Kronecker Products, Tensor Decompositions and 3D Imaging Applications

Misha E. Kilmer
Tufts University

James Nagy
Emory University

Lisa Perrone
Hawaii Pacific University

Outline

- Preconditioning for discrete ill-posed problems
- The matrix approximation problem
- The role of tensors
- Theoretical Results
- Numerical results
- Conclusions and future work

Problem Description

Model:

$$Kf = \hat{g} + e = g$$

- K is ill-conditioned, no gap in SV spectrum
- K is triply Toeplitz or triply T+H
  $K^T K$ may have similar structure in reconstruction
- For an $n \times n \times n$ image, K has n^3 columns!
- Noise, e, is white and unknown.

SVD Analysis

Assume we can compute

$$K = [U_1 U_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix},$$

where Σ_1 is $k \times k$ and corresponds to components such that $u_i^T g \approx u_i^T \hat{g}$.

Noise contaminated exact solution:

$$f = V_1 \Sigma_1^{-1} (U_1^T g) + \overbrace{V_2 \Sigma_2^{-1} \underbrace{(U_2^T g)}_{\approx U_2^T e}}^{dominant}$$

Truncated SVD solution: $f_{reg} = V_1 \Sigma_1^{-1} (U_1^T g)$

Iterative Regularization

We cannot compute the SVD of K! But K is structured, so we can compute matrix-vector products using FFT's quickly $O(n^3 \log n)$ flops if image is $n \times n \times n$.

This means we should use an iterative regularization scheme (eg. CGLS, MRNSD). We stop iterating before the solution converges to the exact solution of the system.

Cost is $O(Nj)$ where a matvec costs $O(N)$ and j is the number of iterations until semi-convergence.

Ideal Preconditioning

Major difficulty is that j can be large! Consider the left preconditioned system

$$M^{-1}Kf = M^{-1}g.$$

If we could compute the SVD of K, the ideal preconditioner would be

$$M = [U_1 U_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}.$$

Then, semi-convergence in 1 iteration since:

- The singular values of $M^{-1}K$ are 1 and Σ_2.
- $M^{-1}g$ looks similar to the TSVD solution so no noise introduced.

Goal

Since we cannot compute the SVD of K, the main goal of our work is to approximate K by a matrix for which we can compute (efficiently) and store (implicitly) the SVD, and use the SVD of the approximation to construct M.

Key: Using the structure of K, the matrix approximation problem can be reduced to a computationally tractable problem involving 3-way arrays (tensors).

Kronecker Products

$B \otimes C$ is the block matrix

$$\begin{bmatrix} b_{11}C & b_{12}C & \cdots & b_{1n}C \\ b_{21}C & b_{22}C & \cdots & b_{2n}C \\ \vdots & \vdots & \vdots & \vdots \\ b_{m1}C & b_{m2}C & \cdots & b_{mn}C \end{bmatrix}.$$

$A \otimes B \otimes C$ is the double-block matrix:

$$\begin{bmatrix} a_{11}B \otimes C & a_{12}B \otimes C & \cdots & a_{1n}B \otimes C \\ a_{21}B \otimes C & a_{22}B \otimes C & \cdots & a_{2n}B \otimes C \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1}B \otimes C & a_{m2}B \otimes C & \cdots & a_{mn}B \otimes C \end{bmatrix}.$$

SVDs of Kronecker Products

If $A = U_a \Sigma_a V_a^T$, $B = U_b \Sigma_b V_b^T$, $C = U_c \Sigma_c V_c^T$, then

$$A \otimes B \otimes C = (U_a \otimes U_b \otimes U_c)(\Sigma_a \otimes \Sigma_b \otimes \Sigma_c)(V_a^T \otimes V_b^T \otimes V_c^T),$$

which is an SVD (under appropriate ordering).

Matrix Approximation Problem

Find A_i, B_i, C_i with the appropriate structure such that

$$\|K - \sum_{i=1}^{s} A_i \otimes B_i \otimes C_i\|_F$$

is minimized. (reason for $s > 1$ will be discussed later)

Related 2D work:

- Kamm & Nagy, '00
- Nagy, & Ng, Perrone, '04
- Perrone, '05
- K. & Nagy, '06

Banded Toeplitz

Complete information about this banded Toeplitz matrix is captured in a central vector of length n:

$$\begin{bmatrix} 1 & 2 & 3 & 0 & 0 \\ 6 & 1 & 2 & 3 & 0 \\ 7 & 6 & 1 & 2 & 3 \\ 0 & 7 & 6 & 1 & 2 \\ 0 & 0 & 7 & 6 & 1 \end{bmatrix}.$$

Doubly Toeplitz

This doubly Toeplitz matrix can be represented by its central column, which we reshape:

$$\left[\begin{array}{ccc|ccc|ccc} a & b & 0 & x & h & 0 & 0 & 0 & 0 \\ r & a & b & l & x & h & 0 & 0 & 0 \\ 0 & r & a & 0 & l & x & 0 & 0 & 0 \\ \hline p & q & 0 & a & b & 0 & x & h & 0 \\ t & p & q & r & a & b & l & x & h \\ 0 & t & p & 0 & r & a & 0 & l & x \\ \hline 0 & 0 & 0 & p & q & 0 & a & b & 0 \\ 0 & 0 & 0 & t & p & q & r & a & b \\ 0 & 0 & 0 & 0 & t & p & 0 & r & a \end{array}\right], \qquad P = \begin{bmatrix} h & b & q \\ x & a & p \\ l & r & t \end{bmatrix}$$

Triply Toeplitz

Same structure as doubly Toeplitz, except now each T_i is a doubly Toeplitz matrix.

Similarly, if banded structure, the matrix is completely represented by its central column, which reshape into a 3rd order tensor P.

Structure

If we assume that the blurring in the 3D image is spatially invariant and that the image satisfies 0 boundary conditions, K will be triply Toeplitz with this banded structure.

If we assume reflexive boundary conditions, K will be triply Toeplitz+Hankel with special banding.

Matrix Approximation Problem Revisited

Note that if A, B, C are banded Toeplitz (or T+H), they are uniquely specified by their respective central columns, vectors a, b, c.

Also, $A \otimes B \otimes C$ is triply Toeplitz (T+H) with special banding structure.

