A hybrid method for computing the smallest eigenvalue of a symmetric and positive definite Toeplitz matrix

Aleksandra Kostić Heinrich Voss*

January 15, 2003

Abstract

In this paper we suggest a hybrid method for computing the smallest eigenvalue of a symmetric and positive definite Toeplitz matrix which takes advantage of two types of methods, Newton's method for the characteristic polynomial and projection methods based on rational interpolation of the secular equation.

Keywords: Toeplitz matrix, eigenvalue problem, hybrid method

AMS Subject Classification: 65F15

1 Introduction

The problem of finding the smallest eigenvalue λ_1 of a real symmetric, positive definite Toeplitz matrix T is of considerable interest in signal processing. Given the covariance sequence of the observed data, Pisarenko [13] suggested a method which determines the sinusoidal frequencies from the eigenvector of the covariance matrix associated with its minimum eigenvalue.

Several methods have been reported in the literature for computing the minimum eigenvalue of T, cf. [2], [4], [5], [6], [7], [8], [9], [10], [11], [12], [15], [16], [17], e.g.

In their seminal paper Cybenko and Van Loan [2] presented the following method: by bisection they determine an initial approximation $\mu_0 \in (\lambda_1, \omega_1)$,

*Hamburg University of Technology, Department of Mathematics, D-21071 Hamburg, Federal Republic of Germany, ({kostic,voss}@tu-harburg.de)

where ω_1 denotes the smallest pole of the secular equation f, and they improve μ_0 by Newton's method for f which converges monotonely and quadratically to λ_1. This approach was improved considerably by Mackens and the second author by replacing Newton's method by a more appropriate root finding method for the secular equation, which is a rational function where the wanted root λ_1 and the smallest pole ω_1 can be very close to each other. Namely, we used a root finding method based on rational Hermitean interpolation in [8], and since this one turned out to be equivalent to a projection method where the eigenvalue problem for T is projected to a two dimensional subspace it could be improved further in [9]. The root finding phase of the approach usually converges very fast whereas the bisection phase can be very costly since no procedure is known how to obtain an initial value $\mu_0 \in (\lambda_1, \omega_1)$ efficiently.

Mastronardi and Boley [11] suggested Newton's method for the characteristic polynomial χ_n of T which can be enhanced by a double step strategy and/or by Hermitean interpolation of the characteristic polynomial (cf. [10]). The advantage of this method upon Cybenko and Van Loan's approach is its conceptual simplicity, and its monotone convergence from below starting with any lower bound of λ_1, for instance with the initial guess $\mu_0 = 0$. However, this convergence usually is slower than that of the root finding method mentioned in the last paragraph.

To conclude, for the secular equation Newton's method requires an expensive preprocessing to obtain a suitable initial value which is not the case for the characteristic polynomial. On the other hand every single step is less expensive for the secular equation, and the convergence usually is faster.

It is interesting to note that the dominant share of the cost for each step of both types of methods is the solution of a Yule–Walker system to evaluate the secular equation and its derivative at a guess λ and the characteristic polynomial of T, respectively. This property suggests to combine both methods in a hybrid approach, and to gather information about one method while performing an iteration step of the other one, thus squeezing the most out of a Yule–Walker solve.

In this paper we propose a method of this type. We start with Newton's method for the characteristic polynomial getting lower bounds of the smallest eigenvalue λ_1. Concurrently we obtain upper bounds of λ_1 from rational interpolations of the secular equation which then yield a good approximation to λ_1 as initial guess for the projection method mentioned above. While evaluating the secular function f in projection steps, at the same time we obtain for free lower bounds of the smallest root ω_1 of the characteristic polynomial of the principal submatrix of T of dimension $n-1$. A further

rational interpolation of f then yields good lower bounds of λ_1 thus gaining an efficient stopping criterion.

The paper is organized as follows. Section 2 summarizes properties of the underlying methods using the characteristic polynomial and the secular equation, respectively. Section 3 suggests the new hybrid method including the improved stopping criterion, and Section 4 demonstrates its efficiency by numerical examples. The paper closes with concluding remarks concerning the use of superfast Toeplitz solvers.

2 Basic approaches

Let $T_n = (t_{|i-j|}) \in \mathbb{R}^{n\times n}$ be a symmetric and positive definite Toeplitz matrix. We consider the problem to determine the smallest eigenvalue (and corresponding eigenvector) of T_n. In this section we summarize two basic methods which will be used in the next section to form a hybrid method.

