

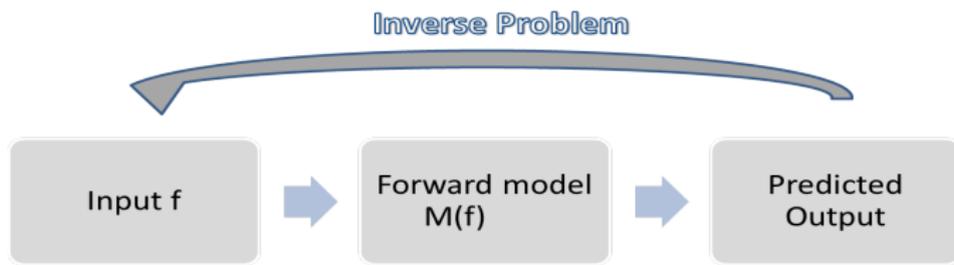
# Ingredients for Computationally Efficient Solution of Large-scale Image Reconstruction Problems

Misha Kilmer  
*Tufts University*

SIAM CS&E 2017, Atlanta, GA

Partially supported by NSF-DMS 1217156, 1217161 and NIH R01-CA 154774

# Motivation



The (unknown)  $\mathbf{f}$  is a (vectorized) 2D or 3D image of, for example:

- pixel intensity
- hydraulic or electric conductivity
- electrical resistivity
- Compton scatter coefficient
- photoelectric absorption
- optical absorption

and  $M$  represents the appropriate (linear or nonlinear) model map.

# Motivation

Forward problem:  $\mathbf{m} = M(\mathbf{f}) + \text{unknown noise}$ .

**Inverse problem:** recovery of vectorized image of the unknown  $\mathbf{f} \in \mathbb{R}^{m \times n(\times k)}$  given the model  $M$  and noisy data  $\mathbf{m}$ .

$$\min_{\mathbf{f} \in \mathbb{R}^{m \times n \times k}} \|\mathbf{m} - M(\mathbf{f})\|^2 + \text{regularization on } \mathbf{f}; + \text{constraints on } \mathbf{f}$$

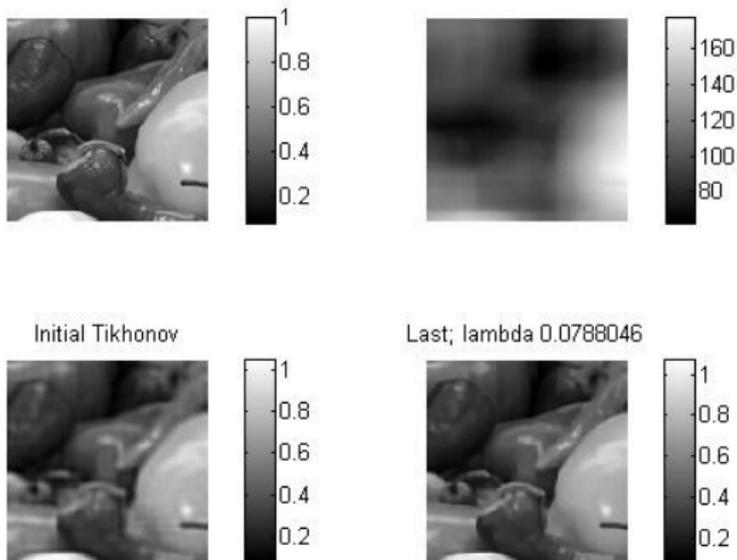
Regularization, often quite sophisticated, used to

- Damp effects noise in the data
- Deals with the lack of, or (near) redundancy in, data
- Enforce a priori knowledge about the solution

Evaluations  $M(\mathbf{f})$  often present a **huge computational bottleneck** for non-linear  $M$ , and are non-trivial for many linear models as well.

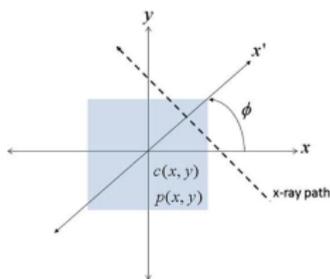
# Motivation - Examples - Linear

Deblurring<sup>1</sup>,  $M(\mathbf{f}) = \mathbf{A}\mathbf{f}$  from discretized convolutional model.



<sup>1</sup>Semerci, K., Miller, "An Adaptive Inner-Outer Iterative Regularization Method for Edge Recovery", ICIAM presentation 2015.

# Motivation - Examples - Linear



Recovering energy-dependent attenuation coefficients<sup>2</sup>

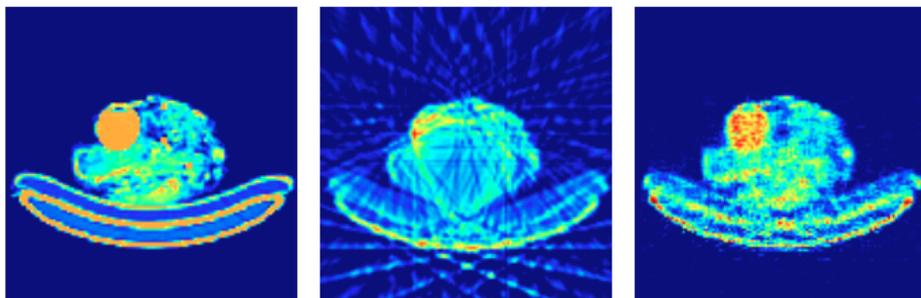
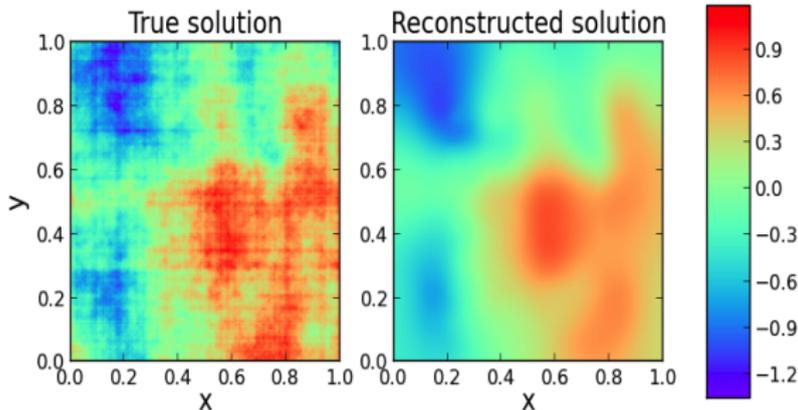
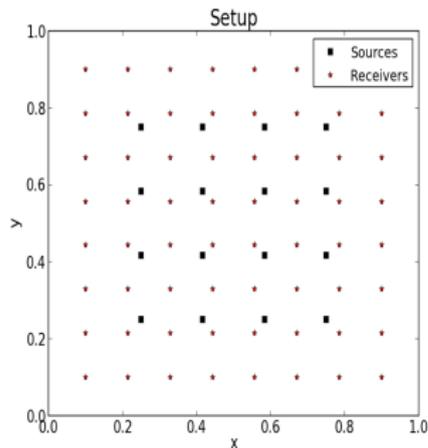


Figure: Left: 85keV phantom; Middle: FBP, Right: with TV regularization

<sup>2</sup>Semerçi, Hao, K., Miller IEEE TIP, 2014

# Motivation – Examples – Nonlinear

Recovery of hydraulic conductivity. Evaluation of  $M(\mathbf{p})$  requires multiple PDE solves and limited data.



