SIAM J. MATRIX ANAL. APPL.
Vol. 23, No. 3, pp. 728–743

NEW BAND TOEPLITZ PRECONDITIONERS FOR ILL-CONDITIONED SYMMETRIC POSITIVE DEFINITE TOEPLITZ SYSTEMS*

D. NOUTSOS† AND P. VASSALOS†

Abstract. It is well known that preconditioned conjugate gradient (PCG) methods are widely used to solve ill-conditioned Toeplitz linear systems $T_n(f)x = b$. In this paper we present a new preconditioning technique for the solution of symmetric Toeplitz systems generated by nonnegative functions f with zeros of even order. More specifically, f is divided by the appropriate trigonometric polynomial g of the smallest degree, with zeros the zeros of f, to eliminate its zeros. Using rational approximation we approximate $\sqrt{f/g}$ by $\frac{p}{q}$, p, q trigonometric polynomials and consider $\frac{p^2 g}{q^2}$ as a very satisfactory approximation of f. We propose the matrix $M_n = B_n^{-1}(q)B_n(p^2 g)B_n^{-1}(q)$, where $B(\cdot)$ denotes the associated band Toeplitz matrix, as a preconditioner whence a good clustering of the spectrum of its preconditioned matrix is obtained. We also show that the proposed technique can be very flexible, a fact that is confirmed by various numerical experiments so that in many cases it constitutes a much more efficient strategy than the existing ones.

Key words. low rank correction, Toeplitz matrix, conjugate gradient, rational interpolation and approximation, preconditioner

AMS subject classifications. 65F10, 65F15

PII. S0895479800376314

1. Introduction.

In this paper we use and analyze band Toeplitz matrices as preconditioners for the solution of the $n \times n$ ill-conditioned symmetric and positive definite Toeplitz system

$$T_n(f)x = b \tag{1.1}$$

by the preconditioned conjugate gradient (PCG) method, where the matrix $T_n(f) \in \mathbb{R}^{n\times n}$ is produced by a real-valued, even, 2π-periodic function defined in the fundamental interval $[-\pi, \pi]$. Then, the (j, k) element of $T_n(f)$ is given by the Fourier coefficient of f, i.e.,

$$T_n(f)_{j,k} = T_{j-k} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-\mathbf{i}(j-k)x}dx, \qquad 1 \le j, k \le n,$$

where $\mathbf{i}$ is the imaginary unit.

Toeplitz matrices arise very often in a wide variety of applications, as, e.g., in the numerical solution of differential equations using finite differences, in statistical problems (linear prediction), in Wiener–Hopf kernels, in Markov chains, in image and signal processing, etc. (see [13], [6], [25]). The generating function f plays a significant role in the location and distribution of the eigenvalues of Toeplitz matrix [13], [7] and in many cases is a priori known. As it is known for the spectrum of $T_n(f)$ there holds $\sigma(T_n(f)) \subseteq [\inf f, \sup f]$.

*Received by the editors August 4, 2000; accepted for publication (in revised form) by L. Eldén August 2, 2001; published electronically January 11, 2002.
http://www.siam.org/journals/simax/23-3/37631.html
†Department of Mathematics, University of Ioannina, GR-451 10, Ioannina, Greece (dnoutsos@cc.uoi.gr, pvassal@cc.uoi.gr). The research of the second author was supported by the Hellenic Foundation of Scholarships (HFS).

Superfast direct methods can solve system (1.1) in $O(n \log^2 n)$ operations, but their stability properties for ill-conditioned Toeplitz matrices are still unclear; see, for instance, [6].

Classical iterative methods such as Jacobi, Gauss–Seidel, and SOR are not effective since the associated spectral radius tends to 1 for large n. The method which is widely used for the solution of such systems is the PCG method. The factors that affect the convergence features of this method are the magnitude of the condition number $\kappa_2(T_n(f))$ and the distribution of the eigenvalues. So a good preconditioner must cluster the eigenvalues of the preconditioned system as much as possible and make the eigenvalues that might lie outside the cluster be bounded by nonzero constants independent of n.

If the generating function is continuous and positive, then problem (1.1) will not be ill-conditioned and the condition number cannot increase proportionally to n, although it can be very large. In this case system (1.1) can be handled by using a preconditioner belonging to some trigonometric matrix algebras (circulant, τ, Hartley; see [24], [2], [3], [23], [14]) or by band Toeplitz preconditioners with weakly increasing bandwidth defined by a polynomial operator $\mathcal{S}_n$, as was proposed in [22]. Theoretically, the latter class of preconditioners seems to perform better as $n \to \infty$ since the number of PCG iterations tends to 1, while in the former cases this number tends to a constant.

When f has a finite number of zeros, each one of finite multiplicity, then system (1.1) is ill-conditioned and the condition number $\kappa_2(T_n(f))$ increases proportionally to n^α where α is the largest number of the multiplicities of the zeros of f [7], [20]. To best handle this case it is necessary to know the number of multiplicities of each one. If this number is not even, then the most suitable technique for this situation [19] fails to make the condition number of the preconditioned matrix independent of its dimension n, and the problem is still open. On the other hand things dramatically change when the multiplicity of each zero is even.

In this case, it was Chan [7] who first proposed as a preconditioner for system (1.1) the Toeplitz band matrix $B_n(g)$ whose generating function g is a trigonometric polynomial that has the same zeros with the same multiplicities as those of f. Next, in [9], not only was g considered as having the zeros of f, but its degree was also increased so that it provided additional degrees of freedom to approximate f and to minimize the relative error $\|\frac{f-g}{g}\|_\infty$ over all trigonometric polynomials g of a fixed degree l. The generating function g is then computed by the Remez algorithm, which can be very expensive from a computational point of view, especially when f has a large number of zeros.

Recently, Serra [21] extended this method by proposing alternative techniques to minimize $\|\frac{f-g}{g}\|_\infty$. More specifically, he chose as g, $z_k g_{l-k}$, where z_k is the trigonometric polynomial of minimum degree k that has all the zeros of f with their multiplicities and g_{l-k} is the trigonometric polynomial of degree $l-k$ which is the best Chebyshev approximation of $\hat{f} = \frac{f}{z_k}$ from the space $\mathcal{P}_{l-k}$ of all trigonometric polynomials of degree at most $l-k$. In addition, in the same work [21], another way was proposed of constructing g_{l-k} by interpolating $\hat{f}$ at the $l-k+1$ zeros of the $(l-k+1)$st degree Chebyshev polynomial of the first kind.

We remark that it has been proved [12] that preconditioners belonging to the aforementioned matrix algebra, when they can be defined, produce weak clustering; i.e., the eigenvalues of the preconditioned matrix are such that for every $\epsilon > 0$ there exists a positive β so that, except for rare exceptions, $O(n^\beta)$ of the eigenvalues lie in the interval $(0, \epsilon)$.

Further preconditioning techniques based on inverses of Toeplitz matrices can be found in [8], [11], [15].

In this paper we extend the previous methods in order to achieve a better clustering for the eigenvalues of the preconditioned matrix and propose a way of constructing a class of preconditioners based on rational approximation or on interpolation to the positive and continuous function $\sqrt{f/z_k}$, with z_k defined previously.

The outline of the present work is as follows. In section 2 we recall some useful issues about the rational approximation, while in section 3 we introduce the technique of constructing the new class of preconditioners based on rational approximation to $\sqrt{f/z_\rho}$ with z_k and analyze the convergence of the PCG method. In section 4 we study the flexibility and possible modification of our method, analyze its cost per iteration, and compare it with that of previous techniques. Finally, in section 5, results of illustrative numerical experiments are exhibited and concluding remarks are made.

