

RECURSIVE-BASED PCG METHODS FOR TOEPLITZ SYSTEMS WITH NONNEGATIVE GENERATING FUNCTIONS*

MICHAEL K. NG[†], HAI-WEI SUN[‡], AND XIAO-QING JIN[§]

*Dedicated to Professor Raymond Chan
for his inspired guidance in studying Toeplitz matrices*

Abstract. In this paper, we consider the solutions of symmetric positive definite, but ill-conditioned, Toeplitz systems $A_n x = b$. Here we propose to solve the system by the recursive-based preconditioned conjugate gradient method. The idea is to use the inverse of A_m (the principal submatrix of A_n) with the *Gohberg–Semencul formula* as a preconditioner for A_n . The inverse of A_m can be generated recursively by using the formula until m is small enough. The construction of the preconditioners requires only the entries of A_n and does not require the explicit knowledge of the generating function f of A_n . We show that if f is a nonnegative, bounded, and piecewise continuous even function with a finite number of zeros of even order, the spectra of the preconditioned matrices are uniformly bounded except for a fixed number of outliers. Hence the conjugate gradient method, when applied to solving the preconditioned system, converges very quickly. Numerical results are included to illustrate the effectiveness of our approach.

Key words. Toeplitz matrices, Gohberg–Semencul formula, recursive-based method, preconditioned conjugate gradient method, preconditioners

AMS subject classifications. 65F10, 65F15, 65T10

PII. S1064827500378155

1. Introduction. An n -by- n matrix A_n with entries a_{ij} is said to be Toeplitz if $a_{ij} = a_{i-j}$. Toeplitz systems of the form $A_n y = b$ occur in a variety of applications in mathematics and engineering [10]. In this paper, we consider the solution of symmetric positive definite Toeplitz systems. There are a number of specialized fast direct methods for solving such systems in $O(n^2)$ operations. The original references for these algorithms are Levinson [27], Durbin [15], and Trench [41]. Superfast algorithms of complexity $O(n \log^2 n)$ operations for Toeplitz systems were proposed by different groups of researchers, for instance, Bitmead and Anderson [5], Brent, Gustavson, and Yun [6], Morf [28], de Hoog [21], Ammar and Gragg [1], and Huckle [24]. The key to these fast and superfast direct methods is to solve the Toeplitz system recursively.

1.1. Fast direct Toeplitz solvers. In the Levinson–Durbin method, we begin with the solution of the 1-by-1 system and then increase the order, using the lower-

*Received by the editors September 15, 2000; accepted for publication (in revised form) September 2, 2002; published electronically April 1, 2003.

<http://www.siam.org/journals/sisc/24-5/37815.html>

[†]Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong (mng@maths.hku.hk). The research of this author was supported in part by Hong Kong Research Grants Council grants HKU 7132/00P and HKU 7130/02P, and HKU CRCG grants 10202720, 10203501, and 10203907.

[‡]Department of Applied Mathematics, Guangdong University of Technology, Guangzhou 510090, People's Republic of China. The research of this author was supported by Guangdong Provincial Natural Science Foundation of China grant 974176, Natural Science Foundation of China grant 19971018, and HKRGC grant CUHK 4207/97P.

[§]Faculty of Science and Technology, University of Macau, Macau, China. The research of this author was supported by UM Research grants RG009/98-99S/JXQ/FST and RG010/99-00S/JXQ/FST.

order solution to obtain the next higher-order solution recursively. We try to obtain the so-called *inverse Choleski factorization* of A_n :

$$(1.1) \quad R_n^t A_n R_n = D_n,$$

where R_n is unit upper triangular and D_n is diagonal. This is done via a block UDL decomposition of A_m^{-1} at each iteration:

$$(1.2) \quad A_m^{-1} = \begin{bmatrix} I_{m-1} & x_{m-1} \\ 0^t & 1 \end{bmatrix} \begin{bmatrix} A_{m-1}^{-1} & 0 \\ 0^t & (\delta^{(m-1)})^{-1} \end{bmatrix} \begin{bmatrix} I_{m-1} & 0 \\ x_{m-1}^t & 1 \end{bmatrix}, \quad m = 2, \dots, n.$$

The vector x_{m-1} and the scalar $\delta^{(m-1)}$ are called the *Szegö vector* and *Schur parameter*, respectively. Once we have R_n and D_n , the solution of the system $A_n y = b$ is given by $R_n D_n^{-1} R_n^t b$, which can be computed easily by using the recurrence relation.

We note that the fast Toeplitz solver can be divided into two phases: the factorization phase to obtain the inverse Choleski factors R_n and D_n of A_n , and the solution phase to find $A_n^{-1} b$. Both phases require $O(n^2)$ operations. However, once we have the factors, the solution phase can be reduced to $O(n \log n)$ operations by using the celebrated *Gohberg–Semencul formula* [17]:

$$(1.3) \quad A_m^{-1} = \frac{1}{\delta^{(m-1)}} [L_m^{(1)} (L_m^{(1)})^t - L_m^{(2)} (L_m^{(2)})^t], \quad m = 2, \dots, n,$$

where $\delta^{(m-1)}$ is described in (1.2), and $L_m^{(1)}$ and $L_m^{(2)}$ are lower triangular Toeplitz matrices given by

$$L_m^{(1)} = \begin{bmatrix} l_1 & 0 & \cdots & 0 & 0 \\ l_2 & l_1 & 0 & & 0 \\ \vdots & l_2 & l_1 & \ddots & \vdots \\ l_{m-1} & & \ddots & \ddots & 0 \\ l_m & l_{m-1} & \cdots & l_2 & l_1 \end{bmatrix} \quad \text{and} \quad L_m^{(2)} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ l_m & 0 & 0 & & 0 \\ \vdots & l_m & 0 & \ddots & \vdots \\ l_3 & & \ddots & \ddots & 0 \\ l_2 & l_3 & \cdots & l_m & 0 \end{bmatrix},$$

respectively, with

$$x_m = [l_1, l_2, \dots, l_m]^t.$$

We note that $\delta^{(m-1)} = l_1$, and, since A_n is symmetric positive definite, we must have $\delta^{(m-1)} > 0$; see [19]. The Gohberg–Semencul formula provides an explicit representation of A_m^{-1} in terms of its first column. We wish to find x_m by solving

$$(1.4) \quad A_m x_m = e_m,$$

where

$$e_m = [1, 0, \dots, 0]^t$$

is the first unit m -vector. From (1.3), we see that $A_n^{-1} b$ can be obtained via Toeplitz matrix-vector multiplications, i.e., the use of fast Fourier transforms (FFTs), in $O(n \log n)$ operations; see, for instance, [38]. We replace the lower and upper tri-

angular factors in (1.3) with circulant and skew-circulant matrices and thus further reduce the cost of the solution phase; see [2] or section 4.

In [1], Ammar and Gragg used the generalized Schur algorithm (a doubling procedure) to compute x_m for A_m^{-1} , where $m = 2, 4, \dots, n/2$, and use the Gohberg–Semencul formula to obtain $A_n^{-1}b$. A detail implementation in [1] shows that the cost is $O(n \log^2 n)$ operations. In [28, 5, 24], their superfast methods are based on repeatedly dividing the original problem into two subproblems half the size, namely the leading principal submatrix and the related Schur complement. All occurring matrices are represented explicitly by proper generating vectors of their displacement characterization. Recently, Huckle [24] showed that the resulting superfast method takes about $81.25n \log^2 n + O(n \log n)$ operations.

1.2. Outline. The method of the superfast Toeplitz solver can be viewed as an approach to use $A_{m/2}^{-1}$ to generate A_m^{-1} . In this paper, we will consider our preconditioners for A_m from the same viewpoint. The main aim of this paper is to construct our preconditioners for A_m by using $A_{m/2}^{-1}$ with the Gohberg–Semencul formula (cf. (1.3)). The inverse of A_m can be generated recursively by using the formula until m is small enough. Finally, $A_n^{-1}b$ can be obtained again by using the Gohberg–Semencul formula. We remark that the construction of the preconditioners requires only the entries of A_m and does not require explicit knowledge of the generating function f of A_m . We show that if f is a nonnegative, bounded, and piecewise continuous even function with a finite number of zeros of even order, the spectra of the preconditioned matrices are uniformly bounded except for a fixed number of outliers. Hence the conjugate gradient method, when applied to solving these m -by- m preconditioned Toeplitz systems, converges in $O(\log m)$ iterations. Using this result, we will show that the complexity of solving such n -by- n Toeplitz systems is of $O(n \log^2 n)$ operations. In contrast, the nonpreconditioned systems will have condition numbers growing like $O(n^{2p_{\max}})$, where p_{\max} is the maximal order of the zeros of f , and hence the complexity for solving the systems will be of order $O(n^{p_{\max}+1} \log n)$.

For the case where the generating function f is strictly positive and continuous, we will show that the conjugate gradient method, when applied to solving these preconditioned Toeplitz systems, converges superlinearly. Therefore, the complexity of solving such n -by- n Toeplitz systems is of $O(n \log n)$ operations.

The outline of this paper is as follows. In section 2, we review the definition of Toeplitz matrices and describe our recursive-based preconditioned conjugate gradient (PCG) method by using the Gohberg–Semencul formula. In section 3, we analyze the spectra of the preconditioned matrices. In section 4, we compare our method with other iterative methods. Numerical results are given in section 5 to illustrate the effectiveness of our approach. Finally, some concluding remarks are given.

