mathbf{H}(\mathbf{m})^{-1}\mathbf{Q}\|_F^2 \geq \epsilon$ **do**

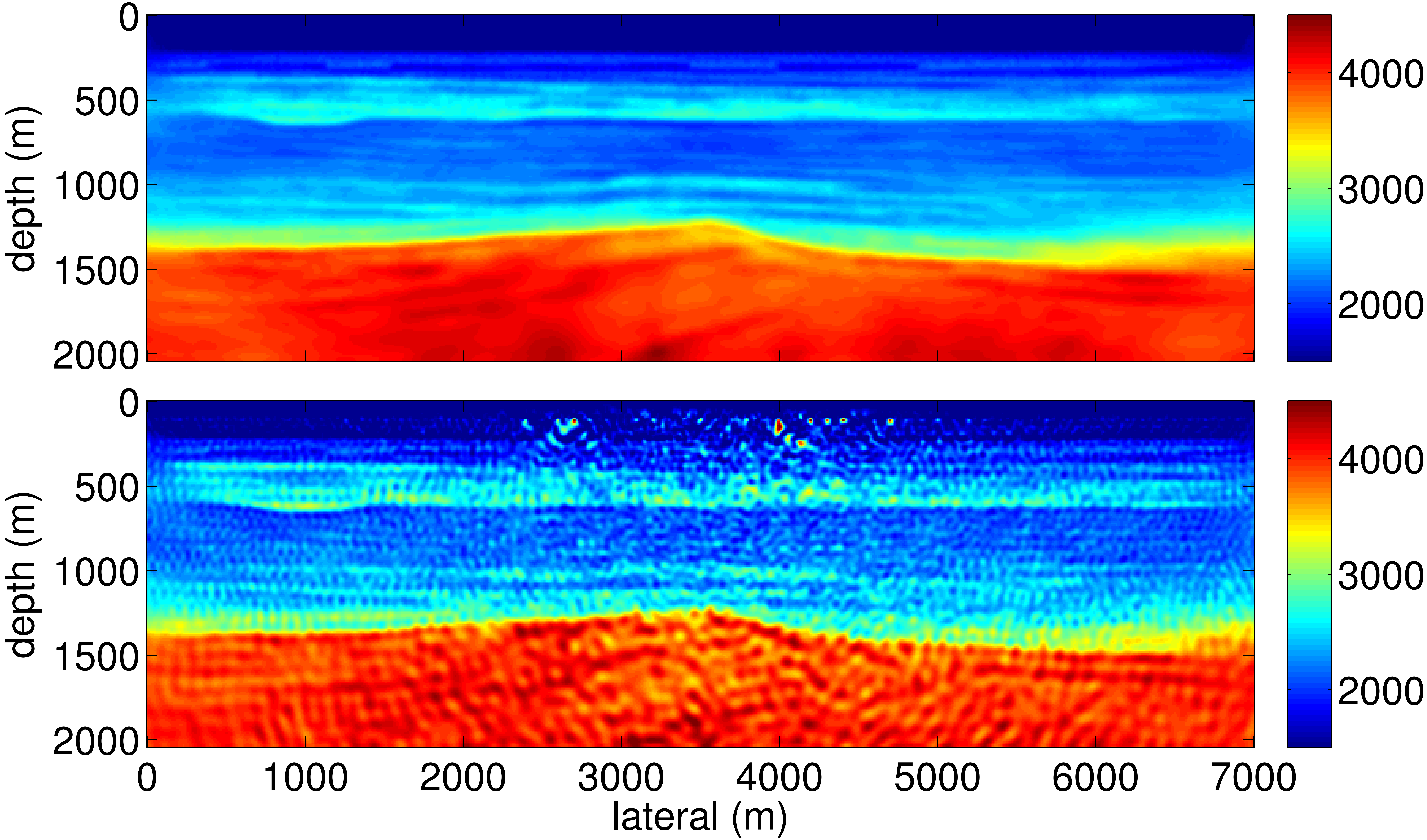
$\quad \delta\mathbf{m}^k := \mathbf{S}^{\mathrm{T}} \arg\min_{\mathbf{x}} \|\delta\mathbf{D} - \mathbf{J}\mathbf{S}^{\mathrm{T}}\mathbf{x}\|_F^2 \quad \text{s.t.} \quad \|\mathbf{x}\|_1 \leq \tau^k$

$\quad \mathbf{m}^{k+1} \longleftarrow \mathbf{m}^k + \alpha^k\delta\mathbf{m}^k$; `// update with linesearch`

$\quad k \longleftarrow k+1;$

end

Inversion results



the modified Gauss-Newton method with L1 constraint

the modified Gauss-Newton method with L2 constraint

Observation & question

- modified Gauss-Newton only promotes sparsity on individual updates
- modified Gauss-Newton does NOT change the FWI objective function

why the *sum* of all modified Gauss-Newton updates is sparse?

SLIM

Least-squares optimization problem

unconstrained objective function

$$\min_{\mathbf{m}} \Phi(\mathbf{m}) := \left\{ \frac{1}{2} \|\mathbf{d} - \boldsymbol{\mathcal{F}}[\mathbf{m}]\|_2^2 \right\}$$

Gauss-Newton update

$$\delta\mathbf{m} = \arg\min_{\delta\mathbf{m}} \|\delta\mathbf{d} - \mathbf{J}[\mathbf{m}_k]\delta\mathbf{m}\|_2$$

the modified Gauss-Newton update

$$\delta\mathbf{m} = \mathbf{S}^{\mathrm{H}} \arg\min_{\delta\mathbf{x}} \|\delta\mathbf{d} - \mathbf{J}[\mathbf{m}_k]\mathbf{S}^{\mathrm{H}}\delta\mathbf{x}\|_2^2 \quad \text{subject to} \quad \|\delta\mathbf{x}\|_{\ell_1} \leq \tau$$

lasso problem

Least-squares optimization with sparse constraint

objective function with sparse constraint

$$\min_{\mathbf{x}} \Phi(\mathbf{x}) := \left\{ \frac{1}{2} \|\mathbf{d} - \boldsymbol{\mathcal{F}}[\mathbf{S}^{\mathrm{H}}\mathbf{x}]\|_2^2 \right\} \quad \text{subject to} \quad \|\mathbf{x} - \mathbf{x}_0\|_{\ell_1} \leq \tau$$

Gauss-Newton update

$$\delta\mathbf{m} = \mathbf{S}^{\mathrm{H}} \arg\min_{\delta\mathbf{x}} \|\delta\mathbf{d} - \nabla\boldsymbol{\mathcal{F}}[\mathbf{m}_k]\mathbf{S}^{\mathrm{H}}\delta\mathbf{x}\|_2^2 \quad \text{subject to} \quad \|\delta\mathbf{x} + \mathbf{x}_k - \mathbf{x}_0\|_{\ell_1} \leq \tau$$

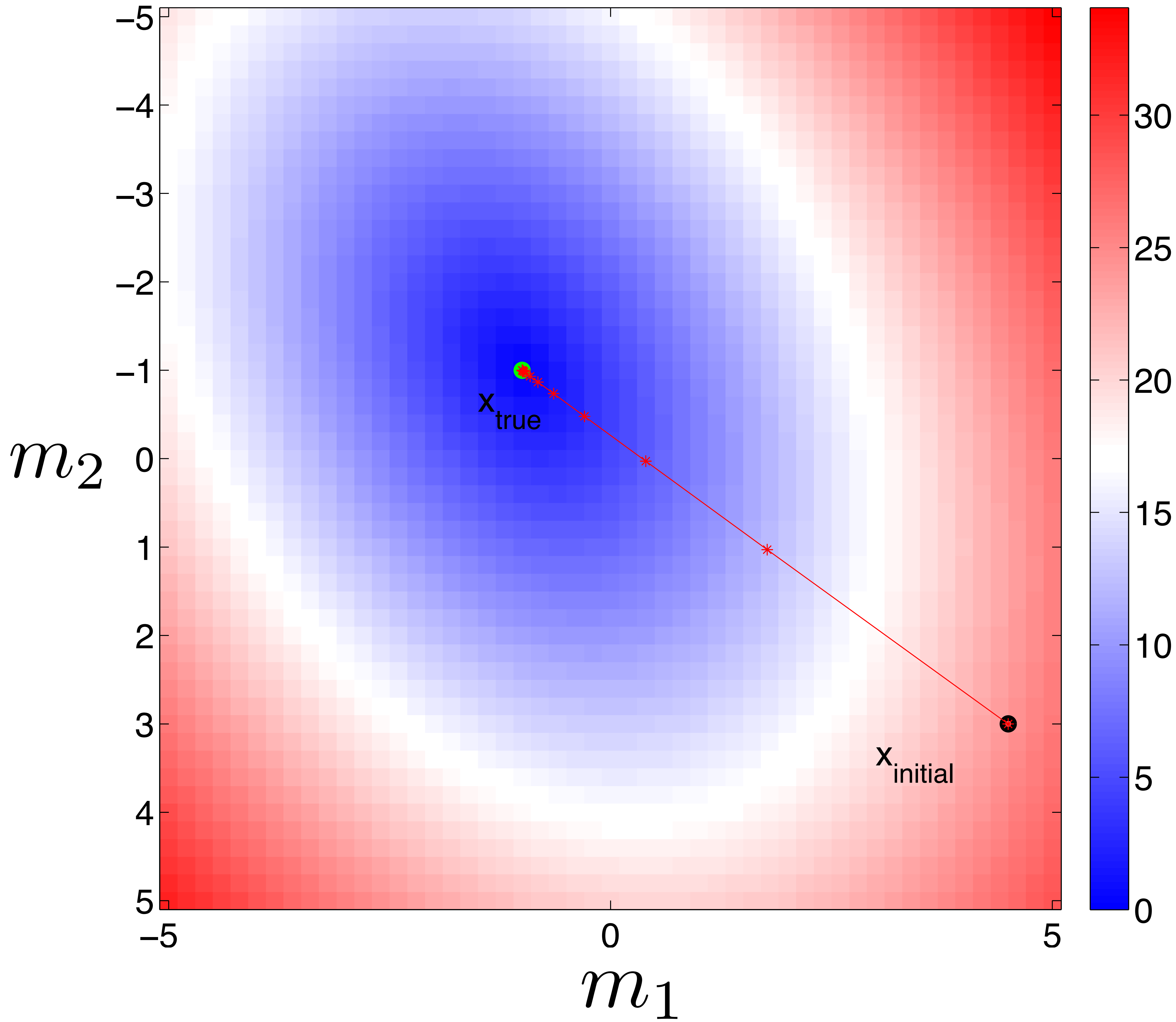
Convex problem with unique solution

$$\Phi(\mathbf{m}) := \left\{ \frac{1}{2} \|\mathbf{d} - \mathbf{A}\mathbf{m}\|_2^2 \right\}$$