2. Preliminaries. In what follows we assume that the generating function f is defined in $[-\pi, \pi]$, is 2π-periodic, continuous, nonnegative, and has zeros of even order.

We define by z_k a trigonometric polynomial of minimum degree k containing all the zeros of f with their multiplicities. Then we define $r_{lm} = \frac{p_l}{q_m}$ as the best rational approximation of $\hat{f} = \sqrt{f/z_k}$ in the uniform norm, i.e.,

$$\|\hat{f} - r_{lm}\|_\infty = \min_{r \in \mathcal{R}(l,m)} \|\hat{f} - r\|_\infty,$$

where $\mathcal{R}(l, m)$ denotes the set of rational functions r, with $p \in \mathcal{P}_l$, $q \in \mathcal{P}_m$, and r irreducible, that is, p and q have no zeros in common.

It is known that when f belongs to some special class of functions [16] then the order of magnitude of the maximum error of an approximation from the space $\mathcal{R}(l, m)$ is better than the corresponding error in the space $\mathcal{P}(l + m)$. In general, we hope that by taking advantage of the flexible nature of rational functions, this set will be a stronger tool than its competitor, the polynomial one. For example, it is obvious that polynomials are not suitable for approximating functions having sharp peaks near the center of their ranges and are slowly varying when $|x|$ increases. Such behavior can be obtained by continuous functions which are not differentiable at some points. However, it is easy to overcome this difficulty by using rational functions.

The next theorem establishes the fact that rational approximation of continuous functions in $[-\pi, \pi]$ is always possible and unique.

THEOREM 2.1. *Let f be in $C[-\pi, \pi]$. Then there exists $r^* \in \mathcal{R}(l, m)$ such that*

$$\|f - r^*\|_\infty < \|f - r\|_\infty$$

for all $r \in \mathcal{R}(l, m)$, $r \neq r^*$.

Proof. See [18, pp. 121, 125] for the proof. ☐

3. Construction of the preconditioner. Let f be a 2π-periodic nonnegative function belonging to $C[-\pi, \pi]$ with zeros $x_1, x_2, \ldots, x_s$ of multiplicities $2\mu_1, 2\mu_2, \ldots,$ $2\mu_s$, respectively, and $2\mu_1 + 2\mu_2 + \cdots + 2\mu_s = \rho$. First, we define

$$z_\rho = \prod_{i=1}^{s} (1 - \cos(x - x_i))^{\mu_i},$$

which is the trigonometric polynomial of minimum degree ρ having all the zeros of f. By dividing f by z_ρ, all its zeros are eliminated and the ratio $\frac{f}{z_\rho}$ becomes a real positive function.

Then, we define the function $\hat{f} = \sqrt{f/z_\rho}$ and approximate it with the rational trigonometric function $r_{lm} = \frac{p_l}{q_m}$, where l, m are the degrees of the numerator and the denominator, respectively. Since $\frac{p_l}{q_m}$ is the best rational approximation of $\sqrt{f/z_\rho}$ for certain l and m, we are led to the conclusion that $\frac{p_l^2}{q_m^2}$ may be a good approximation of $\frac{f}{z_\rho}$. This means that there exists a small $\epsilon > 0$ such that

$$\left\| \frac{f}{z_\rho} - \frac{p_l^2}{q_m^2} \right\|_\infty < \epsilon$$

or, equivalently, that there exists a small $\delta > 0$ such that

$$\left\| \frac{q_m^2}{z_\rho p_l^2} f - 1 \right\|_\infty < \delta.$$

The last inequality means that the values of $\frac{q_m^2}{z_\rho p_l^2} f$ are clustered in a small region near the constant number 1. In terms of matrices, this means that taking $T_n(\frac{z_\rho p_l^2}{q_m^2})$ as a preconditioner matrix for the solution of (1.1), the eigenvalues of $T_n^{-1}(\frac{z_\rho p_l^2}{q_m^2})T_n(f)$ are clustered in a small region near 1 [7] and the PCG method will become very fast. Unfortunately, because this matrix is a full Toeplitz matrix, is hard to construct, and is costly to invert, it is useless as a preconditioner. Instead, we are led to the idea of separating the numerator and the denominator of the ratio $\frac{z_\rho p_l^2}{q_m^2}$ and use as a preconditioner matrix the product of three matrices. More specifically, the preconditioner we propose for the solution of system (1.1) is

$$M_n = B_{nm}^{-1}(q) B_{n\hat{l}}(p^2 z_\rho) B_{nm}^{-1}(q), \qquad \hat{l} = 2l + \rho, \tag{3.1}$$

where the second index in the matrices represents their halfbandwidth, while the first one represents their dimension. The notation $B_{nm}(\cdot)$ will be used instead of $T_n(\cdot)$ for band Toeplitz matrices to emphasize their bandness. The following statements prove the basic assumptions a preconditioner must satisfy and also describe the spectrum of the preconditioned matrix $M_n^{-1}T_n$.

THEOREM 3.1. *The matrix M_n is symmetric and positive definite for every n.*

Proof. Its symmetry is implied directly from the definition (3.1). On the other hand, the eigenvalues of $B_{n\hat{l}}(p^2 z_\rho)$ belong to the interval $(\min p_l^2 z_\rho, \max p_l^2 z_\rho)$, where $0 = \min p_l^2 z_\rho < \max p_l^2 z_\rho \leq \ 2^\rho \max p_l^2$. Therefore, $B_n(p_l^2 z_\rho)$ is symmetric and positive definite. Furthermore, q_m has no zeros in $[-\pi, \pi]$ because it results from the rational approximation to a function which is strictly positive in $[-\pi, \pi]$. So, $B_{nm}(q)$ is symmetric and invertible. Then, for every $x \in \mathbb{R}^n,\ x \neq 0$, we have

$$x^T M_n x = x^T B_{nm}^{-1}(q) B_{n\hat{l}}(p^2 z_\rho) B_{nm}^{-1}(q) x = y^T B_{n\hat{l}}(p^2 z_\rho) y > 0,$$

where $y = B_{nm}^{-1}(q)x$. Hence M_n is symmetric and positive definite. ∎

Theorem 3.1 suggests that the matrix M_n can be taken as a preconditioner matrix. It then remains to study the convergence rate of the PCG method or, equivalently, how the eigenvalues of the matrix $M_n^{-1}T_n$ are distributed. For this, we give without proof the following lemma and then state and prove our main result in Theorem 3.2.

LEMMA 3.1. *Suppose $A, B \in \mathbb{R}^{n\times n}$ are symmetric matrices such that*

$$A = B + \epsilon cc^T,$$

where $c \in \mathbb{R}^n$, $c^Tc = 1$. If $\epsilon > 0$, then

$$\lambda_1(B) \leq \lambda_1(A) \leq \lambda_2(B) \leq \cdots \leq \lambda_n(B) \leq \lambda_n(A),$$

while if $\epsilon \leq 0$, then

$$\lambda_1(A) \leq \lambda_1(B) \leq \lambda_2(A) \leq \cdots \leq \lambda_n(A) \leq \lambda_n(B),$$

provided that the eigenvalues are labeled in nondecreasing order of magnitude. In either case

$$\lambda_k(A) = \lambda_k(B) + t_k\epsilon, \qquad k = 1, 2, \ldots, n,$$

where $t_k \geq 0$, $k = 1, 2, \ldots, n$, and $\sum_{k=1}^n t_k = 1$.

