Deflation and balancing preconditioners for Krylov subspace methods applied to nonsymmetric matrices*

Yogi A. Erlangga† and Reinhard Nabben‡

Abstract

For quite some times, deflation preconditioner has been proposed and used to accelerate the convergence of Krylov subspace methods. For symmetric positive definite linear systems, convergence of CG combined with deflation has been analyzed and compared with other preconditioners, e.g. with the abstract balancing preconditioner *[Nabben and Vuik, SIAM J. Sci. Comput., 27 (2006), pp. 1742–1759]*. In this paper, we extend the convergence analysis to nonsymmetric linear systems in the context of GMRES iteration, and compare it with the abstract nonsymmetric balancing preconditioner. We show that under certain conditions, the 2-norm of residuals produced by GMRES combined with deflation is never larger than the 2-norm of residuals produced by GMRES combined with the balancing preconditioner. Numerical experiments are done to nonsymmetric linear systems arising from a finite volume discretization of the convection-diffusion equation, and the numerical results confirm our theoretical results.

Key words. Deflation, balancing preconditioner, nonsymmetric matrix, GMRES, convection-diffusion.

1 Introduction

For a linear system

$$Au = b, \qquad A \in \mathbb{R}^{n \times n}, \tag{1}$$

where A is a large but sparse nonsymmetric, nonsingular matrix, GMRES [18], among others, is a popular method to iteratively solve it. Such a system

*This was supported by the *Deutsche Forschungsgemeinschaft* (DFG), Project Number NA248/2-2

†TU Berlin, Institut für Mathematik, MA 3-3, Strasse des 17. Juni 136, D-10623 Berlin, Germany, phone: +49-30-31429291, fax: +49-30-31429621(erlangga@math.tu-berlin.de)

‡TU Berlin, Institut für Mathematik, MA 3-3, Strasse des 17. Juni 136, D-10623 Berlin, Germany, phone: +49-30-31429291, fax: +49-30-31429621 (nabben@math.tu-berlin.de)

is encountered, for example, when a discretization is applied to the steady convection-diffusion equation. For starting vector u_0, GMRES constructs a sequence of vectors (called Arnoldi vectors) using Arnoldi orthogonalization [1], which forms the basis for the Krylov subspace, i.e. the subspace

$$\mathcal{K}^k(A, r_0) = \mathrm{span}\{r_0, Ar_0, A^2r_0, \ldots, A^{k-1}r_0\}, \quad r_0 = b - Au_0. \tag{2}$$

The approximate solution at the k-th iteration, denoted by u_k, is then contained in the affine subspace $u_0 + \mathcal{K}^k(A, r_0)$, i.e.,

$$u_k \in u_0 + \mathcal{K}^k(A, r_0). \tag{3}$$

In case of GMRES, u_k minimizes the 2-norm of the residual over the subspace.

In many applications, however, GMRES exhibits slow convergence. Since all Arnoldi vectors are needed during orthogonalization, slow convergence increases the number of Arnoldi vectors being used and stored in the computer memory. This makes GMRES often impractical. A simple remedy to the memory requirement is by restarting GMRES after j iterations, as already suggested in [18], denoted by GMRES(j) throughout. The suitable value of j, the restarting parameter, is generally not known, and an inappropriate value of j may lead GMRES to stagnation.

Morgan [11] proposed a remedy in the context of GMRES(j) by reusing informations already had during the first j iterations. Vectors related to the converged eigenvectors available during the first j GMRES iterations are added to the subspace before restarting; thus, the subspace is *augmented*. Adding these vectors removes (or *deflates*) the corresponding (small) eigenvalues from the spectrum. Related work can also be found in [8, 6, 3]. See also [5] for a unified overview on this class of methods.

A similar idea has also been used in the context of preconditioned conjugate gradient (CG) methods for symmetric positive definite (spd) systems; see e.g. [14, 7]. As the convergence of CG is related to the condition number of the spd linear system to solve, deflation is used to improve the condition number by shifting some of the smallest eigenvalues to zero. Since the corresponding eigenvectors no longer have components during the iterations [14], CG will converge faster. Here, we can speak of the *effective* condition number after deflation, which is never larger than the *original* condition number. In [12, 13], the deflation-based preconditioner is analyzed and compared with the *abstract* form of coarse grid correction preconditioner [15] and the *abstract* form of balancing preconditioner [9]. In theory, the deflation vectors are not necessarily invariant vectors, and more generally, can also be related to the prolongation matrix in the multigrid language.

It is somewhat worthwhile to extend the analysis to nonsymmetric systems. This is the aim of this paper. In addition, we compare deflation with the balancing preconditioner as well. For this purpose, we define the deflation preconditioners as

$$P_D = I - AZE^{-1}Y^T, \quad Q_D = I - ZE^{-1}Y^TA, \quad E = Y^TAZ, \tag{4}$$

where P_D and Q_D are related to the left and right preconditioner, respectively. One can easily show that P_D and Q_D are projectors, i.e., $P_D^2 = P_D$ and $Q_D^2 = Q_D$. Here, Z and Y are suitable deflation subspaces of dimension $n \times r$, which $r \ll n$, and hence E is presumably easy to compute and invert.

In deflation, the solution of (1) is computed as follows. We decompose the solution u into

$$u = (I - Q_D)u + Q_D u = ZE^{-1}Y^T b + Q_D u. \tag{5}$$

As the first term in the right-hand side is easily computed, the factor $Q_D u$ is then obtained by computing $\widetilde{u}$ from

$$P_D A \widetilde{u} = P_D b \tag{6}$$

and then premultiplying it with Q_D. To solve (6) we apply a Krylov subspace method for nonsymmetric systems, e.g. GMRES or Bi-CGSTAB [20]. In case of (6), however, the system is singular. A singular system can still however be solved as long as it is consistent (i.e., $b \in \mathcal{R}(A)$). This is actually the case for (6) because the same projection is applied to both sides. Furthermore, Brown and Walker [2] noted that the least-square problems in GMRES will give solution without breakdown if $\mathcal{N}(A) = \mathcal{N}(A^T)$ or if $\mathcal{N}(A) \cap \mathcal{R}(A) = \{0\}$, even though $\mathcal{N}(A) \neq \mathcal{N}(A^T)$.

For symmetric systems the balancing preconditioner was proposed by Mandel [9]. It is used in domain decomposition methods, and has been analyzed by several other authors in [10, 4, 17, 16, 19]. For nonsymmetric systems we consider the *abstract* balancing preconditioner of the form

$$P_B = Q_D M^{-1} P_D + ZE^{-1}Y^T, \tag{7}$$

with M a nonsingular and possibly nonsymmetric preconditioning matrix. For symmetric positive definite cases (Q_D is replaced by P_D^T in (7)), this preconditioner has already been compared with deflation in [13]. With respect to preconditioning with M, the deflated preconditioning system can be written as

$$M^{-1} P_D A u = M^{-1} P_D b. \tag{8}$$

As mentioned above, in general some assumptions have to be satisfied to guarantee that GMRES will converge for nonsingular systems. However, we will prove that GMRES applied to (6) and (8) will converge without any further assumption.

We first compare spectral properties of deflation and the balancing preconditioner. We prove that $P_B A$ and $M^{-1} P_D A$ have the same spectra except for the first r eigenvalues. With these informations, bounds of GMRES convergence can be derived. These are presented in Section 2. In Section 3, GMRES residuals for $M^{-1} P_D A$ and $P_B A$ are compared. With special starting vector, a relation between residuals of GMRES combined with deflation and the balancing preconditioner can be established for arbitrary full ranked Z and Y. We prove that the preconditioned residual obtained by using the deflation method is less than or equal to the preconditioned residual obtained by using the abstract balancing method. Numerical examples are shown in Section 4 for the convection-diffusion equation. Finally, conclusions are drawn in Section 5.

2 Spectral properties

In this section we evaluate spectral properties of $P_D A$ and their connections with convergence bound of GMRES, and then compare them with $P_B A$. We start the discussion with the case where M is the identity matrix, i.e., equivalently, the case without preconditioning. The last part of this section is then devoted to the case with any nonsingular preconditioner M. Before doing so, we recall some properties related to P_D and Q_D in the next lemma, whose proofs are easily shown by direct computation.

Lemma 2.1 *Let P_D and Q_D be defined as in (4). For any $Z, Y \in \mathbb{R}^{n\times r}$ with rank r, the following equalities hold.*

(i) $P_D A Z = 0$, $Y^T A Q_D = 0$,

(ii) $Q_D Z = 0$, $Y^T P_D = 0$,

(iii) $P_D A = A Q_D$.

Next we present the convergence bound of GMRES for unpreconditioned system $Au = b$ due to Saad and Schultz [18], with its proof.

Assumption 2.2 *$A \in \mathbb{R}^{n\times n}$ is nonsymmetric and diagonalizable, with spectral decomposition $A = X\Lambda X^{-1}$. Here, $X = [x_1 \dots x_n]$ are right eigenvectors of A and $\Lambda = diag(\lambda_1, \dots, \lambda_n)$, satisfying $Ax_i = \lambda_i x_i$, $i = 1, \dots, n$. The eigenvalues λ_i are assumed to be real and nondefective, and $0 < \lambda_i < \lambda_j$, for $i < j$.*

Theorem 2.3 *Let A satisfy Assumption 2.2. Then, at the k-th iteration, GMRES applied to $Au = b$ with starting vector u_0 produces residual which satisfies the inequality*

$$\|r_k\|_2 \le \kappa_2(X)\epsilon^k \|r_0\|_2, \tag{9}$$

where $\kappa_2(X) = \|X\|_2 \|X^{-1}\|_2$ is the condition number of X, $r_0 = b - Au_0$, and

$$\epsilon^k = \min_{p\in\mathbb{P}^k, p(0)=1} \max_{i=1,\dots,n} |p(\lambda_i)|, \tag{10}$$

with any polynomial p whose degree is not larger than $k-1$, and satisfying the constraint $p(0) = 1$.