2. Recursive-based method (RBM). To begin with, let \mathcal{C}^+ be the set of all nonnegative, bounded, and piecewise continuous even functions defined on $[-\pi, \pi]$. The Fourier coefficients of a function f in \mathcal{C}^+ are given by

$$a_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) e^{-ik\theta} d\theta, \quad k = 0, \pm 1, \pm 2, \dots$$

Clearly $a_k = a_{-k}$ for all k . Let A_n be the n -by- n symmetric Toeplitz matrix with the (i, j) th entry given by a_{i-j} , $i, j = 0, \dots, n-1$. The function f is called the generating function of the matrices A_n . We say that θ_0 is a zero of f of order ν if $f(\theta_0) = 0$ and

ν is the smallest positive integer such that $f^{(\nu)}(\theta_0) \neq 0$ and $f^{(\nu+1)}(\theta)$ is continuous in a neighborhood of θ_0 . By Taylor's theorem,

$$f(\theta) = \frac{f^{(\nu)}(\theta_0)}{\nu!}(\theta - \theta_0)^\nu + O((\theta - \theta_0)^{\nu+1})$$

for all θ in that neighborhood. If f is nonnegative, $f^{(\nu)}(\theta_0) > 0$ and ν must be even.

In the following, we denote the essential infimum and the essential supremum of f by f_{\min} and f_{\max} , respectively. From the assumptions, we see that $f_{\max} \neq f_{\min}$. Then by using the fact that

$$(2.1) \quad z^t A_n z = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \sum_{j=1}^n z_j e^{i(j-1)\theta} \right|^2 f(\theta) d\theta$$

for any $f \in \mathcal{C}^+$ and any n -vector $z = (z_1, \dots, z_n)^t$, Chan [7, Lemma 1] proved that

$$(2.2) \quad \lambda_{\min}(A_n) > f_{\min}.$$

Here $\lambda_{\min}(A_n)$ is the smallest eigenvalue of A_n . Since f is nonnegative and is not identically zero, A_n is positive definite for all n .

In this paper, we consider that f attains f_{\min} at finitely many points in $[-\pi, \pi]$ and that f is smooth around these points. More precisely, we assume that $f(\theta) - f_{\min}$ has finitely many zeros in $[-\pi, \pi]$ and that the orders p_j of these zeros are finite and positive. Notice that the matrix A_n is unchanged when f is redefined at finitely many points. Thus we can always assume without loss of generality that f is continuous at those minimum points. We remark that the condition number of A_n generated by such an f grows like $O(n^{2p_{\max}})$, where $p_{\max} = \max_j p_j$; see [32]. However, if $f_{\min} > 0$, then the condition number of A_n generated by such an f is bounded. The systems $A_n y = b$ will be solved by the RBM which has employed the PCG method with A_m^{-1} as preconditioners.

Our idea is to partition a Toeplitz matrix A_m as the following form:

$$(2.3) \quad A_m = \begin{bmatrix} A_{m'} & T_{m', m-m'} \\ T_{m', m-m'}^t & A_{m-m'} \end{bmatrix},$$

where $A_{m'}$ and $A_{m-m'}$ are principal submatrices of A_m and $T_{m', m-m'} = (a_{m'+(l-k)})_{k,l}$ is also an m' -by- $(m - m')$ Toeplitz matrix. We propose to use

$$(2.4) \quad R_m = \begin{bmatrix} A_{m'} & 0 \\ 0 & A_{m-m'} \end{bmatrix}$$

as the preconditioner for A_m with $m > m'$.

Given a Toeplitz system $A_n x_n = e_n$, we define a sequence of subsystems on different levels:

$$A_m x_m = e_m, \quad l \leq m \leq n,$$

where e_m is the first unit m -vector and l is the smallest size of the subsystem. The matrix is inverted exactly at the level l . If A_n is symmetric positive definite, so are $A_{m'}$, $A_{m-m'}$, A_m^{-1} , and $A_{m-m'}^{-1}$. The procedure of our RBM is given as follows:

PROCEDURE **RBM-I** (A_m, m, m').

If $m \leq l$, then $x_m := A_m^{-1}e_m$;

else

(i) construct $A_{m'}^{-1}$ by $x_{m'}$ via the Gohberg-Semencul formula (1.3); $x_{m'} := \mathbf{RBM-I}$ ($A_{m'}, m', m'_1$), where m'_1 and $m' - m'_1$ are the sizes of Toeplitz systems in the next level.

(ii) construct $A_{m-m'}^{-1}$ by $x_{m-m'}$ via the Gohberg-Semencul formula (1.3); $x_{m-m'} := \mathbf{RBM-I}$ ($A_{m-m'}, m-m', m'_2$), where m'_2 and $m-m'-m'_2$ are the sizes of Toeplitz systems in the next level.

(iii) solve the linear system $A_m x_m = e_m$ by using the PCB method with R_m defined as in (2.4) as the preconditioner.

After we obtain the vector x_n , we see from (1.3) that the solution $A_n^{-1}b$ can be obtained via Toeplitz matrix-vector multiplications in $O(n \log n)$ operations. The advantage of the RBM is that it not only solves n -by- n Toeplitz systems $A_n x_n = e_n$, but it also solves other m -by- m Toeplitz systems $A_m x_m = e_m$. We remark that some signal processing applications [22, 39] require us to solve different sizes of Toeplitz systems to obtain the optimal filter.

Below we give a more efficient implementation of the RBM when $m' = m/2$. Let

$$J_m = \begin{bmatrix} 0 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \end{bmatrix}$$

be the m -by- m anti-identity matrix. We note that

$$J_{m/2} A_{m/2} J_{m/2} = A_{m/2}, \quad J_{m/2} T_{m/2}^t = T_{m/2} J_{m/2} = H_{m/2},$$

and here $T_{m/2} = T_{m/2, m/2}$ in (2.3) and $H_{m/2}$ is a Hankel matrix. It follows that

$$(2.5) \quad \begin{bmatrix} I_{m/2} & 0 \\ 0 & J_{m/2} \end{bmatrix} \begin{bmatrix} A_{m/2} & T_{m/2} \\ T_{m/2}^t & A_{m/2} \end{bmatrix} \begin{bmatrix} I_{m/2} & 0 \\ 0 & J_{m/2} \end{bmatrix} = \begin{bmatrix} A_{m/2} & H_{m/2} \\ H_{m/2} & A_{m/2} \end{bmatrix}.$$

Next we transform the matrix in the right-hand side of (2.5) into two Toeplitz-plus-Hankel matrices:

$$(2.6) \quad \begin{aligned} & \frac{1}{2} \begin{bmatrix} I_{m/2} & I_{m/2} \\ I_{m/2} & -I_{m/2} \end{bmatrix} \begin{bmatrix} A_{m/2} & H_{m/2} \\ H_{m/2} & A_{m/2} \end{bmatrix} \begin{bmatrix} I_{m/2} & I_{m/2} \\ I_{m/2} & -I_{m/2} \end{bmatrix} \\ & = \begin{bmatrix} A_{m/2} + H_{m/2} & 0 \\ 0 & A_{m/2} - H_{m/2} \end{bmatrix}. \end{aligned}$$

From (2.6), the solution of the linear system $A_m x_m = e_m$ can be found by solving

$$(2.7) \quad (A_{m/2} + H_{m/2})y_{m/2}^{(1)} = e_{m/2}$$

and

$$(2.8) \quad (A_{m/2} - H_{m/2})y_{m/2}^{(2)} = e_{m/2},$$

respectively. We recall that $e_{m/2} = [1, 0, \dots, 0]^t$ is the first unit $m/2$ -vector. By using the transformation

$$(2.9) \quad x_m = \begin{bmatrix} I_{m/2} & 0 \\ 0 & J_{m/2} \end{bmatrix} \begin{bmatrix} I_{m/2} & I_{m/2} \\ I_{m/2} & -I_{m/2} \end{bmatrix} \begin{bmatrix} y_{m/2}^{(1)} \\ y_{m/2}^{(2)} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} y_{m/2}^{(1)} + y_{m/2}^{(2)} \\ J_{m/2}(y_{m/2}^{(1)} - y_{m/2}^{(2)}) \end{bmatrix},$$

the linear system $A_m x_m = e_m$ can be solved. Therefore, we need only to determine $y_{m/2}^{(1)}$ and $y_{m/2}^{(2)}$ in (2.7) and (2.8), respectively. We note that

$$\begin{bmatrix} I_{m/2} & I_{m/2} \\ I_{m/2} & -I_{m/2} \end{bmatrix}^{-1} = \frac{1}{2} \begin{bmatrix} I_{m/2} & I_{m/2} \\ I_{m/2} & -I_{m/2} \end{bmatrix}$$

and

$$\begin{bmatrix} I_{m/2} & 0 \\ 0 & J_{m/2} \end{bmatrix}^{-1} = \begin{bmatrix} I_{m/2} & 0 \\ 0 & J_{m/2} \end{bmatrix}.$$

Hence the matrices $A_{m/2} \pm H_{m/2}$ are both positive definite. The procedure of our **RBM-I** is rewritten as follows.

PROCEDURE RBM-II (A_m, m).

if $m \leq l$, then $x_m := A_m^{-1} e_m$;

else

(i) construct $A_{m/2}^{-1}$ by $x_{m/2}$ via the Gohberg-Semencul formula (1.3);

$x_{m/2} := \mathbf{RBM-II}$ ($A_{m/2}, m/2$);

(ii) solve linear systems $(A_{m/2} + H_{m/2})y_{m/2}^{(1)} = e_{m/2}$ and

$(A_{m/2} - H_{m/2})y_{m/2}^{(2)} = e_{m/2}$ by using the PCG method with $A_{m/2}^{-1}$ as the preconditioner.

(iii) $x_m := [y_{m/2}^{(1)} + y_{m/2}^{(2)}; J_{m/2}(y_{m/2}^{(1)} - y_{m/2}^{(2)})]/2$.

In the next section, we analyze the performance of our proposed preconditioners R_m .

3. The spectra of the preconditioned matrices. In this section, we study the spectra of the preconditioned matrices $R_m^{-1} A_m$ on different levels. We will need the following theorem.

