SIAM J. MATRIX ANAL. APPL.
Vol. 34, No. 2, pp. 495–518

A FRAMEWORK FOR DEFLATED AND AUGMENTED KRYLOV SUBSPACE METHODS*

ANDRÉ GAUL†, MARTIN H. GUTKNECHT‡, JÖRG LIESEN†, AND REINHARD NABBEN†

Abstract. We consider deflation and augmentation techniques for accelerating the convergence of Krylov subspace methods for the solution of nonsingular linear algebraic systems. Despite some formal similarity, the two techniques are conceptually different from preconditioning. Deflation (in the sense the term is used here) "removes" certain parts from the operator making it singular, while augmentation adds a subspace to the Krylov subspace (often the one that is generated by the singular operator); in contrast, preconditioning changes the spectrum of the operator without making it singular. Deflation and augmentation have been used in a variety of methods and settings. Typically, deflation is combined with augmentation to compensate for the singularity of the operator, but both techniques can be applied separately. We introduce a framework of Krylov subspace methods that satisfy a Galerkin condition. It includes the families of orthogonal residual and minimal residual methods. We show that in this framework augmentation can be achieved either explicitly or, equivalently, implicitly by projecting the residuals appropriately and correcting the approximate solutions in a final step. We study conditions for a breakdown of the deflated methods, and we show several possibilities to avoid such breakdowns for the deflated minimum residual (MINRES) method. Numerical experiments illustrate properties of different variants of deflated MINRES analyzed in this paper.

Key words. Krylov subspace methods, augmentation, deflation, subspace recycling, CG, MINRES, GMRES, RMINRES

AMS subject classifications. 65F10, 65F08

DOI. 10.1137/110820713

1. Introduction.

There are numerous techniques to accelerate the speed of convergence of Krylov subspace methods for solving large linear algebraic systems

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

where $\mathbf{A} \in \mathbb{C}^{N\times N}$ is nonsingular and $\mathbf{b} \in \mathbb{C}^N$. The most widely used technique is *preconditioning*. Here the system (1.1) is modified using left- or right-multiplications with a *nonsingular* matrix (called the preconditioner). A typical goal of preconditioning is to obtain a modified matrix that is in some sense close to the identity matrix. For surveys of preconditioning techniques we refer to the books by Greenbaum [26, Part II] and Saad [46, Chapters 9–14] and the survey of Benzi [3].

Here we consider two approaches for convergence acceleration that are called *deflation* and *augmentation*. Let us briefly describe the main ideas of the two techniques. In deflation the system (1.1) is multiplied (at least implicitly) with a suitably chosen projection, and the general goal is to "eliminate" components that supposedly slow down convergence. Typically these are components that correspond to small

*Received by the editors January 11, 2011; accepted for publication (in revised form) February 13, 2013; published electronically May 14, 2013.
http://www.siam.org/journals/simax/34-2/82071.html

†Institut für Mathematik, Technische Universität Berlin, D-10623 Berlin, Germany (gaul@math.tu-berlin.de, liesen@math.tu-berlin.de, nabben@math.tu-berlin.de). These authors were supported by the DFG Forschungszentrum MATHEON. The work of the third author was also supported by the Heisenberg Program of the DFG.

‡Seminar for Applied Mathematics, ETH Zurich, CH-8092 Zurich, Switzerland (mhg@math.ethz.ch). This author was supported by the DFG Forschungszentrum MATHEON and the Mercator Visiting Professorship Program of the DFG.

eigenvalues. Multiplication by the projection turns the system (1.1) into a consistent singular one, which is then solved by a Krylov subspace method. We need to mention, however, that techniques have been proposed that move small eigenvalues of **A** to some large common value, say, to the value 1; see [1, 17, 31]. Some authors refer to these techniques as "deflation" too. In augmentation techniques the search space of the Krylov subspace method, which is at the same time the Galerkin test space, is "enlarged" by a suitably chosen subspace. A typical goal is to add information about the problem to the search space that is slowly revealed in the Krylov subspace itself, e.g., eigenvectors corresponding to small eigenvalues.

Deflation and augmentation techniques can be combined with conventional preconditioning techniques. Then the projection and augmentation parameters have to be adapted to the preconditioned matrix. In this paper, we assume that (1.1) is already in preconditioned form, i.e., **A** is the preconditioned matrix and **b** the preconditioned right-hand side. Details of preconditioning techniques will thus not be addressed here.

We will now give a brief overview of existing deflation and augmentation strategies. For a more comprehensive presentation we refer to section 9 of the survey article by Simoncini and Szyld [49]. The first deflation and augmentation techniques in the context of Krylov subspace methods appeared in the papers of Nicolaides [41] and Dostál [12]. Both proposed deflated variants of the CG method [29] to accelerate the speed of convergence for symmetric positive definite (spd) matrices **A** arising from discretized elliptic partial differential equations. Since these early works deflation and augmentation have become widely used tools. Several authors working in different fields of numerical analysis applied them to many Krylov subspace methods, and they use a variety of techniques to determine a deflation subspace. A review of all applications is well beyond this introduction. We concentrate in the following on some—but not all—key contributions.

For nonsymmetric systems Morgan [36] and also Chapman and Saad [6] extracted approximate eigenvectors of **A** from the Krylov subspace generated by the GMRES method [47], and then they augmented the Krylov subspace with these vectors; for related references we refer to [22]. A comparable approach in the context of the CG method for spd matrices **A** was described by Saad, Yeung, Erhel, and Guyomarc'h [48]. De Sturler [10] introduced the GCRO method, which involves an outer GCR iteration [15, 16] and an inner deflated GMRES method where the space used for deflation depends on the outer iteration. This method has been extended to GCROT in [11] to incorporate truncation strategies when restarts are necessary. In [33] Kolotilina used a twofold deflation technique for simultaneously deflating the r largest and the r smallest eigenvalues by an appropriate deflating subspace of dimension r. An analysis of acceleration strategies (including augmentation) for minimal residual methods was given by Saad [45] and for restarted methods by Eiermann, Ernst, and Schneider [14]. The latter work analyzes minimal residual (MR) and orthogonal residual (OR) methods in a general framework that allows approximations from arbitrary correction spaces. By using multiple correction spaces forming a direct sum, several cases of augmentation and deflation are discussed. The analysis concentrates on (nearly) **A**-invariant augmentation spaces.

In [37] Morgan proposed a block-GMRES method for multiple right-hand sides that deflates approximated eigenvectors when GMRES is restarted. A similar method for solving systems with multiple shifts and multiple right-hand sides has been introduced by Darnell, Morgan, and Wilcox [9]. Giraud et al. [25] recently developed a flexible GMRES variant with deflated restarting where the preconditioner may vary

from one iteration to the next. In [42] Olshanskii and Simoncini studied spectral properties of saddle point matrices preconditioned with a block-diagonal preconditioner and applied a deflated minimum residual (MinRes) method to the resulting symmetric and indefinite matrix in order to alleviate the influence of a few small outlying eigenvalues. Theoretical results for deflated GMRes based on an exactly $\mathbf{A}$-invariant subspace have been presented in [61].

In addition to deflation/augmentation spaces based on approximative eigenvectors, other choices have been studied. Mansfield [34] showed how Schur complement-type domain decomposition methods can be seen as a series of deflations. Nicolaides [41] constructed a deflation technique based on piecewise constant interpolation from a set of r subdomains, and he pointed out that deflation might be effectively used with a conventional preconditioner. In [35] Mansfield used the same "subdomain deflation" in combination with damped Jacobi smoothing and obtained a preconditioner that is related to the two-grid method. Baker, Jessup, and Manteuffel [2] proposed a GMRes method that is augmented upon restarts by approximations to the error.

In [38, 39, 40] Nabben and Vuik described similarities between the deflation approach and domain decomposition methods for arbitrary deflation spaces. This comparison was extended to multigrid methods in [54, 53].

This brief survey indicates that in principle deflation or augmentation can be incorporated into every Krylov subspace method. However, some methods may suffer from mathematical shortcomings like breakdowns or numerical problems due to round-off errors. The main goal of this paper is not to add further examples to the existing collection but to introduce first a suitable framework for a whole family of such augmented and deflated methods (section 2) and then to prove some results just assuming this framework (section 3). The framework focuses on Krylov subspace methods whose residuals satisfy a certain Galerkin condition with respect to a true or formal inner product. In section 3, we mathematically characterize the equivalence of two approaches for realizing such methods and discuss them along with potential pitfalls. We then discuss known approaches to deflate CG (section 4), GMRes (section 5), and MinRes (section 6) in the light of our general equivalence theorem. Among other results, this will show that a recent version of deflated MinRes, which is part of the "recycling" MinRes (RMinRes) method suggested by Wang, de Sturler, and Paulino [57], can break down and will show how these breakdowns can be avoided by either adapting the right-hand side or the initial guess. We do not focus on specific implementations or algorithmic details but on the mathematical theory of these methods. For the numerical application in section 6.3 we draw on the most robust MinRes implementation that is available.

2. A framework for deflated and augmented Krylov methods. In this section we describe a general framework for deflation and augmentation, which simultaneously covers several Krylov subspace methods whose residuals satisfy a Galerkin condition. Given an initial guess $\mathbf{x}_0 \in \mathbb{C}^N$, a positive integer n, an n-dimensional subspace $\mathcal{S}_n$ of $\mathbb{C}^N$, and a nonsingular matrix $\mathbf{B} \in \mathbb{C}^{N\times N}$, let us first consider an approximation $\mathbf{x}_n$ to the solution $\mathbf{x}$ of the form

$$\mathbf{x}_n \in \mathbf{x}_0 + \mathcal{S}_n, \tag{2.1}$$

so that the corresponding residual

$$\mathbf{r}_n := \mathbf{b} - \mathbf{A}\mathbf{x}_n \in \mathbf{r}_0 + \mathbf{A}\mathcal{S}_n$$

satisfies

$$\mathbf{r}_n \perp \mathbf{B}\mathcal{S}_n. \tag{2.2}$$

If $\mathbf{B}^{\mathsf{H}}\mathbf{A}$ is Hermitian positive definite (Hpd), then $\mathbf{B}^{\mathsf{H}}\mathbf{A}$ induces an inner product $\langle\cdot,\cdot\rangle_{\mathbf{B}^{\mathsf{H}}\mathbf{A}}$, a corresponding norm $\|\cdot\|_{\mathbf{B}^{\mathsf{H}}\mathbf{A}}$, and an orthogonality $\perp_{\mathbf{B}^{\mathsf{H}}\mathbf{A}}$. Imposing (2.1) and (2.2) can then be seen to be equivalent to solving the following minimization problem:

$$\text{find} \quad \mathbf{x}_n \in \mathbf{x}_0 + \mathcal{S}_n \quad \text{s.t.} \quad \|\mathbf{x}-\mathbf{x}_n\|_{\mathbf{B}^{\mathsf{H}}\mathbf{A}} = \min_{\mathbf{y}\in\mathbf{x}_0+\mathcal{S}_n} \|\mathbf{x}-\mathbf{y}\|_{\mathbf{B}^{\mathsf{H}}\mathbf{A}}. \tag{2.3}$$

Note that due to $\mathbf{r}_n = \mathbf{A}(\mathbf{x}-\mathbf{x}_n)$ the condition (2.2) can be written as orthogonality condition for the error $\mathbf{x}-\mathbf{x}_n$:

$$(\mathbf{x}-\mathbf{x}_n) \perp_{\mathbf{B}^{\mathsf{H}}\mathbf{A}} \mathcal{S}_n. \tag{2.4}$$

The following two cases where $\mathbf{B}^{\mathsf{H}}\mathbf{A}$ is Hpd are of particular interest:

(1) $\mathbf{B} = \mathbf{I}$ if $\mathbf{A}$ itself is Hpd;

(2) $\mathbf{B} = \mathbf{A}$ for general nonsingular $\mathbf{A}$.

The case (1) is the one where (2.2) is a typical Galerkin condition: $\mathbf{A}$ is Hpd and the residual $\mathbf{r}_n$ is orthogonal to the linear search space $\mathcal{S}_n$ for $\mathbf{x}_n - \mathbf{x}_0$. In (2.3) we then have

$$\|\mathbf{x}-\mathbf{x}_n\|_{\mathbf{A}} = \|\mathbf{r}_n\|_{\mathbf{A}^{-1}}, \tag{2.5}$$

so while the error is minimal in the $\mathbf{A}$–norm, the residual is minimal in the $\mathbf{A}^{-1}$–norm.

In this paper we will refer to (2.2) also in case (2) as a Galerkin condition, because the search space and the test space are still essentially the same. However, in this case

$$\|\mathbf{x}-\mathbf{x}_n\|_{\mathbf{A}^{\mathsf{H}}\mathbf{A}} = \|\mathbf{r}_n\|_2, \tag{2.6}$$

so (2.2) implies that the 2–norm of the residual is minimized. Consequently, in both cases a minimization property holds.

If the search space $\mathcal{S}_n$ is the nth Krylov subspace generated by $\mathbf{A}$ and the initial residual $\mathbf{r}_0 := \mathbf{b} - \mathbf{A}\mathbf{x}_0$, i.e., if

$$\mathcal{S}_n = \mathcal{K}_n(\mathbf{A},\mathbf{r}_0) := \mathsf{span}\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \ldots, \mathbf{A}^{n-1}\mathbf{r}_0\}, \tag{2.7}$$

then, in case (1), conditions (2.1)–(2.2) mathematically characterize the CG method [29]. It is the prototype of an OR method characterized by (2.1) and (2.2) with $\mathbf{B} = \mathbf{I}$.

In case (2), conditions (2.1)–(2.2) with $\mathcal{S}_n = \mathcal{K}_n(\mathbf{A},\mathbf{r}_0)$ mathematically characterize the GCR [15] and GMRES [47] methods and, for Hermitian $\mathbf{A}$, the MINRES [43] method. If $\mathbf{A}$ is even Hpd, we can resort to Stiefel's conjugate residual (CR) method [52]. All these are prototype MR methods characterized by (2.1) and (2.2) with $\mathbf{B} = \mathbf{A}$.

OR and MR methods often come in pairs defined by the properties of $\mathbf{A}$, the Krylov search space, and, to some extent, the fundamental structure of the algorithms. Examples of such pairs are CG/CR, GCG/GCR, full orthogonalization method (FOM)/GMRES, and CGNE/CGNR. It has been pointed out many times, see, e.g., [4, 8, 13, 14, 27], that the residuals of these pairs of OR/OM methods and in particular the residual norms are related in a simple fashion.

