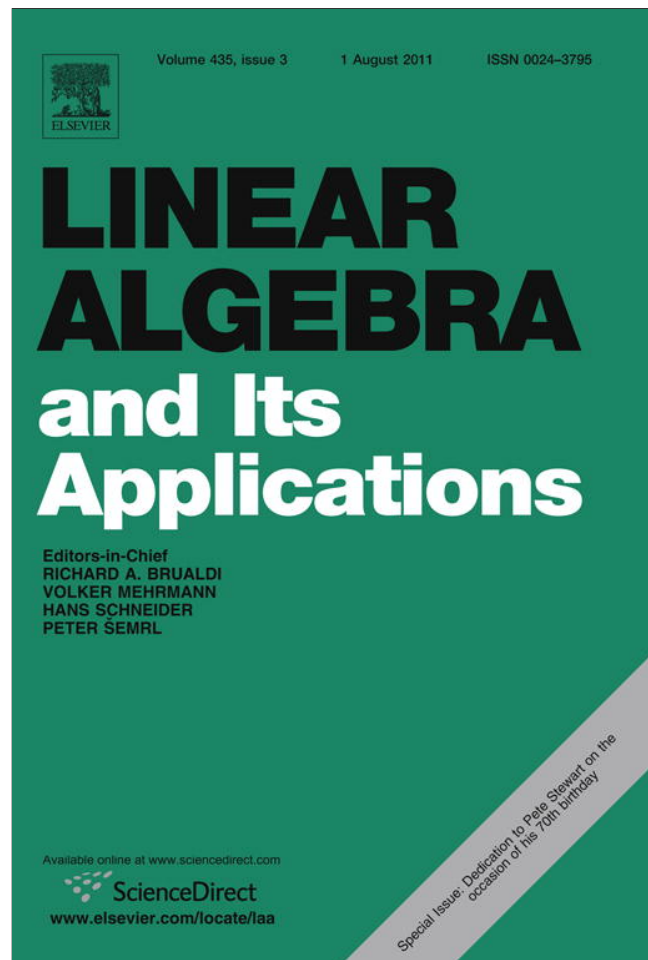


Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



Contents lists available at ScienceDirect

## Linear Algebra and its Applications

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

## Bounding the spectrum of large Hermitian matrices

Yunkai Zhou <sup>a,\*</sup>, Ren-Cang Li <sup>b,2</sup><sup>a</sup> Department of Mathematics, Southern Methodist University, Dallas, TX 75275, USA<sup>b</sup> Department of Mathematics, University of Texas at Arlington, Arlington, TX 76019, USA

## ARTICLE INFO

Submitted by V. Mehrmann

Dedicated to Prof. G.W. Stewart on the occasion of his 70th birthday

## Keywords:

Hermitian matrix  
Largest eigenvalue  
Upper bound  
Lanczos

## ABSTRACT

Estimating upper bounds of the spectrum of large Hermitian matrices has long been a problem with both theoretical and practical significance. Algorithms that can compute tight upper bounds with minimum computational cost will have applications in a variety of areas. We present a practical algorithm that exploits  $k$ -step Lanczos iteration with a safeguard step. The  $k$  is generally very small, say 5–8, regardless of the large dimension of the matrices. This makes the Lanczos iteration economical. The safeguard step can be realized with marginal cost by utilizing the theoretical bounds developed in this paper. The bounds establish the theoretical validity of a previous bound estimator that has been successfully used in various applications. Moreover, we improve the bound estimator which can now provide tighter upper bounds with negligible additional cost.

© 2010 Elsevier Inc. All rights reserved.

## 1. Introduction

Various researchers have studied cost-effective ways to approximate the largest eigenvalue of Hermitian matrices. O'Leary et al. [11] used Rayleigh-quotient iteration; Parlett et al. [12], and Kuczyński and Woźniakowski [10] employed Lanczos iteration with random initial vectors. The focus in [11,12,10] is to make rather accurate estimate of the largest eigenvalue of Hermitian positive definite matrices. Because the estimation is expected to have high accuracy, the iteration steps often cannot be very small.

In contrast, in several applications we cannot afford to estimate the largest eigenvalue to high accuracy because it can be too costly, especially when the estimation procedure has to be done repeatedly

\* Corresponding author.

E-mail addresses: [yzhou@smu.edu](mailto:yzhou@smu.edu) (Y. Zhou), [rcli@uta.edu](mailto:rcli@uta.edu) (R.-C. Li).<sup>1</sup> Supported in part by the National Science Foundation under Grant Nos. CMMI-0727194 and OCI-0749074.<sup>2</sup> Supported in part by the National Science Foundation under Grant Nos. DMS-0702335 and DMS-0810506.

in a single simulation. Instead, we focus on obtaining practical upper bounds of the spectrum with low cost. Our upper bound estimator is also based on the Lanczos iteration [13,16]. One major application of our estimator is for the Chebyshev-filtered subspace iteration method [18,19] in real-space Density Functional Theory (DFT) calculations, where an estimator for the largest eigenvalues of some Hermitian matrices has to be repeatedly called. The Hermitian matrices arise from real-space DFT calculations [2] are indefinite, sparse, and usually of very large scale.

The upper bound plays a crucial role for the efficiency of the Chebyshev filters. It is understood that the tighter the upper bound the better. But there is a tradeoff between tightness of a bound and the computational cost for it. The large dimension implies that it is not cost-effective to let the Lanczos iteration run until a Ritz value converges to the largest eigenvalue. Instead we wish to obtain an upper bound within just a few Lanczos steps, say less than 8 steps, regardless of the dimension of the matrices. To achieve this goal we need to look deeper into the bounds from Lanczos iteration and provide necessary safeguard steps.

An inexpensive upper bound estimator was proposed in [18]. The estimator plays a critical role in the nonlinear Chebyshev filtered subspace iteration method that is now the default solver in the real-space DFT package called PARSEC [2,3]. It has been successfully used for a wide range of DFT calculations, including highly challenging problems with dimension over several millions, where several thousand eigenpairs need to be computed [3,17,19]. The application of the estimator is not limited to DFT calculations, e.g., it plays a part in the Chebyshev–Davidson algorithm [20] for solving Hermitian eigenvalue problems of large dimensions.

However, the upper bound estimator in [18] lacks a rigorous proof. In this paper we will analyze the bound in more details, provide certain conditions under which we can rigorously prove that the estimator provides an upper bound. We also develop tighter upper bounds based on our analysis. Moreover, we will construct practical heuristics to guarantee that the estimator will provide an upper bound even if the conditions do not hold. This is important to ensure robustness of the upper bound estimator.

The rest of this article is organized as follows. Section 2 provides a theoretical analysis for practical bounds. Numerical examples, including those from DFT calculations and two artificial ones purposely constructed to test our bounds, are presented in Section 3 to illustrate the effectiveness of the proposed bounds in the previous section. We conclude this article by a few remarks along with possible applications other than DFT calculations in Section 4

## 2. Theoretical study of upper bounds

As explained in [18], the Chebyshev-filtered subspace iteration method requires an upper bound  $\beta$  that satisfies  $\beta \geq \lambda_{\max}(A)$ , where  $A$  denotes the size  $n \times n$  Hermitian matrix and  $\lambda_{\max}(A)$  its largest eigenvalue. Theoretically, any consistent norm  $\|A\|$ , in particular the  $\ell_1$ - or  $\ell_\infty$ -operator norm provides an upper bound on  $\lambda_{\max}(A)$ . One may also use Gerschgorin's disk theorem to find other upper bounds. Nevertheless, as pointed out in [18, Section 4.3], upper bounds obtained as such are often too crude to be effective for the nonlinear filtered subspace iteration algorithm that calls for the bound estimates. Another obstacle associated with these methods occurs when  $A$  is not stored as a matrix but instead accessed via a matrix-vector product subroutine. This is usually the case in DFT calculations, either in a real-space or in a plane-wave setting.

