GMRES convergence bounds for eigenvalue problems

Melina A. Freitag[*]   Patrick Kürschner[†]   Jennifer Pestana[‡]

May 25, 2017

Abstract

The convergence of GMRES for solving linear systems can be influenced heavily by the structure of the right hand side. Within the solution of eigenvalue problems via inverse iteration or subspace iteration, the right hand side is generally related to an approximate invariant subspace of the linear system. We give detailed and new bounds on (block) GMRES that take the special behavior of the right hand side into account and explain the initial sharp decrease of the GMRES residual. The bounds motivate the use of specific preconditioners for these eigenvalue problems, e.g. tuned and polynomial preconditioners, as we describe. The numerical results show that the new (block) GMRES bounds are much sharper than conventional bounds and that preconditioned subspace iteration with either a tuned or polynomial preconditioner should be used in practice.

Key words. GMRES, convergence analysis, inexact inverse iteration, inexact subspace iteration, Krylov subspace methods, block Krylov methods, preconditioning

MSC classification. 15A18, 65F08, 65F10, 65F15, 65N25

1 Introduction

We consider the convergence of GMRES [24] for solving sequences of linear systems of the form

$$Bx_i = y_i, \tag{1}$$

where $B = A - \sigma I$, $A \in \mathbb{C}^{n \times n}$ and σ is a fixed or variable scalar shift. Throughout, we assume that A is diagonalizable with eigenvalues γ_j, $j = 1, \ldots, n$, so that $B = A - \sigma I$ has eigenvalues $\lambda_j = \gamma_j - \sigma$, $j = 1, \ldots, n$. Thus, we can write $B = Z \Lambda Z^{-1}$, where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n) \in \mathbb{C}^{n \times n}$ and $Z = [z_1, \ z_2, \ \ldots, \ z_n]$ is a matrix of eigenvectors. Without loss of generality, we let y_i be an approximation of the eigenvector z_1 (with both vectors normalized to have unit norm). Our main aim is to describe accurately, using simple bounds, the convergence of GMRES when the right-hand side is an approximate eigenvector; to clearly explain why this convergence behavior is lost for many preconditioners; and how it can be recovered by choosing or modifying the preconditioner appropriately.

---

[*]Department of Mathematical Sciences, University of Bath, Claverton Down, BA2 7AY, United Kingdom, `m.freitag@maths.bath.ac.uk`

[†]Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany, `kuerschner@mpi-magdeburg.mpg.de`.

[‡]Department of Mathematics and Statistics, University of Strathclyde, Glasgow, G1 1XQ, United Kingdom `jennifer.pestana@strath.ac.uk`

Systems of the form (1) arise, for example, when computing an eigenvector and corresponding eigenvalue of a matrix $A \in \mathbb{C}^{n \times n}$ using inverse iteration (see Algorithm 1). In this case the shift is chosen to be close to an eigenvalue of A and the approximate eigenvector y_i is the ith iterate of inverse iteration. If the system (1) is solved exactly (e.g. using a direct solver) one can show that, under certain generic assumptions, the right-hand side y_i converges at least linearly to an eigenvector of A (see, e.g. [21, 15] and references therein). However, for very large matrix sizes n, direct solvers for (1) are infeasible and an iterative method is used to solve (1) inexactly. In this case, one can show that if the solve tolerance is chosen appropriately (e.g. it is reduced proportionally to the eigenvalue residual) we still obtain the same convergence rates for this inexact version of inverse iteration [11, 10].

Algorithm 1: Inexact inverse iteration

Input: x_0, $i_{\max}$
for $i = 0, \ldots, i_{max} - 1$ **do**
    Choose σ and τ_i
    Find y_i such that $\|(A - \sigma I)y_i - x_i\| \leq \tau_i$
    Set $x_{i+1} = y_i / \|y_i\|$ and $\lambda_{i+1} = x_{i+1}^T A x_{i+1}$
    Evaluate $r_{i+1} = (A - \lambda_{i+1} I)x_{i+1}$ and test for convergence
end
Output: $x_{i_{\max}}$, $\lambda_{i_{\max}}$

Here we are concerned with the the convergence behavior of iterative methods for solving (1). It is well known that when the right-hand side is an approximate eigenvector of B, Krylov subspace methods converge faster than for generic right-hand sides. This was first observed for Hermitian positive matrices by Simoncini and Eldén [25], who considered linear solves in Rayleigh quotient iteration. Further results in this direction include convergence bounds in [6] for MINRES (see [19]) applied within inverse iteration for symmetric matrices, and Theorem 2.2 in [12], which characterizes GMRES convergence for non-Hermitian generalized eigenvalue problems using the Schur factorization. For more general systems that do not necessarily come from eigensolvers, research into GMRES convergence has also highlighted the influence of the right-hand side [2, 8], and its representation in the eigenvector basis [18, 28].

In this work, we show that if A is diagonalizable the convergence theory developed in [28] yields an insightful explanation for the GMRES behavior in inverse iteration with different types of preconditioned inner solves. Moreover, we explain why the GMRES residual often decreases sharply in the first iteration [31]. A more detailed description of this phenomenon is given in Section 2.2. Regarding the preconditioned situation, it is well known that so called tuned preconditioners can significantly reduce the number of required GMRES steps. We show that using polynomial preconditioners for the inner solves can be similarly beneficial.

The rest of this paper is structured as follows. In Section 2 we present the GMRES bounds from [28] and show why the GMRES residual norm often has a rapid initial decrease. In Section 3 we apply these bounds to GMRES for inverse iteration without preconditioning, and with standard, tuned and polynomial preconditioners. Extensions to generalized eigenvalue problems and block GMRES are discussed in Section 4, while numerical experiments can be found in Section 5.

Throughout, k denotes the iteration count for the GMRES algorithm and i is the iteration count for the outer iteration (which is inverse iteration here). In our analysis j will denote the index of eigenvalues and eigenvectors, e.g. $Bz_j = \lambda_j z_j$ and is not to be mixed up with the iteration indices.

2 Analysis of the GMRES convergence bound

In this section we apply results from [28] to the system in (1) to give a lower bound on the number of GMRES iterations required to achieve a given residual reduction and explain why GMRES residuals rapidly decrease in the first iteration when $y_i \approx z_1$, i.e., when y_i is a good approximation of an eigenvector of A. The GMRES convergence bounds we use to achieve this are related to those in [6] for Hermitian problems and to more general results in [12] for non-Hermitian generalized eigenvalue problems.

We begin by stating the relevant results from [28]. The key idea of these bounds is to express y_i in the eigenvector basis. Accordingly we introduce

$$w^{[i]} = Z^{-1} y_i / \|y_i\|_2 = \begin{bmatrix} w_1^{[i]} & w_2^{[i]} & \ldots & w_n^{[i]} \end{bmatrix}^T = \begin{bmatrix} w_1^{[i]} & w^{(2)[i]} \end{bmatrix}^T,$$

where $w^{(2)[i]} \in \mathbb{C}^{n-1}$. If $y_i \approx z_1$ it is reasonable to expect that $|w_1^{[i]}| \gg \|w^{(2)[i]}\|_2$ and this is observed in practice (see results in corresponding columns of Tables 2–4). Note that in the rest of this section, we drop the outer iteration index i from w and related quantities for clarity. For simplicity, let us assume that the initial guess for GMRES is the zero vector, so that $r_0 = y_i$. Since y_i is normalized at every step to have unit norm, $\|r_0\| = \|y_i\| = 1 \; \forall i$.

We are now in a position to recall Theorem 2.2 from [28], which states that

$$\|r_k\|_2 \le \|Z\|_2 \min_{\substack{q \in \Pi_k, \\ q(0)=1}} \|Wq(\Lambda)e\|_2 = \|Z\|_2 \min_{\substack{q \in \Pi_k, \\ q(0)=1}} \left(\sum_{j=1}^{n} |w_j q(\lambda_j)|^2 \right)^{\frac{1}{2}}, \qquad (2)$$

where $W = \mathrm{diag}\,(w_1, \ldots, w_n)$. This bound highlights the role of the eigenvalues of B weighted by the corresponding component of w in the GMRES convergence. In particular, whenever $\|w^{(2)}\|_2$ is small enough, $|w_j q(\lambda_j)| \ll |w_1 q(\lambda_1)|$, $j = 2, \ldots, n$ and the minimizing polynomial will focus first on $|w_1 q(\lambda_1)|$. Thus, fast convergence of GMRES is linked both with the spectrum of B and with the quality of the eigenvector approximation y_i.

A slightly different bound is obtained if, similarly to [6, 12], we replace the minimizing polynomial in (2) by $\tilde{q}_k(\lambda) = (1 - \lambda/\lambda_1) q_{k-1}(\lambda)$, where q_{k-1} is a polynomial of degree $k-1$ for which $q_{k-1}(0) = 1$. This is subject of Theorem 2.3 in [28], which states that for $k > 1$,

$$\|r_k\|_2 \le \|Z\|_2 \min_{\substack{q \in \Pi_{k-1}, \\ q(0)=1}} \left(\sum_{j=2}^{n} |\tilde{w}_j q(\lambda_j)|^2 \right)^{\frac{1}{2}} \qquad (3a)$$

$$\le \|Z\|_2 \|\tilde{w}\|_2 \min_{\substack{q \in \Pi_{k-1}, \\ q(0)=1}} \max_{\lambda \in \{\lambda_2, \ldots, \lambda_n\}} |q(\lambda)|, \qquad (3b)$$

where $\tilde{w}_j = w_j(1 - \lambda_j/\lambda_1)$, $j = 2, \ldots, n$. Starting from (2), a similar approach to the proof of (3) shows that, additionally,

$$\|r_1\|_2 \le \|Z\|_2 \left(\sum_{j=2}^{n} |\tilde{w}_j|^2 \right)^{\frac{1}{2}} = \|Z\|_2 \|\tilde{w}\|_2. \qquad (4)$$

In contrast to (2), neither (3) nor (4) involves w_1 and so they emphasize the fact that the relative residuals may be small when $\|\tilde{w}\|_2$ is small. (Note that we may always normalize $\|Z\|_2$, but that this affects w, and hence $\|\tilde{w}\|_2$).

