DELFT UNIVERSITY OF TECHNOLOGY

REPORT 08-05

Theoretical Comparison of Two-Level Preconditioners based on Multigrid and Deflation

J.M. Tang, S.P. MacLachlan, R. Nabben, C. Vuik

ISSN 1389-6520

Reports of the Department of Applied Mathematical Analysis

Delft 2008

THEORETICAL COMPARISON OF TWO-LEVEL PRECONDITIONERS BASED ON MULTIGRID AND DEFLATION

J.M. TANG *, S.P. MACLACHLAN †, R. NABBEN ‡, AND C. VUIK *

Abstract. It is well-known that two-level preconditioned conjugate gradient (PCG) methods provide efficient techniques for solving large and sparse linear systems whose coefficient matrices are symmetric and positive definite (SPD). A two-level PCG method combines traditional and projection-type preconditioners to get rid of the effect of both small and large eigenvalues of the coefficient matrix. In the literature, various two-level preconditioners are known, coming from the fields of deflation, domain decomposition and multigrid. At first glance, these methods seem to be different; however, from an abstract point of view, they are closely related. In [J.M. Tang, R. Nabben, C. Vuik and Y.A. Erlangga, DUT Report 07-04, Delft University of Technology, Delft, 2007], a theoretical and numerical comparison have been carried out for some of these two-level PCG methods. However, the standard multigrid V(1,1)-cycle preconditioner was excluded from that analysis, since this preconditioner has different properties and requires a different treatment than methods discussed in that paper.

The aim of this paper is to relate the two-level PCG method, with a standard multigrid V(1,1)-cycle (MG) preconditioner in its abstract form, to the deflation and abstract balancing Neumann-Neumann methods as analyzed in [J.M. Tang, R. Nabben, C. Vuik and Y.A. Erlangga, DUT Report 07-04, Delft University of Technology, Delft, 2007]. The MG preconditioner is expected to be more effective than these two-level preconditioners, but we show that this is not always the case. For common choices of the parameters, MG does lead to larger error reductions in each iteration, but the work per iteration is much more expensive, which makes this comparison somewhat unfair. We show that, for special choices of fine-level preconditioners in the deflation or abstract balancing Neumann-Neumann methods, the work for each iteration with these preconditioners is approximately the same as that for the MG preconditioner, and the convergence of the resulting two-level PCG methods are also expected to be approximately the same. Numerical experiments are used to emphasize the theoretical results.

Key words. deflation, domain decomposition, multigrid, conjugate gradients, two-grid schemes, spd matrices, two-level preconditioning.

1. Introduction. The Conjugate Gradient (CG) method is a well-known iterative method for solving large linear systems of equations,

$$Ax = b, \quad A \in \mathbb{R}^{n \times n}, \tag{1.1}$$

whose coefficient matrix, A, is sparse, symmetric, and positive definite (SPD). The convergence rate of CG is naturally bounded in terms of the condition number of A; after j iterations of CG,

$$||x - x_j||_A \leq 2||x - x_0||_A \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^j, \tag{1.2}$$

where x_0 is the starting vector and $|| \cdot ||_A$ is the $A-$norm of a vector. In addition, $\kappa = \kappa(A) = \frac{\lambda_{\max}(A)}{\lambda_{\min}(A)}$ denotes the spectral (effective) condition number of A,

*Delft University of Technology, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft Institute of Applied Mathematics, J.M. Burgerscentrum, Mekelweg 4, 2628 CD Delft, The Netherlands ({j.m.tang, c.vuik}@tudelft.nl). Part of the work has been funded by the Dutch BSIK/BRICKS project.

†Tufts University, Department of Mathematics, 503 Boston Avenue, Medford, MA 02155, USA, (scott.maclachlan@tufts.edu). This research was partly supported by the European Community's Sixth Framework Programme, through a Marie Curie International Incoming Fellowship, MIF1-CT-2006-021927.

§Technische Universität Berlin, Institut für Mathematik, MA 3-3, Straße des 17. Juni 136, D-10623 Berlin, Germany (nabben@math.tu-berlin.de). Part of the work of this author has been funded by the Deutsche Forschungsgemeinschaft (DFG), Project NA248/2-2.

where $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ are the largest and smallest nonzero eigenvalue of A, respectively. If κ is large, it is often more favorable to solve a preconditioned system instead of (1.1) directly, where the preconditioned system is defined as $\widehat{A}\hat{x} = \hat{b}$, with $\widehat{A} = M^{-1/2}AM^{-1/2}$, $\hat{x} = M^{1/2}x$, $\hat{b} = M^{-1/2}b$, and $M^{-1} \in \mathbb{R}^{n\times n}$ is an SPD preconditioner. This can be transformed into the system,

$$M^{-1}Ax = M^{-1}b, \tag{1.3}$$

and the resulting method is called a preconditioned CG (PCG or PREC) method. The preconditioner, M^{-1}, should be chosen such that $M^{-1}A$ has a more clustered spectrum or a smaller condition number than A. Furthermore, systems $My = z$ should be cheap to solve, relative to the improvement that they provide in the convergence rate. The design and analysis of such preconditioners for PCG are of great interest. Traditional preconditioners, such as diagonal scaling, basic iterative methods, approximate inverse preconditioning, and incomplete Cholesky preconditioners, are widely used. These preconditioners, however, appear to be less effective for applications with highly refined grids and problems with large coefficient ratios in the original PDEs. In these applications, traditional PCG may suffer from slow convergence due to the presence of relatively small or large eigenvalues, which have a harmful influence on the condition number of the coefficient matrix.

1.1. Two-Level Preconditioning. In addition to traditional preconditioners, a second kind of preconditioner can be incorporated to improve the performance of PCG, so that the resulting approach gets rid of the effect of both small and large eigenvalues. This combined preconditioning is also known as 'two-level preconditioning', and the resulting iterative method is called a 'two-level PCG method', abbreviated as '2L-PCG'. In this case, CG, in combination with a preconditioner based on multigrid, domain decomposition, or deflation techniques, can be interpreted as examples of 2L-PCG methods, since these methods rely on preconditioning on two levels, see [19] and the references therein for more details. The general linear system that is the basis of the 2L-PCG methods can be expressed as

$$\mathcal{P}Ax = \mathcal{P}b, \quad \mathcal{P} \in \mathbb{R}^{n\times n}, \tag{1.4}$$

where $\mathcal{P}$ is the two-level preconditioner. If $\mathcal{P} = M^{-1}$ is taken, we simply obtain the standard PCG method. At first glance, 2L-PCG methods with $\mathcal{P}$ derived from deflation, domain decomposition and multigrid seem to be different. However, it has been shown in [19] that some of these methods are closely related or even equivalent in their abstract forms.

1.2. Aims of this Paper. In this paper, we focus on the comparison between abstract balancing Neumann-Neumann (BNN) [12], deflation (DEF) [17], and multigrid (MG) [2,9,21,25] preconditioners. In [19], it has been shown that BNN and DEF have almost the same spectral properties, and that these properties are quite similar to those of the multigrid V(0,1)- and V(1,0)-cycle preconditioners, even though these are typically not considered as allowable preconditioners for CG.

Here, DEF and BNN will be compared to the 2L-PCG method with a multigrid V(1,1)-cycle preconditioner, denoted as the MG method. The MG method arises from considering the acceleration of a stationary iteration, $x_{j+1} := x_j + \bar{M}^{-1}(b - Ax_j)$, using a projection technique, where $\bar{M}^{-1}$ denotes a preconditioner that can be possibly nonsymmetric. Typically, these methods are analyzed not as preconditioners, but in terms of their error-propagation operators, which can be written as

$(I - \bar{M}^{-T}A)P^T(I - \bar{M}^{-1}A)$, where P is a projection matrix that is based on Z consisting of so-called projection vectors, see Section 2 for the exact definition. This MG method is not analyzed and compared with the other methods in [19], since it has completely different spectral properties and requires a different theoretical treatment, because of the more general choice of traditional preconditioner allowed within MG. The aim of this research is to fill this gap and compare the abstract versions of MG, DEF and BNN.

Of course, the MG method [2, 9, 21, 25] and its properties [1, 5, 8, 13, 16] are well-known. Our intention is not to reproduce these results (although some known results needed for the comparison are briefly reviewed), but to compare and connect MG to other well-known 2L-PCG methods. Intuitively, we expect MG to have better convergence properties than the other 2L-PCG methods when $\bar{M}^{-1} = M^{-1}$, since it is the only 2L-PCG method with two applications of the traditional preconditioners (in the pre- and post-smoothing steps), in addition to a single coarse-grid correction step within one iteration. DEF, on the other hand, has optimal convergence properties in terms of its spectral properties compared with the other 2L-PCG methods (except MG), see [19]. Therefore, it is sufficient for the comparison to show that MG has more favorable spectral properties than DEF, if MG is indeed superior to DEF. Hence, we will often base the analysis on the comparison of DEF and MG. However, the comparison between MG and BNN is, in some cases, easier to perform, so BNN is also used in the analysis.

The main questions of our research are the following.

- Is there an obvious relation between the eigenvalues of DEF and MG?
- Is it true, in general, that MG is expected to converge faster than DEF?
- Denoting the condition numbers of MG and DEF by κ_{MG} and κ_{DEF}, respectively, for which parameters of M^{-1} and Z do we have $\kappa_{\text{MG}} < \kappa_{\text{DEF}}$ unconditionally?
- For which parameters of M^{-1} and Z are MG and DEF expected to converge in approximately the same way?

Note that while the condition number of preconditioned systems is an imperfect indicator of the convergence properties of CG, it is the only analysis tool available with sufficient generality to compare the techniques considered here.

Some spectral analysis for κ_{MG} has been carried out in [3]. In that paper, projection vectors are based on exact eigenvectors of $M^{-1}A$ and more pre- and post-smoothing steps are allowed per iteration. The resulting two-level preconditioner is called a 'multiplicative two-grid spectral preconditioner'. It has been shown that this preconditioner can be effective for many practical applications, where sequences of linear systems have to be solved. In this paper, we restrict ourselves to the standard multigrid V(1,1)-cycle preconditioner using a general matrix Z, while eigenvectors are sometimes used to illustrate the theoretical results. Note that Z and Z^T are known as the prolongation and restriction operator, respectively, in the multigrid field.

This paper is organized as follows. In Section 2, DEF, BNN and MG are described. Then, some spectral properties of MG are presented in Section 3. Thereafter, in Section 4, MG and DEF will be compared by investigating their spectral properties using the identity matrix as the preconditioner. This is then generalized for general preconditioners in Section 5; it is shown there that MG can be less effective than DEF. In Section 6, we show that MG is superior to DEF for more sophisticated preconditioners. Subsequently, Section 7 is devoted to the comparison of MG, BNN and DEF with the same cost per iteration. For special choices of preconditioners,

we show that they are almost spectrally equivalent. Section 8 is devoted to some numerical experiments in order to illustrate the theoretical results. Conclusions are presented in Section 9.

2. Two-Level Preconditioned CG Methods (2L-PCG). In this section, we describe the 2L-PCG methods that are examined in this paper. We start with the following definition, which we assume to hold throughout this paper.

DEFINITION 2.1. *Suppose that an SPD coefficient matrix, $A \in \mathbb{R}^{n\times n}$, and a deflation-subspace matrix, $Z \in \mathbb{R}^{n\times k}$, with full rank and $k < n$ are given. Then, we define the invertible Galerkin matrix, $E \in \mathbb{R}^{k\times k}$, the correction matrix, $Q \in \mathbb{R}^{n\times n}$, and the deflation matrix, $P \in \mathbb{R}^{n\times n}$, as follows:*

$$P := I - AQ, \quad Q := ZE^{-1}Z^T, \quad E := Z^TAZ,$$

where I is the $n \times n$ identity matrix. In addition, $\bar{M}^{-1} \in \mathbb{R}^{n\times n}$ is an arbitrary preconditioning matrix and $M^{-1} \in \mathbb{R}^{n\times n}$ is an SPD preconditioning matrix.

Note that the difference between M^{-1} and $\bar{M}^{-1}$ is that M^{-1} is assumed to be symmetric, positive definite and nonsingular, whereas $\bar{M}^{-1}$ might be nonsymmetric, singular, or even indefinite, so that it is basically the pseudo-inverse of $\bar{M}$. Preconditioner M^{-1} is applied in deflation-like methods, whereas the more general preconditioner, $\bar{M}^{-1}$, is applied solely in multigrid methods, where a general smoothing operator is allowable.

The following lemma will be frequently used, see [19, 24] for more details.

LEMMA 2.2. *Suppose that $A \in \mathbb{R}^{n\times n}$ and $Z \in \mathbb{R}^{n\times k}$ are given. Let Q and P be as in Definition 2.1. Then,*

- $Q = Q^T$;
- $(I - P^T)x = Qb$;
- $AP^T = PA$;
- $QAQ = Q$;
- $P^TZ = PAZ = 0$;
- $QAZ = Z$.

Proof. See, e.g., [19, 24]. □

The deflation method (DEF) can be described as follows. In order to solve $Ax = b$, we decompose x as $x = (I - P^T)x + P^Tx$, where $(I - P^T)x = Qb$ can be computed immediately. For P^Tx, we solve the deflated system,

$$PA\tilde{x} = Pb, \tag{2.1}$$

where $\tilde{x} = x + y$ with $y \in \mathcal{N}(PA)$. Since $P^T\tilde{x} = P^Tx$ follows from Lemma 2.2, the unique solution, x, can be obtained via (2.1) and $x = Qb + P^T\tilde{x}$. Moreover, the deflated system can also be solved using a preconditioner, M^{-1}, giving

$$M^{-1}PA\tilde{x} = M^{-1}Pb, \tag{2.2}$$

see [14, 24] for details. Hence, the two-level preconditioner corresponding to DEF is

$$\mathcal{P}_{\text{DEF}} = M^{-1}P. \tag{2.3}$$

In order to derive the BNN and MG preconditioners, we consider the multiplicative combination of preconditioners. This combination can be explained by considering the stationary iterative methods induced by the preconditioner. Assuming that

C_1 and C_2 are SPD preconditioners, we combine $x_{j+\frac{1}{2}} := x_j + C_1(b - Ax_j)$ and $x_{j+1} := x_{j+\frac{1}{2}} + C_2(b - Ax_{j+\frac{1}{2}})$ to obtain $x_{j+1} = x_j + \mathcal{P}_{m_2}(b - Ax_j)$, with

$$\mathcal{P}_{m_2} := C_1 + C_2 - C_2AC_1, \tag{2.4}$$

which is the multiplicative operator consisting of two preconditioners. In addition, C_1 and C_2 could be combined with another SPD preconditioner, C_3, in a multiplicative way, yielding

$$\mathcal{P}_{m_3} = C_1 + C_2 + C_3 - C_2AC_1 - C_3AC_2 - C_3AC_1 + C_3AC_2AC_1. \tag{2.5}$$

If one substitutes $C_1 = Q,\ C_2 = M^{-1}$ and $C_3 = Q$ into (2.5), we obtain

$$\mathcal{P}_{\text{BNN}} = P^T M^{-1} P + Q, \tag{2.6}$$

which is the two-level preconditioner corresponding to the abstract balancing Neumann-Neumann (BNN) method. In [19], we have shown that BNN has the same spectral properties as the 2L-PCG methods based on multigrid V(0,1)- and V(1,0)-cycle preconditioners. In practice, BNN is always implemented based on these 'reduced' preconditioners, so that the amount of work per iteration is comparable to that of DEF, see also [12, 20].