The upper bound estimator proposed in [18] performs the following: Run  $k$ -step Lanczos [13,16] on  $A$  with a random starting vector to get

$$AQ_k = Q_k T_k + f_k e_k^T, \quad (2.1)$$

where  $Q_k$  is  $n \times k$  and has orthonormal columns,  $T_k$  is  $k \times k$  and tri-diagonal,  $Q_k^* f_k = 0$ , the superscript  $*$  denotes conjugate transpose (it becomes just transpose in the real case), and  $e_k$  is the  $k$ th column of the  $k \times k$  identity matrix. Then take  $\lambda_{\max}(T_k) + \|f_k\|_2$  as an upper bound on  $\lambda_{\max}(A)$ . While no analysis was presented in [18] to theoretically guarantee that  $\lambda_{\max}(T_k) + \|f_k\|_2$  was indeed an upper bound, we observed that a very small  $k$ , say  $4 \leq k \leq 10$ , was often enough to provide an upper bound.

It is not hard to see that the method can fail if the starting vector is unfortunately taken from an invariant subspace that is orthogonal to  $A$ 's eigenspace corresponding to  $\lambda_{\max}(A)$ . But the initial

vector is randomly generated, nowadays any random number generator (see [9]) used would make this scenario a probability zero incidence. Hence in practice we can safely assume that the random initial vector is not orthogonal to the eigenspace corresponding to  $\lambda_{\max}(A)$ .

Now we shall present conditions that guarantee  $\lambda_{\max}(T_k) + \|f_k\|_2$  to be an upper bound on  $\lambda_{\max}(A)$ . Then we proceed to develop some refined bounds.

Denote the eigenvalues of  $A$  and  $T_k$  by

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n \quad \text{and} \quad \mu_1 \leq \mu_2 \leq \dots \leq \mu_k,$$

respectively.

In exact arithmetic,  $Q_k$  in (2.1) has orthonormal columns, i.e.,  $Q_k^* Q_k = I_k$ . Numerically when round-off errors are taken into consideration,  $Q_k$ 's columns are nearly orthonormal for  $k$  not too big, and (2.1) should be replaced by [13, p. 295]

$$AQ_k = Q_k T_k + f_k e_k^T + F_k, \tag{2.1a}$$

where  $F_k = \mathcal{O}(\epsilon)$  records all the roundoff errors and  $\epsilon$  is machine unit roundoff. It is well-known that the orthogonality among  $Q_k$ 's columns begins to deteriorate as  $T_k$ 's eigenvalues start to converge to  $A$ 's [13, Section 13.3]. But as pointed out in [18], for the purpose of bounding  $\lambda_{\max}(A)$ , only very few steps, i.e., small  $k$ , suffices. Because of that, the Lanczos process would usually stop far before  $\lambda_{\max}(T_k)$  approximates one of  $A$ 's eigenvalues so accurately and, as a consequence,

$$\|f_k\|_2 \gg \|F_k\|_2 = \mathcal{O}(\epsilon), \quad \|f_k e_k^T + F_k\|_2 \approx \|f_k\|_2.$$

Moreover, the randomly selected initial vector most likely ensure that  $\lambda_{\max}(T_k)$  would be closer to  $\lambda_{\max}(A)$  than to other eigenvalues of  $A$ . Therefore it is reasonable to assume that after a few Lanczos steps,

$$|\lambda_n - \lambda_{\max}(T_k)| = \min_i |\lambda_i - \lambda_{\max}(T_k)|. \tag{2.2}$$

It is also reasonable to assume that for  $k$  not too big,  $Q_k$ 's columns are nearly orthonormal.

**Lemma 1** (Kahan [7], Cao et al. [1]). *There exist  $k$  eigenvalues of  $A$  :  $\lambda_{i_1} \leq \lambda_{i_2} \leq \dots \leq \lambda_{i_k}$  such that*

$$|\lambda_{i_j} - \mu_j| \leq \frac{\|f_k e_k^T + F_k\|_2}{\sigma_{\min}(Q_k)} \approx \frac{\|f_k\|_2}{\sigma_{\min}(Q_k)}$$

for  $1 \leq j \leq k$ , where  $\sigma_{\min}(Q_k)$  is the smallest singular value of  $Q_k$ .

This lemma holds regardless of the assumption (2.2) and orthogonality among  $Q_k$ 's columns. It also suggests the negligible effect of  $F_k$ , comparing to that of  $\sigma_{\min}(Q_k)$ , on the accuracy of  $\mu_j$  as approximations to some of the  $\lambda_i$ . But if  $Q_k$ 's columns are nearly orthonormal, then  $\sigma_{\min}(Q_k) \approx 1$ .

**Theorem 1.** *If (2.2) holds, then*

$$\lambda_{\max}(A) \leq \lambda_{\max}(T_k) + \frac{\|f_k e_k^T + F_k\|_2}{\sigma_{\min}(Q_k)} \approx \lambda_{\max}(T_k) + \frac{\|f_k\|_2}{\sigma_{\min}(Q_k)}. \tag{2.3}$$

**Proof.** With (2.2), one can take  $i_k = n$  in Lemma 1. Then

$$\lambda_{\max}(A) = \lambda_n \leq \mu_k + \frac{\|f_k e_k^T + F_k\|_2}{\sigma_{\min}(Q_k)} \approx \lambda_{\max}(T_k) + \frac{\|f_k\|_2}{\sigma_{\min}(Q_k)},$$

as expected.  $\square$

A sharper bound than (2.3) is given in the next theorem.

**Theorem 2.** Suppose (2.2) holds. Let  $T_k z = \mu_k z$  and  $\|z\|_2 = 1$ . Then

$$\lambda_{\max}(A) \leq \lambda_{\max}(T_k) + \frac{\|f_k e_k^T z + F_k z\|_2}{\|Q_k z\|_2} \approx \lambda_{\max}(T_k) + \frac{|e_k^T z| \|f_k\|_2}{\|Q_k z\|_2}. \quad (2.4)$$

**Proof.** Multiply Eq. (2.1) by  $z$  from the right yields

$$A Q_k z - \mu_k Q_k z = f_k e_k^T z + F_k z.$$

With (2.2), we have [13, p. 73]

$$\lambda_{\max}(A) = \lambda_n \leq \mu_k + \frac{\|f_k e_k^T z + F_k z\|_2}{\|Q_k z\|_2} \approx \lambda_{\max}(T_k) + \frac{|e_k^T z| \|f_k\|_2}{\|Q_k z\|_2},$$

as expected.  $\square$

We make the following remarks:

1. Since  $|e_k^T z| \leq 1$  and  $\|Q_k z\|_2 \geq \sigma_{\min}(Q_k)$ , (2.3) is a consequence of (2.4).
2. We argue that for a small  $k$ ,  $Q_k$ 's columns are usually nearly orthonormal, which ensures both  $\sigma_{\min}(Q_k)$  and  $\|Q_k z\|_2$  are 1 or almost 1. Therefore the right-hand sides of (2.3) and of (2.4) are essentially

$$\lambda_{\max}(T_k) + \|f_k\|_2, \quad (2.5)$$

$$\lambda_{\max}(T_k) + |e_k^T z| \|f_k\|_2, \quad (2.6)$$

respectively, for the practical purpose. This, in a way, justifies the validity of using  $\lambda_{\max}(T_k) + \|f_k\|_2$  as an upper bound on  $\lambda_{\max}(A)$  in [18].