3

2.1 Inner iterations vs outer iterations

In [12] it was shown that the term

$$\min_{\substack{q\in\Pi_{k-1},\ \lambda\in\{\lambda_2,...,\lambda_n\} \\ q(0)=1}} \max |q(\lambda)| \tag{5}$$

can often be bounded by an expression of the form

$$S(1/C)^{k-1} \tag{6}$$

(by, e.g. [12, Proposition 2.3]) where C and S depend on the spectrum of B. The authors of [12] used this bound to show that the number of GMRES iterations required to reach a desired tolerance, i.e., to satisfy $\|r_k\|_2/\|r_0\|_2 \leq \tau$, can also be bounded. Combining (3) and (6) gives us a simple alternative lower bound on the number of GMRES iterations:

$$k \geq 1 + \frac{1}{\log(C)}\left[\log(S) + \log\left(\frac{\|Z\|_2\|\tilde{w}\|_2}{\tau}\right)\right]. \tag{7}$$

Since both C and S depend only on the spectrum of A and the shift σ, they are independent of the inverse iteration step i. Thus, if we can bound $\|Z\|_2\|\tilde{w}\|_2/\tau$ independently of the inverse iteration step then the above bound on k suggests that the number of GMRES iterations should not increase as inverse iteration proceeds. We will observe that this occurs if GMRES is applied either without a preconditioner, or with specially chosen preconditioners (see Tables 2 and 3).

2.2 Initial decrease in GMRES residuals

Another phenomenon that often arises when solving linear systems with approximate eigenvectors is that $\|r_1\|_2 \ll \|r_0\|_2$ (see e.g. Figure 1). The bound (3) shows clearly why this occurs.

First note that

$$|\tilde{w}_j| = \left|1 - \frac{\lambda_j}{\lambda_1}\right||w_j| = \frac{|\gamma_1 - \gamma_j|}{|\gamma_1 - \sigma|}|w_j| \leq |w_j| \max_{j\in[2,...,n]} \frac{|\gamma_1 - \gamma_j|}{|\gamma_1 - \sigma|} = C_1|w_j|,$$

where the constant C_1 depends only on the spectrum of A and the shift σ. Accordingly, $\|\tilde{w}\|_2 \leq C_1\|w^{(2)}\|$, and (4) becomes

$$\|r_1\|_2 \leq C_2\|w^{(2)}\|_2, \tag{8}$$

where $C_2 = \|Z\|_2 \max_{j\in[2,...,n]} |\lambda_1 - \lambda_j|/|\lambda_1 - \sigma|$.

However, the values of w_j, $j = 2,\ldots n$ are very small, and indeed are zero if the right-hand side is an exact eigenvector. Hence, in the latter stages of inverse iteration there is a sharp decrease in the bound for the relative GMRES residual norms. Of course, this is only true if the right-hand side is an approximate eigenvector of B, i.e. in the unpreconditioned case, or with specially chosen preconditioners. An example of this phenomenon is given in Figure 1.

3 Convergence bounds for preconditioned GMRES

Having considered a general analysis for GMRES convergence that involves the right-hand side vector, we now investigate what this analysis tells us about solving the linear systems in inverse iteration, both with and without preconditioning.

3.1 No preconditioner

When GMRES is applied to (1) without a preconditioner, the coefficient matrix is $B = A - \sigma I$, $A \in \mathbb{C}^{n \times n}$ where, assuming A is diagonalizable, $B = Z\Lambda Z^{-1}$ with $\Lambda \in \mathbb{C}^{n \times n}$. The right-hand side y_i is an approximation of z_1 (and both vectors are normalized to have unit norm). Thus it is possible to write

$$y_i = \alpha_i z_1 + Z_2 p_i, \tag{9}$$

where $\alpha_i \in \mathbb{C}$, $Z_2 = [z_2, \ldots, z_n] \in \mathbb{C}^{n \times n-1}$ and $p_i \in \mathbb{C}^{n-1}$. We assume that α_i and p_i are chosen such that $\|y_i\| = 1 \ \forall i$. The scalar α_i measures the deviation of y_i from z_1 and, under generic assumptions for exact solves, inverse iteration converges, that is $\alpha_i \to 1$ and $\|p_i\| \to 0$ as $i \to \infty$.

The bound in (2), and in particular the vector w, depends on the outer iteration i. Using (9) we obtain

$$
\begin{aligned}
w^{[i]} &= Z^{-1} y_i = Z^{-1}(\alpha_i z_1 + Z_2 p_i) \\
&= \alpha_i e_1 + E_{n-1} p_i,
\end{aligned}
$$

where $e_1 \in \mathbb{R}^n$ and $E_{n-1} \in \mathbb{R}^{n,n-1}$ contain the first and the last $n-1$ columns of the identity matrix, respectively. Recall that $w^{[i]}$ denotes the vector w at outer iteration i. With $\alpha_i \to 1$ and $\|p_i\| \to 0$ as $i \to \infty$ for a convergent outer iteration we obtain $w^{[i]} \to e_1$. Hence in the limit $q(\lambda)$ only needs to be minimized over λ_1.

Noting that $\|Z\|_2$ can be normalized, $\|w^{[i]}\|_2 \to 1$ and $\|\tilde{w}^{[i]}\|_2 \to 0$, we see that the terms $\|Z\|_2 \|w^{[i]}\|_2$ and $\|Z\|_2 \|\tilde{w}^{[i]}\|_2$ in (2) and (3) can be bounded by an arbitrarily small constant for large enough i. Therefore, as the outer iteration progresses, small relative GMRES residual norms are rapidly obtained. Experimentally we see that the number of inner iterations is bounded because $\|\tilde{w}^{[i]}\|_2 \to 0$ at a rate proportional to the decrease of τ. This is reflected by the lower bound in (7), which is constant if $\|\tilde{w}^{[i]}\|_2 \propto \tau$ (cf. Lemma 3.11 in [9]). As the outer iterations progress (i.e. for larger i) there is an initial decrease in the relative GMRES residual norms, as suggested by (8), because $\|\tilde{w}^{[i]}\|_2 \approx 0$.

3.2 Standard preconditioner

We now investigate the changes that occur when a preconditioner P is introduced. Usually GMRES is used with a preconditioner to cluster the eigenvalues of the system matrix. The aim of this is to reduce the bound on the term (5) and hence improve the convergence bound (3). We will see that this usually comes at the expense of increasing $\|\tilde{w}\|_2$, so that the number of GMRES iterations actually grows as inverse iteration proceeds (although this number is generally still lower than the number of GMRES iterations needed without a preconditioner).

Without loss of generality we consider a right preconditioner for (1), that is

$$BP^{-1}\tilde{x}_i = y_i, \tag{10}$$

where $P^{-1}\tilde{x}_i = x_i$. Although the eigenvalues of BP^{-1} may be clustered, the eigenvectors will typically differ from those of B. Thus, although y_i is an approximate eigenvector of B, it is not usually an approximate eigenvector of the coefficient matrix BP^{-1}. This causes the number of GMRES iterations to increase as inverse iteration progresses, as we now show.

In general we have $BP^{-1} \neq ZDZ^{-1}$ (with D a diagonal matrix), that is, BP^{-1} does not have the same eigenvectors as B. Hence, with $BP^{-1} = \bar{Z}D\bar{Z}^{-1}$ and using the decomposition of y_i we obtain

$$w^{[i]} = \bar{Z}^{-1} y_i = \alpha_i \bar{Z}^{-1} z_1 + \bar{Z}^{-1} Z_2 p_i \to \alpha_i \bar{Z}^{-1} z_1,$$

as $i \to \infty$ since $\|p_i\| \to 0$, but $\alpha_i \bar{Z}^{-1} z_1 \neq e_1$ in general.

Hence, there is no reason for $\|w^{(2)^{[i]}}\|_2$ to be small. Instead, $\|w^{(2)^{[i]}}\|_2 \to c$ for some constant c as i increases (see corresponding column in Table 2) and we observe that for some problems $\|w^{(2)^{[i]}}\|_2 > |w_1^{[i]}|$, i.e., $c > 1$. Since $\|w^{(2)^{[i]}}\|_2$ does not approach 0, but τ decreases as the outer iteration progresses, we would expect from (7) that the number of inner iterations increases with increasing i, and indeed this is what our experiments in Section 5 show. Moreover, we find that $\|r_1\|_2$ is not significantly smaller than $\|r_0\|_2$ at later outer iterations. However, for large enough i the terms $\|Z\|_2 \|w^{[i]}\|_2$ or $\|Z\|_2 \|\tilde{w}^{[i]}\|_2$ can be bounded by constants that are are still small relative to $\kappa_2(Z) = \|Z\|_2 \|Z^{-1}\|_2$, and the bounds (2) and (3) can still capture GMRES convergence behavior, especially at later outer iterations (see Figure 1).

To recover bounded numbers of inner iterations when preconditioning we must ensure that the preconditioner preserves the relevant direction given by the right-hand side. For this, we may construct a preconditioner such that

(a) $BP_i^{-1} = \bar{Z}_i D_i \bar{Z}_i^{-1}$ such that (in the limit for large i) $\bar{Z}_i^{-1} z_1 \to e_1$, or

(b) BP^{-1} has the same eigenvectors as B.

The next two subsections show how these preconditioners can be constructed.

3.3   Tuned preconditioner

Assume we have an a preconditioner P such that BP^{-1} is diagonalizable. Then $BP^{-1} = Z\Lambda Z^{-1} P^{-1} = \bar{Z} D \bar{Z}^{-1}$, where D is a diagonal matrix. Since we want to recover bounded GMRES iteration counts, we want y_i to be an approximate eigenvector of BP^{-1}. To achieve this, it is not necessary that BP^{-1} has all the same eigenvectors as B, but it should have the same eigenvector in the direction of the sought eigenvector z_1, that is, we want to enforce $\bar{Z}^{-1} z_1 = e_1$.

If we enforce this condition, we obtain

$$Z\Lambda Z^{-1} P^{-1} z_1 = \bar{Z} D \bar{Z}^{-1} z_1 = \bar{Z} D e_1 = d_{11} z_1.$$

Since $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n) = \text{diag}(\gamma_1 - \sigma, \ldots, \gamma_n - \sigma)$, and assuming that $\gamma_j \neq \sigma$, $\forall j$, the above requirement shows that P^{-1} needs to satisfy

$$P^{-1} z_1 = \frac{d_{11}}{\lambda_1} z_1. \tag{11}$$

Hence, choosing a preconditioner P such that (11) holds, or equivalently

$$P z_1 = \frac{\lambda_1}{d_{11}} z_1,$$

yields $\bar{Z}^{-1} z_1 = e_1$. Note that $d_{11} \neq 0$ (otherwise $\lambda_j = 0$ and $\gamma_j = \sigma$ for some j, a case we excluded).

Clearly, (11) is infeasible as a preconditioner, since z_1, λ_1 and d_{11} are unknown during the iteration. Hence, we propose to use a preconditioner P_i at each iteration i that satisfies

$$P_i y_i = \lambda^{(i)} y_i = (\gamma^{(i)} - \sigma) y_i, \tag{12}$$

where $\gamma^{(i)}$ is an eigenvalue approximation for γ_1; it can be obtained as part of the the inverse iteration algorithm.

