ALGEBRAIC ANALYSIS OF AGGREGATION-BASED MULTIGRID

ARTEM NAPOV* AND YVAN NOTAY†

Abstract. Convergence analysis of two-grids methods based on coarsening by (unsmoothed) aggregation is presented. For diagonally dominant symmetric (M-)matrices, it is shown that the analysis can be conducted locally; that is, the convergence factor can be bounded above by computing separately for each aggregate a parameter which in some sense measures its quality. The procedure is purely algebraic and can be used to control a posteriori the quality of automatic coarsening algorithms. Assuming the aggregation pattern sufficiently regular, it is further shown that the resulting bound is asymptotically sharp for a large class of elliptic boundary value problems, including problems with variable and discontinuous coefficients. In particular, the analysis of typical examples shows that the convergence rate is insensitive to discontinuities under some reasonable assumptions on the aggregation scheme.

Key words. multigrid, algebraic multigrid, two-grid cycle, aggregation, convergence analysis

AMS subject classifications. 65F10, 65N55, 65F50

1. Introduction.

We consider multigrid methods [15, 5, 16] for solving large sparse $n \times n$ linear systems

$$A\mathbf{x} = \mathbf{b} \tag{1.1}$$

with symmetric positive definite (SPD) system matrix A. Multigrid methods are based on the recursive use of a two-grid scheme. A basic two-grid method combines the action of a smoother, often a simple iterative method, and a coarse grid correction, which corresponds to the solution of the residual equations on a coarser grid. The convergence depends on the interplay between this two components and, when simple smoothers are used, it relies essentially on the *coarsening*; that is, on the way the fine grid equations are approximated by the coarse system.

Here we consider coarsening by aggregation. In such schemes, the fine grid unknowns are grouped into disjoint sets, and each set is associated with a unique coarse grid unknown. Piecewise constant prolongation is then a common choice, which means that the solution of the residual equation computed on the coarse grid is transferred back to the fine grid by assigning the value of a given coarse variable to all fine grid variables associated with it. This makes the coarse grid matrix easy to compute and usually as sparse as the original fine grid matrix.

Aggregation schemes are not new and trace back to [1, 3]. They did not receive much attention till recently because of the difficulty to obtain optimal solvers on their basis with V- or W-cycle [14, p.p. 522-524]. However, in two-grid setting, its application to model constant coefficient discrete PDE problems can lead to a level-independent convergence [6]. Moreover, the level-dependent convergence in multi-level setting may be cured by using more sophisticated K-cycles, in which Krylov subspace acceleration is used at each level [11].

Now, the (Fourier) analysis developed in [6] only addresses constant coefficient problems with artificial (periodic) boundary conditions. Although there are numerical evidences that aggregation based methods can be robust in presence of varying or discontinuous coefficients [10], this yet remains to be proved. On the other hand, it is also lacking an analysis which would not only allow to assess a given aggregation scheme for a problem at hand, but could also serve as a guideline in the development of aggregation algorithms, in much the same way the coarsening strategies used in classical AMG methods may be derived from the objective to keep reasonably bounded some convergence measure of the resulting two-grid scheme [2, 12, 13, 14].

This paper is a short version of [8], where these gaps are filled by developing a convergence analysis which relates the global convergence to "local" quantities associated with each aggregate. This analysis is based on a general algebraic result which requires only an appropriate splitting of the system matrix A, and we show how this splitting can be constructed in a systematic way when the matrix is diagonally dominant. Further, the needed local quantities are easy to compute solving an eigenvalue problem of the size of the aggregate. They can also be assessed analytically in a number of cases. This assessment reveals that the convergence is to a large extent insensitive to variations or discontinuities in PDE coefficients if aggregation satisfy some reasonable requirements.

Moreover, the bounds deduced in this way can often be shown asymptotically sharp provided that one assumes a simplified smoothing scheme with only one damped Jacobi pre- or post-smoothing step. Hence, we

*Université Libre de Bruxelles, Service de Métrologie Nucléaire (C.P. 165-84), 50 Av. F.D. Roosevelt, B-1050 Brussels, Belgium. anapov@ulb.ac.be, homepages.ulb.ac.be/∼anapov. Research supported by the Belgian FRNS ("Aspirant").

†Université Libre de Bruxelles, Service de Métrologie Nucléaire (C.P. 165-84), 50 Av. F.D. Roosevelt, B-1050 Brussels, Belgium. ynotay@ulb.ac.be, homepages.ulb.ac.be/∼ynotay. Research supported by the Belgian FRNS ("Directeur de recherches").

do not only develop a qualitative analysis, but also a quantitative one, complementary to Fourier analysis: this latter allows to assess the benefit of more smoothing steps or increasing smoother quality, but is restricted to constant coefficient problems on rectangular grids.

The remainder of this paper is organized as follows. The general framework of aggregation-based two-grid methods is introduced in Section 2. The algebraic analysis is developed in Section 3, and illustrated in Sections 4 and 5 on PDE problems with, respectively, continuous and discontinuous coefficients.

2. Aggregation-based two-grid schemes. The coarsening procedure is based on the agglomeration of the unknowns of the system (1.1) into n_c non-empty disjoint sets called *aggregates*. The size of k-th aggregate is denoted by $n^{(k)} > 0$. For the sake of simplicity, we assume that every node belongs to an aggregate; a more general situation where some nodes remain outside an aggregation procedure is discussed in [8]. We also assume, without loss of generality, that the ordering of the unknowns is such that those belonging to $(k+1)$-th aggregate have higher indices that those belonging to k-th aggregate, $k = 1, ..., n_c - 1$.

The aggregates are the variables of the next (coarse) level in the multigrid hierarchy. Once they are defined, the $n \times n_c$ prolongation matrix is given by

$$(P)_{ij} = \begin{cases} 1 & \text{if } i \text{ belongs to } j\text{-th aggregate }, \quad j = 1, ..., n_c \\ 0 & \text{otherwise}. \end{cases}$$

Hence, setting $\mathbf{1}_m = (1\ 1\ \cdots\ 1)^T$, with m being the vector size, we have

$$P = \begin{pmatrix} \mathbf{1}_{n^{(1)}} & & \\ & \ddots & \\ & & \mathbf{1}_{n^{(n_c)}} \end{pmatrix}. \tag{2.1}$$

In what follows, we assume a slightly more general form of (2.1)

$$P = \begin{pmatrix} \mathbf{p}^{(1)} & & \\ & \ddots & \\ & & \mathbf{p}^{(n_c)} \end{pmatrix} \tag{2.2}$$

with $\mathbf{p}^{(k)}$ being a vector of size $n^{(k)}$. We shall see, however, that for the considered examples the choice $\mathbf{p}^{(k)} = \mathbf{1}_{n^{(k)}}$ is often the best (or even the only reasonable) choice.