3. There is no assumption made in Theorems 1 and 2 for  $\lambda_{\max}(A)$  to be a simple eigenvalue. It can be a multiple eigenvalue.
4. Although our goal is to have upper bound estimates for  $\lambda_{\max}(A)$ , one readily has  $\lambda_{\max}(A) \geq \lambda_{\max}(T_k)$  [13,15], giving a lower bound for  $\lambda_{\max}(A)$ .
5. Similar statements hold for  $\lambda_{\min}(A) = \lambda_1$ . By applying the results above to  $-A$ , we can have economical ways to bound  $\lambda_{\min}(A)$ , both from below and from above. In fact, with (2.1a) if

$$|\lambda_1 - \lambda_{\min}(T_k)| = \min_i |\lambda_i - \lambda_{\min}(T_k)|,$$

then after safely ignoring the effect of roundoff error  $F_k$ , we have, similarly to Theorems 1 and 2,

$$\lambda_{\min}(T_k) - \frac{\|f_k\|_2}{\sigma_{\min}(Q_k)} \leq \lambda_{\min}(A) \leq \lambda_{\min}(T_k),$$

$$\lambda_{\min}(T_k) - \frac{|e_k^T z| \|f_k\|_2}{\|Q_k z\|_2} \leq \lambda_{\min}(A) \leq \lambda_{\min}(T_k),$$

where  $z$  is the unit eigenvector of  $T_k$  associated with  $\lambda_{\min}(T_k)$ .

It is impractical to verify Assumption (2.2) in actual computations because  $A$ 's eigenvalues  $\lambda_i$  are unknown. But the existing convergence theory (in exact arithmetic) [8,13,14] does suggest that if the initial random vector has nontrivial component in the direction of  $A$ 's eigenvector associated with  $\lambda_{\max}(A)$ , then  $\lambda_{\max}(T_k)$  usually converges to  $\lambda_{\max}(A)$  faster than to any other eigenvalues of  $A$ . In floating point arithmetics, the theory is also supported by years of extensive numerical practices.

Recall that the goal is to obtain a true upper bound. The bound (2.5)  $\lambda_{\max}(T_k) + \|f_k\|_2$  has reliably produced an upper bound on  $\lambda_{\max}(A)$  in all our tests from DFT. The numerical performance of bound (2.5) appeared not dependent on hypothesis (2.2). In fact we tried matrices from fields other than

DFT, we also constructed examples difficult for the Lanczos to achieve convergence for the largest eigenvalue, including matrices whose eigenvalues are highly clustered near the largest eigenvalue while smallest eigenvalues dominate in magnitude, but bound (2.5) always provides an upper bound in just a few Lanczos steps.

However, in cases when hypothesis (2.2) fails, occasionally the bound (2.6)  $\lambda_{\max}(T_k) + |e_k^T z| \|f_k\|_2$  derived from Theorem 2 may underestimate  $\lambda_{\max}(A)$ . When it does not underestimate, it gives the sharpest upper bound among all. For the task of robustly providing an upper bound, we propose the following modification to (2.6), which utilizes not just the eigenvector associated with  $\mu_k$ , but instead all the eigenvectors of  $T_k$ . We arrive at the following bound

$$\lambda_{\max}(A) \leq \lambda_{\max}(T_k) + \max_z |e_k^T z| \|f_k\|_2, \tag{2.7}$$

where  $z$  is any unit eigenvector of  $T_k$ . Bound (2.7) works surprisingly well in our tests. One explanation is that the right-hand side of (2.7) is always no smaller than that of (2.6), which means (2.2) in Theorem 2 may be relaxed. Another explanation is that, the scaling factor for  $\|f_k\|_2$  in Theorem 2 is  $\frac{|e_k^T z|}{\|Q_k z\|_2}$ , hence using a scaling factor no less than  $|e_k^T z|$  may provide enhanced safeguard for the rare cases in floating point arithmetic that  $\frac{1}{\|Q_k z\|_2}$  may be close to 1 but smaller than 1. But we feel that a more satisfactory understanding on the choice of  $\max_z |e_k^T z|$  may require further study.

Intuitively, among all the eigenvectors of  $T_k$ , those associated with Ritz values close to  $\lambda_{\max}(T_k)$  (i.e.,  $\mu_{k-2}, \mu_{k-1}, \mu_k$ ) can be more important than those associated with smaller Ritz values. Therefore we can sharpen (2.7) as

$$\lambda_{\max}(A) \leq \lambda_{\max}(T_k) + \max_{z \in \{z_{k-2}, z_{k-1}, z_k\}} |e_k^T z| \|f_k\|_2, \tag{2.8}$$

where  $z_{k-2}, z_{k-1}, z_k$  are the unit Ritz vectors associated with  $\mu_{k-2}, \mu_{k-1}, \mu_k$ , respectively.

We avoid computing any term that contains the Lanczos vectors  $Q_k$  in all of the bounds (2.5)–(2.8). This is mainly for saving computational cost. Since  $Q_k$  is of size  $n$  by  $k$  where  $n$  is usually huge, estimating either  $\sigma_{\min}(Q_k)$  or  $\|Q_k z\|_2$  involves non-negligible cost. Moreover,  $Q_k$  does not need to be stored in the Lanczos bound estimator. As discussed above, when  $k$  is small and convergence just starts to happen to a couple of digits at most, we can practically use  $\sigma_{\min}(Q_k) \approx \|Q_k z\|_2 \approx 1$ .

Clearly, if hypothesis (2.2) holds, the bound (2.6) from Theorem 2 is the sharpest, followed by (2.8), then (2.7), and then (2.5). The latter three provide increasingly stronger safeguard to (2.6) in case (2.2) does not hold.

We note that computing eigenvectors of a  $k \times k$  Hermitian tri-diagonal matrix  $T_k$  for small  $k$  constitutes only negligible cost, therefore the cost for bounds (2.6), (2.8), and (2.7) are essentially the same as that of (2.5).

### 3. Numerical study of the bounds

In this section we provide numerical study of the bounds (2.5)–(2.8). We will see that the bound (2.6) works nicely most of the time, especially for problems from real applications. This is mainly due to the efficiency of the Lanczos method that can quickly approximate the exterior eigenvalues, which makes (2.2) valid most of the time, even for a small  $k$ .

The presented results, Figs. 3.1–3.5, are selected from a large number of tests using matrices<sup>3</sup> from small scale real-space DFT calculations. The dimension of each matrix is noted on the title of each figure. The bounds as functions of the Lanczos step  $k$  shown in the figures are representative behaviors of each bound.

In the legend of each figure, bnd1, bnd2, bnd3, and bnd4 refer to (2.5), (2.6), (2.7), and (2.8), respectively. The maxeig refers to  $\lambda_{\max}(A)$ . The bound (2.6) is often so sharp that we opt not to use a special symbol for bnd2 so that it does not severely block the line for maxeig.