Mastronardi and Boley suggested to determine the smallest eigenvalue λ_1 of T_n by Newton's method for the characteristic polynomial of T_n.

Denote the j-dimensional principal submatrix of T_n by T_j, and the first column of T_{j+1} omitting the diagonal entry t_0 by $t^{(j)} := (t_1\,,\, t_2\,,\, \ldots\,,\, t_{j-1}\,,\, t_j)^T$. If λ is not an eigenvalue of T_j then

$$T_{j+1}-\lambda I_{j+1} = \begin{pmatrix} t_0-\lambda+(t^{(j)})^T y^{(j)}(\lambda) & -(y^{(j)}(\lambda))^T \\ 0 & I_j \end{pmatrix} \begin{pmatrix} 1 & 0^T \\ t^{(j)} & T_j-\lambda I_j \end{pmatrix}$$

where $y^{(j)}(\lambda)$ denotes the solution of the j-th Yule-Walker system

$$(T_j-\lambda I_j)y^{(j)}(\lambda) = -t^{(j)}.$$

Hence, if λ is not in the spectrum of any of the principal submatrices T_j, $j = 1,\ldots,n-1$, then the characteristic polynomials $\chi_j(\lambda) := \det(T_j-\lambda I_j)$ of T_j satisfy the recurrence relation

$$\chi_{j+1}(\lambda) = \chi_j(\lambda)(t_0-\lambda+(t^{(j)})^T y^{(j)}(\lambda)) =: \chi_j(\lambda)\beta_j(\lambda), \quad 1\le j\le n-1,$$

and for the derivatives it holds that

$$\chi'_{j+1}(\lambda) = \chi'_j(\lambda)\beta_j(\lambda) - \chi_j(\lambda)(1+\|y^{(j)}(\lambda)\|_2^2).$$

Durbin's algorithm (cf.[3], p. 194 ff) for the Yule–Walker system

$$(T_{n-1}-\lambda I_{n-1})y^{(n-1)}(\lambda) = -t^{(n-1)}$$

yields the parameters $\beta_j(\lambda)$ and the vectors $y^{(j)}(\lambda)$, $j = 1, \ldots, n-1$. Hence, the characteristic polynomial $\chi_n(\lambda)$ can be evaluated using Durbin's algorithm, and the cost of one evaluation is $2n^2+\mathcal{O}(n)$ flops. With n^2 additional flops to determine $\|y^{(j)}(\lambda)\|_2^2$, $j = 1, \ldots, n-1$ one can evaluate the derivative $\chi_n'(\lambda)$ as well. Therefore, one Newton step for the characteristic polynomial requires $3n^2 + \mathcal{O}(n)$ flops.

For $\lambda < \lambda_1$ the characteristic polynomial $\chi_n(\lambda)$ is convex and monotonically decreasing, and thus for every initial value $\mu_0 < \lambda_1$ (for instance $\mu_0 = 0$) Newton's method converges from the left to the smallest eigenvalue of T_n. For the same reason the secant method converges monotonically increasing to λ_1 if the initial guesses μ_0, μ_1 satisfy $\mu_0 < \mu_1 \leq \lambda_1$.

Newton's method and the secant method very often converge slowly in the beginning if the initial value μ_0 is far from the smallest root of χ_n or if the slope of χ_n is very large. In our numerical examples we observed up to 45 Newton steps to determine the smallest eigenvalue up to a relative error 10^{-6}, and the secant method converges even slowlier. The global behaviour of both methods can be improved considerably by a double step strategy.

Theorem 1 (Stoer and Bulirsch [14], p. 274)
Let $\chi(\lambda)$ be a polynomial of degree $n > 2$, all roots of which are real, $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. Let ξ_1 be the smallest root of $\chi'(\lambda)$. Then for every $\mu < \lambda_1$ the numbers

$$\mu' := \mu - \frac{\chi(\mu)}{\chi'(\mu)},\ \nu := \mu - 2\frac{\chi(\mu)}{\chi'(\mu)},\ \nu' := \nu - \frac{\chi(\nu)}{\chi'(\nu)}$$

are well defined, and they satisfy $\nu < \xi_1$, and $\mu' \leq \nu' \leq \lambda_1$.

Theorem 1 suggests the following acceleration of Newton's method: Double steps

$$\mu_{k+1} := \mu_k - 2\frac{\chi_n(\mu_k)}{\chi_n'(\mu_k)}$$

are performed until $\mu_{k+1} > \lambda_1$, which is signaled by the change of sign of the Newton increment. Then the method switches to the original Newton method. An analogous result holds for the secant method.