<sup>3</sup>Images from Saibaba, Bakhos, Kitanidis, SISC, 2013

# Motivation - Examples - Nonlinear

Invert for saturation, combining ERT and hydrologic measurement data<sup>4</sup>.

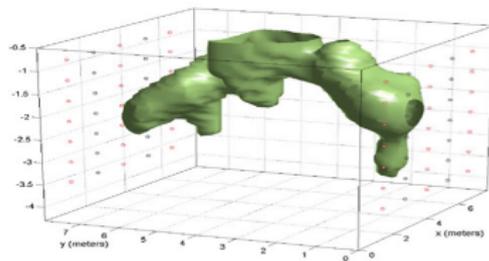
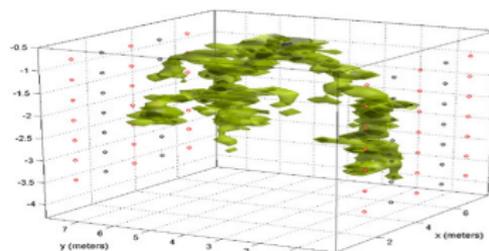
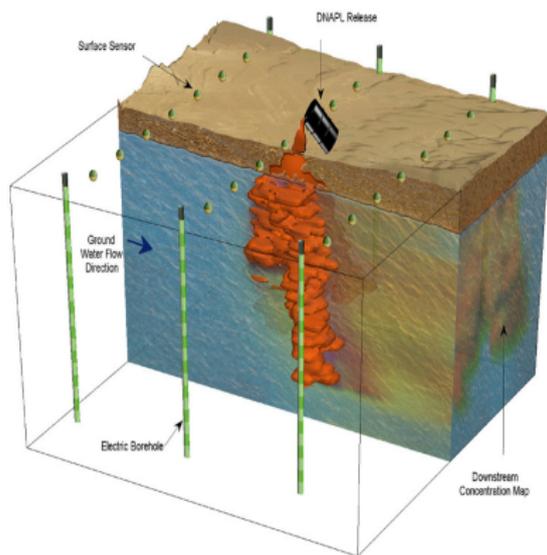
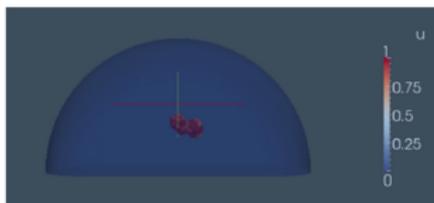
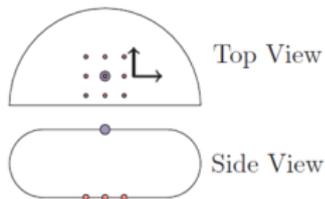
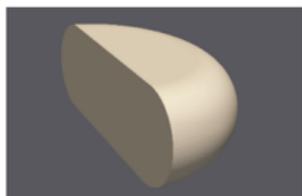


Figure: Left: Model, with source/detector locations; Right top: truth; Bottom: a PaLS joint reconstruction.

<sup>4</sup>Aghasi, et. al., Inverse Problems 29 (2013).

## Motivation - Examples - Nonlinear

Invert for absorption coefficient in breast tissue, given limited, discrete measurements of photon flux, possibly at multiple wavelengths<sup>5</sup>.



<sup>5</sup>Images from Saibaba, et al., SISC 2015.

# General Inverse Problem Formulation

Recover vectorized image of the unknown  $\mathbf{f} \in \mathbb{R}^{m \times n (\times k)}$

$$\min_{\mathbf{f} \in \mathbb{R}^{m \times n \times k}} \|\mathbf{m} - M(\mathbf{f})\|^2 + \lambda^2 \Gamma(\mathbf{f})$$

## Practical Considerations

- Often data limited; voxel based inversion realistic?
- Choice of regularization term(s)
- Choice of parameter(s)
- Evaluation of  $M(\mathbf{f}_k)$  and Jacobian during optimization may require solution of many, large scale forward model solves (discretized PDEs), so **a single optimization step is expensive**

# Many Interesting Subproblems for the CS&E Community

Subproblems requiring expertise across CS&E disciplines:

- Image representation
  - ▶ A new basis or learn a dictionary
  - ▶ Level Set, parametric level sets, or other shape based
- Regularization (e.g.)
  - ▶ Enforce sparsity (in right representation)
  - ▶ Geometric constraints or multiway data constraints (tensors)
  - ▶ Regularization parameter selection (learning)
- Approximate forward model
  - ▶ Krylov subspace methods, recycling, preconditioning
  - ▶ Multigrid
  - ▶ Reduced order models and randomized approaches
- Optimization
  - ▶ Sophisticated algorithms tailored to inverse problems
- Uncertainty quantification

# Many Interesting Subproblems for the CS&E Community

Subproblems requiring expertise across CS&E disciplines:

- Image representation
  - ▶ A new basis or **learn a dictionary**
  - ▶ Level Set, **parametric level sets**, or other shape based
- Regularization (e.g.)
  - ▶ Enforce sparsity (in right representation)
  - ▶ Geometric constraints or multiway data constraints (tensors)
  - ▶ Regularization parameter selection (learning)
- Approximate forward model
  - ▶ Krylov subspace methods, recycling, preconditioning
  - ▶ Multigrid
  - ▶ **Reduced order models and randomized approaches**
- Optimization
  - ▶ Sophisticated algorithms tailored to inverse problems
- Uncertainty quantification

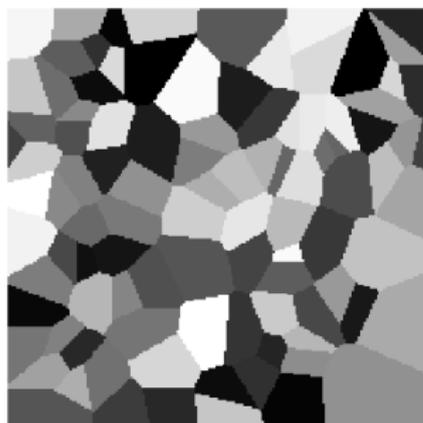
# Main Ingredients

Multi-pronged approach to solving reconstruction problems:

