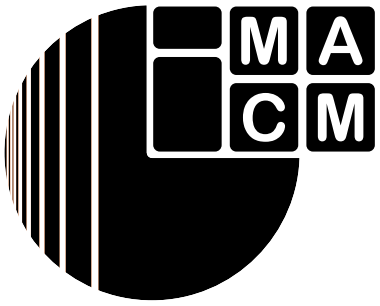

Bergische Universität Wuppertal

Fachbereich Mathematik und Naturwissenschaften

Institute of Mathematical Modelling, Analysis and Computational Mathematics (IMACM)

Preprint BUW-IMACM 16/27

K. Kahl, H. Rittich

The Deflated Conjugate Gradient Method: Convergence, Perturbation and Accuracy

October 2016

http://www.math.uni-wuppertal.de

The Deflated Conjugate Gradient Method: Convergence, Perturbation and Accuracy

K. Kahl*, H. Rittich*

October 17, 2016

Deflation techniques for Krylov subspace methods have seen a lot of attention in recent years. They provide means to improve the convergence speed of these methods by enriching the Krylov subspace with a deflation subspace. The most common approach for the construction of deflation subspaces is to use (approximate) eigenvectors, but also more general subspaces are applicable.

In this paper we discuss two results concerning the accuracy requirements within the deflated CG method. First we show that the effective condition number which bounds the convergence rate of the deflated conjugate gradient method depends asymptotically linearly on the size of the perturbations in the deflation subspace. Second, we discuss the accuracy required in calculating the deflating projection. This is crucial concerning the overall convergence of the method, and also allows to save some computational work.

To show these results, we use the fact that as a projection approach deflation has many similarities to multigrid methods. In particular, recent results relate the spectra of the deflated matrix to the spectra of the error propagator of twogrid methods. In the spirit of these results we show that the effective condition number can be bounded by the constant of a weak approximation property.

Keywords

conjugate gradients, deflation, multigrid, convergence, perturbation

2008 MSC: 65F08, 65F10

1 Introduction

Consider solving the linear system of equations

$$Ax = b, \tag{1}$$

*Fakultät für Mathematik und Naturwissenschaften, Bergische Universität Wuppertal, 42097 Wuppertal, Germany
e-mail: kkahl@math.uni-wuppertal.de (K. Kahl), rittich@math.uni-wuppertal.de (H. Rittich)

where $A \in \mathbb{K}^{n\times n}$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$) is self-adjoint and positive definite and $x, b \in \mathbb{K}^n$. In this paper we are interested in the case where the matrix A is large and sparse. The conjugate gradient (CG) method [14, 16, 27] is an iterative method which is often well suited to solve these systems. The speed of convergence of the CG method depends on the distribution of its eigenvalues and the right hand side. Estimates for the speed of convergence in terms of the condition of the matrix A exist [27, 39]. When the condition number κ is large it can become mandatory to precondition the linear system such that a satisfactory speed of convergence can be guaranteed.

One possibility to precondition the CG method is via deflation as introduced by Nicolaides [22] and Dostal [8] (see also [2, 7, 9, 13, 20, 28, 31]). The basic idea of deflation is to "hide" certain parts of the spectrum of the matrix A from the CG method itself, such that the CG iteration "sees" a system that has a much smaller condition number than A. The part of the spectrum that is hidden from CG is determined by the *deflation subspace* $\mathcal{S} \subseteq \mathbb{K}^n$ and the improvement of the convergence rate of the deflated CG method hinges solely on the choice of $\mathcal{S}$.

One viable and widely used approach for deflation consists of spanning $\mathcal{S}$ by the eigenvectors corresponding to the smallest eigenvalues [28]. This then hides the smallest eigenvalues from the spectrum of A. On the other hand, already in [22] a different choice of $\mathcal{S}$ has been used and further examples for other deflation subspaces can be found in [9, 20].

Different two-level approaches, including deflation and twogrid methods [6, 15, 37], have been compared in [21, 36] and equivalences between these apparently different methods have been established. In continuation of these efforts to connect different two-level approaches, we show how to use theory developed for multigrid methods to analyze the deflated CG method.

After giving a short introduction to the deflated CG method in Section 2 we give an overview on how to apply multigrid theory in the deflation context in Section 3.

Based on these theoretical considerations we answer two different questions concerning accuracy that arise in the deflation approach.

First, in Section 4 we analyze the situation where the deflation subspace $\mathcal{S}$ is only known up to a perturbation and give a bound on the effective condition number with respect to the size of the perturbation. This result is of particular interest in the case where the deflation subspace is spanned by eigenvectors corresponding to the smallest eigenvalues of A. In many practical applications eigenvectors are unknown and need to be approximated and thus the resulting deflation subspace can be thought of as a perturbation of the deflation subspace that uses the exact eigenvectors. This result can in particular be applied to the analysis of methods like eigCG [32] and similar methods [2, 7, 31] where the eigenvectors are approximated numerically.

Second, we consider the situation when the dimension of $\mathcal{S}$ is large. In this case special care is needed within a deflation type method as the deflating projection now involves the solution of a large linear system with the matrix V^*AV, where the columns of V form a basis of $\mathcal{S}$. We discuss the accuracy requirements on the solution of this system to ensure proper convergence of the deflated CG method based on results from [29, 38] on inexact Krylov subspace methods. This reveals a way to reduce computational work needed to perform the deflated CG method.

Some numerical experiments confirming and illustrating the theory are given in Section 6.

We conclude this introduction by fixing basic notation used throughout the paper. For $z \in \mathbb{K}^n$ its residual $r \in \mathbb{K}^n$ is given by $r = b - Az$, the error by $e = x - z$ where x is the solution of (1). Note that $Ae = r$.

Let $\langle v, w\rangle = \sum_{i=1}^{n} \overline{w_i}\, v_i$ be the euclidean inner product of v and w, $\|v\|_2 = \langle v, v\rangle^{1/2}$ the euclidean norm of v. Since A is self-adjoint and positive definite the A-inner product and the A-norm exist and are given by

$$\langle v, w\rangle_A := \langle Av, w\rangle \quad \text{and} \quad \|v\|_A := \langle v, v\rangle_A^{1/2}\,.$$

Let $q_1, q_2, \ldots, q_n$ be an orthonormal basis of eigenvectors of the matrix A, s.t. $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are the corresponding eigenvalues.

2 Review of Deflated CG

Let x_0 be an initial guess, $r_0 = b - Ax_0$ and the ith Krylov subspace be denoted by $\mathcal{K}_i(A, r_0) := \operatorname{span}\{r_0, Ar_0, A^2r_0, \ldots, A^{i-1}r_0\}$. The i-th CG iterate is determined such that $x_i \in x_0 + \mathcal{K}_i(A, r_0)$ and the error $e_i = x - x_i$ is minimized in the A-norm (cf. [27]), i.e.,

$$\|e_i\|_A = \min\{\|x - z\|_A : z \in x_0 + \mathcal{K}_i(A, r_0)\}\,. \tag{2}$$

Note that (2) is just the distance in the A-norm between x and the affine subspace $x_0 + \mathcal{K}_i(A, r_0)$. The convergence of the method can be slow in case of unfavorable spectral properties of the matrix A [39]. The idea of deflation is to modify the CG method such that the iterates x_i are equal to the solution x on a given subspace $\mathcal{S} \subseteq \mathbb{K}^n$ in the sense that the A-orthogonal projections of x_i and x onto $\mathcal{S}$ coincide (and are thus identical for all i). For proper choices of $\mathcal{S}$ this will improve the speed of convergence since the affine subspace which contains the iterates may now be much closer to the solution x than the original one. We give a rigorous description of the deflated CG method in the remainder of this section.

