

A MULTIGRID FOR IMAGE DEBLURRING WITH TIKHONOV REGULARIZATION ^{*}

M. DONATELLI[†]

Abstract. In the resolution of certain image deblurring problems with given boundary conditions we obtain two-level structured linear systems. In the case of shift-invariant point spread function with Dirichlet (zero) boundary conditions, the blurring matrices are block Toeplitz matrices with Toeplitz blocks. If the periodic boundary conditions are used, then the involved structures become block-circulant-circulant-blocks. Furthermore Gaussian-like point spread functions usually lead to numerically banded matrices which are ill-conditioned since they are associated to generating functions that vanish in a neighborhood of (π, π) . We solve such systems by applying a multigrid method. The proposed technique shows an optimality property, i.e., its cost is of $O(N)$ arithmetic operations (like matrix-vector product), where N is the size of the linear system. In the case of images affected by noise we use two Tikhonov regularization techniques to reduce the noise effects.

Key words. Point spread function (PSF); Toeplitz and circulant matrices; ill-conditioning; multigrid methods; Tikhonov and Riley regularization.

1. INTRODUCTION. Basically, image deblurring problems lead to the resolution of a minimization problem. If \mathbf{g} denotes the observed image, the original image \mathbf{f} can be resumed by solving the minimum least square problem

$$\min_{\mathbf{f} \in \mathbf{R}^N} \left\| \tilde{B}\mathbf{f} - \mathbf{g} \right\|_2, \quad (1.1)$$

where the rectangular matrix \tilde{B} denotes the blurring operator, coming from the discretization of the Point Spread Function (PSF). In this work we consider shift-invariant PSF like Gaussian function endowed with a numerical support which is much smaller than the size of the image and rapidly decreasing to zero. The formation of the blurred image \mathbf{g} needs of data outside the area spanned by \mathbf{f} , hence we need to introduce certain assumptions on the unknown boundary data of \mathbf{f} called boundary conditions (BCs), which lead to a square system $B\mathbf{f} = \mathbf{g}$. Clearly B depends on the choice of the BCs.

In literature [3] Dirichlet boundary conditions (D-BCs) and periodic BCs (P-BCs) are widely used. With the D-BCs a null (black) border is added outside the edges of the observed image (they are well suited in astronomical image processing). In this case, since the PSF is shift-invariant, the matrix $B = T_{n_1, n_2}(z)$ is a block Toeplitz matrix with Toeplitz blocks (BTTB), possibly banded at each level if the PSF shows compact support or numerically banded in the general case. The symbol $T_{n_1, n_2}(z)$ denotes a $n_1 \times n_1$ blocks Toeplitz matrix with $n_2 \times n_2$ Toeplitz blocks, which can be explicitly represented as

$$T_{n_1, n_2}(z) = \begin{pmatrix} A_0 & A_{-1} & \dots & A_{1-n_2} \\ A_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & A_{-1} \\ A_{n_1-1} & \dots & A_1 & A_0 \end{pmatrix}, \quad (1.2)$$

^{*}This is a preprint of an article published in Numer. Linear Algebra Appl., 12 (2005), pp. 715–729. Copyright ©Wiley.

[†]Dipartimento di Fisica e Matematica, Università dell’Insubria - Sede di Como, Via Valleggio 11, 22100 Como, Italy

$$A_j = \begin{pmatrix} a_{j,0} & a_{j,-1} & \cdots & a_{j,1-n_2} \\ a_{j,1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{j,-1} \\ a_{j,n_2-1} & \cdots & a_{j,1} & a_{j,0} \end{pmatrix}$$

for $j = 1 - n_1, \dots, n_1 - 1$ and such that the values $a_{j,k}$ are the Fourier coefficients of the bivariate generating function z : $a_{j,k} = \frac{1}{4\pi^2} \int_{\Omega} z(x,y) e^{-i(jx+ky)} dx dy$, $i^2 = -1$, $\Omega = [-\pi, \pi] \times [-\pi, \pi]$. If the function z is real and even with respect to both its variables, then T_{n_1, n_2} is doubly symmetric, i.e. block symmetric with symmetric blocks. With P-BCs the image is periodically extended outside the edges. Therefore, the matrix B becomes block circulant with circulant blocks (BCCB). If F_n denotes the Fourier matrix of dimension n , a circulant matrix $C_n(f)$ can be factorized as $C_n(f) = F_n D_n(f) F_n^H$, where $D_n(f) = \text{diag}_{0 \leq j \leq n-1} (f(x_j))$ with $x_j = 2\pi j/n$. Consequently $C_n(f)$ can be diagonalized using fast Fourier transforms, with $O(n \log(n))$ complex operations also in the banded case. Furthermore, denoting with a_j the Fourier coefficient of f , $C_n(f)$ can be also expressed as $\sum_{j=-\lfloor \frac{n-1}{2} \rfloor}^{\lfloor \frac{n}{2} \rfloor} a_j Z_n^j$, where $(Z_n)_{s,t} = \delta(t - s + 1) \bmod n$ and δ is the Kronecker delta ($\delta(\alpha) = 1$ if $\alpha = 0$ and $\delta(\alpha) = 0$ elsewhere).

Recently, new types of BCs, namely Neumann and anti-reflective BCs, have been proposed [19, 15]. They provide a better reconstruction than the previous popular BCs and a lower computational cost based on discrete cosine or sine transforms.

We underline that image deconvolution leads to very large linear systems. Therefore it is desirable to solve them using iterative methods, e.g. Conjugate Gradient (CG), even suitably preconditioned. In this paper we propose a new approach, based on a multigrid method (MGM) for algebraic problems defined in [1]. In this last work is proven its optimality for monodimensional algebra matrices like circulant and tau (matrices diagonalized by sine transform), while its optimality for Toeplitz matrices is experimentally observed. In [2] the previous theoretical results are extended to the multidimensional case. These works take origin from two previous papers by Fiorentino and Serra Capizzano [11, 12] which led to several generalization for multilevel Toeplitz matrices, see e.g. [13, 8, 9]. For instance the MGM proposed in [9] considers generating functions with zeros of order almost two and it is proved to give a level-independent convergence rate. We remark that in the case of zeros of order almost two our MGM is very the same of those proposed in [9], hence our MGM can be considered an extension of this proposal and more in general of that in [12]. Furthermore, in [6] a Two-Grid method is proposed for the algebra of matrices diagonalized by cosine transforms. Therefore, even if in this paper we consider only D-BCs and P-BCs, a similar strategy can be applied to Neumann and anti-reflective BCs. We stress that the considered MGM is useful for speeding up the convergence of virtually any classic iterative method used as smoother requiring low computational costs.

