A Tuned Preconditioner for Inexact Inverse Iteration Applied to Hermitian Eigenvalue Problems

Melina A Freitag* Alastair Spence†

November 4, 2005

Abstract

In this paper we consider the computation of an eigenvalue and corresponding eigenvector of a large sparse Hermitian positive definite matrix using either inexact inverse iteration with a fixed shift or inexact Rayleigh quotient iteration. For such problems the large sparse linear systems arising at each iteration are often solved approximately by means of symmetrically preconditioned MINRES. We consider preconditioners based on the incomplete Cholesky factorisation and derive a new tuned Cholesky preconditioner which shows considerable improvement over the standard preconditioner. This improvement is analysed using the convergence theory for MINRES. We also compare the spectral properties of the tuned preconditioned matrix with those of the standard preconditioned matrix. In particular, we provide both a perturbation result and an interlacing result, and these results show that the spectral properties of the tuned preconditioner are similar to those of the standard preconditioner. For Rayleigh quotient shifts, comparison is also made with a technique introduced by Simoncini and Eldén [*Inexact Rayleigh quotient-type methods for eigenvalue computations*, BIT, 42 (2002), pp. 159-182] which involves changing the right hand side of the inverse iteration step. Several numerical examples are given to illustrate the theory described in the paper.

Key words. Inexact inverse iteration, preconditioned iterative methods, Hermitian matrices.

AMS subject classifications. Primary 65F15, Secondary 65F10.

*Department of Mathematical Sciences, University of Bath, Claverton Down, BA2 7AY, United Kingdom. email: {m.freitag, a.spence}@maths.bath.ac.uk.

†

1 Introduction

We consider the problem of computing an eigenvalue and the corresponding eigenvector of a Hermitian positive definite matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$, that is

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}, \quad \lambda \in \mathbb{R}, \mathbf{x} \in \mathbb{C}^n \setminus \{\mathbf{0}\}, \tag{1}$$

using inexact inverse iteration with either a fixed shift or Rayleigh quotient shifts. We assume that the matrix $\mathbf{A}$ is very large and sparse and so to exploit the structure iterative techniques, in particular, preconditioned MINRES, may be used to solve the linear shifted systems

$$(\mathbf{A} - \sigma\mathbf{I})\mathbf{y} = \mathbf{x} \tag{2}$$

arising in inverse iteration, where the shift σ is chosen to be close to any eigenvalue. Hence, the linear system is solved to a prescribed tolerance only. It is well known that using exact solves inverse iteration with a fixed shift achieves linear convergence (see for example [2] or [13]). Also, a famous result due to Ostrowski [12] (see also Parlett [13]) shows that Rayleigh quotient iteration yields cubic convergence for a close enough starting guess. Theory for inexact inverse iteration with a fixed shift can be found in [6], [9] and [16] where it was proved that for an appropriately chosen decreasing tolerance inexact inverse iteration with a fixed shift achieves linear convergence. The convergence theory for inexact Rayleigh quotient iteration is given in [16] and [1]. In these papers it was proved that inexact Rayleigh quotient iteration with fixed tolerance achieves quadratic convergence whereas with an appropriately chosen decreasing tolerance it recovers the cubic convergence attained when exact solves are used.

Inexact inverse iteration is an example of an "inner-outer" iterative method. The outer iteration is the basic iterative inverse iteration algorithm, requiring the solve of (2) with a fixed or variable shift σ at each step, with the inner iteration being the inexact iterative solution of (2). Inner-outer methods have been used in other settings, see, for example, Golub and Ye [5].

In order to reduce the number of inner iterations needed to solve (2), preconditioning becomes necessary. Since $\mathbf{A}$ is Hermitian positive definite, it is natural to use an incomplete Cholesky factorisation of $\mathbf{A}$ to construct a symmetrically preconditioned form of (2). Specifically, if $\mathbf{L}\mathbf{L}^*$ is an incomplete Cholesky factorisation of $\mathbf{A}$ then one applies the iterative solver (for example, MINRES) to

$$\mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*}\tilde{\mathbf{y}} = \mathbf{L}^{-1}\mathbf{x}, \quad \mathbf{y} = \mathbf{L}^{-*}\tilde{\mathbf{y}}, \tag{3}$$

rather than to (2). For a fixed shift, it is known that (see for example [1],[9]) the number of inner iterations used by a Krylov solver applied to (3) increases steadily as the outer iteration proceeds. We shall show that with a simple rank-one change to the preconditioner, which we call "tuning" the preconditioner, this steady increase in the number of inner iterations can be stopped, and indeed considerable improvements in the total inner iteration count can be achieved.

In [3] it was suggested to tune the preconditioner in order to improve the performance of the inner solves in inexact inverse iteration for a generalised nonsymmetric eigenproblem, but [3] was concerned with the outer convergence rate and no analysis of the tuned preconditioner was given. In this paper we first derive the tuning for the Hermitian problem and motivate its use by showing that the right hand side of the tuned preconditioned system has the desirable property of being close to the direction of the sought eigenvector. Then we provide careful spectral analysis that explains the differences between the iteration matrices for the tuned and untuned cases. This involves the formulation of a nonstandard eigenvalue perturbation problem, which is analysed by a modification of the Bauer-Fike theorem (see [4]) and a novel interlacing property (in the spirit of [17, p. 94 ff] and [4]). These results show that the spectral properties of the tuned preconditioner are similar to those of the standard preconditioner.

For the case of Rayleigh quotient shifts, the idea from Simoncini & Eldén [15] is to modify the right hand side of the preconditioned system (3) so that the right hand side is close to an approximate null-vector of the iteration matrix (see Section 5.1). This new strategy reduces the number of inner iterations for each solve of (2), but destroys the cubic outer convergence for Rayleigh quotient iteration, achieving only quadratic outer convergence. (Note that this strategy requires that the shifts tend to the desired eigenvalue and so is not an option when the shift is fixed.) We shall see that a method based on the tuned preconditioner maintains the cubic convergence of the original Rayleigh quotient iteration but also achieves cheap inner solves as in the modified right hand side approach of Simoncini & Eldén [15].

In Section 2 of this paper we discuss the theory of inexact inverse iteration with a fixed shift, the convergence theory for MINRES, and then go over the use of the standard incomplete Cholesky preconditioner. We also discuss the application of these results to the solution of the shifted systems in inexact inverse iteration. Then in Section 3 we introduce the idea of a "tuned" incomplete Cholesky preconditioner and prove an important Lemma about the right-hand side of the tuned preconditioned system. We describe how to implement the tuned preconditioning and compare the numerical performance of standard and tuned preconditioners and find that the tuned preconditioner requires fewer inner iterations than the standard preconditioner. In Section 4 we provide a detailed analysis of the spectra of both the tuned and untuned iteration matrices and discuss the consequences for MINRES as iterative solver for the inner iterations. In Section 5 inexact Rayleigh quotient iteration with its properties is revised and the tuned preconditioner is applied to this variable shift method. We compare the numerical performance of the standard and tuned preconditioners. Again, the tuned preconditioner is superior to the the standard preconditioner. Numerical results are also presented comparing the performance of the tuned preconditioner with the approach of Simoncini & Eldén [15]. Section 6 is a short section summarising the paper.

We denote the eigenpairs of $\mathbf{A}$ by $(\lambda_j, \mathbf{x}_j)$, $j = 1, \ldots, n$, and use $\| \cdot \| = \| \cdot \|_2$.

2 Inexact inverse iteration with fixed shifts

In this section we revise the theory for inexact inverse iteration with a fixed shift for the calculation of a simple eigenvalue of the standard Hermitian eigenvalue problem (1), and then go on to discuss the use of MINRES as the iterative solver and preconditioning. A fixed shift method is unlikely to be of interest on its own, but it might well be used to provide a good starting guess for the eigenvector to feed into the Rayleigh quotient iteration. Also, results for fixed shifts are of interest when using subspace based methods, like the Lanczos method.

The following algorithm is a version of inexact inverse iteration with a fixed shift to find any well-separated simple eigenvalue of a Hermitian matrix.

Algorithm 2.1 (Inexact inverse iteration with a fixed shift)**.** *Given* σ *and* $\mathbf{x}^{(0)}$ *with* $\|\mathbf{x}^{(0)}\|= 1$. *For* $i = 0, 1, 2, \ldots$

(1) Choose $\tau^{(i)}$,

(2) Solve $(\mathbf{A} - \sigma\mathbf{I})\mathbf{y}^{(i)} = \mathbf{x}^{(i)}$ *inexactly, that is,*

$$\| (\mathbf{A} - \sigma\mathbf{I})\mathbf{y}^{(i)} - \mathbf{x}^{(i)} \| \leq \tau^{(i)},$$

(3) Compute approximate eigenvector $\mathbf{x}^{(i+1)} = \dfrac{\mathbf{y}^{(i)}}{\|\mathbf{y}^{(i)}\|}$,

(4) Compute approximate eigenvalue $\lambda^{(i+1)} = \mathbf{x}^{(i+1)^*}\mathbf{A}\mathbf{x}^{(i+1)}$,

(5) Evaluate eigenvalue residual $\mathbf{r}^{(i+1)} = (\mathbf{A} - \lambda^{(i+1)}\mathbf{I})\mathbf{x}^{(i+1)}$,

(6) Test for convergence.

The following theorem states the convergence theory for inexact inverse iteration with a fixed shift. It follows directly from Theorem 2.2 in [1], where a detailed proof is given (see also Lemma 2.2 in [6]).

Theorem 2.1 (Convergence of inexact inverse iteration with a fixed shift)**.** *Let (1) be the standard eigenvalue problem for a Hermitian matrix* $\mathbf{A}$ *and consider the application of Algorithm 2.1 to find a simple eigenpair* $(\lambda_1, \mathbf{x}_1)$ *of* $\mathbf{A}$. *Assume* σ *is closer to* λ_1 *than to any other eigenvalue of* $\mathbf{A}$, *and that* $\mathbf{x}^{(0)}$ *is close enough to the desired* $\mathbf{x}_1$. *Then, if a decreasing tolerance is chosen for the inexact solves in the inverse iteration Algorithm 2.1, say* $\tau^{(i)} = C_1 \| \mathbf{r}^{(i)} \|$ *in step (1), then linear convergence is achieved for small enough* $\tau^{(0)}$ *and* C_1.