Five dotted vertical lines are drawn for  $k = 4, 5, 6, 7, 8$ , as references to show that the safeguarded bounds quickly become upper bounds of  $\lambda_{\max}(A)$  even after only 4 or 5 Lanczos iterations.

<sup>3</sup> The matrices are available at the University of Florida Sparse Matrix Collection [4] under the group name "PARSEC".

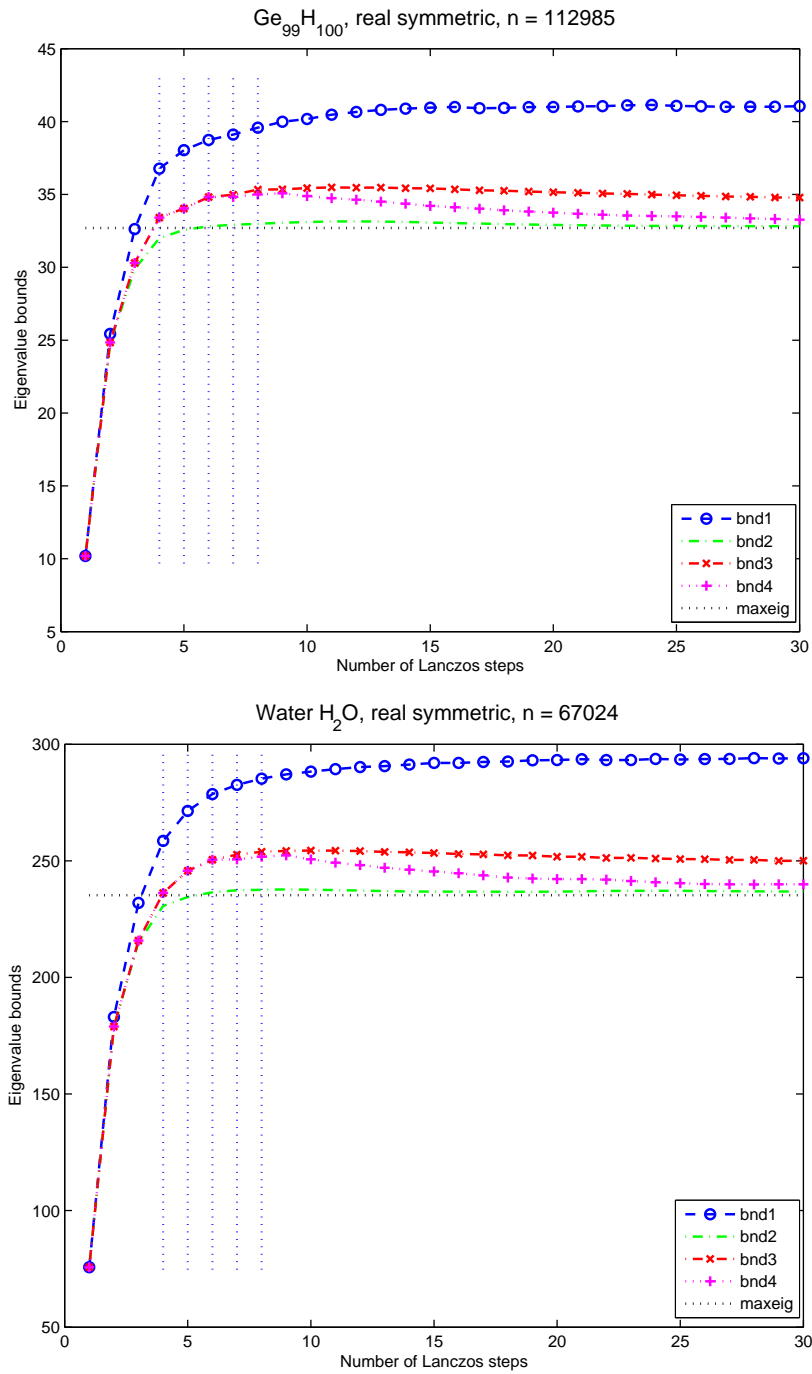


Fig. 3.1. Behavior of bounds on a hydrogen passivated germanium clusters Ge<sub>99</sub>H<sub>100</sub> and a water molecule.

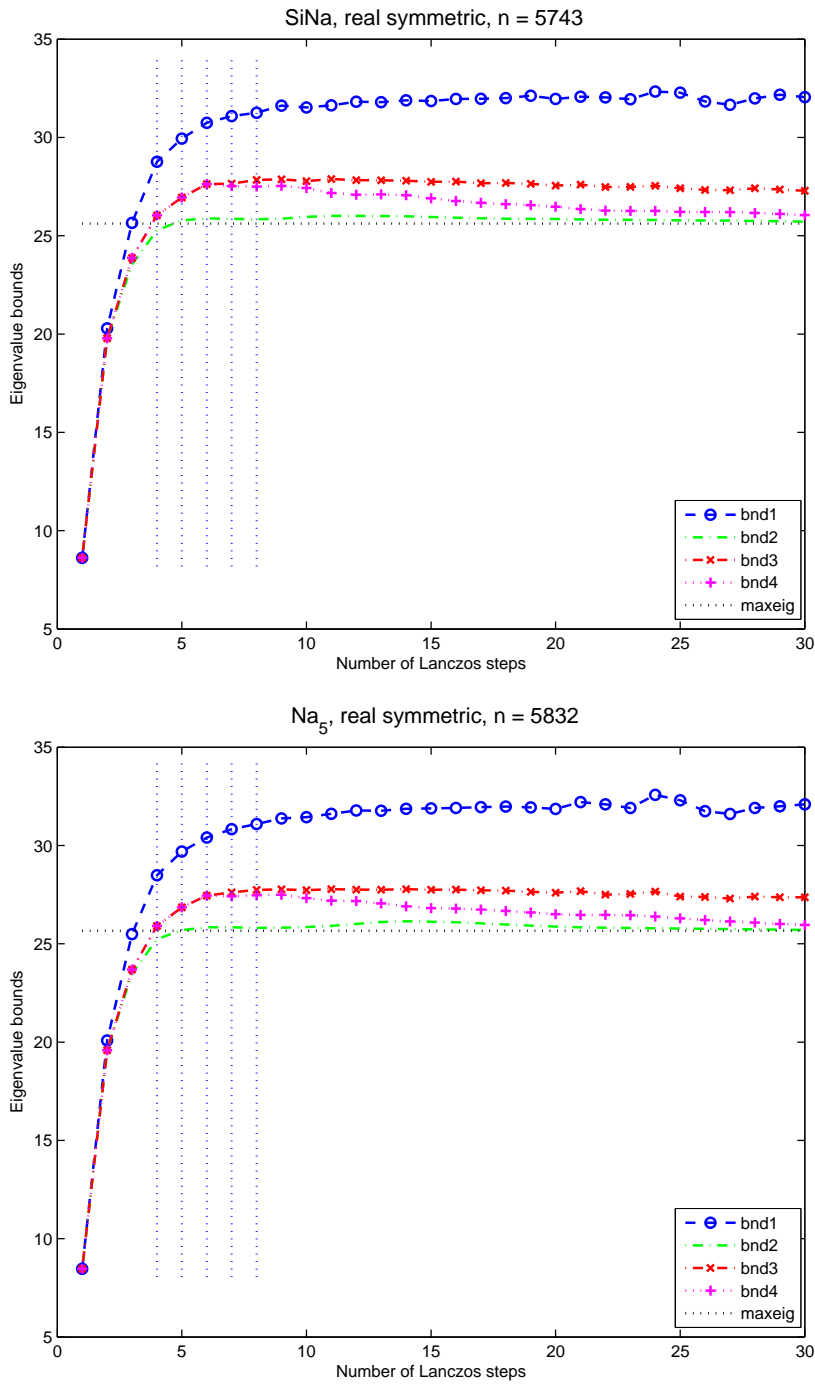


Fig. 3.2. Behavior of bounds on SiNa and Na<sub>5</sub>.