Let $\mathcal{S}^{\perp_A} = (A\mathcal{S})^{\perp}$ be the A-orthogonal complement of a given subspace $\mathcal{S} \subseteq \mathbb{K}^n$. We can split the solution x into a component in $\mathcal{S}$ and a component in $\mathcal{S}^{\perp_A}$ via the A-orthogonal projection $\pi_A(\mathcal{S}) \in \mathbb{K}^{n\times n}$ onto $\mathcal{S}$, i.e.,

$$x = (I - \pi_A(\mathcal{S}))x + \pi_A(\mathcal{S})x\,. \tag{3}$$

Let $V \in \mathbb{K}^{n\times m}$ be a matrix such that its columns form a basis of the subspace $\mathcal{S}$. Since

$$\pi_A(\mathcal{S})x = V(V^*AV)^{-1}V^*Ax = V(V^*AV)^{-1}V^*b \tag{4}$$

we can compute $\pi_A(\mathcal{S})x$—the second term in the right hand side of (3)—without explicit knowledge of x. The first term of (3) can be computed from a solution $\hat{x}$ of the singular linear system

$$A(I - \pi_A(\mathcal{S}))\hat{x} = (I - \pi_A(\mathcal{S}))^*b, \tag{5}$$

which we call the *deflated (linear) system*. For the sake of completeness we prove this in the following lemma. Its statements can be found, e.g., in [9, 20].

Lemma 2.1. *Using the definitions from above we have:*

(i) The following equalities hold

$$A(I - \pi_A(\mathcal{S})) = (I - \pi_A(\mathcal{S}))^*A = (I - \pi_A(\mathcal{S}))^*A(I - \pi_A(\mathcal{S}))\,. \tag{6}$$

(ii) *The matrix $A(I - \pi_A(\mathcal{S}))$ is self-adjoint and positive semi-definite.*

(iii) *The deflated system (5) is consistent, i.e., the right hand side $(I-\pi_A(\mathcal{S}))^* b$ is in the range of $A(I - \pi_A(\mathcal{S}))$. This implies that the system has at least one solution.*

(iv) *If $\hat{x}$ is a solution of the deflated system (5) then*

$$(I - \pi_A(\mathcal{S}))\hat{x} = (I - \pi_A(\mathcal{S}))x\,, \tag{7}$$

where x is the solution of the linear system $Ax = b$.

Proof. Since $\pi_A(\mathcal{S}) = V(V^*AV)^{-1}V^*A$ we have

$$A(I - \pi_A(\mathcal{S})) = A - AV(V^*AV)^{-1}V^*A = (I - \pi_A(\mathcal{S}))^*A\,,$$

and since $I - \pi_A(\mathcal{S})$ is a projection we also get

$$A(I - \pi_A(\mathcal{S})) = A(I - \pi_A(\mathcal{S}))(I - \pi_A(\mathcal{S})) = (I - \pi_A(\mathcal{S}))^*A(I - \pi_A(\mathcal{S}))\,.$$

This proves (i) and also shows that $A(I-\pi_A(\mathcal{S}))$ is self-adjoint. Using (6) and due to $\langle B^*ABx, x\rangle = \langle ABx, Bx\rangle = \langle Bx, Bx\rangle_A$, $B \in \mathbb{K}^{n\times n}$ we have

$$\langle A(I - \pi_A(\mathcal{S}))x, x\rangle = \langle (I - \pi_A(\mathcal{S}))x, (I - \pi_A(\mathcal{S}))x\rangle_A = \|(I - \pi_A(\mathcal{S}))x\|_A^2 \geq 0$$

which gives (ii).

Again due to (6) and the fact that A has full rank we have

$$\operatorname{range}\big(A(I - \pi_A(\mathcal{S}))\big) = \operatorname{range}\big((I - \pi_A(\mathcal{S}))^*A\big) = \operatorname{range}\big((I - \pi_A(\mathcal{S}))^*\big)\,.$$

Hence the system (5) is consistent which proves (iii). To show (iv) we use (5) and (6) yielding

$$A(I - \pi_A(\mathcal{S}))\hat{x} = (I - \pi_A(\mathcal{S}))^*b = (I - \pi_A(\mathcal{S}))^*Ax = A(I - \pi_A(\mathcal{S}))x\,.$$

Multiplying with A^{-1} from the left we obtain (7). □

To sum up, given a solution $\hat{x}$ for the deflated system (5) we can compute the first part of the splitting (3) by (7) and the second part by using the formula (4). Therefore we obtain the solution x for the original system as

$$\begin{aligned} x &= (I - \pi_A(\mathcal{S}))x + \pi_A(\mathcal{S})x \\ &= (I - \pi_A(\mathcal{S}))\hat{x} + V(V^*AV)^{-1}V^*b\,. \end{aligned}$$

This relation allows us to compute approximations to the solution of the original system by computing an approximation to the solution $\hat{x}$ of the deflated system (5). Since $A(I - \pi_A(\mathcal{S}))$ is positive semi-definite we can apply the CG method. The fact that the matrix is singular is no impediment to the standard CG iteration as long as (5) is consistent (cf. [17]), which has been shown to be the case in Lemma 2.1.

For the purpose of analyzing the method we think of deflated CG as applying the standard CG algorithm to the deflated system (5) with the matrix $A(I - \pi_A(\mathcal{S}))$. There are various other mathematically equivalent formulations of deflated CG (for an overview see [13]) for which our analysis holds as well.

Let $\mu_1 \geq \cdots \geq \mu_n \geq 0$ be the eigenvalues of the self-adjoint and positive semi-definite matrix $A(I - \pi_A(\mathcal{S}))$. Let $k \in \mathbb{N}$ denote the largest index such that $\mu_k \neq 0$. The errors of the CG iterates then satisfy

$$\|e_i\|_A \leq 2 \left(\frac{\sqrt{\kappa_{\text{eff}}} - 1}{\sqrt{\kappa_{\text{eff}}} + 1} \right)^i \|e_0\|_A \quad \text{for } i = 0, 1, 2, \ldots, \tag{8}$$

where $\kappa_{\text{eff}} = \frac{\mu_1}{\mu_k}$, see [9, 28]. We call κ_{eff} the effective condition number of the deflated matrix $A(I - \pi_A(\mathcal{S}))$ to distinguish it from the condition number κ of the original matrix A. Thus a bound on the convergence rate of deflated CG can be obtained by estimating the largest and smallest *non-zero* eigenvalue of the matrix $A(I - \pi_A(\mathcal{S}))$.

3 Convergence Analysis

We estimate the effective condition number κ_{eff} of the matrix $A(I - \pi_A(\mathcal{S}))$ in terms of a quantity which arises in the *weak approximation property* used in multigrid theory.

Definition 3.1. *A subspace $\mathcal{S} \subseteq \mathbb{K}^n$ fulfills the* weak approximation property with constant *$K \geq 0$ if K is the* smallest *number s.t.*

$$\|x - \pi(\mathcal{S})x\|_2^2 \leq \frac{K}{\|A\|_2} \|x\|_A^2 \quad \textit{for all } x \in \mathbb{K}^n \,. \tag{9}$$

Here, $\pi(\mathcal{S})$ denotes the ℓ_2-orthogonal projection onto $\mathcal{S}$.

