Multigrid Inversion Algorithms with Applications to Optical Diffusion Tomography

Seungseok Oh, Adam B. Milstein, Charles A. Bouman, and Kevin J. Webb

School of Electrical and Computer Engineering

Purdue University

Supported by the National Science Foundation

Inverse Problems

- Forward model

$$y = f(x) + noise$$

- Inverse problem: Determine X from Y
- Applications include: image restoration, tomography, remote sensing, machine vision

- Computation of $f(x)$ can be very difficult
- Inversion of $f(x)$ can be more difficult
  - Need to search for x which solves equation
  - Can be formulated as optimization problem
  - Optimization may have local minima
  - Convergence may be slow

Our Approach: Multigrid Inversion

- Formulate a series of inverse problems at different scales

$$y^{(k)} = f^{(k)}(x^{(k)}) + noise$$

- Move between scales to solve problem
  - Coarse-to-fine: Fine is more accurate
  - Fine-to-coarse: Coarse is less accurate!
  - Cost functionals are not consistent
  - Dynamically adjust cost functionals for consistency

- Advantages:
  - Designed for *nonlinear* inverse problems
  - Both inverse and *forward* model scales change
  - Coarse scale iterations can be applied at any time
  - Rapid and robust convergence

Gaussian measurement model

- We use a Gaussian measurement model

$$\log p(y|x) = -\frac{1}{2\alpha}||y - f(x)||_\Lambda^2 - \frac{P}{2}\log(2\pi\alpha|\Lambda|^{-1})$$

where

x : unknown image

y : measurement

$f(x)$: forward model

α : measurement noise factor (assumed unknown)

Λ : measurement covariance

P : number of (real valued) dimensions to y

Regularized Inverse

- Joint MAP estimation of x and α yields

$$\begin{aligned}\hat{x} &= \arg\min_x \min_\alpha \{-\log p(y|x,\alpha) + S(x)\} \\ &= \arg\min_x \min_\alpha \left\{ \frac{1}{2\alpha} ||y - f(x)||_\Lambda^2 + \frac{P}{2} \log(2\pi\alpha|\Lambda|^{-1}) + S(x) \right\} \\ &= \arg\min_x \left\{ \frac{P}{2} \log ||y - f(x)||_\Lambda^2 + S(x) \right\}\end{aligned}$$

where $S(x) = -\log p(x)$ is a stabilizing functional

- Estimation of α makes convergence more robust!

The Optimization Problem

- Function to be minimized is

$$c(x; y) = \frac{P}{2} \log ||y - f(x)||_{\Lambda}^{2} + S(x)$$

  - Forward model may be difficult to compute
  - For nonlinear problems $c(x)$ is generally not convex
  - Cost function may have local minima
  - Fixed grid optimization can be slow

Fixed-grid Optimization

$$x_{update} \leftarrow \text{Fixed\_Grid\_Update}(x_{init}, c(\cdot; y))$$

where

$$c(x; y) = \frac{P}{2} \log ||y - f(x)||^2_{\Lambda} + S(x)$$

- Shortcomings
  - All operations are performed at the finest scale
  - Forward model is always evaluated at the finest scale
  - Convergence speed depends on spectral characteristics of error
  - Very sensitive to initial condition
  - Tends to become trapped in local minima

Multigrid Cost Functionals

- Cost functional at scale q

$$c^{(q)}(x^{(q)}; y^{(q)}, r^{(q)}) = \frac{P}{2} \log ||y^{(q)} - f^{(q)}(x^{(q)})||_\Lambda^2 + S^{(q)}(x^{(q)}) - r^{(q)} x^{(q)}$$

$f^{(q)}(\cdot)$ - coarse scale **forward** model

$x^{(q)}$ - coarse scale solution

$S^{(q)}(\cdot)$ - coarse scale stabilizing functional

$y^{(q)}$ - coarse scale measurement

$r^{(q)}$ - adjustment factor at scale q

Coarse Scale Correction

Fixed grid update

$$x^{(q)} \leftarrow \text{Fixed\_Grid\_Update}(x^{(q)}, c^{(q)}(\,\cdot\,; y^{(q)}, r^{(q)}))$$

Decimate result

$$x_{init}^{(q+1)} \leftarrow I_{(q)}^{(q+1)} x^{(q)}$$

Compute $y^{(q+1)}$ (... But how?)

