Electronic Transactions on Numerical Analysis.
Volume 44, pp. 25–52, 2015.

ISSN 1068–9613.

ETNA
Kent State University
http://etna.math.kent.edu

ANALYSIS OF SMOOTHED AGGREGATION MULTIGRID METHODS BASED ON TOEPLITZ MATRICES*

MATTHIAS BOLTEN†, MARCO DONATELLI‡, AND THOMAS HUCKLE§

Abstract. The aim of this paper is to analyze multigrid methods based on smoothed aggregation in the case of circulant and Toeplitz matrices. The analysis is based on the classical convergence theory for these types of matrices and yields optimal choices of the smoothing parameters for the grid transfer operators in order to guarantee optimality of the resulting multigrid method. The developed analysis allows a new understanding of smoothed aggregation and can also be applied to unstructured matrices. A detailed analysis of the multigrid convergence behavior is developed for the finite difference discretization of the 2D Laplacian with nine point stencils. The theoretical findings are backed up by numerical experiments.

Key words. multigrid methods, Toeplitz matrices, circulant matrices, smoothed aggregation-based multigrid

AMS subject classifications. 15B05, 65F10, 65N22, 65N55

1. Introduction. In this paper we consider smoothed aggregation (SA) multigrid methods for solving the linear system

$$Ax = b, \tag{1.1}$$

where $x, b \in \mathbb{C}^N$ and A is an ill-conditioned symmetric positive definite $N \times N$ matrix. Mainly, we analyze the case of multilevel Toeplitz matrices, while some numerical results will be presented also for the discretization of non-constant coefficient partial differential equations (PDEs) based on a local stencil analysis.

On the one hand, the development of multigrid methods for τ-matrices and Toeplitz matrices goes back to [10], the two-level case being considered in [11]. Using the same ideas, methods for circulant matrices were developed later in [25]. While these works provide the basis to develop and analyze multigrid methods for Toeplitz matrices and matrices from different matrix algebras including the τ- and circulant algebra, they do not provide a proof of optimality of the multigrid cycle in the sense that the convergence rate is bounded by a constant $c < 1$ independent of the number of levels used in the multigrid cycle. Such a proof was added later in [1, 2]. In [9], two-grid optimality is proved in the case of a cutting greater than two for Toeplitz matrices. This analysis can be useful for 1D aggregation methods and will be extended to multidimensional problems in this paper.

The theory that is used to build up the two-grid and multigrid methods and to prove their convergence is based on the classical variational multigrid theory as it is presented in, e.g., [16, 18, 19, 22].

Aggregation based multigrid goes back at least to [4], where the so-called aggregation/disaggregation methods [7, 17] have been used in a multigrid setting. The idea of aggregation based multigrid is to avoid a C/F-splitting, i.e., a partitioning of the unknowns into variables that are present on the coarse and the fine level and variables that are present on the fine level only. Rather than grouping the unknowns together into *aggregates*, these aggregates form one variable on the coarse level each. Pure aggregation can be improved by incorporating smoothing [28] in the prolongation and/or the restriction leading to faster

*Received August 5, 2013. Accepted October 28, 2014. Published online on February 4, 2015. Recommended by Johannes Kraus.
†Bergische Universität Wuppertal, Fachbereich Mathematik und Naturwissenschaften, Gaußstraße 20, 42097 Wuppertal, Germany (bolten@math.uni-wuppertal.de).
‡Università dell'Insubria, Via Vallegio 11, 22100 Como, Italy (marco.donatelli@uninsubria.it).
§Technische Universität München, Boltzmannstraße 3, 85748 Garching, Germany (huckle@in.tum.de).

convergence. Recent results on the convergence of aggregation-based multigrid methods can be found in [5, 6, 20, 21].

In this paper, firstly we extend the two-grid optimality results in [9] to multidimensional problems. Using these new convergence results, we provide an analysis of aggregation operators for multilevel Toeplitz matrices. We show that pure aggregation provides only two-grid optimality, but, according to the literature [20], this is not enough for V-cycle optimality. Therefore, we study a simple smoothing aggregation strategy based on a damping factor chosen as the value that provides the best convergence rate. In contrast to previous analysis in the literature [6, 20, 21], our analysis uses a symbolic approach to discuss convergence and to choose the optimal damping factors. A detailed study for the finite difference discretization of the 2D Laplacian with nine point stencils shows that our symbolic approach can be easily performed and implemented and, at the same time, is also very effective. In particular, we show how to design the smoothed aggregation incorporating more than one smoother or allowing nonsymmetric projection such that it leads to fast convergence without increasing the bandwidth of the coarser systems. Finally, numerical results are provided also in the non-constant coefficient case using the local stencil of the operator.

The outline of the paper is as follows. In Section 2 we introduce Toeplitz and circulant matrices, multigrid methods, and some well-known results on multigrid methods for Toeplitz matrices. The main theoretical results appear in Section 3, where the aggregation and the smoothed aggregation optimality conditions are studied in the case of circulant matrices. In Section 4 we discuss how the results obtained in the circulant case can be applied to Toeplitz matrices or to the discretization of non-constant coefficients partial differential equations. Special attention is devoted to the discretization of the 2D Laplacian by nine points stencils in Section 4.3. A wide range of numerical experiments is presented in Section 5, and some concluding remarks complete the paper in Section 6.

2. Preliminary.

In this section we introduce some well-known results on Toeplitz matrices and multigrid methods.

2.1. Toeplitz and circulant matrices.

A Toeplitz matrix $T_n \in \mathbb{C}^{n\times n}$ is a matrix with constant entries on the diagonals, i.e., T_n is of the form

$$T_n = \begin{bmatrix} t_0 & t_{-1} & \cdots & t_{1-n} \\ t_1 & t_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_{-1} \\ t_{n-1} & \cdots & t_1 & t_0 \end{bmatrix}. \tag{2.1}$$

As a consequence, the matrix entries are completely determined by the $2n-1$ values $t_{-n+1}, \ldots, t_{n-1}$. There exists a close relationship between a Toeplitz matrix and its generating symbol $f : \mathbb{R} \to \mathbb{C}$, a 2π-periodic function given by

$$f(x) = \sum_{j=-\infty}^{\infty} t_j \mathrm{e}^{\mathrm{i}2\pi jx}, \qquad t_j = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\mathrm{e}^{-\mathrm{i}2\pi jx}dx, \tag{2.2}$$

with the matrix entries t_j on the diagonals taken as Fourier coefficients of f. The generating symbol f always induces a sequence $\{\mathcal{T}_n(f)\}_{n=1}^{\infty}$ of Toeplitz matrices $\mathcal{T}_n(f)$. In the case of f being a trigonometric polynomial, the resulting Toeplitz matrices are band matrices for n large enough. There are various theoretical results on sequences of Toeplitz matrices and their generating symbol. Most important for the analysis of iterative methods for Toeplitz

ETNA
Kent State University
http://etna.math.kent.edu

matrices is the fact that the distribution of the eigenvalues of the Toeplitz matrix is given by the generating symbol in the limit case $n \to \infty$; cf. [27].

Circulant matrices are of a very similar form. A circulant matrix is a Toeplitz matrix with the additional property $t_{-k} = t_{n-k}$, $k = 1, 2, \ldots$, i.e.,

$$C_n = \begin{bmatrix} t_0 & t_{n-1} & \cdots & t_1 \\ t_1 & t_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_{n-1} \\ t_{n-1} & \cdots & t_1 & t_0 \end{bmatrix}.$$

C_n is diagonalized by the Fourier matrix F_n, where

$$(F_n)_{j,k} = \frac{1}{\sqrt{n}} \mathrm{e}^{-\frac{2\pi \mathrm{i}}{n} jk}, \qquad j, k = 0, \ldots, n-1,$$

i.e.,

$$C_n = F_n \operatorname{diag}(\lambda^{(n)}) F_n^H, \tag{2.3}$$

for $\lambda^{(n)} = (\lambda_0^{(n)}, \ldots, \lambda_{n-1}^{(n)})$ given by $\lambda_j^{(n)} = f(2\pi j/n)$, $j = 0, \ldots, n-1$. Allowing negative indices to denote the diagonals above the main diagonal as in the Toeplitz case, i.e., in (2.1), results in demanding $t_k = t_{k \bmod n}$. Using the generating symbol f in (2.2) similarly to the Toeplitz case, a sequence $\{\mathcal{C}_n(f)\}_{n=1}^{\infty}$ of matrices $\mathcal{C}_n(f)$ is defined. In contrast to the Toeplitz case, the circulant matrices form a matrix algebra as they are diagonalized by the Fourier matrix F_n.

The concept of Toeplitz and circulant matrices can easily be extended to the block case, i.e., the case where the matrix entries are not elements of the field of complex numbers but rather of the ring of $m \times m$ matrices. In this case the generating symbol becomes a matrix-valued 2π-periodic function, and the matrices are called block Toeplitz and block circulant matrices, respectively. The aforementioned properties of the matrices transfer to this case, e.g., a block circulant matrix with block size $m \times m$ and n blocks on the main diagonal is block diagonalized by $F_n \otimes I_m$, where $\otimes$ denotes the Kronecker product and I_m denotes the identity matrix of size $m \times m$. The analysis of multigrid methods with more general blocks is beyond the scope of this article; for further details see, e.g., [15].

An interesting special type of block matrices that we will deal with is the case where the blocks themselves are again Toeplitz/circulant. The resulting matrix will be called block Toeplitz Toeplitz block (BTTB) or block circulant circulant block (BCCB), and it can be described by a bivariate 2π-periodic generating symbol f. This is related to the two-dimensional case $d = 2$. In the general d-level case, the generating symbols are 2π-periodic functions $f : \mathbb{R}^d \to \mathbb{C}$ having Fourier coefficients

$$t_j = \frac{1}{(2\pi)^d} \int_{[-\pi,\pi]^d} f(x) \mathrm{e}^{-\mathrm{i}\langle j | x \rangle}\, dx, \qquad j = (j_1, \ldots, j_d) \in \mathbb{Z}^d,$$

where $\langle \cdot | \cdot \rangle$ denotes the usual scalar product between vectors. From the coefficients t_j, one can build the sequence $\{\mathcal{C}_n(f)\}$, $n = (n_1, \ldots, n_d) \in \mathbb{N}^d$, of multilevel circulant matrices of size $N = \prod_{r=1}^{d} n_r$. Defining the d-dimensional Fourier matrix $F_n = F_{n_1} \otimes \cdots \otimes F_{n_d}$, the matrix $\mathcal{C}_n(f)$ can be written again in the form (2.3), where now the eigenvalues $\lambda^{(n)}$ are defined by

$$\lambda_j^{(n)} = f\left(\frac{2\pi j_1}{n_1}, \ldots, \frac{2\pi j_d}{n_d}\right), \qquad j_i = 0, \ldots, n_i - 1, \quad i = 1, \ldots, d,$$

ordered according to the tensor product structure of the eigenvectors.

ETNA
Kent State University
http://etna.math.kent.edu

2.2. Multigrid methods. A multigrid method is a method to solve a linear system of equations. When traditional stationary iterative methods like Jacobi iteration are used to solve a linear system, they perform poorly when the system becomes more ill-conditioned, e.g., when the mesh width of the discretization of a PDE is decreased. The reason for this behavior is that error components belonging to large eigenvalues are damped efficiently, while error components belonging to small eigenvalues get reduced slowly. In the discretized PDE example, the first correspond to the rough error modes, while the latter correspond to the smooth error modes. For this reason methods like Jacobi iteration are known as "smoothers".

