CONVERGENCE ANALYSIS OF ITERATIVE SOLVERS IN INEXACT RAYLEIGH QUOTIENT ITERATION*

FEI XUE[†] AND HOWARD C. ELMAN[‡]

Abstract. We present a detailed convergence analysis of preconditioned MINRES for approximately solving the linear systems that arise when Rayleigh Quotient Iteration is used to compute the lowest eigenpair of a symmetric positive definite matrix. We provide insight into the "slow start" of MINRES iteration in both a qualitative and quantitative way, and show that the convergence of MINRES mainly depends on how quickly the unique negative eigenvalue of the preconditioned shifted coefficient matrix is approximated by its corresponding harmonic Ritz value. By exploring when the negative Ritz value appears in MINRES iteration, we obtain a better understanding of the limitation of preconditioned MINRES in this context and the virtue of a new type of preconditioner with "tuning". Finally we show that tuning based on a rank-2 modification can be applied with little additional cost to guarantee positive definiteness of the tuned preconditioner.

Key words. Rayleigh Quotient Iteration, harmonic Ritz value, MINRES, tuned preconditioner

1. Introduction. There has been considerable interest in recent years in developing and analyzing eigensolvers that involve at each step (outer iteration) a shift-invert matrix-vector product implemented by solving the shifted linear system iteratively (inner iteration). Inexact inverse iteration is the most simple algorithm of this type and the best understood one. Early papers on the convergence of inexact inverse iteration with fixed shift include [8] and [9], where the main concern is to choose a decreasing sequence of stopping tolerances for inner solvers to maintain linear convergence of the outer iteration. Analysis of inexact Rayleigh Quotient Iteration (RQI) in [13] and [17] shows how the inexactness of the inner solve can affect the convergence of the outer iteration. More recent work focuses on improving the convergence of inner iterations as well as the relation between the inner and outer iterations. Reference [16] introduces some new perspectives on preconditioning in this setting, namely, that faster convergence of inner iterations can be obtained by modifying the right hand side of the preconditioned linear system. Refined analysis of this approach in [1],[2] and [5] shows how different formulations of the linear system, with variable shift and different inner stopping criteria, can affect the convergence of the inner and outer iterations. An alternative preconditioning approach called "tuning" is analyzed in [6] for non-symmetric eigenvalue problems and in [7] for symmetric ones. A tuned preconditioner is a low rank modification of an ordinary preconditioner. Tuning forces the preconditioning operator to behave in the same way as the system matrix on the current approximate eigenvector. The ideas in [16] and [7] are, respectively, to modify the right hand side, or to modify the preconditioner (tune it), so that the preconditioned right hand side approximates the eigenvector of the preconditioned coefficient matrix; in either case, the inner iteration counts can be greatly reduced.

In this paper, we give a detailed analysis of MINRES, both with and without preconditioning, for the inner iteration of RQI for symmetric eigenvalue problems, and we introduce some new approach for preconditioning. It is suggested in [16] that fast convergence of MINRES iteration in this setting implies fast improvement of the eigenvector approximation by the MINRES iterate. Hence by analyzing MINRES convergence, we know how quickly the angle between the MINRES iterate and the true eigenvector decreases as the MINRES iteration proceeds. This perspective has not been emphasized in the literature, and it is adopted in the paper to compare the performance of different versions of MINRES to solve the linear systems in RQI.

We study the convergence of three versions of MINRES used in RQI: unpreconditioned MINRES, preconditioned MINRES with symmetric positive definite preconditioner Q, and preconditioned MINRES with a tuned variant of Q. Using the properties of the harmonic Ritz values and their connection with the MINRES residual polynomial, we provide new insight into the limitations of preconditioning without tuning and show how tuning leads to a major improvement. We show that the convergence of unpreconditioned MINRES and preconditioned MINRES with tuning depends on the angle between the current outer iterate and the true eigenvector as well as the reduced condition number of the (preconditioned) shifted coefficient matrix. In addition, a new tuning strategy based on a rank-2 modification is introduced to guarantee positive definiteness of the tuned preconditioner.

*This work was supported by the U. S. Department of Energy under grant DEFG0204ER25619, and by the U. S. National Science Foundation under grant CCF0726017.

[†]Applied Mathematics and Scientific Computation Program, Department of Mathematics, University of Maryland, College Park, MD 20742. (petrinet@math.umd.edu).

[‡]Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742. (elman@cs.umd.edu)

The paper is organized as follows. Section 2 reviews some preliminary facts and results. Section 3, the major part of the paper, gives detailed convergence analysis of the three versions of MINRES used in RQI. A rank-2 tuning is introduced in Section 4 as an improvement of the rank-1 tuning of [7]. Numerical experiments are given in Section 5, followed by the conclusion in Section 6.

2. Preliminaries. We want to compute the lowest eigenpair of a symmetric positive definite matrix by Rayleigh Quotient Iteration. Consider the eigenvalue problem $Av = \lambda v$, where A is symmetric positive definite with eigenvalues $0 < \lambda_1 < \lambda_2 \leq ... \leq \lambda_n$. Let $V = [v_1, v_2, ..., v_n] = [v_1, V_2]$ be the matrix of orthonormal eigenvectors and let $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_n)$ so that $V^T AV = \Lambda$. Algorithm 1 describes a typical version of inexact Rayleigh Quotient Iteration to find a simple eigenpair.

Algorithm 1: Inexact Rayleigh Quotient Iteration
Given $x^{(0)}$ with $\|x^{(0)}\| = 1$
For i = 0,1,..., until convergence
1. Compute the Rayleigh Quotient $\sigma^{(i)} = x^{(i)T}Ax^{(i)}$
2. Choose $\tau^{(i)}$ and solve $(A - \sigma^{(i)}I)y^{(i)} = x^{(i)}$ inexactly such that $\|x^{(i)} - (A - \sigma^{(i)}I)y^{(i)}\| \leq \tau^{(i)}$
3. Update $x^{(i+1)} = y^{(i)}/\|y^{(i)}\|$
4. Test for convergence
End For

From here through the end of the paper, we drop the superscripts (i) that denote the count of the outer iteration, because we are interested in the convergence of inner iterations. Suppose a normalized outer iterate x is close to v_1 such that $x = \sum_{k=1}^{n} v_k c_k = v_1 \cos\varphi + u \sin\varphi$, where $u \perp v_1$ and $\|u\| = 1$; $\varphi = \angle(x, v_1)$, so that $\cos\varphi = c_1 = v_1^T x$, and $\sin\varphi = \|[0, V_2]^T x\| = \sqrt{c_2^2 + \cdots + c_n^2}$ is small. The Rayleigh quotient associated with x is $\sigma = x^T Ax = \lambda_1 + \sum_{k=2}^{n} (\lambda_k - \lambda_1)c_k^2 = \lambda_1 + (\bar{\lambda} - \lambda_1)\sin^2\varphi$, where $\bar{\lambda} \in [\lambda_2, \lambda_n]$ is a weighted average of $\lambda_2, ..., \lambda_n$. Assume that λ_1 is well-separated from λ_2, and φ is so small that $|\lambda_1 - \sigma| = O(\sin^2\varphi) \ll |\lambda_2 - \sigma| = O(1)$; hence v_1 is the dominant eigenvector of $(A - \sigma I)^{-1}$.

