Electronic Transactions on Numerical Analysis.
Volume 51, pp. 512–528, 2019.
Copyright © 2019, Kent State University.
ISSN 1068–9613.
DOI: 10.1553/etna_vol51s512

ETNA
Kent State University and
Johann Radon Institute (RICAM)

ADAPTIVE MULTILEVEL KRYLOV METHODS*

RENÉ KEHL†, REINHARD NABBEN‡, AND DANIEL B. SZYLD§

Abstract. Inexact (variable) preconditioning of Multilevel Krylov methods (MK methods) for the solution of linear systems of equations is considered. MK methods approximate the solution of the local systems on a subspace using a few, but fixed, number of iteration steps of a preconditioned flexible Krylov method. In this paper, using the philosophy of inexact Krylov subspace methods, we use a theoretically-derived criterion to choose the number of iterations needed on each level to achieve a desired tolerance. We use this criterion on one level and obtain an improved MK method. Inspired by these results, a second ad hoc method is also explored. Numerical experiments for the Poisson, Helmholtz, and the convection-diffusion equations illustrate the efficiency and robustness of this adaptive Multilevel Krylov method.

Key words. Multilevel Krylov methods, flexible GMRES, inexact Krylov subspace methods, inexact preconditioning

AMS subject classifications. 65F10, 65F50, 65N22, 65N55

1. Introduction.

We consider the iterative solution of linear systems

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

where A is a large, sparse, and possibly nonsymmetric matrix. Krylov subspace methods in combination with multilevel preconditioners are the methods of choice for the approximate solution of (1.1); see, e.g., [5, 19, 29].

Erlangga and Nabben [8, 9, 10] introduced Multilevel Krylov methods (MK methods). This type of multilevel method was inspired by the idea of shifting some eigenvalues that are close to zero farther away from zero, e.g., to have value one. This shifting is performed in the spirit of deflation techniques, i.e., one has to solve smaller systems on a subspace or coarse-grid systems. This in turn leads to the multilevel structure. MK methods, as originally presented, approximate the solution of the subspace systems by performing just a few iterations of a preconditioned flexible Krylov method such as flexible GMRES [22].

In [8], the MK methods are analyzed, and their potential is demonstrated for the 2D Poisson and the 2D convection-diffusion equations. The latter is an example with a nonsymmetric matrix of coefficients. By using a simple piecewise constant interpolation to construct the subspace or coarse-grid system, an h-independent convergence for the Poisson equation and an almost h- and Pe-independent convergence for the convection-diffusion equation are observed, where h is the mesh width and Pe is the Péclet number.

The Multilevel Krylov technique was combined with the shifted Laplacian preconditioner [11, 12, 13] to solve high wavenumber 2D Helmholtz equations [9]. There, the linear system is preconditioned by the shifted Laplacian first, and then the shifting of some eigenvalues is performed. Numerical results show that not only the convergence is almost independent of h but also only mildly dependent on the wavenumber. The theoretical results in [15] help to explain this convergence rate. In [23, 24] the shift is performed first, followed by the shifted Laplacian preconditioner. This approach can be seen as a preconditioned Multilevel Krylov

*Received June 24, 2017. Accepted October 16, 2019. Published online on December 16, 2019. Recommended by Johannes Kraus. The work of D. Szyld is supported in part by the U.S. National Science Foundation under grants DMS-1115520 and DMS-1418882 and the U.S. Department of Energy under grant DE-SC 0016578.

†Weierstraß-Institut, Mohrenstraße 39, D-10117 Berlin, Germany (rene.kehl@wias-berlin.de).

‡TU Berlin, Institut für Mathematik, MA 3-3, Straße des 17. Juni 136, D-10623 Berlin, Germany (nabben@math.tu-berlin.de).

§Department of Mathematics, Temple University (038-16), 1805 N. Broad Street, Philadelphia, Pennsylvania 19122-6094, USA (szyld@temple.edu).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

method. This combination leads to a powerful method for 3D realistic Helmholtz equations with high wavenumbers, further illustrating the potential of MK methods.

The idea of applying Krylov subspace methods on different levels or using them as preconditioners was also considered elsewhere. Kraus [16] provided for the first time a convergence analysis showing h-independent convergence rates of nonlinear algebraic multilevel iteration (nonlinear AMLI) methods; see further [17]. Notay and Vassilevski [21] followed this with a similar recursive Krylov-based multigrid cycle, which uses flexible Krylov methods to solve the coarse-grid systems in a classical (algebraic) multigrid step that includes and depends on smoothing steps. Elman, Ernst, and O'Leary [4] use Krylov subspace methods in a combined way: the method solves the coarse-grid system at the intermediate levels of multigrid, and the Krylov subspace method is also used as a replacement for the smoother. In [25], flexible Krylov subspace methods that are preconditioned by Krylov subspace methods themselves are studied but not in the context of a multilevel structure.

In the Multilevel Krylov methods described earlier, the number of iterations performed at each level is fixed *a priori*. The choice of this number results from experiments run with similar systems. In other words, one can consider these methods as Krylov subspace methods with inexact preconditioning, in which the level of inexactness is prescribed in an *ad hoc* manner. Our analysis here instead gives insight into the evolution of the residual, depending on the accuracy of the coarse-grid solutions and hence on the number of iterations on these grids. Inspired by the analysis of inexact Krylov methods in [26, 27], in this paper we consider varying the number of iterations by relaxing the tolerance of each application of the inexact preconditioner, while obtaining the same overall accuracy. We show that for the MK methods, the number of iteration steps for the subspace systems can be reduced during the fine-level iteration in an adaptive way.

The numerical results in [8, 9, 23] demonstrate that the number of iterations on the second level is important. Typically 8 iterations on the second level are used, then 2 on the next level, while only 1 iteration on all other levels. Here we produce a theoretically-derived criterion that can be used to adapt the number of iterations on each level. We thus use this criterion to reduce the number of iterations on the second level. Another ad hoc variant inspired by adaptive MK is also presented. Examples for Poisson, Helmholtz, and convection-diffusion equations illustrate the efficiency and robustness of this technique.

The paper is organized as follows. In Section 2, we briefly describe the Multilevel Krylov methods. In Section 3, we begin with a few comments on general Krylov methods, then we describe inexact Krylov subspace methods and the new results for adaptive Multilevel Krylov methods. Numerical results in Section 4 illustrate the performance of both the proposed adaptive method and the ad hoc variant. Experiments related to the parameters used are shown in Appendix A.

2. Multilevel Krylov methods. We begin by briefly describing the MK methods; for more details see [7, 10]. For the sake of simplicity, in what follows we mainly discuss the two-level method; the multilevel method is obtained by a recursive application of the two-level method on the corresponding coarse-grid solve. We start with a rectangular matrix $Z \in \mathbb{R}^{n\times r}$, with $r \leq n$, i.e., describing an operator

$$Z : \mathbb{R}^r \mapsto \mathbb{R}^n.$$

In the multilevel language the operator Z is the *prolongation* or *interpolation* operator, which transfers variables from a coarse level to a fine one. The *restriction* operator

$$Z^T : \mathbb{R}^n \mapsto \mathbb{R}^r$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

does the opposite; it transfers some variables from a fine level to a coarse one. The so-called coarse-level (grid) matrix or subspace matrix is then defined by the Galerkin operator

$$E = Z^T AZ \in \mathbb{R}^{r\times r}.$$

Here we assume that E is nonsingular. Note that instead of Z^T any other $r \times n$ matrix Y^T can be used (see [7, 18] for details). We define the operator

$$P_D := I - AZE^{-1}Z^T.$$

Note that $P_D^2 = P_D$, thus P_D is the (oblique) projection operator onto the nullspace of Z^T along the range of AZ. We then consider the (projected or) preconditioned system

$$P_D Ax = P_D b. \tag{2.1}$$

A two-level Krylov method is obtained if a Krylov subspace method such as GMRES is applied to (2.1), this being the first level, while the solution of the linear system with the coefficient matrix E is the second level.

