SIAM J. Matrix Anal. Appl.
Vol. 31, No. 3, pp. 877–899

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 initial stagnation 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." A comparison of MINRES with SYMMLQ in this context is also given. 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

AMS subject classifications. 65F18, 65F15, 65F10

DOI. 10.1137/080712908

1. Introduction.

There has been considerable interest in recent years in developing and analyzing eigensolvers with inner-outer structure for computing eigenvalues of matrices closest to some specified value. These algorithms usually involve at each step (outer iteration) a shift-invert matrix-vector product implemented by solving the shifted linear system iteratively (inner iteration). The use of inner iteration becomes mandatory if the matrices are too large for factorization-based exact shift-invert matrix-vector products to be practical. 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 [9] and [11], 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 [21] and [16] 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 [19] 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 nonsymmetric eigenvalue problems and in [7] for symmetric problems. A preconditioner with tuning is a low rank modification of an ordinary preconditioner.

*Received by the editors January 10, 2008; accepted for publication (in revised form) by Y. Saad May 5, 2009; published electronically August 5, 2009. 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.
http://www.siam.org/journals/simax/31-3/71290.html

†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).

Tuning forces the preconditioning operator to behave in the same way as the system matrix on the current approximate eigenvector.

To understand the modified approaches for preconditioning mentioned above, one needs to note that when ordinary preconditioned MINRES is used to solve the linear system arising in RQI, the preconditioned right-hand side is generally far from a good approximate eigenvector of the preconditioned shifted coefficient matrix, and the convergence theory of MINRES indicates that the counts of inner iteration steps needed to reach a prescribed relative tolerance will grow considerably as the outer iteration proceeds. The ideas in [19] 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 approaches for preconditioning. Note that when MINRES is applied to the linear system in RQI, as the inner iteration proceeds, (1) the residual of the linear system decreases (convergence of inner iterations), and (2) the angle between the MINRES iterate and the true eigenvector we are computing decreases gradually to that between the true solution of the linear system and the eigenvector. Results in [19] show that during the course of an RQI (outer) iteration, the initial stagnation of MINRES iteration may be accompanied by considerable improvement of the eigenvector approximation by the MINRES iterate. By analyzing MINRES behavior in depth, we know *how quickly* the angle between the MINRES iterate and the target eigenvector decreases as the MINRES iteration proceeds. This perspective has not been emphasized in the literature, and it is adopted in the paper as the main criterion to compare the performance of different versions of MINRES in this setting.

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. We analyze the initial stagnation of MINRES in this context, as remarked in [19], 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. By probing into the quality of approximations to the true eigenvector by a sequence of Krylov subspaces, 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. We then introduce a tuning strategy based on a rank-2 modification which guarantees positive definiteness of the tuned preconditioner.

The paper is organized as follows. Section 2 reviews some preliminary facts for later discussions. Section 3 gives detailed convergence analysis of the inner iteration for the three versions of MINRES and provides some comments on the different performance of MINRES and SYMMLQ in this setting. A rank-2 modification tuning is introduced in section 4 as an improvement of the rank-1 modification tuning of [7]. Numerical experiments supporting the analysis are given in section 5. We summarize the paper in section 6.

2. Preliminaries.

We want to compute the lowest eigenpair of a symmetric positive definite matrix by RQI. Consider the eigenvalue problem

$$Av = \lambda v, \tag{2.1}$$

where A is symmetric positive definite with eigenvalues $0 < \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n$. Let $V = [v_1, v_2, \ldots, v_n] = [v_1, V_2]$ be the matrix of orthonormal eigenvectors, and let $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ so that $V^T A V = \Lambda$. Algorithm 1 describes a typical version of inexact RQI to find a simple eigenpair.

ALGORITHM 1: Inexact RQI.

Given $x^{(0)}$ with $\|x^{(0)}\| = 1$
For $i = 0, 1, \ldots,$ until convergence
1. Compute the Rayleigh quotient $\sigma^{(i)} = x^{(i)T} A x^{(i)}$
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

$$x = \sum_{k=1}^{n} v_k c_k = v_1 \cos\varphi + u \sin\varphi, \tag{2.2}$$

where u is a unit vector orthogonal to v_1, φ is the angle between x and 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

$$\sigma = x^T A x = c^T \Lambda c = \lambda_1 + \sum_{k=2}^{n} (\lambda_k - \lambda_1) c_k^2 = \lambda_1 + \left(\bar\lambda - \lambda_1\right) \sin^2\varphi, \tag{2.3}$$

where $\bar\lambda = \sum_{i=2}^{n} (\frac{c_i^2}{\sin^2\varphi}) \lambda_i \in [\lambda_2, \lambda_n]$ is a weighted average of $\lambda_2, \ldots, \lambda_n$ uniquely determined by u. 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 $(A - \sigma I)^{-1}$, and the cubic convergence of RQI (see [18, p. 76]) is easily established.

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

$$B U_m = U_m T_m + \beta_{m+1} u_{m+1} e_m^T = U_{m+1} \overline{T}_m, \tag{2.4}$$

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, which depends on the approximation from the Krylov subspace range(U_m) to the associated eigenvector of B as m increases.

We will use a major theorem from [17], 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. For convenience, we use Matlab notation $w(1)$ to denote the first entry of the vector w.

THEOREM 2.1. *Suppose* MINRES *is applied to solve the system $By = b$. At the mth* MINRES *iteration step with the corresponding Lanczos decomposition in* (2.4)*,* *the* MINRES *iterate is*

$$y_m = U_m M_m^{-2} T_m e_1 \beta_1, \tag{2.5}$$

where $M_m^2 = \overline{T}_m^T \overline{T}_m$, $\beta_1 = \|b\|$. *The residual of the linear system is*

$$r_m = b - By_m = U_{m+1} w w(1) \beta_1, \qquad \|r_m\| = |w(1)|\beta_1, \tag{2.6}$$

where $\|w\| = 1$, $w^T \overline{T}_m = 0^T$,

$$|w(1)| = \beta_{m+1} |f_m(1)| / \left(1 + \beta_{m+1}^2 \|f_m\|^2\right)^{1/2}, \tag{2.7}$$

and $f_m = T_m^{-1} e_m$. *Moreover, the residual can be written as*

$$r_m = \chi_m(B) b / \chi_m(0), \tag{2.8}$$

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* $M_m^2 - \xi T_m$. *It can be shown that* $BU_m M_m^{-1}$ *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 [19] 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 $\|r_m\| = \|x - (A - \sigma I)y_m\|$ remains still close to 1 for quite large m. We call this phenomenon *initial stagnation* and describe it in the theorem below. To make the exposition smooth, we defer the proof to Appendix A.

THEOREM 3.1. *Suppose unpreconditioned* MINRES *is used to solve* $(A-\sigma I)y = x$ *in* RQI, *where* $x = v_1 \cos\varphi + u \sin\varphi$ (*see* (2.2)). *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 1. This residual norm estimate shows qualitatively that the initial stagnation 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 $\angle(y_m, v_1)$ decreases with m. Theorem 4.1 of [19] establishes the fact that although the MINRES iteration appears to stagnate in its initial steps,

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

$\angle(y_m, v_1)$ may decrease considerably during these iterations. We restate the theorem 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 $0 < |\mu_1| < |\mu_2| \leq \cdots \leq |\mu_n|$. 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*

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

or approximately,

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

Proof. The result (3.1) is established in [19]. 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 [17], 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 $p_m(B) = \chi_m(B)/\chi_m(0)$. That is,

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

Therefore, the residual vector can be represented as

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

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 mth factor is slightly bigger than 1. In fact, as $\mu_1 < 0$ and $\xi_k^{(m)} > 0$ $(k = 2, \ldots, n)$,

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

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

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

where the inequality comes from the relation $\sum_{i=2}^{n}(c_i/\sin\varphi)^2 = 1$. The estimation of $|1-p_m(\mu_1)|$ can be simplified using (3.5):

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

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

Remark 2. The above theorem shows that, as also observed in [19], improvements of the approximate eigenvector can be obtained during the period of initial stagnation of MINRES. In fact, since $p_m(\mu_1)(\cos\varphi)v_1$ is the dominant term in r_m (see (3.4)), MINRES is almost stagnant when $p_m(\mu_1)$ stays close to 1 during the initial steps. However, in such a scenario, $|1-p_m(\mu_1)| \approx |\mu_1/\xi_1^{(m)}|$ may be increasing from a minuscule number (say, 10^{-10}) to a moderately small number (say, 10^{-3} or 10^{-2}). As this quantity appears in the denominator in (3.1), $\tan\angle(y_m, v_1)$ may decrease significantly even though the MINRES residual remains close to 1.