Recall that there is a connection between the Lanczos algorithm for eigenvalues of a symmetric matrix B and the MINRES and SYMMLQ methods for solving systems $By = b$. Given the starting vector $u_1 = b/\|b\|$, the Lanczos algorithm leads to

$$BU_m = U_m T_m + \beta_{m+1} u_{m+1} e_m^T = U_{m+1} \overline{T}_m \tag{2.1}$$

where the tridiagonal matrix $T_m = \mathrm{tridiag}[\beta_j, \alpha_j, \beta_{j+1}]$ $(1 \leq j \leq m)$ comes from the well-known three-term recurrence formula. Our analysis mainly results from the convergence of the leftmost harmonic Ritz value to the leftmost eigenvalue of B as m increases.

We will use a major theorem from [14], which characterizes the MINRES iterate and establishes a connection between the residual polynomial and the harmonic Ritz values. Our analysis builds on this theorem and the interlacing property of Ritz and harmonic Ritz values.

THEOREM 2.1. *Suppose* MINRES *is applied to solve the system $By = b$. At the m-th* MINRES *iteration step with the corresponding Lanczos decomposition in* (2.1), *the* MINRES *iterate is*

$$y_m = U_m M_m^{-2} T_m e_1 \beta_1, \quad \text{where } M_m^2 = \overline{T}_m^T \overline{T}_m,\ \beta_1 = \|b\|. \tag{2.2}$$

The residual of the linear system is

$$r_m = b - By_m = \chi_m(B)b/\chi_m(0), \tag{2.3}$$

where $\chi(\lambda) = \prod_{i=1}^{m}(\lambda - \xi_i^{(m)}) = \det[\lambda I_m - T_m^{-1} M_m^2]$ is the residual polynomial whose roots are the harmonic Ritz values $\xi_i^{(m)}$, defined as eigenvalues of the pencil $M_m^2 - \xi T_m$. $BU_m M_m^{-1}$ has orthonormal columns and $1/\xi_i^{(m)}$ are the eigenvalues of $H_m = (BU_m M_m^{-1})^T B^{-1} (BU_m M_m^{-1}) = M_m^{-T} T_m M_m^{-1}$.

3. Convergence of MINRES in inexact RQI. In this section, we analyze the convergence of the three versions of MINRES for solving the linear system in RQI. We consider in turn unpreconditioned MINRES, preconditioned MINRES with an ordinary symmetric positive definite preconditioner Q (without tuning), and preconditioned MINRES with a tuned variant of Q.

The analysis is based on properties of harmonic Ritz values. To fix notation in the following subsections, we use θ for Ritz values, ξ for harmonic Ritz values, quantities with hat for the preconditioned system without tuning and those with tilde for the preconditioned system with tuning. B and b are respectively the shifted system matrix and right hand side of the (preconditioned) system in step 2 of algorithm 1.

3.1. Unpreconditioned MINRES. It is observed in [16] that the convergence of unpreconditioned MINRES for $(A-\sigma I)y = x$ can be very slow when the Rayleigh quotient σ is close to λ_1, i.e., when $\varphi = \angle(x, v_1)$ is small enough. That is, the residual norm $\|r_m\| = \|x-(A-\sigma I)y_m\|$ is still close to 1 for quite large m. It is shown that as a result, $\angle(y_m, v_1)$ (the angle between the MINRES iterate and the true eigenvector) also decreases quite slowly. This *slow start* phenomenon is described in the theorem below. A proof is given in [19] (Appendix A).

THEOREM 3.1. *Suppose unpreconditioned* MINRES *is used to solve* $(A-\sigma I)y = x$ *in* RQI, *where* $x = v_1\cos\varphi + u\sin\varphi$. *Assume that* u *has components of* m *eigenvectors of* A *so that* MINRES *will not give the exact solution at the first* m *steps. For any such fixed* u, $\lim_{\varphi\to 0}\|r_k\| = 1$ *for any* $k \le m$. *Moreover, for any given* $k \le m$, *if* φ *is small enough*[1], *then* $1-\|r_k\| = O(\sin^2\varphi)$.

Remark. This residual norm estimate shows qualitatively that the slow start of the inner iteration is more pronounced as the outer iterate x becomes closer to the true eigenvector v_1. For any given $k \le m$, the theorem shows that $\|r_k\|$ tends to be closer to 1 as φ becomes smaller.

In the context of using MINRES in RQI to compute (λ_1, v_1), we are more interested in how quickly $\angle(y_m, v_1)$ decrease with m. Theorem 4.1 of [16] establishes the fact that $\angle(y_m, v_1)$ decreases quickly only if MINRES converges rapidly. We restate the theorem with our notation, and expand on the result by showing that the leftmost harmonic Ritz value $\xi_1^{(m)}$ plays a critical role in the behavior of $\angle(y_m, v_1)$.

THEOREM 3.2. *Let* (μ_i, v_i) *be the eigenpairs of the shifted matrix* $B = A-\sigma I$, *with eigenvalues ordered as* $0 < |\mu_1| < |\mu_2| \le \ldots \le |\mu_n|$. *Let* x *be a unit norm approximation to* v_1 *with small* $\varphi = \angle(x, v_1)$. *Let* y_m *be the* MINRES *approximate solution in* $\mathcal{K}_m(B, x)$ *and* $r_m = x - By_m = p_m(B)x$ *be the associated linear residual. If* $|p_m(\mu_1)| < 1$, *then*

$$\tan\angle(y_m, v_1) \le \frac{|\mu_1|}{|\mu_2|}\frac{1}{|1-p_m(\mu_1)|}\left(1+\frac{(\|r_m\|^2-|p_m(\mu_1)|^2\cos^2\varphi)^{1/2}}{\sin\varphi}\right)\tan\varphi \tag{3.1}$$

or approximately,

$$\tan\angle(y_m, v_1) \le \frac{|\xi_1^{(m)}|}{|\mu_2|}(1+\max_{2\le i\le n}|p_m(\mu_i)|)\tan\varphi \tag{3.2}$$

Proof. The result (3.1) is established in [16]. For (3.2), first recall that as φ is small, $B = A-\sigma I$ has the unique negative eigenvalue $\mu_1 = \lambda_1 - \sigma = O(\sin^2\varphi)$ and the smallest positive eigenvalue $\mu_2 = \lambda_2 - \sigma = O(1)$. Recall also the interlacing property mentioned in [14], that the Ritz values $\{\theta_k^{(m)}\}$ interlace the harmonic Ritz values $\{\xi_k^{(m)}\}\cup\{0\}$. Since $\det[T_2] = -\beta_2^2 = \theta_1^{(2)}\theta_2^{(2)} < 0$, we have $\xi_1^{(2)} < \theta_1^{(2)} < 0 < \theta_2^{(2)} < \xi_2^{(2)}$. To analyze the convergence of MINRES, recall from Theorem 2.1 that the harmonic Ritz values $\xi_k^{(m)}$ are zeros of the residual polynomial. That is,