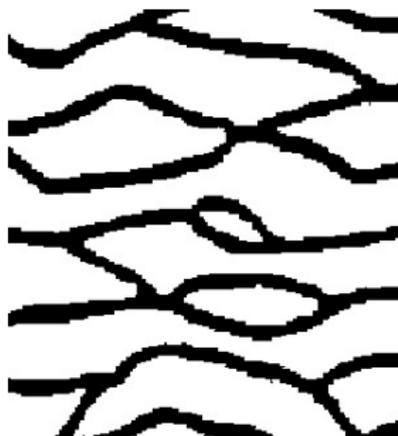
- 1 Determine a “natural” **model for the image space** that provides robustness to noise. Two options:
  - ▶ Find dictionary  $W$ ,  $\mathbf{f} = W\mathbf{p}$ , solve for  $\mathbf{p}$
  - ▶ Low order parameterization of image space,  $\mathbf{f} = \mathbf{f}(\mathbf{p})$ , solve for  $\mathbf{p}$ .
- 2 Determine **surrogate forward operator or model** so function/Jacobian evaluations becomes less costly. Two options:
  - ▶ Randomized techniques
  - ▶ Reduced order modeling

## Priors in the Form of Training Images

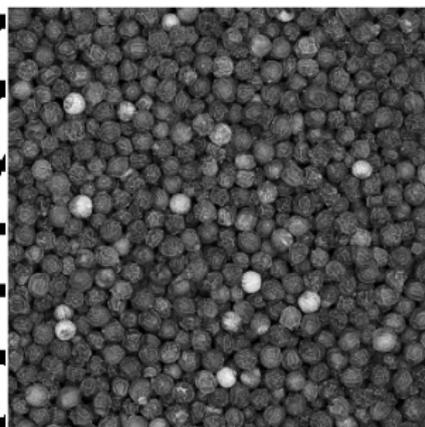
Assume prior information (feature-based) available from training images<sup>6</sup>.



Grains



Geostatistical



Peppers

---

<sup>6</sup>The grains simulation is by Post Doc. Jakob S. Jorgensen, DTU. The Geostatistical image is by Ph.D Knud S. Cordua. The Peppers photo is courtesy of Prof. Samuli Siltanen, University of Helsinki.

# Two-step Matrix-based Dictionary Learning Approach

- Generate  $t$ ,  $p \times r$  subimages from training image
- Vectorize ea. subimage, form matrix  $\mathbf{Y} \in \mathbb{R}^{pr \times t}$
- **Compute** a NMF  $\mathbf{Y} \approx \mathbf{D}\mathbf{H}$ , where  $\mathbf{D}$  has  $s \ll t$  columns.
- Construct “global matrix”  $\mathbf{W}$  from  $\mathbf{D}$ .
- Using  $\mathbf{f} \approx \mathbf{W}\mathbf{p}$ , **solve**

$$\min_{\mathbf{p} \geq 0} \|\mathbf{m} - M(\mathbf{W}\mathbf{p})\|_2^2 + \lambda R(\mathbf{f}) + \mu G(\mathbf{p})$$

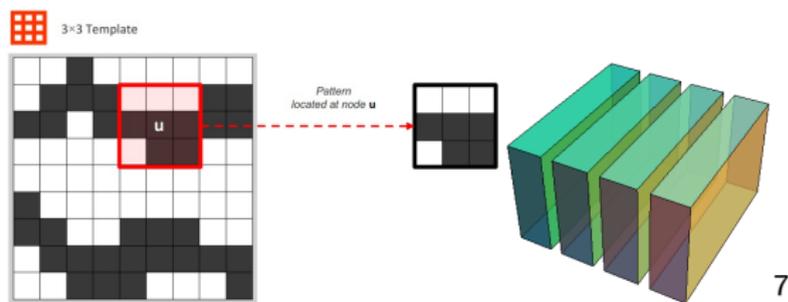
See, for example, Soltani, Andersen, and Hansen, “Tomographic image reconstruction using training images,” J. Comp. Appl. Math., 313 (2017) and references therein.

# Dictionary Image models

Forces the solution to be comprised of desired features

- Some regularization is built in to the image model; optimization now over coefficients
- Additional regularization (non-negativity, smoothness) also applied

Generalization to tensors: Soltani, K., Hansen, BIT, 2016.



---

<sup>7</sup> Left: M Honarkhah, "Stochastic Simulation of patterns using distance-based pattern modeling" Ph.D thesis, Stanford University, 2011. Right: Anika Rounds, M.S. Thesis, Tufts University, 2014.

## Tensor Patch Dictionary Results<sup>8</sup>

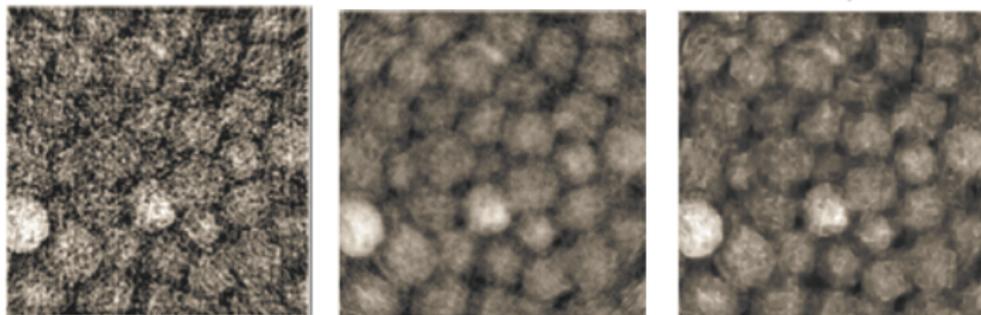


Figure: Left: FBP; Middle: Tikhonov; Right: Tensor Dictionary

---

<sup>8</sup>Soltani, K., Hansen, "A tensor-based dictionary learning approach to tomographic image reconstruction," BIT, 2016.

