

Bergische Universität Wuppertal

Fachbereich Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational  
Mathematics (IMACM)

Preprint BUW-IMACM 13/10

Matthias Bolten, Marco Donatelli, Thomas Huckle

**Analysis of smoothed aggregation multigrid methods  
based on Toeplitz matrices**

July 2013

<http://www-ai.math.uni-wuppertal.de>

## ANALYSIS OF SMOOTHED AGGREGATION MULTIGRID METHODS BASED ON TOEPLITZ MATRICES

MATTHIAS BOLTEN\*, MARCO DONATELLI†, AND THOMAS HUCKLE‡

**Abstract.** Aim of the 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 results in optimal smoothing parameters that have to be chosen for the smoothing of 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 for 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

$$Ax = b,$$

where  $x, b \in \mathbb{C}^N$  and  $A$  is an ill-conditioned symmetric positive definite  $N \times N$  matrix. Mainly, we analyse 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 development of multigrid methods for  $\tau$ -matrices and Toeplitz matrices goes back to [1], the two level case was considered in [2]. Using the same ideas methods for circulant matrices were developed later in [3]. While these works provide the basis to develop and analyze multigrid methods for Toeplitz matrices and matrices from different matrix algebras, including the  $\tau$ - and circulant algebra, they did not provide a prove of optimality of the multigrid cycle, in the sense that the convergence rate is bounded by a constant  $c < 1$  independent on the number of levels used in the multigrid cycle. This prove was added later in [4, 5]. In [6] a 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. [7, 8, 9, 10].

Aggregation based multigrid goes back at least to [11], where the so-called aggregation/disaggregation methods [12, 13] 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 that the unknowns are grouped together into *aggregates*, these aggregates form one variable on the coarse level, each. The pure aggregation can be improved by incorporating smoothing [14] in the prolongation and/or the restriction leading to faster

---

\*Bergische Universität Wuppertal, Fachbereich Mathematik und Naturwissenschaften, Gaußstraße 20, 42097 Wuppertal, Germany (bolten@math.uni-wuppertal.de).

†Università dell’Insubria, Sede di Como, 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 [15, 16, 17, 18].

In this paper, firstly we extend the two-grid optimality results in [6] to multidimensional problems. Using these new convergence results we can provide an analysis of aggregation operators for multilevel Toeplitz matrices. According to the literature [16], we show that the pure aggregation provides only two-grid optimality but it is not enough for V-cycle. Therefore, we study a simple smoothing aggregation strategy based on a damping factor chosen as the value that provides the smallest convergence rate. In contrast to previous analysis in literature [16, 17, 18] our analysis uses a symbolic approach to discuss the convergence and to choose the optimum 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, but, at the same time, it 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 are 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 nonconstant coefficients partial differential equations. A special attention is devoted in Section 4.3 to the discretization of the 2D Laplacian by nine points stencils. A wide range of numerical experiments is presented in Section 5 and some conclusive 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

$$(2.1) \quad T_n = \begin{pmatrix} t_0 & t_{-1} & \cdots & t_{-n+1} \\ t_1 & t_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_{-1} \\ t_{n-1} & \cdots & t_1 & t_0 \end{pmatrix}.$$

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

$$(2.2) \quad f(x) = \sum_{j=-\infty}^{\infty} t_j e^{i2\pi jx}, \quad t_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-i2\pi jx} dx.$$

with entries  $t_j$  on the diagonals given as the 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 matrices is the fact that the distribution of the eigenvalues of the Toeplitz matrix is given by the generating symbol in the limit case  $n \rightarrow \infty$ , cf. [19].

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

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

$C_n$  is diagonalized by the Fourier matrix  $F_n$ , where

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

i.e.

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

for  $\lambda^{(n)} = (\lambda_0^{(n)}, \dots, \lambda_{n-1}^{(n)})$  given by  $\lambda_j^{(n)} = f(2\pi j/n)$ ,  $j = 0, \dots, n-1$ . 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  $\{C_n(f)\}_{n=1}^{\infty}$  of matrices  $C_n(f)$  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  $2\pi$ -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. [20].

An interesting special type of block matrices that we will deal with is the case where the blocks itself are Toeplitz/circulant, again. 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  $2\pi$ -periodic generating symbol  $f$ . This is related to the two-dimensional case  $d = 2$ . In the general  $d$ -level case the generating symbols are  $f : \mathbb{R}^d \rightarrow \mathbb{C}$  a  $2\pi$  periodic functions having Fourier coefficients

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

where  $\langle \cdot | \cdot \rangle$  denotes the usual scalar product between vectors. From the coefficients  $t_j$  one can build the sequence  $\{C_n(f)\}$ ,  $n = (n_1, \dots, n_d) \in \mathbb{N}^d$ , 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  $C_n(f)$  can be written again in the form (2.3) where now  $\lambda^{(n)} = \lambda^{(n_1)} \times \cdots \times \lambda^{(n_d)}$ .

**2.2. Multigrid methods.** A multigrid method is a method to solve a linear system of equations. When traditional stationary iterative methods like Jacobi are used to solve a linear system, they perform poorly when the system gets more ill-conditioned, e.g. when the mesh width is decreased in the discretization of a PDE. The reason for this 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 are known as “smoothers”.

To construct a multigrid method various components have to be chosen. To construct a multigrid method the coefficient matrix of the linear system (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  $R_i : \mathbb{C}^{N_i} \rightarrow \mathbb{C}^{N_{i+1}}$ ,  $i = 0, \dots, l_{\max} - 1$  and  $P_i : \mathbb{C}^{N_{i+1}} \rightarrow \mathbb{C}^{N_i}$ ,  $i = 0, \dots, l_{\max} - 1$  are needed, furthermore a hierarchy of operators  $A_i \in \mathbb{C}^{N_i \times N_i}$ ,  $i = 1, \dots, l_{\max}$  has to be defined. On each level appropriate smoothers  $\mathcal{S}_i$  and  $\tilde{\mathcal{S}}_i$  and the numbers of smoothing steps  $\nu_1$  and  $\nu_2$  have to be chosen, we limit ourselves to stationary iterative methods although other smoother like Krylov-subspace methods can be used, as well. After  $\nu_1$  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

