



ELSEVIER

Linear Algebra and its Applications 358 (2003) 95–114

---

---

LINEAR ALGEBRA  
AND ITS  
APPLICATIONS

---

---

[www.elsevier.com/locate/laa](http://www.elsevier.com/locate/laa)

# A geometric theory for preconditioned inverse iteration III: A short and sharp convergence estimate for generalized eigenvalue problems

Andrew V. Knyazev<sup>a,\*</sup>, Klaus Neymeyr<sup>b</sup>

<sup>a</sup>*Department of Mathematics, Campus Box 170, University of Colorado at Denver, P.O. Box 173364, Denver, CO 80217-3364, USA*

<sup>b</sup>*Mathematisches Institut der Universität Tübingen, Auf der Morgenstelle 10, D-72076 Tübingen, Germany*

Received 20 April 2001; accepted 23 July 2001

Submitted by B.N. Parlett

---

## Abstract

In two previous papers by Neymeyr [Linear Algebra Appl. 322 (1–3) (2001) 61; 322 (1–3) (2001) 87], a sharp, but cumbersome, convergence rate estimate was proved for a simple preconditioned eigensolver, which computes the smallest eigenvalue together with the corresponding eigenvector of a symmetric positive definite matrix, using a preconditioned gradient minimization of the Rayleigh quotient. In the present paper, we discover and prove a much shorter and more elegant (but still sharp in decisive quantities) convergence rate estimate of the same method that also holds for a generalized symmetric definite eigenvalue problem. The new estimate is simple enough to stimulate a search for a more straightforward proof technique that could be helpful to investigate such a practically important method as the locally optimal block preconditioned conjugate gradient eigensolver.

© 2002 Elsevier Science Inc. All rights reserved.

*AMS classification:* 65F15

*Keywords:* Symmetric generalized eigenvalue problem; Preconditioning; Preconditioned eigensolver; Gradient; Steepest descent; Conjugate gradient; Matrix-free; Rayleigh; Ritz; Davidson; Eigenvector; Iterative method

---

\* Corresponding author.

*E-mail addresses:* [andrew.knyazev@cudenver.edu](mailto:andrew.knyazev@cudenver.edu) (A.V. Knyazev), [neymeyr@na.uni-tuebingen.de](mailto:neymeyr@na.uni-tuebingen.de) (K. Neymeyr).

*URL:* [www:http://www-math.cudenver.edu/~aknyazev](http://www-math.cudenver.edu/~aknyazev) (A.V. Knyazev).

## 1. Introduction

Let  $A$  and  $T$  be real symmetric positive definite  $n$ -by- $n$  matrices. We consider the problem of computing the smallest eigenvalue  $\lambda_1$  and the corresponding eigenvector  $u_1$  of matrix  $A$  by preconditioned iterative methods, where  $T$  will play the role of the preconditioner, e.g., [21]. Such eigensolvers are matrix-free, i.e., neither  $A$ , nor the preconditioner  $T$  needs to be available as a matrix, and are designed to solve efficiently and accurately extremely large and ill-conditioned eigenvalue problems.

Let  $\|\cdot\|_A$  denote the  $A$ -based vector norm  $\|\cdot\|_A = (\cdot, A\cdot)$  as well as the corresponding induced operator norm. For our theoretical estimates, we assume that the preconditioner  $T$  approximates the matrix  $A$  such that

$$\|I - T^{-1}A\|_A \leq \gamma, \quad 0 \leq \gamma < 1. \quad (1)$$

In general, as both matrices  $A$  and  $T$  are symmetric positive definite, the following always holds:

$$\delta_0(u, Tu) \leq (u, Au) \leq \delta_1(u, Tu), \quad 0 < \delta_0 \leq \delta_1. \quad (2)$$

The ratio  $\delta_1/\delta_0$  can be viewed as the spectral condition number  $\kappa(T^{-1}A)$  of the preconditioned matrix  $T^{-1}A$  and measures how well the preconditioner  $T$  approximates, up to a scaling, the matrix  $A$ . A smaller ratio  $\delta_1/\delta_0$  typically ensures faster convergence. For mesh problems, matrices  $A$  and  $T$  are called *spectrally equivalent* if the ratio is bounded from above uniformly in the mesh size, see [9].

Assumption (1) leads to (2) with  $\delta_0 = 1 - \gamma$  and  $\delta_1 = 1 + \gamma$ . Vice versa, assumption (2) leads to (1), but only if  $T$  is properly scaled. Namely, if  $T$  satisfies (2), then optimally scaled  $2T/(\delta_0 + \delta_1)$  (replacing the original  $T$ ) satisfies (1) with

$$\gamma = \frac{\kappa(T^{-1}A) - 1}{\kappa(T^{-1}A) + 1}. \quad (3)$$

Our convergence estimates in the present paper for methods with optimal scaling will be based on assumption (2) and will use  $\gamma$  given by (3). We note that some preconditioned eigensolvers, e.g., the steepest descent method we will discuss later, implicitly provide the optimal scaling of the preconditioner. In the rest of the paper, we will assume (1), unless explicitly stated otherwise, in order to be consistent with the previous papers by Neymeyr [31,32].

It is well known that the minimum of the Rayleigh quotient

$$\lambda(u) = \frac{(u, Au)}{(u, u)}, \quad \text{where } u \in \mathbb{R}^n, \quad u \neq 0, \quad (4)$$

is  $\lambda_1$  and the corresponding stationary point is the eigenvector  $u_1$  of  $A$ . Gradient preconditioned eigensolvers generate a sequence of nonzero vectors, which minimizes the Rayleigh quotient, using its gradient, computed in the  $T$ -based scalar product  $(\cdot, \cdot)_T = (\cdot, T\cdot)$ , see, e.g., [9]

$$\nabla_T \lambda(u) = \frac{2}{(u, u)_T} T^{-1}(Au - \lambda(u)u). \quad (5)$$

The simplest method in this class, a two-term gradient minimization, can be written as

$$u^{(i+1)} = u^{(i)} - \omega^{(i)} T^{-1}(Au^{(i)} - \lambda(u^{(i)})u^{(i)}), \quad (6)$$

where  $\omega^{(i)}$  is a scalar step size. Our main goal is to analyze the error reduction of one step of the method,

$$u' = u - \omega T^{-1}(Au - \lambda u), \quad (7)$$

where we discard upper indexes and denote  $u' = u^{(i+1)}$ ,  $u = u^{(i)}$ ,  $\omega = \omega^{(i)}$ , and  $\lambda = \lambda(u^{(i)})$ .

We will consider two choices of  $\omega$  here. The first case is an a priori fixed choice  $\omega = 1$ . This choice is evidently affected by a preconditioner scaling.

The second choice corresponds to the well-known, e.g., [9,21], preconditioned steepest descent for the Rayleigh quotient, where  $\omega$  is chosen to minimize the Rayleigh quotient on the two-dimensional subspace  $\text{span}\{u, T^{-1}(Au - \lambda u)\}$  by means of the Rayleigh–Ritz method. This leads to a 2-by-2 generalized eigenvalue problem that can be solved explicitly by using formulas for roots of the corresponding characteristic equation, which is in this case quadratic. Interestingly, the optimal  $\omega$  in general may be negative, or even infinite, see [18]; in the latter case,  $u' = T^{-1}(Au - \lambda u)$ . We emphasize again that such a choice of  $\omega$  implicitly determines the optimal preconditioner scaling constant; thus (3) can be used in convergence rate estimates in this case.

In the following section, we provide summaries of some pioneering and recent results and give an informal description of the main results of the present paper in the context of the previous work.

## 2. A short survey

The trivial choice  $T = I$  of the preconditioner, see [14,15], suffers from poor convergence for ill-conditioned matrices, cf. [4,12,25,39,46]. Preconditioned gradient methods with a general preconditioner  $T$  for symmetric eigenvalue problem have been studied, e.g., by Samokish [42], Petryshyn [38], Godunov et al. [13], D'yakonov and Orekhov [10], D'yakonov [8], Knyazev [17,20] as well as in the monograph [9] and in a recent survey [21], which include extensive bibliography. Such preconditioned eigensolvers have been used in practice, e.g., for band structure calculations [6,7], thin elastic structures [37], and a real-space ab initio method for electronic structure calculations in terms of nonorthogonal orbitals defined on a grid [11]. In the latter paper, a multigrid preconditioner is employed to improve the steepest descent directions used in the iterative minimization of the energy functional.