$$p_m(\mu) = \prod_{k=1}^{m}(1-\mu/\xi_k^{(m)}). \tag{3.3}$$

Therefore, the residual vector can be represented as

$$\begin{aligned} r_m = p_m(B)x &= p_m(B)\sum_{i=1}^{n} c_i v_i = \sum_{i=1}^{n} p_m(\mu_i)c_i v_i = p_m(\mu_1)\cos\varphi\ v_1 + \sin\varphi\sum_{i=2}^{n}(c_i/\sin\varphi)p_m(\mu_i)v_i \\ &= \cos\varphi\prod_{k=1}^{m}(1-\mu_1/\xi_k^{(m)})v_1 + \sin\varphi\sum_{i=2}^{n}\omega_i\prod_{k=1}^{m}(1-\mu_i/\xi_k^{(m)})v_i, \end{aligned} \tag{3.4}$$

where $\mu_i = \lambda_i - \sigma$ and $\omega_i = c_i/\sin\varphi$ is such that $\sum_{i=2}^{n}\omega_i^2 = 1$. As $\sin\varphi$ is small and $\cos\varphi \approx 1$, it is clear that to make $\|r_m\|$ small, $p_m(\mu_1) = \prod_{k=1}^{m}(1-\mu_1/\xi_k^{(m)})$, the product of m factors, has to be small. This condition is satisfied if and only if the first factor $(1-\mu_1/\xi_1^{(m)})$ is small, because the product of the second through the m-th factor is slightly bigger than 1. In fact, as $\mu_1 < 0$ and $\xi_k^{(m)} > 0$ $(k = 2, \ldots, n)$,

$$1 < \prod_{k=2}^{m}(1-\mu_1/\xi_k^{(m)}) \approx 1-\sum_{k=2}^{m}(\mu_1/\xi_k^{(m)}) < 1+(m-1)|\mu_1|/\mu_2 = 1+O(\sin^2\varphi). \tag{3.5}$$

[1] How small is small enough depends on k; for bigger k, this threshold tends to be smaller

Here we use the first order approximation of the product based on the facts that $\mu_1/\mu_2 = O(\sin^2\varphi) \ll 1$ and, from the interlacing property, that $\xi_2^{(m)}$ approximates μ_2 from above as m increases.

To get the new bound in (3.2), we need to estimate $\|r_m\|^2 - |p_m(\mu_1)|^2\cos^2\varphi$ and $|1-p_m(\mu_1)|$ in (3.1). Since $\{v_i\}$ are orthonormal, we know from (3.4) that

$$\|r_m\|^2 - |p_m(\mu_1)|^2\cos^2\varphi = \sin^2\varphi\sum_{i=2}^{n}(c_i/\sin\varphi)^2 p_m(\mu_i)^2 \le \sin^2\varphi \max_{2\le i\le n} p_m(\mu_i)^2, \tag{3.6}$$

where the inequality comes from the relation $\sum_{i=2}^{n}(c_i/\sin\varphi)^2 = 1$. Using (3.5), we have

$$|1-p_m(\mu_1)| = |1-\prod_{k=1}^{m}(1-\mu_1/\xi_k^{(m)})| \approx |1-(1-\mu_1/\xi_1^{(m)})| = |\mu_1/\xi_1^{(m)}|. \tag{3.7}$$

The new bound (3.2) is easily established from the above two estimates and (3.1). □

Remark. Note that $\max_{2\le i\le n}|p_m(\mu_i)|$ in (3.2) might not have significant effect on the behavior of $\angle(y_m, v_1)$ when m is not too large. Intuitively, if B has a wide spectrum (which is often the case if it is unpreconditioned), $\max_{2\le i\le n}|p_m(\mu_i)|$ does not decrease considerably for small and moderate m since many eigenvalues μ_i cannot indeed be approximated by any harmonic Ritz value $\xi_k^{(m)}$; it becomes small only when m is large enough so that each eigenvalue μ_i is well-approximated by some harmonic Ritz value. Therefore, $\angle(y_m, v_1)$ decreases with m mainly due to the fact that $\xi_1^{(m)}$ approximates $\mu_1 < 0$ from below ($|\xi_1^{(m)}|$ decreases to $|\mu_1|$). That is, the convergence of MINRES for $By = x$ and the decrease of $\angle(y_m, v_1)$ both depend on how quickly $1-\mu_1/\xi_1^{(m)}$ decreases to 0.

To explore this point, we use the relation between Ritz values and the reciprocals of harmonic Ritz values. It is shown in [14] that for a Lanczos decomposition in (2.1), the reciprocals of the harmonic Ritz values of B are Ritz values of B^{-1} from an orthonormal basis of range(BU_m). Hence the convergence of $\xi^{(m)}$ to μ_1 depends on the convergence of the extreme Ritz value $1/\xi_1^{(m)}$ of B^{-1} to the corresponding eigenvalue $1/\mu_1$, which in turn depends on the convergence of angles between the Krylov subspace range(BU_m) and the eigenvector v_1 of B^{-1} associated with $1/\mu_1$. Since the columns of BU_m form a basis of $B\mathcal{K}_m(B,x)$, when $\angle(v_1, B\mathcal{K}_m(B,x))$ is small, the eigenvalue $1/\mu_1$ of B^{-1} can be well-approximated by the extreme Ritz value of B^{-1}, namely $1/\xi_1^{(m)}$, obtained from an orthonormal basis of $B\mathcal{K}_m(B,x) = \mathcal{K}_m(B, Bx)$.

The following two lemmas from Chapter 4 of [18] show the quality of the approximation from $B\mathcal{K}_m(B,x)$ to v_1, and hence lead to our main theorem, which describes how quickly $\xi_1^{(m)}$ approximates μ_1.

LEMMA 3.3. *Suppose B is symmetric and has an orthonormal eigenpairs (μ_i, v_i), with its eigenvalues ordered so that $\mu_1 < \mu_2 \le \cdots \le \mu_n$. Then $\tan\angle(v_1, \mathcal{K}_k(B,u)) \le \frac{\tan\angle(v_1,u)}{c_{k-1}(1+2\eta)}$, where $\eta = \frac{\mu_1-\mu_n}{\mu_n-\mu_2} < -1$, and $c_k(1+2\eta) = (1+2\sqrt{\eta+\eta^2})^k + (1+2\sqrt{\eta+\eta^2})^{-k}$ is the k-th order Chebyshev polynomial of the first kind.*

LEMMA 3.4. *Let (λ, v) be an eigenpair of a symmetric matrix C. Suppose U_φ is a set of orthonormal column vectors for which $\varphi = \angle(v, \mathrm{range}(U_\varphi))$ is small. Then the Rayleigh quotient $H_\varphi = U_\varphi^T C U_\varphi$ has an eigenvalue λ_φ such that $|\lambda - \lambda_\varphi| \le \|E_\varphi\|$,where $\|E_\varphi\| \le \frac{\sin\varphi}{\sqrt{1-\sin^2\varphi}}\|C\| = \tan\varphi\|C\|$.*

Let $u = Bx$ in Lemma 3.3 and $C = B^{-1}$ in Lemma 3.4. Recalling that μ_1 is the eigenvalue of B closest to zero so that $\|B^{-1}\| = 1/|\mu_1|$, we immediately have the following main theorem.