Proof. See Wilkinson [26, pp. 97–98] for the proof. □

THEOREM 3.2. *Let $\lambda_i(M_n^{-1}T_n)$, $i = 1(1)n$, and denote the eigenvalues of $M_n^{-1}T_n$ and m the degree of the denominator q_m of the rational approximation. Then, at least $n - 4m$ eigenvalues of the preconditioned matrix lie in $(h_{\min}, h_{\max})$, at most $2m$ are greater than $h_{\max}$, and at most $2m$ are in $(0, h_{\min})$, where $h = \frac{fq^2}{p^2 z_\rho}$.*

Proof. Obviously the matrix

$$M_n^{-1}T_n = B_{nm}(q)B_{n\hat{l}}^{-1}(p^2z_\rho)B_{nm}(q)T_n(f)$$

is similar to the matrix

$$B_{n\hat{l}}^{-\frac{1}{2}}(p^2z_\rho)B_{nm}(q)T_n(f)B_{nm}(q)B(p^2z_\rho)_{n\hat{l}}^{-\frac{1}{2}}. \tag{3.2}$$

Then, since $B_{nm}(q)$ is a band matrix with halfbandwidth m, the matrix $B_{nm}(q)T_n(f)$ differs from $T_n(qf)$ only in the m first and last rows, and the matrix $B_{nm}(q)T_n(f)B_{nm}(q)$ differs from $T_n(q^2f)$ only in the first and last m rows and columns. So it can be written as a sum of a Toeplitz matrix and a low rank correction matrix, i.e.,

$$B_{nm}(q)T_n(f)B_{nm}(q) = T_n(q^2f) + \Delta, \tag{3.3}$$

where Δ is a symmetric "border" matrix with nonzero elements only in the first and last m rows and columns. So rank$(\Delta) \leq 4m$ is independent of n. Then, from (3.2) and (3.3) we obtain that

$$\begin{aligned}\overbrace{B_{n\hat{l}}^{-\frac{1}{2}}(p^2z_\rho)B_{nm}(q)T_n(f)B_{nm}(q)B_{n\hat{l}}^{-\frac{1}{2}}(p^2z_\rho)}^{\mathrm{E}} &= \overbrace{B_{n\hat{l}}^{-\frac{1}{2}}(p^2z_\rho)T_n(q^2f)B_{n\hat{l}}^{-\frac{1}{2}}(p^2z_\rho)}^{\widetilde{\mathrm{E}}} \\ &\quad + B_{n\hat{l}}^{-\frac{1}{2}}(p^2z)\Delta B_{n\hat{l}}^{-\frac{1}{2}}(p^2z).\end{aligned} \tag{3.4}$$

Since a matrix product does not have rank larger than that of each of the factors involved, there exist $\alpha_i > 0$, $c_i \in \mathbb{R}^n$, $i = 1(1)m_+$, and $\beta_i > 0$, $d_i \in \mathbb{R}^n, i = 1(1)m_-$, with $m_+ + m_- \leq 4m$, such that (3.4) can be written as

$$E - \widetilde{E} = \sum_{i=1}^{m_+} \alpha_i c_i c_i^T - \sum_{i=1}^{m_-} \beta_i d_i d_i^T.$$

So applying successively $m_+ + m_-$ times Lemma 3.1 gives

$$h_{\min} \le \lambda_i(E) \le h_{\max}, \qquad m_- < i \le n - m_+,$$

and the theorem is proved. ◻

It is clear from the previous analysis and statements that contrary to what happens with other band Toeplitz preconditioners, the one we propose of the "premultiplier" matrix $B_{nm}(q)$ may make some of the eigenvalues lie outside the approximation interval $[h_{\min}, h_{\max}]$. We will prove now that the spectral radius of the preconditioned matrix is bounded by a constant number independent of n. For this, first, we state and prove the following lemma.

LEMMA 3.2. *Let B_n be an $n \times n$ symmetric and positive definite band Toeplitz matrix with halfbandwidth s. Then the $k \times k$ principal and trailing submatrices of B_n^{-1} as well as the $k \times k$ submatrices consisting from the first k rows and the last k columns (right upper corner) or from the last k rows and the first k columns (left lower corner) of B_n^{-1} are componentwise bounded for every fixed k independent of n.*

Proof. For principal and trailing submatrices, this property has been proved in [10] for $k = s$. We will prove the validity of this property for $k = s + 1$ and the proof of every fixed k can be completed by induction. From the fundamental relation

$$\sum_{l=1}^{s+1} b_{1l}(B_n^{-1})_{lj} = \delta_{1j},$$

where δ_{1j} is the Kronecker δ, we obtain successively that

$$(B_n^{-1})_{s+1,j} = \frac{1}{b_{1,s+1}} \left(\delta_{1j} - \sum_{l=1}^{s} b_{1l}(B_n^{-1})_{lj} \right), \qquad j = 1, 2, \ldots, s. \tag{3.5}$$

Since all the elements in the right-hand side of (3.5) are bounded, so are the elements $(B_n^{-1})_{s+1,j}$, $j = 1, 2, \ldots, s$. From the symmetry of B_n^{-1} we obtain that the elements $(B_n^{-1})_{j,s+1}$, $j = 1, 2, \ldots, s$, are also bounded. One more application of (3.5) for $j = s + 1$ gives us that the element $(B_n^{-1})_{s+1,s+1}$ is bounded, and the proof for the principal submatrices is complete. Since B_n^{-1} is a persymmetric matrix the elements of the trailing matrix are the same as those of the principal one in reverse order. So the $k \times k$ trailing matrix is also bounded.

It remains to prove the validity of the property for the submatrices in the right upper corner and in the left lower corner of B_n^{-1}. These matrices are transposes of each other due to the symmetry of B_n^{-1}. From the positive definiteness of B_n^{-1} we have that

$$|(B_n^{-1})_{ij}| < \frac{(B_n^{-1})_{ii} + (B_n^{-1})_{jj}}{2}, \qquad i = 1, \ldots, k, \quad j = n - k + 1, \ldots, n.$$

The elements in the right-hand side are the diagonal elements of the $k \times k$ principal and trailing submatrices, respectively, which are bounded, and the proof is complete. ◻

The following theorem proves that the eigenvalues of $M^{-1}T$ have an upper bound.