Remark 1. Instead of (12) one could use

- $P_i y_i = y_i$, since the direction is important for the preconditioner;

- $P_i y_i = B y_i$, since $B y_i = (\gamma^{(i)} - \sigma) y_i + r_i = \lambda^{(i)} y_i + r_i$ with $r_i = (A - \gamma^{(i)} I) y_i$ the eigenvalue residual, and in the limit (12) holds;

- $P_i y_i = A y_i$, since $A y_i = \gamma^{(i)} y_i + r_i$, which, in the limit, lies in the same direction.

The action of P_i^{-1} on a vector v can easily be obtained as an update of $P^{-1} v$ using the Sherman-Morrison formula, that is

$$P_i^{-1} v = P^{-1} v - \frac{P^{-1} y_i - (\lambda^{(i)})^{-1} y_i}{y_i^T P^{-1} y_i} y_i^T P^{-1} v.$$

This can be implemented using only one extra linear solve with P (to compute $P^{-1} y_i$) per outer iteration. Note that we assume that $\lambda^{(i)} \neq 0$.

Theorem 2. Consider inverse iteration, where at each inner iteration a preconditioned system of the form $B P_i^{-1} \tilde{x}_i = y_i$ is solved with $P_i^{-1} \tilde{x}_i = x_i$, and P_i satisfies (12). Assume that $B P_i^{-1}$ is diagonalizable $\forall i$, that is $B P_i^{-1} = Z_i D_i Z_i^{-1}$ with D_i diagonal. Further, assume that B is nonsingular, that is $\lambda^{(i)} = \gamma^{(i)} - \sigma \neq 0 \ \forall i$. Then

(i) $B P_i^{-1} y_i = y_i + \frac{r_i}{\gamma^{(i)} - \sigma}$, where $r_i = (A - \gamma^{(i)} I) y_i$ is the eigenvalue residual (that is, in the limit $B P_i^{-1}$ has an eigenvalue at 1);

(ii) $w^{[i]} \to e_1$ as $i \to \infty$.

Proof. (i) Using the definition of P_i^{-1} we obtain

$$B P_i^{-1} y_i = (\gamma^{(i)} - \sigma)^{-1} B y_i = y_i + \frac{r_i}{\gamma^{(i)} - \sigma},$$

where we have used the fact that $B y_i = (A - \sigma I) y_i = (\gamma^{(i)} - \sigma) y_i + r_i$.

(ii) We need to compute $w^{[i]} = Z_i^{-1} y_i$ as $i \to \infty$. Consider the eigendecomposition $B P_i^{-1} = Z_i D_i Z_i^{-1}$ as $i \to \infty$. Then $y_i \to z_1$ and, using part (i) we have

$$B P_i^{-1} y_i \to z_1,$$

since $r_i \to 0$. Therefore, in the limit $Z_i \to [z_1, \tilde{Z}_2]$, as $i \to \infty$, for some $\tilde{Z}_2$ and hence,

$$w^{[i]} = Z_i^{-1} y_i \to [z_1, \tilde{Z}_2]^{-1} z_1 = e_1.$$

$\square$

We obtain $w^{[i]} \to e_1$. Since $\|Z\|_2$ can be normalized, $\|w^{[i]}\|_2 \to 1$, and $\|\tilde{w}^{[i]}\|_2 \to 0$, the terms $\|Z\|_2 \|w^{[i]}\|_2$ and $\|Z\|_2 \|\tilde{w}^{[i]}\|_2$ in (2) and (3) can be bounded by small constants for large enough i. Experimentally we see that the number of inner iterations is bounded because $\|w^{[i]}\|_2 \to 0$ at a rate proportional to τ (cf. Theorem 6.22 in [9]). Additionally, as the outer iterations progress there is an initial decrease in the relative GMRES residual norms, that is, $\|r_1\|_2 \ll \|r_0\|_2$, because $\|\tilde{w}^{[i]}\|_2 \approx 0$.

3.4 Polynomial preconditioner

As an alternative to tuning we can construct a preconditioner such that $B P^{-1}$ has the same eigenvectors as B. If we choose

$$P^{-1} = p(B),$$

7

where $p(B)$ is a polynomial of B, then $p(B) = Zp(\Lambda)Z^{-1}$ and $BP^{-1} = Z\Lambda p(\Lambda)Z^{-1}$ has the same eigenvectors as B. In this case the theory for the unpreconditioned case applies.

We consider the (right) polynomially preconditioned system $Bp(B)\tilde{x}_i = y_i$, where $x_i = p(B)\tilde{x}_i$ and

$$p(z) = \sum_{h=0}^{d} \mu_h z^h \in \Pi_d.$$

Since $p(B) = Zp(\Lambda)Z^{-1}$, y_i is still an approximate eigenvector of the preconditioned coefficient matrix $Bp(B)$ and it holds that

$$\|r_k\|_2 \le \|Z\|_2 \min_{\substack{q_k \in \Pi_k \\ q_k(0)=1}} \left(\sum_{j=1}^{n} |w_j q_k(\lambda_j p(\lambda_j))|^2 \right)^{\frac{1}{2}}.$$

Hence, all the weights w of the unpreconditioned version are preserved. Typically, $p(B)$ is chosen such that

$$B(p(B)) \approx I \quad \text{or} \quad g(B) := I - Bp(B) \approx 0.$$

In the latter choice $g \in \Pi_{d+1}$ is the residual polynomial, which can be written as

$$g(z) = \prod_{h=1}^{d+1} (1 - \nu_h z) = 1 - \sum_{h=1}^{d+1} \mu_{h-1} z^h.$$

There are different strategies to choose the polynomial coefficients μ_h, which can be determined from the ν_h recursively via

$$g_h(z) = g_{h-1}(z) - \nu_h z g_{h-1}(z), \quad h = 1, \ldots, d+1, \ g_0(z) = 1,$$

see, e.g. [1] for more details. A thorough discussion of various choices for the polynomial coefficients is beyond the scope of this study and we restrict ourselves to a few selected, often used approaches. More information on the choice of the polynomial can be found, e.g., in [13, 3, 29, 1, 4, 16] and the references therein.

A basic common choice [1] is to take the ν_h as reciprocal Chebychev nodes with respect to an interval $[a, b]$:

$$\phi_h = \frac{2h - 1}{2(d + 1)}, \quad \nu_h = 2 \left(b + a - (b - a) \cos(\pi \phi_h) \right)^{-1}, \quad h = 1, \ldots, d+1. \quad (13)$$

Here, a, b denote the smallest and largest eigenvalues (or approximations thereof) of $\Lambda(B) \subset \mathbb{R}$. For complex spectra, a, b are the foci of the ellipse enclosing $\Lambda(B)$. The values a, b can be obtained from approximate eigenvalues of A or B, e.g. by employing `eigs` with the 'SM' and 'LM' flags, or using the Ritz values generated by the Arnoldi process. Alternatively one can compute a very small number of (harmonic) Ritz values [29, 1], which can be generated by a few steps of an (harmonic) Arnoldi process. When using this approach, it is important to ensure that a, b have the same sign but in some of our examples this not the case and either a or b lie a little bit on the wrong side of the origin. In these cases we slightly shifted a, b, e.g. if $b > 0$ but $a < 0$ is small we set $b \leftarrow b - 2a$, $a \leftarrow -a$.

For complex spectra a more sophisticated approach is proposed in [29]. The key idea is to consider the norm

$$\|g\|^2 = \frac{1}{L} \int_{\Gamma} g(z)\overline{g(z)} |\mathrm{d}z| \quad (14)$$

induced by the scalar product $\langle f, g \rangle = \frac{1}{L} \int_{\Gamma} f(z)\overline{g(z)}|\mathrm{d}z|$. Here Γ is a piecewise linear contour approximating the shape of the spectrum of B and L is its arc length. The coefficients μ_i in the polynomial are determined such that the above norm is minimized in a least-squares sense. Using (harmonic) Ritz values allows us to easily generate the piecewise linear contour, e.g. using the convex hull of the Ritz values. However, this approach also allows us to approximate the spectrum by non-convex objects such as a boomerang-shaped region, which can be beneficial in some cases. Note that if B is real, it is sufficient to incorporate only the part of Γ with nonnegative imaginary parts. More details on the implementation of this strategy are given in [29].

A polynomial preconditioner preserves the eigenvector basis and hence appears to be an ideal preconditioner to use within GMRES for linear systems arising within inverse iteration (using the theory in this paper). However, we would like to note that a limitation of good polynomial preconditioners is that we require knowledge about the location of the spectrum of B. Since we are actually seeking an eigenvalue, the necessary information to construct a good polynomial preconditioner might be hard to obtain. We already mentioned the use of spectral estimates, which can be obtained by, e.g., (harmonic) Arnoldi processes. Since we are using inexact inverse iteration with GMRES as inner solver, one strategy deserving further study is to use the upper Hessenberg matrix generated at outer iteration step $i - 1$ to acquire spectral estimates for outer iteration step i. In a similar way, the generated basis matrices of the Krylov subspaces can be used to acquire the coefficients by solving a least square problem along the lines of [16].

For completeness we mention a third polynomial preconditioning strategy. Since iterative Krylov methods for linear systems work implicitly with polynomials of B, we could in principle use any of these methods as a preconditioner for the inner iteration. In other words, we could use (a small number of steps of) a Krylov method within a Krylov method, which is GMRES here. However, since Krylov methods also depend on the right-hand side, which determines the first basis vector in the Krylov subspace, the preconditioner will change with every iteration of GMRES in our consideration. Therefore, a flexible preconditioned Krylov method, such as flexible GMRES [23] must be applied, which is beyond the scope of this study. Similar approaches can be found in, e.g. [30, 7, 5]. The advantage of this strategy is that one does not have to worry about the generation of the polynomial coefficients.

4 Extensions and related issues

In this section several extensions to inverse iteration for the standard eigenvalue problems are considered. First we extend our analysis of GMRES within inverse iteration to generalized eigenvalue problems. Then, we obtain bounds for block GMRES arising in inverse subspace iteration for the standard eigenvalue problem.

4.1 Generalized eigenvalue problems

Consider inverse iteration for the generalized eigenvalue problem $Ax = \lambda Mx$, where at least one of the matrices A and M is nonsymmetric. We shall also assume that the desired eigenvalue is finite in the case that M is singular. For generalized eigenvalue problems, the linear systems we wish to solve in inverse iteration are of the form $(A - \sigma M)x_i = My_i$, where we normalize y_i so that $\|My_i\|_2 = 1$. The eigenvalue approximation can then be obtained by a generalization of the Rayleigh quotient, e.g. $(x_i^T M^T A x_i)/\|Mx_i\|^2$. Let us now assume that $A - \sigma M$ is diagonalizable, i.e. $A - \sigma M = Z\Lambda Z^{-1}$ where Z is a matrix of eigenvectors and Λ is a diagonal matrix

of eigenvalues. We also assume without loss of generality that y_i approximates z_1, the first column of Z.