$$\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 6 & -3 \end{bmatrix}$$

$$\mathbf{d} = \begin{bmatrix} -6 \\ -3 \end{bmatrix}$$

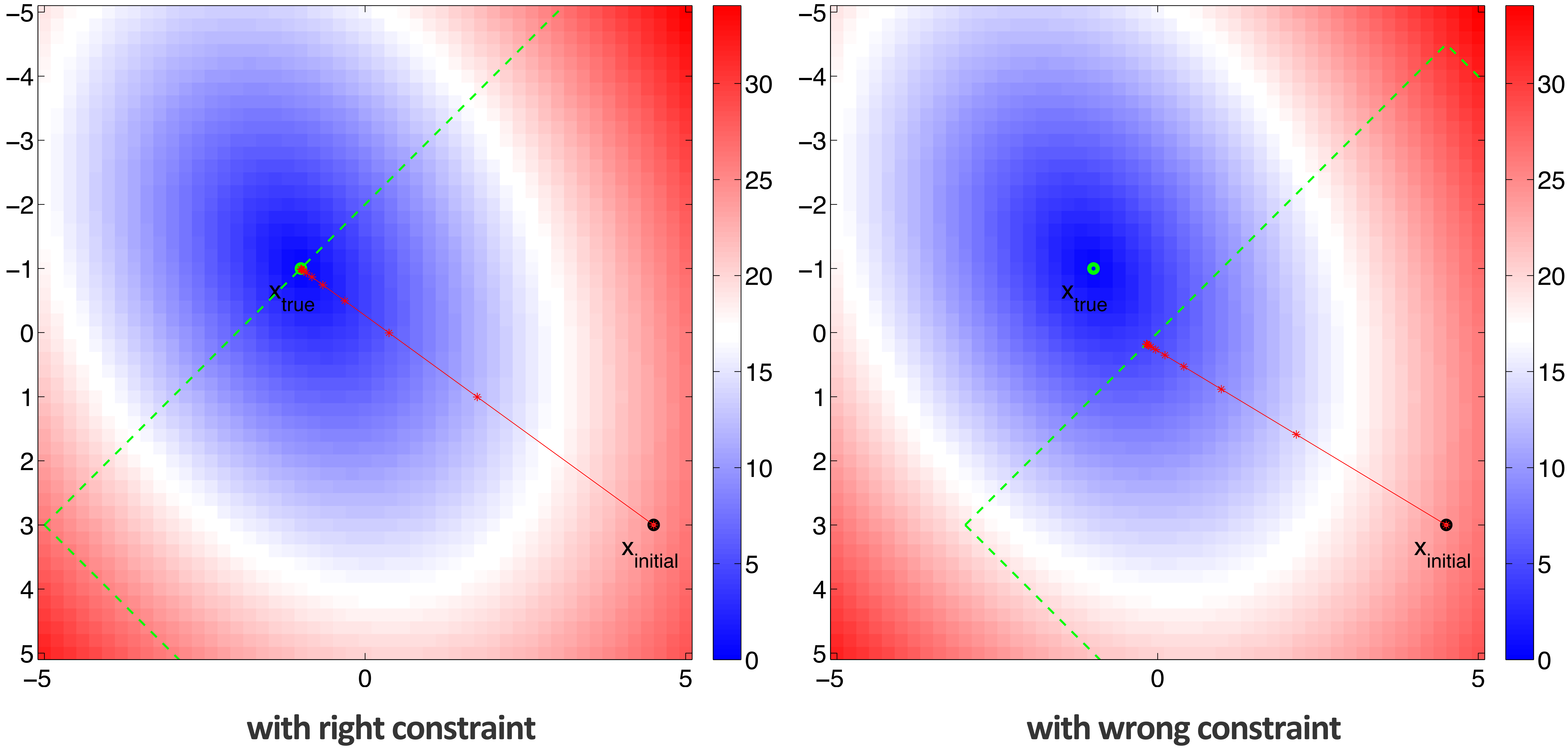
$$\mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}$$



GN with unconstrained objective

GN with sparse constrained objective function

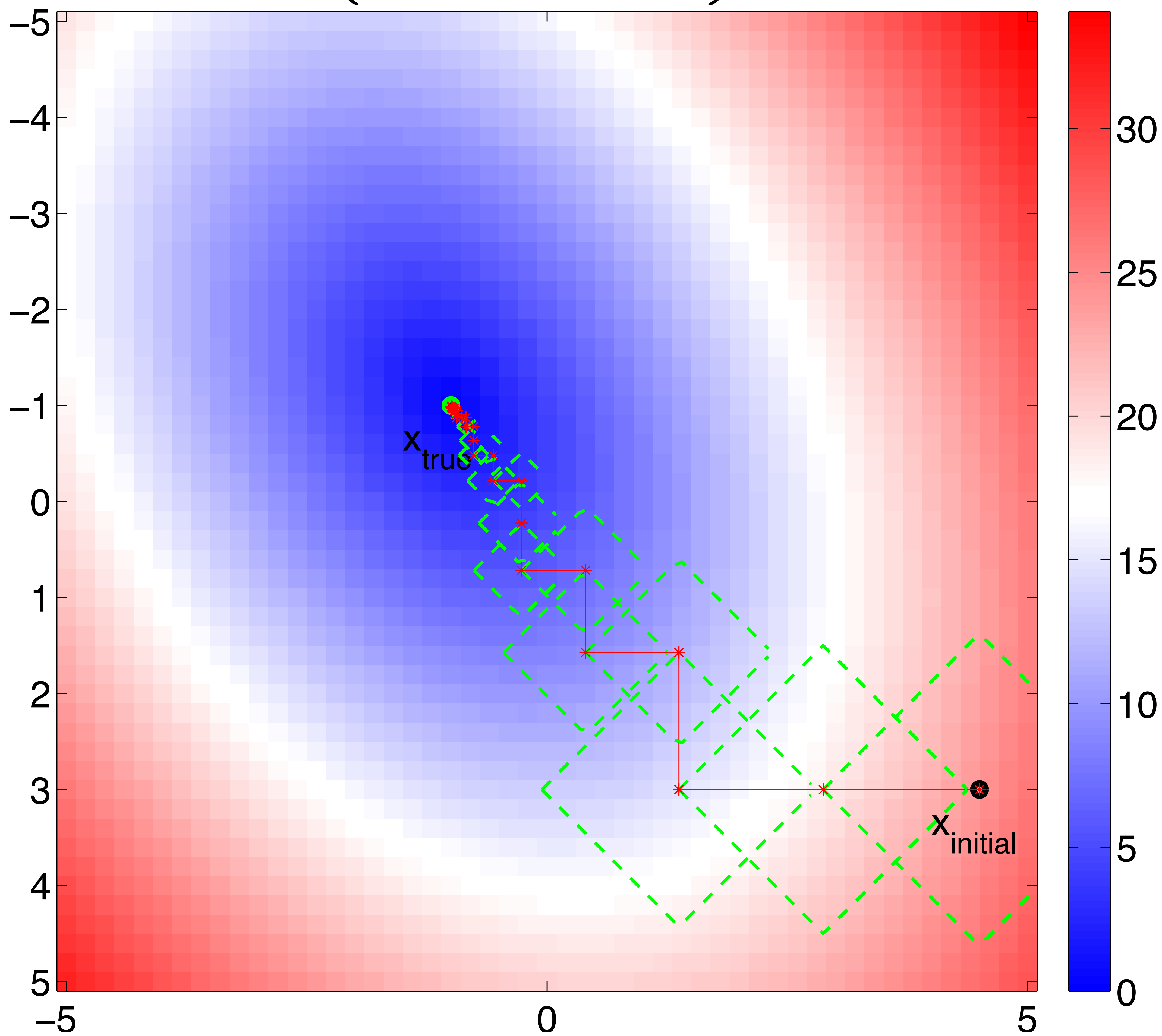
$$\min_{\mathbf{x}} \Phi(\mathbf{x}) := \left\{ \frac{1}{2} \|\mathbf{d} - \boldsymbol{\mathcal{F}}[\mathbf{S}^{\mathrm{H}}\mathbf{x}]\|_2^2 \right\} \quad \text{subject to} \quad \|\mathbf{x} - \mathbf{x}_0\|_{\ell_1} \leq \tau$$



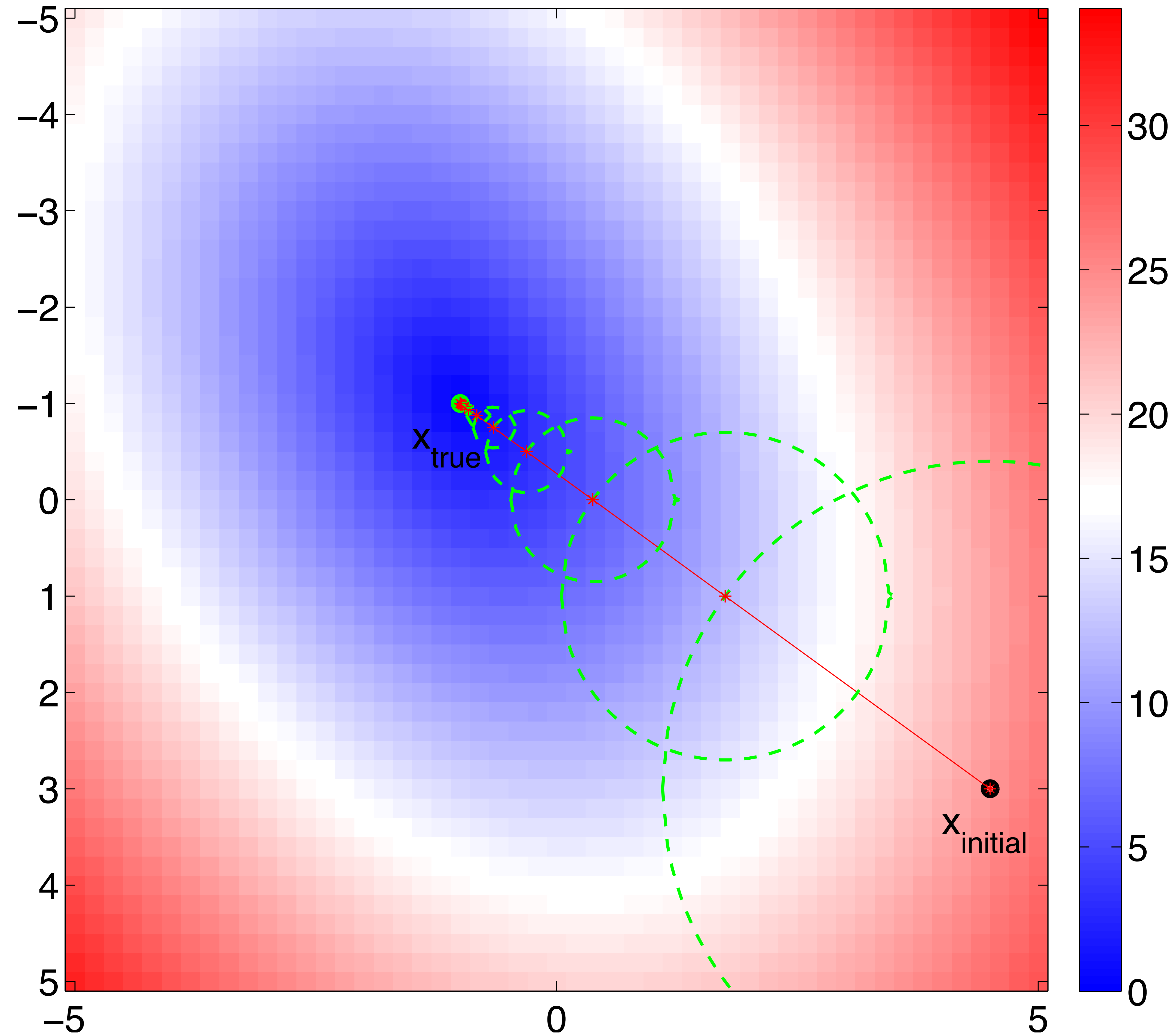
The modified Gauss-Newton

$$\min_{\mathbf{m}} \Phi(\mathbf{m}) := \left\{ \frac{1}{2} \|\mathbf{d} - \boldsymbol{\mathcal{F}}[\mathbf{m}]\|_2^2 \right\}$$

$$\delta\mathbf{m} = \mathbf{S}^{\mathrm{H}} \arg\min_{\delta\mathbf{x}} \|\delta\mathbf{d} - \mathbf{J}[\mathbf{m}_k]\mathbf{S}^{\mathrm{H}}\delta\mathbf{x}\|_2^2 \quad \text{subject to} \quad \|\delta\mathbf{x}\|_{\ell_1} \leq \tau$$



