

On the Approximate Cyclic Reduction Preconditioner

by

Arnold Reusken
Institut für Geometrie und Praktische Mathematik
RWTH Aachen
Templergraben 55
D-52056 Aachen, Germany
e-mail: reuskenigpm.rwth-aachen.de

Abstract. We present a preconditioning method for the iterative solution of large sparse systems of equations. The preconditioner is based on ideas both from ILU preconditioning and from multigrid. The resulting preconditioning technique requires the matrix only. A multilevel structure is obtained by using maximal independent sets for graph coarsening. A Schur complement approximation is constructed using a sequence of point Gaussian elimination steps. The resulting preconditioner has a transparent modular structure similar to the algorithmic structure of a multigrid V -cycle.

Key words. algebraic multigrid, cyclic reduction, Schur complement approximation, independent set orderings.

AMS subject classification. 65F10, 65N20.

1 Introduction

Multigrid methods are very efficient iterative solvers for the large systems of equations resulting from discretizing partial differential equations (cf. [18], [34] and the references therein). An important principle of multigrid is that a basic iterative method which yields appropriate local corrections is applied on a hierarchy of discretizations with different characteristic mesh sizes. This multilevel structure is of main importance for the efficiency of multigrid.

Another class of efficient iterative solvers consists of Krylov subspace methods combined with ILU preconditioning (cf. [8], [27] and the references therein). These methods only need the matrix and are in general easier to implement than multigrid methods. Also the Krylov subspace methods are better suitable for a "black-box" approach. On the other hand, for discretized partial differential equations the Krylov methods with ILU preconditioning are often less efficient than multigrid methods.

In the multigrid field there have been developed methods which have a multilevel structure but require only the matrix of the linear system. These are called *algebraic* multigrid methods. Approaches towards algebraic multigrid are presented in [7], [13], [16], [26], [33]. In all these methods one tries to mimic the multigrid principle. First one introduces a "reasonable" coarse "grid" space. Then a prolongation operator is chosen and for the restriction

one usually takes the adjoint of the prolongation. The operator on the coarse grid space is defined by a Galerkin approach. With these components, a standard multigrid approach (smoothing + coarse grid correction) is applied. These algebraic multigrid methods can be used in situations where a grid (hierarchy) is not available. Also these methods can be used for developing black-box solvers.

Recently there have been developed ILU type of preconditioners with a multilevel structure, cf. [5], [6], [23], [28], [29]. The multilevel structure is induced by a level wise numbering of the unknowns.

Recently, in [3], [4], [25] a few new hybrid methods have been presented, which use ideas both from ILU (incomplete Gaussian elimination) and from multigrid. In [3] a multigraph variant of the well-known HBMG (cf. [1]) is presented, based on the interpretation of the HBMG as an incomplete factorization method. In this method a recursive definition (typical for multigrid methods) and the concept of levels are avoided. The method in [4] is based on an incomplete Gaussian elimination process using levels, combined with smoothing in *all* the unknowns on each level (i.e. an algebraic variant of classical multigrid). A multilevel incomplete Gaussian elimination process, with an algorithmic structure which very similar to the structure of the HBMG (i.e. using levels and with smoothing on the newly added nodes only) is presented in [25]. In the present paper we reconsider the method in [25]. The presentation of the method differs from the presentation used in [25] in a few important aspects. In particular, the technique for Schur complement approximation, which is crucial for the efficiency of the method, is put in a more general setting. In this more general setting it is possible, at least for the two-level case, to prove interesting properties with respect to stability and approximation quality. Such theoretical results are not given in [25]. This theoretical background gives a further explanation of why this multilevel incomplete Gaussian elimination technique might result in an efficient *and robust* preconditioner.

The preconditioner that we present in this paper is based on the recursive application of a two-level method, as in cyclic reduction or in a multigrid V-cycle method. For the definition of a two level structure we use two important concepts: a reduced graph and a maximal independent set. For a given matrix graph $G_A(V, E)$ (V : vertices; E : edges) the reduced graph $G_A(V, E_s)$, with $E_s \subset E$, is obtained by deleting all "weak" edges in the given graph. Such a graph reduction is motivated by a multigrid heuristic (cf. [26], [33]): if a simple (point) smoother is used then, for enhancing robustness, one should coarsen in the direction of the "strong" connections. A red-black ("fine-coarse" in multigrid) partitioning of the vertex set V is constructed by computing a maximal independent set M of the *reduced* graph $G_A(V, E_s)$. Related coarsening techniques using maximal independent sets are presented in [9], [10], [14], [28], [25]. We note that in [25] the set of coarse graph vertices is given by M , whereas in the present paper the set of coarse graph vertices is equal to $V \setminus M$. In experiments we observed that for the efficiency of the preconditioner this difference is of minor importance. However, for the choice that is used in this paper the resulting preconditioner appears to be easier to analyze. The red-black partitioning

yields a corresponding block-representation of the given matrix \mathbf{A} :

$$\mathbf{PAP}^T = \begin{bmatrix} \mathbf{A}_{bb} & \mathbf{A}_{br} \\ \mathbf{A}_{rb} & \mathbf{A}_{rr} \end{bmatrix}, \quad (1)$$

with \mathbf{P} a suitable permutation matrix. The construction of the red-black partitioning is such that, under reasonable assumptions on \mathbf{A} , the \mathbf{A}_{rr} block is guaranteed to be strongly diagonally dominant. Hence, the systems with matrix \mathbf{A}_{rr} which occur in the (approximate) block UL-decomposition (cf. (4)) can be solved accurately with low costs, using a basic iterative solver. A main topic in this paper is the construction of a reasonable approximation $\tilde{\mathbf{S}}_{bb}$ of the Schur complement $\mathbf{S}_{bb} := \mathbf{A}_{bb} - \mathbf{A}_{br}\mathbf{A}_{rr}^{-1}\mathbf{A}_{rb}$. This approximation is obtained by replacing the *block* Gaussian elimination which results in the Schur complement (cf. (4)) by a sequence of *point* Gaussian elimination steps. We will prove some interesting stability and approximation properties of this Schur complement approximation. We also give a rather detailed presentation of how the preconditioner can be implemented. We will explain that if one starts with a particular implementation of the classical cyclic reduction method for a tridiagonal matrix then an implementation of the approximate cyclic reduction preconditioner can be obtained with only little additional effort.

The rest of this paper is organized as follows. In Section 2 we recall the classical cyclic reduction method for a tridiagonal matrix. In Section 3 we discuss how one can generalize this simple cyclic reduction technique such that it is applicable in a much more general setting. For this we present and analyze a general red-black partitioning method and a Schur complement approximation technique. The presentation and analysis is done in a linear algebra framework. In Section 4 we apply the general results of Section 3 to a few typical examples from the field of discretized partial differential equations. In Section 5 we present the approximate cyclic reduction preconditioner and in Section 6 we discuss some implementation issues related to this preconditioner. Finally, in Section 7 we present results of a few numerical experiments.

2 Cyclic reduction for a tridiagonal matrix

We recall the classical method of cyclic reduction. This method can be used, for example, for solving a linear system with a tridiagonal matrix or with a special block tridiagonal matrix (cf. [17], [20], [30], [32]). We explain the cyclic reduction principle by considering an $n \times n$ linear system with a tridiagonal matrix:

$$\mathbf{Ax} = \mathbf{b}, \quad \mathbf{A} = \begin{bmatrix} a_1 & b_1 & & & \\ c_1 & a_2 & b_2 & & \emptyset \\ & \ddots & \ddots & \ddots & \\ \emptyset & & \ddots & \ddots & b_{n-1} \\ & & & c_{n-1} & a_n \end{bmatrix}, \quad a_i \neq 0 \quad \text{for all } i. \quad (2)$$

Reordering the unknowns based on an obvious red-black (or "odd-even") structure results in a permuted system with a matrix of the form

$$\mathbf{PAP}^T = \begin{bmatrix} \mathbf{A}_{bb} & \mathbf{A}_{br} \\ \mathbf{A}_{rb} & \mathbf{A}_{rr} \end{bmatrix}, \quad (3)$$

in which $[\mathbf{A}_{bb} \ \mathbf{A}_{br}]$ represents the equations in the unknowns with a black label and $[\mathbf{A}_{rb} \ \mathbf{A}_{rr}]$ represents the equations in the unknowns with a red label. Note that, because \mathbf{A} is tridiagonal, the diagonal blocks $\mathbf{A}_{bb}, \mathbf{A}_{rr}$ are diagonal matrices. Gaussian elimination in the red points results in a reduced system with dimension (approximately) $\frac{1}{2}n$. In matrix notation this corresponds to block UL-decomposition:

$$\mathbf{PAP}^T = \begin{bmatrix} \mathbf{I} & \mathbf{A}_{br}\mathbf{A}_{rr}^{-1} \\ \emptyset & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{S}_{bb} & \emptyset \\ \mathbf{A}_{rb} & \mathbf{A}_{rr} \end{bmatrix}, \quad \mathbf{S}_{bb} := \mathbf{A}_{bb} - \mathbf{A}_{br}\mathbf{A}_{rr}^{-1}\mathbf{A}_{rb}. \quad (4)$$

The reduced system has a matrix \mathbf{S}_{bb} (Schur complement) which is tridiagonal, and thus the same approach can be applied to \mathbf{S}_{bb} . So the basic cyclic reduction idea is to reduce significantly the dimension of the problem repeatedly until one has a relatively small problem that can be solved easily. This small system is then solved and the previously eliminated (red) unknowns are found by a simple back-substitution process. Note that cyclic reduction is equivalent to Gaussian elimination applied to a permuted system of equations and that different implementations are possible (cf. [17], [32]).

When solving a system as in (2) with cyclic reduction, one usually adapts the righthand side in the reduction phase. For example, in the first reduction step the original system is transformed to

$$\begin{bmatrix} \mathbf{S}_{bb} & \emptyset \\ \mathbf{A}_{rb} & \mathbf{A}_{rr} \end{bmatrix} \mathbf{P}\mathbf{x} = \begin{bmatrix} \mathbf{I} & -\mathbf{A}_{br}\mathbf{A}_{rr}^{-1} \\ \emptyset & \mathbf{I} \end{bmatrix} \mathbf{P}\mathbf{b}. \quad (5)$$

In such a situation we do not need to store the coefficients of \mathbf{A}_{br} . In our approach (cf. Section 5), however, we will need both the upper and the lower triangular part of the UL-decomposition (as in ILU preconditioners). Thus we consider a cyclic reduction algorithm in which the block UL-decomposition as in (4) is computed. This UL-decomposition is then used to solve the system with a backward-forward elimination process. For the generalization of this simple cyclic reduction method to linear systems with general (non tridiagonal) sparse matrices we need a technique which yields reasonable sparse Schur complement approximations. Such a technique will be presented in the next section.

3 Schur complement approximation

We consider large sparse matrices which typically arise from discretization of partial differential equations. In this section we present and analyze a technique for constructing Schur

complement approximations of such matrices. This technique is used in the preconditioner that is presented in Section 5.

Let \mathbf{A} be a given regular $n \times n$ -matrix (cf. examples in Section 7). For the analysis below we assume stability of \mathbf{A} in the sense that

$$\mathbf{A} \text{ is a weakly diagonally dominant M-matrix,} \quad (6)$$

i.e. \mathbf{A} is an M-matrix and $\sum_{j \neq i} |a_{ij}| \leq a_{ii}$ for all i . The matrix \mathbf{A} induces an *ordered directed graph* $G_A(V, E)$, consisting of an ordered set of vertices $V = \{1, 2, \dots, n\}$ and a set E of ordered pairs of vertices called *arcs*. This set E consists of all pairs (i, j) for which $a_{ij} \neq 0$. A directed graph will also be called a *digraph*.

We briefly recall a few notions from graph theory. If (i, j) is an element of E then i is said to be *adjacent to* j and j is said to be *adjacent from* i . Two vertices $i \neq j$ are said to be *independent* if $(i, j) \notin E$ and $(j, i) \notin E$. A subset M of V is called an independent set if every two vertices in M are independent. M is called a *maximal independent set* of vertices if M is independent but no proper superset of M is independent. Note that a maximal independent set is in general not unique. For a vertex $i \in V$, its *neighbourhood* $N(i)$ is defined by $N(i) = \{j \in V \mid j \neq i \text{ and } (i, j) \in E\}$. For $i \in V$ its *degree*, $deg(i)$, is the number of elements in the neighbourhood of i , that is, $deg(i) = |N(i)|$. A vertex i is called an *isolated* vertex if $deg(i) = 0$. Note that an isolated vertex can be adjacent *from* other vertices in V . By $\Delta(G_A)$ we denote the maximum degree, i.e. $\Delta(G_A) = \max\{deg(v) \mid v \in V\}$.

We will construct a red-black partitioning of the vertex set V :

$$V = V_r \cup V_b \quad , \quad V_r \cap V_b = \{\emptyset\} \quad ,$$

which then induces a block representation of \mathbf{A} (cf. (3)):

$$\mathbf{PAP}^T = \begin{bmatrix} \mathbf{A}_{bb} & \mathbf{A}_{br} \\ \mathbf{A}_{rb} & \mathbf{A}_{rr} \end{bmatrix} \quad , \quad (7)$$

with Schur complement $\mathbf{S}_{bb} = (\mathbf{PAP}^T)/\mathbf{A}_{rr} = \mathbf{A}_{bb} - \mathbf{A}_{br}\mathbf{A}_{rr}^{-1}\mathbf{A}_{rb}$. Our goal is to construct a partitioning and a Schur complement approximation $\tilde{\mathbf{S}}_{bb}$ of \mathbf{S}_{bb} such that:

- systems with matrix \mathbf{A}_{rr} can be solved with relatively low costs,
- $\tilde{\mathbf{S}}_{bb}$ is a stable matrix (e.g. M-matrix),
- $\tilde{\mathbf{S}}_{bb}$ has a (sparsity) structure comparable to that of \mathbf{A} ,
- $\text{cond}(\tilde{\mathbf{S}}_{bb}^{-1}\mathbf{S}_{bb})$ is "small".

In Section 3.1 we discuss the construction of the red-black vertex partitioning. In Section 3.2 we treat the Schur complement approximation.