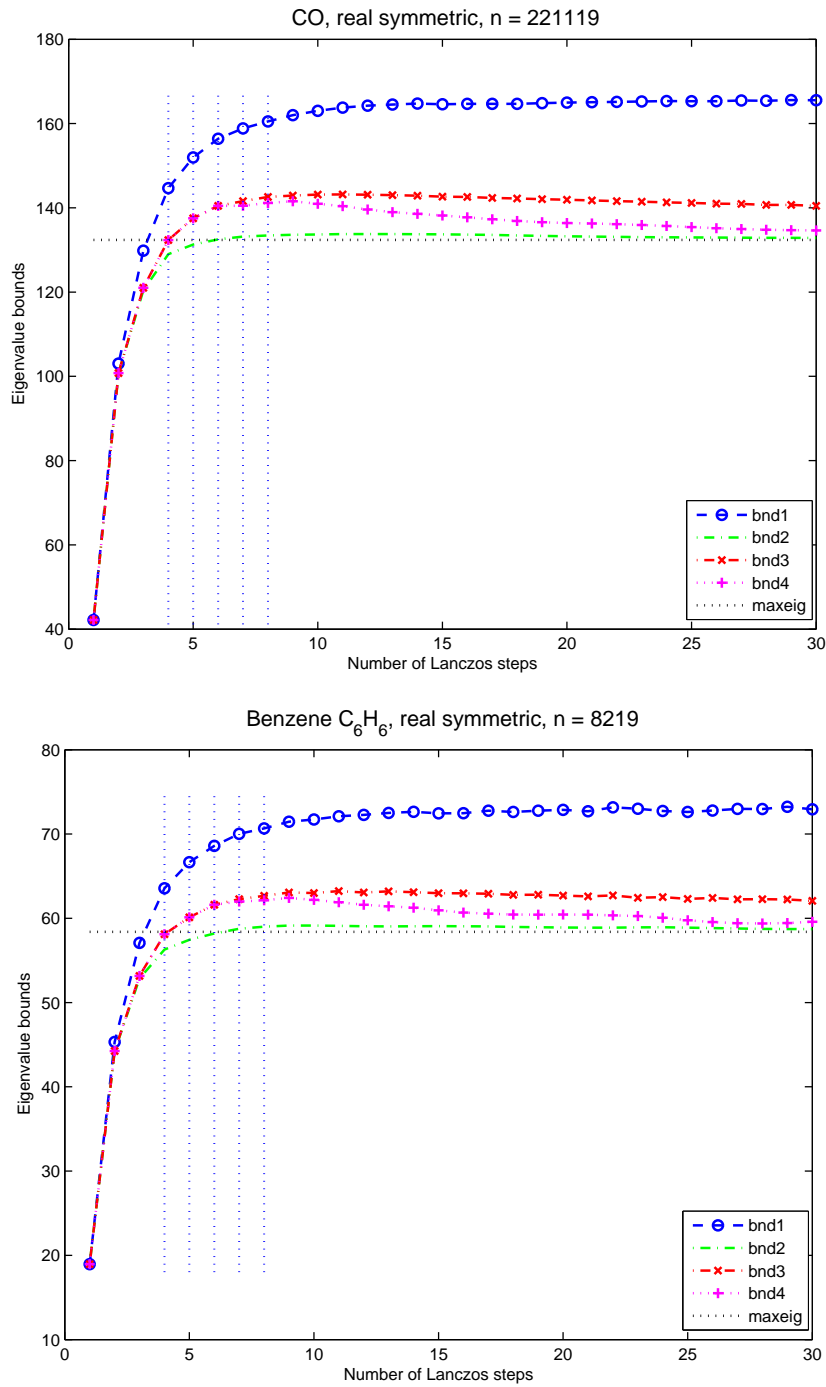
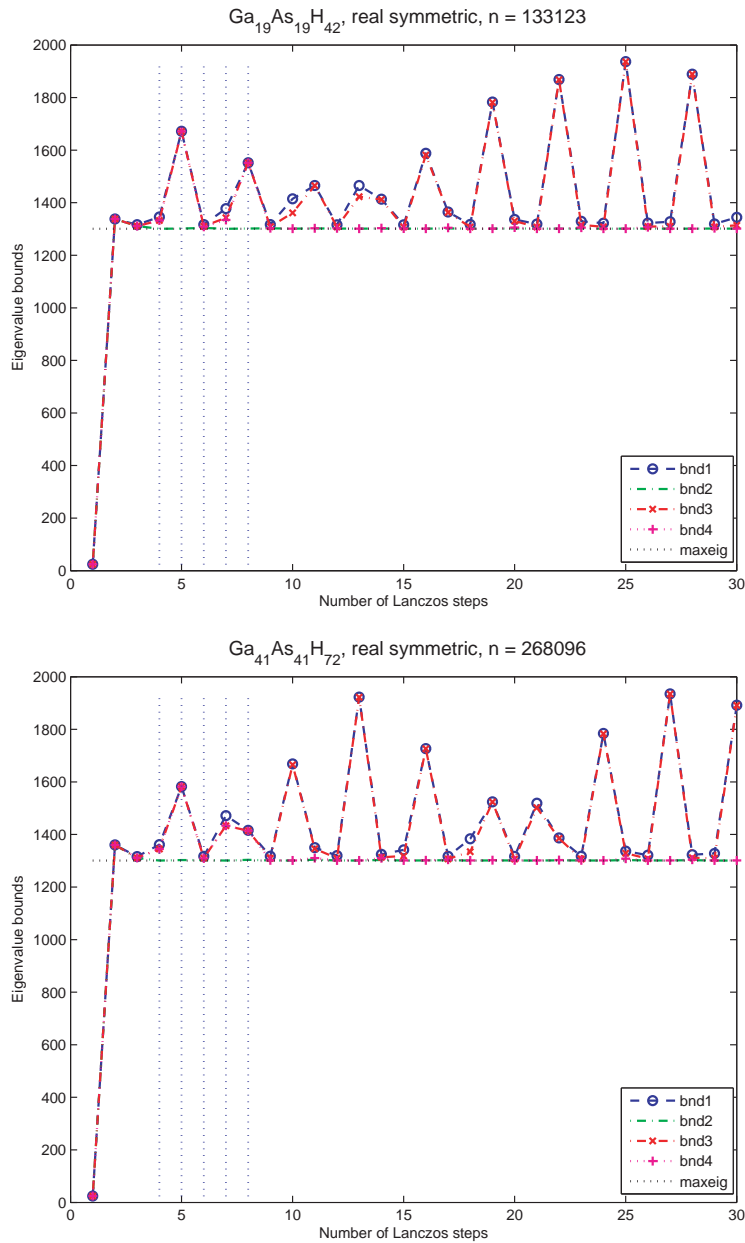


Fig. 3.3. Behavior of bounds on a carbon monoxide CO molecule and the Benzene C<sub>6</sub>H<sub>6</sub> molecule.



