Virginia Tech
Invent the Future®

Tomography with Random and Optimized Sources and Detectors

Eric de Sturler, Selin Sariaydin, Misha Kilmer

Department of Mathematics, Virginia Tech
sturler@vt.edu
www.math.vt.edu/people/sturler/index.html

Collaboration with Serkan Gugercin, Christopher Beattie, Zlatko Drmac

SIAM Conference on Applied Linear Algebra
Atlanta, Georgia, October 26-30, 2015

Overview

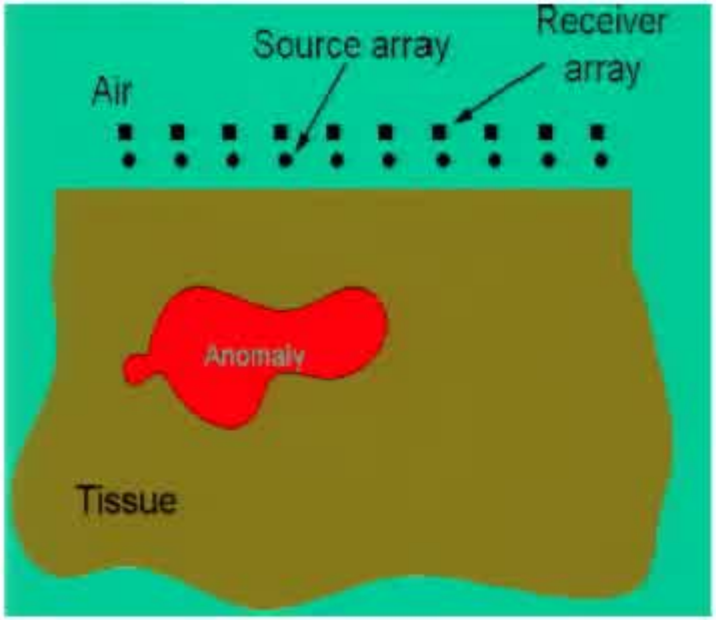
1. Diffuse Optical Tomography and Nonlinear Parametric Inversion
2. A Stochastic Approach Using Simultaneous Random Sources and Detectors
3. Improving the Stochastic Approach Using Optimized Sources and Detectors
4. Conclusions and Future Work

This material is based upon work supported by the National Science Foundation under DMS-1217156 and DMS-1217161.
Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Diffuse Optical Tomography

- Tissue illuminated by near-infrared, frequency modulated light
- Light detected in array(s)
- Tumors have different optical properties than healthy tissue
- Recover images of optical properties, $D(\mathbf{x})$ and $\mu(\mathbf{x})$, in $\nabla \cdot (D(\mathbf{x})\nabla\eta) - (\mu(\mathbf{x}) + \frac{\imath\omega}{\nu}\tilde{I})\eta = b_j$
- Problem is ill-posed and underdetermined, and data is noisy

- Sources 2mm apart: 2500 sources
- Detectors $\approx$ 0.5mm apart. Implies $> O(10^6)$ forward unknowns
- m_s sources and m_ω frequency modulations: solution of $m_s \cdot m_\omega$ discretized PDEs for function evaluation. Additional solves for Jacobian.



Parametric Inversion

- Parameterize D and μ by expansion in compactly supported radial basis functions (CSRBF) and map to near bivalued function using approximate Heavyside function and level set [1]
- Solve for (relatively) modest number of parameters
- This parameterization yields regularized solution
- Assume for simplicity a single frequency. Nonlinear residual:

$$\begin{bmatrix} \mathbf{r}_1(\mathbf{p}) \\ \vdots \\ \mathbf{r}_{m_s}(\mathbf{p}) \end{bmatrix} = \begin{bmatrix} \mathbf{C}^T \mathbf{A}^{-1}(\mathbf{p})\mathbf{b}_1 - \mathbf{d}_1 \\ \vdots \\ \mathbf{C}^T \mathbf{A}^{-1}(\mathbf{p})\mathbf{b}_{m_s} - \mathbf{d}_{m_s} \end{bmatrix}$$

