# Bounds for the Minimum Eigenvalue of a Symmetric Toeplitz Matrix

Heinrich Voss

*Technical University Hamburg–Harburg, Section of Mathematics, D–21071 Hamburg, Federal Republic of Germany, e-mail: voss@tu-harburg.de*

**Abstract**

In a recent paper Melman [12] derived upper bounds for the smallest eigenvalue of a real symmetric Toeplitz matrix in terms of the smallest roots of rational and polynomial approximations of the secular equation $f(\lambda) = 0$, the best of which being constructed by the $(1,2)$-Padé approximation of $f$. In this paper we prove that this bound is the smallest eigenvalue of the projection of the given eigenvalue problem onto a Krylov space of $T_n^{-1}$ of dimension 3. This interpretation of the bound suggests enhanced bounds of increasing accuracy. They can be substantially improved further by exploiting symmetry properties of the principal eigenvector of $T_n$.

*Keywords.* Toeplitz matrix, eigenvalue problem, symmetry

# 1 Introduction

The problem of finding the smallest eigenvalue of a real symmetric, positive definite Toeplitz matrix (RSPDT) is of considerable interest in signal processing. Given the covariance sequence of the observed data, Pisarenko [14] suggested a method which determines the sinusoidal frequencies from the eigenvector of the covariance matrix associated with its minimum eigenvalue.

The computation of the minimum eigenvalue $\lambda_1$ of an RSPDT $T_n$ was considered in, e.g. [2], [7], [8], [9], [10], [11], [13], [16]. Cybenko and Van Loan [2] presented an algorithm which is a combination of bisection and Newton's method for the secular equation. By replacing Newton's method with a root finding method based on rational Hermitian interpolation of the secular equation, Mackens and the present author in [10] improved this approach substantially. In [11] it was shown that the algorithm from [10] is equivalent to a projection method where in every step the eigenvalue problem is projected onto a two-dimensional space. This interpretation suggested a further enhancement to the method of Cybenko and Van Loan. Finally, by exploiting symmetry properties of the principal eigenvector, the methods in [10] and [11] were accelerated in [16].

If the bisection scheme in a method of the last paragraph is started with a poor upper bound for $\lambda_1$, a large number of bisection steps may be necessary to get a suitable initial value for the subsequent root finding method. Usually the dominant share of the cost occurs in the bisection phase, and a good upper bound for $\lambda_1$ is of predominant importance. Cybenko and Van Loan [2] presented an upper bound for $\lambda_1$ which can be obtained from the data determined in Durbin's algorithm for the Yule-Walker system. Dembo [3] derived tighter bounds by using (linear and quadratic) Taylor expansions of the secular equation. In a recent paper Melman [12] improved these bounds in two ways, first by considering rational approximations of the secular equation and, secondly, by exploiting symmetry properties of the principal eigenvector in a similar way as in [16]. Apparently, because of the somewhat complicated nature of their analysis, he restricted his investigations to rational approximations of at most third order.

In this paper we prove that Melman's bounds obtained by first and third order rational approximations can be interpreted as the smallest eigenvalues of projected problems of dimension 2 and 3, respectively, where the matrix $T_n$ is projected onto a Krylov space of $T_n^{-1}$. This interpretation again proves the fact that the smallest roots of the approximating rational functions are upper bounds of the smallest eigenvalue, avoiding the somewhat complicated analysis of the rational functions. Moreover, it suggests a method to obtain improved bounds in a systematic way by increasing the dimension of the Krylov space.

The paper is organized as follows. In Section 2 we briefly sketch the approaches of Dembo and Melman and prove that Melman's bounds can be obtained from a projected eigenproblem. In Section 3 we consider secular equations characterizing the smallest odd and even eigenvalue of $T_n$ and take advantage of symmetry properties of the principal eigenvector to improve the eigenvalue bounds. Finally, in Section 4 we present numerical results.

# 2 Rational approximation and projection

Let

$$T_n = (t_{|i-j|})_{i,j=1,\ldots,n} \in \mathbb{R}^{(n,n)}$$

be a real and symmetric Toeplitz matrix. We denote by $T_j \in \mathbb{R}^{(j,j)}$ its $j$-th principal submatrix, and by $t$ the vector $t = (t_1, \ldots, t_{n-1})^T$. If $\lambda_1^{(j)} \le \lambda_2^{(j)} \le \ldots \le \lambda_j^{(j)}$ are the eigenvalues of $T_j$ then the interlacing property $\lambda_{j-1}^{(k)} \le \lambda_{j-1}^{(k-1)} \le \lambda_j^{(k)}$, $2 \le j \le k \le n$, holds.