Compute $r^{(q+1)}$ (... But how?)

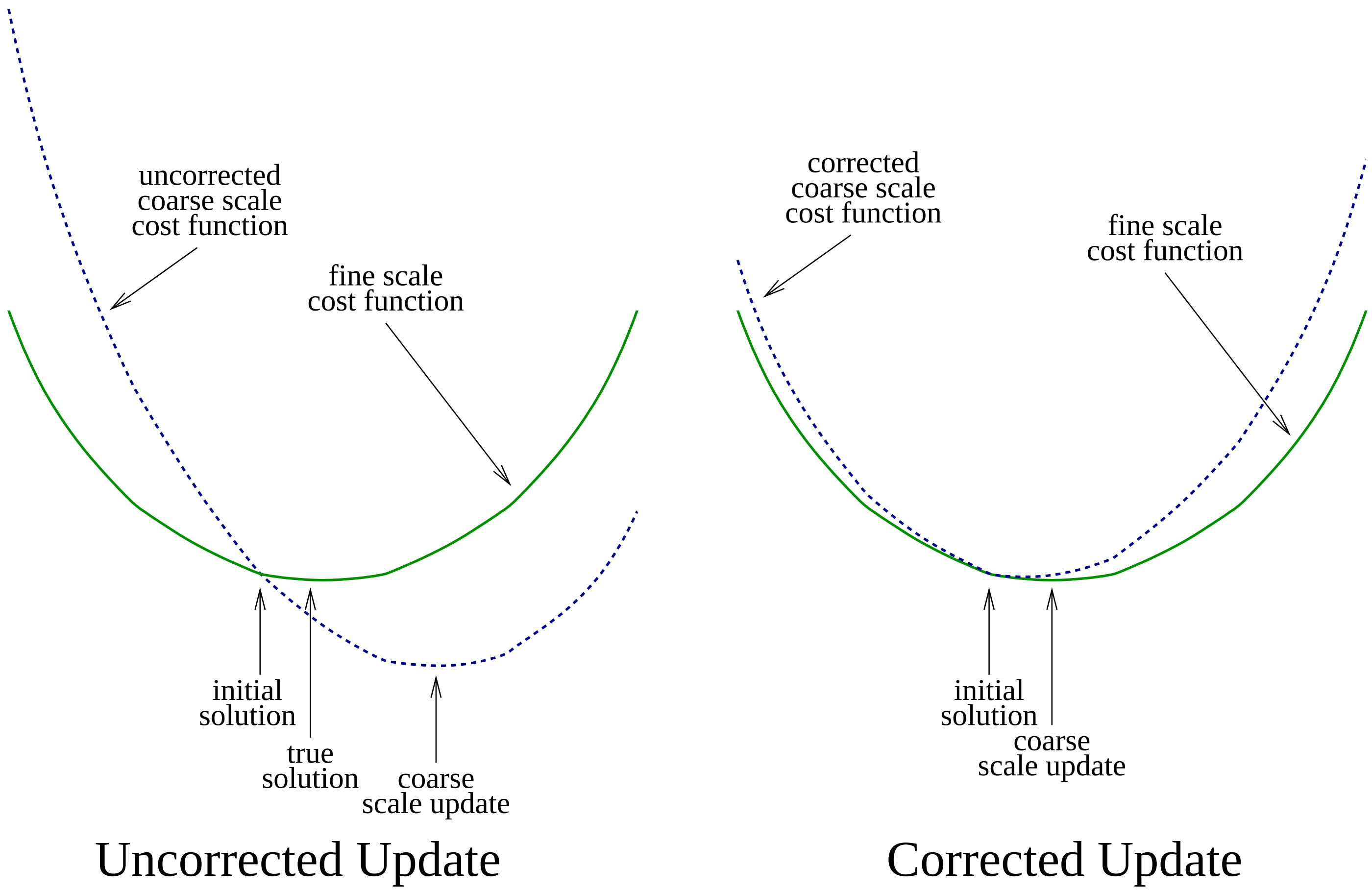
Coarse grid update

$$x_{update}^{(q+1)} \leftarrow \text{Fixed\_Grid\_Update}(x_{init}^{(q+1)}, c^{(q+1)}(\,\cdot\,; y^{(q+1)}, r^{(q+1)}))$$

Interpolate correction

$$x^{(q)} \leftarrow x^{(q)} + I_{(q+1)}^{(q)}(x_{update}^{(q+1)} - x_{init}^{(q+1)})$$

Choosing Consistent Cost Functionals



- Coarse scale cost should:
  - Upper bound fine scale cost functional
  - Be tangent to fine scale cost functional at initial solution

Choosing $y^{(q+1)}$ and $r^{(q+1)}$

- This is VERY important

- Match error in data term at coarse and fine scales

$$y^{(q+1)} \quad \leftarrow \quad y^{(q)} - \left[f^{(q)}(x^{(q)}) - f^{(q+1)}(I_{(q)}^{(q+1)} x^{(q)}) \right]$$

- Match derivatives in cost function at coarse and fine scales

$$\begin{aligned} r^{(q+1)} \quad \leftarrow \quad & \nabla \tilde{c}^{(q+1)}(x^{(q+1)}) \Big|_{x^{(q+1)} = I_{(q)}^{(q+1)} x^{(q)}} \\ & - \left(\nabla \tilde{c}^{(q)}(x^{(q)}) - r^{(q)} \right) I_{(q+1)}^{(q)} \end{aligned}$$

