On the regularizing power of multigrid-type algorithms

Marco Donatelli

Stefano Serra Capizzano

Università dell'Insubria, Dip. Fisica e Matematica - Sede di Como
Via Valleggio 11 - 22100 Como, Italy
(donatelli@uninsubria.it)

Outline

- Image restoration using boundary conditions (BC)
- Spectral properties of the coefficient matrices
- Multi-Grid Methods (MGM)
- Two-Level (TL) regularization
- Multigrid regularization
- Numerical experiments
- Direct Multigrid regularization

Image restoration with BCs

Using boundary conditions (BC), the restored image $\mathbf{f}$ is obtained solving:

$$A\mathbf{f} = \mathbf{g} + \mathbf{n}$$

- $\mathbf{g}$ = blurred image
- $\mathbf{n}$ = noise (random vector)
- A = two-level matrix depending on PSF and BC

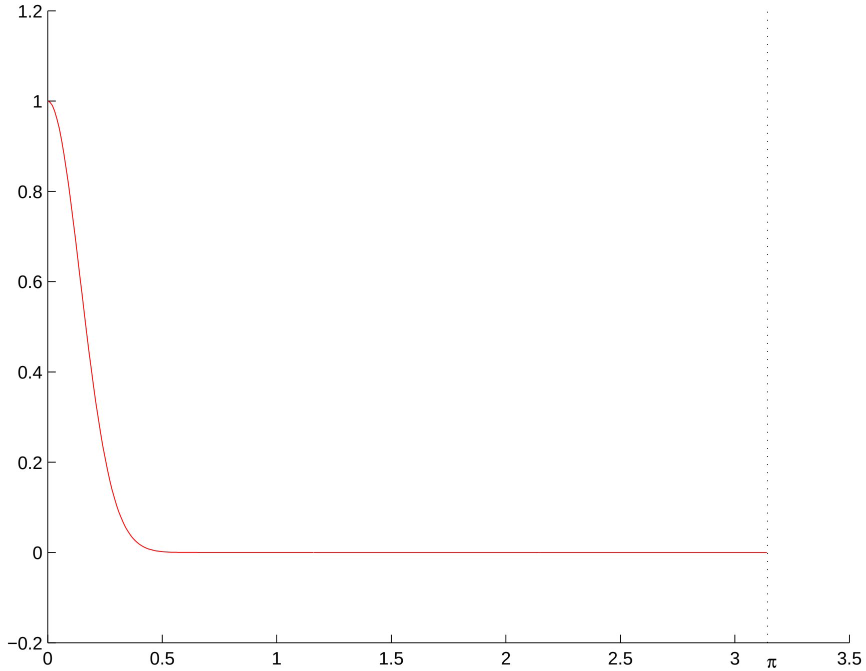
BC	A
Dirichlet	Toeplitz
periodic	circulant
Neumann (reflective)	DCT III
anti-reflective	DST I + low-rank

Generating function of PSF

- 1D problem with gaussian PSF:

$\mathbf{x} = -5 : 0.1 : 5$	101 points
$\mathbf{a} = \mathrm{e}^{-\mathbf{x}^2}$	PSF's coefficients
$\mathbf{a} = (a_{-50}, \ldots, a_0, \ldots, a_{50}),\ a_i = a_{-i}$	
$z(y) = \sum_{i=-50}^{50} a_i \mathrm{e}^{-\mathrm{i}y}$	generating function