Proof Let p be a polynomial of degree no larger than $k-1$ with constraint $p(0) = 1$, and u be a vector in $\mathcal{K}^k$ associated with the residual $b - Au = p(A)r_0$. Then, for $A = X\Lambda X^{-1}$,

$$\|b - Au\|_2 = \|P(A)r_0\|_2 = \|Xp(\Lambda)X^{-1}r_0\|_2 \le \|X\|_2\|X^{-1}\|_2\|\|p(\Lambda)\|_2\|r_0\|_2.$$

Since Λ is a diagonal matrix, $\|p(\Lambda)\|_2 = \max_{i=1,\ldots,n} |p(\lambda_i)|$. Consider now u_k, extracted from $\mathcal{K}^k$ but now related to GMRES approximation. Since u_k minimizes the 2-norm of the residual over $u_0 + \mathcal{K}^k$, then for any polynomial p

$$\|b - Au_k\|_2 \leq \|b - Au\|_2 \leq \|X\|_2 \|X^{-1}\|_2 \|\|r_0\|_2 \max_{i=1,\ldots,n} |p(\lambda_i)|. \tag{11}$$

Choosing a polynomial which minimizes the right-hand side one has

$$\|b - Au_k\|_2 \leq \|X\|_2 \|X^{-1}\|_2 \|\|r_0\|_2 \min_{p \in \mathbb{P}^k, p(0)=1} \max_{i=1,\ldots,n} |p(\lambda_i)|,$$

leading to the theorem, with $\|X\|_2 \|X^{-1}\|_2 =: \kappa_2(X)$. □

A further result is obtained by considering the min-max problem above as the shifted and scaled Chebyshev polynomial in the interval $[\lambda_1, \lambda_n]$,

$$\hat{C}_k(t) = \frac{C_k\left(1 + 2\frac{\lambda_1 - t}{\lambda_n - \lambda_1}\right)}{C_k\left(1 + 2\frac{\lambda_1 - \gamma}{\lambda_n - \lambda_1}\right)}, \tag{12}$$

with $\gamma < \lambda_1$. In this case, setting $\gamma = 0$, the constraint, results in

$$\min_{p \in \mathbb{P}^k, p(0)=1} \max_{i=1,\ldots,n} |p(\lambda_i)| = \min_{p \in \mathbb{P}^k, p(0)=1} \max_{\lambda \in [\lambda_1, \lambda_n]} |p(\lambda_i)| = \frac{1}{C_k\left(2\frac{\mu}{\lambda_n - \lambda_1}\right)}, \tag{13}$$

where $\mu = (\lambda_1 + \lambda_n)/2$. As $C_k(t)$ can alternatively be written as

$$C_k(t) = \frac{1}{2}\left[\left(t + \sqrt{t^2 - 1}\right)^k + \left(t + \sqrt{t^2 - 1}\right)^{-k}\right] \geq \frac{1}{2}\left(t + \sqrt{t^2 - 1}\right)^k, \tag{14}$$

we see, after some manipulations, that $C_k(2\mu/(\lambda_n - \lambda_1)) \geq 1 - 1/(\sqrt{\kappa} + 1)$, where $\kappa = \lambda_n/\lambda_1$. We then have the following corollary.

Corollary 2.4 *Let A be defined as in Assumption 2.2 and let $\kappa := \lambda_n/\lambda_1$. After k iterations GMRES produces residual satisfying the following bound:*

$$\|r_k\|_2 \leq 2\kappa_2(X)\left(1 - \frac{2}{\sqrt{\kappa} + 1}\right)^k \|r_0\|_2. \tag{15}$$

This convergence bound is similar to the convergence bound of conjugate gradient (CG) methods, except that the bound is now represented by the 2-norm of residuals. If A is symmetric positive definite (spd), then $\kappa_2(X) = 1$. For more general cases, $\kappa_2(X)$ is not known, and is too expensive to compute. Furthermore, $\kappa = \lambda_n/\lambda_1$ is not the condition number as usually referred to as in case of spd systems. Under Assumption 2.2 we may still, however, associate κ in (15) with the quality of eigenvalues clustering. The smaller the value is, the more clustered the eigenvalues are.

2.1 Unpreconditioned case; $M = I$

In this section we analyze spectral properties of $P_D A$ and compare them with $P_B A$. M is set equal to the identity matrix ($M = I$). In the first part of this section, and some parts in this paper, we need the following assumption.

Assumption 2.5 *We set $Z = [v_1 \dots v_r]$, where $Av_i = \lambda_i v_i$, $i = 1, \dots, r$. Also, we set $Y = [w_1 \dots w_r]$ the left eigenvector matrix of A, determined from $w_i^T A = \lambda_i w_i^T$ and chosen such that $Y^T Z = I_r$, where I_r is the identity matrix of dimension r. For the left eigenvectors, $W = [w_1 \dots w_n]$.*

Theorem 2.6 *Let Z and Y be defined as in Assumption 2.5. Let $M = I$ in P_B. Then*

$$\sigma(P_D A) = \{0, \dots, 0, \lambda_{r+1}, \dots, \lambda_n\},$$
$$\sigma(P_B A) = \{1, \dots, 1, \lambda_{r+1}, \dots, \lambda_n\}.$$

Proof Under Assumption 2.5, obviously $E = Y^T AZ = \mathrm{diag}(\lambda_1, \dots, \lambda_r) =: \Lambda_r$. For deflation, we see that for $i = 1, \dots, r$

$$P_D A v_i = \left(I - AZ\Lambda_r^{-1} Y^T\right) A v_i = \lambda_i v_i - Z\Lambda_r Y^T v_i = 0,$$

because $Y^T Z = I$. Similarly, for $i = r+1, \dots, n$, $P_D A v_i = \lambda_i v_i - Z\Lambda_r Y^T v_i = \lambda_i v_i$, because $Y^T v_i = 0$. This leads to the first result.

For the balancing preconditioner, one can also proceed with the same procedure as above. In this case

$$P_B A v_i = (I - ZE^{-1} Y^T A)(I - AZE^{-1} Y^T) A v_i + ZE^{-1} Y^T A v_i. \tag{16}$$

By expanding (16) and making use of orthogonality of eigenvectors, we have

$$P_B A v_i = v_i, \quad i = 1, \dots, r. \tag{17}$$

Again, due to orthogonality we also have $P_B A v_i = \lambda_i v_i$, for $i = r+1, \dots, n$. □

So, the action of P_D and P_B on A leads to almost the same spectra except that the smallest r eigenvalues are shifted towards 0 and 1, respectively. Furthermore, $P_D A$ and $P_B A$ share the same eigenvectors, which are equal to the eigenvectors of A.

Next we consider the situation where Z and Y are chosen arbitrarily. This is important because for a large matrix, computing eigenvectors is expensive. Furthermore, in the balancing preconditioner, special technique is employed to construct Z and Y; so they are arbitrary. For the purpose of analysis, we only assume that Z and Y have rank r. First, we will see what the consequence of arbitrary Z and Y is on the spectrum of $P_D A$ and $P_B A$.

Lemma 2.7 *Let Z and Y be any rectangular matrices with rank r. Then, for $i = 1 \dots r$,*

$$\lambda_i = 0 \in \sigma(P_D A),$$
$$\lambda_i = 1 \in \sigma(P_B A),$$

with $Z = [z_1 \dots z_r]$ the corresponding eigenvectors.

Proof In this case, $P_DAZ = 0$ and the lemma is proved for P_DA. For P_BA, one observes that

$$P_BAZ = Q_DP_DAZ + ZE^{-1}Y^TAZ = Z, \tag{18}$$

because of Lemma 2.1 and the definition of E. □

For an arbitrary choice of Z and Y, no conclusion can be drawn for the rest of the eigenvalues. However, we can still compare the spectrum of P_DA and P_BA, if we assume that $\sigma(P_DA)$ is known. In this case, for $r+1 \leq i \leq n$ denote μ_i as the eigenvalues of P_DA. Following Lemma 2.7, we then have

$$\sigma(P_DA) = \{0, \ldots, 0, \mu_{r+1}, \ldots, \mu_n\}. \tag{19}$$

Theorem 2.8 *Let Z and Y be any full ranked rectangular matrices. If $\sigma(P_DA) = \{0, \ldots, 0, \mu_{r+1}, \ldots, \mu_n\}$, and $P_DA\widetilde{v}_i = \mu_i\widetilde{v}_i$, with the corresponding eigenvector $\widetilde{v}_i$, then,*

$$\sigma(P_BA) = \{1, \ldots, 1, \mu_{r+1}, \ldots, \mu_n\}. \tag{20}$$

Proof For P_DA, we have

$$P_DA\widetilde{v}_i = P_D^2A\widetilde{v}_i = P_DAQ_D\widetilde{v}_i = \mu_i\widetilde{v}_i. \tag{21}$$

Since $\mu_i\widetilde{v}_i \neq 0$, we obtain $Q_D\widetilde{v}_i \neq 0$. Observe that

$$P_BAQ_D\widetilde{v}_i = Q_DP_DAQ_D\widetilde{v}_i + ZE^{-1}Y^TAQ_D\widetilde{v}_i = \mu_iQ_D\widetilde{v}_i, \tag{22}$$

because $Y^TAQ_D = 0$. This proves the lemma. □

Theorem 2.8 states that for any choice of Z and Y, the spectrum of P_DA and P_BA are similar. In this case, for P_BA, the corresponding eigenvectors are $Q_D\widetilde{v}_i$, with $\widetilde{v}_i$ the eigenvectors of P_DA.

Next we provide a GMRES convergence bound for P_DA and P_DB. Here we restrict our discussion to the case satisfying Assumptions 2.2 and 2.5. We note that, because of Assumptions 2.2 and 2.5,

(i) $X = [Z\, X_{n-r}]$, $X_{n-r} = [x_{r+1} \ldots x_n]$, $W = [Y\, W_{n-r}]$, $W_{n-r} = [w_{r+1} \ldots w_n]$, and

(ii) $E = Y^TAZ = Y^TZ\Lambda_r = \Lambda_r$, where $\Lambda_r = \text{diag}(\lambda_1, \ldots, \lambda_r)$.

For deflation we have the following lemma.