THEOREM 3.1. *Let B_m be an m -by- m symmetric matrix and C_m be an m -by- m symmetric positive definite matrix. Let the eigenvalues λ_j of $C_m^{-1} B_m$ be arranged in increasing order, i.e., $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_m$. Suppose \mathcal{S}_k is a given k -dimensional subspace of \mathbb{R}^m , where $1 \leq k \leq m$. If there exists a constant c_1 independent of m such that*

$$\frac{x^t B_m x}{x^t C_m x} \geq c_1 \quad \forall x \in \mathcal{S}_k,$$

then

$$\lambda_m \geq \lambda_{m-1} \geq \dots \geq \lambda_{m-k+1} \geq c_1.$$

Moreover, if there exists a constant c_2 such that

$$\frac{x^t B_m x}{x^t C_m x} \leq c_2 \quad \forall x \in \mathcal{S}_k,$$

then

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_k \leq c_2.$$

Proof. The theorem can be proved by using Theorem 4.3.21 in [26, p. 191]. \square

Using the above theorem, we can derive the following corollary.

COROLLARY 3.2. *Let B_m be an m -by- m symmetric matrix and C_m be an m -by- m symmetric positive definite matrix. Suppose Ω_p is an $(m - p)$ -dimensional subspace of \mathbb{R}^m , where $(1 \leq p \leq m/2)$. If there exists a constant $\gamma > 0$ such that*

$$(3.1) \quad \left| \frac{x^t B_m x}{x^t C_m x} \right| \leq \gamma \quad \forall x \in \Omega_p,$$

then at least $m - 2p$ eigenvalues of $C_m^{-1} B_m$ lie in the interval $[-\gamma, \gamma]$.

Proof. Let $c_1 = -\gamma$ and $c_2 = \gamma$ in Theorem 3.1, respectively; then we have

$$-\gamma \leq \lambda_{p+1} \leq \dots \leq \lambda_{m-p} \leq \gamma$$

by using the notation in Theorem 3.1. \square

3.1. The spectra of $R_m^{-1} A_m$. Next we have our main theorem which states that the spectra of the preconditioned matrices are essentially bounded if f is a function in \mathcal{C}^+ with multiple zeros of even order. We recall that f is not necessary to be a continuous function on $[-\pi, \pi]$.

THEOREM 3.3. *Let $f \in \mathcal{C}^+$ and have μ zeros of order $2p_j$ at $\theta_j \in [-\pi, \pi]$ for $1 \leq j \leq \mu$. Let $p = \sum_{j=1}^{\mu} p_j$. Then there exist two positive constants α, β ($\alpha < \beta$), independent of m , such that for all $m > p + 1$, at most $p + 1$ eigenvalues of $R_m^{-1} A_m$ are outside the interval $[\alpha, \beta]$.*

Proof. By definition of zeros (see section 2), such f can be written in the general form

$$(3.2) \quad f(\theta) = \prod_{j=1}^{\mu} (\theta - \theta_j)^{2p_j} h(\theta),$$

where h is a strictly positive function in \mathcal{C}^+ and continuous at $\{\theta_j\}_{j=1}^{\mu}$. It follows that there exist two positive constants γ_1 and γ_2 (independent of m) such that

$$\gamma_1 \leq \frac{f(\theta)}{g_p(\theta)} \leq \gamma_2 \quad \forall \theta \in [-\pi, \pi],$$

where

$$g_p(\theta) = \prod_{j=1}^{\mu} \sin^{2p_j} \left(\frac{\theta - \theta_j}{2} \right);$$

see [9].

Let $z = (z_0, z_1, \dots, z_{m'-1}, z_{m'}, \dots, z_{m-1})^t \in \mathbb{R}$. We denote

$$q_1(\theta) = \sum_{k=0}^{m'-1} z_k e^{ik\theta}$$

and

$$q_2(\theta) = \sum_{k=m'}^{m-1} z_k e^{ik\theta} = e^{im'\theta} \sum_{k=0}^{m-m'-1} z_{m'+k} e^{ik\theta}.$$

We can get from [7, Theorem 1] that

$$z^t A_m z = \int_{-\pi}^{\pi} f(\theta) |q_1(\theta) + q_2(\theta)|^2 d\theta$$

and

$$z^t R_m z = \int_{-\pi}^{\pi} f(\theta) |q_1(\theta)|^2 d\theta + \int_{-\pi}^{\pi} f(\theta) |q_2(\theta)|^2 d\theta.$$

Therefore, we have

$$(3.3) \quad z^t A_m z = z^t R_m z + \int_{-\pi}^{\pi} f(\theta) q_1(\theta) \overline{q_2(\theta)} d\theta + \int_{-\pi}^{\pi} f(\theta) q_2(\theta) \overline{q_1(\theta)} d\theta.$$

Next we study the spectrum of $R_m^{-1} A_m$ or $I_m + R_m^{-1}(A_m - R_m)$.

For simplicity, we let $s = \lceil p/2 \rceil$. We assume

$$\Omega_{2s} = \{z = (z_0, z_1, \dots, z_{m-1})^t \in \mathbb{R}^m \mid z_k = 0, k = m' - s, m' - s + 1, \dots, m' + s - 1\}.$$

We see that Ω_{2s} is a subspace of \mathbb{R}^m with the dimension $m - 2s$. For $z \in \Omega_{2s}$, we obtain

$$(3.4) \quad \int_{-\pi}^{\pi} f(\theta) q_1(\theta) \overline{q_2(\theta)} d\theta = \int_{-\pi}^{\pi} f(\theta) e^{-i(2s+1)\theta} \left(\sum_{k=0}^{m'-s-1} z_k e^{-i(m'-s-1+k)\theta} \right) \cdot \left(\sum_{k=0}^{m-m'-s-1} z_{k+m'+s} e^{-ik\theta} \right) d\theta$$

and

$$(3.5) \quad \int_{-\pi}^{\pi} f(\theta) q_2(\theta) \overline{q_1(\theta)} d\theta = \int_{-\pi}^{\pi} f(\theta) e^{i(2s+1)\theta} \left(\sum_{k=0}^{m'-s-1} z_k e^{i(m'-s-1-k)\theta} \right) \cdot \left(\sum_{k=0}^{m-m'-s-1} z_{k+m'+s} e^{ik\theta} \right) d\theta.$$

Notice that $2s \geq p$ and

$$(3.6) \quad \int_{-\pi}^{\pi} e^{ik\theta} = 0 \quad \forall k \neq 0;$$

hence we have

$$\int_{-\pi}^{\pi} g_p(\theta) q_1(\theta) \overline{q_2(\theta)} d\theta = \int_{-\pi}^{\pi} g_p(\theta) e^{-i(2s+1)\theta} \left(\sum_{k=0}^{m'-s-1} z_k e^{-i(m'-s-1+k)\theta} \right) \cdot \left(\sum_{k=0}^{m-m'-s-1} z_{k+m'+s} e^{-ik\theta} \right) d\theta = 0 \quad \forall z \in \Omega_{2s}$$

and

$$\int_{-\pi}^{\pi} g_p(\theta)q_2(\theta)\overline{q_1(\theta)}d\theta = \int_{-\pi}^{\pi} g_p(\theta)e^{i(2s+1)\theta} \left(\sum_{k=0}^{m'-s-1} z_k e^{i(m'-s-1-k)\theta} \right) \cdot \left(\sum_{k=0}^{m-m'-s-1} z_{k+m'+s} e^{ik\theta} \right) d\theta = 0 \quad \forall z \in \Omega_{2s},$$

i.e.,

$$(3.7) \quad \int_{-\pi}^{\pi} g_p(\theta)q_1(\theta)\overline{q_2(\theta)}d\theta = \int_{-\pi}^{\pi} g_p(\theta)q_2(\theta)\overline{q_1(\theta)}d\theta = 0.$$

It follows by (3.3) and (3.7) that

$$\begin{aligned} & \left| \frac{z^t(A_m - R_m)z}{z^t R_m z} \right| \\ & \leq \frac{|\int_{-\pi}^{\pi} [f(\theta) - \gamma_1 g_q(\theta)]q_1(\theta)\overline{q_2(\theta)}d\theta| + |\int_{-\pi}^{\pi} [f(\theta) - \gamma_1 g_q(\theta)]q_2(\theta)\overline{q_1(\theta)}d\theta|}{\int_{-\pi}^{\pi} f(\theta)(|q_1(\theta)|^2 + |q_2(\theta)|^2)d\theta} \\ & \leq \frac{2 \int_{-\pi}^{\pi} [f(\theta) - \gamma_1 g_q(\theta)]|q_1(\theta)||q_2(\theta)|d\theta}{\int_{-\pi}^{\pi} f(\theta)(|q_1(\theta)|^2 + |q_2(\theta)|^2)d\theta} \\ & \leq \frac{\int_{-\pi}^{\pi} [f(\theta) - \gamma_1 g_q(\theta)](|q_1(\theta)|^2 + |q_2(\theta)|^2)d\theta}{\int_{-\pi}^{\pi} f(\theta)(|q_1(\theta)|^2 + |q_2(\theta)|^2)d\theta} \leq 1 - \frac{\gamma_1}{\gamma_2} \quad \forall z \in \Omega_{2s}. \end{aligned}$$

The results follow by using Corollary 3.2. \square

Using the similar argument as in the proof of the above theorem, we show that if f is a continuous function in \mathcal{C}^+ with multiple zeros of even order, then the spectra of the preconditioned matrices are indeed clustered around 1.

THEOREM 3.4. *Let f be a continuous function in \mathcal{C}^+ and have μ zeros of order $2p_j$ at $\theta_j \in [-\pi, \pi]$ for $1 \leq j \leq \mu$. Let $p = \sum_{j=1}^{\mu} p_j$. For any given $\epsilon > 0$, there exists a positive integer M such that for all $m > M$, at most M eigenvalues of $R_m^{-1}A_m$ are outside the interval $[1 - \epsilon, 1 + \epsilon]$.*

Proof. By definition of zeros, $f^{(2p_j)}$ is continuous at $\{\theta_j\}_{j=1}^{\mu}$; therefore h is continuous at $\{\theta_j\}_{j=1}^{\mu}$. Since f is continuous on $[-\pi, \pi]$, the function h in (3.2) is also continuous on $[-\pi, \pi]$. For any given $\epsilon > 0$, there exists a trigonometric polynomial q of degree M' such that

$$|h(\theta) - q(\theta)| \leq \frac{\epsilon h_{\min}}{2} \quad \forall \theta \in [-\pi, \pi],$$

where h_{\min} is the minimum value of the function h . It follows that

$$1 - \frac{\epsilon}{2} \leq \frac{f(\theta)}{g_p(\theta)q(\theta)} \leq 1 + \frac{\epsilon}{2} \quad \forall \theta \in [-\pi, \pi].$$

The theorem follows by using the similar argument in the previous theorem by letting $\gamma_1 = 1 - \frac{\epsilon}{2}$ and $\gamma_2 = 1 + \frac{\epsilon}{2}$. Here $M = p + M' + 1$. \square

3.2. The spectra of $I_m \pm A_m^{-1}H_m$. When $m' = m/2$, we need to study the spectra of the preconditioned matrices $I_m \pm A_m^{-1}H_m$ on different levels.

