A Hybrid LSMR Algorithm for Large-Scale Tikhonov Regularization

Julianne Chung* Katrina Palmer†

*Department of Mathematics, Virginia Tech

†Department of Mathematical Sciences, Appalachian State University

SIAM Applied Linear Algebra Conference

October 29, 2015

Mathematical Problem

$$\mathbf{b} = \mathbf{A}\mathbf{x}_{\text{true}} + \varepsilon$$

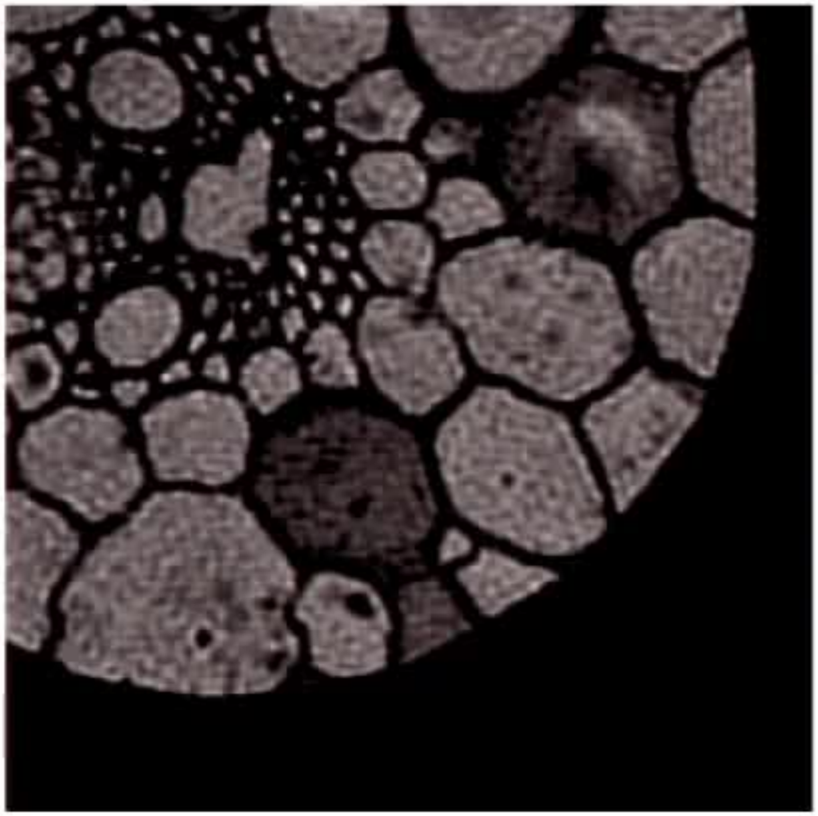
where

- $\mathbf{b} \in \mathbb{R}^n$ - observed data
- $\mathbf{x}_{\text{true}} \in \mathbb{R}^n$ - desired solution
- $\mathbf{A} \in \mathbb{R}^{n \times n}$ - models the forward processes
- $\varepsilon \in \mathbb{R}^n$ - noise, statistical properties may be known

Goal: Given $\mathbf{b}$ and $\mathbf{A}$, compute approximation of $\mathbf{x}_{\text{true}}$

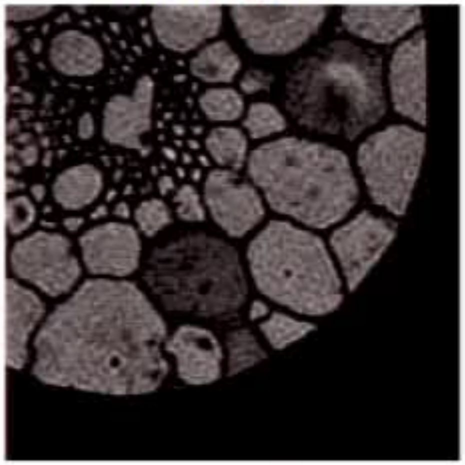
Application: Image Deblurring

- Given: Blurred image, $\mathbf{b}$, and some information about the blurring, $\mathbf{A}$
- Goal: Compute approximation of true image, $\mathbf{x}_{\mathrm{true}}$