- Theorem: If the difference between cost functionals is convex, then multigrid iterations generate monotone decreasing cost.

Stabilizing Functional

- We need would like

$$S^{(q)}(x^{(q)}) \stackrel{\sim}{=} S(x) \ .$$

- For the generalized Gaussian Markov random field model

$$S^{(q)}(x^{(q)}) = \frac{1}{p(\sigma^{(q)})^p} \sum_{\{i,j\} \in \mathcal{N}} b_{i-j} \left| x_i^{(q)} - x_j^{(q)} \right|^p$$

where $\sigma^{(q)} = 2^{q(1-\frac{3}{p})} \cdot \sigma^{(0)}$

Application: Optical Diffusion Tomography (ODT)

- Measure light that passes through a highly scattering medium
- Determine unknown absorption and/or diffusion cross-section of material
- Obeys the frequency-domain diffusion equation:

$$\nabla \cdot [D(r)\nabla\phi_k(r)] + [-\mu_a(r) - j\omega/c]\,\phi_k(r) = -\delta(r - a_k)$$

- Nonlinear forward model $E[y] = f(x)$:

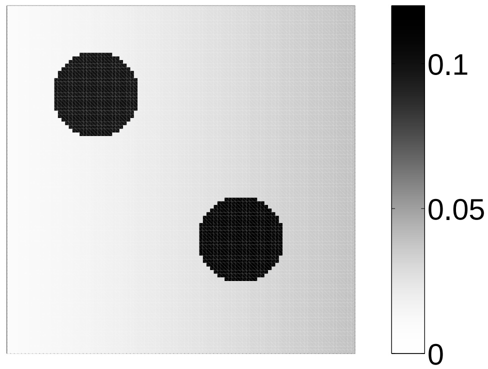
  y : complex measurement of $\phi_k(r)$