We briefly sketch the approaches of Dembo and Melman. To this end we additionally assume that $T_n$ is positive definite. If $\lambda$ is not in the spectrum of $T_{n-1}$ then block Gauss elimination of the variables $x_2, \dots, x_n$ of the system

$$\begin{pmatrix} t_0 - \lambda & t^T \\ t & T_{n-1} - \lambda I \end{pmatrix} x = 0$$

that characterizes the eigenvalues of $T_n$ yields

$$(t_0 - \lambda - t^T (T_{n-1} - \lambda I)^{-1} t) x_1 = 0.$$

We assume that $\lambda_1^{(n)} < \lambda_1^{(n-1)}$. Then $x_1 \neq 0$, and $\lambda_1^{(n)}$ is the smallest positive root of the secular equation

$$f(\lambda) := -t_0 + \lambda + t^T (T_{n-1} - \lambda I)^{-1} t = 0 \tag{1}$$

which may be rewritten in modal coordinates as

$$f(\lambda) = -t_0 + \lambda + \sum_{j=1}^{n-1} \frac{(t^T v^j)^2}{\lambda_j^{(n-1)} - \lambda} = 0. \tag{2}$$

where $v^j$ denotes the eigenvector of $T_{n-1}$ corresponding to $\lambda_j^{(n-1)}$

From

$$f(0) = -t_0 + t^T T_{n-1}^{-1} t = -(1, -t^T T_{n-1}^{-1}) \left( \begin{matrix} t_0 & t^T \\ t & T_{n-1} \end{matrix} ) \right) \begin{pmatrix} 1 \\ -T_{n-1}^{-1} t \end{pmatrix} < 0$$

and $f^{(j)}(\lambda) > 0$ for every $j \in \mathbb{N}$ and every $\lambda \in [0, \lambda_1^{(n-1)}]$ it follows that the Taylor polynomial $p_j$ of degree $j$ such that $f^{(k)}(0) = p_j^{(k)}(0)$, $k = 0, 1, \dots, j$, satisfies

$$f(\lambda) \geq p_j(\lambda) \text{ for every } \lambda < \lambda_1^{(n-1)} \quad \text{and} \quad p_j(\lambda) \leq p_{j+1}(\lambda) \text{ for every } \lambda \geq 0.$$

Hence, the smallest positive root $\mu_j$ of $p_j$ is an upper bound of $\lambda_1^{(n)}$ and $\mu_{j+1} \leq \mu_j$. For $j = 1$ and $j = 2$ these upper bounds were presented by Dembo [3], for $j = 3$ it is contained in Melman [12].

Improved bounds were obtained by Melman [12] by approximating the secular equation by rational functions. The idea of a rational approximation of the secular equation is not new. Dongarra and Sorensen [4] used it in a parallel divide and conquer method for symmetric eigenvalue problems, while in [10] it was used in an algorithm for computing the smallest eigenvalue of a Toeplitz matrix.

Melman considered rational approximations

$$r_j(\lambda) = -t_0 + \lambda + \rho_j(\lambda)$$

of $f$ where

$$\rho_1(\lambda) := \frac{a}{b - \lambda}, \quad \rho_2(\lambda) := a + \frac{b}{c - \lambda}, \quad \rho_3(\lambda) := \frac{a}{b - \lambda} + \frac{c}{d - \lambda},$$

and the parameters $a, b, c, d$ are determined such that

$$\rho_j^{(k)}(0) = \frac{d^k}{d\lambda^k} t^T (T_{n-1} - \lambda I)^{-1} t\Big|_{\lambda=0} = k!\, t^T T_{n-1}^{-(k+1)} t,\ k = 0, 1, \ldots, j. \tag{3}$$

Thus $\rho_1$, $\rho_2$ and $\rho_3$, respectively, are the $(0,1)$-, $(1,1)$- and $(1,2)$-Padé approximations of $\phi(\lambda) := t^T(T_{n-1} - \lambda I)^{-1} t$ (cf. Braess [1])

For the rational approximations $r_j$ it holds that (cf. Melman [12], Theorem 4.1)

$$r_1(\lambda) \le r_2(\lambda) \le r_3(\lambda) \le f(\lambda) \quad \text{for } \lambda < \lambda_1^{(n-1)},$$

and with the arguments from Melman one can infer that for $j = 2$ and $j = 3$ the inequality $r_{j-1}(\lambda) \le r_j(\lambda)$ even holds for every $\lambda$ less than the smallest pole of $r_j$. Hence, if $\mu_j$ denotes the smallest positive root of $r_j(\lambda) = 0$ then

$$\lambda_1^{(n)} \le \mu_3 \le \mu_2 \le \mu_1.$$

The rational approximations $r_1(\lambda)$ and $r_3(\lambda)$ to $f(\lambda)$ are of the form of a secular equation of an eigenvalue problem of dimensions 2 and 3, respectively. Hence, there is some evidence that the roots of $r_1$ and $r_3$ are eigenvalues of projected eigenproblems. In the following we prove that this conjecture actually holds true. Notice that our approach does not presume that the matrix $T_n$ is positive definite.

**Lemma 2.1** *Let $T_n$ be a real symmetric Toeplitz matrix such that 0 is not in the spectrum of $T_n$ and $T_{n-1}$. Let $e^1 := (1, 0, \ldots, 0)^T \in \mathbb{R}^n$, and denote by $V_\ell := span\{e^1, T_n^{-1}e^1, \ldots, T_n^{-\ell}e^1\}$ the Krylov space of $T_n^{-1}$ corresponding to the initial vector $e^1$. Then*

$$\left\{ e^1, \begin{pmatrix} 0 \\ T_{n-1}^{-1} t \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ T_{n-1}^{-\ell} t \end{pmatrix} \right\} \tag{4}$$

*is a basis of $V_\ell$, and the projected eigenproblem of $T_n x = \lambda x$ onto $V_\ell$ can be written as*

$$\tilde{B} y := \begin{pmatrix} t_0 & s^T \\ s & B \end{pmatrix} y = \lambda \begin{pmatrix} 1 & 0^T \\ 0 & C \end{pmatrix} y =: \tilde{C} y \tag{5}$$