The convergence of Krylov subspace methods is driven by the eigenvalues of the coefficient matrix—for (2.1), $\sigma(P_D A)$—as well as the corresponding eigenvectors in the nonsymmetric case; see, e.g., [19, 27]. Related to this feature, the following observation is useful to gain some insight into the convergence of the Multilevel Krylov method. As shown in [20], for any full rank Z, the spectrum of $P_D A$ is

$$\sigma(P_D A) = \{0, \ldots, 0, \mu_{r+1}, \ldots .\mu_n\}. \tag{2.2}$$

We first note that the values of μ_i, $i = r+1, \ldots, n$, depend on the specific choice of Z. The operator P_D is known as the deflation operator; see, e.g., [8, 14]. If A is symmetric positive definite and the columns of Z consist of orthogonal eigenvectors of A, then the eigenvalues μ_j of the preconditioned operator in (2.2) are eigenvalues of A. Hence, some of the eigenvalues of A are deflated out of the spectrum of A, or, in other words, they are shifted to zero. A Krylov subspace method that works for singular matrices can be used to solve (2.1). Typically, only the eigenvalues $\mu_{r+1}, \ldots, \mu_n$ are those which enter into the spectral error bounds for such a Krylov subspace method.

In order to obtain the zeros in (2.2), the linear systems corresponding to the product with E^{-1}, i.e., the subspace systems, have to be computed exactly. If they are approximated or solved inexactly, i.e., if the product with E^{-1} is replaced by one with $E^{-1} + P$, where P is a small perturbation, then the zeros in (2.2) are expected to be small numbers. This would slow down the convergence of the Krylov method.

To allow inexact subspace solves, one can consider the operator

$$P_N := I - AZE^{-1}Z^T + \lambda ZE^{-1}Z^T, \tag{2.3}$$

where λ can be any real number, e.g., the number one. This operator was first introduced in [7] and for $\lambda = 1$ is equivalent to the ADEF operator introduced in [28]. Note that P_N is no longer a projection. We now have that (in exact arithmetic)

$$\sigma(P_N A) = \{\lambda, \ldots, \lambda, \mu_{r+1}, \ldots .\mu_n\}.$$

Thus, some of the eigenvalues are shifted to λ. If the subspace method is solved inexactly, then only a small perturbation might be added to the eigenvalues. Of course, P_N can be written as

$$P_N := I - (A - \lambda I)ZE^{-1}Z^T, \tag{2.4}$$

so just one subspace solve is needed and not two, as it might appear from (2.3).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

The concept of the Multilevel Krylov method is as follows: We use P_N as preconditioner for a flexible Krylov subspace method, and to solve the subspace system with the matrix E, we use a few steps of the same flexible Krylov subspace preconditioned by a preconditioner of the same form as P_N in (2.4) (just of smaller dimension). To apply this second-level preconditioner, a subspace system has to be solved again, which is done in the same manner. Thus, we obtain recursively the multilevel structure. In [8, 9, 23] typically 8 or 4 iterations are used on the second level, 2 iterations on the third level and 1 iteration on all other levels. This is denoted by MK(8,2,1) or MK(4,2,1), respectively, or more generally MK(p,2,1).

The main purpose of this paper is to adaptively choose p at each (outer) iteration of the Krylov subspace method. This choice is based on a solid theoretical foundation, and it is intended to find an approximation to the solution of the problem with a residual of the same order of magnitude but with less computational effort.

Flexible Krylov methods need right preconditioning. Our purpose is to establish a bound for the residual of the GMRES method. As left preconditioning would give results only for the preconditioned residual, we define the operator

$$Q_N := I - ZE^{-1}Z^TA + \lambda ZE^{-1}Z^T = I - ZE^{-1}Z^T(A - \lambda I) \tag{2.5}$$

and solve the system

$$AQ_N\tilde{x} = b, \qquad \text{with} \quad x = Q_N\tilde{x}. \tag{2.6}$$

We note that since

$$P_NA = AQ_N,$$

the spectral properties of P_NA and AQ_N are the same, and the discussion on shifting 0 to $\lambda \in \sigma(AQ_N)$ is thus valid in this case as well.

Finally, we note that the techniques described above can be combined with another (say ordinary) preconditioner in two ways. The first alternative as used in [9, 10] is the following: The matrix A can be preconditioned so that A is replaced by AM or MA, where M is the preconditioner. Secondly, the shifted systems P_NA or AQ_N can be combined with a preconditioner so that we obtain $(MP_D + \lambda ZE^{-1}Z^T)A$ [7] and $A(P_D^TM + \lambda ZE^{-1}Z^T)$, respectively. More details about the implementation of MK methods can be found in [10].

3. Inexact Krylov methods. Our goal is to solve the system (1.1) more efficiently by (right) preconditioning with a given nonsingular matrix Q_N as in (2.5) that might not be given explicitly, i.e., we solve (2.6) approximately. In the iteration process, when applying Q_N, linear systems with E have to be approximately solved. This is done by a few iteration steps of a preconditioned flexible Krylov subspace method. Thus the preconditioner Q_N varies from step to step, and Q_N cannot be applied exactly (i.e., it is not given explicitly, and in fact it is no longer linear). In this paper, we choose flexible GMRES (FGMRES) [22] as the inner and outer solver of the linear systems in (2.6).

In the rest of this section, we begin with a very brief review on general Krylov subspace methods followed by a discussion on inexact methods and their application to the Multilevel Krylov method.

3.1. General Krylov subspace methods. Let A be nonsingular and x_0 an initial vector for the solution of (1.1). Krylov subspace methods find at the m-th iteration an approximation x_m in the space $x_0 + K_m(A, v_1)$, where $K_m(A, v_1)$ is the m-th Krylov subspace, i.e., the space spanned by the vectors $v_1, Av_1, \ldots, A^{m-1}v_1$ satisfying some optimality condition, where v_1 is usually chosen as the normalized initial residual $r_0 = b - Ax_0$, i.e., $r_0 = \beta v_1$,

ETNA
Kent State University and
Johann Radon Institute (RICAM)

$\beta = \|r_0\|_2$. An orthonormal basis $\{v_1, \ldots, v_m\}$ of this Krylov subspace is usually obtained by the Arnoldi procedure, in which the main operation is the matrix-vector product Av_i. These basis vectors are collected as columns of the matrix V_m. This gives rise to the Arnoldi relation

$$AV_m = V_{m+1}H_m, \tag{3.1}$$

where $H_m \in \mathbb{R}^{(m+1)\times m}$ is an upper Hessenberg matrix, and we have assumed that $\|v_k\|_2 = 1$, for all $k \leq m+1$. For GMRES and its variants, the optimality condition corresponds to minimizing the norm of the residual, i.e.,

$$\begin{aligned} &\min_{x \in x_0 + K_m(A, v_1)} \|r_0 - Ax\|_2 \\ &\quad = \min_{y \in \mathbb{R}^m} \|\beta e_1 - H_m y\|_2, \end{aligned} \tag{3.2}$$

where e_1 is the first Euclidean vector. In the remainder of this article $\|\cdot\|$ will describe the Euclidean norm, which is the one to be optimized.

