LOCALIZATION OF THE EIGENVALUES OF TOEPLITZ MATRICES USING ADDITIVE DECOMPOSITION, EMBEDDING IN CIRCULANTS, AND THE FOURIER TRANSFORM

Paulo Jorge S. G. Ferreira*

**Dep. de Electrónica e Telecomunicações / INESC, Universidade de Aveiro, 3800 Aveiro, Portugal. Fax +351-34-21740, e-mail* `pjf@inesca.pt`.

Abstract. This paper explores the relationship between Toeplitz and circulant matrices. Upper and lower bounds for all eigenvalues of hermitian Toeplitz matrices are given, capitalizing on the possibility of embedding a Toeplitz matrix in a circulant, and of expressing any $n \times n$ Toeplitz matrix as a sum of two matrices with known eigenvalues. The bounds can be simultaneously found using a single discrete Fourier transform evaluation. Thus, the total computation time is $O(\log_2 n)$ per bound. Simulation results indicate that the bounds are sharper than a few other known bounds.

Key Words. Computational methods; digital signal processing; eigenvalues; fast Fourier transforms; matrix algebra; numerical methods.

1. INTRODUCTION

The class of Toeplitz matrices is extremely important, for a number of theoretical and practical reasons. These matrices arise naturally in a variety of problems, including trigonometric moment problems, optimum filtering, linear prediction and spectral estimation.

The eigenanalysis of Toeplitz matrices is a task often required in signal processing and control. Finding the eigenvalues of a Toeplitz matrix is a problem with important applications in spectral estimation (Kay and Marple, 1981; Pisarenko, 1973), to quote a specific example. It is not surprising that considerable efforts have been made towards the derivation of bounds for the extreme eigenvalues of Toeplitz matrices. See Hertz (1992) and references therein.

This paper explores the close relationship between Toeplitz and circulant matrices. A very simple embedding result, and a somewhat less direct additive decomposition of perhaps theoretical interest, enable the derivation of bounds for *all* eigenvalues of symmetric Toeplitz matrices.

The bounds have a number of desirable characteristics, including low computational complexity, which is only $O(\log_2 n)$ operations per bound. This is a consequence of the use of the fast Fourier transform algorithm. Despite this, the bounds are sharper than a few other bounds (Slepian and Landau, 1978; Hertz, 1992), and often give tight estimates for the eigenvalues.

2. THE EMBEDDING

Throughout this paper, M_n denotes the set of $n \times n$ matrices over the complex field. The notations $\mathbf{T} \geq 0$ and $\mathbf{T} > 0$ mean that $\mathbf{T}$ is nonnegative definite and positive definite, respectively.

An arbitrary Toeplitz matrix $\mathbf{T} \in M_n$ with elements $T_{ik} = t_{i-k}$ $(0 \leq i, k < n)$ can be embedded in a circulant $\mathbf{C} \in M_m$ if $m \geq 2n-1$. In general, this leads to a matrix $\mathbf{C}$ which will contain $\ell = m - 2n + 1$ arbitrary parameters α_i. The embedding is defined by the elements c_i of the first row of $\mathbf{C}$,

$$c_i = \begin{cases} t_i & 0 \leq i < n, \\ \alpha_{i-n} & n \leq i < n+\ell, \\ t_{i-m} & n+\ell \leq i < m. \end{cases} \tag{1}$$

It can be verified that the matrix $\mathbf{C}$ so defined is indeed circulant.

The results given in this paper could account for hermitian Toeplitz matrices over the complex field, but, for simplicity, only real $n \times n$ symmetric Toeplitz matrices will be considered. These matrices are defined by their first rows t_i $(0 \leq i < n)$. The modifications for the hermitian Toeplitz case are trivial and will be omitted.

