A Survey of Preconditioners for Ill-Conditioned Toeplitz Systems

Raymond H. Chan, Michael K. Ng, and Andy M. Yip

Abstract. In this paper, we survey some of latest developments in using preconditioned conjugate gradient methods for solving mildly ill-conditioned Toeplitz systems where the condition numbers of the systems grow like $O(n^\nu)$ for some $\nu > 0$. This corresponds to Toeplitz matrices generated by functions having zeros of order ν. Because of the ill-conditioning, the number of iterations required for convergence in the conjugate gradient method will grow like $O(n^{\nu/2})$. Different preconditioners proposed for these Toeplitz matrices are reviewed. The main result is that the total complexity of solving an ill-conditioned Toeplitz system is of $O(n \log n)$ operations.

1. Introduction

An n-by-n matrix A_n is said to be Toeplitz if

$$A_n = \begin{bmatrix} a_0 & a_{-1} & \cdots & a_{2-n} & a_{1-n} \\ a_1 & a_0 & a_{-1} & & a_{2-n} \\ \vdots & a_1 & a_0 & \ddots & \vdots \\ a_{n-2} & & \ddots & \ddots & a_{-1} \\ a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \end{bmatrix}, \tag{1.1}$$

i.e., A_n is constant along its diagonals. Toeplitz systems of the form $A_n\mathbf{x} = \mathbf{b}$ occur in a variety of applications in mathematics and engineering [**8**].

Strang in [**25**] proposed using the preconditioned conjugate gradient method with circulant matrices as preconditioners for solving Toeplitz systems. The number of operations per iteration is of order $O(n \log n)$ as circulant systems can be solved efficiently by fast Fourier transforms. Several successful circulant preconditioners have been introduced and analyzed; see for instance [**13, 6**]. In these papers, the given Toeplitz matrix A_n is assumed to be generated by a generating function f, i.e., the diagonals a_j of A_n are given by the Fourier coefficients of f. It was shown that if f is a positive function in the Wiener class (i.e., the Fourier coefficients of f are absolutely summable), then these circulant preconditioned systems converge superlinearly [**6**]. However, if f has zeros, the corresponding Toeplitz systems will

1991 *Mathematics Subject Classification.* Primary 65F10, 65F15, 65T10.

Chan was supported in part by Hong Kong Research Grants Council Grant No. CUHK 4207/97P and CUHK DAG Grant No. 2060143.

Ng was supported in part by Research Grants Council Grant No. HKU 7147/99P and HKU CRCG Grant No. 10202720.

be ill-conditioned. In fact, for the Toeplitz matrices generated by a function with a zero of order 2ν, their condition numbers grow like $O(n^{2\nu})$, see [**23**]. Hence the number of iterations required for convergence will increase like $O(n^{\nu})$, see [**2**, p.24]. Tyrtyshnikov [**28**] has proved that the Strang [**25**] and the T. Chan [**13**] preconditioners both fail in this case. In this paper, we will survey results in using the preconditioned conjugate gradient method for solving Toeplitz systems generated by functions with zeros and give some insight in how to construct effective preconditioners.

The outline of the paper is as follows. In §2, we present the relationship between the spectrum of Toeplitz matrices and its generating function. In §3 and §4, we study band-Toeplitz and circulant-type matrices as preconditioners respectively for ill-conditioned Toeplitz systems. These two types of preconditioners require the explicit knowledge of the generating functions of the Toeplitz matrices. In §5, we consider the case where the Toeplitz matrix is given rather than the generating function and present circulant preconditioners that incorporate the idea of band-Toeplitz preconditioner. Finally, some concluding remarks are given in §6.