**Fig. 3.4.** Behavior of bounds on two hydrogen passivated gallium arsenide clusters Ga<sub>19</sub>As<sub>19</sub>H<sub>42</sub> and Ga<sub>41</sub>As<sub>41</sub>H<sub>72</sub>. These two examples have a distinct feature: three largest eigenvalues dominate in magnitude and are clustered. As seen from this plot, the safeguards used for bnd1 and bnd3 are conservative and may give too large upper bound at some steps. Clearly for these two bounds more Lanczos steps do not translate into sharper bounds. While bnd2 and bnd4 very much collapse onto maxeig after  $k = 4$ , which also shows that more Lanczos steps are not necessary if an appropriate safeguard is applied.

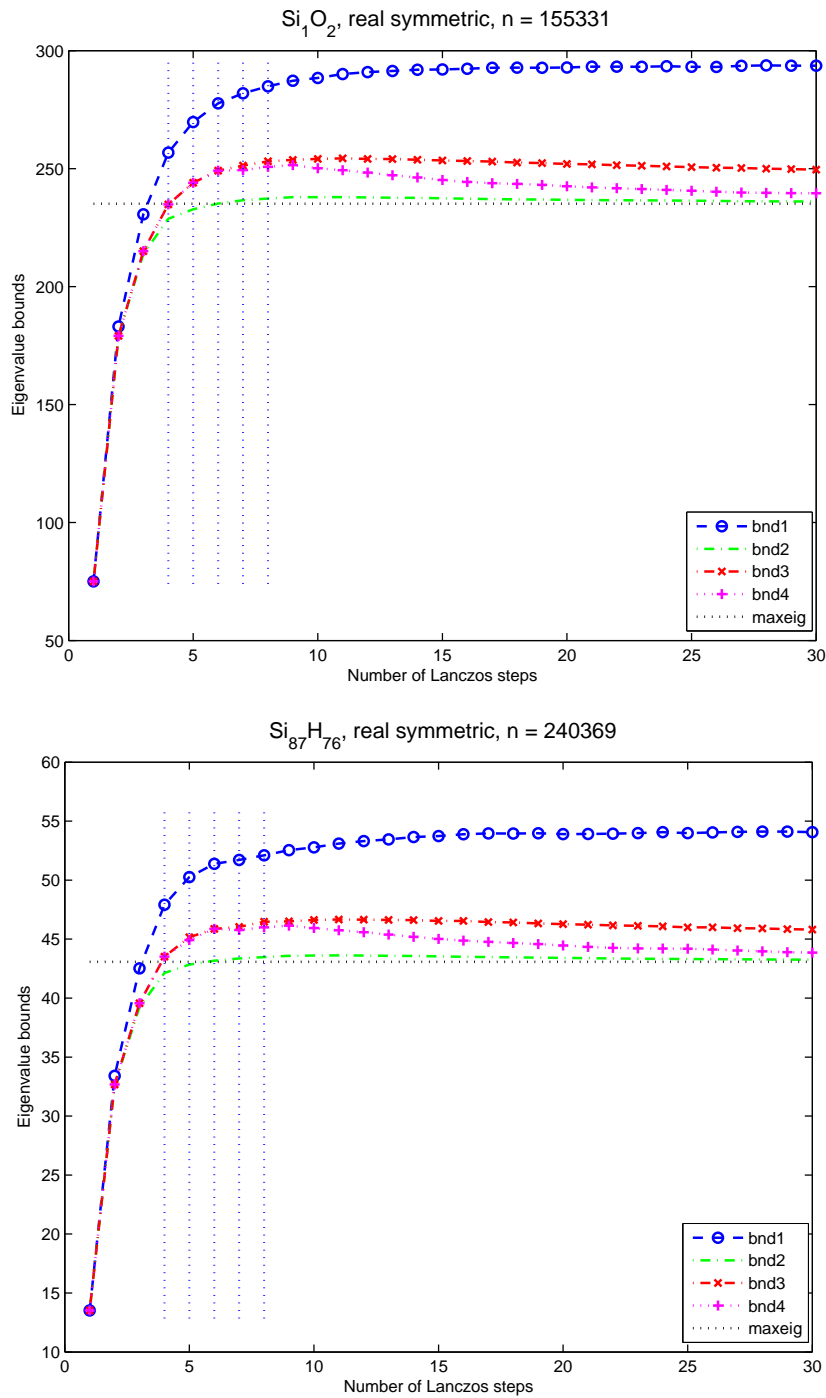
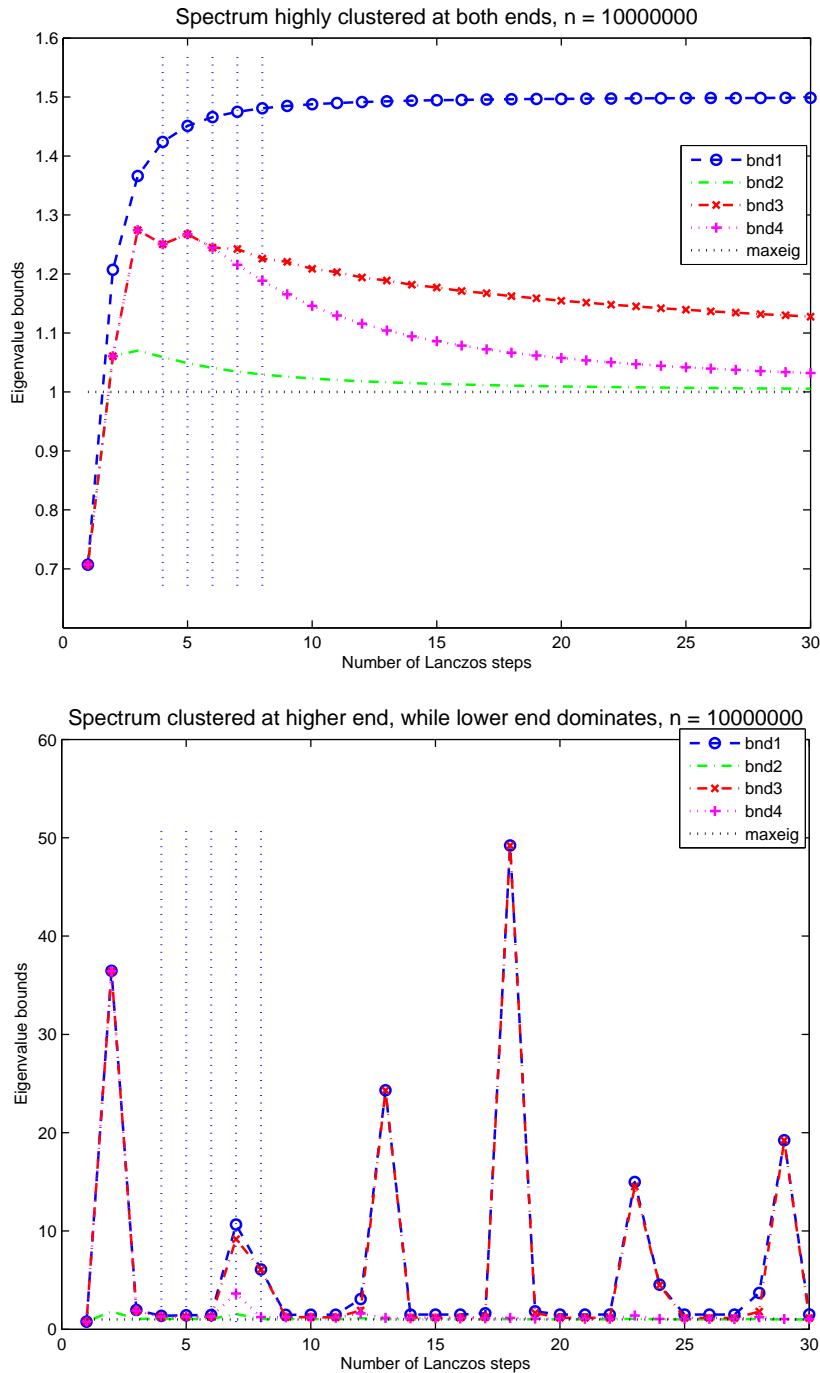
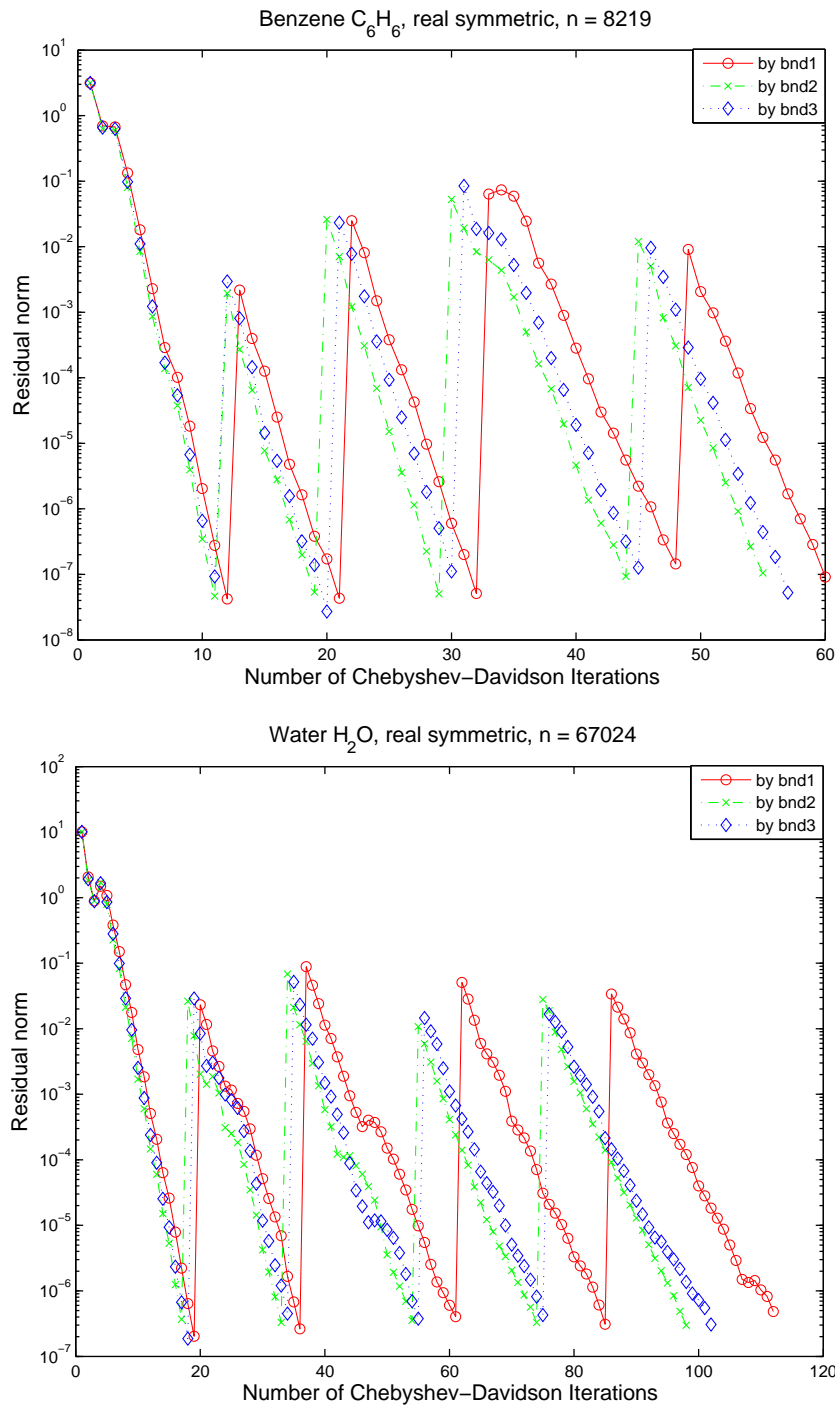


