

THE CONVERGENCE OF HARMONIC RITZ VALUES,  
HARMONIC RITZ VECTORS,  
AND REFINED HARMONIC RITZ VECTORS

ZHONGXIAO JIA

ABSTRACT. This paper concerns a harmonic projection method for computing an approximation to an eigenpair  $(\lambda, x)$  of a large matrix  $A$ . Given a target point  $\tau$  and a subspace  $\mathcal{W}$  that contains an approximation to  $x$ , the harmonic projection method returns an approximation  $(\mu + \tau, \tilde{x})$  to  $(\lambda, x)$ . Three convergence results are established as the deviation  $\epsilon$  of  $x$  from  $\mathcal{W}$  approaches zero. First, the harmonic Ritz value  $\mu + \tau$  converges to  $\lambda$  if a certain Rayleigh quotient matrix is uniformly nonsingular. Second, the harmonic Ritz vector  $\tilde{x}$  converges to  $x$  if the Rayleigh quotient matrix is uniformly nonsingular and  $\mu + \tau$  remains well separated from the other harmonic Ritz values. Third, better error bounds for the convergence of  $\mu + \tau$  are derived when  $\tilde{x}$  converges. However, we show that the harmonic projection method can fail to find the desired eigenvalue  $\lambda$ —in other words, the method can miss  $\lambda$  if it is very close to  $\tau$ . To this end, we propose to compute the Rayleigh quotient  $\rho$  of  $A$  with respect to  $\tilde{x}$  and take it as a new approximate eigenvalue.  $\rho$  is shown to converge to  $\lambda$  once  $\tilde{x}$  tends to  $x$ , no matter how  $\tau$  is close to  $\lambda$ . Finally, we show that if the Rayleigh quotient matrix is uniformly nonsingular, then the refined harmonic Ritz vector, or more generally the refined eigenvector approximation introduced by the author, converges. We construct examples to illustrate our theory.

1. INTRODUCTION

We are concerned with computing a simple eigenpair  $(\lambda, x)$  of a large matrix  $A$  by a general harmonic projection method, which for approximating  $(\lambda, x)$  is to produce a sequence of subspaces  $\mathcal{W}_k$  that contain increasingly accurate approximations to  $x$ . The harmonic Ritz values and Ritz vectors were first studied by Freund [1], Manteuffel and Otto [14] and Morgan [15], but the word ‘harmonic’ was introduced by Paige, Parlett and van der Vorst [17] for symmetric matrices and by Sleijpen and van der Vorst [20] for general matrices. Suppose we want to compute eigenvalues near a target point  $\tau$  in the complex plane. Then these possible interior eigenvalues of  $A$  are transformed into exterior eigenvalues of the matrix  $(A - \tau I)^{-1}$ , so a standard projection method working on  $(A - \tau I)^{-1}$  can be used. However, it is often the case that in practical applications factoring  $A - \tau I$  is too expensive

---

Received by the editor June 7, 2002 and, in revised form, December 23, 2003.

2000 *Mathematics Subject Classification*. Primary 15A18, 65F15, 65F30.

*Key words and phrases*. Harmonic projection, refined harmonic projection, harmonic Ritz value, harmonic Ritz vector, refined harmonic Ritz vector, refined eigenvector approximation, convergence.

Work supported by Special Funds for the State Major Basic Research Projects (G1999032805).

and even impractical. The harmonic projection method is primarily used to settle this problem and to compute interior eigenvalues and eigenvectors of large matrices without factoring  $A - \tau I$ ; see [15, 16, 22, 24] for more details.

In the simplest form the harmonic projection method works as follows (see, e.g., [16], where the method is also referred to as the harmonic Rayleigh–Ritz method. Here we drop the iteration subscript).

1. Compute an orthonormal basis  $W$  for  $\mathcal{W}$ .
- (1) 2. Compute  $B = W^H(A - \tau I)W$  and  $C = W^H(A - \tau I)^H(A - \tau I)W$ .
3. Let  $(\mu, z)$  be an eigenpair of  $B^{-1}C$ , where  $\mu + \tau \cong \lambda$ .
4. Take  $(\mu + \tau, \tilde{x}) = (\mu + \tau, Wz)$  as an approximate eigenpair of  $A$ .

Steps 3 and 4 can be used to compute approximations to other eigenpairs of  $A$ . In the sequel, without loss of generality we suppose for brevity that  $\tau = 0$  unless stated otherwise when making a theoretical analysis.

The matrix  $B^H$  is called the Rayleigh quotient matrix and  $B^{-1}C$  the harmonic Rayleigh quotient matrix of  $A$  with respect to  $W$  and the target  $\tau$ , respectively. The number  $\mu + \tau$  is called a harmonic Ritz value and the vector  $\tilde{x} = Wz$  a harmonic Ritz vector. The rationale for the method is that if  $x \in \mathcal{W}$ , then there is an eigenpair  $(\lambda - \tau, z)$  of  $B^{-1}C$  with  $x = Wz$ . Therefore, it would be intuitively expected that if  $x$  were nearly in  $\mathcal{W}$ , then there should be an eigenpair  $(\mu, z)$  of  $B^{-1}C$  with  $\mu + \tau$  near  $\lambda$  and  $Wz$  near  $x$ .

There are a lot of convergence results available on standard projection methods. When  $A$  is non-Hermitian, we established a priori error bounds for Ritz values and Ritz vectors in terms of the deviation of  $x$  from  $\mathcal{W}$  [3]. The results show that the continuity argument holds and the Ritz values converge. The Ritz vectors, on the other hand, behave more erratically and perhaps fail to converge; that is, the continuity argument for the Ritz vectors does not hold. For a detailed analysis, see [6, 8]. This led us to introduce certain refined eigenvector approximations for which the continuity argument is valid [4, 5]. A number of refined projection methods were studied and a few refined algorithms were developed [4, 5, 7, 9]. We also proposed a refined harmonic Arnoldi method and developed an implicitly restarted algorithm for computing interior eigenpairs of  $A$  [10], where certain refined shifts were presented that were shown to be better than the commonly used exact shifts. Extensive numerical experiments were made, and comparisons were drawn on the refined algorithms available and the corresponding conventional counterparts, indicating that the refined algorithms were often considerably superior to their conventional counterparts. The principle of refined projection methods has, very recently, been extended to compute a partial singular value decomposition of a large matrix. We proposed a refined bidiagonalization Lanczos method and developed an implicitly restarted bidiagonalization Lanczos algorithm, which was more efficient than the implicitly restarted bidiagonalization Lanczos algorithm [13]. In a general setting, a unified convergence theory of general refined projection methods was established when  $A$  is diagonalizable [8] and when  $A$  is a general matrix [12]. Some other theoretical comparisons are made between the refined Ritz vector and the Ritz vector in [11], where a lower and an upper bound are established for the sine of the angle between the two vectors and some relationships are derived between the residual norm of the refined Ritz pair and that of the Ritz pair. The theoretical results show that the Ritz vector can be far away from the refined Ritz vector and the ratio of the residual norm of the Ritz pair and that of the refined

Ritz pair can be arbitrarily large. Some examples are constructed to illustrate the theory. It is pointed out in [11] that the results in it hold for the harmonic and refined harmonic Ritz vectors. Stewart's book [22] and van der Vorst's monograph [24] have given an excellent account of the refined projection methods.

Although there are some papers and scattered results on harmonic projection methods for some special projection subspaces, no general convergence results have been hitherto established for the harmonic method in a unified manner. How to carry out this important task is the goal of this paper. We attempt to establish a general convergence theory for the harmonic method. Our aim is to derive bounds or equalities in terms of the deviation  $\epsilon$  of  $x$  from the subspace  $\mathcal{W}$ ; for a precise definition of  $\epsilon$ , see (3). Here we should point out that the size of  $\epsilon$  must be established from the properties of the underlying algorithm that determines  $\mathcal{W}$ .