A fact related to the OR/MR residual connection is that the iterates and residuals of an MR method can be found from those of the corresponding OR method by a smoothing process introduced by Schönauer; see [58, 60, 27, 28]. The reverse process also exists [27]. Again these processes hold for the residuals of the deflated system and, since they are identical, for those of the explicit augmentation approach.

If $\mathbf{B}^{\mathsf{H}}\mathbf{A}$ is not Hpd, the minimization property (2.3) no longer makes sense, but we may still request that the orthogonality condition (2.2), or, equivalently, (2.4), hold. Resulting algorithms may then break down since an approximate solution $\mathbf{x}_n$ satisfying the conditions may not exist for some n. Nevertheless, such methods are occasionally applied in practice. In particular, the choice

(3) $\mathbf{B} = \mathbf{I}$ and $\mathbf{A}$ nonsingular

covers the FOM of Saad [44, 46], which is sometimes also referred to as Arnoldi method for linear algebraic systems.

For minimizing the error $\mathbf{x}_n - \mathbf{x}$ in the 2-norm one has to choose

(4) $\mathbf{B} = \mathbf{A}^{-\mathsf{H}}$ and $\mathbf{A}$ nonsingular.

Since multiplication by $\mathbf{A}^{-\mathsf{H}}$ is not feasible, these methods only work for particular search spaces; the simplest choice is

$$\mathcal{S}_n = \mathbf{A}^{\mathsf{H}}\mathcal{K}_n\left(\mathbf{A}^{\mathsf{H}}, \mathbf{r}_0\right). \tag{2.8}$$

Unlike the normal Krylov search space of (2.7), this one has the drawback that the (exact) solution of the system need not be in one of these spaces, i.e., even in exact arithmetic convergence is not guaranteed. One interesting example based on this choice is the generalized minimum error (GMErr) method of Weiss [59]. Earlier, for spd matrices, such a method was proposed by Fridman [24], and an alternative algorithm was mentioned by Fletcher [21]. Symmetric indefinite systems can be treated in this way with the SymmLQ algorithm of Paige and Saunders [43]; see also Freund [23] for a review of methods featuring this optimality criterion and yet another algorithm called ME to achieve it.

Finally, we can easily incorporate the CGNR method [29] for solving overdetermined linear systems in the setting of (2.1) and (2.2) by choosing the appropriate Krylov search space. Given such a system

$$\mathbf{E}\mathbf{x} = \mathbf{f} \tag{2.9}$$

with a full-rank $M \times N$–matrix $\mathbf{E}$ (where $M \geq N$), the corresponding normal equations are $\mathbf{E}^{\mathsf{H}}\mathbf{E}\mathbf{x} = \mathbf{E}^{\mathsf{H}}\mathbf{f}$, i.e., $\mathbf{A}\mathbf{x} = \mathbf{b}$ with $\mathbf{A} := \mathbf{E}^{\mathsf{H}}\mathbf{E}$ and $\mathbf{b} := \mathbf{E}^{\mathsf{H}}\mathbf{f}$. Since $\mathbf{A}$ is Hpd, we can apply the CG method which corresponds to case (1) and

$$\mathcal{S}_n = \mathcal{K}_n\left(\mathbf{E}^{\mathsf{H}}\mathbf{E}, \mathbf{E}^{\mathsf{H}}\mathbf{s}_0\right) \tag{2.10}$$

with $\mathbf{s}_0 := \mathbf{f} - \mathbf{E}\mathbf{x}_0$. In this situation we have to distinguish between the residuals $\mathbf{r}_n := \mathbf{b} - \mathbf{A}\mathbf{x}_n = \mathbf{E}^{\mathsf{H}}\mathbf{f} - \mathbf{E}^{\mathsf{H}}\mathbf{E}\mathbf{x}_n$ of the normal equations and the residuals $\mathbf{s}_n := \mathbf{f} - \mathbf{E}\mathbf{x}_n$ of the given system (2.9). The CGNR method allows one to keep track of both. The latter residuals satisfy

$$\mathbf{s}_n \in \mathbf{s}_0 + \mathbf{E}\mathcal{K}_n(\mathbf{E}^{\mathsf{H}}\mathbf{E}, \mathbf{E}^{\mathsf{H}}\mathbf{f}), \qquad \mathbf{s}_n \perp \mathbf{E}\mathcal{K}_n(\mathbf{E}^{\mathsf{H}}\mathbf{E}, \mathbf{E}^{\mathsf{H}}\mathbf{f}), \tag{2.11}$$

and they can be seen to minimize the 2-norm of $\mathbf{s}_n$. Note that it can be viewed as an MR method with a possibly nonsquare $\mathbf{B} = \mathbf{E}$; see [27].

A method that also fits into our framework, though with some modifications, is the CGNE method, also called Craig's method [7], which can also be used for solving

underdetermined linear algebraic systems (2.9) with a full-rank $M \times N$–matrix $\mathbf{E}$. The search space for $\mathbf{x}_n$ in this case is (2.10), but the Galerkin condition becomes $\mathbf{s}_n \perp \mathcal{K}_n(\mathbf{E}\mathbf{E}^{\mathsf{H}}, \mathbf{s}_0)$.

Since we are aiming at a *general framework*, let us for the moment consider an arbitrary, possibly singular matrix $\widehat{\mathbf{A}} \in \mathbb{C}^{N\times N}$ and an arbitrary vector $\widehat{\mathbf{v}} \in \mathbb{C}^N$, such that the Krylov subspace $\mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{v}})$ has dimension n.

Instead of a search space of the form $\mathcal{S}_n = \mathcal{K}_n(\mathbf{A}, \mathbf{r}_0)$ we focus from now on augmented Krylov subspaces of the form

$$\mathcal{S}_n := \mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{v}}) + \mathcal{U}. \tag{2.12}$$

We suppose that $\mathcal{U}$ has dimension k, $0 < k < N$, and denote by $\mathbf{U} \in \mathbb{C}^{N\times k}$ a matrix whose columns form a basis of $\mathcal{U}$, and by $\mathbf{V}_n \in \mathbb{C}^{N\times n}$ one whose columns form a basis of $\mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{v}})$, so that (2.1) can be written as

$$\mathbf{x}_n = \mathbf{x}_0 + \mathbf{V}_n\mathbf{y}_n + \mathbf{U}\mathbf{u}_n \tag{2.13}$$

for some vectors $\mathbf{y}_n \in \mathbb{C}^n$ and $\mathbf{u}_n \in \mathbb{C}^k$. Of course, $\mathbf{U}$ may be redefined when an algorithm like GMRES is restarted, but we will not account for that in our notation.

Assuming the general structure of the search space $\mathcal{S}_n$ in (2.12) we will now investigate augmented Galerkin-type methods that still satisfy (2.1) and (2.2).

3. A general equivalence theorem. Our goal in this section is to show that augmentation can be achieved either explicitly as in (2.12), or implicitly, namely, by projecting the residuals appropriately and correcting the approximate solutions in a final step. Our main result is stated in Theorem 3.2 below.

In order to satisfy (2.2), the residual $\mathbf{r}_n = \mathbf{b} - \mathbf{A}\mathbf{x}_0$ must be orthogonal to both $\mathbf{B}\mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{v}})$ and $\mathbf{B}\mathcal{U}$, hence it must satisfy the pair of orthogonality conditions

$$\mathbf{r}_n \perp \mathbf{B}\,\mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{v}}) \quad \text{and} \quad \mathbf{r}_n \perp \mathbf{B}\,\mathcal{U}. \tag{3.1}$$

Let us concentrate on the second condition of (3.1), which can be written as

$$\mathbf{0} = \mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{r_n} = \mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\left(\mathbf{r_0} - \mathbf{A}\mathbf{V_n}\mathbf{y_n} - \mathbf{A}\mathbf{U}\mathbf{u_n}\right) = \mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\left(\mathbf{r_0} - \mathbf{A}\mathbf{V_n}\mathbf{y_n}\right) - \mathbf{E_B}\mathbf{u_n},$$

where

$$\mathbf{E_B} := \mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{A}\mathbf{U} \in \mathbb{C}^{k\times k}. \tag{3.2}$$

Clearly, if $\mathbf{B}^{\mathsf{H}}\mathbf{A}$ is Hpd, then $\mathbf{E_B}$ is Hpd too—though in a smaller space—and thus nonsingular. In the following derivation, $\mathbf{B}^{\mathsf{H}}\mathbf{A}$ need not be Hpd, but we must then assume that $\mathbf{E_B}$ is nonsingular. Then the second orthogonality condition is equivalent to

$$\mathbf{u}_n = \mathbf{E_B}^{-1}\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\left(\mathbf{r}_0 - \mathbf{A}\mathbf{V}_n\mathbf{y}_n\right). \tag{3.3}$$

Substituting this into (2.13) gives

$$\begin{aligned}\mathbf{x}_n &= \mathbf{x}_0 + \mathbf{V}_n\mathbf{y}_n + \mathbf{U}\left(\mathbf{E_B}^{-1}\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\left(\mathbf{r}_0 - \mathbf{A}\mathbf{V}_n\mathbf{y}_n\right)\right)\\ &= \left(\mathbf{I} - \mathbf{U}\mathbf{E_B}^{-1}\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{A}\right)\left(\mathbf{x}_0 + \mathbf{V}_n\mathbf{y}_n\right) + \mathbf{U}\mathbf{E_B}^{-1}\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{b}, \qquad (3.4)\\ \mathbf{r}_n &= \mathbf{r}_0 - \mathbf{A}\mathbf{V}_n\mathbf{y}_n - \mathbf{A}\mathbf{U}\left(\mathbf{E_B}^{-1}\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\left(\mathbf{r}_0 - \mathbf{A}\mathbf{V}_n\mathbf{y}_n\right)\right)\\ &= \left(\mathbf{I} - \mathbf{A}\mathbf{U}\mathbf{E_B}^{-1}\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\right)\left(\mathbf{r}_0 - \mathbf{A}\mathbf{V}_n\mathbf{y}_n\right). \qquad (3.5)\end{aligned}$$

To simplify the notation we define the $(N \times N)$–matrices

$$\begin{aligned} \mathbf{M_B} &:= \mathbf{U}\mathbf{E_B}^{-1}\mathbf{U}^{\mathsf{H}} = \mathbf{U}\left(\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{A}\mathbf{U}\right)^{-1}\mathbf{U}^{\mathsf{H}}, \\ \mathbf{P_B} &:= \mathbf{I} - \mathbf{A}\mathbf{M_B}\mathbf{B}^{\mathsf{H}}, \\ \mathbf{Q_B} &:= \mathbf{I} - \mathbf{M_B}\mathbf{B}^{\mathsf{H}}\mathbf{A}. \end{aligned} \tag{3.6}$$

Using these matrices (3.4) and (3.5) take the form

$$\mathbf{x}_n = \mathbf{Q_B}\left(\mathbf{x}_0 + \mathbf{V}_n\mathbf{y}_n\right) + \mathbf{M_B}\mathbf{B}^{\mathsf{H}}\mathbf{b}, \tag{3.7}$$

$$\mathbf{r}_n = \mathbf{P_B}\left(\mathbf{r}_0 - \mathbf{A}\mathbf{V}_n\mathbf{y}_n\right). \tag{3.8}$$

Note that imposing the second orthogonality condition in (3.1) on the residual $\mathbf{r}_n$ has determined the vector $\mathbf{u}_n$, which has therefore "disappeared" in (3.7)–(3.8). We next state some basic properties of the matrices $\mathbf{P_B}$ and $\mathbf{Q_B}$. The proof of these properties is straightforward and is therefore omitted.

LEMMA 3.1. *Let* $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{N\times N}$ *and* $\mathbf{U} \in \mathbb{C}^{N\times k}$ *be such that* $\mathbf{E_B} := \mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{A}\mathbf{U}$ *is nonsingular (which implies that* $\operatorname{rank}\mathbf{U} = k$*). Then the matrices in* (3.6) *are well defined and the following statements hold:*

1. $\mathbf{P_B}^2 = \mathbf{P_B}$, $\mathbf{P_B}\mathbf{A}\mathbf{U} = \mathbf{0}$, *and* $\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{P_B} = \mathbf{0}$, *i.e.,* $\mathbf{P_B}$ *is the projection onto* $(\mathbf{B}\mathcal{U})^{\perp}$ *along* $\mathbf{A}\mathcal{U}$.
2. $\mathbf{Q_B}^2 = \mathbf{Q_B}$, $\mathbf{Q_B}\mathbf{U} = \mathbf{0}$, *and* $\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{A}\mathbf{Q_B} = \mathbf{0}$, *i.e.,* $\mathbf{Q_B}$ *is the projection onto* $(\mathbf{A}^{\mathsf{H}}\mathbf{B}\mathcal{U})^{\perp}$ *along* $\mathcal{U}$.
3. $\mathbf{P_B}\mathbf{A} = \mathbf{P_B}\mathbf{A}\mathbf{Q_B} = \mathbf{A}\mathbf{Q_B}$.
4. $\mathbf{P_A} = \mathbf{P_A}^{\mathsf{H}}$, *i.e.,* $\mathbf{P_A}$ *is an orthogonal projection.*

It remains to impose the first orthogonality condition in (3.1), which will determine the vector $\mathbf{y}_n$. To this end, let

$$\widehat{\mathbf{x}}_n := \mathbf{x}_0 + \mathbf{V}_n\mathbf{y}_n \in \mathbf{x}_0 + \mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{v}}),$$

so that by (3.7) $\mathbf{x}_n = \mathbf{Q_B}\widehat{\mathbf{x}}_n + \mathbf{M_B}\mathbf{B}^{\mathsf{H}}\mathbf{b}$. Using the definition of $\mathbf{P_B}$ in (3.6) and statement 3 of Lemma 3.1, this orthogonality condition reads

$$\mathbf{r}_n = \mathbf{b} - \mathbf{A}\mathbf{x}_n = \mathbf{b} - \mathbf{A}\mathbf{Q_B}\widehat{\mathbf{x}}_n - \mathbf{A}\mathbf{M_B}\mathbf{B}^{\mathsf{H}}\mathbf{b} = \mathbf{P_B}(\mathbf{b} - \mathbf{A}\widehat{\mathbf{x}}_n) \perp \mathbf{B}\mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{v}}).$$

We summarize these considerations in the following theorem.