Setting

$$w^{[i]} = Z^{-1}y_i = \begin{bmatrix} w_1^{[i]} & w^{(2)\,[i]} \end{bmatrix}^T$$

we would again expect that when y_i is a good approximation of z_1 that $|w_1^{[i]}| \gg \|w^{(2)\,[i]}\|_2$. Again, for clarity, in the rest of this section, we drop the outer iteration index i from w and related quantities where appropriate. The GMRES bounds in this case becomes

$$\|r_k\|_2 = \min_{\substack{q \in \Pi_k, \\ q(0)=1}} \|q(A-\sigma M)My_i\|_2 = \min_{\substack{q \in \Pi_k, \\ q(0)=1}} \|Zq(\Lambda)(Z^{-1}MZ)w\|_2 = \min_{\substack{q \in \Pi_k, \\ q(0)=1}} \|Zq(\Lambda)f\|_2, \tag{15}$$

where $f := Z^{-1}MZw$. Note that if w approximates the first unit vector, then f approximates the first column of $Z^{-1}MZ$. In this case an analogous bound to (2) is

$$\|r_k\|_2 \le \|Z\|_2 \min_{\substack{q \in \Pi_k, \\ q(0)=1}} \|Fq(\Lambda)e\|_2 = \|Z\|_2 \min_{\substack{q \in \Pi_k, \\ q(0)=1}} \left(\sum_{j=1}^{n} |f_j q(\lambda_j)|^2 \right)^{\frac{1}{2}}, \tag{16}$$

where $F := \mathrm{diag}\,(f_1, \ldots, f_n)$. Thus, as in Section 2, we have a weighted polynomial approximation. However, the weights may all be large because My_i is not close to an eigenvector of $A - \sigma M$ in general.

Considering (3), we find that an analogous bound is

$$\|r_k\|_2 \le \|Z\|_2 \min_{\substack{q \in \Pi_{k-1}, \\ q(0)=1}} \left(\sum_{j=2}^{n} |\tilde{f}_j q(\lambda_j)|^2 \right)^{\frac{1}{2}} \le \|Z\|_2 \|\tilde{f}\|_2 \min_{\substack{q \in \Pi_{k-1}, \\ q(0)=1}} \max_{\lambda \in \{\lambda_2, \ldots, \lambda_n\}} |q(\lambda)|, \tag{17}$$

where $\tilde{f}_j = (1 - \lambda_j/\lambda_1)f_j$. Again, there is no reason for $\|\tilde{f}\|_2$ to be small. We conclude that for generalized eigenproblems, unpreconditioned GMRES may not show a residual reduction similar to the case $M = I$.

One way to reintroduce this behavior is by using the tuned preconditioner $P_i = P(I - y_i y_i^H) + My_i y_i^H$, which satisfies $P_i y_i = My_i$ [12]. We stress again that application of the tuned preconditioner requires only one extra matrix-vector product $P^{-1}My_i$ per outer iteration [12]. We also note that in the absence of a good preconditioner, the choice $P = I$ should at least recover the behavior observed for inverse iteration with unpreconditioned GMRES applied to the standard eigenvalue problem. It is also obvious that the nice properties of polynomial preconditioners do not hold for the generalized eigenproblem, since the weight vectors f_j will be the same as in the unpreconditioned case. One could add tuning to a polynomial preconditioner, but from numerical experience this strategy was not competitive.

4.2   Block GMRES

The linear system in (1) can be generalized to a block linear system, that is, a linear system with multiple right-hand sides of the form

$$BX_i = Y_i, \tag{18}$$

where $B = A - \sigma I$ as before, and $Y_i, X_i \in \mathbb{C}^{n \times u}$, $u \ll n$. Systems of this form arise when seeking an invariant subspace using inverse subspace iteration [22], a block version of the inverse iteration. If these systems are solved by block-GMRES, we obtain similar bounds to (2) (see also Theorem 2.2 in [28]), and can gain insight into the convergence behavior of block-GMRES as the inverse subspace iterations progress.

Lemma 3. Suppose that B is diagonalizable, that is $B = Z\Lambda Z^{-1}$, and block-GMRES is used to solve the linear system with multiple right-hand sides of the form $BX = Y$, $Y \in \mathbb{C}^{n \times u}$. The residual $Y - BX_k$ associated with the approximate solution X_k, obtained with k iterations of block-GMRES starting with $X_0 = 0$ is such that

$$\|Y - BX_k\|_F \le \|Z\|_2 \min_{G_i \in \mathbb{C}^{u \times u}} \left\| W + \sum_{i=1}^{k} \Lambda^i W G_i \right\|_F$$

$$= \|Z\|_2 \min_{\substack{q_{\ell,j} \in \Pi_k, \\ q_{\ell,j}(0) = \delta_{\ell,j}}} \left\| \begin{bmatrix} \sum_{\ell=1}^{u} q_{\ell,1}(\Lambda) w_\ell \\ \vdots \\ \sum_{\ell=1}^{u} q_{\ell,u}(\Lambda) w_\ell \end{bmatrix} \right\|_2,$$

where $W = Z^{-1} Y \in \mathbb{C}^{n \times u}$, w_ℓ is the ℓth column of W, Π_k is the space of polynomials of at most degree k and $\delta_{\ell,j}$ is the Kronecker delta.

Proof. Since $X_k \in \mathrm{span}\{Y, BY, \ldots, B^{k-1}Y\}$ and block-GMRES minimizes the residual we have

$$\|Y - BX_k\|_F \le \min_{G_i \in \mathbb{C}^{u \times u}} \left\| Y + \sum_{i=1}^{k} B^i Y G_i \right\|_F$$

$$= \min_{G_i \in \mathbb{C}^{u \times u}} \left\| (I \otimes Z) \, \mathrm{vec} \left(W + \sum_{i=1}^{k} \Lambda^i W G_i \right) \right\|_2$$

$$\le \|Z\|_2 \min_{G_i \in \mathbb{C}^{u \times u}} \left\| W + \sum_{i=1}^{k} \Lambda^i W G_i \right\|_F,$$

where we have used vectorization and the properties of the Kronecker product in the second step. If we introduce the additional matrix $G_0 = I_u$ then

$$\left(W + \sum_{i=1}^{k} \Lambda^i W G_i \right) e_j = \sum_{i=0}^{k} \Lambda^i W G_i e_j = \sum_{i=0}^{k} \Lambda^i \sum_{\ell=1}^{u} g_{\ell j}^{(i)} w_\ell = \sum_{\ell=1}^{u} q_{\ell,j}(\Lambda) w_\ell,$$

where e_j is the jth unit vector, w_ℓ is the ℓth column of W, $g_{\ell j}^{(i)}$ is the (ℓ, j)th element of G_i and $q_{\ell,j}(\lambda) = \sum_{i=0}^{k} \lambda^i g_{\ell j}^{(i)}$, $j, \ell = 1, \ldots, u$. Clearly $q_{\ell,j} \in \Pi_k$. Additionally, the condition $G_0 = I_u$ means that $q_{\ell,j}(0) = \delta_{\ell,j}$.

Thus,

$$\left\| W + \sum_{i=1}^{k} \Lambda^i W G_i \right\|_F = \left\| \begin{bmatrix} \sum_{\ell=1}^{u} q_{\ell,1}(\Lambda) w_\ell, & \ldots, & \sum_{\ell=1}^{u} q_{\ell,u}(\Lambda) w_\ell \end{bmatrix} \right\|_F$$

$$= \left\| \begin{bmatrix} \sum_{\ell=1}^{u} q_{\ell,1}(\Lambda) w_\ell \\ \vdots \\ \sum_{\ell=1}^{u} q_{\ell,u}(\Lambda) w_\ell \end{bmatrix} \right\|_2,$$

from which the result follows. □

A consequence of Lemma 3 is the following block-GMRES bound.

Theorem 4. Suppose that B is diagonalizable, that is $B = Z\Lambda Z^{-1}$, and block-GMRES is used to solve the linear system with multiple right-hand sides of the form $BX = Y$, $Y \in \mathbb{C}^{n \times u}$. The residual $Y - BX_k$ associated with the approximate

solution X_k, obtained with k iterations of block-GMRES starting with $X_0 = 0$, is such that

$$\|Y - BX_k\|_2 \le \|Z\|_2 \min_{\substack{q \in \Pi_k, \\ q(0)=1}} \left(\sum_{\ell=1}^{u} \sum_{j=1}^{n} |w_{j\ell} q(\lambda_j)|^2 \right)^{\frac{1}{2}} \tag{19}$$

where $Z^{-1}Y = W$, and $w_{j\ell}$ is the (j, ℓ)th entry of W.

Proof. In Lemma 3, replace the minimizing polynomial q by $\hat{q}$, where $\hat{q}_{\ell,j}(\lambda) \equiv 0$, $j \neq \ell$ and

$$\hat{q}_{11}(\lambda) = \hat{q}_{22}(\lambda) = \cdots = \hat{q}_{uu}(\lambda) \equiv \hat{q}(\lambda).$$

(Note that this is equivalent to choosing $G_i = g_i I$, g_i constant, $i = 1, \ldots, k$.) Then,

$$\|Y - BX_k\|_F \le \|Z\|_2 \min_{\substack{q \in \Pi_k, \\ q(0)=1}} \left\| \begin{bmatrix} q(\Lambda)w_1 \\ \vdots \\ q(\Lambda)w_u \end{bmatrix} \right\|_2 = \|Z\|_2 \min_{\substack{q \in \Pi_k, \\ q(0)=1}} \left(\sum_{\ell=1}^{u} \sum_{j=1}^{n} |w_{j\ell} q(\lambda_j)|^2 \right)^{\frac{1}{2}}.$$