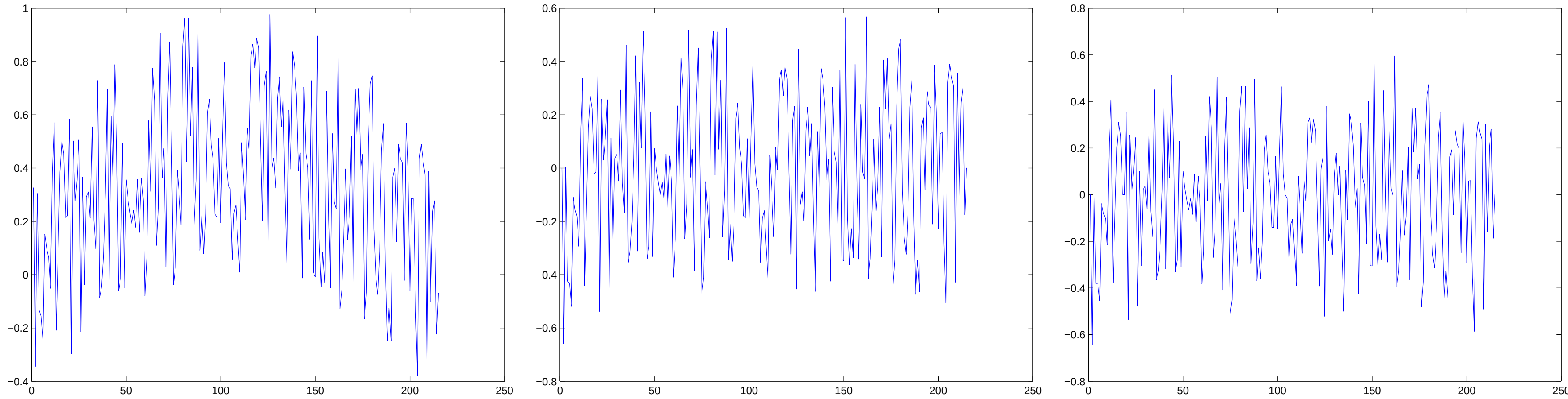
The eigenvalues of $A(z)$ are about a uniform sampling of z in $[0, \pi]$



- The ill-conditioned subspace is mainly constituted by the high frequencies.

Smoothing

- Iterative regularizing methods (e.g. Landweber, CG, ...) firstly reduce the error in the low frequencies (well-conditioned subspace).
- Example: $\mathbf{f} = \sin(x), \; x \in [0, \pi]$ and $\mathbf{g} = A\mathbf{f}$. Solving the linear system $A\tilde{\mathbf{f}} = \mathbf{g}$ by Richardson



Initial error | After 1 iteration | After 5 iterations

- The error is highly oscillating after ten iterations as well.

Multigrid structure

- Idea: project the system in a subspace of lower dimension, solve the resulting system in this space and interpolate the solution in order to improve the previous approximation in the greater space.
- The j-th iteration of the Two-Grid Method(TGM) for the system $A\mathbf{x} = \mathbf{b}$:

$$\boxed{\begin{array}{l} \textbf{(1)}\ \tilde{\mathbf{x}} = Smooth(A, \mathbf{x}^{(j)}, \mathbf{b}, \nu) \\ \textbf{(2)}\ \mathbf{r}_1 = \mathbf{P}(\mathbf{b} - A\tilde{\mathbf{x}}) \\ \textbf{(3)}\ A_1 = \mathbf{P}A\mathbf{P}^H \\ \textbf{(4)}\ \mathbf{e}_1 = A_1^{-1}\mathbf{r}_1 \\ \textbf{(5)}\ \mathbf{x}^{(j+1)} = \mathbf{x}^{(j)} + \mathbf{P}^H\mathbf{e}_1 \end{array}}$$

- Multigrid (MGM): the step (4) becomes a recursive application of the algorithm.

Algebraic Multigrid (AMG)

- The AMG uses information on the coefficient matrix and no geometric information on the problem.
- Different classic smoothers have a similar behavior: in the initial iterations they are not able to reduce effectively the error in the subspace generated by the eigenvectors associated to small eigenvalues (ill-conditioned subspace)

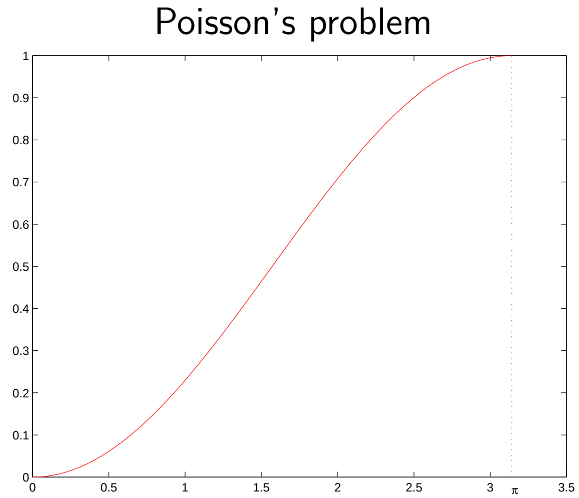
$$\Downarrow$$

  the projector is chosen in order to project the error equation in such subspace.
- A good choice for the projector leads to MGM with a rapid convergence.
- For instance, for Toeplitz and algebra of matrices, see [Aricò, Donatelli, Serra Capizzano, SIMAX, Vol. 26–1 pp. 186–214.].