THEOREM 3.3. *Under the assumptions of Theorem 3.2 there exists a constant c, independent of n, such that $\rho\left(M_n^{-1}T_n(f)\right) \le c$ for every n.*

Proof. We begin the proof by using some relations connecting the spectral radii and the Rayleigh quotients of symmetric matrices. The fact that all the matrices are positive definite is also used.

$$
\begin{aligned}
\rho\left(M_n^{-1}T_n(f)\right) &= \rho\left(B_{nm}(q)B_{n\hat{l}}^{-1}(p^2z_\rho)B_{nm}(q)T_n(f)\right)\\
&= \rho\left(B_{n\hat{l}}^{-\frac{1}{2}}(p^2z_\rho)B_{nm}(q)T_n(f)B_{nm}(q)B_{n\hat{l}}^{-\frac{1}{2}}(p^2z_\rho)\right)\\
&= \max_{x\neq 0}\frac{x^TB_{n\hat{l}}^{-\frac{1}{2}}(p^2z_\rho)B_{nm}(q)T_n(f)B_{nm}(q)B_{n\hat{l}}^{-\frac{1}{2}}(p^2z_\rho)x}{x^Tx}\\
&= \max_{x\neq 0}\left(\frac{x^TT_n(f)x}{x^TB_{nm}^{-1}(q)B_{n\hat{l}}(p^2z_\rho)B_{nm}^{-1}(q)x}\cdot\frac{x^TB_{n\hat{l}}(p^2z_\rho)x}{x^TB_{n\hat{l}}(p^2z_\rho)x}\right)\\
&= \max_{x\neq 0}\left(\frac{x^TT_n(f)x}{x^TB_{n\hat{l}}(p^2z_\rho)x}\cdot\frac{x^TB_{n\hat{l}}(p^2z_\rho)x}{x^TB_{nm}^{-1}(q)B_{n\hat{l}}(p^2z_\rho)B_{nm}^{-1}(q)x}\right)\\
&\leq \max_{x\neq 0}\frac{x^TT_n(f)x}{x^TB_{n\hat{l}}(p^2z_\rho)x}\cdot\max_{x\neq 0}\frac{x^TB_{n\hat{l}}(p^2z_\rho)x}{x^TB_{nm}^{-1}(q)B_{n\hat{l}}(p^2z_\rho)B_{nm}^{-1}(q)x}\\
&= M_1\max_{x\neq 0}\frac{x^TB_{nm}(q)B_{n\hat{l}}(p^2z_\rho)B_{nm}(q)x}{x^TB_{n\hat{l}}(p^2z_\rho)x}\\
&= M_1\max_{x\neq 0}\frac{x^T\left(B_{n\hat{l}+2m}(q^2p^2z_\rho)+\Delta\right)x}{x^TB_{n\hat{l}}(p^2z_\rho)x}\\
&\leq M_1\left(M_2+\max_{x\neq 0}\frac{x^T\Delta x}{x^TB_{n\hat{l}}(p^2z_\rho)x}\right)\\
&\leq M_1\left(M_2+\rho\left(B_{n\hat{l}}^{-1}(p^2z_\rho)\Delta\right)\right).
\end{aligned}
\tag{3.6}
$$

In (3.6) we have taken

$$M_1=\max_{x\neq 0}\frac{x^TT_n(f)x}{x^TB_{n\hat{l}}(p^2z_\rho)x}=\rho\left(B_{n\hat{l}}^{-1}(p^2z_\rho)T_n(f)\right)$$

and

$$M_2=\max_{x\neq 0}\frac{x^TB_{n\hat{l}+2m}(q^2p^2z_\rho)x}{x^TB_{n\hat{l}}(p^2z_\rho)x}=\rho\left(B_{n\hat{l}}^{-1}(p^2z_\rho)B_{n\hat{l}+2m}(q^2p^2z_\rho)\right),$$

which are bounded, since the generating functions $\frac{f}{p^2z_\rho}$ and $\frac{q^2p^2z_\rho}{p^2z_\rho}=q^2$, respectively, are bounded functions in $[-\pi,\pi]$. In (3.6), the matrix product $B_{nm}(q)B_{n\hat{l}}(p^2z_\rho)B_{nm}(q)$ was written as the band Toeplitz matrix $B_{n\hat{l}+2m}(q^2p^2z_\rho)$, generated by the function $q^2p^2z_\rho$, plus the low rank correction matrix Δ.

It is known [5] that the matrix Δ is given by

$$
\begin{aligned}
\Delta = {}& B_{nm}(q)H(q)H(p^2z_\rho)+B_{nm}(q)H^R(q)H^R(p^2z_\rho)\\
&+H(q)H(qp^2z_\rho)+H^R(q)H^R(qp^2z_\rho),
\end{aligned}
$$

where $H(q)$, $H(p^2z_\rho)$, and $H(qp^2z_\rho)$ are Hankel matrices produced by the trigonometric polynomials q, p^2z_ρ, and qp^2z_ρ, respectively, while H^R denotes the matrix obtained from H by reversing the order of its rows and columns.

It is obvious that Δ is a low rank correction matrix that has nonzero elements only in the upper left and lower right triangles, as illustrated below:

$$\Delta = \begin{pmatrix} * & \cdots & * & 0 & \dots & 0 \\ \vdots & \ddots & 0 & \ddots & 0 & \vdots \\ * & 0 & \ddots & 0 & & 0 \\ 0 & \ddots & 0 & \ddots & 0 & * \\ \vdots & 0 & & 0 & \ddots & \vdots \\ 0 & \dots & 0 & * & \dots & * \end{pmatrix}.$$

It is clear that the elements of Δ are bounded and the size of the triangles depends only on the bandwidths m and $\hat{l}$ and are independent of n.

It remains to prove that $\rho(B_{n\hat{l}}^{-1}(p^2 z_\rho)\Delta)$ is bounded. For this, we write the matrices in the following block forms:

$$B_{n\hat{l}}^{-1}(p^2 z_\rho) = \begin{pmatrix} B_1 & * & B_2 \\ * & * & * \\ B_2^T & * & B_1^R \end{pmatrix}, \qquad \Delta = \begin{pmatrix} D & & \\ & O & \\ & & D^R \end{pmatrix},$$

where B_1, B_2 are $k \times k$ matrices if D has k nonzero antidiagonals.

Since the only nonzero columns of the matrix $B_{n\hat{l}}^{-1}(p^2 z_\rho)\Delta$ are its first k and last k ones, the nonidentically zero eigenvalues of $B_{n\hat{l}}^{-1}(p^2 z_\rho)\Delta$ will be the eigenvalues of the matrix

$$\begin{pmatrix} B_1 D & B_2 D^R \\ B_2^T D & B_1^R D^R \end{pmatrix}.$$

In view of Lemma 3.2 this matrix is bounded, and so are its eigenvalues, which proves the present statement. $\square$

So, the eigenvalues that are greater than $h_{\max}$ have an upper bound.

To study the behavior of the eigenvalues that lie in the interval $(0, h_{min})$ we prove the following Lemma.

LEMMA 3.3. *The smallest eigenvalue of the matrix $M_n^{-1}T_n(f)$ has a bound from below a constant number $c_1 > 0$, independent of n, iff the smallest eigenvalue of the matrix $B_{n\rho}^{-1}(z_\rho)B_{nm}(q)B_{n\rho}(z_\rho)B_{nm}(q)$ has lower bound a constant number $c_2 > 0$, independent of n.*

Proof. As in Theorem 3.3 we use the relation connecting the eigenvalues of a symmetric positive definite matrix with the Rayleigh quotient:

$$\begin{aligned}
&\min_i \lambda_i \left(M_n^{-1}T_n(f)\right) = \min_i \lambda_i \left(B_{n\hat{l}}^{-1}(p^2 z_\rho)B_{nm}(q)T_n(f)B_{nm}(q)\right) \\
&= \min_{x\neq 0}\left(\frac{x^T B_{nm}(q)T_n(f)B_{nm}(q)x}{x^T B_{n\hat{l}}(p^2 z_\rho)x}\cdot\frac{x^T B_{nm}(q)B_{n\rho}(z_\rho)B_{nm}(q)x}{x^T B_{nm}(q)B_{n\rho}(z_\rho)B_{nm}(q)x}\cdot\frac{x^T B_{n\rho}(z_\rho)x}{x^T B_{n\rho}(z_\rho)x}\right) \\
&\geq \min_{x\neq 0}\frac{x^T T_n(f)x}{x^T B_{n\rho}(z_\rho)x}\cdot\min_{x\neq 0}\frac{x^T B_{nm}(q)B_{n\rho}(z_\rho)B_{nm}(q)x}{x^T B_{n\rho}(z_\rho)x}\cdot\min_{x\neq 0}\frac{x^T B_{n\rho}(z_\rho)x}{x^T B_{n\hat{l}}(p^2 z_\rho)x} \\
&\geq \min\frac{f}{z_\rho}\cdot\min\frac{1}{p^2}\cdot\min_{x\neq 0}\frac{x^T B_{nm}(q)B_{n\rho}(z_\rho)B_{nm}(q)x}{x^T B_{n\rho}(z_\rho)x}.
\end{aligned}$$