THEOREM 3.5. *Suppose unpreconditioned* MINRES *is used to solve $By = x$ in Rayleigh Quotient Iteration where $B = A - \sigma I$ and $x = v_1\cos\varphi + u\sin\varphi$. Let $\xi_1^{(m)}$ be the leftmost (also the unique negative) harmonic Ritz value. Then*

$$\frac{1}{\xi_1^{(m)}} - \frac{1}{\mu_1} \le \frac{1}{|\mu_1|}\frac{\tan\angle(v_1,Bx)}{c_{k-1}(1+2\eta)}, \quad \text{or equivalently,}\ 1 - \frac{\mu_1}{\xi_1^{(m)}} \le \frac{\tan\angle(v_1,Bx)}{c_{k-1}(1+2\eta)}. \tag{3.8}$$

This theorem shows that unpreconditioned MINRES converges quickly for the linear system in RQI if $\tan\angle(v_1, Bx)$ is small and/or $c_{k-1}(1+2\eta)$ increases with k rapidly. In fact, the numerator provides insight into the slow start of MINRES iteration. Note that $Bx = (A-\sigma I)\sum_{i=1}^{n} c_i v_i = (\lambda_1-\sigma)\cos\varphi v_1 + \sum_{i=2}^{i=n}(\lambda_i - \sigma)c_i v_i$, and hence

$$\tan\angle(v_1, Bx) = \frac{\|[(\lambda_2-\sigma)c_2, ..., (\lambda_n-\sigma)c_n]\|}{|(\lambda_1-\sigma)\cos\varphi|} = O\left(\frac{1}{\sin\varphi\cos\varphi}\right). \tag{3.9}$$

Therefore, for fixed η, as the outer iteration proceeds and x becomes closer to v_1 (φ becomes smaller), (3.8) and (3.9) indicate that more MINRES iterations are needed to make $\xi_1^{(m)}$ close to μ_1 and $1-\mu_1/\xi_1^{(m)}$ considerably smaller than 1. Hence, it takes longer to see an obvious reduction of the dominant component v_1 in r_m so that $\|r_m\|$ is reduced considerably.

To see how rapidly the denominator $c_{k-1}(1+2\eta)$ increases with k, one can see from Lemma 3.3 that the Chebyshev polynomial behaves like $(1+2\sqrt{\eta+\eta^2})^{k-1}$ asymptotically. Hence we define $(1+2\sqrt{\eta+\eta^2})^{-1}$ as the *asymptotic convergence factor* (between 0 and 1). Note that as $\eta < -1$, bigger $|\eta|$ corresponds to smaller asymptotic convergence factor, which implies faster convergence of $\xi_1^{(m)}$ to μ_1, and hence indicates that MINRES converges more quickly. One caveat mentioned in Chapter 4 in [18] is that the bound of angles in Lemma 3.3 might be unfavorable when the algebraically smallest eigenvalues of B are clustered together so that $|\eta|$ could be very close to 1, whereas the actual convergence of the angles might be much faster. Nonetheless, bigger $|\eta|$ is still a reliable predictor of faster convergence of the harmonic Ritz values. In fact, η is closely related to the reduced condition number $\kappa = \mu_n/\mu_2$ of the coefficient matrix since $|\eta| = |\frac{\mu_1-\mu_n}{\mu_n-\mu_2}| = 1 + \frac{\mu_2-\mu_1}{\mu_n-\mu_2}$, and

$$1+\frac{1}{\kappa-1} = 1+\frac{\mu_2}{\mu_n-\mu_2} < |\eta| < 1+\frac{2\mu_2}{\mu_n-\mu_2} = 1+\frac{2}{\kappa-1}. \tag{3.10}$$

Hence smaller κ corresponds to bigger $|\eta|$ and smaller asymptotic convergence factor, and is helpful to make $1-\mu_1/\xi_m^{(1)}$ decrease to 0 more rapidly. This agrees with the result in [7] that smaller κ tends to make MINRES converge more quickly.

We end this subsection with a comment on the assumption in Theorem 3.2 that $p_m(\mu_1) < 1$, which might not always be true for small m. However, this has minimal impact on our convergence analysis. See [19] (Appendix B) for details.

3.2. Preconditioned MINRES with no tuning. It is observed in [16] that solving $(A-\sigma I)y = x$ by MINRES with a symmetric positive definite preconditioner is considerably slower than one might expect based on performance of such preconditioners in the usual setting of linear system solution.

More specifically, let $Q \approx A$ be some symmetric positive definite preconditioner of A, for example, an incomplete Cholesky factorization. We then need to solve

$$\hat{B}\hat{y} \equiv L^{-1}(A-\sigma I)L^{-T}\hat{y} = L^{-1}x, \tag{3.11}$$

where $\hat{y} = L^T y$ and $LL^T = Q$. Let $\hat{\mu}_1 < 0$ be the eigenvalue of $\hat{B}$ closest to zero and $\hat{v}_1$ be the corresponding eigenvector. It follows from (3.3) that a necessary condition of MINRES convergence for the preconditioned system is that for any nonnegligible eigenvector component in the right hand side, the corresponding eigenvalue must be well-approximated by some harmonic Ritz value. Though the right hand side $L^{-1}x$ is not close to $\hat{v}_1$, it usually still has a large component of $\hat{v}_1$. Therefore, it is possible to eliminate the component of $\hat{v}_1$ in $\hat{r}_m$ (hence making $\|\hat{r}_m\|$ small enough) only if the leftmost harmonic Ritz value $\hat{\xi}_1^{(m)}$ approximates $\hat{\mu}_1 < 0$ well enough. However, the following theorem suggests that the number of MINRES steps required for this good approximation to appear tends to increase as the outer iteration proceeds with $\hat{B}$ becoming more singular.

THEOREM 3.6. *Consider the preconditioned system $\hat{B}\hat{y} \equiv L^{-1}(A-\sigma I)L^{-T}\hat{y} = L^{-1}x$ arising in RQI. Let the eigenvalues of $\hat{B}$ be ordered as $\hat{\mu}_1 < \hat{\mu}_2 \le \ldots \le \hat{\mu}_n$, and let the m-step Lanczos decomposition be $\hat{B}\hat{U}_m = \hat{U}_m\hat{T}_m + \hat{\beta}_{j+1}\hat{u}_{j+1}e_j^T$. Then a necessary condition for $\hat{T}_m$ to be indefinite is satisfied if*

$$m \ge \frac{\log(\sqrt{\hat{\mu}_n/|\hat{\mu}_1|}\tan\angle(\hat{v}_1, L^{-1}x))}{\log(1+2\sqrt{\hat{\eta}+\hat{\eta}^2})} + 1, \quad \text{where } \hat{\eta} = \frac{\hat{\mu}_1-\hat{\mu}_n}{\hat{\mu}_n-\hat{\mu}_2}. \tag{3.12}$$

Proof. Recall that the eigenvalues of $B = A-\sigma I$ satisfy $\mu_1 < 0 < \mu_2$, and by the Sylvester inertia law for $\hat{B} = L^{-1}BL^{-T}$, we have $\hat{\mu}_1 < 0 < \hat{\mu}_2$. Using the eigendecompositions $\hat{B} = \hat{V}\mathrm{diag}(\hat{\mu}_1, \ldots, \hat{\mu}_n)\hat{V}^T$ and $\hat{T}_m = \hat{S}_m\hat{\Theta}_m\hat{S}_m^T = \hat{U}_m^T\hat{B}\hat{U}_m$, [14] shows that

$$\hat{\Theta}_m = (\hat{U}_m\hat{S}_m)^T\hat{B}(\hat{U}_m\hat{S}_m) = \hat{W}_m^T\mathrm{diag}(\hat{\mu}_1, \ldots, \hat{\mu}_n)\hat{W}_m \tag{3.13}$$

where $\hat{W}_m = \hat{V}^T\hat{U}_m\hat{S}_m$ satisfies $\hat{W}_m^T\hat{W}_m = I$. The Ritz value $\hat{\theta}$ is a weighted average of the eigenvalues $\hat{\mu}_i$.