THEOREM 3.2. *Let the assumptions of Lemma* 3.1 *hold and let* $\widehat{\mathbf{A}} \in \mathbb{C}^{N\times N}$, $\widehat{\mathbf{v}} \in \mathbb{C}^N$, *and* $n \in \mathbb{N}$ *be such that the Krylov subspace* $\mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{v}})$ *has dimension* n. *Furthermore, let* $\mathbf{b}, \mathbf{x}_0 \in \mathbb{C}^N$ *be arbitrary.*

Then, with $\mathcal{U} := \mathsf{im}\,(\mathbf{U})$ *and the definitions from* (3.6) *the two pairs of conditions,*

$$\begin{aligned} &\mathbf{x}_n \in \mathbf{x}_0 + \mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{v}}) + \mathcal{U}, \\ &\mathbf{r}_n := \mathbf{b} - \mathbf{A}\mathbf{x}_n \perp \mathbf{B}\mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{v}}) + \mathbf{B}\mathcal{U} \end{aligned} \tag{3.9}$$

and

$$\begin{aligned} &\widehat{\mathbf{x}}_n \in \mathbf{x}_0 + \mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{v}}), \\ &\widehat{\mathbf{r}}_n := \mathbf{P_B}(\mathbf{b} - \mathbf{A}\widehat{\mathbf{x}}_n) \perp \mathbf{B}\mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{v}}) \end{aligned} \tag{3.10}$$

are equivalent for $n \geq 1$ *in the sense that*

$$\mathbf{x}_n = \mathbf{Q_B}\widehat{\mathbf{x}}_n + \mathbf{M_B}\mathbf{B}^{\mathsf{H}}\mathbf{b} \quad \textit{and} \quad \mathbf{r}_n = \widehat{\mathbf{r}}_n. \tag{3.11}$$

We call (3.9) the *explicit* deflation and augmentation approach because the augmentation space $\mathcal{U}$ is explicitly included in the search space. The equivalent conditions (3.10) show that the explicit inclusion of $\mathcal{U}$ can be omitted when instead we first construct the iterate $\widehat{\mathbf{x}}_n \in \mathbf{x}_0 + \mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{v}})$ so that the projected residual $\widehat{\mathbf{r}}_n = \mathbf{P_B}(\mathbf{b} - \mathbf{A}\widehat{\mathbf{x}}_n)$ satisfies the given orthogonality condition and then apply the affine correction (3.11) to $\widehat{\mathbf{x}}_n$, whose projected residual equals the one of $\mathbf{x}_n$. We call this second option the *implicit* deflation and augmentation approach.

Note that the theorem makes no assumption on relations between $\widehat{\mathbf{A}}$, $\widehat{\mathbf{v}}$, and $\mathcal{U}$. The only assumption on the augmentation space $\mathcal{U}$ is that the matrix $\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{A}\mathbf{U}$ is nonsingular. (Clearly, if this holds for one basis of $\mathcal{U}$ it holds for all.) Moreover, in the theorem $\widehat{\mathbf{A}}$ and $\widehat{\mathbf{v}}$ are arbitrary except for the assumption that $\mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{v}})$ has dimension n.

In practice, $\widehat{\mathbf{A}}$ and $\widehat{\mathbf{v}}$ should be somehow related to $\mathbf{A}$, however. One specific choice is suggested by Theorem 3.2, in particular (3.10). If

$$\widehat{\mathbf{A}} := \mathbf{P_B}\mathbf{A}, \quad \widehat{\mathbf{v}} := \widehat{\mathbf{r}}_0 := \mathbf{P_B}\mathbf{r}_0 = \mathbf{P_B}(\mathbf{b} - \mathbf{A}\mathbf{x}_0), \quad \text{and} \quad \widehat{\mathbf{b}} := \mathbf{P_B}\mathbf{b},$$

then (3.10) becomes

$$\begin{aligned} &\widehat{\mathbf{x}}_n \in \mathbf{x}_0 + \mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{r}}_0), \\ &\widehat{\mathbf{r}}_n := \widehat{\mathbf{b}} - \widehat{\mathbf{A}}\widehat{\mathbf{x}}_n \perp \mathbf{B}\mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{r}}_0), \end{aligned} \tag{3.12}$$

which is a formal Galerkin condition for the (consistent and singular) deflated system $\widehat{\mathbf{A}}\widehat{\mathbf{x}} = \widehat{\mathbf{b}}$. Based on the Jordan form of $\mathbf{A}$ we show in the following theorem what the Jordan form of $\widehat{\mathbf{A}} = \mathbf{P_B}\mathbf{A}$ looks like when (1) $\mathcal{U}$ is a right invariant subspace or (2) $\mathbf{B}\mathcal{U}$ is a left invariant subspace of $\mathbf{A}$.

THEOREM 3.3. *Suppose that the matrix $\mathbf{A} \in \mathbb{C}^{N\times N}$ has a partitioned Jordan decomposition of the form*

$$\mathbf{A} = \mathbf{S}\mathbf{J}\mathbf{S}^{-1} = \begin{bmatrix} \mathbf{S}_1 & \mathbf{S}_2 \end{bmatrix} \begin{bmatrix} \mathbf{J}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2 \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{S}}_1^{\mathsf{H}} \\ \widehat{\mathbf{S}}_2^{\mathsf{H}} \end{bmatrix}, \tag{3.13}$$

where $\mathbf{S}_1, \widehat{\mathbf{S}}_1 \in \mathbb{C}^{N\times k}$, $\mathbf{S}_2, \widehat{\mathbf{S}}_2 \in \mathbb{C}^{N\times(N-k)}$, $\mathbf{J}_1 \in \mathbb{C}^{k\times k}$, and $\mathbf{J}_2 \in \mathbb{C}^{(N-k)\times(N-k)}$. Then the following assertions hold:

(1) *If $\mathcal{U} = \mathsf{im}\,(\mathbf{S}_1)$, $\mathbf{U} \in \mathbb{C}^{N\times k}$ is any matrix satisfying $\mathsf{im}\,(\mathbf{U}) = \mathcal{U}$, and $\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{A}\mathbf{U}$ is nonsingular, then*

$$\widehat{\mathbf{A}} = \mathbf{P_B}\mathbf{A} = \begin{bmatrix} \mathbf{U} & \mathbf{P_B}\mathbf{S}_2 \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2 \end{bmatrix} \begin{bmatrix} \mathbf{U} & \mathbf{P_B}\mathbf{S}_2 \end{bmatrix}^{-1} \tag{3.14}$$

$$\text{with} \quad \begin{bmatrix} \mathbf{U} & \mathbf{P_B}\mathbf{S}_2 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{B}\mathbf{U}(\mathbf{U}^{\mathsf{H}}\mathbf{B}\mathbf{U})^{-1} & \widehat{\mathbf{S}}_2 \end{bmatrix}^{\mathsf{H}}.$$

(2) *If $\mathbf{B}\mathcal{U} = \mathsf{im}\,(\widehat{\mathbf{S}}_1)$, $\mathbf{U} \in \mathbb{C}^{N\times k}$ is any matrix satisfying $\mathsf{im}\,(\mathbf{U}) = \mathcal{U}$, and $\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{A}\mathbf{U}$ is nonsingular, then*

$$\widehat{\mathbf{A}} = \mathbf{P_B}\mathbf{A} = \begin{bmatrix} \mathbf{U} & \mathbf{S}_2 \end{bmatrix} \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_2 \end{bmatrix} \begin{bmatrix} \mathbf{U} & \mathbf{S}_2 \end{bmatrix}^{-1} \tag{3.15}$$

$$\text{with} \quad \begin{bmatrix} \mathbf{U} & \mathbf{S}_2 \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{B}\mathbf{U}(\mathbf{U}^{\mathsf{H}}\mathbf{B}\mathbf{U})^{-1} & \mathbf{Q}_{\mathbf{B}}^{\mathsf{H}}\widehat{\mathbf{S}}_2 \end{bmatrix}^{\mathsf{H}}.$$

In particular, in both cases the spectrum $\Lambda(\widehat{\mathbf{A}})$ of $\widehat{\mathbf{A}}$ is given by $\Lambda(\widehat{\mathbf{A}}) = \{0\} \cup \Lambda(\mathbf{J}_2)$.

Proof. (1) From Lemma 3.1 we can see that

$$\mathbf{P_B A U} = \mathbf{0} \quad \text{and} \quad \left(\mathbf{B U}(\mathbf{U}^{\mathsf{H}}\mathbf{B U})^{-1}\right)^{\mathsf{H}} \mathbf{P_B A} = \mathbf{0}.$$

By construction, there exists a nonsingular matrix $\mathbf{R} \in \mathbb{C}^{k\times k}$ with $\mathbf{AU} = \mathbf{UR}$. Hence $\mathbf{P_B} = \mathbf{I} - \mathbf{U}(\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{U})^{-1}\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}$ and from $\widehat{\mathbf{S}}_2^{\mathsf{H}}\mathbf{S}_1 = \mathbf{0}$ we conclude that $\widehat{\mathbf{S}}_2^{\mathsf{H}}\mathbf{U} = \mathbf{0}$ and thus

$$\widehat{\mathbf{S}}_2^{\mathsf{H}}\mathbf{P_B A} = \widehat{\mathbf{S}}_2^{\mathsf{H}}\mathbf{A} = \mathbf{J}_2\widehat{\mathbf{S}}_2^{\mathsf{H}}.$$

Furthermore,

$$\mathbf{P_B A}(\mathbf{P_B S}_2) = \mathbf{P_B A S}_2 - \mathbf{P_B AU}(\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{U})^{-1}\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{S}_2 = (\mathbf{P_B S}_2)\mathbf{J}_2$$

and the proof of (1) is complete after recognizing that

$$\begin{bmatrix} (\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}}\mathbf{U})^{-1}\mathbf{U}^{\mathsf{H}}\mathbf{B}^{\mathsf{H}} \\ \widehat{\mathbf{S}}_2^{\mathsf{H}} \end{bmatrix} \begin{bmatrix} \mathbf{U} & \mathbf{P_B S}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}.$$

The proof of (2) is analogous to (1). ∎

Results like the previous theorem motivate the term deflation, which means "making something smaller," since the multiplication with the operator $\mathbf{P_B}$ "removes" certain eigenvalues from the operator $\mathbf{A}$ by "moving them to zero." Special cases of the results shown in Theorem 3.3 have appeared in the literature: in particular, for spd matrices $\mathbf{A}$ and $\mathbf{B} = \mathbf{I}$ in the works of Frank and Vuik [22] and Nabben and Vuik [38, 39], and for nonsymmetric $\mathbf{A}$ and $\mathbf{B} = \mathbf{I}$ in the articles by Erlangga and Nabben [19, 20] and Yeung, Tang, and Vuik [61].

4. Hpd matrices and CG. This section presents some well-known results for deflation and augmentation techniques within the framework described in section 2 in the case where $\mathbf{A}$ is Hpd. The first proposed deflated Krylov subspace methods for Hpd matrices are the deflated CG variants of Nicolaides [41] and Dostál [12]. With a full-rank matrix $\mathbf{U} \in \mathbb{C}^{N\times k}$, $\mathcal{U} = \mathsf{im}\,(\mathbf{U})$ and $\mathbf{B} = \mathbf{I}$ both essentially apply the CG method to the deflated system

$$\widehat{\mathbf{A}}\widehat{\mathbf{x}} = \widehat{\mathbf{b}}, \quad \text{where} \quad \widehat{\mathbf{A}} := \mathbf{P_I A}, \quad \widehat{\mathbf{b}} := \mathbf{P_I b}. \tag{4.1}$$

Here, $\mathbf{P_I} = \mathbf{I} - \mathbf{AU}(\mathbf{U}^{\mathsf{H}}\mathbf{AU})^{-1}\mathbf{U}^{\mathsf{H}}$ is the projection onto $\mathcal{U}^{\perp}$ along $\mathbf{A}\mathcal{U}$ as defined in (3.6) when $\mathbf{B} = \mathbf{I}$. Moreover, $\mathbf{Q_I} = \mathbf{P_I}^{\mathsf{H}}$ is then the projection onto $(\mathbf{A}\mathcal{U})^{\perp}$ along $\mathcal{U}$. Note that all the matrices in (3.6) are well defined because $\mathbf{E_I} = \mathbf{U}^{\mathsf{H}}\mathbf{AU}$ is Hpd if $\mathbf{A}$ is Hpd. Clearly, the deflated matrix $\widehat{\mathbf{A}}$ is Hermitian but singular, since $\mathbf{P_I}$ is a nontrivial projection if $0 < k < N$. In fact, this matrix $\widehat{\mathbf{A}}$ is positive semidefinite, since

$$\mathbf{v}^{\mathsf{H}}\widehat{\mathbf{A}}\mathbf{v} = \mathbf{v}^{\mathsf{H}}\mathbf{P_I A v} = \mathbf{v}^{\mathsf{H}}\mathbf{P_I}^2\mathbf{A v} = \mathbf{v}^{\mathsf{H}}\mathbf{P_I}(\mathbf{P_I A})\mathbf{v} = \mathbf{v}^{\mathsf{H}}\mathbf{P_I}(\mathbf{P_I A})^{\mathsf{H}}\mathbf{v} = \mathbf{v}^{\mathsf{H}}\mathbf{P_I A P_I}^{\mathsf{H}}\mathbf{v} \geq 0$$

holds for any $\mathbf{v} \in \mathbb{C}^N$. The system (4.1) is consistent since it results from a left-multiplication of the nonsingular system $\mathbf{Ax} = \mathbf{b}$ by $\mathbf{P_I}$. (We note that in [41, 12] the application of the projection $\mathbf{P_I}$ to $\mathbf{b}$ is carried out implicitly by adapting the initial guess such that the initial residual is orthogonal to $\mathcal{U} = \mathsf{im}\,(\mathbf{U})$.) The solution $\mathbf{x}$ of $\mathbf{Ax} = \mathbf{b}$ thus also solves $\widehat{\mathbf{A}}\mathbf{x} = \widehat{\mathbf{b}}$, but in (4.1) we replaced $\mathbf{x}$ by $\widehat{\mathbf{x}}$ to indicate the nonuniqueness of the solution. In fact, the general solution is $\widehat{\mathbf{x}} = \mathbf{x} + \mathbf{h}$ with $\mathbf{h} \in \mathcal{U}$

since $\mathbf{P_I A h} = \mathbf{A Q_I h} = \mathbf{0}$ if and only if $\mathbf{Q_I h} = \mathbf{0}$, that is, $\mathbf{h} \in \mathcal{U}$; see statement 2 of Lemma 3.1. The application of $\mathbf{Q_I}$ in the final correction (3.11) will annihilate $\mathbf{h}$. Note that a deflation version including the final correction (3.11) is used by Frank and Vuik [22] and Nabben and Vuik [38, 39].