Proof. Following Parlett [13], if we write

$$\mathbf{x}^{(i)} = \cos\theta^{(i)}\mathbf{x}_1 + \sin\theta^{(i)}\mathbf{x}_\perp^{(i)}, \quad \mathbf{x}_\perp^{(i)} \perp \mathbf{x}_1, \tag{4}$$

with $\|\mathbf{x}_1\| = \|\mathbf{x}_\perp^{(i)}\| = 1$ and $\theta^{(i)} = \angle(\mathbf{x}_\perp^{(i)}, \mathbf{x}_1)$, then the eigenvalue residual defined by

$$\mathbf{r}^{(i)} = (\mathbf{A} - \lambda^{(i)}\mathbf{I})\mathbf{x}^{(i)}, \tag{5}$$

with $\lambda^{(i)} = \mathbf{x}^{(i)^*}\mathbf{A}\mathbf{x}^{(i)}$ satisfies (see [13, Theorem 11.7.1])

$$|\sin\theta^{(i)}||\lambda_2 - \lambda^{(i)}| \leq \|\mathbf{r}^{(i)}\| \leq |\sin\theta^{(i)}||\lambda_n - \lambda_1|. \tag{6}$$

Thus the choice of $\tau^{(i)} = C_1 \ \| \mathbf{r}^{(i)} \|$ asks that the solve tolerance in step (2) of Algorithm 2.1 decreases with the error angle $\theta^{(i)}$. From [1, Lemma 2.1] we have

$$\tan\theta^{(i)} \leq \frac{|\lambda_1 - \sigma|}{|\lambda_2 - \sigma|}\frac{|\sin\theta^{(i)}| + \tau^{(i)}}{|\cos\theta^{(i)}| - \tau^{(i)}},$$

which, with the choice of $\tau^{(i)}$ yields linear convergence for small enough C_1. $\square$

Note that for the special case of a tolerance $\tau^{(i)} = 0$ we obtain the well known linear convergence achieved by exact inverse iteration.

2.1 Convergence theory of MINRES

In order to understand the performance of the inner iteration part of the inexact inverse iteration algorithm we revise some convergence theory of MINRES. For simplicity we shall consider only the case of a simple extreme eigenvalue, since the convergence theory for MINRES is easiest. However the ideas in this paper are applicable for any simple interior eigenvalue, and we show numerical results for an interior eigenvalue in Example 3.1.

First, we quote a theorem about the convergence of MINRES when applied to

$$\mathbf{B}\mathbf{z} = \mathbf{b} \tag{7}$$

for the case of interest here. This is a special case of Theorem 3.1 of [1], but similar results are well known in the literature (see, for example [7] and [8]).

Theorem 2.2. *Suppose that the symmetric matrix* $\mathbf{B}$ *has eigenvalues* $\mu_1, \ldots, \mu_n$ *with corresponding eigenvectors* $\mathbf{w}_1, \ldots, \mathbf{w}_n$. *Let* μ_1 *be well separated from* $\{\mu_j\}_{j=2}^n$ *and assume all the elements of* $\{\mu_j\}_{j=2}^n$ *have the same sign. Furthermore, let* $\kappa^1 = \dfrac{\max_{j=2,\ldots,n}|\mu_j|}{\min_{j=2,\ldots,n}|\mu_j|}$ *be the reduced condition number of* $\mathbf{B}$, *assume* $\max_{j=1,\ldots,n}|\mu_1 - \mu_j| = |\mu_1 - \mu_n|$ *and define* $\mathcal{P}^\perp$ *to be the orthogonal projection along* $\mathbf{w}_1$ *onto span*$\{\mathbf{w}_2, \ldots, \mathbf{w}_n\}$. *If* $\mathbf{z}_k$ *is the result of applying MINRES to (7) with starting value* $\mathbf{z}_0 = \mathbf{0}$ *then*

$$\|\mathbf{b} - \mathbf{B}\mathbf{z}_k\| \leq 2\max_{j=2,\ldots,n}\frac{|\mu_1 - \mu_j|}{|\mu_1|}\left(\frac{\sqrt{\kappa^1}-1}{\sqrt{\kappa^1}+1}\right)^{k-1}\|\mathcal{P}^\perp\mathbf{b}\|. \tag{8}$$

If, in addition the number of inner iterations satisfies

$$k \geq 1 + \frac{\sqrt{\kappa^1}}{2}\left(\log 2\frac{|\mu_1 - \mu_n|}{|\mu_1|} + \log\frac{\|\mathcal{P}^\perp\mathbf{b}\|}{\tau}\right) \tag{9}$$

then $\|\mathbf{b} - \mathbf{B}\mathbf{z}_k\| \leq \tau$.

We now apply this theorem to the solution of $(\mathbf{A}-\sigma\mathbf{I})\mathbf{y}^{(i)} = \mathbf{x}^{(i)}$, with $\mathbf{B} = \mathbf{A} - \sigma\mathbf{I}$, $\mathbf{b} = \mathbf{x}^{(i)}$, $\mu_j = \lambda_j - \sigma$, $\mathbf{w}_j = \mathbf{x}_j$ and $\mathcal{P}^{\perp}\mathbf{x}^{(i)} = \sin\theta^{(i)}\mathbf{x}_{\perp}^{(i)}$ (using (4)). Thus if $k^{(i)}$ denotes the number of inner iterations used by MINRES to solve $(\mathbf{A}-\sigma\mathbf{I})\mathbf{y}^{(i)} = \mathbf{x}^{(i)}$ inexactly as in step (2) of Algorithm 2.1, then

$$k^{(i)} \geq 1 + \frac{\sqrt{\kappa^1}}{2}\left(\log 2\frac{|\lambda_1-\lambda_n|}{|\lambda_1-\sigma|} + \log\frac{|\sin\theta^{(i)}|}{\tau^{(i)}}\right). \tag{10}$$

With $|\lambda_1 - \sigma|$ fixed and $\tau^{(i)} = C_1\|\mathbf{r}^{(i)}\|$, we see, using (6), that the right hand side of (10) is bounded independent of i. Hence we infer that the number of inner iterations used by MINRES will not increase as the outer iteration proceeds. This nice property is not maintained when preconditioning is applied as we discuss next.

2.2 Preconditioned inexact inverse iteration with a fixed shift

In this subsection we consider the application of a preconditioner in the solution of the linear system in step (2) of Algorithm 2.1.

Let $\mathbf{A}$ in the standard eigenvalue problem (1) be Hermitian positive definite and consider the incomplete Cholesky factorisation $\mathbf{LL}^*$, that is,

$$\mathbf{A} = \mathbf{LL}^* + \mathbf{E}, \tag{11}$$

where $\mathbf{E}$ is the symmetric error matrix associated with the incomplete decomposition of $\mathbf{A}$. Then, instead of solving $(\mathbf{A}-\sigma\mathbf{I})\mathbf{y}^{(i)} = \mathbf{x}^{(i)}$ in step (2) of Algorithm 2.1 inexactly, we solve the Hermitian system

$$\mathbf{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbf{L}^{-*}\tilde{\mathbf{y}}^{(i)} = \mathbf{L}^{-1}\mathbf{x}^{(i)}, \quad \mathbf{y}^{(i)} = \mathbf{L}^{-*}\tilde{\mathbf{y}}^{(i)}. \tag{12}$$

to a tolerance $\tau^{(i)}\|\mathbf{L}\|^{-1}$ so that $\|\mathbf{x}^{(i)} - (\mathbf{A}-\sigma\mathbf{I})\mathbf{y}^{(i)}\| \leq \tau^{(i)}$. This does not change the linear outer rate of convergence of the inexact inverse iteration algorithm. However, the right hand side $\mathbf{L}^{-1}\mathbf{x}^{(i)}$ is no longer close to the eigenvector corresponding to the eigenvalue of $\mathbf{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbf{L}^{-*}$ closest to zero and this changes the inner iteration behaviour as the outer iteration proceeds. Now, apply Theorem 2.2 with $\mathbf{B} = \mathbf{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbf{L}^{-*}$, $\mathbf{b} = \mathbf{L}^{-1}\mathbf{x}^{(i)}$, $\tau = \dfrac{\tau^{(i)}}{\|\mathbf{L}\|}$ and with $\kappa_{\mathbf{L}}^1$ denoting the corresponding reduced condition number of $\mathbf{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbf{L}^{-*}$, to obtain the following bound on $k^{(i)}$:

$$k_{\mathbf{L}}^{(i)} \geq 1 + \frac{\sqrt{\kappa_{\mathbf{L}}^1}}{2}\left(\log 2\frac{|\mu_1-\mu_n|}{|\mu_1|} + \log\frac{\|\mathcal{P}^{\perp}\mathbf{L}^{-1}\mathbf{x}^{(i)}\|\|\mathbf{L}\|}{\tau^{(i)}}\right) \tag{13}$$

The key point to note is that there is no reason for $\|\mathcal{P}^{\perp}\mathbf{L}^{-1}\mathbf{x}^{(i)}\|$ to behave like $\sin\theta^{(i)}$ as is the case when there is no preconditioning. So, using $\|\mathcal{P}^{\perp}\mathbf{L}^{-1}\mathbf{x}^{(i)}\| \leq \|\mathbf{L}^{-1}\|$, (13) provides

$$k_{\mathbf{L}}^{(i)} \geq 1 + \frac{\sqrt{\kappa_{\mathbf{L}}^1}}{2}\left(\log 2\frac{|\mu_1-\mu_n|\|\mathbf{L}\|\|\mathbf{L}^{-1}\|}{|\mu_1|} + \log\frac{1}{\tau^{(i)}}\right), \tag{14}$$

and the right hand side increases with i for a decreasing $\tau^{(i)}$. This indicates that there will be growth in the number of inner iterations used by MINRES to solve (12). This is indeed observed in practice as is seen in Table 1(a)

In order to recover the reassuring property of a constant number of inner iterations for preconditioned MINRES, a different approach has to be chosen. Simoncini & Eldén [15] alter the right hand side in (12), but for outer convergence this strategy requires that the shift tends to the desired eigenvalue as is the case for Rayleigh quotient iteration.

In this paper we try the alternative approach of changing the preconditioner to recover the nice property of a constant number of inner iterations at each outer step. This idea is explained in the next section.

Remark 2.1. *In this paper we shall assume that a good preconditioner for* $\mathbf{A}$ *is also a good preconditioner for* $\mathbf{A} - \sigma\mathbf{I}$*. This is the approach taken in [15] and it is likely to be the case if* $\mathbf{A}$ *arises from a discretised partial differential equation where also a tailor-made preconditioner for* $\mathbf{A}$ *may be available.*

3 The tuned preconditioner

In this section we introduce an idea for a new preconditioner to be applied to $(\mathbf{A} - \sigma\mathbf{I})\mathbf{y}^{(i)} = \mathbf{x}^{(i)}$, so that the linear outer convergence is retained, but which provides the advantage of cheap inner solves. This approach is motivated by the tuned preconditioner that was introduced in [3] for the nonsymmetric generalised eigenproblem but needs a more careful treatment to retain the Hermitian structure. Additionally, in this section and in Section 4 we are able to provide theoretical results for the tuned preconditioner that are not available in the nonsymmetric case discussed in [3].