Geometric Multigrid

- The MGM is an optimal solver for elliptic PDE

For elliptic PDE the ill-conditioned subspace is made by low frequencies (complementary with respect to the gaussian blur).



- For the projector a simple and powerful choice is:

$$\mathbf{P} = \frac{1}{4}\begin{bmatrix} 1 & 2 & 1 & & & & \\ & & 1 & 2 & 1 & & \\ & & & & \ddots & \ddots & \\ & & & & 1 & 2 & 1 \end{bmatrix} \qquad \text{(full weighting)}$$

$$\mathbf{P}^T = 2\mathbf{P} \qquad \text{(linear interpolation)}$$

Image restoration and Multigrid

- In the images deblurring the ill-conditioned subspace is related to high frequencies, while the well-conditioned subspace is generated to low frequencies.
- In order to obtain a rapid convergence the algebraic multigrid projects in the high frequencies where the noise “lives” $\Longrightarrow$ noise explosion already at the first iteration (it requires Tikhonov regularization [NLAA in press]).
- In this case the geometric multigrid projects in the well-conditioned subspace (low frequencies) $\Longrightarrow$ it is slowly convergent but it can be a good iterative regularizer.

If we have an iterative regularizing method we can improve its regularizing property using it as smoother in a Multigrid algorithm.

Projector structure

- In order to apply recursively the MGM it is necessary to maintain the same structure at each level (Toeplitz, circulant, ...).
- Projector: $P_i = K_{N_i} T_{N_i}(2 + 2\cos(x))$ s.t. i is the recursion level and

$$T_{N_i}(2+2\cos(x)) = \begin{pmatrix} 2 & 1 & & \\ 1 & 2 & \ddots & \\ & \ddots & \ddots & 1 \\ & & 1 & 2 \end{pmatrix}_{N_i \times N_i}$$

	circulant	Toeplitz & DST – I	DCT – III
$\mathbf{K_{N_i}} \in \mathbb{R}^{N_{i-1} \times N_i}$	$\begin{bmatrix} 1 & 0 & & & & \\ & & 1 & 0 & & \\ & & & \ddots & \ddots & \\ & & & & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 & & & & \\ & & 0 & 1 & 0 & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & 0 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 0 & & & & \\ & & 1 & 1 & 0 & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & 0 & 1 & 1 \end{bmatrix}$

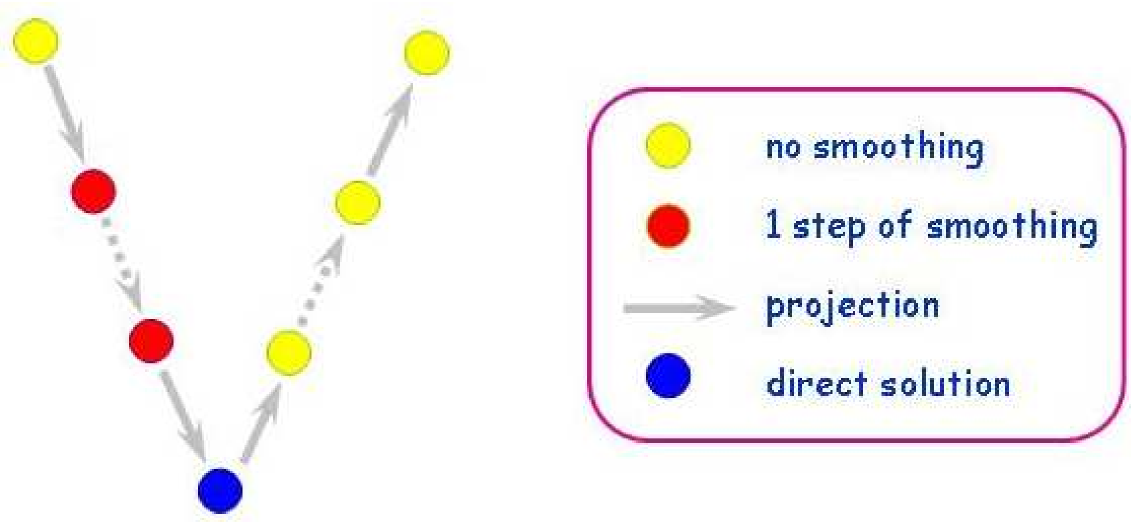
Two-Level (TL) regularization

- Two-Level (TL) regularization (specialization of the TGM):
  1. No smoothing at step (1): $\tilde{\mathbf{x}} = \mathbf{x}^{(j)}$
  2. Step (4): $\mathbf{e}_1 = A_1^{-1}\mathbf{r}_1 \to Smooth(A_1, \mathbf{e}_1, \mathbf{r}_1, \nu)$