2. Toeplitz Matrices and Generating Functions

Let $\mathbf{C}_{2\pi}$ be the space of all 2π-periodic continuous real-valued functions. The Fourier coefficients of a function f in $\mathbf{C}_{2\pi}$ are given by

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

Clearly $a_k = \bar{a}_{-k}$ for all k. Let $A_n[f]$ be the n-by-n Hermitian Toeplitz matrix with the (i,j)th entry given by a_{i-j}, $i,j = 0, \ldots, n-1$. We will use $\mathbf{C}_{2\pi}^{+}$ to denote the space of all nonnegative functions in $\mathbf{C}_{2\pi}$ which are not identically zero. We remark that the Toeplitz matrices $A_n[f]$ generated by $f \in \mathbf{C}_{2\pi}^{+}$ are positive definite for all n, see [**7**, Lemma 1]. Conversely, if $f \in \mathbf{C}_{2\pi}$ takes both positive and negative values, then $A_n[f]$ will be non-definite. In this paper, we only consider $f \in \mathbf{C}_{2\pi}^{+}$.

THEOREM 2.1. *([**17**, p.65]) Let $f \in \mathbf{C}_{2\pi}$. Then the spectrum $\sigma(A_n[f])$ of $A_n[f]$ satisfies*

$$\sigma(A_n[f]) \subseteq [f_{\min}, f_{\max}], \quad \forall n \geq 1, \tag{2.1}$$

where $f_{\min}$ and $f_{\max}$ are the minimum and the maximum values of f respectively. Moreover, if $f_{\max} > f_{\min}$, then

$$f_{\min} < \lambda_{\min}(A_n[f]) \leq \lambda_{\max}(A_n[f]) < f_{\max}. \tag{2.2}$$

In particular, if $f_{\min} > 0$, then $A_n[f]$ is positive definite for all n.

In this paper, we will consider f having zeros. 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. Since f is nonnegative, $f^{(\nu)}(\theta_0) > 0$ and ν must be even. We remark that the condition number of $A_n[f]$ generated by such an f grows like $O(n^{\nu})$, see [**23**].

The convergence rate of the conjugate gradient method is well studied, see Axelsson and Barker [**2**, p.24]. It depends on the condition number of the matrix $A_n[f]$

and how clustered the spectrum of $A_n[f]$ is. If the spectrum is not clustered, as is usually the case for Toeplitz matrices [**17**, p.65], a good estimate of the convergence rate is given in terms of the condition number $\kappa(A_n[f])$ of $A_n[f]$. The estimate can be expressed as

$$\frac{||\mathbf{e_q}||_{A_n[f]}}{||\mathbf{e_0}||_{A_n[f]}} \leq 2\left(\frac{\sqrt{\kappa(A_n[f])}-1}{\sqrt{\kappa(A_n[f])}+1}\right)^q,$$

where $\mathbf{e_q}$ is the error vector at the qth iteration and $||\mathbf{y}||^2_{A_n[f]} \equiv \mathbf{y}^* A_n[f]\mathbf{y}$. Convergence will be slower if f has a zero, in which case the matrices $A_n[f]$ will be ill-conditioned.

One way to speed up the convergence rate of the method is to precondition the Toeplitz system. Thus, instead of solving $A_n\mathbf{x} = \mathbf{b}$, we solve the preconditioned system

$$P_n^{-1}A_n[f]\mathbf{x} = P_n^{-1}\mathbf{b}. \tag{2.3}$$

The matrix P_n, called the preconditioner, should be chosen according to the following criteria:

- P_n should be constructed within $O(n\log n)$ operations.
- $P_n\mathbf{v} = \mathbf{y}$ should be solved in $O(n\log n)$ operations.
- The spectrum of $P_n^{-1}A_n[f]$ should be clustered.

The first two criteria are to keep the operation count per iteration within $O(n\log n)$ as that is the count for the non-preconditioned system. The third criterion comes from the fact that the more clustered the eigenvalues are, the faster the convergence of the method will be, see for instance [**19**, pp. 249-251] and [**2**, pp. 27-28].

In the next three sections, we study different preconditioners for ill-conditioned Toeplitz systems that satisfy the three criteria mentioned above.

3. Toeplitz-type Preconditioners

In [**7**], R. Chan proposed to use band-Toeplitz matrices as preconditioners. The motivation behind using band-Toeplitz matrices is to approximate the generating function f by trigonometric polynomials of fixed degree. The advantage here is that trigonometric polynomials can be chosen to match the zeros of f, so that the preconditioned method still works when f has zeros.