One drawback of Newton's method and the secant method is that they take advantage only of the last iterate and the last two iterates, respectively, but they do not use information gained in previous steps. Improvements employing interpolations of previous iterates were proposed in [10]. However, we do not take advantage of these modifications in the hybrid method considered here.

A different approach for computing the smallest eigenvalue λ_1 of T_n was suggested by Cybenko and Van Loan [2] who determined λ_1 as the smallest

root of the secular function

$$f(\lambda) = \frac{-\chi_n(\lambda)}{\chi_{n-1}(\lambda)} = -t_0 + \lambda + (t^{(n-1)})^T y^{(n-1)}(\lambda) = -\beta_{n-1}(\lambda)$$

which can be evaluated by Durbin's algorithm as well. Since

$$f'(\lambda) = 1 + \|y^{(n-1)}(\lambda)\|^2,$$

the evaluation of $f'(\lambda)$ then essentially comes for free, and for the secular equation a Newton step requires only $2n^2 + \mathcal{O}(n)$ flops.

Let ω_1 denote the smallest eigenvalue of T_{n-1}, i.e. the smallest pole of f. Then f is monotonically increasing and convex for $\lambda < \omega_1$, and therefore Newton's method for the secular equation converges monotonically decreasing for every initial value $\mu_0 \in [\lambda_1, \omega_1)$.

Cybenko and Van Loan [2] suggested to determine a suitable initial value μ_0 by bisection based on Durbin's algorithm. If μ is not in the spectrum of any of the principal submatrices of $T_n - \mu I_n$ then Durbin's algorithm applied to $T_n - \mu I_n$ determines a unit left triangular matrix L and a diagonal matrix D such that

$$L(T_n - \mu I_n)L^T = D := \operatorname{diag}\{1, \delta_1, \ldots, \delta_{n-1}\}.$$

Hence, from Sylvester's law of inertia one gets $\mu < \lambda_1$, if $\delta_j > 0$ for $j = 1, \ldots, n-1$, $\mu \in [\lambda_1, \omega_1)$, if $\delta_j > 0$ for $j = 1, \ldots, n-2$ and $\delta_{n-1} \leq 0$, and $\mu > \omega_1$, if $\delta_j < 0$ for some $j \in \{1, \ldots, n-2\}$.

Hence, for the secular equation Newton's method requires an expensive preprocessing to obtain a suitable initial value which is not the case for the characteristic polynomial. On the other hand every single step is less expensive for the secular equation, and the convergence usually is faster. Since both methods use Durbin's algorithm as a basic building block it is reasonable to combine both methods in a hybrid approach.

Actually, we combine the Newton process for the characteristic polynomial with a modification of Cybenko's and Van Loan's method. The global convergence behaviour of Newton's method for the secular equation usually is not satisfactory since the smallest root λ_1 and the smallest pole ω_1 of the rational function f can be very close to each other. In this situation the initial steps of Newton's method are extremely slow, at least if the initial guess is close to ω_1.

Approximating the secular equation by a suitable rational function the convergence of the method can be improved considerably. f can be rewritten

as

$$f(\lambda) = f(0) + \lambda f'(0) + \lambda^2 \sum_{j=1}^{n-1} \frac{\alpha_j^2}{\omega_j - \lambda} \tag{1}$$

where α_j are real numbers depending on the eigenvectors of T_{n-1} and ω_j denote the eigenvalues of T_{n-1} ordered by magnitude (cf. [8]), and with a shift μ which is not in the spectrum of T_{n-1} it obtains the form

$$f(\lambda) = f(\mu) + (\lambda - \mu) f'(\mu) + (\lambda - \mu)^2 \phi(\lambda; \mu) \tag{2}$$

where

$$\phi(\lambda; \mu) = \sum_{j=1}^{n-1} \frac{\alpha_j^2 \gamma_j^2}{\omega_j - \lambda}, \quad \gamma_j = \frac{\omega_j}{\omega_j - \mu}. \tag{3}$$

The representation (1) of f suggests to replace the linearization of f in Newton's method by a root finding method based on a rational model

$$g(\lambda; \mu) = f(0) + \lambda f'(0) + \lambda^2 \frac{b}{c - \lambda}, \tag{4}$$

where μ is a given approximation to λ_1 and b and c are determined such that

$$g(\mu; \mu) = f(\mu), \; g'(\mu; \mu) = f'(\mu). \tag{5}$$

For this modification the following convergence result was proved in [8].