To construct a multigrid method, various components have to be chosen. In the following, the coefficient matrix of the linear system (1.1) on the finest level is denoted by $A_0 = A$, the multi-index of the size is denoted by $n_0 = n \in \mathbb{N}^d$. The multi-indices of the system sizes on the coarser grids are then denoted by $n_i < n_{i-1}$, $i = 1, \dots, l_{\max}$, where $l_{\max}$ is the maximum number of levels used. Defining $N_i = \prod_{j=1}^d (n_i)_j$, to transfer a quantity from one level to another, restriction operators $R_i : \mathbb{C}^{N_i} \to \mathbb{C}^{N_{i+1}}$, $i = 0, \dots, l_{\text{max}} - 1$, and prolongation operators $P_i : \mathbb{C}^{N_{i+1}} \to \mathbb{C}^{N_i}$, $i = 0, \dots, l_{\text{max}} - 1$, are needed. Furthermore, a hierarchy of operators $A_i \in \mathbb{C}^{N_i \times N_i}$, $i = 1, \dots, l_{\text{max}}$, has to be defined. On each level, appropriate smoothers $\mathcal{S}_i$ and $\tilde{\mathcal{S}}_i$ and the numbers of smoothing steps ν_1 and ν_2 have to be chosen. We limit ourselves to stationary iterative methods although other smoothers like Krylov-subspace methods can be used as well. After ν_1 presmoothing steps using $\mathcal{S}_i$, the residual $r_{n_i} \in \mathbb{C}^{N_i}$ is computed and restricted to the coarse grid; the result is $r_{n_{i+1}}$. On the coarse grid the error is computed by solving

$$A_{i+1} e_{n_{i+1}} = r_{n_{i+1}},$$

and in the multigrid case this is done by a recursive application of the multigrid method. The resulting error is interpolated back to obtain the fine-level error e_{n_i}, and the current iterate is updated using this error. Afterwards, the iterate is improved by postsmoothing. When only one recursive call is applied, like in this paper, the whole iteration is called a V-cycle. The process of correcting the current iterate using the coarse level is known as *coarse-grid correction*, which has the iteration matrix

$$M_i = I - P_i A_{i+1}^{-1} R_i A_i.$$

In summary, the multigrid method $\mathcal{MG}_i$ is given by Algorithm 1.

Algorithm 1 Multigrid cycle $x_{n_i} = \mathcal{MG}_i(x_{n_i}, b_{n_i})$.

$x_{n_i} \leftarrow \mathcal{S}_i^{\nu_1}(x_{n_i}, b_{n_i})$
$r_{n_i} \leftarrow b_{n_i} - A_i x_{n_i}$
$r_{n_{i+1}} \leftarrow R_i r_{n_i}$
$e_{n_{i+1}} \leftarrow 0$
if $i + 1 = l_{\text{max}}$ **then**
  $e_{n_{l_{\text{max}}}} \leftarrow A_{l_{\text{max}}}^{-1} r_{n_{l_{\text{max}}}}$
else
  $e_{n_{i+1}} \leftarrow \mathcal{MG}_{i+1}(e_{n_{i+1}}, r_{n_{i+1}})$
end if
$e_{n_i} \leftarrow P_i e_{n_{i+1}}$
$x_{n_i} \leftarrow x_{n_i} + e_{n_i}$
$x_{n_i} \leftarrow \tilde{\mathcal{S}}_i^{\nu_2}(x_{n_i}, b_{n_i})$

To show convergence of a multigrid method, usually, R_i is chosen to be the adjoint of P_i, and the coarse-grid operator A_{i+1} is chosen as the Galerkin coarse-grid operator $P_i^H A_i P_i$. The classical algebraic convergence analysis is based on two properties, the smoothing property and the approximation property, which are coupled together by an appropriately chosen norm $\|\cdot\|_*$, where in the classical algebraic multigrid theory, the $AD^{-1}A$-norm with $D = \mathrm{diag}(A)$ is chosen, cf. [22], and in the circulant case, the A^2-norm turns out to be helpful; cf. [2].

DEFINITION 2.1 (Smoothing properties). *An iterative method $\mathcal{S}_i$ with iteration matrix S_i fulfills the* presmoothing property *if there exists an $\alpha > 0$ such that for all $v_{n_i} \in \mathbb{C}^{N_i}$ it holds that*

$$\|S_i v_{n_i}\|_{A_i}^2 \leq \|v_{n_i}\|_{A_i}^2 - \alpha \|S_i v_{n_i}\|_*^2. \tag{2.4}$$

Analogously, it fulfills the postsmoothing property *if there exists a $\beta > 0$ such that*

$$\|\tilde{S}_i v_{n_i}\|_{A_i}^2 \leq \|v_{n_i}\|_{A_i}^2 - \beta \|v_{n_i}\|_*^2. \tag{2.5}$$

The following theorem is useful to prove convergence of two-grid methods since the forthcoming condition (2.7) is usually weaker and easier to verify than the *approximation property*

$$\|M_i v_{n_i}\|_{A_i}^2 \leq \gamma \|v_{n_i}\|_*^2. \tag{2.6}$$

THEOREM 2.2 ([22]). *Let $A_i \in \mathbb{C}^{N_i \times N_i}$ be a positive definite matrix, and let $\tilde{\mathcal{S}}_i$ be the postsmoother with iteration matrix $\tilde{S}_i$ fulfilling the postsmoothing property (2.5) for $\beta > 0$. Assume that $R_i = P_i^H$, $A_{i+1} = P_i^H A_i P_i$, and that there exists $\gamma > 0$ independent of N_i such that*

$$\min_{y \in \mathbb{C}^{N_{i+1}}} \|x - P_i y\|_{D_i}^2 \leq \gamma \|x\|_{A_i}^2, \qquad \forall x \in \mathbb{C}^{N_i}, \tag{2.7}$$

where D_i is the main diagonal of A_i. Then $\gamma \geq \beta$ and

$$\|\tilde{S}_i M_i v_{n_i}\|_{A_i} \leq \sqrt{1 - \beta/\gamma}\, \|v_{n_i}\|_{A_i}, \qquad \forall v_{n_i} \in \mathbb{C}^{N_i}.$$

2.3. Multigrid methods for circulant and Toeplitz matrices. In the following, we introduce multigrid methods for circulant matrices and briefly review the convergence results for these methods as our analysis of aggregation based methods is based on these results. After that, we provide an overview over the modifications necessary to deal with Toeplitz matrices in a conceptually very similar way.

Let f_i be the symbol of A_i; in this paper we assume $f_i \geq 0$ thus A_i is positive definite[1]. In general, to design a multigrid method, the smoother, a coarse level with fewer degrees of freedom, the prolongation, and the restriction have to be chosen appropriately. Here, the common choice for both, pre- and postsmoothing is relaxed Richardson iteration, i.e., $\mathcal{S}_i$ is chosen as

$$\mathcal{S}_i(x_{n_i}, b_{n_i}) = \underbrace{(I - \omega_i A_i)}_{=S_i} x_{n_i} + \omega_i b_{n_i}, \tag{2.8}$$

and $\tilde{\mathcal{S}}_i$ is chosen like this but with a different $\tilde{\omega}_i$. Note that for Toeplitz matrices, relaxed Richardson iteration is equivalent to relaxed Jacobi iteration since the diagonal of the coefficient

[1] A_i could be singular for circulant matrices if f vanishes at a grid point. In such case a rank-one correction like in [2] could be considered, but it is not necessary in practice; see [3].

matrix is a multiple of the identity. Using appropriate relaxation parameters ω_i and $\tilde{\omega}_i$, this smoother fulfills the presmoothing property (2.4), respectively the postsmoothing property (2.5) as stated by the following theorem, which can be found in [1, Proposition 3].

THEOREM 2.3 ([1]). *Let $A_i = \mathcal{C}_{n_i}(f_i)$, where $f_i : \mathbb{R}^d \to \mathbb{R}$. Let $\mathcal{S}_i$ be defined in* (2.8) *with $\omega_i \in \mathbb{R}$ and $\tilde{S}_i$ be defined in the same way as S_i but with the parameter $\tilde{\omega}_i \in \mathbb{R}$. Then if $\omega_i, \tilde{\omega}_i \in (0, 2/\|f_i\|_\infty)$, the smoothing properties* (2.4) *and* (2.5) *are fulfilled with $\|\cdot\|_* = \|\cdot\|_{A^2}$.*

Regarding the choice of the coarse level, for circulant matrices usually we assume that the number of unknowns in each "direction" is divisible by 2, i.e., $(n_i)_j \bmod 2 = 0$ for $j = 1, \ldots, d$. Then on the coarse level we choose every other degree of freedom, effectively dividing the number of unknowns by 2^d when moving from level i to level $i+1$. This corresponds to standard coarsening in geometric multigrid. Other coarsenings, e.g., by a factor different from 2 [9] or corresponding to semi-coarsening [12, 14] are derived and used in a straightforward way. The reduction from the fine level to the coarse level is described with the help of a *cutting matrix* $K_{n_i} \in \mathbb{C}^{n_{i+1} \times n_i}$ ([23]), which on the fine level in the case of a 1-level circulant matrix of even size is given by

$$K_{n_i} = \begin{bmatrix} 1 & 0 & & & & & \\ & & 1 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 1 & 0 \end{bmatrix}.$$

The effect of this cutting matrix is that every even variable is skipped when it is transferred to the coarse level. Regarding the action of the cutting matrix on the Fourier matrix, we obtain

$$K_{n_i} F_{n_i} = \frac{1}{\sqrt{2}} [1, 1] \otimes F_{n_{i+1}} = \frac{1}{\sqrt{2}} F_{n_{i+1}} ([1, 1] \otimes I_{n_{i+1}}) \tag{2.9}$$

in the 1-level case ([24]). In the d-level case the cutting matrix is defined by the Kronecker product

$$K_{n_i} = K_{(n_i)_1} \otimes \cdots \otimes K_{(n_i)_d}. \tag{2.10}$$

Combining (2.9) with (2.10) and due to the properties of the Kronecker product, we have

$$\begin{aligned} K_{n_i} F_{n_i} &= K_{(n_i)_1} F_{(n_i)_1} \otimes \cdots \otimes K_{(n_i)_d} F_{(n_i)_d} \\ &= \frac{1}{\sqrt{2^d}} \left(F_{(n_{i+1})_1} \left([1, 1] \otimes I_{(n_{i+1})_1} \right) \right) \otimes \cdots \otimes \left(F_{(n_{i+1})_d} \left([1, 1] \otimes I_{(n_{i+1})_d} \right) \right) \\ &= \frac{1}{\sqrt{2^d}} F_{n_{i+1}} \Theta_{n_{i+1}}, \end{aligned} \tag{2.11}$$

where $\Theta_{n_{i+1}} = ([1, 1] \otimes I_{(n_{i+1})_1}) \otimes \cdots \otimes ([1, 1] \otimes I_{(n_{i+1})_d})$. With the help of the cutting matrix, the prolongation is now defined as

$$P_i = \mathcal{C}_{n_i}(p_i) K_{n_i}^T$$

given some generating symbol p_i, and the restriction is defined as the adjoint of the prolongation, i.e., $R_i = P_i^H$. To study the approximation property, we first define the set $\Omega(x)$ of all "corners" of x given by

$$\Omega(x) = \{y : y_j \in \{x_j, x_j + \pi\}\}$$

and the set $\mathcal{M}(x)$ of all "mirror points" ([11]) of x as

$$\mathcal{M}(x) = \Omega(x)\backslash\{x\}.$$

To obtain optimality, i.e., level independent multigrid convergence, the generating symbol p_i of the prolongation has to fulfill certain properties. For that purpose let $x^0 \in [-\pi, \pi)^d$ be the single isolated zero of the generating symbol f_i of the system matrix on level i. Choose p_i such that

$$\limsup_{x \to x^0} \max_{y \in \mathcal{M}(x)} \left| \frac{p_i(y)}{f_i(x)} \right| < +\infty, \qquad i = 0, \ldots, l_{\max} - 1, \tag{2.12}$$

and such that for all $x \in [-\pi, \pi)^d$ we have

$$0 < \sum_{y \in \Omega(x)} |p_i|^2(y), \qquad i = 0, \ldots, l_{\max} - 1. \tag{2.13}$$

Using (2.12) and (2.13) (cf. [1]), the approximation property (2.6) can be verified for circulant matrices with a constant independent of the level i, and the V-cycle optimality can be proved.

THEOREM 2.4 ([1]). *Let $A_i = \mathcal{C}_{n_i}(f_i)$ with f_i being the d-variate nonnegative generating symbol of A_i having a single isolated zero in $[-\pi, \pi)^d$. If smoothers are chosen according to Theorem 2.3 and the projectors $P_i = \mathcal{C}_{n_i}(p_i) K_{n_i}^H$ and $R_i = P_i^H$ such that p_i fulfills (2.12) and (2.13), then*

$$\|MGM\|_A \leq \xi < 1,$$

where MGM is the V-cycle iteration matrix and ξ is independent of $l_{\max}$.

If the order of the zero x^0 of the generating symbol is $2q$, a natural choice for p_i is

$$p_i(x) = c \cdot \prod_{j=1}^{d} \left(\cos(x_j^0) + \cos(x_j) \right)^q,$$

with a constant c.