On the other hand, we could also use $\bar{M}^{-1}$ twice instead of Q, i.e., $C_1 = \bar{M}^{-T}, C_2 = Q$ and $C_3 = \bar{M}^{-1}$ in (2.5). We use the general preconditioner, $\bar{M}^{-1}$, instead of M^{-1}, because $\bar{M}^{-1}$ is not required to be symmetric nor invertible to define $\mathcal{P}_{m_3}$. The resulting two-level preconditioner, well-known as the multigrid V(1,1)-cycle preconditioner, is then explicitly given by

$$\mathcal{P}_{\text{MG}} = \bar{M}^{-T}P + P^T\bar{M}^{-1} + Q - \bar{M}^{-T}PA\bar{M}^{-1}, \tag{2.7}$$

and the resulting method is called MG, see [2, 9, 21, 25]. Eq. (2.7) is only used for the analysis of MG, but is never implemented using this explicit form as the action of $\mathcal{P}_{\text{MG}}$ can be computed with only a single multiplication, each involving $\bar{M}^{-1}$, $\bar{M}^{-T}$, and Q.

$\mathcal{P}_{\text{MG}}$ is obviously symmetric, since

$$\begin{aligned} \mathcal{P}_{\text{MG}}^T &= (\bar{M}^{-T}P)^T + (P^T\bar{M}^{-1})^T + Q^T - (\bar{M}^{-T}PA\bar{M}^{-1})^T \\ &= P^T\bar{M}^{-1} + \bar{M}^{-1}P + Q + \bar{M}^{-T}PA\bar{M}^{-1} \\ &= \mathcal{P}_{\text{MG}}, \end{aligned}$$

but it is not necessarily positive semi-definite, see Section 3.2. Another common and equivalent way to derive $\mathcal{P}_{\text{MG}}$ is to consider the error-propagation operator based on the multigrid V(1,1)-cycle:

$$V := (I - \mathcal{P}_{\text{MG}}A) = (I - \bar{M}^{-T}A)P^T(I - \bar{M}^{-1}A). \tag{2.8}$$

This is often written as

$$V := S^*P^TS, \quad S := I - \bar{M}^{-1}A, \tag{2.9}$$

where S^* denotes the adjoint of S with respect to the A-inner product, i.e., $(Sy_1, y_2)_A = (y_1, S^*y_2)_A$ for $y_1, y_2 \in \mathbb{R}^n$. Matrices S and S^* are known as the pre- and post-smoothers, respectively, and P^T is the MG coarse-grid correction operation. If (2.8) is multiplied out, we obtain the explicit expression of $\mathcal{P}_{\text{MG}}$ as given in (2.7).

It can be observed that the two-level preconditioner corresponding to DEF is included as a term in the two-level preconditioner of MG if $\bar{M}^{-1} = M^{-1}$ is taken (cf. Eqs. (2.3) and (2.7)). Hence, we might expect that MG is always more effective than DEF. For common choices of M, $\bar{M}$ and Z, this is indeed the case, see, e.g., Section 8.2. However, it is not true in all cases, see Sections 4 and 5.

To summarize, the abbrevations and the two-level preconditioners corresponding to the proposed 2L-PCG methods are presented in Table 2.1.

Name	**Method**	**Two-level preconditioner, $\mathcal{P}$**
PREC	Traditional Preconditioned CG	M^{-1}
DEF	Deflation	$M^{-1}P$
BNN	Abstract Balancing	$P^TM^{-1}P + Q$
MG	Multigrid V(1,1)-cycle	$\bar{M}^{-1}P + P^T\bar{M}^{-1} + Q - \bar{M}^{-1}PA\bar{M}^{-1}$

TABLE 2.1
List of two-level PCG methods which will be compared in this paper.

REMARK 2.3. *We emphasize that the parameters of the two-level PCG methods that will be compared can be arbitrary, so that the comparison between these methods is based on their abstract versions. This means that the results of the comparison are valid for any full-rank matrix Z, SPD matrices A, M^{-1}, and matrix $\bar{M}^{-1}$.*

3. Spectral Properties of MG. In this section, we present some results related to the spectral properties of the MG method. We first prove a result analogous to [15, Thm. 2.5], demonstrating that the MG preconditioner also clusters a number of eigenvalues at 1. Thereafter, we discuss necessary and sufficient conditions for the MG preconditioner to be SPD. Note that while these are natural concerns from a preconditioning point of view, these questions are not commonly considered for MG methods, which are often applied as stationary iterations and not used as preconditioners in all cases, unlike DEF.

First, we present some notation in Definition 3.1.

DEFINITION 3.1. *Let B be an arbitrary matrix and $\mathcal{S}$ be a subspace. Then,*

- *the null space and column space of B are denoted by $\mathcal{N}(B)$ and $\mathcal{R}(B)$, respectively;*
- *the spectrum of B is denoted by $\sigma(B)$;*
- *$\dim\mathcal{S}$ denotes the dimension of $\mathcal{S}$;*
- *if B is SPD, then the SPD square root of B will be denoted by $B^{1/2}$;*
- *B is called convergent in the A-norm (or A-norm convergent) if $||B||_A < 1$.*

In addition, the eigenvalues of $\mathcal{P}_{MG}A$ (or $\mathcal{P}_{MG}$) not treated by coarse-grid correction are always sorted increasingly. In other words, if $\mathcal{P}_{MG}A$ has k unit eigenvalues, then the spectrum of $\mathcal{P}_{MG}A$ is denoted by

$$\{1, \ldots, 1, \mu_{k+1}, \ldots, \mu_n\},$$

where $\mu_{k+1} \leq \ldots \leq \mu_n$.

The following lemma will be useful in this section.

LEMMA 3.2. *Let $S := I - \bar{M}^{-1}A$ and $S^* := I - \bar{M}^{-T}A$ be as given in (2.9). Then,*

$$\dim\mathcal{N}(S) = \dim\mathcal{N}(S^*).$$

Proof. Note first that S^* is similar to S^T, since $S^* = A^{-1}S^TA$. Hence, the eigenvalues of S^* and S^T are the same (including multiplicity), so that

$$\dim \mathcal{N}(S^*) = \dim \mathcal{N}\left(S^T\right).$$

The rank-nullity theorem says that

$$\dim \mathcal{R}(S) + \dim \mathcal{N}(S) = n.$$

On the other hand, the fundamental theorem of linear algebra (see, e.g., [18]) gives an orthogonal decomposition of

$$\mathbb{R}^n = \mathcal{R}(S) \oplus \mathcal{N}\left(S^T\right), \tag{3.1}$$

implies that

$$\dim \mathcal{N}\left(S^T\right) = n - \dim \mathcal{R}(S) = \dim \mathcal{N}(S).$$

□

3.1. Unit Eigenvalues of $\mathcal{P}_{\text{MG}}A$. We first show that, if S is invertible, $\mathcal{P}_{\text{MG}}A$ has at least k unit eigenvalues. Then, a more general result is given without this condition.

THEOREM 3.3. *Let $\mathcal{P}_{MG}$ and S be as defined in (2.7) and (2.9), respectively. If S is invertible, then $\mathcal{P}_{MG}A$ has k unit eigenvalues.*

Proof. It suffices to show that $\mathcal{P}_{\text{MG}}AW = W$ for a full-rank matrix, $W \in \mathbb{R}^{n\times k}$.

Note first that

$$\begin{aligned}
\mathcal{P}_{\text{MG}}A &= (\bar{M}^{-T}P + P^T\bar{M}^{-1} + Q - \bar{M}^{-T}PA\bar{M}^{-1})A \\
&= \bar{M}^{-T}(I - AQ)A + (I - QA)\bar{M}^{-1}A + QA - \bar{M}^{-T}(I - AQ)A\bar{M}^{-1}A \\
&= \bar{M}^{-1}A + \bar{M}^{-T}AS + S^*QAS.
\end{aligned}$$

Multiplying this by $S^{-1}Z$ leads to

$$\begin{aligned}
\mathcal{P}_{\text{MG}}AS^{-1}Z &= \bar{M}^{-1}AS^{-1}Z + \bar{M}^{-T}AZ + S^*QAZ \\
&= \bar{M}^{-1}AS^{-1}Z + Z \\
&= S^{-1}Z,
\end{aligned}$$

using the fact that $QAZ = Z$. Since Z has rank k and S^{-1} is invertible, $W := S^{-1}Z$ also has rank k. □

The condition that S is invertible is required by the proof of Theorem 3.3, but a similar result holds without this condition, see Theorem 3.4.

THEOREM 3.4. *Let S and S^* be as given in (2.9). Let $\mathcal{P}_{MG}$ be as given in (2.7). Suppose that*

$$\dim \mathcal{N}(S) = \dim \mathcal{N}(S^*) = m, \quad m \in \mathbb{N}. \tag{3.2}$$

Then, $\mathcal{P}_{MG}A$ has one as an eigenvalue, with geometric multiplicity at least k and at most $k + 2m$.

Proof. In the following, we use the factorization of $I - \mathcal{P}_{\text{MG}}A = S^*P^TS$ as given in Eqs. (2.8) and (2.9). Considering Eq. (2.9), there are three ways for a vector, $v \neq 0$, to be in $\mathcal{N}(I - \mathcal{P}_{\text{MG}}A)$:

(i) $v \in \mathcal{N}(S)$, so that $Sv = 0$;

(ii) $Sv \in \mathcal{N}(P^T)$, yielding $P^T S v = 0$;
(iii) $P^T S v \in \mathcal{N}(S^*)$, so that $S^* P^T S v = 0$.
We treat each case separately.

(i) The geometric multiplicity of the zero eigenvalue of $I - \mathcal{P}_{\text{MG}} A$ must be at least m, due to Eq. (3.2). This accounts exactly for all contributions to $\mathcal{N}(I - \mathcal{P}_{\text{MG}} A)$ from null space vectors of the first type.

(ii) Counting the geometric multiplicity of vectors of the second type is only slightly more complicated. We can orthogonally partition $\mathbb{R}^n$ as (cf. Eq. 3.1)

$$\mathbb{R}^n = \mathcal{R}(S) \oplus \mathcal{N}\left(S^T\right).$$

Since $\dim \mathcal{R}(S) = n - m$, it must be the case that

$$\dim \mathcal{N}\left(S^T\right) = m. \tag{3.3}$$

Now, consider the intersection of $\mathcal{R}(Z)$ with subspaces $\mathcal{R}(S)$ and $\mathcal{N}\left(S^T\right)$:

$$\mathcal{Z}_1 := \mathcal{R}(Z) \cap \mathcal{R}(S), \quad \mathcal{Z}_2 := \mathcal{R}(Z) \cap \mathcal{N}\left(S^T\right),$$

and let $\dim \mathcal{Z}_1 = k_1$ and $\dim \mathcal{Z}_2 = k_2$. Note that necessarily $k_1 + k_2 = k$, and that k_2 is no bigger than m, because of (3.3). Since $\mathcal{N}(P^T) = \mathcal{R}(Z)$, we have

$$\dim \mathcal{N}(S) = k_1,$$

which is the contribution to the dimension of the null space by vectors of the second type. Since $k_1 + k_2 = k$ for $k_2 \leq m$, the total dimension of the null space arising from vectors of the first and second type must satisfy

$$k \leq k_1 + m \leq k + m.$$

(iii) Similarly, we can determine the dimension of the null space of the third type. Note first that (cf. Eq. (3.1))

$$\mathbb{R}^n = \mathcal{R}\left(P^T S\right) \oplus \mathcal{N}\left(S^T P\right).$$

Let $\mathcal{M} := \mathcal{N}(S^*)$, and define

$$\mathcal{M}_1 = \mathcal{M} \cap \mathcal{R}\left(P^T S\right), \quad \mathcal{M}_2 = \mathcal{M} \cap \mathcal{N}\left(S^T P\right).$$

Then, the number of unit eigenvalues of the third type is

$$m_1 = \dim \mathcal{M}_1 \leq \dim \mathcal{M} = m.$$

Thus,

$$\dim \mathcal{N}(\mathcal{P}_{\text{MG}} A) = m + k_1 + m_1,$$

which can be bounded by

$$k \leq m + k_1 + m_1 \leq k + 2m.$$

Since counting the geometric multiplicity of zero eigenvalues of $I - \mathcal{P}_{\text{MG}} A$ is trivially equal to the geometric multiplicity of unit eigenvalues of $\mathcal{P}_{\text{MG}} A$, the proof is complete. □

REMARK 3.5.

- *If zero is not an eigenvalue of S, then it is also not an eigenvalue of S^* (which is similar to S^T). Thus, Theorem 3.4 then says that $\mathcal{P}_{MG}A$ has exactly k unit eigenvalues. This result is clearly stronger than Theorem 3.3, which states that $\mathcal{P}_{MG}A$ has at least k unit eigenvalues.*
- *Since $\bar{M}^{-1}$ is nonsymmetric, the geometric and algebraic multiplicity of the zero eigenvalue of S (or, equivalently, the unit eigenvalues of $\bar{M}^{-1}A$) should be distinguished, since they might differ.* [1]
- *In a similar manner as Theorem 3.4, it can be shown that $\mathcal{P}_{BNN}A$ has at least k and at most $2k+m$ unit eigenvalues.*

It has already been shown in [15,19] that DEF corresponds to a coefficient matrix that has exactly k zero eigenvalues, whereas the matrix associated with BNN has at least k unit eigenvalues. Theorem 3.4 shows that the matrix corresponding to MG also has at least k unit eigenvalues.