To illustrate the embedding (1) two examples will be given, the first being based on the Toeplitz matrix

$$\mathbf{T} = \begin{bmatrix} t_0 & t_1 & t_2 \\ t_1 & t_0 & t_1 \\ t_2 & t_1 & t_0 \end{bmatrix}.$$

It is convenient to define a matrix $\mathbf{S}$ by

$$\mathbf{S} = \begin{bmatrix} \alpha_0 & t_2 & t_1 \\ t_2 & \alpha_0 & t_2 \\ t_1 & t_2 & \alpha_0 \end{bmatrix},$$

where α_0 is an arbitrary parameter. Embedding $\mathbf{T}$ in a circulant of order six, using (1), leads to the matrix $\mathbf{C}$

$$\begin{aligned}\mathbf{C} &= \left[\begin{array}{ccc|ccc} t_0 & t_1 & t_2 & \alpha_0 & t_2 & t_1 \\ t_1 & t_0 & t_1 & t_2 & \alpha_0 & t_2 \\ t_2 & t_1 & t_0 & t_1 & t_2 & \alpha_0 \\ \hline \alpha_0 & t_2 & t_1 & t_0 & t_1 & t_2 \\ t_2 & \alpha_0 & t_2 & t_1 & t_0 & t_1 \\ t_1 & t_2 & \alpha_0 & t_2 & t_1 & t_0 \end{array}\right]. \\ &= \left[\begin{array}{c|c} \mathbf{T} & \mathbf{S} \\ \hline \mathbf{S} & \mathbf{T} \end{array}\right] \end{aligned} \tag{2}$$

For the second example, let the desired order of the circulant be eight. Applying (1) yields the circulant

$$\left[\begin{array}{ccc|ccccc} t_0 & t_1 & t_2 & \alpha_0 & \alpha_1 & \alpha_0 & t_2 & t_1 \\ t_1 & t_0 & t_1 & t_2 & \alpha_0 & \alpha_1 & \alpha_0 & t_2 \\ t_2 & t_1 & t_0 & t_1 & t_2 & \alpha_0 & \alpha_1 & \alpha_0 \\ \hline \alpha_0 & t_2 & t_1 & t_0 & t_1 & t_2 & \alpha_0 & \alpha_1 \\ \alpha_1 & \alpha_0 & t_2 & t_1 & t_0 & t_1 & t_2 & \alpha_0 \\ \alpha_0 & \alpha_1 & \alpha_0 & t_2 & t_1 & t_0 & t_1 & t_2 \\ t_2 & \alpha_0 & \alpha_1 & \alpha_0 & t_2 & t_1 & t_0 & t_1 \\ t_1 & t_2 & \alpha_0 & \alpha_1 & \alpha_0 & t_2 & t_1 & t_0 \end{array}\right]$$

In this case, the number of arbitrary parameters α_i would be $\ell = 3$. However, a symmetry constraint has been enforced upon the circulant, and reduced that number down to two. In general, a symmetry constraint reduces the number of free parameters from ℓ to $(\ell+1)/2$ or $\ell/2$, depending on whether ℓ is odd or even.

The idea of embedding a Toeplitz matrix in a circulant is a very natural one. As far as the author knows, it was used by Gray (1972) and Dembo *et al.* (1989), although in a different context and for quite different purposes.

3. EIGENSTRUCTURE

A number of aspects pertaining to the eigenvalues and eigenvectors of the block-matrices appearing in (2) are discussed in this section. The results will subsequently be used to obtain bounds for the eigenvalues of symmetric Toeplitz matrices.

In the following, it is assumed that $m = 2n$. This implies that $\mathbf{C}$ has even order and only one free parameter.

It can be verified that the matrix $\mathbf{F} \in M_{2n}$

$$\mathbf{F} = \frac{1}{\sqrt{2}} \begin{bmatrix} +\mathbf{I} & +\mathbf{I} \\ +\mathbf{I} & -\mathbf{I} \end{bmatrix},$$

where $\mathbf{I} \in M_n$, is unitary and reduces $\mathbf{C} \in M_{2n}$ to block-diagonal form, with block-diagonal matrices $\mathbf{T}+\mathbf{S}$ and $\mathbf{T}-\mathbf{S}$.

Note that $\mathbf{F}$ is the block-analog of the Fourier matrix of order two, which diagonalizes all order two circulants, and that $\mathbf{C}$ is block-circulant, as can be seen from (2).

