Quadratic-penalty based full-space methods for waveform inversion

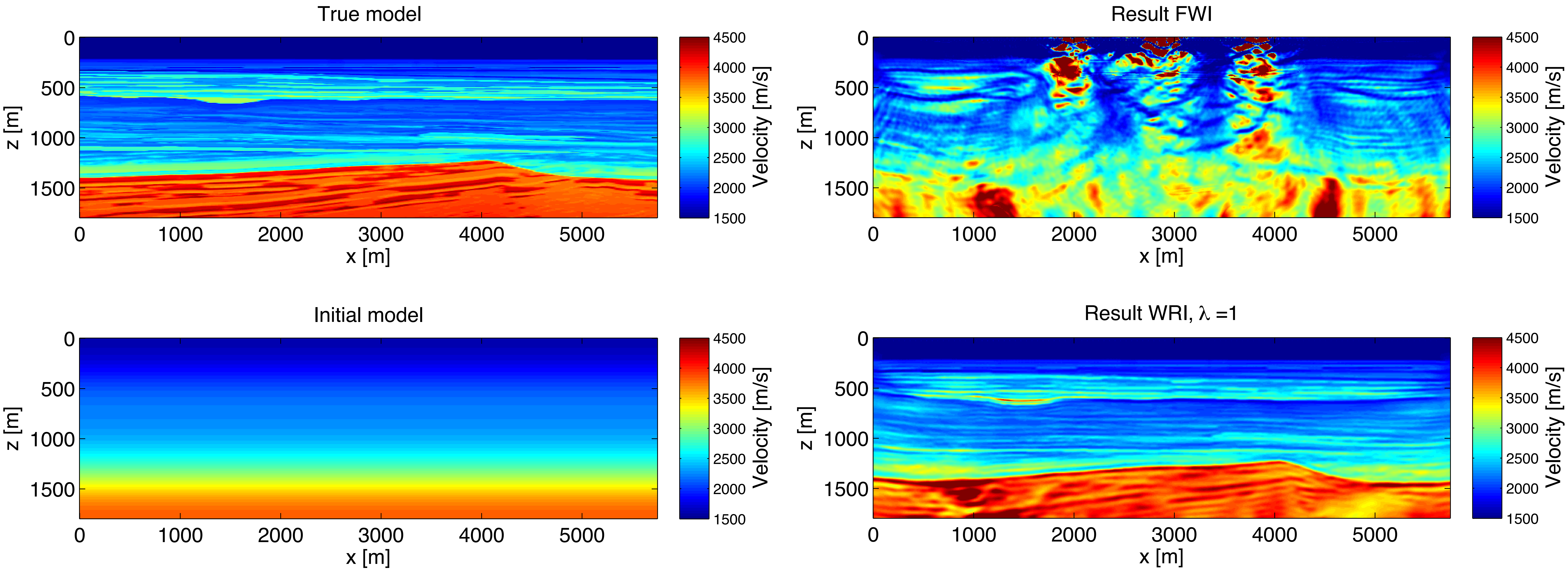
Bas Peters
SINBAD Fall Consortium meeting, 2014.
Whistler, BC.

SLIM
University of British Columbia

Waveform inversion

Works well if initial model is good

Waveform inversion – poor start model



Example from **[Peters et al. 2013]**

Wavefield reconstruction inversion

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

Less sensitive to

- starting models compared to FWI
- missing low-frequency data

Avoids cycle-skipping problems by virtue of extending the search space.

But, requires reasonable accurate solve of the augmented wave equation...

Wavefield reconstruction inversion

Still limitations on quality of start models & missing low-frequencies

Can we do better?

Maybe full-space methods?

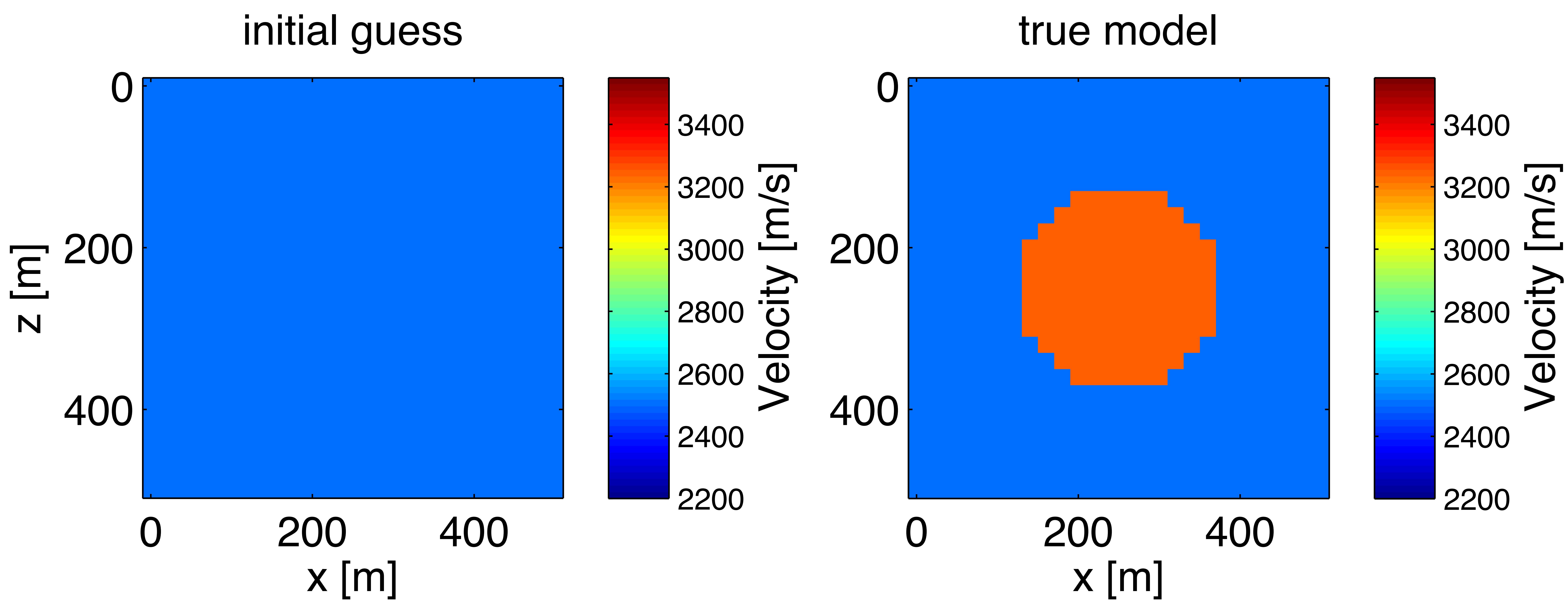
- at the "expense" of storing 2 copies of monochromatic wavefields

But, at the gain of no longer insisting on accurate solves...

