Constrianed Minimax Approximation and
Optimal Preconditioners for Toeplitz Matrices

*Raymond H. Chan*
Department of Mathematics
University of Hong Kong
Hong Kong

and

*Ping Tak Peter Tang*
Mathematics and Computer Science Division
Argonne National Laboratory
Argonne, Illionis

*Abstract* Good preconditioner is extremely important in order for the conjugate gradient method to converge quickly. In the case of Toeplitz matrices, a number of recent studies were made to relate approximation of functions to good preconditioners. In this paper, we present a new result relating the quality of the Toeplitz preconditioner $C$ for the matrix $T$ to the Chebyshev norm $\|(f-g)/f\|_\infty$, where $f$ and $g$ are the generating functions for $T$ and $C$, respectively. In particular, the construction of band-Toeplitz preconditioners becomes a linear minimax approximation problem. The case when $f$ has zeros (but non-negative) is especially interesting and the corresponding approximation problem becomes contrained. We show how the Remez algorithm can be modified to handle the constraints. Numerical experiments confirming the theoretical results will be presented.

*Keywords* Minimax approximation, Remez algorithm, conjugate gradient, Toeplitz matrix, preconditioner.

# 1 Introduction

In this paper we consider an application of constrained minimax approximation to finding preconditioners for symmetric Toeplitz systems generated by a $2\pi$-periodic function.

First, some background material for the sake of completeness. A $n$-by-$n$ symmetric Toeplitz marix $T_n$,

$$T_n = \begin{pmatrix} t_0 & t_1 & \cdots & t_{n-1} \\ t_1 & t_0 & \cdots & t_{n-2} \\ \vdots & \ddots & \ddots & \vdots \\ t_{n-1} & \cdots & t_1 & t_0 \end{pmatrix}, \quad \text{(constant along diagonals)},$$

is said to be generated by a $2\pi$-periodic function $f$ if

$$t_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}, \qquad k = 0, 1, \ldots, n-1,$$

we concentrate in the case where $f$ is even, real-valued, and nonnegative, corresponding to $T_n$ being real, symmetric, and nonnegative definite.

Toeplitz systems of the form $T_n x = b$ occur in a variety of applications, especially in signal processing and control theory. Moreover, in a number of situations, the generating function $f$ is known. Examples are the kernels of the Wiener-Hopf equations (see Gohberg and Fel'dman [7, p. 82]), the spectral density functions in stationary stochastic process (see Grenander and Szegö [9, p. 171]), and the point-spread functions in image deblurring (see Oppenheim [11, p. 200]).

Although direct methods for solving $T_n x = b$ with effort of the order of $n^2$ [14] or even order $n \log^2 n$ [1] exists, rather strong conditions are required to guarantee stability [2]. Thus, recently, Strang proposed using the conjugate gradient (CG) method [12] to solve success in solving $T_n x = b$ by CG in order $n \log n$ work: First, there is a way to perform matrix-vector product $T_n \cdot y$ in order $n \log n$ operations. Second, Strang found a "preconditioning marix" $S_n$, which is circulant, such that the eigenvalue spectrum of $S_n^{-1} T_n$ clusters around 1 and that solving systems of the form $S_n a = b$ requires little work — in this case, order $n \log n$. For an overview of CG and preconditioning, see [8].

Since Strang's work, a number of different circulant preconditioners are proposed and studied (see [3, 4, 5] and the references therein). All these circulant preconditioners assume the generating function $f$ be strictly positive, $f > 0$, and experiments confirm that none of them gives fast convergence when $f$ has zero(s) in $[-\pi, \pi]$.