  x : image of unknown absorption $\mu_a(r)$ and diffusion $D(r)$

PURDUE UNIVERSITY

Simulation Experiment

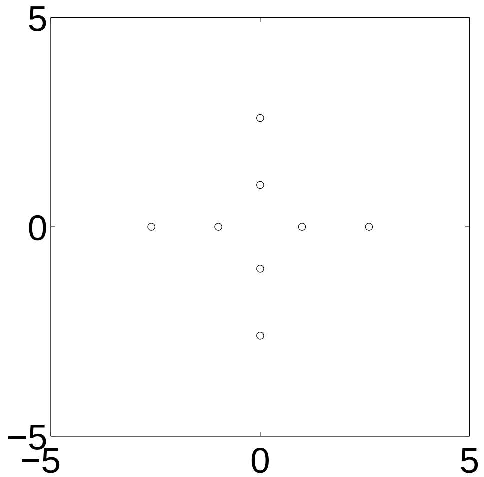
- Phantom



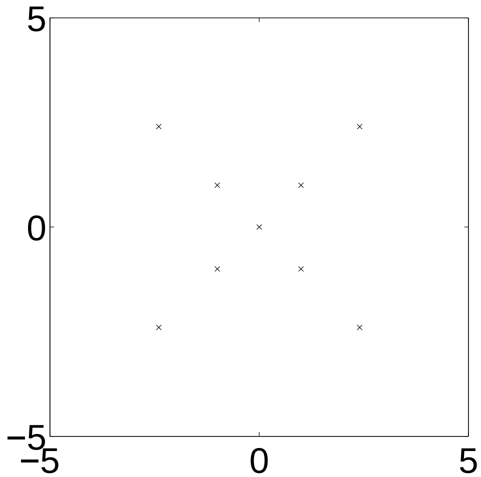
  - 10cm $\times$ 10cm $\times$ 10cm cube
  - Linearly varying background from $\mu = 0.01\text{cm}^{-1}$ to 0.04cm^{-1}
  - Two spherical inhomogeneities with diameters of 1.85cm and densities of $\mu = 0.10\text{cm}^{-1}$ and $\mu = 0.12\text{cm}^{-1}$
  - Diffusion coefficient, D, is constant

- Model
  - 100MHz modulation frequency and 35dB average SNR
  - GGMRF with $p = 1.2$ and $\sigma = 0.018\text{cm}^{-1}$