Remark 3. 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 because $\xi_1^{(m)}$ approximates $\mu_1 < 0$ from below ($|\xi_1^{(m)}|$ decreases to $|\mu_1|$). The behavior of MINRES for $By = x$ and the decrease of $\angle(y_m, v_1)$ both depend on how quickly $\xi_1^{(m)}$ approaches μ_1.

To explore this point, we need to use the relation between Ritz values and the reciprocals of harmonic Ritz values. It is shown in section 5 of [17] that for a Lanczos decomposition in (2.4), 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 [22] show the quality of the approximation from $B\mathcal{K}_m(B,x)$ to v_1 and lead to our main theorem, which describes how quickly $\xi_1^{(m)}$ approximates μ_1 as MINRES iteration proceeds.

LEMMA 3.3. *Suppose B is symmetric and has an orthonormal system of eigenpairs (μ_i, v_i), with its eigenvalues ordered so that $\mu_1 < \mu_2 \le \cdots \le \mu_n$. Then*

$$\tan\angle(v_1, \mathcal{K}_k(B, u)) \le \frac{\tan\angle(v_1, u)}{c_{k-1}(1+2\eta)}, \quad \text{where } \eta = \frac{\mu_1 - \mu_n}{\mu_n - \mu_2} < -1. \tag{3.8}$$

Here $c_k(1+2\eta) = (1+2\sqrt{\eta+\eta^2})^k + (1+2\sqrt{\eta+\eta^2})^{-k}$ is the kth order Chebyshev polynomial of the first kind for $|1+2\eta| > 1$.

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 $\varphi = \angle(v, \text{range}(U_\varphi))$ 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 RQI where $B = A - \sigma I$ and x is defined in (2.2). Let $\xi_1^{(m)}$ be the leftmost (also the unique negative) harmonic Ritz value. Then*

$$\frac{1}{\xi_1^{(m)}} - \frac{1}{\mu_1} \le \frac{1}{|\mu_1|}\frac{\tan\angle(v_1, Bx)}{c_{m-1}(1+2\eta)}, \quad i.e., \quad 1 - \frac{\mu_1}{\xi_1^{(m)}} \le \frac{\tan\angle(v_1, Bx)}{c_{m-1}(1+2\eta)}. \tag{3.9}$$

Suppose m_0 (depending on φ and η) is the smallest integer for which the second upper bound in (3.9) is smaller than 1. Note that as $\xi_1^{(m)} < \mu_1 < 0$ and $1 - \mu_1/\xi_1^{(m)} < 1$ for all m; (3.9) holds trivially for $m < m_0$ because the upper bound is not smaller than 1. Therefore, this theorem describes how quickly $\xi_1^{(m)}$ approaches μ_1 from below ($|\xi_1^{(m)}|$ decreases to $|\mu_1|$) for $m \ge m_0$. It provides insight into MINRES convergence and also sheds some light on the behavior of $\tan\angle(y_m, v_1)$ ($m \ge m_0$) described by (3.2) in Theorem 3.2. These points are elaborated on as follows.

We first analyze the numerator of the upper bound to explore MINRES convergence. 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

$$\begin{aligned} \tan\angle(v_1, Bx) &= \frac{\|[(\lambda_2 - \sigma)c_2, \ldots, (\lambda_n - \sigma)c_n]\|}{|(\lambda_1 - \sigma)\cos\varphi|} \\ &\le \frac{(\lambda_n - \sigma)\sin\varphi}{O(\sin^2\varphi\cos\varphi)} = O\left(\frac{1}{\sin\varphi\cos\varphi}\right). \end{aligned} \tag{3.10}$$

Therefore, for a fixed η, as the outer iteration proceeds and x becomes closer to v_1 (φ becomes smaller), (3.4), (3.9), and (3.10) indicate that m_0 becomes bigger, and more MINRES iterations are needed to make $\xi_1^{(m)}$ close to μ_1 and $1 - \mu_1/\xi_1^{(m)}$ obviously smaller than 1. Hence, MINRES suffers a longer initial stagnation period, as it takes more iterations to significantly reduce the dominant component v_1 in r_m.

To see how rapidly $\tan\angle(y_m, v_1)$ and $\|r_m\|$ decrease for $m \ge m_0$, note that the denominator of the upper bound behaves like $(1+2\sqrt{\eta+\eta^2})^{m-1}$ asymptotically (Lemma 3.3). Hence we define $(1+2\sqrt{\eta+\eta^2})^{-1}$ as the *asymptotic convergence factor* (smaller than 1). Given φ, as bigger $|\eta|$ corresponds to smaller asymptotic convergence factor and smaller m_0, we expect faster convergence of $\xi_1^{(m)}$ to μ_1 for $m \ge m_0$. This indicates that $\tan \angle(y_m, v_1)$ decreases with m ($m \geq m_0$) more quickly and in addition, MINRES will converge more quickly after its initial stagnation period.

Though (3.9) holds trivially for $m < m_0$, one may speculate that $\tan \angle(y_m, v_1)$ still decreases at a rate controlled by η in the initial MINRES steps. This speculation is corroborated to some extent by the following arguments. Reference [16] analyzes the case where the conjugate gradient (CG) method is used to perform the the system solve required by the Jacobi–Davidson method and shows that the convergence of CG for the correction equation simply depends on the effective condition number of $(I - xx^T)(A - \sigma I)(I - xx^T)$, which is essentially the reduced condition number of $A - \sigma I$. On the other hand, [19] shows that when solving $(A - \sigma I)y = x$, Jacobi–Davidson with CG delivers the same inner iterate (up to a constant) as SYMMLQ. This result is extended in [8] for the *preconditioned* solve of non-Hermitian systems when tuning is used for a full orthogonalization method. It can be shown readily that preconditioned SYMMLQ with tuning is equivalent to Jacobi–Davidson with preconditioned CG. Our numerical experiments in section 5 show that when tuning is used (to make the preconditioned solve behave qualitatively like the unpreconditioned solve), the eigenvalue residual curves of the MINRES and SYMMLQ iterates usually go hand in hand. Thus it is reasonable to conclude that $\tan \angle(y_m, v_1)$ decreases at a rate depending only on η. As the numerical experiments show, though φ is quite small in the last outer iteration, the eigenvalue residual of the inner iterate still decreases at a reasonable rate in the initial MINRES steps with tuning.

One caveat mentioned in Chapter 4 in [22] is that the bound of angles in (3.8) might be far from sharp 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 $\xi_1^{(m)}$. 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

$$1 + \frac{1}{\kappa - 1} = 1 + \frac{\mu_2}{\mu_n - \mu_2} < |\eta| < 1 + \frac{2\mu_2}{\mu_n - \mu_2} = 1 + \frac{2}{\kappa - 1}. \tag{3.11}$$

Hence smaller κ corresponds to bigger $|\eta|$ and a smaller asymptotic convergence factor, and it is helpful to make $1 - \mu_1/\xi_1^{(m)}$ 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. Appendix B gives some details on this.