In the context of Hpd matrices the application of the CG method to a deflated system like (4.1) is a commonly used technique; see, e.g., [54] for a survey of results. Finally, we point out that $\widehat{\mathbf{A}}$ as defined in (4.1) is completely determined by $\mathbf{A}$ and the choice of the space $\mathcal{U}$.

According to Nicolaides [41, section 3] and Kaasschieter [30, section 2], the CG method is well defined (in exact arithmetic) for each step n until it terminates with an exact solution when it is applied to a consistent linear algebraic system with a real and symmetric positive semidefinite matrix. This result easily generalizes to complex and Hermitian positive semidefinite matrices.

Mathematically, the nth step of the CG method applied to the deflated system (4.1) with the initial guess $\mathbf{x}_0$ and the corresponding initial residual $\widehat{\mathbf{r}}_0 = \widehat{\mathbf{b}} - \widehat{\mathbf{A}}\mathbf{x_0}$ is characterized by the two conditions

$$\begin{aligned} &\widehat{\mathbf{x}}_n \in \mathbf{x}_0 + \mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{r}}_0), \\ &\widehat{\mathbf{r}}_n = \widehat{\mathbf{b}} - \widehat{\mathbf{A}}\widehat{\mathbf{x}}_n = \mathbf{P_I}(\mathbf{b} - \mathbf{A}\widehat{\mathbf{x}}_n) \perp \mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{r}}_0). \end{aligned}$$

This is nothing but the set of conditions (3.10) in Theorem 3.2 with $\mathbf{B} = \mathbf{I}$. In the sense of relation (3.11) these conditions have been shown to be equivalent to (3.9), namely,

$$\begin{aligned} &\mathbf{x}_n \in \mathbf{x}_0 + \mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{r}}_0) + \mathcal{U}, \\ &\mathbf{r}_n = \mathbf{b} - \mathbf{A}\mathbf{x}_n \perp \mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{r}}_0) + \mathcal{U}, \end{aligned}$$

which is the starting point of the theory for the deflated CG method developed in [48], where the authors also showed the equivalence between CG with explicit augmentation and CG applied to the deflated system (4.1); see section 4 in [48], in particular Theorem 4.6. In a partly similar treatment, Erhel and Guyomarc'h [18] considered an augmented and deflated CG method where the augmentation space $\mathcal{U}$ is itself a Krylov space. It is worth mentioning that both Saad et al. [48, equation (3.12)] and Erhel and Guyomarc'h [18, equation (3.2)] use the initial correction

$$\mathbf{x}_0 := \mathbf{x}_{-1} + \mathbf{M_I r}_{-1} \qquad \text{with} \quad \mathbf{r}_{-1} := \mathbf{b} - \mathbf{A}\mathbf{x}_{-1}$$

to replace a given initial approximation $\mathbf{x}_{-1}$ by one with $\mathbf{r}_0 \perp \mathcal{U}$; in fact, it is easily seen that $\mathbf{r}_0 = \mathbf{P_I r}_{-1}$.

The goal of deflation is to obtain a deflated matrix $\widehat{\mathbf{A}}$ whose effective condition number is smaller than the one of $\mathbf{A}$, for example, by "eliminating" the smallest eigenvalues of $\mathbf{A}$. A detailed analysis of spectral properties of $\mathbf{P_I A}$ and other projection-type preconditioners arising from domain decomposition and multigrid methods was carried out in [39] and [54]. In particular, it was shown in these papers that the effective condition number of $\widehat{\mathbf{A}}$ is less than or equal to the condition number of $\mathbf{A}$ for any augmentation space $\mathcal{U}$. Moreover, if $\Lambda = \Lambda(\mathbf{A})$ is the spectrum of $\mathbf{A}$ and $\mathcal{U}$ is an $\mathbf{A}$-invariant subspace associated with the eigenvalues $\Theta = \{\theta_1, \ldots, \theta_k\} \subset \Lambda$, then the effective 2-norm condition number is

$$\kappa_2(\widehat{\mathbf{A}}) = \frac{\max_{\lambda \in \Lambda \setminus \Theta} \lambda}{\min_{\lambda \in \Lambda \setminus \Theta} \lambda}.$$

In summary, for any augmentation space $\mathcal{U}$, the CG method applied to the (singular) deflated system (4.1) is well defined for any iteration step n, and it terminates with an exact solution $\widehat{\mathbf{x}}$ (in exact arithmetic). Once CG has terminated with a solution $\widehat{\mathbf{x}}$ of the deflated system, we can obtain the uniquely defined solution of the original system using the final correction step

$$\mathbf{x} = \mathbf{Q_I}\widehat{\mathbf{x}} + \mathbf{M_I}\mathbf{b}$$

(cf. (3.11)), which indeed gives

$$\mathbf{A}\mathbf{x} = \mathbf{A}\mathbf{Q_I}\widehat{\mathbf{x}} + \mathbf{A}\mathbf{M_I}\mathbf{b} = \mathbf{P_I}\mathbf{A}\widehat{\mathbf{x}} + \mathbf{A}\mathbf{M_I}\mathbf{b} = (\mathbf{P_I} + \mathbf{A}\mathbf{M_I})\mathbf{b} = \mathbf{b}.$$

This computation is mathematically equivalent to an explicit use of augmentation. Of course, in practice we stop the CG iteration for the deflated system once the solution is approximated sufficiently accurately. We then use the computed approximation $\widehat{\mathbf{x}}_n$ and (3.11) from Theorem 3.2 to obtain an approximation $\mathbf{x}_n$ of the solution of the given system $\mathbf{A}\mathbf{x} = \mathbf{b}$. Note that, according to (3.11), the residual $\widehat{\mathbf{r}}_n = \widehat{\mathbf{b}} - \widehat{\mathbf{A}}\widehat{\mathbf{x}}_n$ of the projected system (4.1) is equal to the residual $\mathbf{r}_n = \mathbf{b} - \mathbf{A}\mathbf{x}_n$ of the original system (1.1).

5. Non-Hermitian matrices and GMRES. In this section we present mostly known results on applying versions of deflated GMRES to a general nonsingular matrix $\mathbf{A}$. We set $\mathbf{B} = \mathbf{A}$ in the framework of section 2 and discuss some choices for $\widehat{\mathbf{A}}$, $\widehat{\mathbf{v}}$, and $\mathcal{U}$.

Morgan [36] and also Chapman and Saad [6] presented variations of GMRES that can be mathematically described by (3.9) with $\widehat{\mathbf{A}} = \mathbf{A}$ and $\widehat{\mathbf{v}} = \mathbf{b} - \mathbf{A}\mathbf{x}_0$. Hence they augmented the search space with an augmentation space $\mathcal{U}$ but neither deflated the matrix nor projected the linear system onto a subspace of $\mathbb{C}^N$.

Erlangga and Nabben [19] used two matrices $\mathbf{Y}, \mathbf{Z} \in \mathbb{C}^{N\times k}$ to define the abstract deflation operator $\mathbf{P_{YZ}} := \mathbf{I} - \mathbf{A}\mathbf{Z}(\mathbf{Y}^{\mathsf{H}}\mathbf{A}\mathbf{Z})^{-1}\mathbf{Y}^{\mathsf{H}}$ for non-Hermitian matrices $\mathbf{A}$. Of course, this choice needs the assumption of nonsingularity of $\mathbf{Y}^{\mathsf{H}}\mathbf{A}\mathbf{Z}$. Requiring $\mathbf{Y}$ and $\mathbf{Z}$ to have full rank obviously is not sufficient. They then applied GMRES to the deflated linear system $\mathbf{P_{YZ}}\mathbf{A}\widehat{\mathbf{x}} = \mathbf{P_{YZ}}\mathbf{b}$.

De Sturler [10] introduced the GCRO method, which is a nested Krylov subspace method involving an outer and an inner iteration. The outer method is the GCR method [15, 16], while the inner iteration uses the projection

$$\mathbf{P_A} = \mathbf{I} - \mathbf{A}\mathbf{M_A}\mathbf{A}^{\mathsf{H}} = \mathbf{I} - \mathbf{A}\mathbf{U}(\mathbf{U}^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{A}\mathbf{U})^{-1}\mathbf{U}^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}$$

to apply several steps of GMRES to the projected (or deflated) linear system

$$\widehat{\mathbf{A}}\widehat{\mathbf{x}} = \widehat{\mathbf{b}}, \quad \text{where} \quad \widehat{\mathbf{A}} := \mathbf{P_A}\mathbf{A}, \quad \widehat{\mathbf{b}} := \mathbf{P_A}\mathbf{b}. \tag{5.1}$$

In GCRO the matrix $\mathbf{U}$ is determined from the corrections of the outer iteration. Clearly, the matrix $\mathbf{E_A} = \mathbf{U}^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{A}\mathbf{U}$ is nonsingular for any matrix $\mathbf{U} \in \mathbb{C}^{N\times k}$ with $\operatorname{rank}\mathbf{U} = k > 0$, so that all matrices in (3.6) are well defined. Note that the projection $\mathbf{P_A}$ is equal to the abstract deflation operator $\mathbf{P_{YZ}}$ of Erlangga and Nabben with the choice $\mathbf{Z} = \mathbf{U}$ and $\mathbf{Y} = \mathbf{A}\mathbf{U}$. For the application of $\mathbf{P_A}$ only the matrix $\mathbf{W} := \mathbf{A}\mathbf{U}$ is needed because $\mathbf{P_A} = \mathbf{I} - \mathbf{W}(\mathbf{W}^{\mathsf{H}}\mathbf{W})^{-1}\mathbf{W}^{\mathsf{H}}$. De Sturler further simplified this in [10, section 2] to $\mathbf{P_A} = \mathbf{I} - \mathbf{C}\mathbf{C}^{\mathsf{H}}$ by choosing a matrix $\mathbf{C} \in \mathbb{C}^{N\times k}$ whose columns form an orthonormal basis of $\mathsf{im}\,(\mathbf{A}\mathbf{U})$.

Here we concentrate on the GMRES method applied to the deflated system (5.1), and we first discuss some known results within the framework presented in section 2. Analogously to the approach for CG described in the previous section, the deflated system (5.1) results from the given system $\mathbf{A}\mathbf{x} = \mathbf{b}$ by a left-multiplication with $\mathbf{P_A}$ which projects onto $(\mathbf{A}\mathcal{U})^{\perp}$ along $\mathbf{A}\mathcal{U}$. Note that $\mathbf{P_A}$ is an orthogonal projection, since $\mathbf{P_A}$ is Hermitian.

If we start GMRES with an initial guess $\mathbf{x}_0$ and the corresponding initial residual $\widehat{\mathbf{r}}_0 = \widehat{\mathbf{b}} - \widehat{\mathbf{A}}\mathbf{x}_0 = \mathbf{P_A}(\mathbf{b} - \mathbf{A}\mathbf{x}_0)$, then the iterate $\widehat{\mathbf{x}}_n$ and the residual $\widehat{\mathbf{r}}_n$ are characterized by the two conditions

$$\widehat{\mathbf{x}}_n \in \mathbf{x}_0 + \mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{r}}_0) \quad \text{and} \quad \widehat{\mathbf{r}}_n = \widehat{\mathbf{b}} - \widehat{\mathbf{A}}\widehat{\mathbf{x}}_n \perp \widehat{\mathbf{A}}\mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{r}}_0).$$

If the columns of $\mathbf{V}_n$ form a basis of $\mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{r}}_0)$, then the second condition means that

$$\begin{aligned}\mathbf{0} &= \mathbf{V}_n^{\mathsf{H}}\widehat{\mathbf{A}}^{\mathsf{H}}\widehat{\mathbf{r}}_n = \mathbf{V}_n^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{P}_{\mathbf{A}}^{\mathsf{H}}\widehat{\mathbf{r}}_n = \mathbf{V}_n^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{P_A}\mathbf{P_A}\,(\mathbf{b} - \mathbf{A}\widehat{\mathbf{x}}_n) = \mathbf{V}_n^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{P_A}\,(\mathbf{b} - \mathbf{A}\widehat{\mathbf{x}}_n) \\ &= \mathbf{V}_n^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\widehat{\mathbf{r}}_n,\end{aligned}$$

or, equivalently,

$$\widehat{\mathbf{r}}_n \perp \mathbf{A}\mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{r}}_0).$$

Note that here the Krylov subspace is multiplied with $\mathbf{A}$ instead of $\widehat{\mathbf{A}}$ and that this condition has precisely the form of the second condition in (3.10). Theorem 3.2 now implies that the mathematical characterization of GMRES applied to the deflated system $\widehat{\mathbf{A}}\widehat{\mathbf{x}} = \widehat{\mathbf{b}}$ is equivalent to the explicit use of augmentation, i.e., the conditions

$$\mathbf{x}_n \in \mathbf{x}_0 + \mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{r}}_0) + \mathcal{U}, \tag{5.2}$$

$$\mathbf{r}_n = \mathbf{b} - \mathbf{A}\mathbf{x}_n \perp \mathbf{A}\mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{r}}_0) + \mathbf{A}\mathcal{U}, \tag{5.3}$$

in the sense that

$$\mathbf{x}_n = \mathbf{Q_A}\widehat{\mathbf{x}}_n + \mathbf{M_A}\mathbf{A}^{\mathsf{H}}\mathbf{b} \quad \text{and} \quad \mathbf{r}_n = \mathbf{b} - \mathbf{A}\mathbf{x}_n = \widehat{\mathbf{b}} - \widehat{\mathbf{A}}\widehat{\mathbf{x}}_n = \widehat{\mathbf{r}}_n. \tag{5.4}$$

As mentioned in the beginning of section 2, conditions (5.2)–(5.3) are equivalent to the minimization problem

$$\text{find} \quad \mathbf{x}_n \in \mathbf{x}_0 + \mathcal{S}_n \quad \text{s.t.} \quad \|\mathbf{b} - \mathbf{A}\mathbf{x}_n\|_2 = \min_{\mathbf{y} \in \mathbf{x}_0 + \mathcal{S}_n} \|\mathbf{b} - \mathbf{A}\mathbf{y}\|_2$$

with the search space $\mathcal{S}_n = \mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{r}}_0) + \mathcal{U}$. In the setting of GCRO, where $\mathbf{U}$ is determined from the GCR iteration, the equivalence between GMRES applied to $\widehat{\mathbf{A}}\widehat{\mathbf{x}} = \widehat{\mathbf{b}}$ and the minimization problem with an explicitly augmented search space has already been pointed out by de Sturler [10, Theorem 2.2]. The GCRO method was extended to an arbitrary rank-k matrix $\mathbf{U}$ in [32, section 2]. In the case where $\mathcal{U}$ is an $\mathbf{A}$-invariant subspace the equivalence is straightforward and has been pointed out by Eiermann, Ernst, and Schneider [14, Lemma 4.3].