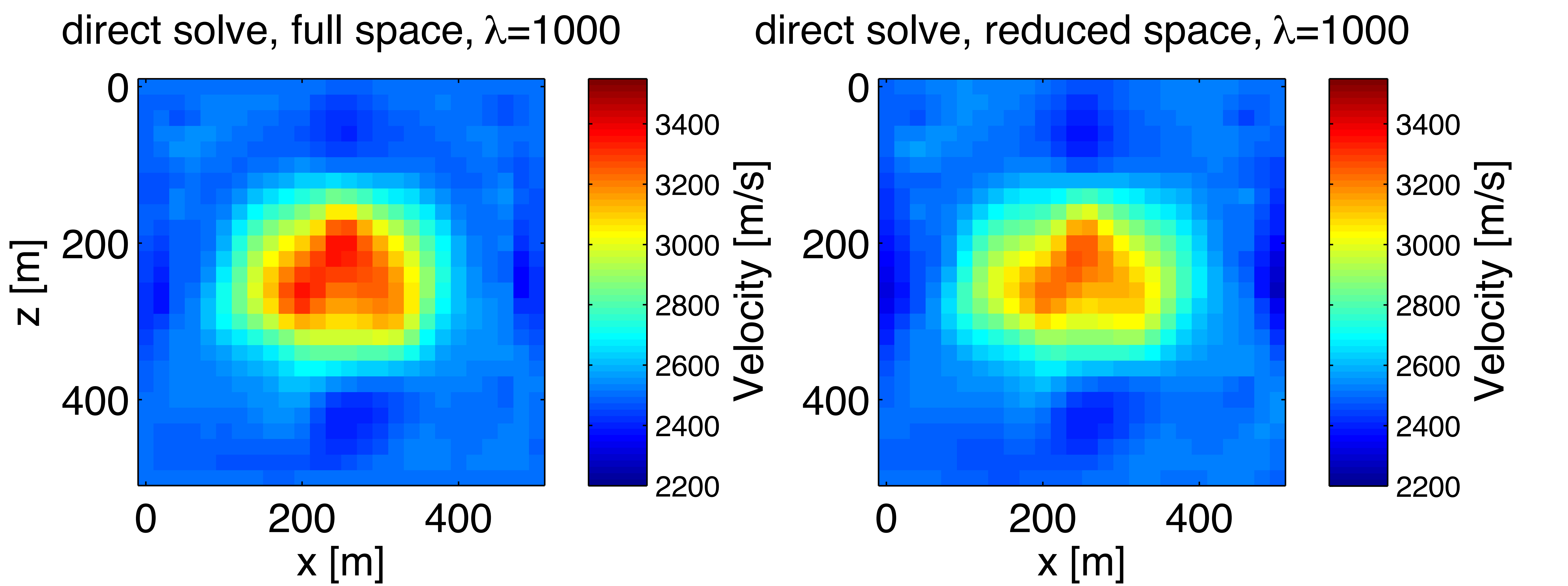
Toy problem

- cross-well setting
- 4 frequencies [6-10] Hz
- 5 simultaneous sources
- 5 receivers