Thus, $\mathbf{C} \geq 0$ if and only if $\mathbf{T}+\mathbf{S} \geq 0$ and $\mathbf{T}-\mathbf{S} \geq 0$, a fact that will be used in the sequel.

Since $\mathbf{S}$ is Toeplitz, both $\mathbf{T}+\mathbf{S}$ and $\mathbf{T}-\mathbf{S}$ are Toeplitz matrices. It is easy to see that $\mathbf{T}+\mathbf{S}$ is circulant whereas $\mathbf{T}-\mathbf{S}$ is not. It is possible to explicitly find the eigenvectors and eigenvalues of these two matrices. To do this, recall that

$$\mathbf{x}_i = [w^0, w^i, w^{2i}, \ldots, w^{(2n-1)i}]^T,$$

where $w = e^{j2\pi/2n}$, and $0 \leq i < 2n$, are orthogonal eigenvectors of $\mathbf{C} \in M_{2n}$. We will denote by λ_i the corresponding eigenvalues, which are just the discrete Fourier transform (DFT) coefficients of the first row of $\mathbf{C}$,

$$\lambda_i = \sum_{k=0}^{2n-1} c_k e^{j2\pi ik/2n}. \tag{3}$$

Partitioning $\mathbf{x}_i$ in two $n \times 1$ blocks

$$\mathbf{x}_i^T = [\mathbf{u}_i^T | \mathbf{v}_i^T]$$

and pre-multiplying by $\mathbf{C}$ leads to

$$\begin{cases} \mathbf{T}\mathbf{u}_i + \mathbf{S}\mathbf{v}_i = \lambda_i \mathbf{u}_i \\ \mathbf{S}\mathbf{u}_i + \mathbf{T}\mathbf{v}_i = \lambda_i \mathbf{v}_i \end{cases}.$$

Using the fact that $\mathbf{v}_i = (-1)^i \mathbf{u}_i$ one obtains

$$\begin{cases} (\mathbf{T}+\mathbf{S})\mathbf{u}_{2i} = \lambda_{2i}\mathbf{u}_{2i} \\ (\mathbf{T}-\mathbf{S})\mathbf{u}_{2i+1} = \lambda_{2i+1}\mathbf{u}_{2i+1} \end{cases}.$$

This shows how the eigenvectors and eigenvalues of $\mathbf{T}+\mathbf{S}$ and $\mathbf{T}-\mathbf{S}$ are related to those of $\mathbf{C}$: the even and odd coefficients of the DFT of the first row of $\mathbf{C}$ are the eigenvalues of $\mathbf{T}+\mathbf{S}$ and $\mathbf{T}-\mathbf{S}$, respectively. The corresponding eigenvectors are $\mathbf{u}_{2i}$ $(0 \leq i < n)$ and $\mathbf{u}_{2i+1}$ $(0 \leq i < n)$, respectively. Note that

$$\mathbf{u}_i = [w^0, w^i, w^{2i}, \ldots, w^{(n-1)i}]^T,$$

where $w = e^{j2\pi/2n}$.

Although the vectors in the sets $\mathbf{u}_{2i}$ $(0 \le i < n)$ and $\mathbf{u}_{2i+1}$ $(0 \le i < n)$ satisfy the orthogonality conditions $\langle \mathbf{u}_{2i}, \mathbf{u}_{2k} \rangle = 0$ and $\langle \mathbf{u}_{2i+1}, \mathbf{u}_{2k+1} \rangle = 0$ $(i \ne k)$, one has $\langle \mathbf{u}_{2i}, \mathbf{u}_{2k+1} \rangle \ne 0$.

4. BOUNDS FOR THE EIGENVALUES

Let $\lambda_0 \ge \lambda_1 \ge \cdots \ge \lambda_{m-1}$ denote the ordered eigenvalues of $\mathbf{C}$, and $\mu_0 \ge \mu_1 \ge \cdots \ge \mu_{n-1}$ denote the ordered eigenvalues of $\mathbf{T}$. For simplicity, assume that $m = 2n$.