- Nonlinear Least Squares Problem:

$$\min_{\mathbf{p}} \frac{1}{2} ||\mathbf{r}(\mathbf{p})||_2^2$$

[1] Aghassi/Kilmer/Miller'11

Nonlinear Problem

The Gauss-Newton (GN) approximation at $\mathbf{p}$

$$m_{GN}(\mathbf{p}+\delta) \approx \frac{1}{2}\mathbf{r}^T\mathbf{r} + \mathbf{r}^T\mathbf{J}\delta + \frac{1}{2}\delta^T\mathbf{J}^T\mathbf{J}\delta$$

where Jacobian: $\frac{\partial \mathbf{r}(\mathbf{p})}{\partial \mathbf{p}_k} = \left[\begin{array}{ccc} \frac{\partial \mathbf{r}(\mathbf{p})}{\partial \mathbf{p}_1} & \dots & \frac{\partial \mathbf{r}(\mathbf{p})}{\partial \mathbf{p}_{m_p}} \end{array}\right]$,

and $\mathbf{J}_{jk}(\mathbf{p}) = \frac{\partial}{\partial \mathbf{p}_k}(\mathbf{C}^T\mathbf{A}^{-1}(\mathbf{p})\mathbf{b}_j) = \mathbf{C}^T\mathbf{A}^{-1}(\mathbf{p})\frac{\partial}{\partial \mathbf{p}_k}\mathbf{A}(\mathbf{p})\mathbf{A}^{-1}(\mathbf{p})\mathbf{b}_j$

Trust region algorithm with REGularized model Solution[2] (TREGS) solves Nonlinear Least-Squares Problem:

- Ill-conditioned Jacobians
- Gauss-Newton step: Regularized solve for $\mathbf{J}\delta\mathbf{p} \approx -\mathbf{r}$
- TSVD of $\mathbf{J}$ for the large singular values
- Combined with selecting large spectral components in $\mathbf{r}$

[2] EdS/Kilmer'11

Recast as a Stochastic Problem

$$\min_{\mathbf{p}} \|\mathbf{r}(\mathbf{p})\|_2^2 \;=\; \min_{\mathbf{p}} \sum_{j=1}^{m_s} \|\mathbf{C}^T\mathbf{A}^{-1}(\mathbf{p})\mathbf{b}_j - \mathbf{d}_j\|_2^2 = \min_{\mathbf{p}} \|\mathbf{C}^T\mathbf{A}^{-1}(\mathbf{p})\mathbf{B} - \mathbf{D}\|_F^2$$

Matrix Form: $\mathbf{R}(\mathbf{p}) = \mathbf{C}^{\mathsf{T}}\mathbf{A}(\mathbf{p})^{-1}\mathbf{B} - \mathbf{D}$

Stochastic optimization problem (Haber, Chung, Herrmann SIOPT'12):

$$\min_{\mathbf{p}} \mathrm{E}_{\mathbf{w}}\left(\mathbf{w}^T\mathbf{R}(\mathbf{p})^T\mathbf{R}(\mathbf{p})\mathbf{w}\right) \;=\; \min_{\mathbf{p}} \mathrm{trace}\; \mathbf{R}(\mathbf{p})^T\mathbf{R}(\mathbf{p}) = \min_{\mathbf{p}} \|\mathbf{R}(\mathbf{p})\|_F^2$$

Rademacher distribution: $\mathbf{w}_j \in \{-1, +1\}$ i.i.d. with equal probability $(1/2)$

Drastically reduces the number of solves in function evaluation but not in Jacobian.

$$\mathbf{J}_{jk}(\mathbf{p}) = \mathbf{C}^T\mathbf{A}^{-1}(\mathbf{p})\frac{\partial}{\partial \mathbf{p}_k}\mathbf{A}(\mathbf{p})\mathbf{A}^{-1}(\mathbf{p})\mathbf{b}_j$$

Simultaneous Random Sources & Detectors

Similar approach for detectors:

$$\min_{\mathbf{p}} \mathrm{E}\left(\mathbf{v}^T \mathbf{R}(\mathbf{p})\mathbf{w}\right)^2 = \min ||\mathbf{R}(\mathbf{p})||_F^2$$

Columns of $\mathbf{B}$ correspond to m_s sources; rows of $\mathbf{C}^T$ to m_d detectors

Replace columns of $\mathbf{B}$ by a few random linear combinations $\mathbf{Bw}_j$ (SAA)
Replace columns of $\mathbf{C}$ by a few random linear combinations $\mathbf{Cv}_j$

$$\mathbf{V}^T \left(\mathbf{C}^T \mathbf{A}(\mathbf{p})^{-1}\mathbf{B} - \mathbf{D}\right)\mathbf{W} = \mathbf{V}^T \mathbf{R}(\mathbf{p})\mathbf{W}$$