To see the condition for $\hat{T}_m$ being indefinite, we need to explore if $\hat{v}_1$ can be well-represented in $\hat{W}_m$ so that $\hat{\mu}_1 < 0$ can be well-approximated by $\hat{\theta}_1^{(m)}$. Consider any, say, the i-th, column of $\hat{U}_m\hat{S}_m$: $t_i =$

$\hat{U}_m\hat{S}_m(:,i) = \hat{v}_1\cos\psi + \hat{u}\sin\psi \in \text{range}(\hat{U}_m)$, where $\psi \geq \angle(\hat{v}_1, \text{range}(\hat{U}_m))$ (recall that $\angle(\hat{v}_1, \text{range}(\hat{U}_m))$ is the smallest angle between $\hat{v}_1$ and any vector in $\text{range}(\hat{U}_m)$), $\hat{u} \in \text{span}\{\hat{v}_2, ..., \hat{v}_n\}$ and $\|\hat{u}\| = 1$. Then

$$\begin{aligned}(3.14)\quad \hat{\theta}_i^{(m)} &= (\hat{V}^T t_i)^T \text{diag}(\hat{\mu}_1, ..., \hat{\mu}_n)(\hat{V}^T t_i) = (\cos\psi e_1 + \sin\psi e_1^{\perp})^T \text{diag}(\hat{\mu}_1, ..., \hat{\mu}_n)(\cos\psi e_1 + \sin\psi e_1^{\perp}) \\ &= \hat{\mu}_1\cos^2\psi + \hat{\mu}^*\sin^2\psi,\end{aligned}$$

where $e_1 = [1, 0, ..., 0]^T$, $\|e_1^{\perp}\| = 1$, and $\hat{\mu}^* = (e_1^{\perp})^T \text{diag}(\hat{\mu}_1, ..., \hat{\mu}_n)(e_1^{\perp}) \in [\hat{\mu}_2, \hat{\mu}_n]$. Hence all Ritz values are positive if and only if $\tan^2\psi > |\hat{\mu}_1|/\hat{\mu}^*$. It follows that, since $\psi \geq \angle(\hat{v}_1, \text{range}(\hat{U}_m))$, $\hat{T}_m$ is positive definite if $\tan^2\angle(\hat{v}_1, \text{range}(\hat{U}_m)) > |\hat{\mu}_1|/\hat{\mu}^*$.

Therefore a necessary condition to make $\hat{T}_m$ indefinite (hence $\theta_1^{(m)} < 0$) is $\tan^2\angle(\hat{v}_1, \text{range}(\hat{U}_m)) < |\hat{\mu}_1|/\hat{\mu}^*$. By Lemma 3.2, since

$$(3.15)\qquad \tan\angle(\hat{v}_1, \text{range}(\hat{U}_m)) < \frac{\tan\angle(\hat{v}_1, L^{-1}x)}{c_{m-1}(1+2\hat{\eta})} < \frac{\tan\angle(\hat{v}_1, L^{-1}x)}{(1+2\sqrt{\hat{\eta}+\hat{\eta}^2})^{m-1}},$$

the necessary condition holds if the last term in the above inequality is smaller than $\sqrt{|\hat{\mu}_1|/\hat{\mu}_n}$. The conclusion follows by taking the logarithm of both sides. □

Remark. This theorem simply suggests that during the initial steps of preconditioned MINRES, the leftmost harmonic Ritz value $\hat{\xi}_1^{(m)}$ will not approximate the negative eigenvalue $\hat{\mu}_1$ of $\hat{B}$, and therefore $\|\hat{r}_m\|$ will not be greatly reduced. In fact, as $\hat{T}_m$ is positive definite for small m, it follows that $\hat{\xi}_1^{(m)} > \hat{\mu}_2 > 0$, by the property of harmonic Ritz values. Therefore (3.4) implies that the component $\hat{v}_1$ in $\hat{r}_m$ is indeed magnified, since all factors of $\prod_{k=1}^{m}(1 - \hat{\mu}_1/\hat{\xi}_k^{(m)})$ are bigger than 1. It is hence impossible for MINRES to perform well during these iterations.

In addition, the number of the "bad" MINRES steps tends to grow as the outer iterate becomes closer to the true eigenvector. In fact, it is shown in [1] (Theorem 9.1) that $\hat{\mu}_1 = O(\sin^2\varphi)$. Since in general $\angle(\hat{v}_1, L^{-1}x) = O(1)$, the bound of m given in the above theorem is like $\log|\frac{C}{\sin\varphi}|/\log(1+2\sqrt{\hat{\eta}+\hat{\eta}^2})$, which increases as the outer iteration proceeds. This estimate of the number of bad MINRES steps clearly shows a major limitation of preconditioned MINRES without tuning when it is used in the setting of RQI. This insight is supported by numerical experiments in section 5.

3.3. Preconditioned MINRES with tuning. One way suggested in [16] to address the fact that preconditioning does not do as well as expected in this setting is to replace the preconditioned system $L^{-1}(A-\sigma I)L^{-T}\hat{y} = L^{-1}x$ by $L^{-1}(A-\sigma I)L^{-T}\hat{y} = L^T x$. This idea comes from the fact that the aim is not to accurately solve the original preconditioned system, but to make the eigenvalue residual associated with MINRES iterate decrease more quickly. The authors show that the modified right hand side $L^T x$ is close to the eigenvector of the system matrix corresponding to the negative eigenvalue and MINRES convergence can be considerably improved. One needs to notice that the recovered MINRES iterate y_m in this case converges to $(A-\sigma I)^{-1}LL^T x$ instead of $(A-\sigma I)^{-1}x$. Though $(A-\sigma I)^{-1}LL^T x$ is not as good as $(A-\sigma I)^{-1}x$ to approximate v_1, it is in practice still better than x. This strategy works because y_m approximates $(A-\sigma I)^{-1}LL^T x$ so fast that for small and moderate m, it is a better approximation to v_1 than its counterpart obtained from the standard use of preconditioned MINRES for $(A-\sigma I)^{-1}x$, though the latter would win when m is large enough.

However, this method is not RQI, and the cubic convergence of the outer iteration is lost. An alternative approach introduced in [7], known as "tuning", entails a rank-1 modification of the Cholesky factor L of the symmetric positive definite preconditioner $Q = LL^T$ so that the tuned preconditioner $\mathbb{Q} = \mathbb{L}\mathbb{L}^T$ satisfies $\mathbb{Q}x = Ax$ (see Section 4 for the construction of $\mathbb{L}$). The preconditioned system with tuning is

$$(3.16)\qquad \mathbb{L}^{-1}(A-\sigma I)\mathbb{L}^{-T}\tilde{y} = \mathbb{L}^{-1}x,$$

leaving the RQI structure unchanged. It is shown in [7] that $\sin\tilde{\varphi} \equiv \sin\angle(\tilde{v}_1, \mathbb{L}^{-1}x) = O(\sin\varphi)$, so that $\tilde{\theta}_1^{(m)}$ and $\tilde{\xi}_1^{(m)}$ are negative at the very beginning of the MINRES iterations, as in the unpreconditioned case. Compared to preconditioned MINRES with no tuning, the overhead of performing "bad" MINRES iterations in which $\hat{\xi}_1^{(m)} > 0$ is avoided with the tuned preconditioner, and convergence of MINRES is hence much faster. Moreover, the cubic convergence of the outer iteration is preserved, since the linear system is not changed by tuning.