If the system matrix A is not circulant but Toeplitz, a few changes are necessary. In the case of a Toeplitz matrix that has a generating symbol f being a trigonometric polynomial of degree at most one, the matrix is in the τ-algebra. Matrices out of the τ-algebra are diagonalized by the matrix

$$(Q_n)_{j,k} = \sqrt{\frac{2}{n+1}} \sin\left(\frac{jk\pi}{n+1}\right), \qquad j, k = 1, \ldots, n.$$

Assuming n_i odd, the cutting matrix K_{n_i} is chosen as

$$K_{n_i} = \begin{bmatrix} 0 & 1 & 0 & & & & & \\ & & & 1 & 0 & & & \\ & & & & & \ddots & & \\ & & & & & & 1 & 0 \end{bmatrix} \tag{2.14}$$

in the τ-case, and the results on multilevel matrices and convergence transfer to this case immediately if Q_n is taken instead of F_n and the appropriate cutting matrix is used. If A is

ETNA
Kent State University
http://etna.math.kent.edu

Toeplitz but the generating symbol is a higher degree trigonometric polynomial of degree δ, the cutting matrix has to be chosen as

$$K_{n_i}(\delta) = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & & & & & & \\ & & & & & 1 & 0 & & & & \\ & & & & & & & \ddots & & & \\ & & & & & & & & 1 & 0 & \cdots & 0 \end{bmatrix}, \tag{2.15}$$

where the first and last δ columns are zero, so the non-constant entries in the first δ and in the last δ rows and columns are not taken into account on the coarser level to guarantee the Toeplitz structure on all levels.

We now focus on the choice of p_i in an aggregation based framework.

3. Aggregation and SA for circulant matrices.

In the following, we start with the definition of simple aggregation based multigrid methods for 1-level circulant matrices corresponding to one dimensional problems. Emphasizing the downside of pure aggregation, we then introduce SA in the circulant setting and finally transfer the results to the d-level case.

3.1. 1-level case.

Let $n = 2^{l_{\max}+1}$. In a 1D aggregation-based multigrid method with aggregates of size 2, this corresponds to a prolongation operator P_i given by

$$P_i^H = \begin{bmatrix} 1 & 1 & & & & & \\ & & 1 & 1 & & & \\ & & & & \ddots & & \\ & & & & & 1 & 1 \end{bmatrix} \in \mathbb{C}^{n_{i+1} \times n_i}.$$

Transferring this to the circulant case yields a prolongation $P_i = \mathcal{C}_{n_i}(p_i) K_{n_i}^T$ with $p_i = a_{1,2}$, where

$$\begin{aligned} a_{1,2} : [-\pi, \pi) &\to \mathbb{C} \\ x &\mapsto a_{1,2}(x) = 1 + \mathrm{e}^{-\mathrm{i}x}. \end{aligned}$$

Note that $\mathcal{C}_{n_i}(p_i)$ is not Hermitian. This projector fulfills (2.13) since

$$\sum_{y \in \Omega(x)} |a_{1,2}(y)|^2 = \sum_{y \in \Omega(x)} \left|1 + \mathrm{e}^{-\mathrm{i}y}\right|^2 = \sum_{y \in \Omega(x)} 2 + 2\cos(y) > 0\,.$$

If the symbol f_i has a single isolated zero of order 2 at the origin, like the Laplacian, this projection does not fulfill (2.12), but it fulfills a weaker condition sufficient for two-grid optimality, namely

$$\limsup_{x \to x^0} \max_{y \in \mathcal{M}(x)} \frac{|p_i(y)|^2}{|f_i(x)|} \le +\infty, \qquad i = 0, \ldots, l_{\max} - 1. \tag{3.1}$$

Hence, the aggregation defines an optimal two-grid method, but it is not strong enough for the optimality of a V-cycle. This agrees with results in [20].

To fulfill the stronger condition (2.12), the prolongation can be improved by smoothing, i.e., applying a step of an iterative method used as a smoother. In the case of Richardson iteration this corresponds to the generating symbol

$$s_{i,\omega}(x) = 1 - \omega f_i(x).$$

ETNA
Kent State University
http://etna.math.kent.edu

Under the assumption that f_i has its single maximum at position $x = \pi$, no additional zero is introduced in $s_{i,\omega}$ when ω is chosen as $\omega = 1/f(\pi)$, and the symbol of the prolongation operator

$$p_i(x) = s_{i,1/f(\pi)}(x)\, a_{1,2}(x)$$

fulfills (2.12) since $s_{i,1/f(\pi)}(\pi) = 0$.

We like to note that if the introduced zero is of second order, it suffices to smooth either the prolongation or the restriction operator, as the symbol of the pure aggregation already has a zero of order 1 at the mirror point $x = \pi$. Since in this case $R_i \neq P_i^H$, the previous theory does not apply. Nevertheless, defining $R_i = K_{n_i}\mathcal{C}_{n_i}(r_i)$, in [8] it is shown that condition (2.13) can be replaced with

$$0 < \sum_{y\in\Omega(x)} r_i(y)p_i(y), \qquad i = 0,\ldots,l_{\text{max}} - 1, \tag{3.2}$$

and the two-grid condition (3.1) with

$$\limsup_{x\to x^0} \max_{y\in\mathcal{M}(x)} \frac{|r_i(y)p_i(y)|}{|f_i(x)|} \leq +\infty, \qquad i = 0,\ldots,l_{\text{max}} - 1.$$

Similarly, assuming that $r_ip_i \geq 0$, condition (2.12) can be replaced with

$$\limsup_{x\to x^0} \max_{y\in\mathcal{M}(x)} \left| \frac{\sqrt{r_i(y)p_i(y)}}{f_i(x)} \right| < +\infty, \qquad i = 0,\ldots,l_{\text{max}} - 1. \tag{3.3}$$

The resulting coarse matrix $A_{i+1} = R_iA_iP_i$ is $A_{i+1} = \mathcal{C}_n(f_{i+1})$ with

$$f_{i+1}(x) = \frac{1}{2} \sum_{y\in\Omega(x/2)} r_i(y)f_i(y)p_i(y), \tag{3.4}$$

and hence it is nonnegative definite for $r_ip_i \geq 0$. Smoothing only the restriction or the prolongation operator, we have

$$r_i(x)p_i(x) = s_{i,\omega}(x)a_{1,2}(x)\overline{a_{1,2}(x)} = s_{i,\omega}(x)(2 + 2\cos(x)).$$

Under the assumption that f has its maximum at π, $s_{i,1/f(\pi)}$ is nonnegative and has a zero of order at least 2 at π. Hence, conditions (3.2) and (3.3) are satisfied, and A_{i+1} is nonnegative definite.

REMARK 3.1. This choice of p_i is only valid for system matrices $A = \mathcal{C}_n(f)$ where the generating symbol has a single isolated zero at $x_0 = 0$. In general for a system matrix with generating symbol f_i having a single isolated zero at x_0, we choose p_i as

$$\begin{aligned} p_i : [-\pi,\pi) &\to \mathbb{C} \\ x &\mapsto p_i(x) = 1 + \mathrm{e}^{-\mathrm{i}(x+x_0)}. \end{aligned}$$

For this prolongation operator we have

$$|p_i(x)|^2 = 2 + 2\cos(x + x_0),$$

so (2.13) and (3.1) are fulfilled, the latter for a single isolated zero x_0 of order 2. The stronger condition (2.12) is fulfilled in the case that f_i has its single maximum at $x_0 + \pi$ by smoothing the operator using ω-Richardson iteration with $\omega = f_i(x_0 + \pi)$.

ETNA
Kent State University
http://etna.math.kent.edu

In general, aggregation with aggregates of sizes g corresponds to using the cutting matrix

$$K_{n_i,g} = \begin{bmatrix} 1 & 0 & \cdots & 0 & & & & & & & & \\ & & & & 1 & 0 & \cdots & 0 & & & & \\ & & & & & & & & \ddots & & & \\ & & & & & & & & & 1 & 0 & \cdots & 0 \end{bmatrix} \tag{3.5}$$

with $g-1$ zero columns after each column containing a one. The prolongation defined by this cutting matrix and the generating symbol $p_i = a_{1,g}$ with

$$\begin{aligned} a_{1,g} : [-\pi, \pi) &\to \mathbb{C} \\ x \mapsto a_{1,g}(x) &= \sum_{k=0}^{g-1} \mathrm{e}^{-\mathrm{i}kx} \end{aligned}$$

is

$$P_i = \mathcal{C}_{n_i}(p_i) K_{n_i,g}^T. \tag{3.6}$$

The effect of the cutting matrix applied to the Fourier matrix is similarly to (2.9) described by

$$K_{n_i,g} F_{n_i} = \frac{1}{\sqrt{g}} e_g^T \otimes F_{n_{i+1}} = \frac{1}{\sqrt{g}} F_{n_{i+1}} (e_g^T \otimes I_{n_{i+1}}),$$

where $e_g^T = [1, \ldots, 1] \in \mathbb{N}^g$, and the set of mirror points consists of the $g-1$ points in $\mathcal{M}_g(x) = \Omega_g(x) \backslash \{x\}$, where

$$\Omega_g(x) = \left\{ y : y = x + \frac{2\pi j}{g} (\mathrm{mod} 2\pi),\ j = 0, 1, \ldots, g-1 \right\}.$$

Assuming $n_0 = n = g^{l_{\max}+1}$, for a given matrix $A_i = \mathcal{C}_{n_i}(f_i)$, the coarse-level matrix $A_{i+1} = P_i^H A_i P_i$, $n_{i+1} = n_i / g$, is given by $A_{i+1} = \mathcal{C}_{n_{i+1}}(f_{i+1})$ with

$$f_{n_{i+1}}(x) = \frac{1}{g} \sum_{y \in \Omega_g(x/g)} |p|^2 f(y), \qquad x \in [-\pi, \pi).$$

For further details see [9], where it is proved that the two-grid convergence follows as in the case $g = 2$ outlined in Section 2.3 with the requirements (3.1) and (2.13) stated on the sets $\mathcal{M}_g$ and Ω_g, respectively. In more detail, the two-grid optimality requires

$$\limsup_{x \to x^0} \max_{y \in \mathcal{M}_g(x)} \frac{|p_i(y)|^2}{|f_i(x)|} \leq +\infty, \qquad i = 0, \ldots, l_{\max} - 1, \tag{3.7}$$

$$0 < \sum_{y \in \Omega_g(x)} |p_i|^2(y), \qquad i = 0, \ldots, l_{\max} - 1, \tag{3.8}$$

for all $x \in [-\pi, \pi)$; see Theorem 5.1 in [9]. The V-cycle optimality for a coarsening factor $g > 2$ is an open problem, but a natural conjecture is that in (2.12), similarly to (3.1), it is enough to replace $\mathcal{M}$ with $\mathcal{M}_g$, namely

$$\limsup_{x \to x^0} \max_{y \in \mathcal{M}_g(x)} \left| \frac{p_i(y)}{f_i(x)} \right| < +\infty, \qquad i = 0, \ldots, l_{\max} - 1. \tag{3.9}$$

As the pure aggregation $p_i = a_{1,g}$ fulfills only (3.7) but not (3.9), the prolongation has to be improved for all mirror points possibly resulting in more than one smoothing parameter ω and thus multiple necessary smoothing steps. Note that the extension of these results to the case of zeros at other positions is possible analogously to the case outlined in Remark 3.1 with the same symbol $p_i(x) = 1 + \mathrm{e}^{-\mathrm{i}(x + x_0)}$.

ETNA
Kent State University
http://etna.math.kent.edu

3.2. Cutting in the d-level case for $d>1$. Using the 1-level case as motivation, prior to introducing aggregation and SA multigrid for d-level circulant matrices with $d \in \mathbb{N}$ (usually associated to d-dimensional problems), we have to extend the theoretical results in [9] to $d > 1$. For that purpose, let $A = \mathcal{C}_n(f)$, where $f : \mathbb{R}^d \to \mathbb{C}$ is a nonnegative function 2π-periodic in each variable, $n \in \mathbb{N}^d$, and $g \in \mathbb{N}^d$ be the size of the aggregates. Assume that $n = g^{l_{\max}+1}$, i.e., $n_j = g_j^{l_{\max}+1}$, $j = 1,\ldots,d$. As before, we define the fine-level operator $A_0 = A$, with $f_0 = f$, and recursively the system size as $n_{i+1} = n_i/g$ (all the multi-index operations in the paper are intended componentwise), the prolongation as in (3.6), where $K_{n_i,g} = K_{(n_i)_1,g_1} \otimes \cdots \otimes K_{(n_i)_d,g_d}$, and the coarse-grid operator as $A_{i+1} = P_i^H A_i P_i$. The set of all corners of $x \in \mathbb{R}^d$ associated to the cutting matrix $K_{n_i,g}$ is

$$\Omega_g(x) = \left\{ y \;\middle|\; y_j \in \left\{ x_j + \frac{2\pi k}{g_j} \,(\mathrm{mod}\, 2\pi) \right\},\ k = 0,\ldots,g_j-1,\ j = 1,\ldots,d \right\}.$$

To simplify the following notation we define $G = \prod_{j=1}^d g_j$.