Once the prolongation P is known, the $n_c \times n$ restriction matrix is set to its transpose and the $n_c \times n_c$ coarse grid matrix is given by the Galerkin formula $A_c = P^T A P$. In order to complete the definition of a two-grid scheme, one also needs to specify the pre- and post-smoother matrices M_1, M_2, as well as the number ν_1 and ν_2 of pre- and post-smoothing steps, respectively. The iteration matrix E_{TG} of the two-grid cycle is then given by

$$E_{TG} = (I - M_2^{-1}A)^{\nu_2}(I - P^T A_c^{-1} P A)(I - M_1^{-1}A)^{\nu_1}, \tag{2.3}$$

where I stands for identity matrix. The main objective of this paper is the analysis of its spectral radius $\rho(E_{TG})$ (that is, its largest eigenvalue in modulus), which governs the convergence of the two-grid scheme.

It is often convenient to define a "global" smoother X via the relation

$$I - X^{-1}A = (I - M_1^{-1}A)^{\nu_1}(I - M_2^{-1}A)^{\nu_2}. \tag{2.4}$$

X has the same effect in one iteration as ν_2 steps of post-smoothing followed by ν_1 steps of pre-smoothing. In what follows, we assume that X is SPD, which does not necessarily requires the symmetry of M_1 and M_2.

3. Algebraic analysis. The starting point of our analysis is a notorious identity for the two-grid convergence rate introduced in [4, Theorem 4.3]. We recall it up to a slight generalization in Theorem 3.1 below. The generalization, that is based on the results in [9], allows for nonsymmetric smoothing scheme, e.g., $\nu_1 = 1$ and $\nu_2 = 0$. It is somehow important because the parameter μ_D for $D = \mathrm{diag}(A)$, which is investigated in the remainder of this paper, appears then directly connected to the convergence factor of a simplified two-grid scheme with only 1 pre- or post-smoothing step.

THEOREM 3.1. *Let A be a $n \times n$ SPD matrix and let P be a $n \times n_c$ matrix of rank $n_c < n$. Let M_1, ν_1 and M_2, ν_2 be such that X, defined by (2.4), is a $n \times n$ SPD matrix and let E_{TG} be the two-grid iteration matrix defined by (2.3).*

Then, setting $\pi_X = P(P^T X P)^{-1} P^T X$, we have

$$\rho(E_{TG}) = \max\left(\lambda_{max}(X^{-1}A) - 1, 1 - \frac{1}{\mu_X}\right), \tag{3.1}$$

where

$$\mu_X = \max_{\mathbf{v}\in\mathbb{R}^n\setminus\{\mathbf{0}\}} \frac{\mathbf{v}^T X(I-\pi_X)\mathbf{v}}{\mathbf{v}^T A\mathbf{v}}\,.$$

Moreover, for any $n\times n$ SPD matrix D, setting $\pi_D = P(P^TDP)^{-1}P^TD$ and

$$\mu_D = \max_{\mathbf{v}\in\mathbb{R}^n\setminus\{\mathbf{0}\}} \frac{\mathbf{v}^T D(I-\pi_D)\mathbf{v}}{\mathbf{v}^T A\mathbf{v}}\,, \tag{3.2}$$

there holds

$$\mu_X \le \left(\max_{\mathbf{v}\in\mathbb{R}^n\setminus\{\mathbf{0}\}} \frac{\mathbf{v}^T X\mathbf{v}}{\mathbf{v}^T D\mathbf{v}}\right)\mu_D\ . \tag{3.3}$$

In particular, if $M_1 = M_2 = \omega^{-1}D$, $\nu_1+\nu_2 = 1$ and $\omega^{-1}\ge\lambda_{\max}(D^{-1}A)$, one has

$$\rho(E_{TG}) = 1 - \frac{\omega}{\mu_D}\,.$$

When D is chosen independently of P, the first factor in the right hand side of (3.3) depends only on the smoothing scheme. If $M_1 = M_2^T = M$ and $\nu_1 = \nu_2$, setting $S = I - M^{-1}A$, one has further

$$\frac{\mathbf{v}^T X\mathbf{v}}{\mathbf{v}^T D\mathbf{v}} \le \sigma^{-1}\ \ \forall\mathbf{v}\in\mathbb{R}^n\setminus\{\mathbf{0}\} \quad\Longleftrightarrow\quad ||S\mathbf{v}||_A^2 \le ||\mathbf{v}||_A^2 - \sigma||\mathbf{v}||_{AD^{-1}A}\ \ \forall\mathbf{v}\in\mathbb{R}^n\,.$$

Hence, when $D = \mathrm{diag}(A)$ (the choice that is privileged in the rest of this work) this quantity is nothing but the inverse of the *smoothing factor* in Ruge-Stüben analysis [14]. On the other hand, the second factor in the right hand side of (3.3) depends on P but not on X, and keeping it bounded amounts to satisfy an *approximation property*.

Now, our analysis is based on the splitting of A as

$$A = A_b + A_r\,, \tag{3.4}$$

where A_b and A_r are both symmetric nonnegative definite and A_b is block diagonal:

$$A_b = \begin{pmatrix} A^{(1)} & & \\ & \ddots & \\ & & A^{(n_c)} \end{pmatrix}, \tag{3.5}$$

where $A^{(k)}$, $k = 1,...,n_c$, is of size $n^{(k)}\times n^{(k)}$. In what follows, $\mathcal{N}(A^{(k)})$ stands for the (possibly nonempty) null space of $A^{(k)}$ and $\mathcal{R}(A^{(k)})$ represents its range.