*where*

$$B = \begin{pmatrix} \mu_1 & \cdots & \mu_\ell \\ \vdots & \cdots & \vdots \\ \mu_\ell & \cdots & \mu_{2\ell-1} \end{pmatrix},\ C = \begin{pmatrix} \mu_2 & \cdots & \mu_{\ell+1} \\ \vdots & \cdots & \vdots \\ \mu_{\ell+1} & \cdots & \mu_{2\ell} \end{pmatrix},\ s = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_\ell \end{pmatrix}$$

*and*

$$\mu_j = t^T T_{n-1}^{-j} t. \tag{6}$$

**Proof** For $\ell = 0$ the Lemma is trivial. Since

$$T_n^{-1} e^1 = \begin{pmatrix} \alpha \\ v \end{pmatrix} \iff \begin{cases} \alpha t_0 + t^T v = 1 \\ \alpha t + T_{n-1} v = 0 \end{cases}$$

for $\ell = 1$ a basis of $V_1$ is given in (4).

Assume that (4) defines a basis of $V_\ell$ for some $\ell \in \mathbb{N}$, then $T_n^{-\ell} e^1$ may be represented as

$$T_n^{-\ell} e^1 = \begin{pmatrix} \beta \\ T_{n-1}^{-1} z \end{pmatrix}, \quad z = \sum_{j=0}^{\ell-1} \gamma_j T_{n-1}^{-j} t.$$

Hence

$$T_n^{-\ell-1} e^1 = T_n^{-1} \begin{pmatrix} \beta \\ T_{n-1}^{-1} z \end{pmatrix} = \beta T_n^{-1} e^1 + T_n^{-1} \begin{pmatrix} 0 \\ T_{n-1}^{-1} z \end{pmatrix} =: \beta T_n^{-1} e^1 + \begin{pmatrix} \delta \\ w \end{pmatrix},$$

where

$$\begin{pmatrix} t_0 & t^T \\ t & T_{n-1} \end{pmatrix} \begin{pmatrix} \delta \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ T_{n-1}^{-1} z \end{pmatrix} \iff \left\{ \begin{array}{c} \delta t_0 + t^T w = 0 \\ \delta t + T_{n-1} w = T_{n-1}^{-1} z \end{array} \right.$$

The second equation is equivalent to

$$w = T_{n-1}^{-2} z - \delta T_{n-1}^{-1} t = \sum_{j=0}^{\ell-1} \gamma_j T_{n-1}^{-j-2} t - \delta T_{n-1}^{-1} t \in \ \mathrm{span}\{T_{n-1}^{-1} t, \ldots, T_{n-1}^{-\ell-1} t\},$$

and (4) defines a basis of $V_{\ell+1}$ for $\ell + 1$.

Using the basis of $V_\ell$ in (4) it is easily seen that eq. (5) is the matrix representation of the projection of the eigenvalue problem $T_n x = \lambda x$ onto the Krylov space $V_\ell$. □

**Lemma 2.2** *Let $B, C, s, \tilde{B}$ and $\tilde{C}$ be defined as in Lemma 2.1. Then the eigenvalues of the projected problem $\tilde{B} y = \lambda \tilde{C} y$ which are not in the spectrum of the subpencil $Bw = \lambda C w$ are the roots of the secular equation*

$$g_\ell(\lambda) := -t_0 + \lambda + s^T (B - \lambda C)^{-1} s. \tag{7}$$

*For $F := (T_{n-1}^{-1} t, \ldots, T_{n-1}^{-\ell} t)$ the secular equation can be rewritten as*

$$g_\ell(\lambda) = -t_0 + \lambda + t^T F (F^T (T_{n-1} - \lambda I) F)^{-1} F^T t. \tag{8}$$

**Proof**: The secular equation in (7) is obtained in the same way as the secular equation $f(\lambda) = 0$ of $T_n x = \lambda x$ at the beginning of this section by block Gauss elimination. The representation (8) is obtained from $B = F^T T_{n-1} F$, $C = F^T F$ and $s = F^T t$. □

**Lemma 2.3** *Let $B, C, s$ be defined in Lemma 2.1, and let*

$$\sigma_\ell(\lambda) = s^T (B - \lambda C)^{-1} s.$$

*Then the $k$-th derivative of $\sigma_\ell$ is given by*

$$\sigma_\ell^{(k)}(\lambda) = k!\, t^T (F (F^T (T_{n-1} - \lambda I) F)^{-1} F^T)^{k+1} t, \ k \geq 0. \tag{9}$$

**Proof**: Let
$$G(\lambda) := (F^T(T_{n-1} - \lambda I)F)^{-1}.$$
Then
$$\frac{d}{d\lambda}G(\lambda) = G(\lambda)F^TFG(\lambda), \tag{10}$$
yields
$$\begin{aligned}
\sigma_\ell'(\lambda) &= t^TFG'(\lambda)F^Tt \\
&= t^TF(F^T(T_{n-1} - \lambda I)F)^{-1}F^TF(F^T(T_{n-1} - \lambda I)F)^{-1}F^Tt \\
&= t^T(F(F^T(T_{n-1} - \lambda I)F)^{-1}F^T)^2t,
\end{aligned}$$
i.e. eq. (9) for $k = 1$.