Note that $\|x - \pi(\mathcal{S})x\|_2$ is the *ℓ_2-distance between x and the subspace $\mathcal{S}$* defined by

$$\operatorname{dist}(\mathcal{S}, x)_2 = \min_{y \in \mathcal{S}} \|x - y\|_2 \,. \tag{10}$$

Up to a scaling by the diagonal entries of A this definition coincides with the definition of the weak approximation property found in the multigrid literature (see, e.g., [4, 5, 26, 35]). It is called "weak" because it is only sufficient for a two-level convergence theory [26, Section 4.5] but not for a multilevel one. Note that any subspace $\mathcal{S}$ fulfills a weak approximation property, however K may be large. In order to guarantee fast twogrid convergence one is interested in subspaces which admit a small value for K. Uniform bounds for K exist for particular families of matrices and certain subspaces and can be directly derived from the results in, e.g., [4, 5, 26, 35]. Such families are typically different levels of discretization of a continuous operator.

To derive the bound for the effective condition number we start with the following auxiliary result.

Lemma 3.2. *Let $\{0\} \neq \mathcal{S} \subseteq \mathbb{K}^n$ be a subspace and $k := n - \dim(\mathcal{S})$. Furthermore, let $v_1, \ldots, v_n$ be a basis of $\mathbb{K}^n$ consisting of eigenvectors corresponding to the eigenvalues $\mu_1 \geq \cdots \geq \mu_n \geq 0$ of $A(I - \pi_A(\mathcal{S}))$.*

Then the vectors $v_1, \dots, v_k, Av_{k+1}, \dots, Av_n$ *form a basis of* $\mathbb{K}^n$ *consisting of eigenvectors corresponding to the eigenvalues* $\mu_1^{-1}, \dots, \mu_k^{-1}, 0, \dots, 0$ *of* $(I - \pi(\mathcal{S}))A^{-1}$. *Thus*

$$\sigma((I - \pi(\mathcal{S}))A^{-1}) = \{\mu_1^{-1}, \dots, \mu_k^{-1}, 0\},$$

where $\sigma((I - \pi(\mathcal{S}))A^{-1})$ *is the spectrum of* $(I - \pi(\mathcal{S}))A^{-1}$.

Proof. See [24, Theorem 2.1], cf. [12, Theorem 3.24]. □

As $\sigma(AB) = \sigma(BA)$ for general matrices $A, B \in \mathbb{K}^{n\times n}$ we obtain the following result from Lemma 3.2.

Corollary 3.3. *Under the same assumptions as in Lemma 3.2 we have*

$$\sigma(A^{-1/2}(I - \pi(\mathcal{S}))A^{-1/2}) = \{\mu_1^{-1}, \dots, \mu_k^{-1}, 0\}.$$

With Corollary 3.3 we are now able to formulate the main theorem of this section.

Theorem 3.4. *Let* $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n \geq 0$ *be the eigenvalues of* $A(I - \pi_A(\mathcal{S}))$ *and* k *the largest integer s.t.* $\mu_k \neq 0$. *Furthermore, let the weak approximation property (9) hold with constant* K. *Then*

$$\mu_1 \leq \|A\|_2 \quad \textit{and} \quad \mu_k = \frac{\|A\|_2}{K}.$$

And consequently the effective condition number κ_{eff} *of the matrix* $A(I - \pi_A(\mathcal{S}))$ *fulfills*

$$\kappa_{\text{eff}} \leq K.$$

Proof. We first show that $\|A\|_2$ is an upper bound for μ_1. As the matrix $A\pi_A(\mathcal{S}) = AV(V^*AV)^{-1}V^*A$ is positive semi-definite and thus

$$\langle A(I - \pi_A(\mathcal{S}))x, x\rangle = \langle Ax, x\rangle - \langle A\pi_A(\mathcal{S})x, x\rangle \leq \langle Ax, x\rangle,$$

we obtain, characterizing eigenvalues by Rayleigh quotients (see, e.g., [40]),

$$\mu_1 = \max_{v\in\mathbb{K}^n\setminus\{0\}} \frac{\langle A(I - \pi_A(\mathcal{S}))v, v\rangle}{\langle v, v\rangle} \leq \max_{x\in\mathbb{K}^n\setminus\{0\}} \frac{\langle Ax, x\rangle}{\langle x, x\rangle} = \lambda_1 = \|A\|_2, \tag{11}$$

where λ_1 was defined as the largest eigenvalue of A.

We now prove that $\frac{\|A\|_2}{K}$ is equal to μ_k. From Corollary 3.3 we see that

$$\mu_k^{-1} = \max_{v\neq 0} \frac{\langle A^{-\frac{1}{2}}(I - \pi(\mathcal{S}))A^{-\frac{1}{2}}v, v\rangle}{\langle v, v\rangle} = \max_{v\neq 0} \frac{\langle (I - \pi(\mathcal{S}))A^{-\frac{1}{2}}v, A^{-\frac{1}{2}}v\rangle}{\langle v, v\rangle}.$$

Substituting v by $A^{\frac{1}{2}}w$ and using $(I - \pi(\mathcal{S}))^2 = (I - \pi(\mathcal{S})) = (I - \pi(\mathcal{S}))^*$ yields

$$\mu_k^{-1} = \max_{v\neq 0} \frac{\langle (I - \pi(\mathcal{S}))w, w\rangle}{\langle A^{\frac{1}{2}}w, A^{\frac{1}{2}}w\rangle} = \max_{v\neq 0} \frac{\langle (I - \pi(\mathcal{S}))w, (I - \pi(\mathcal{S}))w\rangle}{\langle Aw, w\rangle}.$$

And by using the definition of the weak approximation property (9) we have that $\mu_k^{-1} = K/\|A\|_2$. □

Remark 3.5. Theorem 3.4 can also be proven by using the spectral equivalence of the deflated CG method and the CG method preconditioned by a $V(1,0)$-cycle of a multigrid method with a Richardson smoother with weight equal to one [36, Theorem 3.3]. Then Section 1 and Theorem 2.1 in [25] after some algebraic simplifications, yield the above theorem.

4 Perturbation of Deflation Subspaces

We now use Theorem 3.4 to give a bound on the effective condition number when the deflation subspace is perturbed. This is of particular interest when the deflation subspace is spanned by eigenvectors of the matrix A which is common practice, see e.g. [2, 7, 9, 13, 28, 31, 32]. Typically the eigenvectors corresponding to the smallest eigenvalues are unknown and need to be approximated numerically. Thus the question arises how precisely those eigenvectors need to be determined to achieve fast convergence of the deflated CG method. We first answer this question for the perturbation of general deflation subspaces and later on discuss the case of deflation of eigenvectors.

We need to quantify the perturbation of the deflation subspace to give a bound on the effective condition number. The difference between two subspaces $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ can be measured in terms of the largest principle angle θ—also called *subspace gap*—between the subspaces (cf. [14, 30]), i.e.,

$$\|\pi(\mathcal{S}) - \pi(\widetilde{\mathcal{S}})\|_2 = \sqrt{1 - \cos(\theta)^2} = \sin(\theta). \tag{12}$$

Thus we can measure the perturbation as the angle between the perturbed and the unperturbed subspace. This is a suitable measure as it does not depend on the choice of the basis of the subspace. In the following Lemma we give a bound on the weak approximation property constant using the subspace angle θ.