Theorem 1

Let K be triply Toeplitz (banded) and P be the 3D tensor defining the central column of K. Let a_i, b_i, c_i be the central columns of banded Toeplitz matrices A_i, B_i, C_i, respectively. Then

$$\|K - \sum_{i=1}^{s} A_i \otimes B_i \otimes C_i\|_F = \|\bar{P} - \sum_{i=1}^{s} \bar{c_i} \circ \bar{b_i} \circ \bar{a_i}\|_F,$$

where the bar notation implies a (diagonal) weighting on the faces of P and the vectors a_i, b_i, c_i, and $\circ$ implies 3way outer product.

Theorem 2

Similar result when K is triply Toeplitz + Hankel (banded), except the weighting matrix is a very special upper triangular matrix.

Tensor approximation

$$\|K - \sum_{i=1}^{s} A_i \otimes B_i \otimes C_i\|_F = \|\bar{P} - \sum_{i=1}^{s} \bar{c_i} \circ \bar{b_i} \circ \bar{a_i}\|_F,$$

Computing optimal approximation to K requires computing the optimal rank-s approximation to the tensor $\bar{P}$.

- For $s = 1$, unique solution.
- For $s > 1$, is there a solution? (uniqueness requires mild constaints). Can we compute it?
- A good choice for small s may not be known a priori.
- There are no “orthogonality” constraints on the decomposition, and none needed for our application.

Tensor Decompositions

Choices:

- PARAFAC model (N-way Toolbox by R. Bro)
- HOSVD [de Lathauwer, de Moor and Vandewalle, '00]

$$\bar{P} = \sum_{i=1}^{r_1} \sum_{j=1}^{r_2} \sum_{k=1}^{r_3} \delta_{ijk} u_i \circ v_j \circ w_k \approx \sum_{m=1}^{s} \tilde{\delta}_{i_m} u_{i_m} \circ v_{i_m} \circ w_{i_m}.$$

- Other, possibilities with little degradation? Work with Perrone, Martin

Approximation

$$\|K - \sum_{i=1}^{s} A_i \otimes B_i \otimes C_i\|_F = \|\bar{P} - \sum_{i=1}^{s} \bar{c}_i \circ \bar{b}_i \circ \bar{a}_i\|_F$$

- Optimal tensor solution can be calculated if $s = 1$, used to construct A_i, B_i, C_i.
- Settle for HOSVD approximation for $s > 1$.
- To determine s, observe $|\tilde{\delta}_i|$. Gaps/decay for small values of s imply K is well approximated by small number of terms in the Kronecker sum.
- The HOSVD for the tensor should not be confused with the SVD of K.

Preconditioner

- Case $s = 1$, compute SVD from Kronecker product approximation directly and use it to construct M.
- Case $s > 1$, one more level of approximation required (skip details), but M still ends up specified in Kronecker form.

Compute/Storage Summary

For an $n \times n \times n$ image (an SVD of K would cost $O(n^9)$ flops, storage)

- Compute approximation/preconditioner: $<= O(n^4)$ flops
- Store a few columns of SVD for 3, $n \times n$ matrices. (minimal)
- Preconditioner application $O(n^3)$ for small k
- Matvec $O(n^3 \log(n))$