In this paper, we design band-Toeplitz preconditioner $C_n$ via the polynomial minimax approximation

$$\min_p \max_x \left| \frac{f(x) - p(x)}{f(x)} \right|.$$

When $f > 0$, our band-Toeplitz preconditioners lead to the same convergence rate as those circulant ones can offer; but when $f$ has isolated zeros, say, $f(z_1) = f(z_2) = 0$, our preconditioners still prevail to yield fast convergence. In the latter case, we formulate our minimax approximation as a constrained problem:

$$\min_p \max_x \left| \frac{f(x) - p(x)}{f(x)} \right| \quad \text{subject to } p(z_j) = 0,\ j = 1, 2, \ldots, \text{number of zeros.}$$

The rest of the paper is organized as follows. Section 2 reviews the basic theory of generating functions and Toeplitz system that leads naturally to the constructing preconditioners based on polynomial approximations. A spectral analysis of the preconditioned system in terms of the minimax norm of the approximation problem is also presented. Section 3 presents a Simplex/Remez algorithm that finds the constrained or unconstrained minimax polynomial that defines our preconditioner. Section 4 presents some numerical experiments that confirm our analysis.

## 2 Generating Functions and Preconditioners

Let the $n$-by-$n$ Toeplitz marix $T_n$ be generated by $f$, even, real-valued, and nonegative. We also assume $f$ to be continuous. We list several important facts below. Proofs can be found in Grenander and Szegö [9].

1. Let $f([-\pi,\pi]) = [f_{\min}, f_{\max}]$. Then the set of eigenvalues $\lambda(T_n)$ of $T_n$ satisfies

   $$\lambda(T_n) \subseteq [f_{\min}, f_{\max}].i$$

   Moreover,

   $$\lambda_{\min}(T_n) \to f_{\min} \quad \text{and} \quad \lambda_{\max}(T_n) \to f_{\max} \quad \text{as } n \to \infty.$$

2. Given any vector $u = [u_0, u_1, \ldots, u_{n-1}]^T \in \mathbb{R}^n$,

   $$u^T T_n u = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{k=0}^{n-1} u_k e^{ikx} \right|^2 f(x) dx.$$

Using the second fact, the first can be strengthened. If $f_{\min} < f_{\max}$, then $\lambda(T_n) \subseteq (f_{\min}, f_{\max})$ for all finite $n$. The proof is relatively simple. It suffices to note that, provided $f_{\min} < f_{\max}$, then for any $n$, and any $u \in \mathbb{R}^n$ satisfying $u^T u = 1$,

$$f_{\min} < u^T T_n u < f_{\max}.$$

Thus, as long as $f \geq 0$, $T_n$ is positive definite for all finite $n$.

Because $T_n$ is so much determined by $f$, and that a polynomial generating function

$$p_l(x) = \sum_{k=-(l-1)}^{l-1} b_{|k|} e^{ikx} = b_0 + 2(b_1 \cos x + b_2 \cos 2x + \cdots + b_{l-1} \cos(l-1)x)$$

corresponds to band-Toeplitz matrices, one would expect polynomials $p_l$ that approximates $f$, $p_e \approx f$ would give rise to reasonable preconditioners for $T_n$. The next two theorems confirm this idea.

**Theorem 1** *Let $f$ and $g$ be generating functions for $T_n$ and $C_n$, $n = 1, 2, \ldots$, respectively. If*

$$\left\| \frac{f-g}{f} \right\|_\infty \stackrel{\text{def}}{=} \max_{x\in[-\pi,\pi]} \left| \frac{f(x)-g(x)}{f(x)} \right| = h < 1,$$

*then the eigenvalues $\lambda(C_n^{-1}T_n)$ of $C_n^{-1}T_n$ satisfy*

$$\lambda(C_n^{-1}T_n) \subseteq [1/(1+h), 1/(1-h)].$$

**Proof** By the assumptions,

$$f(x)(1-h) \leq g(x) \leq f(x)(1+h) \qquad \text{for all } x \in [-\pi, \pi],$$

and $g \geq 0$. Thus, from previous discussions, we have $T_n$ and $C_n$ positive definite for all $n = 1, 2, 3, \ldots$.