3.2. Inexact Krylov subspace methods. We review here results from [25, 26, 27] on inexact Krylov subspace methods, which we use in our application to Multilevel Krylov methods.

When the matrix-vector multiplication Av_i in the Arnoldi procedure is not performed exactly, it can be interpreted as being of the form $(A + G_i)v_i$, where $G_i \in \mathbb{R}^{n\times n}$, $i = 1, \ldots, m$, can be thought of as perturbations of A. The analogue to (3.1) is an *inexact Arnoldi relation*

$$\begin{aligned} AV_m + [G_1 v_1, \ldots, G_m v_m] &= V_{m+1}H_m, \qquad \text{i.e.,} \\ [(A+G_1)v_1, \ldots, (A+G_m)v_m] &= V_{m+1}H_m. \end{aligned} \tag{3.3}$$

With $\mathcal{E}_m = \sum_{k=1}^m G_k v_k v_k^T$, equation (3.3) can be written in compact form as

$$(A + \mathcal{E}_m)V_m = V_{m+1}H_m.$$

In particular this means that at each iteration, the solution given by the Krylov subspace method with inexact matrix-vector products is an element of a perturbed Krylov subspace $K_m(A + \mathcal{E}_m, v_1)$.

REMARK 3.1. The expression given in (3.2) no longer minimizes the residual norm using the matrix A, as in (3.2), but instead with the perturbed matrix given in (3.3). We refer to the solution of (3.2) as the GMRES (or the FGMRES) minimizer.

One problem that arises during the iteration process is that the residual r_m is no longer an available quantity since the "true residual" should be calculated with help of the unperturbed matrix A, i.e.,

$$r_m = r_0 - AV_m y_m = r_0 - V_{m+1}H_m y_m + \mathcal{E}_m V_m y_m, \tag{3.4}$$

where y_m is the solution of the minimization problem (3.2). Note that neither $\mathcal{E}_m$ nor AV_m are available. The "computed residual" that is carried out by the inexact method is thus

$$\tilde{r}_m = r_0 - (A + \mathcal{E}_m)V_m y_m = r_0 - V_{m+1}H_m y_m = r_m - \mathcal{E}_m V_m y_m. \tag{3.5}$$

Therefore the difference $\|r_m - \tilde{r}_m\|$ can be bounded as follows

$$\|r_m - \tilde{r}_m\| = \|\mathcal{E}_m V_m y_m\| \leq \sum_{k=1}^m \|G_k\| |y_m^{(k)}|,$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

where $y_m^{(k)}$ is the k-th component of the iterate y_m.

The following results come from [26].

LEMMA 3.2. *Assume that m iterations of the inexact Arnoldi method have been carried out, and let y_m be the solution of the GMRES minimization problem (3.2) (see Remark 3.1). Then, for any $k = 1, \ldots, m$,*

$$|y_m^{(k)}| \leq \frac{1}{\sigma_m(H_m)} \|\tilde{r}_{k-1}\|,$$

where $\sigma_m(H_m)$ denotes the smallest singular value of H_m and $\tilde{r}_{k-1}$ is as in (3.5).

THEOREM 3.3. *Let $\varepsilon > 0$. Let r_m be the GMRES residual after m iterations of the inexact Arnoldi method given by (3.4). Under the same notation and hypothesis of Lemma 3.2, if for every $k \leq m$,*

$$\|G_k\| \leq \frac{\sigma_m(H_m)}{m} \frac{1}{\|\tilde{r}_{k-1}\|} \varepsilon, \tag{3.6}$$

then $\|r_m - \tilde{r}_m\| \leq \varepsilon$. Moreover, if

$$\|G_k\| \leq \frac{1}{m\kappa(H_m)} \frac{1}{\|\tilde{r}_{k-1}\|} \varepsilon, \tag{3.7}$$

where $\kappa(H_m)$ denote the condition number of H_m, then $\|(V_{m+1}H_m)^T r_m\| \leq \varepsilon$.

Conditions (3.6) and (3.7) are of the form

$$\|G_k\| \leq c_m \cdot \frac{\varepsilon}{\|\tilde{r}_{k-1}\|} \cdot$$

While in some cases, $c_m \approx 1$ is a reasonable choice for various problems [1, 2], there are cases, where c_m needs to contain further information on A or H_m. In most applications, it is not possible to calculate the exact value of c_m *a priori* since the singular values of H_m at the m-th step are not known beforehand. Nevertheless, there are some estimation strategies one could use; cf. [26, §5, §9]. We discuss this issue further for our application to Multilevel Krylov methods in our numerical experiments in Section 4.

3.3. Application to Multilevel Krylov methods. We begin by considering the application of a right multilevel preconditioner of the form (2.5), with $\lambda > 0$, $Z \in \mathbb{R}^{n \times r}$, $r \ll n$, and $E = Z^T A Z$, the coarse-grid Galerkin matrix. For a given $v_k \in \mathbb{R}^n$, the (approximate) product $Q_N v_k$ proceeds in three steps as follows:

1. Calculate $w_k = Z^T(\lambda I - A)v_k$.
2. Solve $Et = w_k$ inexactly to a given tolerance τ, producing the approximation $\tilde{t}_k$, which leads to the second-level residual $q_k = w_k - E\tilde{t}_k$ so that $\|q_k\|_2/\|w_k\|_2 < \tau$.
3. Calculate $\tilde{z}_k = v_k + Z\tilde{t}_k \approx Q_N v_k$.

For a two-level method, the system in step 2. is solved with a direct solver, i.e., with $\tau = 0$. In the multilevel case, we set $\tilde{A} = E$, $\tilde{Z} \in \mathbb{R}^{r \times s}$, $s \ll r$ and apply a preconditioner of the form of Q_N and perform some iterations with flexible GMRES (FGMRES) starting with the initial vector $t_0 = 0$ on every grid. One either chooses *a priori* the number of FGMRES iterations, as in [8, 9, 10], or by the appropriate choice of the tolerance τ as we propose in this paper. In principle one can continue until the coarsest-grid system becomes sufficiently small. However, in our numerical experiments described in the next section, we use a 5-level method and solve the coarsest system with a direct solver.

In this section, we develop the criteria on how to choose the tolerance τ for each of the linear systems on the second level. This choice will determine the number of FGMRES

ETNA
Kent State University and
Johann Radon Institute (RICAM)

iterations at the second level. As discussed above, the criteria can be applied on further levels as well, but since the MK algorithm uses only one or two steps at other levels, the criterion is not used.

To that end, let $z_k = Q_N v_k$ (the "exact" matrix-vector product), and let $\tilde{z}_k$ be the corresponding approximate matrix-vector product as per the three steps described above. Let us collect these vectors, for $k = 1, \ldots, m$, into a matrix $\tilde{Z}_m := [\tilde{z}_1, \ldots, \tilde{z}_m]$.