Theorem 2: *Let g be given by (4) and (5) where μ is not in the spectrum of T_{n-1}, and denote by $\rho(\mu)$ the smallest positive root of $g(\cdot; \mu)$. Then $\rho(\mu) \geq \lambda_1$.*

If $\mu_0 \in (\lambda_1, \omega_1)$ and $\mu_{k+1} = \rho(\mu_k)$, then the sequence $\{\mu_k\}$ converges monotonically decreasing to λ_1, the convergence is quadratic, and it is faster than Newton's method, i.e. if $\nu_{k+1} = \mu_k - f(\mu_k)/f'(\mu_k)$ then $\lambda_1 \leq \mu_{k+1} \leq \nu_{k+1}$.

In [9] it was shown that $\rho(\mu)$ is the smallest eigenvalue of the projected eigenproblem

$$[q(0), q(\mu)]^T T_n [q(0), q(\mu)] y = \kappa [q(0), q(\mu)]^T [q(0), q(\mu)] y$$

where $q(\mu)$ solves the linear system

$$(T_n - \mu I) q(\mu) = -f(\mu) e_1,$$

and e_1 is the unit vector containing a 1 in its first component and 0 anywhere else. This result suggests to improve the method further by defining μ_{k+1} as the smallest eigenvalue of the projected problem

$$A_k y := Q_k^T T_n Q_k y = \kappa Q_k^T Q_k y =: \kappa B_k y, \quad \text{for } Q_k = [q(0), q(\mu_1), \ldots, q(\mu_k)]$$

where $\mu_1, \ldots, \mu_k \in (0, \omega_1)$ are parameters obtained in the course of the algorithm.

It is interesting to note that every step essentially requires one Yule-Walker solve since

$$q(\mu)^T q(\nu) = \begin{cases} f'(\mu) & \text{for } \mu = \nu \\ (f(\mu) - f(\nu))/(\mu - \nu) & \text{for } \mu \neq \nu \end{cases}$$

and

$$q(\mu)^T T_n q(\nu) = -f(\mu) + \mu q(\mu)^T q(\nu).$$

For this method the following convergence result was proved in [9] by comparison with inverse iteration with Rayleigh quotient shifts.

Theorem 3

Let $\mu_1, \ldots, \mu_{\ell-1}$ be not in the spectrum of T_{n-1}, let $\mu_\ell \in (\lambda_1, \omega_1)$, and for $k \geq \ell$ let μ_{k+1} be the smallest eigenvalue of the projected problem $A_k y = \kappa B_k y$. Then the sequence $\{\mu_k\}$ converges monotonically decreasing to λ_1, and the convergence is at least cubic.

3 A HYBRID METHOD

We are now in the position to draft a hybrid algorithm. We start with a Newton step for the characteristic polynomial χ_n with initial guess $\mu_0 = 0$ followed by a secant step yielding the lower bounds $\mu_0 < \mu_1 < \mu_2 < \lambda_1$. We evaluate f at μ_0, μ_1 and μ_2, where $f(\mu_0)$ and $f(\mu_1)$ come for free, and determine the smallest positive roots ρ_j of $g(\cdot; \mu_j)$ for $j = 1, 2$, and the projected matrices A_2 and B_2.

Since λ_1 is the smallest fixed point of $\rho(\cdot)$ the fixed point μ_3 of the linear interpolation of (μ_1, ρ_1) and (μ_2, ρ_2) should be a reasonable approximation to λ_1. Figure 1 contains the graph of ρ and the linear interpolation with fixed points λ_1 and μ_3, respectively, for 2 examples, a typical one on the left and a less typical one on the right. Although in the example on the right the linear interpolation is far from being a good approximation to ρ its fixed point μ_3 is a reasonable approximation to ρ's fixed point λ_1.