Fig. 3.5. Behavior of bounds on silicon dioxide SiO<sub>2</sub> and hydrogen passivated silicon cluster Si<sub>87</sub>H<sub>76</sub>.



**Fig. 3.6.** Top: Spectrum clustered at both ends. Bottom: Smallest eigenvalues dominate in magnitude. From the figure at bottom, we again observe that the safeguards for  $bnd1$  and  $bnd3$  can be too conservative at certain steps. This again is due to the unbalanced dominance of one end of the spectrum. While  $bnd2$  provides tight upper bounds consistently with  $k$  increasing.



**Fig. 3.7.** Converging five eigenvalues and eigenvectors with a relative tolerance of  $1e-7$  using different upper bounds, with all other conditions held the same.

The plots are for  $k$  from 1 to 30. This is for illustration purpose. In real computations, one seldom needs to go over 15 Lanczos steps. With safeguards, normally 7–9 steps are good enough for an effective upper bound. Moreover, from the plots we see that performing more Lanczos steps does not necessarily improve the bound much. The more cost-effective approach is to perform less than 10 Lanczos steps and apply appropriate safeguards. Another point worth mentioning is that, for  $k < 4$ ,  $\lambda_{\max}(T_k)$  has not become a reasonable approximation to  $\lambda_{\max}(A)$ , hence (2.2) can be completely wrong. In this case the safeguards are not strong enough to provide an upper bound. It is necessary to perform a few more Lanczos steps so that (2.2) becomes reasonable, which means the bound (2.6) becomes right or almost right, then the safeguards in (2.8) or (2.7) will provide a safe upper bound that is sharper than (2.5).

We use a nontrivial example to show that our bounds can handle difficult problems. Fig. 3.6 shows the behavior of bounds for two  $10^7 \times 10^7$  diagonal matrices. (Unitary similarity transformation of a diagonal matrix into a nondiagonal matrix does not affect the behavior of Lanczos, so using diagonal matrices does not lose generality.)

The eigenvalues of the matrix for the left plot of Fig. 3.6 are the Chebyshev zeros on  $[-1, 1]$ , i.e.,

$$\lambda_{n-k+1} = \cos\left(\left(k - \frac{1}{2}\right) \frac{\pi}{n}\right), \quad k = 1, 2, \dots, n; \quad n = 10^7. \quad (3.1)$$

It is known that the eigenvalues are clustered at both ends of the interval  $[-1, 1]$ . But our bounds quickly found upper bounds close to 1 within 4 Lanczos steps, with  $\text{bnd2}$  being really sharp even with a small  $k$  such as  $k = 7$ .