As an example, consider a symmetric diagonally dominant matrix A with positive diagonal entries (in particular, if all off-diagonal entries are nonpositive, the matrix is an M-matrix). The matrices $A^{(k)}$, $k = 1,...,n_c$ can be constructed by restricting the matrix A to the unknowns belonging to the k-th aggregate and then by subtracting the corresponding contribution $C^{(k)} = \mathrm{diag}(c_i)$ from its diagonal, in order to keep

$$A_r = \begin{pmatrix} C^{(1)} & \cdots & * \\ \vdots & \ddots & \vdots \\ * & \cdots & C^{(n_c)} \end{pmatrix} \tag{3.6}$$

diagonally dominant, and, hence, nonnegative definite. Since A is diagonally dominant, the contribution subtracted from the diagonal of each $A^{(k)}$ can be such that either each row of A_b is weakly diagonally dominant; that is

$$(A_b)_{jj} - \sum_{i=1,\,i\neq j}^{n} |(A_b)_{ij}| = 0,\ \ j = 1,...,n\,; \tag{3.7}$$

or such that each row of A_r is weakly diagonally dominant; that is

$$(A_r)_{jj} - \sum_{i=1,\,i\neq j}^{n} |(A_r)_{ij}| = 0,\ \ j = 1,...,n\,; \tag{3.8}$$

or something in between.

Once the splitting is known, the following theorem allows to estimate the "global" approximation property constant μ_D by means of "local" quantities $\mu_D^{(k)}$, $k = 1, ..., n_c$. Because each $\mu_D^{(k)}$ corresponds to a particular aggregate k, it may be seen as a measure of this aggregate's quality.

THEOREM 3.2. *Let $A = A_b + A_r$ be a $n \times n$ SPD matrix, with A_b and A_r symmetric nonnegative definite and A_b having the block-diagonal form (3.5). Let P be a $n \times n_c$ matrix of rank $n_c < n$ and of the form (2.2). Let*

$$D = \begin{pmatrix} D^{(1)} & & \\ & \ddots & \\ & & D^{(n_c)} \end{pmatrix}$$

be a $n \times n$ SPD matrix and define μ_D as in (3.2). Letting

$$\mu_D^{(k)} = \begin{cases} 0 & \text{if } n^{(k)} = 1 \\ \displaystyle\sup_{\mathbf{v^{(k)}} \in \mathbb{R}^{n^{(k)}} \setminus \mathcal{N}(A^{(k)})} \frac{\mathbf{v^{(k)}}^T D^{(k)} (I - \pi_D^{(k)}) \mathbf{v^{(k)}}}{\mathbf{v^{(k)}}^T A^{(k)} \mathbf{v^{(k)}}} & \text{if } n^{(k)} > 1 \,, \end{cases} \tag{3.9}$$

where $k = 1, ..., n_c$ and $\pi_D^{(k)} = \mathbf{p^{(k)}} (\mathbf{p^{(k)}}^T D^{(k)} \mathbf{p^{(k)}})^{-1} \mathbf{p^{(k)}}^T D^{(k)}$, there holds

$$\mu_D \leq \max_{k=1,...,n_c} \mu_D^{(k)} \,. \tag{3.10}$$

Moreover, for $k = 1, ..., n_c$, $\mu_D^{(k)} < \infty$ if and only if $\mathcal{N}(A^{(k)}) \subset \operatorname{span}\{\mathbf{p^{(k)}}\}$, with, in the latter case,

$$\mu_D^{(k)} = \begin{cases} 0 & \text{if } n^{(k)} = 1 \\ \displaystyle\max_{\mathbf{v^{(k)}} \in \mathcal{R}(A^{(k)}) \setminus \{\mathbf{0}\}} \frac{\mathbf{v^{(k)}}^T D^{(k)} (I - \pi_D^{(k)}) \mathbf{v^{(k)}}}{\mathbf{v^{(k)}}^T A^{(k)} \mathbf{v^{(k)}}} & \text{if } n^{(k)} > 1 \,. \end{cases} \tag{3.11}$$

Now, it is clear that the value of $\mu_D^{(k)}$ strongly depends on $\mathbf{p^{(k)}}$. In the theorem below we further indicate the scope of variation of the aggregate's quality if $A^{(k)}$ and $D^{(k)}$ are given, and determine the $\mathbf{p^{(k)}}$ that leads to the best quality.

THEOREM 3.3. *Let $A^{(k)}$ and $D^{(k)}$ be, respectively, a $n^{(k)} \times n^{(k)}$ non-zero symmetric nonnegative definite matrix and a $n^{(k)} \times n^{(k)}$ SPD matrix, with $n^{(k)} > 1$. Let $\mathbf{p^{(k)}}$ be a non-zero vector of size $n^{(k)}$. Let*

$$\mu_D^{(k)} = \sup_{\mathbf{v} \in \mathbb{R}^{n^{(k)}} \setminus \mathcal{N}(A^{(k)})} \frac{\mathbf{v}^T D^{(k)} (I - \pi_D^{(k)}) \mathbf{v}}{\mathbf{v}^T A^{(k)} \mathbf{v}} \,, \tag{3.12}$$

where $\pi_D^{(k)} = \mathbf{p^{(k)}} (\mathbf{p^{(k)}}^T D^{(k)} \mathbf{p^{(k)}})^{-1} \mathbf{p^{(k)}}^T D^{(k)}$ and let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{n^{(k)}}$ be the eigenvalues of ${D^{(k)}}^{-1} A^{(k)}$. Then,

$$\lambda_2^{-1} \leq \mu_D^{(k)} \leq \lambda_1^{-1} \,. \tag{3.13}$$

Moreover, if ${D^{(k)}}^{-1} A^{(k)} \mathbf{p^{(k)}} = \lambda_1 \mathbf{p^{(k)}}$, then

$$\mu_D^{(k)} = \frac{1}{\lambda_2} \,, \tag{3.14}$$

and, assuming $\mu_D^{(k)}$ finite,

$$\mathbf{v} \in \mathcal{R}(A^{(k)}) \text{ and } \mathbf{v}^T D^{(k)} (I - \pi_D^{(k)}) \mathbf{v} = \mu_D^{(k)} \mathbf{v}^T A^{(k)} \mathbf{v} \iff {D^{(k)}}^{-1} A^{(k)} \mathbf{v} = \lambda_2 \mathbf{v} \quad \text{and} \quad \mathbf{v}^T D^{(k)} \mathbf{p^{(k)}} = 0 \,.$$

By way of illustration, consider a symmetric diagonally dominant M-matrix and assume that the splitting $A = A_b + A_r$ is based on the rule (3.7). Then, each $A^{(k)}$ is singular with its null space equal to $\operatorname{span}\{\mathbf{1}_{n^{(k)}}\}$. Theorem 3.2 then shows that one has to use $\mathbf{p^{(k)}} = \mathbf{1}_{n^{(k)}}$ to keep $\mu_D^{(k)}$ finite, in which case, by Theorem 3.3, $\mu_D^{(k)} = \lambda_2 ({D^{(k)}}^{-1} A^{(k)})^{-1}$.