Analogously to the 1-level case, the generating symbol of the system matrix of the coarser level is given as stated in the following lemma.

LEMMA 3.2. *Let $A_i = \mathcal{C}_{n_i}(f_i)$, P_i defined in (3.6), and let $n_{i+1} = g \cdot n_i \in \mathbb{N}^d$, where the multiplication is intended componentwise. Then the coarse-level system matrix $A_{i+1} = P_i^H A_i P_i$ is $A_{i+1} = \mathcal{C}_{n_{i+1}}(f_{i+1})$, where*

$$f_{i+1}(x) = \frac{1}{G} \sum_{y \in \Omega_g(x/g)} |p_i|^2 f_i(y), \qquad x \in [-\pi,\pi)^d. \tag{3.10}$$

Proof. The proof is a generalization of the proof of [25, Proposition 5.1]. First we note that in analogy to (2.11), we have

$$\begin{aligned} K_{n_i,g}F_{n_i} &= K_{(n_i)_1,g_1}F_{(n_i)_1} \otimes \cdots \otimes K_{(n_i)_d,g_d}F_{(n_i)_d} \\ &= \frac{1}{\sqrt{G}} \left(F_{(n_{i+1})_1}\left(e_{g_1}^T \otimes I_{(n_{i+1})_1}\right)\right) \otimes \cdots \otimes \left(F_{(n_{i+1})_d}\left(e_{g_d}^T \otimes I_{(n_{i+1})_d}\right)\right) \\ &= \frac{1}{\sqrt{G}} \left(F_{(n_{i+1})_1} \otimes \cdots \otimes F_{(n_{i+1})_d}\right)\left(\left(e_{g_1}^T \otimes I_{(n_{i+1})_1}\right) \otimes \cdots \otimes \left(e_{g_d}^T \otimes I_{(n_{i+1})_d}\right)\right) \end{aligned}$$

so that

$$K_{n_i}F_{n_i} = \frac{1}{\sqrt{G}} F_{n_{i+1}} \Theta_{n_{i+1},g},$$

where $\Theta_{n_{i+1},g} = (e_{g_1}^T \otimes I_{(n_{i+1})_1}) \otimes \cdots \otimes (e_{g_d}^T \otimes I_{(n_{i+1})_d})$. Hence, for $A_{i+1} = P_i^H A_i P_i$ we have

$$\begin{aligned} P_i^H A_i P_i &= K_{n_i,g}\mathcal{C}_{n_i}^H(p_i)\mathcal{C}_{n_i}(f_i)\mathcal{C}_{n_i}(p_i)K_{n_i,g}^H = K_{n_i,g}F_{n_i}D_{n_i}\left(|p_i|^2 f_i\right)F_{n_i}^H K_{n_i,g}^H \\ &= \frac{1}{G} F_{n_{i+1}}\Theta_{n_{i+1},g}D_{n_i}\left(|p_i|^2 f_i\right)\Theta_{n_{i+1},g}^H F_{n_{i+1}}^H. \end{aligned}$$

Here,

$$D_{n_i}(f) = \mathrm{diag}_{0 \le j \le n_i - e_d}\left(f((x_i)_j)\right),$$

where $(x_i)_j \equiv 2\pi j/n_i = (2\pi j_1/(n_i)_1, \ldots, 2\pi j_d/(n_i)_d)^T$ and $0 \le j \le n_i - e_d$ means that $0 \le j_k \le (n_i)_k - 1$, for $k = 1,\ldots,d$, assuming the standard lexicographic ordering. All

ETNA
Kent State University
http://etna.math.kent.edu

operations and inequalities between multi-indices are intended componentwise. For a given multi-index $k = (k_1, \ldots, k_d)$, $0 \le k_j \le (n_{i+1})_j$, we have

$$\left(\Theta_{n_{i+1},g} x\right)_k = \sum_{l=0}^{g-e_d} x_{k+l},$$

so we obtain

$$\Theta_{n_{i+1},g} D_{n_i}\left(|p_i|^2 f_i\right) \Theta_{n_{i+1},g}^T = \sum_{l=0}^{g-e_d} D_{n_i,g,l}\left(|p_i|^2 f_i\right),$$

where

$$D_{n_i,g,l}(f) = \operatorname{diag}_{n_{i+1}\cdot l \le j' \le n_{i+1}\cdot(l+e_d)-e_d}\left(f((x_i)_{j'})\right).$$

For an example of the multi-index notation in the case $d = g = 2$, we refer to the proof of Proposition 5.1 in [25]. As a result we obtain

$$P_i^H A_i P_i = \frac{1}{G} F_{n_{i+1}} \left(\sum_{l=0}^{g-e_d} D_{n_i,g,l}\left(|p_i|^2 f_i\right) \right) F_{n_{i+1}}^H,$$

and with

$$(x_i)_{j'} = (x_{i+1})_j / g + \pi \cdot l \pmod{2\pi}, \quad 0 \le j \le n_{i+1} - e_d,\ j' = j + n_{i+1}\cdot l,$$

where l is a multi-index and products and sums are intended componentwise, we obtain $P_i^H A_i P_i = \mathcal{C}_{n_{i+1}}(f_{i+1})$, with f_{i+1} defined in (3.10). $\square$

REMARK 3.3. If the two conditions (3.7) and (3.8) are satisfied with $x \in [-\pi, \pi)^d$, we obtain as a consequence of Lemma 3.2 that if x^0 is a zero of f_i, then $g \cdot x^0 \bmod 2\pi$ is a zero of f_{i+1} of the same order.

The two-grid optimality can be obtained similarly to the 1-level case. The following result shows that the two-grid conditions (3.7) and (3.8) are sufficient for condition (2.7).

THEOREM 3.4. *Let $A_i := C_{n_i}(f_i)$, with f_i being a d-variate nonnegative trigonometric polynomial (not identically zero), and let $P_i = C_{n_i}(p_i) K_{n_i,g}^T$ be the prolongation operator with p_i a trigonometric polynomial satisfying condition* (3.7) *for any zero of f_i and globally satisfying condition* (3.8)*. Then, there exists a positive value γ independent of n_i such that inequality* (2.7) *is satisfied.*

Proof. The proof is a combination of [9, Theorem 5.1] and [25, Lemma 6.3], but we report it here for completeness. First, we recall that the main diagonal of A_i is given by $D_i = a_i I_{N_i}$ with $a_i = (2\pi)^d \int_{[-\pi,\pi)^d} f_i(x)dx > 0$ so that $\|\cdot\|_{D_i}^2 = y = a_i\|\cdot\|_2^2$.

In order to prove that there exists a value $\gamma > 0$ independent of n_i such that for any $x \in \mathbb{C}^{N_i}$

$$\min_{y \in \mathbb{C}^{N_{i+1}}} \|x - P_i y\|_{D_i}^2 = a_i \min_{y \in \mathbb{C}^{N_{i+1}}} \|x - P_i y\|_2^2 \le \gamma \|x\|_{A_i}^2,$$

we choose a special instance of y. For any $x \in \mathbb{C}^{N_i}$, let $\overline{y} \equiv \overline{y}(x) \in \mathbb{C}^{N_{i+1}}$ be defined as $\overline{y} = \left[P_i^H P_i\right]^{-1} P_i^H x$. Therefore, (2.7) is implied by

$$\|x - P_i \overline{y}\|_2^2 \le (\gamma / a_i) \|x\|_{A_i}^2, \qquad \forall x \in \mathbb{C}^{N_i},$$

where the latter is equivalent to the spectral matrix inequality

$$W_{n_i}(p_i)^H W_{n_i}(p_i) \le (\gamma/a_i) C_{n_i}(f_i), \tag{3.11}$$

with $W_{n_i}(p_i) = I_{N_i} - P_i \left[P_i^H P_i\right]^{-1} P_i^H$. Given two matrices A and B, the matrix inequality $A \le B$ means that the matrix $B - A$ is Hermitian and positive semi-definite. Since $W_{n_i}(p_i)^H W_{n_i}(p_i) = W_{n_i}(p_i)$, inequality (3.11) can be rewritten as

$$W_{n_i}(p_i) \le (\gamma/a_i) C_{n_i}(f_i). \tag{3.12}$$

Let $\mu = (\mu_1, \dots, \mu_d)$ with $0 \le \mu_r \le (n_{i+1})_r - 1$, $r = 1, \dots, d$, and let $p_i[\mu] \in \mathbb{C}^G$ whose entries are given by the evaluations of p_i over the points of $\Omega(x_\mu^{(n_i)})$ with $x_\mu^{(n_i)} = (2\pi\mu_1/(n_i)_1, \dots, 2\pi\mu_d/(n_i)_d)$. Using the same notation for $f_i[\mu]$, we denote by $\mathrm{diag}(f_i[\mu])$ the diagonal matrix having the vector $f_i[\mu]$ on the main diagonal. There exists a suitable permutation by rows and columns of $F_{n_i}^H W_{n_i}(p_i) F_{n_i}$ such that we can obtain a $G \times G$ block diagonal matrix and the condition (3.12) is equivalent to

$$I_G - \frac{p_i[\mu](p_i[\mu])^T}{\|p_i[\mu]\|_2^2} \le (\gamma/a_i)\mathrm{diag}(f_i[\mu]), \qquad \forall \mu. \tag{3.13}$$

By the Sylvester inertia law [13], the relation (3.13) is satisfied if every entry of

$$\mathrm{diag}(f_i[\mu])^{-1/2} \left(I_G - \frac{p_i[\mu](p_i[\mu])^T}{\|p_i[\mu]\|_2^2} \right) \mathrm{diag}(f_i[\mu])^{-1/2}$$

is bounded in modulus by a constant, which follows from conditions (3.7) and (3.8) as it is shown in detail in the proof of Proposition 4 in [1]. ∎

Since the post-smoothing property holds unchanged, combining Theorem 2.3 and Theorem 3.4 with Theorem 2.2, it follows that the two-grid convergence speed does not depend on the size of the linear system.

3.3. The aggregation operator. In the pure aggregation setting, the generating symbol of the prolongation is given by

$$a_{d,g}(x) = \prod_{j=1}^{d} \sum_{k=0}^{g_j - 1} \mathrm{e}^{-\mathrm{i}kx_j}, \qquad x \in [-\pi, \pi)^d. \tag{3.14}$$

THEOREM 3.5. *For the function $a_{d,g}$ defined in (3.14), there exists a constant c with $0 < c < +\infty$ such that*

$$\limsup_{x \to 0} \max_{y \in \mathcal{M}_g(x)} \frac{|a_{d,g}(y)|}{\sum_{j=1}^{d} x_j^z} = c, \tag{3.15}$$

where

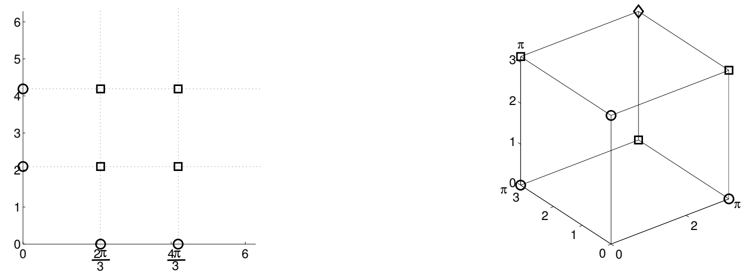
$$z := d - \#\left\{ y_j \mid y_j = 0,\ j = 1, \dots, d \right\}$$

is the number of directions along which $a_{d,g}$ is zero. Furthermore, if f_i has a single isolated zero of order 2 at the origin, $p_i = a_{d,g}$ fulfills (3.8) and (3.7).

Proof. The limit (3.15) follows from the Taylor series of $a_{d,g}$: consider $y \in \mathcal{M}_g(x)$, i.e., $y_j = x_j + \frac{2\pi\ell}{g_j} \pmod{2\pi}$ for $\ell = 0, \dots, g_j - 1$, then the j-th factor of $a_{d,g}(y)$ is

$$\sum_{k=0}^{g_j - 1} \mathrm{e}^{-\mathrm{i}ky_j} = \sum_{k=0}^{g_j - 1} \mathrm{e}^{-\mathrm{i}k(x_j + \frac{2\pi\ell}{g_j})} = \sum_{k=0}^{g_j - 1} \mathrm{e}^{\frac{-\mathrm{i}2\pi k\ell}{g_j}} \mathrm{e}^{-\mathrm{i}kx_j}.$$

ETNA
Kent State University
http://etna.math.kent.edu



$d = 2,\ g = (3, 3)$ $d = 3,\ g = (2, 2, 2)$

FIG. 3.1. *Order of* $y \in \mathcal{M}_g(0)$ *for the aggregation operator* $a_{d,g}$*:* $\circ \to$ *order* $= 1$, $\Box \to$ *order* $= 2$, $\Diamond \to$ *order* $= 3$.