3.2. Positive Definiteness of $\mathcal{P}_{\mathrm{MG}}$. The 2L-PCG process is guaranteed to converge if $\mathcal{P}$, as given (1.4), is SPD or can be transformed into an SPD matrix, see, e.g., [4] for more details. This is certainly satisfied for BNN and DEF, see [19]. Here, we examine this issue for $\mathcal{P}_{\mathrm{MG}}$.

It is easy to see that $\mathcal{P}_{\mathrm{MG}}$ (and, therefore, also $\mathcal{P}_{\mathrm{MG}}A$) is not SPD for all choices of Z and $\bar{M}^{-1}$, as in the next example.

EXAMPLE 3.6. *Suppose that $\bar{M}^{-1} = I$ and $Z = [v_1 \cdots v_k]$, where $\{v_i\}$ is the set of orthonormal eigenvectors corresponding to eigenvalues $\{\lambda_i\}$ of A. Then,*

$$\mathcal{P}_{MG} = P + P^T + Q - PA = 2I - 2ZZ^T + Z\Lambda^{-1}Z^T - A + ZZ^TA, \tag{3.4}$$

where $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_k)$. Multiplying (3.4) by v_i gives us

$$\begin{aligned} \mathcal{P}_{MG}v_i &= 2v_i - 2ZZ^Tv_i + Z\Lambda^{-1}Z^Tv_i - Av_i + ZZ^TAv_i \\ &= 2v_i - 2ZZ^Tv_i + Z\Lambda^{-1}Z^Tv_i - \lambda_i v_i + \lambda_i ZZ^Tv_i. \end{aligned}$$

This implies

$$\mathcal{P}_{MG}v_i = \begin{cases} 2v_i - 2v_i + \frac{1}{\lambda_i}v_i - \lambda_i v_i + \lambda_i v_i = \frac{1}{\lambda_i}v_i, & \text{for } i = 1, \ldots, k; \\ 2v_i - \lambda_i v_i, = (2-\lambda_i)v_i, & \text{for } i = k+1, \ldots, n. \end{cases} \tag{3.5}$$

Hence, if A has eigenvalues $\{\lambda_i\}$, then the spectrum of $\mathcal{P}_{MG}$ is given by

$$\left\{ \frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_k}, 2-\lambda_{k+1}, \ldots, 2-\lambda_n \right\}.$$

That means that this specific operator, $\mathcal{P}_{MG}$, is SPD if and only if $\lambda_n < 2$.

Example 3.6 shows that $\mathcal{P}_{\mathrm{MG}}$ can be indefinite for some Z and $\bar{M}^{-1}$. This highlights an important difference between MG and DEF. Indeed, many preconditioners, M^{-1}, that make sense with DEF lead to indefinite $\mathcal{P}_{\mathrm{MG}}$, while choices of $\bar{M}^{-1}$ that lead to SPD $\mathcal{P}_{\mathrm{MG}}$ might give nonsymmetric operators for $\mathcal{P}_{\mathrm{DEF}}$.

A necessary and sufficient condition for $\mathcal{P}_{\mathrm{MG}}$ to be SPD is given in Theorem 3.7.

[1] A simple example is Gauss-Seidel for the 1D Poisson problem with homogeneous Dirichlet boundary conditions. Take $A = \mathrm{tridiag}(-1, 2, -1)$ and M to be the lower-triangular part of A. Then, S has eigenvalue 0 with algebraic multiplicity $\frac{n}{2}$, assuming that n is even. Since there is only one eigenvector corresponding to this eigenvalue, the geometric multiplicity is 1.

THEOREM 3.7. *Let $\bar{M}^{-1} \in \mathbb{R}^{n\times n}$ and $Z \in \mathbb{R}^{n\times k}$ be as defined in (2.1). Let $\mathcal{P}_{MG}$ be as given in (2.7). A necessary and sufficient condition for $\mathcal{P}_{MG}$ to be SPD is that Z and $\bar{M}^{-1}$ satisfy*

$$\min_{w:\ w\perp AZy\ \forall y} w^T \left(\bar{M}^{-1} + \bar{M}^{-T} - \bar{M}^{-1}A\bar{M}^{-T}\right) w > 0. \tag{3.6}$$

Proof. By definition, $\mathcal{P}_{\text{MG}}$ is positive definite if and only if $u^T\mathcal{P}_{\text{MG}}u > 0$ for all vectors $u \neq 0$. Taking $u := A^{1/2}y$, this means that $\mathcal{P}_{\text{MG}}$ is SPD if and only if

$$y^T A^{1/2}\mathcal{P}_{\text{MG}}A^{1/2}y > 0,$$

for all y, or that $A^{1/2}\mathcal{P}_{\text{MG}}A^{1/2}$ is positive definite. Moreover, $A^{1/2}\mathcal{P}_{\text{MG}}A^{1/2}$ is symmetric and, so, it is SPD if and only if its smallest eigenvalue is greater than 0. This, in turn, is equivalent to requiring that $I - A^{1/2}\mathcal{P}_{\text{MG}}A^{1/2}$ has largest eigenvalue less than 1. But $I - A^{1/2}\mathcal{P}_{\text{MG}}A^{1/2}$ is a similarity transformation of V,

$$A^{1/2}VA^{-1/2} = I - A^{1/2}\mathcal{P}_{\text{MG}}A^{1/2},$$

which can be written as

$$A^{1/2}VA^{-1/2} = (R\widetilde{S})^T(R\widetilde{S}),$$

for

$$R := I - A^{1/2}QA^{1/2}, \quad \widetilde{S} := I - A^{1/2}\bar{M}^{-1}A^{1/2}.$$

Note that the eigenvalues of $(R\widetilde{S})^T(R\widetilde{S})$ are the singular values squared of $R\widetilde{S}$ (see, e.g., [7]), which are also the eigenvalues of $(R\widetilde{S})(R\widetilde{S})^T = R\widetilde{S}\widetilde{S}^TR$. So, the largest eigenvalue of $A^{1/2}VA^{-1/2}$ is less than 1 if and only if the largest eigenvalue of $R\widetilde{S}\widetilde{S}^TR$ is less than one. This happens if and only if

$$\frac{u^TR(\widetilde{S}\widetilde{S}^T)Ru}{u^Tu} < 1, \quad \forall u \neq 0. \tag{3.7}$$

To maximize this ratio, we write

$$u = A^{1/2}Zy_1 + Ry_2,$$

and note that R is the $L2$-orthogonal projection onto the orthogonal complement of the range of $A^{1/2}Z$ [2]. Then,

$$u^TR(\widetilde{S}\widetilde{S}^T)Ru = y_2^TR(\widetilde{S}\widetilde{S}^T)Ry_2,$$

and

$$u^Tu = y_1^TZ^TAZy_1 + y_2^TR^2y_2.$$

[2]Recall that the $L2$-orthogonal projection onto $\mathcal{R}(B)$ is $B(B^TB)^{-1}B^T$, which follows from the following. If we take $v = By$, then $B(B^TB)^{-1}B^Tv = v$. For v such that $v^TBy = 0$, $B(B^TB)^{-1}B^Tv = 0$. The $L2$-orthogonal projection onto $\mathcal{R}(B)^\perp$ is $I - B(B^TB)^{-1}B^T$, which acts in the opposite way.

So, maximizing the ratio over all choices of y_1 means choosing $y_1 = 0$, so that the denominator of (3.7) is as small as possible. Therefore,

$$\frac{u^T R\widetilde{S}\widetilde{S}^T Ru}{u^T u} < 1 \ \forall u \neq 0 \quad \Leftrightarrow \quad \frac{y_2^T R\widetilde{S}\widetilde{S}^T Ry_2}{y_2^T R^2 y_2} < 1 \ \forall y_2 \neq 0. \tag{3.8}$$

Notice that, for any choice of y_1, the ratio in (3.7) is bounded by the ratio with $y_1 = 0$. Thus, if the ratio on the right of (3.8) is bounded below 1 for all y_2, so must be the ratio in Eq. (3.7). But, if the ratio in (3.7) is bounded below 1 for all x, then it is bounded for $x = Ry_2$, which gives the bound at the right-hand side of (3.8).

Equivalently, we can maximize the ratio over $\mathcal{R}(R) = \mathcal{R}(A^{1/2}Z)^{\perp}$ (i.e., the orthogonal complement of the range of $A^{1/2}Z$). So, the largest eigenvalue of $R\widetilde{S}\widetilde{S}^T R$ is less than 1 if and only if

$$\max_{x:x\perp A^{1/2}Zy\forall y} \frac{x^T\widetilde{S}\widetilde{S}^T x}{x^T x} < 1.$$

By computation, we have

$$\widetilde{S}\widetilde{S}^T = I - A^{1/2}\left(\bar{M}^{-1} + \bar{M}^{-T} - \bar{M}^{-1}A\bar{M}^{-T}\right)A^{1/2}.$$

Therefore, this bound is equivalent to requiring

$$\min_{x:x\perp A^{1/2}Zy\forall y} \frac{x^T A^{1/2}\left(\bar{M}^{-1} + \bar{M}^{-T} - \bar{M}^{-1}A\bar{M}^{-T}\right)A^{1/2}x}{x^T x} > 0.$$

Taking $w = A^{1/2}x$, this is, in turn, equivalent to

$$\min_{w:w\perp AZy\forall y} \frac{w^T\left(\bar{M}^{-1} + \bar{M}^{-T} - \bar{M}^{-1}A\bar{M}^{-T}\right)w}{w^T A^{-1}w} > 0,$$

which is equivalent to

$$\min_{w:w\perp AZy\forall y} w^T\left(\bar{M}^{-1} + \bar{M}^{-T} - \bar{M}^{-1}A\bar{M}^{-T}\right)w > 0,$$

because $w^T A^{-1}w > 0$ for all w. □

Hence, a necessary and sufficient condition for $\mathcal{P}_{\text{MG}}$ to be SPD is, thus, given by (3.6). Intuitively, we expect the spectral properties of $\mathcal{P}_{\text{MG}}$ to reflect those of $\bar{M}^{-1}$, with some account for the coarse-grid correction. Eq. (3.6) is particularly interesting in comparison with Theorem 3.8, which gives a necessary and sufficient condition for M^{-1} to define a convergent smoother, see also [6, 26].

THEOREM 3.8. *Let $\bar{M}^{-1} \in \mathbb{R}^{n\times n}$ and $Z \in \mathbb{R}^{n\times k}$ be as defined in (2.1). Let S be as given in (2.9). A necessary and sufficient condition for S to be convergent in the A-norm is*

$$\min_w w^T(\bar{M}^{-1} + \bar{M}^{-T} - \bar{M}^{-1}A\bar{M}^{-T})w > 0. \tag{3.9}$$

Proof. The proof follows as

$$\begin{aligned}
\|S\|_A < 1 &\Leftrightarrow \|I - A^{1/2}\bar{M}^{-1}A^{1/2}\|_2 < 1 \\
&\Leftrightarrow \lambda_{\max}\left((I - A^{1/2}\bar{M}^{-T}A^{1/2})(I - A^{1/2}\bar{M}^{-1}A^{1/2})\right) < 1 \\
&\Leftrightarrow \lambda_{\max}\left(I - A^{1/2}(\bar{M}^{-1} + \bar{M}^{-T} - \bar{M}^{-T}A\bar{M}^{-1})A^{1/2}\right) < 1 \\
&\Leftrightarrow \lambda_{\min}\left(A^{1/2}(\bar{M}^{-1} + \bar{M}^{-T} - \bar{M}^{-T}A\bar{M}^{-1})A^{1/2}\right) > 0 \\
&\Leftrightarrow \min_w w^T\left(\bar{M}^{-1} + \bar{M}^{-T} - \bar{M}^{-T}A\bar{M}^{-1}\right)w > 0,
\end{aligned}$$

where we have used the definition of the A- and 2-norms [3], and the fact that the maximum and minimum Rayleigh quotient are the largest and smallest eigenvalue, respectively, for any symmetric operator. ∎

Theorem 3.8 amounts to the conditions

$$\begin{aligned}
\|S\|_A < 1 \quad &\Leftrightarrow \quad \min_w w^T\widetilde{M}^{-1}w > 0 \\
&\Leftrightarrow \quad \min_w(\bar{M}^{-T}w)^T(\bar{M} + \bar{M}^T - A)(\bar{M}^{-T}w) > 0 \\
&\Leftrightarrow \quad \min_{v=\bar{M}^{-T}w} v^T(\bar{M} + \bar{M}^T - A)v > 0 \\
&\Leftrightarrow \quad \lambda_{\min}(\bar{M} + \bar{M}^T - A) > 0,
\end{aligned}$$

that can be found in [26, Thm. 5.3]. On the other hand, Theorem 3.7 gives

$$\begin{aligned}
\min_{w:w\perp AZy\forall y} w^T\widetilde{M}^{-1}w > 0 \quad &\Leftrightarrow \quad \min_{w:w\perp AZy\forall y}(\bar{M}^{-T}w)^T(\bar{M} + \bar{M}^T - A)(\bar{M}^{-T}w) > 0 \\
&\Leftrightarrow \quad \min_{v:v=\bar{M}^{-T}w, w\perp AZy\forall y} v^T(\bar{M} + \bar{M}^T - A)v > 0,
\end{aligned}$$

where

$$\widetilde{M}^{-1} := \bar{M}^{-1} + \bar{M}^{-T} - \bar{M}^{-T}A\bar{M}^{-1}. \tag{3.10}$$

Necessarily,

$$\min_{v:v=\bar{M}^{-T}w, w\perp AZy\forall y} v^T(\bar{M} + \bar{M}^T - A)v > \min_y y^T\widetilde{M}^{-1}y = \lambda_{\min}(\bar{M} + \bar{M}^T - A) > 0,$$

so the condition for $\mathcal{P}_{\text{MG}}$ to be SPD is weaker than the condition for a convergent S in the A-norm. In other words, The A-norm convergence of S implies both

- convergence of $I - \mathcal{P}_{\text{MG}}A$, and
- $\mathcal{P}_{\text{MG}}$ to be SPD,

but $\mathcal{P}_{\text{MG}}$ can be SPD even if $||S||_A \geq 1$, so long as coarse-grid correction effectively treats amplified modes.

4. Comparison of a Special Case of MG and DEF. In this section, we show that abstract preconditioners in the MG framework do not always lead to better conditioned systems than DEF. Such problems can even be found in the case of $M^{-1} = \bar{M}^{-1} = I$. We start with some spectral bounds on MG and DEF under these

[3] Recall that, for any matrix B and SPD matrix A,

$$||B||_A^2 = \max_x \frac{x^TB^TABx}{x^TAx} = \max_w \frac{w^TA^{-1/2}B^TABA^{-1/2w}}{w^Tw} = ||A^{1/2}BA^{-1/2}||_2^2,$$

with $w := A^{1/2}x$.

assumptions. Thereafter, we perform a comparison between the condition numbers for MG and DEF for simple parameters. Finally, some examples are given to illustrate the difference between the condition numbers associated with MG and DEF.