As in the previous section, we have the "true" residual

$$r_m = r_0 - AQ_N V_m y_m, \tag{3.8}$$

where y_m comes from the solution of the least-squares problem (3.2) in FGMRES. The computed residual is

$$\tilde{r}_m = r_0 - A\tilde{Z}_m y_m. \tag{3.9}$$

Then,

$$r_m - \tilde{r}_m = -A(Q_N V_m - \tilde{Z}_m)y_m = -A\sum_{k=1}^{m} (Q_N v_k - \tilde{z}_k)\, y_m^{(k)}.$$

We now look in detail at the difference in parenthesis. We have

$$\begin{aligned} Q_N v_k - \tilde{z}_k &= v_k + ZE^{-1}w_k - \tilde{z}_k = v_k + ZE^{-1}w_k - (v_k + Z\tilde{t}_k) = Z\left(E^{-1}w_k - \tilde{t}_k\right) \\ &= ZE^{-1}\left(w_k - E\tilde{t}_k\right) = ZE^{-1}q_k, \end{aligned}$$

and hence

$$\|r_m - \tilde{r}_m\| \le \|AZE^{-1}\| \sum_{k=1}^{m} \|q_k\| \cdot |y_m^{(k)}|. \tag{3.10}$$

The bound (3.10) gives the desired connection between the allowable tolerance τ for the residual q_k at the second level so that the residual at the fine level can achieve the desired accuracy. Using this bound and the analysis of the form of the components $y_m^{(k)}$ from [26] given in Lemma 3.2, we obtain our main result.

THEOREM 3.4. *Let $\varepsilon > 0$, and assume that m iterations of the inexact Arnoldi method have been carried out for the preconditioned system $AQ_N\bar{x} = b$, $x = Q_N\bar{x}$. Let y_m be the FGMRES minimizer and $\bar{x}_m = x_0 + V_m y_m$, and let $r_m, \tilde{r}_m$ be given as in (3.8), (3.9), respectively. Let $w_k = Z^T(\lambda I - A)v_k$ and $\tilde{t}_k$ be an approximate solution of $Et = w_k$. If for each iteration $k \le m$, the inner residual $q_k = w_k - E\tilde{t}_k$ satisfies*

$$\|q_k\| \le \frac{\sigma_m(H_m)}{\|AZE^{-1}\|m} \cdot \frac{1}{\|\tilde{r}_{k-1}\|}\varepsilon, \tag{3.11}$$

it follows that $\|r_m - \tilde{r}_m\| \le \varepsilon$. Moreover, if

$$\|q_k\| \le \frac{1}{\|AZE^{-1}\|m\kappa(H_m)} \frac{1}{\|\tilde{r}_{k-1}\|}\varepsilon,$$

then $\|(V_{m+1}H_m)^T r_m\| \le \varepsilon$.

REMARK 3.5. Theorem 3.4 refers to an algorithm where the operator Q_N is applied twice. First it is used as an inexact preconditioner to obtain the solution $\bar{x}$. It is reapplied to

ETNA
Kent State University and
Johann Radon Institute (RICAM)

$\bar{x} = x_0 + V_m y_m$ to calculate the solution x. In practice, we do not keep V_m for the calculation of x, but instead we keep $Z_m = Q_N V_m$ so that we get $x = x_0 + Z_m y_m$.

Theorem 3.4 deals with the residuals of two consecutive levels such that it can be applied successively in the multigrid context. It can be seen as a global strategy for each (inner or outer) iteration process.

As discussed in the previous section, these inner bounds may be difficult to calculate exactly. Here the norm of AZE^{-1} is an additional quantity that has to be estimated. The bound on the inner residual is again of the form

$$\|q_k\| \leq c_m \cdot \frac{\varepsilon}{\|\tilde{r}_{k-1}\|} \cdot \tag{3.12}$$

Due to the proven existence of such a bound, one can use an initial value for c_m and refine that choice if there is no convergence or if the convergence is too slow, until an appropriate tolerance is achieved; see also the comments in [26].

We note that the bound also holds for a residual difference relative to the right-hand side, i.e., use $\tilde{\varepsilon} = \|b\| \cdot \varepsilon$ in Theorem 3.4, and obtain the same result. We also note that if (3.11) holds and the computed residual is below the tolerance, i.e., $\|\tilde{r}_m\| \leq \varepsilon$, then this means that for the unknown true residual the bound

$$\|r_m\| \leq \|\tilde{r}_m\| + \|r_m - \tilde{r}_m\| \leq 2\varepsilon$$

holds at the end of the iteration process, i.e., when the calculated residual $\|\tilde{r}_m\|$ reached the given tolerance.

4. Adaptive MK methods and numerical examples. As mentioned above we use Q_N in (2.5) as preconditioner for a flexible Krylov subspace method (the outer iterations), and to solve the subspace system with the matrix E, we use a few steps of the same flexible Krylov subspace method preconditioned by a preconditioner of the same form as Q_N (just of smaller dimension). To apply this second-level preconditioner, a subspace system has to be solved again, which is done in the same manner. Thus, we obtain recursively the multilevel structure.

The theoretical results of the previous section can be used to devise a strategy to reduce the number of (inner) iterations to solve the systems on each level in the multilevel method during the outer iteration. However, it was observed in [9] that the number of second-level iterations is the most important. The number of iterations on all the other levels can be chosen as 2 or even as 1. So, on these levels there is no need to apply an adaptive technique.

While we can calculate or approximate satisfactorily the value of $\|AZE^{-1}\|$ for some examples, the exact value of $\sigma_m(H_m)$ is not available in advance. The latter quantity comes from the Hessenberg matrix of the inexact Arnoldi equation at the m-th iteration, which is also unavailable. In can be easily seen that the smallest value of c_m for all steps m is attained at the end of the iteration, and we show this for specific examples in Appendix A. We will use this fact and choose a constant value for c_m in our examples.

In the following we will investigate the use of the proposed adaptive MK method on three problems, the 2D-Poisson equation, the convection-diffusion equation, and the Helmholtz equation. Our experiments showed that choosing $c_m = 10$ works out for all of these three problems and the multilevel methods described below.

Thus, here we investigate first an adaptive variant of the MK-method that uses the results of Theorem 3.4. We also investigate an ad hoc static variant, which is inspired by our theoretical results in the sense that we reduce the number of inner iterations as the method progresses but we maintain those numbers fixed. In other words, we are trying to mimic the adaptive paradigm. This strategy may perform well in practice, but it is not guaranteed to converge. It may be useful though for some problems.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

We compare these two relaxed MK methods with two basic MK methods. In [9] the (8,2,1)-method was investigated. We have found that using two steps on all levels except the second instead of one improved this method. Thus, as already mentioned, we use in our comparisons a (8,2,2)-method, i.e., we investigate a five-level method[1] that performs 8 steps on the second level and only 2 iterations on all the other levels, while on the last level the problem is solved directly. The other basic method performs the minimal number of iterations, i.e., two iterations on each level.

More specifically, we consider four classes of experiments:

- **MKadap**: Five-level MK method using 8 inner iterations on the second level in the first outer step of the Krylov method and reducing this number of inner iterations on the subsequent Krylov steps following the strategy given in Theorem 3.4, where the criterion for the inner residual is $\|q_k\| \leq c_m \cdot 10^{-10}/\|\tilde{r}_{k-1}\|$ for an estimated value of c_m. The number of iterations on all other levels is 2.
- **MKstat**: Five-level MK method using 8 inner iterations in the first outer step of the Krylov method, and after a fixed number of outer iterations (say 10), the method changes to 2 inner iterations on the second level. The number of iterations on all other levels is 2.
- **MK(8,2,2)**: Five-level MK method. On the second level we perform 8 iterations and two iterations on every other level.
- **MK(2,2,2)**: Same as the previous case with only 2 iterations per level.