3.2. Preconditioned MINRES with no tuning. It is observed in [19] and [6] 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 a 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

$$\hat{B}\hat{y} \equiv L^{-1}(A - \sigma I)L^{-T}\hat{y} = L^{-1}x, \tag{3.12}$$

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 nearly 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* $\hat{\mu}_1 < \hat{\mu}_2 \le \cdots \le \hat{\mu}_n$, *and let the m-step Lanczos decomposition be* $\hat{B}\hat{U}_m = \hat{U}_m\hat{T}_m + \hat{\beta}_{m+1}\hat{u}_{m+1}e_m^T$. *Then a necessary condition for* $\hat{T}_m$ *to be indefinite is satisfied if*

$$m \ge \frac{\log\left(\sqrt{\hat{\mu}_n/|\hat{\mu}_1|}\tan\angle\left(\hat{v}_1, L^{-1}x\right)\right)}{\log\left(1+2\sqrt{\hat{\eta}+\hat{\eta}^2}\right)} + 1, \quad \text{where } \hat{\eta} = \frac{\hat{\mu}_1 - \hat{\mu}_n}{\hat{\mu}_n - \hat{\mu}_2}. \tag{3.13}$$

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 $\hat{T}_m = \hat{S}_m\hat{\Theta}_m\hat{S}_m^T = \hat{U}_m^T\hat{B}\hat{U}_m$, [17] shows that

$$\hat{\Theta}_m = \left(\hat{U}_m\hat{S}_m\right)^T \hat{B}\left(\hat{U}_m\hat{S}_m\right) = \hat{W}_m^T\mathrm{diag}(\hat{\mu}_1, \ldots, \hat{\mu}_n)\hat{W}_m, \tag{3.14}$$

where $\hat{W}_m = \hat{V}^T\hat{U}_m\hat{S}_m$ has orthonormal columns. In other words, the Ritz value $\hat{\theta}$ is a weighted average of the eigenvalues $\hat{\mu}_i$ (see section 5 of [17]).

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 ith, column of $\hat{U}_m\hat{S}_m$: $t_i = \hat{U}_m\hat{S}_me_i = \hat{v}_1\cos\psi + \hat{u}\sin\psi \in \mathrm{range}(\hat{U}_m)$, where $\psi \ge \angle(\hat{v}_1, \mathrm{range}(\hat{U}_m))$ (recall that $\angle(\hat{v}_1, \mathrm{range}(\hat{U}_m))$ is the smallest angle between $\hat{v}_1$ and any vector in $\mathrm{range}(\hat{U}_m)$), $\hat{u} \in \mathrm{span}\{\hat{v}_2, \ldots, \hat{v}_n\}$, and $\|\hat{u}\| = 1$. Then

$$\begin{aligned}\hat{\theta}_i^{(m)} &= \left(\hat{V}^Tt_i\right)^T \mathrm{diag}(\hat{\mu}_1, \ldots, \hat{\mu}_n)\left(\hat{V}^Tt_i\right) \qquad (1 \le i \le m)\\ &= \left(\cos\psi e_1 + \sin\psi e_1^{\perp}\right)^T \mathrm{diag}(\hat{\mu}_1, \ldots, \hat{\mu}_n)\left(\cos\psi e_1 + \sin\psi e_1^{\perp}\right)\\ &= \hat{\mu}_1\cos^2\psi + \hat{\mu}^*\sin^2\psi,\end{aligned} \tag{3.15}$$

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

Therefore, a necessary condition to make $\hat{T}_m$ indefinite (hence $\theta_1^{(m)} < 0$) is that $\tan^2\angle(\hat{v}_1, \mathrm{range}(\hat{U}_m)) < |\hat{\mu}_1|/\hat{\mu}^*$. By Lemma 3.3, since

$$\tan\le\left(\hat{v}_1, \mathrm{range}\left(\hat{U}_m\right)\right) < \frac{\tan\angle\left(\hat{v}_1, L^{-1}x\right)}{c_{m-1}\left(1+2\hat{\eta}\right)} < \frac{\tan\angle\left(\hat{v}_1, L^{-1}x\right)}{\left(1+2\sqrt{\hat{\eta}+\hat{\eta}^2}\right)^{m-1}}, \tag{3.16}$$

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 4. 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.

Note that (3.2) in Theorem 3.2 cannot be used here to describe $\tan\angle(y_m, v_1)$, where $y_m = L^{-T}\hat{y}_m$ is the recovered iterate from preconditioned MINRES iterate. This is because the right-hand side of (3.12) is in general far from an approximation of $\hat{v}_1$, and there is no obvious relation between the eigenpair of $\hat{B}$ and that of B. Our numerical experiments in section 5 also suggest that no significant improvement of eigenvector approximation can be obtained during the initial MINRES iterations. In the next subsection, we show how tuning solves this difficulty and makes Theorem 3.2 applicable to the preconditioned system.

In addition, the number of "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 $\hat{\mu}_1 = (\lambda_1-\sigma)/\|Lv_1\|^2+O((\lambda_1-\sigma)^2) = O(\sin^2\varphi)$. 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 [19] 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. References [10] and [14] also advocate the use of $L^T x$ as the starting vector of preconditioned Lanczos algorithm to compute a few eigenpairs of symmetric matrices. 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 iteration, 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 $\mathbb{Q}x = Ax$ (the construction of $\mathbb{L}$ is discussed in section 4 below). The preconditioned system with tuning is thus

$$\tilde{B}\tilde{y} \equiv \mathbb{L}^{-1}(A-\sigma I)\mathbb{L}^{-T}\tilde{y} = \mathbb{L}^{-1}x, \tag{3.17}$$

leaving the RQI structure unchanged. Therefore, the cubic convergence of the outer iteration is preserved.

Suppose $\tilde{v}_1$ is the eigenvector of $\tilde{B}$ corresponding to the eigenvalue $\tilde{\mu}_1 < 0$. There is a straightforward relation between $(\tilde{\mu}_1, \tilde{v}_1)$ and the eigenpair (μ_1, v_1) of $B = A - \sigma I$. Note that as $\mathbb{Q}x = Ax$, $\mathbb{Q}v_1 \approx Av_1$ as the RQI outer iterate $x \to v_1$. It follows that

$$(A - \sigma I)v_1 = \mu_1 v_1 = \left(\frac{\mu_1}{\lambda_1}\right)\lambda_1 v_1 = \left(\frac{\mu_1}{\lambda_1}\right)Av_1 \approx \left(\frac{\mu_1}{\lambda_1}\right)\mathbb{Q}v_1 \tag{3.18}$$

and hence

$$\mathbb{L}^{-1}(A - \sigma I)\mathbb{L}^{-T}\left(\mathbb{L}^T v_1\right) \approx \mathbb{L}^{-1}\left(\frac{\mu_1}{\lambda_1}\right)\mathbb{Q}v_1 = \left(\frac{\mu_1}{\lambda_1}\right)\left(\mathbb{L}^T v_1\right). \tag{3.19}$$

The relation $(\tilde{\mu}_1, \tilde{v}_1) \approx (\mu_1/\lambda_1, \mathbb{L}^T v_1)$ is thus established (see Lemma 3.1 in [7]). Similarly, the right-hand side of (3.17) is

$$\begin{aligned} \mathbb{L}^{-1}x \approx \mathbb{L}^{-1}v_1 &= \mathbb{L}^{-1}\left(\frac{1}{\lambda_1}\right)\lambda_1 v_1 = \left(\frac{1}{\lambda_1}\right)\mathbb{L}^{-1}Av_1 \\ &\approx \left(\frac{1}{\lambda_1}\right)\mathbb{L}^{-1}\mathbb{Q}v_1 = \left(\frac{1}{\lambda_1}\right)\mathbb{L}^T v_1 \approx \left(\frac{1}{\lambda_1}\right)\mathbb{L}^T x. \end{aligned} \tag{3.20}$$

In other words, the right-hand side of the preconditioned system with tuning automatically approximates the starting vector $\mathbb{L}^T x$ proposed by [10] and [14] in the preconditioned Lanczos method and by [19] in preconditioned MINRES used in the context of RQI. In addition, it is an approximate eigenvector of the system matrix corresponding to $\tilde{\mu}_1 < 0$. Recall that this is the case for the unpreconditioned system $(A - \sigma I)y = x$. In fact, it is shown in [7] that $\sin\tilde{\varphi} \equiv \sin\angle(\tilde{v}_1, \mathbb{L}^{-1}x) = O(\sin\varphi)$. Therefore, the analysis of unpreconditioned MINRES directly applies to (3.17). The Ritz value $\tilde{\theta}_1^{(m)}$ and harmonic Ritz value $\tilde{\xi}_1^{(m)}$ 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. As a result, MINRES begins to converge earlier and more important, $\tan\angle(y_m, v_1)$ (where $y_m = \mathbb{L}^{-T}\tilde{y}_m$) decreases much more rapidly in the initial steps. See the figures in section 5.