The result follows from $\|Y - BX_k\|_2 \le \|Y - BX_k\|_F$. $\qquad\square$

The key idea in Lemma 3 and Theorem 4 is, again, to express the right-hand side Y in the eigenvector basis of B, i.e.,

$$W = Z^{-1}Y = \begin{bmatrix} W_1 \\ W_2 \end{bmatrix},$$

where $W_1 \in \mathbb{C}^{u \times u}$ and $W_2 \in \mathbb{C}^{(n-u) \times u}$. We write the right-hand side $Y_i \in \mathbb{C}^{n \times u}$ in the form

$$Y_i = Z_1 Y_i^1 + Z_2 Y_i^2,$$

where $Z_1 \in \mathbb{C}^{n \times u}$ and $Z_2 \in \mathbb{C}^{n \times (n-u)}$ are the first u and the last $n-u$ columns of Z, that is $B = [Z_1 \, Z_2] \Lambda [Z_1 \, Z_2]^{-1}$, and $Y_i^1 \in \mathbb{C}^{u \times u}$ is nonsingular and $Y_i^2 \in \mathbb{C}^{(n-u) \times u}$. In the generic situation, we have that inverse subspace iteration converges, e.g. $\mathrm{ran}(Y_i)$ converges to $\mathrm{ran}(Z_1)$ and hence $\|Y_i^2\| \to 0$. Hence for inverse subspace iteration

$$W^{[i]} = Z^{-1}Y_i = Z^{-1}(Z_1 Y_i^1 + Z_2 Y_i^2) = E_1 Y_i^1 + E_2 Y_i^2 = \begin{bmatrix} W_1^{[i]} \\ W_2^{[i]} \end{bmatrix},$$

where E_1 and E_2 are the first u and the last $n - u$ columns of the identity matrix respectively. As $\|Y_i^2\| \to 0$ we have, in the limit,

$$W^{[i]} \approx \begin{bmatrix} W_1^{[i]} \\ 0 \end{bmatrix},$$

where $W_1^{[i]} \in \mathbb{C}^{u \times u}$ and otherwise $\|W_1^{[i]}\| \gg \|W_2^{[i]}\|$, similar to the case where $u = 1$ in the main part of this paper. In light of Lemma 3 and Theorem 4 this means that a lot of entries of $w_\ell^{[i]}$, the ℓth column of $W^{[i]}$ are small or zero. In the limit the bound in (19) becomes

$$\|Y_i - BX_k\|_2 \le \|Z\|_2 \min_{\substack{q \in \Pi_k, \\ q(0)=1}} \left(\sum_{\ell=1}^{u} \sum_{j=1}^{u} |w_{j\ell}^{[i]} q(\lambda_j)|^2 \right)^{\frac{1}{2}}$$

and hence the minimizing polynomial will focus on minimizing over the relevant sought spectrum, e.g. $q(\lambda_1) \ldots, q(\lambda_u)$.

12

This property is violated when a preconditioner is applied in block GMRES, but can be overcome, in a similar way to the case $u = 1$, by a tuned preconditioner, e.g. by using a preconditioner P_i which satisfies $P_i Y_i = B Y_i$, see [22] for details.

It is also possible to use block GMRES within inverse subspace iteration for generalized eigenvalue problems. In this case the linear system that must be solved is of the form $(A - \sigma M)X_i = MY_i$, where again $X_i, Y_i \in \mathbb{C}^{n \times u}$. Assuming that $A - \sigma M = Z \Lambda Z^{-1}$ is diagonalizable, our results carry over to this case, and we obtain the following theorem.

Corollary 5. Suppose that $A - \sigma M$ is diagonalizable, that is $A - \sigma M = Z \Lambda Z^{-1}$, and block-GMRES is used to solve the linear system with multiple right-hand sides of the form $(A - \sigma M)X = MY$, $Y \in \mathbb{C}^{n \times u}$. The residual $MY - (A - \sigma M)X_k$ associated with the approximate solution X_k, obtained with k iterations of block-GMRES starting with $X_0 = 0$, is such that

$$\|MY - (A - \sigma M)X_k\|_2 \leq \|Z\|_2 \min_{\substack{q \in \Pi_k, \\ q(0) = 1}} \left(\sum_{\ell=1}^{u} \sum_{j=1}^{n} |h_{j\ell} q(\lambda_j)|^2 \right)^{\frac{1}{2}} \tag{20}$$

where $Z^{-1}MY = H$, and $h_{j\ell}$ is the (j, ℓ)th entry of H.

Proof. The proof is the same as the proofs of Lemma 3 and Theorem 3 with W replaced by H throughout. $\qquad\square$

As in Section 4.1, we see that although the block GMRES residual can be expressed in terms of a weighted polynomial approximation, there is no reason for any of the weights to be small. However, tuning can also be applied within block GMRES to accelerate convergence, as in [32].

5   Numerical experiments

In this section we consider GMRES convergence within inverse iteration and the bounds previously discussed for nonsymmetric matrices whose properties are summarized in Table 1. The matrix `cd_fd` is taken from [14, 9], whereas `cdde1` and `olm2000` are from the matrix market. In Table 1, the spectral norms and condition numbers of the eigenvector matrices of the unpreconditioned B as well as BP^{-1} for different standard preconditioners are listed. The standard preconditioners P are incomplete LU factorizations with three different drop tolerances θ. To mimic a similar increasing quality of the polynomial preconditioners, the polynomial degrees are set to $d = 5, 10, 15$. (Note that the polynomial preconditioned matrices have the same eigenvector matrix as B.) For the coefficients of $p(B)$, we first generate a small number of Ritz and harmonic Ritz values of B. If these Ritz values are all real or only have small imaginary parts, the coefficient generation (13) based on the reciprocal Chebychev nodes is used, whereas the approach (14) by [29] is used in the case of complex Ritz values. The standard and polynomial preconditioners are kept unchanged during the outer iteration, i.e. they are computed only once at the start. We employ two variants of the tuned preconditioners, which satisfy $P_i y_i = y_i$ and $P_i y_i = A y_i$. We also investigated tuning with $P_i y_i = (\gamma^{(i)} - \sigma) y_i$ as in (12) but the results were similar to the simpler $P_i y_i = y_i$ and so have not been included.

Note that we selected test examples of comparably small sizes in order to be able to compute the eigendecompositions needed for the weight vectors w. The effects and performance gains resulting from the application of tuned preconditioners have been demonstrated with large matrices, e.g. in [27, 17].

Table 1: Matrices used in examples.

Matrix	n	Unpreconditioned		P (10^{-1})		P (10^{-2})		P (10^{-3})	
		$\|Z\|_2$	$\kappa_2(Z)$	$\|Z\|_2$	$\kappa_2(Z)$	$\|Z\|_2$	$\kappa_2(Z)$	$\|Z\|_2$	$\kappa_2(Z)$
`cd_fd`	1024	4.6	1.2e+03	7.6	1.5e+14	5.7	3.4e+06	6.9	7.1e+04
`cdde1`	961	2.8	17	7.4	5.5e+10	5.2	1e+05	4.3	1.8e+03
`olm2000`	2000	1.4	81	11	5.1e+05	7	3.1e+04	2.7	1.3e+03

Table 2: `cdde1` without preconditioning, with untuned ILU preconditioning (drop tolerance 10^{-2}) and with polynomial preconditioning (polynomial degree 10): inverse iteration residual $\|\rho_{i-1}\|_2$, sizes of components of w, constants in (2) and (3) and GMRES iteration numbers at each outer iteration i.

	i	$\lambda^{(i)}$	$\|\rho_{i-1}\|_2$	$|w_1|$	$\|w^{(2)}\|_2$	$\|Z\|_2\|w\|_2$	$\|Z\|_2\|\tilde{w}\|_2$	$\|Z\|_2\|\tilde{w}\|_2/\tau$	it
Unprec.	1	0.099735	0.0033	1.1	1.4	5.1	–	–	54
	2	-0.0028427	0.00095	0.9	0.16	2.6	3.7	1.1e+05	61
	3	-0.0059742	0.00027	1	0.048	2.9	0.65	6.9e+04	51
	4	-0.0049326	8e-05	0.99	0.012	2.8	0.16	6.1e+04	41
	5	-0.0052396	2.2e-05	1	0.0035	2.8	0.045	5.7e+04	34
	6	-0.0051539	6.1e-06	1	0.00094	2.8	0.012	5.6e+04	32
	7	-0.0051775	1.7e-06	1	0.00026	2.8	0.0034	5.5e+04	28
	8	-0.005171	4.6e-07	1	7e-05	2.8	0.00092	5.5e+04	25
Untuned	1	0.099735	0.0033	63	1.4e+02	8.1e+02	–	–	8
	2	-0.0028504	0.00095	3	4.4	28	5	1.5e+05	11
	3	-0.0059693	0.00027	2.7	4.1	26	5.6	5.9e+05	12
	4	-0.0049345	7.9e-05	2.5	4	24	5.4	2e+06	13
	5	-0.0052389	2.2e-05	2.5	4	25	5.5	6.9e+06	14
	6	-0.0051541	6e-06	2.5	4	25	5.5	2.5e+07	15
	7	-0.0051775	1.6e-06	2.5	4	25	5.5	9.1e+07	16
	8	-0.0051711	4.5e-07	2.5	4	25	5.5	3.3e+08	17
Poly. prec.	1	0.099735	0.0033	1.1	1.4	5.1	–	–	9
	2	-0.0028791	0.00094	0.91	0.16	2.6	3.6	1.1e+05	9
	3	-0.0059578	0.00027	1	0.047	2.9	0.64	6.8e+04	6
	4	-0.0049379	7.8e-05	0.99	0.012	2.8	0.16	6e+04	5
	5	-0.005238	2.2e-05	1	0.0034	2.8	0.044	5.7e+04	4
	6	-0.0051543	5.9e-06	1	0.00092	2.8	0.012	5.6e+04	4
	7	-0.0051774	1.6e-06	1	0.00025	2.8	0.0033	5.5e+04	4
	8	-0.0051711	4.5e-07	1	6.9e-05	2.8	0.0009	5.5e+04	3

For each problem we run inverse iteration with an initial eigenvector approximation of $y_0 = \frac{1}{n}[1, \ 1, \ \ldots, \ 1]^T$. We let ρ_i be the inverse iteration residual, so that

$$\rho_i = Ax_i - \lambda_i x_i. \qquad (21)$$

The shifts σ are as in Chapter 6 of [9]. We solve the linear system using GMRES with a zero initial guess and a decreasing tolerance of $\tau = \min\{\delta, \delta\|\rho_{i-1}\|_2\}$. We choose $\delta = 0.1$ except for `cdde1` ($\delta = 0.001$); these are the same values as in Chapter 6 of [9].

At first we have a detailed look at the progress of the outer iteration for the matrix `cdde1` using no preconditioner, ILU with $\theta = 10^{-2}$ and the corresponding two tuned variants, and a polynomial preconditioner with $\deg(p) = d = 10$. Tables 2–3 summarize the changes in the relevant quantities as the outer iteration proceeds, including the quality of the eigenpair approximations, the components in $w^{[i]}$, the constants in the GMRES bounds (2) and (3), and the number of executed GMRES steps. For the tuned preconditioners, also the progress of the spectral norms and condition numbers of the eigenvector matrices is given. Note that for clarity we drop the outer iteration index i from w for the remainder of this section.