Now, for a fixed $n$, consider any $u \in \mathbb{R}^n$.

$$(1-h)\, u^T T_n u \leq u^T C_n u = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{k=0}^{n-1} u_k e^{ikx} \right|^2 g(x) dx \leq (1+h)\, u^T T_n u.$$

Because $C_n$ is positive definite,

$$u^T C_n u \leq (1+h) u^T T_n u \qquad \text{for all } u$$

implies

$$\frac{v^T (C_n^{-1/2} T_n C_n^{-1/2}) v}{v^T v} \geq \frac{1}{1+h} \qquad \text{for all } v.$$

Consequently,

$$\lambda_{\min}(C_n^{-1/2} T_n C_n^{-1/2}) \geq 1/(1+h).$$

Observe that

$$\lambda_{\min}(C_n^{-1/2} T_n C_n^{-1/2}) = \lambda_{\min}(C_n^{-1} T_n)$$

and we have

$$\lambda_{\min}(C_n^{-1} T_n) \geq \frac{1}{1+h}.$$

Similar arguments show

$$\lambda_{\max}(C_n^{-1} T_n) \leq \frac{1}{1-h},$$

and the theorem is proved. ∎

**Theorem 2** *Suppose $f$ and $g$ are both even and that $f, g > 0$. If $\|(f-g)/g\|_\infty = h < 1$, with $(f-g)(x_+) = h \cdot f(x_+)$ and $(f-g)(x_-) = -h \cdot f(x_-)$ for some $x_+, x_-$. Then*

$$\lambda_{\min}(C_n^{-1} T_n) \to \frac{1}{1+h} \quad \text{and} \quad \lambda_{\max}(C_n^{-1} T_n) \to \frac{1}{1-h} \quad \textit{as } n \to \infty.$$

**Proof** Given any $\varepsilon > 0$, we will show that $\lambda_{\max}(C_n^{-1}T_n)$ will be within $O(\varepsilon)$ to $\frac{1}{1-h}$ for $n$ sufficiently large.

Consider $x_+ \geq 0$, $g(x_+) = (1-h)f(x_+)$. By continuity, there is $\delta \geq 0$ such that

$$g(x) \leq (1-h+\varepsilon)f(x) \quad \text{for } |x-x_+| \leq \delta.$$

Now, define a continuous function $\eta$ on $[0,\pi]$ such that $0 \leq \eta \leq 1$, $\eta(x_+) = 1$, and that

$$\eta(x) = 0 \quad \text{for } |x-x_+| > \delta.$$

Extend $\eta$ to $[-\pi,\pi]$ by $\eta(-x) = \eta(x)$. Thus, $\eta$ is a generating function for Toeplitz matrices $A_m$ such that

$$\lambda_{\max}(A_m) \to 1 \quad \text{as } m \to \infty.$$

Hence, there exists a large enough $n$ and a vector $n \in \mathbb{R}^n$ such that

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\sum_{k=0}^{n-1} u_k e^{ikx}\right|^2 \eta(x)dx \geq 1-\varepsilon,$$

and

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\sum_{k=0}^{n-1} u_k e^{ikx}\right|^2 dx = 1.$$

Because of the property of $\eta$, we have

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\sum_{k=0}^{n-1} u_k e^{ikx}\right|^2 \eta(x)dx = \frac{1}{2\pi}\int_{|x-x_+|\leq\delta}\left|\sum_{k=0}^{n-1} u_k e^{ikx}\right|^2 \eta(x)dx \geq 1-\varepsilon.$$

Since $\eta \leq 1$, we also have

$$\frac{1}{2\pi}\int_{|x-x_+|\leq\delta}\left|\sum_{k=0}^{n-1} u_k e^{ikx}\right|^2 dx \geq \frac{1}{2\pi}\int_{| \quad |\leq\delta}\left|\sum_{k=0}^{n-1} u_k e^{ikx}\right|^2 \eta(x)dx \geq 1-\varepsilon.$$