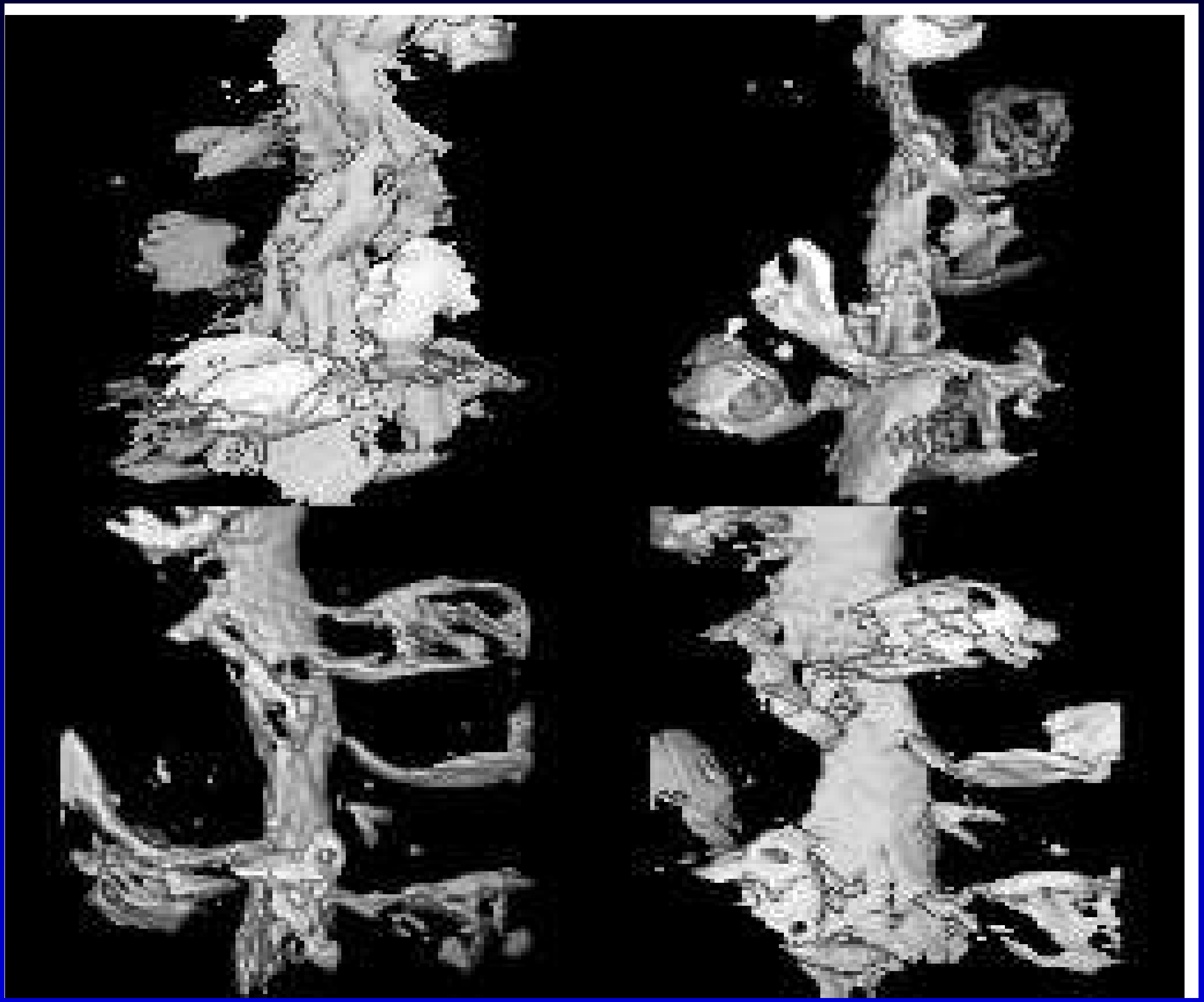
Example 1

Blurring effects caused by optical limits in 3D microscopy, on a stack of 20, 128×103 slices of a dendrite. Yields triply Toeplitz blurring operator.

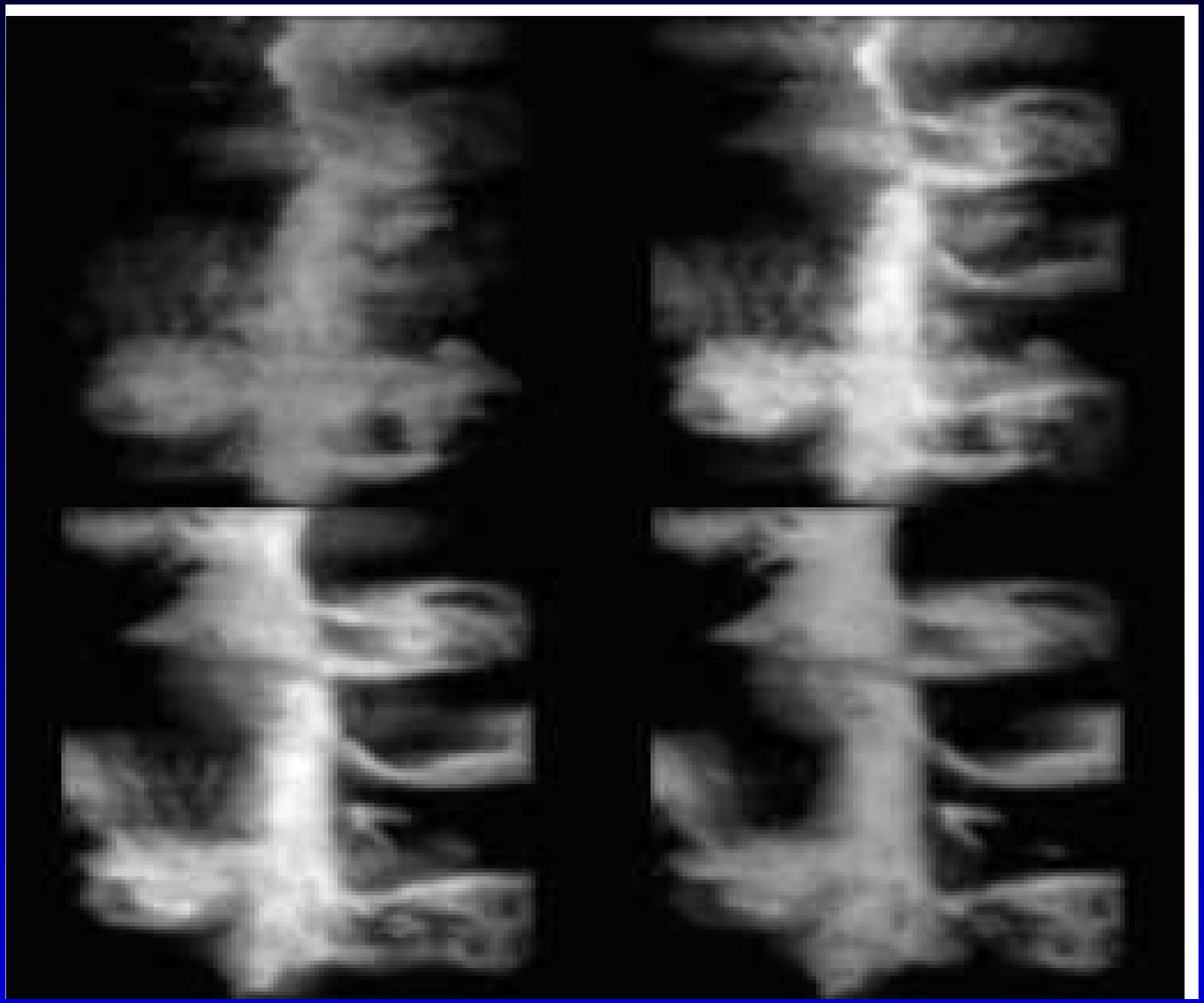
Both examples:

- Matlab 7
- RestoreTools [Nagy, Palmer, Perrone, ‘04]