We briefly discuss the computational cost in the resolution of linear systems arising in image deconvolution with different BCs. The experimental PSFs are usually generated from a ill-conditioned bivariate trigonometric polynomial whose largest degree is much lower than n_1 and n_2 . Therefore the corresponding matrices are ill-conditioned and (doubly) banded, i.e. banded at the block level with banded blocks and the matrix-vector product requires $O(N)$ operations where $N = n_1 n_2$. The inversion of BTTB is known to be very expensive, e.g. the fast direct solvers requires

$O(N^2)$ operations [14], while, concerning the iterative solvers, the most popular preconditioning strategies are far from being optimal in the polynomially ill-conditioned case [18]. The discrete Fourier, cosine, and sine transforms require $O(N \log(N))$ complex or real operations also in the banded case (divide and conquer algorithms not based on Fourier/trigonometric transforms can be applied in linear time for (banded) matrix algebra linear systems but not for multilevel (banded) Toeplitz systems). The MGM that we propose in this paper is optimal also with polynomial ill-conditioned linear systems, furthermore in the banded case it requires only $O(N)$ operations (see [1, 2]) even in the difficult two-level Toeplitz case.

In a wide range of applications \mathbf{g} and B are both corrupted by noise and consequently the reconstruction of the original image \mathbf{f} is a very difficult task since this kind of problems is ill-posed. In this case we need to adopt a regularization technique in order to reduce the noise effects, such as the popular Tikhonov regularization method [21].

In this work we use the Tikhonov technique for P-BCs and we customize it for D-BCs. Since the Tikhonov regularization needs the resolution of the normal equation, with D-BCs we define a method for approximate the coefficient matrix $T_{n_1, n_2}^2(f) + \mu I_N$ of the normal equation, by using $T_{n_1, n_2}(f^2 + \mu)$. Another regularization strategy was proposed by Riley [16]. This approach is suitable only for symmetric and positive definite system matrices K , where it is not necessary to use normal equations and we solve a slightly different minimization problem. In practice the regularization reduces the noise, while the MGM solves the regularized linear system with a linear cost of $O(N)$ arithmetic operations.

The outline of the paper is the following. In section 2 we introduce the MGM for the solution of linear systems with coefficient matrix BCCB and BTTB and we show its optimality. In section 3 we explain, also numerically, that our MGM is not a regularizer for blurred image affected by noise, while in section 4 we propose the Tikhonov and the Riley regularization techniques. In section 5 we report some numerical experiments and comparisons with others MGM proposed in literature for the considered kind of problems. Finally, in section 6 we draw some conclusions.

2. MGM FOR MULTILEVEL STRUCTURED MATRICES. Before describing the MGM used in this paper, we must introduce the τ matrix algebra [4], i.e. the space of matrices diagonalized by the discrete sine transform of type I (DST I). This is essential since the MGM for BTTB is defined extending the one defined for τ matrices. Later we define an MGM for τ and circulant matrices and as last step we generalize it to BTTB using a remarkable relationship between these two classes of matrices.

2.1. The τ algebra (DST I). Let Q_n be the n dimensional DST I with entries

$$[Q_n]_{i,j} = \sqrt{\frac{2}{n+1}} \sin\left(\frac{ji\pi}{n+1}\right), \quad i, j = 1, \dots, n. \quad (2.1)$$

It is known that the matrix Q_n is orthogonal and symmetric, i.e., $Q_n = Q_n^T$ and $Q_n^2 = I$. Moreover, for any n -dimensional vector v , the matrix vector multiplication $Q_n v$ can be computed in $O(n \log n)$ real operations by fast sine transforms. Let $\tau_n = \{Q_n D_n Q_n : D_n = \text{diag}_{j=1, \dots, n}(\lambda_j)\}$, any $A_n \in \tau_n$ can be expressed as $A_n = q(H_n)$, where $H_n = T_n(2 \cos(x))$ and q is a cosine polynomial of degree at most $n-1$. It follows that the eigenvalues λ_j for $j = 1, \dots, n$ have the form $\lambda_j = z(\frac{j\pi}{n+1})$, where $z(x)$ is a cosine polynomial, thus $A_n \in \tau_n$ is denoted by $A_n = \tau_n(z)$. If $a_j = a_{-j}$

denotes the j -th Fourier coefficients of z then the following relationship between τ_n and Toeplitz matrices of the same dimension holds $\tau_n(z)$ is equal to

$$T_n(z) - H_n(z) = \begin{bmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_1 \\ a_{n-1} & \cdots & a_1 & a_0 \end{bmatrix} - \begin{bmatrix} a_2 & \cdots & a_{n-1} & 0 & 0 \\ \vdots & \ddots & & & 0 \\ a_{n-1} & & & & a_{n-1} \\ 0 & & & \ddots & \vdots \\ 0 & 0 & a_{n-1} & \cdots & a_2 \end{bmatrix}. \quad (2.2)$$

We stress that when $z(x)$ is real, even and of finite order b , the Hankel matrix $H_n(z)$ is a negligible spectral correction of $T_n(z)$ with non zero element in north-west and south-east corners. This is the case that occurs in some signal deconvolution applications, where moreover we can assume that z vanishes in a neighborhood of π and is strictly positive elsewhere.

Concerning the bidimensional case (e.g. images), we first mention that τ algebra has a natural version $\tau_{n_1, n_2} = \{(Q_{n_1} \otimes Q_{n_2})D_N(Q_{n_1} \otimes Q_{n_2}) : D_N = \text{diag}_{j=1, \dots, N}(\lambda_j), N = n_1 n_2\}$, where Q_{n_1} and Q_{n_2} are one-level DST I transforms defined as in (2.1). Moreover, in perfect analogy with the τ class in one dimension, we have that $A_{n_1, n_2} \in \tau_{n_1, n_2}$ if and only if $A_{n_1, n_2} = \tau_{n_1, n_2}(z)$ can be (uniquely) written as $\tau_{n_1, n_2}(z) = T_{n_1, n_2}(z) - H_{n_1, n_2}(z)$ that is analogous to equation (2.2) where each element is one-level τ matrix. As in one dimension, matrix operations such as inversion, product, computation of the spectrum and so on can all be done by using a constant number of two-level sine transforms of type I and hence the resulting cost is $O(N \log N)$. When $z(x, y)$ is a polynomial, the matrix-vector product requires $O(N)$ arithmetic operations.