We first show that if $I - 2A$ is positive definite (i.e., if the eigenvalues of A are in the interval $(0, 0.5)$), then the eigenvalues of MG are always larger than those of DEF, see the next lemma.

LEMMA 4.1. *Suppose that $M^{-1} = \bar{M}^{-1} = I$. If $I - 2A$ is positive definite, then $\lambda_i(\mathcal{P}_{MG}A) > \mu_i(\mathcal{P}_{DEF}A)$, where $\{\lambda_i\}$ and $\{\mu_i\}$ are increasing sets of eigenvalues.*

Proof. The proof is similar to [14, Thm. 2.6]. We can write

$$\begin{aligned} A^{1/2}\mathcal{P}_{\text{MG}}A^{1/2} - \mathcal{P}_{\text{DEF}}A &= A^{1/2}(P + P^T + Q - PA)A^{1/2} - PA \\ &= A^{1/2}(I - AQ + I - QA + Q - (I - AQ)A)A^{1/2} - (I - AQ)A \\ &= A^{1/2}\left(2I - 2A + (I - Q)A(I - Q) + AQA - I + A^{1/2}QA^{1/2}\right)A^{1/2} \\ &= A^{1/2}\left(I - 2A + (I - Q)A(I - Q) + AQA + A^{1/2}QA^{1/2}\right)A^{1/2}, \end{aligned}$$

where we have used Lemma 2.2. All of $(I - Q)A(I - Q)$, AQA and $A^{1/2}QA^{1/2}$ are positive (semi-)definite. Since $I - 2A$ is also positive definite, $A^{1/2}\mathcal{P}_{\text{MG}}A^{1/2} - \mathcal{P}_{\text{DEF}}A$ is positive definite as well. Using [10, Cor. 7.7.4], we derive that

$$\lambda_i(A^{1/2}\mathcal{P}_{\text{MG}}A^{1/2}) = \lambda_i(\mathcal{P}_{\text{MG}}A) > \mu_i(\mathcal{P}_{\text{DEF}}A).$$

□

In particular, we have $\lambda_{\max}(\mathcal{P}_{\text{MG}}A) > \mu_{\max}(\mathcal{P}_{\text{DEF}}A)$ from Lemma 4.1. However, this result does not say anything about the connection between the condition numbers associated with MG and DEF, since we cannot show that $\lambda_{\min}(\mathcal{P}_{\text{MG}}A) < \mu_{\min}(\mathcal{P}_{\text{DEF}}A)$, where $\mu_{\min}(\mathcal{P}_{\text{DEF}}A)$ denotes the smallest nonzero eigenvalue of $\mathcal{P}_{\text{DEF}}A$. Lemma 4.1 only tells us that if all μ_i are below 0.5 (i.e., $I - 2A$ is positive definite), then each pair of λ_i and μ_i is nicely ordered.

4.1. Analysis using eigenvectors of A as projection vectors. We will show that the comparison of MG and DEF is still not straightforward, although we take 'simple' parameters. Again, it suffices to restrict ourselves to $M^{-1} = \bar{M}^{-1} = I$. Moreover, we now assume that $Z = [v_1 \cdots v_k]$, where $\{v_i\}$ is the set of orthonormal eigenvectors corresponding to the increasing set of eigenvalues $\{\lambda_i\}$ of A. Then, we know from Example 3.6 that the MG operator is only SPD if $\lambda_i < 2$.

Similarly to Example 3.6, we obtain

$$\begin{aligned} \mathcal{P}_{\text{MG}}Av_i &= 2Av_i - 2ZZ^TAv_i + Z\Lambda^{-1}Z^TAv_i - A^2v_i + ZZ^TA^2v_i \\ &= 2\lambda_i v_i - 2\lambda_i ZZ^Tv_i + \lambda_i Z\Lambda^{-1}Z^Tv_i - \lambda_i^2 v_i + \lambda_i^2 ZZ^Tv_i, \end{aligned}$$

where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_k)$. This implies

$$\mathcal{P}_{\text{MG}}Av_i = \begin{cases} 2\lambda_i v_i - 2\lambda_i v_i + v_i - \lambda_i^2 v_i + \lambda_i^2 v_i = v_i, & \text{for } i = 1, \ldots, k; \\ 2\lambda_i v_i - \lambda_i^2 v_i = \lambda_i(2 - \lambda_i)v_i & \text{for } i = k+1, \ldots, n. \end{cases} \tag{4.1}$$

Hence, if A has eigenvalues $\{\lambda_i\}$, then the spectrum of $\mathcal{P}_{\text{MG}}A$ is given by

$$\{1, \ldots, 1, \lambda_{k+1}(2 - \lambda_{k+1}), \ldots, \lambda_n(2 - \lambda_n)\}. \tag{4.2}$$

We note that $\lambda_i(2 - \lambda_i) \leq 1$ for all $i = k+1, \ldots, n$ since $0 < \lambda_i < 2$, see Figure 4.1. Hence, the condition number of $\mathcal{P}_{\text{MG}}A$ is given by

$$\kappa_{\text{MG}} = \frac{1}{\min\{\lambda_{k+1}(2 - \lambda_{k+1}), \lambda_n(2 - \lambda_n)\}}.$$

On the other hand, for DEF, we know that (see [14])

$$\mathcal{P}_{\text{DEF}} A v_i = \begin{cases} 0, & \text{for } i = 1, \ldots, k; \\ \lambda_i, & \text{for } i = k+1, \ldots, n. \end{cases} \tag{4.3}$$

Therefore,

$$\kappa_{\text{DEF}} = \frac{\lambda_n}{\lambda_{k+1}}.$$

It depends on eigenvalues λ_{k+1} and λ_n of A whether κ_{MG} or κ_{DEF} is more favorable. If $\lambda_{k+1}, \ldots, \lambda_n \to 2$, then obviously $\kappa_{\text{DEF}} < \kappa_{\text{MG}}$. In other words, M^{-1} and Z can be chosen in such a way that MG with an SPD operator is expected to converge slower than DEF, see also Example 4.2.

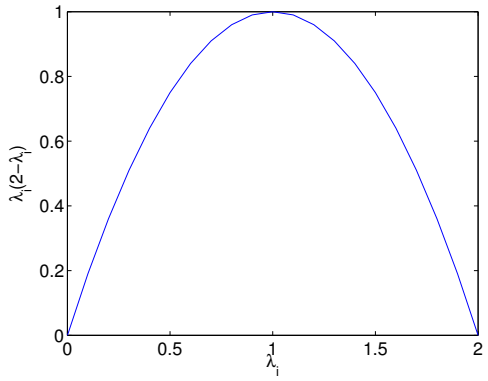


FIG. 4.1. *Function $x(2-x)$ for $x \in [0, 2]$.*

EXAMPLE 4.2. *We first construct a simple example to show that $\kappa_{MG} < \kappa_{DEF}$ does not hold in general, even if $\mathcal{P}_{MG}$ is SPD.*

Let us consider the SPD diagonal matrix, A, given by

$$A = \text{diag}(1, 1.25, 1.5, 1.75).$$

Then, the spectrum of A is given by $\sigma = (1, 1.25, 1.5, 1.75)$, where the corresponding eigenvectors are columns of I: $I = [v_1\ v_2\ v_3\ v_4]$. Hence, $\mathcal{P}_{MG}$ is SPD.

Choose now $Z = [v_1\ v_2]$ and $M^{-1} = I$. Then, the eigenvalues of $\mathcal{P}_{MG}A$ are given by Eq. (4.2):

$$\sigma_{MG} = \{1, 1, \lambda_3(2-\lambda_3), \lambda_4(2-\lambda_4)\} = \{1, 1, 0.4375, 0.75\},$$

whereas (cf. Eq. (4.3))

$$\sigma_{DEF} = \{0, 0, \lambda_3, \lambda_4\} = \{0, 0, 1.5, 1.75\}.$$

This leads immediately to the condition numbers

$$\kappa_{MG} = \frac{1}{\min\{\lambda_{k+1}(2-\lambda_{k+1}), \lambda_n(2-\lambda_n)\}} = \frac{1}{0.4375} = 2.2857, \quad \kappa_{DEF} = \frac{\lambda_n}{\lambda_{k+1}} = \frac{1.75}{1.5} = 1.1667,$$

so that $\kappa_{MG} > \kappa_{DEF}$ obviously holds in this case.

EXAMPLE 4.3. *It is easy to construct examples showing that* $\kappa_{MG} < \kappa_{DEF}$. *For instance, take*

$$A = \mathrm{diag}(0.5, 0.75, 1.0, 1.25),$$

with the same setting of the parameters of MG and DEF as in Example 4.2. Then,

$$\sigma_{MG} = \{1, 1, 1, 0.9375\}, \quad \sigma_{DEF} = \{0, 0, \lambda_3, \lambda_4\} = \{0, 0, 1.0, 1.25\},$$

giving us

$$\kappa_{MG} = \frac{1}{0.9375} = 1.0667, \quad \kappa_{DEF} = \frac{\lambda_n}{\lambda_{k+1}} = \frac{1.25}{1.0} = 1.25,$$

so that $\kappa_{MG} < \kappa_{DEF}$ *holds in this case.*

5. Comparison of MG and DEF for general M^{-1}. In this section, we allow M^{-1} to be chosen arbitrarily, but require that $\bar{M}^{-1} = M^{-1}$ and that Z consist of eigenvectors of $M^{-1}A$. Then, some of the results given in the previous section can be generalized as in the next theorem.

THEOREM 5.1. *Suppose that* $M^{-1} = \bar{M}^{-1}$ *is arbitrary and* $Z = [v_1 \cdots v_k]$ *consists of eigenvectors corresponding to* $\{\lambda_i\}$ *of* $M^{-1}A$. *Let* $\mathcal{P}_{DEF}$ *and* $\mathcal{P}_{MG}$ *be as given in (2.3) and (2.7), respectively. Then,*

(i) $\mathcal{P}_{MG}A$ *has the following eigenvalues:*

$$\begin{cases} 1, & \text{for } i = 1, \ldots, k; \\ \lambda_i(2 - \lambda_i), & \text{for } i = k+1, \ldots, n, \end{cases} \tag{5.1}$$

(ii) $\mathcal{P}_{DEF}A$ *has the following eigenvalues:*

$$\begin{cases} 0, & \text{for } i = 1, \ldots, k; \\ \lambda_i, & \text{for } i = k+1, \ldots, n. \end{cases} \tag{5.2}$$

Proof. The proof follows from [3, Prop. 2] and [24, Sect. 4]. ∎

Note that Eqs. (5.1) and (5.2) are identical to Eqs. (4.2) and (4.3), respectively. Hence, the results from the previous section, based on $M^{-1} = \bar{M}^{-1} = I$ and Z consisting of eigenvectors of A, can be generalized to the case with arbitrary $M^{-1} = \bar{M}^{-1}$ and Z consisting of eigenvectors of $M^{-1}A$. The following example shows that $\kappa_{\mathrm{MG}} < \kappa_{\mathrm{DEF}}$ is not always satisfied (cf. Example 4.2).

EXAMPLE 5.2. *Choose* $\bar{M}^{-1} = M^{-1}$ *and* A *such that* $M^{-1}A$ *is a diagonal matrix given by*

$$M^{-1}A = \mathrm{diag}(1, 1.25, 1.5, 1.75). \tag{5.3}$$

If $Z = [v_1\ v_2]$ *with* v_1 *and* v_2 *to be eigenvectors corresponding to the two smallest eigenvalues of* $M^{-1}A$, *then the eigenvalues of* $\mathcal{P}_{MG}A$ *and* $\mathcal{P}_{DEF}A$ *are*

$$\sigma_{MG} = (1, 1, 0.4375, 0.75), \quad \sigma_{DEF} = (0, 0, 1.5, 1.75).$$

This leads immediately to the condition numbers

$$\kappa_{MG} = \frac{1}{\min\{\lambda_{k+1}(2 - \lambda_{k+1}), \lambda_n(2 - \lambda_n)\}} = 2.2857, \quad \kappa_{DEF} = \frac{\lambda_n}{\lambda_{k+1}} = 1.1667.$$

As in Example 4.2, we have $\kappa_{MG} > \kappa_{DEF}$ for this choice of $\bar{M}^{-1}$, M^{-1} and A.

EXAMPLE 5.3. *It is easy to construct examples showing that $\kappa_{MG} < \kappa_{DEF}$. For instance, take (cf. Example 4.3)*

$$M^{-1}A = \mathrm{diag}(0.5, 0.75, 1.0, 1.25),$$

with the same setting of the parameters of MG and DEF as in Example 5.2. Then, as in Example 4.3, we have

$$\kappa_{MG} = 1.0667 < 1.25 = \kappa_{DEF}.$$

5.1. Comparing MG and DEF. From Theorem 5.1, we know that

$$\kappa_{\mathrm{MG}} = \frac{1}{\min\{\lambda_{k+1}(2-\lambda_{k+1}), \lambda_n(2-\lambda_n)\}}, \quad \kappa_{\mathrm{DEF}} = \frac{\lambda_n}{\lambda_{k+1}}, \tag{5.4}$$

for any $M^{-1} = \bar{M}^{-1}$ when Z consists of eigenvectors of $M^{-1}A$. Hence, determining the method which yields the smallest condition number depends on the eigenvalues, λ_{k+1} and λ_n, of $M^{-1}A$. In this subsection, we present Figure 5.1 from which the best method can be easily determined for given λ_{k+1} and λ_n.

Let $\{\lambda_i\}$ be the eigenvalues of $M^{-1}A$, and suppose that MG is convergent, so that, $0 \le \lambda_j \le 2$ holds for $k < j \le n$ if Z is composed of eigenvectors $v_1, \ldots, v_k$. Furthermore, suppose that the eigenvalues are ordered so that $0 < \lambda_{k+1} \le \lambda_j \le \lambda_n < 2$ for all $k < j \le n$.

