

EXTREME EIGENVALUES OF REAL SYMMETRIC TOEPLITZ MATRICES

A. MELMAN

ABSTRACT. We exploit the even and odd spectrum of real symmetric Toeplitz matrices for the computation of their extreme eigenvalues, which are obtained as the solutions of spectral, or secular, equations. We also present a concise convergence analysis for a method to solve these spectral equations, along with an efficient stopping rule, an error analysis, and extensive numerical results.

1. INTRODUCTION

Spectral equations, or *secular equations*, as they are frequently called (see [13]), are a convenient tool for computing a few eigenvalues of a matrix. These equations are different from the characteristic polynomial (even though equivalent) and contain rational functions, which therefore exhibit singularities. Sometimes these singularities are known explicitly, but this is not always the case, and it is not for the equations that we consider here.

In the present work, we will use such equations for computing the smallest and largest eigenvalues of real symmetric Toeplitz matrices. These matrices appear in quite a large number of applications, chief among them digital signal processing (for an overview of applications see [4]). The computation of the smallest eigenvalue of such matrices was considered in, e.g., [9], [17], [21] and [26], whereas bounds were studied in, e.g., [11], [15] and [24].

All the concepts we need will be defined and explained in the following section, so for the moment let us just say that Toeplitz matrices have two kinds of eigenvalues: even and odd. As our basic tool to compute extreme eigenvalues, we use a set of two equations, one for each type of eigenvalue. This is contrary to [9] and [21], where no such distinction is made and only one secular equation is used. The idea behind these equations is certainly not new, and to put matters in perspective, we note that they appear in [10] in an equivalent form that is less suitable for computation, and are hinted at in [11] without being explicitly stated. No applications of these equations were considered in either paper. In [17], one of these equations is derived in a different way, without taking into account the spectral structure of the submatrices of the matrix, thereby obscuring key properties of the equation. It is used there to compute the smallest even eigenvalue.

More recently, in [23] and in [26], these same equations were derived in a form suitable for computation. In [26], they were then used to compute the smallest

Received by the editor September 22, 1998 and, in revised form, May 24, 1999.

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

Key words and phrases. Toeplitz matrix, extreme eigenvalues, odd and even spectra, spectral equation, secular equation, rational approximation.

eigenvalue of a real symmetric positive definite Toeplitz matrix. The crucial aspect of the aforementioned algorithms is the rootfinder used for solving their respective secular equations. In [9], the rootfinder is simply Newton's method, whereas in [21] and [26], it is a variant on the rational approximation method in [5], which is much more efficient than Newton's method.

In this work, we propose first of all a different and much shorter convergence analysis of the rootfinder in [21] and [26]. A similar technique then leads to a more efficient stopping rule for the algorithm, based on a rational approximation, rather than the polynomial one used in [21] and [26]. We also propose an error analysis, lacking in both [21] and [26], but which is necessary to understand the obtainable accuracy of the computed solution. Finally, we present extensive numerical results, beyond the single class of matrices investigated in [21] and [26], including a class of matrices for which the algorithm works less well than might otherwise be concluded from those papers. In addition to the smallest, we also compute the largest eigenvalue. We will concentrate on positive definite matrices, which occur in most practical applications, even though the approach in this work is valid for any real symmetric Toeplitz matrix.

The paper is organized as follows. In section 2, definitions and basic results on Toeplitz matrices are given. In section 3, we summarize the derivation of the spectral equations and previous results. Section 4 deals with the numerical solution of the spectral equations, while an error analysis and stopping rules are discussed in section 5. In section 6, we present numerical results.

2. PRELIMINARIES

A symmetric matrix $T \in \mathbb{R}^{(n,n)}$ is said to be *Toeplitz* if its elements T_{ij} satisfy $T_{ij} = t_{|j-i|}$ for some vector $\mathbf{t} = (t_0, \dots, t_{n-1})^T \in \mathbb{R}^n$. Many early results about such matrices can be found in, e.g., [3], [6] and [10].

Toeplitz matrices are *persymmetric*, i.e., they are symmetric about their southwest-northeast diagonal. For such a matrix T , this is the same as requiring that $JT^T J = T$, where J is a matrix with ones on its southwest-northeast diagonal and zeros everywhere else (the exchange matrix). It is easy to see that the inverse of a persymmetric matrix is also persymmetric. A matrix that is both symmetric and persymmetric is called *doubly symmetric*.

A *symmetric vector* v is defined as a vector satisfying $Jv = v$, and a *skew-symmetric* or *antisymmetric vector* w as one that satisfies $Jw = -w$. If these vectors are eigenvectors, then their associated eigenvalues are called *even* and *odd*, respectively. Drawing on results in [3], it was shown in [6] that, given a real symmetric Toeplitz matrix T of order n , there exists an orthonormal basis for \mathbb{R}^n , composed of $n - \lfloor n/2 \rfloor$ symmetric and $\lfloor n/2 \rfloor$ skew-symmetric eigenvectors of T , where $\lfloor \alpha \rfloor$ denotes the integral part of α . In the case of simple eigenvalues, this is easy to see from the fact that, if $Tu = \lambda u$, then $T(Ju) = \lambda(Ju)$, because $JTJ = T$ and $J^2 = I$. Therefore u and Ju must be proportional, and therefore u must be an eigenvector of J , which means that either $Ju = u$ or $Ju = -u$. Finally, we note that for $\alpha \in \mathbb{R}$, the matrix $(T - \alpha I)$ is symmetric and Toeplitz, whenever T is.

The identity and exchange matrices are denoted by I and J , respectively, throughout this paper, without specifically indicating their dimensions, which are assumed to be clear from the context.

3. SPECTRAL EQUATIONS

In this section we summarize known results about the various spectral, or “secular”, equations for the eigenvalues of a real symmetric Toeplitz matrix. We will assume from now on that T is positive definite, even though our approach is valid for any real symmetric Toeplitz matrix, as will be briefly explained at the end of this section.

Let us consider the following partition of a symmetric Toeplitz matrix T , which we shall refer to in the future as “partition I”:

$$T = \begin{pmatrix} t_0 & \mathbf{t}^T \\ \mathbf{t} & Q \end{pmatrix},$$

where $t = (t_1, \dots, t_{n-1})^T$ and Q is an $(n-1) \times (n-1)$ symmetric Toeplitz matrix generated by the vector $(t_0, \dots, t_{n-2})^T$. Then the following well-known theorem (see, e.g., [9]) holds:

Theorem 3.1. *The eigenvalues of T that are not shared with those eigenvalues of Q , whose associated eigenspaces are not entirely contained in $\{t\}^\perp$, are given by the solutions of the equation:*

$$(1) \quad -t_0 + \lambda + t^T(Q - \lambda I)^{-1}t = 0.$$

The eigenvectors belonging to the eigenvalues of T that are not in the spectrum of Q are then given by $(1, -(Q - \lambda I)^{-1}t)^T$. \square

We define the function $f(\lambda)$ by

$$(2) \quad f(\lambda) \triangleq -t_0 + \lambda + t^T(Q - \lambda I)^{-1}t.$$

Equation (1) is equivalent to

$$(3) \quad -t_0 + \lambda + \sum_{i=1}^p \frac{c_i^2}{\omega_i - \lambda} = 0,$$

where $\{\omega_i\}_{i=1}^p$ are the p eigenvalues of Q for which the associated eigenspace U_{ω_i} is not entirely contained in the subspace $\{t\}^\perp$, i.e., for which $U_{\omega_i} \not\subset \{t\}^\perp$ (see, e.g., [11] for more details). Denote the orthonormal vectors which form a basis for U_{ω_i} by $\{u_j^{(i)}\}_{j=1}^{d_i}$, where d_i is the dimension of U_{ω_i} . Then the scalars c_i^2 are given by $c_i^2 = \sum_{j=1}^{d_i} (t^T u_j^{(i)})^2$. The rational function in (1), or (3), has p simple poles, dividing the real axis into $p+1$ intervals, on each of which it is monotonely increasing from $-\infty$ to $+\infty$. The $p+1$ solutions $\{\lambda_j\}_{j=1}^{p+1}$ of equation (3) therefore satisfy

$$\lambda_1 < \omega_1 < \lambda_2 < \omega_2 < \dots < \omega_p < \lambda_{p+1},$$

i.e., the eigenvalues ω_i strictly interlace the eigenvalues λ_j ; this is known as Cauchy’s interlacing theorem. These results are well-known, and we refer to, e.g., [9], [11] and [25].

A positive-definite matrix T will therefore certainly have an eigenvalue in the interval $(0, \omega_1)$. In [9], the smallest eigenvalue is then computed from $f(\lambda) = 0$.

It would appear that the previous partition of T is inappropriate, given the persymmetry of Toeplitz matrices. We therefore consider the following, more natural, partition for a matrix (see [11], [23] and [26]), which is both symmetric and

persymmetric and which in the future we shall refer to as “partition II”:

$$T = \begin{pmatrix} t_0 & \tilde{t}^T & t_{n-1} \\ \tilde{t} & G & J\tilde{t} \\ t_{n-1} & (J\tilde{t})^T & t_0 \end{pmatrix},$$

where $\tilde{t} = (t_1, \dots, t_{n-2})^T$ and G is an $(n-2) \times (n-2)$ symmetric Toeplitz matrix generated by the vector $(t_0, \dots, t_{n-3})^T$. In what follows, we denote the even and odd eigenvalues of T by λ_i^e and λ_i^o , and the even and odd eigenvalues of G by μ_i and ν_i , respectively. We then have the following theorem from [23] and [26], which yields two equations: one for even and one for odd eigenvalues of T .