This section describes how to obtain bounds for all eigenvalues of $\mathbf{T}$. Two different methods are used, the first being based on the embedding mentioned in section 2.

It is well-known that the eigenvalues of any of the $(n-1)\times(n-1)$ principal submatrices of a symmetric $n \times n$ matrix $\mathbf{M}$ separate the eigenvalues of $\mathbf{M}$ itself. This is expressed by the classical interlacing inequalities (Horn and Johnson, 1990). Iterating them n times one is lead to the relations

$$\lambda_{n+i} \le \mu_i \le \lambda_i, \quad (0 \le i < n). \tag{4}$$

Inequalities valid for other m could also be readily obtained.

The λ_i are the DFT coefficients of the first row of $\mathbf{C}$ (see (3)). Thus, if n is a power of two, they can be found using the fast Fourier transform (FFT) algorithm, and the number of necessary arithmetic operations will be roughly proportional to $\log 2n$ per bound. Even when n is not a composite number, m can be taken as the first power of two exceeding $2n-1$, and the FFT still applied.

The second method depends on the results given in section 3. The matrix $\mathbf{T}$ is expressed as

$$\mathbf{T} = \frac{1}{2}\left[(\mathbf{T}+\mathbf{S}) + (\mathbf{T}-\mathbf{S})\right],$$

which enables the results on the eigenvalues of $\mathbf{T} \pm \mathbf{S}$ to be used. It can be seen, considering the quadratic form associated with $\mathbf{T}$, that the maximum eigenvalue of $\mathbf{T}$ cannot exceed the sum of the maximum eigenvalues of $\mathbf{T} \pm \mathbf{S}$. Similarly, the minimum eigenvalue of $\mathbf{T}$ can not be less than the sum of the minimum eigenvalues of $\mathbf{T} \pm \mathbf{S}$. Thus

$$\mu_0 \le \frac{1}{2}\left(\max_{0\le i<n} \lambda_{2i} + \max_{0\le i<n} \lambda_{2i+1}\right), \tag{5}$$

$$\mu_n \ge \frac{1}{2}\left(\min_{0\le i<n} \lambda_{2i} + \min_{0\le i<n} \lambda_{2i+1}\right), \tag{6}$$

where now the λ_i stand for the eigenvalues of $\mathbf{C}$ with their natural ordering, expressed by (3).

These are better bounds than the ones which follow from (4), namely,

$$\mu_0 \le \lambda_0 = \max_{0\le i<2n} \lambda_i,$$

$$\mu_n \ge \lambda_{2n} = \min_{0\le i<2n} \lambda_i.$$

In fact, using a theorem due to Weyl and originally derived in the context of integral equations, one can give upper and lower bounds for all eigenvalues of $\mathbf{T}$. The theorem asserts that, if $\mathbf{A}$ and $\mathbf{B}$ are hermitian, and if their eigenvalues $\lambda_i(\mathbf{A})$ and $\lambda_i(\mathbf{B})$ are sorted by increasing order, the ordered eigenvalues of $\mathbf{A}+\mathbf{B}$, $\lambda_k(\mathbf{A}+\mathbf{B})$, satisfy

$$\begin{aligned} \lambda_k(\mathbf{A}) + \lambda_1(\mathbf{B}) &\le \lambda_k(\mathbf{A}+\mathbf{B}), \\ \lambda_k(\mathbf{A}+\mathbf{B}) &\le \lambda_k(\mathbf{A}) + \lambda_n(\mathbf{B}). \end{aligned}$$

For details about the theorem see Horn and Johnson (1990). Setting, for example, $2\mathbf{A} = \mathbf{T}+\mathbf{S}$ and $2\mathbf{B} = \mathbf{T}-\mathbf{S}$, and using the fact that the eigenvalues of $\mathbf{A}$ and $\mathbf{B}$ are the even and odd subsequences of the eigenvalues of the circulant $\mathbf{C}$, leads immediately to upper and lower bounds for the eigenvalues of $\mathbf{T}$,

$$\begin{aligned} 2\lambda_k(\mathbf{T}) &\ge \lambda_k(\mathbf{T}+\mathbf{S}) + \lambda_1(\mathbf{T}-\mathbf{S}), \\ 2\lambda_k(\mathbf{T}) &\le \lambda_k(\mathbf{T}+\mathbf{S}) + \lambda_n(\mathbf{T}-\mathbf{S}). \end{aligned} \tag{7}$$

The bounds (5) and (6) are special cases of these equations.