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

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 V-cycle. The process of correcting the current iterate using the coarse level is known as *coarse grid correction*, it has the iteration matrix

$$(2.4) \quad M_i = I - P_i A_{i+1}^{-1} R_i A_i.$$

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

---

**Algorithm 1** Multigrid cycle  $x_{n_i} = \mathcal{M}\mathcal{G}_i(x_{n_i}, b_{n_i})$

---

```

 $x_{n_i} \leftarrow \mathcal{S}_i^{\nu_1}(x_{n_i}, b_{n_i})$ 
 $r_{n_i} \leftarrow b_{n_i} - A_i x_{n_i}$ 
 $r_{n_{i+1}} \leftarrow R_i r_{n_i}$ 
 $e_{n_{i+1}} \leftarrow 0$ 
if  $i + 1 = l_{\max}$  then
   $e_{n_{l_{\max}}} \leftarrow A_{l_{\max}}^{-1} r_{n_{l_{\max}}}$ 
else
   $e_{n_{i+1}} \leftarrow \mathcal{M}\mathcal{G}_{i+1}(e_{n_{i+1}}, r_{n_{i+1}})$ 
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 that are coupled together by an appropriately chosen

norm  $\|\cdot\|_*$ , where in the classical algebraic multigrid theory the  $AD^{-1}A$ -norm with  $D = \text{diag}(A)$  is chosen, cf. [10], and in the circulant case the  $A^2$ -norm turns out to be helpful, cf. [4].

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*

$$(2.5) \quad \|S_i v_{n_i}\|_{A_i}^2 \leq \|v_{n_i}\|_{A_i}^2 - \alpha \|S_i v_{n_i}\|_*^2.$$

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

$$(2.6) \quad \|\tilde{S}_i v_{n_i}\|_{A_i}^2 \leq \|v_{n_i}\|_{A_i}^2 - \beta \|v_{n_i}\|_*^2.$$

The following theorem is useful to prove two-grid method convergence since the forthcoming condition (2.8) is usually weaker and easier to prove than the *approximation property*

$$(2.7) \quad \|M_i v_{n_i}\|_{A_i}^2 \leq \gamma \|v_{n_i}\|_*^2.$$

THEOREM 2.2 ([10]). *Let  $A_i \in \mathbb{C}^{N_i \times N_i}$  be a positive definite matrix and let  $\tilde{S}_i$  be the postsmoother with iteration matrix  $\tilde{S}_i$  fulfilling the postsmoothing property (2.6) 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*

$$(2.8) \quad \min_{y \in \mathbb{C}^{N_{i+1}}} \|x - P_i y\|_{D_i}^2 \leq \gamma \|x\|_{A_i}^2, \quad \forall x \in \mathbb{C}^{N_i},$$

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

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

**2.3. Multigrid methods for circulant and Toeplitz matrices.** In the following, we will 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 such results. After that, we will 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<sup>1</sup>. In general, to design a multigrid method, the smoother, a coarse level with fewer degrees of freedom, the prolongation and restriction have to be chosen appropriately. Here, the common choice for both, pre- and postsmoothing is relaxed Richardson, i.e.  $\mathcal{S}_i$  is chosen as

$$(2.9) \quad \mathcal{S}_i(x_{n_i}, b_{n_i}) = \underbrace{(I - \omega_i A_i)}_{=S_i} x_{n_i} + \omega_i b_{n_i},$$

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

<sup>1</sup> $A_i$  could be singular for circulant matrices if  $f$  vanishes at a grid point. In such case a rank-one correction like in [4] could be considered, but it is not necessary in practice, see [insert citation to the other paper or a Matthias's paper on singular circulant matrices].

multiple of the identity. Using appropriate relaxation parameters  $\omega_i$  and  $\tilde{\omega}_i$  this smoother fulfills the presmoothing property (2.5) respectively the postsmoothing property (2.6) as stated by the following theorem that can be found as Proposition 3 in [5].

**THEOREM 2.3 ([5]).** *Let  $A_i = \mathcal{C}_{n_i}(f_i)$ , where  $f_i : \mathbb{R}^d \rightarrow \mathbb{R}$  and let  $\mathcal{S}_i$  as defined in (2.9) with  $\omega_i \in \mathbb{R}$ , and  $\tilde{\mathcal{S}}_i$  defined as  $\mathcal{S}_i$  but with parameter  $\tilde{\omega}_i \in \mathbb{R}$ . Then if  $\omega_i, \tilde{\omega}_i \in (0, 2/\|f_i\|_\infty)$ , the smoothing properties (2.5) and (2.6) 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  $j = 1, \dots, d$ . 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 [6] or corresponding to semi-coarsening [21, 22] 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 cut matrix  $K_{n_i} \in \mathbb{C}^{n_{i+1} \times n_i}$  that in the case of a 1-level circulant matrix of even size on the fine level is given by

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

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

$$(2.10) \quad K_{n_i} F_{n_i} = \frac{1}{\sqrt{2}} [1, 1] \otimes F_{n_{i+1}} = \frac{1}{\sqrt{2}} F_{n_{i+1}} ([1, 1] \otimes I_{n_{i+1}})$$

in the 1-level case. In the  $d$ -level case the cut matrix is defined by Kronecker product

$$(2.11) \quad K_{n_i} = K_{(n_i)_1} \otimes \dots \otimes K_{(n_i)_d}.$$

Combining (2.10) with (2.11) and due to the properties of the Kronecker product we have

$$(2.12) \quad \begin{aligned} K_{n_i} F_{n_i} &= K_{(n_i)_1} F_{(n_i)_1} \otimes \dots \otimes K_{(n_i)_d} F_{(n_i)_d} \\ &= \frac{1}{\sqrt{2^d}} (F_{(n_{i+1})_1} ([1, 1] \otimes I_{(n_{i+1})_1})) \otimes \dots \otimes (F_{(n_{i+1})_d} ([1, 1] \otimes I_{(n_{i+1})_d})) \\ &= \frac{1}{\sqrt{2^d}} F_{n_{i+1}} \Theta_{n_{i+1}}, \end{aligned}$$

where  $\Theta_{n_{i+1}} = ([1, 1] \otimes I_{(n_{i+1})_1}) \otimes \dots \otimes ([1, 1] \otimes I_{(n_{i+1})_d})$ . With the help of the cut 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” of  $x$  as

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







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

$$a_{1,g} : [-\pi, \pi) \rightarrow \mathbb{C}$$

$$x \mapsto a_{1,g}(x) = \sum_{k=0}^{g-1} e^{-ikx}$$

is

$$(3.8) \quad P_i = C_{n_i}(p_i)K_{n_i,g}^T.$$