  As smoother a generic regularizing method can be used.
- Since in the finer grid we do not apply the smoother we can project the system $A\mathbf{x} = \mathbf{b}$ instead of the error equation $A\mathbf{e} = \mathbf{r}$.
- The $\mathbf{P}$ = full weighting applied to the observed image $\mathbf{b}$ leads to a reblurring effect followed by a down-sampling (noise damping like a low-pass filter).
- The $\mathbf{P}^T$ = linear interpolation reconstruct exactly the piecewise linear function damping the high oscillation deriving by the noise.

Multigrid regularization

- Applying recursively the Two-Level algorithm, we obtain a Multigrid method.
- V-cycle



- Using a greater number of recursive calls (e.g. W-cycle), the algorithm "works" more in the well-conditioned subspace but it is more difficult to define an early stopping criterium.

Computational cost

- Let $n_0 \times n_0 = n \times n$ be the problem size at the finer level, where $n_0 = n = 2^\alpha$, $\alpha \in \mathbb{N}$, thus at the level j the problem size is $n_j \times n_j$ where $n_j = 2^{\alpha - j}$.
- *Projection* $j \longrightarrow j+1$: $\frac{7}{4} n_j^2$ flops. *Interpolation* $j+1 \longrightarrow j$: $\frac{7}{8} n_j^2$ flops.
- Let $W(n)$ be the computational cost of one smoother iteration for a problem of size $n \times n$ with $W(n) = cn^2 + O(n)$, $c \gg 1$.
  The computational cost at the j-th level is about

$$c_j = W(n_j) + \frac{21}{8} n_j^2 \text{ flops.}$$

- The total cost of one MGM iteration is:

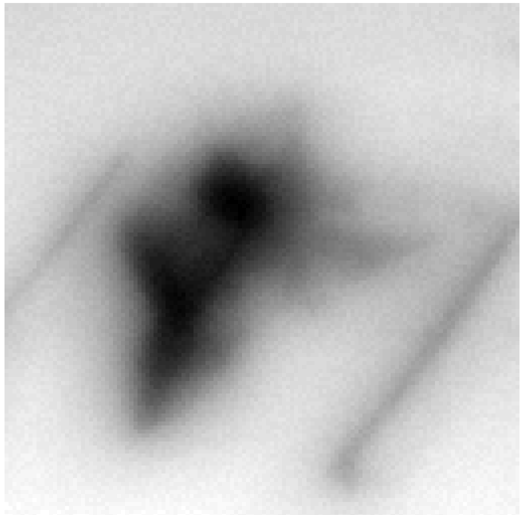
$$\frac{21}{8} n^2 + \sum_{j=1}^{\log_2(n)-1} c_j \; < \; 4n^2 + \frac{4}{3} W\left(\frac{n}{2}\right) \approx \frac{1}{3} W(n).$$

Example 1 (airplane)

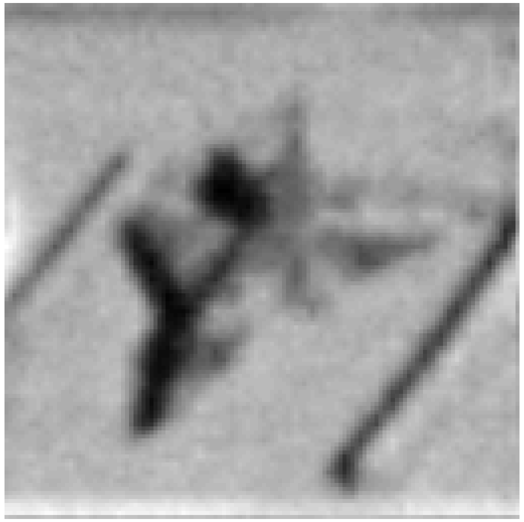
- Periodic BCs
- Gaussian PSF (A spd)
- SNR $= 100$