Note first that if $\lambda_{k+1} = \lambda_n$, then $\mathcal{P}_{\mathrm{MG}}$ consists of at most two different eigenvalues (including their multiplicity). In addition, if $\lambda_{k+1} = 2 - \lambda_n$, then

$$\kappa_{\mathrm{MG}} = \frac{1}{\lambda_{k+1}(2-\lambda_{k+1})} = \frac{1}{\lambda_n(2-\lambda_n)}.$$

Next, the region, $0 < \lambda_{k+1} \le \lambda_n \le 2$, is naturally partitioned into two subdomains, along the line where $\lambda_{k+1}(2-\lambda_{k+1}) = \lambda_n(2-\lambda_n)$, which occurs when $\lambda_{k+1} = 2 - \lambda_n$:

- if $\lambda_{k+1}(2-\lambda_{k+1}) \le \lambda_n(2-\lambda_n)$ (i.e., in the sector of the $\lambda_{k+1} - \lambda_n$ plane to the left of the line $\lambda_{k+1}(2-\lambda_{k+1}) = \lambda_n(2-\lambda_n)$), then

  $$\kappa_{\mathrm{MG}} = \frac{1}{\lambda_{k+1}(2-\lambda_{k+1})}.$$

  Thus, $\kappa_{\mathrm{MG}} < \kappa_{\mathrm{DEF}}$ if and only if

  $$\frac{1}{2-\lambda_{k+1}} < \lambda_n,$$

  or, equivalently, $\lambda_{k+1} \le 2 - \frac{1}{\lambda_n}$;
- if $\lambda_{k+1}(2-\lambda_{k+1}) \ge \lambda_n(2-\lambda_n)$ (i.e., in the sector of the $\lambda_{k+1} - \lambda_n$ plane to the right of the line $\lambda_{k+1}(2-\lambda_{k+1}) = \lambda_n(2-\lambda_n)$), then

  $$\kappa_{\mathrm{MG}} = \frac{1}{\lambda_n(2-\lambda_n)}.$$

  Thus, $\kappa_{\mathrm{MG}} < \kappa_{\mathrm{DEF}}$ if and only if

  $$\lambda_{k+1} \le \lambda_n^2(2-\lambda_n).$$

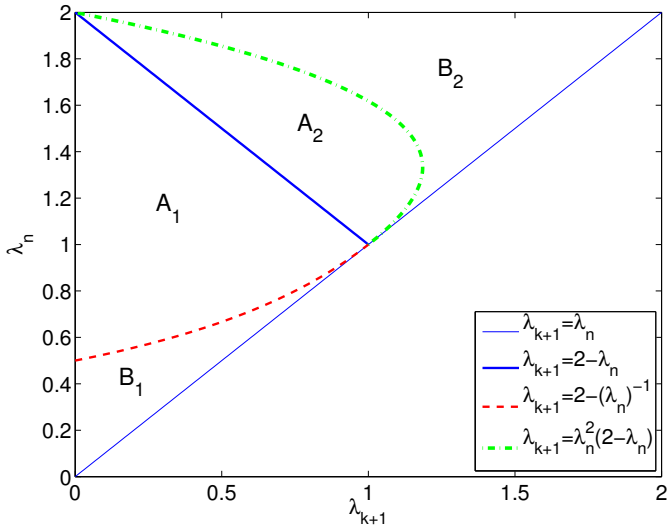


FIG. 5.1. *Regions of the most favorable condition numbers corresponding to MG and DEF for arbitrary* $M^{-1} = \bar{M}^{-1}$*, when* Z *consists of eigenvectors of* $M^{-1}A$*.* $\kappa_{MG} < \kappa_{DEF}$ *holds in Regions* A_1 *and* A_2*, whereas* $\kappa_{DEF} < \kappa_{MG}$ *holds in Regions* B_1 *and* B_2*. The two condition numbers are equal along the red (dotted) and green (dotted-dashed) lines.*

Figure 5.1 depicts these regions graphically. For any given λ_{k+1} and λ_n, the method with smallest condition number follows immediately from this figure. Example 5.4 gives some consequences of Figure 5.1.

EXAMPLE 5.4.

(a) If $\sigma(M^{-1}A) \subseteq (0, 0.5]$*, then we deal with Region* B_1 *and, hence,* $\kappa_{DEF} \leq \kappa_{MG}$*.*

(b) If $\sigma(M^{-1}A) \subseteq (0, 2)$ *with* $\lambda_{k+1} \approx 2 - \lambda_n$*, then we deal with either Region* A_1 *or* A_2 *and* $\kappa_{DEF} > \kappa_{MG}$ *holds.*

Case (a) says that if M^{-1} *is a 'bad' smoother (no eigenvalues of* $I - M^{-1}A$ *are less than* $\frac{1}{2}$*), then MG is expected to converge worse than DEF. On the other hand, Case (b) implies that if* M^{-1} *is a 'good' smoother (all eigenvalues that need to be handled by relaxation are done so with eigenvalues of* $I - M^{-1}A$ *bounded in a neigbourhood on both sides of the origin), then MG converges better than DEF.*

6. Effect of Relaxation Parameters. In the previous section, we have seen that DEF can be more favorable to use than MG in certain circumstances. In this section, it will be shown that MG can always have a lower condition number than DEF, in the case where $M^{-1} = \bar{M}^{-1} = \alpha I$ is used with an optimal choice of α. The relaxation scheme with this choice of M^{-1} is known as Richardson relaxation.

6.1. Analysis of Scaling Relaxation. Instead of considering the original linear system (1.1), we now consider the scaled linear system:

$$\alpha A x = \alpha b, \quad \alpha > 0, \tag{6.1}$$

with $M^{-1} = \bar{M}^{-1} = I$. A subscript, α, will be added to the notation for operators and matrices, if they are for (6.1). So, P_α and $\mathcal{P}_{\mathrm{MG},\alpha}$ denote the deflation matrix and MG-preconditioner based on (6.1), respectively.

Solving the scaled linear system (6.1) with $M^{-1} = \bar{M}^{-1} = I$ is equivalent to solving the preconditioned linear system (1.3) with $M^{-1} = \bar{M}^{-1} = \alpha I$. This follows from the error-propagation operator, V_α:

$$\begin{aligned} V_\alpha := (I - \mathcal{P}_{\text{MG},\alpha}\alpha A) &= (I - I\alpha A)P_\alpha^T(I - I\alpha A) \\ &= (I - (\alpha I)A)P^T(I - (\alpha I)A) \\ &= (I - \bar{M}^{-T}A)P^T(I - \bar{M}^{-1}A) \\ &= V, \end{aligned} \tag{6.2}$$

with $\bar{M}^{-1} = \alpha I$. The parameter, α, can then be regarded as a parameter of the relaxation instead of the linear system. Relaxation methods are typically given by fixed formulas (e.g., $\bar{M} = \text{diag}(A)$ is used in Jacobi iterations and Gauss-Seidel iterations take the lower-triangular part of A), but these formulas are often not optimally scaled. So, the relaxation processes are rescaled, whereas coarse-grid correction remains untouched due to the fact that

$$P_\alpha = I - (\alpha A)Z(Z^T(\alpha A)Z)^{-1}Z^T = I - AZ(Z^TAZ)^{-1}Z^T = P.$$

Therefore, DEF is scaling invariant:

$$\kappa_{\text{DEF},\alpha} = \frac{\lambda_n(M^{-1}P_\alpha\alpha A)}{\lambda_{k+1}(M^{-1}P_\alpha\alpha A)} = \frac{\lambda_n(M^{-1}PA)}{\lambda_{k+1}(M^{-1}PA)} = \kappa_{\text{DEF}}.$$

In contrast, MG is not scaling invariant, and the positive-definiteness property of $\mathcal{P}_{\text{MG},\alpha}$ depends strongly on α, see Theorem 6.1.

THEOREM 6.1. *Assume that $\bar{M}^{-1} = M^{-1} = I$ and $Z = [v_1 \cdots v_k]$, where $\{v_i\}$ is the set of orthonormal eigenvectors corresponding to the increasing set of eigenvalues $\{\lambda_i\}$ of A. Then, $\mathcal{P}_{MG,\alpha}A$ is only SPD if*

$$0 < \alpha < \frac{2}{\lambda_{\max}(A)} = \frac{2}{||A||_2},$$

where $\lambda_{\max}(A)$ is the largest eigenvalue of A not treated by coarse-grid correction.

Proof. We have

$$\mathcal{P}_{\text{MG},\alpha}\alpha A v_i = \alpha\left[2\lambda_i v_i - 2\lambda_i ZZ^Tv_i + \tfrac{\lambda_i}{\alpha}Z\Lambda^{-1}Z^Tv_i - \alpha\lambda_i^2 v_i + \alpha\lambda_i^2 ZZ^Tv_i\right],$$

which is the same expression as given in [3, Prop. 2]. Hence,

$$\mathcal{P}_{\text{MG}}Av_i = \begin{cases} v_i, & \text{for } i = 1,\ldots,k; \\ \alpha\left[\lambda_i(2 - \alpha\lambda_i)\right]v_i, & \text{for } i = k+1,\ldots,n. \end{cases} \tag{6.3}$$

From (6.3), we derive that $\mathcal{P}_{\text{MG},\alpha}A$ is only SPD if $\alpha\left[\lambda_i(2 - \alpha\lambda_i)\right] > 0$, i.e.,

$$0 < \alpha < \frac{2}{\lambda_{\max}} = \frac{2}{||A||_2},$$

for λ_i not treated by coarse-grid correction (cf. Theorems 3.7 and 3.8). ∎

REMARK 6.2.

- *Eq. (6.3) clearly shows that $\kappa_{MG,\alpha} \neq \kappa_{MG}$.*
- *For any A, there exists an α such that $\mathcal{P}_{MG,\alpha}$ is SPD.*

- *The function $\alpha\,[\lambda_i(2-\alpha\lambda_i)]$ has a maximum at $\alpha=\frac{1}{\lambda_i}$, and*

$$\lim_{\alpha\to 0}\kappa_{MG,\alpha}|_{Z^\perp}=\frac{\max_{k+1\le i\le n}\alpha\lambda_i(2-\alpha\lambda_i)}{\min_{k+1\le i\le n}\alpha\lambda_i(2-\alpha\lambda_i)}=\frac{\max_{k+1\le i\le n}\lambda_i}{\min_{k+1\le i\le n}\lambda_i}=\frac{\lambda_n}{\lambda_{k+1}}=\kappa_{DEF}.$$

  This means that, for $\alpha\to 0$, the condition number of $\mathcal{P}_{MG}A$ restricted to the complement of Z (i.e., omitting the unit eigenvalues) is equal to the condition number of $\mathcal{P}_{DEF}A$.

Theorem 6.1 shows that the eigenvalue distribution of MG highly depends on the choice of α. This is further illustrated in the following example.

EXAMPLE 6.3. *We consider the 1-D Laplace equation with the following natural discretization matrix (multiplied by h^2):*

$$A=\begin{bmatrix}2 & -1 & & 0\\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1\\ 0 & & -1 & 2\end{bmatrix}, \tag{6.4}$$

We take $n=10$, $M^{-1}=I$ and $Z\in\mathbb{R}^{10\times 2}$ consisting of two independent random vectors. The scaling parameter, α, will be varied. Note that the eigenvalues of A are in the range of $(0,4]$ and $||A||_2=\lambda_{\max}\approx 4$. Therefore, we expect that $\alpha\le\frac{1}{2}$ gives positive eigenvalues for MG and $\alpha\approx\frac{1}{3}$ is optimal, which can be motivated as follows. The spectrum of $I-\alpha A$ ranges from about 1 (for the smallest eigenvalue of A) to approximately $1-4\alpha$ (for the largest eigenvalue of A). With a classical MG choice of coarse grid, we would expect the coarse-grid correction to accurately correct errors associated with the eigenvectors of eigenvalues between 1 and $1-2\alpha$ (the lower half of the spectrum of A). So, α should be chosen so that convergence of relaxation is as good as possible for the remaining eigenvectors. Since these are bounded between $1-2\alpha$ and $1-4\alpha$, this can be accomplished by choosing α so that $|1-2\alpha|=|1-4\alpha|$, which happens with $\alpha=\frac{1}{3}$.

The spectra of DEF and MG are denoted by

$$\sigma_{DEF}=(0,0,\lambda_1,\ldots,\lambda_8),\quad \sigma_{MG}=(1,1,\mu_1,\ldots,\mu_8).$$

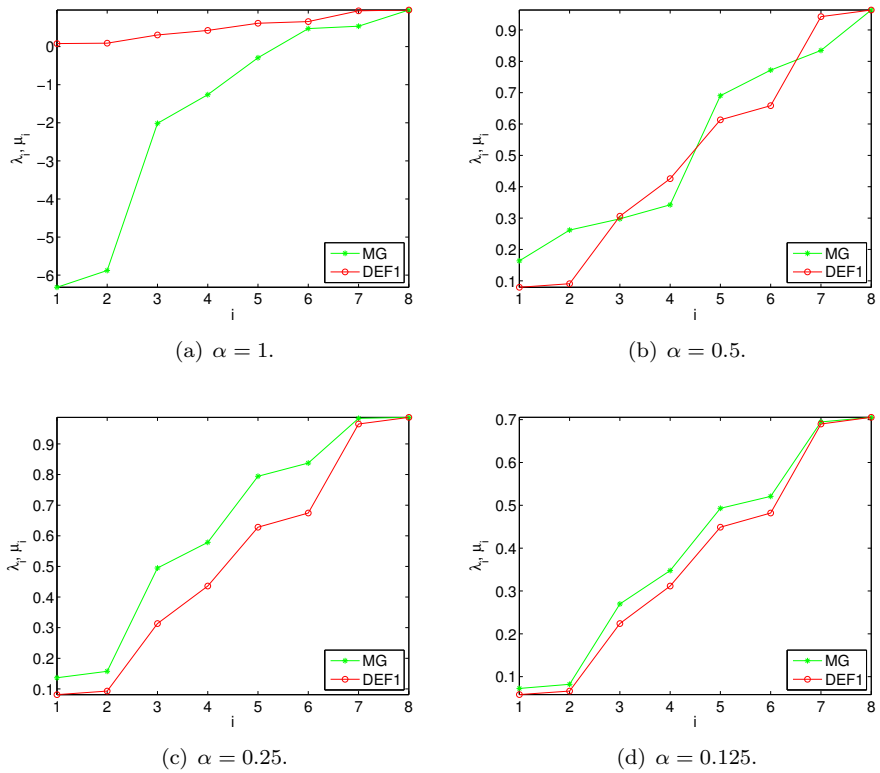
For the sake of convenience, we scale the eigenvalues of DEF by $\frac{\mu_8}{\lambda_8}$. We do not lose generality, since DEF is scaling invariant. We then get

$$\sigma_{DEF}=(0,0,\frac{\lambda_1\mu_8}{\lambda_8},\frac{\lambda_2\mu_8}{\lambda_8},\ldots,\mu_8),\quad \sigma_{MG}=(1,1,\mu_1,\ldots,\mu_8).$$

Now, we plot the eigenvalues corresponding to the eigenvectors restricted to the complement of Z (i.e., the zeros in DEF and ones in MG are omitted), see Figure 6.1.