5. OPTIMUM VALUES FOR THE α_i

The bounds given before depend on the $\ell = m - 2n + 1$ parameters α_i, and so it is natural to ask what values of the α_i lead to best-possible results. This problem is addressed in this section.

Again, the notation λ_i $(0 \le m < m)$ denotes the eigenvalues of the circulant $\mathbf{C}$, with the ordering given by (3). The average value of the eigenvalues is denoted by

$$\bar\lambda = \frac{1}{m}\sum_{k=0}^{m-1} \lambda_k,$$

If *best-possible* is considered to mean that the quantity

$$\xi(\alpha_0, \ldots, \alpha_{\ell-1}) = \sum_{k=0}^{m-1} (\lambda_k - \bar\lambda)^2,$$

is to be minimized, the answer is quite simple. One can then show that the all α_i must be zero. When this happens, the eigenvalues of the circulant are clustered as closely as possible in the Euclidean norm around $\bar\lambda$.

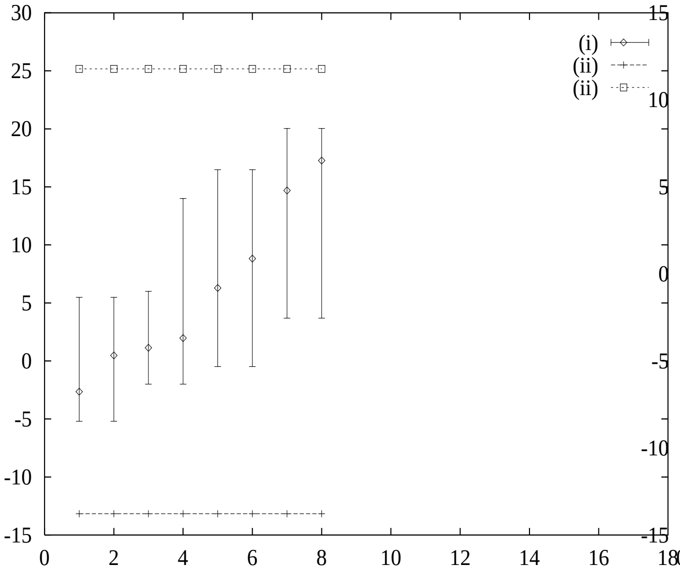

Fig. 1. (i) Bounds obtained using (4); (ii) bounds given in Hertz (1992).

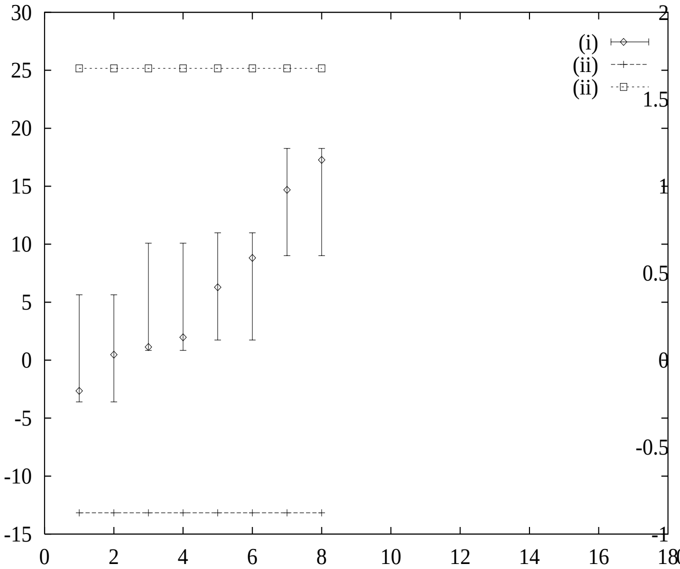

Fig. 2. (i) Bounds obtained using (7); (ii) bounds given in Hertz (1992).