Toy problem

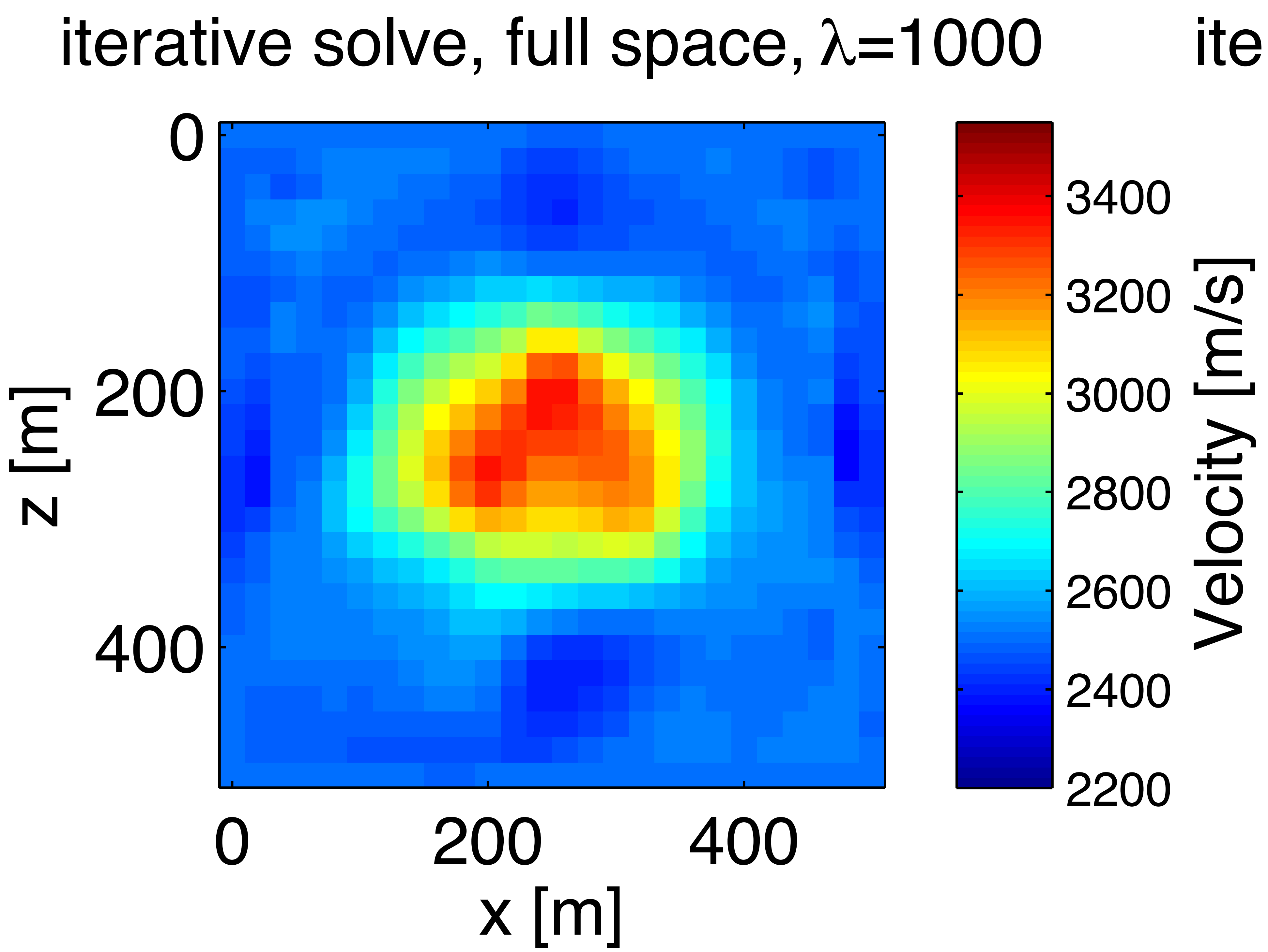

initial guess
true model
0
200
400
0
200
400
z [m]
x [m]
3400
3200
3000
2800
2600
2400
2200
Velocity [m/s]

