ELSEVIER

Linear Algebra and its Applications 324 (2001) 209–226

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Computational error bounds for multiple or nearly multiple eigenvalues

Siegfried M. Rump

Institut für Informatik III, Technical University Hamburg-Harburg, Schwarzenbergstr. 95, D-21071 Hamburg, Germany

Received 7 July 1999; accepted 16 October 2000

Submitted by T. Yamamoto

Abstract

In this paper bounds for clusters of eigenvalues of non-selfadjoint matrices are investigated. We describe a method for the computation of rigorous error bounds for multiple or nearly multiple eigenvalues, and for a basis of the corresponding invariant subspaces. The input matrix may be real or complex, dense or sparse. The method is based on a quadratically convergent Newton-like method; it includes the case of defective eigenvalues, uncertain input matrices and the generalized eigenvalue problem. Computational results show that verified bounds are still computed even if other eigenvalues or clusters are nearby the eigenvalues under consideration. © 2001 Elsevier Science Inc. All rights reserved.

AMS classification: 15A18; 65G10

Keywords: Algebraic eigenvalue problem; Multiple eigenvalues; Defective eigenvalues; Validated bounds

1. Notation

Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Throughout the paper we denote by $A \in M_n(\mathbb{K})$ an $n \times n$ matrix, by $\tilde{X} \in M_{n,k}(\mathbb{K})$ an $n \times k$ approximation to a k-dimensional invariant subspace corresponding to a multiple or a cluster of eigenvalues near some $\tilde{\lambda} \in \mathbb{K}$, such that $A\tilde{X} \approx \tilde{\lambda}\tilde{X}$. The purpose of the paper is to derive bounds for k eigenvalues and a

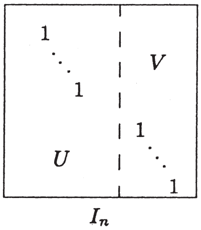
E-mail address: rump@tu-harburg.de (S.M. Rump).

0024-3795/01/$ - see front matter © 2001 Elsevier Science Inc. All rights reserved.
PII: S 0 0 2 4 - 3 7 9 5 (0 0) 0 0 2 7 9 - 2

k-dimensional invariant subspace of A based on the approximations $\tilde{\lambda}$ and $\tilde{X}$. We stress that neither a priori assumptions on the quality of the approximations $\tilde{X}$ and $\tilde{\lambda}$ are made, nor the matrix A is assumed to be symmetric or Hermitian. The computed bounds are always rigorous. If the approximations are too poor and/or a cluster of k eigenvalues near $\tilde{\lambda}$ is not well enough separated from the rest of the spectrum, it may happen that no error bounds can be calculated and a corresponding message is given—rather computing erroneous results.

The computed error bounds are rigorous including all possible computational errors. Note that the letter n is reserved for the dimension of the matrix, and the letter k is reserved for the dimension of the invariant subspace. Also note that the size k of the cluster has to be specified (see also Section 6).

The degree of arbitrariness is removed by freezing k rows of the approximation $\tilde{X}$. If the set of these rows is denoted by v, and by definition $u := \{1, \dots n\} \backslash v$, then throughout the paper we denote by $U \in M_{n,n-k}(\mathbb{R})$ the submatrix of the identity matrix with columns in u. Correspondingly, we define $V \in M_{n,k}(\mathbb{R})$ to comprise of the columns in v out of the identity matrix. Denoting the $n \times n$ identity matrix by I_n, it is $UU^{\mathrm{T}} + VV^{\mathrm{T}} = I_n$, and $V^{\mathrm{T}}\tilde{X} \in M_k(\mathbb{K})$ is the normalizing part of $\tilde{X}$. Note that $U^{\mathrm{T}}U = I_{n-k}$ and $V^{\mathrm{T}}V = I_k$. For example, for $u = \{1, \dots, n-k\}$, $v = \{n-k+1, \dots, n\}$ the situation is as follows.



We will use $n \times k$ interval matrices $\mathbf{X} \in \mathbb{I}M_{n,k}(\mathbb{K})$. Interval quantities are always denoted in bold face. Interval quantities may be represented by infimum and supremum or by midpoint and radius. The fundamental assumption on the interval operations we are using is inclusion isotonicity, that is,

$$\forall F \in \mathbf{F},\ \forall G \in \mathbf{G}: \quad F \circ G \in \mathbf{F} \circ \mathbf{G} \tag{1}$$

for all suitable interval quantities $\mathbf{F}$, $\mathbf{G}$ and suitable operations $\circ \in \{+, -, \cdot\}$. For an introduction to interval arithmetic, see for example [3,16]; cf. also the first paper "Self-validating methods" in this special issue. There are a number of interval packages, among them [1,5,8–11]. An interval package accessible from Matlab is described in [20]. For a fast implementation, which is used in [20], of interval arithmetic for general purpose computers as well as for parallel computers, see [19].

2. A Newton-like iteration

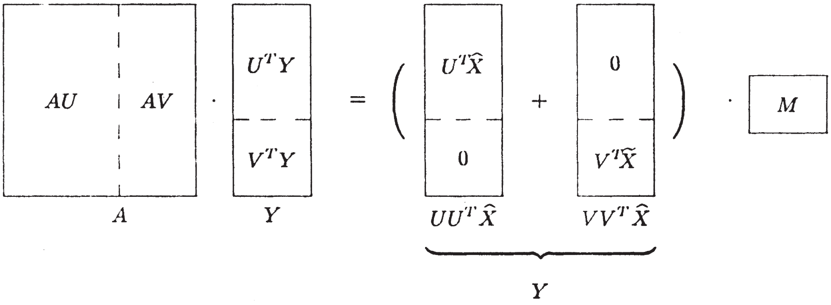
Methods for computing rigorous error bounds are frequently based on some fixed point iteration. A reformulation being suitable for computation with sets (intervals) and for application of Brouwer's fixed point theorem leads to rigorous error bounds.

A number of papers deal with Newton-like improvements of a simple eigenvalue and corresponding eigenvector, among them [22,23]. Dongarra et al. [7] extend the methods to multiple eigenvalues. In the following we present a variant for clusters of eigenvalues and corresponding eigenvectors. It is a numerical method for improving given approximations $\tilde{X}$ and $\tilde{\lambda}$. However, it will turn out to be a suitable basis for a self-validating method to be presented in Section 3.

For given $\tilde{X} \in M_{n,k}(\mathbb{K})$, $\tilde{\lambda} \in \mathbb{K}$, suppose

$$AY = YM \quad \text{for } Y \in M_{n,k}(\mathbb{K}), \quad M \in M_k(\mathbb{K}), \tag{2}$$

such that Y and $\tilde{X}$ coincide in the normalizing part of $\tilde{X} : V^{\mathrm{T}}Y = V^{\mathrm{T}}\tilde{X}$. We collect the unknown quantities $U^{\mathrm{T}}Y$ and M into $\widehat{X} \in M_{n,k}(\mathbb{K})$. In other words, we anticipate computation of $\widehat{X}$ with $U^{\mathrm{T}}\widehat{X} = U^{\mathrm{T}}Y$ and $V^{\mathrm{T}}\widehat{X} = M$. Note that we do not assume M to be diagonal. For $u = \{1, \ldots, n-k\}$, $v = \{n-k+1, \ldots, n\}$ the situation is as follows:



This implies the eigenequation

$$A(UU^{\mathrm{T}}\widehat{X} + VV^{\mathrm{T}}\tilde{X}) = (UU^{\mathrm{T}}\widehat{X} + VV^{\mathrm{T}}\tilde{X})V^{\mathrm{T}}\widehat{X}, \tag{3}$$

such that, according to (2), $Y = UU^{\mathrm{T}}\widehat{X} + VV^{\mathrm{T}}\tilde{X}$ and $M = V^{\mathrm{T}}\widehat{X}$. Consider the following iteration scheme (tr denotes the trace).