Instead of solving (12) in step (2) of Algorithm 2.1 we solve the preconditioned Hermitian system

$$\mathbb{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbb{L}^{-*}\tilde{\mathbf{y}}^{(i)} = \mathbb{L}^{-1}\mathbf{x}^{(i)}, \quad \mathbf{y}^{(i)} = \mathbb{L}^{-*}\tilde{\mathbf{y}}^{(i)}, \tag{15}$$

inexactly, where $\mathbb{L}$ is chosen such that the right hand side of (15) is close to the eigenvector of $\mathbb{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbb{L}^{-*}$ corresponding to the eigenvalue closest to zero. In this case we will then reproduce the inner iteration behaviour observed for unpreconditioned solves. In order to achieve this we require that the preconditioner $\mathbb{L}\mathbb{L}^*$ should satisfy

$$\mathbb{L}\mathbb{L}^*\mathbf{x}^{(i)} = \mathbf{A}\mathbf{x}^{(i)}. \tag{16}$$

Hence, in addition to $\mathbb{L}\mathbb{L}^*$ being close to $\mathbf{A}$ as is usual in preconditioning we ask that $\mathbb{L}\mathbb{L}^*$ acts exactly like $\mathbf{A}$ in the direction of $\mathbf{x}^{(i)}$. It is for this reason we call $\mathbb{L}\mathbb{L}^*$ the "tuned" preconditioner. Note that if (16) holds, then using (5)

$$\mathbb{L}\mathbb{L}^*\mathbf{x}^{(i)} = \lambda^{(i)}\mathbf{x}^{(i)} + \mathbf{r}^{(i)},$$

and rearranging we have, for $\lambda^{(i)} \neq 0$,

$$(\mathbb{L}\mathbb{L}^*)^{-1}\mathbf{x}^{(i)} = \frac{1}{\lambda^{(i)}}\mathbf{x}^{(i)} - \frac{1}{\lambda^{(i)}}(\mathbb{L}\mathbb{L}^*)^{-1}\mathbf{r}^{(i)}. \tag{17}$$

The following Lemma shows that if (16) holds then the right hand side $\mathbb{L}^{-1}\mathbf{x}^{(i)}$ in (15) will be in the approximate direction of the eigenvector of the iteration matrix $\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbb{L}^{-*}$ corresponding to the eigenvalue nearest zero.

Lemma 3.1. *Let $\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbb{L}^{-*}$ have eigenvalues ξ_j and eigenvectors $\hat{\mathbf{w}}_j$, with ξ_1 corresponding to the eigenvalue nearest zero. Let $\mathcal{P}^\perp$ denote the orthogonal projection onto span$\{\hat{\mathbf{w}}_2,\dots,\hat{\mathbf{w}}_n\}$. Assume (16) holds, let $\mathbf{r}^{(i)}$ be defined by (5) and assume $\lambda^{(i)}\neq 0$. Then, for small enough $\|\mathbf{r}^{(i)}\|$ we have*

$$\left|\xi_1-\frac{\lambda^{(i)}-\sigma}{\lambda^{(i)}}\right|\leq C_3\|\mathbf{r}^{(i)}\|, \tag{18}$$

$$\|\mathbb{L}^{-1}\mathbf{x}^{(i)}-\hat{\mathbf{w}}_1\|\leq C_4\|\mathbf{r}^{(i)}\| \tag{19}$$

and

$$\|\mathcal{P}^\perp\mathbb{L}^{-1}\mathbf{x}^{(i)}\|\leq C_4\|\mathbf{r}^{(i)}\| \tag{20}$$

for some C_3 and C_4 independent of i.

Proof. Note that using (5) and (17) we have

$$\begin{aligned}
\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbb{L}^{-*}(\mathbb{L}^{-1}\mathbf{x}^{(i)}) &= \mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})(\mathbb{L}\mathbb{L}^*)^{-1}\mathbf{x}^{(i)}\\
&= \frac{1}{\lambda^{(i)}}\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})(\mathbf{x}^{(i)}-(\mathbb{L}\mathbb{L}^*)^{-1}\mathbf{r}^{(i)})\\
&= \frac{1}{\lambda^{(i)}}\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbf{x}^{(i)}-\frac{1}{\lambda^{(i)}}\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})(\mathbb{L}\mathbb{L}^*)^{-1}\mathbf{r}^{(i)}\\
&= \frac{\lambda^{(i)}-\sigma}{\lambda^{(i)}}(\mathbb{L}^{-1}\mathbf{x}^{(i)})+\frac{1}{\lambda^{(i)}}(\mathbf{I}-\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbb{L}^{-*})\mathbb{L}^{-1}\mathbf{r}^{(i)},
\end{aligned}$$

that is $\mathbb{L}^{-1}\mathbf{x}^{(i)}$ is an approximate eigenvector of $\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbb{L}^{-*}$ with approximate eigenvalue $\dfrac{\lambda^{(i)}-\sigma}{\lambda^{(i)}}$ for small enough $\|\mathbf{r}^{(i)}\|$. Standard perturbation theory for simple eigenvalues of symmetric matrices shows that $\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbb{L}^{-*}$ has a simple eigenvalue ξ_1 near $\dfrac{\lambda^{(i)}-\sigma}{\lambda^{(i)}}$ with corresponding eigenvector $\hat{\mathbf{w}}_1$ near $\mathbb{L}^{-1}\mathbf{x}^{(i)}$ and we obtain (18) and (19). Finally we have

$$\|\mathcal{P}^\perp\mathbb{L}^{-1}\mathbf{x}^{(i)}\|=\|\mathcal{P}^\perp(\hat{\mathbf{w}}_1+(\mathbb{L}^{-1}\mathbf{x}^{(i)}-\hat{\mathbf{w}}_1))\|=\|\mathcal{P}^\perp(\mathbb{L}^{-1}\mathbf{x}^{(i)}-\hat{\mathbf{w}}_1)\|\leq C_4\|\mathbf{r}^{(i)}\|$$

since $\mathcal{P}_1^\perp\hat{\mathbf{w}}_1=0$. ◻

For our purposes, the important result in Lemma 3.1 is (20), which with $\tau^{(i)}$ chosen as in (6) implies that $\|\mathcal{P}^\perp\mathbb{L}^{-1}\mathbf{x}^{(i)}\|=\mathcal{O}(|\sin\theta^{(i)}|)$. This is similar to the corresponding result in the unpreconditioned case and is important when analysing the lower bound for the number of iterations needed by MINRES. We have the following Corollary of Theorem 2.2.

Theorem 3.1. *Assume the conditions of Lemma 3.1 and that (16) holds. Consider the application of MINRES to the inexact solution of*

$$\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbb{L}^{-*}\tilde{\mathbf{y}}^{(i)} = \mathbb{L}^{-1}\mathbf{x}^{(i)}. \tag{21}$$

Assume $\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbb{L}^{-*}$ satisfies the conditions on $\mathbf{B}$ in Theorem 2.2. Denote the reduced condition number by $\kappa_{\mathbb{L}}^1$. Then the number of inner iterations needed by MINRES to solve (21) to a tolerance $\tau^{(i)}\|\mathbb{L}\|^{-1}$, where $\tau^{(i)} = C_1\|\mathbf{r}^{(i)}\|$, satisfies

$$k^{(i)} \geq 1 + \frac{\sqrt{\kappa_{\mathbb{L}}^1}}{2}\left(\log 2\frac{|\xi_1-\xi_n|}{|\xi_1|} + \log\frac{C_4}{C_1}\right) \tag{22}$$

where C_4 is as in (20). Hence the right hand side of (22) is bounded independent of i.

Theorem 3.1 indicates that if we can find a preconditioner that satisfies (16) then we expect no growth in the inner iteration count for MINRES using the tuned preconditioner as the outer iteration proceeds. Numerical results confirming this are given in Table 1(b).

Note, we shall return in Section 4 to the assumption about the eigenvalues of $\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbb{L}^{-*}$ satisfying the conditions on $\mathbf{B}$ in Theorem 2.2. In the rest of this section we discuss the implementation of the tuned preconditioner and then illustrate its performance by a numerical example.

3.1 Implementation of the tuned preconditioner

The following lemma shows that condition (16) can be satisfied using a simple rank-one change to $\mathbf{L}$, the original computed lower triangular Cholesky factor.

Lemma 3.2 (Construction of $\mathbb{L}$)**.** *Let $\mathbf{A}$ be Hermitian positive definite and consider its incomplete Cholesky factorisation $\mathbf{LL}^*$, so that $\mathbf{A} = \mathbf{LL}^* + \mathbf{E}$ where $\mathbf{E}$ is the Hermitian error matrix. Furthermore, let $\mathbf{x}^{(i)}$ be the approximate eigenvector obtained from the ith iteration in the inexact inverse iteration Algorithm. Set $\mathbf{u}^{(i)} = \mathbf{E}\mathbf{x}^{(i)} = \mathbf{A}\mathbf{x}^{(i)} - \mathbf{LL}^*\mathbf{x}^{(i)}$ and $\mathbf{v}^{(i)} = \mathbf{L}^{-1}\mathbf{u}^{(i)}$. Assume ${\mathbf{u}^{(i)}}^*\mathbf{x}^{(i)} \neq 0$ and set*

$$\gamma^{(i)} := \frac{1}{{\mathbf{u}^{(i)}}^*\mathbf{x}^{(i)}}. \tag{23}$$

Furthermore assume

$$1 + \gamma^{(i)}{\mathbf{v}^{(i)}}^*\mathbf{v}^{(i)} \geq 0 \tag{24}$$

and set

$$\alpha^{(i)} = \frac{-1 \pm \sqrt{1+\gamma^{(i)}{\mathbf{v}^{(i)}}^*\mathbf{v}^{(i)}}}{{\mathbf{v}^{(i)}}^*\mathbf{v}^{(i)}}. \tag{25}$$

If $\mathbb{L}$ in (15) is chosen such that

$$\mathbb{L} = \mathbf{L} + \alpha^{(i)}\mathbf{u}^{(i)}{\mathbf{v}^{(i)}}^*, \tag{26}$$

then $\mathbb{L}\mathbb{L}^*\mathbf{x}^{(i)} = \mathbf{A}\mathbf{x}^{(i)}$. Furthermore

$$(\mathbf{I} + \alpha^{(i)}\mathbf{v}^{(i)}{\mathbf{v}^{(i)}}^*)^2 = \mathbf{I} + \gamma^{(i)}\mathbf{v}^{(i)}{\mathbf{v}^{(i)}}^*. \tag{27}$$

Proof. By straightforward calculation of $\mathbb{L}\mathbb{L}^*\mathbf{x}^{(i)}$ with $\mathbb{L}$ given in (26), using $\mathbf{A}\mathbf{x}^{(i)} = \mathbf{E}\mathbf{x}^{(i)} + \mathbf{L}\mathbf{L}^*\mathbf{x}^{(i)}$ and enforcing (16) gives a quadratic equation in $\alpha^{(i)}$. The roots of the quadratic are given by (25). Equality (27) is also obtained by straightforward computation using the property $\gamma^{(i)}\mathbf{v}^{(i)}\mathbf{v}^{(i)^*} = 2\alpha^{(i)}\mathbf{v}^{(i)}\mathbf{v}^{(i)^*} + \alpha^2\mathbf{v}^{(i)}\mathbf{v}^{(i)^*}\mathbf{v}^{(i)}\mathbf{v}^{(i)^*}$. □

Note that in our experiments the choice of the sign in the definition of $\alpha^{(i)}$ was unimportant.

With $\mathbb{L}$ chosen as in Lemma 3.2 equation (16) holds and so the bound in Theorem 3.1 holds, indicating no growth in the number of inner iterations. This is indeed observed in the numerical examples given in Table 1(b).