Source/Detector Positions



Sources

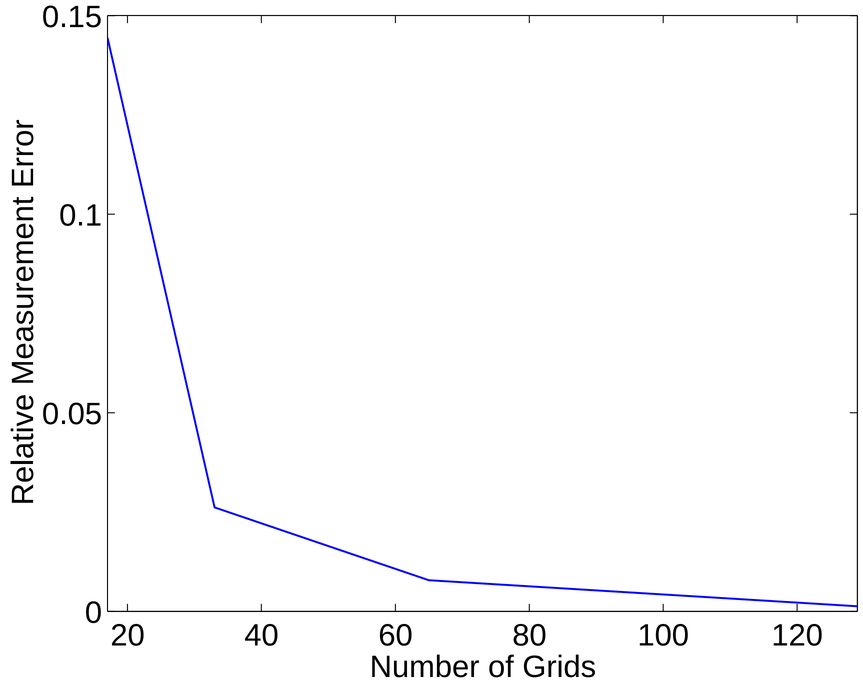


Detectors

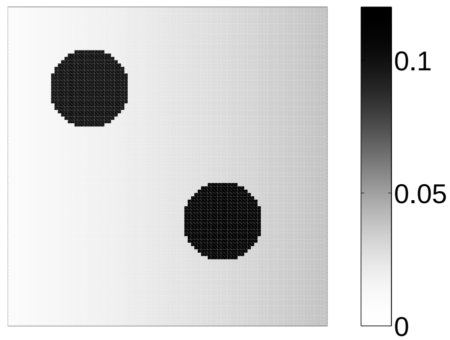
- Sources and detectors on all 6 faces of cube using 100MHz modulation frequency and 35dB average SNR
- **Important:** Source/detector pairs on same face were not used in order to reduce discretization error.

How Much Resolution Do We Need?

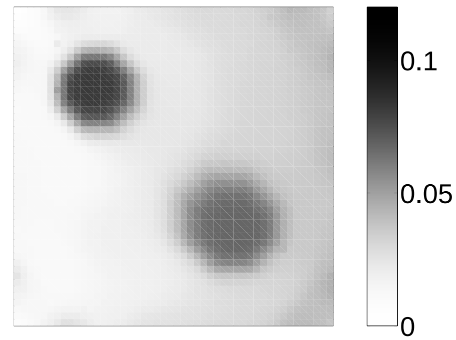
- Relative measurement error versus grid resolution
  - Based on $257 \times 257 \times 257$ reference simulation



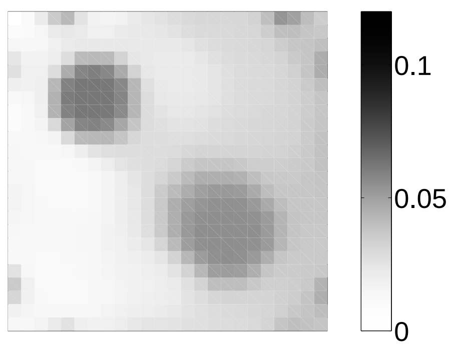
Best Reconstructions at Various Resolutions



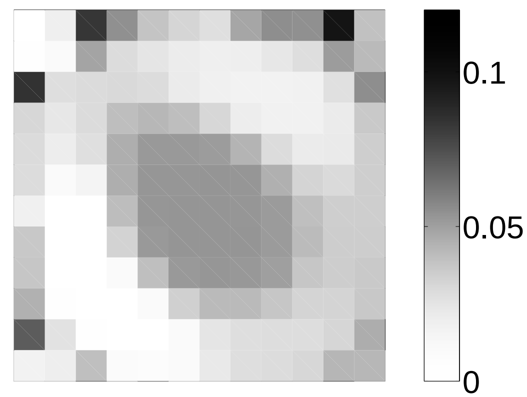
(a) True Phantom



(b) $65 \times 65 \times 65$ Recon.



(c) $32 \times 32 \times 32$ Recon.