Thus

$$\frac{1}{2\pi}\int_{|x-x_+|>\delta}\left|\sum_{k=0}^{n-1} u_k e^{ikx}\right|^2 dx < \varepsilon.$$

Denote the maximum and minimum of $g$ by $g_{\max}$ and $g_{\min}$, respectively. We can now relate $u^T C_n u$ to $u^T T_n u$:

$$\begin{aligned} u^T C_n u &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\sum_{k=0}^{n-1} u_k e^{ikx}\right|^2 g(x)dx \\ &\leq \frac{1}{2\pi}\int_{|x-x_+|\leq\delta}\left|\sum_{k=0}^{n-1} u_k e^{ikx}\right|^2 g(x) + \varepsilon \cdot g_{\max} \end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{2\pi}\int_{|x-x_+|\leq\delta}\left|\sum_{k=0}^{n-1}u_k e^{ikx}\right|^2(1-h+\varepsilon)f(x)dx+\varepsilon g_{\max}\\
&\leq \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\sum_{k=0}^{n-1}u_k e^{ikx}\right|^2(1-h+\varepsilon)f(x)dx+\varepsilon g_{\max}\\
&\leq u^T T_n u\cdot(1-h+\varepsilon)+\varepsilon g_{\max}.
\end{aligned}
$$

Thus,

$$\frac{u^T T_n u}{u^T C_n u}\geq\frac{1}{1-h+\varepsilon}-\frac{\varepsilon\, g_{\max}}{g_{\min}}.$$

Therefore, $\lambda_{\max}(C_n^{-1}T_n)\to 1/(1-h)$ as $n\to\infty$. The proof for $\lambda_{\min}(C_n^{-1}T_n)\to 1/(1+h)$ is similar. ∎

# 3 A Simplex/Remez Algorithm

The previous section suggests that if we wish to use symmetric band-Toeplitz matrices with the first row of the form $[b_0,b_1,\ldots b_{l-1},0,\ldots,0]$, we should seek to solve the following minimax approximation problem,

$$\min_{b_0,b_1,\ldots,b_{l-1}}\max_{x\in[-\pi,\pi]}\left|\frac{f(x)-p_l(x)}{f(x)}\right|,$$

where

$$p_l(x)=b_0+2(b_1\cos x+\ldots+b_{l-1}\cos(l-1)x).$$

Let us first consider the case when $f>0$. Then, the problem is equivalent to

$$\min_{b_0,b_1,\ldots,b_{l-1}}\max_{x\in[-\pi,\pi]}\left|1-q_l(x)\right|,$$

where

$$q_l(x)=\sum_{k=0}^{l-1}b_k\phi_k(x),\quad \phi_0(x)=\frac{1}{f(x)},\ \text{and}\ \phi_k=\frac{2\cos kx}{f(x)}\ \text{for}\ k\geq 1.$$

This problem can be cast as a linear programming like problem:

Minimize $h$

subject to

$$h\geq s(1-q_l(x))\qquad (s,x)\in\{-1,1\}\times[-\pi,\pi].$$

This formulation has a natural dual that can be solved by the Simplex algorithm. This dual is

$$\text{Maximize}\ \sum_{j=1}^{l+1}r(y_j)\cdot c(y_j)\quad y_j=(s_j,x_j)\in\{-1,1\}\times[-\pi,\pi]$$

subject to

$$\mathbf{r} \geq \mathbf{0}$$
$$A\mathbf{r} = [1\,0\,0\,\ldots\,0]^T$$

where the column of the matrix A corresponding to $(s, x) \in \{-1, 1\} \times [-\pi, \pi]$ is

$$[1\, s\phi_1(x)\, s\phi_2(x)\, \ldots\, s\phi_l(x)]^T.$$

Known theory on minimax approximation guarantees existence of a maximum ([6]).

Furthermore, the formulation above is well suited for a Simplex algorithm. In fact, when the underlying approximating family is a Harr system, the Simplex algorithm is exactly the 1-point exchange Remez algorithm (see [13]). For completeness sake, we outline the algorithm.