THEOREM 3.5. *Let $f \in C^+$ and have μ zeros of order $2p_j$ at $\theta_j \in [-\pi, \pi]$ for $1 \leq j \leq \mu$. Let $p = \sum_{j=1}^{\mu} p_j$. Then there exist two positive constants α, β ($\alpha < \beta$), independent of m , such that for all $m > p+1$, at most $p+1$ eigenvalues of $I_m \pm A_m^{-1}H_m$ are outside the interval $[\alpha, \beta]$.*

Proof. By similar arguments in Theorem 3.3, we note that

$$(3.8) \quad z^t H_m z = \sum_{k,l=0}^{m-1} a_{2m-(k+l)-1} z_k z_l = \int_{-\pi}^{\pi} f(\theta) e^{i\theta} \left(\sum_{k=0}^{m-1} z_k e^{i(m-k-1)\theta} \right)^2 d\theta,$$

where z is an m -vector. We can also get from [7, Theorem 1] that

$$(3.9) \quad z^t A_m z = \int_{-\pi}^{\pi} f(\theta) \left| \sum_{k=0}^{m-1} z_k e^{i(m-k-1)\theta} \right|^2 d\theta.$$

In this case, we let $s = \lceil p/2 \rceil$ and consider

$$\Omega_s = \{z = (z_0, z_1, \dots, z_{m-1})^t \in \mathbb{R}^m \mid z_k = 0, k = m-s, \dots, m-1\}.$$

We see that Ω_s is a subspace of \mathbb{R}^m with the dimension $m-s$. Now we consider the form

$$\frac{z^t H_m z}{z^t A_m z} \quad \forall z \in \Omega_s.$$

According to the definition of Ω_s , (3.8), and (3.9), we can get

$$(3.10) \quad \frac{z^t H_m z}{z^t A_m z} = \frac{\int_{-\pi}^{\pi} f(\theta) e^{i(2s+1)\theta} \left(\sum_{k=0}^{m-s-1} z_k e^{i(m-s-1-k)\theta} \right)^2 d\theta}{\int_{-\pi}^{\pi} f(\theta) \left| \sum_{k=0}^{m-s-1} z_k e^{i(m-s-1-k)\theta} \right|^2 d\theta} \quad \forall z \in \Omega_s.$$

By noting that $2s \geq p$ and using (3.6), we have

$$(3.11) \quad \int_{-\pi}^{\pi} g_p(\theta) e^{i(2s+1)\theta} \left(\sum_{k=0}^{m-s-1} z_k e^{i(m-s-1-k)\theta} \right)^2 d\theta = 0.$$

Since $f - \gamma_1 g_p \geq 0$, for all $z \in \Omega_s$, we have

$$(3.12) \quad \begin{aligned} \left| \frac{z^t H_m z}{z^t A_m z} \right| &= \left| \frac{\int_{-\pi}^{\pi} (f(\theta) - \gamma_1 g_p(\theta)) e^{i(2s+1)\theta} \left(\sum_{k=0}^{m-s-1} z_k e^{i(m-s-1-k)\theta} \right)^2 d\theta}{\int_{-\pi}^{\pi} f(\theta) \left| \sum_{k=0}^{m-s-1} z_k e^{i(m-s-1-k)\theta} \right|^2 d\theta} \right| \\ &\leq \frac{\int_{-\pi}^{\pi} |f(\theta) - \gamma_1 g_p(\theta)| \left| \sum_{k=0}^{m-s-1} z_k e^{i(m-s-1-k)\theta} \right|^2 d\theta}{\int_{-\pi}^{\pi} f(\theta) \left| \sum_{k=0}^{m-s-1} z_k e^{i(m-s-1-k)\theta} \right|^2 d\theta} \\ &= 1 - \gamma_1 \frac{\int_{-\pi}^{\pi} g_p(\theta) \left| \sum_{k=0}^{m-s-1} z_k e^{i(m-s-1-k)\theta} \right|^2 d\theta}{\int_{-\pi}^{\pi} f(\theta) \left| \sum_{k=0}^{m-s-1} z_k e^{i(m-s-1-k)\theta} \right|^2 d\theta} \\ &\leq 1 - \frac{\gamma_1}{\gamma_2}. \end{aligned}$$

The result follows by using Corollary 3.2. □

Similar to Theorem 3.4, if f is a continuous function in \mathcal{C}^+ with multiple zeros of even order, then the spectra of the preconditioned matrices are clustered around 1.

THEOREM 3.6. *Let f be a continuous function in \mathcal{C}^+ and have μ zeros of order $2p_j$ at $\theta_j \in [-\pi, \pi]$ for $1 \leq j \leq \mu$. Let $p = \sum_{j=1}^{\mu} p_j$. For any given $\epsilon > 0$, there exists a positive integer M such that for all $m > M$, at most M eigenvalues of $I_m \pm A_m^{-1}H_m$ are outside the interval $[1 - \epsilon, 1 + \epsilon]$.*

We emphasize in Theorems 3.3–3.6 that we do not need explicit knowledge of where the zero of f is in order to construct the preconditioners in our method. By Theorem 3.3 (or Theorem 3.5) and using the fact that the smallest eigenvalue of $I_m \pm A_m^{-1}H_m$ tends to zero at the rate at most $O(m^{p_{\max}})$, the number of PCG iterations required for convergence is of $O(p_{\max} \log m)$; see [3].

Next we consider the case where the generating function f is strictly positive. We note that the spectrum of A_m is contained in $[f_{\min}, f_{\max}]$, where f_{\min} and f_{\max} are the minimum and maximum values of f ; see [7, Lemma 1]. Since $f_{\min} > 0$, A_m is well-conditioned. In this case, the preconditioned system is also well-conditioned. By Theorem 3.4 (or Theorem 3.6), its spectrum is also clustered around 1. In this case, the PCG method, when applied to solving the preconditioned system, will converge superlinearly; see, for instance, [10].

4. Comparisons with other iterative methods.

4.1. Circulant preconditioners. In 1986, Strang [38] and Olkin [29] independently proposed the use of circulant matrices to precondition Toeplitz matrices in conjugate gradient iterations. The details of circulant preconditioners for Toeplitz matrices can be found in [10]. Recently, Chan, Yip, and Ng [13], Chan, Ng, and Yip [11], and Potts and Steidl [30] have proposed using the generalized Jackson kernel to construct circulant preconditioners for Toeplitz systems generated by a nonnegative continuous function with a finite number of zeros of even order. Circulant preconditioners based on Jackson kernels were proposed for the strictly positive case in [12] and the nonnegative case in [33]. The results in [13, 11, 30] have shown that the conjugate gradient method, when applied to solving circulant preconditioned Toeplitz systems generated by nonnegative continuous functions, converges linearly. Therefore, the complexity of solving n -by- n Toeplitz systems generated by nonnegative continuous functions is of $O(n \log n)$ operations.

On the other hand, Yeung and Chan [43] have considered circulant preconditioners for Toeplitz matrices generated by nonnegative, bounded piecewise continuous functions. They have showed that the spectra of these circulant preconditioned Toeplitz matrices cannot be clustered around one. More precisely, they have showed that for all sufficiently small $\epsilon > 0$, the number of eigenvalues of circulant preconditioned Toeplitz matrices that lie outside $(1 - \epsilon, 1 + \epsilon)$ will be at least of $O(\log n)$. The number of iterations required for convergence increases like $O(\log n)$ and hence the convergence rate of the method cannot be superlinear. Therefore, the complexity of solving Toeplitz systems generated by nonnegative, bounded, and piecewise continuous functions is of $O(n \log^2 n)$ operations.

In Table 4.1, we compare the theoretical convergence rates between the RBM and the circulant preconditioning method. We see that the RBM is competitive for ill-conditioned symmetric positive definite Toeplitz systems.

4.2. Banded Toeplitz preconditioners. Banded Toeplitz preconditioners have also been proposed; see [7, 31, 37]. The basic idea behind these preconditioners is to find a function g that matches the zeros of the generating function f . Then the pre-

TABLE 4.1

The comparison of the theoretical results about the convergence rates between the RBM and the circulant preconditioning method.

	Continuous function		Piecewise continuous function	
	Positive	Nonnegative	Positive	Nonnegative
Recursive-based PCG method	superlinearly	converges in $O(\log n)$ iterations	linearly	converges in $O(\log n)$ iterations
Circulant preconditioning method	superlinearly	linearly	converges in $O(\log n)$ iterations	converges in $O(\log n)$ iterations

conditioners are constructed based on the function g . These approaches work when knowledge of the position, number, and order of the zeros of the generating function f is known. In [37], Serra has proposed an algorithm to economically determine the knowledge of the generating function allowing us to choose and define the banded Toeplitz preconditioner.

4.3. Multigrid methods. In the literature, Fiorentino and Serra [16], Chan, Chang, and Sun [8], Huckle [25], Serra [35], and Sun, Jin, and Chang [40] have considered the use of multigrid method for solving ill-conditioned Toeplitz systems. Their idea is to generate a sequence of Toeplitz subsystems on different levels by interpolation and restriction operators. However, their generated Toeplitz matrices are not the principal submatrices of A_n . Compared with their multigrid methods, our method requires no coarse grid corrections at all that may be viewed as a “one-way” multigrid.