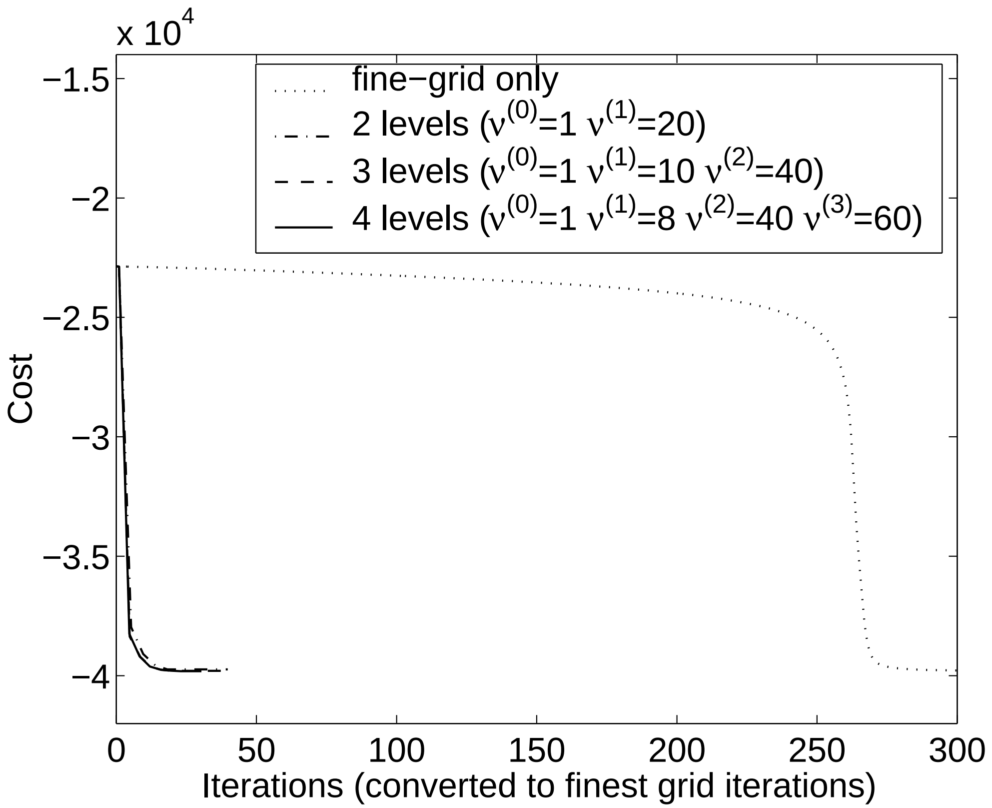


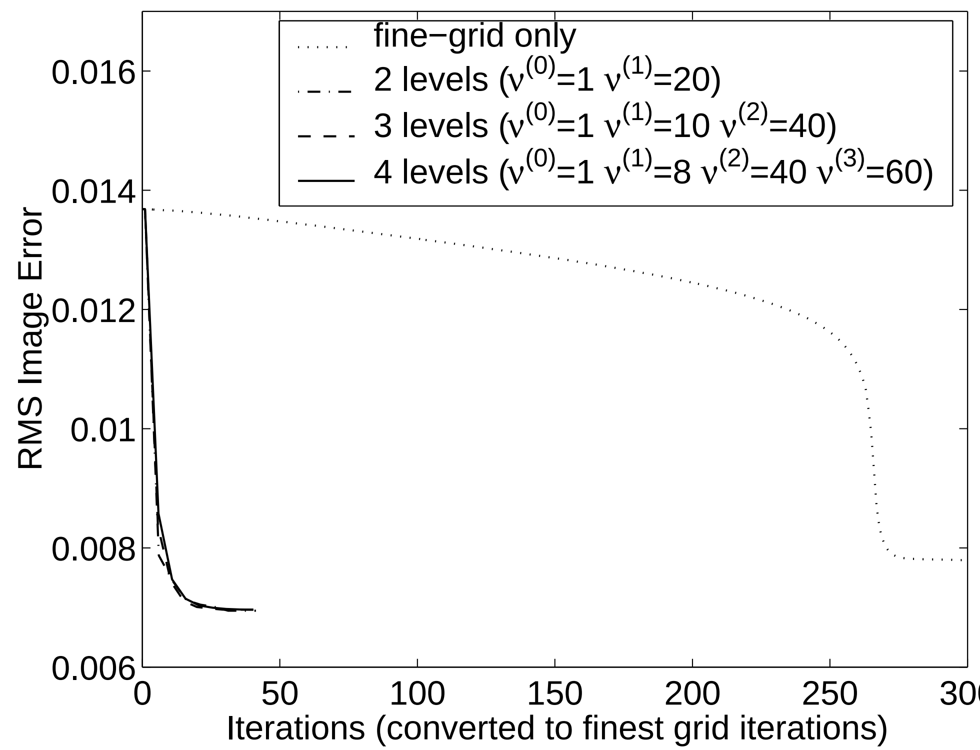
(d) $17 \times 17 \times 17$ Recon.