Preconditioned steepest descent is an obvious way to accelerate the convergence of the basic preconditioned eigensolver (7) with  $\omega = 1$ . There are several more elaborate algorithms, e.g., the recent successive eigenvalue relaxation method of

Ovtchinnikov and Xanthis [36], and preconditioned conjugate gradient algorithms for minimizing the Rayleigh quotient, using an approximate inverse preconditioner, see a recent paper [3] and references there.

The most promising, according to Knyazev [22], Knyazev and Neymeyr [23], preconditioned eigensolver is the locally optimal block preconditioned conjugate gradient (LOBPCG) method suggested and analyzed in [19–22]. For computing the first eigenpair, the new iterate in LOBPCG is determined by the Rayleigh–Ritz method on a three-dimensional subspace, which includes the previous iterate in addition to the current iterate and the preconditioned residual of the two-dimensional trial subspace of the steepest descent method. The LOBPCG converges many times faster than the steepest descent in numerical tests, and is argued in [22] to be practically the optimal method on the whole class of preconditioned eigensolvers. However, no simple comprehensive convergence theory of the LOBPCG, explaining its apparent optimality, is yet known. The reason is that deriving sharp convergence estimates is challenging even for simplest preconditioned eigensolvers such as that described by (7).

While an apparently sharp asymptotic convergence rate estimate for the preconditioned steepest descent method appeared in the very first paper [42], a sharp non-asymptotic convergence rate estimate is not yet known despite major efforts over the decades; see [20] for the review and references. For a simpler method, namely, (7) with  $\omega = 1$ , a sharp nonasymptotic convergence rate estimate has been proved only recently in [31,32]. There, Neymeyr interpreted a preconditioned gradient method with a fixed step size as a *perturbation of a well-known inverse iteration method*, in such a way that the associated system of linear equations was solved approximately by using a *preconditioner*. To highlight this, the method (7) with  $\omega = 1$  was called the *Preconditioned INverse ITERation* (PINVIT). A simple geometric interpretation of the method was discovered that provided a basis for derivation of sharp convergence estimates by Neymeyr [32].

The estimate of Neymeyr [31,32] is sharp, but too cumbersome for a human being. In the present paper, we discover and prove a much shorter and elegant, but still sharp, convergence rate estimate for the same method. The new estimate also holds for a generalized symmetric definite eigenvalue problem. It is simple enough to stimulate a search for a more straightforward proof technique that might finally lead to considerable progress in theory of practically important methods such as LOBPCG [22].

The new convergence factor we obtain here is roughly a square of the one previously derived in [8–10] for the same method in the same terms.

There are several preconditioned eigensolvers, similar to classical subspace iterations, for computing an invariant subspace spanned by a group of eigenvectors corresponding to several smallest eigenvalues of  $A$ ; see, e.g., [5,21,26,27,47] and, for trace minimization methods, see [1,41] and references therein.

In [30], the sharp convergence rate estimate of Neymeyr [31,32] for single-vector preconditioned solver is generalized to cover similar subspace iterations. A sharp

simplification of the estimate of Neymeyr [30] is suggested in [22], but the proof is sketchy and not complete. In the present paper, we fill these gaps in the arguments of Knyazev [22].

Let us finally mention here briefly a number of very recent articles on preconditioned eigensolvers, even though they are not closely related to the subject of the present paper.

Oliveira [35] obtains asymptotic convergence rate estimate of the generalized Davidson method similar to that by Samokish [42] for the preconditioned steepest descent. Sadkane and Sidje [40] discuss the block Davidson method with deflation. Smit and Paardekooper [45] study inexact inverse and Rayleigh quotient iterations, using a perturbation technique somewhat comparable with that used in [31,32], but explicitly based on the error reduction rate of the inner iterations. Basermann [2] applies a block incomplete LU decomposition for preconditioning in the Jacobi–Davidson method JDQR [1,44]. Ng [33] uses for Toeplitz matrices the preconditioned Lanczos method suggested and analyzed in [17,29,43], see also [1]. Morgan [28] gives numerical comparisons of JDQR with the generalized Davidson method, the preconditioned Lanczos methods, and the inexact Rayleigh quotient iterations. Notay [34] suggests a new Jacobi–Davidson type-method, tuned for the symmetric case, called JDCG, and compares it numerically with the original JDQR code and with the revision 3.2 of the LOBPCG code.

The rest of the paper is organized as follows. In Section 3, we derive a new simple and sharp convergence estimate for the PINVIT. Furthermore, we derive an upper estimate for the convergence of preconditioned steepest descent. We extend these results to generalized symmetric definite eigenproblems in Section 4. In Section 5, we present similar convergence estimates for preconditioned subspace iterations.

Numerical results were given in the last section in the preliminary version of the paper published as a technical report [24]. There, we described a bug in revision 3.2 of the LOBPCG code, which was found by analyzing the LOBPCG CPU timings presented in [34], and gave new numbers for revision 3.3 of the LOBPCG, showing a clear superiority of the LOBPCG over JDCG and JDQR, using the same preconditioner, based on an incomplete (with a drop tolerance) Cholesky factorization of the stiffness matrix. Numerical results for the multigrid preconditioning can be found in [23].

### 3. Preconditioned inverse iteration

According to formula (1.5) of Theorem 1.1 in [32], the sharp estimate from above for the Rayleigh quotient of  $u'$ , computed by (7) with  $\omega = 1$  is the following lengthy and, therefore, somewhat unreadable result: if  $\lambda = \lambda(u) \in [\lambda_k, \lambda_{k+1}[$ , then

$$\lambda' = \lambda(u') \leq \lambda_{k,k+1}(\lambda, \gamma), \tag{8}$$

$$\begin{aligned}
& \lambda_{k,k+1}(\lambda, \gamma) \\
&= \lambda \lambda_k \lambda_{k+1} (\lambda_k + \lambda_{k+1} - \lambda)^2 \\
&\quad \times \left( \gamma^2 (\lambda_{k+1} - \lambda) (\lambda - \lambda_k) (\lambda \lambda_{k+1} + \lambda \lambda_k - \lambda_k^2 - \lambda_{k+1}^2) \right. \\
&\quad - 2\gamma \sqrt{\lambda_k \lambda_{k+1}} (\lambda - \lambda_k) (\lambda_{k+1} - \lambda) \\
&\quad \times \sqrt{\lambda_k \lambda_{k+1} + (1 - \gamma^2) (\lambda - \lambda_k) (\lambda_{k+1} - \lambda)} \\
&\quad \left. - \lambda (\lambda_k + \lambda_{k+1} - \lambda) (\lambda \lambda_{k+1} + \lambda \lambda_k - \lambda_k^2 - \lambda_k \lambda_{k+1} - \lambda_{k+1}^2) \right)^{-1}, \quad (9)
\end{aligned}$$

see the theorem below for the exact meaning of notations.

The estimate (8) is sharp in a sense that a preconditioner  $T$  and a vector  $u$  can be found such that the bound for the Rayleigh quotient is attained. Here, we present a concise convergence rate estimate for PINVIT, written in different terms, which is also sharp, but in a different, somewhat weaker sense; see Remark 3 below.