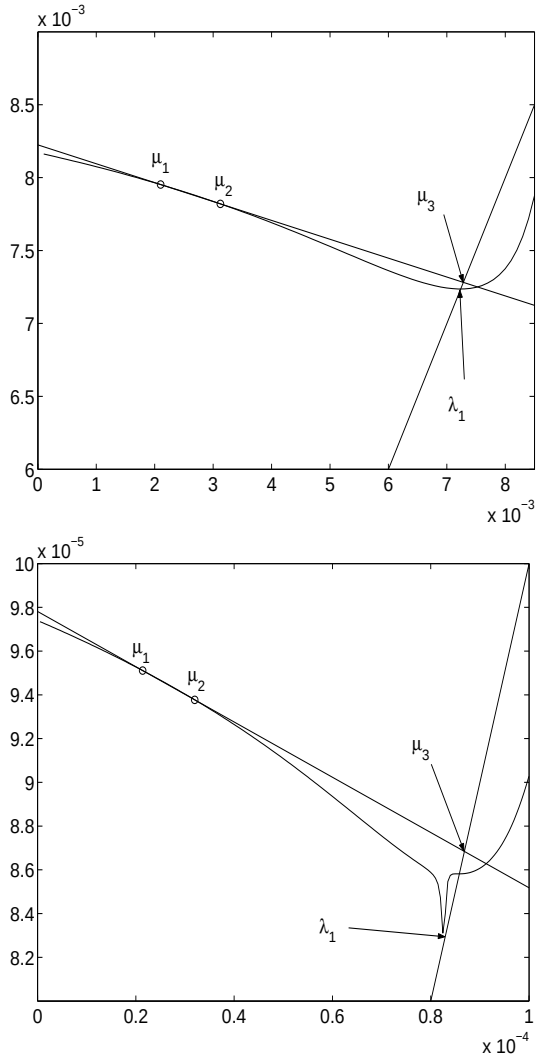

Figure 1: Fixed point of $\rho(\cdot)$ and of linear interpolation of $\rho(\cdot)$

Table 1: Accuracy of interpolation / Distance of root and pole

dim.	relative error of μ_3		rel. distance of λ_1 and ω_1	
	max	average	min	average
32	$9.69E-02$	$1.01E-02$	$6.07E-03$	$3.94E-01$
64	$1.20E-01$	$1.49E-02$	$7.31E-04$	$2.22E-01$
128	$2.87E-01$	$2.03E-02$	$5.86E-04$	$1.32E-01$
256	$1.59E-01$	$1.89E-02$	$4.25E-04$	$7.22E-02$
512	$2.31E-01$	$2.35E-02$	$2.70E-05$	$3.96E-02$
1024	$1.93E-01$	$2.62E-02$	$1.33E-05$	$1.94E-02$
2048	$2.75E-01$	$2.61E-02$	$1.32E-05$	$1.21E-02$

To test the accuracy of the approximation μ_3 to λ_1 we considered Toeplitz matrices

$$T = \sum_{k=1}^{n} \eta_k T_{2\pi\theta_k}, \quad T_\theta = (t_{ij}) = (\cos(\theta(i-j)))$$

where η_k and θ_k are uniformly distributed random numbers taken from $[0,1]$ (cf. Cybenko and Van Loan [2]).

For each of the dimensions $n = 32$, 64, 128, 256, 512, 1024 and $n = 2048$ we considered 100 test problems. On the left hand Table I contains the maximum relative error and the average of relative errors of μ_3.

If μ_3 is less than ω_1 then we update the matrices A_2 and B_2 and continue with the projection method outlined in Section 2. For $\mu_3 \in (\lambda_1, \omega_1)$ we then obtain monotone convergence to λ_1.

It may happen that $\mu_3 > \omega_1$ or that the smallest eigenvalue μ_j of $A_{j-1}y = \mu B_{j-1}y$ for some $j \geq 4$ is greater than ω_1. Since Durbin's algorithm is known to be stable only for a positive definite system matrix (cf. [1]), i.e. since the evaluation of $f(\mu)$ is known to be stable only for $\mu < \omega_1$, we do not take into consideration parameters $\mu_j > \omega_1$ in the projection method. In this case we replace μ_j by the weighted bisection $\mu_j \leftarrow 0.1\tilde{\mu} + 0.9\mu_j$ where $\tilde{\mu} := \max\{\mu_k < \lambda_1 \,:\, 2 \leq k \leq j-1\}$.

The smallest root λ_1 and the smallest pole ω_1 may be very close to each other (Table I on the right contains the minimum relative distance and the average of relative distances in our test examples), and it may happen that the method bounces between lower bounds of λ_1 obtained from the weighted bisection and upper bounds of ω_1 obtained as smallest eigenvalues of projected problems. We remedy this behaviour by the following tie break rule which was introduced in [8] for the method based on root finding by

rational Hermitean interpolation.