4. Discrete PDEs with constant and smoothly varying coefficients.

4.1. Preliminaries. We start considering matrices associated with the 5-point stencil

$$\begin{bmatrix} & -\alpha_y & \\ -\alpha_x & \alpha_d & -\alpha_x \\ & -\alpha_y & \end{bmatrix} \quad \text{with } \alpha_x, \alpha_y > 0 \text{ and } \alpha_d \geq 2(\alpha_x + \alpha_y) \tag{4.1}$$

on a rectangular grid of arbitrary shape. For such matrices we want to assess boxwise aggregates with four nodes per aggregate (as on Figure4.1(a)) and linewise aggregates with two, three and four nodes (as on Figure4.1(b)). We select the splitting $A = A_b + A_r$ satisfying (3.8). The prolongation vector is $\mathbf{p}^{(k)} = \mathbf{1}_{n^{(k)}}$, $k = 1, ..., n_c$ and, as can be checked from (4.2) and (4.4) below, it is an eigenvector of ${D^{(k)}}^{-1} A^{(k)}$ associated with the smallest eigenvalue $\delta_d \alpha_d^{-1}$, where $\delta_d = \alpha_d - 2(\alpha_x + \alpha_y) \geq 0$. Theorem 3.3 then implies that $\mu_D^{(k)} = \lambda_2({D^{(k)}}^{-1} A^{(k)})^{-1} = \alpha_d \lambda_2 (A^{(k)})^{-1}$.

Considering more specifically boxwise aggregates, we have

$$A^{(k)} = \begin{pmatrix} \alpha_x + \alpha_y & -\alpha_x & -\alpha_y & 0 \\ -\alpha_x & \alpha_x + \alpha_y & 0 & -\alpha_y \\ -\alpha_y & 0 & \alpha_x + \alpha_y & -\alpha_x \\ 0 & -\alpha_y & -\alpha_x & \alpha_x + \alpha_y \end{pmatrix} + \delta_d \, I \ , \tag{4.2}$$

and, hence,

$$\mu_D^{(k)} = \frac{2\alpha_x + 2\alpha_y + \delta_d}{2\min(\alpha_x, \alpha_y) + \delta_d} \, , \tag{4.3}$$

whereas for linewise aggregation of size m in x direction

$$A^{(k)} = \begin{pmatrix} \alpha_x & -\alpha_x & & \\ -\alpha_x & 2\alpha_x & \ddots & \\ & \ddots & \ddots & -\alpha_x \\ & & -\alpha_x & \alpha_x \end{pmatrix} + \delta_d \, I \tag{4.4}$$

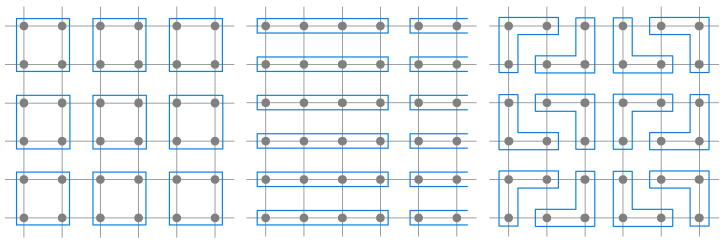
and, hence, the following formula holds for $m = 2, ..., 4$:

$$\mu_D^{(k)} = \frac{2\alpha_x + 2\alpha_y + \delta_d}{(2 - \sqrt{m-2})\alpha_x + \delta_d} \, . \tag{4.5}$$

It follows that linewise aggregates of size 4 oriented in the direction of strong coupling become more attractive than boxwise aggregates whenever $\max(\alpha_x, \alpha_y) > (2 + \sqrt{2}) \min(\alpha_x, \alpha_y)$. Always choosing the best aggregate shape, we have then

$$\mu_D^{(k)} \leq 3 + \sqrt{2} \, . \tag{4.6}$$

Since linewise aggregates of size 3 and 2 have better quality than linewise aggregates of size 4, as can be concluded from (4.5), this upper bound holds for them as well.



(a) (b) (c)

FIGURE 4.1. *Examples of (a) boxwise, (b) linewise and (c) L-shaped aggregation patterns.*

4.2. Constant coefficients. We now discuss more specifically the five point finite difference approximation of

$$\frac{\partial}{\partial x}\left(\alpha_x \frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(\alpha_y \frac{\partial u}{\partial y}\right) + \beta u = f \quad \text{on } \Omega\,, \tag{4.7}$$

with uniform mesh size h in both directions, where the boundary $\partial\Omega$ of the domain $\Omega \in \mathbb{R}^2$ is the union of segments parallel to the x or y axis and connecting the grid nodes. Note that Ω is possibly not convex and may contain holes.

If the PDE coefficients α_x, α_y and β are constant, the above results allow to assess aggregate's quality for some typical aggregate shapes. It is also easy to extend the reasoning to further aggregation schemes, leading to bound above (4.6) by a modest constant if either coefficients are isotropic ($\alpha_x = \alpha_y$) or if one uses linewise aggregation along the strong coupling direction. For instance, if $\alpha_x = \alpha_y$, (4.5) with $m = 3$ also applies to L-shaped aggregates as illustrated on Figure 4.1(c).

Regarding Neumann boundary conditions, only the quality of aggregates that contain boundary nodes is not covered by the above analysis. Again, however, isotropic coefficients and linewise aggregates aligned with strong coupling yield bounds similar to (4.3) and (4.5). For instance, if $\alpha_x = \alpha_y$ and $\beta = 0$, boxwise aggregation near a Neumann boundary result in matrices $A^{(k)}$ and $D^{(k)}$ that have the form analyzed in Lemma 5.1 below, with $\alpha_1 = \alpha_2$ and $\alpha_3 = \alpha_4 = 0$ (boundary aligned with grid lines), $\alpha_2 = \alpha_3 = \alpha_4 = 0$ (resorting corners), or $\alpha_1 = \alpha_2 = \alpha_3$ and $\alpha_4 = 0$ (re-entering corners). As shown in this lemma, one has then $\mu_D^{(k)} \leq 2$ in the two former cases and $\mu_D^{(k)} \leq 2.23$ in the latter, compared to $\mu_D^{(k)} = 2$ away from the boundary.

Note that our analysis does not require all aggregates having the same shape, which in fact seldom occurs with practical aggregation algorithms (see [10] for an example). One should just take care that the global μ_D is not larger than desired because of a few irregular aggregates, which in practice can be prevented by breaking them into smaller pieces.