Algorithm 2.1. *Newton-like iteration for eigenvalue clusters*

```
$X^0 := UU^{\mathrm{T}}\tilde{X} + \tilde{\lambda}V$
for $\nu = 0, 1, \ldots$
  $\lambda_\nu := \mathrm{tr}(V^{\mathrm{T}}X^\nu)/k$,
  $C_\nu := (A - \lambda_\nu I_n)UU^{\mathrm{T}} - (UU^{\mathrm{T}}X^\nu + VV^{\mathrm{T}}\tilde{X})V^{\mathrm{T}}$,
  $X^{\nu+1} := UU^{\mathrm{T}}X^\nu + \lambda_\nu V - C_\nu^{-1} \cdot (A - \lambda_\nu I_n)(UU^{\mathrm{T}}X^\nu + VV^{\mathrm{T}}\tilde{X})$.
```

The iteration will be shown to converge to $\tilde{X}$ under certain conditions. The common "approximation" λ_ν to the cluster is adjusted in every iteration to be the mean value of the eigenvalues of $V^{\mathrm{T}}X^\nu$.

The computation of error bounds for clustered eigenvalues, the goal of this paper, will be based on this iteration. In the following we prove quadratic convergence for non-defective multiple eigenvalues, and geometeric convergence for clusters or defective multiple eigenvalues.

Using (3), $U^{\mathrm{T}}V = V^{\mathrm{T}}U = 0$ and $V^{\mathrm{T}}V = I_k$, we will show

$$C_\nu(X^{\nu+1} - \widehat{X}) = UU^{\mathrm{T}}(X^\nu - \widehat{X})(V^{\mathrm{T}}\widehat{X} - \lambda_\nu I_k). \tag{4}$$

Indeed it is

$$\begin{aligned}
&C_\nu(X^{\nu+1} - \widehat{X})\\
&\quad = C_\nu(X^{\nu+1} - UU^{\mathrm{T}}X^\nu - \lambda_\nu V) + C_\nu(UU^{\mathrm{T}}X^\nu + \lambda_\nu V - \widehat{X})\\
&\quad = -(A - \lambda_\nu I_n)(UU^{\mathrm{T}}X^\nu + VV^{\mathrm{T}}\tilde{X}) + (A - \lambda_\nu I_n)UU^{\mathrm{T}}(X^\nu - \widehat{X})\\
&\qquad -(UU^{\mathrm{T}}X^\nu + VV^{\mathrm{T}}\tilde{X})V^{\mathrm{T}}(\lambda_\nu V - \widehat{X})\\
&\quad = -(A - \lambda_\nu I_n)(UU^{\mathrm{T}}\widehat{X} + VV^{\mathrm{T}}\tilde{X})\\
&\qquad +(UU^{\mathrm{T}}X^\nu + VV^{\mathrm{T}}\tilde{X})(V^{\mathrm{T}}\widehat{X} - \lambda_\nu I_k)\\
&\quad = -(UU^{\mathrm{T}}\widehat{X} + VV^{\mathrm{T}}\tilde{X})(V^{\mathrm{T}}\widehat{X} - \lambda_\nu I_k)\\
&\qquad +(UU^{\mathrm{T}}X^\nu + VV^{\mathrm{T}}\tilde{X})(V^{\mathrm{T}}\widehat{X} - \lambda_\nu I_k)\\
&\quad = UU^{\mathrm{T}}(X^\nu - \widehat{X})(V^{\mathrm{T}}\widehat{X} - \lambda_\nu I_k).
\end{aligned} \tag{5}$$

Using the column sum norm $\|\cdot\|_1$, denote

$$\begin{aligned}
&\alpha_\nu := \|X^\nu - \widehat{X}\|_1,\\
&\hat{\lambda} := \mathrm{tr}(V^{\mathrm{T}}\widehat{X})/k,\\
&\widehat{C} := (A - \hat{\lambda}I)UU^{\mathrm{T}} - (UU^{\mathrm{T}}\widehat{X} + VV^{\mathrm{T}}\tilde{X})V^{\mathrm{T}},\\
&E_\nu := C_\nu - \widehat{C} = (\hat{\lambda} - \lambda_\nu)UU^{\mathrm{T}} + UU^{\mathrm{T}}(\widehat{X} - X^\nu)V^{\mathrm{T}}.
\end{aligned}$$

Then the following theorem shows geometric convergence in the general case.

Theorem 2.2. *With the above notations assume $\widehat{C}$ to be non-singular, and for $\|\cdot\| = \|\cdot\|_1$ assume*

$$q := 2\|\widehat{C}^{-1}\|(\|V^{\mathrm{T}}\widehat{X} - \hat{\lambda}I\| + \|X^0 - \widehat{X}\|) < 1.$$

Then every C_ν is invertible, and the iteration defined by Algorithm 2.1 *converges geometrically to $\widehat{X}$ with*

$$\|X^\nu - \widehat{X}\| \leq q^\nu \cdot \|X^0 - \widehat{X}\|.$$

Proof. It is $|\hat{\lambda} - \lambda_\nu| \leqslant \max|(V^{\mathrm{T}}\widehat{X})_{ii} - (V^{\mathrm{T}}X^\nu)_{ii}| \leqslant \|V^{\mathrm{T}}(\widehat{X} - X^\nu)\| \leqslant \alpha_\nu$ and $\|E_\nu\|_1 \leqslant \max(|\tilde{\lambda} - \lambda_\nu|, \|\widehat{X} - X^\nu\|) \leqslant \|\widehat{X} - X^\nu\| = \alpha_\nu$. We proceed by induction using $\|X^0 - \widehat{X}\| = \alpha_0$, and assume $\|X^\nu - \widehat{X}\| \leqslant q^\nu \cdot \alpha_0$ for some $\nu \in \mathbb{N}$. Using $C_\nu = \widehat{C}(I + \widehat{C}^{-1}E_\nu)$ and $\|\widehat{C}^{-1}E_\nu\| \leqslant \alpha_\nu\|\widehat{C}^{-1}\| \leqslant \alpha_0\|\widehat{C}^{-1}\| < 1/2 < 1$ implies C_ν to be invertible. By (4),

$$\alpha_{\nu+1} = \|X^{\nu+1} - \widehat{X}\| \leqslant \|C_\nu^{-1}\| \cdot \|UU^{\mathrm{T}}(X^\nu - \widehat{X})(V^{\mathrm{T}}\widehat{X} - \lambda_\nu I)\|.$$

Furthermore,

$$\|V^{\mathrm{T}}\widehat{X} - \lambda_\nu I\| \leqslant \|V^{\mathrm{T}}\widehat{X} - \hat{\lambda}I)\| + \alpha_\nu$$

and

$$\|C_\nu^{-1}\| \leqslant \|\widehat{C}^{-1}\|(1 - \|\widehat{C}^{-1}E_\nu\|)^{-1} \leqslant 2\|\widehat{C}^{-1}\|.$$

Putting things together and using $\alpha_\nu \leqslant \alpha_0$ implies

$$\begin{aligned}\alpha_{\nu+1} &= \|X^{\nu+1} - \widehat{X}\| \leqslant 2 \cdot \|\widehat{C}^{-1}\| \cdot \alpha_\nu \cdot (\|V^{\mathrm{T}}\widehat{X} - \hat{\lambda}I\| + \alpha_\nu) \\ &\leqslant q \cdot \alpha_\nu < \alpha_\nu. \qquad \square\end{aligned} \tag{6}$$

The quality of q as defined in Theorem 2.2 decreases with the size of $\|V^{\mathrm{T}}\widehat{X} - \hat{\lambda}I\|$. For a cluster of simple and for multiple but non-defective eigenvalues, q is small. For defective eigenvalues, however, $q < 1$ is hardly satisfied. Nevertheless, practical experience shows that Algortihm 2.1 converges well for reasonably separated and not too defective eigenvalues. This can be theoretically justified. However, we refrain to reformulate Theorem 2.2 because the necessary conditions are a bit involved. For one multiple non-defective eigenvalue we obtain quadratic convergence.