Theorem 3.2. *The even eigenvalues of T that are not shared with those even eigenvalues of G , whose associated eigenspaces are not entirely contained in $\{\tilde{t}\}^\perp$, are the solutions of*

$$(5) \quad -t_0 - t_{n-1} + \lambda + \tilde{t}^T(G - \lambda I)^{-1}(\tilde{t} + J\tilde{t}) = 0,$$

whereas the odd eigenvalues of T that are not shared with those odd eigenvalues of G , whose associated eigenspaces are not entirely contained in $\{\tilde{t}\}^\perp$, are the solutions of

$$(5) \quad -t_0 + t_{n-1} + \lambda + \tilde{t}^T(G - \lambda I)^{-1}(\tilde{t} - J\tilde{t}) = 0.$$

The symmetric and skew-symmetric eigenvectors of T belonging to eigenvalues that are not in the spectrum of G are then given by $(1, -(G - \lambda I)^{-1}(\tilde{t} + J\tilde{t}), 1)^T$ and $(1, -(G - \lambda I)^{-1}(\tilde{t} - J\tilde{t}), -1)^T$, respectively.

Proof. Using partition II of T , the problem $Tx = \lambda x$, or $(T - \lambda I)x = 0$, is equivalent to the following system of equations:

$$(6) \quad (t_0 - \lambda)x_1 + \tilde{t}^T \tilde{x} + t_{n-1}x_n = 0,$$

$$(7) \quad x_1 \tilde{t} + (G - \lambda I)\tilde{x} + (J\tilde{t})x_n = 0,$$

$$(8) \quad t_{n-1}x_1 + (J\tilde{t})^T \tilde{x} + (t_0 - \lambda)x_n = 0,$$

where $\tilde{x} = (x_2, \dots, x_{n-1})^T$ and, as before, $\tilde{t} = (t_1, \dots, t_{n-2})^T$. If λ is not an eigenvalue of G , we have from (7) that

$$(9) \quad \tilde{x} = -(G - \lambda I)^{-1}(x_1 \tilde{t} + x_n(J\tilde{t})).$$

First of all, this means that x_1 and x_n cannot be simultaneously zero without \tilde{x} being zero as well. Substituting (9) into (6) and (8) yields the following system of linear equations in x_1 and x_n :

$$(10) \quad (t_0 - \lambda - \tilde{t}^T(G - \lambda I)^{-1}\tilde{t})x_1 + (t_{n-1} - \tilde{t}^T(G - \lambda I)^{-1}(J\tilde{t}))x_n = 0,$$

$$(11) \quad (t_{n-1} - (J\tilde{t})^T(G - \lambda I)^{-1}\tilde{t})x_1 + (t_0 - \lambda - (J\tilde{t})^T(G - \lambda I)^{-1}(J\tilde{t}))x_n = 0.$$

For this system in x_1 and x_n to have a nontrivial solution, its determinant has to be zero, which means that

$$(12) \quad (t_0 - \lambda - \tilde{t}^T(G - \lambda I)^{-1}\tilde{t})^2 - (t_{n-1} - \tilde{t}^T(G - \lambda I)^{-1}(J\tilde{t}))^2 = 0,$$

where we have used the fact that $J^2 = I$ and the persymmetry of the inverse of a Toeplitz matrix. Thus, the values of λ satisfying (12) are the solutions of the following two equations:

$$(13) \quad -t_0 - t_{n-1} + \lambda + \tilde{t}^T(G - \lambda I)^{-1}(\tilde{t} + J\tilde{t}) = 0,$$

$$(14) \quad -t_0 + t_{n-1} + \lambda + \tilde{t}^T(G - \lambda I)^{-1}(\tilde{t} - J\tilde{t}) = 0.$$

This also means that either $x_1 = x_n$ or $x_1 = -x_n$, corresponding to symmetric or skew-symmetric eigenvectors, respectively, and this could also have been used to derive our equations directly from (10) and (11). Even eigenvalues are therefore obtained by solving equation (4), whereas the solutions of equation (13) yield odd eigenvalues. Since λ was not in the spectrum of G , equation (12) provides the expression for the eigenvectors.

Let us now denote an orthonormal basis of \mathbb{R}^{n-2} , composed of orthonormal eigenvalues of G , by $\{v_i\}_{i=1}^r \cup \{w_j\}_{j=1}^s$, where $r = n - 2 - \lfloor (n - 2)/2 \rfloor$, $s = \lfloor (n - 2)/2 \rfloor$, (v_i, μ_i) are symmetric eigenvector – even eigenvalue pairs and (w_j, ν_j) skew-symmetric eigenvector – odd eigenvalue pairs. This is always possible because of the results in [6], as was explained in the preliminaries. We can therefore write $\tilde{t} = \sum_{i=1}^r a_i v_i + \sum_{j=1}^s b_j w_j$. However, some eigenvalues might have an associated eigenspace that is entirely contained in $\{\tilde{t}\}^\perp$, so that we actually have $\tilde{t} = \sum_{i=1}^{r'} a_i v_i + \sum_{j=1}^{s'} b_j w_j$ with $r' \leq r$ and $s' \leq s$, and the summation does not contain the eigenvectors associated with these particular eigenvalues. This means that

$$J\tilde{t} = \sum_{i=1}^{r'} a_i (Jv_i) + \sum_{j=1}^{s'} b_j (Jw_j) = \sum_{i=1}^{r'} a_i v_i - \sum_{j=1}^{s'} b_j w_j,$$

and therefore

$$\begin{aligned} \tilde{t}^T (G - \lambda I)^{-1} (\tilde{t} + J\tilde{t}) &= 2 \left(\sum_{i=1}^{r'} a_i v_i + \sum_{j=1}^{s'} b_j w_j \right)^T (G - \lambda I)^{-1} \left(\sum_{i=1}^{r'} a_i v_i \right) \\ &= 2 \left(\sum_{i=1}^{r'} a_i v_i + \sum_{j=1}^{s'} b_j w_j \right)^T \left(\sum_{i=1}^{r'} \frac{a_i}{\mu_i - \lambda} v_i \right) \\ &= 2 \sum_{i=1}^{r'} \frac{a_i^2}{\mu_i - \lambda}. \end{aligned}$$

Analogously, we obtain

$$\tilde{t}^T (G - \lambda I)^{-1} (\tilde{t} - J\tilde{t}) = 2 \sum_{j=1}^{s'} \frac{b_j^2}{\nu_j - \lambda}.$$

Equations (4) and (5) now become

$$(15) \quad -t_0 - t_{n-1} + \lambda + 2 \sum_{i=1}^{r'} \frac{a_i^2}{\mu_i - \lambda} = 0$$

and

$$(16) \quad -t_0 + t_{n-1} + \lambda + 2 \sum_{j=1}^{s'} \frac{b_j^2}{\nu_j - \lambda} = 0,$$

which means that no eigenvalues whose associated eigenspaces are entirely contained in $\{\tilde{t}\}^\perp$ appear as poles, so that some of the solutions of (15) or (16) might coincide with these eigenvalues. This concludes the proof. \square

Let us assume once again that all eigenvalues of G are simple (the general case does not differ substantially). We then have $r = r'$ and $s = s'$ in the proof of Theorem 3.2, which shows that the rational functions in each of equations (4) and (5) are of the same form as the function in (1). It is also clear that T will certainly have an even eigenvalue on $(0, \mu_1)$ and an odd one on $(0, \nu_1)$. These equations were also hinted at in [11], without however being proved or explicitly stated. The meaning of Theorem 3.2 is therefore that those even and odd eigenvalues of G whose associated eigenspaces are not completely contained in $\{\tilde{t}\}^\perp$ interlace, respectively, the even and odd eigenvalues of T that are not shared with those eigenvalues of G . This result was obtained in [10] in a different way, along with equivalent forms of equations (15) and (16) that are less suitable for applications.

Finally, because of the orthogonality of the eigenvectors, equations (4) and (5) can be written in a more symmetric way, as shown in the following two equations, which at the same time define the functions $f_e(\lambda)$ and $f_o(\lambda)$:

$$(17) \quad f_e(\lambda) \triangleq -t_0 - t_{n-1} + \lambda + \frac{1}{2}(\tilde{t} + J\tilde{t})^T(G - \lambda I)^{-1}(\tilde{t} + J\tilde{t}) = 0,$$

$$(18) \quad f_o(\lambda) \triangleq -t_0 + t_{n-1} + \lambda + \frac{1}{2}(\tilde{t} - J\tilde{t})^T(G - \lambda I)^{-1}(\tilde{t} - J\tilde{t}) = 0.$$

We note that equation (17) was also obtained in [17], where it was used to compute the smallest eigenvalue, which was known in advance to be even. However, the derivation of the equation is quite different, concentrating exclusively on the smallest eigenvalue and disregarding the spectral structure of the submatrices of T , which obscures important properties of that equation.

Let us now consider the largest eigenvalue of a real symmetric Toeplitz matrix, which we obtain by computing the smallest eigenvalue of a different real symmetric Toeplitz matrix, given an upper bound δ on the largest eigenvalue. This can be accomplished by translating the origin in the spectral equations to δ , replacing the resulting new variable by its opposite and multiplying the equation by (-1) , thus obtaining the exact same type of spectral equation for the matrix $(\delta I - T)$, which is always positive definite. Computing the smallest eigenvalue of this new matrix then yields the largest eigenvalue, since $\lambda_{\min}(\delta I - T) = \delta - \lambda_{\max}(T)$. One possible value for δ is the Frobenius norm of T , defined as (see [14])

$$\|T\|_F = \left(\sum_{i,j=1}^n T_{ij}^2 \right)^{1/2},$$

which for a Toeplitz matrix can be computed in $\mathcal{O}(n)$ flops.