$\Rightarrow$ $65 \times 65 \times 65$ grid is required for accurate reconstruction

$\Rightarrow$ huge computation!

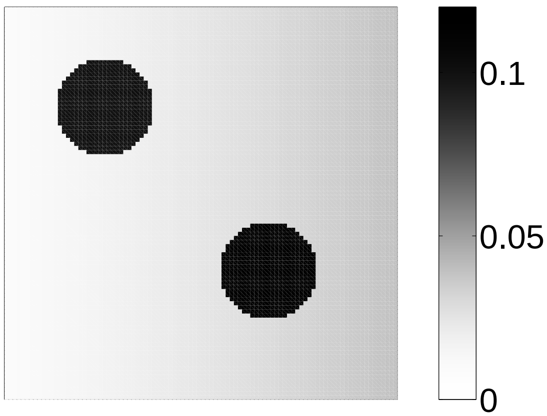
Convergence Speed for $65 \times 65 \times 65$ Reconstruction with Good Initial Condition



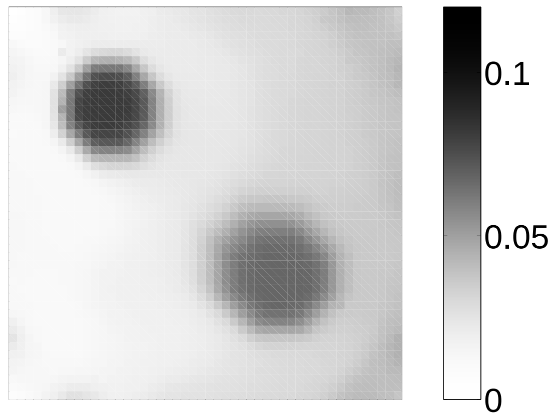


- All iterations in units of a single fixed grid iteration

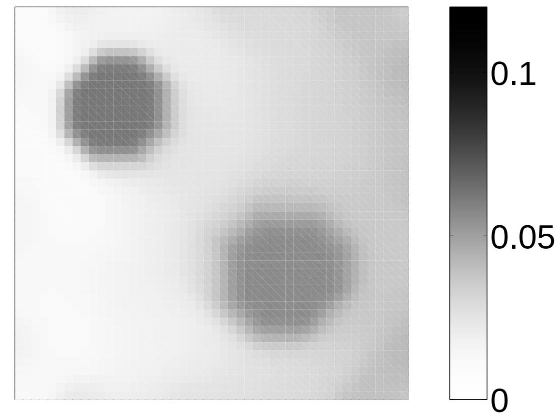
Reconstruction Quality for Multigrid and Fixed Grid Algorithms