When the discrete PDE (4.7) has smoothly varying coefficients, it can be shown (see [8] for details) that vanishing deviation of $A^{(k)}$ and $D^{(k)}$ from constant coefficient case leads to a vanishing perturbation in quality $\mu_D^{(k)}$. Therefore, the results of the this subsection carry over the smoothly variable coefficient case, at least when the mesh size h is small enough.

4.3. Numerical example. We consider the linear system resulting from the 5-point finite difference discretization of (4.7) on $\Omega = [0,1] \times [0,1]$ with Dirichlet boundary conditions and constant coefficients α_x, α_y and $\beta = 0$. The discretization is performed on a uniform rectangular grid of mesh size $h = (N+1)^{-1}$ in both directions.

For the sake of simplicity, we let N be a multiple of 12, which allows that the whole domain is covered with aggregates of the same shape. Using the rule (3.7), the matrices $A^{(k)}$ and $D^{(k)}$ are the same for all aggregates. As a consequence, the quality estimate $\mu_D^{(k)}$ is the same as well.

We consider first an isotropic situation ($\alpha_x = \alpha_y$). The columns from 2 to 7 of Table 4.1 then give the values of μ_D and of its upper bound $\mu_D^{(k)}$ for three types of aggregation pattern, presented on Figure 4.1. Observe that when nodes are added to an aggregate, its quality is not necessarily deteriorated, as can be seen comparing L-shaped and box aggregates. We next consider in columns 8 to 13 an anisotropic situation ($\alpha_x = 10\alpha_y$). One sees that boxwise aggregation is not recommended in this case.

TABLE 4.1
The value of μ_D and of its upper bound (3.10) for different grid sizes.

	$\alpha_x = \alpha_y\,,\ \delta_d = 0$						$\alpha_x = 10\alpha_y\,,\ \delta_d = 0$					
	pairwise		L-shaped		boxwise		linewise (size=3)		linewise (size=4)		boxwise	
N	$\mu_D^{(k)}$	μ_D	$\mu_D^{(k)}$	μ_D	$\mu_D^{(k)}$	μ_D	$\mu_D^{(k)}$	μ_D	$\mu_D^{(k)}$	μ_D	$\mu_D^{(k)}$	μ_D
12	2	1.940	4	2.315	2	1.959	2.2	2.184	3.756	3.638	11	8.431
24	2	1.984	4	2.377	2	1.989	2.2	2.196	3.756	3.744	11	10.185
48	2	1.996	4	2.394	2	1.997	2.2	2.199	3.756	3.753	11	10.778
96	2	1.999	4	2.399	2	1.999	2.2	2.200	3.756	3.755	11	10.943

4.4. Sharpness of the estimate. Numerical results in Table 4.1 indicate that the bound (3.10) on μ_D can be asymptotically sharp for N large enough. Moreover, as shown in Theorem 3.1, if only one Jacobi smoothing iteration is performed, we further have $\rho(E_{TG}) = 1 - \omega\mu_D^{-1}$. Hence, a sharp estimate of μ_D further leads to a sharp estimate of the two-grid convergence rate. The reader can wonder why and when this happens. This is what we look into in the present subsection, starting with the first question for the particular case of boxwise aggregates.

Consider that the setting of the above example holds. Without loss of generality, we assume in addition that $\alpha_x \geq \alpha_y$. First, we recall that ${D^{(k)}}^{-1} A^{(k)} \mathbf{p}^{(k)} = \lambda_1({D^{(k)}}^{-1} A^{(k)})\, \mathbf{p}^{(k)}$, and, hence, the vector $\mathbf{v}_b = (1\,1\,{-1}\,{-1})^T \in \mathcal{R}(A^{(k)})$ that can be checked to satisfy ${D^{(k)}}^{-1} A^{(k)} \mathbf{v}_b = \lambda_2({D^{(k)}}^{-1} A^{(k)})\, \mathbf{v}_b$ reaches, according to Theorem 3.3, the supremum in definition (3.9) of $\mu_D^{(k)}$. Therefore, setting

$$\widetilde{\mathbf{v}} = (\gamma_1\, {\mathbf{v}_b}^T \;\; \gamma_2\, {\mathbf{v}_b}^T \cdots \gamma_{n_c}\, {\mathbf{v}_b}^T)^T\,,$$

we locally reproduces the maximizing vectors for every aggregate. Moreover, setting $\gamma_1 = \gamma_2 = \cdots = \gamma_N = -\gamma_{N+1} = \cdots = -\gamma_{2N} = \gamma_{2N+1} = \cdots = 1$ we further make $\widetilde{\mathbf{v}}$ take the same value at every two connected nodes that belong to different aggregates. Hence, since A_r have the form (3.6) with diagonal blocks being diagonal matrices, there holds $(A_r)_{ij}\,((\widetilde{\mathbf{v}})_i - (\widetilde{\mathbf{v}})_j) = 0$ for all i and j. Therefore, setting $\sigma_i = \sum_{j=1}^n (A_r)_{ij}$ and using

$$\widetilde{\mathbf{v}} A_r \widetilde{\mathbf{v}} = -\sum_{i,j=1}^{n} \frac{1}{2}\,(A_r)_{ij}\,((\widetilde{\mathbf{v}})_i - (\widetilde{\mathbf{v}})_j)^2 + \sum_{i=1}^{n} \sigma_i\,{(\widetilde{\mathbf{v}})_i}^2 \tag{4.8}$$

there holds, since $\sigma_i > 0$ only for the unknowns near the boundary,

$$\widetilde{\mathbf{v}} A_r \widetilde{\mathbf{v}} = \sum_{i=1}^{n} \sigma_i\,{(\widetilde{\mathbf{v}})_i}^2 = 2N(\alpha_x + \alpha_y) = 2N^{-1}(\alpha_x + \alpha_y)\alpha_d^{-1}\widetilde{\mathbf{v}}^T D \widetilde{\mathbf{v}}\,. \tag{4.9}$$