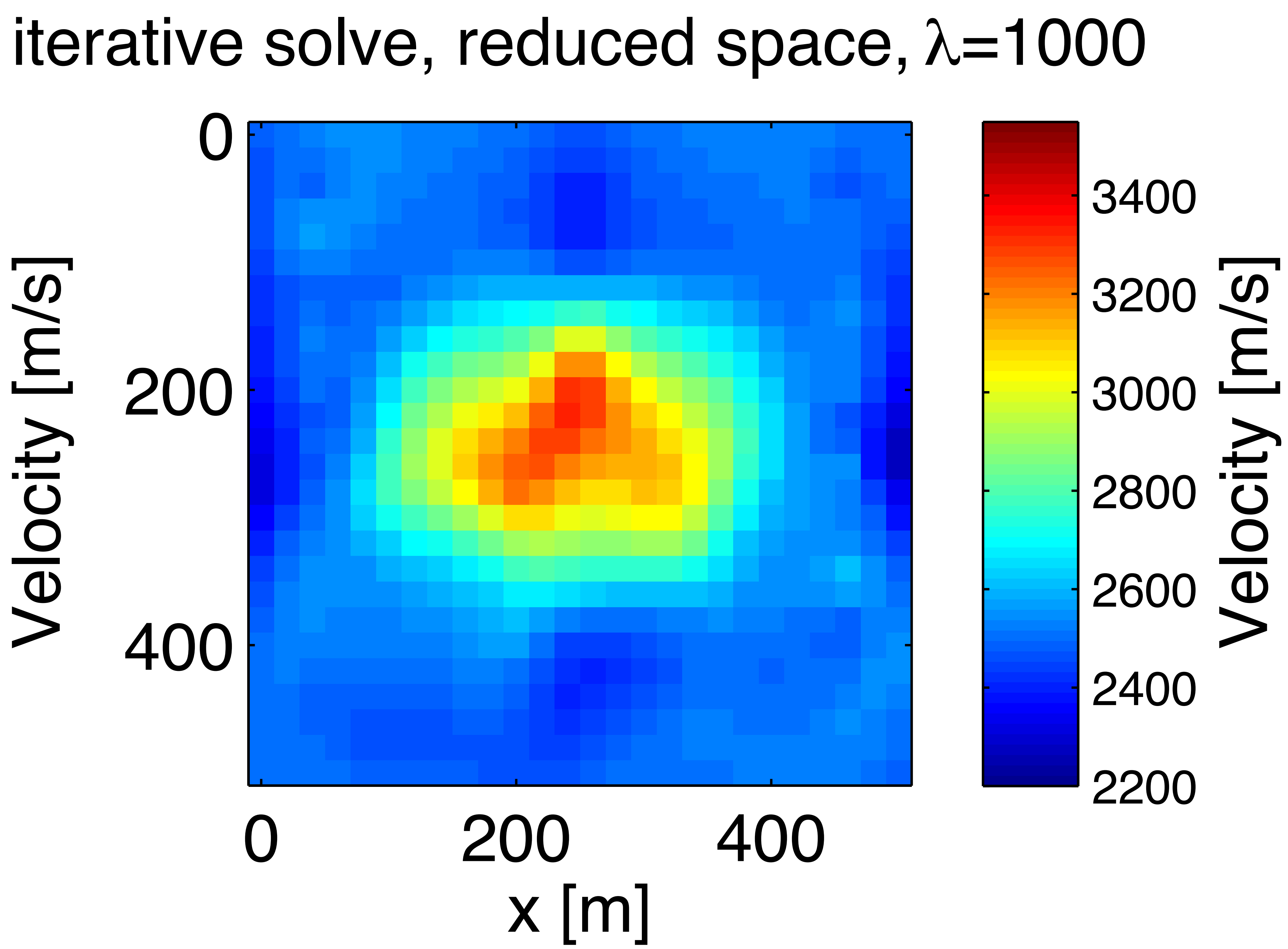
Toy problem



direct solution for least-squares problems

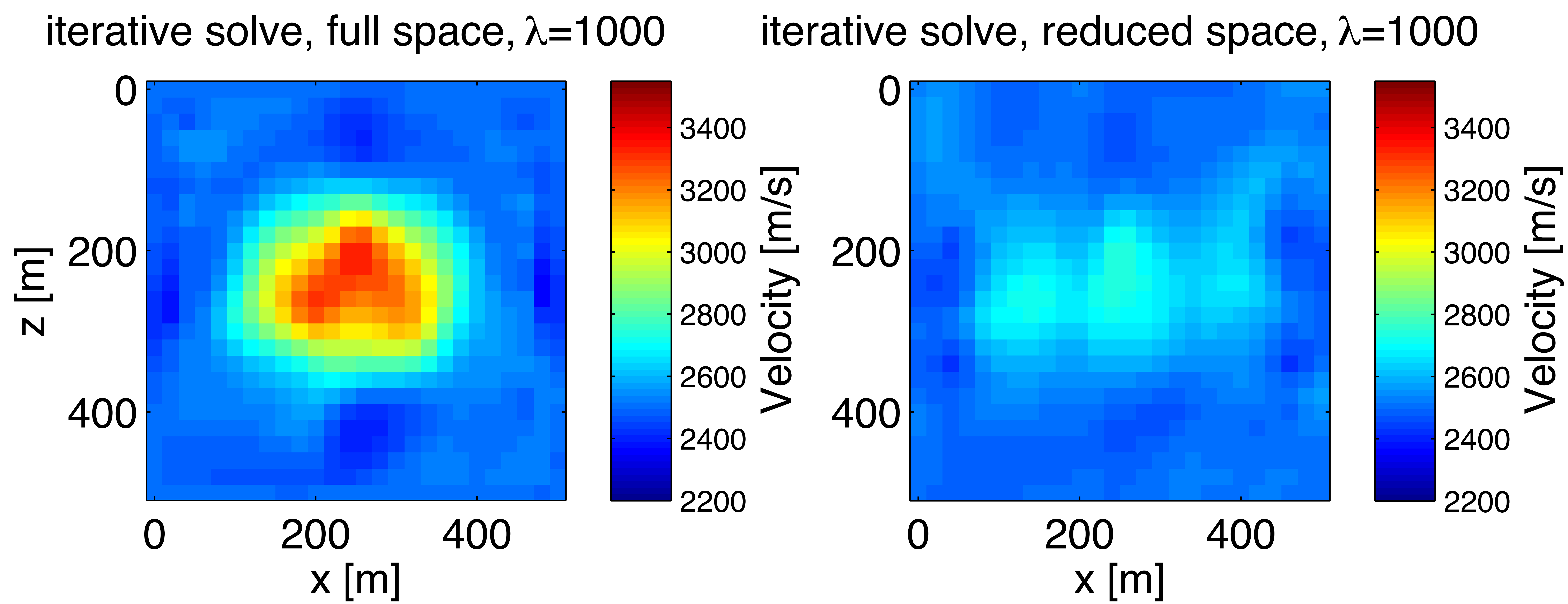
Toy problem





accurate iterative solution for least-squares problems

Toy problem

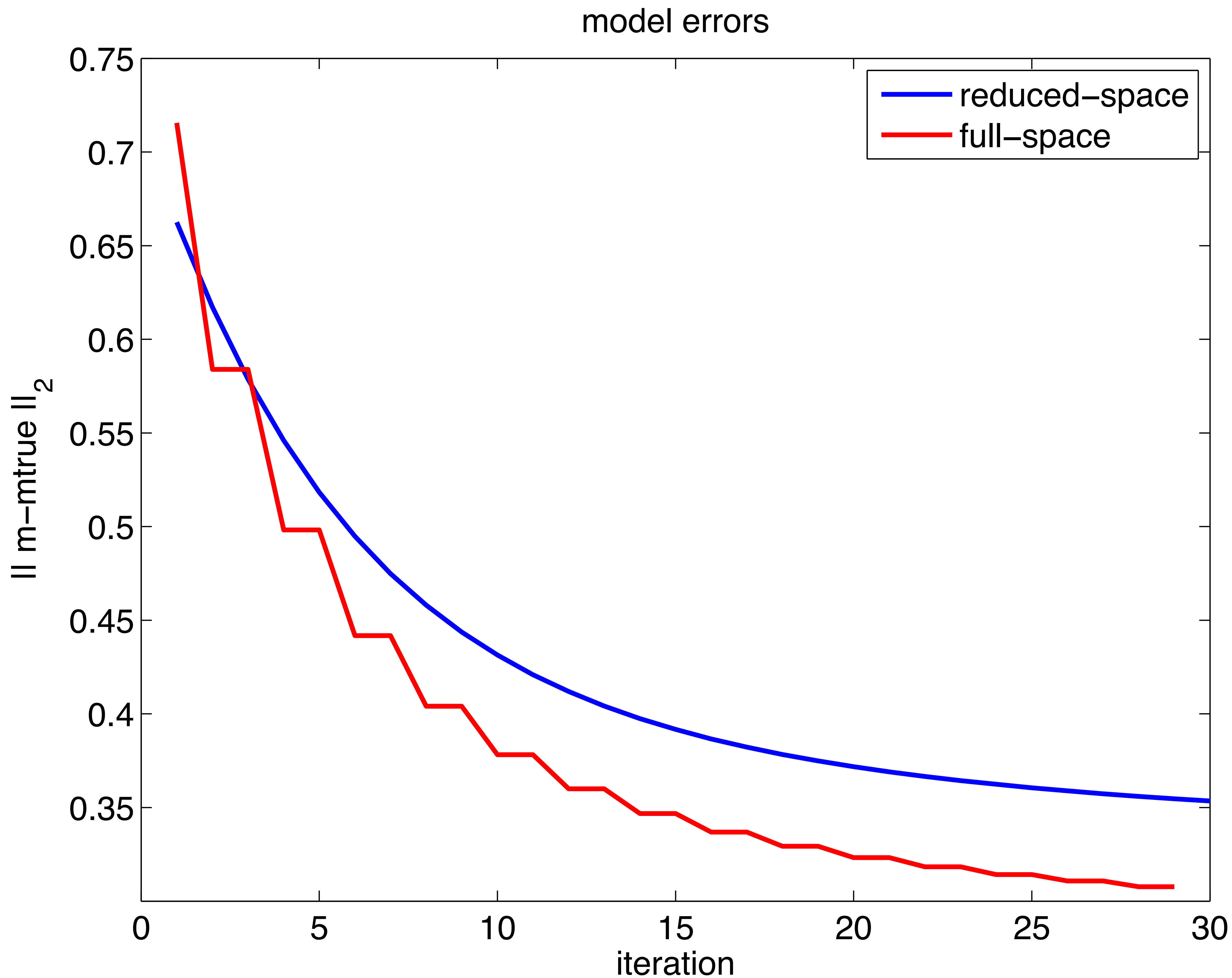


inaccurate iterative solution for least-squares problems

Toy problem

model errors,

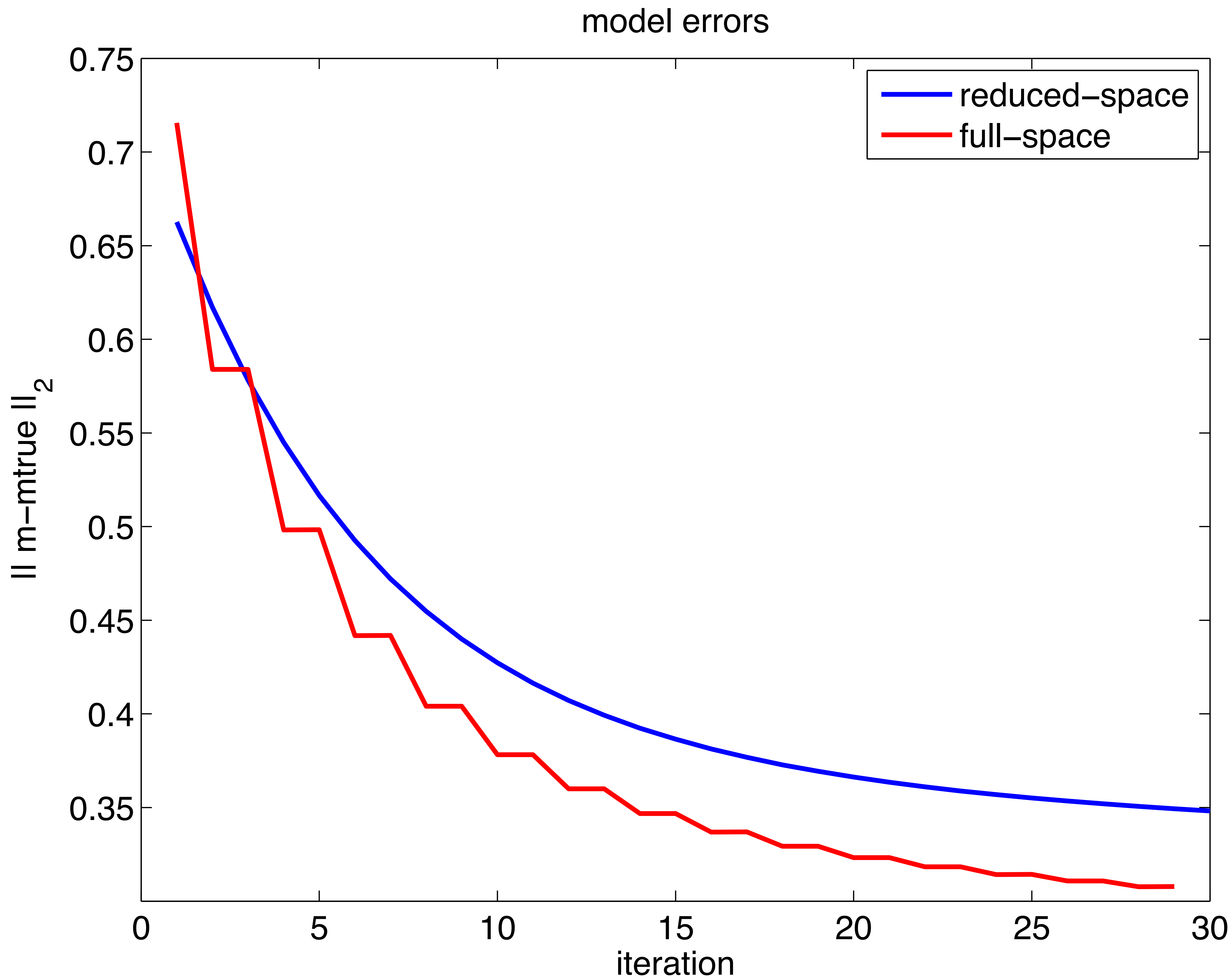
direct solver



Toy problem

model errors,

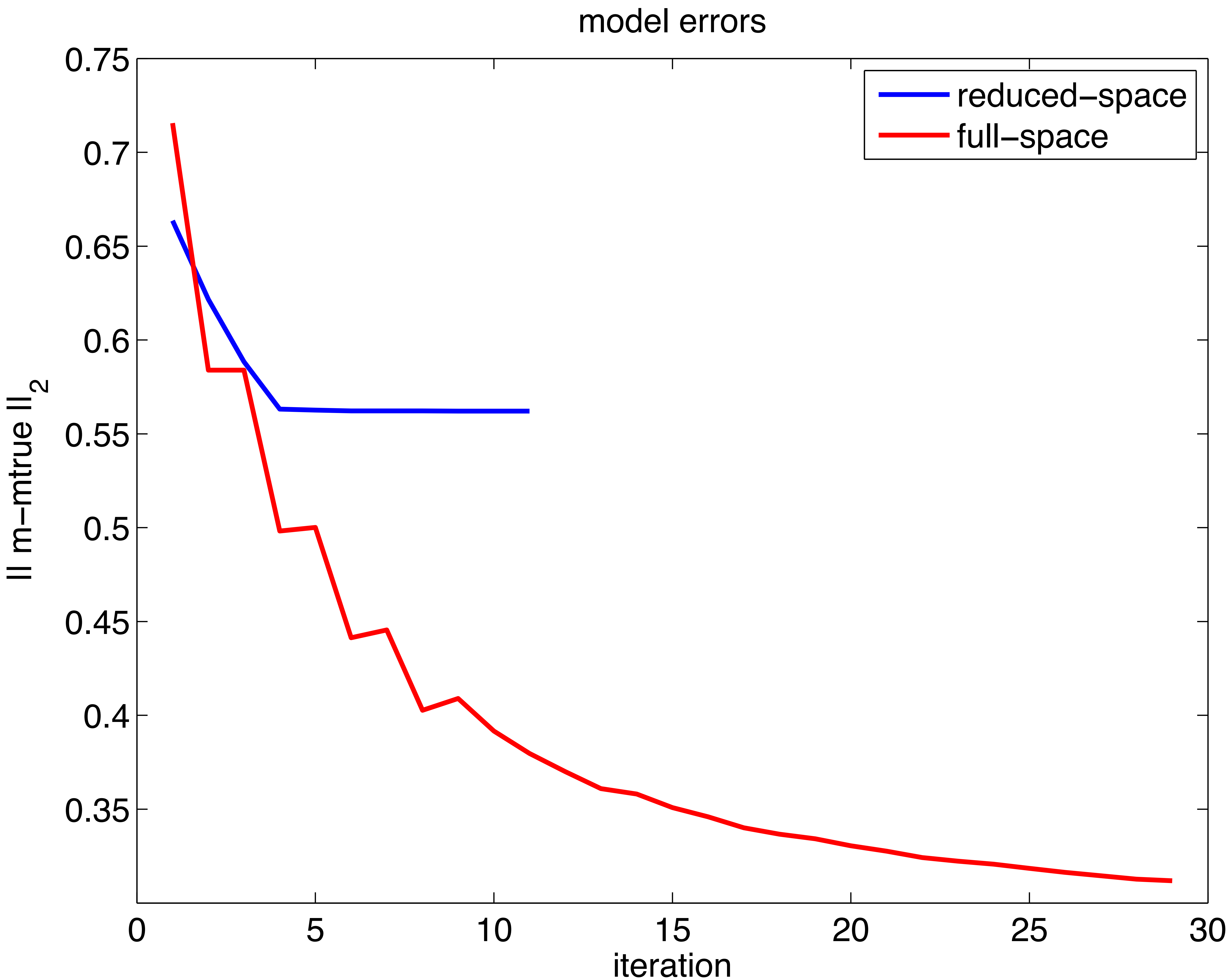
accurate iterative solutions



Toy problem

model errors,

inaccurate iterative solutions



Toy problem

Inexact iterative linear system solves:

- full-space method not very sensitive
- WRI & FWI quite sensitive

Full-space vs reduced-space methods

Bottom line

reduced-space: solve for the fields
update the medium parameters

full-space: update fields & medium parameters

Full-space vs reduced-space methods

Bottom line