The same procedure for computing the smallest eigenvalue of real symmetric positive-definite matrices can be used for general real symmetric matrices as well, provided that a lower bound on the smallest eigenvalue is available. Any known lower bound can be used (see, e.g., [11] or [15]), or one could be obtained by a process where a trial value is iteratively lowered until it falls below the smallest eigenvalue.

For simplicity, we shall assume in what follows that T and Q or T and G do not have their extreme eigenvalues in common, so that we really are computing the extreme eigenvalues of T . When this is not the case, then our equations only lead to the computation of the smallest and largest of those eigenvalues that were defined in Theorem 3.1 and Theorem 3.2.

4. SOLUTION OF THE SPECTRAL EQUATIONS

The functions f , f_e and f_o are increasing convex functions on $(-\infty, \omega_1)$, $(-\infty, \mu_1)$ and $(-\infty, \nu_1)$, respectively. To solve these equations, we will use the method from [9] and [21] for $f(\lambda) = 0$ and adapt it to solve the equations $f_e(\lambda) = 0$ and $f_o(\lambda) = 0$ for the smallest root of T . Once the eigenvalues are computed, it is a simple matter to obtain the eigenvectors, as shown by Theorem 3.1 and Theorem 3.2.

We shall first summarize the method and present its convergence analysis for the equation $f(\lambda) = 0$. Since all our spectral functions are of the same form, it will then be easy to do the same for $f_e = 0$ and $f_o = 0$.

We recall that

$$f(\lambda) = -t_0 + \lambda + t^T(Q - \lambda I)^{-1}t = -t_0 + \lambda + \sum_{j=1}^{n-1} \frac{c_j^2}{\omega_j - \lambda},$$

where $0 < \omega_1 < \omega_2 < \dots < \omega_{n-1}$. We then solve $f(\lambda) = 0$ for its smallest root λ^* , which satisfies $\lambda^* < \omega_1$. The algorithm requires a starting point $\lambda_1 \in (\lambda^*, \omega_1)$. However, the ω_j 's are unknown, so that the algorithm really consists of two phases: phase I, which determines the starting point λ_1 , and phase II, which is the part of the algorithm that will be proved to converge monotonely to λ^* from the right.

To evaluate f , the Yule-Walker (YW) system $(Q - \lambda I)^{-1}w = -t$ needs to be solved (the sign of t is by convention), which can be accomplished recursively by *Durbin's method* (see [11] for the original reference, but a good overview of this and other Toeplitz-related algorithms can be found in [14]). The flop count for this method is $2n^2 + \mathcal{O}(n)$, where we define one flop as in [14], namely a multiplication/division or an addition/subtraction. Once f has been computed, f' is obtained by an inner product in $\mathcal{O}(n)$ flops: $f'(\lambda) = 1 + \|w\|^2$. In addition, Durbin's method provides a factorization (proved in [8]) of $(Q - \lambda I)$ of the form $U^T(Q - \lambda I)U = D$, where D is an $n \times n$ diagonal and U an upper triangular $n \times n$ matrix whose ℓ th column is given by $(Jw^{(\ell-1)}, 1)^T$, where $w^{(\ell)}$ is the solution of the ℓ th dimensional subsystem obtained in the course of solving the aforementioned YW system. The elements of D are obtained as by-products of the solution of the YW system. Sylvester's law of inertia then supplies a means to determine the position of λ relative to ω_1 , e.g., if D contains only positive elements, then $\lambda < \omega_1$. The same can be done for $(T - \lambda I)$ to determine λ 's position relative to λ^* . This is also used in [9], [21], [25] and [26].

Durbin's method is not the only method available for solving the YW equations. There are the so-called superfast solvers (see [1] and [2]), and then there are the standard methods that do not take into account the Toeplitz structure. This is a secondary issue, as we are mainly interested in the solution of the YW equations, regardless of how it was obtained.

In [9], the spectral equation is solved by Newton's method, which is based on a linear approximation. This is however not very appropriate for a rational function, and it would be better to use a rational approximation. This idea was in the present context of eigenvalue computations already used in, e.g., [5] and many others. It was applied to the equations at hand in [21] and [26], and we will do the same here.

The following useful form of $f(\lambda)$ was suggested in [21]:

$$(19) \quad f(\lambda) = f(\mu) + f'(\mu)(\lambda - \mu) + (\lambda - \mu)^2 g(\lambda),$$

where $-\infty < \mu < \omega_1$. The advantage of this form is that, no matter how g is approximated, the resulting approximation to f will always coincide with f at $\lambda = \mu$ up to first derivatives. The function $g(\lambda)$ is of the same general form as the rational part of $f(\lambda)$ and has the same singularities. This can be seen from

$$\begin{aligned} g(\lambda) &= \frac{f(\lambda) - f(\mu)}{(\lambda - \mu)^2} - \frac{f'(\mu)}{\lambda - \mu} \\ &= \frac{(\lambda - \mu) + \sum_{j=1}^{n-1} c_j^2 \left(\frac{1}{\omega_j - \lambda} - \frac{1}{\mu - \lambda} \right)}{(\lambda - \mu)^2} - \frac{f'(\mu)}{\lambda - \mu} \\ &= \frac{1}{\lambda - \mu} \left(\sum_{j=1}^{n-1} \frac{c_j^2}{\omega_j - \mu} \left(\frac{1}{\omega_j - \lambda} - \frac{1}{\mu - \lambda} \right) \right) \\ &= \sum_{j=1}^{n-1} \frac{\frac{c_j^2}{(\omega_j - \mu)^2}}{\omega_j - \lambda} = \sum_{j=1}^{n-1} \frac{\alpha_j^2}{\omega_j - \lambda}, \end{aligned}$$

where $\alpha_j^2 = c_j^2/(\omega_j - \mu)^2$. We note that, not surprisingly, $\lim_{\lambda \rightarrow \mu} g(\lambda) = f''(\mu)/2$.

An iterative method for the solution of $f(\lambda)$ can then be constructed as follows. Phase I starts off by interpolating $f(\lambda)$ at $\lambda = 0$ by $-t_0 + \lambda + a/(b - \lambda)$. As will be shown in Theorem 4.1, the root of this interpolant is an upper bound on λ^* . Bisection and Sylvester's law of inertia are then combined to find a point $\bar{\lambda} < \omega_1$. If $\lambda^* < \bar{\lambda} < \omega_1$, phase I terminates. If not, the procedure is repeated, this time by interpolating $g(\lambda)$ in (19) with $\mu = 0$ at $\lambda = \bar{\lambda}$ by $a/(b - \lambda)$. The root of the approximation to f thus obtained is once again an upper bound on λ^* , and once again bisection and Sylvester's law of inertia are used to find the next $\bar{\lambda}$. This process is repeated with μ in (19) successively equal to the second to last $\bar{\lambda}$, until phase I terminates, at which point we have two points: λ_0 , equal to the last μ obtained, and λ_1 , equal to the last $\bar{\lambda}$, so that $-\infty < \lambda_0 < \lambda^* < \lambda_1 < \omega_1$. In phase II, we interpolate $g(\lambda)$ in (19) with $\mu = \lambda_0$ at the point $\lambda = \lambda_1$ by $a/(b - \lambda)$. The next iterate is then that root λ_2 of the function $f(\lambda_0) + f'(\lambda_0)(\lambda - \lambda_0) + (\lambda - \lambda_0)^2(a/(b - \lambda))$ which satisfies $\lambda^* < \lambda_2 < \lambda_1 < \omega_1$. The iterate λ_2 now plays the role of λ_1 ; λ_0 remains fixed, and the process continues until a suitable stopping criterion is satisfied.

Other initial points, based on higher-order rational approximations, are possible, but we have concluded from extensive numerical experiments that although this leads to fewer iterations in phase I, it causes an increase in the total amount of work because of the computational cost of the higher derivatives, and we consequently do not report these results in section 6.

We define the following abbreviations before stating a conceptual algorithm:

INTPSLV1($f, \alpha, \bar{\lambda}$) indicates the calculation of the root $\bar{\lambda}$ of the interpolant $-t_0 + \lambda + a/(b - \lambda)$ for $f(\lambda)$ at the point $\lambda = \alpha$;

INTPSLV2($f, \alpha, \beta, \bar{\lambda}$) indicates the calculation of the root $\bar{\lambda}$ of the interpolant $f(\alpha) + f'(\alpha)(\lambda - \alpha) + (\lambda - \alpha)^2(a/(b - \lambda))$ for $f(\lambda)$ at the point $\lambda = \beta$;

BSTCT(α, β, λ) stands for bisection of the interval (α, β) , keeping α as its left limit, until the interval (α, λ) is reached;

STP indicates the satisfaction of some stopping rule;

\neg STP is the negation of STP.