Original Image

Inner part 128×128



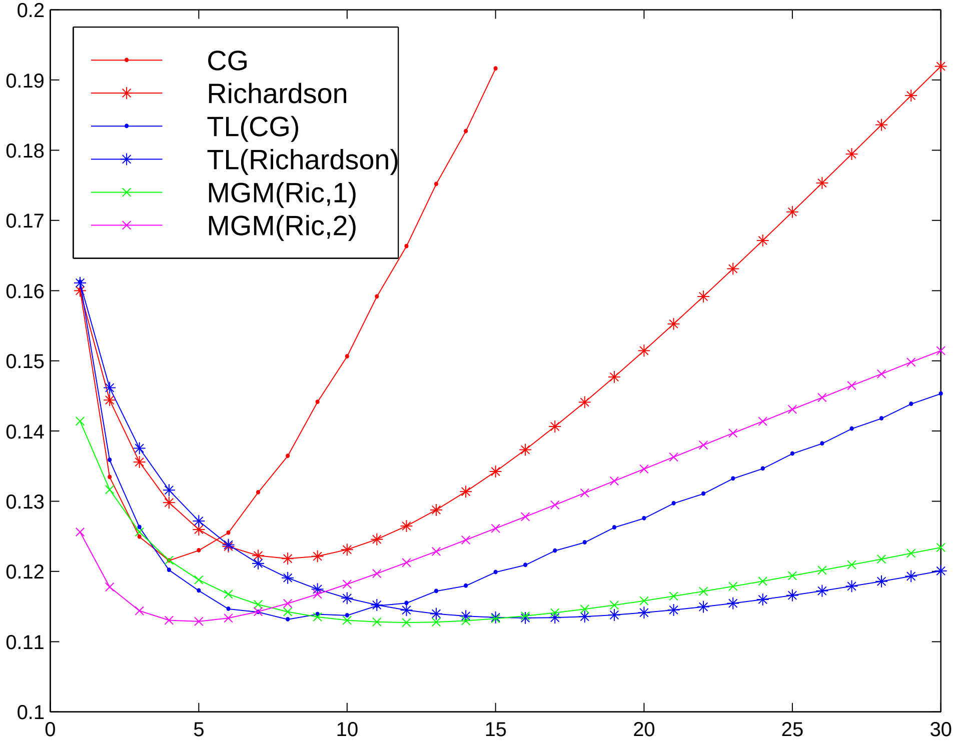
Blurred + SNR $= 100$



Restored with MGM

Restoration error (example 1)

Graph of the relative restoration error $e_j = \|\bar{\mathbf{f}} - \mathbf{f}^{(j)}\|_2 / \|\bar{\mathbf{f}}\|_2$ increasing the number of iterations when solving $A\mathbf{f} = \mathbf{g} + \mathbf{n}$ (RichN = Landweber, CGN = CG for normal equations).



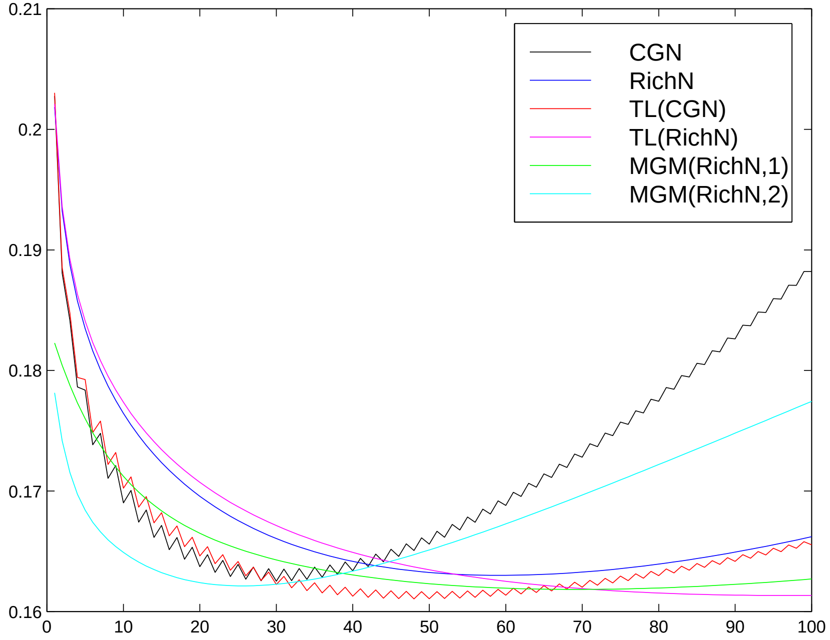
Relative error vs. number of iterations

Method	$\min_{j=1,\ldots} (e_j)$	$\arg\min_{j=1,\ldots} (e_j)$
CG	0.1215	4
Richardson	0.1218	8
TL(CG)	0.1132	8
TL(Rich)	0.1134	16
MGM(Rich, 1)	0.1127	12
MGM(Rich, 2)	0.1129	5
CGN	0.1135	178
RichN	0.1135	352

Minimum restoration error

Example 2 (SNR = 10)

- Same image and PSF but much more noise: SNR $= 10$.
- For CG and Richardson, it is necessary to resort to normal equations.