reduced-space: solve for the fields
update the medium parameters ← alternating strategy

full-space: update fields & medium parameters

Full-space vs reduced-space methods

Bottom line

reduced-space: solve for the fields
update the medium parameters ← alternating strategy

full-space: update fields & medium parameters ← joint updating

Full-space vs reduced-space methods

FWI & WRI are in the reduced-space class –i.e., wave-equations are solved

Full-space is commonly used in a Lagrangian setting.

Because of memory requirements, rarely used in (academic) geophysics. **[EM: E. Haber et al., 2004 ; Seismic: M. J. Grothe et al., 2011]**

Short derivation

Problem formulation:

$$\min_{\mathbf{m},\mathbf{u}} \frac{1}{2}\|\mathbf{P}\mathbf{u}-\mathbf{d}\|_2^2 \quad \text{s.t.} \quad \mathbf{H}(\mathbf{m})\mathbf{u}=\mathbf{q}$$

Short derivation

Quadratic penalty form (WRI):

$$\phi(\mathbf{m},\mathbf{u},\lambda) = \frac{1}{2}\|\mathbf{P}\mathbf{u}-\mathbf{d}\|_2^2 + \frac{\lambda^2}{2}\|\mathbf{H}(\mathbf{m})\mathbf{u}-\mathbf{q}\|_2^2$$

Short derivation

Quadratic penalty form (WRI):

$$\phi(\mathbf{m},\mathbf{u},\lambda) = \frac{1}{2}\|\mathbf{P}\mathbf{u}-\mathbf{d}\|_2^2 + \frac{\lambda^2}{2}\|\mathbf{H}(\mathbf{m})\mathbf{u}-\mathbf{q}\|_2^2$$

$$\begin{pmatrix}\mathbf{P}_1 & & & \\ & \mathbf{P}_2 & & \\ & & \ddots & \\ & & & \mathbf{P}_k\end{pmatrix}\begin{pmatrix}\mathbf{u}_1\\ \mathbf{u}_2\\ \vdots\\ \mathbf{u}_k\end{pmatrix} - \begin{pmatrix}\mathbf{d}_1\\ \mathbf{d}_2\\ \vdots\\ \mathbf{d}_k\end{pmatrix} \qquad \begin{pmatrix}\mathbf{H}_1 & & & \\ & \mathbf{H}_2 & & \\ & & \ddots & \\ & & & \mathbf{H}_k\end{pmatrix}\begin{pmatrix}\mathbf{u}_1\\ \mathbf{u}_2\\ \vdots\\ \mathbf{u}_k\end{pmatrix} - \begin{pmatrix}\mathbf{q}_1\\ \mathbf{q}_2\\ \vdots\\ \mathbf{q}_k\end{pmatrix}$$