An Ill-Posed Inverse Problem

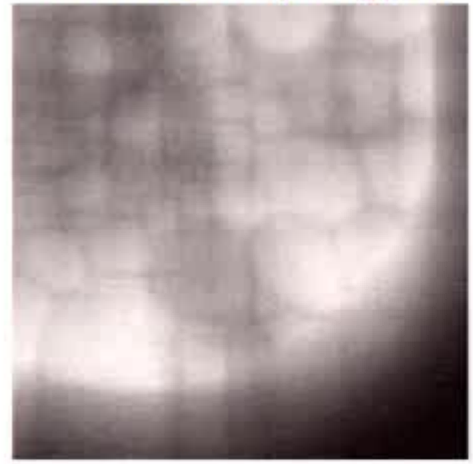
True image: x



Forward Problem

Inverse Problem

Blurred & noisy image: b



Inverse solution: $A^{-1}b$

Naïve inverse solution is corrupted with noise!

Choosing Regularization Parameter λ

- Discrepancy principle: $||(\mathbf{I} - \mathbf{A}\mathbf{A}_\lambda^\dagger)\mathbf{b}||_2 < \delta$
- Generalized cross validation - Golub, Heath and Wahba (1979)

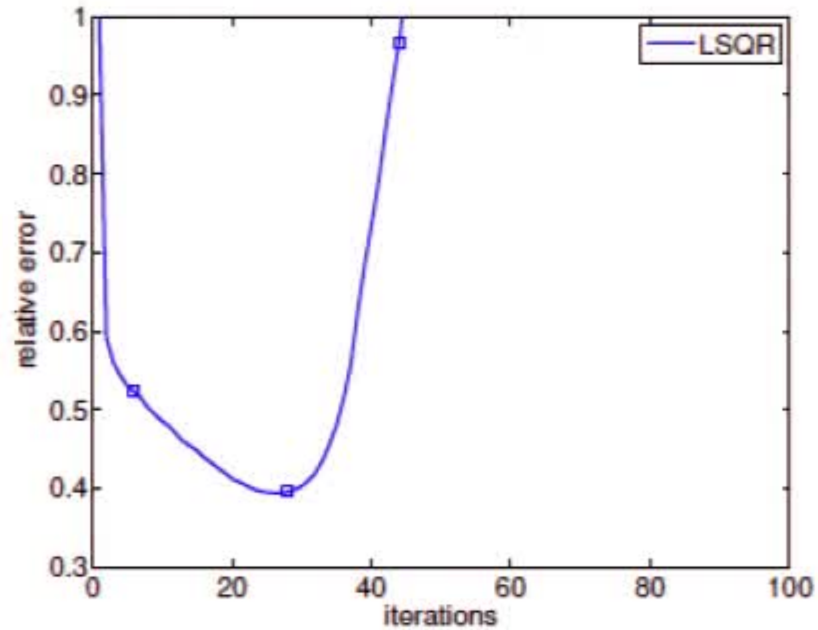
$$G_{\mathbf{A},\mathbf{b}}(\lambda) = \frac{n\|(\mathbf{I} - \mathbf{A}\mathbf{A}_\lambda^\dagger)\mathbf{b}\|_2^2}{\left[\text{trace}(\mathbf{I} - \mathbf{A}\mathbf{A}_\lambda^\dagger)\right]^2}$$

- Unbiased predictive risk estimator (UPRE) - Mallow (1973), Giryes, Elad, Eldar (2011)

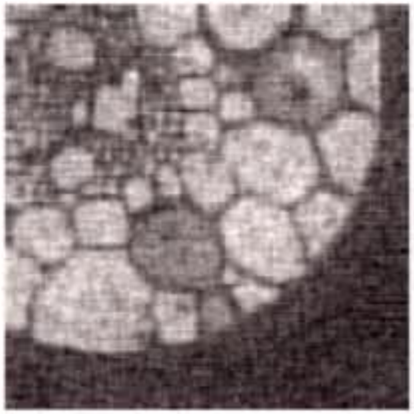
$$U_{\mathbf{A},\mathbf{b}}(\lambda) = \frac{1}{n}\|\mathbf{b} - \mathbf{A}\mathbf{x}_\lambda\|_2^2 + \frac{2\sigma^2}{n}\text{trace}(\mathbf{A}\mathbf{A}_\lambda^\dagger) - \sigma^2 .$$