Algorithm 4.1.

```

INTPSLV1( $f, 0, \lambda_1$ )
BSCT( $0, \lambda_1, \lambda$ ) until  $0 < \lambda < \omega_1$ 
 $\lambda_2 \leftarrow \lambda_1$ 
 $\gamma \leftarrow 0$ 
WHILE  $0 < \lambda_2 < \lambda^*$  and  $\neg$ STP
    INTPSLV2( $f, \gamma, \lambda_2, \sigma$ )
    BSCT( $\gamma, \sigma, \lambda$ ) until  $\gamma < \lambda < \omega_1$ 
     $\gamma \leftarrow \lambda_2$ 
END
 $\lambda_3 \leftarrow \lambda_2$ 
WHILE  $\neg$ STP
    INTPSLV2( $f, \gamma, \lambda_3, \lambda$ )
     $\lambda_3 \leftarrow \lambda$ 
END

```

The following theorem proves the basic properties of the rational approximation, which were already alluded to.

Theorem 4.1. *Let $g(\lambda)$ be defined by (19) with $-\infty < \mu < \omega_1$. Then the interpolant $a/(b - \lambda)$ of $g(\lambda)$ at $\lambda = \lambda_1$, determined by $a/(b - \lambda_1) = g(\lambda_1)$ and $a/(b - \lambda_1)^2 = g'(\lambda_1)$, satisfies the following on the interval $(-\infty, \omega_1)$:*

- (1) $a > 0$ and $\omega_1 < b < \omega_{n-1}$.
- (2) $g(\lambda) \geq \frac{a}{b - \lambda}$.
- (3) $f(\lambda) \geq f(\mu) + f'(\mu)(\lambda - \mu) + (\lambda - \mu)^2 \frac{a}{b - \lambda} \geq f(\lambda_1) + f'(\lambda_1)(\lambda - \lambda_1)$.

Proof. In what follows, we will frequently abbreviate $g(\lambda)$ by g , $g(\lambda_1)$ by g_1 and $g'(\lambda_1)$ by g'_1 . From $a/(b - \lambda_1) = g_1$ and $a/(b - \lambda_1)^2 = g'_1$, we readily obtain $b - \lambda_1 = g_1/g'_1$ and $a = g_1^2/g'_1$. Since $g_1 > 0$, this means that $a > 0$. For b , we have

$$b - \lambda_1 = \frac{\sum_{j=1}^{n-1} \alpha_j^2 / (\omega_j - \lambda_1)}{\sum_{j=1}^{n-1} \alpha_j^2 / (\omega_j - \lambda_1)^2} = \sum_{j=1}^{n-1} \left(\frac{\alpha_j^2 / (\omega_j - \lambda_1)^2}{\sum_{j=1}^{n-1} \alpha_j^2 / (\omega_j - \lambda_1)^2} \right) (\omega_j - \lambda_1).$$

This means that $(b - \lambda_1)$ is a strictly convex combination of the $(\omega_j - \lambda_1)$'s and therefore that $\omega_1 - \lambda_1 < b - \lambda_1 < \omega_{n-1} - \lambda_1$, so that $\omega_1 < b < \omega_{n-1}$, which proves the first part of the theorem.

The proof of the second part is based on the observation that if $a/(b - \lambda)$ interpolates $g(\lambda)$ up to first order, then $(b - \lambda)/a$ interpolates $1/g(\lambda)$ up to first order. Since $(1/g)'' = g^{-3}(2g'^2 - gg'')$, we know from Lemma 2.3 in [22] with $\rho = -1$ that $1/g(\lambda)$ is concave, so that $(b - \lambda)/a \geq 1/g(\lambda)$, implying that $a/(b - \lambda) \leq g(\lambda)$, which concludes the second part of the proof.

Finally, some simple algebra yields

$$\left(f(\mu) + f'(\mu)(\lambda - \mu) + (\lambda - \mu)^2 \frac{a}{b - \lambda} \right)'' = \frac{2a(b - \mu)^2}{(b - \lambda)^3}.$$

Since $a > 0$ and $b > \lambda$ on $(-\infty, \omega_1)$, this approximation to f is, like f itself, convex on $(-\infty, \omega_1)$. This means that it will dominate its own linear interpolant at $\lambda = \lambda_1$

on that same interval. It also coincides with f at $\lambda = \lambda_1$ up to first order, so that on $(-\infty, \omega_1)$:

$$f(\mu) + f'(\mu)(\lambda - \mu) + (\lambda - \mu)^2 \frac{a}{b - \lambda} \geq f(\lambda_1) + f'(\lambda_1)(\lambda - \lambda_1) .$$

This latest expression, combined with the second part, concludes the proof. \square

Since f is increasing and convex on $(-\infty, \omega_1)$, Newton's method for $f(\lambda) = 0$ converges to λ^* from any point in (λ^*, ω_1) . Its order of convergence is at least quadratic, because $f'(\lambda^*) \neq 0$. Theorem 4.1 then shows that Algorithm 4.1 converges to λ^* monotonely from the right at least as fast as Newton's method.

A slightly less general version of Theorem 4.1 was proved in [21], but with a different and much longer proof. Each time the root of an approximation is computed, the algorithm executes $2n^2 + \mathcal{O}(n)$ flops. During a bisection step this can be less, as the step terminates as soon as a negative diagonal element of D (see the previous description of Durbin's algorithm) is found.

We now turn to the equations $f_e(\lambda) = 0$ and $f_o(\lambda) = 0$. In this case, one basically has the same two phases as before, with a few differences to account for the fact that there are now two equations and because it is not known in advance whether the smallest eigenvalue is even or odd. During phase I, bisection is used, with upper bound the smaller of the two upper bounds (one for each equation). In phase II, one computes iterates for both equations, each time picking the smaller of the two to continue, until either f_e or f_o becomes negative. From that point on, the algorithm continues with the equation that has a positive function value.

The convergence analysis is entirely analogous to Algorithm 4.1, as all functions involved have the same rational form and therefore the same properties. The flop count is basically the same as for the previous method, since for any Toeplitz matrix T we have $T(Jy) = -Jt$ if $Ty = -t$. This means that f_e and f_o can be evaluated together at basically the same cost as an evaluation of f . The number of extra calculations that are carried out is negligible compared to the evaluation of a single function value, even for relatively small matrices.

We have the following conceptual algorithm, where the abbreviations have the same meaning as in Algorithm 4.1, with slight and obvious changes to account for the different parameters in the equations.

Algorithm 4.2.

```

INTPSLV1( $f_e, 0, \lambda_1$ )
INTPSLV1( $f_o, 0, \lambda_1$ )
 $\lambda_1 \leftarrow \min\{\lambda_1^e, \lambda_1^o\}$ 
BSCT( $0, \lambda_1, \lambda$ ) until  $0 < \lambda < \min\{\mu_1, \nu_1\}$ 
 $\lambda_2 \leftarrow \lambda_1$ 
 $\gamma \leftarrow 0$ 
WHILE  $0 < \lambda_2 < \lambda^*$  and  $\neg$ STP
    INTPSLV2( $f_e, \gamma, \lambda_2, \sigma_e$ )
    INTPSLV2( $f_o, \gamma, \lambda_2, \sigma_o$ )
     $\sigma = \min\{\sigma_e, \sigma_o\}$ 
    BSCT( $\gamma, \sigma, \lambda$ ) until  $\gamma < \lambda < \min\{\mu_1, \nu_1\}$ 
     $\gamma \leftarrow \lambda_2$ 
     $\lambda_2 \leftarrow \lambda$ 
END
```

```

 $\lambda_3 \leftarrow \lambda_2$ 
WHILE  $\neg$ STP
  WHILE  $f_e > 0$  and  $f_o > 0$  and  $\neg$ STP
    INTPSLV2( $f_e, \gamma, \lambda_3, \lambda^e$ )
    INTPSLV2( $f_o, \gamma, \lambda_3, \lambda^o$ )
     $\lambda_3 \leftarrow \min\{\lambda^e, \lambda^o\}$ 
  END
  IF  $f_e > 0$ 
    WHILE  $\neg$ STP
      INTPSLV2( $f_e, \gamma, \lambda_3, \lambda^e$ )
       $\lambda_3 \leftarrow \lambda^e$ 
    END
  ELSE
    WHILE  $\neg$ STP
      INTPSLV2( $f_o, \gamma, \lambda_3, \lambda^o$ )
       $\lambda_3 \leftarrow \lambda^o$ 
    END
  END
END
END

```

5. ERROR ANALYSIS AND STOPPING RULES

We start by considering the error in the solution of the YW systems by Durbin's method, as this error determines the accuracy with which the corresponding spectral function can be evaluated. Let us begin with the YW system arising in partition I of T . All computed quantities will be denoted by a "hat".

We define the residual η as follows: $\eta = (Q - \lambda I)\hat{w} + t$, where \hat{w} is the computed solution of $(Q - \lambda I)w = -t$, with w denoting the exact solution. We have

$$(Q - \lambda I)w = -t = (Q - \lambda I)\hat{w} - \eta,$$

and therefore $\hat{w} = w + (Q - \lambda I)^{-1}\eta$. We note that for $\lambda < \lambda_{\min}(Q)$ we have $\|(Q - \lambda I)^{-1}\| = (\lambda_{\min}(Q) - \lambda)^{-1}$, which becomes larger as $(\lambda_{\min}(Q) - \lambda)$ becomes smaller. This gap therefore has a clear influence on the accuracy with which \hat{w} can be computed, as was also observed in [9].