Lemma 2.9 *Define $\Lambda_D = diag\{0, \ldots, 0, \lambda_{r+1}, \ldots, \lambda_n\}$, and $\widetilde{r}_{0,D} = P_D(b - A\widetilde{u}_0)$. Under Assumptions 2.2 and 2.5, the Krylov subspace generated after k GMRES iterations applied to $P_DA\widetilde{u} = P_Db$ is*

$$\mathcal{K}^k(P_DA, \widetilde{r}_{0,D}) = span\{\widetilde{r}_{0,D}, X\Lambda_DX^{-1}\widetilde{r}_{0,D}, \ldots, X\Lambda_D^{k-1}X^{-1}\widetilde{r}_{0,D}\}. \tag{23}$$

Proof For $k = 1$, $\widetilde{r}_0 = P_D(b - A\widetilde{u}_{0,D})$. For $k = 2$,

$$\begin{aligned} P_D A &= (I - AZE^{-1}Y^T)A = A - AZE^{-1}Y^T A, \\ &= X(I - \Lambda X^{-1}ZE^{-1}Y^T X)\Lambda X^{-1}. \end{aligned}$$

Note that, because $W^T X = I$, where $A^T W = W\Lambda$,

$$\begin{aligned} X^{-1}ZE^{-1}Y^T X &= W^T ZE^{-1}Y^T X = \begin{bmatrix} Y^T \\ W_{n-r}^T \end{bmatrix} Z\Lambda^{-1}Y^T \begin{bmatrix} Z & X_{n-r} \end{bmatrix}, \\ &= \begin{bmatrix} Y^T Z\Lambda^{-1} \\ W_{n-r}^T Z\Lambda^{-1} \end{bmatrix} \begin{bmatrix} Y^T Z & Y^T X_{n-r} \end{bmatrix} = \begin{bmatrix} \Lambda^{-1} \\ 0 \end{bmatrix} \begin{bmatrix} I_r & 0 \end{bmatrix}, \\ &= \begin{bmatrix} \Lambda_r^{-1} & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned} \tag{24}$$

Thus,

$$\begin{aligned} P_D A &= X\left(I - \Lambda \begin{bmatrix} \Lambda_r^{-1} & 0 \\ 0 & 0 \end{bmatrix}\right)\Lambda X^{-1} = X\left(I - \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}\right)\begin{bmatrix} \Lambda_r & 0 \\ 0 & \Lambda_{n-r} \end{bmatrix} X^{-1}, \\ &= X\Lambda_D X^{-1}. \end{aligned}$$

For $k = 3$, $(P_D A)^2 = P_D A P_D A = X\Lambda_D X^{-1} X\Lambda_D X^{-1} = X\Lambda_D^2 X^{-1}$. By repeating the computation for $l = 4, \ldots, k-1$ the desired result is obtained.

$\square$

Theorem 2.10 *Let Z and Y be defined as in Assumptions 2.2 and 2.5. Denote $\kappa_D = \lambda_n/\lambda_{r+1}$. Then, for any starting vector $\widetilde{u}_0$, GMRES applied to $P_D A\widetilde{u} = P_D b$ generates residuals whose 2-norm is bounded by*

$$\|\widetilde{r}_{k,D}\|_2 \leq 2\kappa_2(X)\left(1 - \frac{2}{\sqrt{\kappa_D} + 1}\right)^k \|\widetilde{r}_{0,D}\|_2, \tag{25}$$

where $\widetilde{r}_{k,D} = P_D(b - Au_k)$.

Proof Here, we have $\widetilde{r} = p(P_D A)\widetilde{r}_{0,D} = Xp(\Lambda_D)X^{-1}\widetilde{r}_{0,D}$, due to Lemma 2.9, where p is a polynomial of degree no larger than $k-1$, with $p(0) = 1$. Hence,

$$\|\widetilde{r}\|_2 \leq \|X\|_2\|X^{-1}\|_2\|p(\Lambda_D)\|_2\|\widetilde{r}_{0,D}\|_2.$$

Noting that $\widetilde{u}_k$ minimizes the 2-norm of residual over $\widetilde{u}_0 + \mathcal{K}^k(P_D A, \widetilde{r}_{0,D})$ and $\|p(\Lambda_D)\|_2 = \max_{i=r+1,\ldots,n}|p(\lambda_i)|$, choosing a polynomial which minimizes the right-hand side leads to

$$\|\widetilde{r}_{k,D}\|_2 := \|P_D(b - A\widetilde{u}_{k,D})\|_2 \leq \|X\|_2\|X^{-1}\|_2\|\widetilde{r}_0\|_2 \min_{p\in\mathbb{P}^k, p(0)=1} \max_{i=r+1,\ldots,n}|p(\lambda_i)|.$$

Taking the shift and scaled Chebyshev polynomial as the trial polynomial with $\lambda \in [\lambda_{r+1}, \lambda_n]$ and repeating the same procedure as in the previous section one arrives at the desired inequality, with $\kappa_D = \lambda_n/\lambda_{r+1}$. $\square$

We see that A and $P_D A$ share the same eigenvectors. Since $\lambda_{r+1} \geq \lambda_1$, $\kappa_D \leq \kappa$, GMRES with deflation preconditioner will asymptotically converge faster than without deflation preconditioner.

Remark 2.11 *Note that this comparison is not fair because $\widetilde{r}$ is the residual of the preconditioned system, and not of the original system. In practice one usually is more interested in the residual of the original system, which is not the by-product of the left preconditioning GMRES. A more detailed residual analysis in Section 3 reveals, however, that in the way the solution is computed, GMRES combined with deflation produces actual residuals which are the same as the preconditioned residuals; see Lemma 3.2.*

For the balancing preconditioner, the Krylov subspace associated with it is written as

$$\begin{aligned} &\mathcal{K}^k(P_B A, P_B(b - Au_0)) = \\ &\text{span}\{P_B(b - Au_0), P_B A P_B(b - Au_0), \ldots, (P_B A)^{k-1} P_B(b - Au_0)\}. \end{aligned} \tag{26}$$

For cases under Assumptions 2.2 and 2.5 we have the following lemma.

Lemma 2.12 *Define $\Lambda_B = \{1, \ldots, 1, \lambda_{r+1}, \ldots, \lambda_n\}$ and $\widetilde{r}_{0,B} = P_B(b - Au_0)$. With Assumptions 2.2 and 2.5,*

$$\mathcal{K}^k(P_B A, \widetilde{r}_{0,B}) = \mathit{span}\{\widetilde{r}_{0,B}, X\Lambda_B X^{-1}\widetilde{r}_{0,B}, \ldots, X(\Lambda_B)^{k-1} X^{-1}\widetilde{r}_{0,N}\}.$$

Proof For $k = 1$, $\widetilde{r}_{0,B} := P_B(b - Au_0) = P_B r_0$. For $k = 2$,

$$\begin{aligned} P_B A &= (Q_D P_D + ZE^{-1}Y^T)A, \\ &= (I - ZE^{-1}Y^T A)(I - AZE^{-1}Y^T)A + ZE^{-1}Y^T A, \\ &= A - ZE^{-1}Y^T AA - AZE^{-1}Y^T A + ZE^{-1}Y^T AAZE^{-1}Y^T A \\ &+ ZE^{-1}Y^T A, \\ &= X(I - X^{-1}ZE^{-1}Y^T AX - \Lambda X^{-1}ZE^{-1}Y^T X \\ &+ X^{-1}ZE^{-1}Y^T AAZE^{-1}Y^T X + X^{-1}ZE^{-1}Y^T X)\Lambda X^{-1}. \end{aligned}$$

From the proof of Theorem 2.9, we then have that

$$\begin{aligned} P_B A &= X\left(I - \begin{bmatrix} \Lambda_r^{-1} & 0 \\ 0 & 0 \end{bmatrix}\Lambda - \Lambda\begin{bmatrix} \Lambda_r^{-1} & 0 \\ 0 & 0 \end{bmatrix} + X^{-1}ZY^T X + \begin{bmatrix} \Lambda_r^{-1} & 0 \\ 0 & 0 \end{bmatrix}\right)\Lambda X^{-1}, \\ &= X\begin{bmatrix} \Lambda_r^{-1} & 0 \\ 0 & I_{n-r} \end{bmatrix}\Lambda X^{-1} = X\begin{bmatrix} I_r & 0 \\ 0 & \Lambda_{n-r} \end{bmatrix}\Lambda X^{-1} =: X\Lambda_B X^{-1}, \end{aligned}$$

because

$$X^{-1}ZY^T X = W^T ZY^T X = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

We can also compute $(P_B A)^k$ for $k > 2$. This leads to the above lemma. □

Making use of Lemma 2.12, we have the GMRES convergence bound for the balancing preconditioner.

Theorem 2.13 *Let Z and Y in P_B satisfy Assumptions 2.2 and 2.5. Define $\kappa_B = \max\{1, \lambda_n\}/\min\{1, \lambda_{r+1}\}$. For any starting vector u_0, the 2-norm of residual of GMRES applied to $P_B A u = P_B b$ satisfies*

$$\|\widetilde{r}_{k,B}\|_2 \leq 2\kappa_2(X)\left(1 - \frac{2}{\sqrt{\kappa_B} + 1}\right)^k \|\widetilde{r}_{0,B}\|_2. \tag{27}$$

Proof Note that $\|\widetilde{r}_B\| = Xp(\Lambda_B)X^{-1}\widetilde{r}_{0,B}$. The proof then follows the same lines as in Theorem 2.10. □

Comparing Theorems 2.10 and 2.13, it is clear that GMRES combined with deflation has convergence bound which is lower than or equal to GMRES combined with the balancing preconditioner. Therefore, we may expect that GMRES applied to deflation-preconditioned linear system will converge faster than GMRES with the balancing preconditioner.

2.2 Case with preconditioner M

The analysis in the previous section can be extended to the case with any nonsingular preconditioner M. One observation can be made for $M^{-1}P_D A$. Since $P_D AZ = 0$, it follows that $M^{-1}P_D AZ = 0$, and hence by preconditioning M the eigenvalues which are deflated to zero remain untouched. The rest of the spectrum, however, changes due to preconditioning. Denote this spectrum by $\sigma(M^{-1}P_D A) = \{0, \ldots, 0, \mu_{r+1}, \ldots, \mu_n\}$. We can derive a comparison between the spectrum of $M^{-1}P_D A$ and $P_B A$. To have this comparison, we need some intermediate results.

Lemma 2.14 *Let A and $Y^T AZ$ be nonsingular. Then we obtain*

$$\sigma(Q_D M^{-1} P_D A) = \sigma(M^{-1}P_D A) = \sigma(Q_D M^{-1} A). \tag{28}$$

Proof We have

$$\sigma(Q_D M^{-1}P_D A) = \sigma(M^{-1}P_D A Q_D) = \sigma(M^{-1}P_D^2 A) = \sigma(M^{-1}P_D A),$$

which proves the first equality. The second equality can also be proved in a similar way. □

Lemma 2.15 *Let A and $Y^T AZ$ be nonsingular. Then,*

$$\sigma((Q_D M^{-1}P_D + ZE^{-1}Y^T)A) = \sigma((M^{-1}P_D + ZE^{-1}Y^T)A). \tag{29}$$

Proof We note that

$$\begin{aligned}(M^{-1}P_D + ZE^{-1}Y^T)A &= M^{-1}P_D^2 A + ZE^{-1}Y^T A \\ &= M^{-1}P_D A Q_D + ZE^{-1}Y^T A \\ &= (M^{-1}P_D A - I)Q_D + I. \end{aligned} \tag{30}$$

Hence

$\lambda \in \sigma((M^{-1}P_D + ZE^{-1}Y^T)A) \Leftrightarrow \lambda = \mu + 1$ for $\mu \in \sigma((M^{-1}P_DA - I)Q_D)$.