Since the functions $\frac{f}{z_\rho}$ and $\frac{1}{p^2}$ have both lower bounds independent of n, the spectrum of the preconditioned matrix has such a bound iff the Rayleigh quotient $\frac{x^T B_{nm}(q) B_{n\rho}(z_\rho) B_{nm}(q) x}{x^T B_{n\rho}(z_\rho) x}$ does. ∎

The above equivalent problem that the matrix $B_{n\rho}^{-1}(z_\rho) B_{nm}(q) B_{n\rho}(z_\rho) B_{nm}(q)$ has a spectrum bounded from below by a positive constant c independent of n remains in this paper an open question for general values of the bandwidths m and ρ. Despite that, strong numerical evidence shows that this holds. To make our conjecture stronger we present the proof for the special cases where $m = 1$ and $\rho = 1, 2$.

THEOREM 3.4. *The matrix $B_{n\rho}^{-1}(z_\rho) B_{nm}(q) B_{n\rho}(z_\rho) B_{nm}(q)$ has its smallest eigenvalue λ_1 bounded from below by a constant number $c > 0$ which is independent of n for $m = 1$ and $\rho = 1, 2$*

Proof. The case $m = \rho = 1$ is quite obvious and is based on the fact that all the tridiagonal symmetric Toeplitz matrices have the same eigenvectors. More specifically, the matrix $B_{n1}(z_1)$ is the Laplace matrix with its eigenvalues and the corresponding normalized eigenvectors being given by

$$\lambda_i = z_1(\theta_i) = 4 \sin^2 \frac{\theta_i}{2}, \qquad x^{(i)} = \sqrt{\frac{2}{n+1}} \left(\sin\theta_i \ \ \sin 2\theta_i \ \ \sin 3\theta_i \ \ \ldots \ \ \sin n\theta_i \right)^T,$$

respectively, where $\theta_i = \frac{\pi i}{n+1}$, $i = 1(1)n$. The matrix $B_{n1}(q)$ is a tridiagonal Toeplitz matrix of the form $tridiag(\beta,\ \alpha,\ \beta)$. Since $B_{n1}(q)$ and $B_{n1}(z_1)$ have the same eigenvectors we can write any arbitrary vector $x \in \mathbb{R}^n$ as a convex combination $x = \sum_{i=1}^n c_i x^{(i)}$, $c_i \in \mathbb{R}$, $i = 1(1)n$. With these assumptions and using the orthogonal properties of $x^{(i)}$'s the Rayleigh quotient gives

$$\begin{aligned} \frac{x^T B_{n1}(q) B_{n1}(z_1) B_{n1}(q) x}{x^T B_{n1}(z_1) x} &= \frac{\left(\sum_{i=1}^n c_i x^{(i)}\right)^T B_{n1}(q) B_{n1}(z_1) B_{n1}(q) \left(\sum_{i=1}^n c_i x^{(i)}\right)}{\left(\sum_{i=1}^n c_i x^{(i)}\right)^T B_{n1}(z_1) \left(\sum_{i=1}^n c_i x^{(i)}\right)} \\ &= \frac{\sum_{i=1}^n c_i^2 q^2(\theta_i) 4 \sin^2 \frac{\theta_i}{2}}{\sum_{i=1}^n c_i^2 4 \sin^2 \frac{\theta_i}{2}} \geq \min_i q^2(\theta_i) \geq \min_{\theta \in [-\pi, \pi]} q^2(\theta). \end{aligned} \tag{3.7}$$

The proof is complete since the function q is strictly positive.

For the case where $(m, \rho) = (1, 2)$ we write the matrix $B_{n2}(z_2)$ as a function of $B_{n1}(z_1)$ and the corresponding Hankel matrices [5], i.e.,

$$B_{n2}(z_2) = (B_{n1}(z_1))^2 + \left(H(z_1) + H^R(z_1)\right)^2,$$

where the notations H and H^R are the same as in Theorem 3.3. For simplicity we denote $H = H(z_1) + H^R(z_1)$, so $H = \mathrm{diag}(-1,\ 0,\ 0, \ldots,\ 0,\ -1)$.

By considering the same convex combination of the vector x, the Rayleigh quotient gives

$$\frac{x^T B_{n1}(q) B_{n2}(z_2) B_{n1}(q) x}{x^T B_{n2}(z_2) x} = \frac{x^T B_{n1}(q) \left(B_{n1}^2(z_1) + H^2\right) B_{n1}(q) x}{x^T \left(B_{n1}^2(z_1) + H^2\right) x} \tag{3.8}$$

$$= \frac{16 \sum_{i=1}^n c_i^2 q^2(\theta_i) \sin^4 \frac{\theta_i}{2} + \frac{2}{n+1} \left(\sum_{i=1}^n c_i q(\theta_i) \sin\theta_i\right)^2 + \frac{2}{n+1} \left(\sum_{i=1}^n c_i q(\theta_i) \sin n\theta_i\right)^2}{16 \sum_{i=1}^n c_i^2 \sin^4 \frac{\theta_i}{2} + \frac{2}{n+1} \left(\sum_{i=1}^n c_i \sin\theta_i\right)^2 + \frac{2}{n+1} \left(\sum_{i=1}^n c_i \sin n\theta_i\right)^2}.$$

First, we suppose that the first term of the denominator in (3.8) is greater than or equal to the second or the third one in order of magnitude. In that case we obtain that the ratio in (3.8), similar to (3.7), has a lower bound the value $\min_{\theta\in[-\pi,\pi]} q^2(\theta)$. Otherwise, we suppose that the second term is greater than the others in order of magnitude. Since the numerator is a sum of quadratic terms, the ratio will tend to zero if all the terms in the numerator decrease with a higher rate. So, we consider the case where the term $\frac{2}{n+1}\left(\sum_{i=1}^n c_i q(\theta_i)\sin\theta_i\right)^2$ has an order of magnitude less than that of $\frac{2}{n+1}\left(\sum_{i=1}^n c_i \sin\theta_i\right)^2$. By substituting $q(\theta) = \alpha + 2\beta\cos\theta_i = \alpha + 2\beta(1 - 2\sin^2\frac{\theta_i}{2})$, we have

$$\sum_{i=1}^n c_i q(\theta_i)\sin\theta_i = (\alpha + 2\beta)\sum_{i=1}^n c_i \sin\theta_i - 4\beta\sum_{i=1}^n c_i \sin^2\frac{\theta_i}{2}\cdot\sin\theta_i,$$

which means that the terms $\sum_{i=1}^n c_i\sin\theta_i$ and $\sum_{i=1}^n c_i \sin^2\frac{\theta_i}{2}\cdot\sin\theta_i$ must have the same orders of magnitude. Applying the Cauchy–Schwarz inequality on the second sum we obtain that

$$\left(\sum_{i=1}^n c_i \sin^2\frac{\theta_i}{2}\cdot\sin\theta_i\right)^2 \le \sum_{i=1}^n c_i^2\sin^4\frac{\theta_i}{2}\cdot\sum_{i=1}^n \sin^2\theta_i = \frac{n+1}{2}\sum_{i=1}^n c_i^2\sin^4\frac{\theta_i}{2}.$$