# Image Parameterization: Breast Tissue Imaging with DOT

Breast tissue made up of adipose, fibroglandular, tumor. **Inverting for every voxel value is unnecessary.**

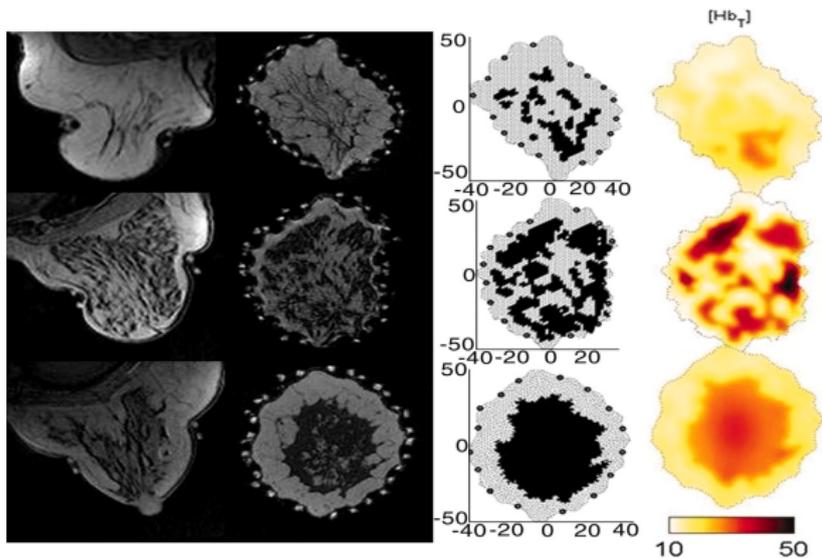


Figure: Ben Brooksby, et. al, PNAS 2006 103 (23) 8828-8833

# Shape-based Approach

Simpliest Model: Model unknown  $f(\mathbf{x})$  as **piecewise continuous**.

$$\chi_D(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in D \\ 0 & \mathbf{x} \in \Omega \setminus D. \end{cases}$$

In a continuous setting, the unknown property  $f(\mathbf{x})$  can be defined over  $\Omega$

$$f(\mathbf{x}) = f_i(\mathbf{x})\chi_D(\mathbf{x}) + f_o(\mathbf{x})(1 - \chi_D(\mathbf{x}))$$

**Goal:** find  $\partial D$  (and parameters defining  $f_i, f_o$ ).

Traditional level sets (e.g. Santosa '96), could be used to specify  $\partial D$ .  
Practical implementation highly non-trivial<sup>9</sup>.

---

<sup>9</sup>van den Doel et al, *J. of Sci. Comp.* 2010; van den Doel & Ascher, *SISC* 2012

## Parametric Level Sets (PaLS)<sup>10</sup>

Let  $\phi$  function of  $m$ -length parameter vector  $\mathbf{p}$  AND  $\mathbf{x}$ :

$$\phi(\mathbf{x}, \mathbf{p}) = \sum_{i=1}^{m_0} \alpha_i \psi_i(\mathbf{x})$$

$\mathbf{p}$  contains  $\alpha_i$  & parameters defining  $i^{\text{th}}$  function.

---

Let  $\psi : \mathbb{R}^+ \rightarrow \mathbb{R}$  denote a sufficiently smooth CSRBF.

$$\phi(\mathbf{x}, \mathbf{p}) := \sum_{j=1}^{m_0} \alpha_j \psi(\|\beta_j(\mathbf{x} - \boldsymbol{\chi}_j)\|^\dagger),$$

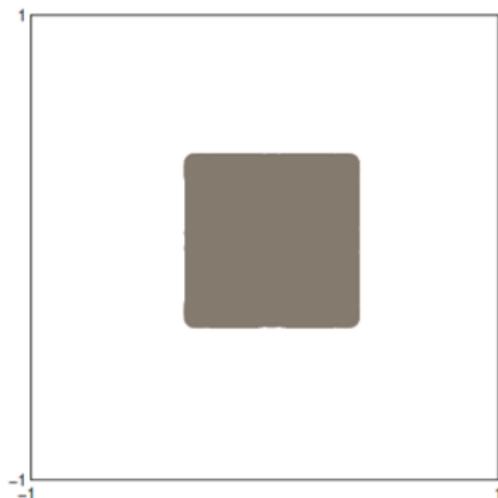
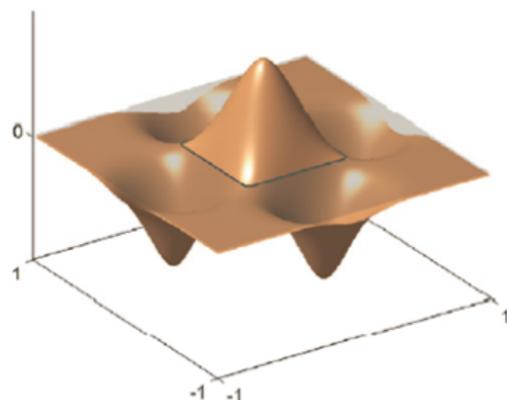
where the  $\boldsymbol{\chi}_j$  are the centers,  $\beta_j$  are dilation factors and

$$\|\mathbf{x}\|^\dagger := \sqrt{\|\mathbf{x}\|_2^2 + \nu^2}$$

Desired parameter vector:  $\mathbf{p} = \left[ \mathbf{a}^T \quad \mathbf{b}^T \quad \boldsymbol{\chi}_x^T \quad \boldsymbol{\chi}_y^T \right]^T$ .

<sup>10</sup>Aghasi, K., Miller, SIIMS, 2011.

# Choice of Functions



CSRBFs: Wendland, *Cambridge Univ. Press*, 2005.

One CSRBF:  $\psi(r) = (\max(0, 1 - r))^2 (2r + 1)$ ;  $r = \sqrt{x^2 + y^2}$ . For a list of other choices, related work, see Aghasi, et al, 2011.

## Optimization Revisited

$$f(\mathbf{x}, \mathbf{p}) = f_i(\mathbf{x})H_\epsilon(\phi(\mathbf{x}, \mathbf{p}) - c) + f_o(\mathbf{x})(1 - H_\epsilon(\phi(\mathbf{x}, \mathbf{p}) - c))$$

---

Assume  $f_i(\mathbf{x}) = f_i$  and  $f_o(\mathbf{x}) = f_o$  and  $\mathbf{f}(\mathbf{p})$  the discretization of  $f(\mathbf{x}, \mathbf{p})$ :

$$\min_{\mathbf{p}} \|\mathbf{m} - M(\mathbf{f}(\mathbf{p}))\|_2$$

Nonlinear LS problem of **relatively small dimension**.

**No additional regularization** except stopping criterion (discrepancy).

Use **TREGS**<sup>11</sup> to solve this, proven more efficient than LM or (modified, regularized) GN.