But,

$\sigma((M^{-1}P_DA - I)Q_D)) = \sigma(Q_D(M^{-1}P_DA - I)) = \sigma(Q_DM^{-1}P_DA - Q_D))$.

Thus

$$\begin{aligned}\lambda &\in \sigma((M^{-1}P_D + ZE^{-1}Y^T)A) \\ &= \sigma(Q_DM^{-1}P_DA - Q_D + I) \\ &= \sigma(Q_DM^{-1}P_DA + ZE^{-1}Y^TA) \\ &= \sigma((Q_DM^{-1}P_D + ZE^{-1}Y^T)A), \qquad (31)\end{aligned}$$

which proves the theorem. □

We have the following theorem.

Theorem 2.16 *Let A, M and Y^TAZ be nonsingular. Then, P_BA is nonsingular. In addition, $P_B \equiv Q_DM^{-1}P_D + ZE^{-1}Y^T$ is also nonsingular.*

Proof Consider $(M^{-1}P_D + ZE^{-1}Y^T)A$. Now assume that there is a vector x such that

$$(M^{-1}P_D + ZE^{-1}Y^T)Ax = 0.$$

Then we have

$$\begin{aligned}0 &= M^{-1}P_DAx + ZE^{-1}Y^TAx \qquad (32)\\ &= M^{-1}AQ_Dx + ZE^{-1}Y^TAx \qquad (33)\\ &= M^{-1}Ax - ZE^{-1}Y^TAx + ZE^{-1}Y^TAx \qquad (34)\\ &= M^{-1}Ax. \qquad (35)\end{aligned}$$

Since A and M^{-1} are nonsingular, x must be zero. Hence

$$0 \notin \sigma((M^{-1}P_D + ZE^{-1}Y^T)A) = \sigma((Q_DM^{-1}P_D + ZE^{-1}Y^T)A),$$

due to Lemma 2.15. Therefore, P_BA is nonsingular. Furthermore, $Q_DM^{-1}P_D + ZE^{-1}Y^T$ is as well nonsingular. □

Theorem 2.17 *Suppose that the spectrum of $M^{-1}P_DA$ is given by:*

$$\sigma(M^{-1}P_DA) = \{0, \ldots, 0, \mu_{r+1}, \ldots, \mu_n\},$$

then

$$\sigma(P_BA) = \sigma(P_{\tilde{B}}A) = \{1, \ldots, 1, \mu_{r+1}, \ldots, \mu_n\}.$$

Conversely, if the spectrum of P_BA is given by

$$\sigma(P_BA) = \{1, \ldots, 1, \mu_{r+1}, \ldots, \mu_n\},$$

then

$$\sigma(M^{-1}P_DA) = \{0, \ldots, 0, \mu_{r+1}, \ldots, \mu_n\}.$$

Proof For $i = 1, \ldots, r$, $P_B AZ = Q_D M^{-1} P_D AZ + ZE^{-1} Y^T AZ = Z$, because $P_D AZ = 0$. Hence, the eigenvectors of $P_B AZ$ which correspond to eigenvalues equal to 1 are the same as those corresponding to eigenvalues equal to 0 of $M^{-1} P_D A$, i.e. $Z = [z_1 \ldots z_r]$.

For $r+1 \leq i \leq n$, suppose that $\widetilde{v}_i$ satisfies $M^{-1} P_D A \widetilde{v}_i = \mu_i \widetilde{v}_i$, where μ_i is the corresponding eigenvalue.

First look at a situation where $M^{-1} P_D A \widetilde{v}_i = 0$, for $r+1 \leq i \leq n$; i.e. some non-zero eigenvalues of $P_D A$ become zero ($\mu_i = 0$) by the action of the preconditioner M. In this case we have

$$P_B A \widetilde{v}_i = Q_D M^{-1} P_D A \widetilde{v}_i + ZE^{-1} Y^T A \widetilde{v}_i = ZE^{-1} Y^T A \widetilde{v}_i. \tag{36}$$

But, for nonsingular M, $M^{-1} P_D A \widetilde{v}_i = 0$. Thus, $P_D A \widetilde{v}_i = 0$. Hence, $A \widetilde{v}_i = AZE^{-1} Y^T A \widetilde{v}_i$, leading to $\widetilde{v}_i = ZE^{-1} Y^T A \widetilde{v}_i$. Substitution of this relation to (36), we thus have $P_B A \widetilde{v}_i = \widetilde{v}_i$. Hence, zero eigenvalues in case of deflation are also shifted to one by the balancing preconditioner, leading to a nonsingular matrix $P_B A$.

Next we assume that $\mu_i \neq 0$ for $r+1 \leq i \leq n$. In this case, we have that $M^{-1} P_D A \widetilde{v}_i = M^{-1} P_D^2 A \widetilde{v}_i = M^{-1} P_D A Q_D \widetilde{v}_i = \mu_i \widetilde{v}_i \neq 0$, implying that $Q_D \widetilde{v}_i \neq 0$. Thus,

$$\begin{aligned} P_B A Q_D \widetilde{v}_i &= Q_D M^{-1} P_D A Q_D \widetilde{v}_i + ZE^{-1} Y^T A Q_D \widetilde{v}_i = Q_D M^{-1} P_D^2 A \widetilde{v}_i \\ &= Q_D M^{-1} P_D A \widetilde{v}_i = \mu_i Q_D \widetilde{v}_i. \end{aligned}$$

Hence, for $i = r+1, \ldots, n$, the eigenvalues of $P_B A$ are the same as the eigenvalues of $M^{-1} P_D A$, with eigenvectors $Q_D \widetilde{v}_i$.

To prove the second statement, we know that for $i = 1, \ldots, r$, $P_B AZ = Z$, which gives, by expanding P_B, $Q_D M^{-1} P_D AZ = 0$. Hence, $0 \in \sigma(Q_D M^{-1} P_D A)$, implying $0 \in \sigma(M^{-1} P_D A)$ due to Lemma 2.14.

For $i = r+1, \ldots, n$, notice that

$$P_B A \widetilde{v}_i = Q_D M^{-1} P_D A \widetilde{v}_i + ZE^{-1} Z^T A \widetilde{v}_i = \mu_i \widetilde{v}_i$$

implies

$$Q_D P_B A \widetilde{v}_i = Q_D M^{-1} P_D A \widetilde{v}_i = Q_D M^{-1} P_D A Q_D \widetilde{v}_i = \mu_i Q_D \widetilde{v}_i,$$

because $Q_D Z = 0$. Thus, μ_i is an eigenvalue of $Q_D M^{-1} P_D A$. But, due to Lemma 2.14, it is also an eigenvalue of $M^{-1} P_D A$. This completes the proof. □

In conclusion, preconditioning by $M^{-1} P_D$ and P_B leads to similar clustering for any choice of full ranked Z and Y and any nonsingular M. Furthermore, $P_B A$ is always nonsingular.

3 Comparison of GMRES residuals

To have more detailed comparisons between $P_D A$ and $P_B A$, in this section we evaluate the approximate solutions built by GMRES and the related residuals.

We first recall that when applied to (1), a Krylov subspace method generates an approximation solution in the Krylov subspace $\mathcal{K}^k(A, r_0)$, defined in (2). In case of left preconditioning, the subspace is now spanned by vectors related to the preconditioned system. For

$$BAu = Bb, \tag{37}$$

with B any preconditioner, the Krylov subspace related to the initial residual $\widetilde{r}_0 = B(b - Au_0)$, where u_0 is the starting vector, is given by

$$\mathcal{K}^k(BA, \widetilde{r}_0) = \text{span}\{\widetilde{r}_0, BA\widetilde{r}_0, \ldots, (BA)^{k-1}\widetilde{r}_0\}. \tag{38}$$

GMRES then minimizes the residual norm

$$\|B(b - A\eta)\|_2, \tag{39}$$

where $\eta \in u_0 + \mathcal{K}^k(BA, \widetilde{r}_0)$. The approximate solution is determined by

$$u_k = u_0 + Bp_{k-1}(BA)\widetilde{r}_0, \tag{40}$$

where p_{k-1} is the polynomial of degree $k-1$, which minimizes the residual norm (39) among all other polynomials of degree $\leq k - 1$.

3.1 Unpreconditioned case, $M = I$

We first consider the balancing preconditioner, $P_B A$.

Theorem 3.1 *Let P_B be defined as in (7), with Z and Y any matrices with rank r. Let M be the identity. The Krylov subspace related to $P_B A$ and $u_{0,B} = 0$ has the following property:*

$$\begin{aligned} &\mathcal{K}^k(P_B A, P_B b) \\ &\subseteq \mathit{span}\{ZE^{-1}Y^T b, Q_D P_D b, Q_D(P_D A)P_D b, \ldots, Q_D(P_D A)^{k-1}P_D b\}. \end{aligned} \tag{41}$$

Proof For $k = 1$, we have $\widetilde{r}_0 = P_B(b - Au_0) = P_B b = Q_D P_D b + ZE^{-1}Y^T b$. For $k = 2$,

$$\begin{aligned} P_B A P_B b &= (Q_D P_D A + ZE^{-1}Y^T A)(Q_D P_D b + ZE^{-1}Y^T b), \\ &= Q_D P_D A Q_D P_D b + Q_D P_D A Z E^{-1} Y^T b + ZE^{-1}Y^T A Q_D P_D b \\ &+ ZE^{-1}Y^T A Z E^{-1} Y^T b, \\ &= Q_D P_D A P_D b + ZE^{-1}Y^T b, \end{aligned}$$

because $AQ_D = P_D A$, $P_D^2 = P_D$, and $P_D AZ = Y^T AQ_D = 0$. One can continue computing $(P_B A)^l P_B b$, for $l = 3, \ldots, k$, leading to the desired result. □

For deflation, we know that the approximate solution at k-th iteration is obtained from the relation $u_{k,D} = (I - Q_D)u + Q_D\widetilde{u}_{k,D} = ZE^{-1}Y^T b + Q_D\widetilde{u}_{k,D}$, where $\widetilde{u}_{k,D}$ is computed iteratively from $P_D A\widetilde{u} = P_D b$. For $\widetilde{u}_0 = 0$, $\widetilde{u}_{k,D}$ lies in the Krylov subspace

$$\mathcal{K}^k(P_D A, P_D b) = \text{span}\{P_D b, P_D A P_D B, \ldots, (P_D A)^{k-1} P_D b\}. \tag{42}$$

Thus,

$$\begin{aligned} u_{k,D} &\in ZE^{-1}Y^T b + Q_D \mathcal{K}^k(P_D A, P_D b), \\ &= ZE^{-1}Y^T b + Q_D \text{span}\{P_D b, P_D A P_D B, \ldots, (P_D A)^{k-1} P_D b\}. \end{aligned} \tag{43}$$

Hence, $u_{k,D}$ and $u_{k,B}$ are members of the same subspace.