We will establish a few convergence results. First, as  $\epsilon \rightarrow 0$ , there is a harmonic Ritz value  $\mu + \tau$  that converges to the eigenvalue  $\lambda$  under the assumption that  $B$  is uniformly nonsingular, that is, the norm of its inverse is uniformly bounded. Second, it is shown that if  $\mu$  remains uniformly separated in an appropriate sense from the other eigenvalues of  $B^{-1}C$  and  $B$  is uniformly nonsingular, then the harmonic Ritz vector converges. So it may be harder for the harmonic Ritz vector to converge than it is for the corresponding harmonic Ritz value. Third, better results for the convergence of  $\mu + \tau$  are derived when  $\tilde{x}$  converges. We show why the harmonic method can produce a spurious harmonic Ritz value and fail to find the desired eigenvalue  $\lambda$  if it is very close to the target  $\tau$ , a phenomenon observed by some researchers [15, 16, 21]. We construct a set of examples to illustrate this phenomenon. We provide a practical way of telling when the harmonic Ritz vector corresponding to a converging harmonic Ritz value is itself converging. Based on a known result of [12], the refined harmonic Ritz vectors, or more generally the refined eigenvector approximation introduced by the author, converges if the approximate eigenvalue converges as  $\epsilon \rightarrow 0$ . Thus the refined eigenvector approximation has the same sufficient conditions for convergence as the corresponding approximate eigenvalue, i.e., the harmonic Ritz value or the Rayleigh quotient, and it needs fewer conditions to converge than the harmonic Ritz vector. Furthermore, we construct an example to indicate that the harmonic projection method cannot tell us how to choose and compute a significant and good harmonic Ritz vector, even though the space  $\mathcal{W}$  contains a good enough approximation to a desired eigenvector, while the refined harmonic projection method works well and delivers an excellent approximation to the desired eigenvector. This happens when the harmonic Ritz value approximating the desired eigenvalue is multiple.

In this paper the norm  $\|\cdot\|$  will denote both the Euclidean vector norm and the subordinate spectral matrix norm. Throughout, we will assume that eigenvectors, harmonic Ritz vectors and refined harmonic Ritz vectors have been normalized to have unit length. We will measure the deviation of a normalized vector  $y$  from a subspace  $\mathcal{X}$  as follows. Let  $X$  be an orthonormal basis for the space  $\mathcal{X}$  and let  $X_\perp$  be an orthonormal basis for the orthogonal complement of  $\mathcal{X}$ . We will measure the deviation of  $y$  from  $\mathcal{X}$  by the quantity

$$(2) \quad \sin \angle(y, \mathcal{X}) = \|X_\perp^H y\| = \|(I - P_{\mathcal{X}})y\|,$$

where  $P_{\mathcal{X}}$  is the orthogonal projector onto  $\mathcal{X}$ . As the notation indicates, this number is the sine of the angle between  $y$  and  $\mathcal{X}$ . For brevity, we will denote the

sine of the angle between our fixed eigenvector  $x$  and our subspace  $\mathcal{W}$  by

$$(3) \quad \epsilon = \sin \angle(x, \mathcal{W}).$$

Let  $X_\perp$  be an orthonormal basis for the orthogonal complement of the space spanned by  $x$ , so that  $(x \ X_\perp)$  is unitary. Then it follows from the relation  $Ax = \lambda x$  that

$$(4) \quad \begin{pmatrix} x^H \\ X_\perp^H \end{pmatrix} A (x \ X_\perp) = \begin{pmatrix} \lambda & h^H \\ 0 & L \end{pmatrix},$$

where  $h^H = x^H A X_\perp$  and  $L = X_\perp^H A X_\perp$ . Because the right-hand side of (4) is block triangular, the eigenvalues of  $A$  consist of  $\lambda$  and the eigenvalues of  $L$ . Since  $\lambda$  is assumed to be simple, the eigenvalues of  $L$  are not equal to  $\lambda$  and hence  $L - \lambda I$  is nonsingular. Define

$$(5) \quad \text{sep}(\lambda, L) = \|(L - \lambda I)^{-1}\|^{-1},$$

where  $\|\cdot\|$  is the spectral matrix norm. The quantity  $\text{sep}$ , which approaches zero as  $\lambda$  approaches an eigenvalue of  $L$ , will play a central role in our theory. For more on the properties of  $\text{sep}$ , see [23].

In the next section we will consider the convergence of harmonic Ritz values. In §3 we will consider the convergence of harmonic Ritz vectors. In §4 we will derive better results for the convergence of  $\mu + \tau$  when  $\tilde{x}$  converges. We show that even if  $\tilde{x}$  tends to  $x$  the harmonic projection method may fail to compute the desired eigenvalue  $\lambda$ , or in other words they can miss  $\lambda$  if it is very close to  $\tau$ . To correct this problem, we propose to compute the Rayleigh quotient  $\rho = \tilde{x}^H A \tilde{x}$  and take it as a new approximate eigenvalue. We show that  $\rho$  converges to  $\lambda$ , independently of how close  $\lambda$  is to  $\tau$ , once  $\tilde{x}$  converges. In §5 we will establish the convergence of refined harmonic Ritz vectors or, more generally, refined eigenvector approximations. The paper concludes with a discussion of our results.

## 2. CONVERGENCE OF HARMONIC RITZ VALUES

It will be shown that the hypothesis  $\epsilon \rightarrow 0$  is not sufficient to insure there is a harmonic Ritz value  $\mu + \tau$  that converges to  $\lambda$ , and the convergence requires that  $B$  be uniformly nonsingular, i.e., the norm of its inverse is uniformly bounded as  $\epsilon \rightarrow 0$ . We will establish this result in two stages. First we will show that if  $\epsilon$  is small and  $\|B^{-1}\|$  is uniformly bounded, then  $\lambda$  is an exact eigenvalue of a matrix  $B^{-1}C + E$  with  $E$  being a small perturbation matrix, assuming without loss of generality that  $\tau = 0$ . We will then use a theorem of Elsner to show that  $B^{-1}C$  must have an eigenvalue  $\mu$  near  $\lambda$ .

**Theorem 2.1.** *Let  $B^{-1}C$  be the harmonic Rayleigh quotient matrix in (1), let  $(\lambda, x)$  be the desired eigenpair of  $A$  with  $\|x\| = 1$ , and let  $\epsilon$ ,  $W$  and  $W_\perp$  be defined as in the introduction. Then there is a matrix  $E$  satisfying*

$$(6) \quad \|E\| \leq \frac{\epsilon}{\sqrt{1 - \epsilon^2}} \|B^{-1}\| (|\lambda| \|A\| + \|A\|^2)$$

and such that  $\lambda$  is an eigenvalue of  $B^{-1}C + E$ .

*Proof.* Let  $y = W^H x$  and  $y_\perp = W_\perp^H x$ . By (2) and (3), we have  $\|y_\perp\| = \epsilon$ . Hence  $\|y\| = \sqrt{1 - \epsilon^2}$ .

From the relation  $Ax - \lambda x = 0$  we have  $A^H Ax - \lambda A^H x = 0$ . Note that  $WW^H + W_\perp W_\perp^H = I$ . Then we have

$$W^H A^H A (W \ W_\perp) \begin{pmatrix} W^H \\ W_\perp^H \end{pmatrix} x - \lambda W^H A^H (W \ W_\perp) \begin{pmatrix} W^H \\ W_\perp^H \end{pmatrix} x = 0.$$

Equivalently,

$$(7) \quad W^H A^H A W y + W^H A^H A W_\perp y_\perp - \lambda W^H A^H W y - \lambda W^H A^H W_\perp y_\perp = 0.$$

Note the definitions of  $B$  and  $C$ . Then left-multiplying the above relation by  $(W^H A^H W)^{-1}$  gives

$$B^{-1} C y - \lambda y = B^{-1} (\lambda W^H A^H W_\perp y_\perp - W^H A^H A W_\perp y_\perp).$$

Normalize  $y$  by setting  $\hat{y} = y/\sqrt{1 - \epsilon^2}$ , and set

$$(8) \quad r = B^{-1} C \hat{y} - \lambda \hat{y} = B^{-1} \frac{\lambda W^H A^H W_\perp y_\perp - W^H A^H A W_\perp y_\perp}{\sqrt{1 - \epsilon^2}}.$$

Then it follows from (8) that

$$\begin{aligned} \|r\| &\leq \frac{\epsilon}{\sqrt{1 - \epsilon^2}} \|B^{-1}\| (\|\lambda\| \|W^H A^H W_\perp\| + \|W^H A^H A W_\perp\|) \\ &\leq \frac{\epsilon}{\sqrt{1 - \epsilon^2}} \|B^{-1}\| (\|\lambda\| \|A\| + \|A\|^2), \end{aligned}$$

because  $W$  and  $W_\perp$  have orthonormal columns that lead to  $\|W^H A^H W_\perp\| \leq \|A^H\| = \|A\|$  and  $\|W^H A^H A W_\perp\| \leq \|A^H A\| = \|A\|^2$ .