So, the order of magnitude of the term $\frac{2}{n+1}\left(\sum_{i=1}^n c_i\sin\theta_i\right)^2$ must be less than or equal to the one of $\sum_{i=1}^n c_i^2\sin^4\frac{\theta_i}{2}$, which is a contradiction. The assumption that the third term is the greater one, in order of magnitude, gives similarly the same contradiction. So, the ratio in (3.8) does not tend to zero as n tends to infinity. □

We remark that the same idea to split the matrix $B_{n\rho}(z_\rho)$ into $(B_{n1}(z_1))^\rho$ plus a sum of Hankel matrices can be used for the proof of the above property in the case of $\rho > 2$. In the case of $m > 1$, first the matrix $B_{nm}(q)$ is written as a sum of the terms $B_{nj}(z_j)$ $j = 0(1)m$, $(B_{n0}(z_0) = I_n)$ and the above idea can be applied. In both cases the analysis becomes more and more complicated. Figures 5.1(b)–(d), 5.2(b), 5.3(b) fully confirm the above properties. Moreover they show that the main interval eigenvalues appear in pairs and the elements of each pair tend to each other as n tends to infinity. In view of this observation, the convergence analysis of the PCG method in [1] assures us that our method will not be seriously affected and the convergence of it will remain superlinear, which is the optimal cost for this method.

4. Computational analysis and modification of the method. In this section we will try to compare, from the computational point of view, our preconditioner with the most recent band Toeplitz preconditioner proposed in [21]. The latter has in general the best performance from all the previous ones, when the generating function f is nonnegative and has zeros of even order.

The main computational cost in every PCG iteration is due to the Toeplitz matrix-vector product $T_n(f)x$ and to the solution of a system with coefficient matrix the preconditioner itself. The first one is the same for both methods and can be computed by means of the fast Fourier transform (FFT) in $10(n\log 2n)$ operations (ops) in a sequential machine, or in $O(\log 2n)$ steps in the parallel PRAM model of computation, when $O(n)$ processors are used. For the inversion of the preconditioners, things slightly change. If we use band Toeplitz preconditioners, then their halfbandwidth $\hat{l}_1$ represents the degree l_1 of the Chebyshev approximation plus the degree ρ of the trigonometric polynomial, which eliminates the zeros of f. The inversion of such type of matrices can be achieved using the LDL^T factorization method in $n(\hat{l}_1^2 + 8\hat{l}_1 + 1)$ ops. We mention that this method is more preferable than the band Cholesky factorization because the latter requires the computation of n square roots, which is quite expensive when n is large.

In the case of our preconditioner the inversion requires two band matrix vector products of total cost $n(8m + 4)$ ops, where m is the halfbandwidth and coincides with the degree of the denominator in the rational approximation. In addition, the inversion of $B_{n\hat{l}_2}$, as in the previous case, can be performed in $n(\hat{l}_2^2 + 8\hat{l}_2 + 1)$ ops, where $\hat{l}_2 = \rho + 2l_2$ and l_2 represents the degree of the numerator of the rational approximation. So the total cost per iteration for this step of the algorithm of the PCG method is about

$$Cost_{it} = n(\hat{l}_2^2 + 8\hat{l}_2 + 8m + 5).$$

We must mention here that more sophisticated techniques reduce the cost of approximating the solution of such systems, up to within an $O(\epsilon)$ error, in $O(n \log m + m \log^2 m \log \log \epsilon^{-1})$ [4]. In both cases, when n is large, the complexity of the method is strongly dominated by that of the first step, which requires $O(n \log 2n)$ ops since $\hat{l}_2, m$ are independent of n. So the methods are essentially equivalent in complexity per iteration. Thus the costs of finding $B_{n\hat{l}_1}^{-1}$ and $B_{nm}B_{n\hat{l}_2}^{-1}B_{nm}$, where $l_1 = l_2 + m$, are comparable.

In case n is not large enough, taking $l_2 = \frac{l_1}{2} - 1$ and making some calculations, we can see that the two preconditioning strategies are approximately equivalent even when $m = \rho l_1$.

According to this observation, if we have two candidates of rational approximations of f with almost the same relative error and degrees (l_1, m_1), (l_2, m_2) with $l_1 + m_1 \approx l_2 + m_2$, it is preferable, from the computation point of view, to choose as the generating function for our preconditioner the one which has the larger m and the smaller l.

Finally, we will focus on the calculation of rational approximation of degree (l, m) of a positive continuous function f. In the recent literature many different strategies that produce this kind of approximation [17] can be found. Each of them is most suitable for certain classes of functions, but the one which is based on the Remez algorithm seems to be, in general, the most appropriate for a large variety of functions. The starting point of this category of algorithms is to construct a rational approximation using rational interpolation, and then this rational approximation is used to generate a better approximation until an alternative set of $m + l + 2$ points is reached. This procedure consists of adjusting the choice of the interpolation points in such a way as to ensure that the relative error decreases. In practice this method can fail in some cases. Usually, problems occur either because the extreme values of the relative error occur more than $m + l + 2$ times, or because the starting rational interpolation has zeros in the interval in which this approximation is sought. The first difficulty is usually overcome by seeking a rational approximation of a different degree or by designing a more robust algorithm. A trick that often works in the latter case is, instead of seeking again for a rational approximation of a different degree, to start with an approximation that is valid over a shorter interval and to use it as a starting point for an approximation on a slightly larger interval. Iterative application of this procedure may enable us to obtain a final approximation in the desired interval.

For the convergence rate of the approximation method we cannot give a theoretical result, but the facts that its computational cost is independent of n and the computations are done only once for a given function make us believe that this issue does not play an important role in the whole procedure.

4.1. Modification of the method. The idea of constructing a preconditioner from a rational approximation of a function can be used in exactly the same way in case of rational interpolation at the Chebyshev points. The advantage of this modification is the simplicity of its calculation. Nevertheless, it is worth noticing that we cannot ensure that this interpolation would not have zeros in the interval of approximation. Despite this, whenever the preconditioning gives us poor results, this technique may give, at least for certain classes of f, results similar to the corresponding ones by the best Chebyshev approximation.

5. Numerical examples and concluding remarks. In this section, we present some numerical examples. The aim of these examples is twofold: (i) to show, by numerical evidence, the correctness of our observations regarding the asymptotical spectral analysis of the preconditioned matrices, and (ii) to compare the convergence rate of our preconditioner with that of the band Toeplitz preconditioner proposed in [21]. We use the latter to compare it with ours because it is the most efficient technique for preconditioning Toeplitz matrices generating by functions with zeros of even order. Our test functions are the following:

$$\text{(i)} \quad f_1(x) = x^4,$$
$$\text{(ii)} \quad f_2(x) = \frac{2x^4}{1+25x^2},$$

and

$$\text{(iii)} \quad f_3(x) = \begin{cases} (x-3)^4(x-1)^2, & 0 \le x \le \pi, \\ (x+3)^4(x+1)^2, & -\pi \le x \le 0. \end{cases}$$

An effort was made to choose functions of different behaviors which produce ill-conditioned matrices T_n. The Toeplitz matrices produced have Euclidean condition numbers of order $O(n^4)$. In our experiments we solve the system $T_n(f)x = b$, where b is the vector having all its components equal to 1. As a starting initial guess of solution the zero vector is used and as a stopping criterion the validity of $\frac{\|r_k\|_2}{\|r_0\|_2} \le 10^{-7}$ is considered, where r_k is the residual vector after k iterations. The construction of matrices and the rational approximations were performed using *Mathematica* in order to have more accurate results, while all the other computations were performed using MATLAB.