For general situations, it is difficult to have a residual comparison between $P_D A$ and $P_B A$. A still useful result, however, can be obtained if A is assumed to be diagonalizable. Before coming into that, we have the following residual relation for deflation.

Lemma 3.2 *Let Z and Y be matrices with rank r. Let $\widetilde{u}_{k,D}$ be the approximate solution from GMRES applied to $P_D A\widetilde{u} = P_D b$, and $u_{k,D} = ZE^{-1}Y^T b + Q_D \widetilde{u}_{k,D}$. Then,*

$$\|b - Au_{k,D}\|_2 = \|P_D(b - A\widetilde{u}_{k,D})\|_2 = \min_{\eta \in \mathcal{K}^k(P_D A, P_D b)} \|P_D(b - A\eta)\|_2. \tag{44}$$

Proof For deflation, at the k-th iteration, the approximate solution $u_{k,D}$ gives the residual $r_{k,D} = b - Au_{k,D} = b - AZE^{-1}Y^T b - AQ_D \widetilde{u}_{k,D} = P_D b - P_D A\widetilde{u}_{k,D} = P_D(b - A\widetilde{u}_{k,D})$, which gives the first equality. The second equality is determined from the fact that GMRES applied to $P_D A\widetilde{u} = P_D b$ produces an approximate solution which minimizes the residual norm, i.e.,

$$\|\widetilde{r}_{k,D}\|_2 := \|P_D(b - A\widetilde{u}_{k,D})\|_2 = \min_{\eta \in \mathcal{K}^k(P_D A, P_D b)} \|P_D(b - A\eta)\|_2.$$

This completes the proof. □

Hence, in case of deflation the preconditioned residual norm computed by GMRES is the same as the actual residual norm. In what follows, the term "actual residual" is always used to refer to the residual of the original system.

Theorem 3.3 *Let $u_{k,D}$ and $\widetilde{u}_{k,D}$ be the deflation iterant, with $\widetilde{u}_{0,D} = 0$. With Assumptions 2.2 and 2.5, for every*

$$u_k \in \mathit{span}\{ZE^{-1}Y^T b, Q_D P_D b, Q_D P_D A P_D b, \ldots, Q_D (P_D A)^{k-1} P_D b\}, \tag{45}$$

we have

$$\|b - Au_{k,D}\|_2 = \|P_D(b - Au_{k,D})\|_2 \le \|b - Au_k\|_2. \tag{46}$$

Proof We decompose $u_k \in \mathcal{K}^k(P_B A, P_B b)$ as

$$u_k = \alpha ZE^{-1}Y^T b + Q_D \eta, \quad \eta \in \text{span}\{P_D b, P_D A P_D b, \ldots, (P_D A)^{m-1} P_D b\}$$

The actual residual for u_k can then be determined, i.e.,

$$r_k = b - Au_k = b - \alpha AZE^{-1}Y^Tb - AQ_D\eta. \tag{47}$$

For deflation, $r_{k,D}$ is related to the residual where $\alpha = 1$. Furthermore, $r = b - Au = b - A(ZE^{-1}Y^Tb + Q_Du) = b - AZE^{-1}Y^Tb - P_DAu = 0$. Subtracting this equality from (47) gives

$$r_k = (1-\alpha)AZE^{-1}Y^Tb + P_DA(u-\eta), \tag{48}$$

and thus,

$$\begin{aligned}\|r_k\|_2 &= \|(1-\alpha)AZE^{-1}Y^Tb + P_DA(u-\eta)\|_2\\ &= (1-\alpha)^2\|AZE^{-1}Y^Tb\|_2 + \|P_DA(u-\eta)\|_2\\ &+ (1-\alpha)b^TYE^{-T}Z^TA^TP_DA(u-\eta)\\ &+ (1-\alpha)(u-\eta)^TA^TP_D^TAZE^{-1}Y^Tb.\end{aligned}$$

Consider the factor $A^TP_D^TAZE^{-1}Y^T$ in the last two terms above. Applying Assumptions 2.2 and 2.5, from the proof of Lemma 2.9, we have that

$$P_DA = X\Lambda_DX^{-1}, \qquad \Lambda_D = \begin{bmatrix}0 & 0\\ 0 & \Lambda_{n-r}\end{bmatrix}.$$

Plug this relation into $A^TP_D^TAZE^{-1}Y^T$ we see that

$$\begin{aligned}YE^{-T}Z^TA^TP_DA &= Y\Lambda_r^{-1}(AZ)^TX\Lambda_DX^{-1} = Y\Lambda_r^{-1}\Lambda_rZ^TX\Lambda_DX^{-1},\\ &= YZ^T[Z\,X_{n-r}]\begin{bmatrix}0 & 0\\ 0 & \Lambda_{n-r}\end{bmatrix}X^{-1},\\ &= Y[I\,0]\begin{bmatrix}0 & 0\\ 0 & \Lambda_{n-r}\end{bmatrix}X^{-1},\\ &= 0.\end{aligned}$$

Hence the last two terms vanish. For $\alpha = 1$, the residual is related to deflation, i.e $r_k|_{\alpha=1} = r_{k,D} = b - Au_{k,D}$. Therefore, for all α and η,

$$\|r_{k,D}\|_2 \le (1-\alpha)^2\|AZE^{-1}Y^Tb\|_2 + \|P_DA(u-\eta)\|_2 = \|r_k\|_2. \tag{49}$$

By using Lemma 3.2, we conclude that $\|r_{k,D}\|_2 = \|P_Dr_{k,D}\|_2 \le \|r_k\|_2$. □

Considering GMRES approximation for $P_BAu = P_Bb$, we have the following theorem.

Theorem 3.4 *Under Assumptions 2.2 and 2.5 and with starting vectors $\widetilde{u}_{0,D} = u_{0,B} = 0$, GMRES with deflation and the balancing preconditioner produces actual residuals such that*

$$\|r_{k,D}\|_2 \le \|r_{k,B}\|_2, \tag{50}$$

where $r_{k,D} = b - Au_{k,D}$ and $r_{k,B} = b - Au_{k,B}$.

Proof Suppose that $u_{k,B} \in \mathcal{K}^k(P_BA, P_BA)$ is obtained by GMRES. Then, u_k is the minimizer of the 2-norm of $P_B(b - A\xi)$, for $\xi \in \mathcal{K}^k(P_BA, P_BA)$. So, we can decompose $u_{k,B}$ as $u_{k,B} = \alpha_B ZE^{-1}Y^Tb + Q_D\eta_B$, $\eta_B \in \mathcal{K}(P_DA, P_Db)$. But, from Lemma 3.3, these α_B and η_B will not give actual residual which is smaller than the actual residual of deflation. $\square$

Next we compare the GMRES residuals for deflation and the balancing preconditioner for special starting vectors. We consider $u_{0,B} = ZE^{-1}Y^Tb$ and $\widetilde{u}_{0,D} = 0$. Such a choice of $u_{0,B}$ has particular reasons in terms of implementation. As one notices from the definition of P_B, with a naive implementation, the balancing preconditioner requires two more matrix vector multiplications than deflation. If A is symmetric positive definite, this choice of $u_{0,B}$ greatly simplifies the CG algorithm and reduces the amount of work of the balancing preconditioner to only one matrix/vector multiplication, which is the same as deflation; see [19]. As shown in [13], for spd systems, such starting vectors lead to exactly the same norm of errors of CG iterant. First, we define the Krylov subspace corresponding to P_BA and starting vector $u_{0,B} = ZE^{-1}Y^Tb$.

Lemma 3.5 *With starting vector $u_{0,B} = ZE^{-1}Y^Tb$, the approximate solution $u_{k,B}$ is contained in the subspace*

$$\mathcal{K}^k(P_BA, \widetilde{r}_{0,B}) = ZE^{-1}Y^Tb + Q_D\,span\{P_Db, P_DAP_Db, \dots, (P_DA)^{k-1}P_Db\},$$

where $\widetilde{r}_{0,B} = Q_DP_Db$.

Proof For $k = 1$, $\widetilde{r}_{0,B} = P_B(b - Au_{0,B}) = P_B(b - AZE^{-1}Y^Tb) = Q_DP_Db$. For $k = 2$,

$$P_BA\widetilde{r}_{0,B} = (Q_DP_DA + ZE^{-1}Y^TA)Q_DP_Db = Q_DP_DAP_Db.$$

Continuing for $(P_BA)^l\widetilde{r}_{0,B}$, $l = 3, \dots, k$, we prove the lemma. $\square$

Theorem 3.6 *With any full ranked Z and Y, and $M = I$, GMRES combined with deflation and the balancing preconditioner with starting vectors $\widetilde{u}_{0,D} = 0$ and $u_{0,B} = ZE^{-1}Y^Tb$ generates solutions which satisfy the inequality*

$$\|b - Au_{k,D}\|_2 \le \|b - Au_{k,B}\|_2. \tag{51}$$

Proof In this case, because of Lemma 3.5,

$$r_{k,B} = b - Au_{k,B} = (I - AZE^{-1}Y^T)b + AQ_D\eta = P_D(b - A\eta), \tag{52}$$

where $\eta \in \text{span}\{P_Db, P_DAP_Db, \dots, (P_DA)^{m-1}P_Db\} \equiv \mathcal{K}^k(P_DA, P_Db)$. Suppose that $\zeta \in \mathcal{K}^k(P_DA, P_Db)$ is such that the 2-norm of $P_D(b - A\zeta)$ is minimized. This can be achieved by the action of GMRES on $P_DA\widetilde{u} = P_Db$, with $\widetilde{u}_{0,D} = 0$. So, in this case, $\zeta = \widetilde{u}_{k,D}$. Hence, for any value η, $\|P_D(b - \eta)\|_2 \ge \|P_D(b - A\zeta)\|_2 = \|P_D(b - A\widetilde{u}_{k,D})\|_2$. Because of (52), $\|r_{k,B}\|_2 \ge \|P_D(b - A\zeta)\|_2$. By using Lemma 3.2, we obtain the above theorem. $\square$

Different from results for spd systems, in this case starting vectors $\widetilde{u}_{0,D} = 0$ and $u_{0,B} = ZE^{-1}Y^Tb$ does not necessarily lead to the same GMRES iterations.

3.2 Preconditioned case

Now we consider the case with nonsingular preconditioner M. First we consider the Krylov subspace related to $P_B A$ and starting vector $u_{0,B} = 0$.