Assume that eq. (9) holds for some $k \in \mathbb{N}$. Then it follows from eq. (10)
$$\begin{aligned}
\sigma_\ell^{(k+1)}(\lambda) &= k!\, t^T\frac{d}{d\lambda}\{(F(F^T(T_{n-1} - \lambda I)F)^{-1}F^T)^{k+1}\}t \\
&= (k+1)!\, t^T(F(F^T(T_{n-1} - \lambda I)F)^{-1}F^T)^k\frac{d}{d\lambda}(F(F^T(T_{n-1} - \lambda I)F)^{-1}F^T)t \\
&= (k+1)!\, t^T(F(F^T(T_{n-1} - \lambda I)F)^{-1}F^T)^kF\frac{d}{d\lambda}G(\lambda)F^Tt \\
&= (k+1)!\, t^T(F(F^T(T_{n-1} - \lambda I)F)^{-1}F^T)^kFG(\lambda)F^TFG(\lambda)F^Tt \\
&= (k+1)!\, t^T(F(F^T(T_{n-1} - \lambda I)F)^{-1}F^T)^{k+2}t,
\end{aligned}$$
which completes the proof. □

**Lemma 2.4** *Let $F := (T_{n-1}^{-1}t, \ldots, T_{n-1}^{-\ell}t)$. Then it holds that*
$$(F(F^TT_{n-1}F)^{-1}F^T)^kt = T_{n-1}^{-k}t \quad \text{for } k = 0, 1, \ldots, \ell, \tag{11}$$
*and*
$$t^T(F(F^TT_{n-1}F)^{-1}F^T)^kt = t^TT_{n-1}^{-k}t \quad \text{for } k = 0, 1, \ldots, 2\ell. \tag{12}$$

**Proof** For $k = 0$ the statement (11) is trivial. Let
$$H := F(F^TT_{n-1}F)^{-1}FT_{n-1}.$$
Then for every $x \in \operatorname{span} F$, $x := Fy$, $y \in \mathbb{R}^\ell$
$$Hx = F(F^TT_{n-1}F)^{-1}F^TT_{n-1}Fy = Fy = x,$$
and $T_{n-1}^{-1}t \in \operatorname{span} F$ yields
$$F(F^TT_{n-1}F)^{-1}F^Tt = HT_{n-1}^{-1}t = T_{n-1}^{-1}t,$$
i.e. eq. (11) for $k = 1$.

If eq. (11) holds for some $k < \ell$ then it follows from $T_{n-1}^{-(k+1)} t \in \operatorname{span} F$

$$\begin{aligned}
(F(F^T T_{n-1} F)^{-1} F^T)^{k+1} t &= (F(F^T T_{n-1} F)^{-1} F^T)(F(F^T T_{n-1} F)^{-1} F^T)^k t \\
&= (F(F^T T_{n-1} F)^{-1} F^T) T_{n-1}^{-k} t \\
&= (F(F^T T_{n-1} F)^{-1} F^T) T_{n-1} T_{n-1}^{-(k+1)} t \\
&= H T_{n-1}^{-(k+1)} t \;=\; T_{n-1}^{-(k+1)} t
\end{aligned}$$

which proves eq. (11).

Eq. (12) follows immediately from eq. (11) for $k = 0, 1, \ldots, \ell$. For $\ell < k \leq 2\ell$ it is obtained from

$$\begin{aligned}
t^T (F(F^T T_{n-1} F)^{-1} F^T)^k t &= ((F(F^T T_{n-1} F)^{-1} F^T)^\ell t)^T ((F(F^T T_{n-1} F)^{-1} F^T)^{k-\ell} t) \\
&= (T_{n-1}^{-\ell} t)^T (T_{n-1}^{-(k-\ell} t) \;=\; t^T T^{-k} t.
\end{aligned}$$

□

We are now ready to prove our main result.

**Theorem 2.5**: *Let $T_n$ be a real symmetric Toeplitz matrix such that $T_n$ and $T_{n-1}$ are nonsingular. Let the matrices $B$ and $C$ be defined in Lemma 2.1, and let*

$$g_\ell(\lambda) = -t_0 + \lambda + s^T (B - \lambda C)^{-1} s =: -t_0 + \lambda + \sigma_\ell(\lambda)$$

*be the secular equation of the projected eigenproblem (5) considered in Lemma 2.1. Then $\sigma_\ell(\lambda)$ is the $(\ell - 1, \ell)$-Padé approximation of the rational function*

$$\phi(\lambda) = t^T (T_{n-1} - \lambda I)^{-1} t.$$

*Conversely, if $\tau_\ell(\lambda)$ denotes the $(\ell - 1, \ell)$-Padé approximation of $\phi(\lambda)$ and $\mu_1^{(\ell)} \leq \mu_2^{(\ell)} \leq \ldots$ are the roots of the rational function $\lambda \mapsto -t_0 + \lambda + \tau_\ell(\lambda)$ ordered by magnitude, then*

$$\lambda_j^{(n)} \leq \mu_j^{(\ell+1)} \leq \mu_j^{(\ell)} \tag{13}$$

*for every $\ell < n$ and $j \in \{1, \ldots, \ell + 1\}$.*

**Proof**: Using modal coordinates of the pencil $Bw = \lambda Cw$ the rational function $\sigma_\ell(\lambda)$ may be rewritten as

$$\sigma_\ell(\lambda) = \sum_{j=1}^{\ell} \frac{\beta_j^2}{\kappa_j - \lambda}$$

where $\kappa_j$ denotes the eigenvalues of this pencil. Hence $\sigma_\ell$ is a rational function where the degree of the numerator and denominator is not greater than $\ell - 1$ and $\ell$, respectively.