3.1 Red-black partitioning

As in algebraic multigrid methods (cf. [26], [33]), for the graph coarsening we distinguish "strong" and "weak" arcs in the digraph. The underlying multigrid heuristic is that if

one uses simple (point) smoothers then, to enhance robustness, one should coarsen in the direction of the "strong" connections.

Every loop in E , i.e. an arc of the form (i, i) , is labeled strong. For every nonisolated vertex $i \in V$ an arc $(i, j) \in E$ with $j \neq i$ is labeled strong if for the corresponding matrix entry a_{ij} we have:

$$|a_{ij}| \geq \beta \max_{j \in N(i)} |a_{ij}| \quad , \quad (8)$$

with $0 \leq \beta < 1$ a given parameter. An arc is labeled weak if it is not strong. Note that for every nonisolated vertex i there is at least one strong arc (i, j) with $j \neq i$. Thus we obtain a partitioning $E = E_s \cup E_w$ of the arcs into strong (E_s) and weak (E_w) arcs. The directed graph consisting of the vertex set V and the set of strong arcs E_s is called the *reduced* digraph and is denoted by $G_A(V, E_s)$. Let M be a maximal independent set of this reduced digraph $G_A(V, E_s)$. In Section 6 we discuss a simple algorithm for computing M with low computational costs. A vertex $i \in V$ is assigned a red (black) label if $i \in M$ ($i \notin M$). The resulting red-black partitioning induces a block representation of \mathbf{A} as in (7). For the analysis of this (matrix) partitioning we introduce some further notation:

$$\mathbf{D}_{rr} = \text{diag}(\mathbf{A}_{rr}) \quad , \quad \mathbf{I}_r : \text{identity matrix of dimension } |M|, \quad (9)$$

$$\hat{\beta} = 0 \text{ if } E_w = \{\emptyset\} \quad , \quad \hat{\beta} = \sup_{(i,j) \in E_w} \frac{|a_{ij}|}{\max_{j \in N(i)} |a_{ij}|} \text{ otherwise.} \quad (10)$$

Note that, by definition, $0 \leq \hat{\beta} \leq \beta < 1$. The construction of the red-black partitioning is such that the submatrix \mathbf{A}_{rr} is strongly diagonally dominant. This is quantified in the following theorem:

Theorem 3.1 *We consider the red-black partitioning as described above and assume that (6) holds. Then the following holds:*

$$\|\mathbf{I}_r - \mathbf{D}_{rr}^{-1} \mathbf{A}_{rr}\|_\infty \leq \frac{(\Delta(G_A) - 1)\hat{\beta}}{(\Delta(G_A) - 1)\hat{\beta} + 1} \quad . \quad (11)$$

Proof . Let $V_r = \{v \in V \mid \text{label}(v) = \text{red}\}$, $V_b = V \setminus V_r$. We renumber the vertices in V such that the red vertices are numbered first. Hence V_r can be represented as $V_r = \{1, 2, \dots, |M|\}$. Let $\mathbf{1}$ be the $|M|$ -vector with all entries equal to 1. Then, with $\mathbf{y} := (\mathbf{I}_r - \mathbf{D}_{rr}^{-1} \mathbf{A}_{rr})\mathbf{1}$ we have

$$\|\mathbf{I}_r - \mathbf{D}_{rr}^{-1} \mathbf{A}_{rr}\|_\infty = \|\mathbf{y}\|_\infty \quad .$$

Consider an arbitrary entry y_k of \mathbf{y} ($k \in V_r$). If k corresponds to an isolated red vertex then $y_k = 0$. We assume that k corresponds to a nonisolated red vertex. Then, with $W_k := N(k) \cap V_r$:

$$|y_k| = \sum_{l \in W_k} \frac{|a_{kl}|}{a_{kk}}$$

$$\begin{aligned}
&\leq \frac{\sum_{l \in W_k} |a_{kl}|}{\sum_{l \in N(k)} |a_{kl}|} \quad (\text{use (6)}) \\
&= \frac{\sum_{l \in W_k} |a_{kl}|}{\sum_{l \in W_k} |a_{kl}| + \sum_{l \in (N(k) \cap V_b)} |a_{kl}|} .
\end{aligned} \tag{12}$$

We use the notation $m_k := \max\{ |a_{kl}| \mid l \in N(k) \}$. Note that for $k \in V_r$, $l \in W_k$ the arc (k, l) is weak. Because k corresponds to a nonisolated vertex, there is at least one strong arc (k, l) , $l \neq k$, and thus we have $|W_k| \leq \Delta(G_A) - 1$. Hence, using (10) we obtain

$$\sum_{l \in W_k} |a_{kl}| \leq (\Delta(G_A) - 1) \hat{\beta} m_k .$$

Using this in (12) yields

$$|y_k| \leq \frac{(\Delta(G_A) - 1) \hat{\beta} m_k}{(\Delta(G_A) - 1) \hat{\beta} m_k + m_k}$$

and proves the estimate (11). \square

Remark 3.2 Note that if we take $\beta = 0$ then the reduced digraph is equal to the original digraph and \mathbf{A}_{rr} is diagonal. Hence the Schur complement $\mathbf{S}_{bb} = \mathbf{A}_{bb} - \mathbf{A}_{br} \mathbf{A}_{rr}^{-1} \mathbf{A}_{rb}$ can be computed exactly. This may seem a favourable situation. However, it is well-known that in general such an exact elimination step yields a Schur complement with a significantly less favourable sparsity structure. Already after a few (one or two) recursive steps we obtain unacceptable (w.r.t. efficiency) fill-in, cf. [28]. In our method (cf. Section 7) we take $\beta \in (0.5, 1)$. Then, in general the reduced digraph contains significantly less arcs than the original one. Due to this the maximal independent set, i.e. the set of red vertices, is much larger than for the case $\beta = 0$. Moreover, the graph coarsening is done mainly in the "direction" of strong connections. This makes it possible (cf. Section 3.2) to construct fairly accurate Schur complement approximations with a sparsity structure which is comparable to that of \mathbf{A} .

3.2 Schur complement approximation

We consider \mathbf{A} such that (6) holds and apply the red-black partitioning of the previous section, resulting in

$$\mathbf{PAP}^T = \begin{bmatrix} \mathbf{A}_{bb} & \mathbf{A}_{br} \\ \mathbf{A}_{rb} & \mathbf{A}_{rr} \end{bmatrix} =: \mathbf{A}^{(0)} . \tag{13}$$

Define \mathbf{D}_{rr} of dimension $|M|$ as in (9) and let $\mathbf{1}$ be the $|M|$ -vector with all entries equal to 1. Let $\tilde{\mathbf{D}}_{rr}$ be the $|M| \times |M|$ diagonal matrix which satisfies

$$\tilde{\mathbf{D}}_{rr} \mathbf{1} = \mathbf{A}_{rr} \mathbf{1} . \tag{14}$$

The result of Theorem 3.1 implies that $\tilde{\mathbf{D}}_{rr}$ is nonsingular. We use the notation \mathbf{I}_r and \mathbf{I}_b for the $|M|$ -dimensional and $(n - |M|)$ -dimensional identity matrix, respectively. The exact

Schur complement $\mathbf{S}_{bb} = \mathbf{A}_{bb} - \mathbf{A}_{br}\mathbf{A}_{rr}^{-1}\mathbf{A}_{rb}$ is the result of a block-Gaussian elimination:

$$\begin{bmatrix} \mathbf{I}_b & -\mathbf{A}_{br}\mathbf{A}_{rr}^{-1} \\ \emptyset & \mathbf{I}_r \end{bmatrix} \mathbf{A}^{(0)} = \begin{bmatrix} \mathbf{S}_{bb} & \emptyset \\ \mathbf{A}_{rb} & \mathbf{A}_{rr} \end{bmatrix}. \quad (15)$$

Note that for any $(n - |M|) \times |M|$ real matrix \mathbf{B}_{br} a left transformation of the form

$$\begin{bmatrix} \mathbf{I}_b & -\mathbf{B}_{br} \\ \emptyset & \mathbf{I}_r \end{bmatrix} \mathbf{A}^{(0)} = \begin{bmatrix} \mathbf{A}_{bb} - \mathbf{B}_{br}\mathbf{A}_{rb} & \mathbf{A}_{br} - \mathbf{B}_{br}\mathbf{A}_{rr} \\ \mathbf{A}_{rb} & \mathbf{A}_{rr} \end{bmatrix} =: \hat{\mathbf{A}}$$

does not affect the Schur complement, i.e. $\mathbf{S}_{bb} = \mathbf{A}^{(0)}/\mathbf{A}_{rr} = \hat{\mathbf{A}}/\mathbf{A}_{rr}$. For the construction of a *sparse* approximation of the decomposition in (15) we consider a sequence of such left transformations with blocks \mathbf{B}_{br} of the form $\mathbf{B}_{br} = \mathbf{A}_{br}\mathbf{E}_{rr}$, with \mathbf{E}_{rr} diagonal. An obvious choice is $\mathbf{E}_{rr} = \mathbf{D}_{rr}^{-1}$. For consistency reasons (discussed below) we also use $\mathbf{E}_{rr} = \tilde{\mathbf{D}}_{rr}^{-1}$. Left transformations using these choices for \mathbf{E}_{rr} can be considered as *point*-Gaussian type of elimination steps. More precisely, for $k \in \mathbb{N}$ we define sequences

$$\mathbf{A}^{(k)} = \begin{bmatrix} \mathbf{A}_{bb}^{(k)} & \mathbf{A}_{br}^{(k)} \\ \mathbf{A}_{rb} & \mathbf{A}_{rr} \end{bmatrix}, \quad \tilde{\mathbf{A}}^{(k)} = \begin{bmatrix} \tilde{\mathbf{A}}_{bb}^{(k)} & \tilde{\mathbf{A}}_{br}^{(k)} \\ \mathbf{A}_{rb} & \mathbf{A}_{rr} \end{bmatrix},$$

as follows, with $\mathbf{A}^{(0)}$ as in (13) :

$$\mathbf{A}^{(k)} = \begin{bmatrix} \mathbf{I}_b & -\mathbf{A}_{br}^{(k-1)}\mathbf{D}_{rr}^{-1} \\ \emptyset & \mathbf{I}_r \end{bmatrix} \mathbf{A}^{(k-1)} \quad \text{for } k \geq 1, \quad (16)$$

$$\tilde{\mathbf{A}}^{(k)} = \begin{bmatrix} \mathbf{I}_b & -\mathbf{A}_{br}^{(k-1)}\tilde{\mathbf{D}}_{rr}^{-1} \\ \emptyset & \mathbf{I}_r \end{bmatrix} \mathbf{A}^{(k-1)} \quad \text{for } k \geq 1. \quad (17)$$

For all k the Schur complements of $\mathbf{A}^{(k)}$ and of $\tilde{\mathbf{A}}^{(k)}$ are equal to \mathbf{S}_{bb} : $\mathbf{A}^{(k)}/\mathbf{A}_{rr} = \tilde{\mathbf{A}}^{(k)}/\mathbf{A}_{rr} = \mathbf{S}_{bb}$. The definitions in (16), (17) yield:

$$\mathbf{A}_{br}^{(k)} = \mathbf{A}_{br}(\mathbf{I}_r - \mathbf{D}_{rr}^{-1}\mathbf{A}_{rr})^k, \quad (18)$$

$$\tilde{\mathbf{A}}_{br}^{(k)} = \mathbf{A}_{br}(\mathbf{I}_r - \mathbf{D}_{rr}^{-1}\mathbf{A}_{rr})^{k-1}(\mathbf{I}_r - \tilde{\mathbf{D}}_{rr}^{-1}\mathbf{A}_{rr}). \quad (19)$$

The result of Theorem 3.1 implies that the blocks $\mathbf{A}_{br}^{(k)}$ and $\tilde{\mathbf{A}}_{br}^{(k)}$ are "small" for k sufficiently large and that, for k large enough, the diagonal blocks $\mathbf{A}_{bb}^{(k)}$ and $\tilde{\mathbf{A}}_{bb}^{(k)}$ might be reasonable approximations of \mathbf{S}_{bb} . In the remainder of this section we analyze these Schur complement approximations $\mathbf{A}_{bb}^{(k)}$ and $\tilde{\mathbf{A}}_{bb}^{(k)}$. In our preconditioner we will use $\tilde{\mathbf{A}}_{bb}^{(2)}$ as an approximation for the Schur complement \mathbf{S}_{bb} (cf. Remark 3.5 and Section 5). For the analysis we consider arbitrary $k \geq 1$.