---

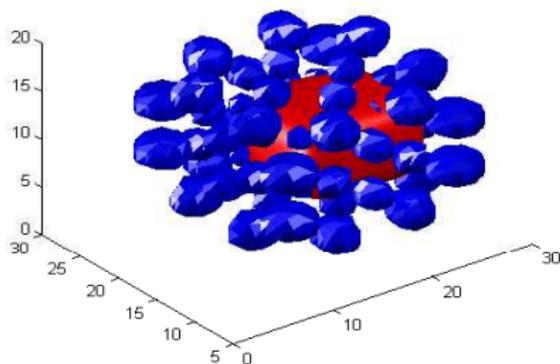
<sup>11</sup>de Sturler and K. "A Regularized Gauss-Newton Trust Region Approach to Imaging in Diffuse Optical Tomography, SISC, 2011

## Motivating Example: Diffuse Optical Tomography

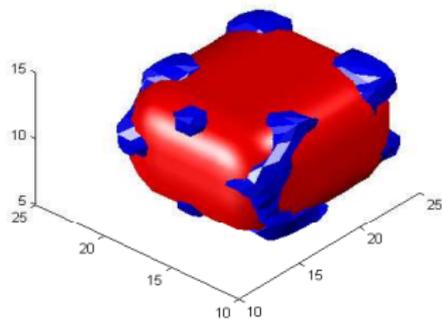
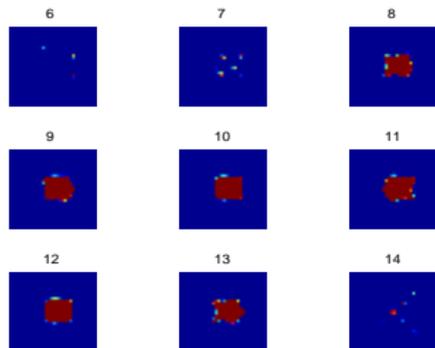
Imaging for **absorption coef** in tissue illuminated by near infrared light.  
Forward model - solve diffusion equation that models photon fluence/flux!

$15 \times 15$  array sources & detectors,  $32 \times 32 \times 21$  grid

$m_0 = 125$  (CSRBFs in  $5 \times 5 \times 5$  grid) .05 percent Gaussian noise



# TREGS: 44 Fev, 29 Jev



# Main Ingredients

Multi-pronged approach to solving reconstruction problems:

- 1 Determine a “natural” model for the image space that provides robustness to noise. Two options:
  - ▶ Find dictionary  $W$ ,  $\mathbf{f} = W\mathbf{p}$ , solve for  $\mathbf{p}$ .
  - ▶ Low order parameterization of image space,  $\mathbf{f} = \mathbf{f}(\mathbf{p})$ , solve for  $\mathbf{p}$ .
- 2 Determine **surrogate forward operator or model** so function/Jacobian evaluations becomes less costly. Two options:
  - ▶ Structure and Randomized approaches
  - ▶ Reduced order modeling

# Approximating the Forward Operator - Linear Case

$M(\mathbf{f}) = \mathbf{A}\mathbf{f}(\mathbf{p})$ . Singular values decay quickly, dense but “structured” e.g.

- Nearly spatially invariant, sums of Kronecker products
- Hierarchical low rank blocks

Solution of inverse problem involves products with, or rank-revealing factorization of,  $\mathbf{A}$ . Generate approximate (hierarchical) low-rank approximations using matrix structure<sup>12</sup> or via matrix randomization techniques<sup>13</sup> and use these as **surrogates**<sup>14</sup> in the optimization problem.

---

<sup>12</sup>See Chung, K, O’Leary, SISC, 37 (2015) and references therein.

<sup>13</sup>Chaillat and Biros, J. Comput. Phys., 231 (2012).

<sup>14</sup>Saibaba, K., Miller, Fantini, SISC, 37 (2015).

# Approximating the Forward Model – Nonlinear Case

Evaluating  $M(\mathbf{p})$  **major computational bottleneck** in applications like hydraulic tomography, DOT.

Function evals involve **transfer function** evals<sup>15</sup>:

$$\Psi(\omega; \mathbf{p}) = \underbrace{\mathbf{C}^T}_{n_{det} \times n} \underbrace{\left( \frac{i\omega}{\nu} \mathbf{E} + \tilde{\mathbf{A}}(\mathbf{p}) \right)^{-1}}_{n \times n} \underbrace{\mathbf{B}}_{n \times n_{src}}$$

Solve  $\min_{\mathbf{p}} \|\mathbf{m} - M(\mathbf{f}(\mathbf{p}))\|_2$ , where

$$M(\mathbf{f}(\mathbf{p}^{(k)})) = \begin{bmatrix} \text{vec}(\Psi(\omega_1, \mathbf{p}^{(k)})) \\ \vdots \\ \text{vec}(\Psi(\omega_\ell, \mathbf{p}^{(k)})) \end{bmatrix}.$$

<sup>15</sup>O'Connell, et al., SISC 2016

# Reducing Costs of Forward Model Evaluation

Ideas (independent of the image representation!):

- 1 Randomization to reduce the number of effective measurements and required model evaluations
- 2 Model reduction techniques to reduce the cost per evaluation
- 3 Combined approach

# Stochastic Sources and Detectors

Alternative representation of residual:  $\|\mathbf{D} - \Psi(0, \mathbf{p}^{(j)})\|_F^2$ .

Motivation: Haber, Chung, Hermann, SIOPT, 2012.

Let  $\Omega \in \mathbb{R}^{n_{src} \times k}$ ,  $k \ll n_{src}$  be drawn from an appropriate distribution, then

$$\frac{1}{k} \|\mathbf{D}\Omega - \Psi(\mathbf{p})\Omega\|_F^2$$

can be used to estimate the cost. Since  $\Psi(\mathbf{p})\Omega = \mathbf{C}^T \left( \tilde{\mathbf{A}}(\mathbf{p})^{-1}(\mathbf{B}\Omega) \right)$ , need solve only  $k$ ,  $n \times n$  linear systems.

The  $k$  columns of  $\mathbf{B}\Omega$  are called the **stochastic sources**.