From Lemma 2.3 and Lemma 2.4 it follows that

$$\sigma_\ell^{(k)}(0) = k!\, t^T (F(F^T T_{n-1} F)^{-1} F^T)^{k+1} t = k!\, t^T T_{n-1}^{-(k+1)} t = \phi^{(k)}(0)$$

for every $k = 0, 1, \ldots, 2\ell - 1$. Hence $\sigma_\ell$ is the $(\ell - 1, \ell)$-Padé approximation of $\phi$.

From the uniqueness of the Padé approximation it follows that $\tau_\ell = \sigma_\ell$. Hence $\mu_1^{(\ell)} \le \mu_2^{(\ell)} \le \dots$ are the eigenvalues of the projection of problem $T_n x = \lambda x$ onto $V_\ell$, and (13) follows from the minimax principle. □

Some remarks are in order:

1. The rational functions $\rho_1$ and $\rho_3$ constructed by Melman [12] coincide with $\sigma_1$ and $\sigma_2$, respectively. Hence, Theorem 2.5 contains the bounds of Melman. Moreover it provides a method to compute these bounds which is much more transparent than the approach of Melman.

2. Obviously the considerations above apply to every shifted problem $T_n - \kappa I$ such that $\kappa$ is not in the spectra of $T_n$ and $T_{n-1}$. Notice that the analysis of Melman [12] is only valid if $\kappa$ is a lower bound of $\lambda_1(T_n)$.

3. In the same way lower bounds of the maximum eigenvalue of $T_n$ can be determined. These generalize the corresponding results by Melman [12] where we do not need an upper bound of the largest eigenvalue of $T_n$.

# 3 Exploiting symmetry of the principal eigenvector

If $T_n \in \mathbb{R}^{(n,n)}$ is a real and symmetric Toeplitz matrix and $E_n$ denotes the $n$-dimensional flipmatrix with ones in its secondary diagonal and zeros elsewhere, then $E_n^2 = I$ and $T_n = E_n T_n E_n$. Hence $T_n x = \lambda x$ if and only if

$$T_n(E_n x) = E_n T_n E_n^2 x = \lambda E_n x,$$

and $x$ is an eigenvector of $T_n$ if and only if $E_n x$ is. If $\lambda$ is a simple eigenvalue of $T_n$ then from $\|x\|_2 = \|E_n x\|_2$ we obtain $x = E_n x$ or $x = -E_n x$. We say that an eigenvector $x$ is symmetric and the corresponding eigenvalue $\lambda$ is even if $x = E_n x$, and $x$ is called skew-symmetric and $\lambda$ is odd if $x = -E_n x$.

One disadvantage of the projection scheme in Section 2 is that it does not reflect the symmetry properties of the principal eigenvector. In this section we present a variant which takes advantage of the symmetry of the eigenvector and which essentially is of equal cost to the method considered in Section 2.

To take into account the symmetry properties of the eigenvector we eliminate the variables $x_2, \dots, x_{n-1}$ from the system

$$\begin{pmatrix} t_0 - \lambda & \tilde{t}^T & t_{n-1} \\ \tilde{t} & T_{n-2} - \lambda I & E_{n-2}\tilde{t} \\ t_{n-1} & \tilde{t}^T E_{n-2} & t_0 - \lambda \end{pmatrix} x = 0 \tag{14}$$

where $\tilde{t} = (t_1, \dots, t_{n-2})^T$.

Then every eigenvalue $\lambda$ of $T_n$ which is not in the spectrum of $T_{n-2}$ is an eigenvalue of the two-dimensional nonlinear eigenvalue problem

$$\begin{pmatrix} t_0 - \lambda - \tilde{t}^T(T_{n-2} - \lambda I)^{-1}\tilde{t} & t_{n-1} - \tilde{t}^T(T_{n-2} - \lambda I)^{-1}E_{n-2}\tilde{t} \\ t_{n-1} - \tilde{t}^T E_{n-2}(T_{n-2} - \lambda I)^{-1}\tilde{t} & t_0 - \lambda - \tilde{t}^T(T_{n-2} - \lambda I)^{-1}\tilde{t} \end{pmatrix} \begin{pmatrix} x_1 \\ x_n \end{pmatrix} = 0. \tag{15}$$

Moreover, if $\lambda$ is an even eigenvalue of $T_n$, then $(1,1)^T$ is the corresponding eigenvector of problem (15), and if $\lambda$ is an odd eigenvalue of $T_n$ then $(1,-1)^T$ is the corresponding eigenvector of system (15).

Hence, if the smallest eigenvalue $\lambda_1^{(n)}$ is even, then it is the smallest root of the rational function

$$f_+(\lambda) := -t_0 - t_{n-1} + \lambda + \tilde{t}^T(T_{n-2} - \lambda I)^{-1}(\tilde{t} + E_{n-2}\tilde{t}), \tag{16}$$

and if $\lambda_1^{(n)}$ is an odd eigenvalue of $T_n$ then it is the smallest root of

$$f_-(\lambda) := -t_0 + t_{n-1} + \lambda + \tilde{t}^T(T_{n-2} - \lambda I)^{-1}(\tilde{t} - E_{n-2}\tilde{t}). \tag{17}$$

Analogously to the proofs given in Section 2, we obtain the following results for the odd and even secular equations.