Theorem 3.3 *We consider the red-black partitioning as described in Section 3.1 and assume (6). For $\mathbf{A}_{bb}^{(k)}$ defined in (16) the following holds:*

$$\mathbf{A}_{bb}^{(k)} \text{ is a weakly diagonally dominant M-matrix for all } k, \quad (20)$$

$$\mathbf{S}_{bb} = \mathbf{A}_{bb}^{(k)} - \mathbf{R}^{(k)} \text{ is a regular splitting for all } k, \quad (21)$$

$$\lim_{k \rightarrow \infty} \mathbf{A}_{bb}^{(k)} = \mathbf{S}_{bb}. \quad (22)$$

Proof. From (6) it follows that \mathbf{S}_{bb} is a weakly diagonally dominant M-matrix and that $\mathbf{A}_{rb} \leq \mathbf{0}$, $\mathbf{A}_{br} \leq \mathbf{0}$ (componentwise inequalities). We introduce $\mathbf{\Delta}_{rr} := \mathbf{I}_r - \mathbf{D}_{rr}^{-1} \mathbf{A}_{rr}$ and note that $\mathbf{\Delta}_{rr} \geq \mathbf{0}$ and, due to Theorem 3.1, $\rho(\mathbf{\Delta}_{rr}) < 1$. Hence we have the representation

$$\mathbf{S}_{bb} = \mathbf{A}_{bb} - \mathbf{A}_{br} \sum_{j=0}^{\infty} \mathbf{\Delta}_{rr}^j \mathbf{D}_{rr}^{-1} \mathbf{A}_{rb}.$$

Using induction and the result in (18) one easily obtains the identity

$$\mathbf{A}_{bb}^{(k)} = \mathbf{A}_{bb} - \mathbf{A}_{br} \sum_{j=0}^{k-1} \mathbf{\Delta}_{rr}^j \mathbf{D}_{rr}^{-1} \mathbf{A}_{rb}. \quad (23)$$

Hence:

$$\mathbf{S}_{bb} = \mathbf{A}_{bb}^{(k)} - \mathbf{R}^{(k)}, \quad \mathbf{R}^{(k)} := \mathbf{A}_{br} \sum_{j=k}^{\infty} \mathbf{\Delta}_{rr}^j \mathbf{D}_{rr}^{-1} \mathbf{A}_{rb}. \quad (24)$$

Note that $\mathbf{R}^{(k)} \geq \mathbf{0}$ and thus $\mathbf{A}_{bb}^{(k)} \geq \mathbf{S}_{bb}$. Since in \mathbf{A}_{bb} all off-diagonal entries are nonpositive and $\mathbf{A}_{br} \sum_{j=0}^{k-1} \mathbf{\Delta}_{rr}^j \mathbf{D}_{rr}^{-1} \mathbf{A}_{rb} \geq \mathbf{0}$ it follows from (23) that all off-diagonal entries of $\mathbf{A}_{bb}^{(k)}$ are nonpositive. We conclude that $\mathbf{A}_{bb}^{(k)}$ is an M-matrix and that the splitting in (24) is regular. From $\mathbf{A}_{bb}^{(k)} \geq \mathbf{S}_{bb}$ and $\mathbf{S}_{bb}(1, 1, \dots, 1)^T \geq \mathbf{0}$ it follows that $\mathbf{A}_{bb}^{(k)}$ is weakly diagonally dominant. Finally, from (24) we obtain that $\lim_{k \rightarrow \infty} \mathbf{R}^{(k)} = \mathbf{0}$ and thus $\lim_{k \rightarrow \infty} \mathbf{A}_{bb}^{(k)} = \mathbf{S}_{bb}$. \square

We conclude that $\mathbf{A}_{bb}^{(k)}$ is a stable approximation of \mathbf{S}_{bb} with $\rho(\mathbf{I}_b - (\mathbf{A}_{bb}^{(k)})^{-1} \mathbf{S}_{bb}) < 1$. However, in typical examples from the pde field one observes that the convergence in (22) is very slow on a certain subspace. In our applications this is a subspace corresponding to "smooth" gridfunctions. In other words, for low values of k , the approximation of $\mathbf{S}_{bb} \mathbf{v}$ by $\mathbf{A}_{bb}^{(k)} \mathbf{v}$ is very poor if \mathbf{v} corresponds to such a smooth gridfunction. This is similar to the phenomenon which causes the slow convergence of basic iterative methods (cf. [19]). In practice we should not use large values of k , since for increasing k the approximation $\mathbf{A}_{bb}^{(k)}$ suffers from serious fill-in.

To improve the approximation quality for low values of k we use a point-Gaussian elimination with matrix $\tilde{\mathbf{D}}_{rr}$ which satisfies the consistency condition (14). This motivates the process in (17) which results in Schur complement approximations $\tilde{\mathbf{A}}_{bb}^{(k)}$. Properties of these approximations are given in Theorem 3.4.

Theorem 3.4 *We consider the red-black partitioning as described in Section 3.1 and assume (6). For $\tilde{\mathbf{A}}_{bb}^{(k)}$ defined in (17) the following holds:*

$$\tilde{\mathbf{A}}_{bb}^{(k)} \text{ has only nonpositive off-diagonal entries for all } k, \quad (25)$$

$$\tilde{\mathbf{A}}_{bb}^{(k)} \text{ is weakly diagonally dominant for all } k, \quad (26)$$

$$\text{if } \tilde{\mathbf{A}}_{bb}^{(k)} \text{ is nonsingular then it is an M-matrix,} \quad (27)$$

$$\tilde{\mathbf{A}}_{bb}^{(k)} \text{ is an M-matrix for } k \text{ sufficiently large,} \quad (28)$$

$$\lim_{k \rightarrow \infty} \tilde{\mathbf{A}}_{bb}^{(k)} = \mathbf{S}_{bb} . \quad (29)$$

$$\text{if } \mathbf{w} \text{ satisfies } \mathbf{A}_{rb}\mathbf{w} = -\mathbf{A}_{rr}\mathbf{1} , \text{ then } \tilde{\mathbf{A}}_{bb}^{(k)}\mathbf{w} = \mathbf{S}_{bb}\mathbf{w} \text{ for all } k , \quad (30)$$

$$\mathbf{I}_b - \tilde{\mathbf{A}}_{bb}^{(k)}\mathbf{S}_{bb}^{-1} = \mathbf{A}_{br}(\mathbf{I}_r - \mathbf{D}_{rr}^{-1}\mathbf{A}_{rr})^{k-1}\tilde{\mathbf{D}}_{rr}^{-1} \begin{bmatrix} \mathbf{A}_{rb} & \tilde{\mathbf{D}}_{rr} \end{bmatrix} (\mathbf{A}^{(0)})^{-1} \begin{bmatrix} \mathbf{I}_b \\ \emptyset \end{bmatrix} \quad (31)$$

Proof. By definition we have

$$\tilde{\mathbf{A}}_{bb}^{(k)} = \mathbf{A}_{bb}^{(k-1)} - \mathbf{A}_{br}^{(k-1)}\tilde{\mathbf{D}}_{rr}^{-1}\mathbf{A}_{rb} . \quad (32)$$

Note that $\mathbf{A}_{bb}^{(k-1)}$ is an M-matrix (Theorem 3.3), $\mathbf{A}_{br}^{(k-1)} \leq \mathbf{0}$ (cf. (18)), $\tilde{\mathbf{D}}_{rr}^{-1} \geq \mathbf{0}$ and $\mathbf{A}_{rb} \leq \mathbf{0}$. Hence the result in (25) holds.

We introduce the n -vector $\mathbf{1}_n = (1, 1, \dots, 1)^T$ and its red-black partitioning $\mathbf{1}_n = \begin{bmatrix} \mathbf{1}_b \\ \mathbf{1}_r \end{bmatrix}$.

From (16) it follows that

$$\mathbf{A}_{bb}^{(k)}\mathbf{1}_b + \mathbf{A}_{br}^{(k)}\mathbf{1}_r = \mathbf{A}_{bb}^{(k-1)}\mathbf{1}_b + \mathbf{A}_{br}^{(k-1)}\mathbf{1}_r - \mathbf{A}_{br}^{(k-1)}\tilde{\mathbf{D}}_{rr}^{-1}(\mathbf{A}_{rb}\mathbf{1}_b + \mathbf{A}_{rr}\mathbf{1}_r) ,$$

which yields $\mathbf{A}_{bb}^{(k)}\mathbf{1}_b + \mathbf{A}_{br}^{(k)}\mathbf{1}_r \geq \mathbf{0}$ for all k . The consistency condition (14) results in $\mathbf{A}_{rb}\mathbf{1}_b + \tilde{\mathbf{D}}_{rr}\mathbf{1}_r = \mathbf{A}_{rb}\mathbf{1}_b + \mathbf{A}_{rr}\mathbf{1}_r \geq \mathbf{0}$ and thus $\tilde{\mathbf{D}}_{rr}^{-1}\mathbf{A}_{rb}\mathbf{1}_b \geq -\mathbf{1}_r$. Using this in (32) we obtain the result in (26):

$$\tilde{\mathbf{A}}_{bb}^{(k)}\mathbf{1}_b = \mathbf{A}_{bb}^{(k-1)}\mathbf{1}_b - \mathbf{A}_{br}^{(k-1)}\tilde{\mathbf{D}}_{rr}^{-1}\mathbf{A}_{rb}\mathbf{1}_b \geq \mathbf{A}_{bb}^{(k-1)}\mathbf{1}_b + \mathbf{A}_{br}^{(k-1)}\mathbf{1}_r \geq \mathbf{0} .$$

For the proof of (27) we use one of the many characterizations of M-matrices (cf. Theorem 5.1 in [15]): if \mathbf{B} is a real square matrix with only nonpositive off-diagonal entries, then \mathbf{B} is an M-matrix if and only if every real eigenvalue of \mathbf{B} is positive. Using (26) and Gershgorin's theorem we conclude that every real eigenvalue of $\tilde{\mathbf{A}}_{bb}^{(k)}$ is nonnegative. If $\tilde{\mathbf{A}}_{bb}^{(k)}$ is nonsingular then 0 cannot be an eigenvalue. Hence all real eigenvalues are positive and (27) holds.

From (31) and Theorem 3.1 we obtain that $\rho(\mathbf{I}_b - \tilde{\mathbf{A}}_{bb}^{(k)}\mathbf{S}_{bb}^{-1}) < 1$ for k sufficiently large and thus $\tilde{\mathbf{A}}_{bb}^{(k)}$ is nonsingular for k sufficiently large. Using (27) we obtain the result in (28).

From (32), (24) and (18) we obtain:

$$\tilde{\mathbf{A}}_{bb}^{(k)} - \mathbf{S}_{bb} = \mathbf{A}_{bb}^{(k-1)} - \mathbf{S}_{bb} - \mathbf{A}_{br}^{(k-1)}\tilde{\mathbf{D}}_{rr}^{-1}\mathbf{A}_{rb}$$

$$\begin{aligned}
&= \mathbf{A}_{br} \sum_{j=k-1}^{\infty} (\mathbf{I}_r - \mathbf{D}_{rr}^{-1} \mathbf{A}_{rr})^j \mathbf{D}_{rr}^{-1} \mathbf{A}_{rb} - \mathbf{A}_{br}^{(k-1)} \tilde{\mathbf{D}}_{rr}^{-1} \mathbf{A}_{rb} \\
&= \mathbf{A}_{br} (\mathbf{I}_r - \mathbf{D}_{rr}^{-1} \mathbf{A}_{rr})^{k-1} \mathbf{A}_{rr}^{-1} \mathbf{A}_{rb} - \mathbf{A}_{br}^{(k-1)} \tilde{\mathbf{D}}_{rr}^{-1} \mathbf{A}_{rb} \\
&= \mathbf{A}_{br} (\mathbf{I}_r - \mathbf{D}_{rr}^{-1} \mathbf{A}_{rr})^{k-1} (\mathbf{I}_r - \tilde{\mathbf{D}}_{rr}^{-1} \mathbf{A}_{rr}) \mathbf{A}_{rr}^{-1} \mathbf{A}_{rb} .
\end{aligned} \tag{33}$$

Combination of the results in (33) and in Theorem 3.1 yields the result in (29).
For \mathbf{w} as in (30) we have

$$(\mathbf{I}_r - \tilde{\mathbf{D}}_{rr}^{-1} \mathbf{A}_{rr}) \mathbf{A}_{rr}^{-1} \mathbf{A}_{rb} \mathbf{w} = -(\mathbf{I}_r - \tilde{\mathbf{D}}_{rr}^{-1} \mathbf{A}_{rr}) \mathbf{1} = \mathbf{0}$$

and thus, using (33), we obtain the result in (30).

From (33) we obtain

$$\mathbf{I}_b - \tilde{\mathbf{A}}_{bb}^{(k)} \mathbf{S}_{bb}^{-1} = -\mathbf{A}_{br} (\mathbf{I}_r - \mathbf{D}_{rr}^{-1} \mathbf{A}_{rr})^{k-1} \tilde{\mathbf{D}}_{rr}^{-1} (\tilde{\mathbf{D}}_{rr} - \mathbf{A}_{rr}) \mathbf{A}_{rr}^{-1} \mathbf{A}_{rb} \mathbf{S}_{bb}^{-1} . \tag{34}$$

From the identity

$$(\mathbf{A}^{(0)})^{-1} = \begin{bmatrix} \mathbf{S}_{bb}^{-1} & \emptyset \\ -\mathbf{A}_{rr}^{-1} \mathbf{A}_{rb} \mathbf{S}_{bb}^{-1} & \mathbf{A}_{rr}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I}_b & -\mathbf{A}_{br} \mathbf{A}_{rr}^{-1} \\ \emptyset & \mathbf{I}_r \end{bmatrix}$$

we obtain

$$[\emptyset \quad \mathbf{I}_r] (\mathbf{A}^{(0)})^{-1} \begin{bmatrix} \mathbf{I}_b \\ \emptyset \end{bmatrix} = -\mathbf{A}_{rr}^{-1} \mathbf{A}_{rb} \mathbf{S}_{bb}^{-1} .$$

Hence

$$\begin{aligned}
-(\tilde{\mathbf{D}}_{rr} - \mathbf{A}_{rr}) \mathbf{A}_{rr}^{-1} \mathbf{A}_{rb} \mathbf{S}_{bb}^{-1} &= (\tilde{\mathbf{D}}_{rr} - \mathbf{A}_{rr}) [\emptyset \quad \mathbf{I}_r] (\mathbf{A}^{(0)})^{-1} \begin{bmatrix} \mathbf{I}_b \\ \emptyset \end{bmatrix} \\
&= [\mathbf{A}_{br} \quad \tilde{\mathbf{D}}_{rr}] (\mathbf{A}^{(0)})^{-1} \begin{bmatrix} \mathbf{I}_b \\ \emptyset \end{bmatrix} - [\emptyset \quad \mathbf{I}_r] \mathbf{A}^{(0)} (\mathbf{A}^{(0)})^{-1} \begin{bmatrix} \mathbf{I}_b \\ \emptyset \end{bmatrix} \\
&= [\mathbf{A}_{br} \quad \tilde{\mathbf{D}}_{rr}] (\mathbf{A}^{(0)})^{-1} \begin{bmatrix} \mathbf{I}_b \\ \emptyset \end{bmatrix} .
\end{aligned}$$

Using this in (34) we obtain the result (31). \square

The results in (25)-(28) show that the Schur complement approximation $\tilde{\mathbf{A}}_{bb}^{(k)}$ has favourable stability properties, comparable to those of $\mathbf{A}_{bb}^{(k)}$ and of \mathbf{A} .