For the spectral function, corresponding to partition I, this yields

$$\hat{f}(\lambda) = -t_0 + \lambda + t^T w + t^T (Q - \lambda I)^{-1} \eta,$$

so that $\hat{f}(\lambda) = f(\lambda) - w^T \eta$. This means that

$$f(\lambda) - \|w\| \|\eta\| \leq \hat{f}(\lambda) \leq f(\lambda) + \|w\| \|\eta\|.$$

Even for $\lambda = \lambda^*$, the computed value of f would only satisfy $\hat{f}(\lambda^*) = w^T \eta$. Since $w = (Q - \lambda I)^{-1}t$, $|w^T \eta|$ could potentially be rather large, depending once again on how small $(\lambda_{\min}(Q) - \lambda)$ is. This is entirely corroborated by the numerical experiments. We found the error bound for Durbin's method from [14], due to G. Cybenko, to be too pessimistic to be useful. We also found that the accuracy of Durbin's method is inferior (sometimes by several orders of magnitude) to, e.g., Cholesky factorization, which is, however, (inevitably) more costly.

A good measure for the accuracy of the computed eigenvalue/eigenvector pair $(\hat{\lambda}, \hat{v})$ is the quantity $\|T\hat{v} - \hat{\lambda}\hat{v}\|/\|\hat{v}\|$, which for partition I becomes

$$(20) \quad \frac{\left\| \begin{pmatrix} t_0 & t^T \\ t & Q \end{pmatrix} \begin{pmatrix} 1 \\ \hat{w} \end{pmatrix} - \hat{\lambda} \begin{pmatrix} 1 \\ \hat{w} \end{pmatrix} \right\|}{(1 + \|\hat{w}\|^2)^{1/2}} = \frac{\left\| \begin{pmatrix} t_0 - \hat{\lambda} + t^T \hat{w} \\ t + (Q - \hat{\lambda}I)\hat{w} \end{pmatrix} \right\|}{(1 + \|\hat{w}\|^2)^{1/2}} \\ = \left(\frac{\hat{f}^2(\hat{\lambda}) + \|\eta\|^2}{1 + \|\hat{w}\|^2} \right)^{1/2}.$$

If $\eta \ll |\hat{f}(\hat{\lambda})|$, then $\|T\hat{v} - \hat{\lambda}\hat{v}\|/\|\hat{v}\| \approx |\hat{f}(\hat{\lambda})|/\|\hat{v}\|$.

From Theorem 5 on p. 141 in [18], we know that there is an eigenvalue λ_T of T such that

$$|\hat{\lambda} - \lambda_T| \leq \frac{\|T\hat{v} - \hat{\lambda}\hat{v}\|}{\|\hat{v}\|}.$$

For a small enough value of the right-hand side expression, we can ensure that $\lambda_T = \lambda^*$.

To get a rough idea of how accurately we can compute the eigenvalues with our algorithms, let us make the rather reasonable assumption that $\|(Q - \lambda I)^{-1}\eta\| \ll \|w\|$, so that $\|w\| \approx \|\hat{w}\|$. Let us also assume that $|\hat{f}(\hat{\lambda})| \approx \|w\| \|\eta\|$, which is about the best we can hope to obtain for $|\hat{f}(\hat{\lambda})|$. Then, by the above, about the best we can hope to obtain for $|\hat{\lambda} - \lambda^*|$ is given by

$$(21) \quad |\hat{\lambda} - \lambda^*| \leq \|\eta\|.$$

Both our assumptions and the error bound are, in general, quite nicely verified by the numerical experiments.

A similar analysis can be carried out for the even and odd spectral equations. Let us consider the even case. We define $\zeta = (G - \lambda I)\hat{y} + \tilde{t}$, where \hat{y} is the computed solution of $(G - \lambda I)y = -\tilde{t}$, with y denoting the exact solution. We have

$$\hat{f}_e(\lambda) = -t_0 - t_{n-1} + \lambda + \tilde{t}^T(y + Jy) + \tilde{t}^T(G - \lambda I)^{-1}(\zeta + J\zeta),$$

so that $\hat{f}_e(\lambda) = f_e(\lambda) + y^T(\zeta + J\zeta)$, which means that

$$f_e(\lambda) - \|y\| \|\zeta + J\zeta\| \leq \hat{f}_e(\lambda) \leq f_e(\lambda) + \|y\| \|\zeta + J\zeta\|.$$

For the odd spectral function, we obtain analogously that

$$f_o(\lambda) - \|y\| \|\zeta - J\zeta\| \leq \hat{f}_o(\lambda) \leq f_o(\lambda) + \|y\| \|\zeta - J\zeta\|.$$

The even eigenvalue/eigenvector pair $(\hat{\lambda}_e, (1, \hat{y} + J\hat{y}, 1)^T)$ satisfies

$$\frac{\left\| \begin{pmatrix} t_0 & \tilde{t}^T & t_{n-1} \\ \tilde{t} & G & J\tilde{t} \\ t_{n-1} & (J\tilde{t})^T & t_0 \end{pmatrix} \begin{pmatrix} 1 \\ \hat{y} + J\hat{y} \\ 1 \end{pmatrix} - \hat{\lambda}_e \begin{pmatrix} 1 \\ \hat{y} + J\hat{y} \\ 1 \end{pmatrix} \right\|}{(2 + \|\hat{y} + J\hat{y}\|^2)^{1/2}} \\ = \frac{\left\| \begin{pmatrix} t_0 + t_{n-1} - \hat{\lambda}_e + \tilde{t}^T(\hat{y} + J\hat{y}) \\ (\tilde{t} + J\tilde{t}) + (G - \hat{\lambda}_e I)(\hat{y} + J\hat{y}) \\ t_0 + t_{n-1} - \hat{\lambda}_e + (J\tilde{t})^T(\hat{y} + J\hat{y}) \end{pmatrix} \right\|}{(2 + \|\hat{y} + J\hat{y}\|^2)^{1/2}} \leq \left(\frac{\hat{f}_e^2(\hat{\lambda}_e) + 2\|\zeta\|^2}{2 + \|\hat{y} + J\hat{y}\|^2} \right)^{1/2}.$$

We obtain an analogous bound for the odd eigenvalue/eigenvector pair by replacing $\hat{f}_e(\hat{\lambda}_e)$ by $\hat{f}_o(\hat{\lambda}_o)$ in the last expression.

Because $(\lambda_{\min}(G) - \lambda) > (\lambda_{\min}(Q) - \lambda)$ for any $\lambda < \lambda_{\min}(Q)$, we can expect better accuracy for the calculation of f_e and f_o than of f , and this is indeed borne out by the numerical experiments.

We now turn to the question of how to terminate the iterations in the solution of the spectral equation. The stopping rule used in [9] and [25] is based on the relative change in the iterates, whereas in [21] and [26] the stopping rule is based on a lower bound on the solution, which leads to slightly better results when compared to the one used in [9] and [25]. We propose a rule, based on a different lower bound. Before we continue, let us briefly review the lower bound in [21]. Throughout the rest of this section we will consider the solution of $f(\lambda) = 0$, as f_e and f_o are of the same form as f and have similar properties.

Denoting the last iterate on the left-hand side of λ^* by λ_0 and the current iterate (which lies on the right-hand side of λ^*) by λ_1 , the bound in [21] is obtained as the root of a quadratic polynomial $p(\lambda)$, on the interval (λ_0, λ_1) . This polynomial is determined by the interpolation conditions $p(\lambda_0) = f(\lambda_0)$, $p'(\lambda_0) = f'(\lambda_0)$ and $p(\lambda_1) = f(\lambda_1)$. It is shown in [21] that the root on (λ_0, λ_1) of $p(\lambda)$ is a lower bound on λ^* . This approach suffers from two shortcomings: first, no use is made of $f'(\lambda_1)$, even though it is computed anyway, and, second, the interpolant is a polynomial, a poor approximation for a rational function.

We propose a different and more accurate bound, which is based on a rational, rather than a polynomial, interpolant, and which does exploit the availability of $f'(\lambda_1)$. It needs, in addition to λ_1 , another point on (λ^*, ω_1) , which we denote by λ_2 . Such a point, together with $g(\lambda_2)$ and $g'(\lambda_2)$, is available from the previous iteration. Let us first write $f(\lambda)$ in the form

$$(22) \quad f(\lambda) = f(\lambda_0) + f'(\lambda_0)(\lambda - \lambda_0) + (\lambda - \lambda_0)^2 g(\lambda),$$

with $\lambda_0 \in (0, \lambda^*)$. To obtain our bound, we interpolate $g(\lambda)$ by a rational function of the form $\phi(\lambda) \triangleq a + b/(c - \lambda)$. As we will show in the following theorem, the function $\phi(\lambda)$ satisfies $\phi(\lambda) \geq g(\lambda)$ on $(-\infty, \lambda_1)$, so that replacing $g(\lambda)$ by $\phi(\lambda)$ in the spectral function leads to a lower bound on λ^* .

Theorem 5.1. *For $-\infty < \lambda_1 < \lambda_2 < \omega_1$, the function*

$$\phi(\lambda) \triangleq a + b/(c - \lambda)$$

such that

$$\begin{aligned} \phi(\lambda_1) &= g(\lambda_1), & \phi'(\lambda_1) &= g'(\lambda_1), \\ \phi''(\lambda_1) &= -2g^{3/2}(\lambda_1) \frac{g'^{-1/2}(\lambda_1) - g'^{-1/2}(\lambda_2)}{\lambda_1 - \lambda_2}, \end{aligned}$$

satisfies $\phi(\lambda) \leq g(\lambda)$ for all λ in $(-\infty, \lambda_1)$.