**Theorem 3.1** *Let $T_n$ be a real symmetric Toeplitz matrix such that 0 is not in the spectrum of $T_n$ and of $T_{n-2}$. Let $t_\pm := \tilde{t} \pm E_{n-2}\tilde{t}$, and let*

$$V_{\ell\pm} := \mathit{span}\left\{e_\pm, T_n^{-1}e_\pm, \dots, T_n^{-\ell}e_\pm\right\}$$

*be the Krylov space of $T_n^{-1}$ corresponding to the initial vector $e_\pm := (1, \dots, \pm 1)^T$. Then*

$$\left\{e_\pm, \begin{pmatrix} 0 \\ T_{n-2}^{-1}t_\pm \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ T_{n-2}^{-\ell}t_\pm \\ 0 \end{pmatrix}\right\}$$

*is a basis of $V_{\ell\pm}$.*

*The projection of the eigenproblem $T_n x = \lambda x$ onto $V_{\ell\pm}$ can be written as*

$$\tilde{B}_\pm y := \begin{pmatrix} t_0 \pm t_{n-1} & s_\pm^T \\ s_\pm & B_\pm \end{pmatrix} y = \lambda \begin{pmatrix} 1 & 0^T \\ 0 & C_\pm \end{pmatrix} y =: \tilde{C}_\pm y \tag{18}$$

*where*

$$B_\pm = \begin{pmatrix} \nu_1^\pm & \dots & \nu_\ell^\pm \\ \vdots & \dots & \vdots \\ \nu_\ell^\pm & \dots & \nu_{2\ell-1}^\pm \end{pmatrix}, \; C_\pm = \begin{pmatrix} \nu_2^\pm & \dots & \nu_{\ell+1}^\pm \\ \vdots & \dots & \vdots \\ \nu_{\ell+1}^\pm & \dots & \nu_{2\ell}^\pm \end{pmatrix}, \; s_\pm = \begin{pmatrix} \nu_1^\pm \\ \vdots \\ \nu_\ell^\pm \end{pmatrix} \tag{19}$$

*and*

$$\nu_j^\pm = 0.5\tilde{t}_\pm^T T_{n-2}^{-j}\tilde{t}_\pm = (\tilde{t} \pm E_{n-2}\tilde{t})^T T_{n-2}^{-j}\tilde{t}. \tag{20}$$

*The eigenvalues of the projected problem (18) which are not in the spectrum of the subpencil $B_\pm w = \lambda C_\pm w$ are the roots of the secular equation*

$$g_\pm(\lambda) = -t_0 \mp t_{n-1} + \lambda + s_\pm^T(B_\pm - \lambda C_\pm)^{-1}s_\pm =: -t_0 \pm t_{n-1} + \lambda + \sigma_{\ell\pm}(\lambda) = 0. \quad (21)$$

*Here, $\sigma_{\ell\pm}(\lambda)$ is the $(\ell-1,\ell)$-Padé approximation of the rational function*

$$\phi_\pm(\lambda) := \tilde{t}^T(T_{n-2} - \lambda I)^{-1}(\tilde{t} \pm E_{n-2}\tilde{t}).$$

*Conversely, if $\tau_{\ell\pm}(\lambda)$ denotes the $(\ell-1,\ell)$-Padé approximation of $\phi_\pm(\lambda)$ and $\mu_{1\pm}^{(\ell)}$ is the smallest root of the rational function*

$$\lambda \mapsto t_0 \pm t_{n-1} - \lambda + \tau_{\ell\pm}(\lambda) = 0,$$

*then*

$$\lambda_1^{(n)} \le \min(\mu_{1+}^{(\ell+1)}, \mu_{1-}^{(\ell+1)}) \le \min(\mu_{1+}^{(\ell)}, \mu_{1-}^{(\ell)}).$$

As in the prvious section, for $\ell = 1$ and $\ell = 2$ Theorem 3.1 contains the bounds which were already presented by Melman [12] using rational approximations of the even and odd secular equations (16) and (17).

# 4 Numerical results

To establish the projected eigenvalue problem (7) one has to compute expressions of the form

$$\mu_j = t^T T_{n-1}^{-j} t, \quad j = 1, \dots, 2\ell.$$

For $\ell = 1$ the quantities $\mu_1$ and $\mu_2$ are obtained from the solution $z^1$ of the Yule-Walker system $T_{n-1}z^1 = -t$ which can be solved efficiently by Durbin's algorithm (cf. [6], p. 195) requiring $2n^2$ flops. Once $z^1$ is known $\mu_1 = t^T z^1$ and $\mu_2 = \|z^1\|_2^2$.

To increase the dimension of the projected problem by one we have to solve the linear system

$$T_{n-1}z^{\ell+1} = z^\ell, \quad (22)$$

and we have to compute two scalar products $\mu_{2\ell+1} = (z^{\ell+1})^T z^\ell$ and $\mu_{2\ell+2} = \|z^{\ell+1}\|_2^2$.