Simultaneous Random Sources & Detectors

In vector form

$$\min_{\mathbf{p}} \|\mathbf{V}^T \mathbf{R}(\mathbf{p}) \mathbf{W}\|_F^2 = \min_{\mathbf{p}} \| \left(\mathbf{W}^T \otimes \mathbf{V}^T \right) \mathbf{r}(\mathbf{p}) \|_2^2$$

With simultaneous sources $\mathbf{W}$ and detectors $\mathbf{V}$
Regularized solve for $\mathbf{J}\delta\mathbf{p} \approx -\mathbf{r}$ and the Gauss-Newton step becomes

$$\left(\mathbf{W}^T \otimes \mathbf{V}^T \right) \mathbf{r}(\mathbf{p} + \delta\mathbf{p}) \approx \left(\mathbf{W}^T \otimes \mathbf{V}^T \right) (\mathbf{r}(\mathbf{p}) + \mathbf{J}\delta\mathbf{p})$$

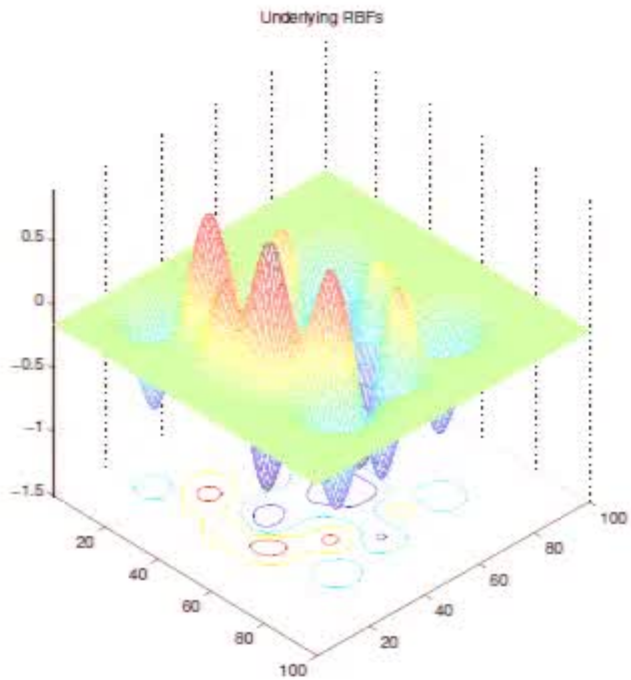
Hence

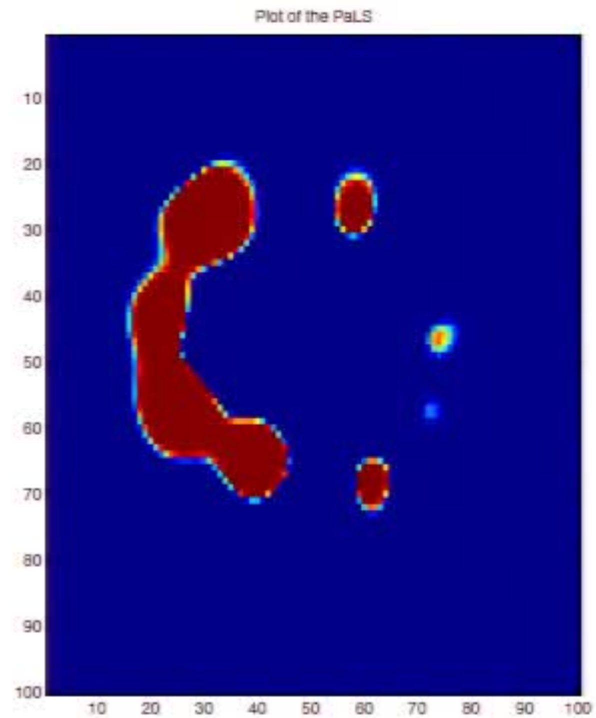
$$(\mathbf{W}^T \otimes \mathbf{V}^T)\mathbf{J}\delta\mathbf{p} \approx -(\mathbf{W}^T \otimes \mathbf{V}^T)\mathbf{r}$$

Use TREGS

A modest number of random sources and detectors is typically enough to get a reasonable solution