with ℓ_1 constraint



with ℓ_2 constraint

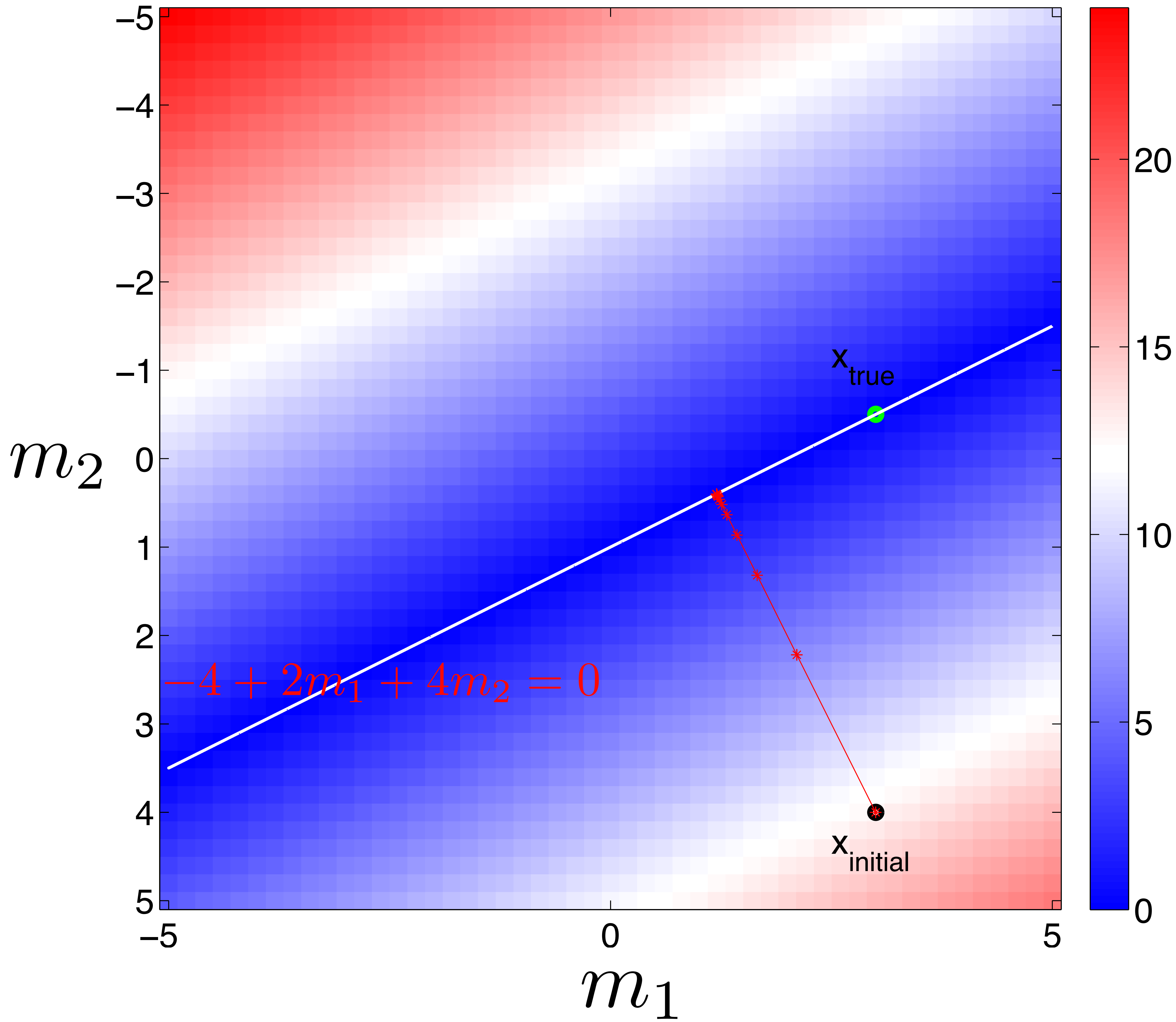
Linear example with multiple solutions (underdetermined)

$$\Phi(\mathbf{m}) := \left\{ \frac{1}{2} \|\mathbf{d} - \mathbf{A}\mathbf{m}\|_2^2 \right\}$$

$$\mathbf{A} = \begin{bmatrix} 2 & 4 \end{bmatrix}$$

$$\mathbf{d} = -4$$

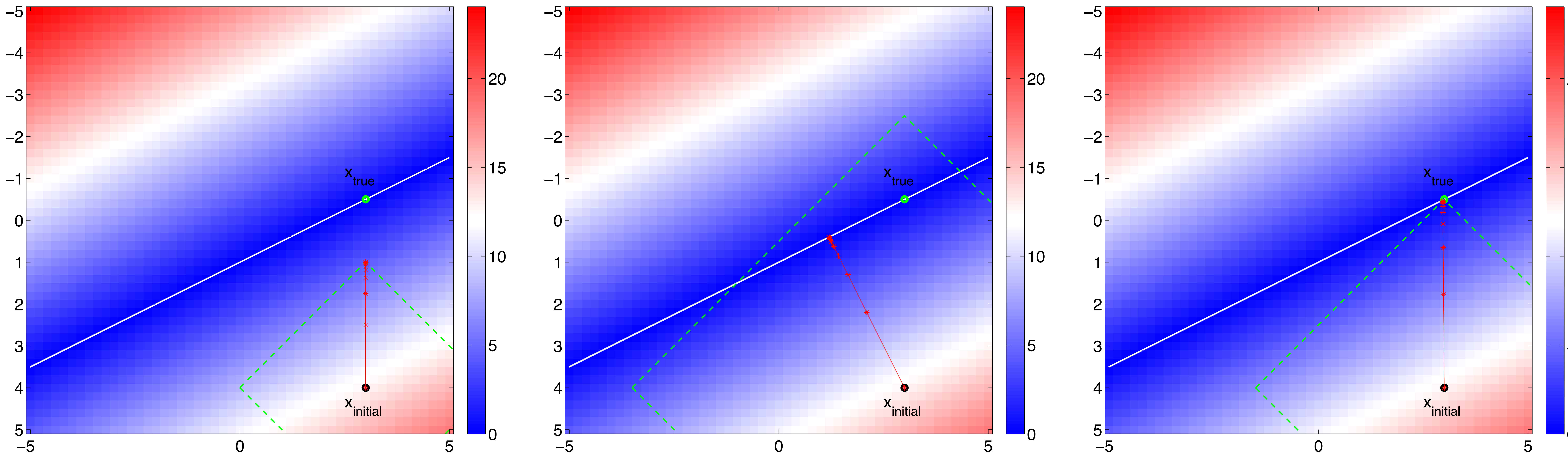
$$\mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}$$



GN with unconstrained objective

GN with sparse constrained objective function

$$\min_{\mathbf{x}} \Phi(\mathbf{x}) := \left\{ \frac{1}{2} \|\mathbf{d} - \boldsymbol{\mathcal{F}}[\mathbf{S}^{\mathrm{H}}\mathbf{x}]\|_2^2 \right\} \quad \text{subject to} \quad \|\mathbf{x} - \mathbf{x}_0\|_{\ell_1} \leq \tau$$



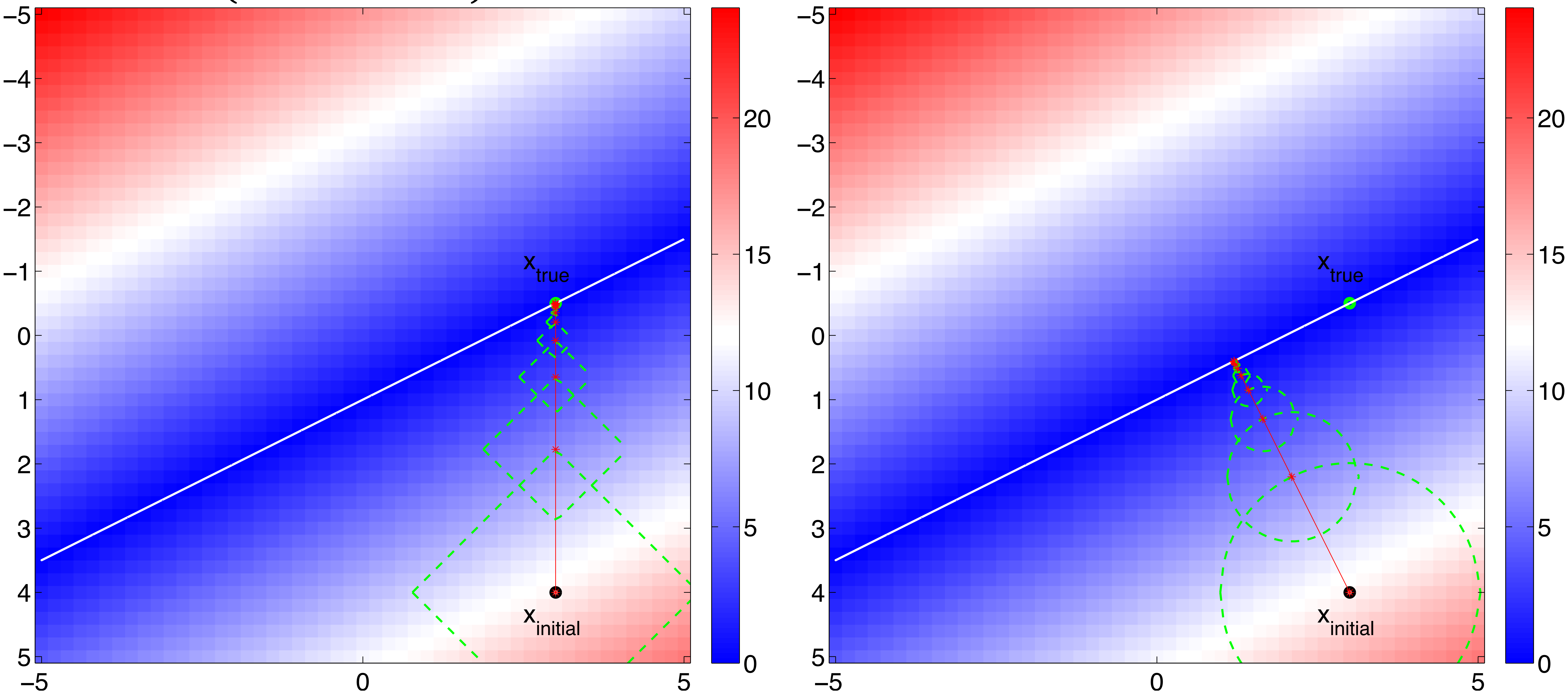
with wrong constraint — $\tau < \tau_{true}$

with wrong constraint — $\tau > \tau_{true}$

with right constraint — $\tau = \tau_{true}$

the modified Gauss-Newton

$$\min_{\mathbf{m}} \Phi(\mathbf{m}) := \left\{ \frac{1}{2} \|\mathbf{d} - \boldsymbol{\mathcal{F}}[\mathbf{m}]\|_2^2 \right\} \qquad \delta\mathbf{m} = \mathbf{S}^{\mathrm{H}} \arg\min_{\delta\mathbf{x}} \|\delta\mathbf{d} - \mathbf{J}[\mathbf{m}_k]\mathbf{S}^{\mathrm{H}}\delta\mathbf{x}\|_2^2 \quad \text{subject to} \quad \|\delta\mathbf{x}\|_{\ell_1} \leq \tau$$