Lemma 4.1. *Let $\mathcal{S} \subseteq \mathbb{K}^n$ and $\widetilde{\mathcal{S}} \subseteq \mathbb{K}^n$ be two subspaces of the same dimension. Assume that the space $\mathcal{S}$ fulfills the weak approximation property with constant K. Then $\widetilde{\mathcal{S}}$ fulfills the weak approximation property with constant*

$$\widetilde{K} \leq \left(K^{1/2} + \sin(\theta) \cdot \sqrt{\kappa(A)}\right)^2,$$

where θ is the largest principal angle between $\mathcal{S}$ and $\widetilde{\mathcal{S}}$.

Proof. Let $x \in \mathbb{K}^n$. Then

$$\begin{aligned} \|x - \pi(\widetilde{\mathcal{S}})x\|_2 &= \|x + \pi(\mathcal{S})x - \pi(\mathcal{S})x - \pi(\widetilde{\mathcal{S}})x\|_2 \\ &\leq \|x - \pi(\mathcal{S})x\|_2 + \|\pi(\mathcal{S})x - \pi(\widetilde{\mathcal{S}})x\|_2 . \end{aligned} \tag{13}$$

The subspace $\mathcal{S}$ fulfills the weak approximation property thus the first term of (13) is bound by

$$\|x - \pi(\mathcal{S})\|_2 \leq K^{1/2} \cdot \frac{\|x\|_A}{\|A\|_2^{1/2}}.$$

We can bound the second term (13) by using the subspace gap (12) to get

$$\begin{aligned} \|\pi(\mathcal{S})x - \pi(\widetilde{\mathcal{S}})x\|_2 &\leq \|\pi(\mathcal{S}) - \pi(\widetilde{\mathcal{S}})\|_2 \cdot \|x\|_2 = \sin(\theta) \cdot \|x\|_2 \\ &\leq \sin(\theta) \cdot \frac{\|x\|_A}{\lambda_n^{1/2}} = \sin(\theta) \cdot \frac{\lambda_1^{1/2}}{\lambda_n^{1/2}} \cdot \frac{\|x\|_A}{\|A\|_2^{1/2}}. \end{aligned}$$

Combining the two estimates yields

$$\|x - \pi(\widetilde{\mathcal{S}})x\|_2 \leq \left(K^{1/2} + \sin(\theta) \cdot \frac{\lambda_1^{1/2}}{\lambda_n^{1/2}}\right) \cdot \frac{\|x\|_A}{\|A\|_2^{1/2}}. \qquad \square$$

Using Theorem 3.4 we get a bound for the smallest eigenvalue of the deflated matrix:

Theorem 4.2. *Let $\mathcal{S} \subseteq \mathbb{K}^n$ and $\widetilde{\mathcal{S}} \subseteq \mathbb{K}^n$ be two subspaces of the same dimension. Let μ_k be the smallest non-zero eigenvalue of $A(I - \pi_A(\mathcal{S}))$ and $\tilde{\mu}_k$ the smallest non-zero eigenvalue of $A(I - \pi_A(\widetilde{\mathcal{S}}))$. Then*

$$\tilde{\mu}_k^{-1} \leq \left(\mu_k^{-1/2} + \lambda_n^{-1/2} \cdot \sin(\theta)\right)^2 = \mu_k^{-1} + O(\theta), \quad \text{for } \theta \to 0 \tag{14}$$

where θ is the largest principal angle between $\mathcal{S}$ and $\widetilde{\mathcal{S}}$.

Recall that Theorem 3.4 stated that the largest eigenvalue of the deflated matrix is smaller than the largest eigenvalue of the original matrix A. Thus in combination with Theorem 4.2 this gives an estimate for the effective condition number.

We now turn our attention to the special case where $\tilde{\mathcal{S}}$ is spanned by inexact eigenvectors. Recall that we defined $q_1, q_2, \ldots, q_n$ to be an orthonormal basis of eigenvectors of the matrix A, s.t. $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ are the corresponding eigenvalues. Furthermore, let $\mathcal{S}$ be the space spanned by the eigenvectors corresponding to the $(n-k)$ smallest eigenvalues, i.e., $\mathcal{S} = \operatorname{range} V$, where

$$V = [q_{k+1} | q_{k+2} | \ldots | q_n] \,.$$

To apply Theorem 4.2 we now determine the value for K, such that the weak approximation property (9) for $\mathcal{S}$ is fulfilled. Let $x \in \mathbb{K}^n$ be given and $x = \sum_{i=1}^n \xi_i q_i$ its expansion in terms of the orthonormal eigenvectors q_i of A. Then the orthogonal projection $\pi(\mathcal{S})x$ of x onto $\mathcal{S}$ fulfills $\pi(\mathcal{S})x = \sum_{i=k+1}^n \xi_i q_i$ and thus

$$\|x - \pi(\mathcal{S})x\|_2^2 = \| \sum_{i=1}^{k} \xi_i q_i + \sum_{i=k+1}^{n} (\xi_i - \xi_i) q_i \|_2^2 = \sum_{i=1}^{k} |\xi_i|^2 \,.$$

This yields

$$\|x\|_A^2 = \sum_{i=1}^{n} |\xi_i|^2 \lambda_i \geq \sum_{i=1}^{k} |\xi_i|^2 \lambda_i \geq \lambda_k \sum_{i=1}^{k} |\xi_i|^2 = \lambda_k \, \|x - \pi(\mathcal{S})x\|_2^2 \,. \tag{15}$$

Hence the weak approximation property (9) holds with $K \leq \frac{\|A\|_2}{\lambda_k} = \frac{\lambda_1}{\lambda_k}$. This is also the smallest possible constant that fulfills the weak approximation property as (15) is an equality for $x = q_k$. Thus $K = \frac{\lambda_1}{\lambda_k}$. Using Theorem 3.4 we obtain $\kappa_{\text{eff}} \leq \frac{\lambda_1}{\lambda_k}$. Furthermore it is known that $\kappa_{\text{eff}} = \frac{\lambda_1}{\lambda_k}$ (see, e.g. [9, Section 1]) hence our bound is sharp.

Now we are in the position to formulate a new result concerning the deviation of the effective condition number from $\frac{\lambda_1}{\lambda_k}$ due to inexact eigenvectors. We apply Theorem 4.2 and Theorem 3.4 to obtain the following proposition.

Proposition 4.3. *Let $q_1, \ldots, q_n$ be an orthonormal basis of eigenvectors corresponding to the eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ of A and let the subspace $\mathcal{S} := \operatorname{span}\{q_{k+1}, \ldots, q_n\}$. Then the effective condition number of the deflated system with respect to the deflation subspace $\widetilde{\mathcal{S}}$ fulfills*

$$\kappa_{\text{eff}} \leq \left(\sqrt{\frac{\lambda_1}{\lambda_k}} + \sqrt{\frac{\lambda_1}{\lambda_n}} \cdot \sin(\theta) \right)^2 = \frac{\lambda_1}{\lambda_k} + O(\theta), \quad \text{for } \theta \to 0 \,, \tag{16}$$

where θ is the largest principal angle between $\mathcal{S}$ and $\widetilde{\mathcal{S}}$.

Proposition 4.3 shows that the effective condition number deviates asymptotically linearly from the unperturbed effective condition number. In addition it indicates that the accuracy of the deflated eigenvectors should increase linearly with the condition number if we aim at keeping the effective condition number within a given factor of its optimal value $\frac{\lambda_1}{\lambda_k}$.