Lemma 3.7 *Let P_B be defined by (7). With the starting vector $u_{0,B} = 0$, the approximate solution after k GMRES iterations lies in the subspace*

$$\mathcal{K}^k(P_B A, P_B b) \subseteq \mathit{span}\{ZE^{-1}Y^T b, Q_D M^{-1} P_D b, Q_D(M^{-1}P_D A)M^{-1}P_D b, \\ \ldots, Q_D(M^{-1}P_D A)^{k-1}M^{-1}P_D b\}.$$

Proof The proof is done by similar computations as in Theorem 3.1. □

For preconditioned deflation, $u_{k,D} = ZE^{-1}Y^T b + Q_D \widetilde{u}_{k,D}$, where $\widetilde{u}_{k,D}$ is now iteratively computed from $M^{-1}P_D A\widetilde{u} = M^{-1}P_D b$. In this case, with $\widetilde{u}_{0,D} = 0$, GMRES minimizes the 2-norm of residuals over the subspace

$$\begin{aligned} &\mathcal{K}^k(M^{-1}P_D A, M^{-1}P_D b) \\ &= \mathrm{span}\{M^{-1}P_D b, (M^{-1}P_D A)M^{-1}P_D b, \ldots, (M^{-1}P_D A)^{k-1}M^{-1}P_D b\}. \end{aligned} \tag{53}$$

Considering the subspace $\mathcal{K}^k(P_B A, P_B b)$, by Lemma 3.7 any vector in $\mathcal{K}^k(P_B A, P_B b)$ can be decomposed as $u_k = \alpha ZE^{-1}Y^T b + Q_D \eta$, where

$$\eta \in \mathrm{span}\{M^{-1}P_D b, M^{-1}P_D A M^{-1}P_D b, \ldots, (M^{-1}P_D A)^{k-1}M^{-1}P_D b\}.$$

Hence,

$$\begin{aligned} r_k &\equiv b - Au_k, \\ &= \alpha ZE^{-1}Y^T b + AQ_D\eta, \\ &= \alpha ZE^{-1}Y^T b + P_D A\eta. \end{aligned}$$

Consider the exact solution $u = ZE^{-1}Y^T b + Q_D u$ and the related residual $r = b - Au = (I - ZE^{-1}Y^T)b + P_D Au = 0$. Subtracting this relation from r_k we have

$$r_k = (1-\alpha)AZE^{-1}Y^T b + P_D A(u - \eta), \tag{54}$$

whose 2-norm is the same as in (48). By imposing Assumptions 2.2 and 2.5 we again have

$$\|r_k\|_2 = (1-\alpha)^2\|AZE^{-1}Y^T b\|_2 + \|P_D A(\eta - u)\|_2. \tag{55}$$

Hence,

Theorem 3.8 *With any nonsingular preconditioner M, and under Assumptions 2.2 and 2.5, for every*

$$u_k \in \mathcal{K}^k(P_B A, P_B b), \tag{56}$$

then

$$\|b - Au_{k,D}\|_2 \le \|b - Au_k\|_2. \tag{57}$$

Proof Setting $\alpha = 1$ in (55) for deflation leads to the theorem. □

If $u_{k,B}$ is the approximate solution of $P_BAu = P_Bb$, we then have the following theorem.

Theorem 3.9 *With starting vectors $\widetilde{u}_{k,D} = u_{k,B} = 0$, any nonsingular preconditioner M, and Assumptions 2.2 and 2.5, GMRES combined with deflation and the balancing preconditioner produces residuals which satisfy the inequality*

$$\|b - Au_{k,D}\|_2 \leq \|b - Au_{k,B}\|_2. \tag{58}$$

Proof Clearly, with the balancing preconditioner, GMRES produces solution $u_{k,B} \in \mathcal{K}^k(P_BA, P_Bb)$. Using Theorem 3.8, the above theorem is proved. □

Since GMRES combined with the balancing preconditioner minimizes the 2-norm of residuals $\|P_B(b-A\eta)\|_2$, this minimization property does not necessarily hold for the 2-norm of the actual residual $\|b - Au_{k,B}\|_2$. This is also the case for deflation, which GMRES minimizes $\|M^{-1}P_D(b - A\eta)\|$ over the subspace $\mathcal{K}^k(M^{-1}P_DA, M^{-1}P_Db)$. However, for deflation we have the following lemma.

Lemma 3.10 *For any Z, Y and nonsingular preconditioning matrix M, residuals related to GMRES combined with deflation and $\widetilde{u}_{0,D} = 0$ satisfy the equality*

$$\begin{aligned} \|M^{-1}r_{k,D}\|_2 &= \|M^{-1}P_D(b - A\widetilde{u}_{k,D})\|_2, \\ &= \min_{\eta \in \mathcal{K}^k(M^{-1}P_DA, M^{-1}P_Db)} \|M^{-1}P_D(b - A\eta)\|_2. \end{aligned} \tag{59}$$

Proof By construction, with $\widetilde{u}_{0,D} = 0$,

$$r_{k,D} =: b - Au_{k,D} = b - A(ZE^{-1}Y^Tb + Q_D\widetilde{u}_{k,D}) = P_D(b - A\widetilde{u}_{k,D}),$$

where $\widetilde{u}_{k,D} \in \mathcal{K}^k(M^{-1}P_DA, M^{-1}P_Db)$. Premultiplying the above equality by M^{-1}, we have

$$\begin{aligned} \|M^{-1}r_{k,D}\|_2 &= \|M^{-1}P_D(b - A\widetilde{u}_{k,D})\|_2 \\ &= \min_{\eta \in \mathcal{K}^k(M^{-1}P_DA, M^{-1}P_Db)} \|M^{-1}P_D(b - A\eta)\|_2. \end{aligned}$$

Clearly, optimality property only holds for the preconditioned residuals. □

Finally, we consider the balancing preconditioner with starting vector $u_{0,B} = ZE^{-1}Y^Tb$ and deflation with $\widetilde{u}_{0,D} = 0$.

Lemma 3.11 *With starting vector $u_{0,B} = ZE^{-1}Y^Tb$ and any nonsingular matrix M, GMRES combined with the balancing preconditioner produces solution at the k-th iteration which lies in the subspace $ZE^{-1}Y^Tb + \mathcal{K}^k(P_BA, \widetilde{r}_{0,B})$, where*

$$\begin{aligned} &\mathcal{K}^k(P_BA, \widetilde{r}_{0,B}) = \\ &Q_D\,span\{M^{-1}P_Db, M^{-1}P_DAM^{-1}P_Db, \ldots, (M^{-1}P_DA)^{k-1}M^{-1}P_Db\}, \end{aligned} \tag{60}$$

where $\widetilde{r}_{0,B} = Q_DM^{-1}P_Db$.

Proof For $k = 1$, $\widetilde{r}_{0,B} = P_B(b - Au_{0,B}) = Q_D M^{-1} P_D b$. The proof is done by recursive computations similar to the proof of Lemma 3.1 for $(P_B A)^l \widetilde{r}_{0,B}$, $l = 2, \dots, k$. □

Theorem 3.12 *With any full ranked matrices Z and Y, any nonsingular preconditioning matrix M, and starting vector $u_{0,B} = ZE^{-1}Y^T b$ and $\widetilde{u}_{0,D} = 0$, GMRES combined with deflation and the balancing preconditioner produces solutions whose residuals satisfy the inequality*

$$\|M^{-1}(b - Au_{k,D})\|_2 \leq \|M^{-1}(b - Au_{k,B})\|_2. \tag{61}$$

Proof For the balancing preconditioner, the solution of GMRES at the k-th iteration is

$$u_{k,B} \in ZE^{-1}Y^T b + \mathcal{K}^k(P_B A, \widetilde{r}_{0,B}), \quad \widetilde{r}_{0,B} = P_B(b - A\widetilde{u}_{0,B}), \tag{62}$$

or, by Lemma 3.11,

$$u_{k,B} = ZE^{-1}Y^T b + Q_D \eta, \qquad \eta \in \mathcal{K}^k(M^{-1}P_D A, M^{-1}P_D b). \tag{63}$$

whose residual is $r_{k,B} = b - Au_{k,B} = P_D(b - A\eta)$. Thus, $M^{-1}r_{k,B} = M^{-1}P_D(b - \eta)$, implying $\|M^{-1}r_{k,B}\|_2 = \|M^{-1}P_D(b - \eta)\|_2$. If ζ minimizes $\|M^{-1}P_D(b - A\zeta)\|_2$ over the subspace $\mathcal{K}^k(M^{-1}P_D A, M^{-1}P_D b)$, then

$$\|M^{-1}P_D(b - \eta)\|_2 \geq \min_{\zeta \in \mathcal{K}^k(M^{-1}P_D A, M^{-1}P_D b)} \|M^{-1}P_D(b - A\zeta)\|_2. \tag{64}$$

In this case, ζ can be obtained by GMRES applied to $M^{-1}P_D A\widetilde{u} = M^{-1}P_D b$ with zero starting vector. Thus, $\zeta = \widetilde{u}_{k,D}$. By using Lemma 3.10, we finally get

$$\|M^{-1}r_{k,B}\|_2 := \|M^{-1}P_D(b - \eta)\|_2 \geq \|M^{-1}P_D(b - A\widetilde{u}_{k,D})\|_2 =: \|M^{-1}r_{k,D}\|_2,$$

where $r_{k,D} = b - Au_{k,D}$ and $r_{k,B} = b - Au_{k,B}$. □

Since the abstract balancing preconditioner is nonsingular, GMRES will converge. Thus, Theorem 3.12 (and as well Theorem 3.6) guarantees that GMRES preconditioned by the singular deflation preconditioner will also converge without any further assumption.

In case with preconditioner M a similar result as Theorem 3.12 can not be obtained for the actual residual. In fact, as is shown in numerical results in the next section, the 2-norm of the actual residual of GMRES combined with the balancing preconditioner is not necessarily smaller than the 2-norm of the actual residual related to deflation.

4 Numerical examples

In this section we perform numerical experiments to confirm our theoretical results. We base our numerical experiments on the linear systems arising from finite volume discretization of the steady-state convection-diffusion equation

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} - \nabla\left(\frac{1}{Pe}\nabla u\right) = f, \qquad \text{in } \Omega = (0,1)^2 \tag{65}$$

with $Pe > 0$ the Péclet number, and f the forcing term. We first consider the one-dimensional version of (65), in which eigenvalues and eigenvectors of the corresponding linear system can be computed cheaply. Later in this section, a 2D convection-diffusion problem is also discussed.