Iterative Regularization

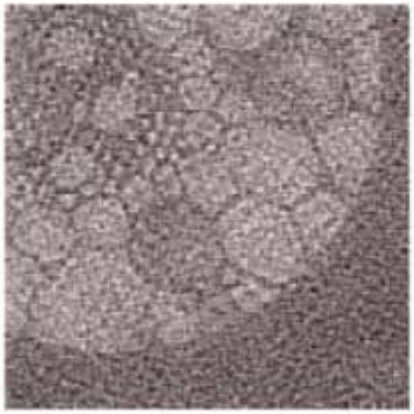
Apply standard iterative method to least squares problem, $\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2$, and terminate early



Iteration 6



Iteration 28



Iteration 44

Previous Work on Hybrid Methods

Regularization embedded in iterative method:

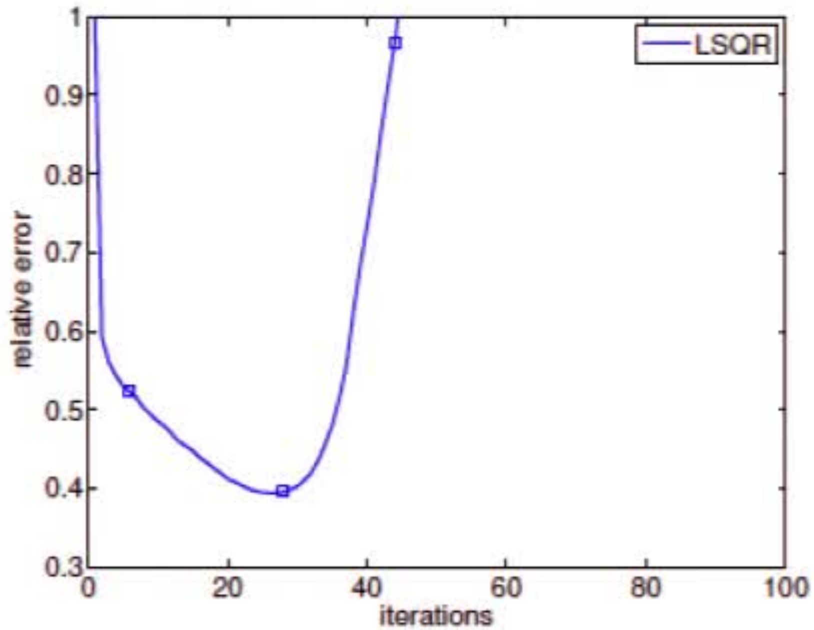
- O'Leary and Simmons, SISSC, 1981.
- Björck, BIT 1988.
- Björck, Grimme, and Van Dooren, BIT, 1994.
- Larsen, PhD Thesis, 1998.
- Hanke, BIT 2001.
- Kilmer and O'Leary, SIMAX, 2001.
- Kilmer, Hansen, Espanol, 2006.
- Bazan, Borges, 2010.
- Renaut, Hnětynková, Mead, 2010.

Use iterative method to solve regularized problem:

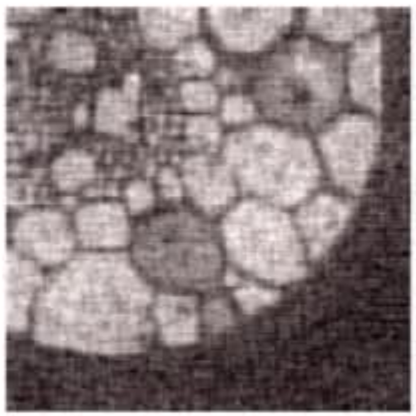
- Golub, Von Matt, Numer. Math.,1991.
- Calvetti, Golub, Reichel, BIT, 1999.
- Frommer, Maass SISC, 1999.

Iterative Regularization

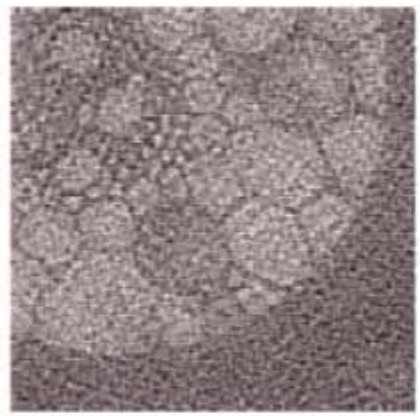
Apply standard iterative method to least squares problem, $\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2$, and terminate early



Iteration 6



Iteration 28



Iteration 44

LSQR Projected Problem

After k steps of GK bidiagonalization, LSQR *projected* problem:

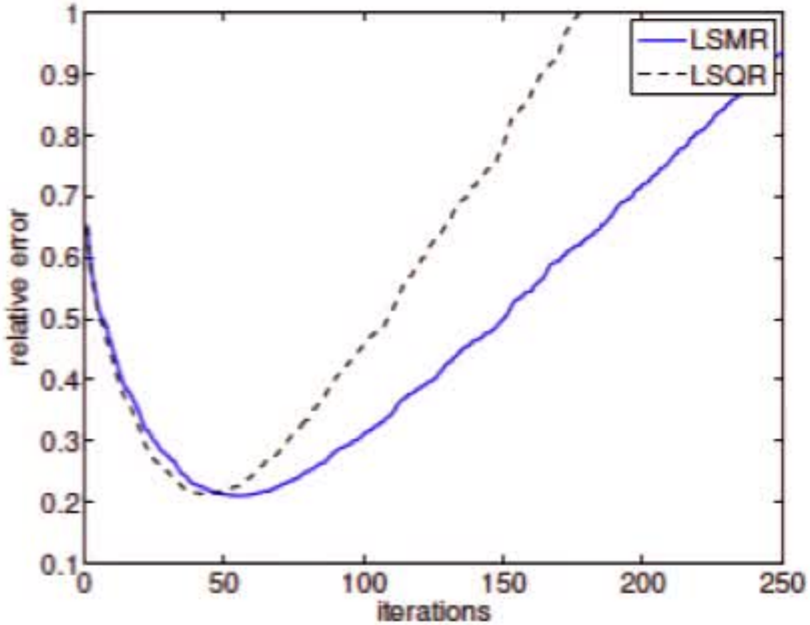
$$\min_{\mathbf{x} \in R(\mathbf{V}_k)} ||\mathbf{A}\mathbf{x} - \mathbf{b}||_2 = \min_{\mathbf{y}} ||\mathbf{B}_k \mathbf{y} - \mathbf{U}_{k+1}^T \mathbf{b}||_2$$
$$= \min_{\mathbf{y}} ||\mathbf{B}_k \mathbf{y} - \beta \mathbf{e}_1||_2$$

where $\mathbf{x}_k = \mathbf{V}_k \mathbf{y}$

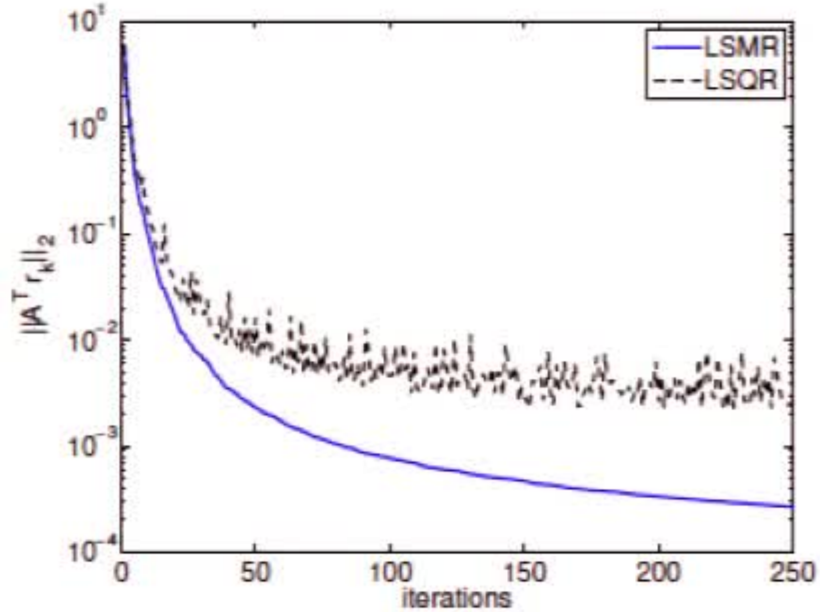
Remarks:

- Ill-posed problem $\Rightarrow$ $\mathbf{B}_k$ may be very ill-conditioned.
- $\mathbf{B}_k$ is much smaller than $\mathbf{A}$
- Standard techniques (e.g. GCV) to find λ and stopping point

LSQR vs LSMR for Ill-posed Problems



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



$\|\mathbf{A}^\top \mathbf{r}_k\|_2$

Remarks:

- LSQR and LSMR converge to the same solution
- LSMR exhibits delayed semiconvergence