Since

$$\sum_{k=0}^{g_j-1} \mathrm{e}^{\frac{-\mathrm{i}2\pi k\ell}{g_j}} = \begin{cases} g_j & \text{if } \ell = 0, \\ 0 & \text{otherwise,} \end{cases}$$

the j-th factor in (3.14) has an infinite Taylor series with the constant term equal to zero only if $\ell \neq 0$.

If f_i has a single isolated zero of order 2 at the origin, then

$$\limsup_{x\to 0} \frac{f_i(x)}{\sum_{j=1}^d x_j^2} = \hat{c}, \qquad 0 < \hat{c} < +\infty,$$

and hence $p_i = a_{d,g}$ fulfills (3.7).

Regarding (3.8), let x be such that $|a_{d,g}|^2(x) = 0$. If x lies on the axes, then $0 \in \Omega_g(x)$ and $|a_{d,g}|^2(0) > 0$. If x does not lie on the axes, then there exists a $y \in \Omega_g(x)$ that lies on an axis and fulfills $|a_{d,g}|^2(y) > 0$. $\Box$

Figure 3.1 gives a visual representation of the behavior of $p_i = a_{d,g}$ at $\mathcal{M}_g(0)$ for two examples. The previous Theorem 3.5 states that if the symbol f has a zero at the origin of order two, then the two-grid method is optimal. On the other hand, the V-cycle cannot be optimal since $p_i = a_{d,g}$ vanishes only with order one at the mirror points located along the cardinal axes. For the same reason, when f vanishes at the origin with a zero of order greater than two, e.g., for the biharmonic problem, also the aggregation two-grid method cannot be optimal. To overcome this weakness of the aggregation operator, smoothing techniques for the projector are usually employed. A simple strategy of this kind will be analyzed in the next section.

3.4. Smoothing the projector by weighted Richardson iteration. The order of the zeros at the points where $p_i = a_{d,g}$ is zero in one direction only can be improved by applying smoothing. For that purpose we again use an ω-Richardson smoother. In the d-level case the generating symbol of this smoother is given by

$$\begin{aligned} s_{i,\omega} &: [-\pi, \pi)^d \to \mathbb{C} \\ &x \to s_{i,\omega}(x) = 1 - \omega f_i(x). \end{aligned} \tag{3.16}$$

ETNA
Kent State University
http://etna.math.kent.edu

LEMMA 3.6. *Assume that $f_i \geq 0$ has a single isolated zero of order 2 at the origin and that f_i attains the maximum only at all $y \in \mathcal{M}_g(0)$ lying on the axes, and let $\tilde{y}$ be one of these points. Then the symbol of the smoothed prolongation given by*

$$p_i(x) = s_{i,1/f(\tilde{y})}(x)\, a_{d,g}(x)$$

fulfills (3.9) *and* (3.8).

Proof. Since $\tilde{y}$ is a point of maximum for f_i, the function $s_{i,1/f(\tilde{y})}$ is nonnegative and vanishes for $y \in \mathcal{M}_g(0)$ lying on the axes with order at least one. From Theorem 3.5, $a_{d,g}$ vanishes at $y \in \mathcal{M}_g(0)$ with order one if y lies on the axes and with order at least two otherwise. Therefore, $p_i = s_{i,1/f(\tilde{y})}\, a_{d,g}$ vanishes with order at least two for all $y \in \mathcal{M}_g(0)$, and hence it fulfills (3.9).

Regarding (3.8), the assumptions on f_i imply that $s_{i,1/f(\tilde{y})}(y) = 0$ only for $y \in \mathcal{M}_g(0)$ lying on the axes. Hence, $\{x \,|\, s_{i,1/f(\tilde{y})}(x) = 0\} \subset \{x \,|\, a_{d,g}(x) = 0\}$ and $p_i = a_{d,g} s_{i,1/f(\tilde{y})}$ fulfills (3.8) since it is already satisfied by $p_i = a_{d,g}$ thanks to Theorem 3.5. ∎

In general ω should be chosen to improve the projector where the aggregation operator is less effective, that is, at the mirror points located along the cardinal axes, i.e., the points belonging to

$$\mathcal{A}_g(0) = \big\{y \in \mathcal{M}_g(0),\ \#\{y_j \,|\, y_j \neq 0,\ j = 1, \dots, d\} = 1\big\}, \tag{3.17}$$

the set of mirror points where only one component does not vanish. Therefore, ω is obtained by imposing that $s_{i,\omega}(y) = 0$ for a certain $y \in \mathcal{A}_g(0)$. If different points in $\mathcal{A}_g(0)$ lead to different values of ω, then more smoothing steps with different ω's should be added to the aggregation. For a detailed discussion see Section 4.3.

Again, if the smoother introduces a zero of order two, it is sufficient to smooth either the prolongation or the restriction operator generalizing the results in [8] to $g > 2$. Moreover, like in Remark 3.1, the aggregation operator for a zero at a position $x^0 \neq 0 \in \mathbb{R}^d$ is defined by

$$p_i(x) = \prod_{j=1}^{d} \sum_{k=0}^{g_j - 1} \mathrm{e}^{-\mathrm{i}k(x_j + x_j^0)}, \qquad x \in [-\pi, \pi)^d.$$

4. Analysis and design of SA for some classes of matrices. Firstly, we observe that the theoretical results obtained in the previous section to design SA for circulant matrices can be applied in a straightforward way to Toeplitz matrices. Subsequently, we study in detail matrices arising from the finite difference discretization of some PDEs.

As noted at the end of Section 2.3, the circulant case can be applied and extended to the Toeplitz case. In analogy to (2.14), the cutting matrix $K_{n_i,g}$ given by (3.5), in the Toeplitz case reads as

$$K_{n_i,g} = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & & & & & & & & \\ & & & & & & & 1 & 0 & \cdots & 0 & & & & \\ & & & & & & & & & & & \ddots & & & \\ & & & & & & & & & & & & 1 & 0 & \cdots & 0 \end{bmatrix},$$

where the first and last $(g-1)/2$ columns are zero. The multilevel counterpart is formed with the help of Kronecker products. In the case that the degree of the trigonometric polynomial is smaller than the maximum of all components of g, the Toeplitz structure is kept on the coarser levels. This is a general advantage of multigrid methods that use reductions of the system size greater than 2. If the degree is higher, the cutting matrix can be padded with zeros as in (2.15).

ETNA
Kent State University
http://etna.math.kent.edu

In the following we consider the finite difference discretization of PDEs, in particular the 2D Laplacian, with constant coefficients. Nevertheless, the analysis can be used to design a SA multigrid also in the case of non-constant coefficients. Indeed, while non-constant coefficients do not lead to circulant or Toeplitz matrices, circulant or Toeplitz matrices can be used as a local model by freezing the coefficients and analyzing the resulting stencils by the methods derived for the constant coefficient case. This approach is employed in [8] and is similar to local Fourier analysis (LFA) for multigrid methods, which is used to analyze geometric multigrid methods. For a detailed review of LFA, see [29]. The developed theory can be used to choose different smoothers based on the local stencil within the smoothing process in general SA multigrid methods. Hence, the used smoother is of the form $I - \Omega_i A_i$ with a diagonal matrix Ω_i, where each Ω_i is related to a frozen local stencil. This strategy will be employed in Section 5.6.

4.1. Symmetric projection for the 2D Laplacian. Now we turn to the finite difference discretization of the 2D Laplacian with constant coefficients. In this case we are able to formulate some results based on the developed theory. In the following we allow only the same coarsening in x and y direction, and therefore we denote the coarsening g by only one integer, $g = 2, 3, 4$, or 5.

LEMMA 4.1. *Let f be an even trigonometric polynomial obtained by an isotropic discretization of the 2D Laplacian. If $g = 2$ or $g = 3$, there always exists a smoother $s_{i,\omega}$ defined in* (3.16) *with a unique ω such that the resulting projection $p_i = s_{i,\omega}\, a_{2,g}$ fulfills* (3.9). *In particular,*

i) for $g = 2$ we obtain $\omega = 1/f(0, \pi)$,
ii) for $g = 3$ we obtain $\omega = 1/f(0, \frac{2\pi}{3})$.

Proof. The function f is nonnegative and vanishes only at the origin with order two. The isotropic discretization leads to a symmetry of f such that $f(0, z) = f(z, 0)$ that is inherited by $s_{0,\omega}$. From (3.17) it holds that

$$\mathcal{A}_{(2,2)}(0) = \{(0, \pi),\, (\pi, 0)\} \quad \text{and} \quad \mathcal{A}_{(3,3)}(0) = \{(0, \tfrac{2\pi}{3}),\, (0, \tfrac{4\pi}{3}),\, (\tfrac{2\pi}{3}, 0),\, (\tfrac{4\pi}{3}, 0)\}.$$

Therefore, ω has to be chosen such that $s_{0,\omega}(0, \pi) = 1 - \omega f(0, \pi) = 0$ for $g = 2$ and $s_{0,\omega}(0, 4\pi/3) = s_{0,\omega}(0, 2\pi/3) = 1 - \omega f(0, 2\pi/3) = 0$ for $g = 3$. The coarse symbols f_i, $i > 0$, preserve the properties of f thanks to Lemma 3.2 and Remark 3.3. ∎

In the case that every fourth point is taken in each direction, i.e., the number of unknowns is reduced by a factor of 16, we obtain a similar result:

LEMMA 4.2. *Let f be an even trigonometric polynomial obtained by an isotropic discretization of the 2D Laplacian. If $g = 4$, then we need two smoothers with two different ω values given by $\omega_1 = 1/f(0, \pi/2)$ and $\omega_2 = 1/f(0, \pi)$ such that the resulting projection $p_i = s_{i,\omega_1} s_{i,\omega_2}\, a_{2,g}$ fulfills* (3.9). *For $g = 5$, the same results holds for $\omega_1 = 1/f(0, 2\pi/5)$ and $\omega_2 = 1/f(0, 4\pi/5)$.*

Proof. The proof is analogous to that of Lemma 4.1 using the sets $\mathcal{A}_{(4,4)}$ and $\mathcal{A}_{(5,5)}$. Two different values of ω are necessary in view of $\cos(\pi/2) = \cos(3\pi/2) \neq \cos(\pi)$ and $\cos(2\pi/5) = \cos(8\pi/5) \neq \cos(4\pi/5) = \cos(6\pi/5)$. ∎

For anisotropic stencils, even with standard coarsening, two ω values are needed.