For all $\ell \geq 1$, we define

$$g_\ell(\theta) = (2-2\cos\theta)^\ell = \left[2\sin\left(\frac{\theta}{2}\right)\right]^{2\ell}, \tag{3.1}$$

which has a unique zero of order 2ℓ at $\theta = 0$. We note that $A_n[g_1]$ is the discrete Laplacian given by

$$\text{tridiag}\ (-1, 2, -1)$$

with eigenvalues given by

$$\lambda_j(A_n[g_1]) = 4\sin^2\left(\frac{\pi j}{2n+2}\right), \quad j = 1, \ldots, n. \tag{3.2}$$

The matrix $A_n[g_\ell]$ will be used as our preconditioners. It is therefore necessary that the diagonals of $A_n[g_\ell]$ can be found easily. We remark that

$$2-2\cos\theta = -\frac{1}{z}(1-z)^2 = -(\frac{1}{z}+2-z), \tag{3.3}$$

where $z = e^{i\theta}$. Hence by the binomial theorem,

$$(2 - 2\cos\theta)^\ell = \sum_{k=-\ell}^{\ell} t_k^{(\ell)} z^k, \tag{3.4}$$

where

$$t_j^{(\ell)} = t_{-j}^{(\ell)} = (-1)^j \begin{pmatrix} 2l \\ \ell + j \end{pmatrix},$$

are the binomial coefficients of $(-1)^\ell(1-z)^{2\ell}$. Hence the diagonals of $A_n[g_\ell]$ can be obtained easily from the Pascal triangle. It is clear that $A_n[g_\ell]$ is a symmetric band-Toeplitz matrix of band-width $(2\ell + 1)$.

Next we analyze the spectra of the preconditioned matrices when the generating function has a zero. The following theorem states that the spectra of preconditioned matrices are uniformly bounded.

THEOREM 3.1. *Let $f \in \mathbf{C}_{2\pi}^+$ and have a zero of order 2ν at θ_0. Then the condition number $\kappa(A_n[g_\nu]^{-1}A_n[f])$ of $A_n[g_\nu]^{-1}A_n[f]$ is uniformly bounded for all $n > 0$.*

PROOF. We can assume without loss of generality that $\theta_0 = 0$. We note that the function $\tilde{f} \equiv f(\theta + \theta_0)$ has a zero at $\theta = 0$ and

$$A_n[\tilde{f}] = V_n^* A_n[f] V_n,$$

where $V_n = \text{ diag } (1, e^{-i\theta_0}, e^{-2i\theta_0}, \cdots, e^{-i(n-1)\theta_0})$.

By assumption, there exists a neighborhood N of 0 such that f is continuous in N. Define

$$F(\theta) = \frac{f(\theta)}{(2 - 2\cos\theta)^\nu}.$$

Clearly F is continuous and positive for $\theta \in N \setminus \{0\}$. Since

$$\lim_{\theta \to 0} F(\theta) = \frac{f^{(2\nu)}(0)}{(2\nu)!}$$

is positive, F is a continuous positive function in N. Since f is continuous and positive in $[-\pi, \pi] \setminus N$, we see that F is a continuous function with a positive essential infimum in $[-\pi, \pi]$. Hence there exist constants $c_1, c_2 > 0$, such that $c_1 \leq F(\theta) \leq c_2$ almost everywhere in $[-\pi, \pi]$. Thus we have

$$c_1 \leq \frac{\mathbf{u}^* A_n[f] \mathbf{u}}{\mathbf{u}^* A_n[g_\nu] \mathbf{u}} \leq c_2$$

for any n-vector $\mathbf{u}$. Hence $\kappa(A_n[g_\nu]^{-1}A_n[f]) \leq c_2/c_1$, which is independent of n. $\square$