Obviously, the values of $\lambda^{(i)}$, $\|\rho_i\|_2$ at each outer iteration reveal that using different preconditioners does not hamper the convergence of the eigenpairs. It is also apparent that $w \to e_1$, as the outer iteration i proceeds, in the unpreconditioned
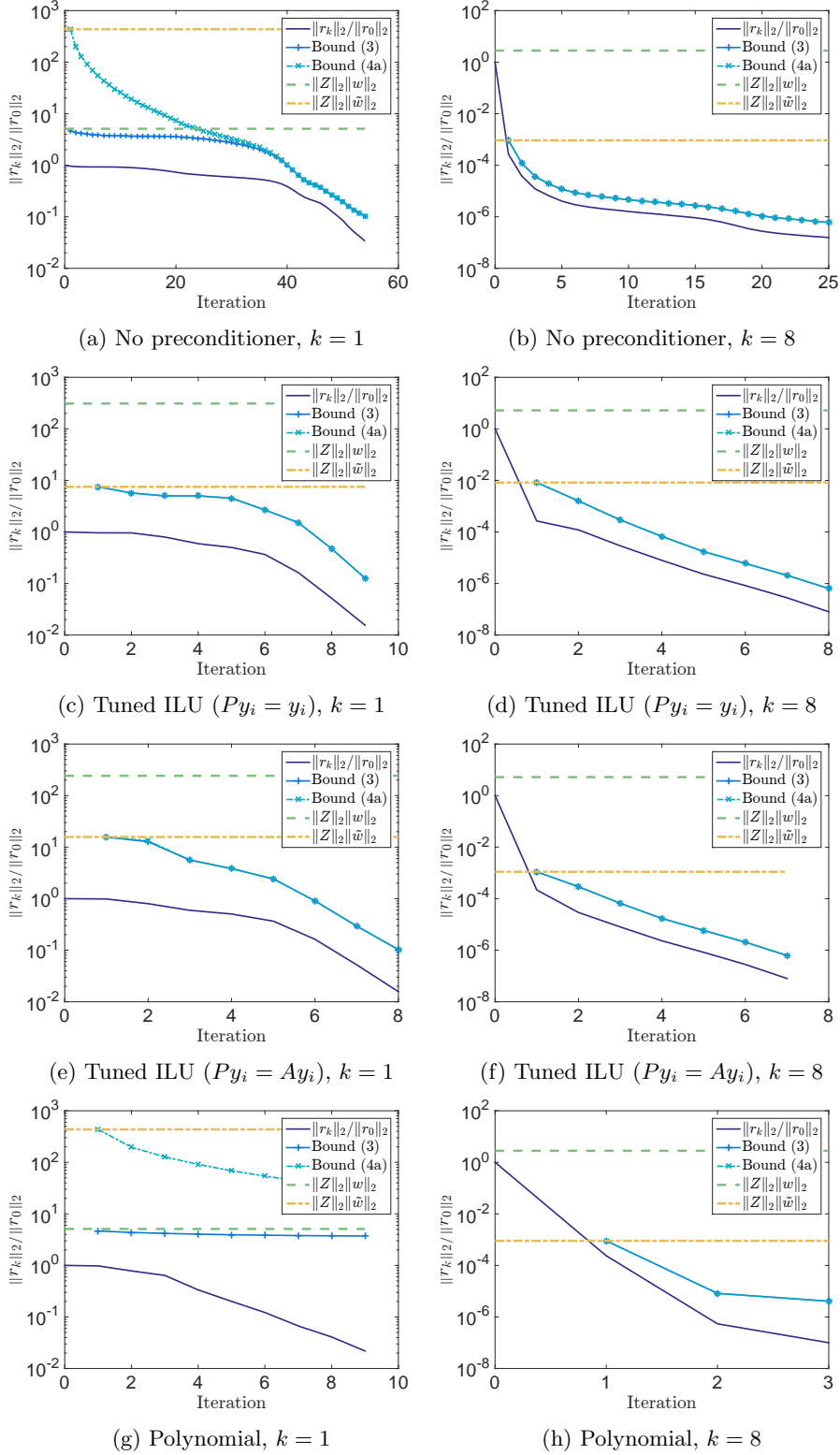
Figure 1: Relative residual norms, bounds (2) and (3a), and the constants $\|Z\|_2\|w\|_2$ and $\|Z\|_2\|\tilde{w}\|_2$ for `cdde1` for different preconditioners. The ILU drop tolerance is 10^{-2} and $\deg(p) = 10$.

Table 3: `cdde1` with tuned preconditioning ($Py_i = Iy_i$ and $Py_i = Ay_i$): inverse iteration residual $\|\rho_{i-1}\|_2$, eigenvector matrix norm $\|Z\|_2$ and condition number $\kappa_2(Z)$, sizes of components of w, constants in (2) and (3) and GMRES iteration numbers at each outer iteration i. The ILU drop tolerance is 10^{-2}.

	i	$\lambda^{(i)}$	$\|\rho_{i-1}\|_2$	$\|Z\|_2$	$\kappa_2(Z)$	$\|w_1\|$	$\|w^{(2)}\|_2$	$\|Z\|_2\|w\|_2$	$\|Z\|_2\|\tilde{w}\|_2$	$\|Z\|_2\|\tilde{w}\|_2/\tau$	it
	1	0.099735	0.0033	4.9	1.9e+04	17	60	3.1e+02	–	–	9
	2	-0.00285	0.00095	5.2	2.3e+04	0.98	0.73	6.3	5.4e+02	1.7e+07	11
	3	-0.005968	0.00027	5.2	2.5e+04	1	0.012	5.2	7.4	7.8e+05	9
I	4	-0.0049348	7.9e-05	5.2	2.4e+04	1	0.0028	5.2	1.4	5.2e+05	8
	5	-0.0052389	2.2e-05	5.2	2.4e+04	1	0.00076	5.2	0.4	5e+05	8
	6	-0.0051541	6e-06	5.2	2.4e+04	1	0.00021	5.2	0.11	5e+05	8
	7	-0.0051775	1.6e-06	5.2	2.4e+04	1	5.8e-05	5.2	0.03	5e+05	8
	8	-0.0051711	4.5e-07	5.2	2.4e+04	1	1.6e-05	5.2	0.0083	5e+05	8
	1	0.099735	0.0033	5	2.5e+04	15	46	2.4e+02	–	–	8
	2	-0.0028498	0.00095	5.2	2e+04	2.5	5.9	33	5.6	1.7e+05	10
	3	-0.0059681	0.00027	5.2	2.5e+04	1	0.18	5.3	0.7	7.4e+04	8
A	4	-0.0049348	7.9e-05	5.2	2.4e+04	1	0.048	5.2	0.2	7.5e+04	7
	5	-0.0052389	2.2e-05	5.2	2.4e+04	1	0.013	5.2	0.054	6.8e+04	7
	6	-0.0051541	6e-06	5.2	2.4e+04	1	0.0035	5.2	0.015	6.8e+04	7
	7	-0.0051775	1.6e-06	5.2	2.4e+04	1	0.00095	5.2	0.004	6.7e+04	7
	8	-0.0051711	4.5e-07	5.2	2.4e+04	1	0.00026	5.2	0.0011	6.7e+04	7

case (Table 2). This behavior is somehow destroyed by the standard preconditioner (Table 2) but mimicked by the two tuned preconditioners (Table 3) and exactly matched for the polynomial one (Table 2). Using either tuned or polynomial preconditioners also has a positive effect on the constants in (2) and (3), which are larger for the standard preconditioner. Most importantly, the number of GMRES iteration steps is notably reduced and remains at an approximately constant level for tuned and polynomial preconditioners. The increasing trend for the standard preconditioner is also evident.

A visual illustration of these observations is given in Figures 1 and 2. Figure 1 shows the behavior of the GMRES residuals for no preconditioner, standard, tuned, and polynomial preconditioners. As discussed in Section 2.2 we observe the initial decrease of both the GMRES residual and the bound in (3a) (when no preconditioner or the tuned preconditioner is used), which is particularly prominent in the later stages of the iteration. Figure 2 shows the history of components of w as the outer iteration proceeds in the two top plots. The bottom plots show the required number of GMRES steps against the outer iteration (left plot) as well as the eigenvalue residual norm against the cumulative sum of inner GMRES steps (right plot). The significant reduction of inner iterations by tuned and polynomial preconditioners is apparent.

Figures 3 and 4 show similar plots for the matrices `cd_fd` and `olm2000` using the same preconditioning settings. For `cd_fd`, using a standard preconditioner results in a drastic increase in the magnitude of the components of the weight vector w (top plots of Figure 3). The beneficial effects of tuned and polynomial preconditioners are similar to the previous examples. For the matrix `olm2000`, the tuned preconditioners lead to a slight increase of the outer iteration steps (14 compared to 11 for the other choices). However, the amount of work in terms of the number of required inner iteration steps is still smaller than for the other variants (bottom plots of Figure 4). The polynomial preconditioner seems to be of lesser quality compared to other preconditioners for the matrix `olm2000`, leading to more inner iterations. Increasing the polynomial degree did not lead to improvements. It seems that for this example, the basic strategies mentioned in Section 3.4 to select the coefficients of the preconditioning polynomial are not sufficient. For these cases, this highlights an advantage of tuned preconditioners over polynomial preconditioning, especially
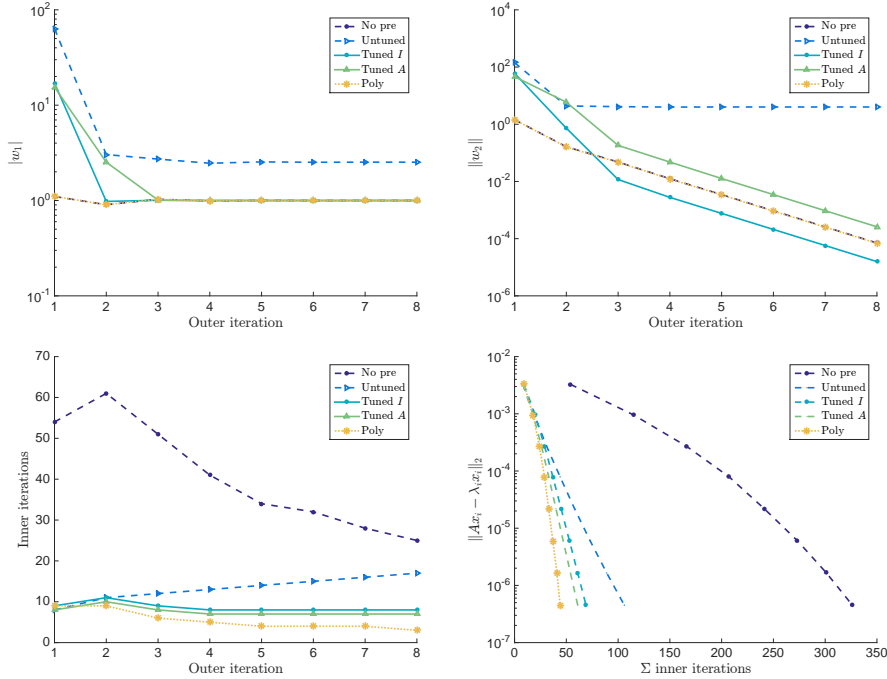
Figure 2: `cdde1`: $|w_1|$, $\|w_2\|$, number of GMRES iterations and outer residual. The ILU drop tolerance is 10^{-2} and $\deg(p) = 10$.

for tuned preconditioners that are built from standard preconditioning approaches, e.g. incomplete factorizations, which can be constructed in a much more automatic and straightforward manner.