Again the deflated matrix $\widehat{\mathbf{A}}$ is singular, and we have to discuss whether the application of GMRES to the deflated system yields (in exact arithmetic) a well-defined sequence of iterates that terminates with a solution. This turns out to be significantly more difficult than in the case of the CG method. Properties of GMRES applied to singular systems have been analyzed by de Sturler [10] and by Brown and Walker [5]. The following result is an extension of [5, Theorem 2.6].

THEOREM 5.1. *Consider an arbitrary matrix* $\widehat{\mathbf{A}} \in \mathbb{C}^{N\times N}$ *and a vector* $\widehat{\mathbf{b}} \in \mathsf{im}\,(\widehat{\mathbf{A}})$ (*i.e., the linear algebraic system* $\widehat{\mathbf{A}}\widehat{\mathbf{x}} = \widehat{\mathbf{b}}$ *is consistent*). *Then the following two conditions are equivalent:*

1. *For every initial guess* $\mathbf{x}_0 \in \mathbb{C}^N$ *the* GMRES *method applied to the system* $\widehat{\mathbf{A}}\widehat{\mathbf{x}} = \widehat{\mathbf{b}}$ *is well defined at each iteration step* n *and it terminates with a solution of the system.*
2. $\mathsf{ker}\,(\widehat{\mathbf{A}}) \cap \mathsf{im}\,(\widehat{\mathbf{A}}) = \{\mathbf{0}\}$.

Proof. It has been shown in [5, Theorem 2.6] that condition 2 implies condition 1. We prove the reverse by contradiction. We assume that $\mathsf{ker}\,(\widehat{\mathbf{A}})\cap\mathsf{im}\,(\widehat{\mathbf{A}}) \neq \{\mathbf{0}\}$, and we will construct an initial guess for which GMRES does not terminate with the solution. For a nonzero vector $\mathbf{y} \in \mathsf{ker}\,(\widehat{\mathbf{A}}) \cap \mathsf{im}\,(\widehat{\mathbf{A}})$ there exists a nonzero vector $\widehat{\mathbf{y}} \in \mathbb{C}^N$, such that $\mathbf{y} = \widehat{\mathbf{A}}\widehat{\mathbf{y}}$, and since $\widehat{\mathbf{A}}\widehat{\mathbf{x}} = \widehat{\mathbf{b}}$ is consistent, there exists a vector $\widehat{\mathbf{x}} \in \mathbb{C}^N$ with $\widehat{\mathbf{b}} = \widehat{\mathbf{A}}\widehat{\mathbf{x}}$. Then the initial guess $\mathbf{x}_0 := \widehat{\mathbf{x}} - \widehat{\mathbf{y}}$ gives $\mathbf{r}_0 = \widehat{\mathbf{b}} - \widehat{\mathbf{A}}\mathbf{x}_0 = \widehat{\mathbf{b}} - \widehat{\mathbf{A}}\widehat{\mathbf{x}} + \widehat{\mathbf{A}}\widehat{\mathbf{y}} = \mathbf{y}$. But since $\mathbf{y} \in \mathsf{ker}\,(\widehat{\mathbf{A}})$, we obtain $\widehat{\mathbf{A}}\mathbf{r}_0 = \mathbf{0}$, so that the GMRES method terminates at the first iteration with the approximation $\mathbf{x}_0$, for which $\mathbf{r}_0 = \mathbf{y} \neq \mathbf{0}$. Thus, for this particular initial guess $\mathbf{x}_0$ the GMRES method cannot determine the solution of $\widehat{\mathbf{A}}\widehat{\mathbf{x}} = \widehat{\mathbf{b}}$. ∎

The situation that the GMRES method terminates without finding the exact solution is often called a *breakdown* of GMRES. The above proof leads to the following characterization of all initial guesses that lead to a breakdown of GMRES at the first iteration.

COROLLARY 5.2. *Let* $\widehat{\mathbf{A}} \in \mathbb{C}^{N\times N}$ *and* $\widehat{\mathbf{x}}, \widehat{\mathbf{b}} \in \mathbb{C}^{N\times N}$ *such that* $\widehat{\mathbf{A}}\widehat{\mathbf{x}} = \widehat{\mathbf{b}}$. *Then the* GMRES *method breaks down at the first iteration for all initial guesses*

$$\mathbf{x}_0 \in \mathscr{X}_0 := \left\{\widehat{\mathbf{x}} - \widehat{\mathbf{y}} \;\middle|\; \widehat{\mathbf{A}}\widehat{\mathbf{y}} \in \mathsf{ker}\,(\widehat{\mathbf{A}}) \setminus \{\mathbf{0}\}\right\}.$$

We next have a closer look at condition 2 in Theorem 5.1. If we had $\mathsf{ker}\,(\widehat{\mathbf{A}}) = \mathsf{ker}\,(\widehat{\mathbf{A}}^{\mathsf{H}})$, then $\mathsf{im}\,(\widehat{\mathbf{A}})^{\perp} = \mathsf{ker}\,(\widehat{\mathbf{A}}^{\mathsf{H}})$ would imply

$$\{\mathbf{0}\} = \mathsf{im}\,(\widehat{\mathbf{A}})^{\perp} \cap \mathsf{im}\,(\widehat{\mathbf{A}}) = \mathsf{ker}\,(\widehat{\mathbf{A}}^{\mathsf{H}}) \cap \mathsf{im}\,(\widehat{\mathbf{A}}) = \mathsf{ker}\,(\widehat{\mathbf{A}}) \cap \mathsf{im}\,(\widehat{\mathbf{A}}),$$

so that condition 2 would hold. Thus condition 2 in Theorem 5.1 is fulfilled for any Hermitian matrix $\widehat{\mathbf{A}}$. For a general non-Hermitian matrix, however, it seems difficult to determine a deflated matrix with $\mathsf{ker}\,(\widehat{\mathbf{A}}) = \mathsf{ker}\,(\widehat{\mathbf{A}}^{\mathsf{H}})$. However, for the deflated system (5.1) we can derive another condition that is equivalent with condition 2 (and hence condition 1) in Theorem 5.1.

COROLLARY 5.3. *For the deflated system* (5.1), *condition* 2 *in Theorem* 5.1 *is satisfied if and only if* $\mathcal{U}\cap(\mathbf{A}\mathcal{U})^{\perp} = \{\mathbf{0}\}$. *In particular, the latter condition is satisfied when* $\mathcal{U}$ *is an exactly* $\mathbf{A}$*-invariant subspace, i.e., when* $\mathbf{A}\mathcal{U} = \mathcal{U}$.

Proof. Using the properties of the projection $\mathbf{P}_{\mathbf{A}}$ from Lemma 3.1 and the fact that $\mathbf{A}$ is nonsingular, we obtain

$$\mathsf{ker}\,(\widehat{\mathbf{A}}) = \mathsf{ker}\,(\mathbf{P}_{\mathbf{A}}\mathbf{A}) = \mathbf{A}^{-1}\mathsf{ker}\,(\mathbf{P}_{\mathbf{A}}) = \mathcal{U},$$
$$\mathsf{im}\,(\widehat{\mathbf{A}}) = \mathsf{im}\,(\mathbf{P}_{\mathbf{A}}\mathbf{A}) = \mathsf{im}\,(\mathbf{P}_{\mathbf{A}}) = (\mathbf{A}\mathcal{U})^{\perp}.$$

If $\mathbf{A}\mathcal{U} = \mathcal{U}$, then $\mathcal{U} \cap (\mathbf{A}\mathcal{U})^{\perp} = \{\mathbf{0}\}$ holds trivially. ∎

For a nonsingular matrix, condition 2 in Theorem 5.1 always holds trivially, and hence a breakdown of GMRES can only occur if the method is applied to a linear algebraic system with a singular matrix. (This fact has been known since the method's introduction in 1986 [47].) Breakdowns have also been analyzed by de Sturler [10]

in the context of the GCRO method. (See the end of section 6.1 below for further comments.) We want to point out that it is unlikely that a random initial guess lies in the subspace $\mathcal{X}_0$ specified in Corollary 5.2 for which the GMRES method breaks down in the first step. However, a general $\mathcal{U}$ may lead to a breakdown. To illustrate the problem of breakdowns in our context, we give an example that is adapted from [5, Example 1.1].

Example 5.4. Consider a linear algebraic system with

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

so that the unique solution is given by the vector $[0,1]^{\mathsf{T}}$. Let the augmentation space be defined by $\mathbf{U}_1 = [1,0]^{\mathsf{T}}$, and then

$$\mathbf{P_A} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \widehat{\mathbf{A}} = \mathbf{P_A}\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \widehat{\mathbf{b}} = \mathbf{P_A}\mathbf{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

If $\mathbf{x}_0$ is the zero vector, then $\widehat{\mathbf{r}}_0 = \widehat{\mathbf{b}}$ and $\widehat{\mathbf{A}}\widehat{\mathbf{r}}_0 = \mathbf{0}$, and thus GMRES applied to the deflated system terminates at the very first iteration with the approximation $\mathbf{x}_0$. Since $\widehat{\mathbf{A}}\mathbf{x}_0 \neq \widehat{\mathbf{b}}$, this is a breakdown of GMRES. Furthermore, applying the correction (3.7) to $\widehat{\mathbf{x}}_0 = \mathbf{x}_0$ does not yield the solution of the original system $\mathbf{A}\mathbf{x} = \mathbf{b}$ because

$$\mathbf{Q_A}\mathbf{x}_0 + \mathbf{M_A}\mathbf{A}^{\mathsf{H}}\mathbf{b} = \mathbf{M_A}\mathbf{A}^{\mathsf{H}}\mathbf{b} = \mathbf{U}_1\mathbf{U}_1^{\mathsf{H}}\mathbf{A}^{\mathsf{H}}\mathbf{b} = \mathbf{0} \neq \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Corollary 5.3 states that the GMRES method applied to the deflated system (5.1) cannot break down if $\mathcal{U}$ is an $\mathbf{A}$-invariant subspace. The following example shows that care has also to be taken with approximate $\mathbf{A}$-invariant subspaces.

Example 5.5. Let $\alpha > 0$ be a small positive number. Then $\mathbf{v} := [0,1,\alpha]^{\mathsf{T}}$ is an eigenvector of the matrix

$$\mathbf{A} := \begin{bmatrix} 0 & 1 & -\alpha^{-1} \\ 1 & 0 & \alpha^{-1} \\ 0 & 0 & 1 \end{bmatrix}$$

corresponding to the eigenvalue 1. Instead of $\mathbf{v}$ we use the perturbed vector $\mathbf{U}_2 := [0,1,0]^{\mathsf{T}}$ as a basis for the deflation space $\mathcal{U} = \mathsf{im}\,(\mathbf{U}_2)$ and obtain

$$\mathbf{A}\mathbf{U}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{P_A} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{P_A}\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & \alpha^{-1} \\ 0 & 0 & 1 \end{bmatrix}.$$

For $\mathbf{x}, \mathbf{b} \in \mathbb{C}^3$ with $\mathbf{A}\mathbf{x} = \mathbf{b}$ the GMRES method then breaks down in the first step for all $\mathbf{x}_0 \in \{\mathbf{x} + \beta[1,0,0]^{\mathsf{T}} \mid \beta \neq 0\}$. Note that $\|\mathbf{U}_2 - \mathbf{v}\|_2 = \alpha$ can be chosen arbitrarily small. A better measure for the quality of an approximate invariant subspace would be the largest principal angle between $\mathcal{U}$ and $\mathbf{A}\mathcal{U}$.

6. Hermitian matrices and variants of MINRES. We will now apply the results presented in sections 2 and 5 to the case where $\mathbf{A}$ is Hermitian, nonsingular, and possibly indefinite. For a Hermitian matrix the GMRES method considered in section 5 is mathematically equivalent to the MINRES method, which is based on the Hermitian Lanczos algorithm and thus uses efficient three-term recurrences.

6.1. The RMinRes method. This subsection discusses the RMinRes method developed by Wang, de Sturler, and Paulino [57]. This method fits into the framework of section 2, and the results presented in section 5 apply. Wang, de Sturler, and Paulino were interested in solving sequences of linear algebraic systems that exhibit only small changes from one matrix in the sequence to the next one, and they suggested to reuse information from previous solves. The RMinRes method consists of two main parts that can basically be analyzed separately: an augmented and deflated MinRes solver which is based on GCRO and an extraction procedure for the augmentation and deflation data. In the second part Wang, de Sturler, and Paulino determined harmonic Ritz vectors that correspond to harmonic Ritz values close to zero and used these approximate eigenspaces for augmenting the Krylov subspace. Here, we omit the extraction of the augmentation and deflation space and concentrate on the method for solving the systems. We refer to this as the solver part of the RMinRes method. We point out that the extracted spaces can be arbitrary if there are no restrictions on the changes of the matrices in the sequence of linear algebraic systems. However, the RMinRes method has been presented in [57] with an application in topology optimization where the extracted approximated eigenvectors of one matrix are still good approximations to eigenvectors of the next matrix. Furthermore, we will not address the preconditioning technique outlined in [57] and assume that the given linear algebraic system is already in the preconditioned form.