**Theorem 1.** *Let  $u \in \mathbb{R}^n$  and let  $\lambda = \lambda(u) \in [\lambda_1, \lambda_n[$  be its Rayleigh quotient, where  $\lambda_1 \leq \dots \leq \lambda_n$  are the eigenvalues of  $A$ . The preconditioner is assumed to satisfy (1) for some  $\gamma \in [0, 1[$ . If  $\lambda = \lambda(u) \in [\lambda_k, \lambda_{k+1}[$ , then it holds for the Rayleigh quotient  $\lambda' = \lambda(u')$  with  $u'$  computed by (7) with  $\omega = 1$  that either  $\lambda' < \lambda_k$  (unless  $k = 1$ ), or  $\lambda' \in [\lambda_k, \lambda[$ . In the latter case,*

$$\frac{\lambda' - \lambda_k}{\lambda_{k+1} - \lambda'} \leq (q(\gamma, \lambda_k, \lambda_{k+1}))^2 \frac{\lambda - \lambda_k}{\lambda_{k+1} - \lambda}, \quad (10)$$

where

$$q(\gamma, \lambda_k, \lambda_{k+1}) = \gamma + (1 - \gamma) \frac{\lambda_k}{\lambda_{k+1}} = 1 - (1 - \gamma) \left( 1 - \frac{\lambda_k}{\lambda_{k+1}} \right) \quad (11)$$

is the convergence factor.

**Proof.** Evidently, having the estimate (8), we only need to show that the maximum for all  $\lambda \in [\lambda_k, \lambda_{k+1}[$  of the function

$$\frac{\lambda_{k,k+1}(\lambda, \gamma) - \lambda_k}{\lambda_{k+1} - \lambda_{k,k+1}(\lambda, \gamma)} \frac{\lambda_{k+1} - \lambda}{\lambda - \lambda_k}, \quad (12)$$

where  $\lambda_{k,k+1}(\lambda, \gamma)$  is explicitly given in (9), is exactly  $(q(\gamma, \lambda_k, \lambda_{k+1}))^2$ . It is easy to check that the function takes this value, when  $\lambda = \lambda_k$ , however we are not able to find a simple proof that it is the maximal value, using the expression for  $\lambda_{k,k+1}(\lambda, \gamma)$  from (9). Instead, we will use a different, though equivalent, representation of  $\lambda_{k,k+1}(\lambda, \gamma)$  from Theorem 1.1 in [32], which provides the “mini-dimensional analysis” in  $\text{span}\{u_k, u_{k+1}\}$ , see also Theorem 5.1 in [31]. We adopt the notations of the latter theorem and set for convenience  $k = 1$  and  $k + 1 = 2$ , without a loss of generality.

It is shown in [31] that the set of all iterates  $E_\gamma$ , when one fixes the vector  $u$  and chooses all preconditioners  $T$  satisfying (1), is a ball, in the  $A$ -based scalar product. In the two-dimensional subspace  $\text{span}\{u_k, u_{k+1}\}$ , the intersection  $\text{span}\{u_k, u_{k+1}\} \cap E_\gamma$  is a disk. The quantity  $r$  will denote the radius of the disk, and  $y$  and  $x$  will be Cartesian coordinates of its center with respect to a Cartesian system of coordinates, given by the  $A$ -orthonormal eigenvectors  $u_1$  and  $u_2$  of  $A$ , which span  $\text{span}\{u_k, u_{k+1}\}$ , correspondingly. Neymeyr [31] obtains the following formulas:

$$x = \sqrt{\frac{\lambda(\lambda - \lambda_1)}{\lambda_2(\lambda_2 - \lambda_1)}}, \quad y = \sqrt{\frac{\lambda(\lambda_2 - \lambda)}{\lambda_1(\lambda_2 - \lambda_1)}}, \quad r = \gamma \sqrt{\frac{(\lambda - \lambda_1)(\lambda_2 - \lambda)}{\lambda_1 \lambda_2}}.$$

According to Neymeyr [31], the unique maximum of the Rayleigh quotient on the whole  $E_\gamma$  is actually attained on the disc  $\text{span}\{u_k, u_{k+1}\} \cap E_\gamma$  and is given by  $\lambda_{1,2}(\lambda, \gamma)$  defined by formula (5.6) of [31], reproduced here:

$$\lambda_{1,2}(\lambda, \gamma) = \frac{\eta^2 + \xi^2}{\eta^2/\lambda_1 + \xi^2/\lambda_2}, \tag{13}$$

where

$$(\eta, \xi) = \left( \sqrt{l^2 - \xi^2}, \frac{x l^2 + r y l}{x^2 + y^2} \right)$$

are the coordinates of the point of the maximum and  $l$  is its Euclidean norm; moreover,

$$l = \sqrt{x^2 + y^2 - r^2}.$$

Formula (9) is then derived from (13) in [31].

For our present proof, the geometric meaning of quantities is not at all important. The only important fact is that (13) provides a formula for  $\lambda_{1,2}(\lambda, \gamma)$  for known  $x, y$  and  $r$ , which, in their turn, are explicitly given as functions of  $\gamma, \lambda, \lambda_1$ , and  $\lambda_2$  only. The rest of the proof is nothing but simple, though somewhat tedious, algebraic manipulations.

Directly from (13), we have

$$\frac{\lambda_{12} - \lambda_1}{\lambda_2 - \lambda_{12}} = \frac{\xi^2 \lambda_1}{\eta^2 \lambda_2} = \frac{\lambda_1}{\lambda_2} \frac{(x l + r y)^2}{(x^2 + y^2)^2 - (x l + r y)^2},$$

where in the denominator

$$(x^2 + y^2)^2 - (x l + r y)^2 = (y l - x r)^2.$$

Here,  $y l - x r$  is positive because of  $y > r$ .

Explicit expressions for  $x$  and  $y$  give

$$\frac{\lambda_2 - \lambda}{\lambda - \lambda_1} = \frac{y^2 \lambda_1}{x^2 \lambda_2}.$$

Therefore, the convergence factor, defined by

$$\frac{\lambda_{12} - \lambda_1}{\lambda_2 - \lambda_{12}} \frac{\lambda_2 - \lambda}{\lambda - \lambda_1} = \frac{\lambda_1^2 y^2}{\lambda_2^2 x^2} \frac{(xl + ry)^2}{(yl - rx)^2} =: q^2[\lambda],$$

is equal to

$$q[\lambda] = \frac{\lambda_1 y (xl + ry)}{\lambda_2 x (yl - rx)} = \frac{\lambda_1}{\lambda_2} \frac{1 + yr/xl}{1 - xr/yl} > 0. \quad (14)$$

Direct computation shows that

$$\frac{yr}{xl} = \gamma(\lambda_2 - \lambda) \left( \frac{\lambda_2}{\lambda_1} \right)^{1/2} z^{-1/2}$$

and

$$\frac{xr}{yl} = \gamma(\lambda - \lambda_1) \left( \frac{\lambda_1}{\lambda_2} \right)^{1/2} z^{-1/2}$$

with  $z = \gamma^2(\lambda_1 - \lambda)(\lambda_2 - \lambda) + \lambda(\lambda_1 + \lambda_2 - \lambda) > 0$ . Hence,

$$q[\lambda] = \frac{\sqrt{\lambda_1/\lambda_2} z^{1/2} + \gamma(\lambda_2 - \lambda)}{\sqrt{\lambda_2/\lambda_1} z^{1/2} - \gamma(\lambda - \lambda_1)}. \quad (15)$$

We note again that value of  $q[\lambda]$  squared in (15) must be the same as that of expression (12) with  $\lambda_{1,2}(\lambda, \gamma)$  given by (9)—it is just written in a more civilized way.