System (22) can be solved efficiently in one of the following two ways. Durbin's algorithm for the Yule-Walker system supplies a decomposition $LT_{n-1}L^T = D$ where $L$ is a lower triangular matrix (with ones in its diagonal) and $D$ is a diagonal matrix. Hence, for every $\ell$ the solution of eq. (22) requires $2n^2$ flops. This method for (22) is called Levinson-Durbin algorithm.

For large dimensions $n$ eq. (22) can be solved using the Gohberg-Semencul formula for the inverse $T_{n-1}^{-1}$ (cf. [5])

$$T_{n-1}^{-1} = \frac{1}{1 - y^T t(1:n-2)}(GG^T - HH^T) \quad (23)$$

where

$$G := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ y_1 & 1 & 0 & \dots & 0 \\ y_2 & y_1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_{n-2} & y_{n-3} & y_{n-4} & \dots & 1 \end{pmatrix} \text{ and } H := \begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ y_{n-2} & 0 & 0 & \dots & 0 \\ y_{n-3} & y_{n-2} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_1 & y_2 & y_3 & \dots & 0 \end{pmatrix}$$

are Toeplitz matrices and $y$ denotes the solution of the Yule-Walker system $T_{n-2}y = t(1:n-2)$.

The advantages associated with eq. (23) are at hand. Firstly, the representation of the inverse of $T_{n-1}$ requires only $n$ storage elements. Secondly, the matrices $G$, $G^T$, $H$ and $H^T$ are Toeplitz matrices, and hence the solution $T_{n-1}z^\ell$ can be calculated in only $O(n \log n)$ flops using fast Fourier transform. Experiments show that when $n \geq 512$ this approach is actually more efficient than the Levinson-Durbin algorithm.

In the method of Section 3 we also have to solve a Yule-Walker system $T_{n-2}z^1 = \tilde{t}$ by Durbin's algorithm, and increasing the dimension of the projected problem by one we have to solve one general system $T_{n-2}z^{\ell+1} = z^\ell$ using the Levinson-Durbin algorithm or the Gohberg-Semencul formula. Moreover, two vector additions $z^{\ell+1} \pm E_{n-2}z^{\ell+1}$ and 4 scalar products have to be determined, and 2 eigenvalue problems of very small dimensions have to be solved. To summarize, again $2n^2 + O(n)$ flops are required to increase the dimension of the projected problem by one.

If the gap between the smallest eigenvalue $\lambda_1^{(n)}$ and the second eigenvalue $\lambda_2^{(n)}$ is large, the sequence of vectors $\begin{pmatrix} 1 \\ z^\ell \end{pmatrix}$ converges very fast to the principal eigenvector of $T_n$ and the matrix $C$ becomes nearly singular. In three of 600 examples that we considered the matrix $C$ even became (numerically) indefinite. However, in all of these examples the relative error of the eigenvalue approximation of the previous step was already $10^{-8}$. In a forthcoming paper we will discuss a stable version of the projection methods in Sections 2 and 3.

**Example** To test the bounds we considered the following class of Toeplitz matrices

$$T = m \sum_{k=1}^{n} \eta_k T_{2\pi\theta_k} \tag{24}$$

where $m$ is chosen such that the diagonal of $T$ is normalized to $t_0 = 1$,

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

and $\eta_k$ and $\theta_k$ are uniformly distributed random numbers in the interval $[0,1]$ (cf. Cybenko and Van Loan [2]).

Table 1 contains the average of the relative errors of the bounds of Section 2 in 100 test problems for each of the dimensions $n = 32$, 64, 128, 256, 512 and 1024. Table 2 shows the corresponding results for the bounds of Section 3. In both tables

TABLE 1. Average of relative errors; bounds of Section 2

| dim | $\ell = 1$ | $\ell = 2$ | $\ell = 3$ | $\ell = 4$ |
|---|---|---|---|---|
| 32 | $1.05\ E+0$ | $4.29\ E-2$ | $8.38\ E-3$ | $1.82\ E-3$ |
| 64 | $1.64\ E+0$ | $6.41\ E-2$ | $1.38\ E-2$ | $4.20\ E-3$ |
| 128 | $2.76\ E+0$ | $7.60\ E-2$ | $1.88\ E-2$ | $5.09\ E-3$ |
| 256 | $5.03\ E+0$ | $9.25\ E-2$ | $1.78\ E-2$ | $6.20\ E-3$ |
| 512 | $7.51\ E+0$ | $1.10\ E-1$ | $2.47\ E-2$ | $6.80\ E-3$ |
| 1024 | $1.65\ E+1$ | $1.05\ E-1$ | $2.43\ E-2$ | $6.60\ E-3$ |

TABLE 2. Average of relative errors; bounds of Section 3

| Dimension | $\ell = 1$ | $\ell = 2$ | $\ell = 3$ | $\ell = 4$ |
|---|---|---|---|---|
| 32 | $5.18\ E-1$ | $8.33\ E-3$ | $8.54\ E-4$ | $3.20\ E-5$ |
| 64 | $9.39\ E-1$ | $2.30\ E-2$ | $1.25\ E-3$ | $3.65\ E-4$ |
| 128 | $1.79\ E+0$ | $2.40\ E-2$ | $1.61\ E-3$ | $6.41\ E-5$ |
| 256 | $3.27\ E+0$ | $4.25\ E-2$ | $4.58\ E-3$ | $7.15\ E-4$ |
| 512 | $5.11\ E+0$ | $5.43\ E-2$ | $4.19\ E-3$ | $8.77\ E-4$ |
| 1024 | $1.11\ E+1$ | $5.45\ E-2$ | $4.81\ E-3$ | $7.42\ E-4$ |