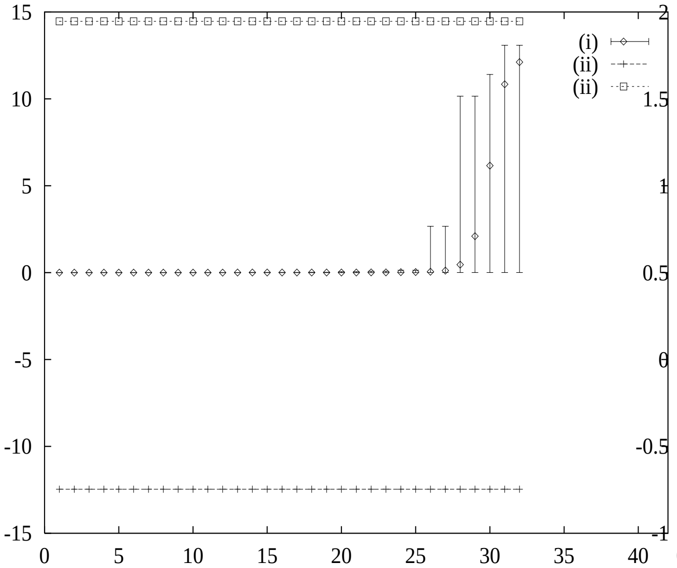

Fig. 3. (i) Bounds obtained using (4); (ii) bounds given in Hertz (1992).

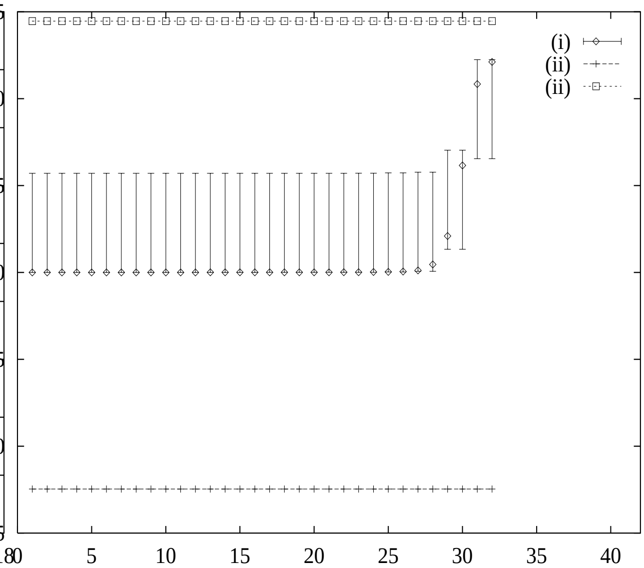

Fig. 4. (i) Bounds obtained using (7); (ii) bounds given in Hertz (1992).

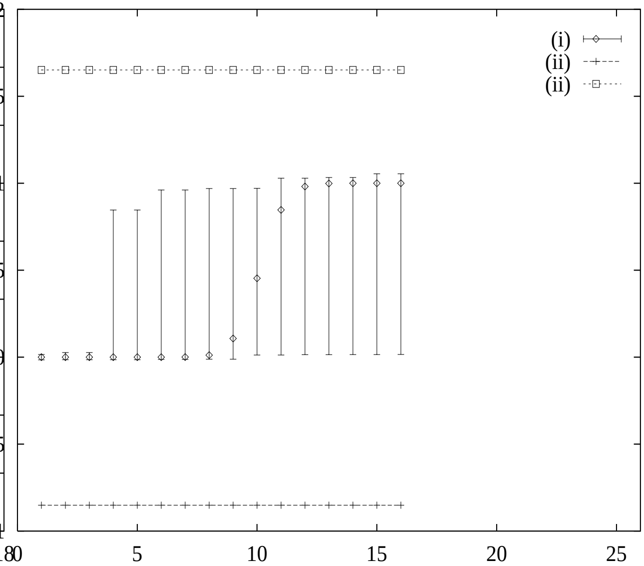

Fig. 5. (i) Bounds obtained using (4); (ii) bounds given in Hertz (1992).