Corollary 2.3. *With the above notations assume $\widehat{C}$ to be non-singular, and assume $V^{\mathrm{T}}\widehat{X}$ to be diagonal with k-fold eigenvalue $\tilde{\lambda}$, i.e. $V^{\mathrm{T}}\widehat{X} = \tilde{\lambda}I$. If*

$$2\|\widehat{C}^{-1}\| \cdot \|X^0 - \widehat{X}\| < 1,$$

then every C_ν is invertible, and the iteration defined by Algorithm 2.1 *converges quadratically to $\widehat{X}$ with*

$$\|X^{\nu+1} - \widehat{X}\| \leqslant 2\|\widehat{C}^{-1}\| \cdot \|X^\nu - \widehat{X}\|^2.$$

Proof. Proceeding like in the proof of Theorem 2.2 and inserting $\|V^{\mathrm{T}}\widehat{X} - \hat{\lambda}I\| = 0$ into (6) yields the result. □

As a difference to iterations presented in [7] we note the following. Over there the next iterate is computed as the solution of a linear system, whereas in Algorithm 2.1 the solution of a linear system with system matrix C_ν is the *correction* to the previous iterate. This may be of computational advantage.

3. Error bounds for eigenvalue clusters

There are a number of papers considering the problem of computing rigorous error bounds for a *simple* eigenvalue and corresponding eigenvector, among them [12,14,17,23]. To the author's knowledge there is only one paper for computing error bounds for *double* eigenvalues [4], and no method for rigorous error bounds of multiple eigenvalues of general matrices.

To develop such a method we need the following lemma for intervals, which was given in a more general form in [17]. For completeness, we give a simple proof which is due to Alefeld [2] (int denotes the interior of a set, ρ the spectral radius).

Lemma 3.1. *Let* $\mathbf{Z}, \mathbf{X} \in \mathbb{IK}^n$ *and* $\mathbf{C} \in \mathbb{I}M_n(\mathbb{K})$ *be given. Suppose (using interval operations)*

$$\mathbf{Z} + \mathbf{C} \cdot \mathbf{X} \subseteq \operatorname{int}(\mathbf{X}). \tag{7}$$

Then every $C \in \mathbf{C}$ *is convergent, i.e.* $\rho(C) < 1$.

Proof. For every fixed $Z \in \mathbf{Z}$, $C \in \mathbf{C}$, the inclusion isotonicity (1) implies $Z + C \cdot \mathbf{X} \subseteq \operatorname{int}(\mathbf{X})$. For the midpoint vector mX and radius vector rX of $\mathbf{X}$, it is $\mathbf{X} = [mX - rX, mX + rX] = \{x \in \mathbb{R}^n\colon mX - rX \leqslant x \leqslant mX + rX\}$ with entrywise comparisons. Therefore, using entrywise absolute values, it is $C \cdot \mathbf{X} = C \cdot mX + [-|C| \cdot rX, |C| \cdot rX]$. Hence, (7) implies $mX - rX < Z + C \cdot mX - |C| \cdot rX \leqslant Z + C \cdot mX + |C| \cdot rX < mX + rX$, and therefore $|C| \cdot rX < rX$. By Perron–Frobenius theory, $\rho(C) \leqslant \rho(|C|) < 1$. □

With these preliminaries we can prove the following theorem.

Theorem 3.2. *Let* $A \in M_n(\mathbb{K})$, $\tilde{X} \in M_{n,k}(\mathbb{K})$, $\tilde{\lambda} \in \mathbb{K}$, $R \in M_n(\mathbb{K})$ *and* $\mathbf{X} \in \mathbb{I}M_{n,k}(\mathbb{K})$ *be given, and let* U, V *partition the identity matrix as defined in Section* 1. *Define*

$$\begin{aligned} f(\mathbf{X}) :=& -R(A\tilde{X} - \tilde{\lambda}\tilde{X}) \\ &+ \{I - R((A - \tilde{\lambda}I)UU^{\mathrm{T}} - (\tilde{X} + UU^{\mathrm{T}} \cdot \mathbf{X})V^{\mathrm{T}})\} \cdot \mathbf{X}. \end{aligned} \tag{8}$$

Suppose

$$f(\mathbf{X}) \subseteq \operatorname{int}(\mathbf{X}). \tag{9}$$

Then there exists $\widehat{M} \in M_k(\mathbb{K})$ *with* $\widehat{M} \in \tilde{\lambda}I_k + V^{\mathrm{T}}\mathbf{X}$ *such that the Jordan canonical form of* $\widehat{M}$ *is identical to a* $k \times k$ *principal submatrix of the Jordan canonical form of* A, *and there exists* $\widehat{Y} \in M_{n,k}(\mathbb{K})$ *with* $\widehat{Y} \in \tilde{X} + UU^{\mathrm{T}}\mathbf{X}$ *such that* $\widehat{Y}$ *spans the corresponding invariant subspace of* A. *It is* $A\widehat{Y} = \widehat{Y}\widehat{M}$.

Proof. The continuous mapping $f : \mathbb{K}^n \to \mathbb{K}^n$ induced by (8) maps by (9) the nonempty, convex and compact set $\mathbf{X}$ into itself. Therefore, Brouwer's fixed point theorem implies existence of a fixed point $\widehat{X} \in \mathbb{K}^n$ with $f(\widehat{X}) = \widehat{X}$ and $\widehat{X} \in \mathbf{X}$. Inserting into (8) yields

$$-R\{(A\tilde{X} - \tilde{\lambda}\tilde{X}) + (A - \tilde{\lambda}I)UU^{\mathrm{T}}\widehat{X} - (\tilde{X} + UU^{\mathrm{T}}\widehat{X})V^{\mathrm{T}}\widehat{X}\} = 0. \tag{10}$$

Furthermore, (8), (9) and Lemma 3.1 imply $I - R \cdot (\ )$ to be convergent and henceforth R and every matrix $B \in \mathbf{B}$ with $\mathbf{B} := (A - \tilde{\lambda}I)UU^{\mathrm{T}} - (\tilde{X} + UU^{\mathrm{T}} \cdot \mathbf{X})V^{\mathrm{T}} \in \mathbb{I}M_n(\mathbb{K})$ to be non-singular. Collecting terms in (10) yields

$$A(\tilde{X} + UU^{\mathrm{T}}\widehat{X}) = (\tilde{X} + UU^{\mathrm{T}}\widehat{X})(\tilde{\lambda}I_k + V^{\mathrm{T}}\widehat{X}),$$

i.e.

$$A\widehat{Y} = \widehat{Y}\widehat{M} \quad \text{for } \widehat{Y} := \tilde{X} + UU^{\mathrm{T}}\widehat{X} \ \text{ and } \ \widehat{M} := \tilde{\lambda}I_k + V^{\mathrm{T}}\widehat{X}.$$

Finally, $(A - \tilde{\lambda}I)UU^{\mathrm{T}} - (\tilde{X} + UU^{\mathrm{T}}\widehat{X})V^{\mathrm{T}} \in \mathbf{B}$ is non-singular and has k columns equal to $-\widehat{Y}$. Therefore, $\widehat{Y}$ is a basis for an invariant subspace of A. For $\widehat{M} = ZJZ^{-1}$ denoting the Jordan canonical form, $A\widehat{Y} = \widehat{Y}\widehat{M}$ implies $A(\widehat{Y}Z) = (\widehat{Y}Z)J$. The theorem is proved. □

For a practical implementation we have to specify the quantities in use. Note that there are no a priori assumptions on $\tilde{X}$, $\tilde{\lambda}$, R; the only assumption is the inclusion property (9). Especially, there are no a priori requirements on the quality of the approximations $\tilde{X}$ and $\tilde{\lambda}$.