It is important to note that replacing $\mathbf{L}$ by $\mathbb{L}$ involves minimal extra computational work. Indeed for the implementation of $\mathbb{L}$ rather than $\mathbf{L}$ at each (i) (that is, at each outer iteration) only a single extra back substitution with $\mathbf{L}$ is needed for the tuned preconditioner $\mathbb{L}$. This is proved using the Sherman-Morrison formula (see [2, p. 95]) for the inverse of a matrix with a rank-one change.

Lemma 3.2 assumes $\mathbf{u}^{(i)^*}\mathbf{x}^{(i)} \neq 0$. If $\mathbf{u}^{(i)}$ is close to zero then $\mathbf{A}\mathbf{x}^{(i)}$ is close to $\mathbf{L}\mathbf{L}^*\mathbf{x}^{(i)}$ and no tuning is necessary. It is less clear how to interprete $\mathbf{u}^{(i)^*}\mathbf{x}^{(i)} = 0$ if $\mathbf{u}^{(i)}$ is not zero. If $\mathbf{u}^{(i)^*}\mathbf{x}^{(i)} > 0$ then (24) is obviously satisfied. For $\mathbf{u}^{(i)^*}\mathbf{x}^{(i)} < 0$, (24) asks that

$$\mathbf{v}^*\mathbf{v} \leq -\mathbf{u}^{(i)^*}\mathbf{x}^{(i)},$$

or, equivalently

$$\mathbf{x}^{(i)^*}\mathbf{E}^*(\mathbf{L}\mathbf{L}^*)^{-1}\mathbf{E}\mathbf{x}^{(i)} \leq -\mathbf{x}^{(i)^*}\mathbf{E}\mathbf{x}^{(i)}, \tag{28}$$

which is satisfied if $\|\mathbf{E}\|$ is small enough, that is, if the drop tolerance is small enough.

3.2 Numerical example

We now present a numerical example to illustrate the theory in this section.

Example 3.1 (Problem from the Matrix Market library [10]). *Here we consider the matrix* `nos5.mtx` *from the Matrix Market library [10]. It is a real symmetric positive definite matrix of size* 468×468 *with* 5172 *nonzero entries. Its first four eigenvalues are given by*

	1st	*2nd*	*3rd*	*4th*
eigenvalue	52.8995	67.5430	115.5912	131.5636.

We consider the fixed shift strategy and seek the eigenvalue closest to $\sigma = 100$, *which is given by* $\lambda_3 \approx 115.5912$. *We compare the costs of the following two different methods:*

(a) Standard incomplete Cholesky preconditioner: Algorithm 2.1 with step (2) implemented by solving (12), where $\mathbf{L}\mathbf{L}^*$ *is the incomplete Cholesky factorisation of* $\mathbf{A}$.

(b) Tuned incomplete Cholesky preconditioner: Algorithm 2.1 with step (2) implemented by solving (15), where $\mathbb{L}$ *is given by (26).*

For the inexact solves we use preconditioned MINRES with

$$\tau^{(i)} = \min\{\tau, 0.1\ \|\mathbf{r}^{(i)}\|\}, \quad \tau = 0.1. \tag{29}$$

and for the incomplete Cholesky decomposition we use a drop tolerance of 0.1*. Then the number of nonzero entries in* $\mathbf{L}$ *is* 1032 *and* $\|\mathbf{E}\| \approx 7.7e+04$ *(with* $\|\mathbf{A}\| \approx 5.8e+05$*), so the standard preconditioner is very poor. We use a random starting guess* $\mathbf{x}^{(0)}$ *and a fixed shift of* $\sigma = 100$*. The computations stop once the eigenvalue residual satisfies* $\|\mathbf{r}^{(i)}\| < 10^{-8}$.

Table 1: Iteration numbers for Example 3.1 using *fixed shift* $\sigma = 100$. Total number of iterations and number of inner iterations using either the standard incomplete Cholesky preconditioner (a) or the tuned incomplete Cholesky preconditioner (b). The numbers of inner iterations per outer iteration are shown columnwise.

	Standard incomplete Cholesky (a)				*Tuned incomplete Cholesky (b)*			
	51	56	64	72	56	30	13	20
	50	57	65	73	43	32	29	21
	51	58	66	74	34	29	19	23
	51	59	67	75	28	30	22	24
	49	60	68	76	17	21	21	20
	51	61	69	77	11	22	24	19
	52	62	70	77	15	22	19	21
	53	63	71	78	25	24	23	22
total	2026 inner iterations (32 outer it.)				779 inner iterations (32 outer it)			

Table 1 shows the number of inner iterations used by methods (a) and (b) and Figure 1 gives a visual account of the same data. Both methods have a relatively high number of outer iterations, since we have only linear convergence with convergence rate $\dfrac{\lambda_3 - \sigma}{\lambda_2 - \sigma} \approx 0.4804$. As we would expect there is almost no difference between methods (a) and (b) as regards the number of outer iterations and the overall outer convergence rate. These results are not presented here since we are primarily interested in the inner iterations used by preconditioned MINRES. The tuned preconditioner is clearly much better than the standard preconditioner in terms of the total number of iterations. The number of inner iterations per outer iteration grows steadily for the standard incomplete Cholesky preconditioner as expected, whilst it stays roughly constant for the tuned preconditioner (indeed there is a decrease observed). Also, the tuned preconditioner requires about half the total number of inner iterations than the standard preconditioner. Indeed this advantage is seen in other numerical experiments not reproduced here, and overall, it appears that the tuned preconditioner has a clear advantage over the standard incomplete Cholesky preconditioner.

4 Spectral analysis for the tuned preconditioner

In this section we compare the spectral properties of $\mathbf{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbf{L}^{-*}$ and $\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbb{L}^{-*}$. In particular we shall show that there is a close relationship between the respective spectra, and so if $\mathbf{L}\mathbf{L}^*$ is a good preconditioner for $\mathbf{A}$ then $\mathbb{L}\mathbb{L}^*$ will also be a good preconditioner. Specifically, we make the comparison using both a perturbation analysis and an interlacing analysis.

First recall that for $j = 1, \ldots, n$, $\mathbf{A}$ has eigenpairs $(\lambda_j, \mathbf{x}_j)$, $\mathbf{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbf{L}^{-*}$ has eigenpairs $(\mu_j, \mathbf{w}_j)$ and $\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbb{L}^{-*}$ has eigenpairs $(\xi_j, \hat{\mathbf{w}}_j)$. If we consider the problem of finding the smallest eigenvalue of $\mathbf{A}$, say λ_1, using a shift σ between λ_1 and λ_2 (the next smallest eigenvalue of $\mathbf{A}$) then $\mathbf{A}-\sigma\mathbf{I}$ has one negative eigenvalue, $\lambda_1-\sigma$, and $n-1$ positive eigenvalues, $\{\lambda_j-\sigma\}_{j=2}^n$. Sylvester's Inertia Theorem readily shows that both $\mathbf{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbf{L}^{-*}$ and $\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbb{L}^{-*}$ have one negative eigenvalue and $n-1$ positive eigenvalues. Thus, in this case, the assumption in Theorem 3.1 that $\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbb{L}^{-*}$ satisfies the conditions on $\mathbf{B}$ in Theorem 2.2 is satisfied. We emphasise that our theory is applicable to an interior eigenvalue, in which case Theorem 2.2 can be altered to apply to a $\mathbf{B}$ with a more general spectrum.

In this section it is convenient to omit all superscripts concerned with the outer iteration steps of inverse iteration.

Consider the two eigenvalue problems

$$\mathbf{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbf{L}^{-*}\mathbf{w} = \mu\mathbf{w} \tag{30}$$

and

$$\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbb{L}^{-*}\hat{\mathbf{w}} = \xi\hat{\mathbf{w}}. \tag{31}$$

Using the relation

$$\mathbb{L} = \mathbf{L} + \alpha\mathbf{u}\mathbf{v}^* = \mathbf{L}(\mathbf{I} + \alpha\mathbf{L}^{-1}\mathbf{u}\mathbf{v}^*) = \mathbf{L}(\mathbf{I} + \alpha\mathbf{v}\mathbf{v}^*) \tag{32}$$

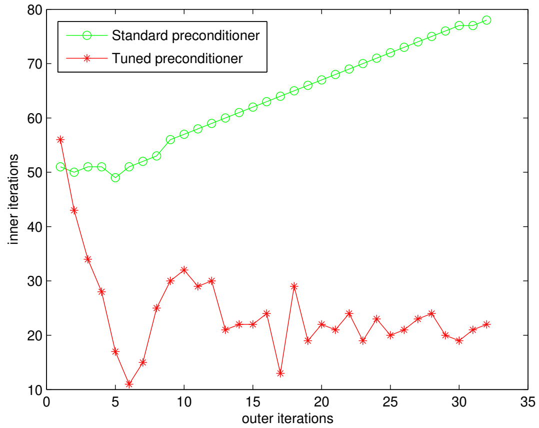

Figure 1: Inner iterations against outer iterations for methods (a) and (b)

(31) becomes

$$(\mathbf{I} + \alpha \mathbf{v}\mathbf{v}^*)^{-1}\mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*}(\mathbf{I} + \alpha \mathbf{v}\mathbf{v}^*)^{-1}\hat{\mathbf{w}} = \xi\hat{\mathbf{w}}. \tag{33}$$

Hence, setting $(\mathbf{I} + \alpha \mathbf{v}\mathbf{v}^*)^{-1}\hat{\mathbf{w}} = \mathbf{w}'$ and using (27) we find that (31) is equivalent to the generalised eigenvalue problem

$$\mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*}\mathbf{w}' = \xi(\mathbf{I} + \gamma \mathbf{v}\mathbf{v}^*)\mathbf{w}'. \tag{34}$$

In the following section 4.1 we will present a perturbation result comparing the eigenvalues ξ of (34) to the eigenvalues μ of (30), which is a modification of the theorem by Bauer and Fike (see, for example [4, Theorem 7.2.2]).

4.1 Perturbation Theory

The following theorem yields a perturbation result for the eigenvalues μ and ξ of (30) and (34).

Theorem 4.1 (Perturbation Property)**.** *Assume σ is not an eigenvalue $\mathbf{A}$. Define $\mathbf{S} = \mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*}$ and consider the two eigenvalue problems*

$$\mathbf{S}\mathbf{w} = \mu\mathbf{w} \tag{35}$$

and

$$\mathbf{S}\mathbf{w}' = \xi(\mathbf{I} + \gamma \mathbf{v}\mathbf{v}^*)\mathbf{w}'. \tag{36}$$

Then μ and ξ are nonzero. Also, let ξ be a solution of (36). Then

$$\min_{\mu\in\Lambda(\mathbf{S})} \left|\frac{\mu - \xi}{\xi}\right| \leq |\gamma \mathbf{v}^*\mathbf{v}|. \tag{37}$$

Proof. If σ is not an eigenvalue of $\mathbf{A}$ then $\mathbf{A} - \sigma\mathbf{I}$ is nonsingular, ans Sylvester's Inertia Theorem shows that μ and ξ in (35) and (36) respectively cannot be zero.