4.4. Computational cost. In this subsection, we analyze the computational costs of our method and compare it with other iterative methods.

The main cost for each subsystem depends on the matrix-vector multiplications $A_m v$ and $R_m v$ vector v in each PCG iteration. For each subsystem, $A_m v$ can be computed in two $2m$ -length FFTs; see, for instance, [10]. For the preconditioner R_m , we need to compute two matrix-vector products $A_{m'}^{-1} v$ and $A_{m-m'}^{-1} v$; see (2.4). The details of fast structured matrix-vector multiplications can be found in [20].

By using the Gohberg–Semencul formula, we have

$$A_k^{-1} = \frac{1}{\delta^{(k-1)}} [L_k^{(1)}(L_k^{(1)})^t - L_k^{(2)}(L_k^{(2)})^t] \quad \forall k \geq 2,$$

where $L_k^{(1)}$ and $L_k^{(2)}$ are lower and upper triangular Toeplitz matrices, respectively, given in (1.3). Because of the displacement structure of Toeplitz matrices, we have

$$J_k A_k^{-1} J_k = A_k^{-1}, \quad J_k (L_k^{(i)})^t J_k = L_k^{(i)}, \quad J_k L_k^{(i)} J_k = (L_k^{(i)})^t, \quad i = 1, 2.$$

It follows that

$$A_k^{-1} = \frac{1}{\delta^{(k-1)}} [L_k^{(1)}(L_k^{(1)})^t - L_k^{(2)}(L_k^{(2)})^t] = \frac{1}{\delta^{(k-1)}} [(L_k^{(1)})^t L_k^{(1)} - (L_k^{(2)})^t L_k^{(2)}].$$

We also note from (1.3) that $L_k^{(1)} + (L_k^{(2)})^t$ is a circulant matrix and $L_k^{(1)} - (L_k^{(2)})^t$ is a skew-circulant matrix. Thus the matrix $[L_k^{(1)}(L_k^{(1)})^t - L_k^{(2)}(L_k^{(2)})^t]$ can be expressed as the products of circulant and skew-circulant matrices:

$$\begin{aligned} & [L_k^{(1)}(L_k^{(1)})^t - L_k^{(2)}(L_k^{(2)})^t] \\ &= \frac{1}{2} \{ [L_k^{(1)} + (L_k^{(2)})^t] [(L_k^{(1)})^t - L_k^{(2)}] + [(L_k^{(1)})^t + L_k^{(2)}] [L_k^{(1)} - (L_k^{(2)})^t] \}. \end{aligned}$$

Since

$$[(L_k^{(1)})^t + L_k^{(2)}][L_k^{(1)} - (L_k^{(2)})^t]J_k = J_k[L_k^{(1)} + (L_k^{(2)})^t][(L_k^{(1)})^t - L_k^{(2)}],$$

we have

$$\begin{aligned} & [L_k^{(1)}(L_k^{(1)})^t - L_k^{(2)}(L_k^{(2)})^t]v \\ &= [L_k^{(1)} + (L_k^{(2)})^t][(L_k^{(1)})^t - L_k^{(2)}]v + [L_k^{(1)} + (L_k^{(2)})^t][(L_k^{(1)})^t - L_k^{(2)}]J_kv. \end{aligned}$$

Therefore, the matrix-vector multiplication $A_k^{-1}v$ is just equal to

$$\frac{1}{2^{\delta(k-1)}}(Real(z) + J_k Imag(z)),$$

where

$$z = [L_k^{(1)} + (L_k^{(2)})^t][(L_k^{(1)})^t - L_k^{(2)}](v + iJ_kv),$$

and $Real(z)$ and $Imag(z)$ denote the real and imaginary parts of z . The circulant and skew-circulant matrix-vector products can be computed efficiently by using FFTs. Therefore, $A_k^{-1}v$ can also be computed in roughly the same amount of time by two $2k$ -length FFTs. Thus the total cost per iteration for each subsystem is about two $2m$ -length, $2m'$ -length, and $2(m - m')$ -length FFTs. In particular, when $m' = m/2$, the total cost per iteration for each subsystem is about four $2m$ -length FFTs.

In comparison, the circulant PCG method requires two $2m$ -length FFTs and two m -length FFTs per iteration for the multiplication of m -by- m Toeplitz and circulant matrix-vector products; see [10]. A more efficient implementation of circulant preconditioners can be found in [23]. The band-Toeplitz PCG method requires two $2m$ -length FFTs and one band-solver where the bandwidth depends on the order of the zeros [7]. For the multigrid method for solving Toeplitz systems, the total cost per iteration is about eight $2m$ -length FFTs. Thus when $m' = m/2$, the cost per iteration of using RBM is about $4/3$ times that required by the circulant PCG method [10], 2 times that required by the band-Toeplitz PCG method, and half of that required by the multigrid method.

Next we estimate the total cost of the RBM for solving an n -by- n symmetric positive definite Toeplitz system with $n = 2^\nu$. For simplicity, we consider the case of $m' = m/2$ in each level. Let l be the smallest size of the subsystem in the RBM (cf. section 2). Therefore, our scheme will solve k subsystems with

$$l = \frac{n}{2^{\nu-k}}.$$

According to Theorem 3.5, the number of iterations required for convergence for the $2^{\nu-j}$ -by- $2^{\nu-j}$ subsystem is about $c_j(\nu - j)$ ($1 \leq j \leq k$). Here c_j are constants independent of n . It follows that the total cost of the RBM is about

$$8 \sum_{j=1}^k c_j(\nu - j) \text{fcost}(2^{\nu-j+1}),$$

where $\text{fcost}(m)$ denotes the cost of an m -length FFT. The cost of an m -length FFT is roughly the twice the cost of an $m/2$ -length FFT. Hence the total cost of the RBM is roughly bounded by

$$(4.1) \quad 16 \max_{1 \leq j \leq k} c_j \text{fcost}(2n)\nu \approx O(n \log^2 n).$$

TABLE 4.2

The comparison of the theoretical complexities between the RBM and the circulant preconditioning method.

	Continuous function		Piecewise continuous function	
	Positive	Nonnegative	Positive	Nonnegative
RBM	$O(n \log n)$	$O(n \log^2 n)$	$O(n \log n)$	$O(n \log^2 n)$
Circulant preconditioning	$O(n \log n)$	$O(n \log n)$	$O(n \log^2 n)$	$O(n \log^2 n)$

When the original Toeplitz system is well-conditioned (the number of iterations required for convergence for $2^{\nu-j}$ -by- $2^{\nu-j}$ subsystem is about c_j ($1 \leq j \leq k$)), the total cost of the RBM is $O(n \log n)$ operations. In Table 4.2, we compare the theoretical complexities between the RBM and the circulant preconditioning. We see that the RBM is better than the circulant preconditioning when the generating function is positive and piecewise continuous.

We note that the RBM not only solves n -by- n Toeplitz systems, but also solves Toeplitz systems of sizes $n/2$ -by- $n/2$, $n/4$ -by- $n/4$, \dots , l -by- l together. However, if we have a procedure that costs at most $qn \log n$ operations (q is a fixed universal constant) for solving n -by- n Toeplitz systems, then the cost of solving any set of Toeplitz linear systems of size m -by- m with $m \in \{n, n/2, n/4, \dots, 1\}$ is of order $2qn \log n$ plus lower order terms. In particular, by using the circulant preconditioning method, the total cost of solving all Toeplitz systems of sizes n -by- n , $n/2$ -by- $n/2$, $n/4$ -by- $n/4$, \dots , l -by- l is of $O(n \log n)$ operations which is less expensive than that of using the RBM (cf. (4.1)).

In the next section, we will demonstrate the effectiveness of the RBM by some numerical examples. We will show that the total computational cost of the RBM for solving all the systems together is less than that of the circulant preconditioning method (cf. Table 5.7). We will also compare the performance of the RBM with the other iterative methods such as the banded Toeplitz preconditioning method and the multigrid method.

5. Numerical examples. In this section, we apply the RBM in section 2 to solve symmetric positive definite Toeplitz systems $A_n x_n = e_n$ for $n = 128, 256, 512, 1024$, and 2048 . The initial guess is the zero vector. The stopping criteria at the finest level and at the coarser levels are $\|r_q\|_2 / \|r_0\|_2 \leq 1 \times 10^{-7}$ and $\|r_q\|_2 / \|r_0\|_2 \leq \tau$, respectively, where r_q is the residual vector after q iterations. In the tests, the coarsest level is $n = 64$, and different stopping criteria for τ ($= 1 \times 10^{-3}, 1 \times 10^{-4}, 1 \times 10^{-7}$) are used.

In Tables 5.1–5.4, we give the number of iterations required for convergence by using R_n as the preconditioner for solving n -by- n Toeplitz systems; see the column under $R(\tau)$. We remark that the preconditioners are constructed recursively. For instance, when we solve $A_{512} x_{512} = e_1$, the preconditioners are constructed by solving two Toeplitz systems $A_{256} x_{256} = e_1$ and $A_{128} x_{128} = e_1$ using the PCG method with the stopping criteria being τ , and using the direct solver for $A_{64} x_{64} = e_1$.