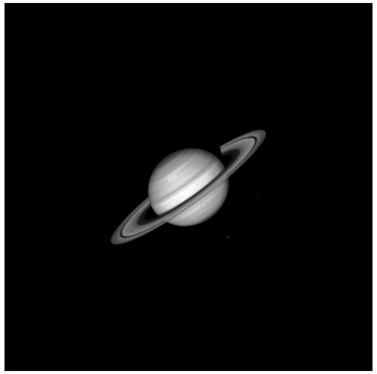
Relative error vs. number of iterations

Method	$\min_{j=1,\dots} (e_j)$	$\arg\min_{j=1,\dots} (e_j)$
CGN	0.1625	30
RichN	0.1630	59
TL(CGN)	0.1611	48
TL(RichN)	0.1613	97
MGM(RichN,1)	0.1618	69
MGM(RichN,2)	0.1621	26
MGM(Rich,1)	0.1648	3
MGM(Rich,2)	0.1630	1

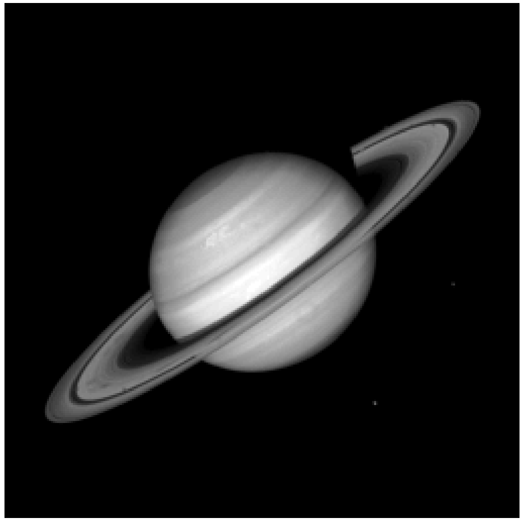
Minimum relative error

Example 3 (Saturn)

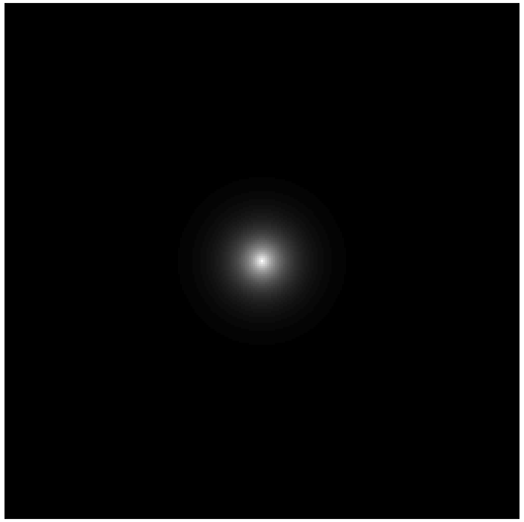
- Periodic BCs (exacts)
- Gaussian PSF ($\lambda(A) \approx -10^{-4}$)
- SNR $=$ 50