Write equation (36) as

$$(\mathbf{S} - \xi\mathbf{I})\mathbf{w}' = \xi\gamma\mathbf{v}\mathbf{v}^*\mathbf{w}'.$$

Now, let $\mu \neq \xi$ (for $\mu = \xi$ the result (37) follows immediately). Then $\mathbf{S} - \xi\mathbf{I}$ is nonsingular and

$$\mathbf{w}' = \xi(\mathbf{S} - \xi\mathbf{I})^{-1}\gamma\mathbf{v}\mathbf{v}^*\mathbf{w}'.$$

Taking norms we obtain

$$\|\mathbf{w}'\| \leq |\xi|\|(\mathbf{S} - \xi\mathbf{I})^{-1}\||\gamma|\|\mathbf{v}\mathbf{v}^*\|\|\mathbf{w}'\|$$

and hence

$$1 \leq |\xi|\frac{1}{\min_{\mu\in\Lambda(\mathbf{S})} |\mu - \xi|}|\gamma|\|\mathbf{v}^*\mathbf{v}\|,$$

yielding (37) after rearrangement. □

Corollary 4.1. *Interchanging the roles of (35) and (36) we have*

$$\min_{\xi \in \Lambda(\mathbf{S},(\mathbf{I}+\gamma \mathbf{v}\mathbf{v}^*))} \left| \frac{\xi - \mu}{\mu} \right| \leq |\gamma \mathbf{v}^* \mathbf{v}| , \tag{38}$$

where $\Lambda(\mathbf{S}, (\mathbf{I} + \gamma \mathbf{v}\mathbf{v}^*))$ *is the spectrum of the generalised eigenproblem (36).*

In Section 4.3 we use this perturbation result to estimate the change in the condition number of the system matrix of (15) compared to the condition number of the system matrix of (12), which is important for the performance of the iterative solver.

4.2 Interlacing property

The following two Lemmata lead to an interlacing result (Theorem 4.2) between the eigenvalues μ of (35) and ξ of (36), which leads to an interlacing result between the eigenvalues of the matrices in (30) and (31). Here we use ideas from Wilkinson (see [17]) and Golub and van Loan [4, Lemma 8.5.2, Theorem 8.5.3].

Lemma 4.1. *Consider the eigenvalue problems*

$$\mathbf{L}^{-1}(\mathbf{A} - \sigma \mathbf{I})\mathbf{L}^{-*}\mathbf{w} = \mu \mathbf{w} \tag{39}$$

and

$$\mathbb{L}^{-1}(\mathbf{A} - \sigma \mathbf{I})\mathbb{L}^{-*}\hat{\mathbf{w}} = \xi \hat{\mathbf{w}}, \tag{40}$$

where $\mathbb{L}$ *is given by (26). Then we can rewrite the second equation as*

$$\mathbf{D}\mathbf{t} = \xi(\mathbf{I} + \gamma \mathbf{z}\mathbf{z}^*)\mathbf{t} \tag{41}$$

or

$$(\mathbf{D} - \xi \gamma \mathbf{z}\mathbf{z}^*)\mathbf{t} = \xi \mathbf{t}, \tag{42}$$

where $(\mathbf{I} + \alpha \mathbf{v}\mathbf{v}^*)\mathbf{Q}\mathbf{t} = \hat{\mathbf{w}}$*,* $\mathbf{z} = \mathbf{Q}^*\mathbf{v}$ *with* $\mathbf{v}$ *as in Lemma 3.2 and* $\mathbf{S} = \mathbf{Q}\mathbf{D}\mathbf{Q}^*$ *is the symmetric real Schur decomposition of* $\mathbf{S} = \mathbf{L}^{-1}(\mathbf{A} - \sigma \mathbf{I})\mathbf{L}^{-*}$*, i.e.* $\mathbf{D}$ *is a diagonal matrix* $\mathbf{D} = diag(\mu_1, \ldots, \mu_n)$ *containing the eigenvalues of* $\mathbf{S}$*.*

Proof. We already know from (30) and (34) that with $\mathbf{S} = \mathbf{L}^{-1}(\mathbf{A} - \sigma \mathbf{I})\mathbf{L}^{-*}$ we can rewrite equations (39) and (40) as

$$\mathbf{S}\mathbf{w} = \mu \mathbf{w} \tag{43}$$

and

$$\mathbf{S}\mathbf{w}' = \xi(\mathbf{I} + \gamma \mathbf{v}\mathbf{v}^*)\mathbf{w}', \tag{44}$$

Then, by using the real symmetric Schur decomposition of $\mathbf{S} = \mathbf{Q}\mathbf{D}\mathbf{Q}^*$, where $\mathbf{D} = \mathrm{diag}(\mu_1, \ldots, \mu_n)$ and setting $\mathbf{Q}^*\mathbf{w}' = \mathbf{t}$ and $\mathbf{Q}^*\mathbf{v} = \mathbf{z}$ we obtain (41), that is

$$\mathbf{D}\mathbf{t} = \xi(\mathbf{I} + \gamma \mathbf{z}\mathbf{z}^*)\mathbf{t}.$$

□

Remark 4.1. *Note, that in (44) the matrix of the right hand side* $\mathbf{I}+\gamma\mathbf{v}\mathbf{v}^*$ *has all eigenvalues* 1 *except for one eigenvalue at* $1+\gamma\mathbf{v}^*\mathbf{v}$. *Thus the matrix* $(\mathbf{I}+\gamma\mathbf{v}\mathbf{v}^*)$ *is positive definite if* $1+\gamma\mathbf{v}^*\mathbf{v}>0$ *(see condition (24)).*

In [4] the efficient computation of the eigenvalues and eigenvectors of a diagonal plus rank-1 matrix was described establishing also an interlacing property between the eigenvalues of the diagonal matrix and the one which is changed by a rank-1 update (see also [17]).

Here, problem (41) is a generalised eigenvalue problem rather than a standard eigenproblem with rank-1 change but we shall prove that for this problem an interlacing property also holds.

The proofs of the following Lemma and Theorem follow the lines of the proofs of [4, Lemma 8.5.2] and [4, Theorem 8.5.3].

Lemma 4.2. *Suppose* $\mathbf{D} = diag(\mu_1,\ldots,\mu_n)\in\mathbb{R}^{n\times n}$ *has the property that* $\mu_1 < \ldots < \mu_n$. *Assume that* $\gamma\neq 0$ *and that* $\mathbf{z}$ *has no zero components. If*

$$(\mathbf{D}-\xi\gamma\mathbf{z}\mathbf{z}^*)\mathbf{t} = \xi\mathbf{t}, \quad \mathbf{t}\neq 0$$

then $\mathbf{z}^*\mathbf{t}\neq 0$ *and* $\mathbf{D}-\xi\mathbf{I}$ *is nonsingular.*

Proof. If ξ were an eigenvalue of $\mathbf{D}$ then $\xi=\mu_j$ for some j and hence with $\mathbf{e}_j$ being the jth canonical vector we have

$$0 = \mathbf{e}_j^*[(\mathbf{D}-\xi)\mathbf{t} - \xi\gamma(\mathbf{z}^*\mathbf{t})\mathbf{z}] = \xi\gamma(\mathbf{z}^*\mathbf{t})\mathbf{z}_j.$$

Since γ, $\mathbf{z}_i$ and ξ are nonzero (if ξ were zero then $\mathbf{D}$ would be singular and σ would be an eigenvalue of $\mathbf{A}$) we must have $\mathbf{z}^*\mathbf{t}=0$ and so $\mathbf{D}\mathbf{t}=\xi\mathbf{t}$. However $\mathbf{D}$ has distinct eigenvalues μ_j and therefore $\mathbf{t}\in\mathrm{span}\{\mathbf{e}_j\}$. But then $0=\mathbf{z}^*\mathbf{t}=\mathbf{z}_j$, yielding a contradiction. Thus ξ is not an eigenvalue of $\mathbf{D}$ and hence $\mathbf{D}-\xi\mathbf{I}$ is nonsingular and $\mathbf{z}^*\mathbf{t}\neq 0$. ☐

We use this result to prove the following theorem.

Theorem 4.2 (Interlacing Property)**.** *Consider the two eigenvalue problems*

$$\mathbf{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbf{L}^{-*}\mathbf{w} = \mu\mathbf{w} \tag{45}$$

and

$$\mathbb{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbb{L}^{-*}\hat{\mathbf{w}} = \xi\hat{\mathbf{w}}, \tag{46}$$

and assume condition (24) holds. Suppose $\mathbf{D} = diag(\mu_1,\ldots,\mu_n)\in\mathbb{R}^{n\times n}$ *and that the diagonal entries satisfy* $\mu_1<\ldots<\mu_n$. *Let* γ *be defined as in (23). Furthermore let* $\mathbf{z}$ *and* $\mathbf{t}$ *be defined as in Lemma 4.1. Assume that* $\gamma\neq 0$ *and that* $\mathbf{z}$ *has no zero components. Let*

$$\mathbf{D}\mathbf{t}_j = \xi_j(\mathbf{I}+\gamma\mathbf{z}\mathbf{z}^*)\mathbf{t}_j, \tag{47}$$

where ξ_j *are the eigenvalues, with* $\xi_1\leq\ldots\leq\xi_n$ *and* $\mathbf{t}_j$ *are the corresponding eigenvectors. Also, let* $\mu_1<\ldots<\mu_p<0<\mu_{p+1}<\ldots<\mu_n$, *where* p *is the number of negative eigenvalues of* $\mathbf{L}^{-1}(\mathbf{A}-\sigma\mathbf{I})\mathbf{L}^{-*}$. *Then*

(a) *The* ξ_j *are the* n *zeros of* $f(\xi) = 1 - \xi\gamma\mathbf{z}^*(\mathbf{D} - \xi\mathbf{I})^{-1}\mathbf{z}$.

(b) *The eigenvector* $\mathbf{t}_j$ *is a multiple of* $(\mathbf{D} - \xi_j\mathbf{I})^{-1}\mathbf{z}$.