Remark 3.5 For $k = 2$ we can represent $\tilde{\mathbf{A}}_{bb}^{(k)}$ as

$$\tilde{\mathbf{A}}_{bb}^{(2)} = \begin{bmatrix} \mathbf{I}_b & -\mathbf{A}_{br} \mathbf{D}_{rr}^{-1} \end{bmatrix} \mathbf{A}^{(0)} \begin{bmatrix} \mathbf{I}_b \\ -\tilde{\mathbf{D}}_{rr}^{-1} \mathbf{A}_{rb} \end{bmatrix} , \tag{35}$$

which can be compared to the representation

$$\mathbf{S}_{bb} = \begin{bmatrix} \mathbf{I}_b & -\mathbf{A}_{br}\mathbf{A}_{rr}^{-1} \end{bmatrix} \mathbf{A}^{(0)} \begin{bmatrix} \mathbf{I}_b \\ * \end{bmatrix} = \begin{bmatrix} \mathbf{I}_b & * \end{bmatrix} \mathbf{A}^{(0)} \begin{bmatrix} \mathbf{I}_b \\ -\mathbf{A}_{rr}^{-1}\mathbf{A}_{rb} \end{bmatrix},$$

with $*$ arbitrary. Hence, in multigrid terminology, the coarse graph approximation $\tilde{\mathbf{A}}_{bb}^{(2)}$ of \mathbf{S}_{bb} is obtained using a Galerkin approach with matrix-dependent prolongation $\mathbf{p}_A = \begin{bmatrix} \mathbf{I}_b \\ -\tilde{\mathbf{D}}_{rr}^{-1}\mathbf{A}_{rb} \end{bmatrix}$ and restriction $\mathbf{r}_A = \begin{bmatrix} \mathbf{I}_b & -\mathbf{A}_{br}\mathbf{D}_{rr}^{-1} \end{bmatrix}$. Note that $\mathbf{p}_A\mathbf{w} = \begin{bmatrix} \mathbf{I}_b \\ -\mathbf{A}_{rr}^{-1}\mathbf{A}_{rb} \end{bmatrix} \mathbf{w}$ for \mathbf{w} as in (30). The resulting coarse graph approximation is stable (in the sense of Theorem 3.4) and satisfies the consistency condition as in (30). If we replace $\tilde{\mathbf{D}}_{rr}$ by \mathbf{D}_{rr} in (35) we obtain a representation for $\mathbf{A}_{bb}^{(2)}$. From numerical experiments with discretized partial differential equations it follows that the use of different diagonal approximations of \mathbf{A}_{rr} ($\tilde{\mathbf{D}}_{rr}$ and \mathbf{D}_{rr}) in the matrix-dependent prolongation and restriction is of main importance. If we use \mathbf{D}_{rr} in both \mathbf{p}_A and \mathbf{r}_A (i.e. $\mathbf{A}_{bb}^{(2)}$) then we obtain a stable approximation, however, the approximation is very poor on a subspace of smooth grid functions (lack of consistency). If we use $\tilde{\mathbf{D}}_{rr}$ in both \mathbf{p}_A and \mathbf{r}_A then for certain problems (e.g. convection-diffusion problems with strong convection) the approximation is poor due to instabilities. We note that the use of different diagonal approximations $\tilde{\mathbf{D}}_{rr} \neq \mathbf{D}_{rr}$ (hence $\mathbf{p}_A \neq \mathbf{r}_A$), which guarantees a consistency and stability property, has a drawback with respect to symmetry. If the original matrix is symmetric then the Schur complement approximations $\tilde{\mathbf{A}}_{bb}^{(k)}$ will be nonsymmetric.

Certain multigrid approaches are based on Schur complement approximation using suitable basis transformations (cf. [1], [2]). For $k = 2$ the coarse graph matrix $\tilde{\mathbf{A}}_{bb}^{(2)}$ is, in a natural way, related to the hierarchical basis transformation

$$\begin{bmatrix} \mathbf{I}_b & -\mathbf{A}_{br}\mathbf{D}_{rr}^{-1} \\ \emptyset & \mathbf{I}_r \end{bmatrix} \mathbf{A}^{(0)} \begin{bmatrix} \mathbf{I}_b & \emptyset \\ -\tilde{\mathbf{D}}_{rr}^{-1}\mathbf{A}_{rb} & \mathbf{I}_r \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{A}}_{bb}^{(2)} & -\mathbf{A}_{br}(\mathbf{I}_r - \mathbf{D}_{rr}^{-1}\mathbf{A}_{rr}) \\ (\mathbf{I}_r - \mathbf{A}_{rr}\tilde{\mathbf{D}}_{rr}^{-1})\mathbf{A}_{rb} & \mathbf{A}_{rr} \end{bmatrix}. \quad (36)$$

Note that this involves matrix-dependent basis transformations (as in [2]).

The multigrid convergence analysis of Hackbusch (cf. [18]) is based on the approximation property and the smoothing property. The approximation property is of main importance for a proper reduction by the multigrid method of smooth error components. In [12], [18] it is shown that this approximation property is closely related to a regularity property, which holds for a certain class of elliptic pde's. Below, in Theorem 3.6, we introduce a sort of algebraic regularity term.

For the formulation of Theorem 3.6 we first introduce norm notations. The number of red vertices is given by $|M|$ and the number of black vertices is given by $m := n - |M|$. On \mathbb{R}^n , $\mathbb{R}^{|M|}$ and \mathbb{R}^m we assume norms denoted by $\|\cdot\|$, $\|\cdot\|_r$ and $\|\cdot\|_b$, respectively. In the general setting of this section we do not specify these norms. We assume that these norms are compatible in the sense that

$$\|\mathbf{w}\|_r = \left\| \begin{bmatrix} \mathbf{0} \\ \mathbf{w} \end{bmatrix} \right\| \quad \text{for all } \mathbf{w} \in \mathbb{R}^{|M|}, \quad \|\mathbf{w}\|_b = \left\| \begin{bmatrix} \mathbf{w} \\ \mathbf{0} \end{bmatrix} \right\| \quad \text{for all } \mathbf{w} \in \mathbb{R}^m. \quad (37)$$

For ease of notation we drop the r, b in $\|\cdot\|_r, \|\cdot\|_b$, i.e. all three norms on $\mathbb{R}^n, \mathbb{R}^{|\mathcal{M}|}$ and \mathbb{R}^m are denoted by $\|\cdot\|$. We also use associated matrix norms denoted by $\|\cdot\|$.

Theorem 3.6 *We consider the red-black partitioning as described in Section 3.1 and assume (6). For $\tilde{\mathbf{A}}_{bb}^{(k)}$ defined in (17) the following holds:*

$$\|\mathbf{I}_b - \tilde{\mathbf{A}}_{bb}^{(k)} \mathbf{S}_{bb}^{-1}\| \leq \|\mathbf{A}_{br}(\mathbf{I}_r - \mathbf{D}_{rr}^{-1} \mathbf{A}_{rr})^{k-1} \tilde{\mathbf{D}}_{rr}^{-1}\| \|[\mathbf{A}_{rb} \quad \tilde{\mathbf{D}}_{rr}](\mathbf{A}^{(0)})^{-1}\|. \quad (38)$$

Proof. From (31) we obtain

$$\mathbf{I}_b - \tilde{\mathbf{A}}_{bb}^{(k)} \mathbf{S}_{bb}^{-1} = \mathbf{A}_{br}(\mathbf{I}_r - \mathbf{D}_{rr}^{-1} \mathbf{A}_{rr})^{k-1} \tilde{\mathbf{D}}_{rr}^{-1} [\mathbf{A}_{rb} \quad \tilde{\mathbf{D}}_{rr}](\mathbf{A}^{(0)})^{-1} \begin{bmatrix} \mathbf{I}_b \\ \emptyset \end{bmatrix}.$$

The assumption (37) yields $\left\| \begin{bmatrix} \mathbf{I}_b \\ \emptyset \end{bmatrix} \right\| = 1$. Combination of these results yields (38). \square

Remark 3.7 We briefly comment on a relationship to the discrete regularity theory in multigrid convergence analyses. In general, the first term in the righthand side of (38) can be controlled using linear algebra arguments only (cf. diagonal dominance result in Theorem 3.1). This is similar to the analysis of the smoothing property in multigrid convergence theory, which is also based on linear algebra arguments only (cf. [18]). For an interpretation of the term

$$\|[\mathbf{A}_{rb} \quad \tilde{\mathbf{D}}_{rr}](\mathbf{A}^{(0)})^{-1}\| \quad (39)$$

in the righthand side of (38) we define the seminorm $|w| := \|[\mathbf{A}_{rb} \quad \tilde{\mathbf{D}}_{rr}]w\|$ for $w \in \mathbb{R}^n$. Bounds for $\|(\mathbf{A}^{(0)})^{-1}\|$, i.e. $(\mathbf{A}^{(0)})^{-1} : (\mathbb{R}^n, \|\cdot\|) \rightarrow (\mathbb{R}^n, \|\cdot\|)$, correspond to the classical notion of stability. A bound for (39) is equivalent to a bound for

$$(\mathbf{A}^{(0)})^{-1} : (\mathbb{R}^n, \|\cdot\|) \rightarrow (\mathbb{R}^n, |\cdot|). \quad (40)$$

In our applications the matrix $[\mathbf{A}_{rb} \quad \tilde{\mathbf{D}}_{rr}]$ corresponds to a difference operator which is similar to the underlying differential operator and the norm $|\cdot|$ measures differences (smoothness). In multigrid convergence theory one can find results (cf. [18], Chapter 6) in which, for discretized elliptic boundary value problems, bounds for $(\mathbf{A}^{(0)})^{-1}$ as an operator between spaces with different smoothness properties (as in (40)) are derived. Such results can be considered as discrete counterparts of results in regularity theory for continuous elliptic boundary value problems, in which for an operator $L : H_0^1(\Omega) \rightarrow H^{-1}(\Omega)$ boundedness of $L^{-1} : H^{-1+s} \rightarrow H^{1+s}(\Omega) \cap H_0^1(\Omega)$ with $s > 0$ is analyzed. If, for example, L corresponds to the Laplacian on the unit square with zero Dirichlet boundary conditions, then the boundedness of $L^{-1} : L^2(\Omega) \rightarrow H^2(\Omega) \cap H_0^1(\Omega)$ is a classical regularity result. A corresponding discrete regularity result is presented in [18] Section 6.3.2.

Related to the robustness of our preconditioning technique (cf. experiments in Section 7) it is important to note that the seminorm $|\cdot|$ is strongly problem dependent. In Section 4 we will indicate that due to this, for certain interesting problems which have poor regularity in the classical multigrid sense, one can still expect a small bound for (40). We emphasize

that, in general, one can not expect to derive reasonable bounds for (40) using linear algebra arguments only. As in multigrid analyses, in this derivation we need that the matrix \mathbf{A} is the discrete analogon of a differential operator.

Remark 3.8 In our implementation we use $\tilde{\mathbf{A}}_{bb}^{(2)}$ as a Schur complement approximation. As explained above, this approximation has favourable stability and consistency properties. By construction this approximation is sparse. However, due to the occurrence of some fill-in, $\tilde{\mathbf{A}}_{bb}^{(2)}$ is in general less sparse than \mathbf{A} . Recursive application of the same technique may result in relatively dense Schur complement approximations on very coarse graphs. Our applications (pde's) are such that fill-in between two vertices which have a long mutual distance in the graph is very small compared to the corresponding diagonal entry. Hence we introduce a parameter $\text{MSIZE} \in \mathbb{N}$. Typically, $\text{MSIZE} \in (2\Delta(\mathbf{A}), 3\Delta(\mathbf{A}))$. If a certain vertex in the digraph of $\tilde{\mathbf{A}}_{bb}^{(2)}$ has degree $d > \text{MSIZE}$ then we modify the corresponding row of $\tilde{\mathbf{A}}_{bb}^{(2)}$ using a common lumping technique: we add the $d - \text{MSIZE}$ in absolute value smallest off-diagonal entries to the diagonal entry and then replace these off-diagonal entries by zeros. Using this modification we obtain Schur complement approximations (on all coarser graphs) for which the corresponding digraph has a maximum degree that is bounded by MSIZE .

Assume that (6) holds. Then for $\tilde{\mathbf{A}}_{bb}^{(2)}$ we have the (stability) properties (25), (26). The lumping technique that we use preserves these properties. In the generic case the Schur complement approximation after lumping is (still) nonsingular and then (cf. proof of (27)) this Schur complement approximation is a sparse weakly diagonally dominant M-matrix. Hence the method and the analysis of this section can be applied recursively.

4 Examples

In this section we analyze a few model problems from the pde field for which we can quantify the bound in Theorem 3.6 (cf. Remark 3.7) We consider a diffusion equation, an anisotropic diffusion equation and a convection-diffusion equation. In all three examples below we treat a standard finite difference discretization of a constant coefficient problem on a square domain $\Omega = (0, 1)^2$. We use a square mesh with mesh size denoted by h and assume periodic boundary conditions.

We analyze the bound (38) in the Euclidean norm $\|\cdot\|_2$, i.e. we consider the two terms

$$\|\mathbf{A}_{br}(\mathbf{I}_r - \mathbf{D}_{rr}^{-1}\mathbf{A}_{rr})^{k-1}\tilde{\mathbf{D}}_{rr}^{-1}\|_2 \quad (41)$$

and

$$\|[\mathbf{A}_{rb} \quad \tilde{\mathbf{D}}_{rr}](\mathbf{A}^{(0)})^{-1}\|_2. \quad (42)$$

The analysis of the first term is based on simple linear algebra arguments (e.g. Gershgorin theorem). For this analysis the restriction to constant coefficient problems with periodic boundary conditions is not relevant. A similar analysis can be applied for other problems (e.g. varying coefficients). For a simple treatment of the second term (42) the standard

Fourier analysis is applied. For this the restriction to constant coefficient problems with periodic boundary conditions is crucial.

We note that for the first example (Poisson equation) the case with homogeneous Dirichlet boundary conditions can be analyzed along the same lines. The results for this Dirichlet case are essentially the same as for the case with periodic boundary conditions.

For the parameter β used in the red-black partitioning (cf. (8)) we take $\beta = 0.7$.