Figure 6.1 shows that MG gives positive eigenvalues for $\alpha\le\frac{1}{2}$. In this case, MG is better conditioned, since its smallest eigenvalue is larger than the smallest eigenvalue of DEF. Moreover, for α close to $\frac{1}{2}$, the eigenvalues of MG and DEF are not obviously related. For smaller α, the eigenvalues of MG are located above those of DEF. In addition, they converge to the eigenvalues of DEF (except for the zero and unit eigenvalues).

Subsequently, the following lemma is useful for general Z and shows that Richardson relaxation is always convergent for a sufficiently small α. This result can also be found in [26].



(a) $\alpha = 1$. (b) $\alpha = 0.5$.

(c) $\alpha = 0.25$. (d) $\alpha = 0.125$.

FIG. 6.1. *Eigenvalues of DEF and MG corresponding to the eigenvectors restricted to the complement of Z.*

LEMMA 6.4. *Suppose that $M^{-1} = \bar{M}^{-1} = \alpha I$. Richardson relaxation is convergent if*

$$0 < \alpha < \frac{2}{||A||_2}.$$

Proof. If Richardson relaxation is convergent, then the equality

$$||I - \alpha A||_2 < 1$$

should be satisfied. Since $I - \alpha A$ is symmetric, we have

$$||I - \alpha A||_2 = \max\left\{|\lambda_{\min}(I - \alpha A)|, |\lambda_{\max}(I - \alpha A)|\right\}.$$

This implies

$$-1 < \lambda_{\min}(I - \alpha A) \leq \lambda_{\max}(I - \alpha A) < 1. \tag{6.5}$$

Combining (6.5) with

$$\lambda_{\min}(I - \alpha A) = 1 - \alpha\lambda_{\max}(A), \quad \lambda_{\max}(I - \alpha A) = 1 - \alpha\lambda_{\min}(A)$$

yields $\alpha < \frac{2}{||A||_2}$ and $\alpha > 0$, using the fact that $||A||_2 = \lambda_{\max}(A)$ for an SPD matrix, A. ∎

In Theorem 6.1, we have already seen that the above theorem holds if eigenvectors are used in Z. Moreover, for multigrid, we typically try to choose α close to $\frac{1}{||A||_2}$, which guarantees that the slow-to-converge modes of relaxation are only those associated with the small eigenvalues of A. A better choice of α is possible if we make assumptions on how the eigenvectors of A associated with small eigenvalues are treated by coarse-grid correction. It is also be possible to get an explicit expression for the optimal α, see the next subsection.

6.2. Optimal Choice of α. The best value of α depends on Z, so the optimal α, denoted as α_{opt}, can only be determined if the choice of Z is fixed. In this case, the job of relaxation is specifically to reduce errors that are conjugate to the range of Z. The best choice of α is the one that minimizes the 'spectral radius' of relaxation over the complement of the range of interpolation, i.e.,

$$\min_{x, y^T Z^T A x = 0 \ \forall y} \frac{|x^T (I - \alpha A) x|}{x^T x}.$$

If we restrict ourselves to Z consisting of eigenvectors of A, parameter α_{opt} is easily determined such that it gives the most favorable condition number for MG, see the next theorem.

THEOREM 6.5. *Suppose that $M^{-1} = \bar{M}^{-1} = \alpha I$ and $\{\lambda_i\}$ is the increasingly-sorted set of eigenvalues of $M^{-1}A$ with corresponding eigenvectors $\{v_i\}$. Let Z be decomposed of k orthonormal eigenvectors from $\{v_i\}$. Moreover, let $\mathcal{P}_{MG}$ be as given in (2.7) such that $\mathcal{P}_{MG}A$ is SPD. Then, $\kappa(\mathcal{P}_{MG,\alpha}A)$ is minimized for*

$$\alpha_{opt} = \frac{2}{\lambda_{k+1} + \lambda_n}. \tag{6.6}$$

Proof. Note first that, by choosing $\bar{M}^{-1} = M^{-1} = \alpha I$, the error-propagation operator for MG can be written as (cf. Eq. (3.4)).

$$\begin{aligned} I - \mathcal{P}_{\text{MG}} A &= (I - \alpha A) P^T (I - \alpha A) \\ &= I - \left(2\alpha I + Q - \alpha QA - \alpha AQ - \alpha^2 A + \alpha^2 AQA\right) A \\ &= 2\alpha I + Z\Lambda^{-1} Z^T - 2\alpha Z Z^T - \alpha^2 A + \alpha^2 Z \Lambda Z^T. \end{aligned}$$

So, applying $\mathcal{P}_{\text{MG}}$ to an eigenvector, v_i, of A gives (cf. Eq. (3.5))

$$\mathcal{P}_{\text{MG}} v_i = \begin{cases} \frac{1}{\lambda_i} v_i, & \text{for } i = 1, \ldots, k; \\ \alpha(2 - \alpha\lambda_i), & \text{for } i = k+1, \ldots, n. \end{cases}$$

Thus, $\mathcal{P}_{\text{MG}}A$ has eigenvalue 1 with algebraic multiplicity k, and $n-k$ eigenvalues of the form $\alpha\lambda_i(2 - \alpha\lambda_i)$, for $i = k+1, \ldots, n$.

Let $\{\sigma_i\}$ be the set of eigenvalues of $\mathcal{P}_{\text{MG}}A$, which are positive and sorted increasingly, so that its condition number is given by $\frac{\sigma_n}{\sigma_1}$. By assumption, $\alpha\lambda_i(2 - \alpha\lambda_i) > 0$ for all $i = k+1, \ldots, n$ and, by calculation, $\alpha\lambda_i(2 - \alpha\lambda_i) < 1$ for all α and λ_i. Thus,

$$\sigma_1 = \min_{i \in [k+1, n]} \{\alpha\lambda_i(2 - \alpha\lambda_i)\}, \quad \sigma_n = 1.$$

Now, assume that the eigenvalues, λ_i, for $i = k+1, \ldots, n$, are more-or-less evenly distributed between λ_{k+1} and λ_n (we do not lose generality by this assumption). Then, we can approximate the best choice of α by doing a continuous optimization

over α and $\lambda \in [\lambda_{k+1}, \lambda_n]$. For fixed α, the minimal value of $\alpha\lambda(2-\alpha\lambda)$ is obtained either at one of the endpoints in λ or at the critical point, where $\frac{d}{d\lambda}(\alpha\lambda(2-\alpha\lambda)) = 0$, which occurs when $\lambda = \frac{1}{\alpha}$, when this is in the admissible range of λ's. However, the function is always concave down: $\frac{d^2}{d\lambda^2}(\alpha\lambda(2-\alpha\lambda)) = -2\alpha^2 < 0$, which implies that the minimum must occur at an end point. So,

$$\min_{\lambda\in[\lambda_{k+1},\lambda_n]} \alpha\lambda(2-\alpha\lambda) = \min\left\{\alpha\lambda_{k+1}(2-\alpha\lambda_{k+1}), \alpha\lambda_n(2-\alpha\lambda_n)\right\}. \tag{6.7}$$

Subsequently, we want to maximize this minimum eigenvalue (since the minimum eigenvalue should be as large as possible),

$$\max_{\alpha} \min\left\{\alpha\lambda_{k+1}(2-\alpha\lambda_{k+1}), \alpha\lambda_n(2-\alpha\lambda_n)\right\}.$$

This is achieved when we choose α so that

$$\alpha\lambda_{k+1}(2-\alpha\lambda_{k+1}) = \alpha\lambda_n(2-\alpha\lambda_n),$$

which occurs when $\alpha = \frac{2}{\lambda_{k+1}+\lambda_n}$. ☐

COROLLARY 6.6. *Let the conditions of Theorem 6.5 be satisfied. For α_{opt}, the smallest eigenvalue of $\mathcal{P}_{MG}A$ is equal to*

$$\frac{4\lambda_{k+1}\lambda_n}{(\lambda_{k+1}+\lambda_n)^2}. \tag{6.8}$$

The condition number of $\mathcal{P}_{MG}A$ is then given by

$$\kappa_{MG} = \frac{1}{\frac{4\lambda_{k+1}\lambda_n}{(\lambda_{k+1}+\lambda_n)^2}} = \frac{(\lambda_{k+1}+\lambda_n)^2}{4\lambda_{k+1}\lambda_n}, \tag{6.9}$$

so that $\kappa_{MG} \leq \kappa_{DEF}$.

Proof. Eq. (6.8) follows by substituting the optimal weighting parameter, α_{opt}, into (6.7). Then, this implies Eq. (6.9). Finally, $\kappa_{\text{MG}} \leq \kappa_{\text{DEF}}$ follows from the fact that

$$\frac{(\lambda_{k+1}+\lambda_n)^2}{4\lambda_{k+1}\lambda_n} \leq \frac{\lambda_n}{\lambda_{k+1}} \quad \Leftrightarrow \quad (\lambda_{k+1}+\lambda_n)^2 \leq (2\lambda_n)^2,$$

which is always true, since $\lambda_{k+1} \leq \lambda_n$. ☐

REMARK 6.7.

- *The condition numbers corresponding to MG and DEF are the same if the spectrum of A is 'flat' (i.e., if $\lambda_{k+1} = \lambda_n$). But, in general, MG gives a more favorable condition number than deflation, using the optimal parameter, α_{opt}.*
- *Examples 4.2 and 5.2 have shown that $\kappa_{MG} \geq \kappa_{DEF}$ can happen in general. However, according to Theorem 6.5, these examples can never be constructed if α_{opt} is used.*
- *In practice, approximations to α are fairly easy to compute, although the exact eigenvalue distribution is usually unknown. Eigenvalue λ_n can be approximated using Gershgorin's theorem, while λ_{k+1} can be taken to be approximately $\frac{\lambda_1+\lambda_n}{2}$ with $\lambda_1 \approx 0$, since $k \approx \frac{n}{2}$ usually holds in multigrid.*

- *An optimal weighting parameter, α_{opt}, can also be considered for general preconditioners, $\bar{M}^{-1}$; however, it is often much more difficult to express α_{opt} explicitly, as it depends on the spectral properties of $\bar{M}^{-1}A$, which may not be known. In general, the optimal choice of α is such that relaxation converges as quickly as possible on the modes that are not being treated by the coarse-grid correction phase of the cycle. Thus, if the spectral picture of $\bar{M}^{-1}A$ is known well-enough to approximate the eigenvalues corresponding to λ_{k+1} and λ_n, a similar choice of α_{opt} as in Eq. (6.8) may be possible.*

7. Symmetrizing the Smoother. In the previous section, we have seen that MG can be expected to converge in fewer iterations than DEF for simple choices of M^{-1} and $\bar{M}^{-1}$. Numerical experiments show that MG also requires fewer iterations than DEF for more sophisticated choices of preconditioners, see Section 8. However, the fact that MG requires fewer iterations than DEF for many preconditioners does not mean that it is more efficient, since each iteration of MG is more expensive, due to the choice of two smoothing steps. In order to make a fairer comparison between DEF and MG, we now consider DEF using the preconditioning version of the symmetrized smoother:

$$\begin{aligned} S^*S &= (I - \bar{M}^{-T}A)(I - \bar{M}^{-1}A) \\ &= (I - \bar{M}^{-T}(\bar{M} + \bar{M}^T - A)\bar{M}^{-1}A) \\ &= I - \widetilde{M}^{-1}A, \end{aligned} \tag{7.1}$$

with

$$\widetilde{M}^{-1} := \bar{M}^{-T}(\bar{M} + \bar{M}^T - A)\bar{M}^{-1} = \bar{M}^{-1} + \bar{M}^{-T} - \bar{M}^{-T}A\bar{M}^{-1}. \tag{7.2}$$

Note that $\widetilde{M}^{-1}$, as defined here, is the same as in Eq. (3.10). Then, we use

$$M^{-1} := \widetilde{M}^{-1} \tag{7.3}$$

as the preconditioner in DEF, since this choice allows implementation in such a way that each iteration of BNN, DEF and MG has similar cost. In this section, we will compare the spectra associated with MG and DEF using (7.3). First, we show that these spectra are almost equal for simple choices of $\bar{M}^{-1}$ with Z consisting of eigenvectors of A. Moreover, DEF gives a somewhat more favorable condition number than MG in this setting. Thereafter, we will generalize this latter result for general $\bar{M}^{-1}$ and Z.

7.1. Spectral Analysis of MG and DEF. Assume that $\bar{M}^{-1} = \alpha I$ and the columns of Z are eigenvectors of A. Then, we can show that the spectra corresponding to MG and DEF are almost the same if Eq. (7.3) is used, see the next theorem.