Next we consider the multiple zero case. Let f be a nonnegative periodic function defined in $[-\pi, \pi]$ with zeros attained at $\{\theta_i\}_{i=1}^k$. Let the order of θ_i be $2\nu_i$ and we order them such that $\nu_1 \leq \cdots \leq \nu_k$. Let $\nu = \sum_{i=1}^k \nu_i$. We define

$$g(\theta) = \prod_{i=1}^{k} [2 - 2\cos(\theta - \theta_i)]^{\nu_i}.$$

The matrix $A_n[g]$ will be used as our preconditioner for $A_n[f]$. To compute the diagonals t_j of B_n, we note that

$$
\begin{aligned}
g(\theta) \;=\; \prod_{j=1}^{k}[(1-e^{i(\theta-\theta_j)})(1-e^{-i(\theta-\theta_j)})]^{\ell_j} &= \prod_{j=1}^{k}[(1-ze^{-i\theta_j})(1-z^{-1}e^{i\theta_j})]^{\ell_j} \\
&= \frac{(-1)^l}{z^l}\prod_{j=1}^{k}(e^{i\theta_j}-2z+z^2e^{-i\theta_j})^{\ell_j} \\
&= \sum_{j=-l}^{l} a_j z^j, \qquad (3.5)
\end{aligned}
$$

where $z = e^{i\theta}$. Thus the diagonals of $A_n[g]$ can be obtained by expanding the product in (3.5). Notice that $A_n[g]$ is a Toeplitz matrix of band-width equals to $(2\ell+1)$. By repeating the arguments in Theorem (3.1), we have

THEOREM 3.2. *There exist constants $c_1, c_2 > 0$, such that*

$$c_1 \le \frac{f(\theta)}{g(\theta)} \le c_2, \quad \forall \theta \in [-\pi,\pi].$$

In particular, $\kappa(A_n[g]^{-1}A_n[f]) \le c_2/c_1$ for all $n > 0$.

We remark that this preconditioner has improved the condition number from $\kappa(A_n[f]) = O(n^{2\nu})$ to $\kappa(A_n[g_\nu]^{-1}A_n[f]) = O(1)$. However, we emphasize that the spectrum of $A_n[g_\nu]^{-1}A_n[f]$ in general will not be clustered around 1 although they are uniformly bounded. The main drawback of using these band-Toeplitz matrices as preconditioners is that when f is positive, these preconditioned systems converge much slower than those preconditioned by circulant preconditioners.

In [**6**], R. Chan and P. Tang designed other kinds of band-Toeplitz preconditioners such that their preconditioned systems converge at the same rate as the circulant preconditioned systems even when f is positive. Their idea is to increase the band-width of the band-Toeplitz preconditioner to get extra degrees of freedom, which enable them not only to match the zeros in f, but also to minimize the relative error $||(f-g)/f||_\infty$ in approximating f by trigonometric polynomials g. The minimizer, which is a trigonometric polynomial, is found by a version of the Remez algorithm proposed by Tang [**26**].

THEOREM 3.3. *Let f be the generating function of Toeplitz matrices and g_ℓ be the minimizer of $||(f-g)/f||_\infty$ over all trigonometric polynomials of degree ℓ. If*

$$\left\| \frac{f-g_\ell}{f} \right\|_\infty = \alpha < 1,$$

then the Toeplitz matrix $A_n[g_\ell]$ is positive definite and

$$\kappa(A_n[g_\ell]^{-1}A_n[f]) \le \frac{1+\alpha}{1-\alpha}, \quad n = 1,2,3,\ldots.$$

PROOF. By assumption, we have

$$f(x)(1-\alpha) \le g_\ell(x) \le f(x)(1+\alpha) \quad \forall x \in [-\pi,\pi].$$

Clearly, $g_\ell(x)$ is nonnegative. In particular, by Theorem 2.1, $A_n[g_\ell]$ is positive definite for all n. Then we have

$$\mathbf{u}^* A_n[f]\mathbf{u} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\left|\sum_{j=0}^{n-1} u_j e^{ijx}\right|^2 dx$$

and

$$\mathbf{u}^* A_n[g_\ell]\mathbf{u} = \frac{1}{2\pi}\int_{-\pi}^{\pi} g_\ell(x)\left|\sum_{j=0}^{n-1} u_j e^{ijx}\right|^2 dx$$

for an arbitrary (complex) n-vector $\mathbf{u}^* = [u_0, u_1, \cdots, u_{n-1}]$. Hence, we get

$$(1-\alpha)\mathbf{u}^* A_n[f]\mathbf{u} \le \mathbf{u}^* A_n[g_\ell]\mathbf{u} \le (1+\alpha)\mathbf{u}^* A_n[f]\mathbf{u}.$$