# Randomize, then Optimize Approach

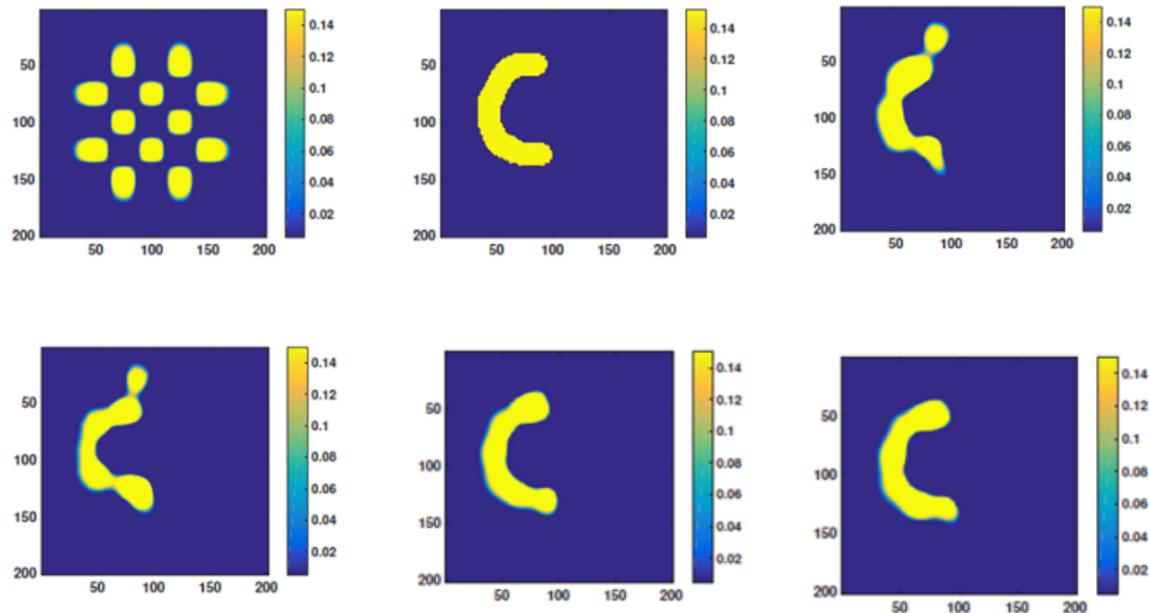
Recent work for the DOT problem<sup>16</sup> takes this idea further.

- Where we need to find  $(\tilde{\mathbf{A}}(\mathbf{p}))^{-*}\mathbf{C}$ , we can replace  $\mathbf{C}$  by  $k$  **stochastic detectors**  $\mathbf{C}\Gamma$  (i.e.  $\Gamma$  is  $n_{det} \times k$ ,  $k \ll n_{det}$ ).
- Compute the solution to the optimization problem using only stochastic sources and detectors.
- This solution largely not sufficient (stagnation).
- Add few ‘optimized’ sources & detectors as optimization progresses to keep only  $\approx k$  solves of the forward and adjoint at each step.

---

<sup>16</sup>Sariyadin, de Sturler, K, 2017

# Randomized + Optimized Results<sup>17</sup>



<sup>17</sup>From Sariyadin, de Sturler, K, "Randomized approach to nonlinear inversion combining simultaneous random and optimized sources and detectors" (2017)

# Model Order Reduction

Still requires some FOM solves throughout the whole optimization.

Use a cheaper-to-evaluate approx.  $\Psi_r(\omega, \mathbf{p})$  such that

$$\Psi(\omega, \mathbf{p}^{(k)}) \approx \Psi_r(\omega, \mathbf{p}^{(k)})$$

where

$$\Psi(\omega, \mathbf{p}) = \mathbf{C}^T \underbrace{\left( \frac{i\omega}{\nu} \mathbf{E} + \tilde{\mathbf{A}}(\mathbf{p}) \right)^{-1} \mathbf{B}}_{n_{src}, n \times n \text{ system solves}}$$

$$\Psi_r(\omega, \mathbf{p}) = \mathbf{C}_r^T \underbrace{\left( \frac{i\omega}{\nu} \mathbf{E}_r + \tilde{\mathbf{A}}_r(\mathbf{p}) \right)^{-1} \mathbf{B}_r}_{n_{src}, r \times r \text{ system solves}}$$

Replace simulated data at  $(\omega_\ell, \mathbf{p}^{(k)})$  with  $\text{vec}(\Psi_r(\omega_\ell, \mathbf{p}^{(k)}))$ .

# Reduced Order Modeling

Use interpolatory parametric model reduction:

- Find  $\Psi_r$  easy to evaluate giving high-fidelity approx to  $\Psi$  over parameters and frequencies **of interest**
- Require the same of  $\nabla_{\mathbf{p}}\Psi_r$
- Use projection matrix  $\mathbf{V}$  with **only  $r$  columns** to obtain the surrogate model

E.g. for  $\omega = 0$ :

$$\Psi_r(0; \mathbf{p}) = \mathbf{C}_r^T \tilde{\mathbf{A}}_r^{-1}(\mathbf{p}) \mathbf{B}_r$$

$$\text{with } \tilde{\mathbf{A}}_r = \underbrace{\mathbf{V}^T \tilde{\mathbf{A}} \mathbf{V}}_{r \times r}; \mathbf{B}_r = \mathbf{V}^T \mathbf{B}; \mathbf{C}_r = \mathbf{V}^T \mathbf{C}.$$

# Construction of $\mathbf{V}$

- Construct  $\mathbf{V}$  to (ideally) cover the entire parameter space of interest
- Need to solve some full order model (FOM) systems to find  $\mathbf{V}$

$$\tilde{\mathbf{A}}(\mathbf{p}^{(k)})\mathbf{X}_k = \mathbf{B}, \quad \tilde{\mathbf{A}}^*(\mathbf{p}^{(k)})\mathbf{Y}_k = \mathbf{C}$$

- Choose  $\mathbf{p}_k$  from initial + optimization steps on the FOM.
- Well informed interpolation points & dual use of computations.
- Solve  $K$  full order model systems, take SVD of  $[\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_K; \mathbf{Y}_0, \dots, \mathbf{Y}_K]$ , use some left singular vectors for  $\mathbf{V}$ <sup>18</sup>
  - ▶ Computes redundant information (wasted computation)
  - ▶ Cost of SVD and choice of truncation

---

<sup>18</sup>E. de Sturler, et al., SISC 2015

# Global Basis Construction

Recent work for DOT<sup>19</sup>: Use the information from the full order model solves to dynamically and directly specify  $\mathbf{V}$ .

$$\tilde{\mathbf{A}}(\mathbf{p}^{(k)})\mathbf{X}_k = \mathbf{B}, \quad \tilde{\mathbf{A}}^*(\mathbf{p}^{(k)})\mathbf{Y}_k = \mathbf{C}$$

- Solve for  $\mathbf{X}_0$  ( $\mathbf{Y}_0$ ).
- For fixed  $k > 0$ , solve for all RHS using **two-level** recycled MINRES.
  - ▶ Some FOM solves must be done
- Incrementally update  $\mathbf{V}$  from the **recycle spaces** plus **augmentation vectors** based on MINRES run.
- Efficient, avoids extra solves, SVD computation/truncation.