Remark 4.4. *For some applications the largest principal angle is not a natural measure for the perturbation of the deflation subspace. Under the following assumption there is a simple bound for the largest principal angle. Let $V \in \mathbb{K}^{n\times n-k}$ with orthonormal columns and $\mathcal{S} := \operatorname{range} V$. Furthermore, let $E \in \mathbb{K}^{n\times n-k}$ and $\widetilde{\mathcal{S}} := \operatorname{range}(V+E)$. Then the largest principal angle θ between $\mathcal{S}$ and $\widetilde{\mathcal{S}}$ fulfills*

$$\sin(\theta) \leq \|E\|_2 .$$

In other words if the deflation subspace is given by an orthonormal basis and we can bound the norm of the difference of this orthonormal basis and a basis of the perturbed deflation subspace then we can bound θ.

Proof. We have [34, Theorem 5.5], [18, I. Theorem 6.34] that

$$\sin(\theta) = \|\pi(\mathcal{S}) - \pi(\widetilde{\mathcal{S}})\|_2 = \|(I - \pi(\widetilde{\mathcal{S}}))\pi(\mathcal{S})\|_2 .$$

It then follows that

$$\sin(\theta) = \max_{\substack{\|x\|=1 \\ x\in\mathcal{S}}} \min_{y\in\widetilde{\mathcal{S}}} \|x - y\|_2 \leq \max_{\substack{\|x\|=1 \\ x\in\mathcal{S}}} \|x - (V+E)V^*x\|_2 .$$

Every $x \in \mathcal{S}$ with $\|x\|$ can be written as $x = Vz$ with $\|z\| = 1$. Thus

$$\begin{aligned}\sin(\theta) &\leq \max_{\|z\|=1} \|Vz - (V+E)V^*Vz\|_2 = \max_{\|z\|=1} \|Vz - Vz + Ez\|_2 \\ &= \max_{\|z\|=1} \|Ez\|_2 = \|E\|_2 .\end{aligned}$$

□

5 Accuracy of the Deflating Projection

The deflated CG method involves the solution of the *inner linear system*

$$(V^*AV)z_{i+1} = V^*Ar_{i+1} \tag{17}$$

in every iteration. In the situation where $\mathcal{S}$ is of large dimension, it can be desirable to solve the inner system (17) inexactly by an iterative method, the *inner iteration*. This is for example the case when many eigenvectors are to be deflated, or, more generally when the deflation subspace is large (and represented by a basis of sparse vectors thus making the use of an inner iteration more attractive, see [9, 20]). In this context, it has been observed in [21] that deflation methods are quite sensitive to the accuracy of the inner iteration.

Given a stopping criterion for the outer iteration, i.e.,

$$\|r_i\|_2 \leq \tau \|b\|_2 =: \varepsilon \tag{18}$$

for some $0 < \tau \ll 1$, we now want to specify a stopping criterion for the inner iteration which is of the form

$$\|r_i^c\|_2 \leq \tau^c \, \|b^c\|_2 \, . \tag{19}$$

The experiments in [21] suggest that it is sufficient to set a *fixed tolerance*

$$\tau^c = \varepsilon \cdot c \quad \text{with} \quad 0 < c \leq 1. \tag{20}$$

In our experiments we observed that this accuracy requirement is only crucial for the first iterations and can be relaxed later on. This can be explained by the theory for inexact Krylov subspace methods from [29, 38] as follows.

We first observe that within the deflated CG method the matrix $A(I - \pi_A(\mathcal{S}))$ is only used to compute matrix vector products, requiring the solution of the inner system (17). Due to the fact that the deflated matrix can be written as

$$A(I - \pi_A(\mathcal{S})) = A - AV(V^*AV)^{-1}V^*A \, .$$

We can interprete the inexact calculation of $(V^*AV)^{-1}$ as the inexact matrix vector product by replacing

$$(V^*AV)^{-1} \quad \text{by} \quad (V^*AV)^{-1} + \Delta_i$$

for some perturbation matrix Δ_i—different in every iteration. Then

$$\begin{aligned} A - AV((V^*AV)^{-1} + \Delta_i)V^*A) &= A - AV(V^*AV)^{-1}V^*A - \underbrace{AV\Delta_i V^*A}_{=:E_i} \\ &= A(I - \pi_A(\mathcal{S})) - E_i \, . \end{aligned}$$

Hence, the inaccurate solution of (17) is essentially equivalent to an inaccurately computed matrix vector product. (For a more general treatment of the perturbation of projections see [33].)

We start our discussion by considering the full orthogonalization method (FOM) [27] which is equivalent to CG in exact arithmetic. (The first vector should be computed with full accuracy.) Assume we run FOM where the ith matrix vector product is replaced by a product with the matrix $A(I - \pi_A(\mathcal{S})) - E_i$. According to [29] the method converges to the desired tolerance ε if

$$\|E_i\| < C \tfrac{1}{\|r_i\|} \varepsilon$$

for some constant $C > 0$ (see also [38]) and it has been observed in [3] that the rate of convergence does not change by much even when choosing $C = 1$.

In particular we obtain due to $\|E_i\| \leq \|AV\|^2 \cdot \|\Delta_i\|$ that the norm of E_i is bounded by a factor proportional to the norm of Δ_i. In turn the norm of Δ_i can be bounded by a value that is proportional to the error in the computation of $(V^*AV)^{-1}b^c$ and this error can be bounded by a value proportional to the norm of the residual. Hence there exists a constant $c > 0$ such that

$$\tau^c = c \tfrac{1}{\|r_i\|} \varepsilon$$

is sufficient for FOM to converge to the desired tolerance. We will refer to this as the *adaptive tolerance*, where r_i is the residual of the outer iteration.

Equation (5) means that the relative tolerance for the inner iteration can be relaxed while the outer iteration advances.

In exact arithmetic FOM is equivalent to the CG method. When using inexact CG the loss of orthogonality can become a problem. Thus additional orthogonalization might be required as in the flexible CG method [23].

6 Numerical Experiments

This section contains numerical experiments to illustrate the developed theory and the quality of our convergence estimates. We use the following three test matrices called *Simple*, *Poisson* and *NOS1*.

Simple The *Simple* matrix is the diagonal matrix $A = \operatorname{diag}(10^{-2}, 1, \ldots, 1) \in \mathbb{R}^{100\times 100}$ with eigenvalues $\lambda_{100} = 10^{-2}$ and $\lambda_1 = \cdots = \lambda_{99} = 1$.

The convergence rate of the (deflated) CG method is invariant under simultaneous unitary transformations of the matrix, right hand side, and initial guess. Thus without loss of generality we can consider diagonal matrices.

The condition number of the Simple matrix is 100. If the deflation subspace is spanned by the eigenvector corresponding to the smallest eigenvalue e_1 then the condition number of the deflated matrix is 1. Any deflation subspace that is orthogonal to e_1 yields an effective condition number of 100.

If we pick a perturbation orthogonal to e_1 and increase its size step by step, the condition number should gradually go from 1 to 100. This will give us a first impression of the behavior of the bounds.

Poisson The *Poisson*(N) matrix is the $N^2 \times N^2$ matrix arising from the finite element discretization of Poisson's equation using quadratic bilinear elements on a uniform $N \times N$ grid. It is one of the simplest, non-trivial matrices that appear in real world applications.

NOS1 The *NOS1* matrix is the 237×237 matrix from [1]. According to its description, it is a "finite element approximation to [the] biharmonic operator on a beam with one end free and one end fixed." We choose this as our third example as it has a large condition number of about $2 \cdot 10^7$.