If $\mu_j < \lambda_1$ then we evaluate $f(\mu_j)$ and $f'(\mu_j)$ to update A_{j-1} and B_{j-1}, and we determine μ_{j+1} as smallest eigenvalue of $A_j y = \mu B_j y$. It is obvious that we get for free a further upper bound $\tilde{\mu}_{j+1} = \mu_j - f(\mu_j)/f'(\mu_j)$ from Newton's method, where both approximations, $\tilde{\mu}_{j+1}$ and μ_{j+1} correspond to quadratically or even cubically convergent processes. If the relative distance of these two bounds is not small then μ_j cannot be a good approximation to λ_1, and it is not unlikely that $\mu_{j+1} > \omega_1$. Hence, if $|\mu_{j+1} - \tilde{\mu}_{j+1}|/\mu_{j+1} > 0.01$ then we replace μ_{j+1} by $\mu_{j+1} \leftarrow 0.1\mu_j + 0.9\mu_{j+1}$.

Until now we applied the characteristic polynomial of T_n only to construct an initial guess for the projection method. While evaluating $f(\mu_k)$ in subsequent steps of the projection method we make further use of the characteristic polynomial $\chi_{n-1}(\mu_k)$ of the principal submatrix T_{n-1} (which comes for free) to terminate the method efficiently. The termination is based on the following Theorem which yields a lower bound of λ_1 using a further rational interpolation of the secular equation. In distinction to the rational approximation of f in (4) and (5) here the pole is a fixed lower bound of ω_1.

Theorem 4: *Let $\kappa \in (0, \lambda_1)$, $\mu \in (\kappa, \omega_1)$ and $p \in (\kappa, \omega_1)$. Let*

$$h(\lambda) := f(\mu) + f'(\mu)(\lambda - \mu) + (\lambda - \mu)^2 \frac{b}{p - \lambda},$$

where b is determined such that the interpolation condition $h(\kappa) = f(\kappa)$ holds.

Then $b > 0$, i.e. h is strictly monotonically increasing and strictly convex in $(0, p)$, and the unique root of h in $(0, p)$ is a lower bound of λ_1.

Proof: From equation (2) and from the interpolation condition $h(\kappa) = f(\kappa)$ we obtain

$$b = (p - \kappa)\phi(\kappa; \mu) > 0.$$

That the unique root $\tilde{\lambda}$ of h in $(0, p)$ is a lower bound of λ_1 is obvious for $p \leq \lambda_1$. For $p > \lambda_1$ we have to show $h(\lambda_1) > 0$. This follows from equations (2) and (3):

$$
\begin{aligned}
h(\lambda_1) &= f(\mu) + f'(\mu)(\lambda_1 - \mu) + (\lambda_1 - \mu)^2 \frac{b}{p - \lambda_1} \\
&= f(\lambda_1) - (\lambda_1 - \mu)^2 \left(\phi(\lambda_1; \mu) - \frac{(p - \kappa)\phi(\kappa; \mu)}{p - \lambda_1} \right) \\
&= \frac{(\lambda_1 - \mu)^2}{p - \lambda_1} \left((p - \kappa)\phi(\kappa; \mu) - (p - \lambda_1)\phi(\lambda_1; \mu) \right) \\
&= \frac{(\lambda_1 - \mu)^2}{p - \lambda_1} \sum_{j=1}^{n-1} \alpha_j^2 \gamma_j^2 \left(\frac{p - \kappa}{\omega_j - \kappa} - \frac{p - \lambda_1}{\omega_j - \lambda_1} \right) \\
&= \frac{(\lambda_1 - \mu)^2}{p - \lambda_1} \sum_{j=1}^{n-1} \alpha_j^2 \gamma_j^2 \frac{(\omega_j - p)(\lambda_1 - \kappa)}{(\omega_j - \kappa)(\omega_j - \lambda_1)} > 0. \quad \square
\end{aligned}
$$

Since the algorithm eventually enters the interval (λ_1, ω_1) we obtain a suitable lower bound p of ω_1 in the course of the algorithm. We can even do a little better. ω_1 is the smallest root of the characteristic polynomial χ_{n-1} of T_{n-1}, all roots of which are real. Hence, given two lower bounds μ_k and μ_{k+1} of ω_1 a secant step for χ_{n-1} with these parameters yields a lower bound of ω_1, too, which is bigger than $\max\{\mu_k, \mu_{k+1}\}$. Notice that evaluating f by Durbin's algorithm at some μ we get $\chi_{n-1}(\mu)$ for free.

Putting these considerations together we end up with Algorithm 1.