(c) *If* $\gamma > 0$, *then*

$$\mu_1 < \xi_1 < \mu_2 < \xi_2 < \ldots < \mu_p < \xi_p < 0$$

and

$$0 < \xi_{p+1} < \mu_{p+1} < \xi_{p+2} < \mu_{p+2} < \ldots < \xi_n < \mu_n,$$

while, if $\gamma < 0$ *then*

$$\xi_1 < \mu_1 < \xi_2 < \mu_2 < \ldots < \xi_p < \mu_p < 0$$

and

$$0 < \mu_{p+1} < \xi_{p+1} < \mu_{p+2} < \xi_{p+2} < \ldots < \mu_n < \xi_n.$$

Proof. From Lemma 4.1 we know that we can reduce problems (45) and (46) to (47). If $(\mathbf{D} - \xi\gamma\mathbf{z}\mathbf{z}^*)\mathbf{t} = \xi\mathbf{t}$, then

$$(\mathbf{D} - \xi\mathbf{I})\mathbf{t} - \xi\gamma(\mathbf{z}^*\mathbf{t})\mathbf{z} = 0. \tag{48}$$

From Lemma 4.2 we know that $(\mathbf{D} - \xi\mathbf{I})$ is nonsingular. Thus

$$\mathbf{t} \in \operatorname{span}((\mathbf{D} - \xi\mathbf{I})^{-1}\mathbf{z})$$

thereby establishing (b). Applying $\mathbf{z}^*(\mathbf{D} - \xi\mathbf{I})^{-1}$ to both sides of equation (48) we get

$$\mathbf{z}^*\mathbf{t}(1 - \xi\gamma\mathbf{z}^*(\mathbf{D} - \xi\mathbf{I})^{-1}\mathbf{z}) = 0.$$

By Lemma (4.2), $\mathbf{z}^*\mathbf{t} \neq 0$ and so this shows that if ξ is an eigenvalue of the generalized problem (47) then $f(\xi) = 0$, establishing (a). To show the interlacing property (c) we need to look more carefully at the equation

$$f(\xi) = 1 - \xi\gamma\sum_{j=1}^{n}\frac{\mathbf{z}_j^2}{\mu_j - \xi}.$$

In order to find the roots of $f(\xi) = 0$ the following equality has to be satisfied

$$\frac{1}{\xi} = \gamma\sum_{j=1}^{n}\frac{\mathbf{z}_j^2}{\mu_j - \xi}. \tag{49}$$

Hence, the roots of $f(\xi) = 0$ can be found by determining the intersection points of

$$f_1(\xi) = \frac{1}{\xi} \quad \text{and} \quad f_2(\xi) = \gamma\sum_{j=1}^{n}\frac{\mathbf{z}_j^2}{\mu_j - \xi}.$$

Note that the derivative of $f_2(\xi)$ is given by

$$f_2'(\xi) = \gamma\sum_{j=1}^{n}\frac{\mathbf{z}_j^2}{(\mu_j - \xi)^2}$$

and thus the derivative is either strictly positive or strictly negative, depending on the sign of γ. Also, note that for $\xi \to \pm\infty$ we get $f_2(\xi) \to 0$. Furthermore, since $\gamma \neq 0$ and the μ_j are distinct, $f(\xi)$ has n zeroes. Depending on the sign of γ we have the following situations:

If $\gamma > 0$, then $f_2'(\xi) > 0$, that is $f_2(\xi)$ is monotonely increasing between its poles at $\xi = \mu_j$, where μ_j are the eigenvalues of (35). The hyperbola $f_1(\xi)$ is monotonely decreasing for all ξ in $(-\infty, 0)$ and $(0, \infty)$. The plot in Figure 2 illustrates the

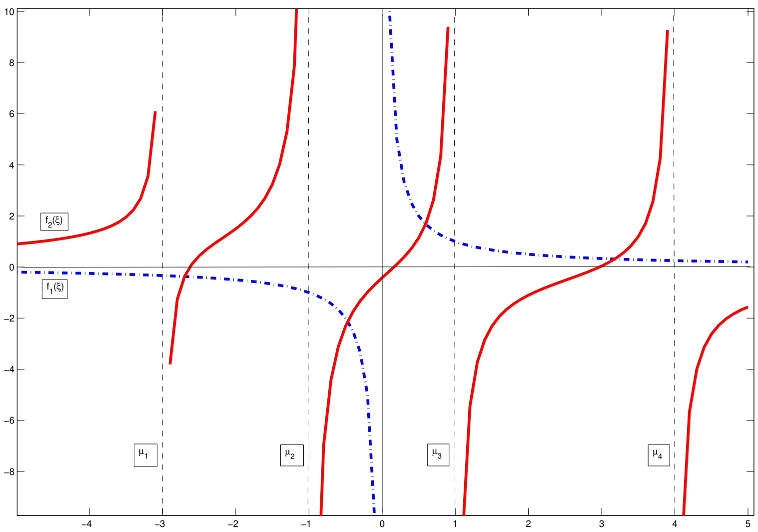

Figure 2: Intersection points of $f_1(\xi)$ and $f_2(\xi)$ for $\gamma > 0$

situation for $n = 4$ and $p = 2$. We see that due to the monotonicity properties of $f_1(\xi)$ and $f_2(\xi)$ there is exactly one intersection point of $f_1(\mu)$ and $f_2(\xi)$ between each of the poles at $\xi = \mu_j$ except in the interval containing zero. In this case there is an intersection point between μ_p and zero and a second intersection point between zero and μ_{p+1}. Next, we show that there are no intersection points $\xi > \mu_n$ and $\xi < \mu_1$, that is that the intersection points are shifted towards the origin with respect to the poles. For $\xi \to \pm\infty$ we get $f_2(\xi) \to 0$ and since $f_2(\xi)$ is monotonely increasing $f_2(\xi)$ approaches zero from below (for $\xi \to \infty$) or from above (for $\xi \to -\infty$). The decreasing hyperbola $f_1(\xi)$ does exactly the opposite and therefore the two curves cannot intersect for $\xi > \mu_n$ and $\xi < \mu_1$.

On the other hand, if $\gamma < 0$, then $f_2'(\xi) < 0$ and therefore $f_2(\xi)$ is monotonely decreasing between its poles at $\xi = \mu_j$ and the hyperbola $f_1(\xi)$ is monotonely decreasing for all ξ in $(-\infty, 0)$ and $(0, \infty)$. The plot in Figure 3 illustrates the situation for $n = 4$ and $p = 2$. Again we observe that due to the monotonicity properties of $f_1(\xi)$ and $f_2(\xi)$ there is exactly one intersection point of $f_1(\xi)$ and $f_2(\xi)$ between each of the poles at $\xi = \mu_j$ with the exception that there is no intersection between the poles $\mu_p < 0$ and $\mu_{p+1} > 0$. Next, we show that there are two further intersection points, one for $\xi > \mu_n$ and one for $\xi < \mu_1$, and hence the intersection points

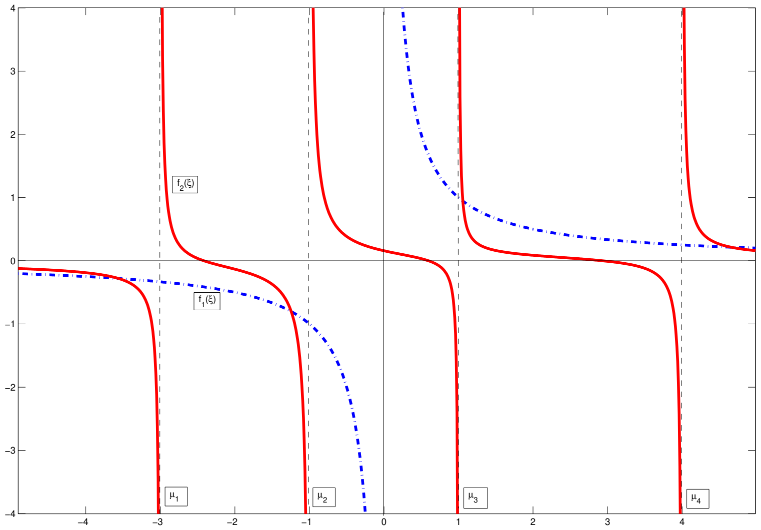

Figure 3: Intersection points of $f_1(\xi)$ and $f_2(\xi)$ for $\gamma < 0$

are shifted away from the origin with respect to the poles. Consider $\xi \to \infty$. Both functions $f_1(\xi)$ and $f_2(\xi)$ are monotonely decreasing and approaching zero. In order to show that they intersect we need to show that $f_1(\xi) > f_2(\xi)$ for $\xi \to \infty$, since, obviously close to the pole $\xi = \mu_n + \delta$, $\delta \to 0$, $f_1(\xi) < f_2(\xi)$. Hence, for $f_1(\xi) > f_2(\xi)$ for $\xi \to \infty$, we have to show

$$\frac{1}{\xi} > \gamma \sum_{j=1}^{n} \frac{\mathbf{z}_j^2}{\mu_j - \xi} \quad \text{for} \quad \xi \to \infty$$

which is equivalent to

$$1 > \gamma\xi \sum_{j=1}^{n} \frac{\mathbf{z}_j^2}{\mu_j - \xi} \quad \text{for} \quad \xi \to \infty.$$

Taking the limit we obtain

$$1 > -\gamma \sum_{j=1}^{n} \mathbf{z}_j^2 = -\gamma \mathbf{z}^* \mathbf{z}$$

and using $\mathbf{Q}^*\mathbf{v} = \mathbf{z}$ this is equivalent to

$$1 + \gamma \mathbf{v}^* \mathbf{v} > 0$$

the condition (24) in Lemma 3.2. In order to show that $f_1(\xi) < f_2(\xi)$ for $\xi \to -\infty$ a similar analysis applies. Thus we have shown that the eigenvalues are shifted away from the origin for $\gamma < 0$. $\square$

Hence, we see that for $\gamma > 0$ the eigenvalues ξ are moved towards the origin, interlacing the eigenvalues μ, whereas for $\gamma < 0$ the eigenvalues ξ are moved away from the origin interlacing the eigenvalues μ.

Theorem 4.2 is proved in the special case of no multiple eigenvalues μ and no zero components of $\mathbf{z}$. Just as in [4, Theorem 8.5.4] these restrictions are easily removed.

Theorem 4.3. *Consider the two eigenvalue problems*

$$\mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*}\mathbf{w} = \mu\mathbf{w} \tag{50}$$

and

$$\mathbb{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbb{L}^{-*}\hat{\mathbf{w}} = \xi\hat{\mathbf{w}}, \tag{51}$$

and assume condition (24) holds. Suppose $\mathbf{D} = diag(\mu_1, \ldots, \mu_n) \in \mathbb{R}^{n\times n}$ *and let* γ *be defined as in (23). Furthermore let* $\mathbf{z}$ *and* $\mathbf{t}$ *be defined as in Lemma 4.1. Assume that* $\gamma \neq 0$ *and let*

$$\mathbf{D}\mathbf{t}_j = \xi_j(\mathbf{I} + \gamma\mathbf{z}\mathbf{z}^*)\mathbf{t}_j, \tag{52}$$

where ξ_j *are the eigenvalues, with* $\xi_1 \leq \ldots \leq \xi_n$ *and* $\mathbf{t}_j$ *are the corresponding eigenvectors. Also, let* $\mu_1 \leq \ldots \leq \mu_p < 0 < \mu_{p+1} \leq \ldots \leq \mu_n$*, where* p *is the number of negative eigenvalues of* $\mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*}$ *. Then the same interlacing result as in Theorem 4.2 (c) holds, except that the strict inequalities change to equalities for* $\mathbf{z}_j = 0$ *and in case of multiple eigenvalues* μ_j *of* $\mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*}$*.*

Proof. We only need to show the result for $\mathbf{z}_j = 0$ and in case of multiple μ_j. For other cases the result follows from Theorem 4.2.