The bottom plot of Fig. 3.6 uses a modified matrix: we multiply the smallest 100 eigenvalues from (3.1) by 100 and keep the rest unchanged. Now the eigenvalues are in  $[-100, 1]$ , with the smallest eigenvalues dominating in magnitude and having favorable gaps. Lanczos method should have hard time approximating the largest eigenvalue  $\lambda_n \approx 1$ , relative to converging the smallest ones. But even for this problem, our bounds still found sharp upper bounds in about 5 steps.

We also implemented the new bounds in the Chebyshev–Davidson algorithm [20]. As expected, a sharper upper bound improves the overall performance of this algorithm. Fig. 3.7 contains two typical examples that show the advantage of sharper upper bounds. But we warn that if a bound underestimates the largest eigenvalue, it will result in ineffective Chebyshev filters. In this case unwanted part of spectrum will be magnified instead of dampened, which can lead to very slow convergence or even non-convergence. Appropriate safeguards to ensure upper bounds are essential for constructing effective filters.

#### 4. A practical estimator and concluding remarks

The focus of this note is on cost-effective estimates of an upper bound for the spectrum of large Hermitian matrices. We proposed using very few steps Lanczos iteration with a safeguard step. Four closely related bounds are discussed. The safeguards used in the bounds (2.6), (2.8), (2.7), and (2.5) are of increasing strength. A natural question now is: which bound to use in practice? We provide the following answer: If one is interested in a safe upper bound that can be obtained within as few Lanczos steps as possible, then (2.5) is the choice. Since the cost for computing all these bounds are similar to computing (2.5), another viable and potentially better choice is to combine some of the bounds. E.g., one can compute (2.6) and (2.7) for a small  $k$  and return the average value of the two as an upper bound; or, one can run a small  $k$  step Lanczos and compute (2.8) at each step, then return the largest one as an upper bound. Notice that hypothesis (2.2) has been the main theoretical concern, if in a situation that a vector close to the eigenvector related to  $\lambda_{\max}(A)$  is available, then one can start the Lanczos iteration using this vector instead of a random vector. This will make (2.2) valid in very few Lanczos steps, then (2.6) is the best choice. Such situation exists, for example, when the bound estimator is called within a loop and the previous iteration provides an approximate eigenvector corresponding to the largest eigenvalue.

The above discussion shows that there are a few practical combinations of the bounds (2.5)–(2.8) for constructing an upper bound estimator. We list one of them in Algorithm 1. The  $\text{tol}$  in Algorithm 1 is a user specified tolerance. Clearly, if one prefers a more conservative upper bound, then (2.5) can be return on line (c) in Algorithm 1.

**Algorithm 1.** Estimating an upper bound of a Hermitian matrix  $A$ .

If a vector close to the eigenvector associated with  $\lambda_{\max}(A)$  is available,  
 Run 5 steps Lanczos with this vector as initial and return bound (2.6).  
 Else, pick an integer  $K$  with  $4 < K \leq 8$ ,

- (a) Run 4 steps Lanczos with a random initial vector
- (b) For  $k = 5$  to  $K$  Do
  - Compute the  $k$ -th step Lanczos decomposition; Compute  $|e_k^T z| \|f_k\|$ ;
  - If  $(|e_k^T z| \|f_k\| < tol)$ , return bound (2.8), stop.
- (c) Return the average of bounds (2.6) and (2.7).

The bound estimators proposed here can be extended to situations where one needs to estimate extreme singular values through Lanczos bidiagonalization [5, p. 495]. One possible application is the computation of (nearly) optimal scaling parameters for calculating the polar factor of a matrix by the scaled Newton iteration [6, p. 205].

**Acknowledgements**

We wish to express our sincere appreciation to the two anonymous referees who carefully read through the paper and provided many constructive comments that improve the paper.

**References**

- [1] Z.-H. Cao, J.-J. Xie, R.-C. Li, A sharp version of Kahan's theorem on clustered eigenvalues, *Linear Algebra Appl.* 245 (1996) 147–155.
- [2] J.R. Chelikowsky, N. Troullier, Y. Saad, Finite-difference-pseudopotential method: electronic structure calculations without a basis, *Phys. Rev. Lett.* 72 (1994) 1240–1243.
- [3] J.R. Chelikowsky, A.T. Zayak, T.-L. Chan, M.L. Tiago, Y. Zhou, Y. Saad, Algorithms for the electronic and vibrational properties of nanocrystals, *J. Phys. Condens. Matter* 21 (2009) 06427:1–06427:7.
- [4] T.A. Davis, The University of Florida Sparse Matrix Collection, TOMS, submitted for publication.
- [5] G.H. Golub, C.F. Van Loan, *Matrix Computations*, third ed., Johns Hopkins University Press, Baltimore, Maryland, 1996.
- [6] N.J. Higham, *Functions of Matrices: Theory and Computation*, SIAM Press, Philadelphia, 2008.
- [7] W. Kahan, Inclusion Theorems for Clusters of Eigenvalues of Hermitian Matrices, Technical Report, Computer Science Department, University of Toronto, 1967.
- [8] S. Kaniel, Estimates for some computational techniques in linear algebra, *Math. Comp.* 20 (1966) 369–378.
- [9] D.E. Knuth, *The Art of Computer Programming II: Seminumerical Algorithms*, third ed., Addison-Wesley, Reading, Massachusetts, 1997.
- [10] J. Kuczyński, H. Woźniakowski, Estimating the largest eigenvalues by the power and Lanczos algorithms with a random start, *SIAM J. Matrix Anal. Appl.* 13 (1992) 1094–1122.
- [11] D.P. O'Leary, G.W. Stewart, J.S. Vandergraft, Estimating the largest eigenvalue of a positive definite matrix, *Math. Comp.* 33 (1979) 1289–1292.
- [12] B.N. Parlett, H. Simon, L.M. Stringer, On estimating the largest eigenvalue with the Lanczos algorithm, *Math. Comp.* 38 (1982) 153–165.
- [13] B.N. Parlett, *The Symmetric Eigenvalue Problem*, SIAM, Philadelphia, 1998.
- [14] Y. Saad, On the rates of convergence of the Lanczos and the block-Lanczos methods, *SIAM J. Numer. Anal.* 15 (1980) 687–706.
- [15] G.W. Stewart, J.G. Sun, *Matrix Perturbation Theory*, Academic Press Inc., Boston, 1990.
- [16] G.W. Stewart, *Matrix Algorithms II: Eigensystems*, SIAM Press, Philadelphia, 2001.
- [17] M.L. Tiago, Y. Zhou, M. Alemany, Y. Saad, J.R. Chelikowsky, Evolution of magnetism in iron from the atom to the bulk, *Phys. Rev. Lett.* 97 (2006) 147201-1–147201-4.
- [18] Y. Zhou, Y. Saad, M.L. Tiago, J.R. Chelikowsky, Self-consistent-field calculations using Chebyshev-filtered subspace iteration, *J. Comput. Phys.* 219 (2006) 172–184.
- [19] Y. Zhou, Y. Saad, M.L. Tiago, J.R. Chelikowsky, Parallel self-consistent-field calculations using Chebyshev-filtered subspace acceleration, *Phys. Rev. E* 74 (2006) 066704-1–066704-8.
- [20] Y. Zhou, Y. Saad, A Chebyshev–Davidson algorithm for large symmetric eigenvalue problems, *SIAM J. Matrix Anal. Appl.* 29 (2007) 954–971.