On the other hand, note that ${\mathbf{p}^{(k)}}^T D^{(k)} \mathbf{v}_b = \mathbf{0}$ implies $\pi_D \widetilde{\mathbf{v}} = \mathbf{0}$, and, hence,

$$\mu_D \geq \frac{\widetilde{\mathbf{v}}^T D(I-\pi_D)\widetilde{\mathbf{v}}}{\widetilde{\mathbf{v}}^T A_b \widetilde{\mathbf{v}} + \widetilde{\mathbf{v}}^T A_r \widetilde{\mathbf{v}}} = \frac{\widetilde{\mathbf{v}}^T D \widetilde{\mathbf{v}}}{\widetilde{\mathbf{v}}^T A_b \widetilde{\mathbf{v}} + \widetilde{\mathbf{v}}^T A_r \widetilde{\mathbf{v}}} = \frac{\widetilde{\mathbf{v}}^T D \widetilde{\mathbf{v}}}{{\mu_D^{(k)}}^{-1}\,\widetilde{\mathbf{v}}^T D \widetilde{\mathbf{v}} + \widetilde{\mathbf{v}}^T A_r \widetilde{\mathbf{v}}} = \frac{\mu_D^{(k)}}{1 + \mu_D^{(k)}\,\frac{\widetilde{\mathbf{v}}^T A_r \widetilde{\mathbf{v}}}{\widetilde{\mathbf{v}}^T D \widetilde{\mathbf{v}}}}\,. \tag{4.10}$$

It then follows from (4.9) that $\mu_D \to \mu_D^{(k)}$ for $N \to \infty$. Essentially the same results can be proven for the linewise aggregates of size $m \leq 4$, see [8] for details.

A more general analysis of the sharpness ("when this happens?" part of the above question) is also possible and is performed in [8]. In particular, it is shown there that the lower bound on μ_D (like the one given by (4.10)) can be computed using only aggregates belonging to a given subset $\bar{\Omega}_h$ of unknowns. The main contribution to the left hand side of (4.8) then comes from the set $\partial\bar{\Omega}_h$ of "boundary" unknowns of $\bar{\Omega}_h$. One than often has, denoting by $|\bar{\Omega}_h|$ (resp. $|\partial\bar{\Omega}_h|$) the size of the set $\bar{\Omega}_h$ (resp. $\partial\bar{\Omega}_h$),

$$\widetilde{\mathbf{v}}^T A_r \widetilde{\mathbf{v}} = \sum_{i=0}^{n} \sigma_i (\widetilde{\mathbf{v}})_i^2 = \mathcal{O}(|\partial\bar{\Omega}_h|)\,,$$

and

$$\widetilde{\mathbf{v}}^T D \widetilde{\mathbf{v}} = \mathcal{O}(|\bar{\Omega}_h|)\,.$$

Therefore, if the subset $\bar{\Omega}_h$ corresponds to a subdomain covered with boxwise (as on Figure 4.1(a)) or linewise (as on Figure 4.1(b)) aggregation pattern and associated to constant coefficients α_x, α_y and β, the bound (4.10) can be shown to hold for $\mu_D^{(k)}$ associated to aggregates in $\bar{\Omega}_h$. Hence, one proves that the asymptotical ($h \to 0$) value of μ_D satisfies

$$\mu_D \geq \mu_D^{(k)}\,, \qquad \text{aggregate } k \subset \bar{\Omega}_h\,, \tag{4.11}$$

if the quotient $|\partial\bar{\Omega}_h|/|\bar{\Omega}_h|$ vanishes for h small enough. Moreover, if $\bar{\Omega}_h$ contains the aggregates of the worst quality; that is, if $\max_{\ell=1,\ldots,n_c} \mu_D^{(\ell)} = \mu_D^{(k)}$ for aggregates $k \subset \bar{\Omega}_h$, then (4.11) proves the asymptotical sharpness.

5. Discrete PDEs with discontinuous coefficients.

5.1. Preliminaries. As in the previous section, our analysis is based on the aggregates' quality, which in turn involves the computation of the second smallest eigenvalue of small matrices. The following lemma is helpful in this respect.

LEMMA 5.1. *Let*

$$A_d = \frac{1}{2}\begin{pmatrix} 4\alpha_1 & -2\alpha_1 & -2\alpha_1 & \\ -2\alpha_1 & 3\alpha_1+\alpha_2 & & -\alpha_1-\alpha_2 \\ -2\alpha_1 & & 3\alpha_1+\alpha_3 & -\alpha_1-\alpha_3 \\ & -\alpha_1-\alpha_2 & -\alpha_1-\alpha_3 & 2\alpha_1+\alpha_2+\alpha_3 \end{pmatrix} \tag{5.1}$$

and

$$D_d = \mathrm{diag}\,(4\alpha_1 \;\; 2(\alpha_1+\alpha_2) \;\; 2(\alpha_1+\alpha_3) \;\; (\alpha_1+\alpha_2+\alpha_3+\alpha_4))\;, \tag{5.2}$$

where $\alpha_1 > 0$ *and* α_2, α_3, $\alpha_4 > 0$. A_d *is positive semi-definite, and let* $\lambda_2(D_d^{-1}A_d)$ *be the smallest nonzero eigenvalue of* $D_d^{-1}A_d$.

Then,

$$\lambda_2(D_d^{-1}A_d) \geq \frac{5-\sqrt{17}}{8}\,, \tag{5.3}$$

and, if $\alpha_1 = \alpha_2$ *and* $\alpha_3 = \alpha_4$, *there holds*

$$\lambda_2(D_d^{-1}A_d) = \min\left(\frac{1}{2}\,, \frac{3\alpha_1+\alpha_3}{4(\alpha_1+\alpha_3)}\right)\,. \tag{5.4}$$

Moreover, if $\alpha_1 \geq \alpha_2$, α_3, α_4, *one has*

$$\lambda_2(D_d^{-1}A_d) \geq \beta \tag{5.5}$$

with $\beta = \lambda_2(D_d^{-1}A_d)\,(\approx 0.449)$ *being evaluated for* $\alpha_1 = \alpha_2 = \alpha_4 = 1$ *and* $\alpha_3 = 0$.

Furthermore,

$$\lambda_2(D_d^{-1}A_d) \geq \frac{1}{2} \quad \textit{if} \quad \left\{ \begin{array}{r} \alpha_1 \geq \alpha_2 = \alpha_3 = \alpha_4 \\ \textit{or } \alpha_1 = \alpha_2 \geq \alpha_3 = \alpha_4 \\ \textit{or } \alpha_1 = \alpha_2 = \alpha_3 \geq \alpha_4 \end{array} \right. \tag{5.6}$$

5.2. Analysis. We consider the PDE (4.7) with piecewise constant isotropic coefficients ($\alpha_x(x,y) = \alpha_y(x,y)$) and $\beta = 0$, and assume Dirichlet boundary conditions. As in the previous section, we consider the five point finite difference approximation with uniform mesh size h in both directions (mesh box integration scheme [7]), and assume that the boundary $\partial\Omega$ of $\Omega \subset \mathbb{R}^2$ is the union of segments parallel to the x or y axis and connecting the grid nodes. We aim at assessing boxwise aggregation as illustrated on Figure 4.1(a), which was shown relevant for isotropic coefficients in the previous section.