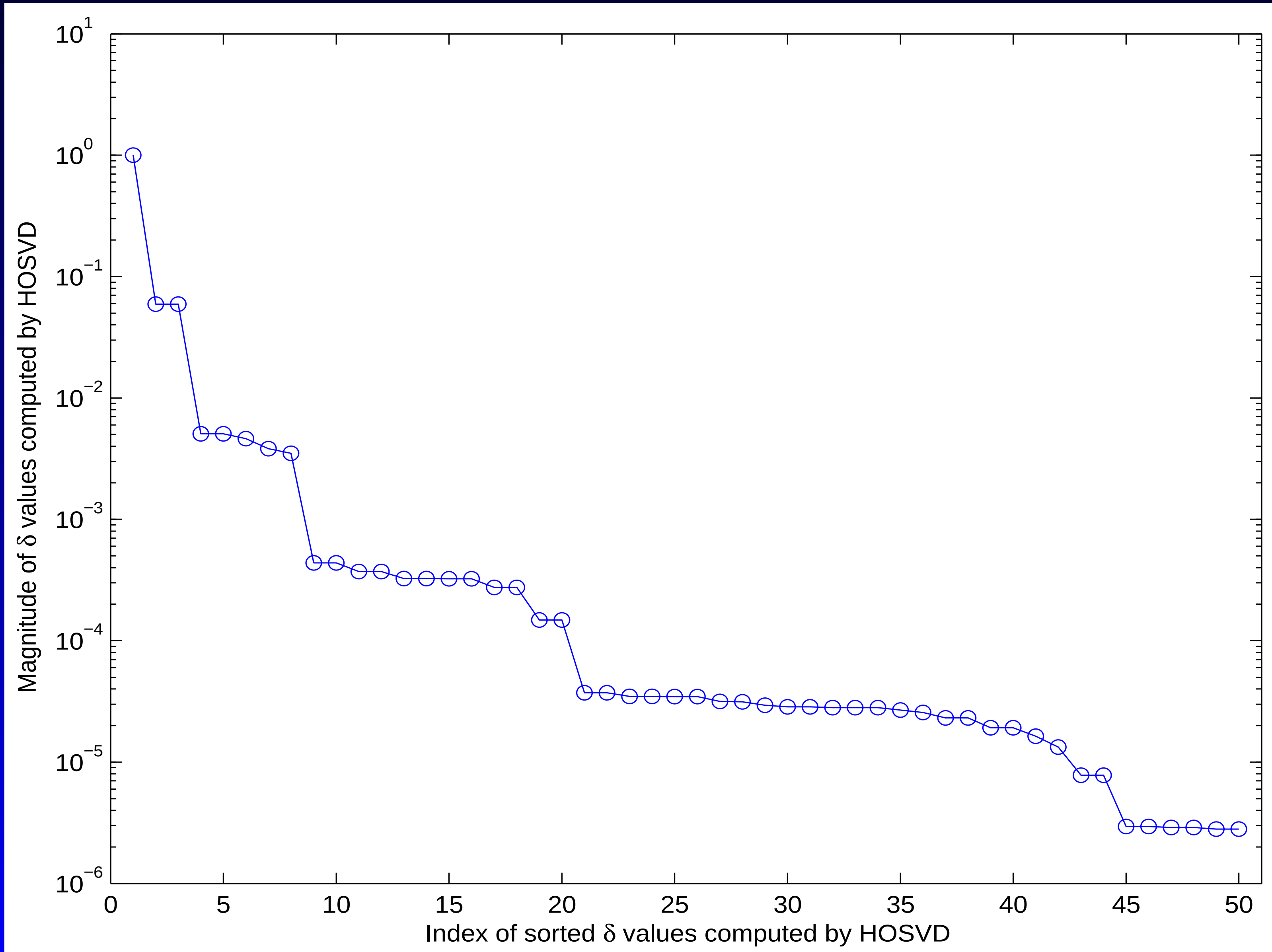
True Image Slices (1,7,13,19)