We now want to eliminate dependence of the convergence factor  $q[\lambda]$  on  $\lambda$ , by finding a sharp upper bound  $q$ , independent of  $\lambda$ . For that, let us show

$$q'[\lambda] < 0,$$

which is equivalent to

$$\gamma \sqrt{\lambda_1 \lambda_2} (\lambda_2 - \lambda_1) < (\lambda_2 - \lambda_1) z^{1/2} + \left( \frac{d}{d\lambda} z^{1/2} \right) \{ \lambda_2 (\lambda_2 - \lambda) + \lambda_1 (\lambda - \lambda_1) \}.$$

Taking the square of both sides and inserting  $z$  and  $(d/d\lambda)z^{1/2}$ , we observe after factorization that the last inequality holds provided that the following quantity

$$(1 - \gamma^2)(\lambda_2 - \lambda_1)^2 (\lambda_1 + \lambda_2 - \lambda)^2 [(1 + \gamma)\lambda_1 + (1 - \gamma)\lambda_2] \\ \times [(1 - \gamma)\lambda_1 + (1 + \gamma)\lambda_2]$$

is positive, which is trivial under our assumptions  $0 \leq \gamma < 1$  and  $0 < \lambda_1 \leq \lambda < \lambda_2$ . Thus,  $q[\lambda]$  takes its largest value, when  $\lambda = \lambda_1$ :

$$q[\lambda_1] = \gamma + (1 - \gamma) \frac{\lambda_1}{\lambda_2} = \frac{\lambda_1}{\lambda_2} + \gamma \left( 1 - \frac{\lambda_1}{\lambda_2} \right) \\ = 1 - (1 - \gamma) \left( 1 - \frac{\lambda_1}{\lambda_2} \right). \quad \square$$

**Remark 2.** What does it mean if  $\lambda \in [\lambda_k, \lambda_{k+1}[$ ,  $\lambda' < \lambda_k$ , and  $k > 1$ ? It means that on this iterative step the eigenvalue approximation has passed the level of  $\lambda_k$  on its monotonic way down and will converge to an eigenvalue with a smaller index. The theorem can be used again for the next iterate, with a properly adjusted  $k$ . Under assumptions of the theorem it is impossible to predict a priori if  $\lambda' < \lambda_k$  without making some extra assumptions.

In practice, we typically observe convergence to the smallest eigenvalue  $\lambda_1$  with possible slowdowns when passing larger eigenvalues. It is a normal behavior for a gradient minimization method as every eigenvector, except those corresponding to the extreme eigenvalues, is a saddle point of the minimizing function, the Rayleigh quotient. No extra assumptions, which would guarantee the convergence to the smallest eigenvalue  $\lambda_1$  for an arbitrary initial guess with  $k > 1$  when a general preconditioner is used, are presently known even for the simplest preconditioned eigensolver we analyze in Theorem 1. By analogy with the theory of classical inverse iterations, it might seem natural to suppose that the assumption that the initial vector is not orthogonal to the first eigenvector is necessary and sufficient to assure the convergence to  $\lambda_1$ . Counterexamples show that both necessary and sufficient parts of this supposition are wrong.

**Remark 3.** It follows directly from the proof of the theorem above that the true convergence factor in the estimate (10) depends on  $\lambda$ , but this dependence is not decisive. We eliminate  $\lambda$  to make the estimate much shorter.

Thus, our upper bound (11) of the convergence factor does not depend on  $\lambda$  and is sharp, as a function of the decisive quantities  $\gamma$ ,  $\lambda_k$ ,  $\lambda_{k+1}$  only. The estimate (10) is also asymptotically sharp, when  $\lambda \rightarrow \lambda_k$ , as it then turns into the sharp estimate (8).

**Remark 4.** The preconditioned steepest descent for the Rayleigh quotient when  $\omega$  is computed to minimize the Rayleigh quotient on the two-dimensional subspace  $\text{span}\{u, T^{-1}(Au - \lambda u)\}$ , evidently produces a smaller value  $\lambda'$  compared to that when  $\omega$  is chosen a priori. Thus, the convergence rate estimate (10) with the convergence factor (11) holds for the preconditioned steepest descent method, too. Moreover, we can now assume (2) instead of (1) and use (3) as we already discussed in Section 1, which leads to

$$1 - \gamma = \frac{2}{\kappa(T^{-1}A) + 1}. \tag{16}$$

The convergence factor estimate for the preconditioned steepest descent obtained here is not sharp, see Remark 12 below.

An alternative recent estimate can be found in: E. Ovtchinnikov, Convergence estimates for the generalized Davidson method for symmetric eigenvalue problems I, The preconditioning aspect (accepted by SIAM J. Numer. Anal., 2002).

#### 4. Generalized symmetric definite eigenvalue problems

We now consider a generalized *symmetric definite* eigenvalue problem of the form  $(A - \lambda B)u = 0$  with real symmetric  $n$ -by- $n$  matrices  $A$  and  $B$ , assuming that  $A$  is positive definite. This describes a regular matrix pencil  $A - \lambda B$  with a discrete spectrum (set of eigenvalues  $\lambda$ ). It is well known that such a generalized eigenvalue problem has all real eigenvalues  $\lambda_i$  and corresponding (right) eigenvectors  $u_i$ , satisfying  $(A - \lambda_i B)u_i = 0$ , can be chosen orthogonal in the following sense:  $(u_i, Au_j) = (u_i, Bu_j) = 0$ ,  $i \neq j$ . In some applications, the matrix  $B$  is simply the identity,  $B = I$ , and then we have the standard symmetric eigenvalue problem with matrix  $A$ , which has  $n$  real positive eigenvalues

$$0 < \lambda_{\min} = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n = \lambda_{\max}.$$

We already discussed the case  $B = I$  in the previous section.

In general, when  $B \neq I$ , the pencil  $A - \lambda B$  has  $n$  real, some possibly infinite, eigenvalues. If  $B$  is nonsingular, all eigenvalues are finite. If  $B$  is positive semi-definite, some eigenvalues are infinite, all other eigenvalues are positive, and we consider the problem of computing the smallest  $m$  eigenvalues of the pencil  $A - \lambda B$ . When  $B$  is indefinite, it is convenient to consider the pencil  $\mu A - B$  with eigenvalues

$$\mu = \frac{1}{\lambda}, \quad \mu_{\min} = \mu_n \leq \dots \leq \mu_1 = \mu_{\max},$$

where we want to compute the largest  $m$  eigenvalues,  $\mu_1, \dots, \mu_m$ , and corresponding eigenvectors.

We first consider the case  $B > 0$ , when we may still use  $\lambda$ 's. We naturally redefine the Rayleigh quotient (4) to

$$\lambda(u) = \frac{(u, Au)}{(u, Bu)}, \quad \text{where } u \in \mathbb{R}^n, \quad u \neq 0, \quad (17)$$

and replace method (7) with the following:

$$u' = u - \omega T^{-1}(Au - \lambda(u)Bu), \quad (18)$$

still assuming that the preconditioner  $T$  approximates  $A$  according to (1).

A different popular approach to deal with a generalized eigenvalue problem, e.g., utilized in the ARPACK based MATLAB code EIGS.m, relies on explicit factorizations of the matrix  $B$ ,  $A$ , or their linear combination. It cannot, of course, be used in a matrix-free environment, when all matrices are only available as matrix-vector-multiply (MVM) functions.

The method (18) is not new. It was previously studied, e.g., by D'yakonov and Orekhov [8–10]. Here, we easily derive a new sharp convergence estimate for it, using our previous result for  $B = I$ .

**Theorem 5.** *Let  $B > 0$ . Let  $u \in \mathbb{R}^n$  and let  $\lambda = \lambda(u) \in [\lambda_1, \lambda_n[$  be its Rayleigh quotient, where  $\lambda_1 \leq \dots \leq \lambda_n$  are the eigenvalues of  $B^{-1}A$ . The preconditioner is assumed to satisfy (1) for some  $\gamma \in [0, 1[$ . If  $\lambda = \lambda(u) \in [\lambda_k, \lambda_{k+1}[$ , then it holds for the Rayleigh quotient  $\lambda' = \lambda(u')$  with  $u'$  computed by (18) with  $\omega = 1$  that either  $\lambda' < \lambda_k$  (unless  $k = 1$ ), or  $\lambda' \in [\lambda_k, \lambda[$ . In the latter case, the convergence estimate (10) holds with the convergence factor (11).*

**Proof.** As  $B > 0$ , the bilinear form  $(\cdot, \cdot)_B = (\cdot, B\cdot)$  describes a scalar product, in which matrices  $B^{-1}T$  and  $B^{-1}A$  are symmetric positive definite. Let us make all the following substitutions at once:

$$(\cdot, \cdot)_B \Rightarrow (\cdot, \cdot), \quad B^{-1}A \Rightarrow A, \quad B^{-1}T \Rightarrow T.$$

Then, formula (18) turns into (7) and the generalized eigenvalue problem for the pencil  $A - \lambda B$  becomes a standard eigenvalue problem for the matrix  $B^{-1}A$ . Thus, we can use Theorem 5 that gives us the present theorem after the back substitution to the original terms of the present section.  $\square$

Remarks 2–4 hold with evident modifications for  $B > 0$ .