In Tables 5.1, 5.2, and 5.3 we report the number of iterations needed until convergence is achieved in each case; B_n^{*l} denotes the optimal band Toeplitz preconditioner [21] which is generated by the trigonometric polynomial $z_\rho g_l$, with g_l being the best Chebyshev approximation of $\frac{f}{z_\rho}$ out of $\mathcal{P}_l$, $\hat{B}_n^l$ is the band Toeplitz preconditioner where $\hat{g}_l$ is the interpolation polynomial at the Chebyshev points, $M_n^{l,m}$ denotes our main proposed preconditioner obtained by the best rational approximation procedure of degree (l,m), and $R_n^{l,m}$ denotes the preconditioner that results after applying rational interpolation of degree (l,m).

TABLE 5.1
Number of iterations for $f_1(x)$.

n	B_n^{*1}	$\hat{B}_n^1$	B_n^{*3}	$\hat{B}_n^3$	B_n^{*4}	$\hat{B}_n^4$	$M_n^{0,1}$	$R_n^{0,1}$	$M_n^{1,1}$	$R_n^{1,1}$	$M_n^{1,2}$	$R_n^{1,2}$
16	9	8	9	7	7	6	8	7	6	6	5	5
32	10	10	11	8	9	7	10	9	7	7	6	6
64	13	12	11	10	9	8	11	11	9	9	8	8
128	15	15	12	11	10	10	12	13	11	11	10	10
256	16	16	12	13	10	10	13	13	12	12	11	11
512	16	16	13	13	10	11	13	14	13	13	11	12

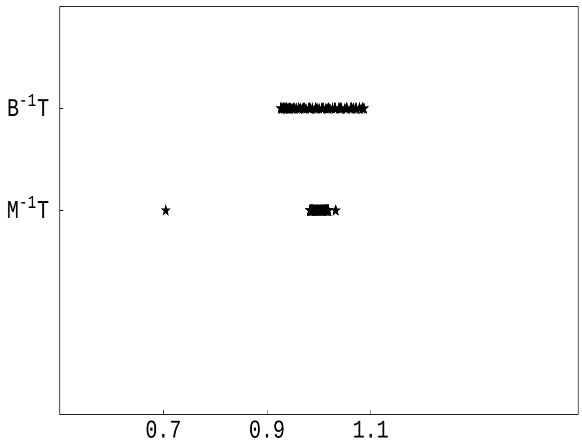
TABLE 5.2
Number of iterations for $f_2(x)$.

n	B_n^{*3}	B_n^{*4}	B_n^{*5}	B_n^{*6}	$M_n^{1,1}$	$R_n^{2,2}$
16	8	8	7	8	8	6
32	13	13	12	11	11	7
64	19	18	15	13	12	9
128	24	19	17	14	12	11
256	25	21	18	15	13	13
512	27	22	18	16	14	14

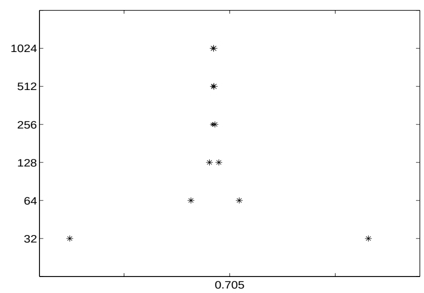
TABLE 5.3
Number of iterations for $f_3(x)$.

n	B_n^{*3}	B_n^{*5}	B_n^{*7}	$M_n^{1,2}$	$R_n^{(1,2)}$
16	9	7	7	9	8
32	17	14	13	18	11
64	34	28	22	21	14
128	65	48	36	21	20
256	111	69	54	23	24
512	152	93	66	23	27

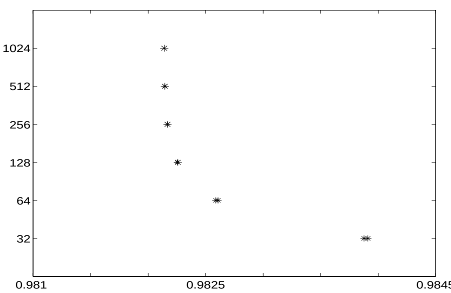
In Figures 5.1(a), 5.2(a), 5.3(a), the spectra of the matrices $M_n^{-1}T_n(f_i)$, $i = 1, 2, 3$, are illustrated, while in Figures 5.1(b)-(d), 5.2(b), 5.3(b) we focus on the behavior of the pairs of eigenvalues of the matrix lying outside the interval $[h_{\min}, h_{\max}]$ for different values of n. The boundness and the convergence in pairs is obvious in all figures. Especially, we stress the case of Figures 5.1 and 5.3, where as we expected from the theory at most eight eigenvalues would lie outside the interval $[h_{\min}, h_{\max}]$, but in practice, for the first test function, only three pairs of eigenvalues lie outside this interval, one of which (the second lower pair) moves very close to the lower bound $h_{\min} = 0.98214$, while, for the third test function, only two pairs lie outside this interval. Finally, we remark that in the case of f_3 and for $n = 512$, the preconditioning by band Toeplitz B^{*3} "clusters" the eigenvalues of the preconditioned matrix in $[0.5, 584.3]$, B^{*5} does so in $[0.36, 104.7]$, while $M^{1,2}$ collects the main mass of them in $[0.67, 1.65]$ and $R^{1,2}$ collects it in $[0.95, 14.25]$.



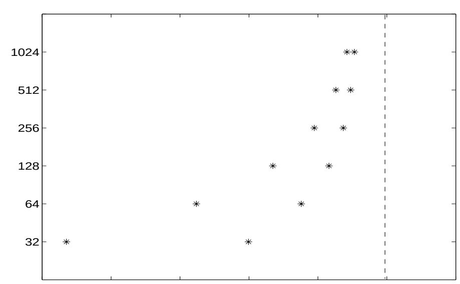
(a) The main mass of the eigenvalues of the preconditioned matrices.



(b) The lower extreme pair.

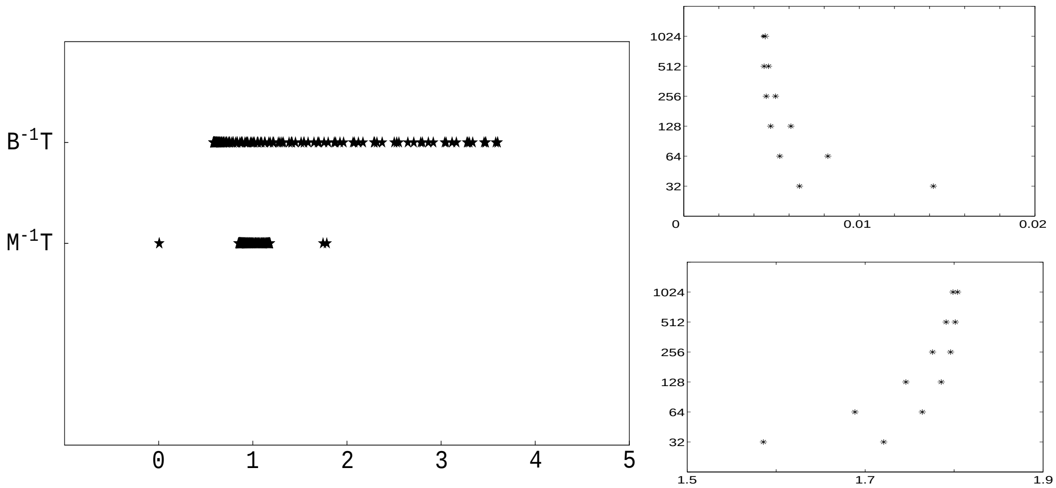


(c) The second upper pair.