LEMMA 4.3. *Let f be an anisotropic discretization of the 2D Laplacian. If $g = 2$, we need two different ω values given by $\omega_1 = 1/f(\pi, 0)$ and $\omega_2 = 1/f(0, \pi)$ such that the resulting projection $p_i = s_{i,\omega_1} s_{i,\omega_2}\, a_{2,g}$ fulfills* (3.9). *For $g = 3$, we also need two ω values, namely $\omega_1 = 1/f(2\pi/3, 0)$ and $\omega_2 = 1/f(0, 2\pi/3)$. For $g = 4$ and $g = 5$, four ω values are necessary.*

ETNA
Kent State University
http://etna.math.kent.edu

Proof. Due to the anisotropic discretization, $f(\pi, 0) \neq f(0, \pi)$ in general, and hence twice the number of ω's with respect to the isotropic case is required in Lemmas 4.1 and 4.2 . ∎

4.2. Non-symmetric projection for the 2D Laplacian. The SA projection is defined by applying the aggregation prolongation $C_n(a_{d,g})K_n^T$ in the restriction and the prolongation step and the additional smoothers $S_j := I - \omega_j \mathrm{diag}(A)^{-1}A$, $j = 1, \dots, k'$. In the symmetric application we include each S_j in the restriction and the prolongation generating in total $k = 2k'$ smoothing factors. In the nonsymmetric application we include each S_j only once, either in the restriction or in the prolongation resulting in $k = k'$. Hence, the coarse system is related to the matrix

$$K_n C_n^H(a_{d,g}) S_{k'} S_1 A S_1 \dots S_{k'} C_n(a_{d,g}) K_n^T$$

in the symmetric case and in the nonsymmetric application, e.g., to

$$K_n C_n^H(a_{d,g}) S_1..S_l A S_{l+1} \dots S_{k'} C_n(a_{d,g}) K_n^T \ .$$

THEOREM 4.4. *To maintain the original block tridiagonal structure also on the coarse levels, the overall number k of smoothers that can be included in both restriction and prolongation is restricted by $k < g$. Therefore, if we incorporate the smoothing only in the restriction* or *the prolongation, $k' = k < g$ smoothers are allowed in SA. If we use symmetric projection with $R_i^T = P_i$, then we have to satisfy $k' < g/2$, respectively, $2k' = k < g$:*

g	2	3	4	5
allowed k' for nonsymmetric case	1	2	3	4
allowed k' for symmetric application	0	1	1	2

Proof. The projection has block bandwidth given by $g - 1$ upper diagonals, and the matrix A and the smoothers are block tridiagonal with 1 upper diagonal. Hence, applying k smoothers leads to $g + k$ upper diagonals. Picking out every g-th diagonal gives a block tridiagonal 9-point stencil if $k < g$. ∎

THEOREM 4.5. *To arrive at the right number of zeros in the restriction/prolongation such that* (3.3) *holds, the number k' of smoothers necessary on the whole is given by:*

g	2	3	4	5
necessary k' in the isotropic case	1	1	2	2
necessary k' in the anisotropic case	2	2	4	4

Proof. The symmetric application of the aggregation gives the right order of zeros at all mirror points that are not lying on the coordinate axes. Following the analysis in Lemmas 4.1 and 4.2, the smoothers, respectively ω_j, have to be chosen to add zeros on $f(0, 2\pi j/g)$, $f(2\pi j/g, 0)$, $j = 1, \dots, g-1$. Because of the identities $\cos(2\pi/3) = \cos(4\pi/3)$, $\cos(2\pi/4) = \cos(6\pi/4)$ and $\cos(2\pi/5) = \cos(8\pi/5)$, $\cos(4\pi/5) = \cos(6\pi/5)$, in the isotropic case, many of the mirror points coincide, and it is only necessary to smooth the restriction or prolongation to satisfy (3.3). For the anisotropic case, we have to consider the two axes x and y separately and hence to double the number of smoothers like in Lemma 4.3. ∎

ETNA
Kent State University
http://etna.math.kent.edu

To achieve both goals for the order of zeros and the block tridiagonal structure, combining Theorems 4.4 and 4.5, we can apply the SA according to the following cases:

1. isotropic case and nonsymmetric projection for all g,
2. isotropic case and symmetric projection for $g = 3$ or $g = 5$,
3. anisotropic case and nonsymmetric projection for $g = 3$ or $g = 5$,
4. anisotropic case and symmetric projection for no g.

4.3. SA for the 2D Laplacian with 9-point stencils. Now we want to discuss exemplarily and in detail the application of the smoothed aggregation technique to the 2D Laplacian with 9-point stencils. We design the projections such that on all levels we derive again 9-point stencils, and we use smoothers in the projection to get zeros of order at least 2 at all mirror points besides the origin according to condition (3.3). Therefore, our analysis will be focused on obtaining a stable stencil according to the following definition.

DEFINITION 4.6. *A stencil associated to a symbol f_i is* stable *if there exist r_i and p_i that satisfy* (3.3) *and $f_{i+1} = \alpha_i f_i$ with $\alpha_i > 0$.*

Of course, if the stencil f at the finest level is stable, then the same holds for all f_i at the coarser levels $i = 1, \ldots, l_{\max}$.

Applying the nonsymmetric projection, e.g., by including the smoothers only in the prolongation or in the restriction, the coarse matrix will again be symmetric because of the cutting procedure, but the coarse system might get indefinite. Therefore, we have to analyze the resulting coarse-grid matrix and determine when it is symmetric positive definite. An obvious criterion that we use here is the M-matrix property.

According to items 1–3 at the end of the previous section, we study in detail items 1 and 2 for the isotropic stencil

$$\frac{1}{4+4c}\begin{bmatrix} -c & -1 & -c \\ -1 & 4+4c & -1 \\ -c & -1 & -c \end{bmatrix}, \qquad c \geq 0, \tag{4.1}$$

which is associated to the symbol

$$f(x,y) = \big(2 - \cos(x) - \cos(y) + c(2 - \cos(x+y) - \cos(x-y))\big)/(2+2c)\,, \tag{4.2}$$

and item 3 for the anisotropic case

$$f(x,y) = (1-\cos(x)) + b\,(1-\cos(y))\,, \qquad b > 0. \tag{4.3}$$

Firstly, we compute stable stencils for the isotropic case with nonsymmetric projection (item 1) for $g = 2, \ldots, 5$.

THEOREM 4.7. *For $g = 2$ and nonsymmetric smoothing, the stencil* (4.1) *with $c = 1/\sqrt{2}$ is stable. Moreover the coarse system is a block tridiagonal M-matrix for all $c > 0$.*

Proof. From the symbol (4.2), only one $\omega = (1+c)/(1+2c)$ is necessary to ensure that $1 - \omega f(0,\pi) = 0$ and hence to satisfy (3.3). Using the function

$$g(x,y) = f(x,y)(1-\omega f(x,y))(1+\cos(x))(1+\cos(y)),$$

from (3.4) it follows that

$$f_1(x,y) = \frac{1}{4}\big(g(\tfrac{x}{2},\tfrac{x}{2}) + g(\tfrac{x}{2}+\pi,\tfrac{x}{2}) + g(\tfrac{x}{2},\tfrac{x}{2}+\pi) + g(\tfrac{x}{2}+\pi,\tfrac{y}{2}+\pi)\big)\,.$$

This function can be evaluated at $(0,0)$, $(0,\pi)$, and (π,π) leading to

$$f_1(0,0) = 0, \qquad f_1(0,\pi) = \frac{1+2c}{4(1+c)}, \qquad f_1(\pi,\pi) = \frac{c}{1+2c}\,.$$

These function values are related to a 9-point stencil, respectively the trigonometric polynomial

$$f_1(x,y) = \sigma - \delta(\cos(x) + \cos(y)) - \epsilon \cos(x)\cos(y)$$

with

$$\sigma = \frac{1+6c+6c^2}{8(1+2c)(1+c)}\,, \qquad \delta = \frac{c}{4(1+2c)}\,, \qquad \epsilon = \frac{1+2c+2c^2}{8(1+2c)(1+c)},$$

resulting in the coarse-grid stencil

$$\frac{1}{8(1+2c)(1+c)}\begin{bmatrix} -1/4 - c/2 - c^2/2 & -c-c^2 & -1/4 - c/2 - c^2/2 \\ -c-c^2 & 1+6c+6c^2 & -c-c^2 \\ -1/4 - c/2 - c^2/2 & -c-c^2 & -1/4 - c/2 - c^2/2 \end{bmatrix},$$

which gives an M-matrix for all $c > 0$.

For a stable stencil the functions f and f_1 have to be equivalent up to a scalar factor, or $2c\delta = \epsilon$, which is satisfied for $c = \frac{1}{\sqrt{2}}$. $\square$

The following theorems can be shown using the same technique, where the coarse symbol f_1 is computed generalizing (3.4) to $g > 2$ like in Lemma 3.2.

THEOREM 4.8. *For $g = 3$ and nonsymmetric smoothing, the stencil* (4.1) *with $c = 1/\sqrt{2}$ is stable. Moreover, the coarse stencil*

$$\frac{1}{18(1+2c)(1+c)}\begin{bmatrix} -3-4.5c-3c^2 & 3/2-9c-12c^2 & -3-4.5c-3c^2 \\ 3/2-9c-12c^2 & 6+54c+60c^2 & 3/2-9c-12c^2 \\ -3-4.5c-3c^2 & 3/2-9c-12c^2 & -3-4.5c-3c^2 \end{bmatrix}$$

defines a block tridiagonal M-matrix for $c > \frac{-3+\sqrt{17}}{8} \approx 0.140388$.

THEOREM 4.9. *For $g = 4$ and nonsymmetric smoothing, the stencil* (4.1) *with $c = 0$ or $c = 1$ is stable. Moreover, the coarse stencil*

$$\frac{1}{8(1+c)(1+2c)^2}\times \begin{bmatrix} -5c-8c^2-5c^3 & -2-2c-8c^2-6c^3 & -5c-8c^2-5c^3 \\ -2-2c-8c^2-6c^3 & 8+28c+64c^2+44c^3 & -2-2c-4c^2-6c^3 \\ -5c-8c^2-5c^3 & -2-2c-8c^2-6c^3 & -5c-8c^2-5c^3 \end{bmatrix}$$

defines a block tridiagonal M-matrix for all $c > 0$.

THEOREM 4.10. *For $g = 5$ and nonsymmetric smoothing, the stencil* (4.1) *with $c = 1.910044687..$ and $c = 0.2296814707..$ is stable. Moreover, the coarse stencil*

$$\frac{1}{20(1+c)(1+2c)^2}\times \begin{bmatrix} 2-13c-24c^2-16c^3 & -9-4c-12c^2-8c^3 & 2-13c-24c^2-16c^3 \\ -9-4c-12c^2-8c^3 & 28+68c+144c^2+96c^3 & -9-4c-12c^2-8c^3 \\ 2-13c-24c^2-16c^3 & -9-4c-12c^2-8c^3 & 2-13c-24c^2-16c^3 \end{bmatrix}$$

defines a block tridiagonal M-matrix for $c > 0.1234139034$.

ETNA
Kent State University
http://etna.math.kent.edu

Consider now the isotropic case and symmetric projection (item 2).

THEOREM 4.11. *For $g = 3$ and symmetric smoothing, the stencil (4.1) with $c = 1$ or $c = 0$ is stable. Moreover, the coarse stencil*

$$\frac{1}{12(1+2c)^2(1+c)} \times \begin{bmatrix} -7c-12c^2-8c^3 & -3-4c-12c^2-8c^3 & -7c-12c^2-8c^3 \\ -3-4c-12c^2-8c^3 & 12+44c+96c^2+64c^3 & -3-4c-12c^2-8c^3 \\ -7c-12c^2-8c^3 & -3-4c-12c^2-8c^3 & -7c-12c^2-8c^3 \end{bmatrix}$$

defines a block tridiagonal M-matrix for all $c > 0$.

THEOREM 4.12. *For $g = 5$ and symmetric smoothing, the stencil (4.1) with $c = 0.1991083336..$ or $c = 0.8931363030..$ is stable. Moreover, the coarse-grid matrix is a block tridiagonal M-matrix for $c > 0.1475660601...$*

Finally, we consider the anisotropic case and nonsymmetric projection (item 3).

THEOREM 4.13. *For $g = 3$ and nonsymmetric smoothing, the anisotropic stencil of the symbol (4.3) is stable, and the coarse-grid matrix is again an M-matrix for all $b > 0$.*

Proof. We need two ω values, $\omega_1 = \frac{2(1+b)}{3b}$ and $\omega_2 = \frac{2(1+b)}{3}$. These lead to

$$f_1(\pi,\pi) = 2\,, \qquad f_1(0,\pi) = \frac{2b}{1+b}\,, \qquad f_1(\pi,0) = \frac{2}{1+b}\,.$$

Therefore, the coarse-grid symbol is

$$f_1(x,y) = \alpha(1-\cos(x)) + \beta(1-\cos(y)), \quad \text{with} \quad \beta = \frac{b}{1+b}\,,\ \alpha = \frac{1}{1+b}\,. \qquad \square$$

5. Numerical examples. All numerical tests were obtained using MATLAB R2014a. We implemented the outlined method based on the developed theory for circulant and Toeplitz d-level matrices with generating symbols with second order zeros at the origin. The optimal ω was chosen automatically on each level by computing the values of the symbol at all the critical mirror points lying on the axes. Two steps each of Richardson iteration were used as pre- and postsmoother. The coarsest-grid was of size g^d in the circulant case and 1 in the case of Toeplitz matrices. For even cut sizes g we consider the circulant case only to allow for a meaningful geometric interpretation of the resulting aggregation method. We report the number of iterations required to achieve a reduction of the residual by a factor of 10^{-10}, the operator complexity, and the asymptotic convergence rate given by the residuals of the last two cycles.

5.1. 2-level isotropic examples. We consider stencils of the general form (4.1)

$$\frac{1}{4+4c}\begin{bmatrix} -c & -1 & -c \\ -1 & 4+4c & -1 \\ -c & -1 & -c \end{bmatrix}. \tag{5.1}$$

For $c = 0$ this yields the second-order accurate 5-point finite difference discretization of the Laplacian with the stencil

$$\frac{1}{4}\begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix}, \tag{5.2}$$

ETNA
Kent State University
http://etna.math.kent.edu

TABLE 5.1
Results for the circulant case for the 5-point Laplacian (5.2) for $g = 2$ and nonsymmetric smoothing.