4 NUMERICAL EXPERIMENTS

We applied the hybrid method to the test examples mentioned in the last section. Table II contains the average number of flops and the average number of Durbin steps needed to determine the smallest eigenvalue in 100 test problems where the iteration was terminated if the error was guaranteed to be less than 10^{-6}. For comparison we added the cost of the projection method in [9], and a Newton type method for the characteristic polynomial in [10] which uses previous iterations in a Hermitean interpolation. Notice that every step of the projection method and the hybrid method requires $2n^2 + \mathcal{O}(n)$ flops whereas every step of the Newton type method requires $3n^2 + \mathcal{O}(n)$ flops.

We already mentioned in Section 2 that Newton's method and the secant iteration for the characteristic polynomial can be accelerated using double steps, and of course these can be introduced in the initial phase of the hybrid method, too. However, then it may happen that after the first double

Algorithm 1 Hybrid method for computing the smallest eigenvalue

Initial guess $\mu_0 = 0$
Determine μ_1 by Newton's method and μ_2 by secant method for χ_n
  simultaneously for $j = 1, 2$ determine roots ρ_j of $g(\cdot; \mu_j)$
  simultaneously determine matrices A_2 and B_2
  store best known lower bound $\lambda_\ell = \mu_2$ and
  minimum eigenvalue of $A_2 y = \mu B_2 y$ as upper bound λ_u of λ_1
  determine lower bound p of ω_1 by secant step for χ_{n-1}
Determine fixed point μ_3 of linear interpolation of (μ_j, ρ_j), $j = 1, 2$
for $k = 3, 4, \ldots$ until convergence **do**
  Evaluate $f(\mu_k)$, $f'(\mu_k)$, $\chi_{n-1}(\mu_k)$ by Durbin's algorithm
  and decide whether $\mu_k < \lambda_1$ or $\mu_k \in (\lambda_1, \omega_1)$ or $\mu_k > \omega_1$
  if $\mu_k > \omega_1$ **then**
    $\lambda_u = \min\{\lambda_u, \mu_k\}$
    $\mu_k \leftarrow 0.1 * \lambda_\ell + 0.9 * \lambda_u$
  else
    update $A_{k-1} \to A_k$ and $B_{k-1} \to B_k$
    determine smallest eigenvalue ν of $A_k y = \mu B_k y$
    $\lambda_u = \min\{\lambda_u, \nu\}$
    determine new lower bound p of ω_1 by secant step for χ_{n-1}
    and new lower bound λ_ℓ of λ_1 using Theorem 4
    if $\mu_k \in (\lambda_1, \omega_1)$ **then**
      $\mu_{k+1} = \lambda_u$
    else
      $\lambda_{\text{newton}} = \mu_k - f(\mu_k)/f'(\mu_k)$
      if $|\lambda_{\text{newton}} - \nu|/\nu| < 0.01$ **then**
        $\mu_{k+1} = \nu$
      else
        $\mu_{k+1} = 0.1 * \lambda_\ell + 0.9 * \nu$
      end if
    end if
  end if
end for

Table 2: Average number of flops and Durbin calls

dim.	projection		Newton type		hybrid	
	flops	steps	flops	steps	flops	steps
32	$1.226E04$	4.49	$1.759E04$	4.01	$1.418E4$	4.92
64	$4.716E04$	5.08	$6.171E04$	4.36	$5.248E4$	5.27
128	$1.871E05$	5.41	$2.113E05$	4.10	$2.047E5$	5.46
256	$8.242E05$	6.13	$8.861E05$	4.42	$8.295E5$	5.69
512	$3.582E06$	6.75	$3.968E06$	5.00	$3.456E6$	6.02
1024	$1.523E07$	7.22	$1.627E07$	5.15	$1.463E7$	6.44
2048	$6.588E07$	7.83	$6.593E07$	5.23	$5.728E7$	6.31

Table 3: Cost of hybrid method and its modification

dim.	hybrid		modification	
	flops	steps	flops	steps
32	$1.418E4$	4.92	$1.308E4$	4.48
64	$5.248E4$	5.27	$4.992E4$	4.98
128	$2.047E5$	5.46	$1.880E5$	4.97
256	$8.295E5$	5.69	$7.666E5$	5.22
512	$3.456E6$	6.02	$3.180E6$	5.50
1024	$1.463E7$	6.44	$1.339E7$	5.85
2048	$5.728E7$	6.31	$5.459E7$	5.99

Newton step $\mu_1 > \omega_1$ or $\lambda_1 < \mu_1 < \omega_1$. In both cases we continue directly with the projection method where for $\mu_1 > \omega_1$ we first replace μ_1 by $\mu_1 \leftarrow 0.5\mu_1$, i.e. the single Newton step. Similarly, if $\mu_1 < \lambda_1$ and the iterate μ_2 of a double secant step satisfies $\mu_2 > \lambda_1$ then we continue directly with the projection method, and again we replace μ_2 by the result of a single secant step if $\mu_2 > \omega_1$. If $\mu_1 < \mu_2 < \lambda_1$ then we determine μ_3 by linear interpolation of ρ and proceed as in the hybrid method above. These modifications yield a further improvement of the method. Table III contains the average number of flops and Durbin calls for the hybrid method and its modification.