Example 1 (Poisson equation). We consider the standard five-point discretization of the Poisson equation with stencil

$$[\mathbf{A}] = \frac{1}{h^2} \begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix}. \quad (43)$$

In this situation all edges in the digraph are labeled strong and the reduced graph is equal to the original graph. One possible maximal independent set is obtained from standard red-black coloring of the grid. Then \mathbf{A}_{rr} is diagonal and $\mathbf{D}_{rr} = \tilde{\mathbf{D}}_{rr} = \mathbf{A}_{rr}$. The Schur complement approximation is exact (i.e. $\mathbf{A}_{bb}^{(k)} = \tilde{\mathbf{A}}_{bb}^{(k)} = \mathbf{S}_{bb}$ for all k) and results in a Schur complement with stencil

$$[\mathbf{S}_{bb}] = \frac{1}{2h^2} \begin{bmatrix} -\frac{1}{2} & -1 & -\frac{1}{2} \\ -1 & 6 & -1 \\ -\frac{1}{2} & -1 & -\frac{1}{2} \end{bmatrix}. \quad (44)$$

We now define $\mathbf{A} := \mathbf{S}_{bb}$ and apply one further coarsening step to this Schur complement matrix. The edges in the digraph corresponding to the matrix entries $-\frac{1}{2}$ are labeled weak. Hence the reduced digraph has a structure which corresponds to the five-point stencil

$$\begin{bmatrix} * \\ * & * & * \\ * \end{bmatrix} \text{ and, as in the case of the five-point stencil in (43), one possible maximal}$$

independent set (of the reduced digraph) is obtained by standard red-black coloring of the grid. This then results in stencils

$$[\mathbf{A}_{rr}] = \frac{1}{2h^2} \begin{bmatrix} -\frac{1}{2} & & -\frac{1}{2} \\ & 6 & \\ -\frac{1}{2} & & -\frac{1}{2} \end{bmatrix}, \quad [\mathbf{A}_{rb}] = \frac{1}{2h^2} \begin{bmatrix} & -1 & \\ -1 & & -1 \\ & -1 & \end{bmatrix}, \quad [\mathbf{A}_{br}] = \frac{1}{2h^2} \begin{bmatrix} & -1 & \\ -1 & & -1 \\ & -1 & \end{bmatrix}.$$

Hence $\mathbf{D}_{rr} = \frac{6}{2h^2} \mathbf{I}_r$, $\tilde{\mathbf{D}}_{rr} = \frac{4}{2h^2} \mathbf{I}_r$.

We first consider the term (41). Using $\|\mathbf{A}_{br}\|_2^2 \leq \|\mathbf{A}_{br}\|_1 \|\mathbf{A}_{br}\|_\infty \leq \frac{16}{4h^4}$, $\|\tilde{\mathbf{D}}_{rr}\|_2 = \frac{2}{h^2}$ and a Gershgorin theorem, we obtain

$$\|\mathbf{A}_{br}(\mathbf{I}_r - \mathbf{D}_{rr}^{-1} \mathbf{A}_{rr})^{k-1} \tilde{\mathbf{D}}_{rr}^{-1}\|_2 \leq \|\mathbf{I}_r - \mathbf{D}_{rr}^{-1} \mathbf{A}_{rr}\|_2^{k-1} \leq \left(\frac{1}{3}\right)^{k-1}.$$

With respect to the term (42) we note that $[\mathbf{A}_{rb} \quad \tilde{\mathbf{D}}_{rr}](\mathbf{A}^{(0)})^{-1} = [\emptyset \quad \mathbf{I}_r] \tilde{\mathbf{A}}(\mathbf{A}^{(0)})^{-1}$, where $\tilde{\mathbf{A}}$ and $\mathbf{A}^{(0)}$ have stencils

$$[\tilde{\mathbf{A}}] = \frac{1}{2h^2} \begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix}, \quad [\mathbf{A}^{(0)}] = \frac{1}{2h^2} \begin{bmatrix} -\frac{1}{2} & -1 & -\frac{1}{2} \\ -1 & 6 & -1 \\ -\frac{1}{2} & -1 & -\frac{1}{2} \end{bmatrix}.$$

A straightforward Fourier analysis yields $\|[\mathbf{A}_{rb} \quad \tilde{\mathbf{D}}_{rr}](\mathbf{A}^{(0)})^{-1}\|_2 \leq 1$. Theorem 3.6 now yields

$$\|\mathbf{I}_b - \tilde{\mathbf{A}}_{bb}^{(k)} \mathbf{S}_{bb}^{-1}\|_2 \leq \left(\frac{1}{3}\right)^{k-1}.$$

We conclude that, already for $k = 2$, $\tilde{\mathbf{A}}_{bb}^{(k)}$ is a good preconditioner for \mathbf{S}_{bb} *uniformly* in h . One easily verifies that in this example a similar bound, which is independent of h , does not hold for $\mathbf{A}_{bb}^{(2)}$.

Example 2 (Anisotropic diffusion equation). We consider the five-point discretization of an anisotropic diffusion equation with stencil

$$[\mathbf{A}] = \frac{1}{h^2} \begin{bmatrix} & -\varepsilon & \\ -1 & 2 + 2\varepsilon & -1 \\ & -\varepsilon & \end{bmatrix}, \quad \text{with } 0 < \varepsilon < \frac{1}{2}. \quad (45)$$

The edges in the digraph corresponding to the matrix entries $-\varepsilon$ are labeled weak. Hence the reduced digraph has a structure which corresponds to the three-point stencil $[* \ * \ *]$. One possible maximal independent set (of the reduced digraph) is obtained by standard semi-coarsening (i.e. coarsening by a factor 2 in the x -direction only) of the grid. This then results in stencils

$$[\mathbf{A}_{rr}] = \frac{1}{h^2} \begin{bmatrix} -\varepsilon \\ 2 + 2\varepsilon \\ -\varepsilon \end{bmatrix}, \quad [\mathbf{A}_{rb}] = \frac{1}{h^2} \begin{bmatrix} -1 & \cdot & -1 \end{bmatrix}, \quad [\mathbf{A}_{br}] = \frac{1}{h^2} \begin{bmatrix} -1 & \cdot & -1 \end{bmatrix}.$$

Hence $\mathbf{D}_{rr} = \frac{2+2\varepsilon}{h^2} \mathbf{I}_r$, $\tilde{\mathbf{D}}_{rr} = \frac{2}{h^2} \mathbf{I}_r$. Along the same lines as in Example 1 we obtain for the term (41):

$$\|\mathbf{A}_{br}(\mathbf{I}_r - \mathbf{D}_{rr}^{-1} \mathbf{A}_{rr})^{k-1} \tilde{\mathbf{D}}_{rr}^{-1}\|_2 \leq \|\mathbf{I}_r - \mathbf{D}_{rr}^{-1} \mathbf{A}_{rr}\|_2^{k-1} \leq \left(\frac{\varepsilon}{\varepsilon + 1}\right)^{k-1}.$$

Related to the term (42) we note that

$$[\mathbf{A}_{rb} \quad \tilde{\mathbf{D}}_{rr}](\mathbf{A}^{(0)})^{-1} = [\emptyset \quad \mathbf{I}_r] \tilde{\mathbf{A}}(\mathbf{A}^{(0)})^{-1},$$

where $\tilde{\mathbf{A}}$ and $\mathbf{A}^{(0)}$ have stencils

$$[\tilde{\mathbf{A}}] = \frac{1}{h^2} \begin{bmatrix} & 0 & \\ -1 & 2 & -1 \\ & 0 & \end{bmatrix}, \quad [\mathbf{A}^{(0)}] = \frac{1}{h^2} \begin{bmatrix} & -\varepsilon & \\ -1 & 2 + 2\varepsilon & -1 \\ & -\varepsilon & \end{bmatrix}.$$

Note that the algebraic regularity is measured using the difference operator $\tilde{\mathbf{A}}$ which contains differences *only in the direction of the strong edges* (cf. Remark 3.7). In the classical multigrid convergence analysis there is a severe deterioration of regularity for $\varepsilon \downarrow 0$ (cf. [31]). Here, however, due to the problem dependent measure of regularity, a Fourier analysis yields

$$\|[\mathbf{A}_{rb} \quad \tilde{\mathbf{D}}_{rr}](\mathbf{A}^{(0)})^{-1}\|_2 \leq 1.$$

From Theorem 3.6 we obtain

$$\|\mathbf{I}_b - \tilde{\mathbf{A}}_{bb}^{(k)} \mathbf{S}_{bb}^{-1}\|_2 \leq \left(\frac{\varepsilon}{\varepsilon + 1}\right)^{k-1}.$$

We conclude that, already for $k = 2$, $\tilde{\mathbf{A}}_{bb}^{(k)}$ is a good preconditioner for \mathbf{S}_{bb} , *uniformly* in ε and h .

Remark 4.1 Note that $\tilde{\mathbf{A}}(\mathbf{A}^{(0)})^{-1}$ is bounded uniformly in the parameters h, ε ("algebraic regularity"), but that $(\mathbf{A}^{(0)})\tilde{\mathbf{A}}^{-1}$ is not uniformly bounded. If we define $\tilde{\mathbf{S}}_{bb}$ to be the Schur complement on $\tilde{\mathbf{A}}$, one easily verifies that also $\mathbf{S}_{bb}\tilde{\mathbf{S}}_{bb}^{-1}$ is not uniformly bounded. Hence, this example shows that simply taking the Schur complement of the reduced matrix (obtained by lumping all small off-diagonal entries to the diagonal) as a Schur complement approximation, is not a satisfactory approach.

Example 3 (Convection-diffusion equation). We consider the five-point discretization of a convection-diffusion equation with stencil

$$[\mathbf{A}] = \frac{1}{h} \begin{bmatrix} & & -\varepsilon & & \\ & -1 - \varepsilon & 1 + 4\varepsilon & -\varepsilon & \\ & & & & \\ & & & & -\varepsilon \\ & & & & \end{bmatrix}, \quad \text{with } 0 < \varepsilon < 2. \quad (46)$$

Note that if $\varepsilon > \frac{7}{3}$ then all edges are labeled strong and graph coarsening is done as for the pure diffusion stencil (43) of Example 1. For the case $0 < \varepsilon < 2$, i.e. strong convection, the edges in the digraph corresponding to the matrix entries $-\varepsilon$ are labeled weak. Hence the reduced digraph has a structure which corresponds to the two-point stencil $[* \quad * \quad \cdot]$. One possible maximal independent set (of the reduced digraph) is obtained by standard semi-coarsening (i.e. coarsening by a factor 2 in the x -direction only) of the grid. This then results in stencils

$$[\mathbf{A}_{rr}] = \frac{1}{h} \begin{bmatrix} & -\varepsilon & \\ & 1 + 4\varepsilon & \\ & & -\varepsilon \end{bmatrix}, \quad [\mathbf{A}_{rb}] = \frac{1}{h} \begin{bmatrix} & & -\varepsilon \\ -1 - \varepsilon & & \\ & & \end{bmatrix}, \quad [\mathbf{A}_{br}] = \frac{1}{h} \begin{bmatrix} -1 - \varepsilon & & \\ & & \\ & & -\varepsilon \end{bmatrix}.$$

Hence $\mathbf{D}_{rr} = \frac{1+4\varepsilon}{h}\mathbf{I}_r$, $\tilde{\mathbf{D}}_{rr} = \frac{1+2\varepsilon}{h}\mathbf{I}_r$. As in the previous two examples, we obtain:

$$\|\mathbf{A}_{br}(\mathbf{I}_r - \mathbf{D}_{rr}^{-1}\mathbf{A}_{rr})^{k-1}\tilde{\mathbf{D}}_{rr}^{-1}\|_2 \leq \|\mathbf{I}_r - \mathbf{D}_{rr}^{-1}\mathbf{A}_{rr}\|_2^{k-1} \leq \left(\frac{2\varepsilon}{4\varepsilon + 1}\right)^{k-1}.$$

Related to the term (42) we note that

$$[\mathbf{A}_{rb} \quad \tilde{\mathbf{D}}_{rr}](\mathbf{A}^{(0)})^{-1} = [\emptyset \quad \mathbf{I}_r] \tilde{\mathbf{A}}(\mathbf{A}^{(0)})^{-1},$$

where $\tilde{\mathbf{A}}$ and $\mathbf{A}^{(0)}$ have stencils

$$[\tilde{\mathbf{A}}] = \frac{1}{h} \begin{bmatrix} 0 & & & \\ -1 - \varepsilon & 1 + 2\varepsilon & & -\varepsilon \\ & & 0 & \\ & & & \end{bmatrix}, \quad [\mathbf{A}^{(0)}] = \frac{1}{h} \begin{bmatrix} & & & -\varepsilon \\ -1 - \varepsilon & 1 + 4\varepsilon & & -\varepsilon \\ & & & -\varepsilon \\ & & & \end{bmatrix}.$$

As in Example 2, the regularity is measured using a difference operator which contains differences only in the direction of the strong edges. A simple Fourier eigenvalue analysis yields

$$\|\tilde{\mathbf{A}}(\mathbf{A}^{(0)})^{-1}\|_2 \leq 1,$$

and hence $\|[\mathbf{A}_{rb} \quad \tilde{\mathbf{D}}_{rr}](\mathbf{A}^{(0)})^{-1}\|_2 \leq 1$. Thus we obtain

$$\|\mathbf{I}_b - \tilde{\mathbf{A}}_{bb}^{(k)} \mathbf{S}_{bb}^{-1}\|_2 \leq \left(\frac{2\varepsilon}{4\varepsilon + 1}\right)^{k-1}.$$

We conclude that, already for $k = 2$, $\tilde{\mathbf{A}}_{bb}^{(k)}$ is a good preconditioner for \mathbf{S}_{bb} , *uniformly* in ε and h .

5 Approximate cyclic reduction preconditioner

In this section we present the approximate cyclic reduction preconditioner. In the presentation we distinguish two phases: a *decomposition phase* (construction of the preconditioner) and a *solution phase* (application of the preconditioner). In the decomposition phase we only need the matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$. In the solution phase we need the right hand side \mathbf{b} and the decomposition resulting from the decomposition phase.

Decomposition phase. We assume a sparse matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$. The corresponding ordered digraph has a vertex set that is represented as $\{1, 2, \dots, n\}$. *Dimbound*, with $1 < \text{Dimbound} < n$ is a given integer (used in D5 below). Set $i := 1$, $\mathbf{A}_1 := \mathbf{A}$, $m_0 := n$.