As in section 5, we set $\mathbf{B} = \mathbf{A}$ and consider first the resulting deflated system of the form (5.1),

$$\widehat{\mathbf{A}}\widehat{\mathbf{x}} = \widehat{\mathbf{b}}, \quad \text{where} \quad \widehat{\mathbf{A}} := \mathbf{P_A}\mathbf{A}, \quad \widehat{\mathbf{b}} := \mathbf{P_A}\mathbf{b}.$$

If we apply MinRes to this linear algebraic system with an initial guess $\mathbf{x}_0$ and the corresponding initial residual $\widehat{\mathbf{r}}_0 = \widehat{\mathbf{b}} - \widehat{\mathbf{A}}\mathbf{x}_0 = \mathbf{P_A}(\mathbf{b} - \mathbf{A}\mathbf{x}_0)$, then the iterate $\widehat{\mathbf{x}}_n$ and the residual $\widehat{\mathbf{r}}_n$ are characterized by the two conditions

$$\widehat{\mathbf{x}}_n \in \mathbf{x}_0 + \mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{r}}_0) \quad \text{and} \quad \widehat{\mathbf{r}}_n = \widehat{\mathbf{b}} - \widehat{\mathbf{A}}\widehat{\mathbf{x}}_n \perp \widehat{\mathbf{A}}\mathcal{K}_n(\widehat{\mathbf{A}}, \widehat{\mathbf{r}}_0). \tag{6.1}$$

This is essentially the approach of Kilmer and de Sturler [32, section 2]. Olshanskii and Simoncini [42] recently used a different approach where the MinRes method is applied to the deflated system $\mathbf{P_I}\mathbf{A}\widehat{\mathbf{x}} = \mathbf{P_I}\mathbf{b}$ with the special initial guess $\mathbf{x}_0 = \mathbf{U}(\mathbf{U}^{\mathsf{H}}\mathbf{A}\mathbf{U})^{-1}\mathbf{U}^{\mathsf{H}}\mathbf{b}$. We note that the presentation in [42] is slightly different but the above can be seen with minor algebraic modifications to the relations in and preceding Proposition 3.1 in [42].

An attentive reader has certainly noticed that the deflated matrix $\widehat{\mathbf{A}} = \mathbf{P_A}\mathbf{A} = \mathbf{A} - \mathbf{A}\mathbf{M_A}\mathbf{A}^2$ is in general *not* Hermitian, even when $\mathbf{A}$ is Hermitian. However, as pointed out in [32, footnote on p. 2153] and [57, footnote on p. 2446], a straightforward computation shows that

$$\mathcal{K}_n(\mathbf{P_A}\mathbf{A}, \mathbf{P_A}\mathbf{v}) = \mathcal{K}_n(\mathbf{P_A}\mathbf{A}\mathbf{P_A}, \mathbf{P_A}\mathbf{v}) \tag{6.2}$$

holds for every vector $\mathbf{v} \in \mathbb{C}^N$ because $\mathbf{P_A}$ is a projection. The matrix $\mathbf{P_A}\mathbf{A}\mathbf{P_A}$ is obviously Hermitian (since $\mathbf{A}$ and $\mathbf{P_A}$ are Hermitian), and hence the Krylov subspaces we work with are also generated by a Hermitian matrix. It is therefore possible to implement a MinRes-like method for the deflated system, which is based on three-term recurrences and which is characterized by the conditions (6.1). As presented in section 5, these conditions combined with the correction step (5.4) are equivalent to the explicit use of augmentation, i.e., conditions (5.2)–(5.3).

The latter conditions are the basis of the solver part of the RMinRes method by Wang, de Sturler, and Paulino in [57, section 3]. We summarize the above and give two mathematically equivalent characterizations of the RMinRes solver applied to the system $\mathbf{A}\mathbf{x} = \mathbf{b}$ with an initial guess $\mathbf{x}_0$:

1. The original approach used in [57] incorporates explicit augmentation, which means to construct iterates $\mathbf{x}_n$ satisfying the two conditions

$$\begin{aligned} &\mathbf{x}_n \in \mathbf{x}_0 + \mathcal{K}_n(\mathbf{P_A}\mathbf{A}, \mathbf{P_A}\mathbf{r}_0) + \mathcal{U}, \\ &\mathbf{r}_n = \mathbf{b} - \mathbf{A}\mathbf{x}_n \perp \mathbf{A}\mathcal{K}_n(\mathbf{P_A}\mathbf{A}, \mathbf{P_A}\mathbf{r}_0) + \mathbf{A}\mathcal{U}. \end{aligned} \tag{6.3}$$

2. A mathematically equivalent approach is to apply MinRes to the deflated system

$$\mathbf{P_A}\mathbf{A}\widehat{\mathbf{x}} = \mathbf{P_A}\mathbf{b} \tag{6.4}$$

and correct the resulting iterates $\widehat{\mathbf{x}}_n$ according to $\mathbf{x}_n = \mathbf{Q_A}\widehat{\mathbf{x}}_n + \mathbf{M_A}\mathbf{A}\mathbf{b}$.

Note that on an algorithmic level the second approach exhibits lower computational cost since the correction in the space $\mathcal{U}$ is only carried out once at the end, while the RMinRes solver requires one update per iteration.

Since the solver part of RMinRes is mathematically equivalent to MinRes (and GMRes) applied to the deflated system, Corollary 5.3 also applies to RMinRes. In particular, the method can break down for specific initial guesses if (and only if) $\mathcal{U} \cap (\mathbf{A}\mathcal{U})^{\perp} \neq \{\mathbf{0}\}$. Breakdowns cannot occur if $\mathcal{U}$ is an exact $\mathbf{A}$-invariant subspace, but this is an unrealistic assumption in practical applications. Note that the matrix $\mathbf{A}$ in Example 5.4 is Hermitian, thus it also serves as an example for a breakdown of the RMinRes solver. That the RMinRes method can break down may already be guessed from the fact that this method is based on the GCRO method and thus potentially suffers from the breakdown conditions for GCRO derived in [10]. However, the possibility of breakdowns has not been mentioned in [57], and in the example of a GCRO breakdown given in [10] the matrix $\mathbf{A}$ is not Hermitian. Hence this example cannot be used in the context of the RMinRes method, which is intended for Hermitian matrices.

In the next subsection we show how to suitably modify the RMinRes approach to avoid breakdowns.

6.2. Avoiding breakdowns in deflated MinRes. We have seen in section 5 that if $\mathsf{ker}\,(\widehat{\mathbf{A}}) = \mathsf{ker}\,(\widehat{\mathbf{A}}^{\mathsf{H}})$, then condition 1 in Theorem 5.1 is satisfied. Consequently, if we can determine a Hermitian deflated matrix $\widehat{\mathbf{A}}$ and a corresponding consistent deflated system, MinRes applied to this system cannot break down for any initial guess.

Using the projections $\mathbf{P_A}$ and $\mathbf{Q_A}$ from (3.6) we decompose the solution $\mathbf{x}$ of $\mathbf{A}\mathbf{x} = \mathbf{b}$ as

$$\mathbf{x} = \mathbf{P_A}\mathbf{x} + (\mathbf{I} - \mathbf{P_A})\mathbf{x} = \mathbf{P_A}\mathbf{x} + \mathbf{A}\mathbf{M_A}\mathbf{A}\mathbf{x} = \mathbf{P_A}\mathbf{x} + \mathbf{A}\mathbf{M_A}\mathbf{b}, \tag{6.5}$$

$$\mathbf{x} = \mathbf{Q_A}\mathbf{x} + (\mathbf{I} - \mathbf{Q_A})\mathbf{x} = \mathbf{Q_A}\mathbf{x} + \mathbf{M_A}\mathbf{A}^2\mathbf{x} = \mathbf{Q_A}\mathbf{x} + \mathbf{M_A}\mathbf{A}\mathbf{b}. \tag{6.6}$$

Using (6.6), the system $\mathbf{A}\mathbf{x} = \mathbf{b}$ becomes $\mathbf{A}(\mathbf{Q_A}\mathbf{x} + \mathbf{M_A}\mathbf{A}\mathbf{b}) = \mathbf{b}$. With the definition of $\mathbf{P_A}$ and $\mathbf{A}\mathbf{Q_A} = \mathbf{P_A}\mathbf{A}$ (cf. Lemma 3.1) we see that this is equivalent to

$$\mathbf{P_A}\mathbf{A}\mathbf{x} = \mathbf{P_A}\mathbf{b}.$$

We now substitute for $\mathbf{x}$ from (6.5) and obtain $\mathbf{P_A A}(\mathbf{P_A x} + \mathbf{A M_A b}) = \mathbf{P_A b}$ which is equivalent to

$$\mathbf{P_A A P_A x} = \mathbf{P_A Q_A^H b}. \tag{6.7}$$

We can show the following result for the MinRes method applied to this symmetric system.

THEOREM 6.1. *For each initial guess* $\mathbf{x}_0 \in \mathbb{C}^N$ *the* MinRes *method applied to the system* (6.7) *yields* (*in exact arithmetic*) *a well-defined iterate* $\overline{\mathbf{x}}_n$ *at every step* $n \geq 1$ *until it terminates with a solution. Moreover, the sequence of iterates*

$$\mathbf{x}_n := \mathbf{Q_A}\left(\mathbf{P_A}\overline{\mathbf{x}}_n + \mathbf{A M_A b}\right) + \mathbf{M_A A b} \tag{6.8}$$

is well defined. It terminates (*in exact arithmetic*) *with the exact solution* $\mathbf{x}$ *of the original linear system* $\mathbf{Ax} = \mathbf{b}$*, and its residuals are given by* $\mathbf{r}_n = \mathbf{b} - \mathbf{A}\mathbf{x}_n = \mathbf{P_A Q_A^H b} - \mathbf{P_A A P_A}\overline{\mathbf{x}}_n$.

Proof. The first part follows from the fact that the system (6.7) is a consistent system with a Hermitian matrix $\mathbf{P_A A P_A}$, so that we can apply Theorem 5.1. It remains to show the second part. The nth residual of the original system $\mathbf{Ax} = \mathbf{b}$ is given by

$$\begin{aligned}
\mathbf{r}_n = \mathbf{b} - \mathbf{A}\mathbf{x}_n &= \mathbf{b} - \mathbf{A}\left(\mathbf{Q_A}\left(\mathbf{P_A}\overline{\mathbf{x}}_n + \mathbf{A M_A b}\right) + \mathbf{M_A A b}\right) \\
&= \mathbf{b} - \mathbf{A Q_A}\left(\mathbf{P_A}\overline{\mathbf{x}}_n + \mathbf{A M_A b}\right) - \mathbf{A M_A A b} \\
&= \left(\mathbf{I} - \mathbf{A M_A A}\right)\mathbf{b} - \mathbf{P_A A}\left(\mathbf{P_A}\overline{\mathbf{x}}_n + \mathbf{A M_A b}\right) \\
&= \mathbf{P_A b} - \mathbf{P_A A P_A}\overline{\mathbf{x}}_n - \mathbf{P_A A}^2\mathbf{M_A b} \\
&= \mathbf{P_A}\left(\mathbf{I} - \mathbf{A}^2\mathbf{M_A}\right)\mathbf{b} - \mathbf{P_A A P_A}\overline{\mathbf{x}}_n \\
&= \mathbf{P_A Q_A^H b} - \mathbf{P_A A P_A}\overline{\mathbf{x}}_n.
\end{aligned}$$

We see that $\mathbf{r}_n$ is equal to the nth MinRes residual for the system (6.7). In particular, this implies that the exact solution of (1.1) is given by (6.8) once an exact solution $\overline{\mathbf{x}}_n$ of (6.7) has been determined by MinRes. ∎

When MinRes is applied to the deflated system (6.7), the Hermitian iteration matrix $\mathbf{P_A A P_A}$ can again be replaced by the non-Hermitian matrix $\mathbf{P_A A}$ (cf. section 6.1).

The following theorem shows that a modification of the initial guess suffices to make the solver part of the RMinRes method mathematically equivalent to MinRes applied to the system (6.7).

THEOREM 6.2. *We consider the following two approaches:*

1. *The solver part of the* RMinRes *method applied to* $\mathbf{Ax} = \mathbf{b}$ *with the initial guess* $\widehat{\mathbf{x}}_0 := \mathbf{P_A x}_0 + \mathbf{A M_A b}$ *and resulting iterates* $\mathbf{x}_n$ *and residuals* $\mathbf{r}_n = \mathbf{b} - \mathbf{A}\mathbf{x}_n$.
2. *The* MinRes *method applied to* (6.7) *with the initial guess* $\mathbf{x}_0$ *and resulting iterates* $\overline{\mathbf{x}}_n$ *and residuals* $\overline{\mathbf{r}}_n := \mathbf{P_A Q_A^H b} - \mathbf{P_A A P_A}\overline{\mathbf{x}}_n$.

Both approaches are equivalent in the sense that $\mathbf{x}_n = \mathbf{Q_A}(\mathbf{P_A}\overline{\mathbf{x}}_n + \mathbf{A M_A b}) + \mathbf{M_A A b}$ *and* $\mathbf{r}_n = \overline{\mathbf{r}}_n$.