For comparison, we also give the number of iterations by the full multigrid method (M) [8], the PCG method with no preconditioner (I), the Strang circulant preconditioner (S) [38], the T. Chan circulant preconditioner (T) [14], the Jackson kernel-type preconditioner (J) of order 6 [13], and also the banded preconditioner (B) [7]. Moreover, the double asterisk ** in the tables signifies that more than 200 iterations are required. We also note that the Strang preconditioner in general is not positive def-

TABLE 5.1
Number of iterations for well-conditioned systems.

n	$\theta^4 + 1$							
	I	S	C	J	B	$R(1 \times 10^{-3})$	$R(1 \times 10^{-4})$	$R(1 \times 10^{-7})$
128	71	7	8	7	15	5	5	5
256	78	7	7	7	17	5	5	5
512	80	7	7	7	17	5	5	5
1024	81	7	7	7	17	5	4	4
2048	82	7	7	7	17	4	4	4

TABLE 5.2
Number of iterations for functions with order 2 zeros.

n	θ^2								
	I	S	C	J	B	M	$R(1 \times 10^{-3})$	$R(1 \times 10^{-4})$	$R(1 \times 10^{-7})$
128	170	-	16	10	10	12	5	5	5
256	**	-	20	10	10	12	5	5	5
512	**	-	24	9	10	12	5	5	5
1024	**	-	32	9	10	12	5	5	5
2048	**	-	43	9	10	12	6	5	5

n	$(\theta^2 - 1)^2$								
	I	S	C	J	B	M	$R(1 \times 10^{-3})$	$R(1 \times 10^{-4})$	$R(1 \times 10^{-7})$
128	**	9	30	14	11		6	6	6
256	**	10	27	14	12		6	6	6
512	**	8	36	13	12		6	6	6
1024	**	12	46	13	12		6	6	6
2048	**	13	52	13	12		6	6	6

n	$\theta^2(\pi^2 - \theta^2)^2$								
	I	S	C	J	B	M	$R(1 \times 10^{-3})$	$R(1 \times 10^{-4})$	$R(1 \times 10^{-7})$
128	119	10	17	11	13		6	6	6
256	**	13	20	11	14		6	6	6
512	**	15	26	11	14		6	6	6
1024	**	17	33	12	15		6	6	6
2048	**	19	46	12	16		6	6	6

n	$J(\theta)$								
	I	S	C	J	B	M	$R(1 \times 10^{-3})$	$R(1 \times 10^{-4})$	$R(1 \times 10^{-7})$
128	81	19	17	12	14	8	8	8	8
256	173	24	21	13	15	8	8	8	8
512	**	46	27	13	15	8	9	9	9
1024	**	81	34	14	15	8	9	9	9
2048	**	105	51	16	15	8	9	9	9

inite; see [10]. When some of the eigenvalues are negative, we denote the iteration number by “-” as the PCG method does not apply to nondefinite systems and the solution thus obtained may be inaccurate.

The first test function in Table 5.1 is a positive function and therefore corresponds to a well-conditioned system. Notice that the number of iterations for the nonpreconditioned systems tends to a constant when n is large, indicating that the convergence is linear. In this case, we see that all preconditioners work well and the convergence is fast. We also remark that the full multigrid method cannot be applied to this test function since the order of the minimum point of $\theta^4 + 1$ is 4; see [8].

TABLE 5.3
Number of iterations for functions with order 4 zeros.

n	θ^4							
	I	S	C	J	B	$R(1 \times 10^{-3})$	$R(1 \times 10^{-4})$	$R(1 \times 10^{-7})$
128	**	–	71	16	24	7	7	7
256	**	–	161	18	27	8	8	8
512	**	–	167	17	29	8	8	8
1024	**	–	**	18	30	9	10	10
2048	**	–	**	19	31	19	15	11

n	$\theta^4(\pi^2 - \theta^2)$							
	I	S	C	J	B	$R(1 \times 10^{-3})$	$R(1 \times 10^{-4})$	$R(1 \times 10^{-7})$
128	**	–	33	18	24	8	8	8
256	**	–	45	18	26	8	8	8
512	**	–	60	18	29	11	11	11
1024	**	–	82	21	30	12	12	12
2048	**	–	135	22	32	15	14	13

TABLE 5.4
Number of iterations for other functions.

n	$ \theta $							
	I	S	C	J	B	$R(1 \times 10^{-3})$	$R(1 \times 10^{-4})$	$R(1 \times 10^{-7})$
128	56	8	9	8	7	6	6	6
256	77	8	9	8	7	6	6	6
512	110	8	10	8	7	6	6	6
1024	144	8	10	9	7	7	6	6
2048	**	8	10	9	7	7	7	7

n	$ \theta ^3$							
	I	S	C	J	B	$R(1 \times 10^{-3})$	$R(1 \times 10^{-4})$	$R(1 \times 10^{-7})$
128	**	–	41	13	53	7	7	7
256	**	–	62	12	55	8	8	8
512	**	–	98	14	56	8	8	8
1024	**	–	152	15	58	9	9	9
2048	**	–	**	15	59	15	10	10

The four test functions in Table 5.2 are nonnegative functions with single or multiple zeros of order 2 on $[-\pi, \pi]$. They are θ^2 , $(\theta^2 - 1)^2$, $\theta^2(\pi^2 - \theta^2)^2$, and

$$J(\theta) = \begin{cases} \theta^2, & |\theta| \leq \pi/2, \\ 1, & |\theta| > \pi/2. \end{cases}$$

Thus the condition numbers of the Toeplitz matrices are growing like $O(n^2)$, and hence the number of iterations required for convergence without using any preconditioners is increasing like $O(n)$. Moreover, since the zero of $(\theta^2 - 1)^2$ is not at $\theta = 0$, the multigrad method cannot be applied directly; see [8]. Recently, Huckle [25] proposed a method to handle this case. We remark that the construction of circulant preconditioners and our preconditioners requires only the entries of A_n and does not require explicit knowledge of the zero of the generating function.

We note from Table 5.2 that for these generating functions, the number of iterations required for convergence by using T. Chan preconditioner increases as n increases. The PCG with the Strang, Jackson circulant preconditioners and banded

TABLE 5.5
Numbers of outliers of the preconditioned matrices for θ^2 .

n	$R(1 \times 10^{-3})$		$R(1 \times 10^{-4})$		$R(1 \times 10^{-7})$		J	
	$\epsilon = 0.1$	$\epsilon = 0.01$	$\epsilon = 0.1$	$\epsilon = 0.01$	$\epsilon = 0.1$	$\epsilon = 0.01$	$\epsilon = 0.1$	$\epsilon = 0.01$
64	1	2	1	2	1	2	5	12
128	1	2	1	2	1	2	5	12
256	1	2	1	2	1	2	5	12
512	1	2	1	2	1	2	5	12

TABLE 5.6
Number of outliers of the preconditioned matrices for θ^4 .

n	$R(1 \times 10^{-3})$		$R(1 \times 10^{-4})$		$R(1 \times 10^{-7})$		J	
	$\epsilon = 0.1$	$\epsilon = 0.01$	$\epsilon = 0.1$	$\epsilon = 0.01$	$\epsilon = 0.1$	$\epsilon = 0.01$	$\epsilon = 0.1$	$\epsilon = 0.01$
64	2	3	2	3	2	3	17	35
128	2	4	2	3	2	3	16	36
256	3	5	2	3	2	3	16	36
512	4	8	3	4	2	3	16	36

preconditioners, the multigrid method, and the RBM perform quite well. However, the number of iterations required by the RBM is less than those required by the other methods. We find that for different τ (the stopping criteria at the coarser levels), the numbers of iterations required by the RBM are about the same.

We also consider two functions with zeros of order 4. They are θ^4 and $\theta^4(\pi^2 - \theta^2)$. The condition number of Toeplitz matrices will increase like $O(n^4)$, and the matrices will be very ill-conditioned even for moderate n . We remark again that the full multigrid method [8] cannot be applied to solving Toeplitz systems generated by a function with zeros of order 4. Recently, Serra [35] also presented a new multigrid scheme to handle generating function with zeros of higher order. However, the numerical results are not given. We see from Table 5.3 that both the Strang and T. Chan circulant preconditioners do not work. In fact, Tyrtyshnikov [42] has proved that the Strang [38] and the T. Chan [14] preconditioners both fail in this case. Other transform based preconditioners also fail; see [4]. The RBM, the Jackson circulant preconditioners, and the banded preconditioners perform quite well. However, the number of iterations required by the RBM is less than those required by the other methods. However, when τ is large (i.e., the stopping criteria at the coarser levels are large), the RBM needs more iterations to converge at the finest level. In Table 5.4, we test functions that our theory does not cover. The first two functions are not differentiable at their zeros. Table 5.4 shows that both the Jackson circulant preconditioners and the RBM perform quite well.

To further compare the clustering property of our proposed preconditioner and the Jackson circulant preconditioner, we give in Tables 5.5 and 5.6 the numbers of outliers of the preconditioned matrices outside the interval $(1 - \epsilon, 1 + \epsilon)$. We see that the spectra of our preconditioned matrices are more clustered than those of Jackson-type kernel preconditioned matrices. The number of outliers keeps the same for the generating function θ^2 . However, we find that for the generating function θ^4 , the number of outliers of our preconditioned matrices increases when τ is large.

Next we compare the computational cost of the RBM and the circulant preconditioning method with the Jackson kernel-type preconditioner. In Figures 5.1 and 5.2, we show the number flops required to solve $A_n x_n = e_n$ for $n = 128, 256, 512$,