For 1D convection-diffusion problem, we consider an artifical 1D problem with jump in Pe. Problems with jump in Pe lead to linear systems with very large difference between the largest and the smallest eigenvalue (in magnitude). In our case, we set

$$Pe(x) = \begin{cases} 1, & 0 \leq x < 0.8, \\ 200, & 0.8 \leq x \leq 1. \end{cases} \tag{66}$$

The boundary conditions are $u(0) = 0$ and $u(1) = 1$, which resembles extremely thin boundary layer flows near $x = 1$, and $f = 0$. The 1D convection-diffusion equation is discretized by using the cell-centered finite volume discretization. The convective flux term is approximated by the central discretization. In order to avoid wiggly numerical solutions, the grid is refined at the vicinity of $x = 1$, keeping the mesh Péclet number less than 2 for stability reason. (At this moment we are, however, not concerned with the accuracy of the approximate solutions, and are interested more on the validity of the theoretical results for a specific problem.) In the subdomain where $Pe = 1$, 40 cells are used. In total 200 cells are used. The resultant linear system has real and simple eigenvalues, with 18 of them having value less than one. The largest eigenvalue is ~ 399.9, while the smallest one is ~ 0.12, giving a ratio of an order of 10^3.

Figure 1 shows convergence of GMRES measured by the 2-norm of relative residuals (left) and by the 2-norm of actual residuals (right), where Z and Y are the eigenvectors matrices and zero starting vector is used. For number of deflation vectors less than 20, preconditioning by deflation and the balancing preconditioner leads to almost identical GMRES convergence. Deflation becomes significantly more superior than the balancing preconditioner as 20 eigenvectors are used (or 20 smallest eigenvalues are deflated). In this case, the effective smallest eigenvalue for deflated system is ~ 1.30, while for the balancing preconditioner is 1. The right figure shows that the 2-norm of the actual residual of deflation is always smaller than that of the balancing preconditioner. Figure 2 shows the preconditioned residuals and the actual residuals in 2-norm of deflation preconditioner. As clearly seen, the two are the same.

Next, special starting vector is used for the balancing preconditioner, i.e. $u_{0,B} = ZE^{-1}Y^Tb$. The convergence history is shown in Figure 3 for different numbers of deflation vectors. Again, eigenvectors are used in Z and Y. In this, this choice of starting vectors leads to almost identical GMRES convergence except for the last few iterations, where deflation produces smaller two-norm of actual residuals than the balancing preconditioner.

The last set of numerical tests from the 1D convection-diffusion equation is given for arbitrary Z and Y. Here, we choose Z based on subdomain structuring proposed in [14] and used in [7]. Suppose that the domain Ω with index set $\mathcal{I} = \{i | u_i \in \Omega\}$ is partitioned into m non-overlapping subdomain Ω_j, $j = 1, \ldots, m$,

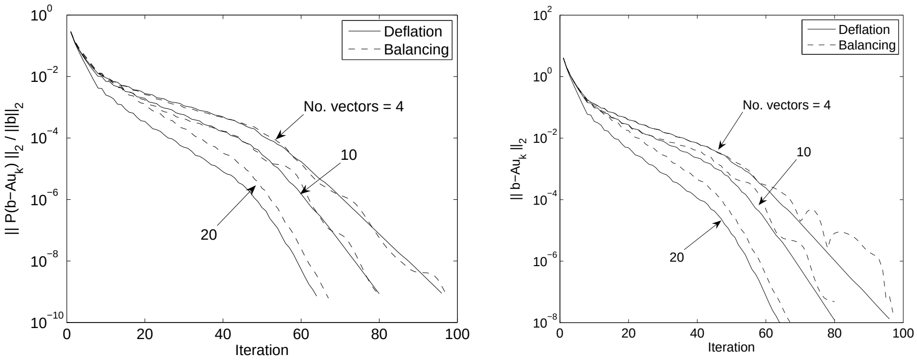

Figure 1: One-dimensional convection-diffusion equation with jumps. Shown are residuals of the preconditioned system with zero starting vectors, $M = I$, and Z and Y consisting of eigenvectors of A. Left: preconditioned relative residuals. Right: actual residuals.

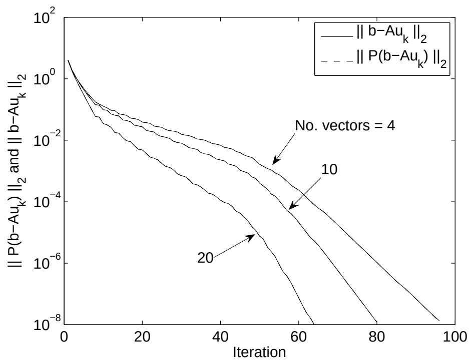

Figure 2: One-dimensional convection-diffusion equation with jumps. Shown are preconditioned residuals and the actual residuals with zero starting vector, $M = I$, and Z and Y consisting of eigenvectors of A

with respective index $\mathcal{I}_j = \{i \in \mathcal{I} | u_i \in \Omega_j\}$. Then, Z is defined by

$$z_{ij} = \begin{cases} 1, & i \in \mathcal{I}_j, \\ 0, & i \notin \mathcal{I}_j, \end{cases} \tag{67}$$

and Y is set equal to Z; $Y = Z$. Particularly in this example we first partition Ω in to two subdomains of the same Péclet number. Based on this partition, partitioning is done further until the number of deflation vectors needed is reached.

Figure 4 shows GMRES convergence with zero starting vector for both preconditioners. In this case, GMRES combined with deflation converges faster

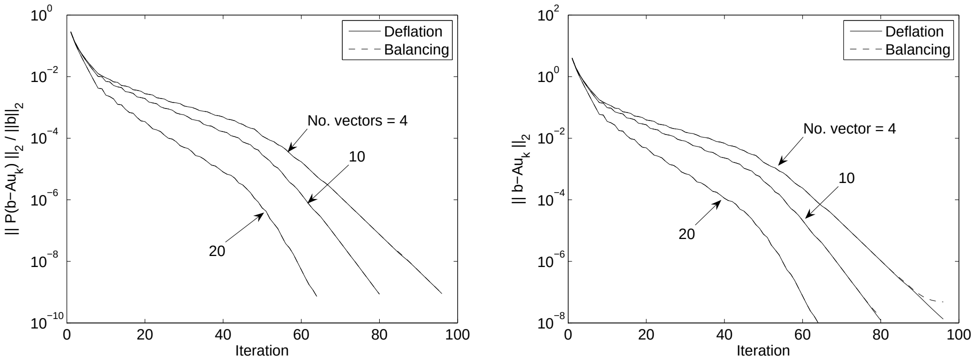

Figure 3: One-dimensional convection-diffusion equation with jumps. Shown are residuals of the preconditioned system with zero starting vector for deflation and $u_{0,B} = ZE^{-1}Y^Tb$ for the balancing preconditioner, $M = I$, and Z and Y consisting of eigenvectors of A. Left: preconditioned relative residuals. Right: actual residuals.

than GMRES with the balancing preconditioner. The right figure of Figure 4 also indicates that the norm of the actual residuals of deflation is always smaller than that of the balancing preconditioner. Similar results are also obtained for $u_{0,B} = ZE^{-1}Y^Tb$ the starting vector in case of the balancing preconditioner. Even though both preconditioners result in almost identical convergence, clearly deflation still produces smaller norms of residuals compared to the balancing preconditioner.

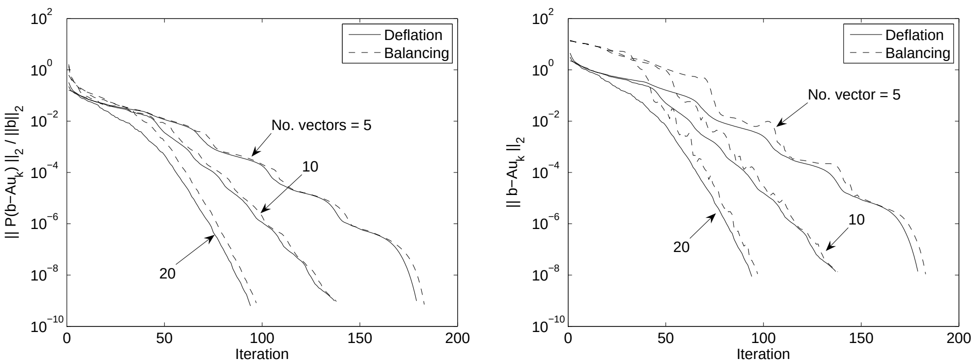

Figure 4: One-dimensional convection-diffusion equation with jumps. Shown are residuals of the preconditioned system with zero starting vector, $M = I$, and Z and Y are arbitrary. Left: preconditioned relative residuals. Right: actual residuals.

Figures 6 and 7 show convergence if a preconditioner M is also incorpo-

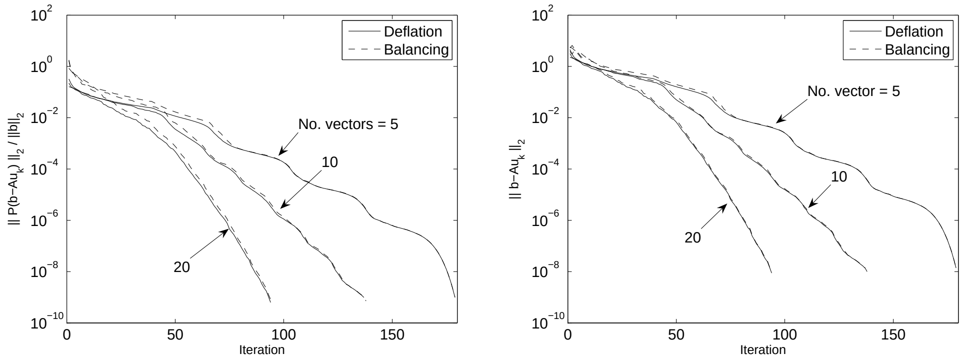

Figure 5: One-dimensional convection-diffusion equation with jumps. Shown are residuals of the preconditioned system with zero starting vector for deflation and $u_{0,B} = ZE^{-1}Y^Tb$ for the balancing preconditioner, $M = I$, and Z and Y consisting of eigenvectors of A. Left: preconditioned relative residuals. Right: actual residuals.

rated. In this case M is preconditioning matrix based on the diagonal scaling. In Figure 6, zero starting vectors are used, while in Figure 7 a special starting vector is used for the balancing preconditiner. The convergence for both preconditioners are similar, with deflation generating preconditioned residual (i.e., $M^{-1}r_k$) whose 2-norm is smaller than the balancing preconditioner. For the actual residual, this conclusion does not necessarily hold (right figures). There, we observe at some steps that the balancing preconditioner produces smaller 2-norm of the actual residuals than deflation.