with ℓ_1 constraint **with ℓ_2 constraint**

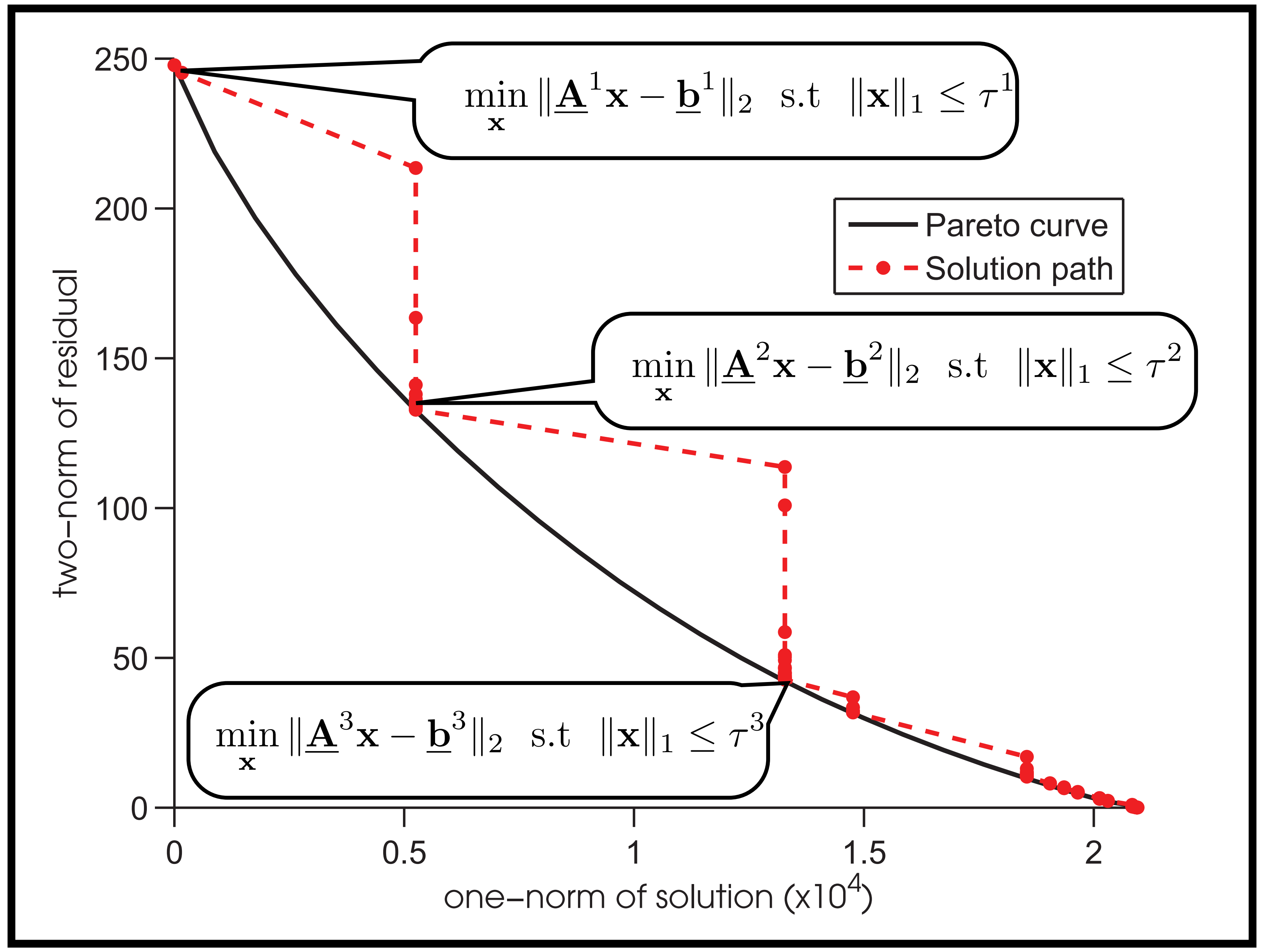
How to choose sparsity level for the modified GN

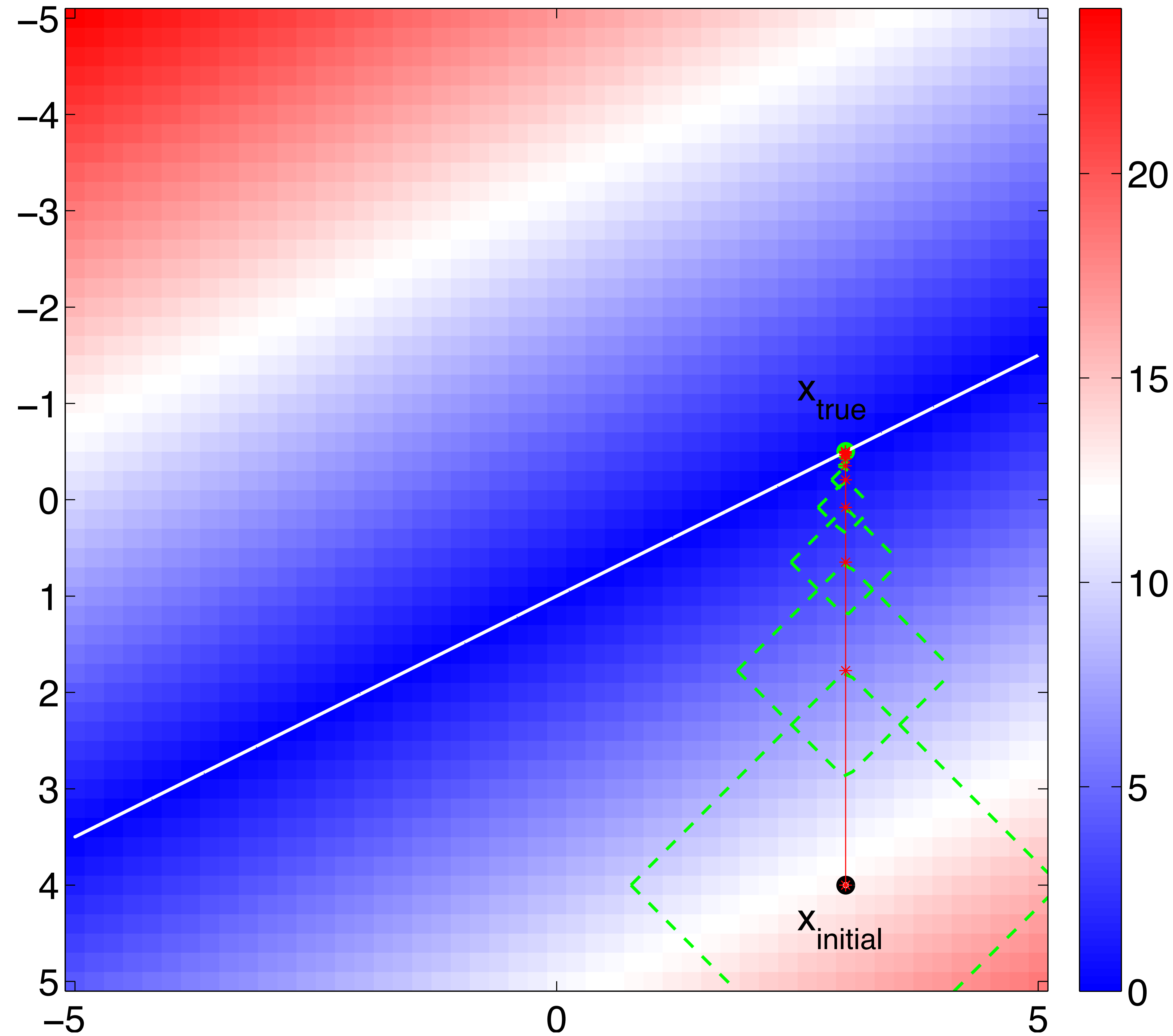
$$\min_{\mathbf{m}} \Phi(\mathbf{m}) := \left\{\frac{1}{2}\|\mathbf{d} - \boldsymbol{\mathcal{F}}[\mathbf{m}]\|_2^2\right\} \qquad \delta\mathbf{m} = \mathbf{S}^{\mathrm{H}} \arg\min_{\delta\mathbf{x}} \|\delta\mathbf{d} - \mathbf{J}[\mathbf{m}_k]\mathbf{S}^{\mathrm{H}}\delta\mathbf{x}\|_2^2 \quad \text{subject to} \quad \|\delta\mathbf{x}\|_{\ell_1} \leq \tau$$

BPDN problem

$$\delta\mathbf{m} = \mathbf{S}^{\mathrm{H}} \arg\min_{\delta\mathbf{x}} \|\delta\mathbf{x}\|_1$$

$$\text{subject to} \quad \|\delta\mathbf{d} - \mathbf{J}\mathbf{S}^{\mathrm{H}}\delta\mathbf{x}\|_2^2 \leq \sigma$$





Pareto reference

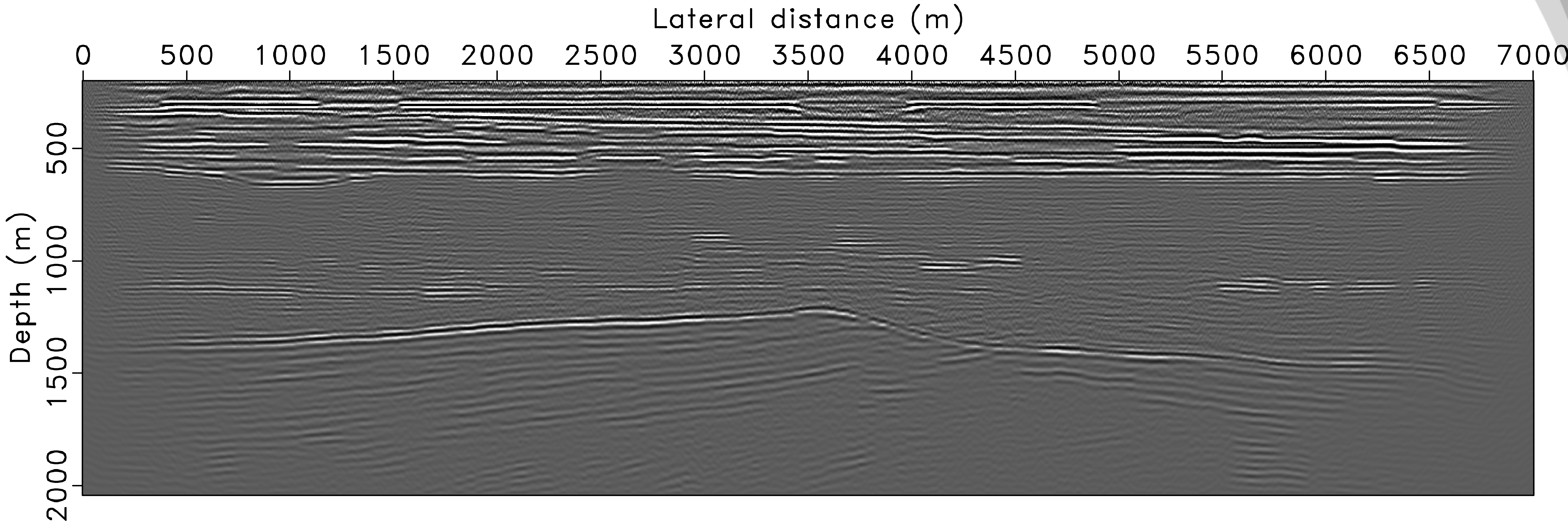
Least-squares migration

$$\delta \mathbf{m} = \arg\min_{\delta \mathbf{m}} \|\delta \mathbf{u} - \mathbf{J}\delta \mathbf{m}\|_2^2$$

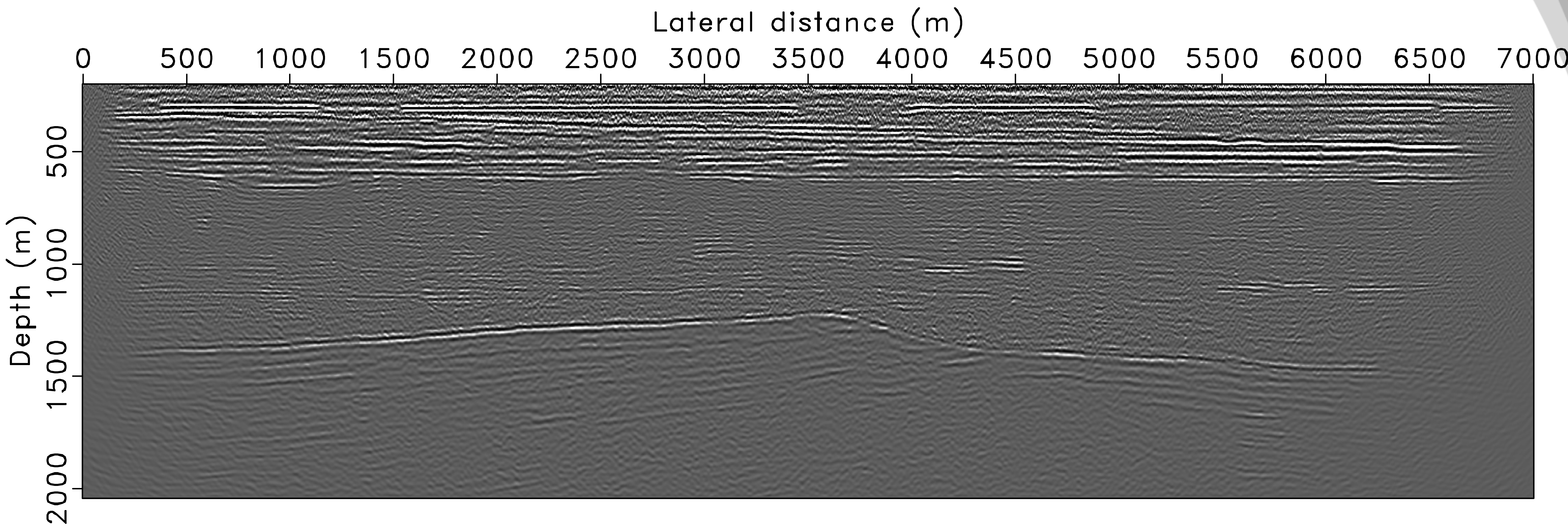
- 10 random frequencies (20Hz-50Hz)
- 17 randomly selected shots out of 350 shots
- LASSO problems determined by SPGL1