Proof. We first consider the function $\psi(\lambda) \triangleq \bar{a} + \bar{b}/(\bar{c} - \lambda)$, determined by the conditions

$$(23) \quad \psi(\lambda_1) = g(\lambda_1), \quad \psi'(\lambda_1) = g'(\lambda_1), \quad \psi''(\lambda_1) = g''(\lambda_1).$$

In [23], it was shown that $\bar{c} > \omega_1$ for such an approximation, so this means that $\psi'^{-1/2}(\lambda) \equiv (\bar{c} - \lambda)/\sqrt{\bar{b}}$ approximates $g'^{-1/2}(\lambda)$ linearly. Computing $(g'^{-1/2})''$, we

obtain

$$(24) \quad (g'^{-1/2})'' = \frac{1}{2}g'^{-\frac{5}{2}} \left(\frac{3}{2}g''^2 - g'g''' \right).$$

The right-hand side in (24) is negative by Lemma 2.3 in [22] with $\rho = -2$, which means that $g'^{-1/2}$ is concave on $(-\infty, \omega_1)$, so that $(\bar{c} - \lambda)/\sqrt{\bar{b}} \geq g'^{-1/2}(\lambda)$ on that same interval. Since \bar{b} , $(\bar{c} - \lambda)$ and $g'(\lambda)$ are positive, this is equivalent to $\bar{b}/(\bar{c} - \lambda)^2 \leq g'(\lambda)$. We then have, for $\lambda \geq \lambda_1$,

$$\int_{\lambda_1}^{\lambda} \frac{\bar{b} d\sigma}{(\bar{c} - \lambda)^2} \geq \int_{\lambda_1}^{\lambda} g'(\sigma) d\sigma,$$

and therefore

$$\frac{\bar{b}}{\bar{c} - \lambda} - \frac{\bar{b}}{\bar{c} - \lambda_1} \geq g(\lambda) - g(\lambda_1).$$

Adding and subtracting \bar{a} in the left-hand side of this inequality and using $\psi(\lambda_1) = g(\lambda_1)$ yields $\psi(\lambda) \geq g(\lambda)$ on $(-\infty, \omega_1)$. Since this derivation is based on the concavity of $g'^{-1/2}$, all inequalities still hold on $(-\infty, \lambda_1)$ if we replace the linear approximation at $\lambda = \lambda_1$ by the one obtained by approximating $(g'^{-1/2})'(\lambda_1)$ with $(g'^{-1/2}(\lambda_1) - g'^{-1/2}(\lambda_2))/(\lambda_1 - \lambda_2)$ for some point λ_2 such that $\lambda_2 \in (\lambda_1, \omega_1)$. Because $(g'^{-1/2})' = (-1/2)g'^{-3/2}g''$, this is equivalent to replacing $g''(\lambda_1)$ in the interpolation conditions (23) by

$$-2g'^{3/2}(\lambda_1) \frac{g'^{-1/2}(\lambda_1) - g'^{-1/2}(\lambda_2)}{\lambda_1 - \lambda_2},$$

which leads to the coefficients a , b and c that define ϕ . We can therefore conclude that $\phi(\lambda) \geq g(\lambda)$ on $(-\infty, \lambda_1)$. \square

It is easily verified that the function

$$f(\lambda) = f(\lambda_0) + f'(\lambda_0)(\lambda - \lambda_0) + (\lambda - \lambda_0)^2 \phi(\lambda)$$

is convex and increasing on $[\lambda_0, \lambda_1]$, so that its root on that interval is a lower bound on λ^* . Consequently, this procedure can be used as a stopping rule, once at least two iterates on the right-hand side of λ^* are available. The root can easily be computed by solving a cubic equation, or with Newton's method and λ_1 as a starting point, with a negligible number of flops compared to one evaluation of f .

We now formally define the two stopping rules.

(1) Stopping Rule 1. The stopping rule from [21]: stop when

$$\frac{|\lambda - B1|}{(1/2)(\lambda + B1)} \leq \epsilon_1,$$

where ϵ_1 is a tolerance, λ is an iterate obtained in phase II of the algorithm, and $B1$ is the corresponding bound, obtained in [21]: $\lambda_0 < B1 \leq \lambda^* < \lambda_1$.

(2) Stopping Rule 2. The stopping rule, based on Theorem 5.1: stop when

$$\frac{|\lambda - B2|}{(1/2)(\lambda + B2)} \leq \epsilon_2,$$

where ϵ_2 is a tolerance, λ is an iterate obtained in phase II of the algorithm, and $B2$ is the corresponding bound, obtained by the rational interpolant in Theorem 5.1: $\lambda_0 < B2 \leq \lambda^* < \lambda_1$. It is used as soon as at least two iterates on the right-hand

side of λ^* are available. Until then (in fact just for the first iteration in phase II), stopping rule 1 is used.

6. NUMERICAL RESULTS

Let us first formally define the methods that will be compared in the numerical experiments.

(1) **CVLM**. “The Cybenko-Van Loan method”. This is Algorithm 4.1. When equipped with stopping rule 1 or 2 from the previous section, we denote it CVLM_1 or CVLM_2 , respectively.

(2) **SPM**. “The Spectral Parity method”. This is Algorithm 4.2. It comes, like the CVLM, equipped with stopping rule 1 or 2, and is denoted accordingly SPM_1 or SPM_2 . When computing the largest eigenvalue, both methods use the Frobenius norm to obtain an upper bound as explained at the end of section 3.

We have tested these methods on three classes of positive semi-definite matrices. For each class and for each of the dimensions $n = 100, 200, 400, 800$, we have run the following experiments: 200 experiments comparing the two stopping rules for the smallest eigenvalue, 200 experiments computing the smallest and largest eigenvalues for the three methods CVLM_1 (the method in [21]), SPM_1 (Algorithm 4.2 with the stopping rule from [21]) and SPM_2 (Algorithm 4.2 with the new stopping rule), and 200 experiments comparing the amount of work done per phase for the methods CVLM_1 and SPM_2 . The results appear in Tables 1 through 9. In Table 1 through Table 6 one finds, in the first column underneath each dimension, the average values (with their standard deviations in parentheses) of the number of flops required to compute a single eigenvalue in terms of the number of flops required for one full step in the CVLM_1 method, which is $2n^2 + 3n$ flops. This expression takes into account the number of flops required by Durbin’s method and the inner products that must be computed. The second column underneath each dimension represents the average value of $-\log_{10}(\|Tv - \lambda v\|/\|v\|)$ with its standard deviation, i.e., the accuracy of the eigenvalue/eigenvector pair. The accuracy is given for illustrative purposes only, as these values are not directly controlled by the stopping rules.

In Table 7 through Table 9, the entries represent average percentages (with their standard deviations in parentheses) for the CVLM_1 and SPM_2 methods, as follows: the last number in each entry is the amount of work executed in phase II as a percentage of the total work in the computation of one eigenvalue. To provide a clearer picture of what happens in phase I, we have divided this phase into two parts: the part where an attempt is made at finding an upper bound on the eigenvalue which is smaller than the second eigenvalue, and the part where bisection is performed when this attempted bound lies to the right of the second eigenvalue. The percentage of work devoted to these parts is given in the first and second numbers in each entry, respectively. The sum of these two numbers then represents the percentage of work carried out in phase I. We note that, due to roundoff, the percentages may sometimes not add up to 100.

We stress that we have reported the amount of work needed to compute both the eigenvalue and the eigenvector. In general, this entails one full step more than would be required if we computed eigenvalues only. This is so because the stopping rule allows us to determine that the next iterate does indeed satisfy the stopping rule without having to compute another function value. However, if we also want to compute the associated eigenvector, then we need to solve an additional YW system.

Let us now list the three classes of matrices.

(1) **CVL matrices.** These are matrices defined in [9] (whence their name) as

$$T = \mu \sum_{k=1}^n \xi_k T_{2\pi\theta_k},$$

where n is the dimension of T , μ is such that $T_{kk} = 1$, $k = 1, \dots, n$, and

$$(T_\theta)_{ij} = \cos(\theta(i-j)).$$

These matrices are positive semi-definite of rank two. We generated random matrices of this kind by taking the value of θ to be uniformly distributed on $(0, 1)$.

(2) **KMS matrices.** These are the Kac-Murdock-Szegő matrices (see [19]), defined as

$$T_{ij} = \nu^{|i-j|},$$

where $0 < \nu < 1$ and $i, j = 1, \dots, n$, where n is the dimension of the matrix. These matrices are positive definite and are characterized by the fact that their even and odd eigenvalues lie extremely close together. Random matrices of this kind were generated by taking the value of ν to be uniformly distributed on $(0, 1)$.

(3) **UNF matrices.** We define UNF matrices by first defining a random vector v of length n whose components are uniformly distributed on $(-10, 10)$. We then modify that vector by adding to its first component 1.1 times the absolute value of the smallest eigenvalue of the Toeplitz matrix generated by v . Finally, the vector v is normalized by dividing it by its first component, provided that it is different from zero. The Toeplitz matrix generated by this normalized vector is then called a UNF matrix. From their construction, these matrices are positive semi-definite. Theoretically, some of the matrices generated in the experiments might be singular, although we never encountered this situation in practice. We also never encountered a case where the smallest or largest eigenvalue was a double eigenvalue for both T and Q or T and G .

The tolerance for the relative error on the eigenvalue (for both ϵ_1 and ϵ_2 in the definition of the stopping rules) was set equal to 10^{-10} , which cannot always be obtained because of (21), which was found to be very realistic. As a safeguarding measure, the algorithm was stopped if roundoff caused the function value to become negative in phase II.