D1. *Red-black partitioning of the vertex set.* Given the digraph of \mathbf{A}_i we make a red-black partitioning of the vertices. We use the method of Section 3.1. In this method we use a parameter β , with $0 \leq \beta < 1$. This results in n_i vertices with label red and m_i vertices with label black. Note: $m_i + n_i = m_{i-1}$.

D2. *Determine permutation.* We determine a symmetric permutation $p_i : \{1, 2, \dots, m_{i-1}\} \rightarrow \{1, 2, \dots, m_{i-1}\}$ such that applying this permutation to the set of vertices results in an ordering in which all vertices with label red have index $j \in (m_i, m_{i-1}]$ and all vertices with label black have index $j \in [1, m_i]$. Note that since we only have to permute between the sets

$\{j \mid j > m_i \text{ and label}(j) = \text{black}\}$ and $\{j \mid j \leq m_i \text{ and label}(j) = \text{red}\}$, such a permutation can be fully characterized by a permutation $\hat{p}_i : \{m_i + 1, m_i + 2, \dots, m_{i-1}\} \rightarrow \{1, 2, \dots, m_i\}$.

D3. *Determine permuted matrix.* The symmetric matrix corresponding to the permutation p_i of D2 is denoted by \mathbf{P}_i . We determine $\mathbf{P}_i \mathbf{A}_i \mathbf{P}_i$. This matrix has a 2×2 -block representation:

$$\mathbf{P}_i \mathbf{A}_i \mathbf{P}_i = \begin{bmatrix} \mathbf{A}_i^{bb} & \mathbf{A}_i^{br} \\ \mathbf{A}_i^{rb} & \mathbf{A}_i^{rr} \end{bmatrix}, \quad (47)$$

with $\mathbf{A}_i^{rr} \in \mathbb{R}^{n_i \times n_i}$, $\mathbf{A}_i^{bb} \in \mathbb{R}^{m_i \times m_i}$, $\mathbf{A}_i^{rb} \in \mathbb{R}^{n_i \times m_i}$, $\mathbf{A}_i^{br} \in \mathbb{R}^{m_i \times n_i}$.

D4. *Compute Schur complement approximation.* Compute an approximation $\mathbf{A}_{i+1} \in \mathbb{R}^{m_i \times m_i}$ of the Schur complement $\mathbf{P}_i \mathbf{A}_i \mathbf{P}_i / \mathbf{A}_i^{rr}$. We use the approximation resulting after two point-Gaussian type of elimination steps as explained in Remark 3.8. We use a parameter MSIZE.

D5. *Store.* Save $m_i, \hat{p}_i, \mathbf{A}_i^{rr}, \mathbf{A}_i^{rb}, \mathbf{A}_i^{br}$. If $m_i < Dimbound$ then save \mathbf{A}_{i+1} (stop the reduction process) else $i := i + 1$ and goto D1.

If this decomposition process stops with $i = i_{\max}$, we obtain integers $m_1 > m_2 > \dots > m_{i_{\max}}$, permutation vectors \hat{p}_i ($1 \leq i \leq i_{\max}$), sparse matrices $\mathbf{A}_i^{rr}, \mathbf{A}_i^{rb}, \mathbf{A}_i^{br}$ ($1 \leq i \leq i_{\max}$) and the approximate Schur complement on the highest level $\mathbf{A}_{i_{\max}+1}$. We use the following terminology: \hat{p}_i is called the *permutation operator* on level i , \mathbf{A}_i^{rr} is called the *solve operator* on level i , \mathbf{A}_i^{rb} is called the *collect operator* on level i , \mathbf{A}_i^{br} is called the *distribute operator* on level i .

The red vertices on all levels, together with the black vertices on the final level induce a direct sum decomposition $\mathbb{R}^n = \mathbb{R}^{n_1} \oplus \mathbb{R}^{n_2} \oplus \dots \oplus \mathbb{R}^{n_{i_{\max}}} \oplus \mathbb{R}^{m_{i_{\max}}}$. The vertices on level i with label red are assigned the level number i , and the vertices on level i_{\max} with label black are assigned level number $i_{\max} + 1$. The vertices (unknowns) with level number j are called the level j vertices (unknowns). Note that every vertex has a unique level number.

Solution phase. For a clear description of the solution phase we introduce permute, collect, distribute and solve operations. These operations use the corresponding operators which are available from the decomposition phase. We give a description in a pseudo-programming language.

procedure permuteoperation(i : integer; var $\mathbf{x} \in \mathbb{R}^{m_{i-1}}$) (* uses \hat{p}_i *)

 for $j := m_i + 1$ to m_{i-1} do

 if $j \neq \hat{p}_i(j)$ then interchange x_j and $x_{\hat{p}_i(j)}$;

procedure collectoperation(i : integer; var $\mathbf{x} \in \mathbb{R}^{n_i}$; $\mathbf{g} \in \mathbb{R}^{m_i}$) (* uses \mathbf{A}_i^{rb} *)

 compute $\mathbf{x} := \mathbf{x} - \mathbf{A}_i^{rb} \mathbf{g}$;

```

procedure distributeoperation( $i$ : integer; var  $\mathbf{x} \in \mathbb{R}^{m_i}$ ;  $\mathbf{g} \in \mathbb{R}^{n_i}$ ) (* uses  $\mathbf{A}_i^{br}$  *)
  compute  $\mathbf{x} := \mathbf{x} - \mathbf{A}_i^{br} \mathbf{g}$ ;

```

```

procedure solveoperation( $i$ : integer; var  $\mathbf{x} \in \mathbb{R}^{n_i}$ ) (* uses  $\mathbf{A}_i^{rr}$  *)
  solve  $\mathbf{A}_i^{rr} \mathbf{w} = \mathbf{x}$  approximately. We use  $\nu$  Gauss-Seidel iterations with starting vector
   $(\text{diag}(\mathbf{A}_i^{rr}))^{-1} \mathbf{x}$ . The result is written in  $\mathbf{x}$ .

```

```

procedure highestlevelsolve(var  $\mathbf{x} \in \mathbb{R}^{m_{i_{\max}}}$ ) (* uses  $\mathbf{A}_{i_{\max}+1}$  *)
  solve  $\mathbf{A}_{i_{\max}+1} \mathbf{w} = \mathbf{x}$ ;  $\mathbf{x} := \mathbf{w}$ ;

```

Using these procedures it is easy to formulate the backward and forward substitution process, i.e. the solution phase, of the approximate cyclic reduction preconditioner. On each level i ($1 \leq i \leq i_{\max} + 1$) we define `ULsolve` as follows:

```

procedure ULsolve( $i$ : integer; var  $\mathbf{f} \in \mathbb{R}^{m_{i-1}}$ );
  var  $\mathbf{f}_{\text{red}} \in \mathbb{R}^{n_i}$ ;
  begin
    if  $i = i_{\max} + 1$  then highestlevelsolve( $\mathbf{f}$ ) else
    begin
      permuteoperation( $i$ ,  $\mathbf{f}$ );
      partition  $\mathbf{f} = \begin{pmatrix} \mathbf{f}_b \\ \mathbf{f}_r \end{pmatrix}$  with  $\mathbf{f}_r \in \mathbb{R}^{n_i}$ ,  $\mathbf{f}_b \in \mathbb{R}^{m_i}$ ;
      make a copy  $\mathbf{f}_{\text{red}} := \mathbf{f}_r$ ;
      solveoperation( $i$ ,  $\mathbf{f}_{\text{red}}$ );
      distributeoperation( $i$ ,  $\mathbf{f}_b$ ,  $\mathbf{f}_{\text{red}}$ );
      ULsolve( $i + 1$ ,  $\mathbf{f}_b$ );
      collectoperation( $i$ ,  $\mathbf{f}_r$ ,  $\mathbf{f}_b$ );
      solveoperation( $i$ ,  $\mathbf{f}_r$ );
      permuteoperation( $i$ ,  $\mathbf{f}$ );
    end
  end;

```

An approximate solution of $\mathbf{Ax} = \mathbf{b}$ results from the call `ULsolve(1, \mathbf{b})`. The structure of `ULsolve` is similar to the structure of the multigrid V -cycle algorithm as presented in [18]. The distribute and collect operations correspond to the multigrid restriction and prolongation respectively. The solve operation corresponds to the smoother in multigrid. Note, however, that in `ULsolve` we do not use any grid information and that every unknown is involved in the solve operations of precisely one level (as in hierarchical basis multigrid, cf. [1]).

If in the decomposition phase all Schur complement approximations (which are computed in step D4) are stored, then an algebraic version of the classical multigrid method can be implemented. This version then uses smoothing in *all* fine grid points and the convergence rate will be improved. In this approach we can also use a W -cycle instead of a V -cycle.

Note that for this version, in view of efficiency, the "rate of coarsening" has to be controlled. In our opinion the preconditioner as presented in this section is much easier to implement than the algebraic version of classical multigrid, due to the fact that in the former method every unknown is on precisely one level.

6 Implementation issues

In Section 5 we presented the approximate cyclic reduction preconditioner. In Section 7 we use this preconditioner in a standard Krylov subspace method. In this section we briefly discuss a few implementation aspects of the preconditioner.

For the implementation of the preconditioner we first consider the standard cyclic reduction method for a *tridiagonal* matrix (cf. Section 2). This method can be implemented along the lines as explained in Section 5. Hence we first make an implementation of the decomposition phase consisting of the procedures D1-D5. However, for the tridiagonal case, in D1 we use the natural odd-even numbering and in D4 we compute the Schur complement *exactly* (apart from rounding errors), because \mathbf{A}_{rr} is diagonal. We implement the procedures D2, D3, D5 as explained in Section 5. For the solution phase we implement the procedure `ULsolve` as explained in Section 5. In the subroutine `solveoperation` we can take $\nu = 0$ because \mathbf{A}_i^{rr} is diagonal and thus the solution of $\mathbf{A}_i^{rr} \mathbf{w} = \mathbf{x}$ is given by $\mathbf{w} = (\text{diag}(\mathbf{A}_i^{rr}))^{-1} \mathbf{x}$. This then yields an implementation of the classical cyclic reduction method for a tridiagonal matrix. The implementation of this direct solver uses an arbitrary sparse matrix format (e.g. Compressed Row Storage or Ellpack-Itpack format) and can be tested by applying it to tridiagonal matrices. Note that this implementation of cyclic reduction, for a tridiagonal matrix, has a clear modular structure.

The cyclic reduction *preconditioner* for a *general sparse matrix* is obtained by *modifying a few subroutines* in this implementation: in D1 we use a more general red-black partitioning technique (as in Section 3.1), in D4 we use a suitable sparse Schur complement approximation approach and in the subroutine `solveoperation` we use $\nu > 0$ Gauss-Seidel iterations. We emphasize that all other components in the implementation of the classical cyclic reduction method and also the sparse matrix data structure are not altered. In this sense, the implementation of the classical cyclic reduction solver (for a tridiagonal matrix) yields an implementation of the approximate cyclic reduction preconditioner (for a general sparse matrix) with only little additional effort.

The implementation of a Gauss-Seidel method (or any other basic iterative method) in the subroutine `solveoperation` is straightforward. Below we briefly comment on the implementation of the modifications in step D1 and step D4 of the decomposition phase.

In step D1 we first have to label the edges in the digraph with "strong" or "weak". Using the criterion (8) this is straightforward. Then we consider the *reduced* digraph $G_A(V, E_s)$ as explained in Section 3.1 and we want to compute a maximal independent set M of this reduced digraph. Different techniques for constructing a maximal independent set

are known. Several possible algorithms are given in [14], [28]. We sketch our method for constructing a maximal independent set M . The algorithm consists of a graph traversal, i.e. visiting all the vertices of $G_A(V, E_s)$ in a systematic way, and a labeling method. One well-known algorithm for graph traversal is the breadth first search (BFS), cf. [21]. The BFS algorithm starts with a vertex $v \in V$ and marks it as visited. Unvisited vertices adjacent from v are visited next. Then unvisited vertices adjacent from these vertices are visited and so on. This approach is applied to every connected component of $G_A(V, E_s)$. A detailed description can be found in [21].

For the labeling method we initialize with $\text{label}(v) := \text{white}$ for all $v \in V$. Let v_0 be the currently visited vertex in the BFS algorithm. If v_0 is an isolated vertex, i.e. there are no vertices adjacent from v_0 and $\text{label}(v_0) = \text{white}$ we define $\text{label}(v_0) := \text{red}$. If v_0 is not isolated then we apply:

```

if label( $v_0$ ) = white then
  if label( $w$ )  $\in$  {white, black} for all  $w$  adjacent from  $v_0$  then
    label( $v_0$ ) := red;
    label( $w$ ) := black for all  $w$  adjacent from  $v_0$ 
  endif else
    label( $v_0$ ) := black
endif.

```

This results in a red-black partitioning of the vertex set V and the set of red vertices is a maximal independent set of the reduced digraph $G_A(V, E_s)$.

In the modification of step D4 we approximate the Schur complement using the point Gaussian type of elimination technique as explained in Section 3.2. In the preconditioner considered here we use $\tilde{\mathbf{A}}_{bb}^{(2)}$ as a sparse approximation of \mathbf{S}_{bb} . In Remark 3.5 it is explained that we can represent this Schur complement approximation using a Galerkin approach with matrix dependent prolongation and restriction. One could use an implementation based on this representation. Here we discuss another implementation based on a point Gaussian elimination technique as formulated, in linear algebra terms, in (16), (17). In this implementation we distinguish two steps. In the first step we compute $\mathbf{A}^{(1)}$ defined in (16) and in the second step we compute $\tilde{\mathbf{A}}_{bb}^{(2)}$ using (17).