We present various numerical experiments with the two-dimensional Poisson equation, with the two-dimensional Helmholtz equation, and with the convection-diffusion equation. As mentioned before, we use $c_m = 10$ for all experiments and a global tolerance of $\varepsilon = 10^{-10}$. For the MKstat version, we choose to change the number of second-level iterations from 8 to 2 after 10 (outer) iterations.

4.1. 2D Poisson equation. In this section, we present numerical experiments with the 2D Poisson equation

$$-\Delta u = f \qquad \text{in } \Omega = (0,1)^2$$

with inhomogeneous Dirichlet conditions on the boundary $\partial\Omega$. The discretization with a central difference scheme leads to a symmetric positive definite matrix A. For the multilevel method we choose four coarse levels such that each solution vector on a coarse grid has half the size (rounded down) of the corresponding fine-grid solution. The solution on the coarsest grid is carried out with a direct method and the deflation matrix Z resulting from assembling two neighbor points in a lexicographic way, i.e., Z has the following form:

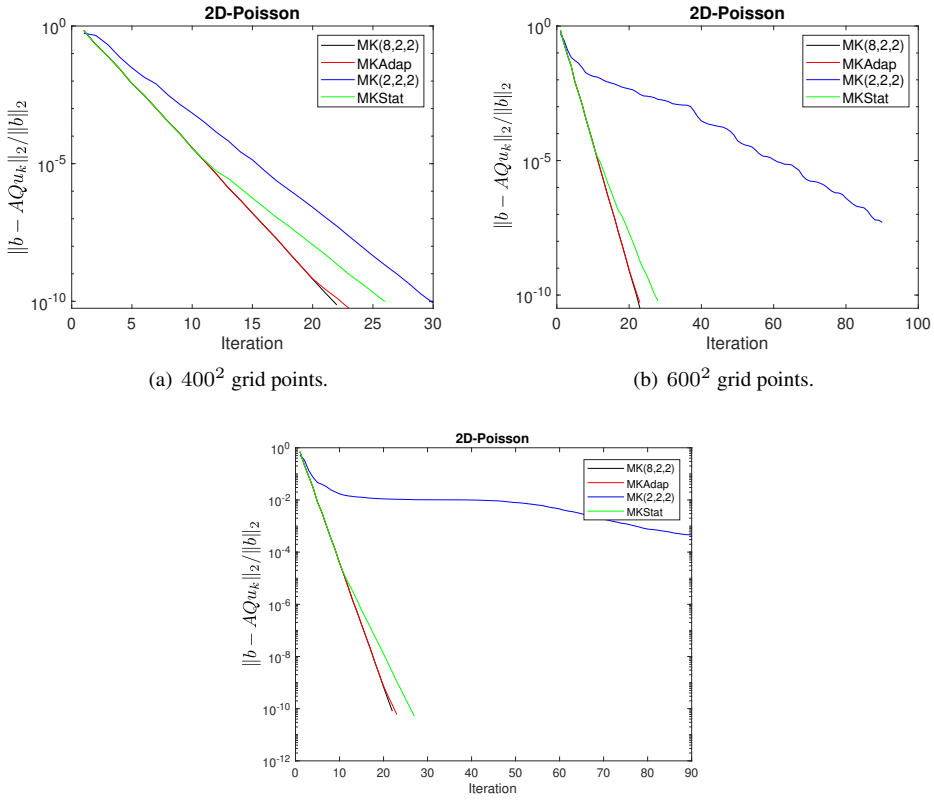
$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

The results of our numerical experiments for this problem for different grid sizes are displayed in Figure 4.1. The number of direct solves on the coarsest grid and the number of

[1] In all cases reported here we use the MK method with five levels. We have explored going deeper with more levels, but this does not improve the performance, as shown in Appendix A.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

iterations to converge to a relative residual below $\varepsilon = 10^{-10}$ are given in Table 4.1, where x is the Matlab backslash solution. In the table, we report the number of coarsest-grid direct solves that have to be performed during the process. This is a useful measure in order to compare these methods. Since the adaptive MK methods can be seen as inexact two-level methods, for completeness, we also perform numerical tests with an exact two-level method.



(a) 400^2 grid points.

(b) 600^2 grid points.

(c) 800^2 grid points.

FIG. 4.1. *Convergence behavior for the Poisson equation using* 400^2, 600^2, *and* 800^2 *grid points. Relative residual norm vs. number of Krylov subspace iterations.*

As it can be observed, for the smallest grid considered, all four versions converge in between 22 and 30 iterations, but for larger grid sizes, the method MK(2,2,2) fails to converge in 90 iterations (set as the maximum allowed)[2]. The original version MK(8,2,2) and the adaptive strategy converge essentially in the same number of iterations, while the ad hoc static version exhibits a certain delay. The computational effort though is varied. The adaptive strategy is about 10% faster than the original MK(8,2,2) method, while the static version is about 50% faster. This difference in performance is easily explained. While each iteration of the original MK(8,2,2) method uses 8 second-level iterations, the static version reduces this to 2 second-level iterations after 10 steps. The adaptive strategy starts reducing the number

[2]We note, however, that we could have equally set this maximum to 100 as we do for the problem in the next section, and the results would not change.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 4.1
Number of calculated coarsest-grid solutions (number of iterations to converge) [iteration number where switching criterion is satisfied] and the error $\|x - x_k\|$ of the different MK methods for the 2D Poisson equation.

Grid	MK(8,2,2)	MKadap	MKstat	MK(2,2,2)	two-level
400^2	704(22) $2.9 \cdot 10^{-10}$	592(23)[17] $5.7 \cdot 10^{-10}$	424(26) $1.3 \cdot 10^{-9}$	240(30) $1.4 \cdot 10^{-9}$	22(22) $5.4 \cdot 10^{-10}$
600^2	736(23) $6.1 \cdot 10^{-10}$	596(23)[17] $7.2 \cdot 10^{-10}$	440(28) $1.5 \cdot 10^{-9}$	– (90) $1 \cdot 10^{-4}$	22(22) $8.0 \cdot 10^{-10}$
800^2	704(22) $4.6 \cdot 10^{-10}$	596(23)[17] $1.0 \cdot 10^{-9}$	432(27) $1.2 \cdot 10^{-9}$	– (90) $2.8 \cdot 10^{-1}$	22(22) $1.1 \cdot 10^{-9}$

of second-level iterations only when the criterion (3.12) is satisfied, which occurs at the 18th (outer) iteration. The number of necessary second-level iterations then decreases during the outer iteration process until it reaches the minimal allowed value of 2; for the Poisson problem on a 400^2-point grid, the method performs 6, 4, and then 2 steps on the second level[3].

It follows that the two new strategies, the adaptive and static versions, perform better than the fixed strategies in this case. We note also that while the static strategy works in this case, this cannot always be guaranteed (as shown in the example in the next section). Compared to the two-level method we have the same number of iterations, while the true (relative) residual norm is below the tolerance of $2 \cdot 10^{-10}$.

We conclude the section by mentioning that we experimented with larger values of c_m[4], and we were able to observe a further reduction in the computational time to about 20% of the original method.