From the different mathematical equivalent formulations of the deflated CG method [13] we chose the one from Saad, Yeung, Erhel and Guyomarc'h [28]. Throughout this section we refer to the *optimal condition number* as the *effective condition number* of the deflated matrix for the unperturbed subspace $\mathcal{S}$. We refer to the *original condition number* as the condition number of the matrix A. We refer to the *effective condition number* as the effective condition number of the deflated matrix for the perturbed subspace $\widetilde{\mathcal{S}}$. Analogously the optimal, original and effective smallest eigenvalues is the smallest non-zero eigenvalue of $A(I - \pi_A(\mathcal{S})$, A and $A(I - \pi_A(\widetilde{\mathcal{S}}))$, respectively.

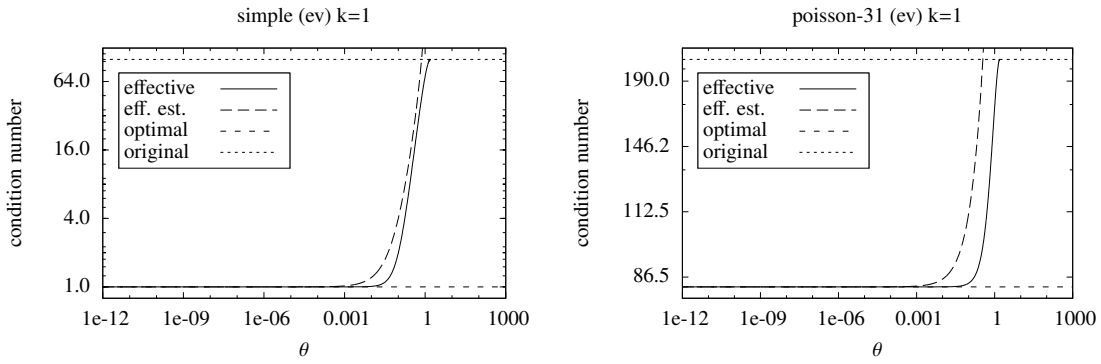

Figure 1: Estimated and computed effective condition number for the Simple (left) and Poisson(31) (right) matrix as a function of the principal angle θ.

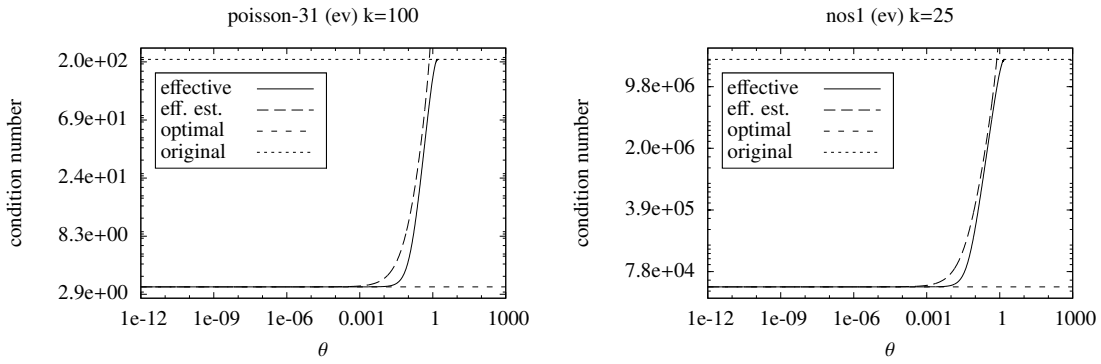

Figure 2: Estimated and computed effective condition for the Poisson(31) matrix, deflating 100 eigenvectors, (left) and NOS1 matrix, deflating 25 eigenvectors, (right) as a function of the principal angle θ.

6.1 Perturbation in Eigenvector Deflation

In this section we illustrate the perturbation theory from Section 4. We denote the eigenvalues of the matrix A under consideration by $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ with corresponding eigenvectors $q_1, \ldots, q_n$.

First, we consider the *Simple* matrix and choose $\mathcal{S} = \text{span}\{q_n\}$. We choose the perturbation vector $E = \alpha q_{n-1}$, $\alpha \in \mathbb{R}$, thus $\tilde{\mathcal{S}} = \text{span}\{q_n + \alpha q_{n-1}\}$ is the perturbed deflation space. This choice of perturbation has a large impact on the effective condition number. Thus it will illustrate the sharpness of the developed bound. We compute the effective condition number of the matrix $A(I - \pi_A(\tilde{\mathcal{S}}))$ and the estimate from Section 4. The results are given in the left part of Figure 1. Here the graph "effective" shows the effective condition number, the graph "eff. est." shows the value from (16). We note that the estimate recovers the behavior of the actual effective condition number quite well.

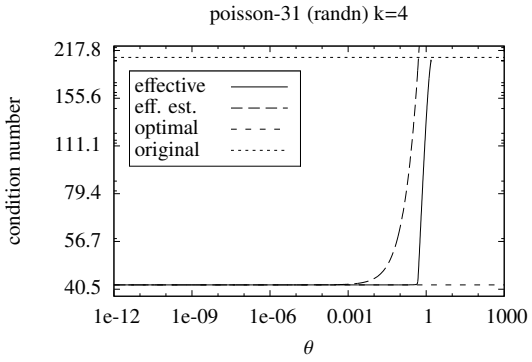

Figure 3: Estimated and computed effective condition for the Poisson(31) matrix, 4 eigenvectors and a random perturbation as a function of the principal angle θ.

We run the same test for the Poisson(31) matrix and report the results in the right part of Figure 1. Qualitatively the results are very similar to those for the matrix Simple with the difference between the estimate and the effective condition number being slightly larger. Deflating the eigenvectors corresponding to the k smallest eigenvalues, i.e., the subspaces $\mathcal{S} = \text{span}\{q_{n-(k-1)}, \ldots, q_n\}$ and $E = \alpha \left[q_{n-2(k-1)} | \ldots | q_{n-(k-1)-1}\right]$ yields similar results as shown in Figure 2; we choose a deflation subspace consisting of approximately $n/10$ eigenvectors corresponding to the smallest eigenvalues of the Poisson(31) and the NOS1 matrix, respectively.

The perturbations chosen so far had a relatively large effect on the effective condition number. To illustrate that we conduct a test with a random perturbation matrix $E = \alpha R$. In here the entries of R are normally $\mathcal{N}(0, 1)$ distributed random numbers. In Figure 3 we can see that the theoretical bound is not as sharp as in the previous examples, but the qualitative behavior is again captured.

Based on the convergence bound (8) of deflated CG one can easily derive an estimate for the number of iterations deflated CG requires to reduce the ℓ_2-norm of the residual to a prescribed tolerance τ in terms of κ_{eff}. To this extend first observe that due to (8)

$$\|r_i\|_2 \leq 2\sqrt{\kappa}\left(\frac{\sqrt{\kappa_{\text{eff}}}-1}{\sqrt{\kappa_{\text{eff}}}+1}\right)^i \|r_0\|_2 \quad \text{for } i = 0, 1, 2, \ldots$$

and thus

$$i \geq \frac{\log(\tau/(2\sqrt{\kappa}))}{\log((\sqrt{\kappa_{\text{eff}}}-1)/(\sqrt{\kappa_{\text{eff}}}+1))}$$

to reduce the ℓ_2-norm of the residual by a factor of τ.

In order to illustrate that our bound for κ_{eff} can be used as a tool to set the required accuracy of eigenvectors used in the deflation subspace, we report in Figure 4 the number of iterations deflated CG required to converge as a function of the perturbation of the smallest $k = 50$ eigenvectors. As one can see the estimate gives a fairly good idea about the accuracy requirement, i.e., it correctly captures the steep increase of the number of iteration. We stop the iteration when the norm of the residual is reduced by a factor of 10^6.