We use the notation $V_r = \{ v \in V \mid \text{label}(v) = \text{red} \}$, $V_b = \{ v \in V \mid \text{label}(v) = \text{black} \}$ and $\overline{N(v)} := N(v) \cup \{v\}$, $v \in V$. Given the matrix $\mathbf{A} = (a_{uv})$, $u \in V$, $v \in \overline{N(u)}$ we describe the implementation of the basic transformation

$$\mathbf{A} := \begin{bmatrix} \mathbf{I}_b & -\mathbf{A}_{br}\mathbf{D}_{rr}^{-1} \\ \emptyset & \mathbf{I}_r \end{bmatrix} \mathbf{A}, \quad (48)$$

used in (16). As indicated in (48), instead of using an iteration index (k in (16)) we overwrite the matrix with new results. Hence \mathbf{A}_{bb} and \mathbf{A}_{br} are overwritten by $\mathbf{A}_{bb} - \mathbf{A}_{br}\mathbf{D}_{rr}^{-1}\mathbf{A}_{rb}$ and $\mathbf{A}_{br} - \mathbf{A}_{br}\mathbf{D}_{rr}^{-1}\mathbf{A}_{rr}$, respectively. For storage of intermediate results we need a sparse

matrix $\mathbf{T}_{br} = (t_{uw})$, $u \in V_b$, $v \in V_r$. In \mathbf{T}_{br} we store $\mathbf{A}_{br}\mathbf{D}_{rr}^{-1}\mathbf{A}_{rr}$. A transformation as in (48), can be implemented as follows:

```

 $\forall u \in V_b \ \forall v \in (N(u) \cap V_r) : t_{uw} = 0$  for all  $w \in (\overline{N(v)} \cap V_r)$  (initialization).
for all  $u \in V_b$  do
  for all  $v \in (N(u) \cap V_r)$  do
    for all  $w \in \overline{N(v)}$  do
      if  $w \in V_b$  then
         $a_{uw} := a_{uw} - \frac{a_{uv}a_{vw}}{a_{vv}}$  (compute  $\mathbf{A}_{bb} - \mathbf{A}_{br}\mathbf{D}_{rr}^{-1}\mathbf{A}_{rb}$ ) (*)
      else
         $t_{uw} := t_{uw} - \frac{a_{uv}a_{vw}}{a_{vv}}$  (compute  $\mathbf{A}_{br}\mathbf{D}_{rr}^{-1}\mathbf{A}_{rr}$ )
      endif
    endfor
  endfor
endfor
endfor
 $\forall u \in V_b \ \forall v \in (N(u) \cap V_r) : a_{uw} := a_{uw} - t_{uw}$  for all  $w \in (\overline{N(v)} \cap V_r)$  (overwrite  $\mathbf{A}_{br}$ )

```

This describes the first step for the computation of the approximate Schur complement. Note that increasing k in (16) corresponds to repeating this procedure, using the updated matrix \mathbf{A} as the input for the next step. In our preconditioner, however, in view of the increase of fill-in we apply this procedure only once ($k = 1$ in (16)). In the second step of the computation of the approximate Schur complement we implement the point Gaussian elimination step (17) with $k = 2$, resulting in $\tilde{\mathbf{A}}_{bb}^{(2)} = \mathbf{A}_{bb}^{(1)} - \mathbf{A}_{br}^{(1)}\tilde{\mathbf{D}}_{rr}^{-1}\mathbf{A}_{rb}$. Since in the first step we have overwritten the given matrix \mathbf{A} by the matrix $\mathbf{A}^{(1)}$ we use the notation $\tilde{\mathbf{A}}_{bb}^{(2)} = \mathbf{A}_{bb} - \mathbf{A}_{br}\tilde{\mathbf{D}}_{rr}^{-1}\mathbf{A}_{rb}$. Let $\tilde{\mathbf{d}}_u$, $u \in V_r$ be the vector which corresponds to the diagonal of $\tilde{\mathbf{D}}_{rr}$. Using this notation the second step, i.e. the computation of $\tilde{\mathbf{A}}_{bb}^{(2)}$ from $\mathbf{A}^{(1)}$, can be implemented similar to (*) above:

```

for all  $u \in V_b$  do
  for all  $v \in (N(u) \cap V_r)$  do
    for all  $w \in (\overline{N(v)} \cap V_b)$  do
       $a_{uw} := a_{uw} - \frac{a_{uv}a_{vw}}{d_{vv}}$ 
    endfor
  endfor
endfor

```

Note that here $N(u)$ ($N(v)$) corresponds to the neighbourhood of u (v) in the digraph of the matrix $\mathbf{A} = \mathbf{A}^{(1)}$ which resulted from the first step.

As a final implementation issue, we briefly discuss the storage of information. The decomposition phase yields $i_{\max} + 1$ levels and every vertex has a unique level number. The matrix $\mathbf{A}_{i_{\max}+1}$ (approximate Schur complement on the highest level) can be stored by assigning

to each level $i_{\max} + 1$ vertex *one row* of $\mathbf{A}_{i_{\max}+1}$. All the information in $\hat{p}_i, \mathbf{A}_i^{rr}, \mathbf{A}_i^{rb}, \mathbf{A}_i^{br}$ ($1 \leq i \leq i_{\max}$) can be stored by assigning to each level i vertex the corresponding \hat{p}_i *entry*, *one row* of \mathbf{A}_i^{rr} , *one row* of \mathbf{A}_i^{rb} and *one column* of \mathbf{A}_i^{br} . The amount of information thus stored at a level i vertex is fully determined by the sparsity of the matrix \mathbf{A}_i . On level $i = 1$ this sparsity is given and on level $i > 1$ it is controlled by the parameter MSIZE as explained in Remark 3.8.

Remark 6.1 The approximate cyclic reduction preconditioner has a clear modular structure. This makes it easy to implement modified versions, in which, for example, we use other red-black partitioning techniques (in D1), another Schur complement approximation (in D4) or other basic iterative solvers (in `solveoperation`). In this paper we restrict ourselves to the basic form as presented in Section 5 and we do not consider modified versions. However, as a result of further research certain modifications might be recommended for certain problem classes.

In the preconditioner we use the parameters: β , MSIZE, *Dimbound* and ν . We do not consider optimization of the efficiency of the preconditioner with respect to these parameters, but use (reasonable) fixed default values (cf. Section 7).

7 Numerical experiments

In this section we show results of a few numerical experiments with the approximate cyclic reduction preconditioner. We use its basic form as presented in Section 5. For the parameters we use the following default values in *all* experiments: $\beta = 0.7$, MSIZE= 14, *Dimbound* = 50 and $\nu = 2$.

In all experiments we use a righthand side $\mathbf{b} \equiv \mathbf{0}$ and a starting vector $\mathbf{x} = (1, 1, \dots, 1)^T$. We consider two methods:

- GMRES(5): standard GMRES method with restart after 5 iterations.
- GMRES(5)+ preconditioning: standard left-preconditioned GMRES(5); we use the approximate cyclic reduction preconditioner.

Since we are mainly interested in the performance of the preconditioner, we present results only for GMRES(5), although for certain problems below the use of other outer iterations (e.g. CG) might have been more efficient. In all figures below the unit on the horizontal axis is one (preconditioned) GMRES(5) iteration, which consists of 5 standard (preconditioned) GMRES iterations.

Experiment 1. We consider the convection-diffusion equation:

$$\begin{cases} -\varepsilon\Delta u + a(x, y)u_x + b(x, y)u_y = f & \text{in } \Omega = (0, 1)^2, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

The functions a, b are defined by: $a(x, y) = 0.1$ if $(x, y) \in (0.5, 0.8)^2$ and $a(x, y) = 100$ otherwise; $b(x, y) = 0.2$ if $(x, y) \in (0.5, 0.8)^2$ and $b(x, y) = 200$ otherwise.

We use a uniform square mesh with mesh size h and a finite difference discretization with stencil

$$[\mathbf{A}] = \frac{\varepsilon}{2h^2} \begin{bmatrix} -\frac{1}{2} & -1 & -\frac{1}{2} \\ -1 & 6 & -1 \\ -\frac{1}{2} & -1 & -\frac{1}{2} \end{bmatrix} + \frac{1}{h} \begin{bmatrix} 0 & 0 & 0 \\ -\frac{a^2}{a+b} & \frac{a^2+ab+b^2}{a+b} & 0 \\ -\frac{ab}{a+b} & \frac{-b^2}{a+b} & 0 \end{bmatrix}.$$

The discretization of the convection term is as in [22]. We consider $h = 1/96$ (i.e. 9025 unknowns). The resulting linear system is solved approximately using GMRES(5) (+ preconditioner). For several values of ε/h the convergence behaviour of GMRES(5) is shown in Figure 1a. As expected, we observe slow convergence and an undesirable dependence of the convergence behaviour on the parameter ε/h . In Figure 1b we show the results for GMRES(5) with approximate cyclic reduction preconditioner. In the preconditioned case we observe a smoother and much faster convergence behaviour. Note that, although the problem appears to be more difficult for smaller values of ε/h , the convergence rate improves significantly if ε/h decreases. For $\varepsilon/h = 1$ the coarsening strategy yields 11 levels with $m_i = 9025, 4513, 2534, 1381, 777, 437, 255, 140, 80, 50, 32$ for $i = 0, 1, \dots, 10$. For $\varepsilon/h = 10^3$ and $\varepsilon/h = 10^{-3}$ we obtain similar results for the coarse graph orders m_i .

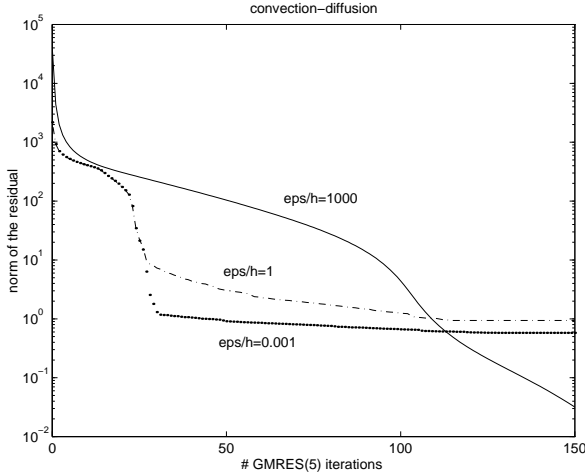


Figure 1a

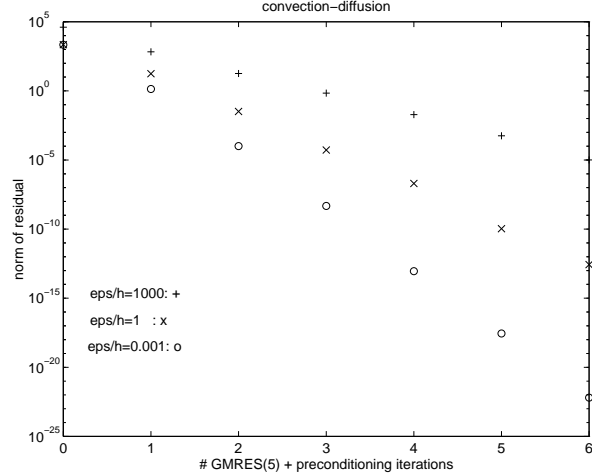


Figure 1b

Experiment 2. We consider a rotated anisotropic equation (cf. [34]):

$$\begin{cases} -(\varepsilon c^2 + s^2)u_{xx} - 2(\varepsilon - 1)csu_{xy} - (\varepsilon s^2 + c^2)u_{yy} = f & \text{in } \Omega = (0, 1)^2, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

with $0 < \varepsilon < 1$, $c = \cos \phi$, $s = \sin \phi$. For the angle $\phi = \phi(x, y)$ we take $\phi = \frac{\pi}{4}$ if $x \leq \frac{1}{2}$ and $\phi = -\frac{\pi}{4}$ if $x > \frac{1}{2}$. We use a standard finite difference discretization on a uniform square

mesh with mesh size h , resulting in a discrete operator with stencils

$$[\mathbf{A}] = \frac{1}{h^2} \begin{bmatrix} \frac{1}{2}(\varepsilon - 1) & -\varepsilon & 0 \\ -\varepsilon & 3\varepsilon + 1 & -\varepsilon \\ 0 & -\varepsilon & \frac{1}{2}(\varepsilon - 1) \end{bmatrix}, \quad [\mathbf{A}] = \frac{1}{h^2} \begin{bmatrix} 0 & -\varepsilon & \frac{1}{2}(\varepsilon - 1) \\ -\varepsilon & 3\varepsilon + 1 & -\varepsilon \\ \frac{1}{2}(\varepsilon - 1) & -\varepsilon & 0 \end{bmatrix}$$

on the left half ($x \leq \frac{1}{2}$) and the right half ($x > \frac{1}{2}$) of the domain, respectively. Note that for $\varepsilon \ll 1$ there are strong anisotropies in different directions. We take $h = 1/96$. For several values of ε the results for GMRES(5) with and without preconditioning are shown in Figure 2b and Figure 2a, respectively. For $\varepsilon = 0.01$ the coarsening strategy yields 10 levels with $m_i = 9025, 4467, 2214, 1142, 586, 306, 173, 101, 57, 30$, for $i = 0, 1, \dots, 9$.

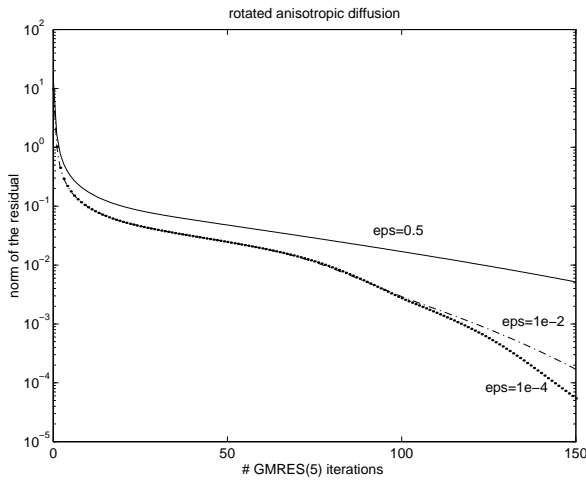


Figure 2a

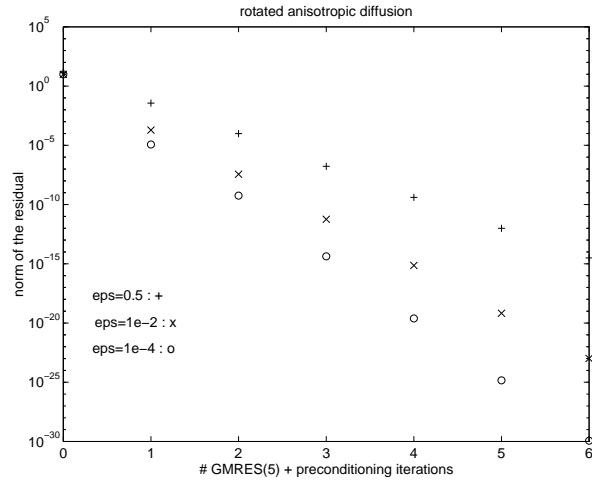


Figure 2b

Experiment 3. We take SHERMAN3 from the Harwell-Boeing collection. This is a symmetric matrix of order 5005 with 20033 nonzero entries. The convergence for GMRES(5) with and without preconditioning is shown in Figure 3b and Figure 3a, respectively.