Since $A_n[g_\ell]$ is positive definite, we finally have

$$\kappa(A_n[g_\ell]^{-1}A_n[f]) \le \frac{1+\alpha}{1-\alpha}.$$

□

The parameter α is given explicitly in the Remez algorithm. It gives an a priori bound on the number of iterations required for convergence.

The main idea behind Theorems 3.1 and 3.3 is to approximate the given nonnegative generating function f by trigonometric polynomials that match the zeros of f. Clearly, any function g that matches the zeros of f and gives rise to Toeplitz matrices that are easily invertible can be considered too. This idea is exploited in Di Benedetto [**3**], Di Benedetto, Fiorentino, and Serra [**5**], and Serra [**22**]. In [**5**], f is first approximated by g_ν as in (3.1), then the quotient f/g_ν is further approximated by a trigonometric polynomial or rational function to enhance the convergence rate.

4. Circulant-type Preconditioners

Circulant matrices, as preconditioners for Toeplitz systems in the preconditioned conjugate gradient method, have been studied extensively since 1986, see [**8**]. We remark that a matrix C is called circulant if it is Toeplitz and the last entry of every row is the first entry of its succeeding row. It has been shown that the circulant preconditioners are good preconditioners for the solutions of Toeplitz systems generated by positive function, see [**8**]. However, when $f_{\min} = 0$, Tyrtyshnikov has proved theoretically [**28**] that Strang's S_n and T. Chan's T_n preconditioners will fail in this case. In fact, he showed that the numbers of outlying eigenvalues of $S_n^{-1}A_n[f]$ and $T_n^{-1}A_n[f]$ are of $O(n^{\nu/(\nu+\mu)})$ and $O(n^{\nu/(\nu+1)})$, respectively. Here, μ is the degree of smoothness of the function f, and ν is the order of f at the zeros. These results were numerically verified in Tyrtyshnikov and Strela [**29**].

In [**20**], Potts and Steidl proposed to use f to generate ω-circulant preconditioners to precondition $A_n[f]$. Circulant matrices belong to the class of $\{\omega\}$-circulant matrices [**15**] which are defined as follows:

Definition 4.1. Let $\omega = e^{i\theta_0}$ with $\theta_0 \in [-\pi, \pi]$. An n-by-n matrix W is said to be an n-by-n *$\{\omega\}$-circulant matrix* if it has the spectral decomposition

$$C_{\omega,n} = \Omega_n^* F_n^* \Lambda_n F_n \Omega_n. \tag{4.1}$$

Here $\Omega_n = \mathrm{diag}[1, \omega^{-1/n}, \ldots, \omega^{-(n-1)/n}]$ and Λ_n is a diagonal matrix containing the eigenvalues of W_n.

Notice that $\{\omega\}$-circulant matrices are Toeplitz matrices with the first entry of each column obtained by multiplying the last entry of the preceding column by ω. In particular, $\{1\}$-circulant matrices are circulant matrices while $\{-1\}$-circulant matrices are skew-circulant matrices.

The diagonal matrix Λ_n in (4.1) can be obtained in $O(n \log n)$ operations by taking the FFT of the first column of $C_{\omega,n}$, i.e.,

$$\Lambda_n \mathbf{1} = F_n \Omega_n W_n \mathbf{e}_1.$$

Once Λ_n is obtained, the products $C_{\omega,n}\mathbf{y}$ and $C_{\omega,n}^{-1}\mathbf{y}$ for any vector $\mathbf{y}$ can be computed via (4.1) in $O(n \log n)$ operations.