2.2. The MGM for τ and circulant matrices. For the sake of simplicity, in this subsection we assume $n_1 = n_2 = n$ reserving the subscript for the levels of the multigrid, but also in general ($n_1 \neq n_2$) everything works unchanged. Let $A \in \mathbf{C}^{N \times N}$ be a Hermitian positive definite matrix, $\mathbf{b} \in \mathbf{C}^N$, m be integer with $0 < m < N$. Fix integers $n_0 = n > n_1 > n_2 > \cdots > n_m > 0$, $N_i = n_i^2$ for $i = 0, \dots, m$, take $P_i \in \mathbf{C}^{N_{i+1} \times N_i}$ full-rank matrices and consider a class \mathcal{S}_i of iterative methods for N_i -dimensional linear systems. The related multigrid method produces the sequence $\{\mathbf{x}^{(k)}\}_{k \in \mathbf{N}}$ according to the rule $\mathbf{x}^{(k+1)} = \mathcal{MGM}(0, \mathbf{x}^{(k)}, \mathbf{b})$, with \mathcal{MGM} recursively defined as follows:

$$\begin{array}{l} \mathbf{x}_i^{(\text{out})} := \mathcal{MGM}(i, \mathbf{x}_i^{(\text{in})}, \mathbf{b}_i) \\ \hline \text{If } (i = m) \text{ Then Solve}(A_m \mathbf{x}_m^{(\text{out})} = \mathbf{b}_m) \\ \quad \text{Else } \begin{array}{l} \mathbf{1} \quad \mathbf{r}_i := \mathbf{b}_i - A_i \mathbf{x}_i^{(\text{in})} \\ \mathbf{2} \quad \mathbf{b}_{i+1} := P_i \mathbf{r}_i \\ \mathbf{3} \quad A_{i+1} := P_i A_i (P_i)^H \\ \mathbf{4} \quad \mathbf{y}_{i+1} := \mathcal{MGM}(i+1, \mathbf{0}_{n_{i+1}}, \mathbf{b}_{i+1}) \\ \mathbf{5} \quad \mathbf{x}_i^{(\text{int})} := \mathbf{x}_i^{(\text{in})} + (P_i)^H \mathbf{y}_{i+1} \\ \mathbf{6} \quad \mathbf{x}_i^{(\text{out})} := \mathcal{S}_i^\nu(\mathbf{x}_i^{(\text{int})}) \end{array} \end{array} \quad (2.3)$$

Step **1** calculates the residue of the proposed solution; steps **2**, **3**, **4** and **5** define the recursive Coarse Grid Correction (CGC) by projection (**2**) of the residual, sub-grid correction (**3**, **4**) and interpolation (**5**), while step **6** performs some (ν) iterations of a “post-smoother”. In the definition of a MGM it is fundamental that the CGC leads to a good approximation of the error in the subspace where the smoother is ineffective.

The MGM has essentially two degrees of indetermination for $i = 0, \dots, m-1$: the S_i (smoothers) and the P_i (projectors). The choice of the projectors P_i , and the calculation of the matrices A_i is performed before the beginning of the MGM procedure (pre-computing phase) with a logarithmic cost in the dimension N . The MGM defined in (2.3) is the simplest multigrid scheme with only one recursive call also known in literature as V -cycle.

We analyze a special instance of the MGM algorithm (2.3) for τ or circulant two-level matrices, studied in [1]. We define $R = \{\tau, C\}$ in order to describe both τ and circulant matrices, hence R_{n_i, n_i} will be τ_{n_i, n_i} in the τ case and C_{n_i, n_i} in the circulant case. In our MGM the smoother is the relaxed Richardson iteration, namely $S_i = I_{N_i} - \omega_i A_i$ (ω_i is relaxing parameter that will be fixed later). The projector is defined as $P_i = U_i \cdot R_{n_i, n_i}(p_i)$ where p_i is a real bivariate polynomial which will be defined in (2.4) and (2.5), while $U_i : R_{n_i, n_i} \rightarrow R_{n_{i+1}, n_{i+1}}$ is the cutting operator. The U_i is defined as $K_i \otimes K_i$ with K_i one-level cutting matrix defined as

$$K_i \in \mathbf{R}^{n_{i+1} \times n_i} \left| \begin{array}{c} \tau \\ \left[\begin{array}{cccc} 0 & 1 & 0 & \\ & & 0 & 1 & 0 \\ & & & \ddots & \ddots & \ddots \\ & & & & 0 & 1 & 0 \end{array} \right] \end{array} \right| \left| \begin{array}{c} \text{circulant} \\ \left[\begin{array}{cccc} 1 & 0 & & \\ & 0 & 1 & 0 \\ & & \ddots & \ddots \\ & & & 0 & 1 & 0 \end{array} \right] \end{array} \right| .$$

For the circulant algebra at each step we halve the size $n_{i+1} = \frac{n_i}{2}$ starting with $n_0 = 2^{k_0}$, $k_0 \in \mathbf{N}$, while for the τ algebra $n_{i+1} = \frac{n_i-1}{2}$ and $n_0 = 2^{k_0} - 1$.

If the coefficient matrix is $A_i = R_{n_i, n_i}(z_i)$ with z_i having a unique zero at $w_0 = (x_0, y_0)$ of order $2q$, the matrix $R_{n_i, n_i}(p_i)$ is chosen with p_i such that for $w = (x, y)$

$$\limsup_{w \rightarrow w_0} \left| \frac{p_i(\hat{w})}{z_i(w)} \right| < +\infty, \quad \hat{w} \in M(w), \quad i = 0, \dots, m-1, \quad (2.4)$$

where

$$0 < \sum_{\hat{w} \in M(w) \cup \{w\}} p_i^2(\hat{w}), \quad i = 0, \dots, m-1, \quad (2.5)$$

with $M(w) = \{(\pi + x, y), (x, \pi + y), (\pi + x, \pi + y)\}$ being the set of the ‘‘mirror points’’ of w (see e.g. [12]).

We stress as the cutting matrix U_i preserves the matrix structure at each level, while $R_{n_i, n_i}(p_i)$ selects the subspace where the smoother is ineffective. In this way, using a Galerkin strategy to compute the matrices at each level, we obtain an MGM according the Ruge-Stüben theory [17]. Furthermore, our MGM is optimal according to the following proposition.