If we now define

$$E = -r \hat{y}^H,$$

then it is easy to verify that  $E$  satisfies (6) and  $(B^{-1} C + E) \hat{y} = \lambda \hat{y}$ . □

When  $A$  is a Hermitian positive or negative definite matrix, it is known that  $B$  is such a matrix too and furthermore by the interlacing property the smallest singular value of  $B$  is no less than that of  $A$ , i.e.,  $\|B^{-1}\| \leq \|A^{-1}\|$ , so that  $B$  is uniformly nonsingular independently of the size of  $\epsilon$ . For a general  $A$ , however, there is no guarantee that  $B$  is nonsingular as  $\epsilon \rightarrow 0$  even though  $A$  is assumed to be nonsingular. When  $A$  is positive (negative) definite, i.e., when the Hermitian part  $\frac{1}{2}(A + A^H)$  of  $A$  is positive (negative) definite, it is easily seen that  $B$  is positive definite too. So  $B$  is surely nonsingular, but  $\|B^{-1}\|$  is not necessarily uniformly bounded as  $\epsilon \rightarrow 0$ . This implies that for a general matrix  $A$ , although it is positive (negative) definite, there is no guarantee that  $E$  must converge to zero if  $\epsilon \rightarrow 0$ . In the sequel, we will see how the convergence of harmonic Ritz values, harmonic Ritz vectors and refined harmonic Ritz vectors to be introduced later depends on  $B$ .

If  $E$  goes to zero as  $\epsilon \rightarrow 0$ , it would appear that we can now deduce the convergence of the harmonic Ritz values from the continuity of eigenvalues of a matrix. However, it is not so simple, and there is a subtle point here. As  $\epsilon$  changes, the space  $\mathcal{W}$  and hence the harmonic Rayleigh quotient  $B^{-1} C$  must also change. Thus there is the possibility that as  $\epsilon \rightarrow 0$  the eigenvalue  $\lambda$  of  $B^{-1} C + E$  becomes ill-conditioned so quickly that no eigenvalue of  $B^{-1} C$  is near  $\lambda$  although  $E$  tends to zero.

Fortunately, a theorem of Elsner [23, p.168] implies that this cannot happen if  $\|B^{-1}\|$  is uniformly bounded. Specifically, Elsner's theorem says that, given

matrices  $A$  and  $\tilde{A}$  of order  $n$ , for any eigenvalue  $\lambda$  of  $A$  there is an eigenvalue  $\tilde{\lambda}$  of  $\tilde{A}$  satisfying

$$|\lambda - \tilde{\lambda}| \leq (\|A\| + \|\tilde{A}\|)^{1-\frac{1}{n}} \|A - \tilde{A}\|^{\frac{1}{n}}.$$

Hence we have the following corollary.

**Corollary 2.2.** *There is an eigenvalue  $\mu$  of  $B^{-1}C$  such that*

$$(9) \quad |\lambda - \mu| \leq (2\|A\| + \|E\|)^{1-\frac{1}{m}} \|E\|^{\frac{1}{m}},$$

where  $m$  is the order of  $B^{-1}C$ .

The right-hand side of (9) depends on  $\|A\|$ ,  $\epsilon$  and  $\|B^{-1}\|$ . Hence we can conclude that, as  $\epsilon \rightarrow 0$ , if  $\|B^{-1}\|$  is uniformly bounded there is always a harmonic Ritz value that converges to  $\lambda$ . The bound (9) will, though attainable, in general be a gross overestimate. We will come back to this point and revisit the convergence of  $\mu$  in §4, where better results on  $|\lambda - \mu|$  will be derived.

### 3. CONVERGENCE OF HARMONIC RITZ VECTORS

Elsner’s theorem says that no eigenvalue of a matrix can be infinitely ill-conditioned in a global sense as  $\epsilon \rightarrow 0$  if  $\|B^{-1}\|$  is uniformly bounded. The effects of a given error on the eigenvalues are limited, and the smaller the error, the smaller the bound. However, the same is decidedly not true of eigenvectors. For example, as two simple eigenvalues approach one another, their eigenvectors become less and less stable. In fact, there are arbitrarily small perturbations that can be made to have arbitrarily large effects simply by bringing the eigenvalues close enough.

Fortunately, the following theorem [12] for an arbitrary pair shows that there is an easily checked condition that insures that the harmonic Ritz vector converges.

**Theorem 3.1.** *Let the (not necessarily harmonic Ritz) pair  $(\nu, \tilde{x})$  be given, and let  $\text{sep}(\lambda, L)$  be defined by (5). If*

$$(10) \quad \text{sep}(\lambda, L) - |\nu - \lambda| > 0,$$

then

$$(11) \quad \sin \angle(x, \tilde{x}) \leq \frac{\|A\tilde{x} - \nu\tilde{x}\|}{\text{sep}(\nu, L)} \leq \frac{\|A\tilde{x} - \nu\tilde{x}\|}{\text{sep}(\lambda, L) - |\nu - \lambda|}.$$

We have seen that  $\mu + \tau \rightarrow \lambda$  as  $\epsilon \rightarrow 0$  if  $\|B^{-1}\|$  is uniformly bounded. If, in addition, the residual norm  $\rho = \|A\tilde{x} - (\mu + \tau)\tilde{x}\|$  approaches zero, the harmonic Ritz vector  $\tilde{x}$  converges to the eigenvector  $x$ . Thus a converging harmonic Ritz value and vanishing residual imply a converging harmonic Ritz vector. These convergence conditions can be easily checked in the course of the harmonic projection methods.

It is instructive to see how the harmonic Ritz vector can fail to converge. Let  $\tau = 0$ , let  $\mu$  be the converging harmonic Ritz value, and let  $z$  be the corresponding eigenvector of  $B^{-1}C$ . Let  $(z \ Z_{\perp})$  be unitary. From the relation  $B^{-1}Cz = \mu z$  it follows that

$$\begin{pmatrix} z^H \\ Z_{\perp}^H \end{pmatrix} B^{-1}C(z \ Z_{\perp}) = \begin{pmatrix} \mu & g^H \\ 0 & G \end{pmatrix}.$$

Since the only assumption about the subspace  $\mathcal{W}$  is that it contains an approximation to the eigenvector  $x$ , the eigenvalues of  $G$  need not be near the eigenvalues of  $A$  other than  $\lambda$ . In fact, all the eigenvalues of  $G$  lie in a disk at center 0 and radius  $\|G\|$ . In particular, an eigenvalue of  $G$  could happen to be equal to  $\mu$ . If this double

eigenvalue is not defective, its eigenvectors will span a two-dimensional subspace, and it will be impossible to tell which is the one that reproduces an approximation to  $x$  without computing each of the residual norms. When an eigenvalue of  $G$  is near but not equal to  $\mu$ , there will be a unique eigenvector associated with  $\mu$ , but there is no guarantee that it will reproduce a significant approximation to  $x$ .

The above qualitative analysis suggests that the harmonic Ritz vector will converge provided the eigenvalue  $\mu$  of  $B^{-1}C$  remains well separated from those of  $G$  and  $\|B^{-1}\|$  is uniformly bounded. The following theorem shows that this is indeed the case.