THEOREM 7.1. *Assume that $\bar{M}^{-1} = \alpha I$, Eqs. (7.2) and (7.3) hold and the columns of Z are some eigenvectors corresponding to eigenvalues of A, which lie in the interval $(0, \frac{2}{\alpha})$. Let $\mathcal{P}_{DEF}$ and $\mathcal{P}_{MG}$ be as given in (2.3) and (2.7), respectively. Then, the eigenvalues of $\mathcal{P}_{DEF}A$ and $\mathcal{P}_{MG}A$ are the same, except that the unit eigenvalues of MG, corresponding to the columns of Z, are zero eigenvalues of DEF.*

Proof. Substituting $\bar{M}^{-1} = \alpha I$ into (7.2) yields

$$\widetilde{M}_\alpha^{-1} := \alpha^2 \left(\frac{2}{\alpha} I - A \right).$$

Because the eigenvalues of A lie in the interval $(0, \frac{2}{\alpha})$, both DEF and MG are positive-definite (cf. Theorem 6.1). This yields

$$\mathcal{P}_{\text{DEF}}A = \widetilde{M}_{\alpha}^{-1}PA = \alpha^2\left(\frac{2}{\alpha}I - A\right)(I - ZZ^T)A = \alpha^2\left(\frac{2}{\alpha}A - \frac{2}{\alpha}ZZ^TA - A^2 + AZZ^TA\right).$$

This implies

$$\mathcal{P}_{\text{DEF}}Av_i = \alpha^2\left(\frac{2}{\alpha}\lambda_i v_i - \frac{2}{\alpha}\lambda_i ZZ^Tv_i - \lambda_i^2 v_i + \lambda_i AZZ^Tv_i\right).$$

Then, we derive

$$\mathcal{P}_{\text{DEF}}Av_i = \begin{cases} \alpha^2\left(\frac{2}{\alpha}\lambda_i v_i - \frac{2}{\alpha}\lambda_i v_i - \lambda_i^2 v_i + \lambda_i^2 v_i\right) &= 0, & \text{for } i = 1,\ldots,k; \\ \alpha^2\left(\frac{2}{\alpha}\lambda_i v_i - \lambda_i^2 v_i\right) &= \alpha\left(\lambda_i(2-\alpha\lambda_i)\right)v_i, & \text{for } i = k+1,\ldots,n. \end{cases}$$

Recall that (see Eq. (6.3))

$$\mathcal{P}_{\text{MG}}Av_i = \begin{cases} v_i, & \text{for } i = 1,\ldots,k; \\ \alpha\left(\lambda_i(2-\alpha\lambda_i)\right)v_i, & \text{for } i = k+1,\ldots,n, \end{cases}$$

which shows that the eigenvalues of $\mathcal{P}_{\text{DEF}}A$ and $\mathcal{P}_{\text{MG}}A$ are the same for $i = k+1,\ldots,n$ and the theorem follows. □

Surprisingly, the eigenvalues not treated by the coarse-grid correction or deflation matrix in MG and DEF are identical. If one considers the condition numbers, then we get

$$\kappa_{\text{DEF}} = \frac{\max_{k+1\le i\le n}\{\lambda_i(2\alpha-\lambda_i)\}}{\min_{k+1\le i\le n}\{\lambda_i(2\alpha-\lambda_i)\}} \le \frac{1}{\min_{k+1\le i\le n}\{\alpha\lambda_i(2-\alpha\lambda_i)\}} = \kappa_{\text{MG}},$$

since $\lambda_i(2\alpha-\lambda_i) \le 1$ for all $i = k+1,\ldots,n$ and $0 < \alpha \le 1$ is assumed. Therefore, DEF has a better condition number than MG.

REMARK 7.2. *A 'reasonable' choice of α often leads to $\max_{k+1\le i\le n}\{\alpha\lambda_i(2\alpha-\lambda_i)\}$ becoming quite close to 1. If this is the case, then the condition numbers for DEF and MG are almost identical. Since we typically choose α to be close to the inverse of the average of λ_{k+1} and λ_n, we expect this to be the case. In other words, DEF is guaranteed to be better conditioned than MG, but only slightly.*

7.2. Spectral Analysis for General $\bar{M}^{-1}$ and Z. In the previous subsection, we have shown that, for special choices of $\bar{M}^{-1}$ and Z, DEF with preconditioner $\widetilde{M}^{-1}$ and MG yield the same eigenvalues for those modes that are not treated by the coarse-grid correction. From [19], it follows that the eigenvalue distributions of BNN with $\widetilde{M}^{-1}$ and MG are exactly the same, resulting in the same condition numbers. Below, we show that this result also holds for general Z and $\bar{M}^{-1}$ such that $\widetilde{M}^{-1}$ is SPD, resulting in the fact that S is certainly convergent in the A-norm by Theorem 3.8.

THEOREM 7.3. *Let $\bar{M}^{-1} \in \mathbb{R}^{n\times n}$ be as given in (2.1) such that $\mathcal{P}_{MG}$ is SPD. In addition, let $M^{-1} = \widetilde{M}^{-1}$ be as defined in (7.2) such that $\mathcal{P}_{BNN}$ is SPD. Then, the eigenvalues of $\mathcal{P}_{MG}A$ and $\mathcal{P}_{BNN}A$ are equal.*

Proof. We show the equivalence of κ_{MG} and κ_{BNN}, by examining the extreme eigenvalues of their error propagation forms,

$$\begin{cases} I - \mathcal{P}_{\text{MG}}A &= S^*P^TS; \\ I - \mathcal{P}_{\text{BNN}}A &= P^T(I - \widetilde{M}^{-1}A)P^T. \end{cases}$$

We examine both methods by making the same similarity transformation,

$$I - \mathcal{P}A \rightarrow A^{\frac{1}{2}}(I - \mathcal{P}A)A^{-\frac{1}{2}},$$

where $A^{1/2}$ is the symmetric square root of the SPD matrix, A. This allows us to make use of the fact that $I - A^{\frac{1}{2}}QA^{\frac{1}{2}}$ is an orthogonal projection in the $L2$-inner product (i.e., $(I - A^{\frac{1}{2}}QA^{\frac{1}{2}})^T(I - A^{\frac{1}{2}}QA^{\frac{1}{2}}) = I - A^{\frac{1}{2}}QA^{\frac{1}{2}}$). Computing the similarity transformed systems, we have

$$\begin{cases} A^{\frac{1}{2}}(I - \mathcal{P}_{\text{MG}}A)A^{-\frac{1}{2}} &= (I - A^{\frac{1}{2}}\bar{M}^{-T}A^{\frac{1}{2}})(I - A^{\frac{1}{2}}QA^{\frac{1}{2}})(I - A^{\frac{1}{2}}\bar{M}^{-1}A^{\frac{1}{2}}), \\ A^{\frac{1}{2}}(I - \mathcal{P}_{\text{BNN}}A)A^{-\frac{1}{2}} &= (I - A^{\frac{1}{2}}QA^{\frac{1}{2}})(I - A^{\frac{1}{2}}\widetilde{M}^{-1}A^{\frac{1}{2}})(I - A^{\frac{1}{2}}QA^{\frac{1}{2}}). \end{cases}$$

We define

$$C := (I - A^{\frac{1}{2}}QA^{\frac{1}{2}})(I - A^{\frac{1}{2}}M^{-1}A^{\frac{1}{2}}),$$

so that we can write

$$\begin{cases} A^{\frac{1}{2}}(I - \mathcal{P}_{\text{MG}}A)A^{-\frac{1}{2}} &= C^TC; \\ A^{\frac{1}{2}}(I - \mathcal{P}_{\text{BNN}}A)A^{-\frac{1}{2}} &= CC^T, \end{cases}$$

using the following the equalities:

$$\begin{cases} (I - A^{\frac{1}{2}}QA^{\frac{1}{2}})^2 &= I - A^{\frac{1}{2}}QA^{\frac{1}{2}}; \\ (I - A^{\frac{1}{2}}QA^{\frac{1}{2}})^T &= I - A^{\frac{1}{2}}QA^{\frac{1}{2}}; \\ (I - A^{\frac{1}{2}}\bar{M}^{-1}A^{\frac{1}{2}})^T &= I - A^{\frac{1}{2}}\bar{M}^{-T}A^{\frac{1}{2}}; \\ I - A^{\frac{1}{2}}\widetilde{M}^{-1}A^{\frac{1}{2}} &= (I - A^{\frac{1}{2}}\bar{M}^{-T}A^{\frac{1}{2}})(I - A^{\frac{1}{2}}\bar{M}^{-1}A^{\frac{1}{2}}). \end{cases}$$

Since $A^{\frac{1}{2}}(I - \mathcal{P}_{\text{MG}}A)A^{-\frac{1}{2}}$ and $A^{\frac{1}{2}}(I - \mathcal{P}_{\text{BNN}}A)A^{-\frac{1}{2}}$ are similar to $I - \mathcal{P}_{\text{MG}}A$ and $I - \mathcal{P}_{\text{BNN}}A$, respectively, and, $\sigma(C^TC) = \sigma(CC^T)$ (see, e.g., [19]), we obtain

$$\sigma(I - \mathcal{P}_{\text{MG}}A) = \sigma(C^TC) = \sigma(I - \mathcal{P}_{\text{BNN}}A),$$

and the theorem follows immediately. □

From Theorem 7.3, we obtain that MG and BNN with $\widetilde{M}^{-1}$ give exactly the same condition number. This also implies that the condition number of MG is not smaller than the condition number of DEF, see the next corollary.

COROLLARY 7.4. *Let $\bar{M}^{-1}$ and $M^{-1} = \widetilde{M}^{-1}$ be as in Theorem 7.3 such that $\mathcal{P}_{DEF}$ is SPD. Then,*

(i) $\kappa_{MG} = \kappa_{BNN}$;

(ii) $\kappa_{DEF} \leq \kappa_{MG}$,

where $\kappa_{MG}, \kappa_{BNN}$ and κ_{DEF} are the condition numbers corresponding to MG, BNN and DEF, respectively.

Proof. (i) This is a trivial consequence of Theorem 7.3.

(ii) This result follows by combining Theorem 7.3 and [15, Thm. 2.7]. □

REMARK 7.5.

- *Ordering the smoothers in the opposite way might lead to a different definition of $\widetilde{M}^{-1}$; this, in turn, could change the eigenvalues of MG and BNN, although an analogous result to Theorem 7.3 still holds for the consistent choice of S and $\widetilde{M}^{-1}$.*
- *Corollary 7.4 shows that BNN, DEF and MG are expected to show comparable convergence behavior for special choices of traditional preconditioners. We note that this result is only valid in exact arithmetic. If coarse-grid systems are solved inaccurately, DEF might have convergence difficulties, while BNN and MG are less sensitive to it, see, e.g., [19].*

8. Numerical Experiments. In this section, we present the results of some numerical experiments, where PREC and the presented 2L-PCG methods (i.e., DEF, BNN and MG) will be compared. We start with a 1-D Laplace-like problem to illustrate the theory as obtained in Section 5. Then, a 2-D bubbly flow problem will be considered to show the performance of DEF, BNN and MG in a more realistic setting. We stress that these examples are chosen to highlight the presented theory and not to present the efficiency of the solvers; in practice, very different choices of $\bar{M}^{-1}$, M^{-1} and Z are used for each method, see [11, 23].

The starting vector for each iterative method will be arbitrary and the iterative process will be terminated if the norm of the relative residual, $\frac{||r_{j+1}||_2}{||r_0||_2}$, falls below a fixed tolerance, $\delta = 10^{-8}$.

8.1. 1-D Laplace-like Problem. Several 1-D Laplace-like problems are considered, with the matrices

$$A = \begin{bmatrix} \beta & \gamma & & \emptyset \\ \gamma & \beta & \ddots & \\ & \ddots & \ddots & \gamma \\ \emptyset & & \gamma & \beta \end{bmatrix}, \quad \beta, \gamma \in \mathbb{R}, \tag{8.1}$$

where we vary the constants β and γ so that each test case corresponds to a different region as shown in Figure 5.1, see Table 8.1. In addition, we choose $\bar{M}^{-1} = M^{-1} = I$ and Z consisting of eigenvectors corresponding to the smallest eigenvalues of A. Right-hand side, b, is chosen randomly. We take $n = 100$ (other values of n lead to approximately the same results) and the number of projection vectors, k, will be varied. The results of the experiment can be found in Table 8.2.

Test Problem	β	γ	Range of λ_i	Region	Expected Fastest Method
(T1)	1.5	-0.125	[1.25, 1.75]	B2	DEF
(T2)	1	-0.05	[0.9, 1.1]	A1 / A2	MG
(T3)	0.25	-0.1	[0.05, 0.45]	B1	DEF
(T4)	1.25	-0.125	[1.0, 1.5]	A1 / A2	MG/DEF

TABLE 8.1
Test cases corresponding to different regions as presented in Figure 5.1.

From Table 8.2, the following observations can be made.

- Table 8.2(a): DEF yields a smaller condition number and is always faster than MG in this case.
- Table 8.2(b): MG yields a smaller condition number and is always faster than DEF.
- Table 8.2(c): Since the condition number associated with DEF is always below that of MG, DEF is expected to be faster than MG. However, that is not the case in this test problem. The two methods converge at the same rate for large k, but MG is faster than DEF for small k. This can be explained by the fact that the spectrum of eigenvalues of MG consists of two clusters, see Figure 8.1(c). If the first cluster of ones is omitted (or are approximated by a Ritz value), then the condition number of the remaining spectrum is more favorable compared to the condition number associated with DEF. For example, in the case of $k = 2$, we have $\kappa_{\text{MG}} = 7.0$ (instead of $\kappa_{\text{MG}} = 10.1$)

(a) $\beta = 1.5,\ \gamma = -0.125$.

	$k=2$		$k=20$		$k=60$	
Method	# It.	κ	# It.	κ	# It.	κ
PREC	11	1.4	11	1.4	11	1.4
DEF	11	1.4	10	1.3	8	1.1
BNN	11	1.7	10	1.7	8	1.7
MG	15	2.3	15	2.3	12	2.3

(b) $\beta = 1,\ \gamma = -0.05$.

	$k=2$		$k=20$		$k=60$	
Method	# It.	κ	# It.	κ	# It.	κ
PREC	9	1.2	9	1.2	9	1.2
DEF	9	1.2	9	1.2	7	1.1
BNN	9	1.2	9	1.2	7	1.1
MG	5	1.01	5	1.01	5	1.01

(c) $\beta = 0.25,\ \gamma = -0.1$.

	$k=2$		$k=20$		$k=60$	
Method	# It.	κ	# It.	κ	# It.	κ
PREC	34	9.0	34	9.0	34	9.0
DEF	34	8.8	24	4.9	11	1.4
BNN	34	19.6	25	11.0	11	3.2
MG	30	10.1	22	5.7	11	1.9

(d) $\beta = 1.25,\ \gamma = -0.125$.