Similar to the unpreconditioned case, the convergence of preconditioned MINRES with tuning and the decrease of $\tan\angle(y_m, v_1)$ basically depend on how quickly $\tilde{\xi}_1^{(m)}$ approaches $\tilde{\mu}_1$ from below. We have the following bound just like (3.9):

$$1 - \frac{\tilde{\mu}_1}{\tilde{\xi}_m^{(1)}} \le \frac{\tan\angle\left(\tilde{v}_1, \tilde{B}\mathbb{L}^{-1}x\right)}{c_{k-1}(1 + 2\tilde{\eta})}, \quad \text{where } \tilde{\eta} = \frac{\tilde{\mu}_1 - \tilde{\mu}_n}{\tilde{\mu}_n - \tilde{\mu}_2} < -1. \tag{3.21}$$

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, by definition, η of the unpreconditioned MINRES is a constant that depends 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.

Preconditioned MINRES with tuning also has an initial stagnation period 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. The initial stagnation is less pronounced for the preconditioned case with tuning because its asymptotic convergence factor is smaller.

3.4. Comparison of SYMMLQ and MINRES used in RQI. To solve the linear systems arising in RQI, a natural alternative to MINRES is SYMMLQ. With extensive numerical tests, Dul in [3] claimed that MINRES improves eigenvector approximation to some prescribed level in considerably fewer iterations than SYMMLQ. Rigorous analysis and comparison of the two methods is not seen in the literature. Here we provide some comments on the two solvers in this context.

Our experience is that MINRES is better than SYMMLQ in general, but the advantage may vary considerably depending on the preconditioned problem. In one of our sample problems with appropriate tuned preconditioner, there is little difference between the two methods, but for ill-conditioned problems without a preconditioner, as shown in [3], SYMMLQ might not even be able to improve the eigenvalue residual in a reasonable number of iterations.

To compare the MINRES iterate y_m^{MR} and SYMMLQ iterate y_m^{SL}, we see that the MINRES linear residual for $By = x$ is $x - By_m^{MR} = p_m^{MR}(B)x$ (by the definition of the MINRES residual polynomial p_m), so that

$$y_m^{MR} = B^{-1}\left(I - p_m^{MR}(B)\right)x = \left(I - p_m^{MR}(B)\right)\left(B^{-1}x\right) = \left(I - p_m^{MR}(B)\right)\sum_{i=1}^{n} b_i v_i$$

$$= \sum_{i=1}^{n}\left(1 - p_m^{MR}(\mu_i)\right)b_i v_i = \sum_{i=1}^{n}\left(1 - \prod_{j=1}^{m}\left(1 - \mu_i/\xi_j^{(m)}\right)\right)b_i v_i, \tag{3.22}$$

where $B^{-1}x = \sum_{i=1}^{n} b_i v_i$ is the true solution. Similarly, for SYMMLQ, we have

$$y_m^{SL} = \sum_{i=1}^{n}\left(1 - p_m^{SL}(\mu_i)\right)b_i v_i = \sum_{i=1}^{n}\left(1 - \prod_{j=1}^{m}\left(1 - \mu_i/\theta_j^{(m)}\right)\right)b_i v_i. \tag{3.23}$$

The above expressions show clearly that the difference between the MINRES and SYMMLQ iterates as approximations to v_1 simply results from the different quality of approximation to the extreme eigenvalue μ_1 and the interior eigenvalues μ_i by harmonic Ritz and Ritz values. Since $\angle(y_m, v_1)$ largely depends on the ratio of the magnitudes of eigenvectors corresponding to interior eigenvalues to that of v_1 contained in y_m, we speculate that the reason for $\angle(y_m^{MR}, v_1) < \angle(y_m^{SL}, v_1)$ is that harmonic Ritz values tend to be better approximations to the interior eigenvalues, though μ_1 is better approximated by the Ritz value $\theta_1^{(m)}$ (see [13], [17], and [20]).

Reference [3] also shows that the curve of eigenvalue residuals of MINRES iterates is generally smooth, whereas that of SYMMLQ iterates tends to be oscillatory. This phenomenon can be explained qualitatively by the fact the interior eigenvalues are susceptible to being impersonated by nonconverged Ritz values. That is, an interior eigenvalue μ_k can be well approximated by some Ritz value at the mth step of the Lanczos process when the angle between the eigenvector v_k and the current Krylov subspace range(U_m) is not small [22]. At the mth SYMMLQ step, a small number of interior eigenvalues μ_k might be impersonated by some "incorrect" Ritz value $\theta_{j(k)}^{(m)}$ (the subscript $j(k)$ is a function of k; $1 < j(k) < m$) so that $1 - \mu_k/\theta_{j(k)}^{(m)}$ is fairly small and hence $1 - \prod_{j=1}^{m}(1 - \mu_k/\theta_j^{(m)})$ decreases dramatically. But in the next SYMMLQ step the impersonation may disappear, and this quantity recovers its magnitude in the step before impersonation. This causes $\angle(y_m^{SL}, v_1)$ to fluctuate considerably. MINRES does not have this problem, however: a harmonic Ritz value would not well approximate an eigenvalue unless the corresponding eigenvector is well represented in range(U_m) (see page 293 of [22] and equation (4.19) of [22]). As a result, $1 - \prod_{j=1}^{m}(1 - \mu_k/\xi_j^{(m)})$ will not fluctuate greatly as m increases, and the decreasing curve of eigenvalue residuals is smoother. We use this observation in section 5 to develop stopping criteria for the inner iterations.

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 L of the ordinary symmetric positive definite preconditioner $Q = LL^T$. We restate Lemma 3.2 from [7] to construct the tuned Cholesky 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 $(L^{-1}w)^T(L^{-1}w)\alpha^2 + 2\alpha - \frac{1}{w^T x} = 0$.*

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

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

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 $\mathbb{Q}$ has to be positive definite for MINRES. It is shown in [7] that two conditions must be satisfied to guarantee positive definiteness, namely,

$$
(Ax - Qx)^T x \neq 0 \quad \textit{and} \quad 1 + \frac{(Ax - Qx)^T Q^{-1}(Ax - Qx)}{(Ax - Qx)^T x} \geq 0. \tag{4.2}
$$

However, it is possible that $(Ax - Qx)^T x$ is 0 or small enough to cause numerical problems. Moreover, it is shown in [7] that in cases where $(Ax - Qx)^T x < 0$, 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 can be 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 (see [15, Chapter 11]). In particular, we can use the BFGS modification

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

It is easy to see that $\mathbb{Q}x = Ax$. Lemma 11.5 in [15] 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 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 equation in Lemma 4.1.

5. Numerical experiments. We compare unpreconditioned MINRES, preconditioned MINRES without tuning, and preconditioned MINRES with tuning for solving the linear system in RQI in numerical experiments on three benchmark eigenvalue problems from Matrix Market [12]. 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. The last one is a generalized symmetric positive semidefinite problem of order 2003 from fluid flows defined by symmetric positive definite $K = bcsst\mathbf{k}13.mtx$ and symmetric positive semidefinite $M = bcsst\mathbf{m}13.mtx$ with rank 1241. The first two examples show the differences among the three versions of MINRES. The third problem suggests that tuning might be used for more complex eigenvalue problems.

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

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

where

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

and stop($eigres_m^{MR}$) and stop($eigres_m^{MR}$) are defined similarly.