For $A\tilde{X} \approx \tilde{\lambda}\tilde{X}$, the matrix R serves as a preconditioner with (8) giving the obvious choice

$$R \approx ((A - \tilde{\lambda}I)UU^{\mathrm{T}} - \tilde{X}V^{\mathrm{T}})^{-1}.$$

Note that Theorem 3.2 computes an inclusion $\mathbf{X}$ for the *error* with respect to $\tilde{\lambda}$ and $\tilde{X}$. This is computationally most advantageous [17,18] in terms of the quality of the error bounds. A choice for $\mathbf{X}$ is a small superset of the correction term $-R(A\tilde{X} - \tilde{\lambda}\tilde{X})$.

To arrive at a computational procedure we define an interval iteration for $\mathbf{X}$ in order *to construct* an inclusion $\mathbf{X}$. In principle, the interval iteration could be $\mathbf{X}^{\nu+1} := f(\mathbf{X}^{\nu})$. However, we have to make sure that inclusion in the interior of $\mathbf{X}^{\nu}$ as in (9) is possible. That needs a certain width of $\mathbf{X}^{\nu}$, and a good way to take care of that is a so-called epsilon inflation [6,15,17,18]. There is a precise theoretical justification for the epsilon-inflation [18].

Moreover, we need to specify the size k of the cluster. If approximations to all eigenvalues are available, one may guess k from their distances. The choice of k may be critical. For example, the proof of Theorem 3.2 implies that an inclusion is not possible if k is smaller than the actual multiplicity of the eigenvalue to be included.

Finally, we need an inclusion of the eigenvalue cluster, that is an inclusion of the eigenvalues of $\widehat{M}$. Since $\widehat{M} \in \tilde{\lambda}I_k + V^{\mathrm{T}}\mathbf{X}$, this could easily be achieved by the union of Gerschgorin circles of $\tilde{\lambda}I_k + V^{\mathrm{T}}\mathbf{X}$. However, for defective eigenvalues this would yield quite pessimistic bounds.

Instead we proceed as follows. For an interval matrix $\mathbf{C} \in \mathbb{I}M(\mathbb{K})$, denote by $|\mathbf{C}| \in M(\mathbb{R})$ the matrix of the entrywise maximum modulus of $\mathbf{C}$. Therefore, $|C_{ij}| \leqslant (|\mathbf{C}|)_{ij}$ for every $C \in \mathbf{C}$. Then,

$$\text{for } r := \rho(|V^{\mathrm{T}}\mathbf{X}|) \text{ there are } k \text{ eigenvalues of } A \text{ in}$$
$$U_r(\tilde{\lambda}) := \{z \in \mathbf{C}\colon |z - \tilde{\lambda}| \leqslant r\}, \tag{11}$$

where ρ denotes the spectral radius, in this case the Perron root of $|V^{\mathrm{T}}\mathbf{X}| \in M_k(\mathbb{R})$. By principle, the inclusion is complex; we comment on this later. To see (11), observe that $\widehat{M} = \tilde{\lambda} I_k + \tilde{M}$ for some $\widetilde{M} \in V^{\mathrm{T}}\mathbf{X}$, that the eigenvalues of $\widehat{M}$ are the eigenvalues of $\widetilde{M}$ shifted by $\tilde{\lambda}$, and for any eigenvalue of μ of $\widetilde{M}$ it follows by Perron–Frobenius theory $|\mu| \leqslant \rho(\widetilde{M}) \leqslant \rho(|\widetilde{M}|) \leqslant \rho(|V^{\mathrm{T}}\mathbf{X}|) = r$.

The Perron root of a non-negative matrix $C := |V^{\mathrm{T}}\mathbf{X}| \in M(\mathbb{R})$ is estimated for every non-trivial non-negative vector x by

$$\rho(C) \leqslant \varphi(x) \quad \text{with} \quad \varphi(x) = \max_{x_i \neq 0} (Cx)_i / x_i.$$

Whatever numerical method is at hand for calculating an approximation of the Perron vector y of C, $\varphi(y)$ is an upper bound for $\rho(C)$. Usually, a few power iterations do a good job.

The advantage of using (11) occurs especially for defective eigenvalues. The matrix $V^{\mathrm{T}}\mathbf{X}$ basically contains error terms except large off-diagonal quantities characterizing the Jordan blocks. If the error terms are of size ε and off-diagonal elements are of size 1, the spectral radius of $|V^{\mathrm{T}}\mathbf{X}|$ is of size $\varepsilon^{1/m}$, where m is the size of the largest Jordan block. Therefore, the inclusion is of size $\varepsilon^{1/m}$, and this corresponds exactly to the sensitivity of defective eigenvalues [21]. In turn, this implies that if the distance of an m-fold defective eigenvalue to the rest of the spectrum is of the order $\varepsilon^{1/m}$, then “numerically” the cluster comprises of at least $m + 1$ eigenvalues and, for $k = m$, no inclusion is possible. This is confirmed by the numerical examples in Section 6. The above considerations define the following algorithm (where $I = I_n$).

Algorithm 3.3. *Verified error bounds for eigenvalue clusters.*

```
$\mathbf{R} = \mathrm{inv}\,((A - \tilde{\lambda} I)UU^{\mathrm{T}} - \tilde{X}V^{\mathrm{T}})$;
$\mathbf{Z} = -R \cdot (A - \tilde{\lambda} I)\tilde{X}$;
$\mathbf{C} = I - R((A - \tilde{\lambda} I)UU^{\mathrm{T}} - \tilde{X}V^{\mathrm{T}})$;
$\mathbf{X} = \mathbf{Z}$; $\alpha = 0$; $\alpha_{\max} = 10$;
repeat
  $\alpha = \alpha + 1$;
  $\mathbf{Y} = \mathbf{X} + 0.1 \cdot [-|\mathbf{Z}| - \eta, |\mathbf{Z}| + \eta]$;
  $\mathbf{X} = \mathbf{Z} + \mathbf{C}\mathbf{Y} + R(UU^{\mathrm{T}}\mathbf{Y}V^{\mathrm{T}}\mathbf{Y})$;
  ready = ($\mathbf{X} \subseteq \mathrm{int}(\mathbf{Y})$);
*until* ready *or* $(\alpha = \alpha_{\max})$;
*if* ready
   compute $r \geqslant \rho(|V^{\mathrm{T}}\mathbf{X}|)$;
   $\mathbf{L} = \{z \in \mathbf{C} : |z - \tilde{\lambda}| \leqslant r\}$
```

All quantities in bold face are interval quantities, and the computations yielding an interval quantity as a result are supposed to use interval opertions. The quantity η in the epsilon-inflation denotes the smallest representable positive floating point number ($\sim 10^{-308}$ in double precision). This is to ensure that the inflation always increases the width [17,18].

We note that in Algorithm 3.3 we use $\tilde{\lambda} I_k + V^{\mathrm{T}} f(\mathbf{X})$ as the inclusion of $\widetilde{M}$, which is not covered by Theorem 3.2. However, the fixed point property shows that defining $\mathbf{X}^0 := \mathbf{X}$ and $\mathbf{X}^{\nu+1} := f(\mathbf{X}^{\nu})$, it is $\widehat{M} \in \tilde{\lambda} I_k + V^{\mathrm{T}}\mathbf{X}^{\nu}$ for all ν. By purpose, we refrained from noting this in Theorem 3.2 because the reader might use this to improve the inclusion. In general, one better improves the initial approximations by few floating point iterations before starting the verification process. This takes less computing time, and since by principle an interval iterate *must* contain a corresponding floating point iterate, it is also superior.

Note that the final inclusion $\mathbf{L}$ for k eigenvalues of A is always complex. This needs to be the case for the following reason. Suppose A has a k-fold real eigenvalue, k even, and Algorithm 3.3 computes an inclusion $\mathbf{L}$. Then for small perturbations of A the algorithm will still end successfully, but eigenvalues may have moved into the complex plane. In other words, the question whether there exists a *real* multiple eigenvalue (of even multiplicity) is an ill-posed problem.