4.2. The convection-diffusion equation. In the following we investigate the behavior of the MK methods for the numerical solution of the convection-diffusion equation

$$\frac{\partial u}{\partial y} - \frac{1}{Pe}\Delta u = 0, \qquad \text{in } \Omega = (0,1)^2,$$

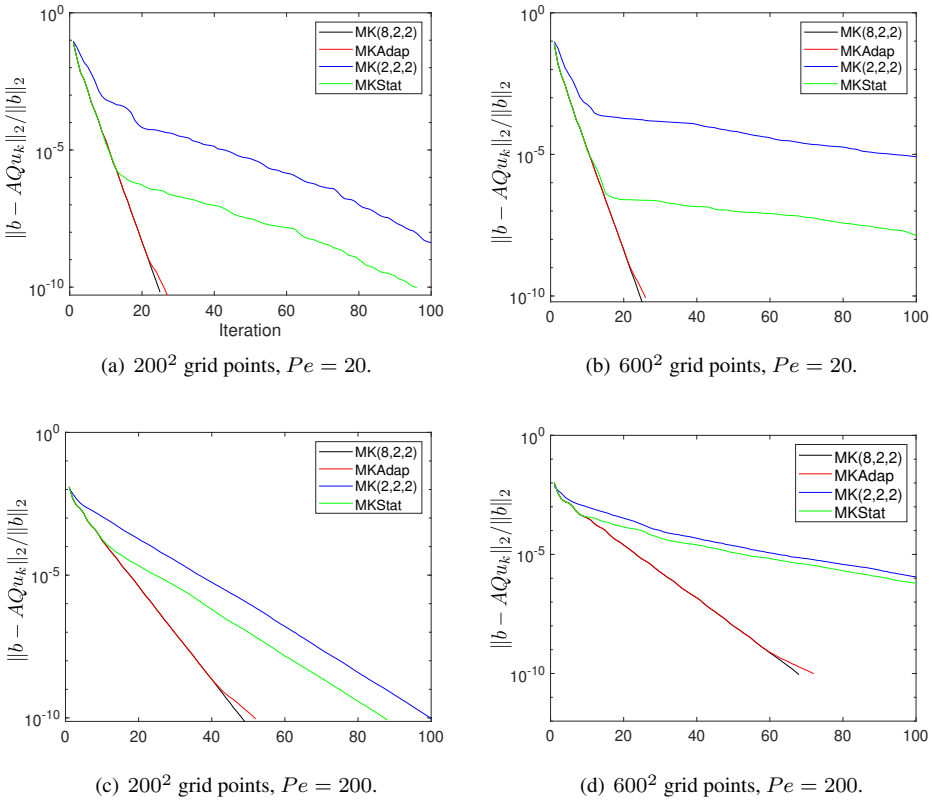
where Pe is the Péclet number. The boundary conditions for the problem are given as follows: $u(x,0) = u(0,y) = 0, u(x,1) = u(1,y) = 1$. The problem is solved in a finite difference context with an upwind scheme as given in [3].

We consider the same four versions of the MK method as described earlier with the same tolerance $\varepsilon = 10^{-10}$. Here we study the convergence behaviour on two square grids, namely 200×200 and 600×600 mesh points, with the unknowns ordered using a downwind scheme for two values of the Péclet number, $Pe = 20$ and $Pe = 200$. Convergence curves for the four methods are reported in Figure 4.2. As before, for completeness, we also perform numerical tests with an exact two-level method.

Similarly to the previous examples, one can observe that the adaptive method converges essentially in the same number of iterations as the original MK(8,2,2) version, while there is a significant reduction of computational cost; this reduction is presented in Table 4.2. In fact, the adaptive method is faster than the regular MK(8,2,2) version by about 15% with the value of $c_m = 10$. On the other hand, the convergence delay of the static ad hoc version is much more marked, leading to non-convergence in some cases. The MK(2,2,2) version fails to converge in 100 iterations[5].

[3] Experiments with a 200^2-point grid are not reported since they are qualitatively similar to those with a 400^2-point grid.
[4] Further comments on the value of c_m can be found in Appendix A.
[5] Experiments with 400^2-grids are qualitatively similar to those reported.

ETNA
Kent State University and
Johann Radon Institute (RICAM)



(a) 200^2 grid points, $Pe = 20$.

(b) 600^2 grid points, $Pe = 20$.

(c) 200^2 grid points, $Pe = 200$.

(d) 600^2 grid points, $Pe = 200$.

FIG. 4.2. *Convergence behavior for the convection-diffusion problem. Relative residual norm vs. number of Krylov subspace iterations.*

As was the case for the Poisson equation, the MK methods (except the MK(2,2,2) method) lead to the same number of iterations as the two-level method.

4.3. 2D-Helmholtz equation. In this section, numerical experiments for a 2D Helmholtz problem in a square domain with constant wavenumber are presented, i.e., we solve the problem

$$\mathcal{A}u := -\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \omega^2(x,y)\right) u(x,y) = g(x,y), \qquad \text{in } \Omega = (0,1)^2.$$

At the boundaries, the first-order approximation to the Sommerfeld (non-reflecting) condition due to Engquist and Majda [6] is imposed. We consider a problem where a source is generated in the middle of the domain.

Before solving the system (2.6) with the flexible GMRES method [22], we apply the (complex) shifted Laplacian preconditioner M to the matrix A, i.e., the discretization of

$$\mathcal{M} := -\frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - (1 - 0.5 \cdot i)\,\omega^2(x,y),$$

such that we solve $\hat{A}Q\bar{x} = b$, $Q\bar{x} = x$, $\hat{A} = AM^{-1}$.

This is the setup used in [9]. Here, we consider numerical experiments with the same environment, i.e., the same construction of the deflation matrices Z, Galerkin matrices E, etc.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 4.2
Number of calculated coarsest-grid solutions (number of iterations to converge) [iteration number where switching criterion is satisfied] and the error $\|x - x_k\|$ of the different MK methods for the convection-diffusion problem, where x is the Matlab backslash solution.

Pe = 20	MK(8,2,2)	MKadap	MKstat	MK(2,2,2)	two-level
200^2	800(25) $6.8 \cdot 10^{-9}$	672(27)[18] $1.1 \cdot 10^{-8}$	1008(96) $1.5 \cdot 10^{-7}$	– (100) $7 \cdot 10^{-6}$	25(25) $6.1 \cdot 10^{-9}$
600^2	800(25) $3.4 \cdot 10^{-8}$	664(26)[19] $8.1 \cdot 10^{-8}$	−(100) $7.2 \cdot 10^{-4}$	– (100) 0.4	25(25) $2.3 \cdot 10^{-8}$
Pe = 200					
200^2	1568(49) $8.0 \cdot 10^{-9}$	1348(51)[38] $1.9 \cdot 10^{-8}$	920(88) $2.2 \cdot 10^{-8}$	800(100) $2.4 \cdot 10^{-8}$	47(47) $6.1 \cdot 10^{-9}$
600^2	2176(68) $3.4 \cdot 10^{-7}$	1992(72)[59] $1.7 \cdot 10^{-7}$	−(100) $0.9 \cdot 10^{-2}$	– (100) $1.1 \cdot 10^{-2}$	56(56) $2.8 \cdot 10^{-8}$

TABLE 4.3
Number of calculated coarsest-grid solutions (number of iterations to converge) of the different MK methods for the 2D-Helmholtz equation.

ω	n	MK(8,2,2)	MKadap	MKstat	MK(2,2,2)
40	250^2	576(18)	484(19)	392(22)	232(29)
70	250^2	672(21)	592(25)	536(40)	464(58)
100	250^2	928(29)	836(29)	920(88)	968(121)
120	300^2	1056(33)	964(33)	1264(131)	1288(161)

The matrix M is the discretization of $\mathcal{M}$. Furthermore, the value for c_m is chosen to be 10 as it was in the prior experiments. We briefly describe some of the elements of this setup but refer the reader for more details to [9].