We elaborate on this strategy as follows: Our aim is to stop MINRES as soon as $\angle(y_m, v_1) \approx \angle(y_{exact}, v_1)$ (the cubic convergence of the outer iteration is thus preserved). The first criterion is adopted by [19], where it is shown to be roughly equivalent to the condition stop($|1 - p_m(\mu_1)|$). This is a necessary condition for $p_m(\mu_1) \ll 1$ (say, $p_m(\mu_1)$ is of order 10^{-3} to 10^{-2}), which in turn implies that MINRES has started to converge. Our experience is that $\angle(y_m, v_1) \approx \angle(y_{exact}, v_1)$ usually holds when MINRES has started to converge. The second criterion is directly connected to the eigenvalue problem: since the right-hand side is dominated by v_1, we expect $\angle(y_m, v_1) \approx \angle(y_{exact}, v_1)$ once the eigenvalue residual stops decreasing. However, with just these two criteria, MINRES might stop prematurely due to a possibly slow approximation process. The criterion stop($eigres_m^{SL}$) helps prevent an early stop, since $eigres_m^{SL}$ tends to be oscillatory until $\angle(y_m, v_1)$ approximates $\angle(y_{exact}, v_1)$ well (see section 3.4), whereas in our experience, $eigres_m^{MR}$ tends to stagnate slightly before this (see Figures 5.1–5.3). Finally, we require the stopping criteria to be satisfied for two successive steps to further ensure that MINRES does not stop prematurely.

One could instead choose a smaller ϵ_{inner} and stop MINRES when the criteria are satisfied for only one step, but this usually makes MINRES continue for quite a few

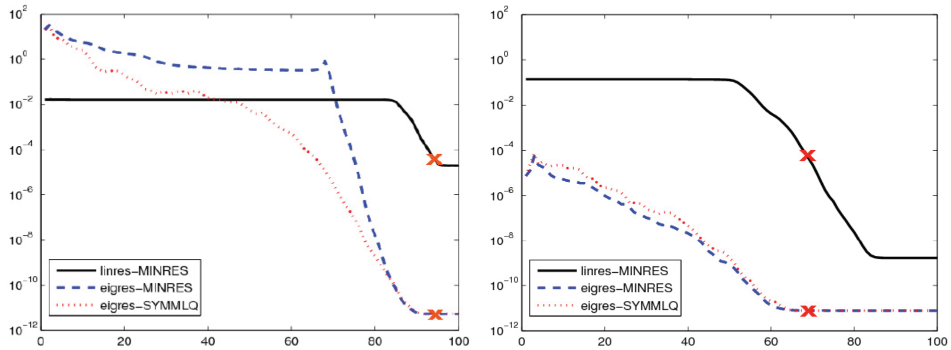

FIG. 5.1. MINRES *linear residual,* MINRES*, and* SYMMLQ *eigenvalue residual in the third outer iteration on Problem* 1*, with drop tolerance* 0.25*. Left: preconditioned solve without tuning. Right: preconditioned solve with rank*-1 *tuning.*

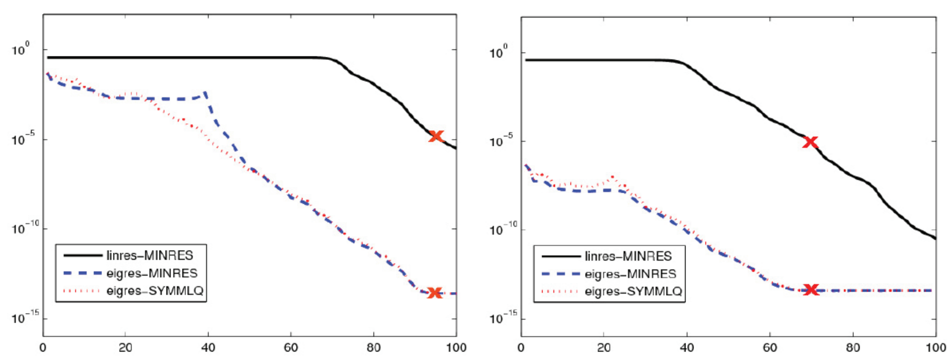

FIG. 5.2. MINRES *linear residual,* MINRES*, and* SYMMLQ *eigenvalue residual in the third outer iteration on Problem* 2*, with drop tolerance* 0.25*. Left: preconditioned solve without tuning. Right: preconditioned solve with rank*-2 *tuning.*

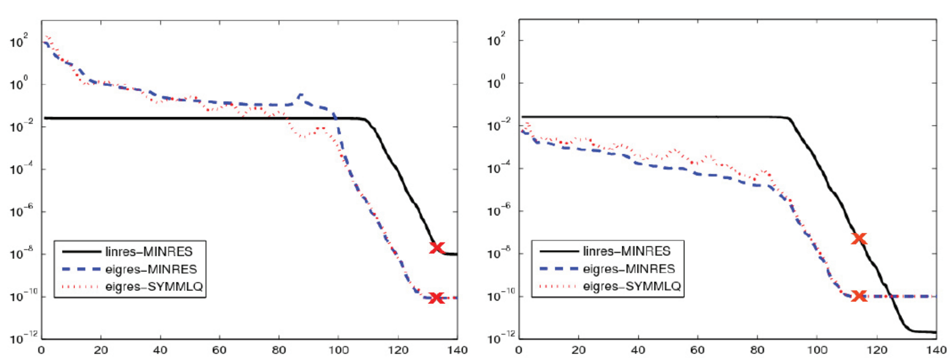

FIG. 5.3. MINRES *linear residual,* MINRES*, and* SYMMLQ *eigenvalue residual in the third outer iteration on Problem* 3*, with drop tolerance* 0.0015*. Left: preconditioned solve without tuning. Right: preconditioned solve with rank*-2 *tuning.*

steps after $\angle(y_m, v_1) \approx \angle(y_{exact}, v_1)$. We take $\epsilon_{inner} = 0.01$ for all the criteria in the tests. The combined criteria guarantee a fair comparison of preconditioned MINRES without and with tuning for solving the linear systems in RQI.

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

	Non	No tuning	Tuning
MINRES iterations	160*	94	68
Negative Ritz shows in	2	64	1
Aymptotic convergence factor	0.9901	0.9189	0.9189
Reduced condition number	8.6172e+3	5.1497e+2	5.1509e+2
Initial angle	3.6915e–3	3.6942e–1	3.9601e–5

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

	Non	No tuning	Tuning
MINRES iterations	200*	95	69
Negative Ritz shows in	2	31	1
Aymptotic convergence factor	0.9984	0.9347	0.9347
Reduced condition number	1.5154e+6	8.2201e+2	8.2201e+2
Initial angle	1.5345e–4	2.3665e–3	1.1692e–6

Note that we choose not to use the residual of the linear system $\|x-(A-\sigma I)y_m\|$ in the stopping criteria, because, as Figures 5.1–5.3 show, it is not possible to specify an extent to which the norm of the linear residual should be decreased for all problems when $\angle(y_m, v_1) \approx \angle(y_{exact}, v_1)$ holds.

5.2. Results and comments. We use the incomplete Cholesky preconditioner from Matlab 7.4 with drop tolerance 0.25 for problem 1 and 2, and 0.0015 for problem 3. 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 on these problems are shown in Figures 5.1–5.3 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 (3.11). In fact, unpreconditioned MINRES fails to satisfy the stopping criteria in the specified maximum number of steps. From now on, we compare only the preconditioned MINRES without and with tuning.

It is obvious from Figures 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

Table 5.3
Numbers of preconditioned MINRES *iteration steps without tuning needed to have* $\theta_1^{(m)} < 0$.

Outer iteration	1	2	3	4
Problem 1	7	19	64	
Problem 2	1	1	31	44
Problem 3	1	8	82	

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), first 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 $\mathbb{Q}^{-1}x \approx A^{-1}x \approx \lambda_1^{-1}v_1$, which is a much better approximation to v_1 than $\mathbb{Q}^{-1}x$.