	$k=2$		$k=20$		$k=60$	
Method	# It.	κ	# It.	κ	# It.	κ
PREC	11	1.5	11	1.5	11	1.5
DEF	12	1.5	11	1.4	8	1.1
BNN	12	1.5	11	1.5	8	1.5
MG	10	1.3	10	1.3	9	1.3

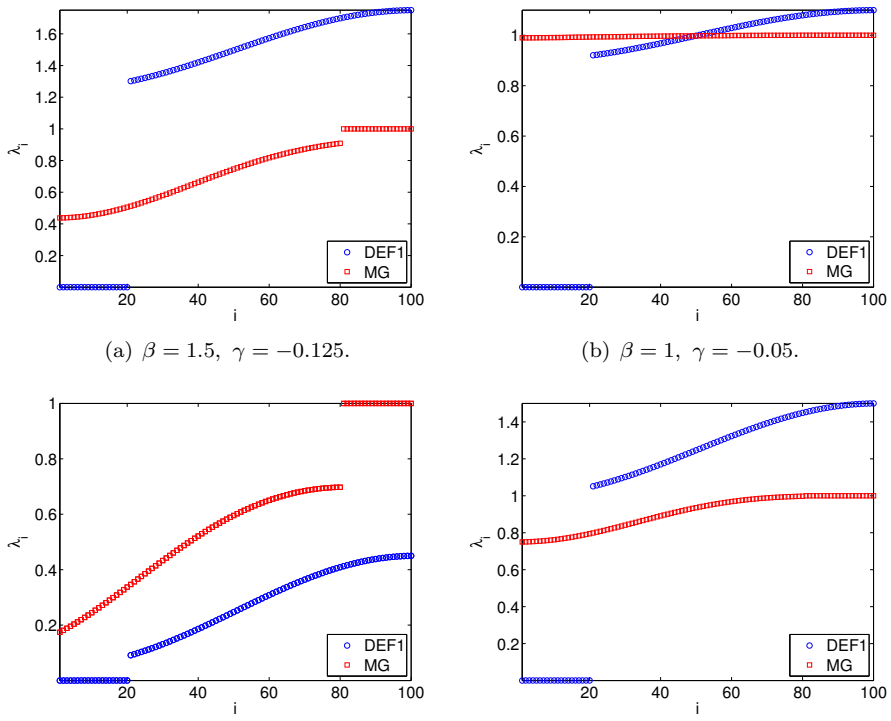
TABLE 8.2

Results of the experiment with test cases as presented for the Laplace-like problem in Table 8.1. The results are presented in terms of number of iterations, # It., and condition number, κ.

when the unit eigenvalues are omitted. Obviously, this would then be the smallest condition number over all of the methods.

- Table 8.2(d): MG has a smaller condition number and is faster than DEF for small k. On the other hand, for large k, DEF has a smaller condition number than MG and performs somewhat better than MG. Indeed, the best method depends on λ_{k+1}.

8.2. 2-D Bubbly Flow Problem. Using pressure-correction techniques for solving the Navier-Stokes equations, the major computational bottleneck in modelling the bubbly flow problem is the solution of the Poisson equation with a discontinuous



(a) $\beta = 1.5,\ \gamma = -0.125$. (b) $\beta = 1,\ \gamma = -0.05$.

(c) $\beta = 0.25,\ \gamma = -0.1$. (d) $\beta = 1.25,\ \gamma = -0.125$.

FIG. 8.1. *Eigenvalues associated with DEF and MG for the test cases with $k = 20$ as presented in Table 8.2.*

coefficient,

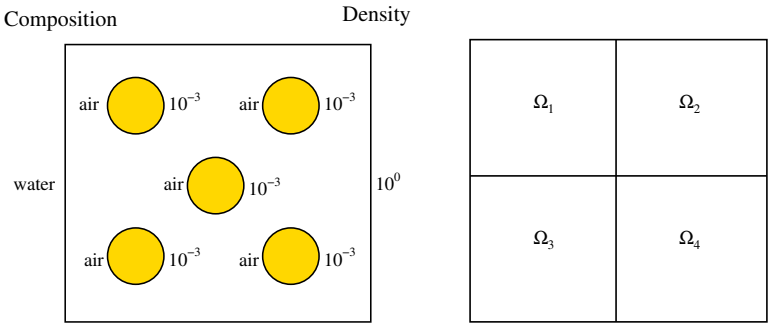
$$-\nabla \cdot \left(\frac{1}{\rho(\mathbf{x})} \nabla p(\mathbf{x}) \right) = 0, \quad \mathbf{x} = (x, y) \in \Omega = (0, 1)^2, \tag{8.2}$$

where p denotes the pressure and ρ is a piecewise-constant density coefficient, see [11, 22] and the references therein. We consider circular air bubbles in Ω filled with water, see Figure 8.2(a) for the geometry. Here, we consider the density contrast to be equal to $\epsilon = 10^3$. A standard second-order finite-difference scheme is applied to discretize (8.2), where we use a uniform Cartesian grid. Nonhomogeneous Neumann boundary conditions are imposed so that the resulting linear system (1.1) is still compatible. Moreover, we choose as preconditioner, M^{-1}, the Incomplete Cholesky decomposition without fill-in, IC(0). We refer to [19] for more details about the setting of this experiment.

Let the open domain, Ω, be divided into subdomains, $\Omega_j,\ j = 1, 2, \ldots, k+1$, such that $\overline{\Omega} = \cup_{j=1}^{k+1} \overline{\Omega}_j$ and $\Omega_i \cap \Omega_j = \emptyset$ for all $i \neq j$. The discretized domain and subdomains are denoted by Ω_h and Ω_{h_j}, respectively. Then, for each Ω_{h_j} with $j = 1, 2, \ldots, k+1$, a projection vector, z_j, is defined as follows:

$$(z_j)_i := \begin{cases} 0, & x_i \in \Omega_h \setminus \overline{\Omega}_{h_j}; \\ 1, & x_i \in \Omega_{h_j}, \end{cases} \tag{8.3}$$

where x_i is a grid point of Ω_h. The subdomains are identical square domains, which



(a) Geometry of the piecewise-constant coefficient, ρ.

(b) Geometry of the subdomains, Ω_j.

FIG. 8.2. *Settings for the bubbly flow problem.*

are chosen independently of the bubbles and the number of them can be varied, see also Figure 8.2(b). It can be shown that the projection vectors approximate slowly varying eigenvectors corresponding to small eigenvalues. Then, we take $Z := [z_1 \; z_2 \; \cdots \; z_k]$. Hence, Z consists of orthogonal, disjoint and piecewise-constant vectors. We remark that the projection vectors are not restricted to choices that are common in DDM and deflation. Typical MG projection vectors could also be taken.

8.2.1. Experiment with $\bar{M}^{-1} = M^{-1}$. The results with $\bar{M}^{-1} = M^{-1}$ are presented in Table 8.3.

From the table, it can be observed that, for all k, DEF and BNN require the same number of iterations, whereas MG is the fastest method in terms of the number of iterations, which is as expected. Recall that this does not necessarily mean that MG is the fastest method with respect to computing time, since each iteration of MG is more expensive than a iteration of DEF. Moreover, note that the difference in performance between the cases with $k = 2^2$ and $k = 4^2$ is small. In these cases, the corresponding projection vectors are bad approximations of the eigenvectors associated with the smallest eigenvalues of A.

	$k = 2^2 - 1$		$k = 4^2 - 1$		$k = 8^2 - 1$	
Method	# It.	$\frac{\|x_{\text{it}} - x\|_2}{\|x\|_2}$	# It.	$\frac{\|x_{\text{it}} - x\|_2}{\|x\|_2}$	# It.	$\frac{\|x_{\text{it}} - x\|_2}{\|x\|_2}$
DEF	149	1.5×10^{-8}	144	3.1×10^{-8}	42	1.8×10^{-8}
BNN	149	1.5×10^{-8}	144	3.1×10^{-8}	42	1.1×10^{-8}
MG	86	1.0×10^{-7}	93	6.5×10^{-8}	32	1.9×10^{-8}

TABLE 8.3

Number of required iterations for convergence and the $2-$*norm of the relative errors of 2L-PCG methods, for the bubbly flow problem with* $n = 64^2$ *and* $\bar{M}^{-1} = M^{-1}$. *PREC requires 137 iterations and leads to a relative error of* 4.6×10^{-7}.

8.2.2. Experiment with Symmetrized Smoother. We perform the same experiment as above, but taking $M^{-1} = \bar{M}^{-1} + \bar{M}^{-T} - \bar{M}^{-T} A \bar{M}^{-1}$, while $\bar{M}^{-1}$ is still the IC(0) preconditioner. In contrast to the previous experiment, the amount of work for each iteration of BNN, MG and DEF is now approximately the same and Theorem 7.3 holds. The results of this experiment are presented in Table 8.4.

	$k = 2^2 - 1$		$k = 4^2 - 1$		$k = 8^2 - 1$	
Method	# It.	$\frac{\|\|x_{\text{it}} - x\|\|_2}{\|\|x\|\|_2}$	# It.	$\frac{\|\|x_{\text{it}} - x\|\|_2}{\|\|x\|\|_2}$	# It.	$\frac{\|\|x_{\text{it}} - x\|\|_2}{\|\|x\|\|_2}$
DEF	87	7.2×10^{-8}	94	1.3×10^{-8}	34	7.6×10^{-9}
BNN	87	7.2×10^{-8}	94	1.3×10^{-8}	34	7.6×10^{-9}
MG	86	1.0×10^{-7}	93	6.5×10^{-8}	32	1.9×10^{-8}

TABLE 8.4

Number of required iterations for convergence and the $2-$*norm of the relative errors of 2L-PCG methods, for the bubbly flow problem with* $n = 64^2$ *and* $M^{-1} = \bar{M}^{-1} + \bar{M}^{-T} - \bar{M}^{-T} A \bar{M}^{-1}$. *PREC requires 137 iterations and leads to a relative error of* 4.6×10^{-7}.

As can be observed in Table 8.3, MG is now comparable with DEF and BNN, as expected from the theory of Section 7. All methods require approximately the same number of iterations and lead to the same accuracy.

9. Conclusions. We compare two-level PCG methods based on deflation (DEF), balancing Neumann-Neumann(BNN) and multigrid V(1,1)-cycle (MG) preconditioners in their abstract forms, which all consist of combinations of traditional and projection-type preconditioners. When simple choices are made for the algorithmic components, each MG iteration is more expensive than a DEF or BNN iteration, due to the more sophisticated form of the two-level preconditioner. At first glance, we would expect MG to be the most effective method; however, we have shown that there exist some traditional and projection preconditioners such that DEF is expected to converge faster than MG.

If Richardson relaxation is used with an optimal weighting as a traditional preconditioner, then it can be proven that MG always gives a more favorable condition number than DEF or BNN. For more sophisticated and effective traditional preconditioners, we still expect MG to be superior to DEF and BNN, although the work per iteration of MG remains more than for the other methods.

For special choices of traditional preconditioners, it can be shown that BNN, DEF and MG require the same amount of work per iteration and their spectra only differ in one cluster of eigenvalues around 0 or 1. Hence, these methods are expected to show comparable convergence behavior.

Numerical experiments with Laplace-like and bubbly flow problems emphasize the theoretical results.

REFERENCES

[1] J. H. Bramble, J. E. Pasciak, J. Wang, and J. Xu. Convergence estimates for multigrid algorithms without regularity assumptions. *Math. Comp.*, 57:23–45, 1991.

[2] W. L. Briggs, V. E. Henson, and S. F. McCormick. *A Multigrid Tutorial*. SIAM Books, Philadelphia, 2000. Second edition.

[3] B. Carpentieri, L. Giraud, and S. Gratton. Additive and multiplicative two-level spectral preconditioning for general linear systems. *SIAM J. Sci. Comput.*, 29(4):1593–1612, 2007.

[4] V. Faber and T. Manteuffel. Necessary and sufficient conditions for the existence of a conjugate gradient method. *SIAM Journal on Numerical Analysis*, 21:352–362, 1984.

[5] R. D. Falgout, P. S. Vassilevski, and L. T. Zikatanov. On two-grid convergence estimates. *Numerical linear algebra with applications*, 12(5–6):471–494, 2005.

[6] A. Frommer, R. Nabben, and D. B. Szyld. Convergence of stationary iterative methods for Hermitian semidefinite linear systems and applications to Schwarz methods. *SIAM J. Matrix Anal. Appl.*, 2008. Submitted.

[7] G. H. Golub and C. F. van Loan. *Matrix Computations*. Johns Hopkins Univ. Press, Baltimore, MD, 1996. Third edition.

[8] W. Hackbusch. Convergence of multi–grid iterations applied to difference equations. *Math. Comp.*, 34:425–440, 1980.
[9] W. Hackbusch. *Multigrid Methods and Applications*. Springer-Verlag, Berlin, 1985.
[10] R. Horn and C. Johnson. *Matrix Analysis*. Cambridge University Press, New York, NY, USA, 1990. USA Edition.
[11] S. P. MacLachlan, J. M. Tang, and C. Vuik. Fast and robust solvers for pressure correction in bubbly flow problems. *J. Comput. Phys. (under revision)*, 2008.
[12] J. Mandel. Balancing domain decomposition. *Comm. Numer. Meth. Engrg.*, 9:233–241, 1993.
[13] S. F. McCormick and J. W. Ruge. Convergence estimates for multigrid algorithms without regularity assumptions. *SIAM J. Numer. Anal.*, 19:924–929, 1982.
[14] R. Nabben and C. Vuik. A comparison of Deflation and Coarse Grid Correction applied to porous media flow. *SIAM J. Numer. Anal.*, 42:1631–1647, 2004.
[15] R. Nabben and C. Vuik. A comparison of deflation and the balancing preconditioner. *SIAM J. Sci. Comput.*, 27:1742–1759, 2006.
[16] R. A. Nicolaides. On some theoretical and practical aspects of multigrid methods. *Math. Comp.*, 33:933–952, 1979.
[17] R. A. Nicolaides. Deflation of conjugate gradients with applications to boundary value problems. *SIAM J. Numer. Anal.*, 24(2):355–365, 1987.
[18] G. Strang. *Introduction to Linear Algebra*. Wellesley-Cambridge Press, 1993.
[19] J. M. Tang, R. Nabben, C. Vuik, and Y. A. Erlangga. Theoretical and numerical comparison of various projection methods derived from deflation, domain decomposition and multigrid methods. DIAM Report 07-04, Delft University of Technology, Delft, 2007. Submitted to J. Sci. Comput.
[20] A. Toselli and O. Widlund. *Domain Decomposition Methods - Algorithms and Theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer, Berlin, 2004.
[21] U. Trottenberg, C. W. Oosterlee, and A. Schüller. *Multigrid*. Academic Press, London, 2000.
[22] S. van der Pijl. Computation of bubbly flows with a mass-conserving level-set method. PhD thesis, Delft University of Technology, 2005.
[23] S. P. van der Pijl, A. Segal, C. Vuik, and P. Wesseling. A mass-conserving Level-Set method for modelling of multi-phase flows. *Int. J. Numer. Methods Fluids*, 47:339–361, 2005.
[24] C. Vuik, A. Segal, and J.A. Meijerink. An efficient preconditioned CG method for the solution of a class of layered problems with extreme contrasts in the coefficients. *J. Comput. Phys.*, 152:385–403, 1999.
[25] P. Wesseling. *An Introduction to Multigrid Methods*. John Wiley & Sons, Chichester, 1992. Corrected Reprint. Philadelphia: R.T. Edwards, Inc., 2004.
[26] D. M. Young. *Iterative Solutions of Large Linear Systems*. Academic Press, New York, 1971.