The same consideration applies to the inclusion of the invariant subspace. For non-defective eigenvalues the $k \times k$ matrix $\tilde{\lambda} I_k + V^{\mathrm{T}}\mathbf{X}$ will *essentially* be diagonal (see Corollary 2.3). But it cannot be proved to be *exactly* diagonal because the problem whether eigenvalues are defective is again ill-posed.

The choice of the normalizing indices (which define U and V, see Section 1) is still free. A simple choice, which is used in our implementation, is to take k rows of $\tilde{X}$ with largest spectral norm. An optimal choice, for example minimizing the width of the inclusion or minimizing $\rho(|\mathbf{C}|)$, is not known.

Finally, we mention that if the entries of A are not exactly representable on the computer, A may be replaced by some $\mathbf{A}$ with $A \in \mathbf{A}$ and obvious changes in Algorithm 3.3. If an inclusion is computed, then using the main principle of inclusion isotonicity (1), assumption (9) in Theorem 3.2 is satisfied for every $\tilde{A} \in \mathbf{A}$, especially for the anticipated matrix A.

Given approximations $\tilde{X}$ and $\tilde{\lambda}$, the computing time of Algorithm 3.3 is governed by the computation of R and $\mathbf{C}$. The interval matrix $\mathbf{C}$ can be computed using $R \cdot (AUU^{\mathrm{T}} - \tilde{X}V^{\mathrm{T}}) + \tilde{\lambda} \cdot RUU^{\mathrm{T}}$, where the two products can be computed by ordinary floating point matrix multiplication in rounding downwards and rounding upwards, respectively. All other operations are $\mathrm{O}(n^2)$ provided $k \ll n$. Thus, the total computing time is $3n^3 + \mathrm{O}(n^2)$. Note that AUU^{T} is just an extraction of $n-k$ columns out of A.

In Section 6, numerical examples will show the robustness of the algorithm, espectially with respect to nearby clusters.

4. Generalized eigenvalue problems

Let $B \in M_n(\mathbb{K})$ be given with $A\tilde{X} \approx \tilde{\lambda}B\tilde{X}$ for an eigenvalue cluster near $\tilde{\lambda}$. Then the following modification of Theorem 3.2 yields verified bounds for the generalized eigenvalue problem.

Theorem 4.1. *Let* $A, B \in M_n(\mathbb{K})$, $\tilde{X} \in M_{n,k}(\mathbb{K})$, $\tilde{\lambda} \in \mathbb{K}$, $R \in M_n(\mathbb{K})$ *and* $\mathbf{X} \in \mathbb{I}M_{n,k}(\mathbb{K})$ *be given, and let* U, V *partition the identity matrix as defined in Section* 1. *Define*

$$\begin{aligned} f(\mathbf{X}) = & -R(A\tilde{X} - \tilde{\lambda}B\tilde{X}) \\ & +\{I - R\cdot((A-\tilde{\lambda}B)UU^{\mathrm{T}} - B(\tilde{X} + UU^{\mathrm{T}}\mathbf{X})V^{\mathrm{T}})\}\cdot\mathbf{X}. \end{aligned}$$

Suppose

$$f(\mathbf{X}) \subseteq \operatorname{int}(\mathbf{X}).$$

Then there exist $\widehat{M} \in M_k(\mathbb{K})$, $\widehat{Y} \in M_{n,k}(\mathbb{K})$ *with* $\widehat{M} \in \tilde{\lambda}I_k + V^{\mathrm{T}}\mathbf{X}$ *and* $\widehat{Y} \in \tilde{X} + UU^{\mathrm{T}}\mathbf{X}$ *such that*

$$A\widehat{Y} = B\widehat{Y}\widehat{M}.$$

The *proof* follows the lines of the proof of Theorem 3.2 and is omitted.

The adaptation of Algorithm 3.3 is straightforward using

$$R \approx ((A-\tilde{\lambda}B)UU^{\mathrm{T}} - B\tilde{X}V^{\mathrm{T}})^{-1}$$

as a preconditioner.

5. Sparse matrices

The previous discussion does not apply to sparse matrices because the preconditioner R, as an approximate inverse, will, in general, be full. Therefore, we seek for an “inverse-free” variant of Theorems 3.2 and 4.1.

Consider (8) and define

$$G := (A - \tilde{\lambda}I)UU^{\mathrm{T}} - \tilde{X}V^{\mathrm{T}}.$$

An approximate inverse of G served as preconditioning matrix R. For the moment, suppose G is invertible and set $R := G^{-1}$. Then

$$\begin{aligned} & I - R((A-\tilde{\lambda}I)UU^{\mathrm{T}} - (\tilde{X} + UU^{\mathrm{T}}\mathbf{X})V^{\mathrm{T}}) \\ & \quad = I - RG + R\cdot UU^{\mathrm{T}}\mathbf{X}V^{\mathrm{T}} = R\cdot UU^{\mathrm{T}}\mathbf{X}V^{\mathrm{T}}, \end{aligned}$$

and inserting into (8) we may replace *f* by

$$f^*(\mathbf{x}) = R(-A\tilde{X} + \tilde{\lambda}\tilde{X} + UU^{\mathrm{T}}\mathbf{X}V^{\mathrm{T}}\mathbf{X}).$$

Hence, condition (9) for f^*, i.e. $f^*(\mathbf{X}) \subseteq \mathrm{int}(\mathbf{X})$, is satisfied if G is invertible and the solution of every linear system $Gy = b,\ b \in -A\tilde{X} + \tilde{\lambda}\tilde{X} + UU^{\mathrm{T}}\mathbf{X}V^{\mathrm{T}}\mathbf{X}$ is included in the interior of **X**. By replacing *f* by f^* in the proof of Theorem 3.2, this proves the following result.

Theorem 5.1. *With the notation of Theorem* 3.2 *and* $G := (A - \tilde{\lambda}I)UU^{\mathrm{T}} - \tilde{X}V^{\mathrm{T}}$ *define*

$$Y := \{y : \exists b \in -A\tilde{X} + \tilde{\lambda}\tilde{X} + UU^{\mathrm{T}}\mathbf{X}V^{\mathrm{T}}\mathbf{X} \text{ and } Gy = b\}.$$

If G is non-singular and

$$Y \subseteq \mathrm{int}(\mathbf{X}), \tag{12}$$

then all assertions of Theorem 3.2 *remain true.*

This reduces the eigenproblem to the solution of a linear system with interval right-hand side, and any algorithm solving this problem for sparse matrices is suitable for computing verified bounds for eigenvalue clusters for sparse matrices. Efficient methods based on a rigorous lower bound for the smallest singular value of the system matrix have been presented in [18]. Those algorithms include the proof of non-singularity of the system matrix.

6. Computational results

For judging computational robustness of Algorithm 3.3 we are interested in the quality of the results with respect to:

1. the dimension of the matrix,
2. the dimension of the invariant subspace,
3. the size of the cluster,
4. the distance to the next eigenvalue or cluster,
5. defective eigenvalues, and
6. many clusters.

Those questions will be addressed in the following. All results are shown for real matrices; the results for complex matrices are completely similar. All computations have been performed using INTLAB [20] in double precision (∼16 decimals), an interval package for use under Matlab V5 [13]. The INTLAB implementation of Algorithm 3.3 is available at our homepage www.ti3.tu-harburg.de/rump/intlab/.

The test matrices are generated by $A = XLX^{-1}$ (multiplication and inversion in floating point), where the anticipated eigenvalues are stored in the diagonal matrix *L*, and *X* is a random matrix with entries uniformly distributed within $[-1, 1]$. All matrices *A* are then scaled to have norm 1. For this floating point matrix, approximate eigenvalues and eigenvectors are computed by the Matlab routine `eig` through `[V, D] = eig(A)`. These approximations are used as input to our verification routine. The accuracy of these approximations corresponds to the sensitivity of the problem. Especially for defective eigenvalues the accuracy may be very poor.