During the whole iteration process, the matrix M^{-1} is never calculated explicitly, but the approximation to the solution of $M^{-1}v$ for a given vector v is carried out by one multigrid step. This also means that we never calculate the Galerkin matrix

$$E = Z^T \hat{A} Z = Z^T A M^{-1} Z$$

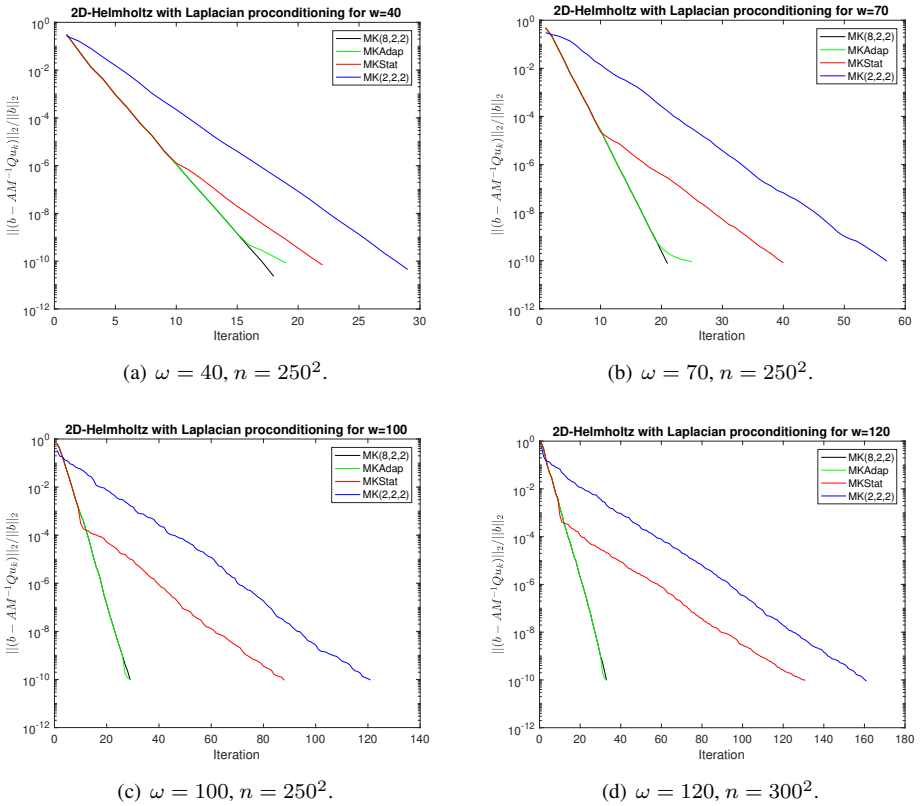
explicitly, which might no longer be a sparse matrix. Instead we approximate E by the product of the coarse-grid versions of A, M^{-1}, and the identity, i.e.,

$$E \approx Z^T A Z (Z^T M Z)^{-1} Z^T Z.$$

In our experiments we use the Jacobi smoother for 2 presmoothing steps of the Laplacian multigrid preconditioning procedure. In Figure 4.3, we show the results for a discretization on a 250^2-point grid for different wavenumbers from 40 to 100 and on a larger 300^2-point grid for $\omega = 120$. In Table 4.3 we give an overview on the number of coarsest-grid direct solves that have to be performed during the process. This is a useful measure to compare these methods concerning their efficiency. It can be observed that the adaptive method is more efficient than the others.

5. Conclusions. We developed a theory of adaptive Multilevel Krylov methods. At each level, the number of iterations of the second level is determined by the residual at the previous step using the criterion (3.12). With this strategy, the theory guarantees convergence of the adaptive MK method to the tolerance. An ad hoc static method which mimics this strategy is

ETNA
Kent State University and
Johann Radon Institute (RICAM)



(a) $\omega = 40$, $n = 250^2$. (b) $\omega = 70$, $n = 250^2$.

(c) $\omega = 100$, $n = 250^2$. (d) $\omega = 120$, $n = 300^2$.

FIG. 4.3. *Convergence behavior for the Helmholtz problem. Relative residual norm vs. number of Krylov subspace iterations.*

also presented which in some cases is also effective. The constant c_m in (3.12) depends on quantities which are not always available. If one finds that given a value of c_m, the method fails to converge, then the value of c_m should be reduced.

Acknowledgement. We thank the two anonymous referees and the handling editor Johannes Kraus for comments and questions which helped improve our presentation.

Appendix A. In this appendix, we further discuss the choice of the parameters used in our computations, namely the quantity c_m in (3.12) and the number of levels used in the MK methods.

From (3.11), we see that the quantity c_m is given by $\sigma_m(H_m)/(m\|AZE^{-1}\|)$, where m here should be interpreted as the maximum number of possible iteration of the Krylov subspace. We have used $m = 90$ for Poisson's problem and $m = 100$ for the convection-diffusion equation; see Tables 4.1 and 4.2. We run our codes for problems with order n between $4 \cdot 10^2$ and 10^5 and computed $\sigma_m(H_m)$, the smallest singular values of the upper Hessenberg matrix in the inexact Arnoldi method from (3.3). These are shown in Figure A.1 (left) and Figure A.2 (left) for the Poisson problem and the convection-diffusion equation (for $Pe = 200$), respectively. In both cases, $\sigma_m(H_m)$ varies within a narrow range and is of the order of 10^{-1}.

We approximated the value of $\|AZE^{-1}\|$ using appropriately the power method up to a precision of 10^{-6}. We plotted the ration $\sigma_m(H_m)/\|AZE^{-1}\|$ for these two problems for

ETNA
Kent State University and
Johann Radon Institute (RICAM)

different grid sizes in Figure A.1 (right) and Figure A.2 (right). As it can be appreciated, the calculation gives a value of c_m of order of $10^{-2}/m$, i.e., of the order of 10^{-4}.

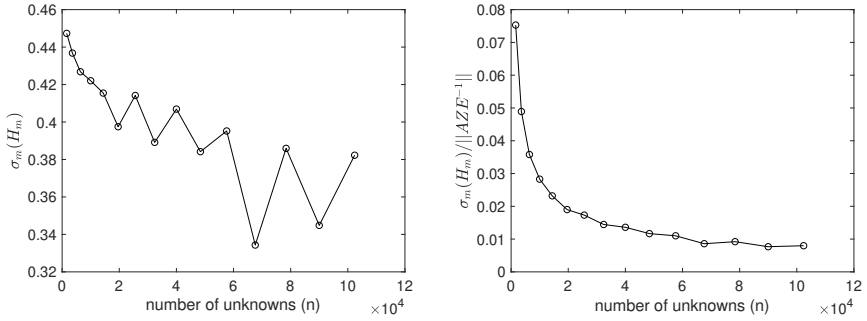


FIG. A.1. *Left: smallest singular value of the Hessenberg matrix. Right: values for* mc_m.

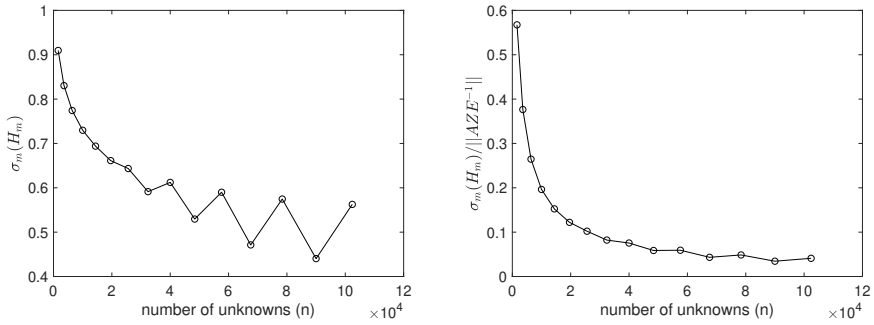


FIG. A.2. *Left: smallest singular value of the Hessenberg matrix. Right: values for* mc_m.