**Theorem 3.2.** *Under the above assumptions, if*

$$\text{sep}(\lambda, G) > 0,$$

then

$$(12) \quad \sin \angle(x, \tilde{x}) \leq \left( 1 + \frac{2\|B^{-1}\|\|A\|^2}{\sqrt{1-\epsilon^2}\text{sep}(\lambda, G)} \right) \epsilon.$$

*Proof.* As in the proof of Theorem 2.1, let  $\hat{y}$  be the normalized value of  $W^H x$ . Then if  $r = B^{-1}C\hat{y} - \lambda\hat{y}$ , we have from  $|\lambda| \leq \|A\|$  and an inequality in the proof of Theorem 2.1 that

$$\|r\| \leq \frac{\epsilon}{\sqrt{1-\epsilon^2}} \|B^{-1}\| (|\lambda|\|A\| + \|A\|^2) \leq \frac{2\epsilon}{\sqrt{1-\epsilon^2}} \|B^{-1}\|\|A\|^2.$$

Let  $(\mu, z)$  be the eigenpair of  $B^{-1}C$  such that  $Wz$  is the harmonic Ritz vector  $\tilde{x}$  used to approximate  $x$ . Then it follows from Theorem 3.1 that

$$\sin \angle(z, \hat{y}) \leq \frac{\|r\|}{\text{sep}(\lambda, G)}.$$

Hence

$$\sin \angle(z, \hat{y}) \leq \frac{2\|B^{-1}\|\|A\|^2}{\text{sep}(\lambda, G)} \frac{\epsilon}{\sqrt{1-\epsilon^2}}.$$

Since  $W$  is orthonormal, we have from the definition of  $\tilde{x}$  and  $\hat{y}$  that

$$\sin \angle(z, \hat{y}) = \sin \angle(Wz, W\hat{y}) = \sin \angle(\tilde{x}, WW^H x) = \sin \angle(\tilde{x}, P_{\mathcal{W}}x).$$

Note that

$$\angle(x, \tilde{x}) \leq \angle(x, P_{\mathcal{W}}x) + \angle(P_{\mathcal{W}}x, \tilde{x}) = \angle(x, \mathcal{W}) + \angle(\tilde{x}, P_{\mathcal{W}}x),$$

where the equality comes from the relation  $\angle(x, P_{\mathcal{W}}x) = \angle(x, \mathcal{W})$ . Hence

$$\sin \angle(x, \tilde{x}) \leq \sin \angle(x, \mathcal{W}) + \sin \angle(\tilde{x}, P_{\mathcal{W}}x)$$

Since  $\sin \angle(x, \mathcal{W}) = \|(I - P_{\mathcal{W}})x\| = \epsilon$  and  $\sin \angle(\tilde{x}, P_{\mathcal{W}}x) = \sin \angle(z, \hat{y})$ , we have

$$\begin{aligned} \sin \angle(x, \tilde{x}) &\leq \epsilon + \frac{2\|B^{-1}\|\|A\|^2}{\text{sep}(\lambda, G)} \frac{\epsilon}{\sqrt{1-\epsilon^2}} \\ &\leq \left( 1 + \frac{2\|B^{-1}\|\|A\|^2}{\sqrt{1-\epsilon^2}\text{sep}(\lambda, G)} \right) \epsilon. \quad \square \end{aligned}$$

By the continuity of  $\text{sep}$ , we have  $\text{sep}(\lambda, G) \geq \text{sep}(\mu, G) - |\mu - \lambda|$ . Hence

$$\sin \angle(x, \tilde{x}) \leq \left( 1 + \frac{2\|B^{-1}\|\|A\|^2}{\sqrt{1-\epsilon^2}(\text{sep}(\mu, G) - |\mu - \lambda|)} \right) \epsilon.$$

Since  $\mu$  approaches  $\lambda$  as  $\epsilon \rightarrow 0$  if  $\|B^{-1}\|$  is uniformly bounded, we see that a sufficient condition for the convergence of the harmonic Ritz vector is that  $\text{sep}(\mu, G)$  be uniformly bounded away from 0, i.e.,  $\mu$  is uniformly well separated from the other harmonic Ritz values of  $A$  with respect to  $\mathcal{W}$ . This condition can also be checked during the computation of the harmonic Ritz vector. Note that  $\text{sep}(\mu, G)$  can be arbitrarily small and even equal zero, no matter how well  $\lambda$  is separated from the other eigenvalues of  $A$ . So it can happen that the bound of (12) does not converge to zero as  $\epsilon \rightarrow 0$ . Therefore,  $\tilde{x}$  may not converge to  $x$ , even though  $\epsilon$  tends to zero and  $\|B^{-1}\|$  is uniformly bounded independently of the size of  $\epsilon$ . The result illustrates that if  $\mu$  is very close to some other eigenvalues of  $B^{-1}C$ , then  $\tilde{x}$  could be an inaccurate approximation to  $x$  even though  $\epsilon$  is very small.

4. CONVERGENCE OF HARMONIC RITZ VALUES REVISITED

As Corollary 2.2 has indicated, the error bound on  $|\lambda - \mu|$ , though attainable, may in general be a gross overestimate. Assuming that the harmonic Ritz vector  $\tilde{x}$  converges to  $x$ , we now derive a much more satisfactory result on the convergence of  $\mu$ .

**Theorem 4.1.** *Let  $(\mu, \tilde{x})$  be the harmonic approximation to  $(\lambda, x)$ , and assume that  $\tilde{x} \rightarrow x$  with  $\sin \angle(x, \tilde{x}) = O(\epsilon)$  and  $|\lambda| \gg \epsilon$ . Then, for a general  $A$ ,*

$$(13) \quad |\lambda - \mu| = O(\epsilon),$$

and, for a Hermitian  $A$ ,

$$(14) \quad |\lambda - \mu| = O(\epsilon^2).$$

*Proof.* By assumption and (1), we have

$$\mu = \frac{\tilde{x}^H A^H A \tilde{x}}{\tilde{x}^H A^H \tilde{x}}.$$

Now let us write  $\tilde{x} = \alpha x + \beta w$ , where  $w \perp x$  and  $\|w\| = 1$ . Then  $|\alpha| = \cos \angle(x, \tilde{x})$  and  $|\beta| = \sin \angle(x, \tilde{x})$ .

We now get

$$(15) \quad \begin{aligned} \mu &= \frac{(\alpha \lambda x + \beta A w)^H (\alpha \lambda x + \beta A w)}{(\alpha \lambda x + \beta A w)^H (\alpha x + \beta w)} \\ &= \frac{|\alpha|^2 |\lambda|^2 + \bar{\alpha} \bar{\lambda} \beta x^H A w + \alpha \lambda \bar{\beta} w^H A^H x + |\beta|^2 \|A w\|^2}{|\alpha|^2 \bar{\lambda} + \alpha \bar{\beta} w^H A^H x + |\beta|^2 w^H A^H w} \end{aligned}$$

$$(16) \quad = \frac{(1 - \epsilon^2) |\lambda|^2 + O(\epsilon)}{(1 - \epsilon^2) \bar{\lambda} + O(\epsilon)},$$

where the bar denotes the complex conjugate of a number. Since it is assumed that  $|\lambda| \gg \epsilon$ , it is easily verified that

$$|\lambda - \mu| = O(\epsilon).$$

If  $A$  is Hermitian, then  $x^H A w = w^H A^H x = 0$ . So (15) becomes

$$(17) \quad \mu = \frac{(1 - \epsilon^2) |\lambda|^2 + O(\epsilon^2)}{(1 - \epsilon^2) \bar{\lambda} + O(\epsilon^2)},$$

from which it follows that

$$|\lambda - \mu| = O(\epsilon^2). \quad \square$$

For a practical algorithm, it is impossible to let  $\epsilon$  tend to zero, and instead it is required that  $\epsilon$  be small, say  $\epsilon = 10^{-6}$  or  $10^{-8}$ . Keeping this in mind, we know, although  $|\lambda|$  or more generally  $|\lambda - \tau|$  is a constant, it can be very small. It is observed from (16) and (17) that if  $|\lambda|$  is of order  $\epsilon$  or smaller, then (13) and (14) are no longer valid. In fact, it can be easily derived from both (16) and (17) that  $|\mu - \lambda| = O(1)$  in this case. This means that the harmonic method may not work well, and it will very possibly produce a spurious or insignificant harmonic Ritz value  $\mu$ . Therefore, whenever  $\tau$  is very close to the desired eigenvalue  $\lambda$ , it is very likely that the method will miss  $\lambda$  even though  $\tilde{x}$  is supposed to converge to  $x$ . Such behavior was observed in experiments, e.g., [15, 16, 21], but no theoretical explanation or justification was given. Here our theoretical analysis shows why such a phenomenon could happen.