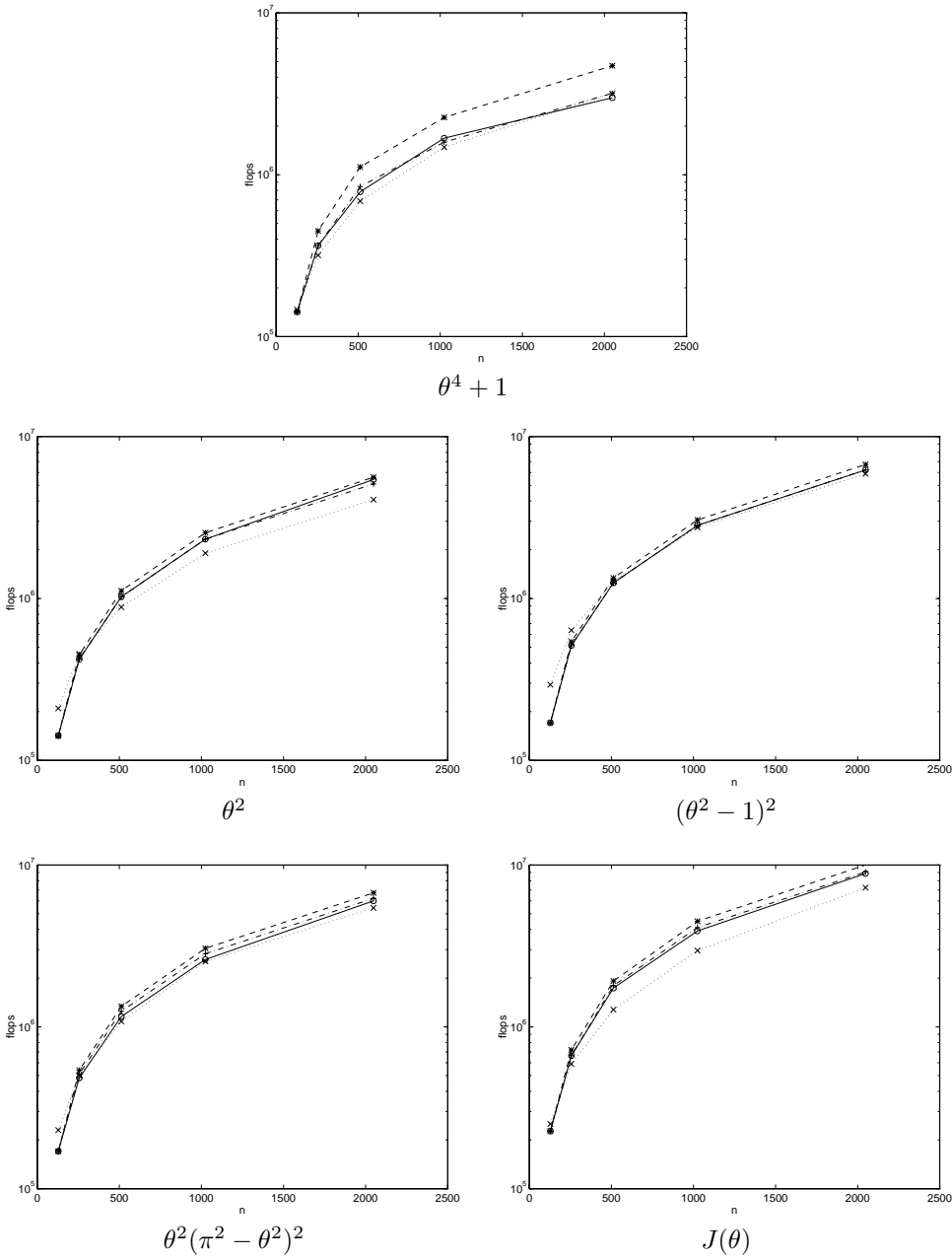


FIG. 5.1. Number of flops required for different preconditioners. J :x..... , $R(1 \times 10^{-3})$: —○— , $R(1 \times 10^{-4})$: - - - + - - - , $R(1 \times 10^{-7})$: - - - * - - - .

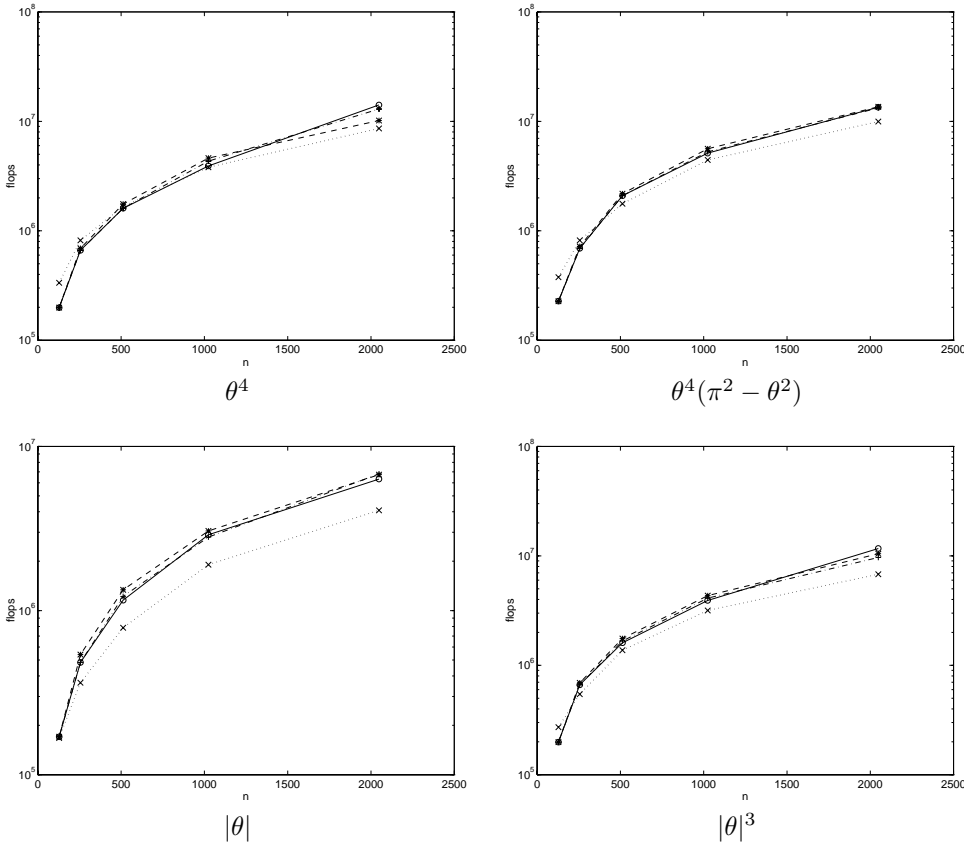


FIG. 5.2. Number of flops required for different preconditioners. J :x..... , $R(1 \times 10^{-3})$: —○— , $R(1 \times 10^{-4})$: -.-.-+.-.-.- , $R(1 \times 10^{-7})$: --- * --- .

1024, and 2048. We see from the figures that the circulant preconditioning method is more efficient than the RBM. The numbers of flops required by the RBM for different generating functions are greater than those by the circulant preconditioning method. However, their differences in the numbers of flops are not significant. We find that the computational cost of the RBM is roughly 1.5 times that of the circulant preconditioning method with the Jackson circulant preconditioner.

In Table 5.7, we list the numbers of flops required to solve Toeplitz systems of sizes 128, 256, 512, 1024, and 2048 together. For circulant preconditioners, these numbers are shown in brackets under the column J . Since the RBM with $\tau = 1 \times 10^{-7}$ not only solves 2048-by-2048 Toeplitz systems but also solves Toeplitz systems of sizes 128, 256, 512, and 1024 together, these numbers are just the same as the numbers shown in Figures 5.1 and 5.2. We find in the table that the total computational cost of the RBM for solving all the systems together is on average about 7/9 times that of the circulant preconditioning method with the Jackson circulant preconditioner. Therefore, the RBM is more efficient if we want to solve several Toeplitz systems of different sizes together.

5.1. Block-Toeplitz systems. In this subsection, we remark that our recursive-based PCG method can be adapted to handle kn -by- kn symmetric positive definite

TABLE 5.7
 Total numbers of million flops required for different generating functions.

Generating function	J	R	Ratio of the computational cost of the RBM over that of the circulant preconditioning method
$\theta^4 + 1$	5.80	4.74	0.82
θ^2	7.53	5.64	0.75
$(\theta^2 - 1)^2$	10.85	6.75	0.62
$\theta^2(\pi^2 - \theta^2)^2$	9.79	6.75	0.69
$J(\theta)$	12.33	10.17	0.82
θ^4	15.24	10.30	0.68
$\theta^4(\pi^2 - \theta^2)$	17.37	13.65	0.79
$ \theta $	7.30	6.86	0.94
$ \theta ^3$	12.16	10.86	0.89

block-Toeplitz matrices

$$T_{k,n} = \begin{bmatrix} A^{(0)} & A^{(1)} & \dots & A^{(n-2)} & A^{(n-1)} \\ A^{(1)} & A^{(0)} & A^{(1)} & & A^{(n-2)} \\ \vdots & A^{(1)} & A^{(0)} & \ddots & \vdots \\ A^{(n-2)} & & \ddots & \ddots & A^{(1)} \\ A^{(n-1)} & A^{(n-2)} & \dots & A^{(1)} & A^{(0)} \end{bmatrix},$$

where each $A^{(i)}$ is a k -by- k symmetric matrix. We are interested in solving the block-Toeplitz linear system

$$T_{k,n}X = B,$$

where X and B are kn -by- k matrices. For multilevel Toeplitz matrices, Serra and Tyrtyshnikov [36] proved that multilevel circulant preconditioners cannot produce a superlinearly convergence rate in general.

To modify the recursive-based PCG method for the block-Toeplitz system, we first note the following results of Gohberg and Heinig [18]:

$$(5.1) \quad T_{k,m}^{-1} = \Psi_{k,m}^{(1)} \Phi_{k,m} (\Psi_{k,m}^{(1)})^t - \Psi_{k,m}^{(2)} \Phi_{k,m} (\Psi_{k,m}^{(2)})^t, \quad m = 2, 3, \dots, n,$$

where $\Psi_{k,m}^{(1)}$ and $\Psi_{k,m}^{(2)}$ are km -by- km lower triangular block Toeplitz matrices given by

$$\Psi_{k,m}^{(1)} = \begin{bmatrix} V^{(1)} & 0 & \dots & 0 & 0 \\ V^{(2)} & V^{(1)} & 0 & & 0 \\ \vdots & V^{(2)} & V^{(1)} & \ddots & \vdots \\ V^{(m-1)} & & \ddots & \ddots & 0 \\ V^{(m)} & V^{(m-1)} & \dots & V^{(2)} & V^{(1)} \end{bmatrix}$$

and

$$\Psi_{k,m}^{(2)} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ V^{(m)} & 0 & 0 & & 0 \\ \vdots & V^{(m)} & 0 & \ddots & \vdots \\ V^{(3)} & & \ddots & \ddots & 0 \\ V^{(2)} & V^{(3)} & \dots & V^{(m)} & 0 \end{bmatrix},$$

TABLE 5.8
Number of Iterations required for the matrix-valued generating function with $k = 3$.