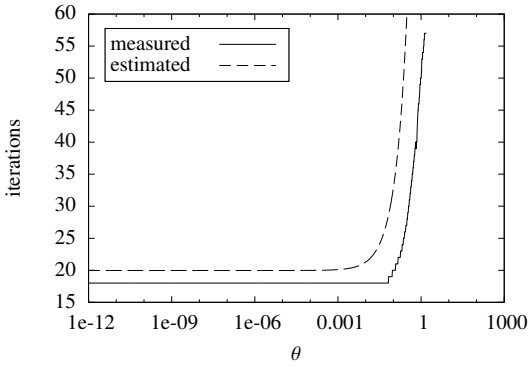

Figure 4: Number of iterations of the deflated CG method for the Poisson(31) matrix in dependence of the size of the perturbation. Deflation of the 50 smallest eigenvectors. Reduction of the norm of the residual by a factor of 10^{-6}.

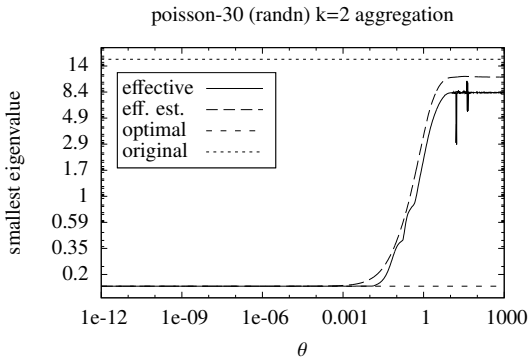

Figure 5: Aggregation with $k = 2$ for Poisson(30), random perturbation. Estimated and measured μ_k^{-1}.

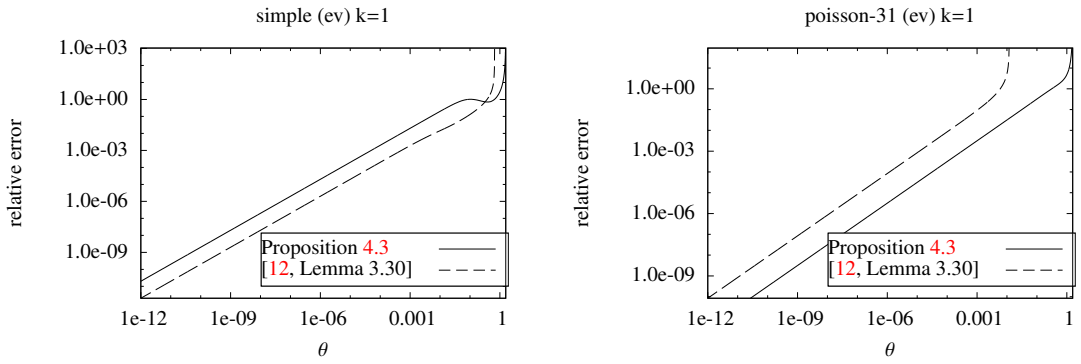

Figure 6: The relative error of the eigenvalue bound from [12, Lemma 3.30] and Proposition 4.3 for the Simple (left) and Poisson(31) (right) matrix.

As a final experiment we consider a deflation subspace that is not directly spanned by eigenvectors. We consider the Poisson(30) matrix and construct V as described in [20]. That is, we take eigenvectors to the smallest two eigenvalues and "chop" them up over 2×2 aggregates, i.e., the 30×30 grid is divided into 2×2 blocks and the $15^2 \cdot 2$ columns of V are simply the entries of the eigenvectors restricted to these blocks. The perturbed basis $\widetilde{V}$ is obtained by repeating the same procedure with the same eigenvectors that are perturbed by two scaled $\mathcal{N}(0, 1)$ random vectors. In Figure 5 we show the estimate for μ_k^{-1} together with the actual μ_k^{-1} in dependence of the largest principal angle between the subspace spanned by the columns of V and $\widetilde{V}$. Again we see that the qualitative behavior is captured well by the estimate, but as expected the gap between estimate and actual value of μ_k^{-1} is larger than in the cases where V is built of orthonormal vectors.

In [12, Lemma 3.30] a bound for the difference of the eigenvalues of the deflated matrix in the perturbed and in the unperturbed case is given. If the effective condition number of the unperturbed deflated matrix is known then a bound for the effective condition number in the perturbed case is easily derived. We compare this bound with the result from Proposition 4.3. To this extend we compare the relative error, i.e.,

$$\frac{|\text{bound} - \kappa_{\text{eff}}|}{\kappa_{\text{eff}}} .$$

of the two bounds in Figure 6. The bound of [12] is better than the bound from Proposition 4.3 for the Simple matrix. For the Poisson(31) matrix the situation is the other way around.

6.2 Accuracy of the Inner System

We now consider the Poisson(31) matrix and choose the columns of V as the 200 eigenvectors corresponding to the smallest eigenvalues of the Poisson(31) matrix. In Figure 7 we show the convergence of the deflated CG iteration for varying inner accuracy requirements, i.e., the accuracy of the solution of $V^*AVe = V^*Ar$ approximated by an inner CG iteration. The left hand

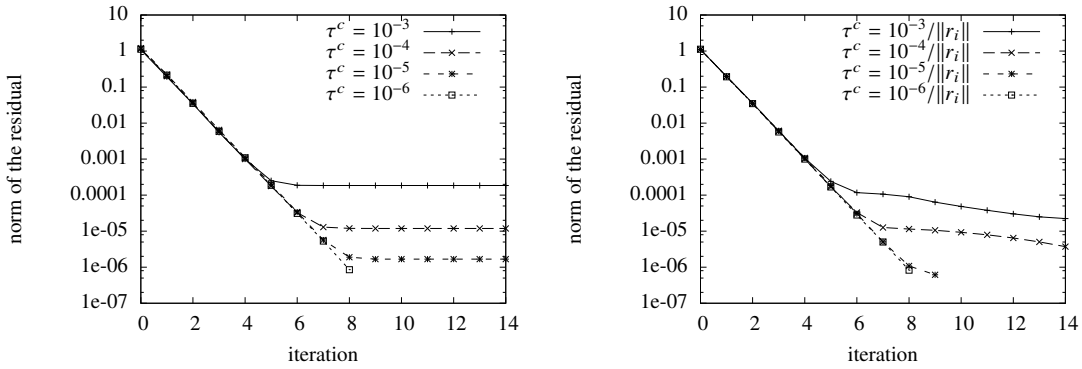

Figure 7: Deflation of 200 eigenvectors of the Poisson(31) matrix with fixed (left) and adaptive (right) inner tolerance.

side of the figure shows the norm of the residual for different choices of τ^c in (19); the right hand side shows these norms for different choices of $\tau^c \|r_i\|_2$; see (5).

From the discussion in section 5 we expect that the adaptive tolerance (5) will produce similar results to the fixed one (20) and this can indeed be observed in Figure 7. Choosing the tolerance adaptively produces even slightly better results.

In numerical experiments for other choices for the deflation subspace V which we do not report here, we observed no qualitative difference of the two choices for the inner tolerance criterion.

Motivated by this observation we run the same test again using the adaptive tolerance (5) as the stopping criterion and monitor the number of inner iterations in each outer iteration. We choose $c = 0.1$ and solve to a tolerance of $\tau = 10^{-6}$. As the results in Table 1 show, less and less iterations of the inner CG method are needed when the outer iteration advances and the adaptive tolerance is used. For the fixed tolerance a high number of inner iterations is required throughout the whole outer iteration. We note that the convergence is not distinguishable from the case where an exact inner solve would be performed.