# dof	# iter.	op. compl.	asymp. conv.
4×4	19	1.1000	0.3164
8×8	18	1.3000	0.3164
16×16	17	1.3750	0.3040
32×32	18	1.3938	0.3101
64×64	18	1.3938	0.3089
128×128	18	1.3996	0.3065
256×256	18	1.3999	0.3074

TABLE 5.2
Results for the circulant case for the 5-point Laplacian (5.2) for linear interpolation and full-weighting.

# dof	# iter.	op. compl.	asymp. conv.
4×4	18	1.2000	0.3164
8×8	18	1.5000	0.3164
16×16	17	1.5750	0.2955
32×32	18	1.5938	0.3096
64×64	18	1.5984	0.3069
128×128	18	1.5996	0.3070
256×256	18	1.5999	0.3074

while for $c = 1$ we obtain the second-order accurate 9-point finite element discretization of the Laplacian given by the stencil

$$\frac{1}{8}\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}. \tag{5.3}$$

We start with the case $g = 2$. To prevent unbounded growth of the operator complexity, we do not consider symmetric prolongation and restriction, but we rather consider a smoothed prolongation operator only. As mentioned above, we only consider the circulant case. The results for the 5-point Laplacian with stencil (5.2) can be found in Table 5.1. For the purpose of comparison we ran the same test with the same parameters but with the common linear interpolation and full-weighting restriction as described in, e.g., [2]. Comparing the results in Table 5.2, we like to emphasize that while the asymptotic convergence rate and the number of iterations needed to reduce the residual by a factor of 10^{-10} are comparable, the operator complexity of the smoothed aggregation-based method is lower, i.e., each multigrid cycle is cheaper. Results for the 9-point stencil (5.3) are found in Table 5.3. As a last example we considered the stencil given by (5.1) with $c = 1/\sqrt{2}$ that was shown to be stable in Theorem 4.7; the results are displayed in Table 5.4.

Next, we consider $g = 3$. In this case symmetric smoothing of prolongation and restriction does not lead to stencil growth, so we first start with this approach. We tested it for the 5- and 9-point Laplacian that are stable due to Theorem 4.11. The results for these stencils in the Toeplitz case can be found in Tables 5.5 and 5.6, and the results for the circulant case are comparable. If non-symmetric smoothing of the prolongation only is applied, the 5-point discretization of the Laplace operator leads to an indefinite stencil from level 2 onwards, so we did not consider it here. Note that it does not fulfill the requirements of Theorem 4.8, so the positive definiteness is not guaranteed anyway. The results for the 9-point stencil (5.3)

ETNA
Kent State University
http://etna.math.kent.edu

TABLE 5.3
Results for the circulant case for the 9-point Laplacian (5.3) for $g = 2$ and nonsymmetric smoothing.

# dof	# iter.	op. compl.	asymp. conv.
4×4	12	1.1111	0.1526
8×8	13	1.2778	0.1944
16×16	12	1.3194	0.1922
32×32	12	1.3299	0.1841
64×64	12	1.3325	0.1862
128×128	12	1.3331	0.1849
256×256	12	1.3333	0.1854

TABLE 5.4
Results for the circulant case for the stable stencil (5.1) with $c = 1/\sqrt{2}$ for $g = 2$ and nonsymmetric smoothing.

# dof	# iter.	op. compl.	asymp. conv.
4×4	13	1.1111	0.1746
8×8	12	1.2778	0.1952
16×16	13	1.3194	0.1982
32×32	13	1.3299	0.1875
64×64	13	1.3325	0.1881
128×128	13	1.3331	0.1850
256×256	13	1.3333	0.1860

are given in Table 5.7; those for the stencil (5.1) with $c = 1/\sqrt{2}$, which is stable due to Theorem 4.8, can be found in Table 5.8. We also considered the 5-point Laplacian (5.2) in the case $g = 4$. In this case the stencil is stable; cf. Theorem 4.9. As in the case $g = 2$, we only present results for the circulant case, which can be found in Table 5.9. Finally, results for the stencil (5.1) with $c = 0.22968147..$ are presented in Table 5.10 for the Toeplitz case with $g = 5$. The stencil is stable due to Theorem 4.10, and the results for the circulant case are similar. In all cases we see a nice convergence behavior that is independent of the number of levels. As expected, the convergence rate deteriorates when more aggressive coarsening is chosen. This could be overcome by adding more smoothing steps or by using more efficient smoothers.

5.2. 2-level anisotropic examples. We consider matrices with the stencil

$$\begin{bmatrix} -\frac{1}{12} & -\frac{6b-2a}{12a+12b} & -\frac{1}{12} \\ -\frac{6a-2b}{12a+12b} & 1 & -\frac{6a-2b}{12a+12b} \\ -\frac{1}{12} & -\frac{6b-2a}{12a+12b} & -\frac{1}{12} \end{bmatrix}, \tag{5.4}$$

yielding the symbol

$$f(x) = 1 - \frac{12a - 4b}{12a + 12b} \cos(x_1) - \frac{12b - 4a}{12a + 12b} \cos(x_2) - \frac{1}{3} \cos(x_1) \cos(x_2).$$

This corresponds to a discretization of an anisotropic PDE. First we consider an example with a slight anisotropy where we choose $a = 1$ and $b = 1.1$. To reduce the growth of the operator complexity we again choose to smooth the prolongation only. The results for the Toeplitz case are presented in Table 5.11; those for the circulant case are similar. If the anisotropy is increased, the convergence rate deteriorates, as expected. The results for $a = 1$ and $b = 2$ can be found in Table 5.12. The consideration of even higher anisotropies is not meaningful as

ETNA
Kent State University
http://etna.math.kent.edu

TABLE 5.5
Results for the Toeplitz case for the 5-point Laplace (5.2) for $g = 3$ and symmetric smoothing.

# dof	# iter.	op. compl.	asymp. conv.
9×9	22	1.1328	0.3679
27×27	32	1.1906	0.5485
81×81	33	1.2129	0.5721
243×243	33	1.2209	0.5729

TABLE 5.6
Results for the Toeplitz case for the 9-point Laplace (5.3) for $g = 3$ and symmetric smoothing.

# dof	# iter.	op. compl.	asymp. conv.
9×9	14	1.0800	0.2308
27×27	20	1.1082	0.3970
81×81	21	1.1191	0.4203
243×243	21	1.1230	0.4217

TABLE 5.7
Results for the Toeplitz case for the 9-point Laplace (5.3) for $g = 3$ and nonsymmetric smoothing.

# dof	# iter.	op. compl.	asymp. conv.
9×9	18	1.0784	0.3083
27×27	23	1.1080	0.4073
81×81	23	1.1191	0.4252
243×243	24	1.1230	0.4374

TABLE 5.8
Results for the Toeplitz case for the stable stencil (5.1) with $c = 1/\sqrt{2}$ for $g = 3$ and nonsymmetric smoothing.

# dof	# iter.	op. compl.	asymp. conv.
9×9	19	1.0784	0.3245
27×27	24	1.1080	0.4306
81×81	25	1.1191	0.4457
243×243	25	1.1230	0.4464

TABLE 5.9
Results for the circulant case for the stable 5-point stencil (5.2) for $g = 4$ and nonsymmetric smoothing.

# dof	# iter.	op. compl.	asymp. conv.
16×16	60	1.0625	0.7377
64×64	58	1.0664	0.7303
256×256	59	1.0667	0.7308

TABLE 5.10
Results for the Toeplitz case for the optimal stencil (5.1) with $c = 0.22968147..$ for $g = 5$ and nonsymmetric smoothing.

# dof	# iter.	op. compl.	asymp. conv.
25×25	65	1.0317	0.7229
125×125	81	1.0395	0.7841
625×625	81	1.0412	0.7845

ETNA
Kent State University
http://etna.math.kent.edu

TABLE 5.11
Results for the Toeplitz case for the anisotropic stencil (5.4) with $a = 1$ and $b = 1.1$ for $g = 3$.

# dof	# iter.	op. compl.	asymp. conv.
9×9	17	1.0784	0.2717
27×27	27	1.1080	0.4604
81×81	28	1.1191	0.4863
243×243	28	1.1230	0.4869

TABLE 5.12
Results for the Toeplitz case for the anisotropic stencil (5.4) with $a = 1$ and $b = 2$ for $g = 3$.

# dof	# iter.	op. compl.	asymp. conv.
9×9	23	1.0784	0.3797
27×27	38	1.1080	0.5903
81×81	40	1.1191	0.6118
243×243	40	1.1230	0.6126

other coarsening strategies like semi-coarsening [26] or the use of stretched aggregates [20] is advisable then.

5.3. 3D example. The stencil of the Laplacian in 3 dimensions using linear finite elements is given by

$$\frac{4}{3}\left[\begin{array}{ccc|ccc|ccc} -4 & -8 & -4 & -8 & 0 & -8 & -4 & -8 & -4 \\ -8 & 0 & -8 & 0 & 128 & 0 & -8 & 0 & -8 \\ -4 & -8 & -4 & -8 & 0 & -8 & -4 & -8 & -4 \end{array}\right].$$

The results for the Toeplitz case with $g = 3$ are displayed in Table 5.13. The results for the circulant case are very similar, so we omit them. The results show that our approach works as expected for higher levels/dimensions as well.

5.4. An example with a dense system matrix. The approach is not limited to sparse circulant and Toeplitz matrix, but it is rather generally applicable to matrices where the generating symbol has an isolated zero of even order. To illustrate this, we present result for the 1-level Toeplitz matrix with generating symbol

$$f(x) = x^2.$$

As the Fourier coefficients of f are given by

$$a_k = \begin{cases} \pi^2/3 & \text{for } k = 0, \\ (-1)^k \frac{2}{k^2} & \text{otherwise,} \end{cases}$$

this results in a sequence of dense matrices. The described smoothed aggregation multigrid method works for this example as expected. Results for $g = 2$ and nonsymmetric smoothing of the prolongation operator are found in Table 5.14. We like to note that all coarse-level matrices are non-Toeplitz matrices that are not just a low-rank perturbation of a Toeplitz matrix due to the dense prolongation operator. While our aggregation-based method works for this problem, a method that is tailored for this kind of problems, like the one presented in [9], is better suited than our general approach.

ETNA
Kent State University
http://etna.math.kent.edu

TABLE 5.13
Results for the Toeplitz case for the finite element discretization of the 3D Laplacian using cubic finite elements for $g = 3$ and symmetric smoothing.

# dof	# iter.	op. compl.	asymp. conv.
$9 \times 9 \times 9$	14	1.0292	0.2212
$27 \times 27 \times 27$	19	1.0421	0.3932
$81 \times 81 \times 81$	21	1.0469	0.4197

TABLE 5.14
Results for the Toeplitz case for matrices with generating symbol $f(x) = x^2$ with $g = 2$.

# dof	# iter.	op. compl.	asymp. conv.
4	16	1.2857	0.2532
8	22	1.3226	0.3758
16	24	1.3307	0.4148
32	25	1.3327	0.4318
64	25	1.3332	0.4375
128	25	1.3333	0.4399
256	25	1.3333	0.4411

5.5. Optimality of ω. To illustrate the optimality of ω resulting from our analysis, we varied the value of ω. We chose the 9-point stencil (5.3) for the Toeplitz case with $g = 3$ as this is a stable stencil according to our theoretical results. We changed the optimal ω obtained with the developed theory by multiplying it by a factor $\alpha \in [0.8, 1.2]$ on each level. In each case we were solving a system of size $3^5 \times 3^5$ using the same right hand side and a zero initial guess. The resulting asymptotic convergence rates are provided in Table 5.15. While the asymptotic convergence rate does not vary much in a neighborhood of the optimal ω, the optimal ω yields the best convergence rate. This shows that the theory is valid, but the methods seem to be relatively robust regarding the choice of the smoothing parameter.