The effect of the cut matrix applied to the Fourier matrix is similar to (2.10) 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  $e_g^T = [1, \dots, 1] \in \mathbb{N}^g$  and the set of mirror points consists of the  $g - 1$  points in  $\mathcal{M}_g(x) = \Omega_g(x) \setminus \{x\}$  where

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

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

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

For further details see [6], 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 (2.14) stated on the sets  $\mathcal{M}_g$  and  $\Omega_g$ , respectively. In more detail, the two-grid optimality requires

$$(3.9) \quad \limsup_{x \rightarrow x^0} \frac{|p_i(y)|^2}{|f_i(x)|} \leq +\infty, \quad y \in \mathcal{M}_g(x), \quad i = 0, \dots, l_{\max} - 1,$$

$$(3.10) \quad 0 < \sum_{y \in \Omega_g(x)} |p_i|^2(y), \quad i = 0, \dots, l_{\max} - 1,$$

for all  $x \in [-\pi, \pi)$ , see Theorem 5.1 in [6]. The V-cycle optimality for a coarsening factor  $g > 2$  is an open problem, but a natural conjecture is that in (2.13), similarly to (3.1), it is enough to replace  $\mathcal{M}$  with  $\mathcal{M}_g$ , namely

$$(3.11) \quad \limsup_{x \rightarrow x^0} \left| \frac{p_i(y)}{f_i(x)} \right| < +\infty, \quad y \in \mathcal{M}_g(x), \quad i = 0, \dots, l_{\max} - 1.$$

As the pure aggregation  $p_i = a_{1,g}$  fulfills only (3.9) but not (3.11), the prolongation has to be improved for all mirror points, possibly resulting in more than one smoothing parameter  $\omega$  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  $p_i(x) = 1 + e^{-i(x+x_0)}$ .

**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,  $d \in \mathbb{N}$  (usually associated to  $d$ -dimensional problems), we have to extend the theoretical results in [6] to  $d > 1$ . For that purpose let  $A = \mathcal{C}_n(f)$ , where  $f : \mathbb{R}^d \rightarrow \mathbb{C}$  is a nonnegative function  $2\pi$ -periodic in each variable,  $n \in \mathbb{N}^d$ ,  $g \in \mathbb{N}^d$  is the size of the aggregates and assume that  $n = g^{l_{\max}+1}$ , i.e.,  $n_j = g_j^{l_{\max}+1}$ ,  $j = 1, \dots, d$ . 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-indices operations in the paper are intended component-wise), the prolongation as in (3.8) where  $K_{n_i,g} = K_{(n_i)_1,g_1} \otimes \dots \otimes K_{(n_i)_d,g_d}$ , 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 cut matrix  $K_{n_i,g}$  is

$$\Omega_g(x) = \left\{ y \mid y_j \in \left\{ x_j + \frac{2\pi k}{g_j} \pmod{2\pi} \right\}, k = 0, \dots, g_j - 1, j = 1, \dots, d \right\}.$$

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 by the following lemma.

LEMMA 3.2. *Let  $A_i = \mathcal{C}_{n_i}(f_i)$ ,  $P_i$  defined in (3.8), and  $n_{i+1} = g \cdot n_i \in \mathbb{N}^d$ , 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*

$$(3.12) \quad f_{i+1}(x) = \frac{1}{G} \sum_{y \in \Omega_g(x/g)} |p_i|^2 f_i(y), \quad x \in [-\pi, \pi]^d.$$

*Proof.* The proof is a generalization of the proof of Proposition 5.1 in [3]. First we note that in analogy to (2.12), we have

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

so

$$(3.13) \quad K_{n_i} F_{n_i} = \frac{1}{\sqrt{G}} F_{n_{i+1}} \Theta_{n_i,g},$$

where  $\Theta_{n_i,g} = (e_{g_1}^T \otimes I_{n_{i,1}}) \otimes \dots \otimes (e_{g_d}^T \otimes I_{n_{i,d}})$ . So, for  $A_{i+1} = P_i^H A_i P_i$  we have

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

Here,

$$D_{n_i}(f) = \text{diag}_{0 \leq j \leq n_i - e_d} (f((x_i)_j)),$$

where  $(x_i)_j \equiv 2\pi j/n_i = (2\pi j_1/(n_i)_1, \dots, 2\pi j_d/(n_i)_d)^T$  and  $0 \leq j \leq n_i - e_d$ . All operations and inequalities between multi-indices are intended component-wise. For a given multi-index  $k = (k_1, \dots, k_d)$ ,  $0 \leq k_j \leq (n_{i+1})_j$  we have

$$(\Theta_{n_i,g} x)_k = \sum_{l=0}^{g-e_d} x_{k+l},$$

so we obtain

$$\Theta_{n_i, g} D_{n_i}(|p_i|^2 f_i) \Theta_{n_i, g}^T = \sum_{l=0}^{g-e_d} D_{n_i, g, l}(|p_i|^2 f_i),$$

where

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

For an example of the multi-index notation in the case  $d = g = 2$  we refer to the proof of Proposition 5.1 in [3]. As result we obtain

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

and with

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

we get  $P_i^H A_i P_i = C_{n_{i+1}}(f_{i+1})$ , with  $f_{i+1}$  defined in (3.12).

□

REMARK 3.3. *If the two conditions (3.9) and (3.10) are satisfied with  $x \in [-\pi, \pi]^d$ , we obtain as consequence of Lemma 3.2 that if  $x^0$  is a zero of  $f_i$  then  $g \cdot x^0 \pmod{2\pi}$  is a zero of  $f_{i+1}$  with the same order.*

The two-grid optimality can be obtained similarly to the 1-level case. The following result shows that two-grid conditions (3.9) and (3.10) are sufficient in order to satisfy the condition (2.8).

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  $p_i$  trigonometric polynomial, satisfying condition (3.9), for any zero of  $f_i$ , and satisfying globally condition (3.10). Then, there exists a positive value  $\gamma$  independent of  $n_i$  such that inequality (2.8) is satisfied.*

*Proof.* The proof is a combination of Theorem 5.1 in [6] and Lemma 6.3 in [3], but we report it here for completeness. First, we recall that the main diagonal of  $A_i$  is given by  $D_i = t_i I_{N_i}$  with  $t_i = \|f_i\|_1 > 0$ , so that  $\|\cdot\|_{D_i}^2 = y = t_i \|\cdot\|_2^2$ .

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

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