Here we assume that the possible discontinuities match the grid lines. Hence, Ω is a union of non overlapping subdomains Ω_i in which the coefficients are constant, and the boundary $\partial\Omega_i$ of each Ω_i is formed by segments aligned with grid lines and having grid nodes as end points. To exclude some uncommon situations, we assume that every two such end points are separated by a distance not less than $2h$ and that each box aggregate contains at least one point which is interior to one subdomains. In practice, this assumption is automatically met if the mesh size is small enough; in fact, it has to be not larger than $h_0/2$, where h_0 is the size of the coarsest mesh that still correctly reproduce the geometry of the problem.

The most general situation corresponding to this setting is then schematized on Figure 5.1(a) where the central aggregate has one node interior to Ω_1 and the opposite node at the intersection of four subdomains: Ω_1, Ω_2, Ω_3 and Ω_4. With the splitting satisfying (3.7), the corresponding aggregate's matrices $A^{(k)}$ and $D^{(k)}$ are given by (5.1) and (5.2), respectively, with α_i, $i = 1, ..., 4$, being the PDE coefficient in the subdomain Ω_i.

Because of the assumption (3.7) and of Theorem 3.3, aggregate's quality $\mu_D^{(k)}$ is the inverse of the second smallest eigenvalue of ${D^{(k)}}^{-1}\,A^{(k)}$. Lemma 5.1 then shows us the following.

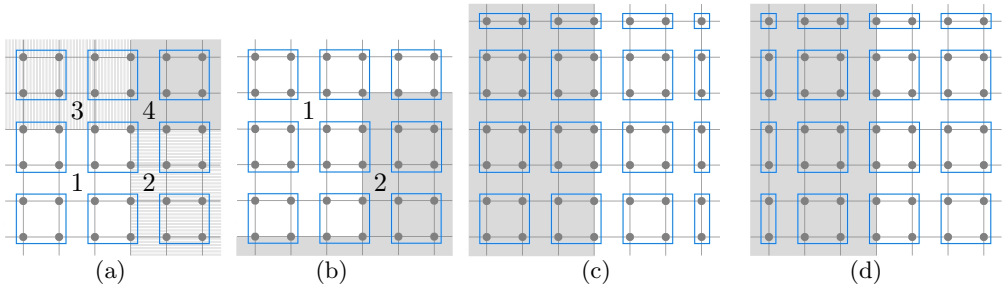


FIGURE 5.1. (*a*) *general box aggregate situation with respect to discontinuities and* (*b*) *discontinuity nodes aggregated with white point nodes;* (*c*) *and* (*d*) *schematize potential aggregation strategies for the numerical example.*

- The approach is robust in all cases, since, by (5.3), $\mu_D^{(k)}$ is always bounded above independently of the relation between the coefficients α_i.
- Nevertheless, from a practical viewpoint, (5.3) allows a significant decrease of aggregate's quality compared with the constant coefficient case. However, according to (5.5), which implies $\mu_D^{(k)} \leq 2.23$ (compared with 2 in constant coefficient case), a major deterioration is avoided when $\alpha_1 \geq \alpha_2, \alpha_3, \alpha_4$. The latter condition is satisfied if nodes belonging to several subdomains Ω_i are always aggregated only with nodes that belong to Ω_i with largest PDE coefficient α_i. Roughly speaking, the rule may be summarized as "aggregate discontinuity nodes with those of the strong coefficient region".
- In many practical cases, no more than two subdomains are involved at a time for a single aggregate, and either $\alpha_1 = \alpha_2 = \alpha_3$, or $\alpha_1 = \alpha_2$ and $\alpha_3 = \alpha_4$, or $\alpha_2 = \alpha_3 = \alpha_4$ hold, as illustrated on Figure (5.1)(b). Then, if the rule above is applied; that is, if α_1 is in addition the largest coefficient, (5.6) applies and shows that there is no deterioration at all compared with the constant coefficient case.

5.3. Numerical example. Consider the PDE (4.7) on a square domain $\Omega = [0,1] \times [0,1]$ with $\beta = 0$,

$$\alpha_x(x,y) = \alpha_y(x,y) = \begin{cases} 1 & \text{if } x \leq 1/2 \\ d(>1) & \text{if } x > 1/2\,. \end{cases}$$

and with Dirichlet boundary conditions. Consider the linear system (1.1) resulting from its five point finite difference discretization (mesh box integration scheme [7]) on the regular grid of mesh size $h = N^{-1}$. Since discontinuities needs to be aligned with grid lines, N has to be even. For simplicity of presentation, we further assume that it is a multiple of 4. The number of unknowns being $(N-1) \times (N-1)$ (there is no unknown for Dirichlet nodes), the grid cannot be covered with box aggregates only and the coarsening is completed by pair and singleton aggregates. Then, the domain may be covered with box aggregates starting from the left bottom corner (as on Figure 5.1(c)) or from the right bottom corner (as on Figure 5.1(d)).

TABLE 5.1
The value of μ_D and of its upper bound (3.10) *for different aggregation strategies and for $d = 10$.*

	strategy (a)		strategy (b)	
N	$\max_{k=1,\ldots,n_c} \mu_D^{(k)}$	μ_D	$\max_{k=1,\ldots,n_c} \mu_D^{(k)}$	μ_D
32	3.385	3.181	2	1.993
64	3.385	3.286	2	1.998
128	3.385	3.336	2	2.000
256	3.385	3.361	2	2.000

Note that the quality of aggregates outside discontinuity is at most 2, as can be concluded in the isotropic case ($\alpha_x = \alpha_y$) from (4.3) (for box aggregates) or from (4.5) with $m = 2$ (for pair aggregates). The bound is therefore determined by the quality of aggregates containing nodes on the discontinuity, which are given for $d = 10$ in Table 5.1. Observe that for the second strategy the convergence estimate is exactly the same as in the constant coefficient case. For box aggregate, this follows from the analysis in the previous subsection: the aggregates then obeys the "strong coefficient" rule stated above. Regarding the first aggregation strategy, note that for box aggregates one has

$$\mu_D^{(k)} = \lambda_2({D^{(k)}}^{-1} A^{(k)})^{-1} = \frac{4(1+d)}{3+d}\,, \tag{5.7}$$

using (5.4) with $\alpha_1 = \alpha_2 = 1$ and $\alpha_3 = \alpha_4 = d$. This is also true in the pairwise case, since then

$$A^{(k)} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\,,\quad D^{(k)} = \begin{pmatrix} 4 & \\ & 2(d+1) \end{pmatrix}\,.$$

Note that (5.7) implies $\mu_D^{(k)} = 3.38$ for $d = 10$ and $\mu_D^{(k)} \to 4$ for $d \to \infty$.