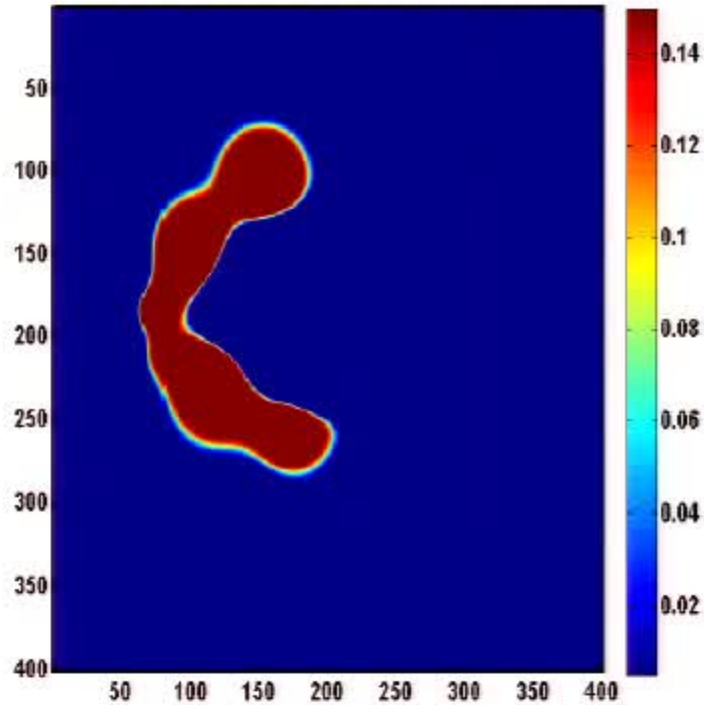
Weighted CSRBFs





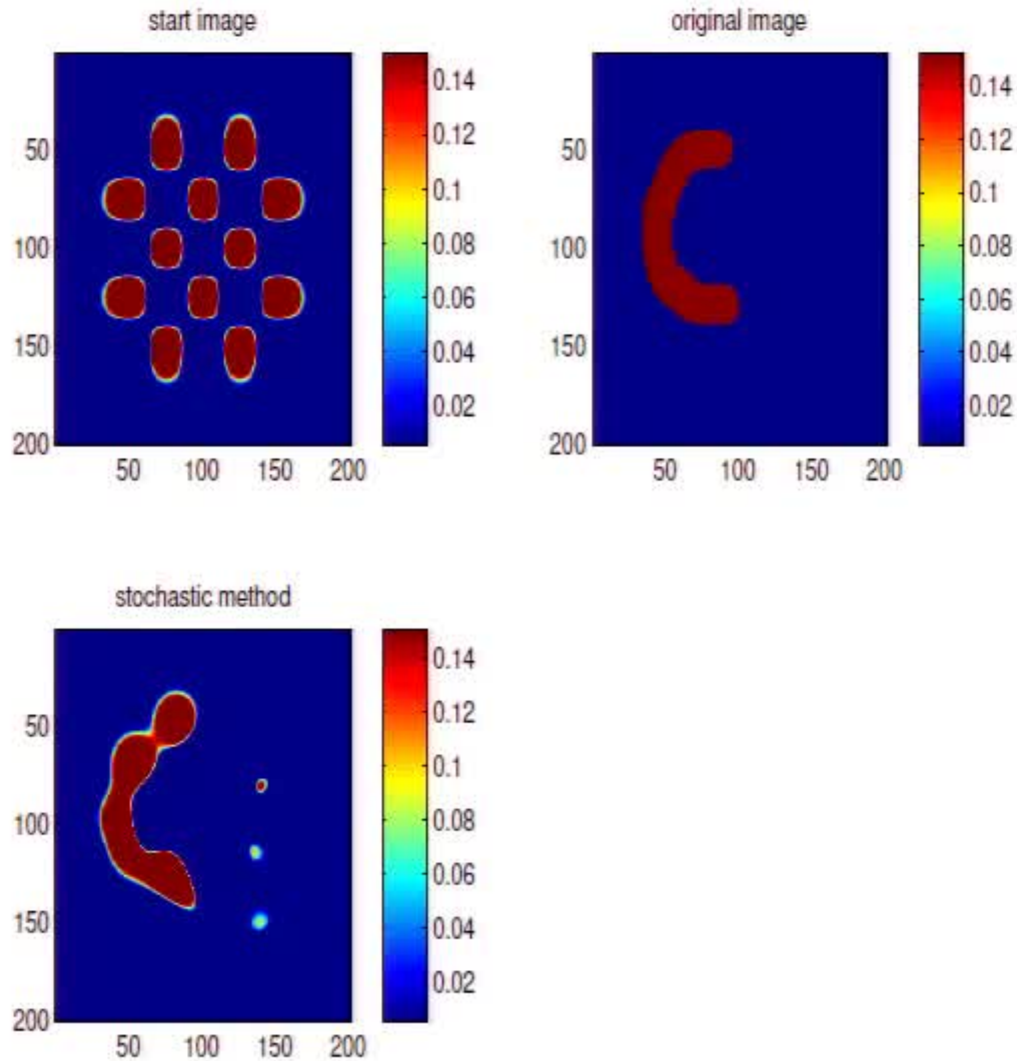
Full Order Model Result

Figure: Image reconstruction results with all sources and all detectors $m_s = m_d = 32$



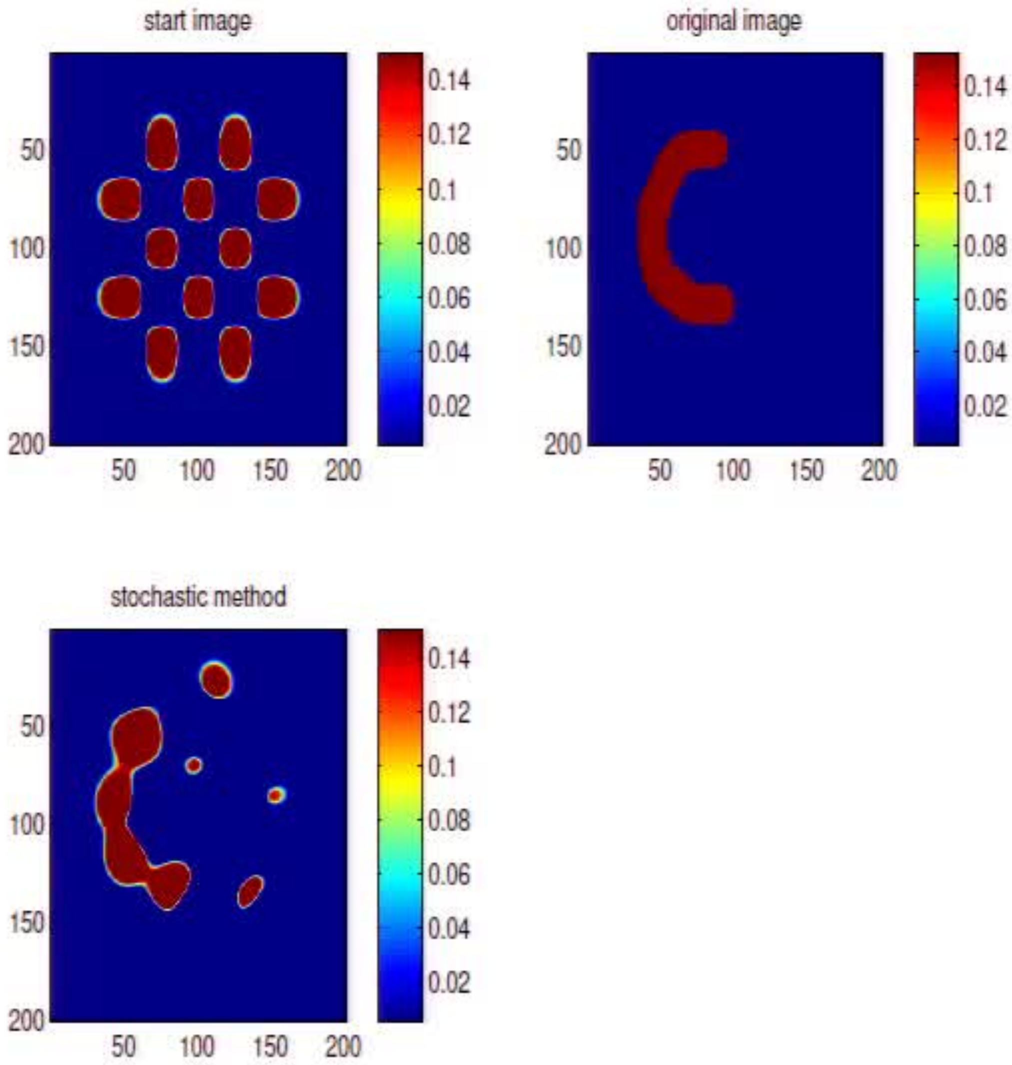
Stochastic Results

Figure: Example 1: Image reconstruction results with $\ell_s = \ell_d = 10$



Stochastic Results

Figure: Example 2: Image reconstruction results with $\ell_s = \ell_d = 10$



How to Improve the Stochastic Approach?