we chose a special instance of  $y$ . For any  $x \in \mathbb{C}^{N_i}$ , let  $\bar{y} \equiv \bar{y}(x) \in \mathbb{C}^{N_{i+1}}$  be defined as  $\bar{y} = [P_i^H P_i]^{-1} P_i^H x$ . Therefore, (2.8) is implied by

$$\|x - P_i \bar{y}\|_2^2 \leq (\gamma/t_i) \|x\|_{A_i}^2, \quad \forall x \in \mathbb{C}^{N_i},$$

where the latter is equivalent to the matrix inequality

$$(3.14) \quad W_{n_i}(p_i)^H W_{n_i}(p_i) \leq (\gamma/t_i) C_{n_i}(f_i),$$

with  $W_{n_i}(p_i) = I_{N_i} - P_i [P_i^H P_i]^{-1} P_i^H$ . Since  $W_{n_i}(p_i)^H W_{n_i}(p_i) = W_{n_i}(p_i)$ , inequality (3.14) can be rewritten as

$$(3.15) \quad W_{n_i}(p_i) \leq (\gamma/t_i) C_{n_i}(f_i).$$

Let  $\mu = (\mu_1, \dots, \mu_d)$  with  $0 \leq \mu_r \leq (n_{i+1})_r - 1$ ,  $r = 1, \dots, d$ , 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  $\text{diag}(f_i[\mu])$  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 (3.15) is equivalent to

$$(3.16) \quad I_G - \frac{p_i[\mu](p_i[\mu])^T}{\|p_i[\mu]\|_2^2} \leq (\gamma/t_i)\text{diag}(f_i[\mu]), \quad \forall \mu.$$

By the Sylvester inertia law [24], the relation (3.16) is satisfied if every entry of

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

is bounded in modulus by a constant, which follows from the conditions (3.9) and (3.10).

□

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

$$(3.17) \quad a_{d,g}(x) = \prod_{j=1}^d \sum_{k=0}^{g_j-1} e^{-ikx_j}, \quad x \in [-\pi, \pi]^d.$$

**THEOREM 3.5.** *For the function  $a_{d,g}$  defined in (3.17) there exists a constant  $c$  with  $0 < c < +\infty$  such that*

$$(3.18) \quad \limsup_{x \rightarrow 0} \frac{|a_{d,g}(y)|}{\sum_{j=1}^d x_j^z} = c, \quad y \in \mathcal{M}_g(x).$$

where  $z = d - \#\{y_j \mid y_j = 0, j = 1, \dots, d\}$  is the number of directions along which  $a_{d,g}$  is zero.

Further on, if  $f_i$  has a single isolated zero of order 2 at the origin,  $p_i = a_{d,g}$  fulfills (3.10) and (3.9).

*Proof.* The limit (3.18) 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

$$\sum_{k=0}^{g_j-1} e^{-iky_j} = \sum_{k=0}^{g_j-1} e^{-ik(x_j + \frac{2\pi\ell}{g_j})} = \sum_{k=0}^{g_j-1} e^{-\frac{i2\pi k\ell}{g_j}} e^{-ikx_j}.$$

Since

$$\sum_{k=0}^{g_j-1} e^{-\frac{i2\pi k\ell}{g_j}} = \begin{cases} g_j & \text{if } \ell = 0, \\ 0 & \text{otherwise,} \end{cases}$$

the  $j$ -th factor in (3.17) 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

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

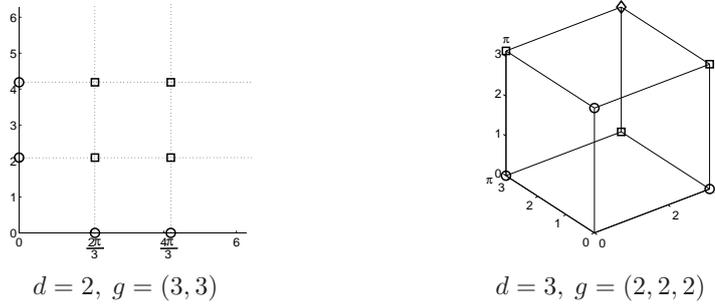


FIG. 3.1. Order of  $y \in \mathcal{M}_g(0)$  for the aggregation operator  $a_{d,g}$ :  $\circ \rightarrow$  order = 1,  $\square \rightarrow$  order = 2,  $\diamond \rightarrow$  order = 3.

and hence  $p_i = a_{d,g}$  fulfills (3.9).

Regarding (3.10), 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 that fulfills  $|a_{d,g}|^2(y) > 0$ .  $\square$

Figure 3.1 gives a visual representation of the behaviour 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 can not 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 can not 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 subsection.

**3.4. Smoothing the projector by weighed Richardson.** The order of the zero 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  $\omega$ -Richardson smoother. In the  $d$ -level case the generating symbol of this smoother is given by

$$(3.19) \quad s_{i,\omega} : [-\pi, \pi)^d \rightarrow \mathbb{C}$$

$$(3.20) \quad x \rightarrow s_{i,\omega}(x) = 1 - \omega f_i(x).$$

LEMMA 3.6. Assume that  $f_i \geq 0$  has a single isolated zero of order 2 at the origin and that  $f_i$  obtains 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 (3.11) and (3.10).

*Proof.* Since  $\tilde{y}$  is 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 (3.11).

Regarding (3.10), the assumptions on  $f_i$  implies 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  $p_i =$



**4.1. Symmetric projection for 2D Laplacian.** Now we turn to finite differences 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 will 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 (3.19) with unique  $\omega$  such that the resulting projection  $p_i = s_{i,\omega} a_{2,g}$  fulfills (3.11). 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 on  $f$  such that  $f(0, z) = f(z, 0)$ , that is inherited by  $s_{0,\omega}$ . From (3.21) it holds  $A_{(2,2)}(0) = \{(0, \pi), (\pi, 0)\}$  and  $A_{(3,3)}(0) = \{(0, \frac{2\pi}{3}), (0, \frac{4\pi}{3}), (\frac{2\pi}{3}, 0), (\frac{4\pi}{3}, 0)\}$ . Therefore,  $\omega$  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 same properties of  $f$  thanks to Lemma 3.2 and Remark 3.3.  $\square$  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  $g = 4$  we need two smoothers with two different  $\omega$  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 (3.11). 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 by the sets  $A_{(4,4)}$  and  $A_{(5,5)}$ . Two different  $\omega$  are necessary in view of  $\cos(\pi/2) = \cos(3\pi/2) \neq \cos(\pi)$  and  $\cos(2\pi/5) = \cos(8\pi/5) \neq \cos(4\pi/5) = \cos(6\pi/5)$ .  $\square$  For anisotropic stencils even with standard coarsening two  $\omega$  are needed.

LEMMA 4.3. *Let  $f$  be an anisotropic discretization of the 2D Laplacian. If  $g = 2$  we need two different  $\omega$  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 (3.11). For  $g = 3$  we need also two  $\omega$ , 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  $\omega$  are necessary.*

*Proof.* Due to the anisotropic discretization  $f(\pi, 0) \neq f(0, \pi)$ , in general, and hence a double number of  $\omega$  with respect to the isotropic case in lemmas 4.1 and 4.2 is required.

$\square$