To cover the general case, when  $B$  may not be definite, we replace  $\lambda$ 's with  $\mu$ 's by switching from the pencil  $A - \lambda B$  to the pencil  $B - \mu A$ . We redefine the Rayleigh quotient (17) to

$$\mu(u) = \frac{(u, Bu)}{(u, Au)}, \quad \text{where } u \in \mathbb{R}^n, \quad u \neq 0, \tag{19}$$

and replace method (18) with the following:

$$u' = u + \omega T^{-1}(Bu - \mu(u)Au), \tag{20}$$

still assuming that the preconditioner  $T$  approximates  $A$  according to (1). We are now interested in the largest eigenvalue  $\mu_1$  of the matrix  $A^{-1}B$ .

The method (20) was previously suggested, e.g., in [16,18] and reproduced in [9], where the latter reference has advantages of being in English and widely available. We now obtain a new sharp convergence estimate for it, using our previous theorem.

**Theorem 6.** *Let  $u \in \mathbb{R}^n$  and let  $\mu = \mu(u) \in ]\mu_n, \mu_1]$  be its Rayleigh quotient, where  $\mu_1 \geq \dots \geq \mu_n = \mu_{\min}$  are the eigenvalues of  $A^{-1}B$ . The preconditioner is assumed to satisfy (1) for some  $\gamma \in [0, 1[$ . If  $\mu = \mu(u) \in ]\mu_{k+1}, \mu_k]$ , then it holds for the Rayleigh quotient  $\mu' = \mu(u')$  with  $u'$  computed by (20) with*

$$\omega = \frac{1}{\mu - \mu_{\min}}$$

that either  $\mu' > \mu_k$  (unless  $k = 1$ ), or  $\mu' \in ]\mu, \mu_k]$ . In the latter case, the convergence estimate

$$\frac{\mu_k - \mu'}{\mu' - \mu_{k+1}} \leq q^2 \frac{\mu_k - \mu}{\mu - \mu_{k+1}}, \quad (21)$$

holds with the convergence factor

$$q = 1 - (1 - \gamma) \frac{\mu_k - \mu_{k+1}}{\mu_k - \mu_{\min}}. \quad (22)$$

**Proof.** We first rewrite the estimate of the previous theorem for  $B > 0$  in terms of  $\mu$ 's:

$$\frac{\mu_k - \mu'}{\mu' - \mu_{k+1}} \leq q^2 \frac{\mu_k - \mu}{\mu - \mu_{k+1}}, \quad q = 1 - (1 - \gamma) \frac{\mu_k - \mu_{k+1}}{\mu_k}. \quad (23)$$

Here we use the fact that

$$\frac{\mu_k - \mu'}{\mu' - \mu_{k+1}} \frac{\mu - \mu_{k+1}}{\mu_k - \mu} = \frac{\lambda' - \lambda_k}{\lambda_{k+1} - \lambda'} \frac{\lambda_{k+1} - \lambda}{\lambda - \lambda_k}$$

and that

$$q = 1 - (1 - \gamma) \left( 1 - \frac{\lambda_k}{\lambda_{k+1}} \right) = 1 - (1 - \gamma) \frac{\mu_k - \mu_{k+1}}{\mu_k}. \quad (24)$$

We now are ready to deal with a general symmetric  $B$ . We use a trick, proposed in [16,18] and replicated in [9]. Namely, we substitute our actual matrix  $B$ , which is not necessarily positive definite, with positive definite matrix  $B_\alpha = B - \alpha A > 0$ , where a scalar  $\alpha < \mu_{\min}$ , and apply the previous estimate (23) to the pencil  $B_\alpha - \mu_\alpha A$  with eigenvalues  $\mu_\alpha = \mu - \alpha$ . This gives (23), but with

$$q = 1 - (1 - \gamma) \frac{\mu_k - \mu_{k+1}}{\mu_k - \alpha}.$$

Finally, we realize that the method itself is invariant with respect to  $\alpha$ , except for the scalar shift that must be now chosen as

$$\omega = \frac{1}{\mu - \alpha}.$$

Moreover, everything depends continuously on  $\alpha < \mu_{\min}$ , so we can take the limit  $\alpha = \mu_{\min}$  as well. This proves estimate (21) with  $q$  given by (22).  $\square$

**Remark 7.** Let us assume that the preconditioner  $T$  is scaled such that

$$\frac{1}{\kappa(T^{-1}A)}(u, Tu) \leq (u, Au) \leq (u, Tu),$$

so  $1 - \gamma = 1/\kappa(T^{-1}A)$ . Then our new estimate (21) with  $q$  given by (22) can be directly compared with estimate (6.4) and (6.5) of Knyazev [20]. Simple algebraic manipulations show that asymptotically, with respect to  $\mu \rightarrow \mu_1$ , or with respect to

$\kappa(T^{-1}A) \rightarrow 1$ , the convergence factors are the same in both estimates. However, in all other cases the new estimate (21), (22) is better.

Remarks 2–4 for general  $B$  turn into the following remarks.

**Remark 8.** If  $\mu \in ]\mu_{k+1}, \mu_k]$ ,  $\mu' > \mu_k$ , and  $k > 1$  the method will converge to an eigenvalue with a smaller index. The theorem can be used again for the next iterate, with a properly adjusted  $k$ .

**Remark 9.** The convergence factor (22) is sharp, as a function of the decisive quantities  $\gamma$ ,  $\mu_k - \mu_{k+1}$ ,  $\mu_k - \mu_{\min}$  only. The estimate (21) is also asymptotically sharp, when  $\mu \rightarrow \mu_k$ , as it then turns into a sharp estimate.

**Remark 10.** The preconditioned steepest ascent for the Rayleigh quotient (19), where  $\omega$  in (20) is computed to maximize the Rayleigh quotient on the two-dimensional subspace  $\text{span}\{u, T^{-1}(Bu - \mu Au)\}$ , evidently produces a larger value  $\mu'$  compared to that when  $\omega$  is chosen a priori. Thus, the convergence rate estimate (21) with the convergence factor (22) holds for the preconditioned steepest ascent method, too. Moreover, we can now assume (2) instead of (1) and use (16), which leads to the formula

$$q = 1 - \frac{2}{\kappa(T^{-1}A) + 1} \frac{\mu_k - \mu_{k+1}}{\mu_k - \mu_{\min}}. \tag{25}$$

**Remark 11.** In the locally optimal preconditioned conjugate gradient method (4.2) of Knyazev [22], the trial subspace is enlarged compared to that of the preconditioned steepest ascent method of Remark 10. Thus, the convergence rate estimate (21) with  $q$  given by (25) holds for the former method, too, assuming (2) and taking (16). Our preconditioner  $T$  was denoted as  $T^{-1}$  in [22].

**Remark 12.** An alternative convergence rate estimate for the preconditioned steepest ascent and for the locally optimal preconditioned conjugate gradient method is given in [17], see formulas (3.9)–(3.12). Estimate (3.9) is written exactly in the same form as our (21), but the convergence factor  $q$  is different. It is given by a lengthy formula, which asymptotically, with respect to  $\mu \rightarrow \mu_1$ , or with respect to  $\kappa(T^{-1}A) \rightarrow 1$ , can be significantly simplified and leads to the following asymptotic convergence factor

$$q = \frac{1 - \xi}{1 + \xi}, \quad \xi = \frac{1}{\kappa(T^{-1}A)} \frac{\mu_k - \mu_{k+1}}{\mu_k - \mu_{\min}} \tag{26}$$

also presented in [19, formula (1.3)].