Interlacing Property

- Let $\mathbf{B} = \mathbf{P} \begin{pmatrix} \mathbf{S} \\ \mathbf{0} \end{pmatrix} \mathbf{Q}^\top$ be SVD, with sing. vals. $s_1 \geq \ldots \geq s_k > 0$
- Eigenvalues of $\mathbf{B}^\top \mathbf{B}$: s_i^2, $i = 1, \ldots, k$
- Matrix
$$\hat{\mathbf{B}}^\top \hat{\mathbf{B}} = \mathbf{Q}^\top (\mathbf{S}^2 \mathbf{S}^2 + \bar{\beta}_{k+1}^2 \mathbf{q}_k \mathbf{q}_k^\top) \mathbf{Q}$$
where $\mathbf{q}_k$ is the kth column of $\mathbf{Q}$
- Using a theorem from Bunch, Nielsen and Sorensen (1978), we get interlacing property:
$$s_k^2 \leq \hat{s}_k \leq \ldots \leq s_2^2 \leq \hat{s}_2 \leq s_1^2 \leq \hat{s}_1$$
where $\hat{s}_1, \ldots, \hat{s}_k$ are sing. vals. of $\hat{\mathbf{B}}$
- In summary, singular values of $\hat{\mathbf{B}}$ approximate the *squares* of the largest and smallest singular values of $\mathbf{A}$