**4.2. Non-symmetric projection for 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 and additional smoothers  $S_j := I - \omega_j \text{diag}(A)^{-1}A, j = 1, \dots, k$ . In the symmetric application we include each  $S_j$  in the restriction and the prolongation. In the nonsymmetric application we include each  $S_j$  only once, either in the restriction or in the prolongation. Hence, the coarse system is related to the matrix

$$K_n C_n^H(a_{d,g}) S_k \dots S_1 A S_1 \dots S_k C_n(a_{d,g}) K_n^T$$

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

$$K_n C_n^H(a_{d,g}) S_1 \dots S_l A S_{l+1} \dots S_k C_n(a_{d,g}) K_n^T .$$

THEOREM 4.4. *To maintain the original block tridiagonal structure also on the coarse levels the 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,  $k < g$  smoothers are allowed in SA; if we use symmetric projection with  $R_i^T = P_i$ , then we have to satisfy  $k < g/2$ :

| $g$                                   | 2 | 3 | 4 | 5 |
|---------------------------------------|---|---|---|---|
| allowed $k$ for nonsymmetric case     | 1 | 2 | 3 | 4 |
| allowed $k$ for symmetric application | 0 | 1 | 1 | 2 |

*Proof.* The injection has block bandwidth given by  $g - 1$  upper diagonals,  $A$  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 block tridiagonal 9-point stencil if  $k < g$ .  $\square$

**THEOREM 4.5.** *To derive the right number of zeros in the restriction/prolongation such that (3.5) holds, the necessary number  $k$  of smoothers 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 on all mirror points that are not lying on the coordinate axes. Following the analysis in lemmas 4.1 and 4.2 the smoothers, resp.  $\omega_j$ , 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 smoothing only the restriction or prolongation is necessary to satisfy (3.5). 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.  $\square$

To obtain both goals on 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 2D Laplacian with 9-point stencils.** Now we want to discuss exemplarily and in detail the application of the smoothed aggregation technique on the 2D Laplacian with 9-point stencils. We design the projections that on all levels we derive again 9-point stencils and that we use smoothers in the projection to get zeros of order at least 2 on all mirror points besides the origin according to condition (3.5). Therefore, our analysis will be focused to obtain a stable stencil according to the following definition.

**DEFINITION 4.6.** *A stencil associated to a symbol  $f_i$  is stable if exist  $r_i$  and  $p_i$  that satisfy (3.5), and  $f_{i+1} = \alpha_i f_i$ , with  $\alpha_i > 0$ . Of course, if the stencil  $f$  at the finest level is stable the same holds for all  $f_i$  at the coarser levels  $i = 1, \dots, l_{\max}$ .*

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 the points 1–3 at the end of the previous subsection, we study in detail

points 1 and 2 for the isotropic stencil

$$(4.1) \quad \frac{1}{4+4c} \begin{bmatrix} -c & -1 & -c \\ -1 & 4+4c & -1 \\ -c & -1 & -c \end{bmatrix}, \quad c \geq 0,$$

which is associated to the symbol

$$(4.2) \quad f(x, y) = (2 - \cos(x) - \cos(y) + c(2 - \cos(x+y) - \cos(x-y)))/(2+2c),$$

and the point 3 for the anisotropic case

$$(4.3) \quad f(x, y) = ((1 - \cos(x)) + b(1 - \cos(y))), \quad b > 0.$$

Firstly, we compute stable stencils for the isotropic case and nonsymmetric projection (point 1) for  $g = 2, \dots, 5$ .

**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  $1 - \omega f(0, \pi) = 0$  and so to satisfy (3.5). Using the function

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

from (3.6) it follows

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

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

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

These function values are related to a 9-point stencil, resp. trigonometric polynomial

$$(4.4) \quad f_1(x, y) = \sigma - \delta(\cos(x) + \cos(y)) - \epsilon \cos(x) \cos(y)$$

with

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

resulting in the coarse grid stencil

$$(4.5) \quad \frac{1}{8(1+2c)(1+c)} \begin{bmatrix} -1/4 - c/2 - c^2/2 & -c - c^2 & -1/4 - c/2 - c^2/2 \\ -c - c^2 & 1 + 6c + 6c^2 & -c - c^2 \\ -1/4 - c/2 - c^2/2 & -c - c^2 & -1/4 - c/2 - c^2/2 \end{bmatrix},$$

which gives a 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 proven using the same technique, where the coarse symbol  $f_1$  is computed generalizing (3.6) 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*

$$(4.6) \quad \frac{1}{18(1+2c)(1+c)} \begin{bmatrix} -3 - 4.5c - 3c^2 & 3/2 - 9c - 12c^2 & -3 - 4.5c - 3c^2 \\ 3/2 - 9c - 12c^2 & 6 + 54c + 60c^2 & 3/2 - 9c - 12c^2 \\ -3 - 4.5c - 3c^2 & 3/2 - 9c - 12c^2 & -3 - 4.5c - 3c^2 \end{bmatrix}$$

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

$$(4.7) \quad \frac{1}{8(1+c)(1+2c)^2} \begin{bmatrix} -5c - 8c^2 - 5c^3 & -2 - 2c - 8c^2 - 6c^3 & -5c - 8c^2 - 5c^3 \\ -2 - 2c - 8c^2 - 6c^3 & 8 + 28c + 64c^2 + 44c^3 & -2 - 2c - 4c^2 - 6c^3 \\ -5c - 8c^2 - 5c^3 & -2 - 2c - 8c^2 - 6c^3 & -5c - 8c^2 - 5c^3 \end{bmatrix}$$

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

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

$$(4.8) \quad \frac{1}{20(1+c)(1+2c)^2} \begin{bmatrix} 2 - 13c - 24c^2 - 16c^3 & -9 - 4c - 12c^2 - 8c^3 & 2 - 13c - 24c^2 - 16c^3 \\ -9 - 4c - 12c^2 - 8c^3 & 28 + 68c + 144c^2 + 96c^3 & -9 - 4c - 12c^2 - 8c^3 \\ 2 - 13c - 24c^2 - 16c^3 & -9 - 4c - 12c^2 - 8c^3 & 2 - 13c - 24c^2 - 16c^3 \end{bmatrix}$$

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