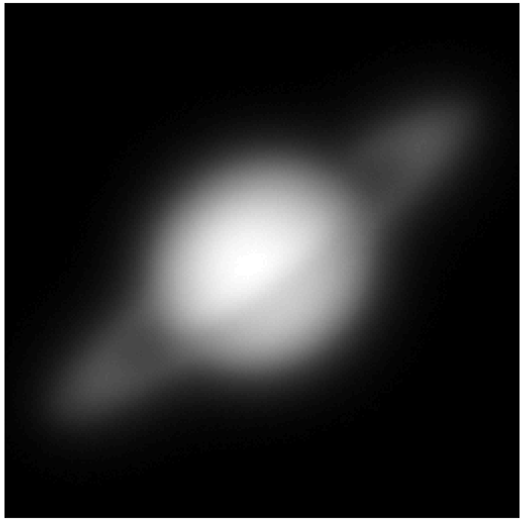
Original image



Inner part 128×128



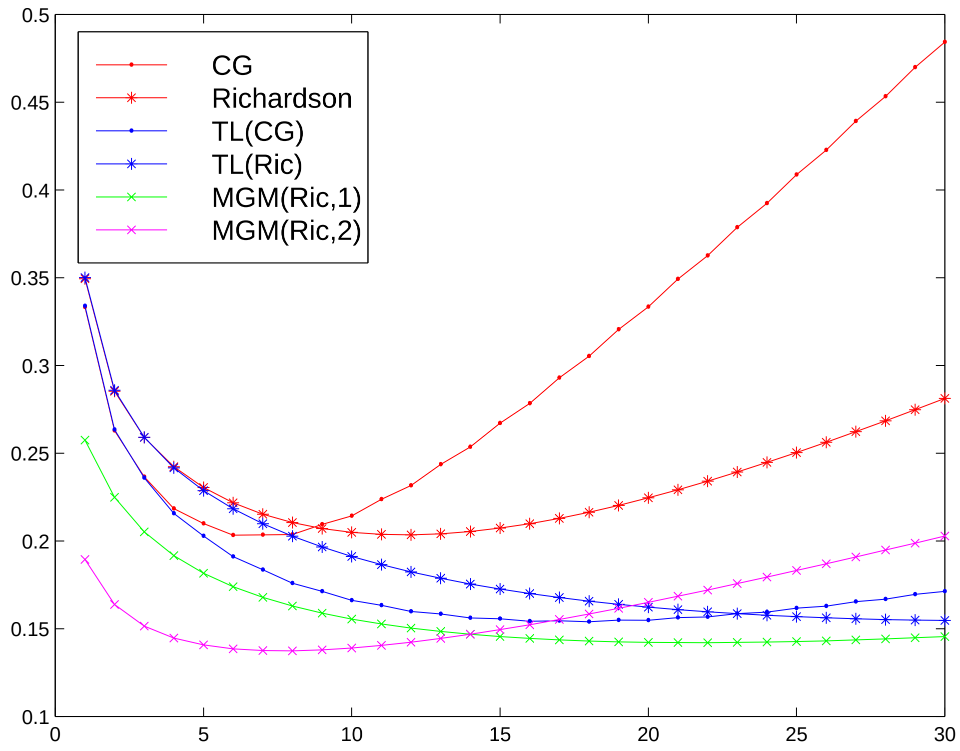
PSF



Blurred $+$ SNR $=$ 50

Restoration error (example 3)

Graph of the relative restoration error $e_j = \|\bar{\mathbf{f}} - \mathbf{f}^{(j)}\|_2 / \|\bar{\mathbf{f}}\|_2$ increasing the number of iterations when solving the linear system $A\mathbf{f} = \mathbf{g} + \mathbf{n}$.