PROPOSITION 2.1 ([2]). *Let $A_0 = R_{n_0, n_0}(z_0)$ with z_0 being nonnegative, if for every $i = 0, \dots, m-1$*

- (i) *the post-smoother is $S_i = I_{N_i} - \omega_i A_i$, where $0 < \omega_i \leq 2/\|z_i\|_\infty$,*
- (ii) *the projector $P_i = U_i R_{n_i, n_i}(p_i)$ is such that p_i satisfies the conditions (2.4) and (2.5),*

then $A_i = R_{n_i, n_i}(z_i)$ with z_i being nonnegative and the MGM defined in (2.3) is optimal.

The generating functions z_i are computed as $z_{i+1} = \frac{1}{4} \sum_{\hat{w} \in M(w) \cup \{w\}} p_i^2 z_i(\hat{w})$, $i = 0, \dots, m-1$, while the zero w_i moves into $w_{i+1} = 2w_i \pmod{2\pi}$, $i = 0, \dots, m-1$,

$n \times n$	0-pre 1-post (Rich)	1-pre (Rich) 1-post (Rich)	1-pre (Rich) 1-post (CG)
32×32	266	90	47
64×64	262	90	49
128×128	264	89	47
256×256	267	88	47

TABLE 2.1

Circulant case: Number of iterations with MGM for the solution of $C_{n,n}(z)\mathbf{x} = \mathbf{b}$ with $z(x, y) = (2 + \cos(x) + \cos(y))^3$ and \mathbf{x} a random vector. The smoothers are Richardson (Rich) and CG. With s -pre (post) are denoted s iterations of pre-smoother (post-smoother).

maintaining the same order at each recursion level (see [12]). In [2] is proved that the optimal choice for ω_i , which leads to the most relevant error reduction, is $1/\|z_i\|_\infty$, $i = 0, \dots, m - 1$.

In the deblurring problems considered in this work, the generating function z_0 is well approximated by a non negative even trigonometric polynomial that vanishes in (π, π) with order $2q$. Therefore, in this case, at the second level the zero moves in the origin and it remains in the origin for each following level; hence, the sequence $p^{(q)}$ of projector's generating functions defined as

$$\{p^{(q)}\}_i = \begin{cases} (2 - 2\cos(x))^q(2 - 2\cos(y))^q, & i = 0, \\ (2 + 2\cos(x))^q(2 + 2\cos(y))^q, & i = 1, \dots, m - 1, \end{cases} \quad (2.6)$$

satisfies the conditions (2.4) and (2.5) and Proposition 2.1 holds.

Usually in a multigrid algorithm also a pre-smoother is used (in algorithm (2.3) only a post-smoother is used in step **6**) applied to $\mathbf{x}_i^{(\text{in})}$ before step **1**. Clearly, if we add a pre-smoother or we increase the number of smoothing iterations, then the convergence rate of the MGM can only increase. Furthermore, different iterative methods like Richardson, Jacobi, Gauss-Seidel and others have similar spectral behavior. To make MGM work, a bad projector has to be cured by a stronger smoother or vice versa. According to the multi-iterative strategy (see [1] for a recent discussion on this topic), if we use Richardson both for the pre-smoother that for the post-smoother, the two relaxation parameters are chosen as $\omega_i^{(\text{pre})} = 1/\|z_i\|_\infty$ and $\omega_i^{(\text{post})} = 2\omega_i^{(\text{pre})}$. Moreover if the pre-smoother or the post-smoother is a stationary iterative method, then the remaining smoother can be also a non-stationary method like CG, in the worst case it does not improve the convergence speed and for the optimality the other classic smoother is sufficient. However, in Table 2.1, we can see as the use of the CG can lead to a faster MGM. In the related example we consider the generating function $z(x, y) = (2 + \cos(x) + \cos(y))^3$ which will be the "kernel" of the PSF proposed in the following sections, it has only one zero in (π, π) of order 6, hence, according to (2.6), the sequence of projector's generating functions $p^{(3)}$ leads to an optimal MGM.

From Table 2.1 the optimality of our MGM in the Circulant case is numerically evident, a constant number of iterations is required for the convergence, but analogous results hold in the τ case as well. The procedures are implemented in Fortran 90 using double precision, $\mathbf{x}^{(0)}$ is the null vector and the MGM was stopped when $\|\mathbf{r}^{(k)}\|/\|\mathbf{g}\| < 10^{-5}$, with $\mathbf{r}^{(k)}$ being the residual to the k -th iteration. The problem is solved directly on a grid of size 8×8 .

The structure of the cutting matrix U_i implies that the bandwidth along each direction is about halved at each recursion level and tends to double the bandwidth

of the projector (all projectors at each level have the same bandwidth). In this way the recursive algorithm is well defined and the number of the nonzero coefficients is smaller than the matrix dimensions. Moreover the number m of recursion levels is lesser than $\min(\log_2(n_1), \log_2(n_2))$. It follows that, since the computational cost at each level is about the same of the smoother, with the proposed smoothers and band matrices, the computational cost of one MGM (V -cycle) iteration is of $O(N)$ arithmetic operations. For more details see [1, 2].

2.3. The MGM for BTTB matrices. We generalize the MGM previously defined for the two-level τ algebra to the two-level Toeplitz class using the two-level generalization of the relation (2.2), which characterizes any Toeplitz matrix as its natural τ preconditioner plus an Hankel correction.

In the 1D case, it is possible to preserve the exact Toeplitz structure at each level without cutting much information defining the projector as $P_i = K_i\{t\}T_{n_i}(p_i)$, where t is defined as the degree of p_0 minus 1 (we stress that the degree of p_i is constant with respect to i) and $K_i\{t\} = \begin{bmatrix} 0_{n_{i+1}-t}^t & K_{n_{i+1}-t}^{n_i-2t} & 0_{n_{i+1}-t}^t \end{bmatrix} \in \mathbf{R}^{(n_{i+1}-t) \times n_i}$, with $0_{\alpha}^{\beta} \in \mathbf{R}^{\alpha \times \beta}$ the null matrix and $K_{n_{i+1}-t}^{n_i-2t} \in \mathbf{R}^{(n_{i+1}-t) \times (n_i-2t)}$ is the usual cutting matrix where we put in evidence the dimensions instead of the recursion levels. To apply the MGM recursively, we must start from dimension $n_0 = 2^{k_0} - 1 - 2t$, $k_0 \in \mathbf{N}$, hence the dimension of problem at each sublevel is $n_i = 2^{k_0-i} - 1 - 2t$. The matrix $K_i\{t\}$ is the cutting matrix that preserves the Toeplitzness at each level cutting the less information as possible.