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.

We show by the third test that tuning can also be used for generalized eigenvalue problems that cannot be converted into standard eigenvalue problems. Since $M = bcsst\mathbf{m}13.mtx$ is singular, one has to solve $(K - \sigma M)y = Mx$ in RQI. Similar to the previous standard problems, the tuned preconditioner $\mathbb{Q}$ is a rank-1 modification of the preconditioner $Q \approx K$ such that $\mathbb{Q}x = Kx$. Our convergence analysis of MINRES may not be applied directly, because the eigenvectors are now M-orthogonal and expressions of the entries of the tridiagonal matrix T_m become less clear. Moreover, the fact that Mx is not close to the "negative" eigenvector of $K - \sigma M$ makes the Ritz value analysis more complicated. However, Figure 5.3 and Table 5.3 show that the pattern observed in the previous two standard eigenvalue problems still holds for this problem.

Tables 5.4–5.5 show some cases when the rank-2 tuning has to be used. In problems 2 and 3, the 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 three 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.

Table 5.4
Number of preconditioned MINRES *iteration steps needed to satisfy the stopping criterion in the third outer iteration for problem* 2.

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 5.5
Number of preconditioned MINRES *iteration steps needed to satisfy the stopping criterion in the third outer iteration for problem* 3.

Drop tolerance	2.5e–4	5.0e–4	7.5e–4	1.0e–3	1.25e–3	1.5e–3
No tuning	76	84	103	112	122	133
Rank-1 tuning	71	73	90	–	–	–
Rank-2 tuning	65	73	89	99	107	115

6. Conclusion.

We have presented a detailed convergence analysis of three versions of MINRES to solve the linear systems in RQI 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 qualitative and quantitative understanding of initial stagnation of MINRES iterations, the main weakness of ordinary preconditioning without tuning in inexact RQI, the virtue of tuning, and the advantage of MINRES over SYMMLQ.

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.

Considering the performance of the three preconditioned MINRES solves on the last test problem, we speculate that our convergence analysis of MINRES on standard eigenvalue problems can be extended to generalized eigenvalue problems.

Appendix A. Proof of initial stagnation.

A.1. Unpreconditioned MINRES (Theorem 3.1).

Proof. Note from (2.2) that $\sin\varphi = \sqrt{c_2^2+\cdots+c_n^2}$ and $u = \sum_{k=2}^{n} \frac{c_k}{\sqrt{c_2^2+\cdots+c_n^2}} v_k$. That is, u is uniquely determined by the ordered set $\{c_k/\sqrt{c_2^2+\cdots+c_n^2}\}$; one can fix u and change only φ by increasing/decreasing $\{c_k\}(2 \le k \le n)$ by a common factor to see qualitatively how MINRES convergence is affected by φ.

One can see from (2.5) that $y_1 = 0$ since $T_1 = [0]$. We now assume that $m \ge 2$.

Recall the spectral decomposition of A, the Rayleigh quotient (2.3), and Lanczos decomposition (2.4). For any $k \le m$, we have

$$
\begin{aligned}
\text{(A.1)} \qquad x^T(A-\sigma I)^k x = c^T(\Lambda-\sigma I)^k c &= \sum_{i=2}^{n}(\lambda_i-\sigma)^k c_i^2 + (\lambda_1-\sigma)^k c_1^2 \\
&= \sin^2\varphi \sum_{i=2}^{n}\left((\lambda_1-\sigma)+(\lambda_i-\lambda_1)\right)^k \left(\frac{c_i^2}{\sin^2\varphi}\right) + (\lambda_1-\sigma)^k\cos^2\varphi \\
&= \sin^2\varphi \sum_{i=2}^{n}\left(\sum_{j=0}^{k} \mathrm{C}_k^j(\lambda_1-\sigma)^j(\lambda_i-\lambda_1)^{k-j}\right)\left(\frac{c_i^2}{\sin^2\varphi}\right) + (\lambda_1-\sigma)^k\cos^2\varphi \\
&= \sin^2\varphi \sum_{j=0}^{k} \mathrm{C}_k^j(\lambda_1-\sigma)^j\left(\sum_{i=2}^{n}(\lambda_i-\lambda_1)^{k-j}\left(\frac{c_i^2}{\sin^2\varphi}\right)\right) + (\lambda_1-\sigma)^k\cos^2\varphi.
\end{aligned}
$$

Letting $l_s = \sum_{i=2}^{n}(\lambda_i - \lambda_1)^s(\frac{c_i^2}{\sin^2\varphi}) \in [\lambda_2 - \lambda_1, \lambda_n - \lambda_1]$ (which depends only on $\{\lambda_i\}$ and u but not on φ) be a weighted average of $\{(\lambda_i - \lambda_1)^s\}(0 \le s \le k)$, and using $\lambda_1 - \sigma = (\lambda_1 - \bar{\lambda})\sin^2\varphi$, we then have

(A.2)

$$\begin{aligned} x^T(A-\sigma I)^k x &= \sum_{j=0}^{k} \mathrm{C}_k^j l_{k-j}\left(\lambda_1 - \bar{\lambda}\right)^j \sin^{2j+2}\varphi + \left(\lambda_1 - \bar{\lambda}\right)^k \sin^{2k}\varphi\cos^2\varphi \\ &= l_k \sin^2\varphi + \mathrm{C}_k^1 l_{k-1}\left(\lambda_1 - \bar{\lambda}\right)\sin^4\varphi + \cdots + \mathrm{C}_k^{k-1} l_1 \left(\lambda_1 - \bar{\lambda}\right)^{k-1}\sin^{2k}\varphi \\ &\quad + \left(\lambda_1 - \bar{\lambda}\right)^k \sin^{2k+2}\varphi + \left(\lambda_1 - \bar{\lambda}\right)^k \sin^{2k}\varphi\cos^2\varphi \\ &= \sin^2\varphi\left(l_k + \cdots + \left(\mathrm{C}_k^{k-1} l_1\left(\lambda_1 - \bar{\lambda}\right)^{k-1} + \left(\lambda_1 - \bar{\lambda}\right)^k\right)\sin^{2k-2}\varphi\right) \\ &= q_{k-1}\left(\sin^2\varphi\right)\sin^2\varphi, \end{aligned}$$

where q_{k-1} is a polynomial of degree $k-1$ whose coefficients depend on $\{\lambda_i\}$ and u; in particular, l_k, the constant term of q_{k-1}, is independent of φ.

We can thus find the first few entries of T_m in closed form. For example,

$$\alpha_1 = x^T(A-\sigma I)x = \sigma - \sigma = 0, \tag{A.3}$$

$$\begin{aligned} \beta_2 &= \|(A-\sigma I)x - \alpha_1 x\| = \sqrt{x^T(A-\sigma I)^2 x} = \sqrt{q_1\left(\sin^2\varphi\right)\sin^2\varphi} \\ &= \sqrt{q_1\left(\sin^2\varphi\right)}\sin\varphi = \sqrt{l_2}\sin\varphi + O\left(\sin^3\varphi\right), \end{aligned} \tag{A.4}$$

$$\begin{aligned} \alpha_2 &= u_2^T(A-\sigma I)u_2 = \left(\frac{(A-\sigma I)x}{\beta_2}\right)^T (A-\sigma I)\left(\frac{(A-\sigma I)x}{\beta_2}\right) \\ &= \frac{x^T(A-\sigma I)^3 x}{\beta_2^2} = \frac{q_2\left(\sin^2\varphi\right)\sin^2\varphi}{q_1\left(\sin^2\varphi\right)\sin^2\varphi} = \frac{q_2\left(\sin^2\varphi\right)}{q_1\left(\sin^2\varphi\right)}, \end{aligned} \tag{A.5}$$

and

$$\begin{aligned} \beta_3 &= \|(A-\sigma I)u_2 - \beta_2 x - \alpha_2 u_2\| = \left\|\frac{(A-\sigma I)^2 x}{\beta_2} - \beta_2 x - \frac{\alpha_2(A-\sigma I)x}{\beta_2}\right\| \\ &= \sqrt{\frac{x^T(A-\sigma I)^4 x}{\beta_2^2} + \beta_2^2 + \alpha_2^2 - 2\beta_2^2 - 2\alpha_2^2 + 2\alpha_2\alpha_1} \\ &= \sqrt{\frac{x^T(A-\sigma I)^4 x}{x^T(A-\sigma I)^2 x} - \left(\frac{x^T(A-\sigma I)^3 x}{x^T(A-\sigma I)^2 x}\right)^2 - x^T(A-\sigma I)^2 x} \\ &= \sqrt{\frac{q_3\left(\sin^2\varphi\right)}{q_1\left(\sin^2\varphi\right)} - \left(\frac{q_2\left(\sin^2\varphi\right)}{q_1\left(\sin^2\varphi\right)}\right)^2 - q_1\left(\sin^2\varphi\right)\sin^2\varphi}. \end{aligned} \tag{A.6}$$