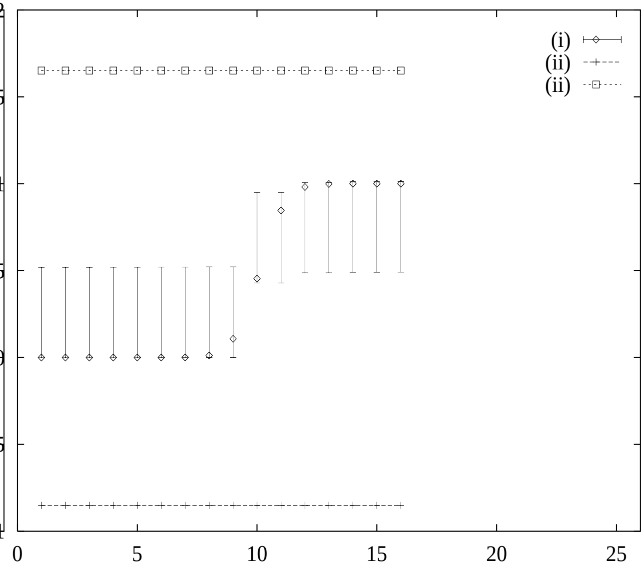

Fig. 6. (i) Bounds obtained using (7); (ii) bounds given in Hertz (1992).

For brevity, a detailed discussion will be omitted. However, note that

$$\frac{\partial \lambda_k}{\partial \alpha_i} = e^{j2\pi k(n+i)/m},$$

and that $\bar{\lambda} = t_0$. It is easy to see that $\partial\xi/\partial\alpha_i = 0$, for $i = 0, 1, \ldots, \ell - 1$, leads to

$$\sum_{k=0}^{m-1} \lambda_k e^{j2\pi k(n+i)/m} = 0,$$

which implies $c_{m-n-i} = 0$, that is, $\alpha_i = 0$. Thus, as stated before, the values which render the bounds best-possible are $\alpha_i = 0$ ($i = 0, 1, \ldots, \ell - 1$), independently of the matrix $\mathbf{T}$.

6. POSITIVE-DEFINITENESS OF C

In this section we give a simple condition on $\mathbf{T}$ which ensures that $\mathbf{C} \geq 0$.

Let $\mathbf{C}$ be defined as in (2), with $\alpha_0 = 0$. Let $\mathbf{b}^T = [t_0, \ldots, t_{n-1}]$ and $\mathbf{c}^T = [t_{n-1}, \ldots, t_0]$. If $\mathbf{T} > 0$, and if $|\mathbf{b}^T\mathbf{T}^{-1}\mathbf{c}| < 1$, then $\mathbf{C} \geq 0$.

The proof is as follows. It has been shown that $\mathbf{C} \geq 0$ if and only if $\mathbf{T} + \mathbf{S} \geq 0$ and $\mathbf{T} - \mathbf{S} \geq 0$. Adapting a result given by Johnson (1990) to the Toeplitz case one finds that $|\mathbf{b}^T\mathbf{T}^{-1}\mathbf{c}| < 1$ implies that $\mathbf{T}$ can be embedded in a $(n+1) \times (n+1)$ Toeplitz positive definite matrix $\mathbf{M}$ with zeros in the upper-right and lower-left corners.

Since $\mathbf{M} > 0$, a classical theorem connected with the trigonometric moment problem (Horn and Johnson, 1990; Riesz and Sz.-Nagy, 1955) asserts the existence of a nondecreasing f such that

$$t_k = \int_{-\pi}^{+\pi} e^{jkx} df(x),$$

with $t_n = 0$. Using this one finds

$$\mathbf{v}^T\mathbf{T}\mathbf{v} = \int_{-\pi}^{+\pi} \left| \sum_k v_k e^{jkx} \right|^2 df(x),$$

$$\mathbf{v}^T\mathbf{S}\mathbf{v} = \int_{-\pi}^{+\pi} e^{jnx} \left| \sum_k v_k e^{jkx} \right|^2 df(x).$$

Since the t_k are real by hypothesis, this can be shown to imply, for $-1 \leq \theta \leq 1$,

$$\mathbf{v}^T(\mathbf{T} \pm \mathbf{S})\mathbf{v} = (1 \pm \theta)\mathbf{v}^T\mathbf{T}\mathbf{v} \geq 0,$$

which proves the assertion.