The convergence analysis of unpreconditioned MINRES directly applies to (3.16). The convergence basically depends on how quickly $\tilde{\xi}_1^{(m)}$ approaches $\tilde{\mu}_1$ from below. We have the following bound:

$$1 - \frac{\tilde{\mu}_1}{\tilde{\xi}_m^{(1)}} \le \frac{\tan\angle(\tilde{v}_1, \tilde{B}\mathbb{L}^{-1}x)}{c_{k-1}(1+2\tilde{\eta})}, \quad \text{where } \tilde{\eta} = \frac{\tilde{\mu}_1 - \tilde{\mu}_n}{\tilde{\mu}_n - \tilde{\mu}_2} < -1. \tag{3.17}$$

We can see that preconditioned MINRES with tuning converges much more quickly than unpreconditioned MINRES because the asymptotic convergence factor of the former is considerably smaller than that of the latter. See section 5 for comparisons of the two quantities. Note that η of the unpreconditioned MINRES is a constant depending only on the eigenvalues of A, whereas $\hat{\eta}$ and $\tilde{\eta}$ may change as the outer iteration proceeds; in our experience, these changes in the preconditioned eigenvalues tend to be small.

Similar to unpreconditioned MINRES in this context, preconditioned MINRES with tuning also has a slow start if the outer iterate x is close to v_1. In Appendix A we show that the relative linear residual $\|\tilde{r}_m\|/\|\mathbb{L}^{-1}x\| = 1 - O(\sin^2\tilde{\varphi})$ holds in the same way as for the unpreconditioned MINRES solve. However, since the asymptotic convergence factor of preconditioned MINRES with tuning is smaller, the slow start is less pronounced than that of unpreconditioned MINRES.

Remark. As stated before, we are interested in how quickly $\angle(y_m, v_1)$ decreases as the MINRES iteration proceeds, where $y_m = L^{-T}\hat{y}_m$ (or $\mathbb{L}^{-T}\tilde{y}_m$) for the preconditioned MINRES without (or with) tuning. By analyzing the convergence of MINRES for the preconditioned systems (3.11) and (3.16), we also have some idea about how rapidly the recovered linear residual $\|r_m\| = \|L\hat{r}_m\| \le \|L\|\|\hat{r}_m\|$ (or $\|\mathbb{L}\|\|\tilde{r}_m\|$) of the original linear system decreases with m. In light of Theorem 3.2, we can hence roughly estimate how quickly $\angle(y_m, v_1)$ decreases in the three versions of MINRES iteration simply by comparing the convergence speed of each MINRES solve.

4. Preconditioner with tuning based on a rank-2 modification. The symmetric preconditioner with tuning defined in [7] is based on a rank-1 modification of the Cholesky factor of the ordinary symmetric positive definite preconditioner $Q = LL^T$. We restate the lemma from [7] to construct the tuned factor.

LEMMA 4.1. *Suppose $Q = LL^T \approx A$ is a symmetric positive definite preconditioner of A. Let x be an approximation of v_1 and define $w = Ax - Qx$. The tuned Cholesky factor $\mathbb{L}$ is defined as $\mathbb{L} = L + \alpha w(L^{-1}w)^T$, where α is the real solution of $(L^{-1}w)^T(L^{-1}w)\alpha^2 + 2\alpha - \frac{1}{w^Tx} = 0$.*

The tuned preconditioner $\mathbb{Q} = \mathbb{L}\mathbb{L}^T$ can also be defined as a symmetric rank-1 modification of Q:

$$\begin{aligned} \mathbb{Q} = \mathbb{L}\mathbb{L}^T &= (L + \alpha w(L^{-1}w)^T)(L + \alpha w(L^{-1}w)^T)^T = LL^T + 2\alpha ww^T + ((L^{-1}w)^T(L^{-1}w))\alpha^2 ww^T \\ &= Q + \frac{ww^T}{w^Tx} = Q + \frac{(Ax-Qx)(Ax-Qx)^T}{(Ax-Qx)^Tx}, \end{aligned} \tag{4.1}$$

such that $\mathbb{Q}x = Ax$. This definition has the advantage enabling $\mathbb{Q}$ to be defined for preconditioners not specified by Cholesky factors.

The tuned preconditioner $\mathbb{Q}$ is appropriate for MINRES only if it is positive definite. It is shown in [7] that two conditions must be satisfied to guarantee positive definiteness, namely

$$(Ax - Qx)^Tx \ne 0, \quad \textit{and} \quad 1 + \frac{(Ax-Qx)^TQ^{-1}(Ax-Qx)}{(Ax-Qx)^Tx} \ge 0. \tag{4.2}$$

In practice, it is possible that $(Ax - Qx)^Tx$ is 0 or small enough to cause numerical problems. Moreover, it is shown in [7] that in cases where $(Ax - Qx)^Tx < 0$, the second condition above is satisfied only if $\|A - Q\|$ is very small. The latter requirement is difficult to enforce except in cases where the Cholesky factor is very dense; for example, Q is the incomplete Cholesky preconditioner with very small drop tolerance.

Positive definiteness of a tuned preconditioner can be enforced with less stringent constraints by using a rank-2 modification of Q. This approach is used to construct approximate Hessians for quasi-Newton methods in optimization ([12], Ch 11). In particular, we can use the BFGS modification

$$\mathbb{Q} = Q - \frac{(Qx)(Qx)^T}{(Qx)^Tx} + \frac{(Ax)(Ax)^T}{(Ax)^Tx}. \tag{4.3}$$

It is easy to see that $\mathbb{Q}x = Ax$. Lemma 11.5 in [12] shows that if the denominator of the last term in (4.3) is positive (which is the case here), $\mathbb{Q}$ is positive definite.

A tuned preconditioner based on the rank-2 modification is slightly more expensive to apply than the that based on the rank-1 modification. One should try the rank-1 modification and turn to the rank-2 version only when the former is not positive definite, i.e., when there is no real solution to the quadratic equation in Lemma 4.1.

5. Numerical Experiments. We compare unpreconditioned MINRES, preconditioned MINRES with no tuning, and preconditioned MINRES with tuning for solving the linear system in RQI, in numerical experiments on two benchmark eigenvalue problems from MatrixMarket [10]. The first problem $nos5.mtx$ is a real symmetric positive definite matrix of order 468 coming from finite element approximation to a biharmonic operator that describes beam bending in a building. The second consists of two matrices $K = bcsst\mathbf{k}08.mtx$ and $M = bcsst\mathbf{m}08.mtx$ of order 1074 that define a generalized symmetric positive definite eigenvalue problem $Kx = \lambda Mx$ used for dynamic modeling of a structure. This generalized problem can be easily transformed to the standard problem $M^{-1/2}KM^{-1/2}(M^{1/2}x) = \lambda(M^{1/2}x)$ where the coefficient matrix can be formed directly because M is a positive definite diagonal matrix.