**Step 0.** *Initialization.* Find an initial feasible basis $B^{(0)} = \{y_1^{(0)}, y_2^{(0)}, \ldots, y_{l+1}^{(0)}\}$ such that the basis marix $A$ is nonsingular and that

$$\mathbf{r}^{(0)} = A^{-1} \cdot [1\,0\,0\,\ldots\,0]^T$$

is nonnegative. Set $i = 1$.

**Step 1.** *Solve for Parameters.* Determine $h^{(i)}$ and $\mathbf{b}^{(i)} = [b_0^{(i)}, b_1^{(i)}, \ldots, b_{l-1}^{(i)}]^T$ such that

$$A^T \begin{bmatrix} h^{(i)} \\ \mathbf{b}^{(i)} \end{bmatrix} = \mathbf{c}^{(i-1)}.$$

Note that this step corresponds to solving for equal alternation in the Remez algorithm because we have

$$1 - \sum_{k=0}^{(i)} b_k^{(i)} \cdot \phi_k(x_j^{(i)}) = s_j^{(i-1)}\, h^{(i)}, \quad j = 1, 2, \ldots, l+1.$$

**Step 2.** *Check for near optimality.* Let

$$M^{(i)} = \max_x \left| 1 - \sum_{k=0}^{l-1} b_k^{(i)} \phi_k(x) \right| = \hat{s} \left( 1 - \sum_{k=0}^{l-1} b_k^{(i)} \phi_k(\hat{x}) \right).$$

If $M^{(i)}$ and $h^{(i)}$ are close enough, stop. Otherwise, go to Step 3.

**Step 3.** *Simplex Exchange.* Apply Simplex exchange on $(\hat{s}, \hat{x})$ with $B^{(i-1)}$ to obtain

$$B^{(i)} = \{y_1^{(i)}, y_2^{(i)}, \ldots, y_{l+1}^{(i)}\}.$$

Increment $i := i + 1$ and go back to Step 1.

It can be shown that, given any $\varepsilon > 0$,

$$0 \leq M^{(i)} - h^{(i)} \leq \varepsilon$$

after a finite number of iterations $i$ (see [13] and also [10]).

Next, we consider the case when $f$ has an isolated zero, say $z$. Because $f \geq 0$, we assume that $f$ is smooth enough that $f'(z) = 0$. To help the numeric algorithm, we explicitly put in the constraints $p_l(z) = p_l'(z) = 0$, that is

$$\begin{aligned} b_0 + 2(b_1 \cos z + \ldots + b_{l-1} \cos(l-1)z) &= 0 \\ b_1 \sin z + 2b_2 \sin 2z + \ldots + (l-1)\sin(l-1)z &= 0 \end{aligned}$$

The linear programming problem can be modified in a straightforward manner.

$$\text{minimize } h$$

subject to

$$\begin{aligned} &h \geq s(1 - q_l(x)) \qquad (s, x) \in \{-1, 1\} \times [-\pi, \pi] \\ &0 \geq p_l(z), \quad 0 \geq -p_l(z) \\ &0 \geq p_l'(z), \quad 0 \geq -p_l'(z). \end{aligned}$$

The corresponding dual becomes

$$\text{Maximize } \sum_{j=1}^{l+1} r(y_j) \cdot c(y_j)$$

subject to

$$\begin{aligned} &\mathbf{r} \geq \mathbf{0} \\ &A\mathbf{r} = [1\,0\,0\,\ldots\,0]^T \end{aligned}$$

where, now, $y_j \in -1, 1 \times [-\pi, \pi]$ (the domain), or $y_j$ corresponds to one of the four constraints. For $y_j = (s, x) \in -1, 1 \times [-\pi, \pi]$ we have $c(y_j) = s$ as before. But for $y_j$ corresponding to a constraint $\gamma(z) \geq 0$, say, $\gamma(z) = -p_l'(z)$, the cost coefficient is zero (regardless of the specific constraint) and the corresponding column of A is given by

$$\left[\frac{\partial \gamma}{\partial h}\ \frac{\partial \gamma}{\partial b_0}\ \cdots\ \frac{\partial \gamma}{\partial b_{l-1}}\right]^T.$$

Clearly, the Simplex algorithm can be handle the constrained problem.