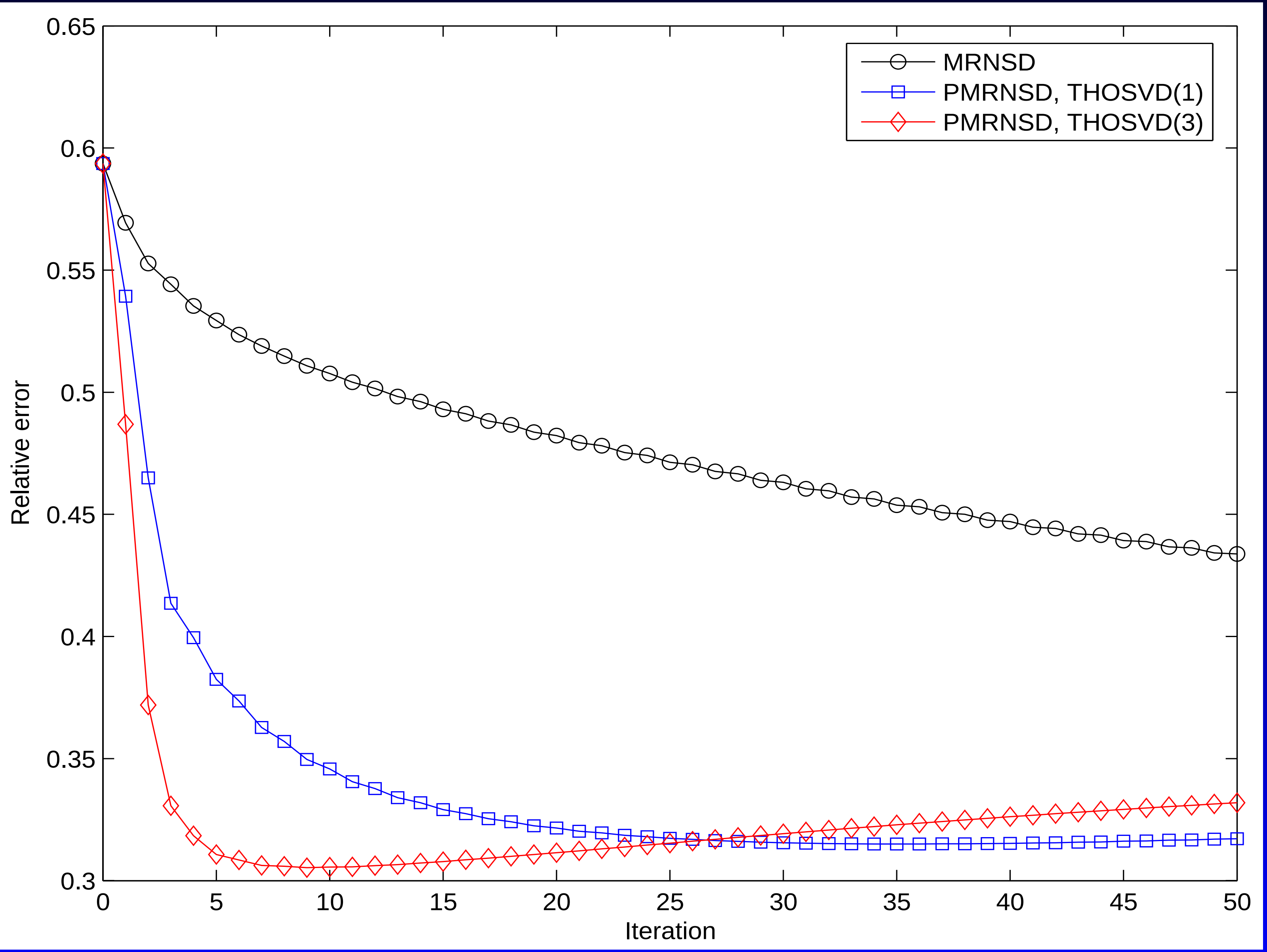
Blurred Image Slices



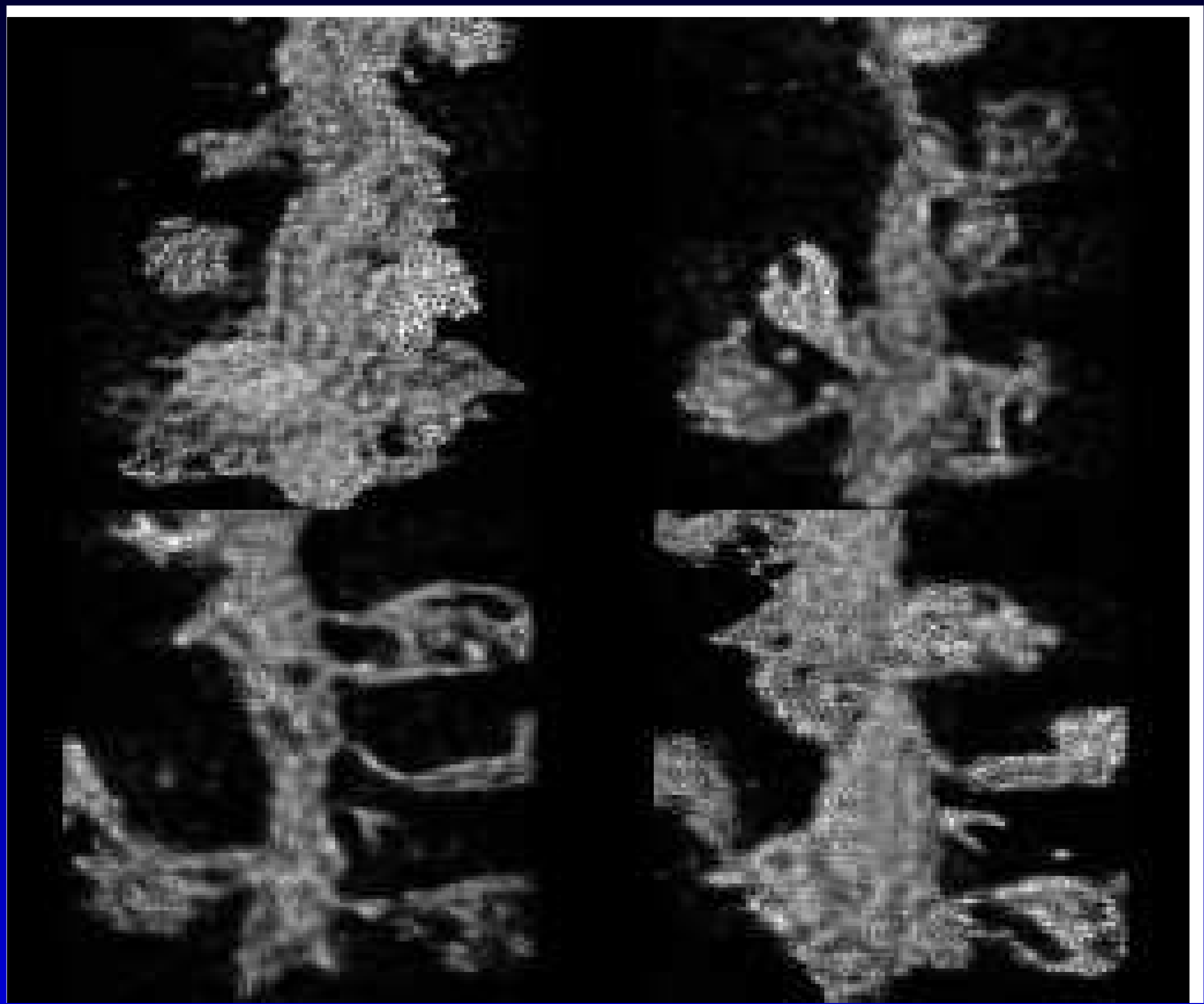
Largest 50 $|\tilde{\delta}_i|$



Relative Errors



Reconstructed Images



Example 2

Blurring effects caused by partial volume averaging in spiral CT

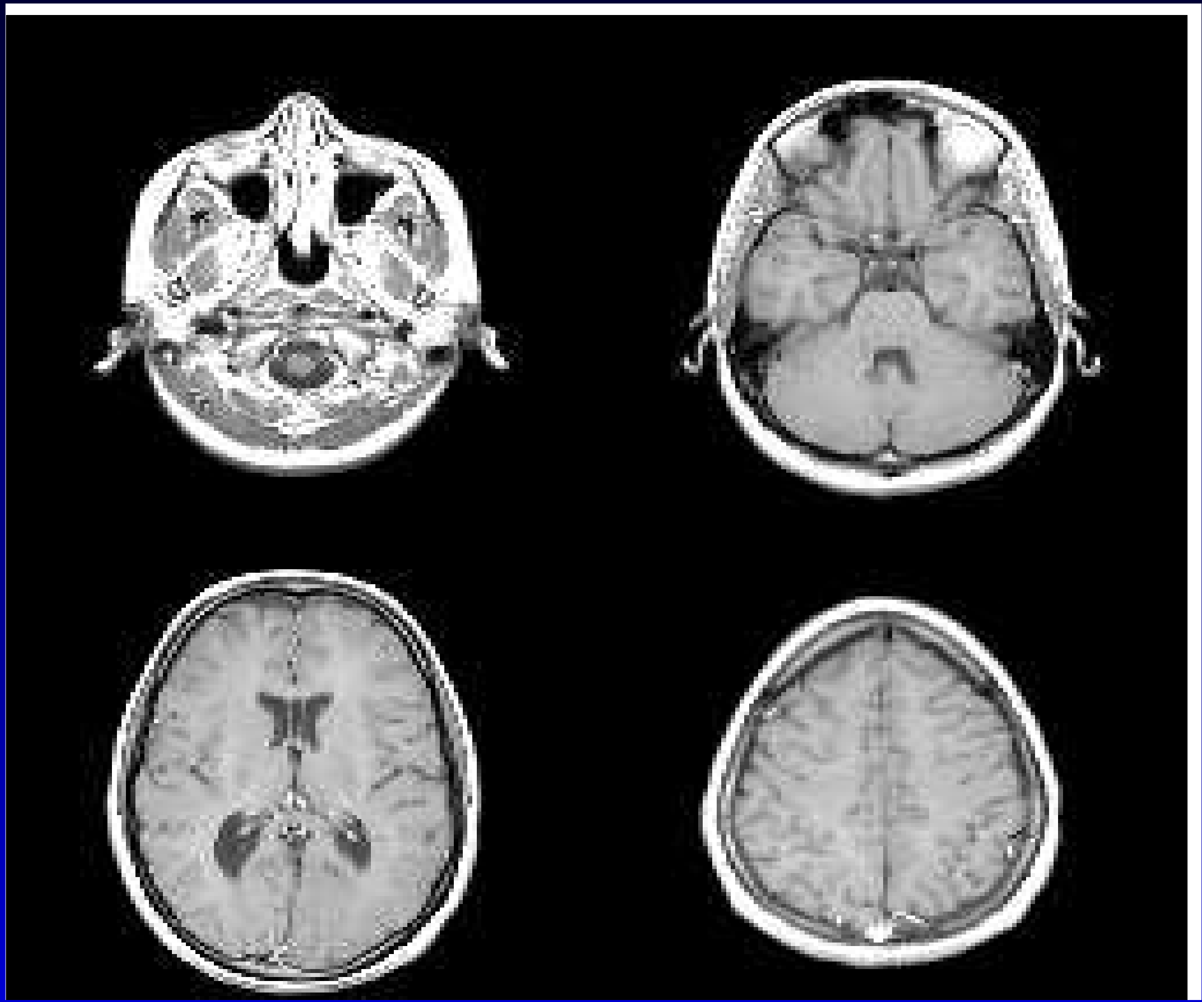