Analogously, in the 2D case, p_i is a suitable bivariate nonnegative polynomial of partial degrees $t_j + 1$, with $j = 1, 2$. Let $U_i\{t\} = K_i\{t_1\} \otimes K_i\{t_2\}$, we define $P_i = U_i\{t\}T_{n_i, n_i}(p_i)$. Therefore, the two-level Toeplitz matrix at the multigrid recursion level $i + 1$ is $T_{n_{i+1}, n_{i+1}}(z_{i+1}) = P_i T_{n_i, n_i}(z_i) (P_i)^H \in \mathbf{R}^{N_{i+1} \times N_{i+1}}$ for $i = 0, \dots, m - 1$, where n_{i+1} is defined as in the one-level case.

Despite of its simplicity, the proposed technique is very effective since the number of elements that we neglect is reasonably low. In the two-dimensional case the size of the information that we lose is proportional to n , but also in this case the MGM maintains an optimal behavior as stressed by the following example. We consider the same generating function $z(x, y) = (2 + \cos(x) + \cos(y))^3$ of the Table 2.1, which vanishes in (π, π) with order six. In [1] is experimentally shown that condition (2.4) is rather strict and that for zeros of order six it is sufficient that the generating function of the projector vanishes with order four. Therefore, in order to cut less elements and avoid computationally expensive projectors, we take the sequence of projector's generating functions $p^{(2)}$ instead of $p^{(3)}$ required by the Proposition 2.1 and hence $t_1 = t_2 = 1$. The problem is solved directly on a grid of size $(7 - 2t_1) \times (7 - 2t_2)$, $t_1 = t_2 = 1$. From Table 2.2 the number of iterations, that the MGM requires to converge, reduces increasing the size of the problems. This stems from the fact that increasing n the percentage of the elements that we neglect, that are about $4n$, decreases respect to the algebraic size of the problem, which is $n^2 \times n^2$.

Finally, we remark that this MGM for two-level Toeplitz matrices can be considered an extension of that proposed in [9]. Indeed, in the case where z_0 vanishes only in (π, π) or in $(0, 0)$ with order almost two, in [9] is proved that the proposed MGM has a level-independent convergence and besides it is exactly the same here defined in the case of the sequence $p^{(1)}$. Furthermore, in this work we consider generating functions having zeros with order $2q$ also greater than two, then we define the choice of the sequence $p^{(q)}$ according to the τ algebra case while the cutting matrix is modified in order to maintain the two-level Toeplitz structure at each recursion level. In Table 2.2

$n \times n$	$p^{(2)}, \xi = 2$		$p^{(1)}, \xi = 0$
	1-pre (Rich)	1-pre (Rich)	1-pre (Rich)
	1-post (Rich)	1-post (CG)	1-post (CG)
$(31 - \xi) \times (31 - \xi)$	171	94	111
$(63 - \xi) \times (63 - \xi)$	167	90	129
$(127 - \xi) \times (127 - \xi)$	157	80	135
$(255 - \xi) \times (255 - \xi)$	132	75	144

TABLE 2.2

Toeplitz case: Number of iterations with MGM for the solution of $T_{n,n}(z)\mathbf{x} = \mathbf{b}$ with $z(x, y) = (2 + \cos(x) + \cos(y))^3$ and \mathbf{x} a random vector. The smoothers are Richardson (Rich) and CG. With s -pre (post) are denoted s iterations of pre-smoother (post-smoother).

we can see that for this example $p^{(1)}$ is less powerful than $p^{(2)}$ and it is not sufficient to obtain an optimal MGM.

3. THE (OPTIMAL) MGM IS NOT A REGULARIZER. For the P-BCs the proposed MGM is effective in the nonsymmetric case as well, while for D-BCs, since all τ matrices are symmetric, it works only for symmetric PSF. For non symmetric blurring functions, let s be the quadrantally symmetric blurring function given by $s_{i,j} = (z_{i,j} + z_{i,-j} + z_{-i,j} + z_{-i,-j})/4$, then $T_{n_1, n_2}(s)$ can be a good preconditioner for $T_{n_1, n_2}(z)$. This strategy has a “structural” interpretation since the blurring function z is replaced by its symmetrization s and if z is close to symmetric then $z - s$ is small and, by linearity, $T_{n_1, n_2}(z) - T_{n_1, n_2}(s)$ is small in norm: consequently it is reasonable to expect that the preconditioner $T_{n_1, n_2}(s)$ should lead to good results.

In the following tests we will use a PSF such that the resulting blur operator is a band approximation of the classical Gaussian blur whose Fourier coefficients are positive, symmetric and decay exponentially and whose generating function is close to zero in a neighborhood of (π, π) and is positive elsewhere. We define it by the “kernel” $F(x, y) = (2 + \cos(x) + \cos(y))^3$, that vanishes with order six in (π, π) and is positive elsewhere. The blur effect is increased multiplying $F(x, y)$ by a $\psi(x, y) > 0$ obtaining a PSF with a larger support. In this way the generating function of PSF is $z(x, y) = F(x, y)\psi(x, y)/c$, where c is a constant that is chosen equal to the sum of the Fourier coefficients of $F(x, y)\psi(x, y)$ like that $\|\mathcal{A}_{n_1, n_2}(z)\|_\infty = 1$ with $\mathcal{A} \in \{T, C\}$ and the PSF calculates a weighted average of the nearest pixels (see Fig. 3.1). This is a good approximation of an experimental Gaussian with a compact support. Furthermore, since this paper is a preliminary work where we apply the MGM proposed to image deconvolution problems, we use a polynomial and not a Gaussian blur in order to satisfy exactly the conditions 2.4 and 2.5 which lead to optimality. The Gaussian blur, like other types of BCs, will be subject of future work. We stress that however the conditioning number of $\mathcal{A}_{n_1, n_2}(z)$ is $O(\min(n_1, n_2)^6)$ (see [5]), then the resulting sequence is highly ill-conditioned.

If the observed image is affected by noise then the latter is amplified in the resolution of the linear system and the deconvolved image is much different from the original. According to Section 1 and to the theory of Toeplitz matrices (see e.g. [21]), if $z(\pi, \pi) = 0$ for i sufficiently large the eigenvalue λ_i of $\mathcal{A}_{n_1, n_2}(z)$ approaches zero and is associated to a highly oscillating eigenvector (high frequency). It follows that the inversion of a small λ_i amplifies the noise coefficients and the deconvolved image is corrupted; indeed a small percentage of noise can produce sensible variations in the deconvolved image.