See "*Efficient least-squares imaging with sparsity promotion and compressive sensing*"

The modified GN with L1 constraint



The modified GN with L2 constraint



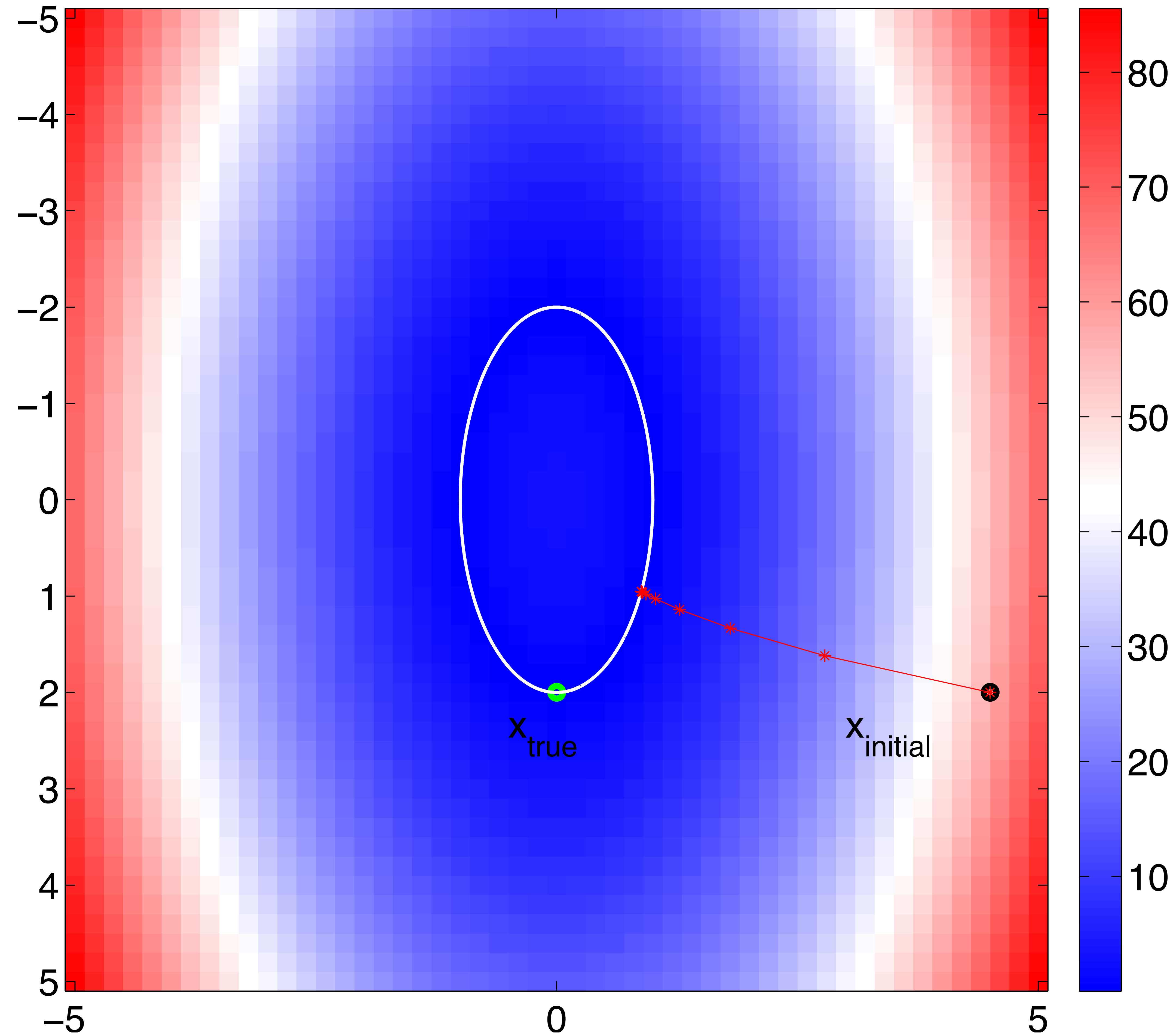
Nonlinear problem with multiple solutions

$$\|\mathbf{d} - \mathbf{m}^T \mathbf{A}^T \mathbf{A} \mathbf{m}\|_2^2$$

$$\mathbf{A} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

$$\mathbf{d} = -4$$

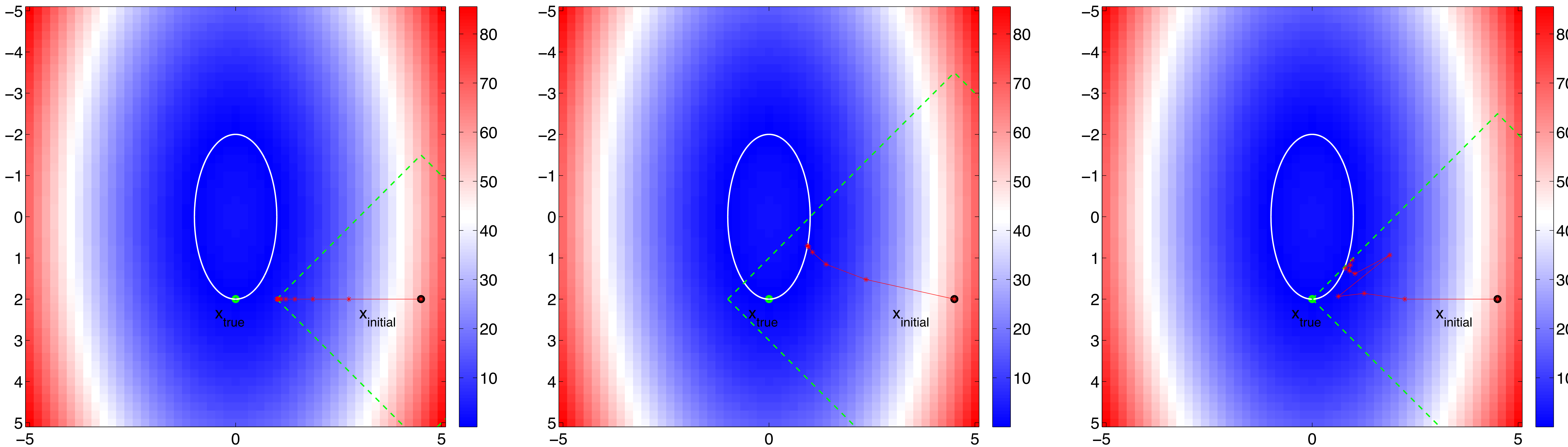
$$\mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}$$



GN with unconstrained objective

GN with sparse constrained objective function

$$\min_{\mathbf{x}} \Phi(\mathbf{x}) := \left\{ \frac{1}{2} \|\mathbf{d} - \boldsymbol{\mathcal{F}}[\mathbf{S}^{\mathrm{H}}\mathbf{x}]\|_2^2 \right\} \quad \text{subject to} \quad \|\mathbf{x} - \mathbf{x}_0\|_{\ell_1} \leq \tau$$



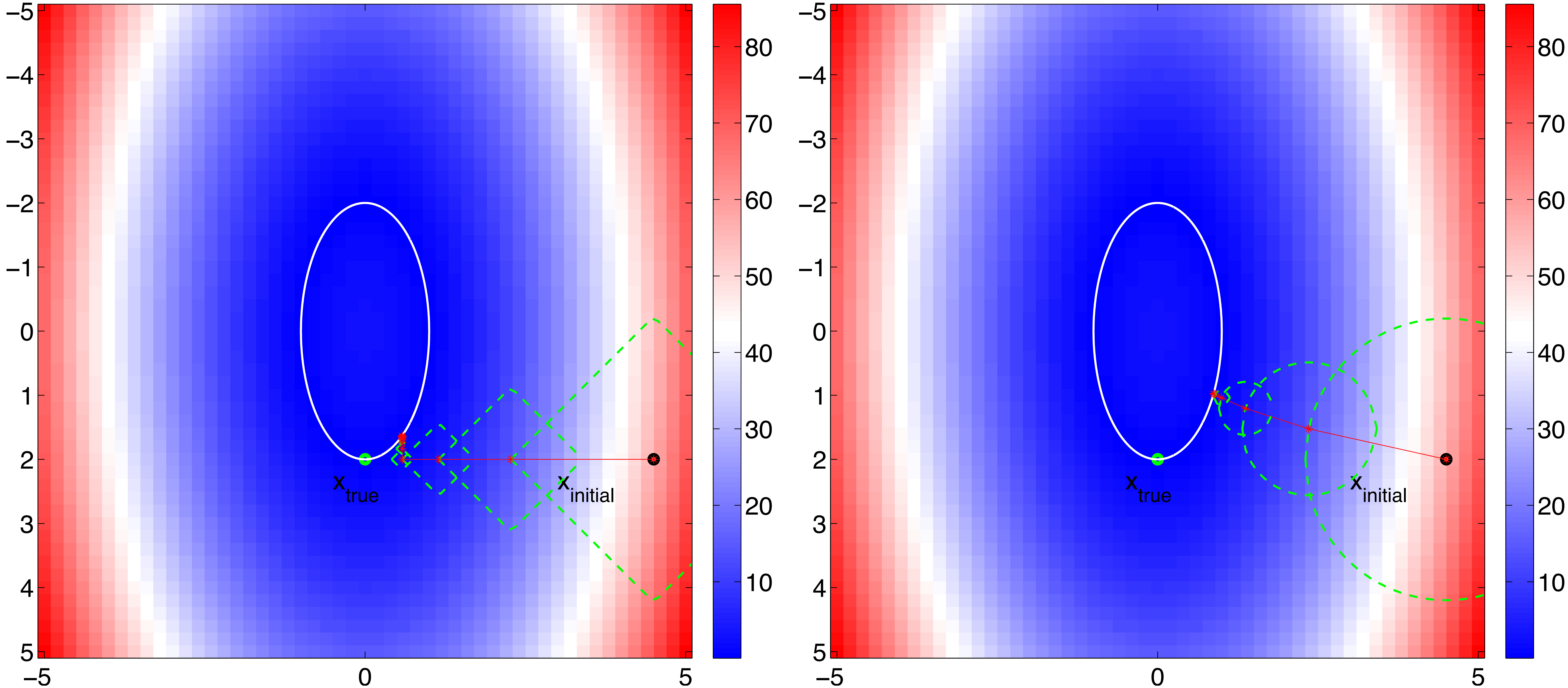
with wrong constraint $\tau < \tau_{true}$

with wrong constraint $\tau > \tau_{true}$

with right constraint $\tau = \tau_{true}$

The modified Gauss-Newton

$$\min_{\mathbf{m}} \Phi(\mathbf{m}) := \left\{ \frac{1}{2} \|\mathbf{d} - \boldsymbol{\mathcal{F}}[\mathbf{m}]\|_2^2 \right\} \qquad \delta\mathbf{m} = \mathbf{S}^{\mathrm{H}} \arg\min_{\delta\mathbf{x}} \|\delta\mathbf{d} - \mathbf{J}[\mathbf{m}_k]\mathbf{S}^{\mathrm{H}}\delta\mathbf{x}\|_2^2 \quad \text{subject to} \quad \|\delta\mathbf{x}\|_{\ell_1} \le \tau$$



Observations

- For convex problems with unique solution, the modified Gauss-Newton method will find the solution as other methods with unconstrained objective.