For all examples and all used preconditioners, Table 4 gives a condensed overview of the obtained results, listing the final magnitudes of the weight components in w, and the total numbers of outer and inner (GMRES) iterations. While the majority of the results are in line with the previous observations, there are some outliers. For the matrices `cd_fd` and `cdde1`, the tuned preconditioners built from the ILU with drop tolerance $\theta = 10^{-1}$ lead to much smaller magnitudes of the weights compared to standard preconditioner, but the components of w do not have the property $|w_1| > \|w^{(2)}\|_2$. Decreasing the drop tolerance to $\theta = 10^{-2}$, 10^{-3} leads to the desired situation $w \to e_1$. Moreover, increasing the degree of the polynomial preconditioner is not always beneficial as we see in all three examples for $d = 15$. With this setting the polynomial preconditioner leads to a worse performance compared to lower polynomial degrees. As we mentioned above, the origin of this behavior can is the basic coefficient generation approaches, which appear to return more and more very tiny coefficients if the degree is increased. Given the difficulties with the polynomial preconditioner in these cases, we recommend to either look for better coefficient selection strategies or use a tuned preconditioner.

We now transform `cd_fd` into a generalized problem by adding an artificial, tridiagonal matrix M having the values 2.5, 5, 2.5 on the lower, main, and upper diagonal, respectively. The shift is $\sigma = 30$ but the remaining settings to execute this test are unchanged. The result are illustrated in Figure 5. As predicted, the two upper plots show that, in contrast to the situation $M = I$, the weight vector in the unpreconditioned and polynomial preconditioned case does not converge to e_1. This weight behavior is only achieved by employing tuned preconditioners. The bottom left plot also shows a increasing number of inner iterations when no or a
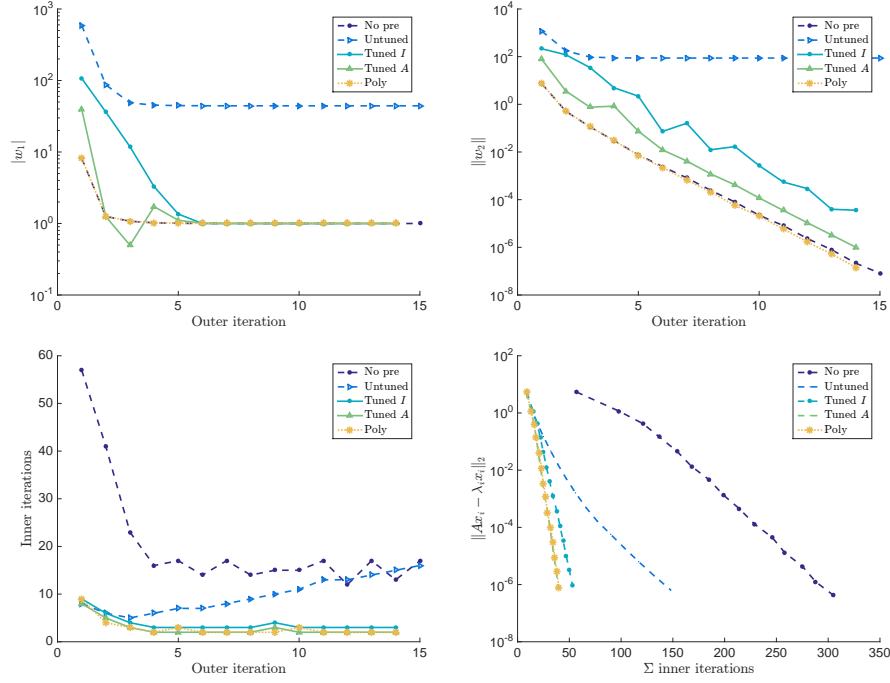
Figure 3: `cd_fd`: $|w_1|$, $\|w_2\|$, number of GMRES iterations and eigenvalue residual. The ILU drop tolerance is 10^{-2} and $\deg(p) = 10$.
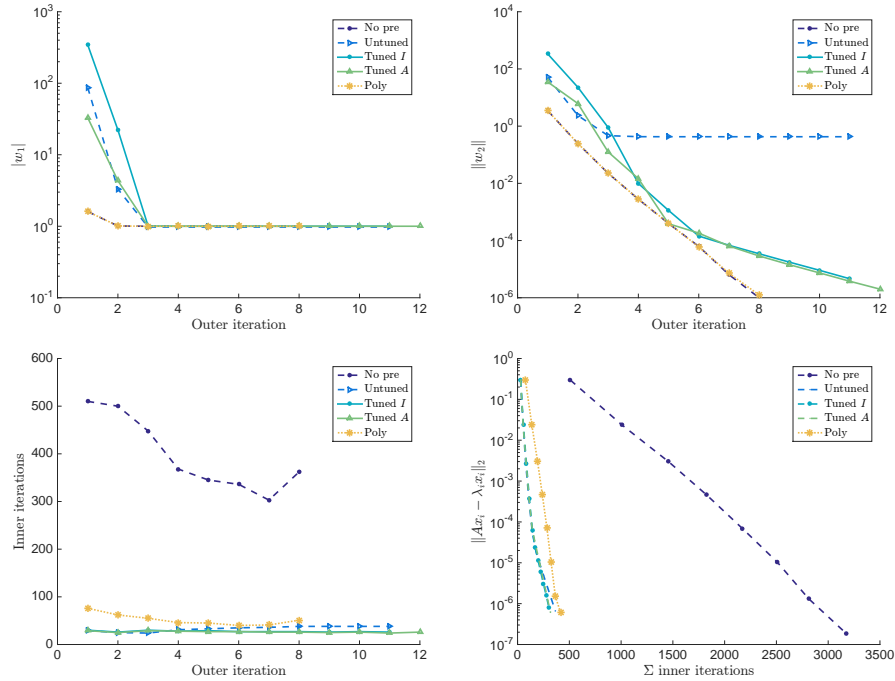

Figure 4: `olm2000`: $|w_1|$, $\|w_2\|$, number of GMRES iterations and eigenvalue residual. The ILU drop tolerance is 10^{-2} and $\deg(p) = 10$.

Table 4: Results for all examples: final weight components $|w_1|$, $\|w^{(2)}\|_2$, and total numbers of outer and inner (GMRES) iterations.

ex.	prec.	θ, d	final $|w_1|$	final $\|w^{(2)}\|_2$	outer	$\sum$ inner
	no prec.		1	8.2e-08	15	305
	std. P	1e-1	1.6735e+08	1.8e+08	13	225
	tuned I	1e-1	58.199	1.1e+02	15	100
	tuned A	1e-1	1.7041	2.3	15	87
	poly	$d = 5$	1	1.6e-07	14	61
cd_fd	std. P	1e-2	44.229	87	15	148
	tuned I	1e-2	1	3.6e-05	14	53
	tuned A	1e-2	1	1e-06	14	39
	poly	$d = 10$	1	1.4e-07	14	40
	std. P	1e-3	1.2081	4.4	15	89
	tuned I	1e-3	1	2.1e-06	14	45
	tuned A	1e-3	1	2.4e-07	14	31
	poly	$d = 15$	1	4.7e-08	20	79
	no prec.		0.99995	7e-05	8	326
	std. P	1e-1	6.0179e+05	1.2e+06	8	190
	tuned I	1e-1	0.99997	1.2	8	123
	tuned A	1e-1	1280.8	1.4e+03	8	115
	poly	$d = 5$	0.99995	6.9e-05	8	67
cdde1	std. P	1e-2	2.5239	4	8	106
	tuned I	1e-2	0.99999	1.6e-05	8	69
	tuned A	1e-2	1	0.00026	8	61
	poly	$d = 10$	0.99995	6.9e-05	8	44
	std. P	1e-3	3.4135	10	8	53
	tuned I	1e-3	1	3.9e-05	8	40
	tuned A	1e-3	0.99994	0.00026	8	32
	poly	$d = 15$	0.99998	3.7e-05	10	51
	no prec.		1	1e-06	8	3172
	std. P	1e-1	3.4649	2.8	14	505
	tuned I	1e-1	1	9.2e-06	14	393
	tuned A	1e-1	1	1.2e-05	14	379
	poly	$d = 5$	1	9.7e-07	8	1934
olm2000	std. P	1e-2	0.97526	0.43	11	365
	tuned I	1e-2	1	4.6e-06	11	303
	tuned A	1e-2	1	2e-06	12	318
	poly	$d = 10$	1	1.2e-06	8	416
	std. P	1e-3	0.97524	0.12	10	313
	tuned I	1e-3	1	4.3e-06	10	260
	tuned A	1e-3	1	2.9e-06	10	255
	poly	$d = 15$	1	1.7e-05	20	1660

polynomial preconditioner is used.

Even though this paper concentrates on GMRES bounds for inverse iteration, we finally show results of block-GMRES convergence for inverse subspace iteration. We consider the matrix cdde1 with the settings from above and we seek an invariant subspace of dimension $u = 6$. The drop tolerance of the incomplete LU factorization used for the preconditioner is 10^{-2} and we apply a tuned preconditioner which satisfies $P_i Y_i = A Y_i$, e.g. $P_i = P + (A - P) Y_i (Y_i^T Y_i)^{-1} Y_i^T$. For the implementation of block GMRES we used an algorithm discussed in [26, 20].