Short derivation

SLIM

Newton's method

Hessian

gradient

$$\begin{pmatrix} \mathbf{P}^*\mathbf{P} + \lambda^2\mathbf{H}^*\mathbf{H} & \nabla^2_{\mathbf{u},\mathbf{m}}\phi \\ \nabla^2_{\mathbf{m},\mathbf{u}}\phi & \lambda^2\mathbf{G}^*_{\mathbf{m}}\mathbf{G}_{\mathbf{m}} \end{pmatrix} \begin{pmatrix} \delta_{\mathbf{u}} \\ \delta_{\mathbf{m}} \end{pmatrix} = - \begin{pmatrix} \mathbf{P}^*(\mathbf{Pu} - \mathbf{d}) + \lambda^2\mathbf{H}^*(\mathbf{Hu} - \mathbf{q}) \\ \lambda^2\mathbf{G}^*_{\mathbf{m}}(\mathbf{Hu} - \mathbf{q}) \end{pmatrix}$$

updates for medium parameters

updates for all fields

Short derivation

Newton's method:

$$\begin{pmatrix} \mathbf{P}^*\mathbf{P} + \lambda^2\mathbf{H}^*\mathbf{H} & \nabla^2_{\mathbf{u},\mathbf{m}}\phi \\ \nabla^2_{\mathbf{m},\mathbf{u}}\phi & \lambda^2\mathbf{G}^*_{\mathbf{m}}\mathbf{G}_{\mathbf{m}} \end{pmatrix} \begin{pmatrix} \delta_{\mathbf{u}} \\ \delta_{\mathbf{m}} \end{pmatrix} = -\begin{pmatrix} \mathbf{P}^*(\mathbf{P}\mathbf{u} - \mathbf{d}) + \lambda^2\mathbf{H}^*(\mathbf{H}\mathbf{u} - \mathbf{q}) \\ \lambda^2\mathbf{G}^*_{\mathbf{m}}(\mathbf{H}\mathbf{u} - \mathbf{q}) \end{pmatrix}$$

Short derivation

Approximate Hessian:

$$\begin{pmatrix} \mathbf{P}^*\mathbf{P} + \lambda^2\mathbf{H}^*\mathbf{H} & 0 \\ 0 & \lambda^2\mathbf{G}_{\mathbf{m}}^*\mathbf{G}_{\mathbf{m}} \end{pmatrix} \begin{pmatrix} \delta_{\mathbf{u}} \\ \delta_{\mathbf{m}} \end{pmatrix} = -\begin{pmatrix} \mathbf{P}^*(\mathbf{P}\mathbf{u} - \mathbf{d}) + \lambda^2\mathbf{H}^*(\mathbf{H}\mathbf{u} - \mathbf{q}) \\ \lambda^2\mathbf{G}_{\mathbf{m}}^*(\mathbf{H}\mathbf{u} - \mathbf{q}) \end{pmatrix}$$

Short derivation

Approximate Hessian:

$$\begin{pmatrix} \mathbf{P}^*\mathbf{P} + \lambda^2\mathbf{H}^*\mathbf{H} & 0 \\ 0 & \lambda^2\mathbf{G}_{\mathbf{m}}^*\mathbf{G}_{\mathbf{m}} \end{pmatrix} \begin{pmatrix} \delta_{\mathbf{u}} \\ \delta_{\mathbf{m}} \end{pmatrix} = - \begin{pmatrix} \mathbf{P}^*(\mathbf{Pu} - \mathbf{d}) + \lambda^2\mathbf{H}^*\big(\mathbf{Hu} - \mathbf{q}\big) \\ \lambda^2\mathbf{G}_{\mathbf{m}}^*\big(\mathbf{Hu} - \mathbf{q}\big) \end{pmatrix}$$

Can be solved inexactly (cheap)!

Algorithm

0. construct initial guess $\mathbf{m}$ for medium and $\mathbf{u}_i$ for each field

while not converged **do**

1. form Hessian and gradient `// form (~free)`
2. ignore the $\nabla^2_{\mathbf{u},\mathbf{m}}\phi, \nabla^2_{\mathbf{m},\mathbf{u}}\phi$ blocks `// approximate`
3. find $\delta\mathbf{m}$ & each $\delta\mathbf{u}_i$ in parallel `// solve`
4. find steplength α using linesearch `// evaluate (~free)`
5. $\mathbf{m} = \mathbf{m} + \alpha\delta\mathbf{m}$ & $\mathbf{u} = \mathbf{u} + \alpha\delta\mathbf{u}$ `// update model and fields`

end

Algorithm

0. construct initial guess $\mathbf{m}$ for medium and $\mathbf{u}_i$ for each field

while not converged **do**

1. form Hessian and gradient `// form (~free)`
2. ignore the $\nabla^2_{\mathbf{u},\mathbf{m}}\phi, \nabla^2_{\mathbf{m},\mathbf{u}}\phi$ blocks `// approximate`
3. find $\delta\mathbf{m}$ & each $\delta\mathbf{u}_i$ in parallel `// solve` ← medium and field updates are independent
4. find steplength α using linesearch `// evaluate (~free)`
5. $\mathbf{m} = \mathbf{m} + \alpha\delta\mathbf{m}$ & $\mathbf{u} = \mathbf{u} + \alpha\delta\mathbf{u}$ `// update model and fields`

end

Algorithm

0. construct initial guess $\mathbf{m}$ for medium and $\mathbf{u}_i$ for each field

while not converged **do**

depend on the updated model and updated fields

1. form Hessian and gradient `// form (~free)`
2. ignore the $\nabla^2_{\mathbf{u},\mathbf{m}}\phi, \nabla^2_{\mathbf{m},\mathbf{u}}\phi$ blocks `// approximate`
3. find $\delta\mathbf{m}$ & each $\delta\mathbf{u}_i$ in parallel `// solve`
4. find steplength α using linesearch `// evaluate (~free)`
5. $\mathbf{m} = \mathbf{m} + \alpha\delta\mathbf{m}$ & $\mathbf{u} = \mathbf{u} + \alpha\delta\mathbf{u}$ `// update model and fields`