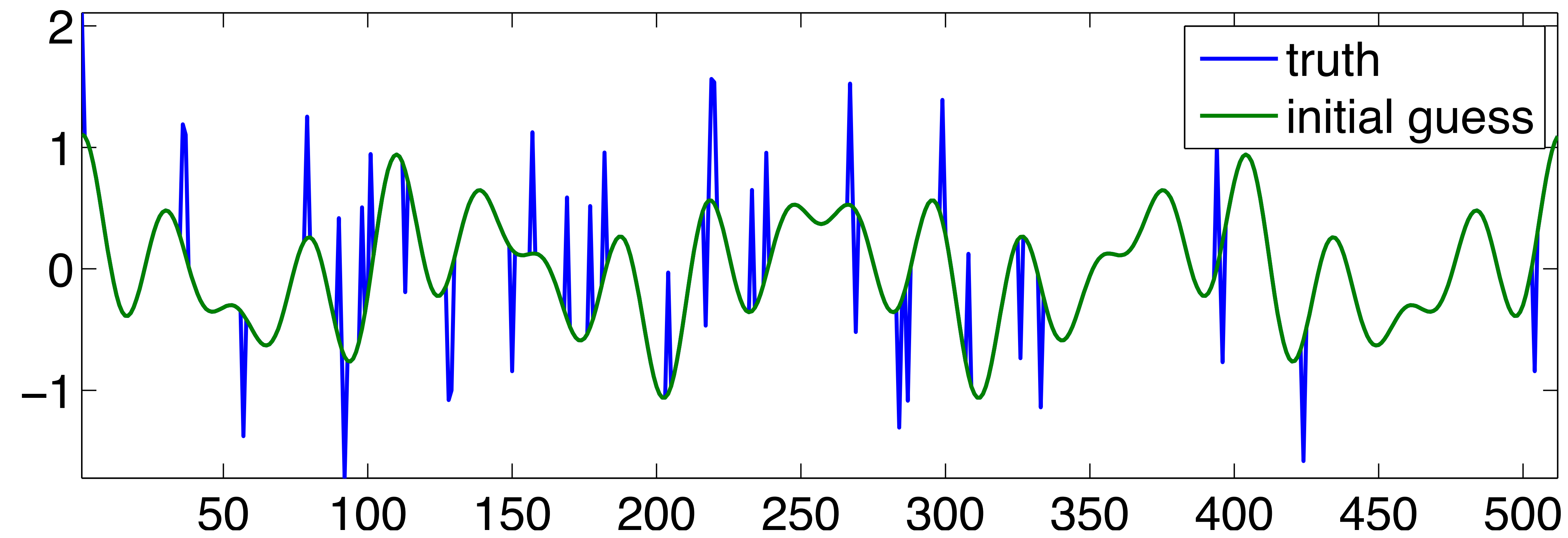
- For problems with multiple solutions, the modified Gauss-Newton method can find a solution with sparse perturbation of the initial guess, if updates share the same support.

The phase retrieval problem

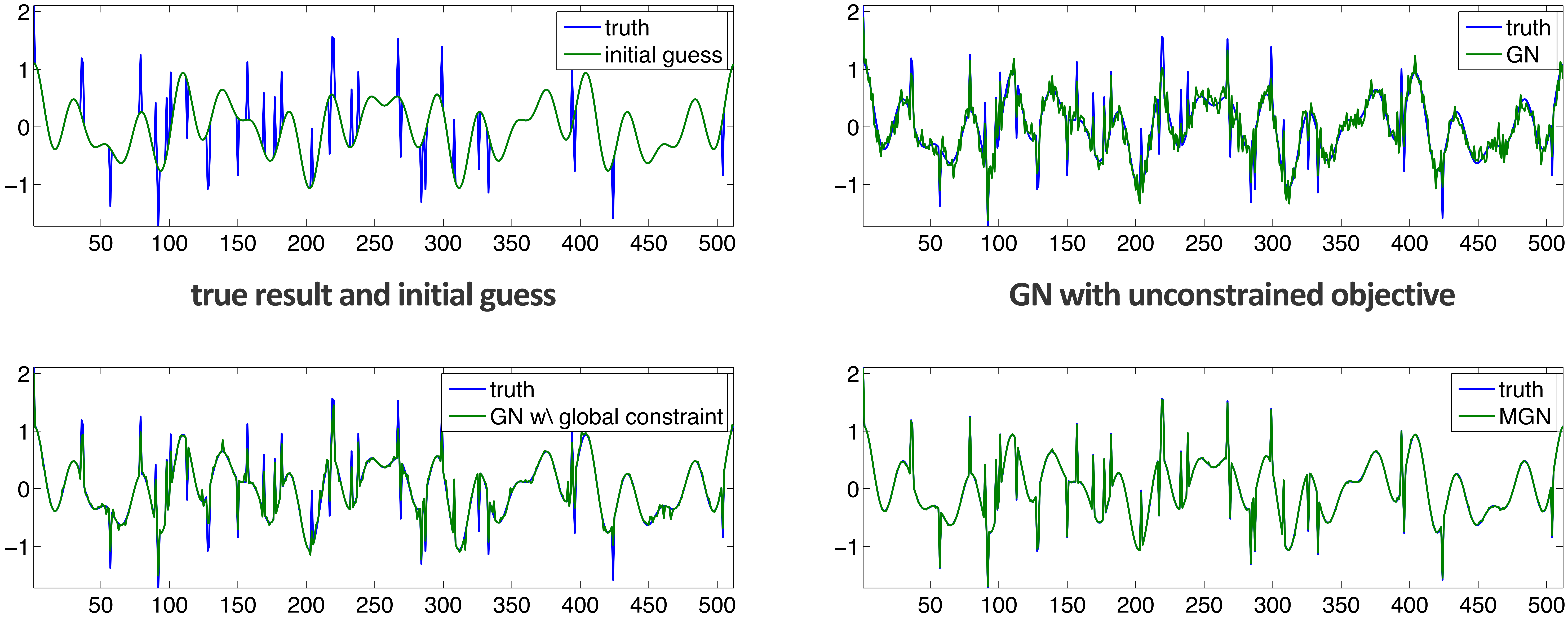
objective function:

$$\Phi(\mathbf{x}) := \left\{ \frac{1}{2} \| \mathbf{d} - \mathrm{diag}(\mathbf{Ax})(\mathbf{Ax}) \|_2^2 \right\}$$

$$\begin{aligned} \mathbf{A} &: 400 \times 512 \text{ matrix} \\ \mathbf{d} &: 400 \times 1 \text{ vector} \\ \mathbf{x} &: 512 \times 1 \text{ unknown vector} \end{aligned}$$