5.4. Sharpness of the estimate. Table 5.1 indicates that, once again, the upper bound (3.10) is seemingly asymptotically exact. In fact, the analysis summarized at the end of Section 4 allows to show that, asymptotically, μ_D can not be smaller than 2 for an isotropic ($\alpha_x = \alpha_y$) PDE (4.7) with $\beta = 0$ and a regular covering by box aggregates in at least one subdomain in which the PDE coefficients are constant. Hence, our analysis is accurate when discontinuity nodes are aggregated with nodes in strong coefficient region, since then $\mu_D^{(k)} \leq 2.23$. If, in addition, $\mu_D^{(k)} \leq 2$, like for the second strategy in the numerical example above, then the bound is asymptotically sharp.

It is more challenging to show the sharpness when $\mu_D^{(k)}$ is significantly larger than 2 for some aggregates along discontinuity, essentially because the proportion of such aggregates is $\mathcal{O}(h)$ or less. In particular, if $\bar{\Omega}_h$ is chosen to be a set of nodes belonging to the (worst quality) discontinuity aggregates, the quotient $|\partial\bar{\Omega}_h|/|\bar{\Omega}_h|$ then does not vanishes for $h \to 0$. Nevertheless, it is interesting to confirm that, as seen in Table 5.1, such a limited amount of low quality aggregates is sufficient to affect the global convergence, and hence that the rule "aggregate discontinuity nodes with those of the strong coefficient region" has some practical relevance. In this view, we prove in [8] the sharpness of our estimate for the numerical example above with the first aggregation strategy (depicted on Figure 5.1(c)), which does not follow the "strong coefficient" rule. It is further explained in [8] how the proof can be extended to more general situations.

6. Closing remarks. We have developed an analysis of aggregation-based two-grid method for SPD linear systems. When the system matrix is diagonally dominant, an upper bound on the convergence factor can be obtained in a purely algebraic way, assessing locally and independently the quality of each aggregate by solving an eigenvalue problem of the size of the aggregate.

We have applied our bound to scalar elliptic PDE problems in two dimensions, showing that aggregation-based two-grid methods are robust if

- in presence of anisotropy, one uses linewise aggregates aligned with the direction of strong coupling;
- in presence of discontinuities, one avoids mixing inside an aggregate nodes belonging to a strong coefficient region or its boundary with nodes interior to a weak coefficient region.

Furthermore, we have shown that the bound is asymptotically sharp when a significant part of the domain is regularly covered by box or line aggregates of the same shape.

Note that we have conducted the analysis in two dimensions for the sake of simplicity. The same type of analysis can be developed for three dimensional problems, leading to similar conclusions.

Our results may also have an impact on practical aggregation schemes. Because of the above mentioned sharpness, it is indeed sensible to expect that aggregation methods can be improved by improving aggregates' quality. And because aggregates' quality is cheap to assess, this parameter can effectively be taken into account in the design of aggregation algorithms. For instance, one may a posteriori check aggregates' quality and break low quality aggregates into smaller pieces. It is also possible, in a greedy-like approach, to decide wether a node (or a group of nodes) should be added to an aggregate according its impact on the aggregate's quality and/or select the neighboring (sets of) nodes that are the most favorable in this respect. These practical aspects are subject to further research.

REFERENCES

[1] D. Braess. Towards algebraic multigrid for elliptic problems of second order. *Computing*, 55:379–393, 1995.

[2] A. Brandt, S. F. McCormick, and J. W. Ruge. Algebraic multigrid (amg) for sparse matrix equations. In D. J. Evans, editor, *Sparsity and its Application*, pages 257–284. Cambridge University Press, Cambridge, 1984.

[3] V. E. Bulgakov. Multi-level iterative technique and aggregation concept with semi-analytical preconditioning for solving boundary-value problems. *Comm. Numer. Methods Engrng.*, 9:649–657, 1993.

[4] Robert D. Falgout, Panayot S. Vassilevski, and Ludmil T. Zikatanov. On two-grid convergence estimates. *Numer. Lin. Alg. Appl.*, 12:471–494, 2005.

[5] W. Hackbusch. *Multi-grid Methods and Applications*. Springer, Berlin, 1985.

[6] A. C. Muresan and Y. Notay. Analysis of aggregation-based multigrid. *SIAM J. Sci. Comput.*, 30:1082–1103, 2008.

[7] S. Nakamura. *Computational Methods in Engineering and Science*. John Wiley & Sons, New York, 1977.

[8] A. Napov and Y. Notay. Algebraic analysis of aggregation-based multigrid. (GANMN 09–04), 2009. `http://homepages.ulb.ac.be/~anapov`(submitted).

[9] Y. Notay. Algebraic analysis of two-grid methods: the nonsymmetric case. *Numer. Lin. Alg. Appl.* published online in Wiley InterScience. DOI: 10.1002/nla.649, 2009.

[10] Y. Notay. An aggregation-based algebraic multigrid method. *Electronic Trans. Numer. Anal.*, 2009. To appear.

[11] Y. Notay and P. S. Vassilevski. Recursive Krylov-based multigrid cycles. *Numer. Lin. Alg. Appl.*, 15:473–487, 2008.

[12] J. W. Ruge and K. Stüben. Algebraic multigrid (AMG). In S. F. McCormick, editor, *Multigrid Methods*, volume 3 of *Frontiers in Applied Mathematics*, pages 73–130. SIAM, Philadelphia, PA, 1987.

[13] K. Stüben. Algebraic multigrid (AMG): experiences and comparisons. *Appl. Math. Comput.*, 13:419–452, 1983.

[14] K. Stüben. *An Introduction to Algebraic Multigrid*, pages 413–532. In Trottenberg et al. [15], 2001. Appendix A.

[15] U. Trottenberg, C. W. Oosterlee, and A. Schüller. *Multigrid*. Academic Press, London, 2001.

[16] P. Wesseling. *An Introduction to Multigrid Methods*. J. Wiley and Sons, Chichester, 1992.