**Step 0.** *Initialization.* Find an initial feasible basis $B^{(0)} = \{y_1^{(0)}, y_2^{(0)}, \ldots, y_{l+1}^{(0)}\}$ such that the basis marix $A$ is nonsingular and that

$$\mathbf{r}^{(0)} = A^{-1} \cdot [1\,0\,0\,\ldots\,0]^T$$

is nonnegative. Set $i = 1$.

**Step 1.** *Solve for Parameters.* Determine $h^{(i)}$ and $\mathbf{b}^{(i)} = [b_0^{(i)}, b_1^{(i)}, \ldots, b_{l-1}^{(i)}]^T$ such that

$$A^T \begin{bmatrix} h^{(i)} \\ \mathbf{b}^{(i)} \end{bmatrix} = \mathbf{c}^{(i-1)}.$$

Note that this step corresponds to solving for equal alternation and possibly for constraints.

**Step 2.** *Check for near optimality.* Let

$$M^{(i)} = \max_x \left| 1 - \sum_{k=0}^{l-1} b_k^{(i)} \phi_k(x) \right| = \hat{s} \left( 1 - \sum_{k=0}^{l-1} b_k^{(i)} \phi_k(\hat{x}) \right).$$

If all the constraints are satisfied and if $M^{(i)}$ and $h^{(i)}$ are close enough, stop. Otherwise, go to Step 3.

**Step 3.** *Simplex Exchange.* If a particular constraint is not satisfied, apply Simplex exchange on $B^{(i-1)}$ with the violated constraint. Otherwise, apply Simplex exchange on $(\hat{s}, \hat{x})$ with $B^{(i-1)}$. After the exchange, we have

$$B^{(i)} = \{y_1^{(i)}, y_2^{(i)}, \ldots, y_{l+1}^{(i)}\}.$$

Increment $i := i + 1$ and go back to Step 1.

# 4 Numerical Experiments

We present experiments of preconditioning on Toeplitz matrices $T_n$ of various sizez $n$ generated by five different functions. For each matrix, we use band-Toeplitz preconditioners $C_{n,l}$ of various half bandwidth $l$. In each iteration of the preconditioned conjugate gradient method, we have to compute matrix vector multiplication of the form $T_n x$ and solution of linear system $C_{n,l} u = v$. The matrix vector product $T_n x$ can be computed in order $n \log n$ operations by embedding $T_n$ into a $2n$-by-$2n$ circulant matrix and then using Fast Fourier Transform (see Strang [12]). The solution of $C_{n,l} u = v$ can be obtained by using efficient band solvers (see Golub and Van Loan [8] or Wright [15]). Typically, we will decompose $C_{n,l}$ into some triangular factors and then use backward and forward solves. The cost of obtaining the triangular factors is of the order $nl^2$, and each subsequent solve costs order $nl$ as the triangular factors will also be banded.

We compare the convergence rate of the band-Toeplitz preconditioner with circulant preconditioner on five different generating functions. They are $\cosh x$, $x^4 + 1$, $1 - e^{-x^2}$, $(x-1)^2(x+1)^2$ and $x^4$. The first two functions are positive while the others have either a single or double zero. The matrices $T_n$ are formed by evaluating the Fourier coefficients of the generating functions.