Good localization from random sources and detectors in modest # steps.
Replace few sources and detectors by optimized sources and detectors.

Compute full Jacobian and residual
Replace a few sources/detectors to maximize Frobenius norm

$$\max \|(\mathbf{W}^T \otimes \mathbf{V}^T)\mathbf{J}\|_F^2 \qquad \text{(think spaces not vectors)}$$

TREGS prefers large sin. values (also link with opt. design of experiments)

Drop: $\mathbf{V}^{m_d \times \ell_d}$ by $\tilde{\mathbf{V}}^{m_d \times \ell_{d-1}}$ such that $\mathcal{R}(\tilde{\mathbf{V}}) \subset \mathcal{R}(\mathbf{V})$

Find unitary $\mathbf{\Gamma}^{\ell_d \times \ell_{d-1}}$ such that $\tilde{\mathbf{V}} = \mathbf{V}\mathbf{\Gamma}^{\ell_d \times \ell_{d-1}}$ and

$$\max_{\mathbf{\Gamma}} \|(\mathbf{W}^T \otimes \tilde{\mathbf{V}}^T)\mathbf{J}\|_F^2$$

Drop Random Sources & Detectors (optimal fashion)

For simplicity consider a single simultaneous source and ℓ_d detectors

$$\left[\mathbf{w}_1^T \otimes \mathbf{V}^T\right] \mathbf{J} = \mathbf{V}^T \widehat{\mathbf{J}}$$

Goal: $\max ||\tilde{\mathbf{V}}^T \widehat{\mathbf{J}}||_F^2$ i.e. $\max(\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_{\ell_d - 1}^2)$

$$\max ||\tilde{\mathbf{V}}^T \widehat{\mathbf{J}}||_F^2 = \max ||\mathbf{\Gamma}^T \mathbf{V}^T \widehat{\mathbf{J}}||_F^2$$

$$\widehat{\mathbf{V}}^T \widehat{\mathbf{J}} = \mathbf{\Phi \Omega \Psi}^T (\mathrm{SVD})$$

Keep the largest $(\ell_d - 1)$ left singular vectors of $\mathbf{V}^T \widehat{\mathbf{J}}$

$$\mathbf{\Gamma} = [\phi_1 \ \ \phi_2 \ \ \cdots \ \ \phi_{\ell_d - 1}]$$

We can also remove one (or more) simultaneous sources

Adding Optimized Sources & Detectors

Find unit $\mathbf{q} \perp \mathbf{V}$ that maximizes

$$\left\| \begin{bmatrix} \mathbf{w}_1^T \otimes \mathbf{V}^T \\ \mathbf{w}_1^T \otimes \mathbf{q}^T \\ \mathbf{w}_2^T \otimes \mathbf{V}^T \\ \mathbf{w}_2^T \otimes \mathbf{q}^T \\ \vdots \\ \mathbf{w}_{\ell_s} \otimes \mathbf{q}^T \end{bmatrix} \mathbf{J} \right\|_F^2 = \left\| \begin{bmatrix} \mathbf{V}^T \\ \mathbf{q}^T \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{J}}_1 & \dots & \widehat{\mathbf{J}}_{\ell_s} \end{bmatrix} \right\|_F^2$$