7. EXAMPLES AND DISCUSSION

Simple as they are, these bounds outperform some other well-known bounds (Slepian and Landau, 1978; Hertz, 1992). In fact, they often are surprisingly accurate.

The data in figs. 1 and 2 refer to a Toeplitz matrix with first row

$$[6, -4, 2, 1, -1, 3, 1, 2].$$

Figures 3 and 4 refer to the autocorrelation matrix of order 32 of a linear combination of sinusoids in additive white noise.

Figures 5 and 6 concern the Toeplitz matrix defined by

$$T_{ik} = \frac{\sin[\pi r(i-k)]}{\pi(i-k)}.$$

This matrix occurs in the recovery of missing samples from band-limited signals, sampled with oversampling factor $0 < r < 1$ (Marks II, 1983; Ferreira, 1992). It is also related to the discrete prolate spheroidal functions (Slepian, 1978).

In all cases, the upper and lower bounds given in Hertz (1992) are also shown.

The accuracy of the bounds suggest a number of possibilities, such as using them to find initial guesses for iterative eigen-routines. Spectral analysis or model identification using the Pisarenko harmonic decomposition method requires finding the smallest eigenvalue of a Toeplitz matrix (Kay and Marple, 1981; Pisarenko, 1973). If the order of the model is unknown, a computationally expensive algorithm results. Can these bounds help in to decide upon the model order? These are two questions, related to the results given herein, which are probably worthy of further investigation.

8. REFERENCES

Dembo, A., C. L. Mallows and L. A. Shepp (1989). Embedding nonnegative definite Toeplitz matrices in nonnegative definite circulant matrices, with application to covariance estimation. *IEEE Transactions on Information Theory* **35**(6), 1206–1212.

Ferreira, P. J. S. G. (1992). Incomplete sampling series and the recovery of missing samples from oversampled band-limited signals. *IEEE Transactions on Signal Processing* **40**(1), 225–227.

Gray, R. M. (1972). On the asymptotic eigenvalue distribution of Toeplitz matrices. *IEEE Transactions on Information Theory* **18**(6), 725–730.

Hertz, D. (1992). Simple bounds on the extreme eigenvalues of Toeplitz matrices. *IEEE Transactions on Information Theory* **38**(1), 175–176.

Horn, R. A. and C. R. Johnson (1990). *Matrix Analysis*. Cambridge University Press. Cambridge.

Johnson, C. R. (1990). Positive definite completions: A guide to selected literature. In: *Signal Processing, Part I: Signal Processing Theory* (L. Auslander et al., Eds.). pp. 167–188. Springer-Verlag. New York.

Kay, S. M. and S. L. Marple (1981). Spectrum analysis — a modern perspective. *Proceedings of the IEEE* **69**(11), 1380–1419.

Marks II, R. J. (1983). Restoring lost samples from an oversampled band-limited signal. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **31**(3), 752–755.

Pisarenko, V. F. (1973). The retrieval of harmonics from a covariance function. *Geophys. J. Roy. Astron. Soc.* **33**, 347–366.

Riesz, F. and B. Sz.-Nagy (1955). *Functional Analysis*. Frederick Ungar Publishing Company. New York. Reprinted by Dover Publications, 1990.

Slepian, D. (1978). Prolate spheroidal wave functions, Fourier analysis and uncertainty — V: The discrete case. *The Bell System Technical Journal* **57**(5), 1371–1429.

Slepian, D. and H. J. Landau (1978). A note on the eigenvalues of Hermitian matrices. *SIAM Journal on Mathematical Analysis* **9**(2), 291–297.