TABLE 3. Average of common logarithm of relative errors; bounds of Section 2

| dim | $\ell = 1$ | $\ell = 2$ | $\ell = 3$ | $\ell = 4$ |
|---|---|---|---|---|
| 32 | −0.10 (0.32) | −1.93 (0.98) | −3.76 (2.50) | −6.31 (3.80) |
| 64 | 0.14 (0.26) | −1.65 (0.89) | −3.15 (2.09) | −5.26 (3.43) |
| 128 | 0.40 (0.19) | −1.63 (1.03) | −3.14 (2.62) | −5.15 (3.65) |
| 256 | 0.64 (0.22) | −1.41 (0.84) | −2.64 (1.81) | −4.42 (3.19) |
| 512 | 0.81 (0.27) | −1.35 (0.77) | −2.23 (1.09) | −3.74 (2.28) |
| 1024 | 1.14 (0.24) | −1.37 (0.76) | −2.37 (1.40) | −3.93 (2.71) |

TABLE 4. Average of common logarithm of relative errors; bounds of Section 3

| dim | $\ell = 1$ | $\ell = 2$ | $\ell = 3$ | $\ell = 4$ |
|---|---|---|---|---|
| 32 | −0.45 (0.37) | −3.38 (1.76) | −6.88 (3.65) | −10.28 (4.02) |
| 64 | −1.29 (0.31) | −2.72 (1.46) | −5.88 (2.93) | −9.27 (3.85) |
| 128 | 0.17 (0.23) | −2.79 (1.80) | −5.84 (3.37) | −9.27 (3.93) |
| 256 | 0.44 (0.24) | −2.32 (1.40) | −5.17 (2.96) | −8.42 (4.11) |
| 512 | 0.65 (0.27) | −2.15 (1.35) | −4.88 (2.93) | −8.10 (4.26) |
| 1024 | 0.97 (0.23) | −2.15 (1.41) | −5.03 (3.05) | −8.12 (4.47) |

the first two columns contain the relative errors of the bounds given by Melman. The experiments clearly show that exploiting symmetry of the principal eigenvector leads to significant improvements of the bounds.

The mean values of the relative errors do not reflect the quality of the bounds. Large bounds are taken into account with a much larger weight than small ones. To demonstrate the average number of correct leading digits of the bounds in Table 3 and Table 4 we present the mean values of the common logarithms of the relative errors. In parenthesis we added the standard deviations.

**Acknowledgement** Thanks are due to Wolfgang Mackens for stimulating discussions.

# References

[1] D. Braess, Nonlinear Approximation Theory. Springer Verlag, Berlin 1986

[2] G. Cybenko and C. Van Loan, Computing the minimum eigenvalue of a symmetric positive definite Toeplitz matrix. SIAM J. Sci. Stat. Comput. 7 (1986), 123 — 131

[3] A. Dembo, Bounds on the extreme eigenvalues of positive definite Toeplitz matrices. IEEE Trans. Information Theory, 34 (1988), 352 — 355

[4] J.J. Dongarra and D.C. Sorensen, A fully parallel algorithm for the symmetric eigenvalue problem. SIAM J. Sci. Stat. Comput, 8 (1987), pp. s139 — s154

[5] I.C. Gohberg and A.A. Semencul, On the inversion of finite Toeplitz matrices and their continuous analogs. Mat.Issled. 2 (1972), 201 — 233

[6] G.H. Golub and C.F. Van Loan, Matrix Computations. 3rd edition. The John Hopkins University Press, Baltimore and London, 1996.

[7] Y.H. Hu and S.-Y. Kung, Toeplitz eigensystem solver. IEEE Trans. Acoustics, Speech, Signal Processing 33 (1985), 1264 — 1271

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

[9] T. Huckle, Circulant and skewcirculant matrices for solving Toeplitz matrices. SIAM J. Matr. Anal. Appl. 13 (1992), 767 — 777

[10] W. Mackens and H. Voss, The minimum eigenvalue of a symmetric positive definite Toeplitz matrix and rational Hermitian interpolation. SIAM J. Matr. Anal. Appl. 18 (1997), 521 — 534

[11] W. Mackens and H. Voss, A projection method for computing the minimum eigenvalue of a symmetric positive definite Toeplitz matrix. Lin. Alg. Appl. 275-276 (1998), 401 — 415

[12] A. Melman, Bounds on the extreme eigenvalues of real symmetric Toeplitz matrices. Stanford University, Report SCCM 98-12

[13] N. Mastronardi and D. Boley, Computing the smallest eigenpair of a symmetric positive definite Toeplitz matrix. To appear in SIAM J. Sci. Comput.

[14] V.F. Pisarenko, The retrieval of harmonics from a covariance function. Geophys. J. R. astr. Soc. 33 (1973), 347 — 366

[15] W.F. Trench, Interlacement of the even and odd spectra of real symmetric Toeplitz matrices. Lin. Alg. Appl. 195 (1993), 59 — 68

[16] H. Voss, Symmetric schemes for computing the minimum eigenvalue of a symmetric Toeplitz matrix. To appear in Lin. Alg. Appl.