As argued, the requirement that  $|\lambda| \gg \epsilon$ , or more generally  $|\lambda - \tau| \gg \epsilon$ , is essential and cannot be dropped for  $\mu + \tau \rightarrow \lambda$ . Alternatively, however, we can choose a better and reliable approximate eigenvalue, which definitely converges to  $\lambda$  when  $\tilde{x}$  converges to  $x$ , no matter how small  $|\lambda - \tau|$  is, as shown below.

**Theorem 4.2.** *Assume that  $\tilde{x} \rightarrow x$  with  $\sin \angle(x, \tilde{x}) = O(\epsilon)$ , and define the Rayleigh quotient  $\rho = \tilde{x}^H A \tilde{x}$ . Then for a general  $A$  and any  $\tau$  we have*

$$(18) \quad |\lambda - \rho| = O(\epsilon),$$

and for a Hermitian  $A$  we have

$$(19) \quad |\lambda - \rho| = O(\epsilon^2).$$

*Proof.* The proof is analogous to that of Theorem 4.1. □

We should point out that there is a close relationship between the above results and (6) and (9). We will present an important result for a general  $\tau$ .

**Theorem 4.3.** *Suppose that  $\lambda$  is the eigenvalue of  $A$  closest to  $\tau$  and  $\epsilon$  as defined in the previous theorem. Then for a general  $A$  we have*

$$(20) \quad \|B^{-1}\| \geq \frac{1}{|\lambda - \tau| + c\epsilon},$$

and for a Hermitian positive definite  $A - \tau I$  we have

$$(21) \quad \frac{1}{|\lambda - \tau| + c\epsilon} \leq \|B^{-1}\| \leq \frac{1}{|\lambda - \tau|},$$

where

$$c = \|A\| + |\tau| - |\lambda - \tau|\epsilon.$$

*Proof.* Suppose that  $x$  is the eigenvector of  $A$  associated with  $\lambda$ . It follows from the definition of  $B$  that

$$\begin{aligned} \min_{\|y\|=1} \|By\| &= \min_{\|y\|=1} \|B^H y\| \\ &= \min_{\|y\|=1} \|W^H (A - \tau I) W y\|. \end{aligned}$$

Since  $W$  is orthonormal and  $y$  is an arbitrary  $m$ -dimensional vector of unit length, we may take a specific  $y$  such that

$$W y = \alpha x + \beta x_{\perp}$$

with  $x \perp x_\perp$ ,  $\|x_\perp\| = 1$  and  $|\alpha| = \sqrt{1 - \epsilon^2}$ ,  $|\beta| = \epsilon$ . Therefore, from  $\|W^H x\| = \sqrt{1 - \epsilon^2}$  we have

$$\begin{aligned} \min_{\|y\|=1} \|By\| &\leq \|W^H(A - \tau I)(\alpha x + \beta x_\perp)\| \\ &= \|\alpha(\lambda - \tau)W^H x + \beta W^H(A - \tau I)x_\perp\| \\ &\leq |\lambda - \tau| |\alpha| \|W^H x\| + |\beta| \|A - \tau I\| \\ &= |\lambda - \tau|(1 - \epsilon^2) + (\|A\| + |\tau|)\epsilon \\ &= |\lambda - \tau| + c\epsilon. \end{aligned}$$

Since

$$\|B^{-1}\| = \frac{1}{\min_{\|y\|=1} \|By\|},$$

it follows from the above that

$$\|B^{-1}\| \geq \frac{1}{|\lambda - \tau| + c\epsilon},$$

which proves (20).

If  $A - \tau I$  is Hermitian positive definite, then  $B = W^H(A - \tau I)W$  is too. Since  $W$  is orthonormal, by the interlacing property we get

$$\|B^{-1}\| \leq \|(A - \tau I)^{-1}\| = \frac{1}{|\lambda - \tau|},$$

which proves (21).  $\square$

The above theorem states that once  $|\lambda - \tau|$  is very small, say of order  $\epsilon$  or even smaller, the upper bound in (6) will not be close to zero, so that  $\|E\|$  is not necessarily small for a very small but fixed  $\epsilon$ . If  $A - \tau I$  is Hermitian positive definite, the bound (21) indicates that  $\|B^{-1}\|$  is uniformly bounded but is definitely of the same order as  $1/|\lambda - \tau|$ . Hence it follows from (9) that in a practical algorithm  $\mu + \tau$  may be an insignificant approximation to  $\lambda$  for a general matrix  $A$ , i.e., the harmonic method will very possibly produce a spurious harmonic Ritz value.

Recall from (1) that  $\tilde{x} = Wz$ , where  $z$  is the eigenvector of the harmonic Rayleigh quotient  $B^{-1}C$  associated with the eigenvalue  $\mu$ . So

$$\rho = \tilde{x}^H A \tilde{x} = \tilde{x}^H (A - \tau I + \tau I) \tilde{x} = z^H B^H z + \tau,$$

which means that it is very cheap to compute  $\rho$ . Theorem 4.2 says that once  $\tilde{x}$  converges then its Rayleigh quotient must converge to the eigenvalue  $\lambda$  *no matter how close  $\tau$  is to  $\lambda$* . Note that Theorem 4.1 means that once  $\tau$  is very close to the eigenvalue  $\lambda$  and  $\mu + \tau$  is used to approximate  $\lambda$ , then the harmonic projection method can fail since  $\mu + \tau$  does not converge to  $\lambda$ , i.e., the method can miss the desired eigenvalue  $\lambda$ . Therefore, the above two theorems suggest that it would be better and safer to use the Rayleigh quotient  $\rho$  of the harmonic Ritz vector  $\tilde{x}$ , rather than the harmonic Ritz value  $\mu + \tau$ , as an approximation to  $\lambda$  whenever the target  $\tau$  is close to the desired eigenvalue of  $A$ . Another merit of using  $\rho$  as an approximation is that the residual of  $(\rho, \tilde{x})$  is minimal over all scalars. Actually, this has been done so in some harmonic projection methods, e.g., [16, 21], though there it was unclear why the methods may miss the desired  $\lambda$  if the harmonic Ritz value  $\mu + \tau$  was used.

Theorems 4.1 and 4.2 show that in the Hermitian case the harmonic Ritz value  $\mu + \tau$  and the Rayleigh quotient  $\rho$  can converge quadratically when the harmonic Ritz vector only converges linearly.

We now present an example with varying  $\tau$  to illustrate our theory.

**Example 1.** Let  $A = \text{diag}(0, 5, -5)$  and suppose that we are required to compute the eigenpair  $(0, e_1)$ . Assume that the orthonormal basis  $W$  of  $\mathcal{W}$  is

$$W = \begin{pmatrix} 1 & 0 \\ 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{pmatrix}.$$

Then taking  $\tau = 0$  gives

$$B = W^H A^H W = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad C = W^H A^H A W = \begin{pmatrix} 0 & 0 \\ 0 & 25 \end{pmatrix}.$$

The eigenvector  $e_1$  is in  $\mathcal{W}$  exactly, so that  $\epsilon = 0$ . However, the harmonic projection method involves the eigenproblem of a singular matrix pencil  $(C, B)$ , which has all numbers as its eigenvalues. So the method fails completely.