One can similarly evaluate α_3 and β_4, though the expressions for other entries become much more complicated. Note that as the coefficients of q_{k-1} are uniquely determined by u, for any fixed u, all the entries in T_m are functions of $\sin^2\varphi$ only. (Obviously, if x consists of $m+1$ eigenvectors of A, then $(A-\sigma I)U_{m+1} = U_{m+1}T_{m+1}$ with $\beta_{m+2} = 0$, and $\|r_{m+1}\| = 0$. This is why we restrict our analysis to how $\|r_k\|$ $(1 \le k \le m)$ is affected as φ goes to zero.)

To show that $\lim_{\varphi\to 0}\|r_k\| = 1$ for all k, with $1 \le k \le m$, we need only to establish the result for $k = m$, since the MINRES residual norm decreases monotonically. In light of Theorem 2.1, the key point is to show that $f_m(1)$ is the unique dominant entry in $f_m = T_m^{-1}e_m$. In fact, the entries of f_m can be evaluated by Cramer's rule as follows:

$$\text{(A.7)}\qquad \begin{aligned}|f_m(1)| &= \frac{|\det[e_m, T_m(1:m,2:m)]|}{|\det[T_m]|} = \frac{\prod_{k=2}^m \beta_k}{\beta_2^2 \det[T_m(3:m,3:m)]}\\ &= \frac{1}{\beta_2}\frac{\prod_{k=3}^m \beta_k}{|\det[T_m(3:m,3:m)]|}\end{aligned}$$

and therefore

$$\text{(A.8)}\qquad \frac{1}{|f_m(1)|^2} = \beta_2^2\left(\frac{|\det[T_m(3:m,3:m)]|}{\prod_{k=3}^m \beta_k}\right)^2,$$

where $T_m(i:m, j:m)$ is the submatrix of T_m consisting of its ith through mth rows and jth through mth columns.

We now show that $\frac{1}{|f_m(1)|^2} = O(\sin^2\varphi)$ and hence $\lim_{\varphi\to 0}\frac{1}{|f_m(1)|^2} = 0$. Simple observation from (A.2) shows that $\lim_{\varphi\to 0} q_{k-1}(\sin^2\varphi) = l_k$ (the constant term of the polynomial) and thus the limit of α_2 and β_3 are l_3/l_2 and $\sqrt{l_4 l_2 - l_3^2}/l_2$, respectively. We can show by induction that all α_k and β_k have some *nonzero* limit (independent of φ) except for $\alpha_1 = 0$ and $\beta_2 \approx \sqrt{l_2}\sin\varphi$. Since all β_k ($3 \le k \le m$) and all entries in $|\det[T_m(3:m,3:m)]|$ have a nonzero limit as $\varphi \to 0$, the term in (A.8) multiplying β_2^2 also has a nonzero limit. Recalling from (A.4) that $\beta_2^2 = q_1(\sin^2\varphi)\sin^2\varphi$, we have $\frac{1}{|f_m(1)|^2} = O(\sin^2\varphi)$ and $\lim_{\varphi\to 0}\frac{1}{|f_m(1)|^2} = 0$.

Note that β_2 in the second column of T_m is the only nonzero entry of the first row and hence $f_m(2) = 0$. Then

$$\text{(A.9)}\qquad \begin{aligned}|f_m(3)| &= \frac{|\det[T_m(1:m,1:2), e_m, T_m(1:m,4:m)]|}{|\det[T_m]|}\\ &= \frac{\beta_2^2\prod_{k=4}^m \beta_k}{\beta_2^2|\det[T_m(3:m,3:m)]|} = \frac{\prod_{k=4}^m \beta_k}{|\det[T_m(3:m,3:m)]|}.\end{aligned}$$

Therefore, $|f_m(3)|$ has a nonzero limit as $\varphi \to 0$, and $\lim_{\varphi\to 0}\frac{f_m(3)^2}{f_m(1)^2} = 0$. One can show in the same way that $\lim_{\varphi\to 0}\frac{f_m(k)^2}{f_m(1)^2} = 0$ ($4 \le k \le m$). Using (2.7), we have

$$\text{(A.10)}\qquad \begin{aligned}\frac{1}{|w(1)|^2} &= \frac{1+\beta_{m+1}^2\|f_m\|^2}{\beta_{m+1}^2|f_m(1)|^2} = \frac{1}{\beta_{m+1}^2|f_m(1)|^2} + 1 + \sum_{k=2}^m \frac{f_m(k)^2}{f_m(1)^2}\\ &= 1 + O\left(\sin^2\varphi\right).\end{aligned}$$

The assertion follows immediately from (2.6). ∎

A.2. Preconditioned MINRES with tuning. We can use the same reasoning to show that $\lim_{\varphi\to 0}\|\tilde{r}_m\|/\|\mathbb{L}^{-1}x\| = 1$ for preconditioned MINRES with tuning.

Let $\tilde{B}\tilde{U}_m = \tilde{U}_m\tilde{T}_m + \tilde{\beta}_{j+1}\tilde{u}_{j+1}e_j^T$ be the m-step Lanczos decomposition. The first Lanczos vector $\tilde{u}_1 = \mathbb{L}^{-1}x/\|\mathbb{L}^{-1}x\| = \tilde{v}_1\cos\tilde{\varphi} + \tilde{u}\sin\tilde{\varphi}$, where $\tilde{u} \perp \tilde{v}_1$ and $\|\tilde{u}\| = 1$. Since the smallest eigenvalue of $\tilde{B}$ is $\tilde{\mu}_1 = O(\sin^2\tilde{\varphi})$ (by Theorem 9.1 of [1]), we have the first entry of $\tilde{T}_m$ as follows:

$$\text{(A.11)}\qquad \begin{aligned}\tilde{\alpha}_1 &= \tilde{u}_1^T\tilde{B}\tilde{u}_1 = (\tilde{v}_1\cos\tilde{\varphi} + \tilde{u}\sin\tilde{\varphi})^T\tilde{B}(\tilde{v}_1\cos\tilde{\varphi} + \tilde{u}\sin\tilde{\varphi})\\ &= \tilde{\mu}_1\cos^2\tilde{\varphi} + \bar{\tilde{\mu}}\sin^2\tilde{\varphi} = O\left(\sin^2\tilde{\varphi}\right),\end{aligned}$$

where $\bar{\tilde{\mu}} \in [\tilde{\mu}_2, \tilde{\mu}_n]$. In light of (A.4) and (A.5), we can show easily that $\tilde{\beta}_2 = O(\sin\tilde{\varphi})$, $\tilde{\alpha}_2$, $\tilde{\beta}_3$, and all other entries have a nonzero limit as $\tilde{\varphi}$ goes to zero (where we recall from the comment right after (3.17) that $\sin\tilde{\varphi} = O(\sin\varphi)$. An analysis of $\tilde{f}_m = \tilde{T}_m^{-1} e_m$ is similar to that of f_m as follows:

$$\text{(A.12)} \qquad \left|\tilde{f}_m(1)\right| = \frac{\left|\det\left[e_m, \tilde{T}_m(1:m, 2:m)\right]\right|}{\left|\det\left[\tilde{T}_m\right]\right|} = \frac{\prod_{k=2}^m \tilde{\beta}_k}{\left|\tilde{\alpha}_1 \det\left[\tilde{T}_m(2:m,2:m)\right] - \tilde{\beta}_2^2 \det\left[\tilde{T}_m(3:m,3:m)\right]\right|} = \frac{O(\sin\tilde{\varphi})}{O\left(\sin^2\tilde{\varphi}\right)} = O\left(\frac{1}{\sin\tilde{\varphi}}\right),$$

$$\text{(A.13)} \qquad \left|\tilde{f}_m(2)\right| = \frac{\left|\det\left[\tilde{T}_m(1:m,1), e_m, \tilde{T}_m(1:m, 3:m)\right]\right|}{\left|\det\left[\tilde{T}_m\right]\right|} = \frac{|\tilde{\alpha}_1| \prod_{k=3}^m \tilde{\beta}_k}{\left|\tilde{\alpha}_1 \det\left[\tilde{T}_m(2:m,2:m)\right] - \tilde{\beta}_2^2 \det\left[\tilde{T}_m(3:m,3:m)\right]\right|} = \frac{O\left(\sin^2\tilde{\varphi}\right)}{O\left(\sin^2\tilde{\varphi}\right)} = O(1).$$

One can show other entries of $\tilde{f}_m$ also have a nonzero limit and hence $\tilde{f}_m(k)/\tilde{f}_m(1) = O(\sin\tilde{\varphi})(2 \le k \le n)$. Exactly the same reasoning as in Appendix A.1 shows that $\lim_{\tilde{\varphi}\to 0} 1/|\tilde{w}(1)|^2 = 1$, and the relative linear residual $\lim_{\tilde{\varphi}\to 0} \|\tilde{r}_m\|/\|\mathbb{L}^{-1}x\| = 1$. Similar to the unpreconditioned solve, if $\tilde{\varphi}$ is small enough, then $1 - \|\tilde{r}_m\|/\|\mathbb{L}^{-1}x\| = O(\sin^2\tilde{\varphi})$.

Appendix B. Assumption of Theorem 3.2. In fact,

$$p_m(\mu_1) = \prod_{k=1}^m \left(1 - \mu_1/\xi_k^{(m)}\right) \approx \left(1 - \mu_1/\xi_1^{(m)}\right)\left(1 - \sum_{k=2}^m \mu_1/\xi_k^{(m)}\right)$$

$$\text{(B.1)} \qquad = 1 - \sum_{k=2}^m \mu_1/\xi_k^{(m)} - \left(\mu_1/\xi_1^{(m)}\right)\left(1 - \sum_{k=2}^m \mu_1/\xi_k^{(m)}\right),$$

which is smaller than 1 if and only if

$$\text{(B.2)} \qquad \xi_1^{(m)} > -\left(\frac{1 - \sum_{k=2}^m \mu_1/\xi_k^{(m)}}{\sum_{k=2}^m 1/\xi_k^{(m)}}\right).$$

On the other hand, we can find the closed form of $\xi_1^{(2)}$ and $\xi_2^{(2)}$ by the definition of harmonic Ritz values. We do this by solving the generalized eigenvalue problem $M_2^2 w = \xi T_2 w$, where, by Theorem 2.1,

$$\text{(B.3)} \qquad M_2^2 = \bar{T}_2^T \bar{T}_2 = \begin{bmatrix} \alpha_1^2 + \beta_2^2 & \beta_2(\alpha_1 + \alpha_2) \\ \beta_2(\alpha_1 + \alpha_2) & \beta_2^2 + \alpha_2^2 + \beta_3^2 \end{bmatrix}, \qquad T_2 = \begin{bmatrix} \alpha_1 & \beta_2 \\ \beta_2 & \alpha_2 \end{bmatrix}.$$

We solve the equivalent problem $T_2^{-1}M^2w = \xi w$ with $\alpha_1 = 0$ and find that

$$\xi_1^{(2)} = \frac{\alpha_2 - \sqrt{\alpha_2^2 + 4\beta_2^2 + 4\beta_3^2}}{2} = \frac{\alpha_2 - \sqrt{\alpha_2^2 + 4\beta_3^2 + O\left(\sin^2\varphi\right)}}{2}, \tag{B.4}$$

where β_2, α_2, and β_3 are given in (A.4) through (A.6). Note that this is a negative number bounded below independent of φ, and $\xi_1^{(m)}$ increases with m to approximate μ_1 from below. Therefore, in the first few MINRES iterations, $p_m(\mu_1) > 1$ if

$$\frac{\alpha_2 - \sqrt{\alpha_2^2 + 4\beta_2^2 + 4\beta_3^2}}{2} < -\left(\frac{1 - \sum_{k=2}^m \mu_1/\xi_k^{(m)}}{\sum_{k=2}^m 1/\xi_k^{(m)}}\right). \tag{B.5}$$

For some problems, $p_m(\mu_1) > 1$ holds in the initial MINRES iteration steps, but it will not take many iterations in practice before $p_m(\mu_1) < 1$, so that Theorem 3.2 can be applied and the bound in (3.2) becomes informative.

Acknowledgments. We thank the reviewers and the editor for their helpful comments and some very useful suggestions.

REFERENCES

[1] J. Berns-Müller and A. Spence, *Inexact inverse iteration and GMRES*, Electron. Trans. Numer. Anal., submitted.

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

[3] F. A. Dul, *MINRES and MINERR are better than SYMMLQ in eigenpair computations*, SIAM J. Sci. Comput., 19 (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*, Electron. Trans. Numer. Anal., 28 (2007), pp. 40–64.

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

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

[8] M. Freitag and A. Spence, *Rayleigh quotient iteration and simplified Jacobi-Davidson method with preconditioned iterative solves*, Linear Algebra Appl., 428 (2008), pp. 2049–2060.

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

[10] G. Golub and Q. Ye, *An inverse free preconditioned Krylov subspace method for symmetric generalized eigenvalue problems*, SIAM J. Sci. Comput., 24 (2002), pp. 312–334.

[11] Y. Lai, K. Lin, and W. Lin, *An inexact inverse iteration for large sparse eigenvalue problems*, Numer. Linear Algebra Appl., 4 (1997), pp. 425–437.

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

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

[14] R. B. Morgan and D. S. Scott, *Preconditioning the Lanczos algorithm for sparse symmetric eigenvalue problems*, SIAM J. Sci. Comput., 14 (1993), pp. 585–593.

[15] S. G. Nash and A. Sofer, *Linear and Nonlinear Programming*, McGraw-Hill, New York, 1996.

[16] Y. Notay, *Convergence analysis of inexact Rayleigh quotient iteration*, SIAM J. Matrix Anal. Appl., 24 (2003), pp. 627–644.

[17] C. Paige, B. Parlett, and H. van der Vorst, *Approximate solutions and eigenvalue bounds from Krylov subspaces*, Numer. Linear Algebra Appl., 2 (1995), pp. 115–134.

[18] B. Parlett, *The Symmetric Eigenvalue Problem*, Classics Appl. Math. 20, SIAM, Philadelphia, 1998.

[19] V. Simoncini and L. Eldén, *Inexact Rayleigh quotient-type methods for eigenvalue computations*, BIT, 42 (2002), pp. 159–182.
[20] G. L. G. Sleijpen and J. van den Eshof, *On the use of harmonic Ritz pairs in approximating internal eigenpairs*, Linear Algebra Appl., 358 (2003), pp. 115–137.
[21] P. Smit and M. H. C. Paardekooper, *The effects of inexact solvers in algorithms for symmetric eigenvalue problems*, Linear Algebra Appl., 287 (1999), pp. 337–357.
[22] G. W. Stewart, *Matrix Algorithms Volume* II: *Eigensystems*, SIAM, Philadelphia, 2001.