We note that in a particular case  $\kappa(T^{-1}A) = 1$  this factor appears in [25]. Asymptotically, when  $\mu \rightarrow \mu_1$  and  $j = 1$ , it is consistent with that obtained in [42]. Moreover, according to Remark 3.7 of [17], the estimate (21) with the factor  $q$  given by (26) holds for the preconditioned steepest ascent nonasymptotically as well, but

under a practically unrealistic assumption that matrices  $A^{-1}B$  and  $T^{-1}(B - \mu A)$  commute for any real  $\mu$ . It is not presently known if the estimate (21) with the factor  $q$  given by (26) holds for the preconditioned steepest ascent under general assumptions of Theorem 6. This is a subject of a current research.

The importance of the convergence factor (26) is that it is smaller than (22), which means that our estimate (21) with the factor  $q$  given by (22) is not sharp for the preconditioned steepest ascent. Numerical tests of Knyazev and Neymeyr [23] using multigrid preconditioning support this conclusion.

## 5. Preconditioned subspace iterations

In this section, we will present a generalization of results of the previous two sections to the case, where  $m$  extreme eigenpairs are computed simultaneously in so-called *subspace*, or *block* iteration methods.

We need to return to the case  $B = I$  again and consider first the following block version of method (6).

Let the current iterate  $U^{(i)}$  be an  $n$ -by- $m$  matrix with columns, approximating  $m$  eigenvectors of  $A$ , corresponding to  $m$  smallest eigenvalues. We assume that

$$(U^{(i)})^T U^{(i)} = I, \quad (U^{(i)})^T A U^{(i)} = \text{diag}(\lambda_1^{(i)}, \dots, \lambda_m^{(i)}) = A^{(i)}.$$

We perform one step of iterations

$$\hat{U}^{(i+1)} = U^{(i)} - T^{-1}(AU^{(i)} - U^{(i)}A^{(i)})\Omega^{(i)}, \quad (27)$$

where  $\Omega^{(i)}$  is an  $m$ -by- $m$  matrix, a generalization of the scalar step size. Finally, we compute the next iterate  $U^{(i+1)}$  by the Rayleigh–Ritz procedure for the pencil  $A - \lambda I$  on the trial subspace given by the column-space of  $\hat{U}^{(i+1)}$  such that

$$\begin{aligned} (U^{(i+1)})^T U^{(i+1)} &= I, \\ (U^{(i+1)})^T A U^{(i+1)} &= \text{diag}(\lambda_1^{(i+1)}, \dots, \lambda_m^{(i+1)}) = A^{(i+1)}. \end{aligned}$$

The preconditioned iterative method (27) with  $\Omega^{(i)} = I$  is analyzed in [5], where a survey on various attempts to analyze this and simplified preconditioned subspace schemes is also given. In this analysis, restrictive conditions on the initial subspace are assumed to be satisfied.

An alternative theory for method (27) with  $\Omega^{(i)} = I$  is developed in [30], based on the sharp convergence rate estimate (8) of Neymeyr [31,32] for single-vector preconditioned solver that we use in the previous two sections. The advantages of the approach of Neymeyr [30] are that:

- it is applicable to any initial subspaces,
- the convergence rate estimate can be used recursively, while the estimate of Bramble et al. [5] cannot,

- the estimates for the convergence of the Ritz values are individually sharp in a sense that an initial subspace and a preconditioner can be constructed so that the convergence rate estimate for a fixed index  $j \in [1, m]$  is attained,
- the convergence rate estimate for a fixed index  $j$  is exactly the same as (8) for the single-vector method (6) with  $\omega^{(i)} = 1$ .

The only serious disadvantage of the estimates of Neymeyr [30] is that they deteriorate when eigenvalues of interest  $\lambda_1, \dots, \lambda_m$  include a cluster. The actual convergence of method (27) in numerical tests is known not to be sensitive to clustering of eigenvalues, and estimates of Bramble et al. [5] do capture this property, essential for subspace iterations.

A sharp simplification of the estimate of Neymeyr [30] is suggested in Theorem 5.1 of [22], but the proof is sketchy and not complete. In this section, we fill these gaps in the arguments of Knyazev [22].

First, we reproduce here the result of Theorem 3.3 of [30]: for a fixed index  $j \in [1, m]$ , if  $\lambda_j^{(i)} \in [\lambda_{k_j}, \lambda_{k_j+1}[$  and the method (27) with  $\Omega^{(i)} = I$  is used, then

$$\lambda_j^{(i+1)} \leq \lambda_{k_j, k_j+1}(\lambda_j^{(i)}, \gamma), \tag{28}$$

where the latter quantity is given by (9). Now, using the fact that the estimate (28) is identical to (8) and that our proof of Theorem 1 provides an equivalent representation of expression (9), we immediately derive the following generalization of Theorem 1 to the block method.

**Theorem 13.** *The preconditioner is assumed to satisfy (1) for some  $\gamma \in [0, 1[$ . For a fixed index  $j \in [1, m]$ , if  $\lambda_j^{(i)} \in [\lambda_{k_j}, \lambda_{k_j+1}[$ , then it holds for the Ritz value  $\lambda_j^{(i+1)}$  computed by (27) with  $\Omega^{(i)} = I$  that either  $\lambda_j^{(i+1)} < \lambda_{k_j}$  (unless  $k_j = j$ ), or  $\lambda_j^{(i+1)} \in [\lambda_{k_j}, \lambda_j^{(i)}[$ .*

*In the latter case,*

$$\frac{\lambda_j^{(i+1)} - \lambda_{k_j}}{\lambda_{k_j+1} - \lambda_j^{(i+1)}} \leq (q(\gamma, \lambda_{k_j}, \lambda_{k_j+1}))^2 \frac{\lambda_j^{(i)} - \lambda_{k_j}}{\lambda_{k_j+1} - \lambda_j^{(i)}}, \tag{29}$$

where

$$q(\gamma, \lambda_{k_j}, \lambda_{k_j+1}) = \gamma + (1 - \gamma) \frac{\lambda_{k_j}}{\lambda_{k_j+1}} = 1 - (1 - \gamma) \left(1 - \frac{\lambda_{k_j}}{\lambda_{k_j+1}}\right) \tag{30}$$

is the convergence factor.

By analogy with Remarks 2–4, we have the following.

**Remark 14.** If  $\lambda_j^{(i)} \in [\lambda_{k_j}, \lambda_{k_j+1}[$ ,  $\lambda_j^{(i+1)} < \lambda_{k_j}$ , and  $k_j > j$ , the  $j$ th approximation will converge to an eigenvalue with a smaller index. The theorem can be used again for the next iterate, with a properly adjusted  $k_j$ .

**Remark 15.** For a fixed index  $j$ , the convergence factor  $q(\gamma, \lambda_{k_j}, \lambda_{k_j+1})$  given by (30) is sharp, as a function of the decisive quantities  $\gamma, \lambda_{k_j}, \lambda_{k_j+1}$  only. The estimate (29) and (30) is also asymptotically sharp, when  $\lambda_j^{(i)} \rightarrow \lambda_{k_j}$ , as it then turns into the sharp estimate (28).

Let us highlight again that, while the convergence factors (30) are sharp individually, when we fix the index  $j$ , they are not sharp collectively, for all  $j = 1, \dots, m$ , neither asymptotically, when the initial subspace is already close to the seeking subspace spanned by the first  $m$  eigenvectors. In the latter case, the estimates of [5] are better.

**Remark 16.** There are several different versions of the preconditioned block steepest descent; see, e.g., [21]. In one of them,  $U^{(i+1)}$  is computed by the Rayleigh–Ritz method of the  $2m$ -dimensional trial subspaces, spanned by columns of  $U^{(i)}$  and  $T^{-1}(AU^{(i)} - U^{(i)}A^{(i)})$ . This leads to Ritz values  $\lambda_j^{(i+1)}$ , which are not larger than those produced by (27) with any  $\Omega^{(i)}$ , in particular, with  $\Omega^{(i)} = I$ . Thus, the convergence rate estimate (29) with the convergence factor (30) holds for this version of the preconditioned block steepest descent method, too. Moreover, we can now assume (2) instead of (1) and use (16).