---

<sup>19</sup>O'Connell, et al., SISC, to appear 2017

# Numerical Experiment

DOT example<sup>20</sup>; invert for the absorption using PaLS for the image model.

## Full Order Model

- Original System Size:  
 $n = 40,401$
- Number of Func Evals: 30 ;  
Num Jac Evals 15
- 1,440, Full systems of size  $n \times n$

## ROM

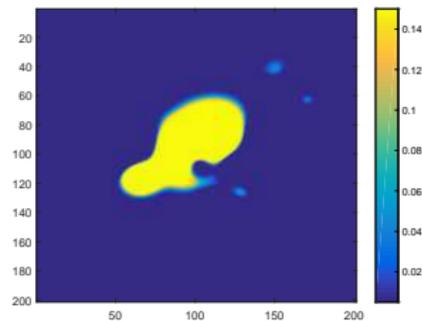
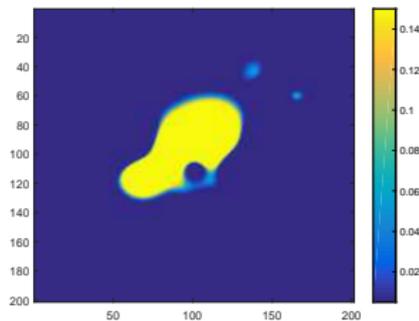
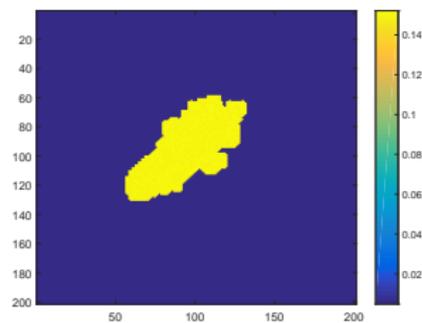
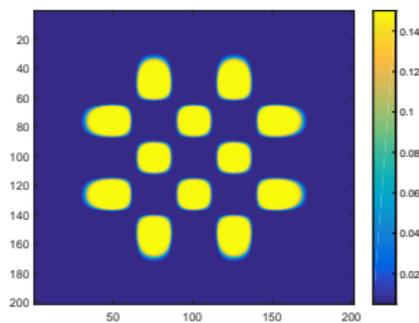
- ROM System Size:  
 $r = 197$
- Number of Func Evals: 28; Num  
Jac Evals 14
- 187,  $n \times n$  systems + 1,344 of  
size  $r \times r$

32 sources and 32 detectors

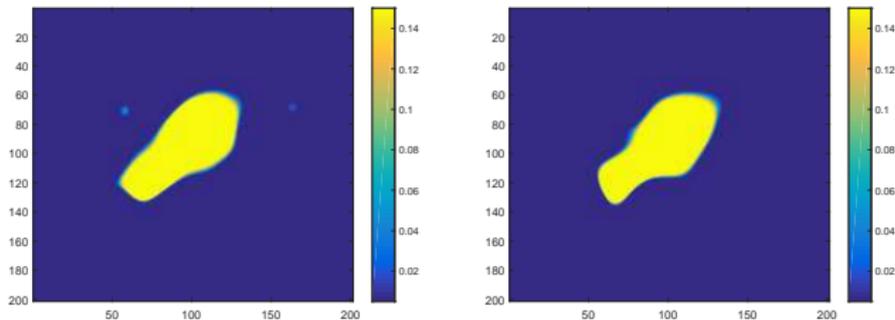
---

<sup>20</sup>O'Connell, et al., SISC, to appear 2017

# Reconstruction Results



## DOT- ROM results with 2 frequencies



Reconstruction results using data at both 0 and 10 MHz. The FOM required 82 function evaluations and 49 Jacobian evaluations, the ROM required 104 function evaluations and 56 Jacobian evaluations.  $r = 265$

# Combining ROMS and Randomization

ROMs reduce the size of the linear systems but not the number of systems per optimization step, whereas randomization reduces the number of systems to solve but does not change their size.

Combined, these approaches will be quite powerful!

# Combining ROMS and Randomization

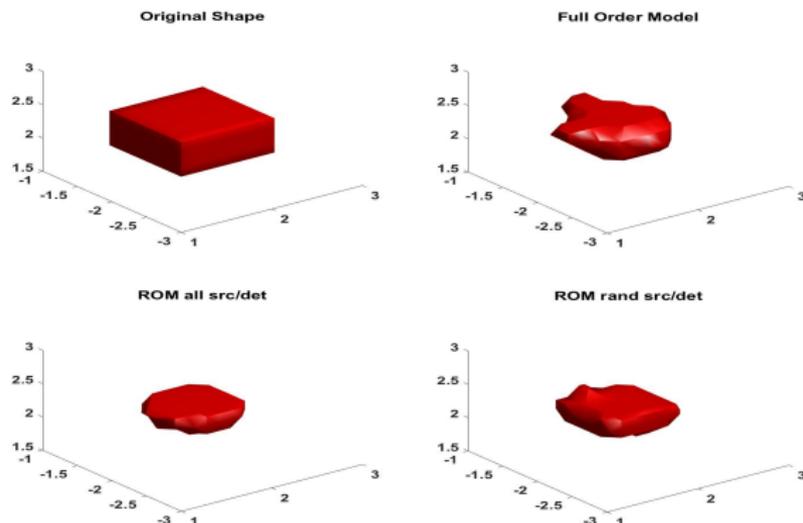
Current work for DOT<sup>21</sup>

- $32 \times 32 \times 32$  ( $n = 32768$ )
- 225 sources (top); 225 detectors (bottom).
- 27 compactly supported radial basis functions
- FOM, all sources and detectors 17550,  $n \times n$  solves.
- ROM from 5 interp points  $\rightarrow$  2250,  $n \times n$  solves.
- Randomized approach to computing ROM basis, 60 stochastic sources & detectors ea. 5 interpolation points, leading to 600 large linear solves

---

<sup>21</sup>Sariaydin, et. al; See Tues. PP1, “Computing Reduced Order Models Using Randomization”

# Combining ROMS and Randomization



FOM: 17,500  $n \times n$

ROM, usual: 2,250  $n \times n$ ,  $r = 568$ , 7,200,  $r \times r$  solves

ROM, stochastic: 600  $n \times n$ ,  $r = 568$ , 9,900,  $r \times r$  solves

Combined approach reduces no. large linear solves by **about a factor 30**

# Updating of ROMS

Recent work by Munster and de Sturler<sup>22</sup> possible to **cheaply** determine when FOM solves are needed to refresh the ROM during optimization.