It should be mentioned that the deviations from the mean are almost always in the same direction for all methods, i.e., in the case where a particular method needs to carry out a higher (or lower) number of operations than the average, then this is almost always true for all other methods as well. The experiments then clearly show (as was also reported in [26]) that exploiting the even and odd spectra yields a better method, both in the number of iterations and in the standard deviation, especially for the CVL and UNF matrices, where the improvement in the average amount of work for the computation of the smallest eigenvalue is at least 25% across dimensions. The smaller standard deviation indicates a more stable root-finding process, as could be expected by the larger distance from the root to the nearest singularity of the spectral function. The magnitude of the improvement diminishes with decreasing distance between the even and odd eigenvalues, as is obviously true for the KMS matrices, where the improvement is not significant.

TABLE 1. Comparison of the methods with $\epsilon_1 = \epsilon_2 = 10^{-10}$ for the smallest eigenvalue of CVL matrices.

Method	Dimension			
	100	200	400	800
CVLM ₁	8.50(2.57) / 14.08(0.72)	9.03(2.36) / 13.58(0.69)	10.15(2.90) / 12.95(0.73)	11.01(2.85) / 12.54(0.88)
CVLM ₂	7.91(2.47) / 13.54(0.87)	8.50(2.26) / 13.31(0.75)	9.63(2.80) / 12.75(0.75)	10.39(2.72) / 12.40(0.80)
SPM ₁	6.21(1.36) / 14.14(0.69)	7.00(1.32) / 13.71(0.67)	7.96(1.74) / 13.11(0.74)	8.63(1.90) / 12.59(0.99)
SPM ₂	5.83(1.26) / 13.90(0.84)	6.64(1.14) / 13.53(0.73)	7.61(1.55) / 13.01(0.76)	8.14(1.56) / 12.54(0.96)

TABLE 2. Comparison of the methods with $\epsilon_1 = \epsilon_2 = 10^{-10}$ for the smallest eigenvalue of KMS matrices.

Method	Dimension			
	100	200	400	800
CVLM ₁	13.52(2.56) / 13.70(0.61)	15.10(2.74) / 13.30(0.61)	16.70(2.86) / 12.92(0.63)	17.77(2.79) / 12.44(0.64)
CVLM ₂	13.42(2.63) / 13.62(0.76)	15.07(2.76) / 13.25(0.67)	16.70(2.87) / 12.92(0.63)	17.77(2.79) / 12.44(0.64)
SPM ₁	12.60(2.50) / 13.87(0.55)	14.29(2.81) / 13.49(0.59)	15.74(2.94) / 13.06(0.63)	17.12(2.75) / 12.59(0.59)
SPM ₂	12.38(2.58) / 13.71(0.81)	14.23(2.87) / 13.47(0.60)	15.74(2.95) / 13.06(0.63)	17.12(2.75) / 12.59(0.59)

TABLE 3. Comparison of the methods with $\epsilon_1 = \epsilon_2 = 10^{-10}$ for the smallest eigenvalue of UNF matrices.

Method	Dimension			
	100	200	400	800
CVLM ₁	8.38(2.64) / 14.54(0.44)	9.93(3.24) / 14.31(0.49)	10.58(2.98) / 13.95(0.49)	11.02(3.28) / 13.66(0.54)
CVLM ₂	7.71(2.58) / 13.46(1.43)	9.21(3.15) / 13.17(1.47)	9.89(2.97) / 12.97(1.39)	10.24(3.32) / 12.45(1.44)
SPM ₁	6.51(0.96) / 14.71(0.46)	7.02(1.27) / 14.40(0.45)	8.24(1.80) / 14.09(0.44)	8.83(1.79) / 13.79(0.53)
SPM ₂	5.98(0.80) / 14.16(1.18)	6.40(1.11) / 13.53(1.33)	7.48(1.71) / 13.13(1.22)	8.16(1.73) / 13.03(1.07)

TABLE 4. Comparison of the methods for minimum and maximum eigenvalues with $\epsilon_1 = \epsilon_2 = 10^{-10}$ for CVL matrices.

Method		Dimension			
		100	200	400	800
M I N	CVLM ₁	8.52(2.58) / 14.08(0.73)	9.10(2.71) / 13.52(0.79)	10.45(3.18) / 12.87(0.65)	11.54(3.07) / 12.47(0.79)
	SPM ₁	6.24(1.39) / 14.14(0.69)	6.90(1.36) / 13.63(0.75)	7.67(1.43) / 13.08(0.81)	8.86(2.10) / 12.53(0.80)
	SPM ₂	5.86(1.28) / 13.89(0.85)	6.51(1.21) / 13.47(0.75)	7.29(1.29) / 12.93(0.77)	8.31(1.90) / 12.43(0.73)
M A X	CVLM ₁	14.37(2.97) / 15.19(0.43)	14.40(2.73) / 15.00(0.42)	15.32(3.19) / 14.77(0.45)	15.83(3.95) / 14.65(0.52)
	SPM ₁	12.15(1.61) / 15.31(0.43)	12.79(1.87) / 15.13(0.43)	13.20(1.64) / 14.95(0.46)	13.05(1.74) / 14.81(0.50)
	SPM ₂	11.22(1.63) / 12.53(2.06)	11.96(1.89) / 12.94(2.04)	12.47(1.60) / 13.53(1.68)	12.45(1.64) / 13.91(1.47)

TABLE 5. Comparison of the methods for minimum and maximum eigenvalues with $\epsilon_1 = \epsilon_2 = 10^{-10}$ for KMS matrices.

Method		Dimension			
		100	200	400	800
M I N	CVLM ₁	13.22(2.70) / 13.78(0.55)	15.30(2.65) / 13.31(0.60)	16.23(2.77) / 12.91(0.61)	17.91(2.81) / 12.38(0.67)
	SPM ₁	12.03(2.39) / 13.94(0.59)	14.28(2.73) / 13.42(0.58)	15.66(2.87) / 12.99(0.58)	17.19(2.73) / 12.57(0.64)
	SPM ₂	11.81(2.47) / 13.81(0.87)	14.23(2.78) / 13.40(0.63)	15.66(2.87) / 12.99(0.58)	17.19(2.73) / 12.57(0.64)
M A X	CVLM ₁	13.01(3.71) / 14.45(0.42)	15.84(4.22) / 14.10(0.43)	17.47(4.18) / 13.66(0.50)	19.21(4.24) / 13.28(0.52)
	SPM ₁	11.93(3.56) / 14.60(0.41)	14.85(4.15) / 14.24(0.45)	16.86(4.18) / 13.84(0.50)	18.66(4.31) / 13.41(0.50)
	SPM ₂	11.63(3.70) / 14.38(0.85)	14.73(4.29) / 14.18(0.62)	16.82(4.27) / 13.83(0.50)	18.64(4.33) / 13.40(0.48)

TABLE 6. Comparison of the methods for minimum and maximum eigenvalues with $\epsilon_1 = \epsilon_2 = 10^{-10}$ for UNF matrices.

Method		Dimension			
		100	200	400	800
M I N	CVLM ₁	8.03(2.34) / 14.52(0.39)	9.67(2.99) / 14.28(0.47)	11.15(3.31) / 13.94(0.50)	11.39(3.69) / 13.60(0.52)
	SPM ₁	6.68(1.13) / 14.70(0.41)	7.14(1.48) / 14.37(0.43)	8.18(1.87) / 14.10(0.48)	8.62(1.67) / 13.72(0.48)
	SPM ₂	6.05(0.94) / 14.10(1.16)	6.54(1.35) / 13.55(1.36)	7.48(1.78) / 13.19(1.22)	7.94(1.64) / 13.07(1.05)
M A X	CVLM ₁	14.79(3.03) / 15.37(0.42)	15.14(2.77) / 15.20(1.00)	15.89(3.48) / 15.07(0.47)	16.72(3.73) / 14.93(0.52)
	SPM ₁	13.12(1.60) / 15.44(1.03)	13.56(1.60) / 15.41(0.43)	13.81(1.70) / 15.25(0.45)	14.21(1.91) / 15.10(0.46)
	SPM ₂	12.27(1.58) / 13.24(2.03)	12.77(1.57) / 13.31(2.01)	13.20(1.58) / 13.81(1.79)	13.72(1.83) / 14.21(1.56)

TABLE 7. Comparison of the work per phase for minimum and maximum eigenvalues with $\epsilon_1 = \epsilon_2 = 10^{-10}$ for CVL matrices.

Method		Dimension			
		100	200	400	800
M	CVLM ₁	33(9) / 17(14) / 50(12)	36(8) / 17(14) / 46(11)	39(10) / 19(16) / 41(11)	42(10) / 20(15) / 38(10)
N	SPM ₂	35(7) / 12(10) / 53(12)	41(6) / 13(8) / 46(10)	47(7) / 13(8) / 40(8)	49(7) / 15(9) / 36(7)
M	CVLM ₁	9(2) / 57(8) / 34(7)	9(1) / 60(7) / 32(6)	8(1) / 62(7) / 31(6)	7(1) / 62(7) / 30(6)
X	SPM ₂	12(2) / 58(6) / 30(5)	10(1) / 63(5) / 27(4)	9(1) / 66(4) / 25(3)	9(1) / 66(4) / 25(3)

TABLE 8. Comparison of the work per phase for minimum and maximum eigenvalues with $\epsilon_1 = \epsilon_2 = 10^{-10}$ for KMS matrices.