True phantom

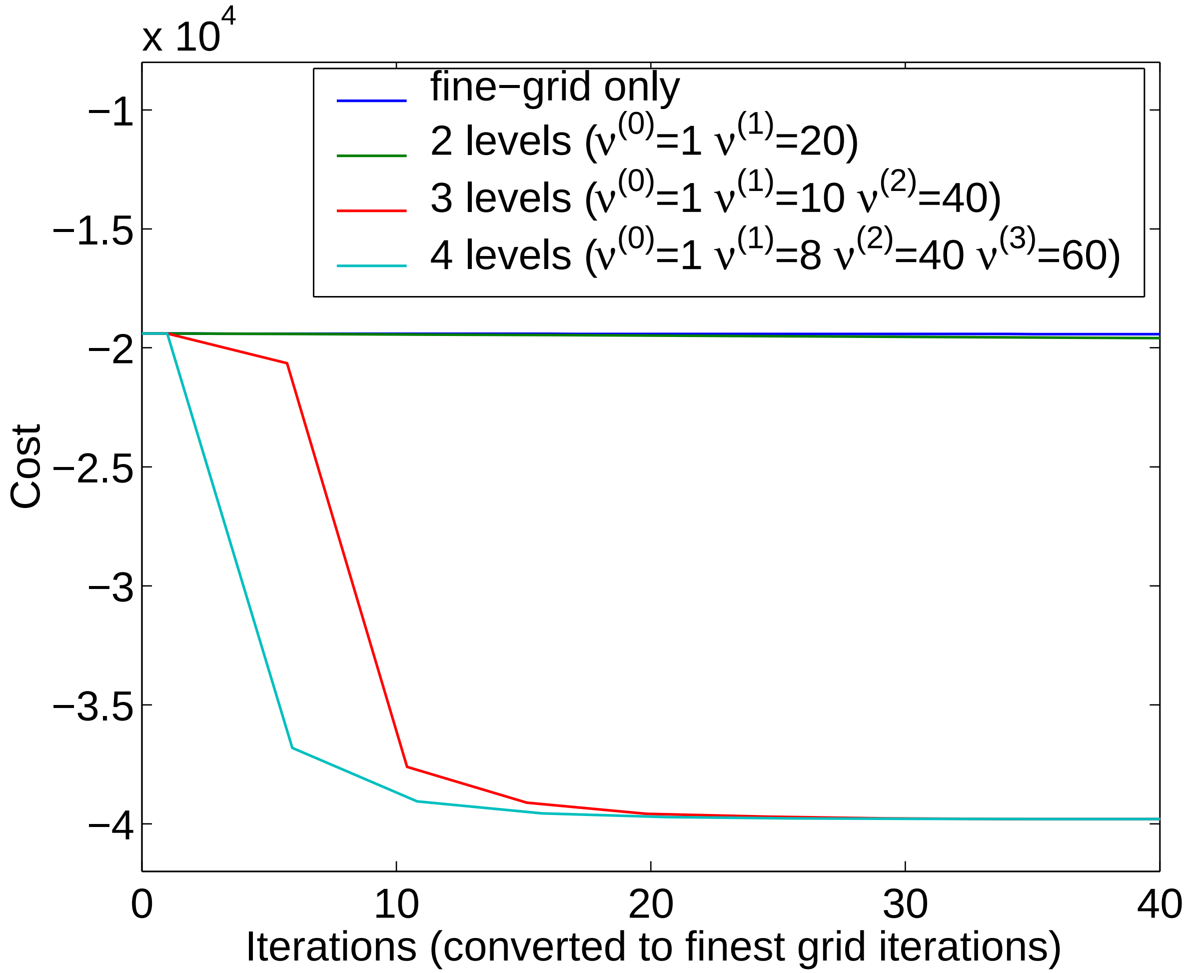


4 level multigrid
(25.6 iter.)



Fixed-grid
(300 iter.)

Convergence Speed for $65 \times 65 \times 65$ Reconstruction with Poor Initial Condition



Concluding Remarks

- About Inverse Problems
  - Nonlinear inverse problems are difficult and of growing importance
  - Grid resolution can be important for both forward and inverse problem
- About Multigrid Inversion
  - Fast and robust convergence
  - Insensitive to initial condition
  - Widely applicable
  - Changes grid resolution for both forward and inverse problems
  - Very stable convergence