Results



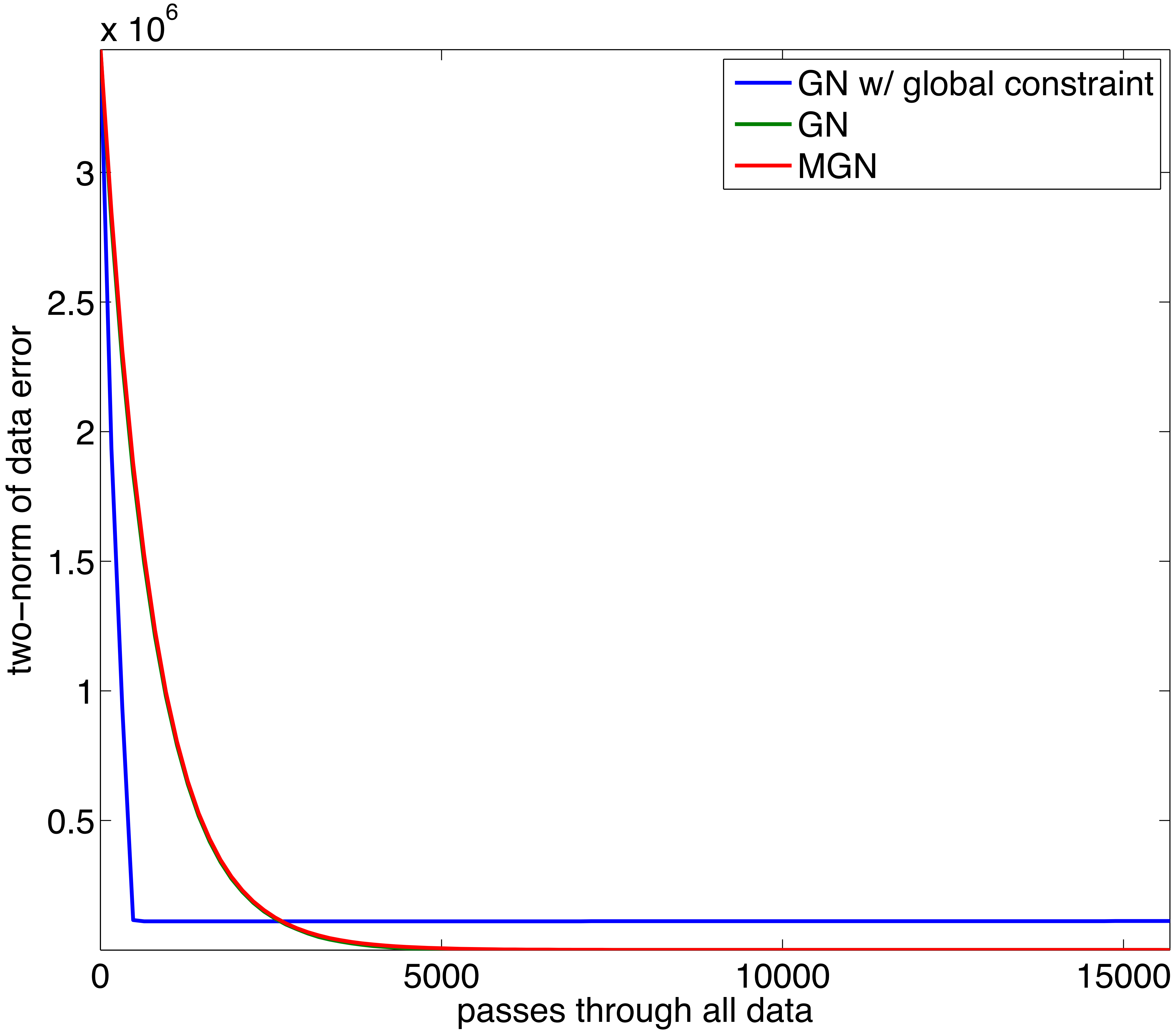
true result and initial guess

GN with unconstrained objective

GN with ℓ_1 constrained objective

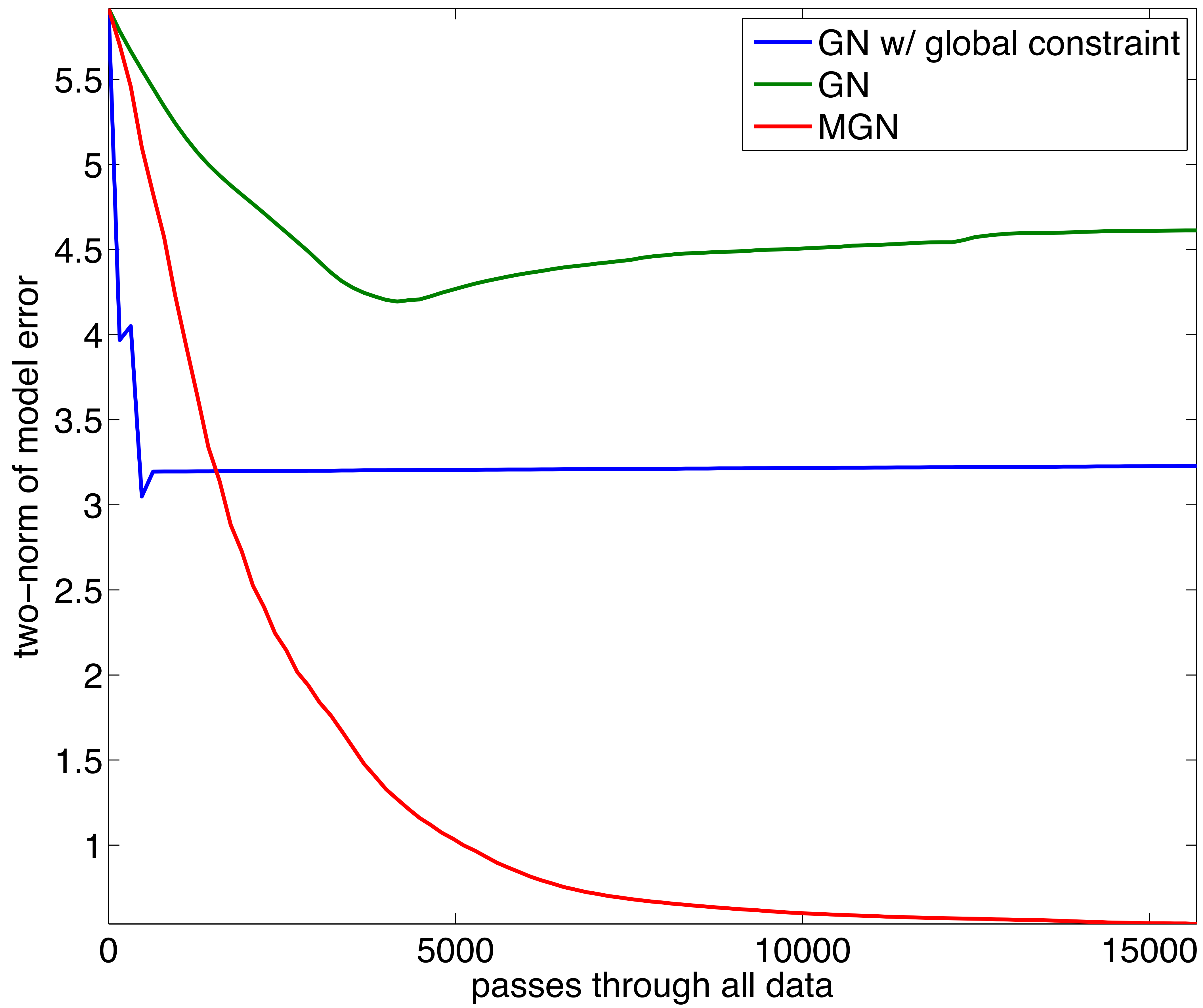
modified Gauss-Newton

Convergency



objective

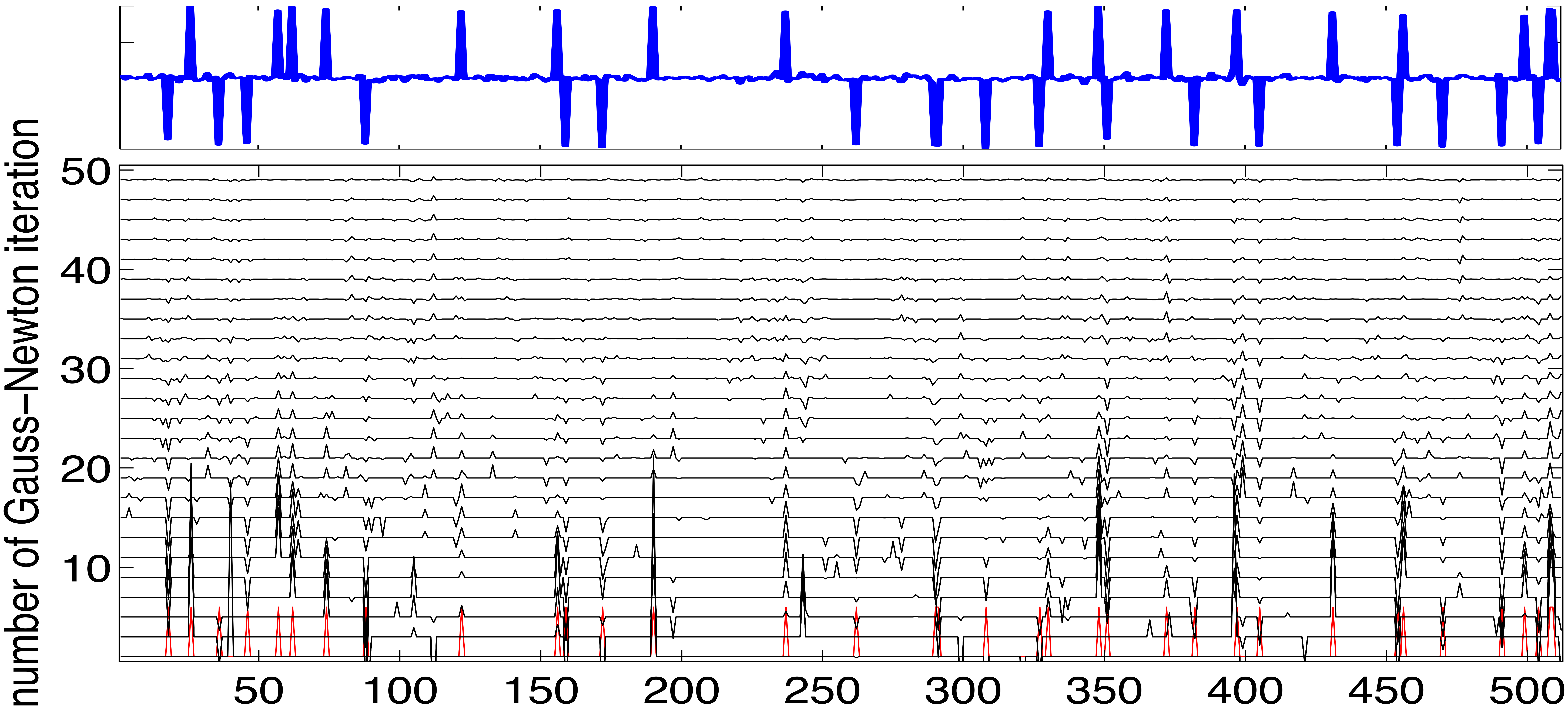
$$\frac{1}{2}\|\mathbf{d} - \mathrm{diag}(\mathbf{A}\mathbf{x}_k)(\mathbf{A}\mathbf{x}_k)\|_2^2$$



relative model error

$$\|\mathbf{x}_k - \mathbf{x}_{true}\|_2$$

Modified Gauss-Newton updates

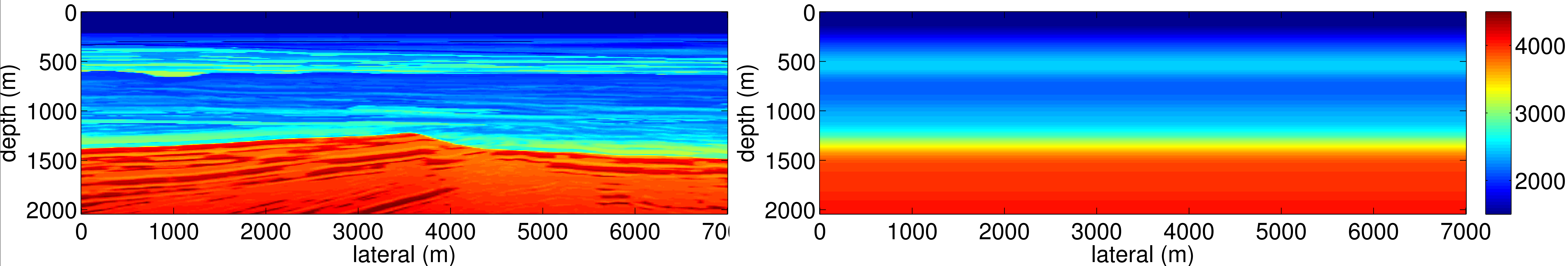


red is the position of sparse support

BG model example

BG Compass model

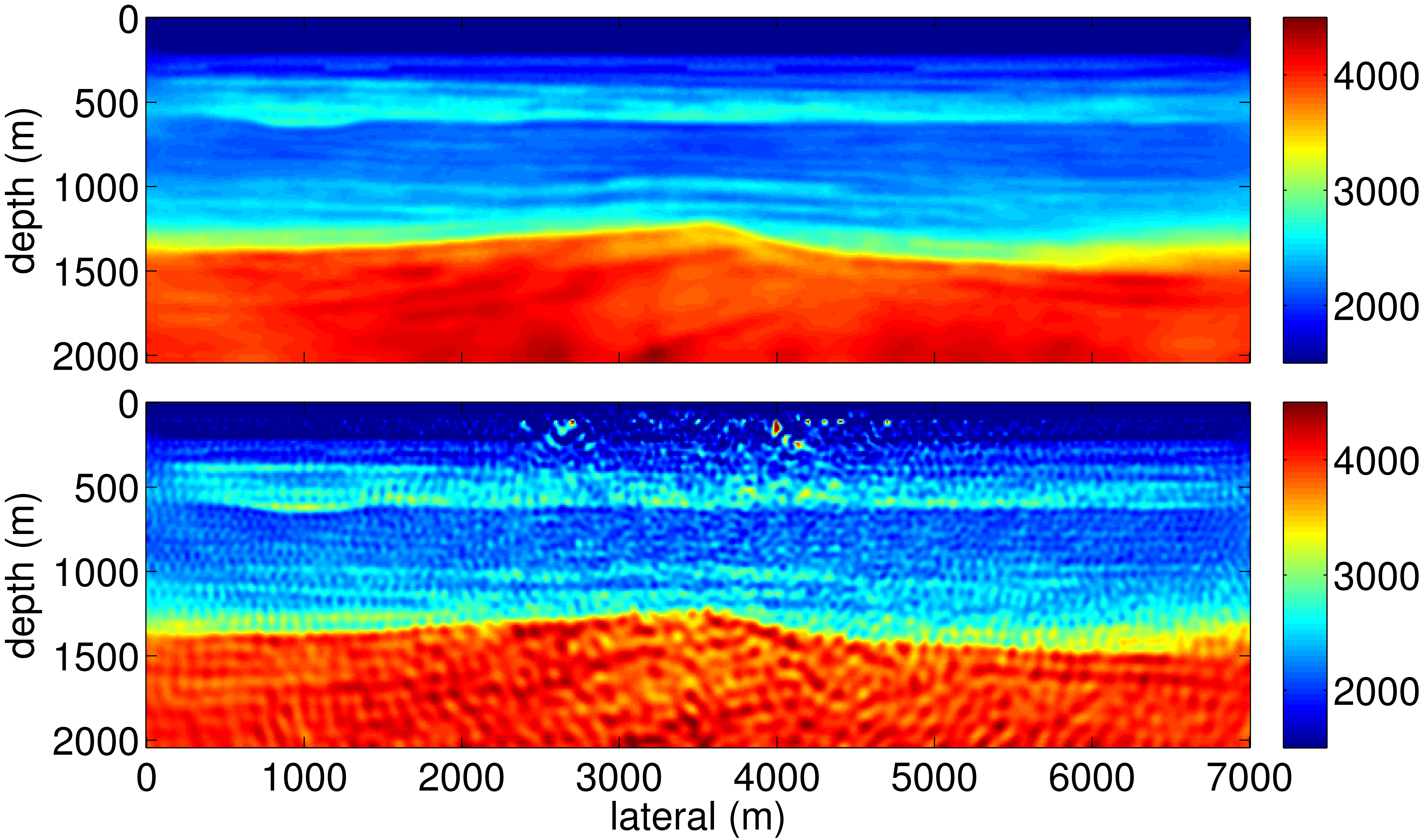
- 2 x 7 km
- 350 shot positions, 700 fixed receivers
- 3-15Hz, 10 frequency bands
- 5 GN updates for each band
- observed data is from time domain finite difference