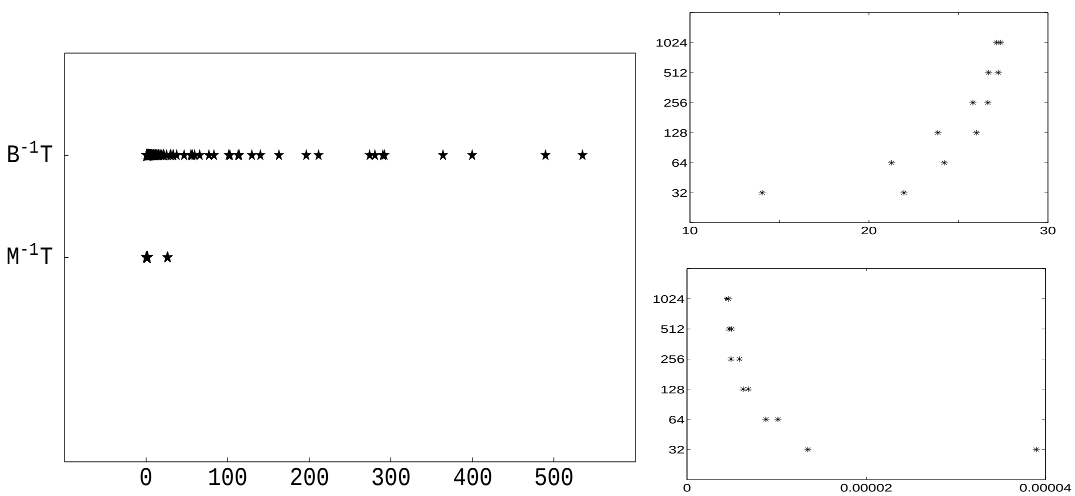
(d) The upper extreme pair.

FIG. 5.1. *Spectra of* $(M_n^{2,2})^{-1}T_n(f_1)$ *and* $(B_n^{*5})^{-1}T_n(f_1)$ *for* $n = 128$ *and behavior of the pairs of eigenvalues that lie outside the interval* $[h_{\min}, h_{\max}]$ *with* $h_{\min} = 0.98214$.



(a) (b)

FIG. 5.2. *Spectra of* $(M_n^{1,1})^{-1}T_n(f_2)$ *and* $(B_n^{*3})^{-1}T_n(f_2)$ *for* $n = 128$ *and behavior of the pairs of eigenvalues that lie outside the interval* $[h_{min}, h_{max}]$.



(a) (b)

FIG. 5.3. *Spectra of* $(M_n^{1,2})^{-1}T_n(f_3)$ *and* $(B_n^{*3})^{-1}T_n(f_3)$ *for* $n = 256$ *and behavior of the pairs of eigenvalues that lie outside the interval* $[h_{min}, h_{max}]$.

REFERENCES

[1] O. Axelsson and G. Lindskog, *On the rate of convergence of the preconditioned conjugate gradient method*, Numer. Math., 48 (1986), pp. 499–523.

[2] D. Bini and F. Di Benedetto, *A new preconditioner for parallel solution of positive definite Toeplitz systems*, in Proceedings of the Second Annual Symposium on Parallel Algorithms and Architectures, Crete, Greece, 1990, pp. 220–223.

[3] D. Bini and P. Favati, *On a matrix algebra related to the discrete Hartley transform*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 500–507.

[4] D. A. Bini and B. Meini, *Effective methods for solving banded Toeplitz systems*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 700–719.

[5] A. Böttcher and B. Silbermann, *Introduction to Large Truncated Toeplitz Matrices*, Springer-Verlag, New York, 1998.

[6] J. R. Bunch, *Stability of methods for solving Toeplitz systems of equations*, SIAM J. Sci. Statist. Comput., 6 (1985), pp. 349–364.

[7] R. H. Chan, *Preconditioners for Toeplitz systems with nonnegative generating functions*, IMA J. Numer. Anal., 11 (1991), pp. 333–345.

[8] R. H. Chan and K. P. Ng, *Toeplitz preconditioners for Hermitian Toeplitz systems*, Linear Algebra Appl., 190 (1993), pp. 181–208.

[9] R. H. Chan and P. T. P. Tang, *Fast band-Toeplitz preconditioners for Hermitian Toeplitz systems*, SIAM J. Sci. Comput., 15 (1994), pp. 164–171.

[10] F. Di Benedetto, *Analysis of preconditioning techniques for ill-conditioned Toeplitz matrices*, SIAM J. Sci. Comput., 16 (1995), pp. 682–697.

[11] F. Di Benedetto, G. Fiorentino, and S. Serra, *C.G. preconditioning of Toeplitz matrices*, Comput. Math. Appl., 25 (1993), pp. 35–45.

[12] F. Di Benedetto and S. Serra, *A unifying approach to abstract matrix algebra preconditioning*, Numer. Math., (1999), pp. 57–90.

[13] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*, 2nd ed., Chelsea, New York, 1984.

[14] X. Q. Jin, *Hartley preconditioners for Toeplitz systems generated by positive continuous functions*, BIT, 34 (1994), pp. 367–371.

[15] T. K. Ku and C. C. Kuo, *A minimum-phase LU factorization preconditioner for Toeplitz matrices*, SIAM J. Sci. Comput., 13 (1992), pp. 1470–1487.

[16] G. Lorentz, *Approximation of Functions*, 2nd ed., Chelsea, New York, 1986.

[17] M. Powell, *Approximation Theory and Methods*, Cambridge University Press, Cambridge, UK, 1982.

[18] T. Rivlin, *Introduction to the Approximation of Functions*, Dover, New York, 1981.

[19] S. Serra, *New PCG based algorithms for the solution of Hermitian Toeplitz systems*, Calcolo, 32 (1995), pp. 154–176.

[20] S. Serra, *On the extreme spectral properties of Toeplitz matrices generated by l^1 functions with several minima (maxima)*, BIT, 36 (1996), pp. 135–142.

[21] S. Serra, *Optimal, quasi-optimal and superlinear band-Toeplitz preconditioners for asymptotically ill-conditioned positive definite Toeplitz systems*, Math. Comput., 66 (1997), pp. 651–665.

[22] S. Serra Capizzano, *Toeplitz preconditioners constructed from linear approximation processes*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 446–465.

[23] S. Serra, *A Korovkin-type theory for finite Toeplitz operators via matrix algebras*, Numer. Math., 82 (1999), pp. 117–142.

[24] G. Strang, *A proposal for Toeplitz matrix calculations*, Stud. Appl. Math., 74 (1986), pp. 171–176.

[25] H. Widom, *Toeplitz matrices*, in Studies in Real and Complex Analysis, Stud. Math. 3, Math. Assoc. Amer., Buffalo, NY, 1965, pp. 179–209.

[26] J. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford Press, Oxford, 1965.