n	$R(1 \times 10^{-3})$	$R(1 \times 10^{-4})$	$R(1 \times 10^{-7})$
64	3	3	3
128	3	3	2
256	2	2	2
512	2	2	2

respectively, and

$$\Phi_{k,m} = \begin{bmatrix} (V^{(1)})^{-1} & & & \\ & (V^{(1)})^{-1} & & \\ & & \ddots & \\ & & & (V^{(1)})^{-1} \end{bmatrix}$$

with

$$T_{k,m} \begin{bmatrix} V^{(1)} \\ V^{(2)} \\ \vdots \\ V^{(m)} \end{bmatrix} = \begin{bmatrix} I \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

We remark that $V^{(i)}$ are k -by- k matrices. Similar to the case of the Toeplitz matrix, in the block case, we consider using

$$R_{k,m} = \begin{bmatrix} T_{k,m'} & 0 \\ 0 & T_{k,m-m'} \end{bmatrix}$$

as the preconditioner for $T_{k,m}$ with $m > m'$. The RBM can be derived similarly as in section 2. In Table 5.8, we show the numbers of iterations required for convergence for the block-Toeplitz systems [34] generated by the following matrix-valued function:

$$\begin{bmatrix} \theta^2 + 1 & |\theta| & \theta^2 \\ |\theta| & \theta^2 + 1 & |\theta| \\ \theta^2 & |\theta| & 2\theta^2 + 1 \end{bmatrix}.$$

In this test, the stopping criteria at the finest level and at the coarser levels are $\|r_q\|_2/\|r_0\|_2 \leq 1 \times 10^{-7}$ and $\|r_q\|_2/\|r_0\|_2 \leq \tau = 1 \times 10^{-3}, 1 \times 10^{-4}, 1 \times 10^{-7}$, respectively. The coarsest level is $n = 32$. Numerical results show that our method is quite efficient. We plan to investigate the block case of the RBM in future research.

In summary, we have developed the recursive-based PCG method using the Gohberg–Semencul formula to solve symmetric positive definite, but ill-conditioned, Toeplitz systems. Theoretical and numerical results are given to illustrate the fast convergence of the method.

Acknowledgment. We would like to thank the referees for their valuable suggestions.

REFERENCES

[1] G. S. AMMAR AND W. B. GRAGG, *Superfast solution of real positive definite Toeplitz systems*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 61–76.

- [2] G. AMMAR AND P. GADER, *New decomposition of the inverse of a Toeplitz matrix*, in Signal Processing, Scattering and Operator Theory, and Numerical Methods, Proceedings of the International Symposium on the Mathematical Theory of Networks and Systems Vol. 3, M. Kaashoek, J. van Schuppen, and A. Ran, eds., Birkhäuser, Boston, 1990, pp. 421–428.
- [3] F. DI BENEDETTO, *Analysis of preconditioning techniques for ill-conditioned Toeplitz matrices*, SIAM J. Sci. Statist. Comput., 16 (1995), pp. 682–697.
- [4] F. DI BENEDETTO AND S. SERRA, *A unifying approach to abstract matrix algebra preconditioning*, Numer. Math., 82 (1999), pp. 57–90.
- [5] R. BITMEAD AND B. ANDERSON, *Asymptotically fast solution of Toeplitz and related systems of linear equations*, Linear Algebra Appl., 34 (1980), pp. 103–116.
- [6] R. BRENT, F. GUSTAVSON, AND D. YUN, *Fast solution of Toeplitz systems of equations and computation of Padé approximants*, J. Algorithms, 1 (1980), pp. 259–295.
- [7] R. CHAN, *Toeplitz preconditioners for Toeplitz systems with nonnegative generating functions*, IMA J. Numer. Anal., 11 (1991), pp. 333–345.
- [8] R. H. CHAN, Q.-S. CHANG, AND H.-W. SUN, *Multigrid method for ill-conditioned symmetric Toeplitz systems*, SIAM J. Sci. Comput., 19 (1998), pp. 516–529.
- [9] R. H. CHAN AND K.-P. NG, *Fast iterative solvers for Toeplitz-plus-band systems*, SIAM J. Sci. Statist. Comput., 14 (1993), pp. 1013–1019.
- [10] R. H. CHAN AND M. K. NG, *Conjugate gradient methods for Toeplitz systems*, SIAM Rev., 38 (1996), pp. 427–482.
- [11] R. CHAN, M. NG, AND A. YIP, *The best circulant preconditioners for Hermitian Toeplitz systems II: The multiple-zero case*, Numer. Math., to appear.
- [12] R. H. CHAN AND M.-C. YEUNG, *Circulant preconditioners constructed from kernels*, SIAM J. Numer. Anal., 29 (1992), pp. 1093–1103.
- [13] R. H. CHAN, A. M. YIP, AND M. K. NG, *The best circulant preconditioners for Hermitian Toeplitz systems*, SIAM J. Numer. Anal., 38 (2001), pp. 876–896.
- [14] T. F. CHAN, *An optimal circulant preconditioner for Toeplitz systems*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 766–771.
- [15] J. DURBIN, *The fitting of time series models*, Rev. Int. Stat., 28 (1960), pp. 233–244.
- [16] G. FIORENTINO AND S. SERRA, *Multigrid methods for symmetric positive definite block Toeplitz matrices with nonnegative generating functions*, SIAM J. Sci. Comput., 17 (1996), pp. 1068–1081.
- [17] I. GOHBERG AND A. SEMENCUL, *On the inversion of finite Toeplitz matrices and their continuous analogs*, Mat. Issled., 2 (1972), pp. 201–233.
- [18] I. GOHBERG AND G. HEINIG, *Inversion of finite-section Toeplitz matrices consisting of elements of non-commutative algebra*, Rev. Roumaine Math. Pures Appl., 19 (1974), pp. 623–663.
- [19] G. HEINIG AND L. ROST, *Algebraic Methods for Toeplitz-like Matrices and Operators*, Birkhäuser Verlag, Basel, Switzerland, 1984.
- [20] G. HEINIG AND L. ROST, *Hartley transform representations of symmetric Toeplitz matrix inverses with application to fast matrix-vector multiplications*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 86–105.
- [21] F. DE HOOG, *A new algorithm for solving Toeplitz systems of equations*, Linear Algebra Appl., 88/89 (1987), pp. 123–138.
- [22] S. HAYKIN, *Adaptive Filter Theory*, Prentice-Hall, London, 1991.
- [23] T. HUCKLE, *Iterative Methods for Toeplitz-like Matrices*, Report SCCM-94-05, Computer Science Department, Stanford University, Stanford, CA, 1994.
- [24] T. HUCKLE, *Superfast solution of linear equations with low displacement rank*, in High Performance Algorithms for Structured Matrix Problems, Adv. Theory Comput. Comput. Math. II, P. Arbenz, M. Paprzycki, A. H. Sameh, and V. Sarin, eds., Nova Sci. Publ., Commack, NY, 1998, pp. 149–162.
- [25] T. HUCKLE, *Multigrid preconditioning and Toeplitz matrices*, Electron. Trans. Numer. Anal., 13 (2002), pp. 81–105.
- [26] R. HORN AND C. JOHNSON, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
- [27] N. LEVINSON, *The Wiener RMS (root mean square) error criterion in filter design and prediction*, J. Math. Phys., 25 (1947), pp. 261–278.
- [28] M. MORF, *Doubling algorithms for Toeplitz and related equations*, in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Process., 1980, pp. 954–959.
- [29] J. OLKIN, *Linear and Nonlinear Deconvolution Problems*, Ph.D. thesis, Rice University, Houston, TX, 1986.
- [30] D. POTTS AND G. STEIDL, *Preconditioners for ill-conditioned Toeplitz systems constructed from positive kernels*, SIAM J. Sci. Comput., 22 (2001), pp. 1741–1761.
- [31] S. SERRA, *Preconditioning strategies for Hermitian Toeplitz systems with nondefinite generat-*

- ing functions*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 1007–1019.
- [32] S. SERRA, *On the extreme eigenvalues of Hermitian (block) Toeplitz matrices*, Linear Algebra Appl., 270 (1998), pp. 109–129.
- [33] S. SERRA, *Korovkin theorems and linear positive gram matrix algebra approximations of Toeplitz matrices*, Linear Algebra Appl., 284 (1998), pp. 307–334.
- [34] S. SERRA, *Spectral and computational analysis of block Toeplitz matrices having nonnegative definite matrix-valued generating functions*, BIT, 39 (1999), pp. 152–175.
- [35] S. SERRA, *Convergence analysis of two-grid methods for elliptic Toeplitz and PDEs matrix-sequences*, Numer. Math., 92 (2002), pp. 433–465.
- [36] S. SERRA CAPIZZANO AND E. TYRTYSHNIKOV, *Any circulant-like preconditioner for multilevel matrices is not superlinear*, SIAM J. Matrix Anal. Appl., 21 (1999), pp. 431–439.
- [37] S. SERRA, *How to choose the best iterative strategy for symmetric Toeplitz systems*, SIAM J. Numer. Anal., 36 (1999), pp. 1078–1103.
- [38] G. STRANG, *A proposal for Toeplitz matrix calculations*, Stud. Appl. Math., 74 (1986), pp. 171–176.
- [39] T. STROHMER, *A Levinson–Galerkin algorithm for regularized trigonometric approximation*, SIAM J. Sci. Comput., 22 (2000), pp. 1160–1183.
- [40] H. SUN, X. JIN, AND Q. CHANG, *Convergence of the multigrid method for ill-conditioned block Toeplitz systems*, BIT, 41 (2001), pp. 179–190.
- [41] W. F. TRENCH, *An algorithm for the inversion of finite Toeplitz matrices*, SIAM J. Appl. Math., 12 (1964), pp. 515–522.
- [42] E. TYRTYSHNIKOV, *Circulant preconditioners with unbounded inverses*, Linear Algebra Appl., 216 (1995), pp. 1–24.
- [43] M. YEUNG AND R. CHAN, *Circulant preconditioners for Toeplitz matrices with piecewise continuous generating functions*, Math. Comp., 61 (1993), pp. 701–718.