Let now  $B \neq I, B > 0$ . Then we assume that

$$(U^{(i)})^T B U^{(i)} = I, \quad (U^{(i)})^T A U^{(i)} = \text{diag}(\lambda_1^{(i)}, \dots, \lambda_m^{(i)}) = A^{(i)}.$$

We perform one step of iterations, cf. (18),

$$\hat{U}^{(i+1)} = U^{(i)} - T^{-1}(A U^{(i)} - B U^{(i)} A^{(i)}) \Omega^{(i)}, \quad (31)$$

and compute the next iterate  $U^{(i+1)}$  by the Rayleigh–Ritz procedure for the pencil  $A - \lambda B$  on the trial subspace given by the column-space of  $\hat{U}^{(i+1)}$  such that

$$\begin{aligned} (U^{(i+1)})^T B U^{(i+1)} &= I, \\ (U^{(i+1)})^T A U^{(i+1)} &= \text{diag}(\lambda_1^{(i+1)}, \dots, \lambda_m^{(i+1)}) = A^{(i+1)}. \end{aligned}$$

Repeating the same arguments as those in the proof of Theorem 5, we conclude that Theorem 13 also trivially holds for the method (31) with  $\Omega^{(i)} = I$  for solving an generalized eigenvalue problem for pencil  $A - \lambda B$ , when  $B > 0$ .

Finally, in the general case, when  $B$  may not be definite, we modify the method (31) for the pencil  $B - \mu A$  the following way: assuming that

$$(U^{(i)})^T A U^{(i)} = I, \quad (U^{(i)})^T B U^{(i)} = \text{diag}(\mu_1^{(i)}, \dots, \mu_m^{(i)}) = M^{(i)},$$

we perform one step of iterations, cf. (20),

$$\hat{U}^{(i+1)} = U^{(i)} + T^{-1}(B U^{(i)} - A U^{(i)} M^{(i)}) \Omega^{(i)}, \quad (32)$$

and compute the next iterate  $U^{(i+1)}$  by the Rayleigh–Ritz procedure for the pencil  $B - \mu A$  on the trial subspace given by the column-space of  $\hat{U}^{(i+1)}$  such that

$$\begin{aligned} (U^{(i+1)})^T A U^{(i+1)} &= I, \\ (U^{(i+1)})^T B U^{(i+1)} &= \text{diag}(\mu_1^{(i+1)}, \dots, \mu_m^{(i+1)}) = M^{(i+1)}. \end{aligned}$$

By analogy with the proof of Theorem 6, we derive:

**Theorem 17.** *The preconditioner is assumed to satisfy (1) for some  $\gamma \in [0, 1[$ . For a fixed index  $j \in [1, m]$ , if  $\mu_j^{(i)} \in ]\mu_{k_j+1}, \mu_{k_j}]$ , then it holds for the Ritz value  $\mu_j^{(i+1)}$  computed by (32) with*

$$\Omega^{(i)} = (M^{(i)} - \mu_{\min} I)^{-1}$$

that either  $\mu_j^{(i+1)} > \mu_{k_j}$  (unless  $k_j = j$ ) or  $\mu_j^{(i+1)} \in ]\mu_j^{(i)}, \mu_{k_j}]$ . In the latter case,

$$\frac{\mu_{k_j+1} - \mu_j^{(i+1)}}{\mu_j^{(i+1)} - \mu_{k_j}} \leq (q(\gamma, \mu_{k_j}, \mu_{k_j+1}))^2 \frac{\mu_{k_j+1} - \mu_j^{(i)}}{\mu_j^{(i)} - \mu_{k_j}}, \tag{33}$$

where

$$q(\gamma, \mu_{k_j}, \mu_{k_j+1}) = 1 - (1 - \gamma) \left( \frac{\mu_{k_j} - \mu_{k_j+1}}{\mu_{k_j} - \mu_{\min}} \right) \tag{34}$$

is the convergence factor.

**Remark 18.** If columns of  $U^{(i+1)}$  are computed by the Rayleigh–Ritz method for the pencil  $B - \mu A$ , as  $m$  Ritz vectors corresponding to the  $m$  largest Ritz values, on the  $2m$ -dimensional trial subspace spanned by columns of  $U^{(i)}$  and  $T^{-1}(BU^{(i)} - AU^{(i)}M^{(i)})$ , the convergence rate estimate (33) with the convergence factor (34) holds for this version of the preconditioned block steepest ascent method, too. Moreover, we can now assume (2) instead of (1) and use (16).

**Remark 19.** In the LOBPCG method of Knyazev [22],  $U^{(i+1)}$  is computed by the Rayleigh–Ritz method on the  $3m$ -dimensional trial subspaces, spanned by columns of  $U^{(i-1)}$ ,  $U^{(i)}$  and  $T^{-1}(BU^{(i)} - AU^{(i)}M^{(i)})$ . Thus, in LOBPCG the trial subspace is enlarged compared to that of the preconditioned block steepest ascent method, described in the previous remark. Therefore, evidently, the convergence rate estimate (33) with the convergence factor given by (34) with (16), assuming (2), holds for the LOBPCG method, too; see Theorem 5.1 of [22].

Remark 19 provides us with the only presently known theoretical convergence rate estimate of the LOBPCG for  $m > 1$ . However, this estimate is, by construction, the same as that for the preconditioned block steepest ascent method, which, in its turn, is the same as that of the PINVIT with the optimal scaling. Numerical comparison of these methods according to Knyazev [20–22], Knyazev and Neymeyr [23] demonstrates, however, that the LOBPCG method is in practice much faster. Therefore, our theoretical convergence estimates of the LOBPCG of the present paper are

not sharp enough yet to explain excellent convergence properties of the LOBPCG in numerical simulations.

## 6. Conclusion

We derive a short and sharp convergence rate estimate for basic preconditioned eigensolvers. The analysis presented here should increase understanding and provide tools for investigation of more efficient preconditioned eigensolvers, such as LOBPCG [22,23], under development, see <http://www-math.cudenver.edu/~aknyazev/research/eigensolvers/>

## References

- [1] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, H. van der Vorst (Eds.), *Templates for the Solution of Algebraic Eigenvalue Problems*, SIAM, Philadelphia, PA, 2000.
- [2] A. Basermann, Parallel block ILUT/ILDLT preconditioning for sparse eigenproblems and sparse linear systems, *Numer. Linear Algebra Appl.* 7 (7–8) (2000) 635–648 (Preconditioning techniques for large sparse matrix problems in industrial applications, Minneapolis, MN, 1999).
- [3] L. Bergamaschi, G. Pini, F. Sartoretto, Approximate inverse preconditioning in the parallel solution of sparse eigenproblems, *Numer. Linear Algebra Appl.* 7 (3) (2000) 99–116.
- [4] W. Bradbury, R. Fletcher, New iterative methods for solution of the eigenproblem, *Numer. Math.* 9 (1966) 259–267.
- [5] J. Bramble, J. Pasciak, A. Knyazev, A subspace preconditioning algorithm for eigenvector/eigenvalue computation, *Adv. Comput. Math.* 6 (1996) 159–189.
- [6] D.C. Dobson, An efficient method for band structure calculations in 2D photonic crystals, *J. Comput. Phys.* 149 (2) (1999) 363–376.
- [7] D.C. Dobson, J. Gopalakrishnan, J.E. Pasciak, An efficient method for band structure calculations in 3D photonic crystals, *J. Comput. Phys.* 161 (2) (2000) 668–679.
- [8] E. D'yakonov, Iteration methods in eigenvalue problems, *Math. Notes* 34 (1983) 945–953.
- [9] E. D'yakonov, *Optimization in Solving Elliptic Problems*, CRC Press, Boca Raton, FL, 1996.
- [10] E. D'yakonov, M. Orekhov, Minimization of the computational labor in determining the first eigenvalues of different operators, *Math. Notes* 27 (1980) 382–391.
- [11] J. Fattebert, J. Bernholc, Towards grid-based  $O(N)$  density-functional theory methods: Optimized nonorthogonal orbitals and multigrid acceleration, *Phys. Rev. B* 62 (3) (2000) 1713–1722.
- [12] Y. Feng, D. Owen, Conjugate gradient methods for solving the smallest eigenpair of large symmetric eigenvalue problems, *Internat. J. Numer. Methods Engrg.* 39 (1996) 2209–2229.
- [13] S. Godunov, V. Ogneva, G. Prokopov, On the convergence of the modified method of steepest descent in the calculation of eigenvalues, *Amer. Math. Soc. Transl. Ser. 2* 105 (1976) 111–116.
- [14] M. Hestenes, W. Karush, A method of gradients for the calculation of the characteristic roots and vectors of a real symmetric matrix, *J. Res. Nat. Bur. Standards* 47 (1951) 45–61.
- [15] L.V. Kantorovich, *Functional analysis and applied mathematics*, *Transl. Uspehi Mat. Nauk* 3 (6(28)) (1949) 89–185 (National Bureau of Standards Report, Washington, 1952).
- [16] A.V. Knyazev, *Computation of eigenvalues and eigenvectors for mesh problems: algorithms and error estimates*, Department of Numerical Mathematics, USSR Academy of Sciences, Moscow, 1986 (in Russian).
- [17] A.V. Knyazev, Convergence rate estimates for iterative methods for a mesh symmetric eigenvalue problem, *Russian J. Numer. Anal. Math. Modelling* 2 (1987) 371–396.