Consider now the isotropic case and symmetric projection (point 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

$$(4.9) \quad \frac{1}{12(1+2c)^2(1+c)} \begin{bmatrix} -7c - 12c^2 - 8c^3 & -3 - 4c - 12c^2 - 8c^3 & -7c - 12c^2 - 8c^3 \\ -3 - 4c - 12c^2 - 8c^3 & 12 + 44c + 96c^2 + 64c^3 & -3 - 4c - 12c^2 - 8c^3 \\ -7c - 12c^2 - 8c^3 & -3 - 4c - 12c^2 - 8c^3 & -7c - 12c^2 - 8c^3 \end{bmatrix}$$

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

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

Finally, we consider the anisotropic case and nonsymmetric projection (point 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  $\omega$ 's,  $\omega_1 = \frac{2(1+b)}{3b}$  and  $\omega_2 = \frac{2(1+b)}{3}$ , those lead to

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

Therefore, the coarse grid symbol is

$$f_1(x, y) = \alpha(1 - \cos(x)) + \beta(1 - \cos(y))$$

with

$$\beta = \frac{b}{1+b}, \quad \alpha = \frac{1}{1+b}.$$

□

**5. Numerical examples.** All numerical tests were obtained using MATLAB R2012a. We implemented the outlined method based on the developed theory for circulant and Toeplitz  $d$ -level matrices with generating symbols with second order zero at the origin. The optimal  $\omega$  was chosen automatically on each level by computing the value of the symbol at all the critical mirror points lying on the axes. We used 3 steps each of the Richardson iteration 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 to yield 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.

| # dof            | # iter. | op. compl. | asympt. conv. |
|------------------|---------|------------|---------------|
| $4 \times 4$     | 13      | 1.1000     | 0.1780        |
| $8 \times 8$     | 12      | 1.3000     | 0.1779        |
| $16 \times 16$   | 12      | 1.3750     | 0.1680        |
| $32 \times 32$   | 12      | 1.3938     | 0.1693        |
| $64 \times 64$   | 12      | 1.3938     | 0.1720        |
| $128 \times 128$ | 12      | 1.3996     | 0.1714        |
| $256 \times 256$ | 12      | 1.3999     | 0.1714        |

TABLE 5.1

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

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

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

For  $c = 0$  this yields the 2nd-order accurate 5-point finite difference discretization of the Laplacian with the stencil

$$(5.2) \quad \frac{1}{4} \begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix},$$

while for  $c = 1$  we obtain the 2nd-order accurate 9-point finite element discretization of the Laplacian given by the stencil

$$(5.3) \quad \frac{1}{8} \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}.$$

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 denoted 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, the results for the 9-point stencil (5.3) are found in Table 5.2. 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 in Table 5.3.

Next, we consider  $g = 3$ . In this case symmetric smoothing of prolongation and restriction does not lead to stencil grow, so we first start with this approach. We tested this approach 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 5.4 and 5.5, 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,

| # dof            | # iter. | op. compl. | asympt. conv. |
|------------------|---------|------------|---------------|
| $4 \times 4$     | 8       | 1.1111     | 0.0596        |
| $8 \times 8$     | 10      | 1.2778     | 0.1071        |
| $16 \times 16$   | 9       | 1.3194     | 0.1062        |
| $32 \times 32$   | 9       | 1.3299     | 0.1060        |
| $64 \times 64$   | 9       | 1.3325     | 0.1067        |
| $128 \times 128$ | 9       | 1.3331     | 0.1046        |
| $256 \times 256$ | 9       | 1.3333     | 0.1053        |

TABLE 5.2

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

| # dof            | # iter. | op. compl. | asympt. conv. |
|------------------|---------|------------|---------------|
| $4 \times 4$     | 9       | 1.1111     | 0.0730        |
| $8 \times 8$     | 10      | 1.2778     | 0.1103        |
| $16 \times 16$   | 9       | 1.3194     | 0.1059        |
| $32 \times 32$   | 9       | 1.3299     | 0.1105        |
| $64 \times 64$   | 9       | 1.3325     | 0.1041        |
| $128 \times 128$ | 9       | 1.3331     | 0.1055        |
| $256 \times 256$ | 9       | 1.3333     | 0.1054        |

TABLE 5.3

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. | asympt. conv. |
|------------------|---------|------------|---------------|
| $9 \times 9$     | 19      | 1.1355     | 0.3606        |
| $27 \times 27$   | 23      | 1.1908     | 0.4296        |
| $81 \times 81$   | 23      | 1.2129     | 0.4368        |
| $243 \times 243$ | 24      | 1.2209     | 0.4376        |

TABLE 5.4

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

| # dof            | # iter. | op. compl. | asympt. conv. |
|------------------|---------|------------|---------------|
| $9 \times 9$     | 13      | 1.0800     | 0.2215        |
| $27 \times 27$   | 16      | 1.1082     | 0.2718        |
| $81 \times 81$   | 16      | 1.1191     | 0.2769        |
| $243 \times 243$ | 16      | 1.1230     | 0.2788        |

TABLE 5.5

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

| # dof            | # iter. | op. compl. | asympt. conv. |
|------------------|---------|------------|---------------|
| $9 \times 9$     | 17      | 1.0800     | 0.2570        |
| $27 \times 27$   | 16      | 1.1082     | 0.2943        |
| $81 \times 81$   | 16      | 1.1191     | 0.2989        |
| $243 \times 243$ | 16      | 1.1230     | 0.3015        |

TABLE 5.6

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