Hybrid LSMR as a Krylov subspace projection method

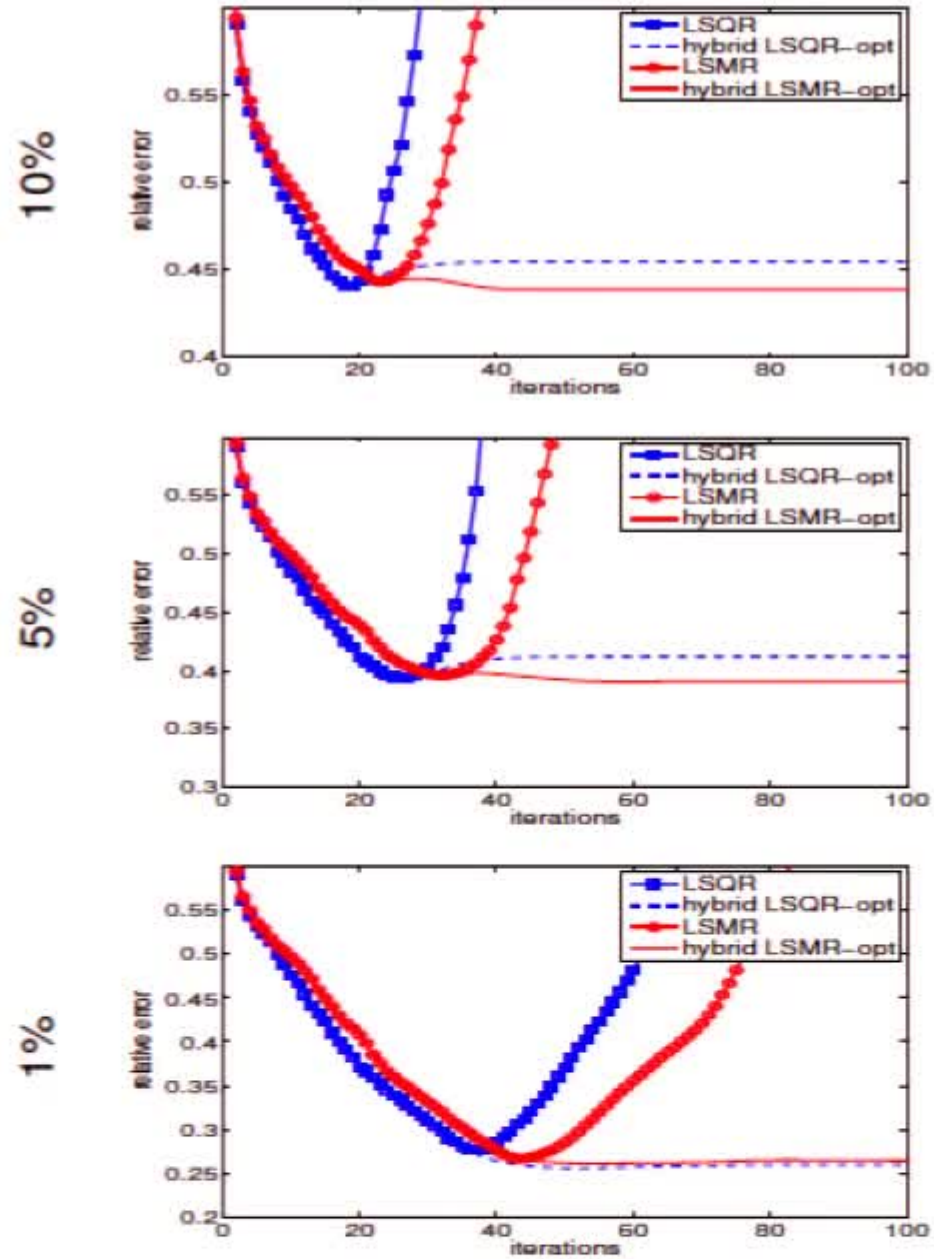
Theorem

Fix $\lambda \geq 0$. Let $\mathbf{y}_k$ be the exact solution to the regularized subproblem,

$$\mathbf{y}_k = \arg\min_{\mathbf{y}} \left\| \hat{\mathbf{B}}_k \mathbf{y} - \bar{\beta}_1 \mathbf{e}_1 \right\|_2^2 + \lambda^2 \|\mathbf{y}\|_2^2 ,$$

where $\hat{\mathbf{B}}_k, \mathbf{V}_k$ are derived from the original problem. Then the kth iterate of hybrid LSMR, $\mathbf{x}_k = \mathbf{V}_k \mathbf{y}_k$, solves the following problem

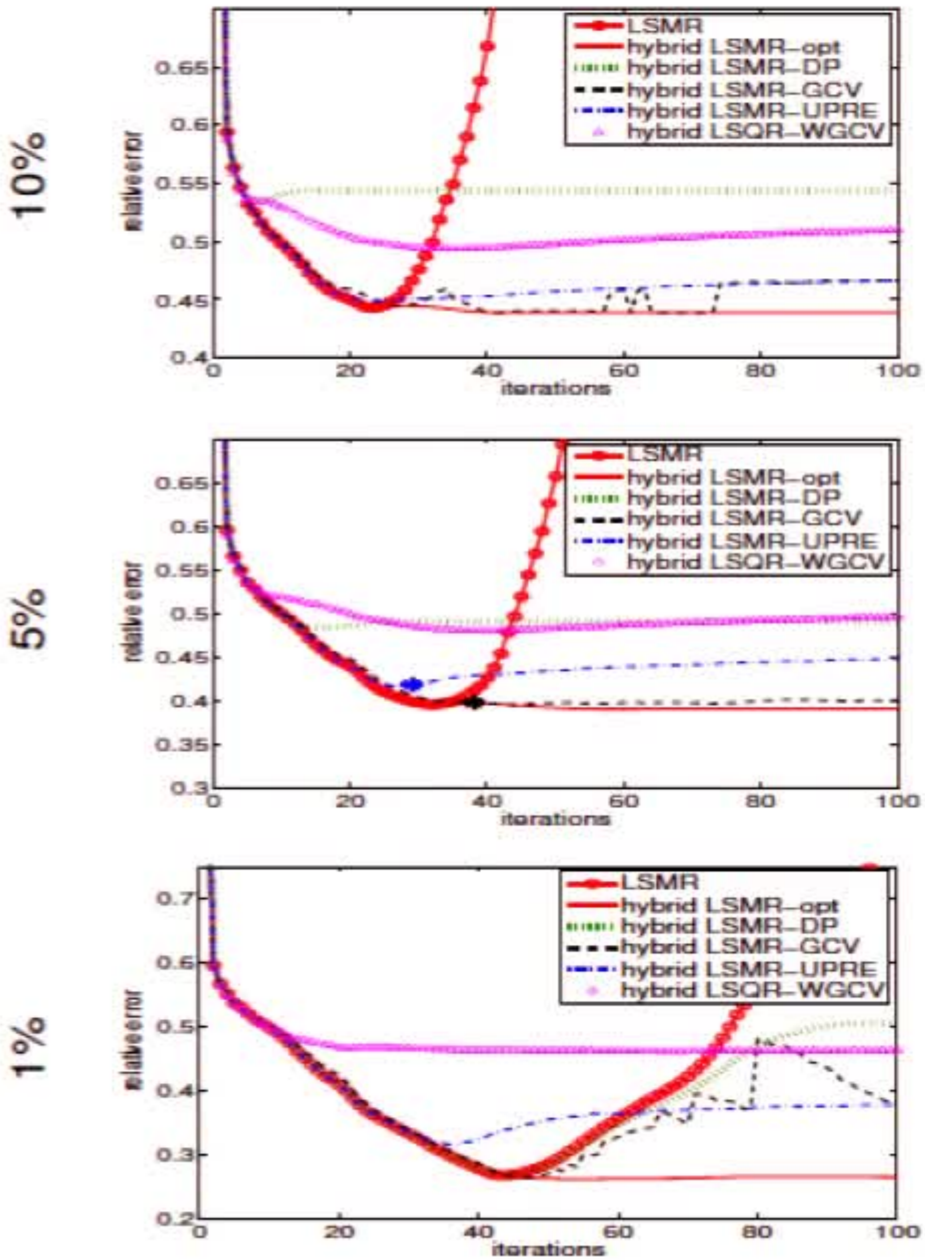
$$\min_{\mathbf{x} \in \mathcal{K}_k(\mathbf{A}^\top \mathbf{A}, \mathbf{A}^\top \mathbf{b})} \left\| \mathbf{A}^\top \mathbf{A} \mathbf{x} - \mathbf{A}^\top \mathbf{b} \right\|_2^2 + \lambda^2 \|\mathbf{x}\|_2^2 .$$