- [18] A.V. Knyazev, On modified gradient methods for spectral problems. *Differentsial'nye Uravneniya* 23(4) (1987) 715–717, 735 (in Russian).
- [19] A.V. Knyazev, A preconditioned conjugate gradient method for eigenvalue problems and its implementation in a subspace, in: *Eigenvertaufgaben in Natur- und Ingenieurwissenschaften und ihre numerische Behandlung*, Oberwolfach, 1990, *Internat. Ser. Numer. Math.*, vol. 96, Birkhäuser, Basel, 1991.
- [20] A.V. Knyazev, Preconditioned eigensolvers—an oxymoron? *Electron. Trans. Numer. Anal.* 7 (1998) 104–123.
- [21] A.V. Knyazev, Preconditioned eigensolvers: practical algorithms, in: Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, H. van der Vorst (Eds.), *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, SIAM, Philadelphia, PA, 2000, pp. 352–368 (Section 11.3. An extended version published as a technical report UCD-CCM 143, 1999, at the Center for Computational Mathematics, University of Colorado at Denver).
- [22] A.V. Knyazev, Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method, *SIAM J. Sci. Comput.* 23 (2) (2001) 517–541.
- [23] A.V. Knyazev, K. Neymeyr, Efficient solution of symmetric eigenvalue problems using multigrid preconditioners in the locally optimal block conjugate gradient method. ETNA, 2001 (accepted).
- [24] A.V. Knyazev, K. Neymeyr, A geometric theory for preconditioned inverse iteration. III: A short and sharp convergence estimate for generalized eigenvalue problems, 2001. Available from <http://www-math.cudenver.edu/ccmreports/rep173.pdf>, CU-Denver.
- [25] A.V. Knyazev, A.L. Skorokhodov, On exact estimates of the convergence rate of the steepest ascent method in the symmetric eigenvalue problem, *Linear Algebra Appl.* 154–156 (1991) 157–245.
- [26] D. Longsine, S. McCormick, Simultaneous Rayleigh-quotient minimization methods for  $Ax = \lambda Bx$ , *Linear Algebra Appl.* 34 (1980) 195–234.
- [27] S. McCormick, T. Noe, Simultaneous iteration for the matrix eigenvalue problem, *Linear Algebra Appl.* 16 (1977) 43–56.
- [28] R.B. Morgan, Preconditioning eigenvalues and some comparison of solvers, *J. Comput. Appl. Math.* 123 (1–2) (2000) 101–115 (*Numerical Analysis 2000*, Vol. III. *Linear Algebra*).
- [29] R.B. Morgan, D.S. Scott, Preconditioning the Lanczos algorithm for of sparse symmetric eigenvalue problems, *SIAM J. Sci. Comput.* 14 (3) (1993) 585–593.
- [30] K. Neymeyr, A geometric theory for preconditioned inverse iteration applied to a subspace, *Math. Comput.* 71 (2002) 197–216.
- [31] K. Neymeyr, A geometric theory for preconditioned inverse iteration. I: Extrema of the Rayleigh quotient, *Linear Algebra Appl.* 322 (2001) 61–85.
- [32] K. Neymeyr, A geometric theory for preconditioned inverse iteration. II: Convergence estimates, *Linear Algebra Appl.* 322 (2001) 87–104.
- [33] M.K. Ng, Preconditioned Lanczos methods for the minimum eigenvalue of a symmetric positive definite Toeplitz matrix, *SIAM J. Sci. Comput.* 21 (6) (2000) 1973–1986 (electronic).
- [34] Y. Notay, Combination of Jacobi–Davidson and conjugate gradients for the partial symmetric eigenproblem, *Numer. Linear Algebra Appl.* 9 (2002) 21–44.
- [35] S. Oliveira, On the convergence rate of a preconditioned sunspace eigensolver, *Computing* 63 (3) (1999) 219–231.
- [36] E. Ovtchinnikov, L. Xanthis, Successive eigenvalue relaxation: a new method for generalized eigenvalue problems and convergence estimates, *Proc. Roy. Soc. London Ser. A* 457 (2001) 441–451.
- [37] E.E. Ovtchinnikov, L.S. Xanthis, Effective dimensional reduction algorithm for eigenvalue problems for thin elastic structures: A paradigm in three dimensions, *Proc. Nat. Acad. Sci. U.S.A.* 97 (3) (2000) 967–971.
- [38] W. Petryshyn, On the eigenvalue problem  $Tu - \lambda Su = 0$  with unbounded and non-symmetric operators  $T$  and  $S$ , *Philos. Trans. Roy. Soc. London Ser. A.* 262 (1968) 413–458.
- [39] G. Rodrigue, A gradient method for the matrix eigenvalue problem  $Ax = \lambda Bx$ , *Numer. Math.* 22 (1973) 1–16.

- [40] M. Sadkane, R.B. Sidje, Implementation of a variable block Davidson method with deflation for solving large sparse eigenproblems, *Numer. Algorithms* 20 (2–3) (1999) 217–240.
- [41] A. Sameh, Z. Tong, The trace minimization method for the symmetric generalized eigenvalue problem, *J. Comput. Appl. Math.* 123 (1–2) (2000) 155–175 (Numerical Analysis 2000, Vol. III. Linear Algebra).
- [42] B. Samokish, The steepest descent method for an eigenvalue problem with semi-bounded operators, *Izv. Vyssh. Uchebn. Zaved. Mat.* 5 (1958) 105–114 (in Russian).
- [43] D.S. Scott, Solving sparse symmetric generalized eigenvalue problems without factorization, *SIAM J. Numer. Anal.* 18 (1981) 102–110.
- [44] G.L.G. Sleijpen, H.A. Van der Vorst, A Jacobi–Davidson iteration method for linear eigenvalue problem, *SIAM J. Matrix Anal. Appl.* 17 (2) (1996) 401–425.
- [45] P. Smit, M.H.C. Paardekooper, The effects of inexact solvers in algorithms for symmetric eigenvalue problems, *Linear Algebra Appl.* 287 (1–3) (1999) 337–357 (Special issue celebrating the 60th birthday of Ludwig Elsner).
- [46] X. Yang, A survey of various conjugate gradient algorithms for iterative solution of the largest/smallest eigenvalue and eigenvector of a symmetric matrix, Collection: Application of conjugate gradient method to electro-magnetic and signal analysis, *Progr. Electromagnetic Res.* 5 (1991) 567–588.
- [47] T. Zhang, G.H. Golub, K.H. Law, Subspace iterative methods for eigenvalue problems, *Linear Algebra Appl.* 294 (1–3) (1999) 239–258.