Of course, in practice it is rare that  $\epsilon = 0$  and  $B$  is singular exactly. But nearly singular may be bad enough. We now perturb  $W$  by taking the element in position  $(2, 1)$  to be  $\eta = 10^{-6}$ . Then it can be justified  $\epsilon = 7.0682e - 7$ . Take  $\tau = \lambda = 0$ . Then  $|\lambda - \tau| = 0$  is much smaller than  $\epsilon$ . We have the orthonormalized  $W$ :

$$W = \begin{pmatrix} -0.99999999999950 & 0.00000070710678 \\ -0.00000100000000 & -0.70710678118602 \\ & 0 & -0.70710678118672 \end{pmatrix},$$

$$B = \begin{pmatrix} 5.0000e - 12 & 3.5355e - 6 \\ 3.5355e - 6 & -5.0013e - 12 \end{pmatrix}, \quad C = \begin{pmatrix} 2.5000e - 11 & 1.7678e - 5 \\ 1.7678e - 5 & 25.0000 \end{pmatrix}.$$

In this case,  $\|B^{-1}\| = 2.8284e + 5$ , and the harmonic Ritz values  $\mu_1 = 5.0000$ ,  $\mu_2 = -5.0000$  and the harmonic Ritz vectors  $\tilde{x}_1$  and  $\tilde{x}_2$  both are very close to  $e_1$ . Indeed,  $\sin \angle(e_1, \tilde{x}_1) = 7.0682e - 7 \approx \epsilon$  and  $\sin \angle(e_1, \tilde{x}_2) = 9.9982e - 7 \approx \epsilon$ . Therefore, none of the harmonic Ritz values is significant, and they are spurious. But two Rayleigh quotients  $\rho_1 = 5.0000e - 12$  and  $\rho_2 = -5.0000e - 12$  are very good approximations to  $\lambda = 0$ . Meanwhile,  $|\rho_1 - \lambda| = O(\epsilon^2)$  and  $|\rho_2 - \lambda| = O(\epsilon^2)$ . So the harmonic projection method misses  $\lambda$  if  $\tau$  is too close to the desired eigenvalue  $\lambda$ . Note that  $A$  is Hermitian. These results confirm our theoretical analysis.

The above  $\tau$  is taken to be exactly  $\lambda$ . We next take  $\tau = 1e - 10$ , so that  $|\lambda - \tau|$  is not zero but considerably smaller than  $\epsilon$ . In this case,  $\|B^{-1}\| = 2.8285e + 5$ , the harmonic Ritz values  $\mu_1 + \tau = -0.1249$ ,  $\mu_2 + \tau = 2.0012e + 2$ . So, still, both of them are spurious and insignificant. But

$$\sin \angle(e_1, \tilde{x}_1) = 7.0714e - 7 \quad \text{and} \quad \sin \angle(e_1, \tilde{x}_2) = 2.8311e - 5,$$

so both  $\tilde{x}_1$  and  $\tilde{x}_2$  are excellent approximations to  $e_1$ . The Rayleigh quotients  $\rho_1 = -1.2492e - 13$  and  $\rho_2 = 2.0012e - 10$  are very accurate approximations to  $\lambda$  with accuracy  $O(\epsilon^2)$ . Again, experiments confirm our theory.

If we take  $\tau = 1e - 7$ , which is of the same order as  $\epsilon$ , the situation becomes slightly different. The two harmonic Ritz values are still spurious, but unlike the above cases, there is only one harmonic Ritz vector  $\tilde{x}_1$  to approximate  $e_1$  with accuracy  $\sin \angle(e_1, \tilde{x}_1) = 7.0714e - 7 = O(\epsilon)$ . The corresponding Rayleigh quotient

$\rho_1 = -1.2490e - 16$  has error  $|\rho_1 - \lambda| = O(\epsilon^2)$ . So we can recover a very good approximation to  $\lambda = 0$  successfully.

If we take  $\tau = 1e - 2$ , so that  $|\lambda - \tau| = 1e - 2 \gg \epsilon$ , according to our theory the harmonic projection method should find an accurate harmonic Ritz value if  $\|B^{-1}\|$  is not large. Actually, we have  $\mu_1 + \tau = -1.2500e - 9$ , a good approximation to  $\lambda = 0$  with accuracy  $O(\epsilon)$ . The harmonic Ritz vector  $\tilde{x}_1$  is very good too, with  $\sin \angle(e_1, \tilde{x}_1) = 7.0714e - 7 = O(\epsilon)$ . However, the Rayleigh quotient  $\rho_1 = 1.0000e - 14$  is much more accurate than the harmonic Ritz value  $\mu_1 + \tau$ . If we take  $\tau = 1$ , for which  $|\lambda - \tau|$  is further away from  $\epsilon$ , we get a more accurate harmonic Ritz value  $\mu_1 + \tau = -1.2000e - 12$  and  $\tilde{x}_1$  with  $\sin \angle(e_1, \tilde{x}_1) = 7.2082e - 7 = O(\epsilon)$ . In this case, the Rayleigh quotient  $\rho_1 = 1.0000e - 12$  has about the same accuracy as the harmonic Ritz value  $\mu_1 + \tau$ .

Finally, it can be verified that in each case of the above the bound of (20) is very tight.

We now explain, from another viewpoint, why the harmonic method may fail to find  $\lambda$  when  $\tau$  is very close to  $\lambda$ . Suppose  $\tau = 0$ . Recall from (1) that the harmonic Ritz values of  $A$  with respect to  $\mathcal{W}$  are the reciprocals of the Ritz values of  $A^{-1}$  with respect to  $A\mathcal{W}$ . Now, if  $A$  has an eigenvalue  $\lambda$  very close to zero, the subspace  $A\mathcal{W}$  is almost deficient in the eigenvector  $x$  associated with  $\lambda$ , even though  $\epsilon = \sin \angle(x, \mathcal{W})$  approaches zero. Based on the convergence theory for the standard projection methods [3, 8, 12], it is known that generally there is no Ritz value of  $A^{-1}$  with respect to  $A\mathcal{W}$  that converges to  $1/\lambda$ , i.e., there is in general no harmonic Ritz value  $\mu$  of  $A$  with respect to  $\mathcal{W}$  that converges to  $\lambda$ . This explains why the harmonic projection method can miss the desired eigenvalue  $\lambda$  when it is very close to the target  $\tau$ .

## 5. CONVERGENCE OF REFINED HARMONIC RITZ VECTORS

Theorem 3.1 only requires that  $\nu$  be sufficiently near  $\lambda$ , and that  $\tilde{x}$  have a sufficiently small residual. For our concern, the harmonic projection methods, as Theorem 3.2 indicates, the harmonic Ritz vector  $\tilde{x}$  converges to the eigenvector  $x$  only conditionally as  $\epsilon \rightarrow 0$ . Under the assumption that  $\|B^{-1}\|$  is uniformly bounded and  $|\lambda - \tau|$  is not small, since the harmonic Ritz value  $\mu + \tau$  is known to converge to  $\lambda$ , this suggests that we retain the harmonic Ritz value and replace the harmonic Ritz vector with a vector  $\hat{x} \in \mathcal{W}$  having the smallest residual, as proposed in [4, 5, 7, 8, 9]. Thus we take  $\hat{x}$  to be the solution of the problem

$$(22) \quad \min_{x \in \mathcal{W}, \|x\|=1} \|(A - (\mu + \tau)I)x\|$$

Alternatively,  $\hat{x} = Wv$ , where  $v$  is the right singular vector of  $(A - (\mu + \tau)I)W$  corresponding to its smallest singular value. We will call such a vector a refined harmonic Ritz vector or, more generally, a refined eigenvector approximation. The method is referred to as a refined harmonic projection method. In particular, the refined harmonic Arnoldi method was proposed and an implicitly restarted refined harmonic Arnoldi algorithm was developed by the author in [10].