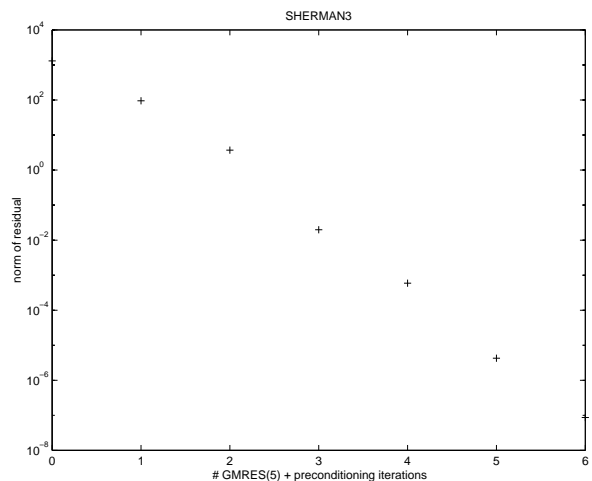
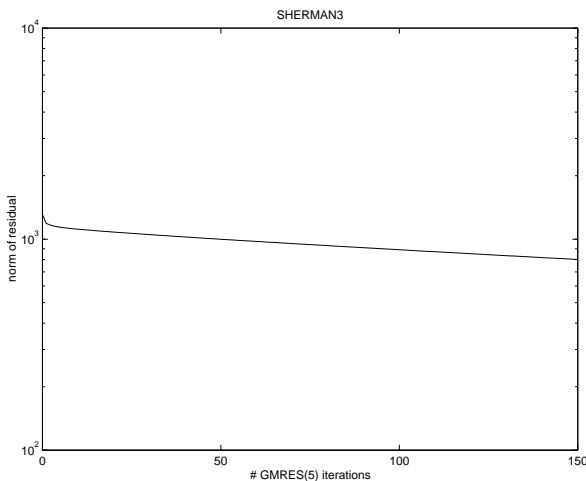


Figure 3a

Figure 3b

Remark 7.1 To give an indication of the arithmetic costs of the preconditioner we consider two typical examples: Experiment 1 with $\frac{\varepsilon}{h} = 1$ and Experiment 3. We only consider the costs for the evaluation of the preconditioner, i.e. one call of `ULsolve(1, b)`. In these (typical) examples the arithmetic work needed for the construction of the preconditioner is less than the work needed in one call of `ULsolve(1, b)`. As a unit of arithmetic work we use a MATVEC, which is the work needed for one matrix-vector multiplication with the given matrix \mathbf{A} . In Experiment 1 with $\frac{\varepsilon}{h} = 1$ the matrix \mathbf{A} contains approximately 81000 nonzero entries. The union of the \mathbf{A}_i^{rr} matrices over all levels contains approximately 53000 nonzero entries. Hence the application of 4 Gauss-Seidel iterations (2 in each call of `solveoperation`) is roughly equivalent to 2.6 MATVEC. The union of the \mathbf{A}_i^{rb} (\mathbf{A}_i^{br}) over all levels contains approximately 55000 (54000) nonzero entries. Hence the total costs for the application of the collect and distribute operations in `ULsolve(1, b)` is comparable to 1.3 MATVEC. Thus, in this example, the total costs in one evaluation of the preconditioner is approximately 3.9 MATVEC. Note that for the preconditioner we have to store 162000 reals, which is comparable to 2 times the amount of storage needed for \mathbf{A} . In Experiment 3 we have a matrix \mathbf{A} with ≈ 20000 nonzero entries. In the union of the \mathbf{A}_i^{rr} we have approximately 16000 nonzero entries. Hence the 4 Gauss-Seidel iterations have costs ≈ 3.2 MATVEC. In the union of the \mathbf{A}_i^{rb} (\mathbf{A}_i^{br}) there are ≈ 15000 (15000) nonzeros. Hence the collect and distribute operations have total costs ≈ 1.5 MATVEC. The total costs, in this example, for one call of `ULsolve(1, b)` are roughly 4.7 MATVEC. The space needed for storage of the preconditioner is comparable to 2.3 times the memory space needed for \mathbf{A} . We note that in all other experiments presented above, the costs of one `ULsolve` evaluation are between 3 and 5 MATVEC.

Experiment 4. We take ORSREG1 from the Harwell-Boeing collection. This is a non-symmetric matrix from oil reservoir simulation of order 2205 with 14133 nonzero entries. The results for GMRES(5) with and without preconditioning are shown in Figure 4b and Figure 4a, respectively. For the coarsening strategy yields 7 levels with $m_i = 2205, 882, 441, 220, 94, 58, 32$, for $i = 0, 1, \dots, 6$.

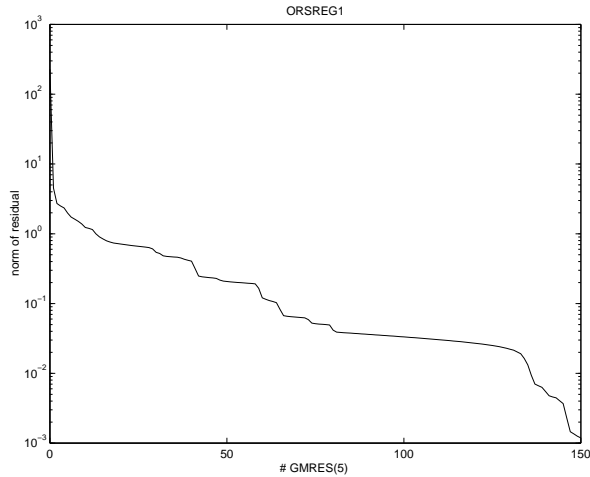


Figure 4a

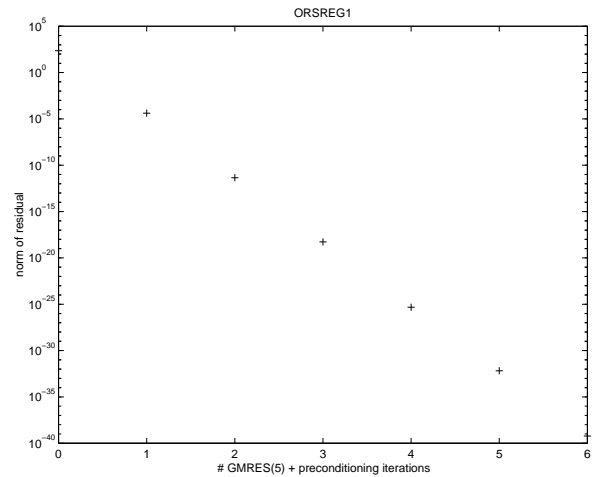


Figure 4b

Remark 7.2 In this paper we have analyzed an approximate cyclic reduction preconditioning technique. In the numerical experiments we considered a basic variant with fixed default parameter values. Clearly there are many possible other variants. Here we mention one particular variant in which we allow the parameter MSIZE (cf. Remark 3.8) to be level dependent. In numerical experiments we observed that allowing more fill-in on higher levels (increasing MSIZE if the level number i increases) may yield a significant improvement in the efficiency of the preconditioner. A systematic study of this and other variants is left for future research.

References

- [1] R.E. Bank, T.F. Dupont and H. Yserentant, *The hierarchical basis multigrid method*, Numer. Math., 52 (1988), pp. 427-458.
- [2] R.E. Bank and S. Gutsch, *Hierarchical basis for the convection-diffusion equation on unstructured meshes*, to appear in the Proceedings of the Ninth International Symposium on Domain Decomposition Methods for Partial Differential Equations (P. Björstad, M. Espedal and D. Keyes, eds.), Wiley, New York.
- [3] R.E. Bank and R.K. Smith, *The hierarchical basis multigraph algorithm*, report Department of Mathematics, University of California at San Diego, 1997.
- [4] R.E. Bank and C. Wagner, *Multilevel ILU Decomposition*, report Department of Mathematics, University of California at San Diego, 1997.
- [5] E.F.F. Botta and A. van der Ploeg, *Preconditioning techniques for matrices with arbitrary sparsity patterns*, in Proceedings of the Ninth International Conference on Finite Elements in Fluids, New Trends and Applications, 1995, pp. 989-998.

- [6] E.F.F. Botta and W. Wubs, *MRILU: it's the preconditioning that counts*, report W-9703, Department of Mathematics, University of Groningen, The Netherlands, 1997.
- [7] D. Braess, *Towards algebraic multigrid for elliptic problems of second order*, Computing, 55 (1995), pp. 379-393.
- [8] A.M. Bruaset, *A survey of preconditioned iterative methods*, Pitman Research Notes in Mathematics, vol. 328, Longman 1995.
- [9] T.F. Chan, S. Go and J. Zou, *Multilevel domain decomposition and multigrid methods for unstructured meshes: algorithms and theory*, Technical Report 95-24, Department of Mathematics, University of California at Los Angeles, 1996.
- [10] T.F. Chan and B. Smith, *Domain decomposition and multigrid methods for elliptic problems on unstructured meshes*, in: Domain Decomposition Methods in Science and Engineering, Proceedings of the Seventh International Conference on Domain Decomposition, D. Keyes and J. Xu, eds., AMS, Providence, 1994, pp. 175-189.
- [11] C. Chartrand and L. Lesniak, *Graphs & digraphs*, Chapman & Hall, London, third edition, 1996.
- [12] N.H. Decker, J. Mandel and S.V. Parter, *On the Role of Regularity in Multigrid Methods*, in: Multigrid Methods, Proceedings of the 3rd Copper Mountain Conference, S. McCormick, J. Dendy, J. Mandel, S. Parter and J. Ruge, eds., Marcel Dekker, New York 1988, pp. 143-156.
- [13] J.E. Dendy, *Black box multigrid*, J. Comput. Phys., 48 (1982), pp. 366-386.
- [14] R. van Driessche and D. Roose, *A graph contraction algorithm for the fast calculation of the Fiedler vector of a graph*, in: Proceedings of the Seventh SIAM Conference on Parallel Processing for Scientific Computing, D.H. Bailey, P.E. Bjørstad, J.R. Gilbert, M.V. Mascagni, R.S. Schreiber, H.D. Simon, V.J. Torczon, L.T. Watson, eds., SIAM, Philadelphia, 1995, pp. 621-626.
- [15] M. Fiedler, *Special matrices and their applications in numerical mathematics*, Martinus nijhoff Publishers, Dordrecht 1986.
- [16] J. Fuhrmann, *Zur Verwendung von Mehrgitterverfahren bei der numerischen Behandlung elliptischer partieller Differentialgleichungen zweiter Ordnung mit variablen Koeffizienten*, PhD thesis, Technische Universität Chemnitz-Zwickau, 1994. Verlag Shaker, Aachen 1995.
- [17] G.H. Golub and C. van Loan, *Matrix Computations*, Johns Hopkins University Press, second edition, 1989.
- [18] W. Hackbusch, *Multigrid Methods and Applications*, Springer, Berlin, Heidelberg, New York, 1985.

- [19] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*, Springer, New York 1994.
- [20] D. Heller, *Some aspects of the cyclic reduction algorithm for block tridiagonal linear systems*, SIAM J. Numer. Anal., 13 (1976), pp. 484-496.
- [21] E. Horowitz and S. Sahni, *Fundamentals of data structures in Pascal*, Pitman, London, 1984.
- [22] W. Layton, *On the principal axes of diffusion in difference schemes for 2D transport problems*, J. Comput. Phys., 90 (1990), pp. 336-347.
- [23] A. van der Ploeg, *Preconditioning for sparse matrices with applications*, PhD thesis, University of Groningen, 1994.
- [24] A. Reusken, *A multigrid method based on incomplete Gaussian elimination*, Numer. Linear Algebra Appl., 3 (1996), pp. 369-390.
- [25] A. Reusken, *Approximate cyclic reduction preconditioning*, to appear in Proceedings of the Fifth European Multigrid conference (W. Hackbusch and G. Wittum, eds.).
- [26] J.W. Ruge and K. Stüben, *Algebraic multigrid*, in Multigrid Methods, S.F. McCormick, ed., SIAM, Philadelphia, 1987, pp. 73-130.
- [27] Y. Saad, *Iterative methods for sparse linear systems*, PWS Publishing Company, Boston 1996.
- [28] Y. Saad, *ILUM: a multi-elimination ILU preconditioner for general sparse matrices*, SIAM J. Sci. Comput., 17 (1996), pp. 830-847.
- [29] Y. Saad and J. Zhang, *BILUM: Block Versions of Multi-Elimination and Multi-Level ILU Preconditioner for General Sparse Linear Systems*, report UMSI 97/126, Department of Computer Science, University of Minnesota, 1997.
- [30] J. Schröder and U. Trottenberg, *Reduktionsverfahren für Differenzgleichungen bei Randwertaufgaben I*, Numer. Math., 22 (1973), pp. 37-68.
- [31] R. Stevenson, *Robustness of multi-grid applied to anisotropic equations on convex domains and on domains with re-entrant corners*, Numer. Math., 66 (1993), pp. 373-398.
- [32] P.N. Swarztrauber, *The methods of cyclic reduction, Fourier analysis and the FACR algorithm for the discrete solution of Poisson's equation on a rectangle*, SIAM Review, 19 (1977), pp. 490-501.
- [33] P. Vanek, J. Mandel and M. Brezina, *Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems*, Computing, 56 (1996), pp. 179-196.
- [34] P. Wesseling, *An introduction to Multigrid Methods*, Wiley, Chichester, 1992.