We note that when $f(x) = 1 - e^{-x^2}$, its Fourier coefficients cannot be evaluated exactly. In this case, we approximate them by

$$\begin{aligned} a_j &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ijx}dx \\ &\approx \frac{1}{2n}\sum_{k=0}^{2n-1} f(\frac{k\pi}{n}-\pi)e^{-ij(k\pi/n-\pi)}, \quad j = 0, \pm 1, \pm 2, \cdots \end{aligned}$$

where the last expression is evaluated by using the Fast Fourier Transform.

In our tests, the vector of all ones is the right hand side vector, the zero vector is the initial guess and the stopping criterion is $||r_q||_2/||r_0||_2 \le 10^{-7}$, where $r_q$ is the residual vector after $q$ iterations. All computations are done by Matlab on a Sun workstation. Tables 1-5 show the numbers of iterations required for convergence with different choices of preconditioners. In the tables, $I$ denotes no preconditioner is used, $C$ is the T. Chan circulant preconditioner [5], and $C_{n,l}$ is the band-Toeplitz preconditioner with half-bandwidth $l$.

We note that for $f$ that are positive, our preconditioner, with half-bandwidth 5, performs at the same rate as the circulant preconditioner. In the cases where $f$ has zeros, our preconditioned systems still converge at a rate that is independent of the size of the matrices. For the circulant preconditioned systems, however, the number of iterations required grows as the size of the matrix increases.

| $n$ | $I$ | $C$ | $B_{n,2}$ | $B_{n,3}$ | $B_{n,4}$ | $B_{n,5}$ |
|---|---|---|---|---|---|---|
| 16 | 9 | 6 | 9 | 7 | 6 | 5 |
| 32 | 16 | 6 | 10 | 7 | 6 | 6 |
| 64 | 21 | 5 | 11 | 8 | 6 | 6 |
| 128 | 23 | 5 | 10 | 8 | 6 | 6 |
| 256 | 24 | 5 | 10 | 7 | 6 | 6 |

Table 1. Numbers of Iterations for $f(x) = \cosh x$.

| $n$ | $I$ | $C$ | $B_{n,2}$ | $B_{n,3}$ | $B_{n,4}$ | $B_{n,5}$ |
|---|---|---|---|---|---|---|
| 16 | 10 | 9 | 9 | 8 | 8 | 7 |
| 32 | 22 | 7 | 16 | 11 | 8 | 7 |
| 64 | 37 | 7 | 22 | 12 | 8 | 7 |
| 128 | 56 | 6 | 25 | 12 | 8 | 7 |
| 256 | 67 | 6 | 26 | 12 | 8 | 7 |

Table 2. Numbers of Iterations for $f(x) = x^4 + 1$.

| $n$ | $I$ | $C$ | $B_{n,2}$ | $B_{n,3}$ | $B_{n,4}$ | $B_{n,5}$ |
|---|---|---|---|---|---|---|
| 16 | 9 | 6 | 9 | 7 | 4 | 3 |
| 32 | 14 | 7 | 15 | 7 | 5 | 3 |
| 64 | 24 | 8 | 17 | 8 | 5 | 3 |
| 128 | 42 | 10 | 17 | 8 | 5 | 3 |
| 256 | 77 | 13 | 17 | 8 | 5 | 3 |

Table 3. Numbers of Iterations for $f(x) = 1 - e^{-x^2}$.

| $n$ | $I$ | $C$ | $B_{n,3}$ | $B_{n,4}$ | $B_{n,5}$ | $B_{n,6}$ |
|---|---|---|---|---|---|---|
| 16 | 11 | 9 | 9 | 9 | 8 | 7 |
| 32 | 27 | 14 | 13 | 11 | 9 | 7 |
| 64 | 74 | 17 | 16 | 11 | 8 | 7 |
| 128 | 193 | 22 | 18 | 11 | 8 | 7 |
| 256 | 465 | 28 | 19 | 11 | 8 | 7 |

Table 4. Numbers of Iterations for $f(x) = (x-1)^2(x+1)^2$.

| $n$ | $I$ | $C$ | $B_{n,3}$ | $B_{n,4}$ | $B_{n,5}$ | $B_{n,6}$ |
|---|---|---|---|---|---|---|
| 16 | 12 | 10 | 9 | 9 | 9 | 7 |
| 32 | 34 | 16 | 15 | 10 | 11 | 9 |
| 64 | 119 | 26 | 21 | 13 | 11 | 9 |
| 128 | 587 | 77 | 24 | 15 | 12 | 10 |
| 256 | $> 1000$ | 179 | 27 | 16 | 12 | 10 |

Table 5. Numbers of Iterations for $f(x) = x^4$.