Method		Dimension			
		100	200	400	800
M	CVLM ₁	8(2) / 68(8) / 24(7)	7(1) / 73(5) / 20(4)	6(1) / 76(5) / 18(4)	6(1) / 78(4) / 16(3)
N	SPM ₂	9(2) / 67(7) / 25(5)	7(2) / 71(6) / 21(5)	6(1) / 75(6) / 19(4)	6(1) / 77(4) / 17(3)
M	CVLM ₁	9(4) / 62(16) / 29(12)	8(5) / 69(15) / 23(11)	6(3) / 76(11) / 18(8)	6(2) / 79(8) / 16(6)
X	SPM ₂	11(5) / 60(17) / 29(12)	9(5) / 67(17) / 24(12)	7(3) / 74(12) / 19(9)	6(3) / 77(9) / 17(6)

TABLE 9. Comparison of the work per phase for minimum and maximum eigenvalues with $\epsilon_1 = \epsilon_2 = 10^{-10}$ for UNF matrices.

Method		Dimension			
		100	200	400	800
M	CVLM ₁	25(5) / 21(14) / 54(11)	24(6) / 27(15) / 49(11)	23(5) / 32(15) / 45(11)	22(5) / 34(15) / 44(11)
N	SPM ₂	31(4) / 17(8) / 52(8)	31(5) / 17(10) / 51(9)	31(6) / 25(14) / 44(10)	29(6) / 31(12) / 40(8)
M	CVLM ₁	8(1) / 61(8) / 32(7)	7(1) / 63(7) / 30(6)	7(1) / 65(7) / 28(7)	6(1) / 67(7) / 27(6)
X	SPM ₂	9(1) / 64(5) / 26(4)	9(1) / 67(4) / 25(3)	8(1) / 68(4) / 24(3)	8(1) / 70(4) / 22(3)

Comparing stopping rules, we can see clearly that stopping rule 2 is better than stopping rule 1, even though the improvement is modest for the classes of matrices that we investigated. The reported accuracy is generally lower for the second stopping rule; this is explained by the fact that this stopping rule is better at detecting when the relative accuracy of the eigenvalue falls below a given value, and therefore stops earlier. This does not mean that the method is less accurate: a higher tolerance in the stopping rule will achieve a higher accuracy for the eigenvalue. The improvement for the KMS matrices is marginal, due to the fact that in this case the algorithm usually terminates after the first iteration in phase II, so that only stopping rule 1 is used.

The computation of the largest eigenvalue presents a very different picture. Because the Frobenius norm provides a much cruder bound for the largest eigenvalue of our matrices than zero for the smallest, a large amount of work is spent in phase I, leading to a much higher overall number of flops. This is typical for any situation where the initial bound is not very good, such as when computing the smallest eigenvalue for a matrix which is not positive-definite. In these cases, the advantages of our method are sharply reduced.

Looking at Tables 7, 8 and 9, we find, in general, that slightly more than half the work is spent in phase I, and that there are no large differences between the methods in this respect. The percentage of work executed in the bisection part of phase I exhibits a large variability. Here too we should mention that the deviations from the mean are usually in the same direction for both methods and once again we found a lower variability for the SPM₂ method than for the CVLM₁ method. As can be expected, the worse the initial bound, the more work is spent in phase I. This is certainly true for the maximum eigenvalue for all three matrix classes, and it would also occur in the computation of the smallest eigenvalue of a matrix which is not positive-definite, if a good lower bound is not available. For KMS matrices, with their closely spaced eigenvalues, more than 80% of the work is spent in phase I, most of it in the bisection part.

We found that the error analysis from the previous section explained the numerical behavior of the algorithms very adequately. The limitations on the obtainable accuracy are mainly caused by the error in Durbin's algorithm. If so desired, iterative refinement can be used to improve accuracy. It is rather immaterial which method is used to solve the YW equations, as this does not influence the performance of the methods relative to one another; it only affects the absolute number of arithmetic operations that are executed, and the attainable accuracy.

REFERENCES

- [1] Ammar, G.S. and Gragg, W.B. (1987): The generalized Schur algorithm for the superfast solution of Toeplitz systems. In *Rational Approximations and its Applications in Mathematics and Physics*, J. Gilewicz, M. Pindor and W. Siemaszko, eds., *Lecture Notes in Mathematics* **1237**, Springer-Verlag, Berlin, pp. 315–330. MR **88c**:65032
- [2] Ammar, G.S. and Gragg, W.B. (1989): Numerical experience with a superfast Toeplitz solver. *Linear Algebra Appl.*, **121**, pp. 185–206. MR **90g**:65030
- [3] Andrew, A.L. (1973): Eigenvectors of certain matrices. *Linear Algebra Appl.*, **7**, pp. 151–162. MR **47**:6726
- [4] Bunch, J.R. (1985): Stability of methods for solving Toeplitz systems of equations. *SIAM J. Sci. Stat. Comput.*, **6**, pp. 349–364. MR **87a**:65073
- [5] Bunch, J.R., Nielsen, C.P., Sorensen, D.C. (1978): Rank-one modification of the symmetric eigenproblem. *Numer. Math.* **31**, pp. 31–48. MR **80g**:65038

- [6] Cantoni, A. and Butler, F. (1976): Eigenvalues and eigenvectors of symmetric centrosymmetric matrices. *Linear Algebra Appl.*, **13**, pp. 275–288. MR **53**:476
- [7] Cuppen, J.J.M. (1981): A divide and conquer method for the symmetric tridiagonal eigenvalue problem. *Numer. Math.*, **36**, pp. 177–195. MR **82d**:65038
- [8] Cybenko, G. (1980): The numerical stability of the Levinson-Durbin algorithm for Toeplitz systems of equations. *SIAM J. Sci. Stat. Comput.*, **1**, pp. 303–319. MR **82i**:65019
- [9] Cybenko, G. and Van Loan, C. (1986): Computing the minimum eigenvalue of a symmetric positive definite Toeplitz matrix. *SIAM J. Sci. Stat. Comput.*, **7**, pp. 123–131. MR **87b**:65042
- [10] Delsarte, P. and Genin, Y. (1983): Spectral properties of finite Toeplitz matrices. In *Mathematical Theory of Networks and Systems, Proc. MTNS-83 International Symposium*, Beer-Sheva, Israel, Lecture Notes in Computer Sci., vol. 58, Springer-Verlag, Berlin, pp. 194–213. MR **86k**:15003
- [11] Dembo, A. (1988): Bounds on the extreme eigenvalues of positive definite Toeplitz matrices. *IEEE Trans. Inform. Theory*, **34**, pp. 352–355. MR **89b**:15028
- [12] Durbin, J. (1960): The fitting of time series model. *Rev. Inst. Int. Stat.*, **28**, pp. 233–243.
- [13] Golub, G.H. (1973): Some modified matrix eigenvalue problems. *SIAM Rev.*, **15**, pp. 318–334. MR **48**:7569
- [14] Golub, G. and Van Loan, C. (1989): *Matrix Computations*. The Johns Hopkins University Press, Baltimore and London. MR **90d**:65055
- [15] Hertz, D. (1992): Simple bounds on the extreme eigenvalues of Toeplitz matrices. *IEEE Trans. Inform. Theory*, **38**, pp. 175–176. MR **92j**:15005
- [16] Hu, Y.H. and Kung, S.Y. (1985) A Toeplitz eigensystem solver. *IEEE Trans. Acoust. Speech Signal Process.*, **34**, pp. 1264–1271. MR **87f**:65044
- [17] Huang, D. (1992): Symmetric solutions and eigenvalue problems of Toeplitz systems. *IEEE Trans. Signal Processing*, **40**, pp. 3069–3074.
- [18] Isaacson, E. and Keller, H.B. (1994): *Analysis of Numerical Methods*. Dover Publications, Inc., New York. CMP 94:14; MR **34**:924 (original ed.)
- [19] Kac, M., Murdock, W.L. and Szegő, G. (1953): On the eigenvalues of certain Hermitian forms. *J. Rat. Mech. Anal.*, **2**, pp. 767–800. MR **15**:538b
- [20] Levinson, N. (1947): The Wiener RMS (root mean square) error criterion in filter design and prediction. *J. Math. and Phys.*, **25**, pp. 261–278. MR **8**:391e
- [21] Mackens, W. and Voss, H. (1997): The minimum eigenvalue of a symmetric positive-definite Toeplitz matrix and rational Hermitian interpolation. *SIAM J. Matrix Anal. Appl.*, **18**, pp. 521–534. MR **98e**:65029
- [22] Melman, A. (1997): A unifying convergence analysis of second-order methods for secular equations. *Math. Comp.*, **66**, pp. 333–344. MR **97e**:65036
- [23] Melman, A. (1998): Spectral functions for real symmetric Toeplitz matrices. *J. Comput. Appl. Math.*, **98**, pp. 233–243. MR **99i**:65036
- [24] Slepian, D. and Landau, H.J. (1978): A note on the eigenvalues of Hermitian matrices. *SIAM J. Math. Anal.*, **9**, pp. 291–297. MR **57**:6052
- [25] Trench, W.F. (1989): Numerical solution of the eigenvalue problem for Hermitian Toeplitz matrices. *SIAM J. Matrix Anal. Appl.*, **10**, pp. 135–146. MR **90i**:65068
- [26] Voss, H. (1999): Symmetric schemes for computing the minimum eigenvalue of a symmetric Toeplitz matrix. *Linear Algebra Appl.*, **287**, pp. 359–371. MR **99m**:65073

BEN-GURION UNIVERSITY, BEER-SHEVA, ISRAEL

Current address: Department of Computer Science, SCCM Program, Stanford University, Stanford, California 94305-9025

E-mail address: melman@sccm.stanford.edu