| # dof            | # iter. | op. compl. | asympt. conv. |
|------------------|---------|------------|---------------|
| $9 \times 9$     | 17      | 1.0800     | 0.2799        |
| $27 \times 27$   | 16      | 1.1082     | 0.2746        |
| $81 \times 81$   | 16      | 1.1191     | 0.2784        |
| $243 \times 243$ | 16      | 1.1230     | 0.2819        |

TABLE 5.7

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. | asympt. conv. |
|------------------|---------|------------|---------------|
| $16 \times 16$   | 40      | 1.0625     | 0.6336        |
| $64 \times 64$   | 39      | 1.0664     | 0.6214        |
| $256 \times 256$ | 40      | 1.0667     | 0.6250        |

TABLE 5.8

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

so the positive definiteness is not guaranteed, anyway: The results for the 9-point stencil (5.3) are in Table 5.6, those for the stencil (5.1) with  $c = 1/\sqrt{2}$ , that is stable due to Theorem 4.8, are in Table 5.7. 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 in the circulant case that can be found in Table 5.8. Finally, results for the stencil (5.1) with  $c = 0.22968147..$  are presented in Table 5.9 for the Toeplitz case with  $g = 5$ . The stencil is stable due to Theorem 4.10, 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

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

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 in Table 5.10, 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 5.11. The consideration of even higher anisotropies is not meaningful, as other coarsening strategies like semicoarsening or the use of stretched aggregates is advisable.

| # dof            | # iter. | op. compl. | asyp. conv. |
|------------------|---------|------------|-------------|
| $25 \times 25$   | 47      | 1.0319     | 0.6751      |
| $125 \times 125$ | 49      | 1.0395     | 0.6924      |
| $625 \times 625$ | 50      | 1.0412     | 0.6975      |

TABLE 5.9

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. | asyp. conv. |
|------------------|---------|------------|-------------|
| $3^2 \times 3^2$ | 16      | 1.0800     | 0.2744      |
| $3^3 \times 3^3$ | 18      | 1.1082     | 0.3346      |
| $3^4 \times 3^4$ | 18      | 1.1191     | 0.3376      |
| $3^5 \times 3^5$ | 19      | 1.1230     | 0.3450      |

TABLE 5.10

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

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

$$\frac{4}{3} \begin{bmatrix} -4 & -8 & -4 \\ -8 & & -8 \\ -4 & -8 & -4 \end{bmatrix} \begin{bmatrix} -8 & & -8 \\ & 128 & \\ -8 & & -8 \end{bmatrix} \begin{bmatrix} -4 & -8 & -4 \\ -8 & & -8 \\ -4 & -8 & -4 \end{bmatrix}.$$

The results for the Toeplitz case with  $g = 3$  are in Table 5.12, the results for the circulant case are very similar, so we omit them. The results show that the approach works as expected for higher levels/dimensions, as well.

**5.4. Optimality of  $\omega$ .** To illustrate the optimality of the  $\omega$  resulting from our analysis, we varied the  $\omega$ . 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  $\omega$  obtained with the developed theory by multiplying it by a factor  $\alpha \in [0.9, 1.1]$  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 5.13. While the asymptotic convergence rate does not vary much in a neighborhood of the optimal  $\omega$ , the optimal  $\omega$  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.

**5.5. Non-constant coefficient case.** The obtained results can be used to define SA methods for the non-constant coefficient case straightforwardly. For that purpose we use Jacobi as smoother, but we introduce a diagonal matrix  $\Omega$  to damp the relaxation. We deal with model problem 3 in [26, p. 131], i.e.

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

where  $\epsilon$  varies, discretized using the stencil

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

| # dof            | # iter. | op. compl. | asympt. conv. |
|------------------|---------|------------|---------------|
| $3^2 \times 3^2$ | 21      | 1.0800     | 0.3846        |
| $3^3 \times 3^3$ | 24      | 1.1082     | 0.4545        |
| $3^4 \times 3^4$ | 26      | 1.1191     | 0.4829        |
| $3^5 \times 3^5$ | 26      | 1.1230     | 0.4841        |

TABLE 5.11

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

| # dof                    | # iter. | op. compl. | asympt. conv. |
|--------------------------|---------|------------|---------------|
| $9 \times 9 \times 9$    | 12      | 1.0293     | 0.1964        |
| $27 \times 27 \times 27$ | 15      | 1.0421     | 0.2611        |
| $81 \times 81 \times 81$ | 15      | 1.0469     | 0.2715        |

TABLE 5.12

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

We chose  $\epsilon$  as

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

and we scaled the matrix symmetrically such that it has ones on the diagonal. We build regular  $3 \times 3$  aggregates, i.e., we deal with the case  $g = 3$ . For the traditional SA approach we smoothed the prolongation and restriction operator with  $\omega$ -Jacobi, in accordance with [14] we choose  $\omega = 2/3$ . For our adaptive approach using the local model we build a local stencil for each grid point and calculated the locally optimal  $\omega$ 's. As the problem is locally anisotropic we got two  $\omega$ 's that were used to build diagonal matrices  $\Omega_1$  and  $\Omega_2$  that are used to smooth 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 5.14 and in Figure 5.1. The choice of smoothing parameters that is achieved is illustrated in Figure 5.2, where the two  $\omega$ 's that are chosen in the  $27 \times 27$  case are plotted. Our modification clearly outperforms the traditional approach. Moreover, the dependence of the iteration count on the system size and thus on the number of levels is weaker, as we see an increase of more than 35% for the plain SA when going from  $81 \times 81$  to  $243 \times 243$ , while the increase in the latter case is only 12%.