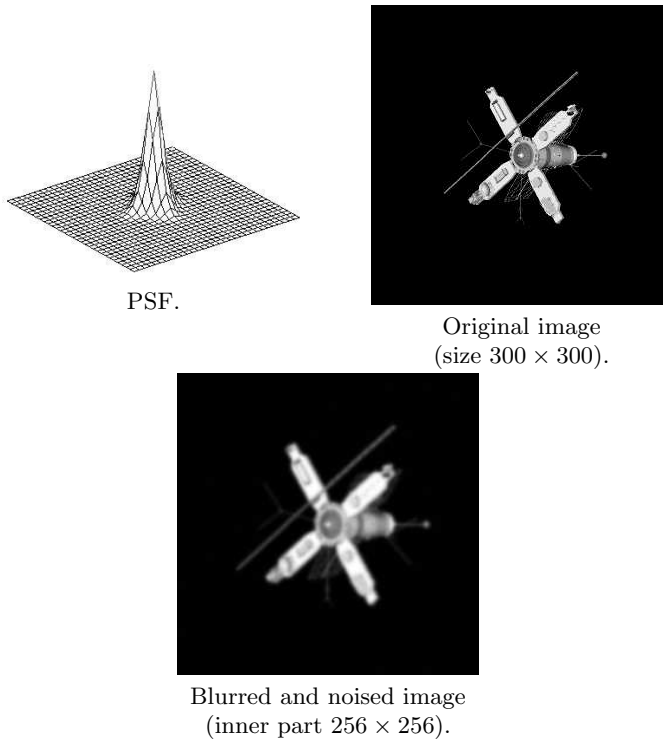


FIG. 3.1. *PSF and sequence images of “satellite”.*

In our tests we add a 2% of noise to the blurred image \mathbf{b} adding to it $\mathbf{n} = 2 \frac{\|\mathbf{b}\|_2}{\|\mathbf{c}\|_2} \mathbf{c}$, where $\mathbf{c} = \text{rand}(N, 1)$ in Matlab notation is a vector of N random components with uniform distribution, i.e. $\mathbf{g} = \mathbf{b} + \mathbf{n}$. Since we can not expect to obtain in the solution more than two digits of precision, the MGM is stopped when $\|\mathbf{r}^{(k)}\| / \|\mathbf{g}\| < 10^{-4}$ with $\mathbf{r}^{(k)}$ being the residual to the k -th iteration. In the MGM we perform two iterations of the pre-smoother that is relaxed Richardson with $\omega_i = 1 / \|z\|_\infty$ and two iterations of the post-smoother that is CG without preconditioning. At each level we increase by one the number of smoother’s iterations which does not increase the computational cost, see [2]. In Fig. 3.1 we can see the original image (the “satellite”¹ [15]) and its blurred and noised version used in the experiments. The original image of size 300 by 300 pixels has a black border greater than the support of the PSF. Therefore all the different kinds of boundary conditions are equivalent from a modelistic viewpoint. In this way the results are independent of border perturbations and we can focus on the properties of the solver (our MGM).

Several iterative methods like Richardson and CG without preconditioning have regularization properties: the error first decreases, reaches a minimum value and after increases. The error shows this behavior since these iterative methods first solve the problem in the low frequencies subspace where there is not the noise and after in the high frequencies subspace (where the noise lives). In Table 3.1 we can see that with our MGM and P-BCs the relative error norm increases already at the first

¹developed at the US Air Force Phillips Laboratory, Lasers and Imagine Directorate, Kirtland Air Force Base, New Mexico

#(Iter.)	1	30	60	85
$\ \text{error}\ _2$	1.876715E+06	2.702400E+06	2.702924E+06	2.702967E+06

TABLE 3.1

Relative error in $\|\cdot\|_2$ for the deconvolution with P-BCs without regularization.

iteration tending to a steady value; the same result holds for D-BCs as well. Therefore, the described MGM does not have any regularization properties like Richardson and CG. Its optimality prevents it to be a good regularizer since at each iteration it approximates the error in the whole frequency space and not only in a subspace: therefore it is disturbed by noise at each iteration. In fact at the first level the coarse grid correction projects the problem in the subspace generated by the high frequencies and solves it also in the components corrupted by noise. In the literature a few techniques are proposed to solve this problem, some of which are also independent of the method used to solve the linear system. The probably most known technique, that we will use in the next section, was defined by Tikhonov in [20].

However, thanks to the previous observations, with different choice of the projector it is possible to define iterative multigrid regularizing methods, like Landweber and CG for normal equations, that filters the noise at each level and that does not require any Tikhonov regularization (see [10] where a blending of a geometric multigrid and a specialized V-cycle for structured matrices is used for regularizing purposes).

4. REGULARIZATION STRATEGIES.

4.1. Tikhonov regularization. The regularization technique proposed by Tikhonov is based on the idea of reconstructing an image like the original and at the same time to control the noise effects. The minimum problem (1.1), now is reformulated as

$$\min_{\mathbf{f} \in \mathbf{R}^N} \left\{ \left\| \tilde{B}\mathbf{f} - \mathbf{g} \right\|_2^2 + \mu \|\mathbf{f}\|_2^2 \right\}. \quad (4.1)$$

If $\mu = 0$ then the solution of (4.1) is also solution of (1.1), but it is distorted by noise. Increasing μ we increase the weight of the part that takes under control the noise effects ($\|\mathbf{f}\|_2^2$), but the solution calculated is a little different from the solution of (1.1), i.e. the computed image is a little bit modified to reduce the noise effects. Therefore, the parameter μ is a weight which allows one to increase or decrease the noise distortion and the quality of the image. It follows that the choice of μ is crucial to obtain a deblurred image as close as possible to the true image. There exist many techniques to define the optimal value of μ and they are related to a theory independent of the method used to solve the linear system: thus in this paper the value of μ is determined experimentally like the value that minimizes the relative error 2-norm, i.e. $\|\mathbf{f} - \mathbf{f}^{(k)}\|_2 / \|\mathbf{f}\|_2$.

The minimization problem (4.1) can be reformulated with BCs and normal equations as

$$(B^T B + \mu I)\mathbf{f} = B^T \mathbf{g}, \quad (4.2)$$

where I is the identity matrix with the same dimensions of B . The numerical solution of (4.2) has the main disadvantage that the conditioning number doubles when $\mu \approx 0$. Furthermore, for D-BCs the matrix $B^T B$ is no longer a BTTB (for P-BCs we do not have this problem thanks to the matrix algebra structure of BCCB).