5.6. The non-constant coefficient case. The obtained results can be used to define SA methods for the non-constant coefficient case straightforwardly. For this purpose we use Jacobi iteration as smoother, but we introduce a diagonal matrix Ω to damp the relaxation in a pointwise manner. We deal with model problem 3 in [26, p. 131], i.e.,

$$\begin{aligned} -\epsilon u_{xx} - u_{yy} &= f \qquad (x, y) \in \Omega = (0, 1)^2, \\ u &= g \qquad (x, y) \in \partial\Omega, \end{aligned}$$

with varying ϵ and discretize the problem using the stencil

$$\frac{1}{h^2} \begin{bmatrix} & -1 & \\ -\epsilon & 2(1+\epsilon) & -\epsilon \\ & -1 & \end{bmatrix}.$$

We choose ϵ as

$$\epsilon(x, y) = \frac{1}{2}(2 + \sin(2\pi x)\sin(2\pi y))$$

and scale the matrix symmetrically such that it has ones on the diagonal. We build regular 3×3 aggregates, i.e., dealing with the case $g = 3$. For the traditional SA approach, we smooth the prolongation and restriction operator with ω-Jacobi iteration; in accordance with [28] we

ETNA
Kent State University
http://etna.math.kent.edu

TABLE 5.15
Asymptotic convergence rate for slight perturbation of ω in Toeplitz case with $g = 3$, system size $3^5 \times 3^5$ and symmetric smoothing of the grid transfer operators.

α	0.8	0.9	1.0	1.1	1.2
asymp. conv.	0.4897	0.4260	0.4217	0.4270	0.4777

TABLE 5.16
Results for the Toeplitz case for the finite element discretization of the 3D Laplacian using cubic finite elements for $g = 3$ and symmetric smoothing.

# dof	9×9	27×27	81×82	243×243
# iter. ($\omega = 2/3$)	27	39	52	68
asymp. conv. ($\omega = 2/3$)	0.4053	0.5415	0.6315	0.6996
# iter. (opt. ω)	21	31	39	45
asymp. conv. (opt. ω)	0.3160	0.5156	0.5717	0.6142

choose $\omega = 2/3$. For our adaptive approach using the local model, we build a local stencil for each grid point and calculate the locally optimal ω's. As the problem is locally anisotropic, we obtained two values of ω that were used to set up two diagonal matrices Ω_1 and Ω_2 for the smoothing of the prolongation and the restriction operator, respectively, by multiplying them by

$$S_i = I - \Omega_i A, \quad i = 1, 2.$$

Nonsymmetric smoothing is used to prevent the operator complexity from growing. While the operator complexity is the same for both approaches, the achieved convergence rates and iteration counts vary. They can be found in Table 5.16 and in Figure 5.1. The choice of smoothing parameters that is achieved is illustrated in Figure 5.2, where the two ω values that are chosen in the 27×27 case are plotted. Our modification clearly outperforms the traditional approach.

6. Conclusion. Aggregation-based multigrid methods for circulant and Toeplitz matrices can be analyzed using the classical theory. The non-optimality of non-SA-based multigrid methods can be explained easily by the lack of fulfillment of (2.12) by the prolongation and restriction operator in that case. Guided by this observation, sufficient conditions for an improvement of the grid transfer operators by an application of Richardson iteration can be derived, including the optimal choice of the parameter. The results carry over from aggregates of size 2^d to larger aggregates. Numerical experiments show that the theory is valid and that it can be used as a local model to choose the appropriate damping in SA even for the non-constant coefficient case. As a result, the application of more than one smoother is recommended in connection with nonsymmetric coarsening in order to match the necessary order of the zeros in the projection without increasing the sparsity of the coarse matrices.

Acknowledgment. We would like to thank the referees for their valuable comments and suggestions. The work of the first author is partly supported by the German Research Foundation (DFG) through the Priority Programme 1648 "Software for Exascale Computing" (SPPEXA), and the work of the second author is partly supported by PRIN 2012 N. 2012MTE38N.

ETNA
Kent State University
http://etna.math.kent.edu

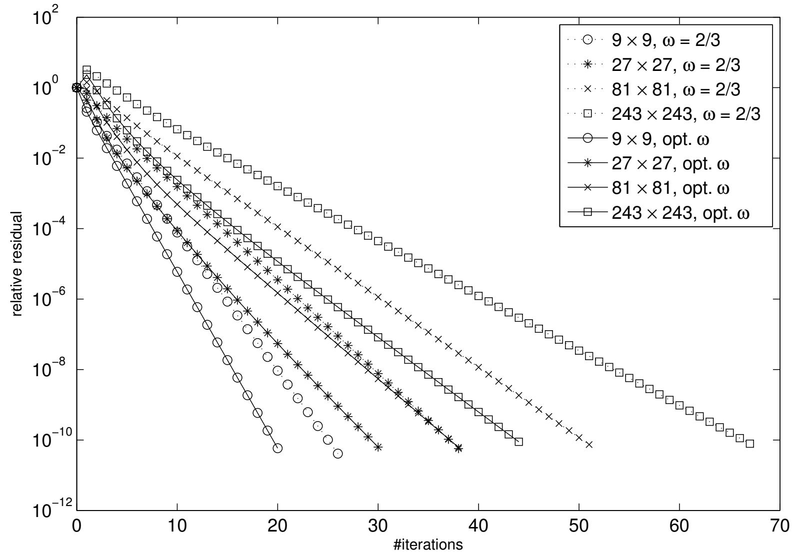


FIG. 5.1. *Convergence history of the standard SA method with $\omega = 2/3$ and the proposed version with adaptively chosen ω based on the local stencil, aggregate size in both cases was 3.*

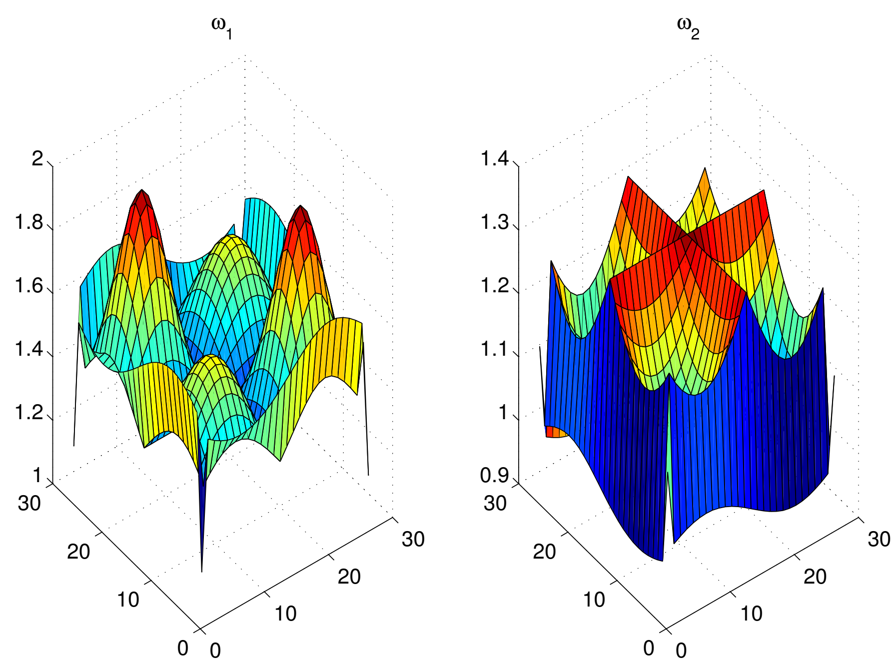


FIG. 5.2. *ω's chosen in the 27×27 case.*

REFERENCES

[1] A. ARICÒ AND M. DONATELLI, *A V-cycle multigrid for multilevel matrix algebras: proof of optimality*, Numer. Math., 105 (2007), pp. 511–547.
[2] A. ARICÒ, M. DONATELLI, AND S. SERRA-CAPIZZANO, *V-cycle optimal convergence for certain (multilevel) structured linear systems*, SIAM J. Matrix Anal. Appl., 26 (2004), pp. 186–214.
[3] M. BOLTEN, M. DONATELLI, T. HUCKLE, AND C. KRAVVARITIS, *Generalized grid transfer operators for multigrid methods applied on Toeplitz matrices*, BIT, in press, doi 10.1007/s10543-014-0512-2 (2014).
[4] D. BRAESS, *Towards algebraic multigrid for elliptic problems of second order*, Computing, 55 (1995), pp. 379–393.
[5] M. BREZINA, R. FALGOUT, S. MACLACHLAN, T. MANTEUFFEL, S. MCCORMICK, AND J. RUGE, *Adaptive smoothed aggregation* $(\alpha\mathrm{SA})$ *multigrid*, SIAM Rev., 47 (2005), pp. 317–346.
[6] M. BREZINA, P. VANĚK, AND P. S. VASSILEVSKI, *An improved convergence analysis of smoothed aggregation algebraic multigrid*, Numer. Linear Algebra Appl., 19 (2012), pp. 441–469.
[7] F. CHATELIN AND W. L. MIRANKER, *Acceleration by aggregation of successive approximation methods*, Linear Algebra Appl., 43 (1982), pp. 17–47.
[8] M. DONATELLI, *An algebraic generalization of local Fourier analysis for grid transfer operators in multigrid based on Toeplitz matrices*, Numer. Linear Algebra Appl., 17 (2010), pp. 179–197.
[9] M. DONATELLI, S. SERRA-CAPIZZANO, AND D. SESANA, *Multigrid methods for Toeplitz linear systems with different size reduction*, BIT, 52 (2012), pp. 305–327.
[10] G. FIORENTINO AND S. SERRA, *Multigrid methods for Toeplitz matrices*, Calcolo, 28 (1991), pp. 283–305.
[11] ———, *Multigrid methods for symmetric positive definite block Toeplitz matrices with nonnegative generating functions*, SIAM J. Sci. Comput., 17 (1996), pp. 1068–1081.
[12] R. FISCHER AND T. HUCKLE, *Multigrid methods for anisotropic BTTB systems*, Linear Algebra Appl., 417 (2006), pp. 314–334.
[13] G. H. GOLUB AND C. F. VAN LOAN, *Matrix computations*, 3rd ed., Johns Hopkins University Press, Baltimore, 1996.
[14] T. HUCKLE AND J. STAUDACHER, *Multigrid preconditioning and Toeplitz matrices*, Electron. Trans. Numer. Anal., 13 (2002), pp. 81–105.
http://etna.math.kent.edu/vol.13.2002/pp81-105.dir
[15] ———, *Multigrid methods for block Toeplitz matrices with small size blocks*, BIT, 46 (2006), pp. 61–83.
[16] J. MANDEL, S. MCCORMICK, AND J. RUGE, *An algebraic theory for multigrid methods for variational problems*, SIAM J. Numer. Anal., 25 (1988), pp. 91–110.
[17] J. MANDEL AND B. SEKERKA, *A local convergence proof for the iterative aggregation method*, Linear Algebra Appl., 51 (1983), pp. 163–172.
[18] S. F. MCCORMICK, *Multigrid methods for variational problems: general theory for the V-cycle*, SIAM J. Numer. Anal., 22 (1985), pp. 634–643.
[19] S. F. MCCORMICK AND J. W. RUGE, *Multigrid methods for variational problems*, SIAM J. Numer. Anal., 19 (1982), pp. 924–929.
[20] A. C. MURESAN AND Y. NOTAY, *Analysis of aggregation-based multigrid*, SIAM J. Sci. Comput., 30 (2008), pp. 1082–1103.
[21] A. NAPOV AND Y. NOTAY, *Algebraic analysis of aggregation-based multigrid*, Numer. Linear Algebra Appl., 18 (2011), pp. 539–564.
[22] J. W. RUGE AND K. STÜBEN, *Algebraic multigrid*, in Multigrid methods, S. F. McCormick, ed., Frontiers Appl. Math. Vol. 3, SIAM, Philadelphia, 1987, pp. 73–130.
[23] S. SERRA-CAPIZZANO, *Convergence analysis of two-grid methods for elliptic Toeplitz and PDEs matrix-sequences*, Numer. Math., 92 (2002), pp. 433–465.
[24] S. SERRA-CAPIZZANO AND C. TABLINO-POSSIO, *Preliminary remarks on multigrid methods for circulant matrices*, in Numerical Analysis and its Applications (Rousse, 2000), L. Vulkov, J. Waśniewski, and P. Yalamov, eds., Lecture Notes in Comput. Sci., Vol. 1988, Springer, Berlin, 2001, pp. 152–159.
[25] ———, *Multigrid methods for multilevel circulant matrices*, SIAM J. Sci. Comput., 26 (2004), pp. 55–85.
[26] U. TROTTENBERG, C. W. OOSTERLEE, AND A. SCHÜLLER, *Multigrid*, Academic Press, San Diego, 2001.
[27] E. TYRTYSHNIKOV, *A unifying approach to some old and new theorems on preconditioning and clustering*, Linear Algebra Appl., 232 (1996), pp. 1–43.
[28] P. VANĚK, J. MANDEL, AND M. BREZINA, *Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems*, Computing, 56 (1996), pp. 179–196.
[29] R. WIENANDS AND W. JOPPICH, *Practical Fourier Analysis for Multigrid Methods*, Chapman & Hall, Boca Raton, 2005.