## 5 Concluding Remarks

By understanding Toeplitz preconditioner from the point of view of minimax approximation of the corresponding generating functions, we can construct band-Toeplitz preconditioners that offer fast convergence rate even when the matrix to be preconditioned has a generating function with a zero. Moreover, our preconditioner with modest bandwidth is also an excellent choice for $f$ without zero. We emphasize that for a given $f$, the entries of the preconditioners is unchanged as $n$ increases. Thus, we need to invoke the Remez algorithm once for each $f$. We note moreover that the Cholesky factors of $C_{n,l}$ can be used to build the Cholesky factors of $C_{n+1,l}$. That can reduce the cost of factorization of the band-Toeplitz preconditioner.

## 6 Acknowledgment

This work of the second author was supported by the Applied Mathematical Sciences subprogram of the Office of Energy Research, U. S. Department of Energy, under Contract W-31-109-Eng-38.

## References

[1] G. Ammar and W. Gragg, *Superfast Solution of Real Positive Definite Toeplitz Systems*, SIAM J. Matrix Appl., V9 (1988), pp. 61–76.

[2] J. Bunch, *Stability of Methods for Solving Toeplitz Systems of Equations*, SIAM J. Sci. Statist. Comput., V6 (1985), pp. 349–364.

[3] R. Chan and M. Yeung, *Circulant Preconditioners for Toeplitz Matrices with Positive Continuous Generating Functions*, Math. Comp., V58 (1992), pp. 233-240.

[4] R. Chan and K. Ng, *Fast Iterative Solvers for Toeplitz-plus-Band Systems*, to appear in SIAM. J. Sci. Stat. Comp.

[5] T. Chan, *An Optimal Circulant Preconditioner for Toeplitz Systems*, SIAM J. Sci. Statist. Comp., V9 (1988), pp. 766–771.

[6] V. F. Demjanov and V. N. Malozemov, *Introduction to Minimax*, John Wiley, New York, 1974.

[7] I. Gohberg and I. Fel'dman, *Convolution Equations and Projection Methods for Their Solutions*, Volume 41, Transaction of Mathematical Monographs, American Mathematical Society, Rhode Island, 1974.

[8] G. Golub and C. Van Loan, *Matrix Computations*, 2nd Ed., The Johns Hopkins University Press, Maryland, 1989.

[9] U. Grenander and G. Szegö, *Toeplitz Form and Its Applications*. 2nd Ed., Chelsea Pub. Co., New York, 1984.

[10] P. Laurent and C. Carasso, *An Algorithm of Successive Minimization in Convex Programming*, R.A.I.R.O. Analyse Numerique, V12 (1978), pp.377–400.

[11] A. Oppenheim, *Applications of Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, New Jersey, 1978.

[12] G. Strang, *A Proposal for Toeplitz Matrix Calculations*, Studies in App. Math., V74 (1986), pp. 171–176.

[13] P. T. P. Tang, *A Fast Algorithm for Linear Complex Chebyshev Approximation*, Math. of Comp., V51 (1988), pp. 721–739.

[14] W. Trench, *An Algorithm for the Inversion of Finite Toeplitz Matrices*, SIAM J. Appl. Math., V12 (1964), pp. 515–522.

[15] Stephen J. Wright, *Parallel Algorithms for Banded Linear Systems*, SIAM J. Sci. Stat. Comp., V12 (1991), pp. 824–842.

ææææææææææææææ