If $\mathbf{z}_j = 0$ then from (41) we obtain

$$\mathbf{D}\mathbf{e}_j = \xi(\mathbf{I} + \gamma\mathbf{z}\mathbf{z}^*)\mathbf{e}_j = \xi\mathbf{e}_j,$$

where $\mathbf{e}_j$ is the jth canonical vector. Hence $\xi_j = \mu_j$ with corresponding eigenvector $\mathbf{e}_j$ which is even better than interlacing. Furthermore, if $\mu_j = \mu_{j+1}$ we can transform the problem to a problem with a zero component of $\mathbf{z}$. Let $\mathbf{U} = \mathbf{G}(j, j+1, \theta)$ be a (orthogonal) Givens rotation in the $(j, j+1)$ plane with the property that $\tilde{\mathbf{z}}_{j+1} = 0$, that is

$$\mathbf{U}\mathbf{z} = [\mathbf{z}_1, \ldots, \tilde{\mathbf{z}}_j, 0, \mathbf{z}_{j+2}, \ldots, \mathbf{z}_n]^* = \tilde{\mathbf{z}}.$$

It is not hard to show that $\mathbf{U}^*\mathbf{D}\mathbf{U} = \mathbf{D}$. Hence

$$\mathbf{U}^*(\mathbf{D} - \xi\gamma\mathbf{z}\mathbf{z}^*)\mathbf{U} = \mathbf{D} - \xi\gamma\tilde{\mathbf{z}}\tilde{\mathbf{z}}^*$$

and using the previous observation for $\tilde{\mathbf{z}}_{j+1} = 0$ we get $\mu_{j+1} = \mu_j$ is an eigenvalue ξ of the generalized problem (47) with corresponding eigenvector $\mathbf{U}\mathbf{e}_{j+1}$. ☐

Remark 4.2. *Combining the results of Theorem 4.1 and 4.2 we obtain one sided bounds for the largest and smallest eigenvalues of* $\mathbb{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbb{L}^{-*}$ *in terms of the largest and smallest eigenvalues of* $\mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*}$*. Thus we also obtain bounds on the condition number of* $\mathbb{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbb{L}^{-*}$ *in terms of the condition number of* $\mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*}$*. Furthermore we can conclude that any eigenvalue clustering properties of* $\mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*}$ *are preserved in* $\mathbb{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbb{L}^{-*}$*.*

Thus we are able to obtain qualitative and quantitative information about the quality of $\mathbb{L}$ *as a preconditioner compared with* $\mathbf{L}$*.*

4.3 Consequences for the Tuned Preconditioner

Here we merely compare the various terms which appear in (13) and (22), which give bounds for the inner iterations in MINRES using the standard and tuned preconditioners respectively.

As before, we assume that in our investigation all the eigenvalues $\mu_2, \dots, \mu_n$ of $\mathbf{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbf{L}^{-*}$ are positive and μ_1, the extremal eigenvalue is negative. Thus (using Silvester's Inertia Theorem) the eigenvalues $\xi_2, \dots, \xi_n$ of $\mathbb{L}^{-1}(\mathbf{A} - \sigma\mathbf{I})\mathbb{L}^{-*}$ are positive and ξ_1 is negative. We have to compare the reduced condition numbers

$$\kappa^1_{\mathbf{L}} = \frac{|\mu_n|}{|\mu_2|} \quad \text{and} \quad \kappa^1_{\mathbb{L}} = \frac{|\xi_n|}{|\xi_2|}$$

as well as the terms

$$\frac{|\mu_1 - \mu_n|}{|\mu_1|} \quad \text{and} \quad \frac{|\xi_1 - \xi_n|}{|\xi_1|}.$$

Situation for $\gamma > 0$

If $\gamma > 0$ then, from Theorem 4.2, the eigenvalues ξ are shifted towards the origin with respect to the eigenvalues μ. Hence

$$\xi_n \leq \mu_n$$

holds and from (37) we get

$$\frac{\mu_2}{1 + |\gamma \mathbf{v}^* \mathbf{v}|} \leq \xi_2.$$

Combining both bounds yields

$$\kappa^1_{\mathbb{L}} = \frac{|\xi_n|}{|\xi_2|} \leq \frac{|\mu_n|}{|\mu_2|}(1 + |\gamma \mathbf{v}^* \mathbf{v}|) = \kappa^1_{\mathbf{L}}(1 + |\gamma \mathbf{v}^* \mathbf{v}|), \tag{53}$$

an upper bound on the change to the reduced condition number due to tuning.

Using a similar consideration we obtain

$$\frac{|\mu_1 - \mu_n|}{|\mu_1|} \leq (1 + |\gamma \mathbf{v}^* \mathbf{v}|)\frac{|\xi_1 - \xi_n|}{|\xi_1|} \tag{54}$$

for $\gamma > 0$.

Situation for $\gamma < 0$

For $\gamma < 0$ a similar discussion also yields (53) and

$$\frac{|\xi_1 - \xi_n|}{|\xi_1|} \leq (1 + |\gamma \mathbf{v}^* \mathbf{v}|)\frac{|\mu_1 - \mu_n|}{|\mu_1|}. \tag{55}$$

4.4 Numerical example

We consider a numerical example to support our theory in this section and compare the reduced condition numbers.

Example 4.1. *We consider the matrix* `nos5.mtx` *from the Matrix Market library [10]. This is a matrix of size* 468*. We use a shift* $\sigma = 55$ *which is close to the smallest eigenvalue of* $(\mathbf{A} - \sigma\mathbf{I})$ *and which leads to exactly one negative eigenvalue of* $(\mathbf{A} - \sigma\mathbf{I})$*. Again, we choose* $\mathbf{x}$ *to be a random perturbation from the eigenvector belonging to the smallest eigenvalue. Note that in this case* $\gamma > 0$ *and condition (24) is ensured.*

Table 2: Results for Example 4.1. The table gives values for $\mathbf{u}^*\mathbf{x} = \frac{1}{\gamma}$, $1 + |\gamma\mathbf{v}^*\mathbf{v}|$, $\kappa^1_{\mathbf{L}}$ and $\kappa^1_{\mathbb{L}}$ and for different drop tolerances

DROP TOLERANCE	$\mathbf{u}^*\mathbf{x}$	$1 + \|\gamma\mathbf{v}^*\mathbf{v}\|$	$\kappa^1_{\mathbf{L}}$	$\kappa^1_{\mathbb{L}}$
0.5	24837.2	3.1230	10393.2	12068.2
0.2	31020.6	2.1342	588.9	595.7

Table 2 shows the results for Example 4.1. For more information we consider the corresponding reduced condition numbers which influence the convergence rates of MINRES.

We observe that the change in the condition numbers and thus the change in the convergence rate is moderate. In fact, it only changes in the third or fourth significant digit. We also observe that the on the perturbation of the reduced condition numbers (53) obtained in this section are not sharp at all and the actual perturbation of the reduced condition number is very small since for both drop tolerances $\kappa^1_{\mathbb{L}} \ll \kappa^1_{\mathbf{L}}(1 + |\gamma\mathbf{v}^*\mathbf{v}|)$.

5 Inexact Rayleigh quotient iteration

First, we summarise the theory of the inexact Rayleigh quotient iteration for the standard Hermitian positive definite eigenvalue problem (1). Rayleigh quotient iteration is a special version of inverse iteration where the variable shift in (2) is chosen to be the Rayleigh quotient

$$\lambda^{(i)} = \rho(\mathbf{x}^{(i)}) = \frac{\mathbf{x}^{(i)^*}\mathbf{A}\mathbf{x}^{(i)}}{\mathbf{x}^{(i)^*}\mathbf{x}^{(i)}}. \tag{56}$$

The following algorithm is a version of inexact inverse iteration with Rayleigh quotient shifts.

Algorithm 5.1 (Inexact Rayleigh quotient iteration)**.** *Given* $\mathbf{x}^{(0)}$ *with* $\|\mathbf{x}^{(0)}\| = 1$*. Compute* $\lambda^{(0)} = \mathbf{x}^{(0)^*}\mathbf{A}\mathbf{x}^{(0)}$*. For* $i = 0, 1, 2, \ldots$

(1) Choose $\tau^{(i)}$*,*

(2) Solve $(\mathbf{A}-\lambda^{(i)}\mathbf{I})\mathbf{y}^{(i)}=\mathbf{x}^{(i)}$ *inexactly, that is,*

$$\| (\mathbf{A}-\lambda^{(i)}\mathbf{I})\mathbf{y}^{(i)}-\mathbf{x}^{(i)} \| \leq \tau^{(i)},$$

(3) Compute approximate eigenvector $\mathbf{x}^{(i+1)}=\dfrac{\mathbf{y}^{(i)}}{\|\mathbf{y}^{(i)}\|}$*,*

(2) Compute approximate eigenvalue $\lambda^{(i+1)}=\mathbf{x}^{(i+1)^*}\mathbf{A}\mathbf{x}^{(i+1)}$*,*

(5) Evaluate eigenvalue residual $\mathbf{r}^{(i+1)}=(\mathbf{A}-\lambda^{(i+1)}\mathbf{I})\mathbf{x}^{(i+1)}$*,*

(6) Test for convergence.

The following theorem states the convergence theory for inexact Rayleigh quotient iteration.

Theorem 5.1 (Convergence of inexact Rayleigh quotient iteration)**.** *Let (1) be the standard eigenvalue problem for a Hermitian matrix* $\mathbf{A}$ *and consider the application of Algorithm 5.1 to find a simple eigenpair* $(\lambda_1, \mathbf{x}_1)$*. Depending on the tolerance* $\tau^{(i)}$ *the following rates of convergence are obtained by inexact Rayleigh quotient iteration (Algorithm 5.1) with a sufficiently close starting guess.*

1. Decreasing tolerance. *If* $\tau^{(i)} \leq C_1 \|\mathbf{r}^{(i)}\|$ *in step (1) Algorithm 5.1, then cubic convergence is achieved by Algorithm 5.1.*
2. Fixed tolerance. *If* $\tau^{(i)}=\tau$ *in step (1) Algorithm 5.1, then quadratic convergence is achieved by Algorithm 5.1.*

For a tolerance $\tau^{(i)}=0$ *(exact RQI) we obtain cubic convergence.*

Proof. Proofs for exact Rayleigh quotient iteration can be found, for example in [13], and for inexact Rayleigh Quotient iteration in [1] and [16]. □

5.1 Preconditioned inexact Rayleigh quotient iteration

In this section we study the application of a preconditioner for the solution of the linear system in step (2) of Algorithm 5.1 and thereby revisit the preconditioners from the previous section.

For the solution by MINRES of the preconditioned system in the inexact Rayleigh quotient method we have for the standard preconditioner

$$\mathbf{L}^{-1}(\mathbf{A}-\lambda^{(i)}\mathbf{I})\mathbf{L}^{-*}\tilde{\mathbf{y}}^{(i)}=\mathbf{L}^{-1}\mathbf{x}^{(i)}, \quad \mathbf{y}^{(i)}=\mathbf{L}^{-*}\tilde{\mathbf{y}}^{(i)}, \tag{57}$$

where $\mathbf{A}=\mathbf{L}\mathbf{L}^*+\mathbf{E}$ as in (11), and

$$\mathbb{L}^{-1}(\mathbf{A}-\lambda^{(i)}\mathbf{I})\mathbb{L}^{-*}\tilde{\mathbf{y}}^{(i)}=\mathbb{L}^{-1}\mathbf{x}^{(i)}, \quad \mathbf{y}^{(i)}=\mathbb{L}^{-*}\tilde{\mathbf{y}}^{(i)}, \tag{58}$$

for the tuned preconditioner given by Lemma 3.2. Clearly the outer rate of convergence of Algorithm 5.1 is unaltered by the choice of $\mathbf{L}$ or $\mathbb{L}$, but we shall see in Example 5.1 that $\mathbb{L}$ shows significant numerical advantages over $\mathbf{L}$, because of the fact that $\mathbb{L}^{-1}\mathbf{x}^{(i)}$ is roughly in the direction of the eigenvector of $\mathbb{L}^{-1}(\mathbf{A} - \lambda^{(i)}\mathbf{I})\mathbb{L}^{-*}$ corresponding to the eigenvalue nearest zero.