*Remark.* As we saw in the previous section, if  $\tau$  is close to the desired eigenvalue  $\lambda$ , then the harmonic projection method may not work well, and it will produce a spurious or insignificant harmonic Ritz value  $\mu + \tau$ . In this case, the method can miss  $\lambda$ . However, we have seen that for a general  $A$  the Rayleigh quotient  $\rho = \tilde{x}^* A \tilde{x}$

is a better and more reliable approximation to  $\lambda$ . If  $\mu + \tau$  is spurious, the refined harmonic Ritz vector  $\hat{x}$  must be an inaccurate approximation to  $x$ . Therefore, for reliability and convergence, it is advisable to compute  $\hat{x}$  satisfying (22) with  $\mu + \tau$  replaced by  $\rho$ .

The following theorem was proved in [12] for an *arbitrary* approximate eigenvalue  $\nu$ , so it works here for our concern.

**Theorem 5.1.** *Let  $L$  be defined as in (4). If*

$$(23) \quad \text{sep}(\nu, L) \geq \text{sep}(\lambda, L) - |\nu - \lambda| > 0,$$

*then*

$$(24) \quad \sin \angle(x, \hat{x}) \leq \frac{\|A - \nu I\| \epsilon + |\lambda - \nu|}{\sqrt{1 - \epsilon^2}(\text{sep}(\lambda, L) - |\nu - \lambda|)}.$$

Note that  $\text{sep}(\lambda, L)$  is a constant depending only on  $A$ . It follows immediately from (24) that if  $\nu \rightarrow \lambda$  then the vector  $\hat{x}$  converges to the eigenvector  $x$ . This means that  $\hat{x}$  and  $\nu$  have the same sufficient condition for convergence. In particular, comparing with Theorem 3.2, we see  $\hat{x}$  is more likely to converge than  $\tilde{x}$  if  $\nu = \mu + \tau \rightarrow \lambda$ . If  $\rho$  is used instead and it converges to  $\lambda$ , then we must have  $\hat{x} \rightarrow x$ .

We now present an example. It illustrates that the refined harmonic Ritz vector is uniquely defined and must be a good approximation to the desired eigenvector provided the approximate eigenvalue  $\nu$  is good and the space  $\mathcal{W}$  contains an accurate approximation to the desired eigenvector, while the harmonic projection method either cannot tell us how to define a proper harmonic Ritz vector or, although it is uniquely determined, is insignificant and has very poor accuracy.

**Example 2.** Let  $A = \text{diag}(1, 0, 1)$ , and suppose we want to compute the eigenpair  $(1, e_1)$ . Assume that we come up with an orthonormal basis  $W$  of  $\mathcal{W}$ :

$$W = \begin{pmatrix} 1 & 0 \\ 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{pmatrix}.$$

The desired eigenvector  $e_1$  is exactly in  $\mathcal{W}$ . Taking  $\tau = 0$ , we get

$$B = C = \begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}.$$

Therefore,  $B^{-1}C = I$ , the  $2 \times 2$  identity matrix. We obtain a double multiple harmonic Ritz value one, which is exactly equal to the desired eigenvalue. However, any nonzero two-dimensional vector  $z$  is an eigenvector of  $B^{-1}C$  corresponding to the eigenvalue one. Since  $B^{-1}C$  itself gives no clue as to which vector to choose, we might take  $z = (1/\sqrt{2}, 1/\sqrt{2})^T$ . In this case, our approximate eigenvector by the harmonic projection method is

$$\tilde{x} = (1/\sqrt{2}, 1/2, 1/2)^T,$$

which is far from the desired eigenvector  $e_1$  and cannot be considered to be a significant approximation to  $e_1$  in any sense. So, the harmonic projection can fail completely even though the space  $\mathcal{W}$  contains the desired eigenvector. This is due to the double multiple eigenvalue of  $B^{-1}C$ . In fact, the harmonic Ritz vector  $\tilde{x}$  will be a poor approximation to  $x$  if  $\mu$  is close to some other eigenvalues of  $B^{-1}C$ , as Theorem 3.2 indicates. Actually, it cannot be uniquely defined if  $\mu$  is a multiple

but nondefective eigenvalue of  $B^{-1}C$ . More generally, as Theorem 3.2 indicates, if  $\mu$  is close to some other eigenvalues of  $B^{-1}C$ , then the harmonic Ritz vector is uniquely defined but may have very poor accuracy and can be an insignificant approximation to  $x$  although  $\epsilon$  is very small.

In contrast, our refined harmonic projection method works well and fixes this problem successfully. Taking  $\mu = 1$ , the refined harmonic Ritz vector  $\hat{x}$  is uniquely determined and is exactly equal to  $e_1$  by solving the problem (22). Furthermore, the residual norm of refined harmonic Ritz pair  $(\mu, \hat{x})$  is zero.

We now perturb  $W$  by taking the elements in positions (2,1) and (1,2) to be  $\eta = 1e - 6$ . Then we have the orthonormalized

$$W = \begin{pmatrix} -0.9999999999950 & 0.00000070710678 \\ -0.00000100000000 & -0.70710678118602 \\ & 0 & -0.70710678118672 \end{pmatrix}$$

and  $\epsilon = \sin \angle(e_1, \mathcal{W}) = 7.0682e - 7$ .

The harmonic projection method gives

$$B = \begin{pmatrix} 1.0000000000011 & 0.00000032687900 \\ 0.00000032687900 & 2.00000036543901 \end{pmatrix},$$

$$C = \begin{pmatrix} 0.99999999999900 & -0.00000070710678 \\ -0.00000070710678 & 0.50000000000146 \end{pmatrix}, \quad B^{-1}C = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}.$$

The harmonic Ritz values, the eigenvalues of  $B^{-1}C$ , are  $\mu_1 = \mu_2 = 1$ , which are *exactly* multiple in finite precision. Still, any two-dimensional vector  $z$  is an eigenvector of  $B^{-1}C$  corresponding to the multiple eigenvalue. So the harmonic projection method fails even though  $\mathcal{W}$  contains a very accurate approximation to  $e_1$  with  $\epsilon = 7.0682e - 7$ . However, if we solve the problem (22) using  $\nu = \mu_1$ , we get that the right singular vector of  $AU - U$  associated with the smallest singular value zero is  $v = (-0.99999999999900, 0.00000141421356)^T$  and the refined harmonic Ritz vector is uniquely given by

$$\hat{x} = \begin{pmatrix} 0.99999999999950 \\ 0 \\ -0.00000100000000 \end{pmatrix},$$

which is an excellent approximation to  $e_1$  with  $\sin \angle(e_1, \hat{x}) = 9.9982e - 7$ . Now, on the one hand, naturally  $\epsilon \leq \sin \angle(e_1, \hat{x})$ ; on the other hand, note that the upper bound of (24) is  $\epsilon/\sqrt{1 - \epsilon^2} = 7.0682e - 7$ . So our bound on  $\sin \angle(e_1, \hat{x})$  is very sharp.

As a by-product, we found that  $B^{-1}C = I$  always holds in finite precision, no matter how  $\eta$ , a perturbation in the entries of positions (2,1) and (1,2) of the original orthonormal basis  $W$ , is. So, we can always get an exact approximation to the eigenvalue one of  $A$ , but at the same time we are always in trouble when using the harmonic projection method to determine and compute a good harmonic Ritz vector. In contrast, the refined harmonic Ritz vector is always uniquely defined and is a very good approximation to the desired eigenvector when  $\epsilon$  is small.