The size k of the generated cluster is an implicit input through the number of columns of the approximation $\tilde{X}$ to an invariant subspace. In a practical computation, k has to be determined by some heuristic. For the following computations, we took $\tilde{\lambda}$ to be the center of the k nearest approximations in D to the anticipated cluster. In our test, a straightforward heuristic specifying some minimum distance to other approximate eigenvalues would for all successfully computed inclusions yield the same result k.

The following is displayed. For the test matrices, 100 samples each, inclusions $\mathbf{L}$ of the eigenvalues and inclusions $(U^{\mathrm{T}}\mathbf{X})_{ij}$ for the corresponding invariant subspaces are computed. Note that we cannot expect better accuracy than $\varepsilon \cdot \|A\| = \varepsilon$, for ε denoting the relative rounding error unit. Therefore, for rad denoting the radius of an interval, we display the average and maximum of rad($\mathbf{L}$) and rad$((U^{\mathrm{T}}\mathbf{X})_{ij})$, respectively, average and maximum taken over all 100 samples and all entries. The last column displays the number of samples (out of 100) where Algorithm 3.3 failed.

Before we address the above questions, we start with the well-known Wilkinson test matrix of dimension 21. Note that the matrix is symmetric and could be treated by other techniques as well. However, it seems to be a must for eigenvalue computations. Eigenvalues λ come in pairs with increasing distance. We treat nine pairs from $\{0.98 \pm 3 \times 10^{-15}\}$ to $\{0.19, 0.16\}$. The remaining three eigenvalues 0.09, 0.02 and -0.10 can hardly be considered as "cluster" (note that the Wilkinson matrix A has been normed to $\|A\|_1 = 1$). Inclusions for all pairs are calculated with results displayed in Table 1.

Although the larger eigenvalues are very close and the smaller eigenvalues are separated up to ±10%, all error bounds were computed successfully (Table 1). In this example we have only one test matrix, therefore the distinction between avge rad $\mathbf{L}$ and max rad $\mathbf{L}$ does not apply.

Table 1
Wilkinson 21 matrix, inclusion of eigenvalue pairs (matrix normed to 1)

λ	rad $\mathbf{L}$	Avge rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Max rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Failed
$0.98 \pm 3.2\mathrm{e}-15$	$3.2\mathrm{e}-15$	$1.7\mathrm{e}-18$	$1.0\mathrm{e}-17$	–
$0.84 \pm 2.6\mathrm{e}-12$	$2.6\mathrm{e}-12$	$2.8\mathrm{e}-18$	$1.0\mathrm{e}-17$	–
$0.73 \pm 3.2\mathrm{e}-10$	$3.2\mathrm{e}-10$	$3.7\mathrm{e}-18$	$1.0\mathrm{e}-17$	–
$0.64 \pm 1.9\mathrm{e}-08$	$1.9\mathrm{e}-08$	$3.3\mathrm{e}-18$	$1.0\mathrm{e}-17$	–
$0.55 \pm 7.5\mathrm{e}-07$	$7.5\mathrm{e}-07$	$3.5\mathrm{e}-18$	$1.0\mathrm{e}-17$	–
$0.45 \pm 2.1\mathrm{e}-05$	$2.1\mathrm{e}-05$	$3.6\mathrm{e}-18$	$1.0\mathrm{e}-17$	–
$0.36 \pm 3.8\mathrm{e}-04$	$3.8\mathrm{e}-04$	$4.0\mathrm{e}-18$	$1.2\mathrm{e}-17$	–
$0.27 \pm 3.7\mathrm{e}-03$	$3.7\mathrm{e}-03$	$5.1\mathrm{e}-18$	$1.5\mathrm{e}-17$	–
$0.18 \pm 1.5\mathrm{e}-02$	$1.5\mathrm{e}-02$	$7.0\mathrm{e}-18$	$2.2\mathrm{e}-17$	–

The 21 eigenvalues can also be treated as simple eigenvalues (which means $k = 1$ in Algorithm 3.3). This is apparently difficult for the largest eigenvalues which are clustered to a relative distance of 3×10^{-15}. The results are shown in Table 2.

For all eigenvalues including the largest ones, rigorous errors bounds are successfully computed. Note that for the largest eigenvalues the invariant subspace becomes ill-conditioned; this can be read off the poorer inclusion.

To the first question, matrices of different dimensions (100 samples each) were generated with k-fold eigenvalue 2 for $k = 10$, and all the other $n - k$ eigenvalues well separated in the unit circle.

Table 3 seems to show a weak dependency on the dimension of the matrices, possibly due to overestimations by interval computations.

Table 2
Wilkinson 21 matrix, inclusion as simple eigenvalues (matrix normed to 1)

λ	rad $\mathbf{L}$	Avge rad$(U^{\mathrm{T}}\mathbf{X})_j$	Max rad$(U^{\mathrm{T}}\mathbf{X})_j$	Failed
0.97692674390031	1.6e − 16	1.7e − 05	1.6e − 04	–
0.97692674390030	1.6e − 16	1.6e − 05	1.6e − 04	–
0.83733442248739	1.6e − 16	3.1e − 08	1.9e − 07	–
0.83733442248227	1.6e − 16	2.8e − 08	1.8e − 07	–
0.73081282934809	1.6e − 16	2.2e − 10	1.2e − 09	–
0.73081282871039	1.6e − 16	2.2e − 10	1.2e − 09	–
0.63672292813897	8.0e − 17	3.9e − 12	2.0e − 11	–
0.63672289078331	8.0e − 17	4.2e − 12	2.1e − 11	–
0.54547582105311	8.0e − 17	9.7e − 14	4.9e − 13	–
0.54547432020519	8.0e − 17	9.1e − 14	4.5e − 13	–
0.45456767500017	8.0e − 17	3.2e − 15	1.6e − 14	–
0.45452567979481	8.0e − 17	3.4e − 15	1.7e − 14	–
0.36403218394917	8.0e − 17	2.2e − 16	1.1e − 15	–
0.36327710921669	4.0e − 17	2.2e − 16	1.0e − 15	–
0.27664539023444	4.0e − 17	2.3e − 17	1.0e − 16	–
0.26918717128961	4.0e − 17	2.3e − 17	1.0e − 16	–
0.19365538357841	4.0e − 17	7.4e − 18	2.5e − 17	–
0.16266557751773	4.0e − 17	8.3e − 18	3.5e − 17	–
0.08613948795721	2.0e − 17	3.9e − 18	1.5e − 17	–
0.02307325609970	1.5e − 17	3.6e − 18	2.0e − 17	–
−0.10231286564727	4.0e − 17	3.6e − 18	2.0e − 17	–

Table 3
Different dimensions $k = 10$

n	Avge rad $\mathbf{L}$	Avge rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Max rad $\mathbf{L}$	Max rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Failed
50	6.3e − 14	7.5e − 15	1.6e − 12	1.5e − 12	–
100	4.2e − 14	7.4e − 15	4.1e − 13	1.1e − 12	–
200	8.9e − 14	2.3e − 14	2.0e − 12	3.1e − 12	–
500	1.7e − 13	7.3e − 14	3.6e − 12	1.1e − 11	–

Table 4 displays sensitivity to different dimensions of the invariant subspace. All matrices are of dimension 100 with a k-fold eigenvalue 2, and $n-k$ eigenvalues in the unit circle. Again, 100 samples each are treated.

Not much dependency of the quality of the results on the dimension of the invariant subspace is visible. There is one exceptional (random) sample for $k=5$.