$$\kappa(x,y,z) = (\kappa_1(x,y,z) + \kappa_2(x,y,z) + \kappa_3(x,y,z))/3\,,$$

where

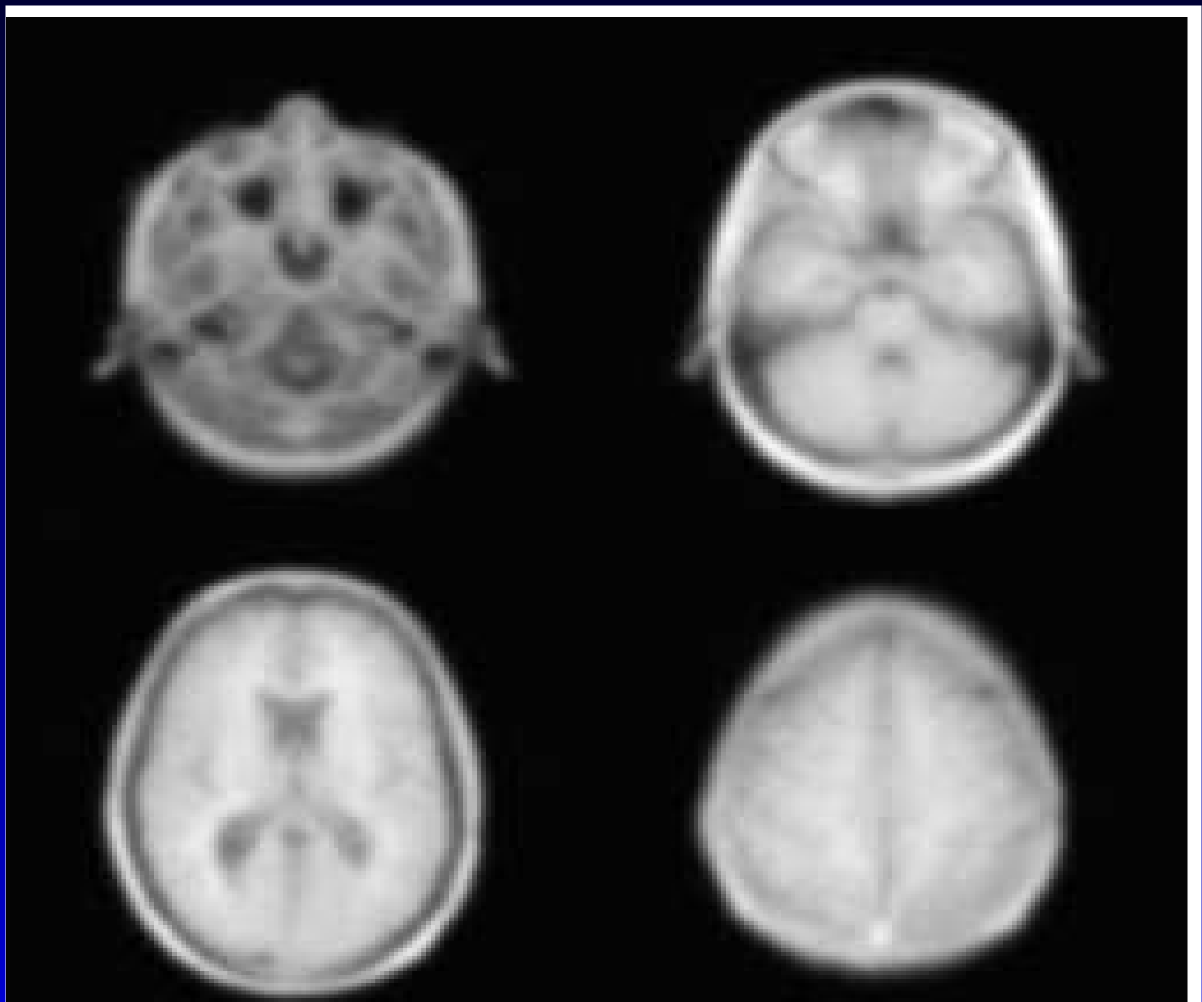
$$\kappa_i(x,y,z) = \frac{1}{\sqrt{(2\pi)^3}\sigma_i^3} e^{-(x^2+y^2+z^2)/2\sigma_i^2}\,,$$

and σ_i were chosen randomly with $1 \le \sigma_i \le 2$.