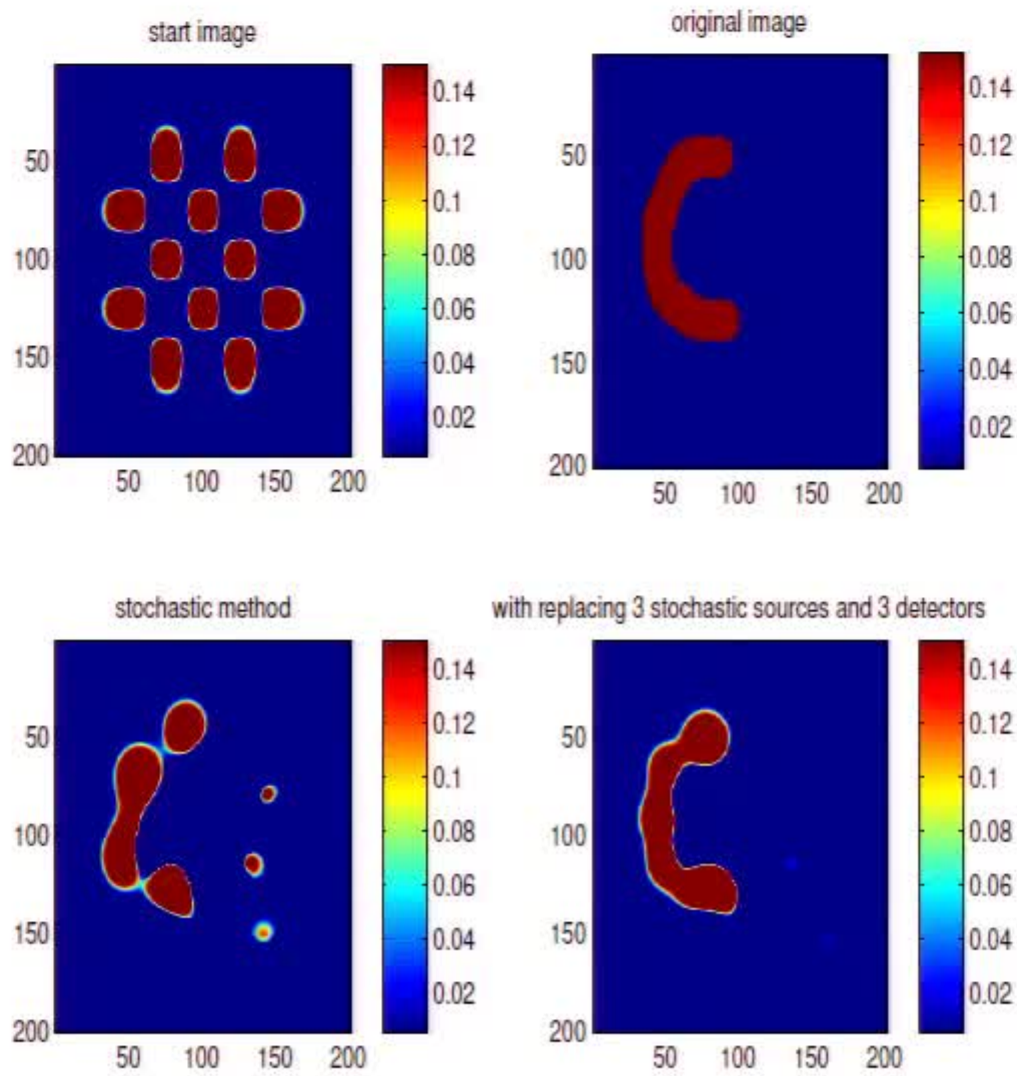
New part: $\left\| \mathbf{q}^T \left[\widehat{\mathbf{J}}_1 \ \dots \ \widehat{\mathbf{J}}_{\ell_s} \right] \right\|_2^2$ and taking $\mathbf{q} = \mathbf{V}_c \mathbf{y}$ $(\mathbf{V}_c \perp \mathbf{V})$ gives

$$\max_{\mathbf{y}} \left\| \mathbf{y}^T \left[\mathbf{V}_c^T \widehat{\mathbf{J}}_1 \ \dots \ \mathbf{V}_c^T \widehat{\mathbf{J}}_{\ell_s} \right] \right\|_2^2$$

Standard problem, can be solved by SVD (also for multiple new det.s)

Examples

Figure: Example 1: Image reconstruction results with $\ell_s = \ell_d = 10$



Results in Iterations

Average number of iterations, Function evaluations, Jacobian evaluations

	Iter	Fevals	Jevals	tol	PDE solves
Only Stochastic Method	12	13	6	3e-7	190
+ Replacing 1 src/1 det	10	9	6	3e-10	214
+ Replacing 2 src/2 det	10	9	5	3e-10	204
Only Stochastic Method	(149)	(150)	(119)	(3e-10)	2690
All srcs/ All dets	16	17	9	3e-10	1120