true model

initial model

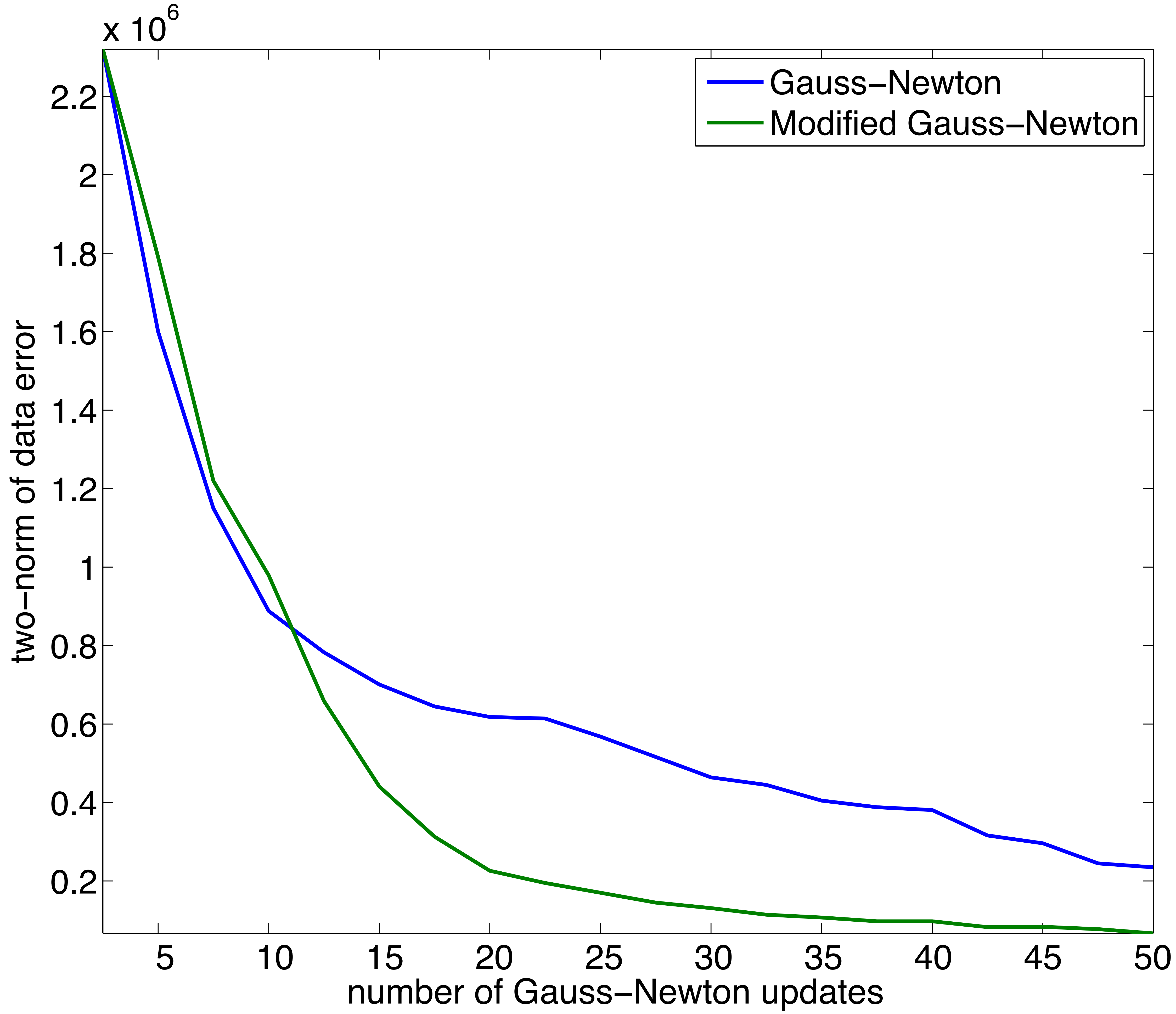
Inversion results



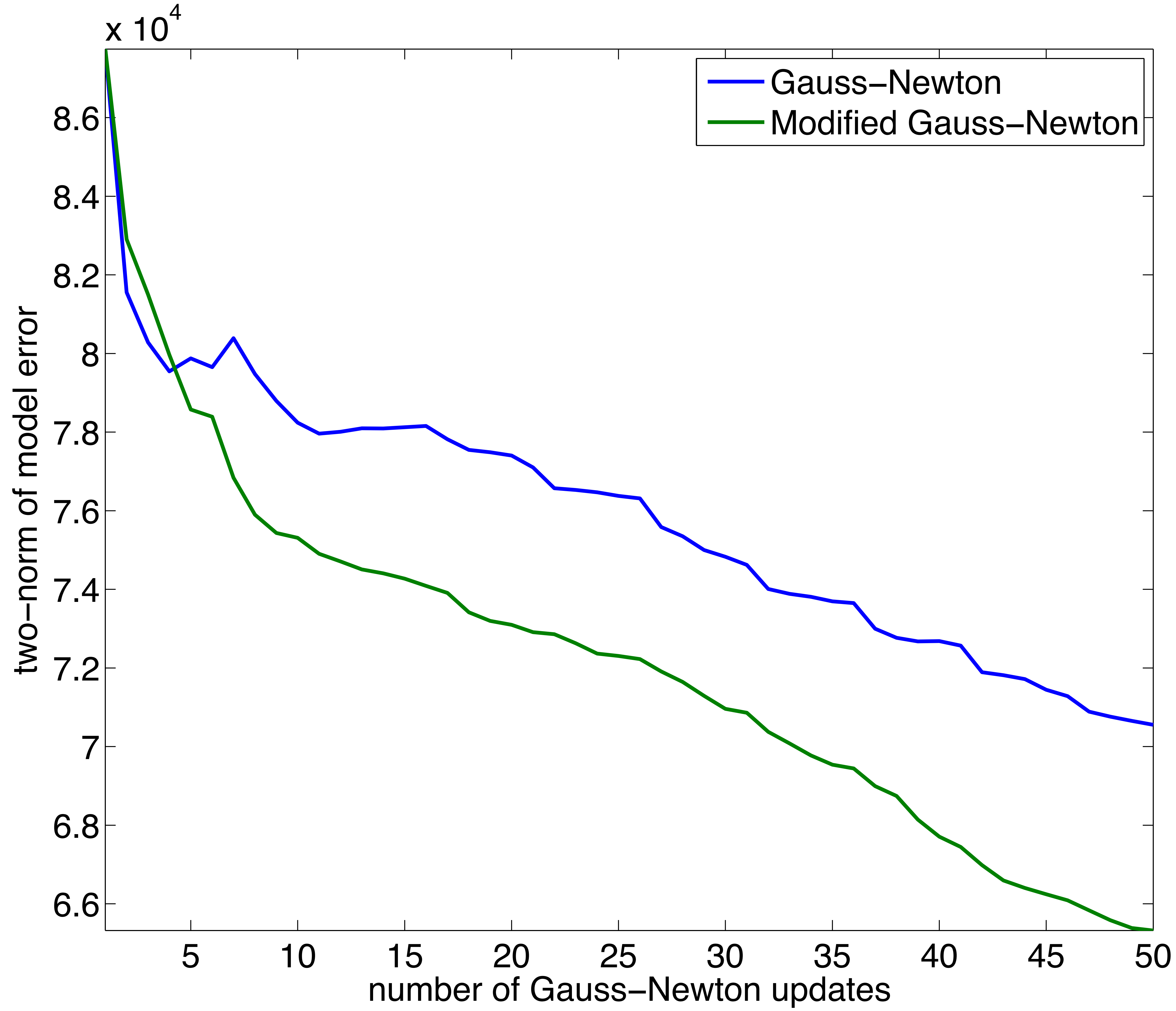
the modified Gauss-Newton method with L1 constraint

the modified Gauss-Newton method with L2 constraint

Convergence

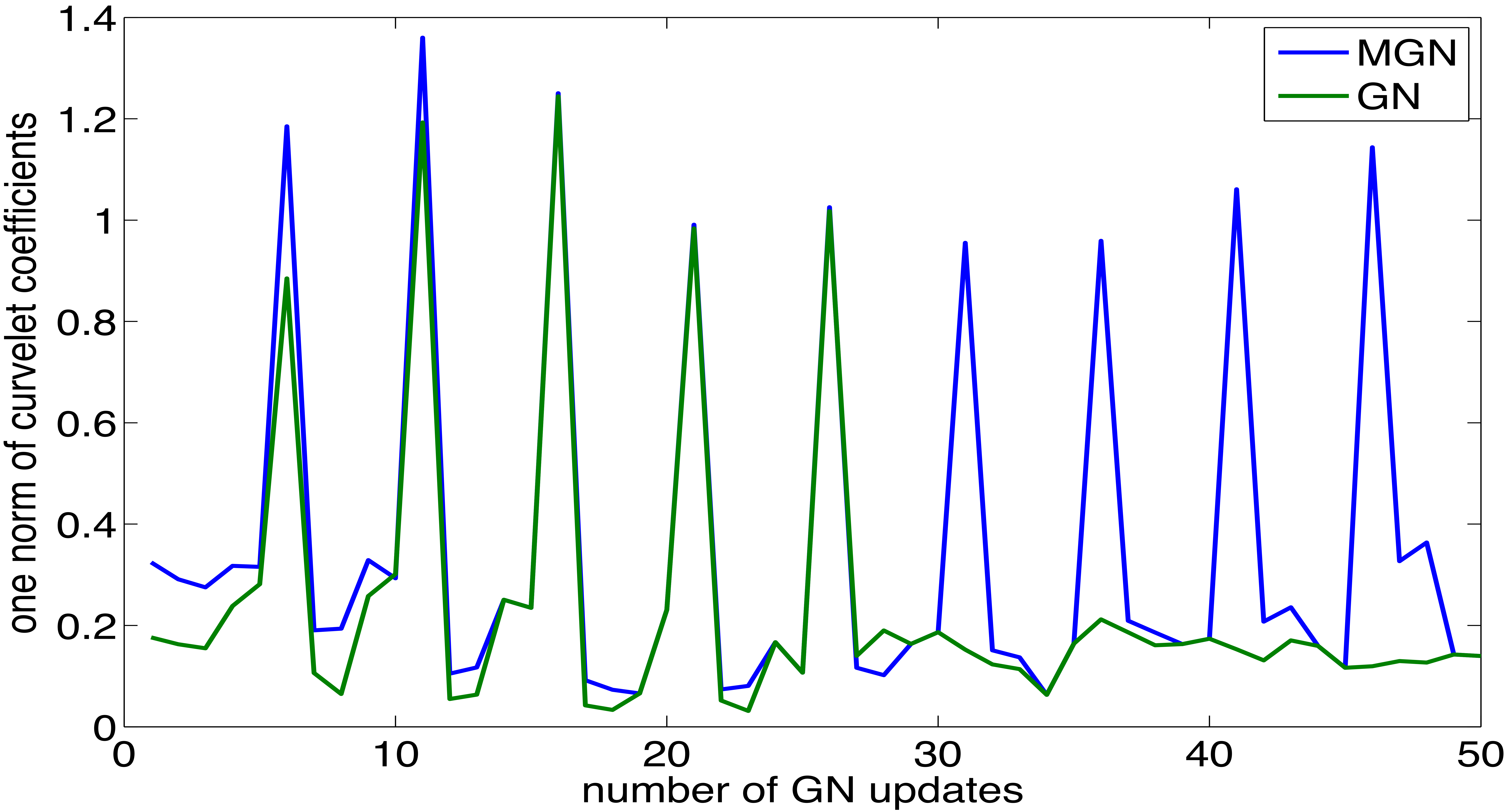


FWI objective



model error

Curvelet coefficients of updates



SLIM

Conclusion

- introduce sparsity constraint by changing the objective function does not necessarily generating the solution with sparse perturbation of the initial guess
- the modified Gauss-Newton method can provides us the solution with sparse perturbation, if all updates share the same sparsity support
- the modified Gauss-Newton works for linear problems (seismic imaging)
- the modified Gauss-Newton works for FWI, if a reasonable initial model is provided

Acknowledgements

Thank you for your attention !

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

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, BGP, BP, CGG, Chevron, ConocoPhillips, ION, Petrobras, PGS, Total SA, WesternGeco, and Woodside.