5.1. Stopping criteria for inner iterations. Efficiency of each solver is evaluated by the MINRES iteration counts needed in a given outer iteration to satisfy some stopping criterion. Note that in MINRES iteration, we can easily monitor the SYMMLQ iterate also because it can be obtained for free [4]. We define $eigres_m^{MR}$ and $eigres_m^{SL}$ to be the eigenvalue residual associated with the MINRES iterate y_m^{MR} and the SYMMLQ iterate y_m^{SL} respectively, and we stop the MINRES iteration when the relative changes of $\|y_m^{MR}\|$, $eigres_m^{MR}$ and $eigres_m^{SL}$ are all small enough. In other words, the stopping criterion is

$$\text{stop}(\|y_m\|) \ \& \ \text{stop}(eigres_m^{MR}) \ \& \ \text{stop}(eigres_m^{SL}), \tag{5.1}$$

where

$$\text{stop}(\|y_m\|) \equiv \frac{|\ \|y_{m-k}\| - \|y_{m-k-1}\|\ |}{\|y_{m-k}\|} < \epsilon_{inner}, \quad k = 0, 1, \tag{5.2}$$

and stop($eigres_m^{MR}$) and stop($eigres_m^{MR}$) are defined similarly. The first two criteria are roughly necessary conditions for MINRES to start to converge. The criterion stop($eigres_m^{SL}$) helps prevent an early stop, since, as observed in [3], $eigres_m^{SL}$ tends to be oscillatory until $\angle(y_m, v_1)$ approximates $\angle(y_{exact}, v_1)$ well. We require the stopping criteria to be satisfied for two successive steps to further ensure that MINRES does not stop prematurely. We take $\epsilon_{inner} = 0.01$ for all the criteria in the tests. These combined criteria basically guarantee that MINRES is stopped as soon as $\angle(y_m, v_1) \approx \angle(y_{exact}, v_1)$.

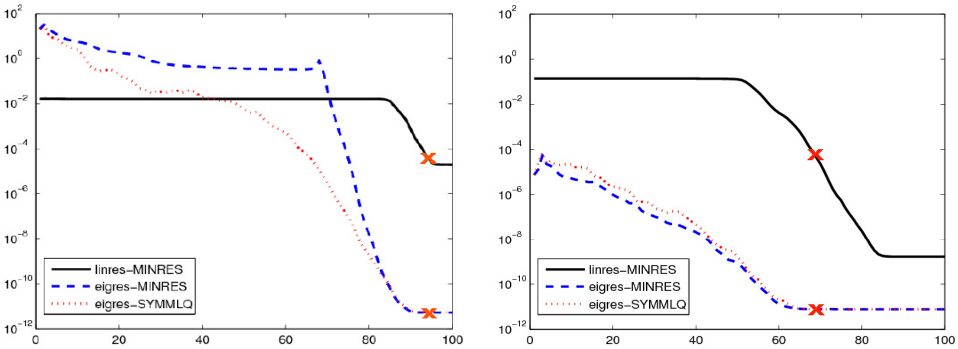

FIG. 5.1. MINRES *linear residual,* MINRES *and* SYMMLQ *eigenvalue residual in the third outer iteration on Problem 1 (drop tol* 0.2 ... *Ri*

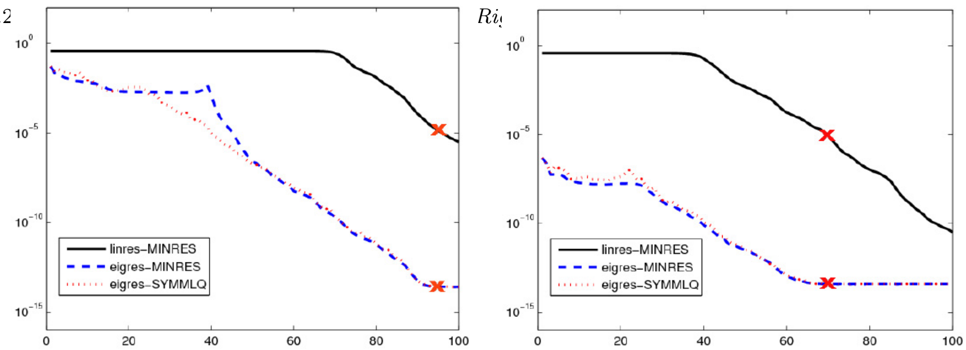

FIG. 5.2. MINRES *linear residual,* MINRES *and* SYMMLQ *eigenvalue residual in the third outer iteration on Problem 2 (drop tol* 0.25*). Left: preconditioned solve without tuning. Right: preconditioned solve with rank-2 tuning.*

5.2. Results and comments. We use the incomplete Cholesky preconditioner from Matlab 7.4 with drop tolerance 0.25. In each test the starting vector $x^{(0)}$ is chosen to be close enough to the target eigenvector v_1 so that the Rayleigh quotient $\sigma^{(0)}$ satisfies $|\lambda_1 - \sigma^{(0)}| < |\lambda_2 - \sigma^{(0)}|$. The results for MINRES in the third outer iteration of RQI are shown in Figures 5.1–5.2 and Tables 5.1–5.2.

Tables 5.1–5.2 show clearly that unpreconditioned MINRES converges slowly; as shown in Section 3.2, this is because $\tan(v_1, Bx) = O(\frac{1}{\sin\varphi\cos\varphi})$, and the asymptotic convergence factor is very close to 1 (i.e., the reduced condition number is big); see (3.9) and (3.10). In fact, unpreconditioned MINRES fails to satisfy the stopping criteria in the specified maximum number of steps.

It is obvious from Figure 5.1–5.2 that preconditioned MINRES with tuning significantly outperforms the version without tuning. The cross marks on the curves indicate the MINRES iteration at which the stopping criteria are satisfied. It takes more steps for preconditioned MINRES without tuning to satisfy the stopping criteria than the version with tuning. The eigenvalue residual curve (dashed lines) of the tuned MINRES iterate is well below that of the untuned one, and the norm of the residual of the linear system (solid lines) also decreases more quickly due to tuning. Moreover, 1) the eigenvalue residual curve decreases slowly in the first dozens of steps of MINRES without tuning, and 2) the eigenvalue residual curve of preconditioned MINRES without tuning starts at a value much larger than the value at which the curve of the version with tuning starts.

Both the phenomena 1) and 2) can be explained by the fact that tuning forces the preconditioning operator to behave like A on the current outer iterate x. The reason for phenomenon 1) is given in Section 3.2: in the initial steps of MINRES without tuning, the negative eigenvalue of the preconditioned coefficient matrix cannot be approximated by any harmonic Ritz value because $\hat{T}_m$ is positive definite, and hence MINRES cannot perform well. Moreover, Table 5.3 shows that the number of these “bad” MINRES steps increases as the outer iteration proceeds, as Theorem 3.6 suggests. To explain phenomenon 2), suppose $\hat{y}_0 = 0$ for the preconditioned MINRES without tuning. It follows that $\hat{y}_1 \in \hat{y}_0 + \mathcal{K}_1(\hat{B}, \hat{b})$ is a multiple of the preconditioned right hand side $\hat{b} = L^{-1}x$, and the recovered iterate $y_1 = L^{-T}\hat{y}_1$ is a multiple of $L^{-T}L^{-1}x = Q^{-1}x$, which is in general far from a good approximation of v_1. Similarly for the preconditioned MINRES with tuning, y_1 is a multiple of $\mathbb{Q}^{-1}x$. Since $\mathbb{Q}$ and A behave in the same way on $x \approx v_1$, it is reasonable to expect that $\mathbb{Q}^{-1}x \approx A^{-1}x \approx \lambda_1^{-1}v_1$ is a better approximation to v_1 than $\mathbb{Q}^{-1}x$.