## 6. CONCLUDING REMARKS

The results of the last four sections have nice and flexible applications. Suppose we have a harmonic method that produces subspaces  $\mathcal{W}$  containing increasingly accurate approximations to our desired eigenvector  $x$ . Corollary 2.2 says that there

is a harmonic Ritz value  $\mu + \tau$  that converges to our distinguished eigenvalue  $\lambda$  if  $\|B^{-1}\|$  is uniformly bounded. Theorem 3.1 says that if the residual  $A\tilde{x} - (\mu + \tau)\tilde{x}$  of the harmonic Ritz vector  $\tilde{x}$  converges to zero, then the harmonic Ritz vector itself converges to  $x$ . However, Theorem 3.2 shows that the harmonic Ritz vector  $\tilde{x}$  can fail to converge to  $x$  if  $\mu$  is arbitrarily close to some other eigenvalues of  $B^{-1}C$ . On the other hand, the harmonic method may produce a spurious harmonic Ritz value and fail to compute  $\lambda$ —in other words, it can miss  $\lambda$  if  $\tau$  is very close to  $\lambda$ . This is in accordance with Corollary 2.2, as the smaller  $|\lambda - \tau|$ , the larger  $\|B^{-1}\|$ . More precisely, if  $\tilde{x}$  converges to  $x$ , Theorem 4.1 says that  $\mu + \tau$  converges to  $\lambda$  if  $\lambda$  is not very close to  $\tau$ , while Theorem 4.2 shows that the Rayleigh quotient  $\rho$  is a better and reliable alternative to approximate  $\lambda$ , and it converges to  $\lambda$  without requiring that  $\tau$  not be very close to  $\lambda$ . Theorem 4.2 suggests that we take the more accurate and reliable Rayleigh quotient  $\rho$  in (22) so that we get a better and more reliable refined eigenvector approximation. If the residual of  $(\mu + \tau, \tilde{x})$  or that of  $(\rho, \tilde{x})$  stagnates, we can switch to a refined harmonic Ritz vector, in which case Theorem 5.1 assures its convergence to  $x$ . We illustrated some of our theory by two examples.

A nice feature of our bounds is that  $\lambda$  is linked to the other eigenvalues of  $A$  only by the quantity  $\text{sep}(\lambda, L)$ , and  $\mu$  is linked to the other eigenvalues of the harmonic Rayleigh quotient  $B^{-1}C$  only by the quantity  $\text{sep}(\mu, G)$ . In particular, this implies that we do not have to assume that the other eigenvalues are simple—i.e., that the matrix in question is diagonalizable.

As is known from [3, 8, 12], the convergence of a Ritz value and its refined Ritz vector only requires that  $\epsilon$  tend to zero, independently of  $B$ . But in the harmonic and refined harmonic projection methods, it is essential to assume that  $\|B^{-1}\|$  is uniformly bounded, or in other words  $B$  is uniformly nonsingular as  $\epsilon \rightarrow 0$ , which surely holds only for  $A - \tau I$  Hermitian positive or negative definite. If  $\tau$  is very close to  $\lambda$ , then  $\|B^{-1}\|$  must be large as  $\epsilon \rightarrow 0$ , as Theorem 4.3 indicates. So we must take a  $\tau$  not very close to  $\lambda$ . In any event, we should take a better and more reliable Rayleigh quotient  $\rho$  as an approximation to  $\lambda$ . Thus, the harmonic projection method may be inferior to its corresponding standard projection method, and accordingly this holds for the refined harmonic projection method and the refined projection method.

#### ACKNOWLEDGMENTS

I thank the anonymous referee very much for his/her valuable comments and suggestions, which helped me think over this topic in greater depth and improve on the presentation considerably.

#### REFERENCES

- [1] R. W. Freund. *Quasi-kernel polynomials and their use in non-Hermitian matrix iterations*, J. Comput. Appl. Math., 43 (1992): 135–158. MR 94a:65018
- [2] G. H. Golub and C. F. van Loan. *Matrix Computations*, the 3rd ed., The Johns Hopkins University Press, Baltimore and London, 1996. MR 97g:65006
- [3] Z. Jia. *The convergence of generalized Lanczos methods for large unsymmetric eigenproblems*, SIAM J. Matrix Anal. Appl., 16 (1995): 843–862. MR 96d:65062
- [4] Z. Jia. *Refined iterative algorithms based on Arnoldi's process for large unsymmetric eigenproblems*, Linear Algebra Appl., 259 (1997): 1–23. MR 98c:65060
- [5] Z. Jia. *A refined iterative algorithm based on the block Arnoldi process for large unsymmetric eigenproblems*, Linear Algebra Appl., 270 (1998): 171–189. MR 98m:65055

- [6] Z. Jia. *Generalized block Lanczos methods for large unsymmetric eigenproblems*, Numer. Math., 80 (1998): 239–266. MR 99f:65059
- [7] Z. Jia. *Polynomial characterizations of the approximate eigenvectors by the refined Arnoldi method and an implicitly restarted refined Arnoldi algorithm*, Linear Algebra Appl., 287 (1999): 191–214. MR 99j:65046
- [8] Z. Jia. *Composite orthogonal projection methods for large matrix eigenproblems*, Science in China, Ser.A, 42 (1999): 577–585. MR 2000b:65056
- [9] Z. Jia. *A refined subspace iteration algorithm for large sparse eigenproblems*, Appl. Numer. Math., 32 (2000): 35–52. MR 2000j:65045
- [10] Z. Jia. *The refined harmonic Arnoldi method and an implicitly restarted refined algorithm for computing interior eigenpairs of large matrices*, Appl. Numer. Math., 42 (2002): 489–512. MR 2003g:65049
- [11] Z. Jia. *Some theoretical comparisons of refined Ritz vectors and Ritz vectors*, Science in China, Ser.A, 47 (Supp.) (2004): 222–233.
- [12] Z. Jia and G. W. Stewart. *An analysis of the Rayleigh–Ritz method for approximating eigenspaces*, Math. Comput., 70 (2001): 637–647. MR 2001g:65040
- [13] Z. Jia and D. Niu. *An implicitly restarted refined bidiagonalization Lanczos algorithm for computing a partial singular value decomposition*, SIAM J. Matrix Anal. Appl., 25 (2003): 246–265.
- [14] T. A. Manteuffel and J. Otto. *On the roots of orthogonal polynomials and residual polynomials associated with a conjugate gradient method*, Numer. Linear Algebra Appl., 1 (1994): 449–475. MR 95i:65051
- [15] R. B. Morgan. *Computing interior eigenvalues of large matrices*, Linear Algebra Appl., 154/156 (1991): 289–309. MR 92e:65050
- [16] R. B. Morgan and M. Zeng. *Harmonic projection methods for large nonsymmetric eigenvalue problems*, Numer. Linear Algebra Appl., 5 (1998): 35–55. MR 99g:65046
- [17] C. C. Paige, B. N. Parlett and H. A. van der Vorst. *Approximate solutions and eigenvalue bounds from Krylov subspaces*, Numer. Linear Algebra Appl., 2 (1995): 115–134. MR 96a:65054
- [18] B. N. Parlett. *The Symmetric Eigenvalue Problem*, SIAM, Philadelphia, PA, 1998. MR 99c:65072
- [19] Y. Saad. *Numerical Methods for Large Eigenvalue Problems*, Manchester University Press, 1992. MR 93h:65052
- [20] G. L. G. Sleijpen and H. A. van der Vorst. *A Jacobi-Davidson iteration method for linear eigenvalue problems*, SIAM J. Matrix Anal. Appl., 17 (1996): 401–425. MR 96m:65042
- [21] G. L. G. Sleijpen, H. A. van der Vorst and E. Meijerink. *Efficient expansion of subspaces in the Jacobi-Davidson method for standard and generalized eigenproblems*, Electr. Trans. Numer. Anal., 7 (1998): 75–89. MR 99i:65040
- [22] G. W. Stewart. *Matrix Algorithms: Vol. II Eigensystems*, SIAM, Philadelphia, PA, 2001.
- [23] G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory*, Academic Press, New York, 1990. MR 92a:65017
- [24] H. A. van der Vorst. *Computational Methods for Large Eigenvalue Problems*, In P.G. Ciarlet and J.L. Lions (eds), Handbook of Numerical Analysis, Volume VIII, North-Holland (Elsevier), Amsterdam, 2002, pp. 3–179. MR 2003e:65058

DEPARTMENT OF MATHEMATICAL SCIENCES, TSINGHUA UNIVERSITY, BEIJING 100084, PEOPLES  
REPUBLIC OF CHINA

*E-mail address:* zjia@math.tsinghua.edu.cn