For Table 5 we choose $n=100$ and $k=5$ with k eigenvalues uniformly and randomly distributed in $[2-e, 2+e]$, and $n-k$ eigenvalues in the unit circle. The uniform approximation to the eigenvalue cluster is $\tilde{\lambda}=2$.

Even a quite large width of the eigenvalue cluster seems to have no influence on the accuracy of the inclusion of the basis of an invariant subspace; in fact, accuracy even improves. The radius of the inclusion of the eigenvalue cluster is smaller than e because the matrix is scaled to norm 1 after generation. With increasing the size e of the cluster it becomes more and more difficult to obtain an inclusion at all. This seems natural because the uniform approximation $\tilde{\lambda}:=2$ becomes more and more poor. However, in the examples where the inclusion failed we may alternatively treat the eigenvalues as *simple* eigenvalues. We did this, and in *all* samples verified error bounds were computed.

Next we investigate the influence of a nearby second cluster. For $n=100$ and $k=5$ we generate a k-fold eigenvalue 2, another k-fold eigenvalue $2+e$, and $n-2k$ eigenvalues in the unit circle.

Table 6 seems to show little influence on the accuracy of the inclusions of the eigenvalue cluster, but with the second cluster moving into the first the inclusion of

Table 4
Different dimensions of invariant subspace, matrix dimension $n=100$

k	Avge rad $\mathbf{L}$	Avge rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Max rad $\mathbf{L}$	Max rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Failed
1	5.4e − 15	9.9e − 16	2.2e − 13	1.0e − 13	–
2	8.3e − 15	1.3e − 14	1.5e − 13	1.9e − 12	–
5	1.5e − 09	8.0e − 15	1.5e − 07	1.0e − 12	–
10	4.8e − 14	2.2e − 14	3.7e − 13	7.6e − 12	–
20	1.7e − 13	1.9e − 14	7.3e − 12	4.0e − 12	–
50	3.3e − 13	7.0e − 15	7.9e − 12	2.2e − 12	–

Table 5
Eigenvalue cluster in $[2-e, 2+e]$, $n=100$, $k=5$

e	Avge rad $\mathbf{L}$	Avge rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Max rad $\mathbf{L}$	Max rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Failed
0	7.7e − 14	1.9e − 14	5.0e − 12	2.5e − 12	–
1e − 13	2.7e − 13	5.0e − 15	1.0e − 12	3.9e − 13	–
1e − 09	3.2e − 08	4.0e − 15	1.0e − 07	3.1e − 13	–
1e − 08	3.3e − 07	3.8e − 15	1.1e − 06	2.4e − 13	3
1e − 07	3.6e − 06	2.1e − 15	2.2e − 05	1.3e − 13	5
1e − 06	4.5e − 05	8.6e − 16	1.4e − 04	2.7e − 14	41
1e − 05	1.4e − 04	5.8e − 16	2.3e − 04	4.1e − 15	92

Table 6
Two k-fold eigenvalues 2 and $2+e$, $n=100$, $k=5$

e	Avge rad $\mathbf{L}$	Avge rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Max rad $\mathbf{L}$	Max rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Failed
1e − 01	5.8e − 14	9.5e − 14	1.5e − 12	8.2e − 12	–
1e − 04	2.0e − 14	7.4e − 11	1.6e − 13	5.9e − 09	1
1e − 06	2.0e − 14	4.3e − 09	1.8e − 13	1.5e − 07	2
1e − 07	1.7e − 14	2.2e − 08	2.7e − 13	4.2e − 07	13
1e − 08	1.4e − 14	1.9e − 07	4.8e − 14	6.9e − 06	33
1e − 09	1.2e − 14	1.3e − 06	4.7e − 14	3.0e − 05	68
1e − 10	9.7e − 15	4.8e − 06	9.7e − 15	1.3e − 05	99

Table 7
Five Jordan blocks of size 2

n	Avge rad $\mathbf{L}$	Avge rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Max rad $\mathbf{L}$	Max rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Failed
50	1.5e − 06	2.0e − 15	8.5e − 05	1.7e − 13	1
100	1.5e − 07	5.3e − 15	1.2e − 06	4.8e − 13	12
200	3.3e − 07	5.7e − 15	8.4e − 06	3.4e − 13	20
500	4.2e − 07	7.1e − 15	1.8e − 06	1.7e − 13	65

the invariant subspace becomes poorer. Again, as we expect, too close clusters cannot be separated by verified bounds. However, for a cluster distance down to about $\varepsilon^{1/2}$ verified bounds are calculated, in agreement to the sensitivity. As before, in those examples where Algorithm 3.3 failed we may alternatively treat the two clusters as *one* cluster of eigenvalues. Doing this, for *all* samples rigorous error bounds were computed.

For examples including defective eigenvalues we proceed as before but generate a direct sum of Jordan blocks. The first example comprises a 10-fold eigenvalue 2 in five Jordan blocks each of size 2, for different dimensions n. Furthermore, the matrix is generated to have one eigenvalue 1 and $n-11$ randomly distributed eigenvalues within $[-1,1]$. Again 100 samples are generated. The results are shown in Table 7. Note that the theoretical sensitivity of the eigenvalues is approximately $\varepsilon^{1/2} = 1.5 \times 10^{-8}$.

In all examples up to now the matrix $A = XLX^{-1}$ was generated by floating point matrix inversion and product. In general, this alters a multiple eigenvalue of L into a cluster of small size. This seems not too important for non-defective eigenvalues. For defective eigenvalues we may want to be sure that the matrix has indeed a multiple *and* defective eigenvalue. We assure this by calculating an inclusion $\mathbf{Y} \in \mathbb{I}M_n(\mathbb{K})$ of X^{-1}, i.e. $X^{-1} \in \mathbf{Y}$. This can be done for example using the algorithms described in [17]. Then, we calculate $\mathbf{A} = X \cdot L \cdot \mathbf{Y}$ using interval multiplications. By the inclusion isotonicity (1), this implies $XLX^{-1} \in \mathbf{A}$. For this interval matrix $\mathbf{A} \in \mathbb{I}M_n(\mathbb{K})$, we apply the mentioned interval version of Algorithm 3.3. Input approximations $\tilde{\lambda}$, $\tilde{X}$ are, as always, calculated by the Matlab routine eig, in this case applied to the midpoint matrix of $\mathbf{A}$. Those approximations are very poor, in fact, the approximation

$\tilde{X}$ of a basis for a k-dimensional invariant subspace is numerically of rank 1. For the same data as before, five Jordan blocks each of size 2 for the eigenvalue 2, the results are shown in Table 8.

The results are surprisingly similar to Table 7, although the input matrix **A** now comprises of (narrow) intervals. This is because the sensitivity of the double eigenvalues is of the order 10^{-8}, and a perturbation of the order 10^{-16} in the matrix elements plays a less important role. Note that the computed inclusions **L** and **X** are valid for *all* $A \in \mathbf{A}$, especially for the matrix with five pairs of defective eigenvalues.

Finally, we increase the size of the Jordan blocks and generate one k-fold eigenvalue 2 of geometric multiplicity 1, and otherwise one eigenvalue 1 and $n-k-1$ randomly chosen eigenvalues in $[-1, 1]$. The matrix $A = XLX^{-1}$ is computed in floating point, the dimension is always $n = 100$.

The increased number of failure reflects the ill-conditionedness of the problem, see Table 9. Note the theoretical sensitivity $\varepsilon^{1/5} \sim 7.4 \times 10^{-4}$ for $k = 5$. The same problem for interval input matrix $\mathbf{A} = X \cdot L \cdot \mathbf{Y}$, where $\mathbf{Y}$ is an inclusion of X^{-1}, produces the results displayed in Table 10. The dimension is again $n = 100$.

As before, there is not too much difference to pure floating point input because the sensitivity of the defective eigenvalue predominates by far the effect of the uncertainty of the matrix elements.

There seems to be room for inprovement in case of defective eigenvalues. One problem is that the approximation of the k-dimensional basis for the invariant delivered by `eig` is extremely ill-conditioned for large k.