	Non	No tuning	Tuning
MINRES iter	160*	94	68
neg Ritz shows in	2	64	1
aymptotic cvg. factor	0.9901	0.9189	0.9189
reduced cond. number	8.6172e+3	5.1497e+2	5.1509e+2
initial angle	3.6915e–3	3.6942e–1	3.9601e–5

Table 5.1
Comparison of three MINRES *methods in the third outer iteration on Problem 1*

	Non	No tuning	Tuning
MINRES iter	200*	95	69
neg Ritz shows in	2	31	1
aymptotic cvg. factor	0.9984	0.9347	0.9347
reduced cond. number	1.5154e+6	8.2201e+2	8.2201e+2
initial angle	1.5345e–4	2.3665e–3	1.1692e–6

Table 5.2
Comparison of three MINRES *methods in the third outer iteration on Problem 2*

Tables 5.1–5.2 provide data supporting the above comparison. First, note that there is little difference in the asymptotic convergence factor and the reduced condition number between the preconditioned MINRES without and with tuning. The difference comes from the last rows in the two tables: the angle between the preconditioned right hand side and the eigenvector of the preconditioned coefficient matrix corresponding to the unique negative eigenvalue is much bigger in the case without tuning than it is in the case with tuning. As explained, it is this very fact that makes the first MINRES iterate with tuning ($\mathbb{Q}^{-1}x$) a much better approximation to v_1 than that without tuning ($Q^{-1}x$). Moreover, for the untuned preconditioner, $\hat{T}_m$ is positive definite in the first 63 steps in Problem 1 and in the first 30 steps in Problem 2. One can see from Figures 5.1–5.2 that the eigenvalue residual curves start to decrease quickly soon after $\hat{T}_m$ becomes indefinite. Tables 5.4 shows some cases when the rank-2 tuning has to be used. In problems 2, the

Outer Iteration	1	2	3	4
Problem 1	7	19	64	
Problem 2	1	1	31	44

TABLE 5.3
Preconditioned MINRES *steps without tuning needed to have* $\theta_1^{(m)} < 0$

Drop tolerance	0.05	0.07	0.1	0.25	0.3	0.35
No Tuning	51	75	82	95	111	139
Rank-1 Tuning	35	51	60	–	–	–
Rank-2 Tuning	36	52	59	69	77	97

TABLE 5.4
Preconditioned MINRES *steps needed to satisfy the stopping criterion in the third outer iteration for Problem 2*

rank-1 tuning makes the tuned preconditioner indefinite when the drop tolerance is above some threshold, and rank-2 tuning works with any drop tolerance. In the two test problems, there is little performance difference between preconditioned MINRES with the rank-1 and the rank-2 tuning. As the drop tolerance increases, the iteration counts of preconditioned MINRES with and without tuning both increase, but the difference between them becomes more pronounced.

6. Conclusion. We have presented a detailed convergence analysis of three versions of MINRES to solve the linear systems in Rayleigh Quotient Iteration to find the lowest eigenpair of a symmetric positive definite matrix. Based on insight about the behavior of Ritz and harmonic Ritz values, our analysis includes understanding of slow start of MINRES iterations, the main weakness of ordinary preconditioning without tuning in this setting, and the virtue of tuning. Using the idea of the BFGS formula in quasi-Newton methods, we propose a tuning method based on a rank-2 modification which guarantees positive definiteness of the symmetric tuned preconditioner. Other rank-2 modification formulas, such as DFP in quasi-Newton methods, could also be used.

REFERENCES

[1] J. Berns-Müller and A. Spence, *Inexact inverse iteration and GMRES*, submitted to Electronic Transactions on Numerical Analysis.

[2] J. Berns-Müller, I. G. Graham, and A. Spence, *Inexact inverse iteration for symmetric matrices*, Linear Algebra and its Applications, Vol. 416, Issues 2–3, (2006), pp. 389-413.

[3] F. A. Dul, *MINRES and MINERR are better than SYMMLQ in eigenpair computations*, SIAM Journal on Scientific Computing, Vol. 19, No. 6, (1998), pp. 1767–1782.

[4] H. C. Elman, D. J. Silvester and A. J. Wathen, *Finite Elements and Fast Iterative Solvers*, Oxford University Press, New York, 2005.

[5] M. Freitag and A. Spence, *Convergence theory for inexact inverse iteration applied to the generalised nonsymmetric eigenproblem*, Electronic Transactions on Numerical Analysis, v. 28, (2007), pp. 40–64.

[6] M. Freitag and A. Spence, *Convergence rates for inexact inverse iteration with application to preconditioned iterative solves*, BIT Numerical Mathematics, v. 47, n.1 (2007), pp. 27–44.

[7] M. Freitag and A. Spence, *A tuned preconditioner for inexact inverse iteration applied to Hermitian eigenvalue problems*, to appear in IMA Journal on Numerical Analysis (2007).

[8] G. Golub and Q. Ye, *Inexact inverse iteration for generalized eigenvalue problems*, BIT Numerical Mathematics, v. 40, n.4 (2000), pp. 671–684.

[9] Y. Lai, K. Lin, and W. Lin, *An inexact inverse iteration for large sparse eigenvalue problems*, Numerical Linear Algebra with Applications, Vol.4, No.5, (1997), pp. 425–437.

[10] *Matrix Market*, http://math.nist.gov/MatrixMarket/

[11] R. B. Morgan, *Computing interior eigenvalues of large matrices, Linear Algebra Appl.*, 154–156 (1991), pp. 289–309.

[12] S. G. Nash and A. Sofer, *Linear and nonlinear programming*, McGraw-Hill, 1996.

[13] Y. Notay, *Convergence analysis of inexact Rayleigh Quotient Iteration*, SIAM Journal on Matrix Analysis and Applications, Vol. 24, No.3, (2003), pp. 627–644.

[14] C. Paige, B. Parlett and H. van der Vorst, *Approximate solutions and eigenvalue bounds from Krylov subspaces*, Numerical Linear Algebra with Applications, Vol 2(2), (1995), pp. 115–134..

[15] B. Parlett, *The symmetric eigenvalue problem*, SIAM, Philadelphia, 1998.

[16] V. Simoncini and L. Eldén, *Inexact Rayleigh quotient-type methods for eigenvalue computations*, BIT Numerical Mathematics, v. 42, n.1 (2002), pp.159–182.

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

[18] G. W. Stewart, *Matrix Algorithms Volumn II: Eigensystems*, SIAM, Philadelphia, 1998.

[19] F. Xue and H. Elman, *Convergence analysis of iterative solvers in inexact Rayleigh Quotient Iteration*, submitted to SIAM Journal on Matrix Analysis and Applications.