In the case of D-BCs, we propose a technique that allows to approximate the solution \mathbf{f} of (4.2) solving a linear system with coefficient matrix BTTB. It is based on the same idea that in Subsection 2.3 was used to define a projector which preserves the toeplitzness of the coefficient matrix at each multigrid level. For the sake of notational simplicity, we first consider the 1D case (Toeplitz matrix) which can be easily extended to the 2D case using tensor arguments. Let

$$(T_n(z)T_n(z) + \mu I_n)\mathbf{f}_n = T_n(z)\mathbf{g}_n \quad (4.3)$$

be the linear system that we have to solve. Since $T_n(z)$ is symmetric, $T_n(z)^2$ is symmetric but it is not Toeplitz unless $T_n(z)$ is a diagonal matrix. If the bandwidth of $T_n(z)$ is $2b + 1$ then $T_n(z)^2 = T_n(z^2) + H_n(z^2)$, where $H_n(z^2)$ has first and last $2b$ nonzero rows and columns. Cutting the matrix $H_n(z^2)$ we obtain a system with a solution “near” \mathbf{f}_n but with Toeplitz coefficient matrix:

$$T_n(z^2 + \mu)\tilde{\mathbf{f}}_n = T_n(z)\mathbf{g}_n, \quad (4.4)$$

where the smaller b the more $\tilde{\mathbf{f}}_n$ is near to \mathbf{f}_n . Since in the problems considered in this paper b is small, solving the system (4.4) we do not lose much information. The coefficient matrix is band Toeplitz positive-definite thus it is possible to use our MGM with good results as stressed in the experimentation of Section 5. However, we can obtain a fast solver, possibly improving the quality of the deblurred image, also solving the linear system (4.3) with CG and using the MGM applied to the linear system (4.4) as preconditioner.

In the 2D case we apply the same strategy along each dimension. Therefore, for both P-BCs and D-BCs, we solve the linear system

$$\mathcal{A}_{n_1, n_2}(z^2 + \mu)\mathbf{f} = \mathcal{A}_{n_1, n_2}(z)\mathbf{g}. \quad (4.5)$$

4.2. Riley regularization. We have assumed that $z(x, y)$ is real, even in both variables and nonnegative, hence $\mathcal{A}_{n_1, n_2}(z)$ is positive definite. It follows that our MGM can be applied without passing to the normal equations and we can apply a regularization technique firstly formulated by Riley in 1955 [16]. With the appropriate BCs, instead of solving the linear system (4.2) we solve $(B + \theta I)\mathbf{f} = \mathbf{g}$, which is equivalent to

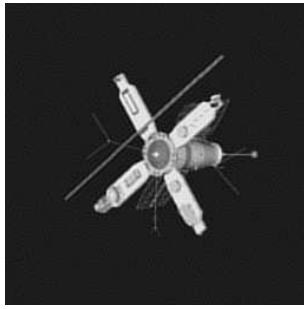
$$\min_{\mathbf{f} \in \mathbf{R}^N} \left\{ \left\| B^{\frac{1}{2}}\mathbf{f} - B^{-\frac{1}{2}}\mathbf{g} \right\|_2^2 + \theta \|\mathbf{f}\|_2^2 \right\}. \quad (4.6)$$

The considered technique can be applied also under weaker assumptions: experimentally, it is observed that the Riley method works in the case where B is only numerically positive definite i.e. when B shows few negative eigenvalues very small in modulus.

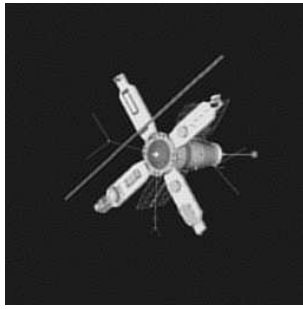
We stress that the factor $B^{-\frac{1}{2}}\mathbf{g}$ can cause a little amplification of the noise as is also verified experimentally in the next section. Usually we have $\theta > \mu$. From (4.6) it follows that, with the considered BCs, we solve

$$\mathcal{A}_{n_1, n_2}(z + \theta)\mathbf{f} = \mathbf{g}. \quad (4.7)$$

5. NUMERICAL EXPERIMENTS. In this section we use the two techniques presented in the previous sections to deblur the “satellite” image with a noise of 2% (see Fig. 3.1).

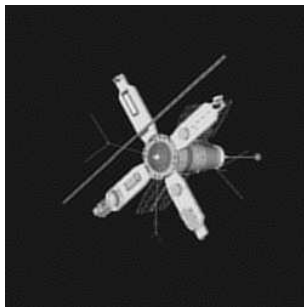


Tikhonov with μ_{opt}
error norm = 0.1578

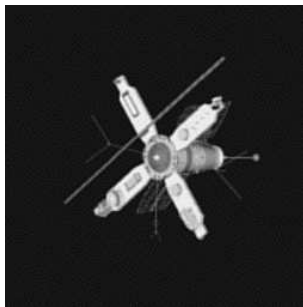


Riley with θ_{opt}
error norm = 0.1948

FIG. 5.1. *Deblurred image with P-BCs.*



Tikhonov with μ_{opt}
error norm = 0.1579



Riley with θ_{opt}
error norm = 0.1957

FIG. 5.2. *Deblurred image with D-BCs.*

For the Tikhonov method, we solve the linear system (4.5), where the parameter μ_{opt} is chosen as the value that minimizes the 2-norm of the relative error. In the same way, for the Riley method, we solve the system (4.7) with θ_{opt} . We observe that $\theta_{opt} > \mu_{opt}$, probably because we have to reduce the noise amplification deriving from $B^{-\frac{1}{2}}\mathbf{g}$. Therefore, the linear system obtained with Riley is better conditioned than with Tikhonov and the MGM converges with a less number of iterations. On the other hand, the factor $B^{-\frac{1}{2}}$ in the minimization problem (4.6) amplifies a little the noise and the image deblurred with Tikhonov is slightly better than the one with Riley. Indeed, using P-BCs, the relative deconvolution error in 2-norm with Tikhonov is 0.1578 while with Riley is 0.1948; the computed images are shown in Fig. 5.1. With D-BCs we obtain similar result with about the same number of iterations (see Fig. 5.2). Indeed the relative deconvolution error in 2-norm with Riley is 0.1957 while with Tikhonov, using the approximation technique presented in Section 4, it is 0.1579. Since the original image has a black border greater than the support of the PSF, both the two types of BCs are equivalent and it is natural to obtain similar results. Moreover, this shows that the “cutting” technique proposed in Section 4 for BTTB matrices is effective also in the 2D case where the rank correction is proportional to n .