Next, we generate many clusters to observe robustness in that practical situation. For given N, we generate a matrix having a k-fold eigenvalue in each of the N

Table 8
Five Jordan blocks of size 2, interval matrix

n	Avge rad $\mathbf{L}$	Avge rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Max rad $\mathbf{L}$	Max rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Failed
50	1.2e − 06	2.9e − 15	5.1e − 05	1.9e − 13	6
100	2.3e − 07	3.9e − 15	2.3e − 06	1.6e − 13	10
200	3.6e − 07	5.8e − 15	2.9e − 06	1.7e − 13	32
500	4.2e − 07	8.2e − 15	1.3e − 06	8.4e − 13	56

Table 9
One Jordan block of different size (approximate matrix)

k	Avge rad $\mathbf{L}$	Avge rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Max rad $\mathbf{L}$	Max rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Failed
2	2.3e − 08	1.8e − 15	3.7e − 07	1.7e − 13	1
3	8.9e − 06	6.5e − 16	2.6e − 04	1.9e − 14	13
4	8.9e − 05	3.4e − 16	3.1e − 04	3.3e − 15	54
5	6.7e − 04	3.7e − 16	7.5e − 04	1.1e − 15	98

Table 10
One Jordan block of different size (interval matrix)

k	Avge rad $\mathbf{L}$	Avge rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Max rad $\mathbf{L}$	Max rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Failed
2	3.4e − 08	2.9e − 15	2.1e − 07	3.1e − 13	1
3	1.2e − 05	9.8e − 16	1.0e − 04	5.1e − 14	24
4	1.6e − 04	5.3e − 16	5.7e − 04	4.8e − 15	71
5	–	–	–	–	100

Table 11
Clusters in N Chebyshev nodes, $k = 5$

n	N	rad $\mathbf{L}$	Avge rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Max rad$(U^{\mathrm{T}}\mathbf{X})_{ij}$	Failed
50	10	7.4e − 15	3.9e − 15	1.2e − 14	–
100	20	1.5e − 14	3.6e − 14	9.0e − 14	–
150	30	5.3e − 14	1.1e − 12	5.8e − 12	–
200	40	2.1e − 14	3.9e − 13	1.1e − 12	–
250	50	4.8e − 14	2.6e − 12	9.3e − 12	–
300	60	2.9e − 14	8.5e − 13	3.3e − 12	–
350	70	4.3e − 14	2.5e − 12	7.9e − 12	–
400	80	1.1e − 13	3.6e − 11	1.8e − 10	–
450	90	3.1e − 14	1.2e − 11	4.1e − 11	–
500	100	3.0e − 14	2.8e − 12	1.1e − 11	–

Chebyshev nodes. The dimension of the matrix is then $n = N \cdot k$. Table 11 displays the results for $k = 5$ and different number of nodes. Inclusion is calculated for the eigenvalue near 1.

As Table 11 shows, the algorithm has no trouble in separating the busiest eigenvalues near 1, and also the others are included without problems. The eigenvalues are computed to high accuracy, only the accuracy of the bounds for the invariant subspaces decrease with narrowing clusters. For $N = 50$ the two largest eigenvalues of the resulting matrix of dimension 500 are 0.9999 and 0.9989.

Finally we compare computing times for dimensions $n = 100$ and $n = 200$ with $k = 10$, a 10-fold eigenvalue. We compare (i) Algorithm 3.3 for error bounds of the cluster with (ii) the Matlab built-in routine `eig` (to approximate all eigenvalues).

	$n = 100$	$n = 200$
Time for Algorithm 2.1	0.35 s	2.36 s
Time for eig	0.46 s	4.12 s

Time is measured on a 300 MHz PC. The computing time does not depend on k for $k \ll n$, for k approaching n it increases by about a factor 2.

Acknowledgements

The author wishes to thank M. Nakao, N. Yamamoto and T. Yamamoto for fruitful discussions and most valuable comments.

References

[1] ACRITH High-Accuracy Arithmetic Subroutine Library, Program Description and User's Guide, IBM Publications, No. SC 33-6164-3, 1986.

[2] G. Alefeld, private communication.

[3] G. Alefeld, J. Herzberger, Introduction to Interval Computations, Academic Press, New York, 1983.

[4] G. Alefeld, H. Spreuer, Iterative improvement of componentwise errorbounds for invariant subspaces belonging to a double or nearly double eigenvalue, Computing 36 (1986) 321–334.

[5] ARITHMOS, Benutzerhandbuch, Siemens AG, Bibl.-Nr. U 2900-I-Z87-1 edition, 1986.

[6] O. Caprani, K. Madsen, Iterative methods for interval inclusion of fixed points, BIT 18 (1978) 42–51.

[7] J.J. Dongarra, C.B. Moler, J.H. Wilkinson, Improving the accuracy of computed eigenvalues and eigenvectors, SIAM J. Numer. Anal. 20 (1) (1983) 23–45.

[8] D. Husung, Precompiler for scientific computation (TPX), Technical Report 91.1, Institut für Informatik III, TU Hamburg-Harburg, 1989.

[9] R. Klatte, U. Kulisch, M. Neaga, D. Ratz, Ch. Ullrich, PASCAL-XSC – Sprachbeschreibung mit Beispielen, Springer, New York, 1991.

[10] R. Klatte, U. Kulisch, A. Wiethoff, C. Lawo, M. Rauch, C-XSC A C++ Class Library for Extended Scientific Computing, Springer, Berlin, 1993.

[11] O. Knüppel, PROFIL/BIAS – a fast interval library, Computing 53 (1994) 277–287.

[12] R. Krawczyk, Fehlerabschätzung reeller Eigenwerte und Eigenvektoren von Matrizen, Computing 4 (1969) 281–293.

[13] MATLAB User's Guide, Version 5. The Math Works, 1997.

[14] G. Mayer, Enclosures for eigenvalues and eigenvectors, in: L. Atanassova, J. Herzberger (Eds.), Computer Arithmetic and Enclosure Methods, IMACS, Elsevier, Amsterdam, 1992, pp. 49–67.

[15] G. Mayer, Success in epsilon-inflation, in: G. Alefeld, et al. (Eds.), Scientific Computing and Validated Numerics, Proceedings of the International Symposium on Scientific Computing, Computer Arithmetic and Validated Numerics SCAN-95, Wuppertal, Germany, 26-29 September 1995; Math. Res., vol. 90, Akademie Verlag, Berlin, 1996, pp. 98–104.

[16] A. Neumaier, Interval Methods for Systems of Equations, Encyclopedia of Mathematics and its Applications, Cambridge University Press, New York, 1990.

[17] S.M. Rump, Kleine Fehlerschranken bei Matrixproblemen, Ph.D. Thesis, Universität Karlsruhe, 1980.

[18] S.M. Rump, Verification methods for dense and sparse systems of equations, in: J. Herzberger (Ed.), Topics in Validated Computations – Studies in Computational Mathematics, Elsevier, Amsterdam, 1994, pp. 63–136.

[19] S.M. Rump, Fast and parallel interval arithmetic, BIT 39 (3) (1999) 539–560.

[20] S.M. Rump, INTLAB – INTerval LABoratory, in: T. Csendes (Ed.), Developments in Reliable Computing, Kluwer Academic Publishers, Dordrecht, 1999, pp. 77–104.

[21] J.H. Wilkinson, The Algebraic Eigenvalue Problem, Clarendon Press, Oxford, 1965.

[22] J.H. Wilkinson, Error bounds for computed invariant subspaces, Technical Report 81–02, Proceedings of Rutishauser Symposium on Numerical Analysis, Eidgenössische Technische Hochschule, Zürich, Switzerland, February 1981.

[23] T. Yamamoto, Error bounds for computed eigenvalues and eigenvectors, Numer. Math. 34 (1980) 189–199.