In the following, we assume that f has only a finite number of zeros. Then we can choose ζ such that $f(2\pi j/n + \zeta) > 0$ for $0 \le j < n$. Using the values $f(2\pi j/n + \zeta)$, we can construct the ω-circulant preconditioner

$$C_{\omega,n}[f] = \Omega_n^* F_n^* \ \mathrm{diag}\left(f(\zeta), f(\frac{2\pi}{n} + \zeta), \cdots f(\frac{2(n-1)\pi}{n} + \zeta) \right) F_n \Omega_n$$

to precondition $A_n[f]$. Here $\omega = e^{i\zeta}$. Potts and Steidl [**20**] proved the following theorem about the clustering property of $C_{\omega,n}[f]^{-1}A_n[f]$.

THEOREM 4.2. *Let $f \in \mathbf{C}_{2\pi}^{+}$ and have a zero of order $2p$ at θ_0. Let $\omega = e^{i\theta_0}$. For any given $\epsilon > 0$, there exist positive integers N_1 and N_2 such that for $n \ge N_2$, $C_{\omega,n}[f]^{-1}A_n[f]$ has at most N_2 eigenvalues outside the interval* $(1-\epsilon, 1+\epsilon)$.

PROOF. Define $s_{2p}(\theta) \equiv \sin^{2p}(\theta - \theta_0)$. By assumption, $f(\theta) = s_{2p}(\theta) g(\theta)$ for some positive continuous function $g(\theta)$ on $[-\pi, \pi]$. By the Weierstrass theorem (see Cheney [**14**, p.144]), given any $\epsilon > 0$, there exists a positive trigonometric polynomial

$$g_k(\theta) = \sum_{j=-k}^{k} \rho_k e^{ik\theta}$$

with $\rho_{-k} = \bar{\rho}_k$ such that $g_k(\theta)$ satisfies the following condition

$$g_k(\theta) - \frac{1}{2}\epsilon g_{\min} \le g(\theta) \le g_k(\theta) + \frac{1}{2}\epsilon g_{\min}. \tag{4.2}$$

It follows that

$$s_{2p}(\theta)\left(g_k(\theta) - \frac{1}{2}\epsilon g_{\min} \right) \le f(\theta) \le s_{2p}(\theta)\left(g_k(\theta) + \frac{1}{2}\epsilon g_{\min} \right).$$

Then we obtain

$$\frac{\mathbf{x}^* A_n[f]\mathbf{x}}{\mathbf{x}^* C_{\omega,n}[f]\mathbf{x}} \le \frac{\mathbf{x}^* A_n[s_{2p}g]\mathbf{x}}{\mathbf{x}^* C_{\omega,n}[f]\mathbf{x}} + \frac{1}{2}\epsilon g_{\min} \frac{\mathbf{x}^* A_n[s_{2p}]\mathbf{x}}{\mathbf{x}^* C_{\omega,n}[f]\mathbf{x}}. \tag{4.3}$$

As $s_{2p}(\theta + \theta_0)$ is a p-th degree trigonometric polynomial, $A_n[s_{2p}(\theta + \theta_0)]$ is a banded Toeplitz matrix with half bandwidth $p+1$. Therefore when $n > 2p$, by the definition of the Strang circulant preconditioner, we have

$$C_{0,n}\left[s_{2p}(\theta + \theta_0)\right] = A_n\left[s_{2p}(\theta + \theta_0)\right] + \begin{bmatrix} 0 & 0 & L_p \\ 0 & 0 & 0 \\ L_p^* & 0 & 0 \end{bmatrix}, \tag{4.4}$$

where L_p is a p-by-p matrix, see [**20, 25**]. By multiplying

$$\Omega = \text{ diag } \left(1, e^{-i\theta_0/n}, \ldots, e^{-i(n-1)\theta_0/n}\right)$$

and Ω^* on (4.4), we get

$$C_{\omega,n}[s_{2p}(\theta)] = \Omega^* C_{0,n}[s_{2p}(\theta+\theta_0)]\Omega = A_n[s_{2p}(\theta)] + R_n(2p), \tag{4.5}$$