TABLE A.1
Number of calculated coarsest-grid solutions (and number of iterations) for 7*- and* 9*-level methods.*

Grid	MK(8,2,2)	MKadap	MKstat	MK(2,2,2)
400^2	2816(22)	2480(24)[18]	2944(65)	1248(39)
800^2	11776(23)	10176(25)[18]	8320(38)	3248(41)

These values could be used for the adaptive parameter, but they were shown to be too restrictive, as our experiments with $c_m = 10$ show. The reason for this discrepancy was already discussed in [26], where it is explained that the bounds used in its derivation such as (3.10) are far from tight.

We end this appendix by exploring the question on whether using more than five levels in the multilevel procedures would be advisable. We compare the solution of the Poisson problem for n with values 400^2 and 800^2 with a 7- and 9-level method, respectively. We perform 8 steps on the second and 2 on every coarser level. The results are given in Table A.1. These results should be compared with those of Table 4.1. It can be observed that for this

ETNA
Kent State University and
Johann Radon Institute (RICAM)

problem, the number of (outer) Krylov iterations does not change with more than 5 levels. On the other hand, the number of coarsest-grid solutions soars, which implies a much larger execution time.

REFERENCES

[1] A. BOURAS AND V. FRAYSSÉ, *Inexact matrix-vector products in Krylov methods for solving linear systems: a relaxation strategy*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 660–678.
[2] A. BOURAS, V. FRAYSSÉ, AND L. GIRAUD, *A relaxation strategy for inner-outer linear solvers in domain decomposition methods*, Tech. Report TR/PA/00/17, CERFACS, Toulouse, France, 2000.
[3] A. N. BROOKS AND T. J. R. HUGHES, *Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations*, Comput. Methods Appl. Mech. Engrg., 32 (1982), pp. 199–259.
[4] H. C. ELMAN, O. G. ERNST, AND D. P. O'LEARY, *A multigrid method enhanced by Krylov subspace iteration for discrete Helmholtz equations*, SIAM J. Sci. Comput., 23 (2001), pp. 1291–1315.
[5] H. C. ELMAN, D. J. SILVESTER, AND A. J. WATHEN, *Finite Elements and Fast Iterative Solvers*, 2nd ed., Oxford University Press, Oxford, 2014.
[6] B. ENGQUIST AND A. MAJDA, *Absorbing boundary conditions for the numerical simulation of waves*, Math. Comp., 31 (1977), pp. 629–651.
[7] Y. A. ERLANGGA AND R. NABBEN, *Deflation and balancing preconditioners for Krylov subspace methods applied to nonsymmetric matrices*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 684–699.
[8] ———, *Multilevel projection-based nested Krylov iteration for boundary value problems*, SIAM J. Sci. Comput., 30 (2008), pp. 1572–1595.
[9] ———, *On a multilevel Krylov method for the Helmholtz equation preconditioned by shifted Laplacian*, Electron. Trans. Numer. Anal., 31 (2008), pp. 403–424. http://etna.ricam.oeaw.ac.at/vol.31.2008/pp403-424.dir/pp403-424.pdf
[10] ———, *Algebraic multilevel Krylov methods*, SIAM J. Sci. Comput., 31 (2009), pp. 3417–3437.
[11] Y. A. ERLANGGA, C. W. OOSTERLEE, AND C. VUIK, *A novel multigrid based preconditioner for the heterogeneous Helmholtz equation*, SIAM J. Sci. Comput., 27 (2006), pp. 1471–1492.
[12] Y. A. ERLANGGA, C. VUIK, AND C. W. OOSTERLEE, *On a class of preconditioners for solving the Helmholtz equation*, Appl. Numer. Math., 50 (2004), pp. 409–425.
[13] ———, *Comparison of multigrid and incomplete LU shifted-Laplace preconditioners for the inhomogeneous Helmholtz equation*, Appl. Numer. Math., 56 (2006), pp. 648–666.
[14] J. FRANK AND C. VUIK, *On the construction of deflation-based preconditioners*, SIAM J. Sci. Comput., 23 (2001), pp. 442–462.
[15] L. GARCÍA RAMOS AND R. NABBEN, *On the spectrum of deflated matrices with applications to the deflated shifted Laplace preconditioner for the Helmholtz equation*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 262–286.
[16] J. KRAUS, *An algebraic preconditioning method for M-matrices: linear versus non-linear multilevel iteration*, Numer. Linear Algebra Appl., 9 (2002), pp. 599–618.
[17] ———, *Additive Schur complement approximation and application to multilevel preconditioning*, SIAM J. Sci. Comput., 34 (2012), pp. A2872–A2895.
[18] R. KEHL, L. GARCÍA RAMOS, AND R. NABBEN, *Projections, deflation and multigrid for non-symmetric matrices*, SIAM J. Matrix Anal. Appl., 41 (2020), to appear.
[19] J. LIESEN AND Z. STRAKOŠ, *Krylov subspace methods. Principles and Analysis*, Oxford University Press, Oxford, 2013.
[20] R. NABBEN AND C. VUIK, *A comparison of deflation and coarse grid correction applied to porous media flow*, SIAM J. Numer. Anal., 42 (2004), pp. 1631–1647.
[21] Y. NOTAY AND P. S. VASSILEVSKI, *Recursive Krylov-based multigrid cycles*, Numer. Linear Algebra Appl., 15 (2008), pp. 473–487.
[22] Y. SAAD, *A flexible inner-outer preconditioned GMRES algorithm*, SIAM J. Sci. Comput., 14 (1993), pp. 461–469.
[23] A. H. SHEIKH, D. LAHAYE, L. GARCÍA RAMOS, R. NABBEN, AND C. VUIK, *Accelerating the shifted Laplace preconditioner for the Helmholtz equation by multilevel deflation*, J. Comput. Phys., 322 (2016), pp. 473–490.
[24] A. H. SHEIKH, D. LAHAYE, AND C. VUIK, *On the convergence of shifted Laplace preconditioner combined with multilevel deflation*, Numer. Linear Algebra Appl., 20 (2013), pp. 645–662.
[25] V. SIMONCINI AND D. B. SZYLD, *Flexible inner-outer Krylov subspace methods*, SIAM J. Numer. Anal., 40 (2002), pp. 2219–2239 (2003).
[26] ———, *Theory of inexact Krylov subspace methods and applications to scientific computing*, SIAM J. Sci. Comput., 25 (2003), pp. 454–477.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[27] ———, *On the occurrence of superlinear convergence of exact and inexact Krylov subspace methods*, SIAM Rev., 47 (2005), pp. 247–272.

[28] J. M. TANG, R. NABBEN, C. VUIK, AND Y. A. ERLANGGA, *Comparison of two-level preconditioners derived from deflation, domain decomposition and multigrid methods*, J. Sci. Comput., 39 (2009), pp. 340–370.

[29] A. TOSELLI AND O. B. WIDLUND, *Domain Decomposition Methods—Algorithms and Theory*, Springer, Berlin, 2005.