5 CONCLUDING REMARKS

We have presented a hybrid method for computing the smallest eigenvalue of a symmetric and positive definite Toeplitz matrix taking advantage of

both types of methods, Newton's method for the characteristic polynomial and projection methods based on rational approximations of the secular equation. At least for high dimensions these methods yield noteworthy improvements over the underlying approaches. We used Durbin's method to solve the occurring Yule-Walker equations and to determine the Schur parameters δ_j requiring $2n^2 + \mathcal{O}(n)$ flops. Of course this could have been done by superfast Toeplitz solvers the complexity of which is only $\mathcal{O}(n \log^2 n)$ operations. Notice however, that this pays only if the dimension n is larger than 512. In a similar way as in [15] the method can be further enhanced taking advantage of symmetry properties of the eigenvectors of a symmetric Toeplitz matrix.

References

[1] Cybenko G. The numerical stability of the Levinson-Durbin algorithm for Toeplitz systems of equations. *SIAM J. Sci. Stat. Comput.* 1980; **1**: 303–309

[2] Cybenko G, Van Loan CF. Computing the minimum eigenvalue of a symmetric positive definite Toeplitz matrix. *SIAM J. Sci. Stat. Comput.* 1986; **7**: 123–131

[3] Golub GH, Van Loan CF. *Matrix Computations* (3rd edn). The John Hopkins University Press: Baltimore and London 1996

[4] Hu YH, Kung SY. Toeplitz eigensystem solver. *IEEE Trans. Acoustics, Speech, Signal Processing* 1985; **33**: 1264–1271

[5] Huckle T. Computing the minimum eigenvalue of a symmetric positive definite Toeplitz matrix with spectral transformation Lanczos method. In *Numerical Treatment of Eigenvalue Problems, Vol. 5* Albrecht J, Collatz L, Hagedorn P, Velte W (eds). Birkhäuser: Basel 1991; 109–115

[6] Huckle T. Circulant and skewcirculant matrices for solving Toeplitz matrices. *SIAM J. Matr. Anal. Appl.* 1992; **13**: 767–777

[7] Kostić A, Voss H. A method of order $1 + \sqrt{3}$ for computing the smallest eigenvalue of a symmetric Toeplitz matrix. *WSEAS Transactions on Mathematics* 2002; **1**: 1–6

[8] Mackens W, Voss H. The minimum eigenvalue of a symmetric positive definite Toeplitz matrix and rational Hermitian interpolation. *SIAM J. Matr. Anal. Appl.* 1997; **18**: 521–534

[9] Mackens W, Voss H. A projection method for computing the minimum eigenvalue of a symmetric positive definite Toeplitz matrix. *Lin. Alg. Appl.* 1998; **275–276**: 401–415

[10] Mackens W, Voss H. Computing the minimum eigenvalue of a symmetric positive definite Toeplitz matrix by Newton type methods. *SIAM J. Sci. Comput.* 2000: **21**: 1650–1656

[11] Mastronardi N, Boley D. Computing the smallest eigenpair of a symmetric positive definite Toeplitz matrix. *SIAM J. Sci. Comput.* 1999; **20**: 1921–1927

[12] Melman A. Extreme eigenvalues of real symmetric Toeplitz matrices. *Math. Comp.* 2000; **70**: 649–669

[13] Pisarenko VF. The retrieval of harmonics from a covariance function. *Geophys. J. R. astr. Soc.* 1973; **33**: 347–366

[14] Stoer J, Bulirsch R. *Introduction to Numerical Analysis*. Springer: New York 1980

[15] Voss H. Symmetric schemes for computing the minimum eigenvalue of a symmetric Toeplitz matrix. *Lin. Alg. Appl.* 1999; **287**: 359–371

[16] Voss H. A symmetry exploiting Lanczos method for symmetric Toeplitz matrices. *Numerical Algorithms* 2000; **25**: 377–385

[17] Voss H. A variant of the inverted Lanczos method. *BIT Numerical Analysis* 2001; **41**: 1111–1120