LSQR vs LSMR

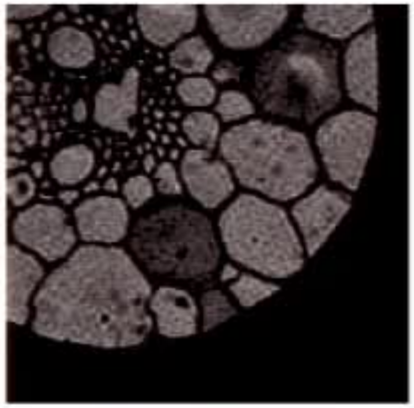
- LSMR exhibits delayed semiconvergence
- hybrid methods use optimal regularization parameter here
- hybrid LSMR can produce lower relative errors, especially for high noise levels

Comparison of Regularization Parameter Selection Methods



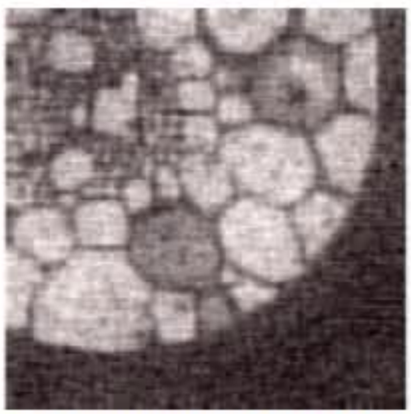
- For DP and UPRE, noise variance estimated using highest frequency of wavelet transform of **b**
- hybrid-LSQR using Weighted-GCV (WGCV) provided for comparison

Reconstructed Images (5% Noise)

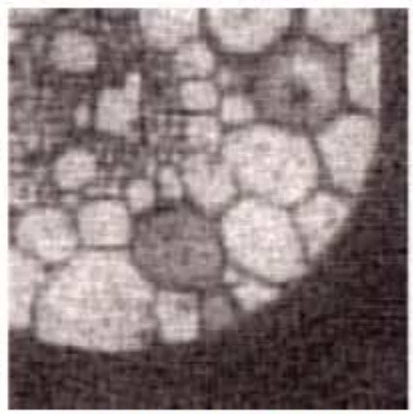


True image

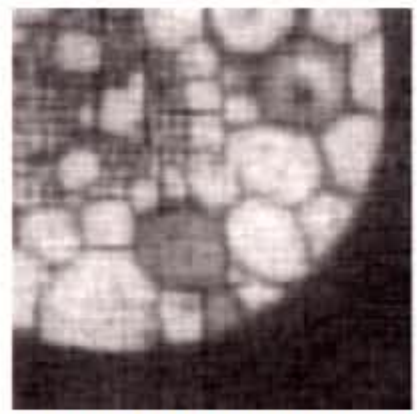
hybrid LSQR-WGCV (40)



“optimal” LSMR (31)



hybrid LSMR-GCV (37)



hybrid LSMR-UPRE (28)