Proof. Let us start with the MinRes method applied to (6.7), which constructs iterates $\overline{\mathbf{x}}_n = \mathbf{x}_0 + \mathbf{V}_n\mathbf{y}_n$, where $\mathbf{V}_n \in \mathbb{C}^{N\times n}$ is of full rank n such that $\mathsf{im}\,(\mathbf{V}_n) = \mathcal{K}_n(\mathbf{P_A A P_A}, \mathbf{P Q^H r}_0)$. Then $\mathbf{P_A V}_n = \mathbf{V}_n$ and the corrected iterates are

$$\begin{aligned}
\mathbf{x}_n &= \mathbf{Q_A}(\mathbf{P_A}(\mathbf{x}_0 + \mathbf{V}_n\mathbf{y}_n) + \mathbf{A M_A b}) + \mathbf{M_A A b} = \mathbf{Q_A}(\widehat{\mathbf{x}}_0 + \mathbf{V}_n\mathbf{y}_n) + \mathbf{M_A A b} \\
&= \mathbf{Q_A}\widehat{\mathbf{x}}_n + \mathbf{M_A A b}
\end{aligned} \tag{6.9}$$

with $\widehat{\mathbf{x}}_n := \widehat{\mathbf{x}}_0 + \mathbf{V}_n\mathbf{y}_n$. For $n > 0$ the nth residual of $\overline{\mathbf{x}}_n$ with respect to the system (6.7) is

$$\begin{aligned}\overline{\mathbf{r}}_n &= \mathbf{P_A}\mathbf{Q_A}^{\mathsf{H}}\mathbf{b} - \mathbf{P_A}\mathbf{A}\mathbf{P_A}\overline{\mathbf{x}}_n = \mathbf{P_A}(\mathbf{Q_A}^{\mathsf{H}}\mathbf{b} - \mathbf{A}\mathbf{P_A}\mathbf{x}_0 - \mathbf{A}\mathbf{V}_n\mathbf{y}_n)\\ &= \mathbf{P_A}(\mathbf{b} - \mathbf{A}(\mathbf{P_A}\mathbf{x}_0 + \mathbf{A}\mathbf{M_A}\mathbf{b} + \mathbf{V}_n\mathbf{y}_n)) = \mathbf{P_A}\mathbf{b} - \mathbf{P_A}\mathbf{A}\widehat{\mathbf{x}}_n =: \widehat{\mathbf{r}}_n.\end{aligned}$$

This is the residual of $\widehat{\mathbf{x}}_n$ with respect to the system (6.4). We also have

$$\widehat{\mathbf{r}}_0 = \mathbf{P_A}\mathbf{b} - \mathbf{P_A}\mathbf{A}\widehat{\mathbf{x}}_0 = \mathbf{P_A}\mathbf{Q_A}^{\mathsf{H}}\mathbf{b} - \mathbf{P_A}\mathbf{A}\mathbf{P_A}\mathbf{x}_0 = \overline{\mathbf{r}}_0,$$

and thus the starting vectors of the Krylov subspace for both methods are equal. Because of (6.2) the Krylov subspaces are also equal. From the definition of the Krylov subspaces we immediately obtain

$$\overline{\mathbf{r}}_n \perp \mathbf{P_A}\mathbf{A}\mathbf{P_A}\mathcal{K}_n(\mathbf{P_A}\mathbf{A}\mathbf{P_A}, \overline{\mathbf{r}}_0) \quad \Longleftrightarrow \quad \widehat{\mathbf{r}}_n \perp \mathbf{P_A}\mathbf{A}\mathcal{K}_n(\mathbf{P_A}\mathbf{A}, \widehat{\mathbf{r}}_0).$$

We can now see that the iterates $\widehat{\mathbf{x}}_n$ are the iterates of MinRes applied to (6.4) with the initial guess $\widehat{\mathbf{x}}_0$. Along with the correction (6.9) this was shown to be equivalent to the RMinRes solver applied to $\mathbf{A}\mathbf{x} = \mathbf{b}$ with the initial guess $\widehat{\mathbf{x}}_0$ (cf. section 6.1). ∎

This means that (in exact arithmetic) breakdowns in the solver part of the RMinRes method can be prevented by either adapting the right-hand side to $\mathbf{Q}^{\mathsf{H}}\mathbf{b}$ and correcting the approximate solution at the end according to Theorem 6.1 or by choosing the adapted initial guess $\widehat{\mathbf{x}}_0$ defined in Theorem 6.2. Both choices do not increase the computational cost significantly since these computations only need to be carried out once. A similar special initial guess has also been used in [54] to obtain a robust deflation-based preconditioner for the CG method; compare the A-DEF2 method in [54, Table 2].

6.3. Numerical experiments. In this subsection, we will show the numerical behavior of selected Krylov subspace methods discussed above. Detailed numerical experiments with the deflated CG method (cf. section 4) and equivalent approaches have been presented in [54]. Here, we will focus on the solver part of the RMinRes method and the deflated MinRes method in order to numerically illustrate the phenomenon of breakdowns that have only been described theoretically so far (cf. sections 6.1 and 6.2). Both methods are implemented in MATLAB with three-term Lanczos recurrences and Givens rotations for solving the least squares problem. All residuals have been computed explicitly in each iteration.

Example 6.3. In this example we use a matrix $\mathbf{A} = \mathbf{W}^{\mathsf{H}}\mathbf{D}\mathbf{W} \in \mathbb{R}^{2m\times 2m}$, $m = 50$, where $\mathbf{D} = \operatorname{diag}(\lambda_1, \ldots, \lambda_{2m})$ with $\lambda_j = \sqrt{j}$, $\lambda_{m+j} = -\sqrt{j}$ for $j = 1, \ldots, m$ and $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_{2m}]$ is a randomly generated orthogonal matrix. We consider a matrix $\mathbf{U} = [u_1, \ldots, u_k]$ whose columns are pairwise orthogonal eigenvectors of $\mathbf{A}$, i.e., $\mathbf{A}\mathbf{U} = \mathbf{U}\mathbf{D_U}$ and $\mathbf{U}^{\mathsf{H}}\mathbf{U} = \mathbf{I}_k$ with a diagonal matrix $\mathbf{D_U} = \operatorname{diag}(\lambda_{j_1}, \ldots, \lambda_{j_k})$ for $0 < j_1 < \cdots < j_k < 2m$. This means that $\mathcal{U} = \mathsf{im}\,(\mathbf{U})$ is an exact $\mathbf{A}$-invariant subspace. Then a straightforward computation reveals that $\mathbf{P_A} = \mathbf{Q_A} = \mathbf{I} - \mathbf{U}\mathbf{U}^{\mathsf{H}}$, which is obviously Hermitian, and

$$\mathbf{P_A}\mathbf{A}\mathbf{P_A} = \mathbf{P_A}\mathbf{A}\mathbf{Q_A} = \mathbf{P_A}^2\mathbf{A} = \mathbf{P_A}\mathbf{A}, \qquad \mathbf{P_A}\mathbf{Q_A}^{\mathsf{H}} = \mathbf{P_A}^2 = \mathbf{P_A}.$$

By comparing the correction steps of RMinRes and deflated MinRes (cf. sections 6.1 and 6.2) and using $\mathbf{P_A}\mathbf{A}\mathbf{M_A} = \mathbf{0}$, we can see that both methods are mathematically equivalent if $\mathcal{U}$ is an exact invariant subspace.

We solve the system $\mathbf{A}\mathbf{x} = \mathbf{b}$ with a random right-hand side $\mathbf{b}$ and the initial guess $\mathbf{x}_0 = \mathbf{0}$. In Figure 6.1 we show the relative residual norms of the solvers

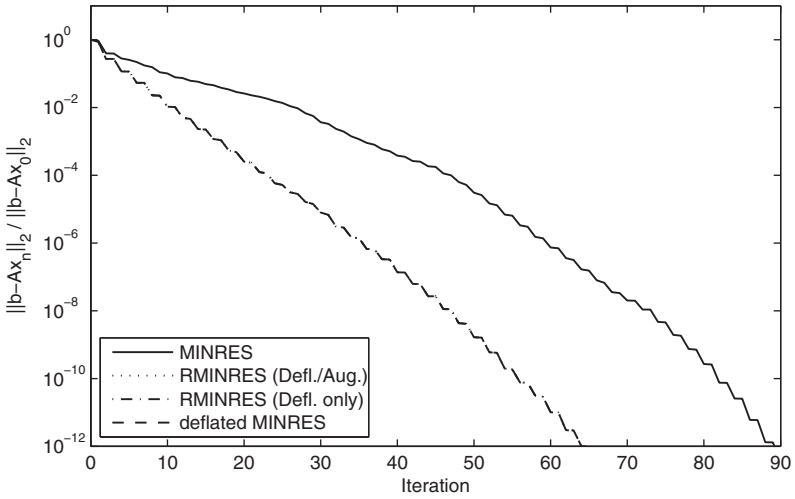


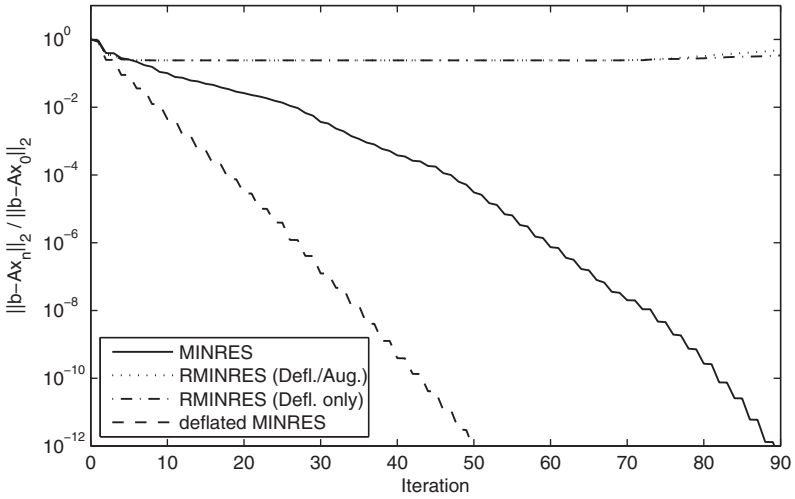
FIG. 6.1. *Convergence history for Example* 6.3. *The convergence curves of both* RMINRES *solver implementations and the deflated* MINRES *method coincide.*

- MINRES (solid line),
- RMINRES with explicit augmentation and deflation (dotted line) according to Wang, de Sturler, and Paulino [57]; cf. (6.3),
- RMINRES with deflation only (dash-dotted line), i.e., the residual norms of MINRES applied to the system $\mathbf{P_A A x} = \mathbf{P_A b}$; cf. (6.4),
- deflated MINRES (dashed line), i.e., the residual norms of MINRES applied to the system $\mathbf{P_A A P_A x} = \mathbf{P_A Q_A^H b}$; cf. section 6.2.

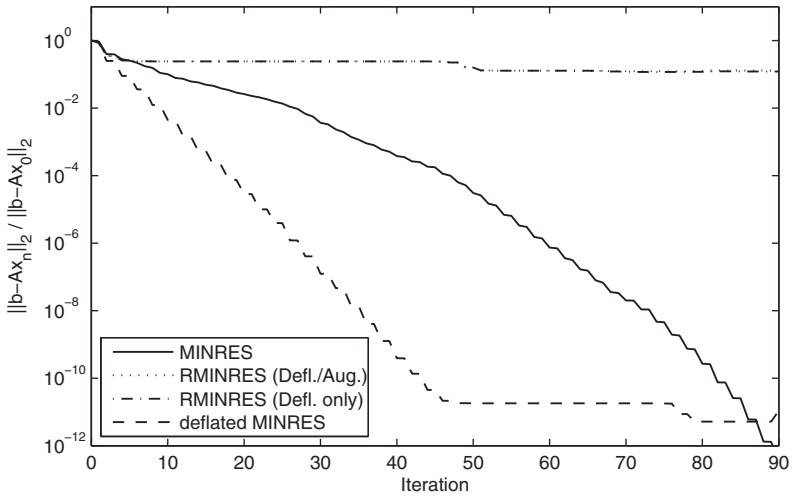
For the last three methods we used the matrix $\mathbf{U} = [\mathbf{w}_1, \ldots, \mathbf{w}_5, \mathbf{w}_{51}, \ldots, \mathbf{w}_{55}]$ which contains the eigenvectors associated with the 10 eigenvalues of $\mathbf{A}$ of smallest absolute value. Thus the deflation space $\mathcal{U}$ has dimension 10. We have shown above that the two implementations of RMINRES and the deflated MINRES method are mathematically equivalent, and in this example the three convergence curves corresponding to these methods indeed coincide; see Figure 6.1.

Example 6.4. We now investigate breakdowns and near-breakdowns of the RMINRES method using a set of artificially constructed examples. Of course, the occurrence of an exact breakdown as in the following examples will be rare in practical applications.

For our construction we use the same matrix $\mathbf{A}$ as in Example 6.3, and we construct a subspace $\mathcal{U}$ for which $\mathcal{U} \cap (\mathbf{A}\mathcal{U})^\perp \neq \{\mathbf{0}\}$. Thus, the condition that guarantees a breakdown-free RMINRES computation is violated; cf. section 6.1. To construct the subspace $\mathcal{U}$ we choose an integer k, $0 < k < m$, and we define $\mathbf{W}_1 = [\mathbf{w}_{i_1}, \ldots, \mathbf{w}_{i_k}]$ and $\mathbf{W}_2 = [\mathbf{w}_{m+i_1}, \ldots, \mathbf{w}_{m+i_k}]$ for indices $0 < i_1 < \cdots < i_k < m$. With $\mathbf{D_U} = \operatorname{diag}(\lambda_{i_1}, \ldots, \lambda_{i_k})$ we obtain $\mathbf{AW}_1 = \mathbf{W}_1\mathbf{D_U}$ and $\mathbf{AW}_2 = -\mathbf{W}_2\mathbf{D_U}$ because of the symmetry of the spectrum of $\mathbf{A}$. We now choose the matrix $\mathbf{U} = \mathbf{W}_1 + \mathbf{W}_2$. Applying $\mathbf{A}$ yields $\mathbf{AU} = (\mathbf{W}_1 - \mathbf{W}_2)\mathbf{D_U}$ and using the fact that $\mathbf{W}$ is unitary shows that $\mathbf{U^H AU} = \mathbf{0}$, or, equivalently, $\mathcal{U} \subset (\mathbf{A}\mathcal{U})^\perp$. The proof of Theorem 5.1 gives us a way to construct an initial guess which leads to an immediate breakdown of RMINRES. For an arbitrary $\mathbf{0} \neq \mathbf{u} \in \mathcal{U}$ we choose $\mathbf{x}_0 = \mathbf{A}^{-1}(\mathbf{b} - \mathbf{u})$. Because of $\mathcal{U} \perp \mathbf{A}\mathcal{U}$ we have $\mathbf{P_A u} = \mathbf{u}$ and the initial residual of RMINRES is $\mathbf{r}_0 = \mathbf{P_A b} - \mathbf{P_A A x}_0 = \mathbf{u}$. The breakdown then occurs in the first iteration because $\mathbf{P_A A r}_0 = \mathbf{P_A A u} = \mathbf{0}$ since



(a) Unperturbed deflation space $\mathbf{U}^{(1)}$



(b) Perturbed deflation space $\mathbf{U}^{(2)} = \mathbf{U}^{(1)} + \mathbf{E}$

FIG. 6.2. *Convergence history for Example* 6.4. *The convergence curves of both* RMINRES *solver implementations coincide.*

$\mathbf{Au} \in \mathbf{A}\mathcal{U} = \mathsf{ker}\,(\mathbf{P_A})$. For these constructed initial guesses the RMINRES method indeed breaks down immediately in numerical experiments, whereas the deflated MINRES method finds the solution after one step. There is no need to plot these results.

Of greater interest are situations with perturbed data. Interestingly, randomly perturbed initial guesses lead to a breakdown of RMINRES with the previously constructed deflation space as well. In Figure 6.2(a) we show the relative residual norms of the solvers listed above applied to the same $\mathbf{A}$ and $\mathbf{b}$ as in the previous example and with the matrix $\mathbf{U}^{(1)} = [\mathbf{w}_1 + \mathbf{w}_{51}, \ldots, \mathbf{w}_{10} + \mathbf{w}_{60}]$.