Figure 6 shows the history of the norms of W_1 and W_2, the splittings of the matrix $W \in \mathbb{C}^{n \times u}$ discussed in Section 4.2, as the outer iteration proceeds. The bottom plots show the required number of block GMRES steps against the outer iteration (left plot) and the residual norm against the cumulative sum of inner block GMRES steps (right plot). As expected the reduction of inner iteration numbers by the tuned preconditioner is apparent. The reduction of $\|W_2\|$ can be seen for the unpreconditioned as well as tuned case, the magnitude of $\|W_1\|$ is close to 2.449, but the overall behavior of the weights is similar to the single vector inverse iteration.
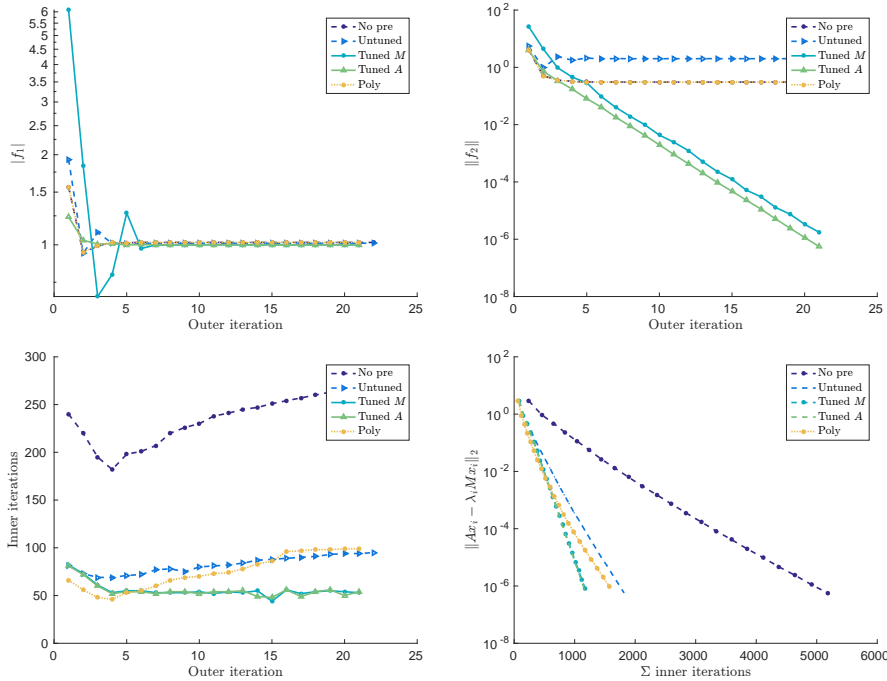
Figure 5: Results for the generalized problem $A =$`cd_fd`, $M =$ tridiag$(2.5, 5, 2.5)$: weight components $|f_1|$, $\|f_2\|$, number of GMRES iterations and eigenvalue residual. The ILU drop tolerance is 10^{-1} and $\deg(p) = 5$.

6 Conclusion

In this paper we discussed the convergence behavior of GMRES (as a prominent iterative method) for solving linear systems that arise during the solution of eigenvalue problems via inverse iteration. We gave detailed bounds on GMRES that take the special behavior of the right hand side into account and explained the initial sharp decrease of the GMRES residual. The bounds gave rise to adapted preconditioners for GMRES when applied to eigenvalue problems, e.g. tuned and polynomial preconditioners. The analysis was extended to inverse iteration for the generalized eigenvalue problem and subspace iteration, where block GMRES bounds were given. The numerical results showed that the new GMRES (block GMRES) bounds are much sharper than conventional bounds and that preconditioned subspace iteration with either a tuned or polynomial preconditioner should be used, where the tuned preconditioner is generally easier to construct. Possible future research perspectives should therefore, e.g., include generation strategies of high quality polynomial coefficients adapted to the outer eigenvalue iteration.

References

[1] M. I. Ahmad, D. B. Szyld, and M. B. van Gijzen. Preconditioned multishift BiCG for H_2-optimal model reduction. Technical Report 12-06-15, Department
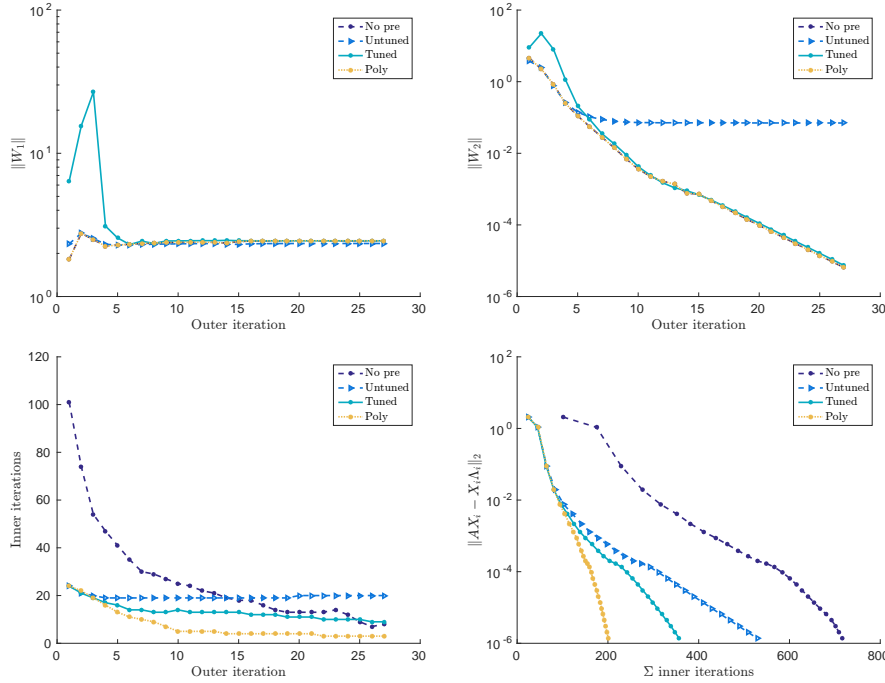
Figure 6: Inverse subspace iteration for `cdde1`, $u = 6$: $\|W_1\|$, $\|W_2\|$, number of block-GMRES iterations and eigenvalue residual. The ILU drop tolerance is 10^{-2}.

of Mathematics, Temple University, June 2012. Revised March 2013 and June 2015.

[2] M. Arioli, V. Pták, and Z. Strakoš. Krylov sequences of maximal length and convergence of GMRES. *BIT*, 38:636–643, 1998.

[3] S. F. Ashby, T. A. Manteuffel, and J. S. Otto. A Comparison of Adaptive Chebyshev and Least Squares Polynomial Preconditioning for Hermitian Positive Definite Linear Systems. *SIAM J. Sci. Statist. Comput.*, 13(1):1–29, 1992.

[4] R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. A. van der Vorst. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd Edition.* SIAM, Philadelphia, PA, 1994.

[5] M. Baumann and M. van Gijzen. Nested Krylov methods for shifted linear systems. *SIAM J. Sci. Comput.*, 37:90–112, 2015.

[6] J. Berns-Müller, I. G. Graham, and A. Spence. Inexact inverse iteration for symmetric matrices. *Linear Algebra Appl.*, 416(2):389–413, 2006.

[7] L. Du, T. Sogabe, and S.-L. Zhang. IDR(s) for solving shifted nonsymmetric linear systems. *J. Comput. Appl. Math.*, 274(0):35–43, 2015.

[8] J. Duintjer Tebbens and G. Meurant. Prescribing the behavior of early terminating GMRES and Arnoldi iterations. *Numer. Algorithms*, 65:69–90, 2014.

[9] M. A. Freitag. *Inner-outer Iterative Methods for Eigenvalue Problems – Convergence and Preconditioning.* PhD thesis, University of Bath, 2007.

[10] M. A. Freitag and A. Spence. Convergence theory for inexact inverse iteration applied to the generalised nonsymmetric eigenproblem. *Electron. Trans. Numer. Anal.*, 28:40–64, 2007.

[11] M. A. Freitag and A. Spence. A tuned preconditioner for inexact inverse iteration applied to Hermitian eigenvalue problems. *IMA J. Numer. Anal.*, 28(3):522–551, 2008.

[12] M. A. Freitag, A. Spence, and E. Vainikko. Rayleigh quotient iteration and simplified Jacobi-Davidson with preconditioned iterative solves for generalised eigenvalue problems. Techn. report, Dept. of Math. Sciences, University of Bath, 2008.

[13] R. Freund. On conjugate gradient type methods and polynomial preconditioners for a class of complex non-hermitian matrices. *Numer. Math.*, 57(1):285–312, 1990.

[14] G. H. Golub and Q. Ye. Inexact inverse iteration for generalized eigenvalue problems. *BIT*, 40(4):671–684, 2000.

[15] I. C. Ipsen. Computing an eigenvector with inverse iteration. *SIAM Rev.*, 39(2):254–291, 1997.

[16] Q. Liu, R. B. Morgan, and W. Wilcox. Polynomial Preconditioned GMRES and GMRES-DR. *SIAM J. Sci. Comput.*, 37(5):S407–S428, 2015.

[17] A. Martinez. Tuned preconditioners for the eigensolution of large SPD matrices arising in engineering problems. *Numer. Linear Algebra Appl.*, 23(3):427–443, 2016.

[18] G. Meurant and J. Duintjer Tebbens. The role eigenvalues play in forming GMRES residual norms with non-normal matrices. *Numer. Algorithms*, 68:143–165, 2015.

[19] C. C. Paige and M. A. Saunders. Solution of Sparse Indefinite Systems of Linear Equations. *SIAM J. Numer. Anal.*, 12(4):617–629, 1975.

[20] M. L. Parks, K. M. Soodhalter, and D. B. Szyld. A block Recycled GMRES method with investigations into aspects of solver performance. *ArXiv e-prints (and submitted for publication)*, 2016.

[21] G. Peters and J. H. Wilkinson. Inverse iteration, ill-conditioned equations and Newton's method. *SIAM Rev.*, 21(3):339–360, 1979.

[22] M. Robbé, M. Sadkane, and A. Spence. Inexact inverse subspace iteration with preconditioning applied to non-hermitian eigenvalue problems. *SIAM J. Matrix Anal. Appl.*, 31(1):92–113, Feb. 2009.

[23] Y. Saad. A Flexible Inner-Outer Preconditioned GMRES Algorithm. *SIAM J. Sci. Comput.*, 14(2):461–469, 1993.

[24] Y. Saad and M. Schultz. GMRES a generalised minimum residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Statist. Comput.*, 7:856–869, 1986.

[25] V. Simoncini and L. Eldén. Inexact Rayleigh quotient-type methods for eigenvalue computations. *BIT*, 42(1):159–182, 2002.

[26] K. Soodhalter. A block MINRES algorithm based on the banded Lanczos method. *Numer. Algorithms*, 69:473–494, 2015.

[27] D. Szyld and F. Xue. Efficient Preconditioned Inner Solves For Inexact Rayleigh Quotient Iteration And Their Connections To The Single-Vector Jacobi–Davidson Method. *SIAM J. Matrix Anal. A.*, 32(3):993–1018, 2011.

[28] D. Titley-Peloquin, J. Pestana, and A. J. Wathen. GMRES convergence bounds that depend on the right-hand-side vector. *IMA J. Numer. Anal.*, 34:462–479, 2014.

[29] M. B. van Gijzen. A polynomial preconditioner for the GMRES algorithm. *J. Comput. Appl. Math.*, 59(1):91–107, 1995.

[30] M. B. van Gijzen, G. L. G. Sleijpen, and J.-P. M. Zemke. Flexible and multi-shift induced dimension reduction algorithms for solving large sparse linear systems. *Numer. Lin. Alg. Appl.*, 22(1):1–25, 2015.

[31] F. Xue and H. C. Elman. Fast inexact subspace iteration for generalized eigenvalue problems with spectral transformation. *Linear Algebra Appl.*, 435(3):601–622, 2011.

[32] F. Xue and H. C. Elman. Fast inexact subspace iteration for generalized eigenvalue problems with spectral transformation. *Linear Algebra and its Applications*, 435(3):601–622, 2011.