end

Full-space vs reduced-space methods

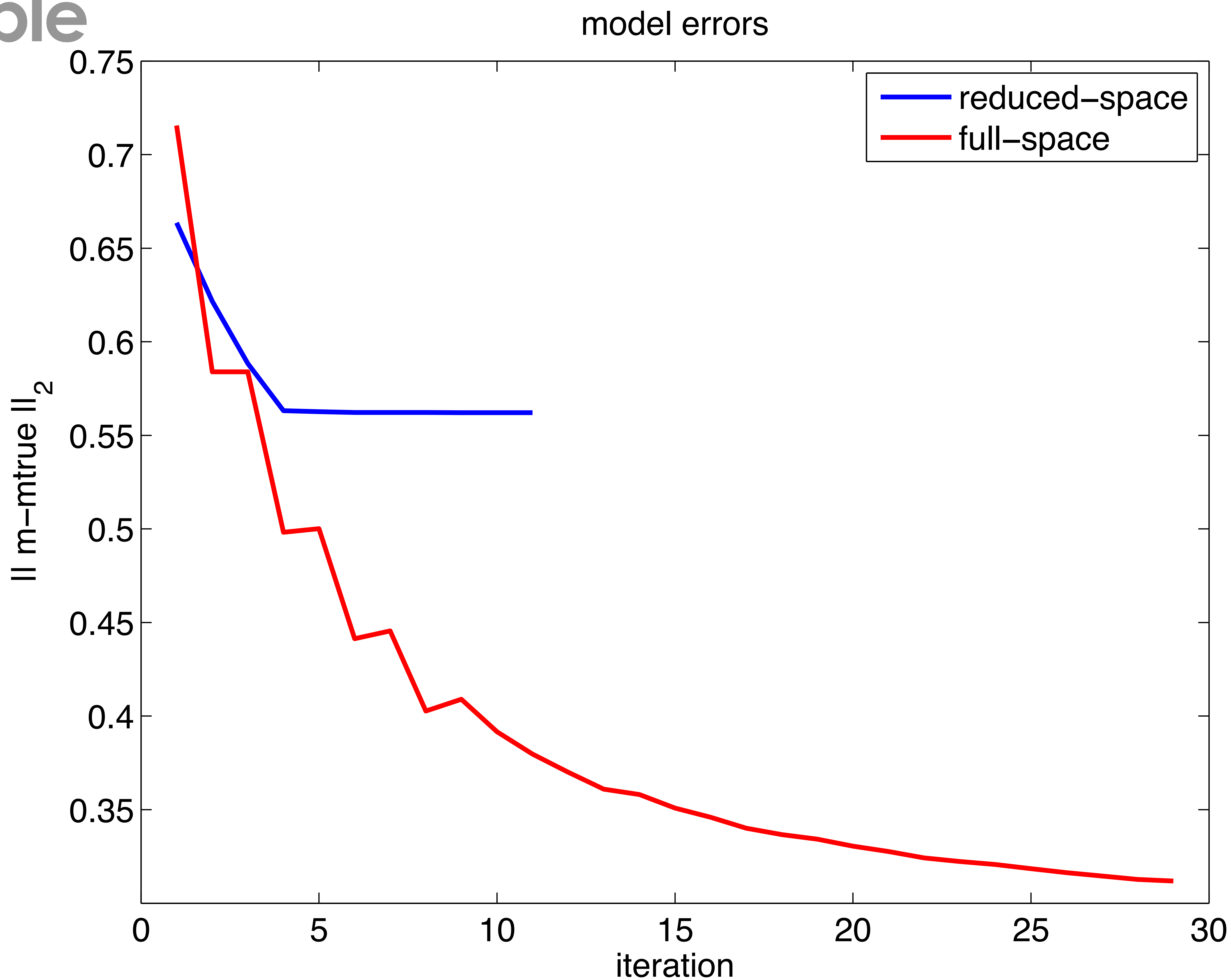
	FWI & WRI	full
Hessian	dense	sparse
Hessian	solve “PDE’s”	~free
gradient	solve “PDE’s”	~free
memory	2 fields per parallel process	all fields in memory (can be distributed over nodes)
function evaluation	solve “PDE’s”	~free

~free = sparse matrix-vector products

Previous example

model errors,

inaccurate iterative solutions



Inexact full-space vs inexact reduced-space

FWI & WRI:

- error in objective function value
- error in gradient
- error in Hessian

(error in gradient, error in Hessian) → error in medium parameter update

Inexact full-space vs inexact reduced-space

FWI & WRI:

- error in objective function value
- error in gradient
- error in Hessian

→ error in medium parameter update

Full-space from WRI:

- objective function value always exact
- gradient always exact
- Hessian always exact

→ 0 iterations → gradient descent
many iterations → Newton's method

Toy examples

Using a direct solver:

- similar reconstruction quality compared to WRI+diagonal Hessian approximation
- need to test on more realistic models.

Memory requirements

save all fields for all frequencies & sources

can be distributed over multiple nodes

Feasible? Need

- parallel computing
- simultaneous sources
- small frequency batches

Computational cost

Independent update computation

No communication between compute nodes to compute updates

1 iteration of WRI ≈ 1 iteration of full-space Newton type quadratic penalty

Conclusions

Constructed a full-space method which:

- updates fields & medium parameters simultaneously
- computational cost ≈ reduced-space methods
- similar parallelism as in FWI & WRI
- many properties are different from FWI & WRI
- promising results with iterative solvers

- con: need to store all fields
- but, less storage needed compared to Lagrangian full-space methods

Current & future work

Test on more realistic examples.
Evaluate reconstruction quality compared to WRI.
Maximize benefit from inexact update computation.

Acknowledgements

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, CGG, Chevron, ConocoPhillips, ION, Petrobras, PGS, Statoil, Total SA, Sub Salt Solutions, WesternGeco, and Woodside.