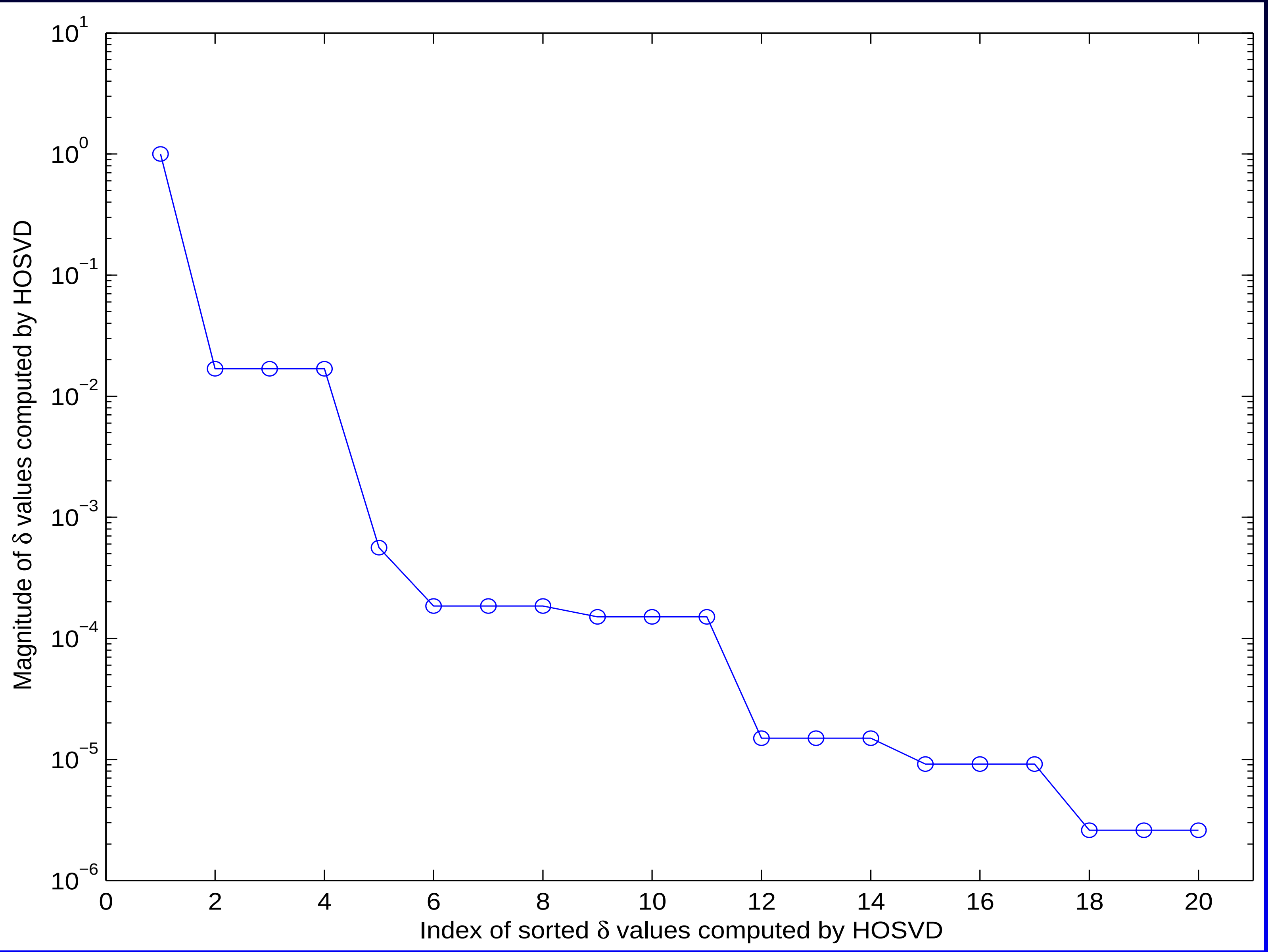
True Image Slices 1,8,15,22



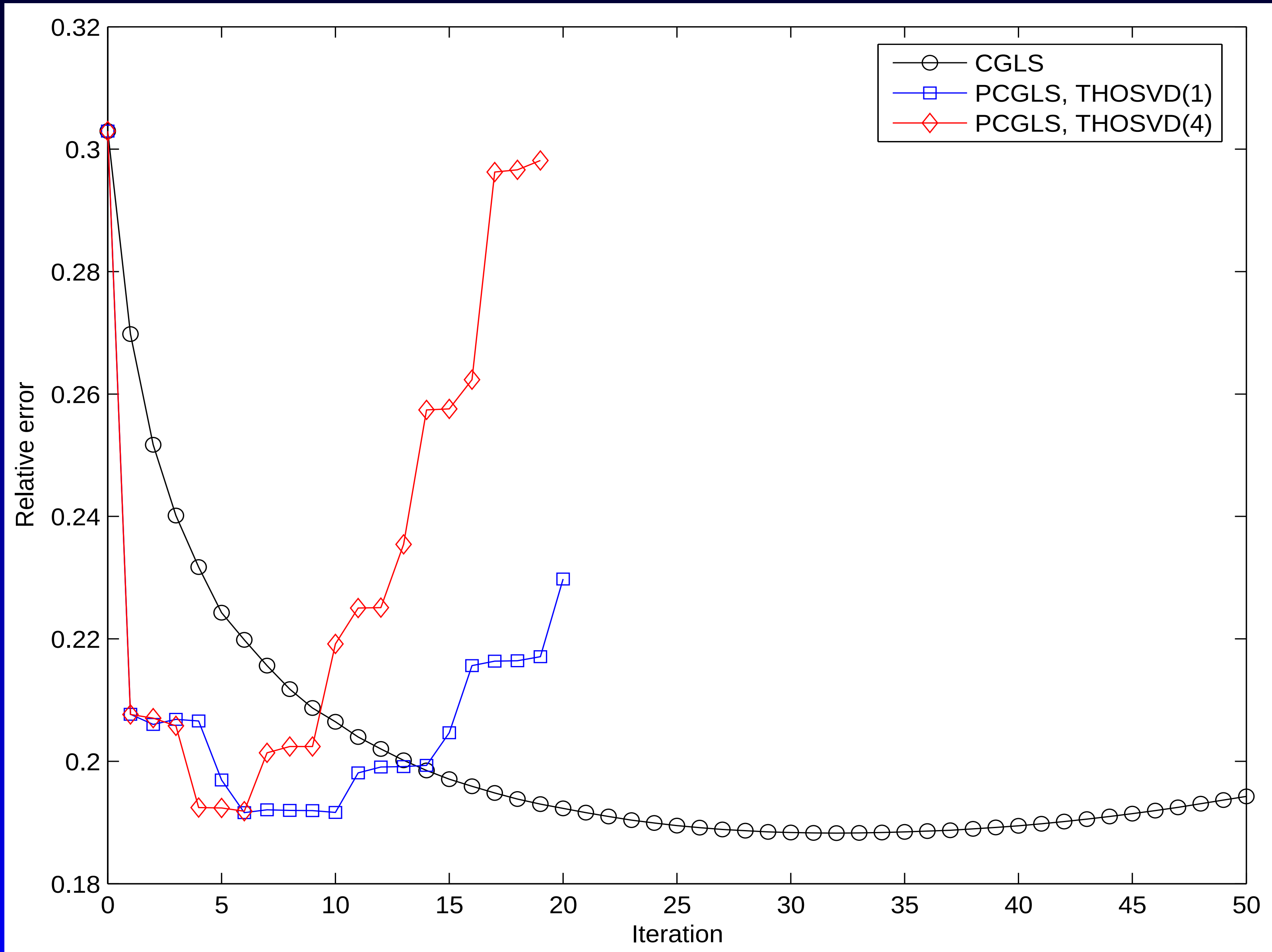
Blurred, Noisy Slices 1,8,15,22



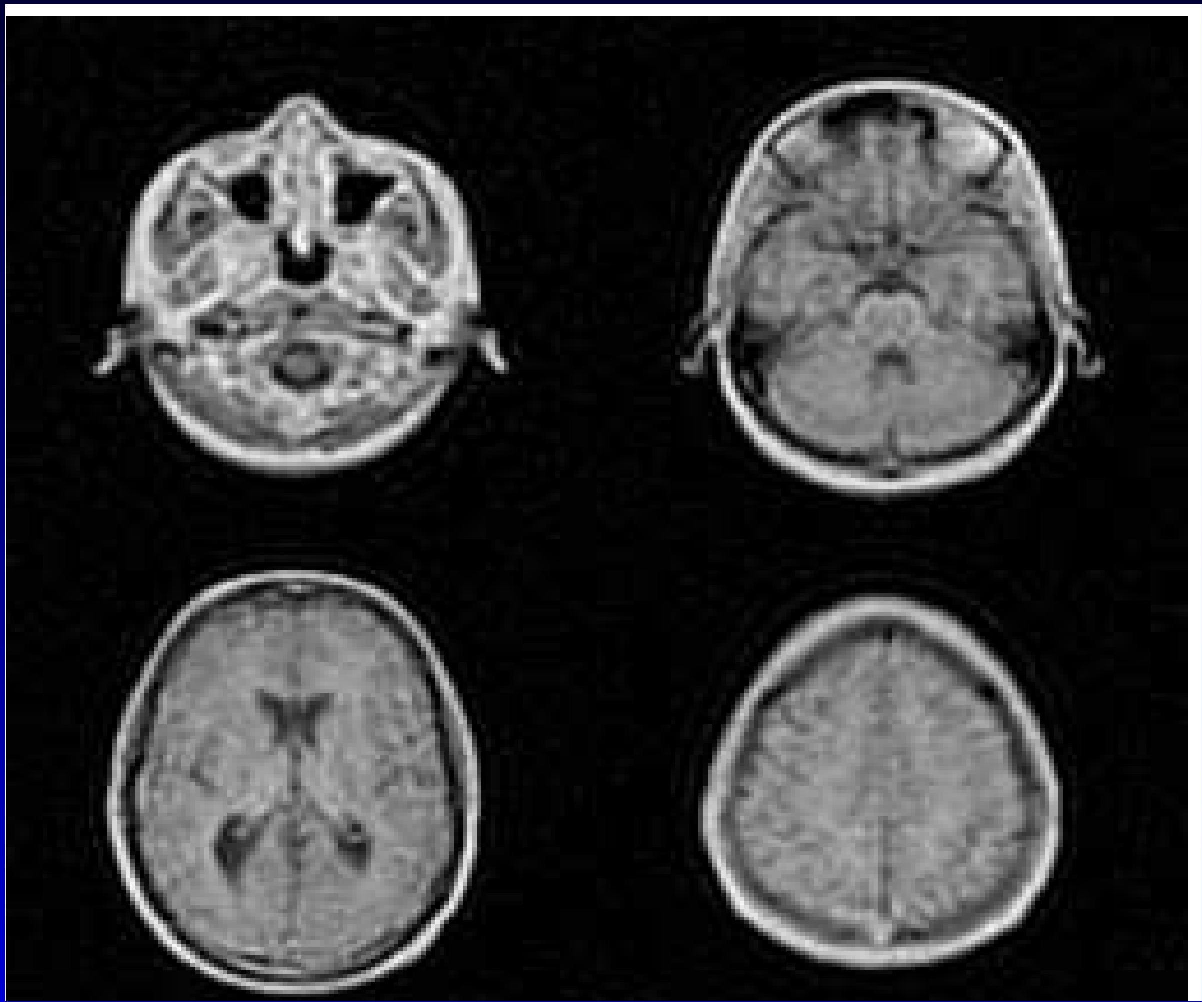
Largest $|\tilde{\delta}_i|$



Convergence History



Restored Images



Conclusions and Future Work

- Derived Kronecker product approximations to operators in image processing applications, used them to compute preconditioners making 3D deblurring computationally tractable.
- Optimal approximation if one term is used.
- Raises interesting questions about tensor decompositions, which are “best”.
- Extension to dense structured matrices (image processing, PDE’s).
- Kronecker approximations do provide useful basis for edge preserving regularization (work with Per Christian Hansen).