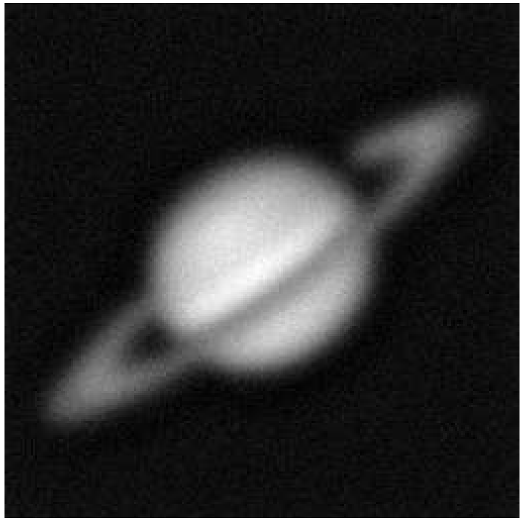


Relative error vs. number of iterations

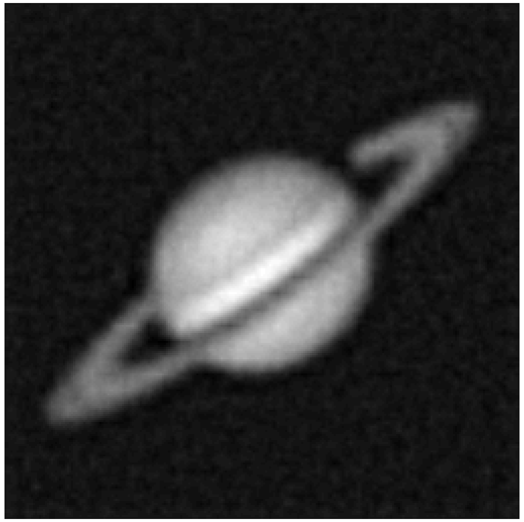
Method	$\min_{j=1,\ldots} (e_j)$	arg $\min_{j=1,\ldots} (e_j)$
CG	0.2033	6
Richardson	0.2035	12
TL(CG)	0.1539	18
TL(Rich)	0.1547	30
MGM(Rich,1)	0.1421	22
MGM(Rich,2)	0.1374	8
CGN	0.1302	2500
MGM(CGN,1)	0.1297	250
MGM(RichN,2)	0.1305	1700

Minimum relative error

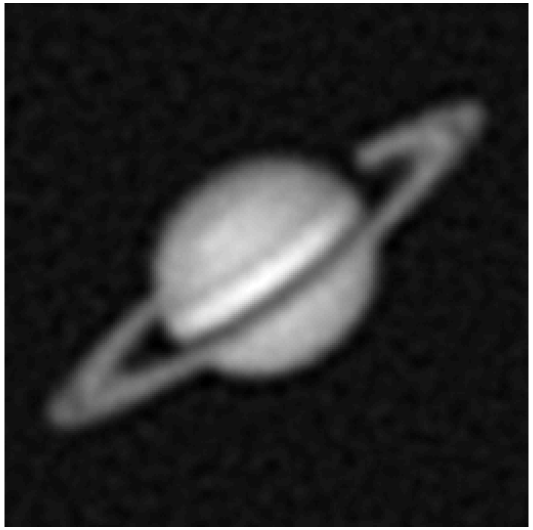
Restored images



CG



MGM(Rich,2)

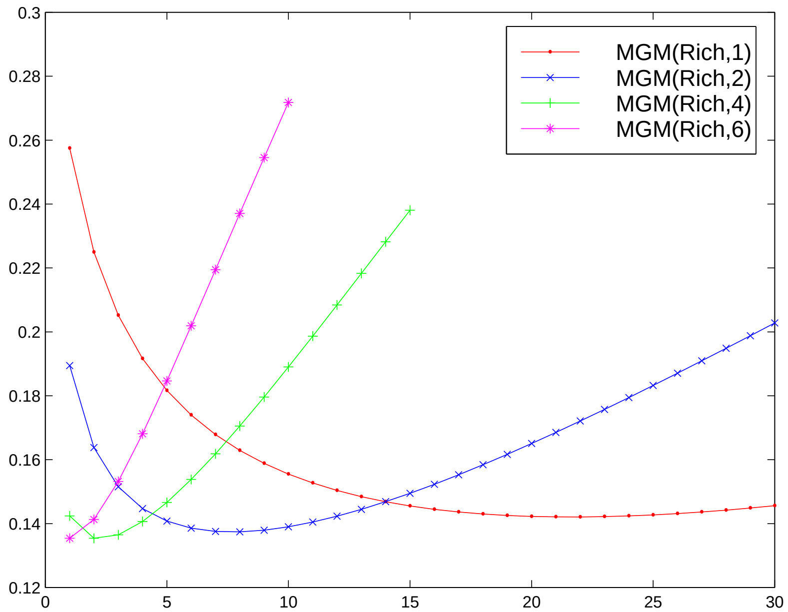


CGN

	CG	MGM(Rich,2)	CGN (normal equation)
Minimum error	0.2033	0.1374	0.1302
Number of iterations	6	8	2500

Direct multigrid regularization

- Trend of the error after only one iteration of MGM(Rich,γ) varying γ.
- It is a direct regularization method with regularization parameter γ.



Relative error vs. number of iterations

γ	e_1
1	0.25747
2	0.18944
3	0.15723
4	0.14241
5	0.13658
6	0.13543
7	0.13674
8	0.13947

The CGN reaches e minimum equal to 0.1302 after 2500 iterations

- The computational cost increase with γ but not so much (e.g. $\gamma = 8 \Rightarrow O(N^{1.5})$).

Conclusions

- The Multigrid (with a regularizing method as smoother) is a good regularizer $\Rightarrow$ we can improve the power of an iterative regularizing method using it as smoother inside a MGM scheme.
- The MGM regularization is robust for small negative eigenvalues as well.
- Usually it is not necessary to resort to normal equations.
- It can lead to several generalizations.