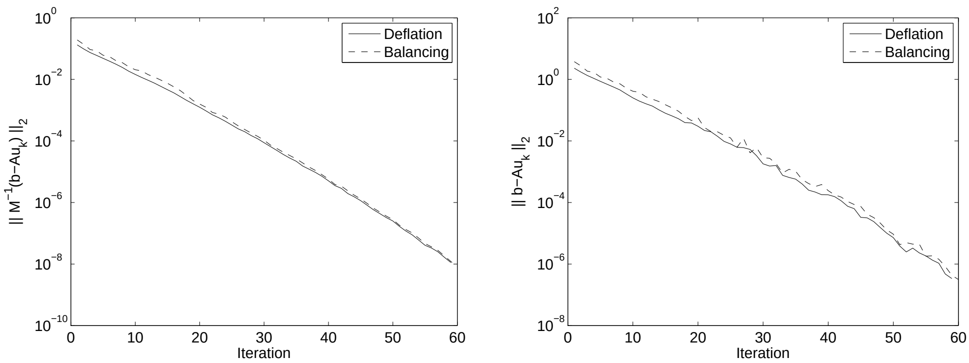

Figure 6: One-dimensional convection-diffusion equation with jumps. Shown are residuals of the preconditioned system with $\widetilde{u}_{0,D} = u_{0,B} = 0$, M the diagonal scaling preconditioner, and Z and Y as in (67). Left: preconditioned residuals (based on M). Right: actual residuals.

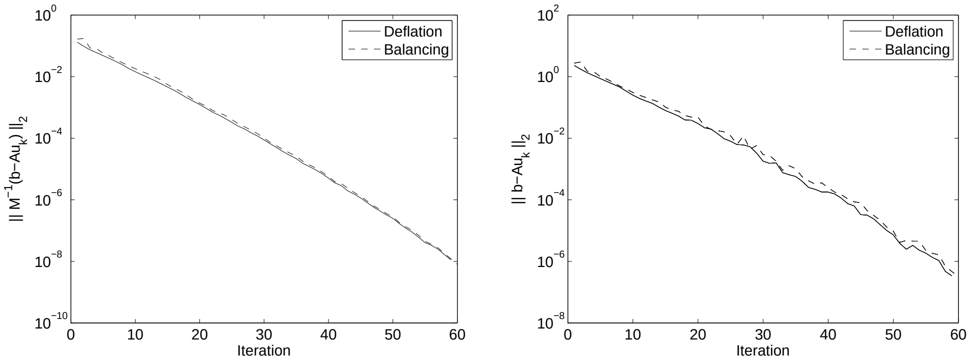

Figure 7: One-dimensional convection-diffusion equation with jumps. Shown are residuals of the preconditioned system with $\widetilde{u}_{0,D} = 0$ and $u_{0,B} = ZE^{-1}Y^Tb$, M the diagonal scaling preconditioner, and Z and Y as in (67). Left: preconditioned residuals (based on M). Right: actual residuals.

For 2D case, we consider convection-diffusion problem with Dirichlet boundary conditions at $x = 1$ and $y = 0$, and homogeneous Neumann boundary conditions at $x = 0$ and $y = 1$; ([21]). The equation is discretized by using finite volume discretization. The convection flux term is approximated by upwind scheme. The Péclet number is 200. The grid is refined in the y-direction in the vicinity of $y = 0$, while in the x-direction the grid size is kept constant. The domain is partitioned into 10×10 subdomains. Deflation vectors are constructed based on (67). Figure 8 shows convergence results for starting vector $\widetilde{u}_{0,D} = 0$ and $u_{0,B} = ZE^{-1}Y^Tb$. In this case, residuals related to deflation and the balancing preconditioners are very similar. Figure 9 (left) shows that the preconditioned residual (based on M) of the balancing preconditioner is never smaller than that of deflation, the actual residual. This is, however, not always the case for the actual residual (Figure 9 (right)).

5 Conclusion

In this paper a comparison between deflation and the balancing preconditioner for nonsymmetric linear systems has been given, within the context of GMRES. Analysis shows that with special starting vectors, deflation generates approximate solutions whose related residuals (with respect to M) are never larger than the balancing preconditioner. If the deflation vectors are chosen to be the eigenvectors, the 2-norm of the actual residuals of deflation is always never larger than the balancing preconditioner. For general deflation vectors we proved that the 2-norm of preconditioned residual of GMRES combined with deflation is smaller than that of GMRES combined with the balancing preconditioner. Numerical experiments confirm the theoretical results. The two examples presented in this

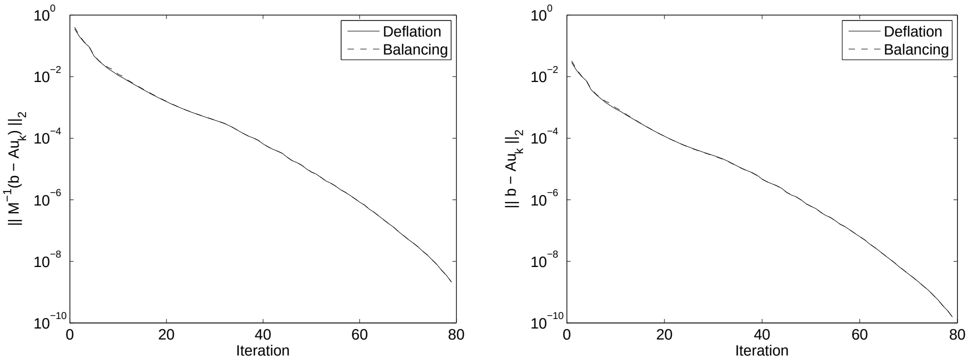

Figure 8: Two-dimensional convection-diffusion equation with constant coefficient ($Pe = 200$). Shown are residuals of the preconditioned system with $\widetilde{u}_{0,D} = 0$ and $u_{0,B} = ZE^{-1}Y^Tb$, M the diagonal scaling preconditioner, and Z and Y as in (67). Left: preconditioned residuals (based on M). Right: actual residuals.

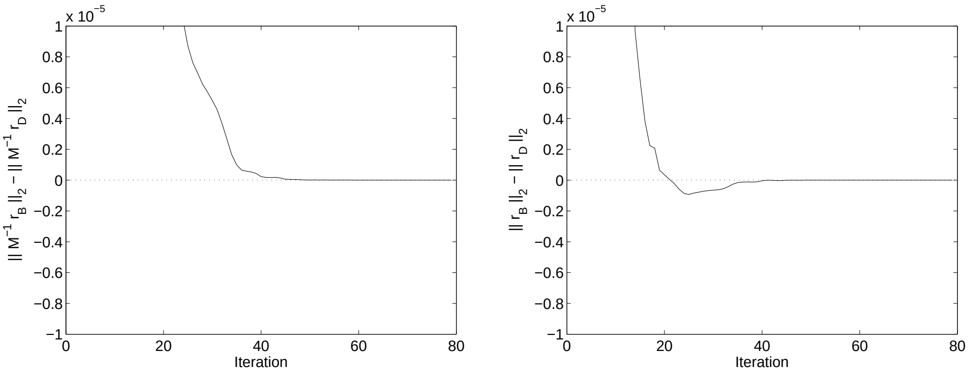

Figure 9: Residual difference from the two-dimensional convection-diffusion equation with constant coefficient ($Pe = 200$). The starting vectors are $\widetilde{u}_{0,D} = 0$ and $u_{0,B} = ZE^{-1}Y^Tb$, M the diagonal scaling preconditioner, and Z and Y as in (67). Left: preconditioned residuals (based on M). Right: actual residuals.

paper suggest that the difference in GMRES iterants for deflation and the balancing preconditioner case becomes more significant if jumps or discontinuities are present.

References

[1] W. E. Arnoldi. The principle of minimized iterations in the solution of the matrix eigenvalue problem. *Quart. Appl. Math.*, 9:17–29, 1951.

[2] P. N. Brown and H. F. Walker. GMRES on (nearly) singular systems. *SIAM J. Matrix Anal. Appl.*, 18:37–51, 1997.

[3] A. Chapman and Y. Saad. Deflated and augmented Krylov subspace techniques. *Numer. Linear Algebra Appl.*, 4:43–66, 1997.

[4] M. Dryja and O. B. Widlund. Schwarz methods of Neumann-Neumann type for three-dimensional eliptic finite element problems. *Comm. Pure Appl. Math.*, 48:121–155, 1995.

[5] M. Eiermann, O. G. Ernst, and O. Schneider. Analysis of acceleration strategies for restarted minimal residual methods. *J. Comput. Appl. Math.*, 123:261–292, 2000.

[6] J. Erhel, K. Burrage, and B. Pohl. Restarted GMRES preconditioned by deflation. *J. Comput. Applied Math.*, 69:261–292, 1996.

[7] J. Frank and C. Vuik. On the construction of deflation-based preconditioners. *SIAM J. Sci. Comput.*, 23:442–462, 2001.

[8] S. A. Kharchenko and A. Y. Yerenim. Eigenvalue translation based preconditioners for the GMRES(k) method. *Numer. Linear Algebra Appl.*, 2:51–77, 1995.

[9] J. Mandel. Balancing domain decomposition. *Comm. Numer. Methods Engrg.*, 9:233–241, 1993.

[10] J. Mandel and M. Brezina. Balancing domain decomposition for problems with large jumps in coefficients. *Math. Comp.*, 216:1387–1401, 1996.

[11] R. B. Morgan. A restarted GMRES method augmented with eigenvectors. *SIAM J. Matrix Anal. Appl.*, 16:1154–1171, 1995.

[12] R. Nabben and C. Vuik. A comparison of deflation and coarse grid correction applied to porous media flow. *SIAM J. Numer. Anal.*, 42:1631–1647, 2004.

[13] R. Nabben and C. Vuik. A comparison of deflation and the balancing preconditioner. *SIAM J. Sci. Comput.*, 27:1742–1759, 2006.

[14] R. A. Nicolaides. Deflation of conjugate gradients with applications to boundary value problems. *SIAM J. Numer. Anal.*, 24:355–365, 1987.

[15] A. Padiy, O. Axelsson, and B. Polman. Generalized augmented matrix preconditioning approach and its application to iterative solution of ill-conditioned algebraic systems. *SIAM J. Matrix Anal. Appl.*, 22:793–818, 2000.

[16] L. F. Pavarino and O. B. Widlund. Balancing Neumann-Neumann methods for incomprible stokes equations. *Comm. Pure Appl. Math.*, 55:302–335, 2002.

[17] A. Quarteroni and A. Valli. *Domain Decomposition Methods for Partial Differential Equations*. Oxford Science Publications, Oxford, UK, 1999.

[18] Y. Saad and M. H. Schultz. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 7:856–869, 1986.

[19] A. Tosseli and O. Widlund. *Domain decomposition methods*. Springer-Verlag, Berlin, 2005.

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

[21] P. Wesseling. *Principles of Computational Fluids Dynamics*. Springer-Verlag, Berlin Heilderberg, 2001.