At the end (Table 5.1) we compare the proposed MGM ($p^{(2)}$) with the MGM proposed in [9] ($p^{(1)}$) and the classic MGM (linear interpolation) used for image deconvolution in recent papers [7, 13]. However in the latter two works the linear interpolation is applied to a PSF that is related to the dense Gaussian kernel that it not

θ	10^{-2}	10^{-3}	10^{-4}	10^{-5}
$p^{(2)}$	15	49	73	73
$p^{(1)}$	15	55	103	233
linear interpolator	15	45	286	> 1000

TABLE 5.1

Number of iterations for the deconvolution of satellite with P-BCs using Riley regularization.

considered in this paper. For dense Gaussian kernel, there is no generating function in the usual sense, or the generating function has a zero of infinite order at $[\pi, \pi]$. Furthermore in [7, 13] the authors used as smoother the CG with preconditioning where the preconditioner is the cosine or circulant optimal preconditioner. Here we use simpler smoothers, and we see as our projector works much better than the linear interpolation when the relaxation parameter decreases. Furthermore, using PCG as smoother and our projector we can improve the convergence behavior. In Table 5.1 the number of iterations is reported for decreasing values of θ , for the three different multigrid methods using Riley regularization (also used in [7, 13]) and P-BCs. We stress as with our MGM the number of iterations tends to a constant value while with the other methods it sensibly increases.

6. CONCLUSION. In this paper we have considered the deconvolution of blurred and noised images with shift-invariant PSF having a small support using Dirichlet or periodic BCs. The proposed MGM can be used to solve rapidly (optimally) the linear system, but it does not have any regularization property in itself. Therefore, when the images are distorted by noise, it is necessary to use classic regularization techniques like Tikhonov. However, with a different choice of the projectors it is possible to obtain a MGM which is a regularizing method instead of a fast solver combined with the Tikhonov regularization (see [10]). A work to be done in the future is to analyze the behavior of our MGM with more “complicated” PSF and more powerful BCs (Neumann and anti-reflective).

ACKNOWLEDGMENT. Sincere thanks to the referees for the pertinent remarks and the useful suggestions that helped me to improve the quality of the paper.

REFERENCES

- [1] A. Aricò, M. Donatelli and S. Serra Capizzano, V-cycle optimal convergence for certain (multi-level) structured, *SIAM J. Matrix Anal. Appl.*, in press.
- [2] A. Aricò and M. Donatelli, Multigrid for multilevel matrix algebra linear systems: proof of optimality, *manuscript*, 2004.
- [3] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*, IOP Publishing, Bristol, 1998.
- [4] D. Bini and M. Capovani, Spectral and computational properties of band symmetric Toeplitz matrices, *Linear Algebra Appl.*, 52/53, pp. 99–125, 1983.
- [5] A. Böttcher and S. Grudsky, On the condition numbers of large semi-definite Toeplitz matrices, *Linear Algebra Appl.*, 279, pp. 285–301, 1998.
- [6] T. H. Chan, S. Serra Capizzano and C. Tablino Possio, Two-Grid Methods for Bandes Linear Systems for DCT III algebra, *Numer. Linear Algebra Appl.*, 00, pp. 1–8, 2002.
- [7] R.H. CHAN, T.F. CHAN, W. WAN, Multigrid for differential-convolution problems arising in image processing, *Proc. Workshop on Scientific Computing*, G. Golub, S.H. Lui, F. Luk, R. Plemmons Eds, Springer, pp. 58–72, 1999.
- [8] R. H. Chan, Q. Chang and H. Sun, Multigrid method for ill-conditioned symmetric Toeplitz systems, *SIAM J. Sci. Comp.*, Vol. 19-2, pp. 516–529, 1998.

- [9] Q. Chan, X. Jin and H. Sun, Convergence of the multigrid method for ill-conditioned block Toeplitz systems, *BIT*, Vol. 41, pp. 179–190, 2001.
- [10] M. Donatelli and S. Serra Capizzano, On the regularizing power of multigrid-type algorithms, *Tech. Rep. N. 501/04 Dip. Matematica Univ. Genova*, 2004.
- [11] G. Fiorentino and S. Serra Capizzano, *Multigrid methods for Toeplitz matrices*, *Calcolo*, Vol. 28, pp. 283–305, 1991.
- [12] G. Fiorentino and S. Serra, Multigrid methods for symmetric positive definite block Toeplitz matrices with nonnegative generating functions, *SIAM J. Sci. Comput.*, Vol. 17-4, pp. 1068–1081, 1996.
- [13] T. Huckle, J. Staudacher, Multigrid preconditioning and Toeplitz matrices, *Electr. Trans. Numer. Anal.*, 13 (2002), pp. 81–105.
- [14] N. Kalouptsidis, G. Carayannis, and D. Manolakis, Fast algorithms for block Toeplitz matrices with Toeplitz entries, *Signal Process.*, 6, pp. 77–81, 1984.
- [15] M. Ng, R. Chan and W. C. Tang, A fast algorithm for deblurring models with Neumann boundary conditions, *SIAM J. Sci. Comput.*, 21, pp. 851–866, 1999.
- [16] J. D. Riley, Solving systems of linear equations with a positive definite, symmetric, but possibly ill-conditioned matrix, *Math. Tables Aids Comput.*, 9, pp. 96–101, 1955, English translation of Dokl. Akad. Nauk. SSSR, 151, pp. 501–504, 1963.
- [17] J. W. Ruge and K. Stüben *Algebraic Multigrid*, in *Frontiers in Applied Mathematics: Multigrid Methods* S. McCormick Ed. SIAM, Philadelphia (PA), pp. 73–130, 1987.
- [18] S. Serra Capizzano, Matrix algebra preconditioners for multilevel Toeplitz matrices are not superlinear, *Linear Algebra Appl.*, 343/344, pp. 303–319, 2002.
- [19] S. Serra Capizzano, A note on anti-reflective boundary conditions and fast deblurring models, *SIAM J. Sci. Comput.*, in press.
- [20] A. N. Tikhonov, Solution of incorrectly formulated problems and regularization method, *Soviet Math. Dokl.*, 4, pp. 1035–1038, 1963.
- [21] E. Tyrtysnikov and N. Zamarashkin, Spectra of multilevel Toeplitz matrices: advanced theory via simple matrix relationships, *Linear Algebra Appl.*, 270, pp. 15–27, 1998.