In a novel strategy, Simoncini & Eldén [15] suggest that the right hand side of (57) be altered so that one solves the modified Hermitian system

$$\mathbf{L}^{-1}(\mathbf{A} - \lambda^{(i)}\mathbf{I})\mathbf{L}^{-*}\tilde{\mathbf{y}}^{(i)} = \mathbf{L}^{*}\mathbf{x}^{(i)}, \quad \mathbf{y}^{(i)} = \mathbf{L}^{-*}\tilde{\mathbf{y}}^{(i)}, \tag{59}$$

in step (2) of Algorithm 5.1 rather than (57). We remark that this strategy has been used before to enhance the Lanczos process [11],[14], however the motivation in Simoncini & Eldén [15] is new. They noted that $\mathbf{L}^{*}\mathbf{x}^{(i)}$ is an approximation to the eigenvector of the coefficient matrix corresponding to the eigenvalue closest to zero (see Section 6 of [15]). This method is analysed in [1] where the advantages of (59) over (57) from the point of view of the inner iteration count are discussed carefully. However (59) only gives quadratic outer convergence, since (59) is equivalent to $(\mathbf{A} - \lambda^{(i)}\mathbf{I})\mathbf{y}^{(i)} = \mathbf{L}\mathbf{L}^{*}\mathbf{x}^{(i)}$ and so the right hand side is altered from the traditional $\mathbf{x}^{(i)}$. Also, if this approach is used there is no advantage in using a decreasing tolerance for the inexact solves applied to (59), since any method based on (59) would normally only converge quadratically if $\mathbf{L}\mathbf{L}^{*}\mathbf{x}^{(i)} \neq \mathbf{x}^{(i)}$ due to the quadratic convergence of the Rayleigh quotient to the desired eigenvalue for a close enough starting guess.

In fact there is a close relationship between tuning the preconditioner and the approach of [15] as we now show. Equation (58) can be written as

$$\mathbb{L}^{-1}(\mathbf{A} - \lambda^{(i)}\mathbf{I})\mathbb{L}^{-*}\tilde{\mathbf{y}}^{(i)} = \mathbb{L}^{*}\mathbb{L}^{-*}\mathbb{L}^{-1}\mathbf{x}^{(i)},$$

and using (17)

$$\mathbb{L}^{-1}(\mathbf{A} - \lambda^{(i)}\mathbf{I})\mathbb{L}^{-*}\tilde{\mathbf{y}}^{(i)} = \frac{1}{\lambda^{(i)}}\mathbb{L}^{*}\mathbf{x}^{(i)} - \frac{1}{\lambda^{(i)}}\mathbb{L}^{-1}\mathbf{r}^{(i)}. \tag{60}$$

We see that (60) has the form of (59), but with a perturbed and scaled right hand side.

In Example 5.2 we compare the numerical performance of (58) and (59).

5.2 Numerical Example

Here we consider a numerical example to compare the methods discussed above.

Example 5.1 (Problem from the Matrix Market library [10]). *Consider the same matrix* `nos5.mtx` *and setup as in Example 3.1 but use Rayleigh quotient shift (56). We seek the third smallest eigenvalue, given by* $\lambda_3 \approx 115.5912$. *The starting approximation* $\mathbf{x}^{(0)}$ *is chosen to be sufficiently close to* $\mathbf{x}_3$. *We compare the costs of the following methods:*

(a) *Standard incomplete Cholesky preconditioner: Algorithm 5.1 with (2) implemented by solving (57), where* $\mathbf{LL}^*$ *is the incomplete Cholesky factorisation of* $\mathbf{A}$.

(b) *Tuned incomplete Cholesky preconditioner: Algorithm 5.1 with (2) implemented by solving (58), where* $\mathbb{L}$ *is given by (26).*

For the inexact solves we use the decreasing tolerance $\tau^{(i)} = \min\{0.1, 0.1\|\mathbf{r}^{(i)}\|\}$. *We use the incomplete Cholesky factorisation of* $\mathbf{A}$ *given by* $\mathbf{LL}^*$ *with drop tolerances* 0.25 *(leading to* 662 *nonzero entries in* $\mathbf{L}$*) and* 0.1 *(leading to* 1032 *nonzero entries in* $\mathbf{L}$*). The computations stop once the eigenvalue residual satisfies*

$$\|\mathbf{r}^{(i)}\| < 10^{-10}.$$

Table 3: Iteration numbers for Example 5.1 using *Rayleigh quotient shifts*. The total number of iterations and number of inner iterations for inexact Rayleigh quotient iteration using either the standard incomplete Cholesky preconditioner (a), or the tuned preconditioner (b).

	Standard preconditioner		*Tuned preconditioner*	
	DROP TOLERANCES			
OUTER ITERATION	0.25	0.1	0.25	0.1
1	68	62	61	56
2	85	76	69	63
3	148	132	90	78
total	301	270	220	197

In Table 3 the results for methods (a) and (b) are given. Both the methods achieve

Table 4: Error propagation $\|\mathbf{A}\mathbf{x}^{(i)} - \lambda^{(i)}\mathbf{x}^{(i)}\|_2$ for Example 5.1 using *Rayleigh quotient shift* for inexact RQI with preconditioned solves using methods (a) and (b)

	Standard preconditioner		*Tuned preconditioner*	
	DROP TOLERANCES			
OUTER ITERATION	0.25	0.1	0.25	0.1
1	17.45	17.45	17.45	17.45
2	6.3e-2	6.2e-2	6.2e-2	6.2e-2
3	1.7e-7	1.3e-7	2.1e-7	1.5e-7
4	8.5e-11	6.3e-11	2.4e-11	2.1e-11

cubic outer convergence but the tuned preconditioner requires fewer number of inner iterations than the standard preconditioner, since the right hand side is in the direction of an approximate null-space of the iteration matrix. Note that in Table 4 cubic convergence can hardly be observed, since it only sets in for close enough solutions, but then the precision of MATLAB is not high enough.

Example 5.2 (Comparison between Simoncini & Eldén [15] and the tuned preconditioner)**.** *We use the same matrix and setup as in Example 5.1 and look for the same eigenvalue. We compare the tuned preconditioner using 58 to the modified right hand side approach of Simoncini & Eldén using (59). We solve both (59) and (58) to the fixed tolerance of* $\tau = 0.01$*, so both methods yield quadratic convergence.*

Table 5: Iteration numbers for Example 5.2 using *Rayleigh quotient shifts*. The total number of iterations and number of inner iterations for inexact Rayleigh quotient iteration using either the the modified right hand side approach by Simoncini & Eldén or the tuned preconditioner

	Simoncini & Eldén		*Tuned preconditioner*	
	Drop Tolerances			
Outer Iteration	0.25	0.1	0.25	0.1
1	67	62	29	26
2	74	66	56	55
3	85	75	71	67
4	63		18	
total	289	203	174	148

From Table 5 we observe that the tuned preconditioner requires fewer inner iterations than the modified right hand side approach.

6 Conclusions and further thoughts

We have analysed the behaviour of preconditioned MINRES solves in inexact inverse iteration using both a fixed shift and Rayleigh quotient shifts for computing the eigenpairs of a given Hermitian positive definite matrix.

We derived a new tuned preconditioner and provided a detailed spectral analysis on both the system matrices arising for the standard preconditioner and for the tuned preconditioner. We analysed the performance of both preconditioners in the context of the convergence of MINRES.

For inexact inverse iteration with a fixed shift we compared the performance of the standard incomplete Cholesky preconditioner with the tuned Cholesky preconditioner. The tuned preconditioner is shown to yield a improvement over the standard preconditioner with regard to the total number of inner iterations.

When Rayleigh quotient shifts are used the tuned preconditioner is again superior to the standard preconditioner. The method based on the tuned preconditioning is also seen to be superior to the method of Simoncini & Eldén. We remark that the method based on the tuned preconditioner is also likely to be superior when the cubic convergence reduces the number of outer iterations needed compared with the method of [15].

In summary, the tuned preconditioner used in an inexact iterative method with decreasing tolerance, combines the advantages of maintaining the outer convergence rate achieved by exact solves, with the efficient inner iteration performance exhibited by the Simoncini & Eldén method.

References

[1] J. Berns-Müller, I. G. Graham, and A. Spence, *Inexact inverse iteration for symmetric matrices*, 2005. Submitted to Linear Algebra Appl.

[2] J. W. Demmel, *Applied Numerical Linear Algebra*, SIAM, Philadelphia, 1997.

[3] M. A. Freitag and A. Spence, *Convergence rates for inexact inverse iteration with application to preconditioned iterative solves*, 2005. Submitted to BIT.

[4] G. Golub and C. V. Loan, *Matrix Computations*, John Hopkins University Press, Baltimore, 3rd ed., 1996.

[5] G. H. Golub and Q. Ye, *Inexact preconditioned conjugate gradient method with inner-outer iteration*, SIAM Journal on Scientific Computing, 21 (1999), pp. 1305–1320.

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

[7] A. Greenbaum, *Iterative methods for solving linear systems*, vol. 17 of Frontiers in Applied Mathematics, SIAM, Philadelphia, PA, 1997.

[8] W. Hackbusch, *Iterative solution of large sparse systems of equations*, Springer-Verlag, Berlin, 1994.

[9] Y.-L. Lai, K.-Y. Lin, and L. Wen-Wei, *An inexact inverse iteration for large sparse eigenvalue problems*, Numerical Linear Algebra with Applications, 1 (1997), pp. 1–13.

[10] *Matrix market.* available online at `http://math.nist.gov/MatrixMarket/`, 2004.

[11] R. B. Morgan and D. S. Scott, *Preconditioning the Lanczos algorithm for sparse symmetric eigenvalue problems*, SIAM Journal on Scientific Computing, 14 (1993), pp. 585–593.

[12] A. M. Ostrowski, *On the convergence of the Rayleigh quotient iteration for the computation of the characteristic roots and vectors. I*, Archive for Rational Mechanics and Analysis, 1 (1957), pp. 233–241.

[13] B. N. Parlett, *The Symmetric Eigenvalue Problem*, vol. 20 of Classics in Applied Mathematics, SIAM, Philadelphia, PA, 1998. Corrected reprint of the 1980 original.

[14] D. S. Scott, *Solving sparse symmetric generalized eigenvalue problems without factorization*, SIAM Journal on Numerical Analysis, 18 (1981), pp. 102–110.

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

[16] P. Smit and M. H. C. Paardekooper, *The effects of inexact solvers in algorithms for symmetric eigenvalue problems*, Linear Algebra and its Applications, 287 (1999), pp. 337–357.

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