$$F_{ROM}(\mathbf{p}^{(k)}) = \|\mathbf{D} - \Psi_r(\mathbf{p}^{(k)})\|_F$$

compared to

$$F_{Full}(\mathbf{p}^{(k)}) = \|\mathbf{D} - \Psi(\mathbf{p}^{(k)})\|_F \text{ estimated by } \|\mathbf{D}\mathbf{w} - \Psi(\mathbf{p}^{(k)})\mathbf{w}\|_F$$

and

$$\Psi(\mathbf{p}^{(k)})\mathbf{w} = \mathbf{C}^T \underbrace{\left( \mathbf{A}(\mathbf{p}^{(k)})^{-1}(\mathbf{B}\mathbf{w}) \right)}_{1 \text{ solve}}$$

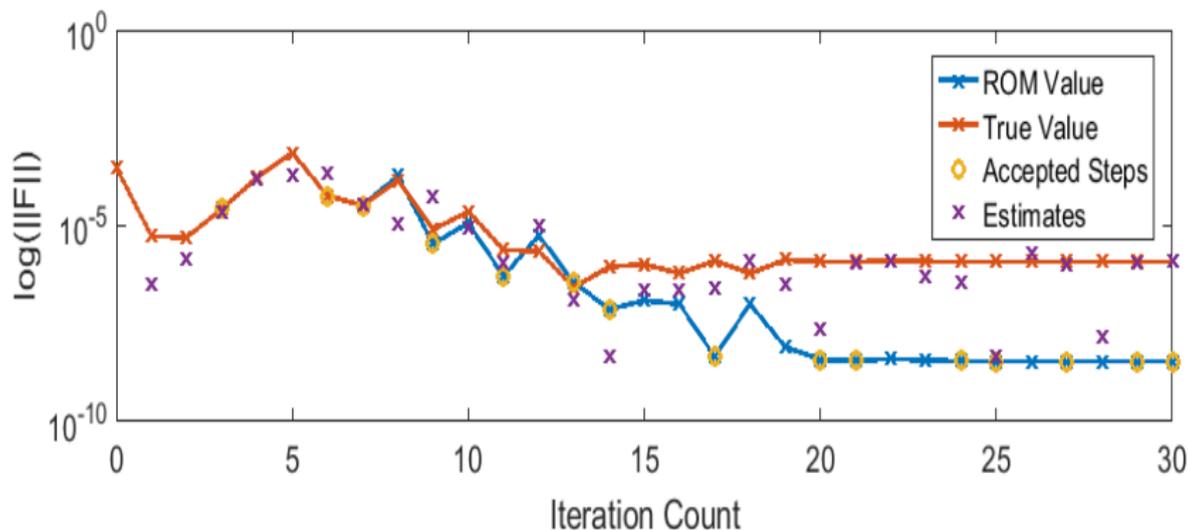
where  $\mathbf{w}$  from Rademacher distribution.

---

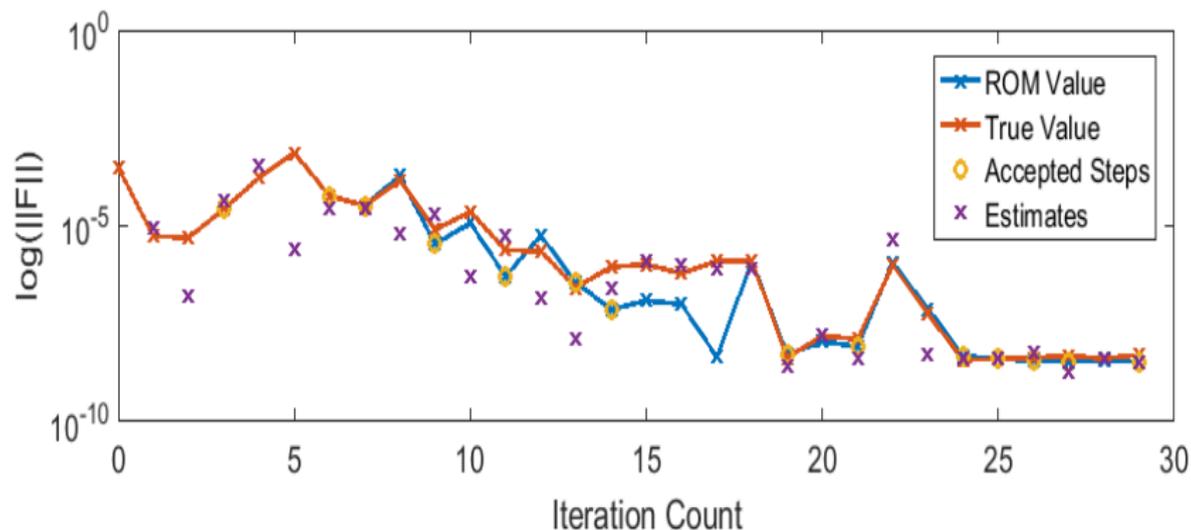
<sup>22</sup> “Nonlinear inversion using parametric model reduction with stochastic error estimates”, 2017

# Updating of ROMs

$n = 40401$ ; 2D DOT PaLS;  $r = 74$ ; 398  $n \times n$  solves and 1120  $r \times r$  solves.



# Updating of ROMs



System Size	Number Solves
40401	461
74	448
78	640

# Summary and Future Directions

Touched on two main ingredients for producing effective and efficient algorithms for reconstruction problems.

- Image Modeling:
  - ▶ Enforcing a priori information directly on the image model can make more robust to noise, may reduce the search space.
- Approximating the Forward Model:
  - ▶ **Complementary scientific computing techniques** working together have best chance of significantly improving reconstruction speed while maintaining accuracy

Related on-going work (Saibaba, K., de Sturler, Miller):  $\Psi(\mathbf{p}) \approx$  low rank; randomized SVD approximation using result of computations equivalent to using stochastic sources and/or detectors.

# Many Interesting Subproblems Remain

- Image representation
  - ▶ Other machine learning techniques
  - ▶ Other parametric representations (like PaLS, that are grid independent)
- Regularization (e.g.)
  - ▶ Sparsity or edge enhancing
  - ▶ Geometric constraints or multiway data constraints (tensors)
- Evaluation and approximation of forward model
  - ▶ Krylov subspace methods, recycling, preconditioning, multigrid
  - ▶ Other randomized techniques
  - ▶ Other ROM approaches – e.g.
    - ★ Wed., MS160,
    - ★ Fri., MS275, MS302 , includes tensor POD, multiscale approaches
- Optimization
  - ▶ Sophisticated, tailored algorithms – MS160, MS187
- Uncertainty quantification and related work
  - ▶ MS165, MS192; and Fri, MS293 10:25