Furthermore, breakdowns also occur when we perturb the deflation space. Figure 6.2(b) shows the results for a perturbed matrix $\mathbf{U}^{(2)} = \mathbf{U}^{(1)} + \mathbf{E}$ with a random $\mathbf{E} \in \mathbb{C}^{100\times 10}$ and $||\mathbf{E}||_2 = 10^{-10}$. The used initial guess is the same perturbed initial guess as in the experiment conducted for Figure 6.2(a).

Note that both RMinRes implementations suffer from a breakdown after a few steps with both matrices $\mathbf{U}^{(1)}$ and $\mathbf{U}^{(2)}$. With the unperturbed matrix $\mathbf{U}^{(1)}$ the deflated MinRes method converges to the solution with a relative residual smaller than 10^{-12}, while in the case of the perturbed matrix $\mathbf{U}^{(2)}$ the method stagnates with a relative residual of order 10^{-11}. This stagnation of deflated MinRes seems to be related to an unfavorable spectrum of $\mathbf{P_A AP_A}$ for these specifically constructed and perturbed matrices like $\mathbf{U}^{(2)}$. It is unlikely that the stagnation is caused by roundoff errors because the stagnation also occurs (up to iteration 100) when full recurrences (GMRes) are used instead of short recurrences (MinRes). Perturbing the matrix $\mathbf{U}$ from Example 6.3, whose columns are exact eigenvectors of $\mathbf{A}$, does not cause stagnation. This behavior is still subject to further research.

Note that the construction of $\mathcal{U} = \mathsf{im}\,(\mathbf{U})$ in Example 6.4 is such that $\mathbf{A}\mathcal{U} \perp \mathcal{U}$ which cannot be achieved with a (nearly) $\mathbf{A}$-invariant subspace if $\mathbf{A}$ is Hermitian. In [57] an approximation to an invariant subspace of a previous matrix in a sequence of linear algebraic systems is used. In this situation care has to be taken that the extracted space is still a good approximation to an invariant subspace of the current matrix. However, in the experiments of [57, section 7] this seems to be fulfilled since stagnation has not been observed.

7. Conclusions. In this paper we first analyzed theoretically the link between basic theoretical properties of deflated and augmented Krylov subspace methods whose residuals satisfy a Galerkin condition, including the minimum residual methods whose inclusion into the class of Galerkin methods requires a replacement of the standard inner product. We proved that augmentation can be achieved without explicitly augmenting the Krylov subspace but instead projecting the residuals appropriately and using a correction formula for the approximate solutions. We discussed this result in detail for the CG method and GMRes/MinRes methods, the main representatives of our class. It turned out that for these methods some of our results had been mentioned before in the literature.

The projections which arise from the augmentation can also be used to obtain a deflated system. We have seen that a left-multiplication of the original system with the corresponding projection yields a deflated system for which the CG method and GMRes/MinRes methods implicitly achieve augmentation. We proved that for nonsingular Hermitian matrices the MinRes method for the deflated system is equivalent to the solver part of the RMinRes method introduced in [57]. While CG never breaks down, GMRes, MinRes, and thus RMinRes may suffer from breakdowns when used with the deflated systems. We stated necessary and sufficient conditions to characterize breakdowns of these minimal residual methods. For Hermitian matrices, we introduced the deflated MinRes method which also uses a Hermitian deflated matrix and proved that it cannot break down. These results were illustrated numerically.

Our framework covers methods based on a specific type of Galerkin condition; see (2.1)–(2.2). It does not include methods based on other conditions, in particular those that in practical methods are realized using the non-Hermitian Lanczos algorithms. Examples for such methods are BiCG [21] and its variants including CGS [50], Bi-CGStab [55], and IDR(s) [51]. Extending our framework to such methods remains a subject of further work.

Moreover, in this paper we did not discuss or recommend practical choices of deflation or augmentation spaces. Finding spaces that lead to an improved convergence behavior of the deflated or augmented method is a highly challenging task that should be attacked with a specific application in mind. Similar to preconditioning,

there exists no single best strategy for choosing deflation or augmentation spaces in practice. Often one deflates (approximations of) eigenvectors corresponding to the smallest eigenvalues of the given matrix. For symmetric or Hermitian positive definite matrices, this strategy can be shown to reduce the effective condition number, which in turn leads to improved convergence bounds and actually faster convergence of the iterative solver; see, e.g., [56]. For nonsymmetric or non-Hermitian matrices, however, the question of effective choices of deflation or augmentation spaces is largely open.

Acknowledgments. The authors wish to thank the anonymous referees for their comments which helped to improve the presentation.

REFERENCES

[1] J. Baglama, D. Calvetti, G. H. Golub, and L. Reichel, *Adaptively preconditioned GMRES algorithms*, SIAM J. Sci. Comput., 20 (1998), pp. 243–269.
[2] A. H. Baker, E. R. Jessup, and T. Manteuffel, *A technique for accelerating the convergence of restarted GMRES*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 962–984.
[3] M. Benzi, *Preconditioning techniques for large linear systems: A survey*, J. Comput. Phys., 182 (2002), pp. 418–477.
[4] P. N. Brown, *A theoretical comparison of the Arnoldi and GMRES algorithms*, SIAM J. Sci. Statist. Comput., 12 (1991), pp. 58–78.
[5] P. N. Brown and H. F. Walker, *GMRES on (nearly) singular systems*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 37–51.
[6] A. Chapman and Y. Saad, *Deflated and augmented Krylov subspace techniques*, Numer. Linear Algebra Appl., 4 (1997), pp. 43–66.
[7] E. J. Craig, *The N-step iteration procedures*, J. Math. Phys., 34 (1955), pp. 64–73.
[8] J. Cullum, *Peaks, plateaus, numerical instabilities in a Galerkin/minimal residual pair of methods for solving $Ax = b$*, Appl. Numer. Math., 19 (1995), pp. 255–278.
[9] D. Darnell, R. B. Morgan, and W. Wilcox, *Deflated GMRES for systems with multiple shifts and multiple right-hand sides*, Linear Algebra Appl., 429 (2008), pp. 2415–2434.
[10] E. de Sturler, *Nested Krylov methods based on GCR*, J. Comput. Appl. Math., 67 (1996), pp. 15–41.
[11] E. De Sturler, *Truncation strategies for optimal Krylov subspace methods*, SIAM J. Numer. Anal., 36 (1999), pp. 864–889.
[12] Z. Dostál, *Conjugate gradient method with preconditioning by projector*, Int. J. Comput. Math., 23 (1988), pp. 315–323.
[13] M. Eiermann and O. Ernst, *Geometric aspects in the theory of Krylov space methods*, Acta Numer., 10 (2001), pp. 251–312.
[14] M. Eiermann, O. G. Ernst, and O. Schneider, *Analysis of acceleration strategies for restarted minimal residual methods*, J. Comput. Appl. Math., 123 (2000), pp. 261–292.
[15] S. C. Eisenstat, H. C. Elman, and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal., 20 (1983), pp. 345–357.
[16] H. C. Elman, *Iterative Methods for Large, Sparse, Nonsymmetric Systems of Linear Equations*, Ph.D. thesis, Yale University, New Haven, CT, 1982.
[17] J. Erhel, K. Burrage, and B. Pohl, *Restarted GMRES preconditioned by deflation*, J. Comput. Appl. Math., 69 (1996), pp. 303–318.
[18] J. Erhel and F. Guyomarc'h, *An augmented conjugate gradient method for solving consecutive symmetric positive definite linear systems*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1279–1299.
[19] 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.
[20] Y. A. Erlangga and R. Nabben, *Multilevel projection-based nested Krylov iteration for boundary value problems*, SIAM J. Sci. Comput., 30 (2008), pp. 1572–1595.
[21] R. Fletcher, *Conjugate gradient methods for indefinite systems*, in Proceedings of the 6th Biennial Dundee Conference on Numerical Analysis, University of Dundee, Dundee, 1975, Lecture Notes in Math. 506, Springer, Berlin, 1976, pp. 73–89.
[22] J. Frank and C. Vuik, *On the construction of deflation-based preconditioners*, SIAM J. Sci. Comput., 23 (2001), pp. 442–462.

[23] R. W. Freund, *On conjugate gradient type methods and polynomial preconditioners for a class of complex non-Hermitian matrices*, Numer. Math., 57 (1990), pp. 285–312.
[24] V. M. Fridman, *The method of minimum iterations with minimum errors for a system of linear algebraic equations with a symmetrical matrix*, USSR Comput. Math. Math. Phys., 2 (1963), pp. 362–363.
[25] L. Giraud, S. Gratton, X. Pinel, and X. Vasseur, *Flexible GMRES with deflated restarting*, SIAM J. Sci. Comput., 32 (2010), pp. 1858–1878.
[26] A. Greenbaum, *Iterative Methods for Solving Linear Systems*, Frontiers in Applied Mathematics 17, SIAM, Philadelphia, 1997.
[27] M. H. Gutknecht and M. Rozložník, *By how much can residual minimization accelerate the convergence of orthogonal residual methods?*, Numer. Algorithms, 27 (2001), pp. 189–213.
[28] M. H. Gutknecht and M. Rozložník, *A framework for generalized conjugate gradient methods—with special emphasis on contributions by Rüdiger Weiss*, Appl. Numer. Math., 41 (2002), pp. 7–22.
[29] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Research Nat. Bur. Standards, 49 (1952), pp. 409–436.
[30] E. F. Kaasschieter, *Preconditioned conjugate gradients for solving singular systems*, J. Comput. Appl. Math., 24 (1988), pp. 265–275.
[31] S. A. Kharchenko and A. Y. Yeremin, *Eigenvalue translation based preconditioners for the* GMRES(k) *method*, Numer. Linear Algebra Appl., 2 (1995), pp. 51–77.
[32] M. E. Kilmer and E. de Sturler, *Recycling subspace information for diffuse optical tomography*, SIAM J. Sci. Comput., 27 (2006), pp. 2140–2166.
[33] L. Y. Kolotilina, *Twofold deflation preconditioning of linear algebraic systems.* I. *Theory*, J. Math. Sci., 89 (1998), pp. 1652–1689.
[34] L. Mansfield, *On the conjugate gradient solution of the Schur complement system obtained from domain decomposition*, SIAM J. Numer. Anal., 27 (1990), pp. 1612–1620.
[35] L. Mansfield, *Damped Jacobi preconditioning and coarse grid deflation for conjugate gradient iteration on parallel computers*, SIAM J. Sci. Statist. Comput., 12 (1991), pp. 1314–1323.
[36] R. B. Morgan, *A restarted GMRES method augmented with eigenvectors*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 1154–1171.
[37] R. B. Morgan, *Restarted block-GMRES with deflation of eigenvalues*, Appl. Numer. Math., 54 (2005), pp. 222–236.
[38] 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.
[39] R. Nabben and C. Vuik, *A comparison of deflation and the balancing preconditioner*, SIAM J. Sci. Comput., 27 (2006), pp. 1742–1759.
[40] R. Nabben and C. Vuik, *A comparison of abstract versions of deflation, balancing and additive coarse grid correction preconditioners*, Numer. Linear Algebra Appl., 15 (2008), pp. 355–372.
[41] R. A. Nicolaides, *Deflation of conjugate gradients with applications to boundary value problems*, SIAM J. Numer. Anal., 24 (1987), pp. 355–365.
[42] M. A. Olshanskii and V. Simoncini, *Acquired clustering properties and solution of certain saddle point systems*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2754–2768.
[43] C. C. Paige and M. A. Saunders, *Solutions of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.
[44] Y. Saad, *Krylov subspace methods for solving large unsymmetric systems*, Math. Comp., 37 (1981), pp. 105–126.
[45] Y. Saad, *Analysis of augmented Krylov subspace methods*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 435–449.
[46] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, 2003.
[47] Y. Saad and M. H. Schultz, *GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.
[48] Y. Saad, M. Yeung, J. Erhel, and F. Guyomarc'h, *A deflated version of the conjugate gradient algorithm*, SIAM J. Sci. Comput., 21 (2000), pp. 1909–1926.
[49] V. Simoncini and D. B. Szyld, *Recent computational developments in Krylov subspace methods for linear systems*, Numer. Linear Algebra Appl., 14 (2007), pp. 1–59.
[50] P. Sonneveld, *CGS, a fast Lanczos-type solver for nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 10 (1989), pp. 36–52.
[51] P. Sonneveld and M. B. van Gijzen, IDR(s)*: A family of simple and fast algorithms for solving large nonsymmetric systems of linear equations*, SIAM J. Sci. Comput., 31 (2008), pp. 1035–1062.
[52] E. Stiefel, *Relaxationsmethoden bester Strategie zur Lösung linearer Gleichungssysteme*, Comm. Math. Helv., 29 (1955), pp. 157–179.

[53] J. M. Tang, S. P. MacLachlan, R. Nabben, and C. Vuik, *A comparison of two-level preconditioners based on multigrid and deflation*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 1715–1739.

[54] 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.

[55] H. A. van der Vorst, *Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 631–644.

[56] C. Vuik, R. Nabben, and J. Tang, *Deflation acceleration for domain decomposition preconditioners*, in Proceedings of the 8th European Multigrid Conference, TU Delft, Delft, The Netherlands, P. Wesseling, C. Oosterlee, and P. Hemker, eds., 2006.

[57] S. Wang, E. de Sturler, and G. H. Paulino, *Large-scale topology optimization using preconditioned Krylov subspace methods with recycling*, Internat. J. Numer. Methods Engrg., 69 (2007), pp. 2441–2468.

[58] R. Weiss, *Convergence Behavior of Generalized Conjugate Gradient Methods*, Ph.D. thesis, University of Karlsruhe, Karlsruhe, Germany, 1990.

[59] R. Weiss, *Error-minimizing Krylov subspace methods*, SIAM J. Sci. Comput., 15 (1994), pp. 511–527.

[60] R. Weiss, *Properties of generalized conjugate gradient methods*, Numer. Linear Algebra Appl., 1 (1994), pp. 45–63.

[61] M. Yeung, J. Tang, and C. Vuik, *On the convergence of GMRES with invariant-subspace deflation*, Report 10-14, Delft University of Technology, Delft, The Netherlands, 2010.