7 Conclusions

We have shown that to get an optimal convergence rate of the deflated CG method the eigenvectors only have to be computed to a limited accuracy. This is particularly interesting in situations where a linear system has to be solved for many right hand sides and the additional overhead of computing approximations, e.g. by using ARPACK [19], of the smallest eigenvectors can be compensated. This is for example the case when computing rational approximations of matrix functions [10] or in the computation of expectation values in statistical physics, e.g., Lattice Gauge theory [11].

Table 1: Number of inner iterations for both tolerance selection strategies ($\tau = 10^{-6}$, $c = 0.1$).

It.	$\|r_i\|_2$	Inner-It.	
		Fixed	Adaptive
0	$8.21 \cdot 10^{0}$	35	35
1	$1.15 \cdot 10^{0}$	27	27
2	$2.15 \cdot 10^{-1}$	27	24
3	$3.93 \cdot 10^{-2}$	27	20
4	$6.41 \cdot 10^{-3}$	27	16
5	$1.13 \cdot 10^{-3}$	27	11
6	$1.96 \cdot 10^{-4}$	27	8
7	$3.25 \cdot 10^{-5}$	27	4
8	$5.70 \cdot 10^{-6}$	27	2
9	$9.15 \cdot 10^{-7}$	27	0

Acknowledgments

We would like to thank the referees for many helpful remarks and suggestions and Andreas Frommer for his help and advice. Furthermore, the work is partly supported by the German Research Foundation (DFG) through the Priority Programme 1648 "Software for Exascale Computing" (SPPEXA) and the Transregional Collaborative Research Centre 55 (SFB/TRR55) "Hadron Physics from Lattice QCD".

References

[1] *Matrix market*. http://math.nist.gov/MatrixMarket/.

[2] A. M. Abdel-Rehim, R. B. Morgan, D. A. Nicely, and W. Wilcox, *Deflated and restarted symmetric Lanczos methods for eigenvalues and linear equations with multiple right-hand sides*, SIAM J. Sci. Comput., 32 (2010), pp. 129–149.

[3] A. Bouras and V. Frayssé, *Inexact matrix-vector products in Krylov methods for solving linear systems: a relaxation strategy*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 660–678.

[4] A. Brandt, *Algebraic multigrid theory: The symmetric case*, Appl. Math. Comput., 19 (1986), pp. 23–56.

[5] M. Brezina, A. J. Cleary, R. D. Falgout, V. E. Henson, J. E. Jones, T. A. Manteuffel, S. F. McCormick, and J. W. Ruge, *Algebraic multigrid based on element interpolation (AMGe)*, SIAM J. Sci. Comput., 22 (2001), pp. 1570–1592.

[6] W. L. Briggs, V. E. Henson, and S. F. McCormick, *A Multigrid Tutorial*, SIAM, 2nd ed., 2000.

[7] T. F. Chan and W. L. Wan, *Analysis of projection methods for solving linear systems with multiple right-hand sides*, SIAM J. Sci. Comput., 18 (1997), pp. 1698–1721.

[8] Z. Dostál, *Conjugate gradient method with preconditioning by projector*, Int. J. Comput. Math., 23 (1988), pp. 315–323.

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

[10] A. Frommer and V. Simoncini, *Matrix functions*, in Model order reduction: theory, research aspects and applications, vol. 13 of Math. Ind., Springer, Berlin, 2008, pp. 275–303.

[11] C. Gattringer and C. Lang B., *Quantum Chromodynamics on the Lattice. An Introductory Presentation*, no. 788 in Lecture Notes in Physics, Springer Heidelberg, 2010.

[12] A. Gaul, *Recycling Krylov subspace methods for sequences of linear systems. Analysis and applications*, PhD thesis, Technischen Universität Berlin, 2014.

[13] A. Gaul, M. H. Gutknecht, J. Liesen, and R. Nabben, *A framework for deflated and augmented Krylov subspace methods*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 495–518.

[14] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, 2nd ed., 1989.

[15] W. Hackbusch, *Multi-Grid Methods and Applications*, Springer-Verlag, 1985.

[16] M. R. Hestenes and E. Stiefel, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Standards, 49 (1952), pp. 409–436 (1953).

[17] E. F. Kaasschieter, *Preconditioned conjugate gradients for solving singular systems*, J. Comput. Appl. Math., 24 (1988), pp. 265–275.

[18] T. Kato, *Perturbation theory for linear operators*, Springer, 2nd ed., 1976.

[19] R. B. Lehoucq, D. C. Sorensen, and C. Yang, *ARPACK users' guide*, vol. 6 of Software, Environments, and Tools, SIAM, Philadelphia, PA, 1998.

[20] M. Lüscher, *Local coherence and deflation of the low quark modes in lattice QCD*, J. High Energy Phys., 2007 (2007), p. 081.

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

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

[23] Y. Notay, *Flexible conjugate gradients*, SIAM J. Sci. Comput., 22 (2000), pp. 1444–1460.

[24] ——, *Algebraic analysis of two-grid methods: The nonsymmetric case*, Numer. Linear Algebra Appl., 17 (2010), pp. 73–96.

[25] Y. Notay and A. Napov, *Further comparison of additive and multiplicative coarse grid correction*, Appl. Numer. Math., 65 (2013), pp. 53–62.

[26] J. W. Ruge and K. Stüben, *Algebraic multigrid*, in Multigrid methods, S. F. McCormick, ed., vol. 3 of Frontiers in Applied Mathematics, SIAM, Philadelphia, PA, 1987, pp. 73–130.

[27] Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, 2nd ed., 2003.

[28] 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.

[29] V. Simoncini and D. B. Szyld, *Theory of inexact Krylov subspace methods and applications to scientific computing.*, SIAM J. Sci. Comput., 25 (2003), pp. 454–477.

[30] V. Simoncini and D. B. Szyld, *On the occurrence of superlinear convergence of exact and inexact Krylov subspace methods*, SIAM Rev., 47 (2005), pp. 247–272.

[31] C. Smith, A. Peterson, and R. Mittra, *A conjugate gradient algorithm for the treatment of multiple incident electromagnetic fields*, IEEE T. Antenn. Propag., 37 (1989), pp. 1490–1493.

[32] A. Stathopoulos and K. Orginos, *Computing and deflating eigenvalues while solving multiple right-hand side linear systems with an application to quantum chromodynamics*, SIAM J. Sci. Comput., 32 (2010), pp. 439–462.

[33] G. W. Stewart, *On the numerical analysis of oblique projectors*, SIAM J. Matrix Anal. Appl., 32 (2011), pp. 309–348.

[34] G. W. Stewart and J.-g. S. Sun, *Matrix Perturbation Theory*, Academic Press, 1990.

[35] K. Stüben, *An introduction to algebraic multigrid*, in Multigrid, U. Trottenberg, C. W. Oosterlee, and A. Schüller, eds., Academic Press, 2001, pp. 413–532.

[36] 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.

[37] U. Trottenberg, C. W. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, 2001.

[38] J. van den Eshof and G. L. G. Sleijpen, *Inexact Krylov subspace methods for linear systems*, SIAM J. Matrix Anal. Appl., 26 (2004), pp. 125–153.

[39] A. van der Sluis and H. A. van der Vorst, *The rate of convergence of conjugate gradients*, Numer. Math., 48 (1986), pp. 543–560.

[40] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, 1965.