**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 (2.13) by the prolongation and restriction operator in that case. Guided by this observation sufficient conditions for an improvement of the grid transfer operators by application of the 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

| $\alpha$      | 0.90   | 0.95   | 1.00   | 1.05   | 1.10   |
|---------------|--------|--------|--------|--------|--------|
| asympt. conv. | 0.2901 | 0.2800 | 0.2777 | 0.2783 | 0.2847 |

TABLE 5.13

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

| # dof                            | $9 \times 9$ | $27 \times 27$ | $81 \times 82$ | $243 \times 243$ |
|----------------------------------|--------------|----------------|----------------|------------------|
| # iter. ( $\omega = 2/3$ )       | 20           | 32             | 45             | 61               |
| asympt. conv. ( $\omega = 2/3$ ) | 0.3246       | 0.4923         | 0.6038         | 0.6820           |
| # iter. (opt. $\omega$ )         | 15           | 20             | 25             | 28               |
| asympt. conv. (opt. $\omega$ )   | 0.2164       | 0.3686         | 0.4251         | 0.4628           |

TABLE 5.14

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

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.

## REFERENCES

- [1] G. Fiorentino, S. Serra, Multigrid methods for Toeplitz matrices, *Calcolo* 28 (1991) 238–305.
- [2] G. Fiorentino, S. Serra, Multigrid methods for symmetric positive definite block Toeplitz matrices with non-negative generating functions, *SIAM J. Sci. Comput.* 17 (5) (1996) 1068–1081.
- [3] S. Serra-Capizzano, C. Tablino-Possio, Multigrid methods for multilevel circulant matrices, *SIAM J. Sci. Comput.* 26 (1) (2004) 55–85.
- [4] A. Aricò, M. Donatelli, S. Serra-Capizzano, V-cycle optimal convergence for certain (multilevel) structured linear systems, *SIAM J. Matrix Anal. Appl.* 26 (1) (2004) 186–214.
- [5] A. Aricò, M. Donatelli, A V-cycle multigrid for multilevel matrix algebras: proof of optimality, *Numer. Math.* 105 (2007) 511–547.
- [6] M. Donatelli, S. Serra-Capizzano, D. Sesana, Multigrid methods for Toeplitz linear systems with different size reduction, *BIT* 52 (2) (2012) 305–327.
- [7] J. Mandel, S. F. McCormick, J. W. Ruge, An algebraic theory for multigrid methods for variational problems, *SIAM J. Numer. Anal.* 25 (1) (1988) 91–110.
- [8] S. F. McCormick, Multigrid methods for variational problems: General theory for the V-cycle, *SIAM J. Numer. Anal.* 22 (4) (1985) 634–643.
- [9] S. F. McCormick, J. W. Ruge, Multigrid methods for variational problems, *SIAM J. Numer. Anal.* 19 (5) (1982) 924–929.
- [10] J. W. Ruge, K. Stüben, Algebraic multigrid, in: S. F. McCormick (Ed.), *Multigrid methods*, Vol. 3 of *Frontiers Appl. Math.*, SIAM, Philadelphia, 1987, pp. 73–130.
- [11] D. Braess, Towards algebraic multigrid for elliptic problems of second order, *Computing* 55 (1995) 379–393.
- [12] F. Chatelin, W. L. Miranker, Acceleration by aggregation of successive approximation methods, *Linear Algebra Appl.* 43 (1982) 17–47.
- [13] J. Mandel, B. Sekerka, A local convergence proof for the iterative aggregation method, *Linear Algebra Appl.* 51 (1983) 163–172.
- [14] P. Vaňek, J. Mandel, M. Brezina, Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems, *Computing* 56 (1996) 179–196.
- [15] M. Brezina, R. D. Falgout, S. P. MacLachlan, T. Manteuffel, S. F. McCormick, J. Ruge, Adaptive smoothed aggregation (asa) multigrid, *SIAM Rev.* 47 (2) (2005) 317–346.
- [16] A. C. Muresan, Y. Notay, Analysis of aggregation-based multigrid, *SIAM J. Sci. Comput.* 30 (2) (2008) 1082–1103.
- [17] A. Napov, Y. Notay, Algebraic analysis of aggregation-based multigrid, *Num. Lin. Alg. Appl.* 18 (2011) 539–564.
- [18] M. Brezina, P. Vaňek, P. S. Vassilevski, An improved convergence analysis of smoothed aggregation algebraic

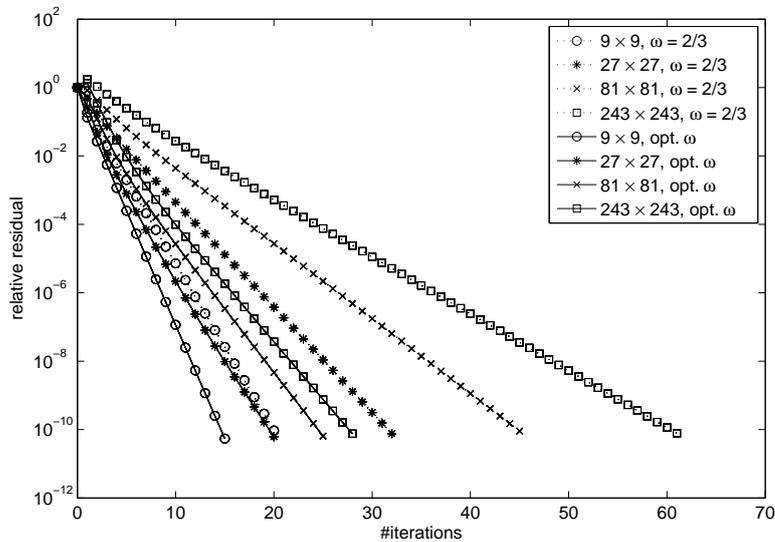


FIG. 5.1. Convergence history of the standard SA method with  $\omega = 2/3$  and the proposed version with adaptively chosen  $\omega$  based on the local stencil, aggregate size in both cases was 3.

- multigrid, Num. Lin. Alg. Appl. 9 (3) (2012) 441–469.
- [19] E. Tyrtshnikov, A unifying approach to some old and new theorems on preconditioning and clustering, Linear Algebra Appl. 232 (1996) 1–43.
- [20] T. Huckle, J. Staudacher, Multigrid methods for block Toeplitz matrices with small size blocks, BIT 46 (2006) 61–83.
- [21] R. Fischer, T. Huckle, Multigrid methods for anisotropic bttb systems, Linear Algebra Appl. 416 (2006) 314–334.
- [22] T. Huckle, J. Staudacher, Multigrid preconditioning and toeplitz matrices, Electron. Trans. Numer. Anal. 13 (2002) 82–105.
- [23] M. Donatelli, An algebraic generalization of local Fourier analysis for grid transfer operators in multigrid based on Toeplitz matrices, Num. Lin. Alg. Appl. 17 (2010) 179–197.
- [24] G. H. Golub, C. F. Van Loan, Matrix Computations, The Johns Hopkins University Press, Baltimore and London, 1996.
- [25] R. Wienands, W. Joppich, Practical Fourier analysis for multigrid methods, Vol. 4 of Numerical Insights, Chapman & Hall/CRC, Boca Raton, 2005.
- [26] U. Trottenberg, C. Oosterlee, A. Schüller, Multigrid, Academic Press, San Diego, 2001.