where $R_n(2p)$ denotes an n-by-n matrix of rank at most $2p$. We also note by [**4**, Theorem 3.1] that

$$A_n[s_{2p}(\theta)g_k] = A_n[s_{2p}(\theta)]A_n[g_k] + R_n(2k+2p). \tag{4.6}$$

By putting (4.5) and (4.6) into (4.3), we get

$$\begin{aligned}\frac{\mathbf{x}^* A_n[f]\mathbf{x}}{\mathbf{x}^* C_{\omega,n}[f]\mathbf{x}} \quad \le \quad & \frac{\mathbf{x}^* C_{\omega,n}[s_{2p}(\theta)]C_{\omega,n}[g]\mathbf{x}}{\mathbf{x}^* C_{\omega,n}[f]\mathbf{x}} + \frac{\mathbf{x}^* R_n(q)\mathbf{x}}{\mathbf{x}^* C_{\omega,n}[f]\mathbf{x}} + \\ & \frac{1}{2}\epsilon g_{\min}\frac{\mathbf{x}^* C_n[s_{2p}(\theta)]\mathbf{x}}{\mathbf{x}^* C_{\omega,n}[f]\mathbf{x}} + \frac{1}{2}\epsilon g_{\min}\frac{\mathbf{x}^* R_n(2p)\mathbf{x}}{\mathbf{x}^* C_{\omega,n}[f]\mathbf{x}}\end{aligned}$$

with $q = 4k + 4p + \min\{2k, 2p\}$. By noting that

$$\frac{\mathbf{x}^* C_n[s_{2p}(\theta)]\mathbf{x}}{\mathbf{x}^* C_{\omega,n}[f]\mathbf{x}} \le \frac{1}{g_{\min}}$$

and using (4.2), we have

$$\frac{\mathbf{x}^*(A_n[f] - R_n(q+2p))\mathbf{x}}{\mathbf{x}^* C_{\omega,n}[f]\mathbf{x}} \le 1 + \epsilon.$$

Similarly, we can show the following inequality

$$1 - \epsilon \le \frac{\mathbf{x}^*(A_n[f] - R_n(q+2p))\mathbf{x}}{\mathbf{x}^* C_{\omega,n}[f]\mathbf{x}}.$$

Hence at most $(q + 2p)$ eigenvalues of $C_{\omega,n}[f]^{-1}A_n[f]$ are not contained in the interval $[1-\epsilon, 1+\epsilon]$. The result follows. $\square$

The clustering of the eigenvalues of $C_{\omega,n}[f]^{-1}A_n[f]$ for f having multiple zeros follows in a similar way. In [**20**], Potts and Steidl have also extended this approach to non-Hermitian Toeplitz matrices and doubly symmetric block Toeplitz matrices with Toeplitz blocks. Besides using fast Fourier transforms, they also employ other fast trigonometric transforms to construct preconditioners. Numerical results have shown that this kind of preconditioner is very effective to ill-conditioned Toeplitz systems.

5. Unknown Generating Functions

The approaches in the last two sections work when the generating function f is given explicitly, i.e., all Fourier coefficients $\{a_j\}_{j=-\infty}^{\infty}$ of f are available. However, when we are given only a finite n-by-n Toeplitz system, i.e., only $\{a_j\}_{|j|<n}$ are given and the underlying f is unknown, then banded-Toeplitz or ω-circulant preconditioners cannot be constructed. In contrast, most well-known circulant preconditioners, such as the Strang and the T. Chan preconditioners, are defined using only the entries of the given Toeplitz matrix. Di Benedetto in [**3**] has proved that the condition numbers of the preconditioned matrices by sine transform preconditioners are uniformly bounded. However, the preconditioners themselves may be

singular or indefinite in general. In this section, we develop a family of positive definite circulant preconditioners that work for ill-conditioned Toeplitz systems and do not require the explicit knowledge of f, i.e., they require only $\{a_j\}_{|j|<n}$ for an n-by-n Toeplitz system.

Our idea is based on the unified approach proposed in Chan and Yeung [**11**], where they showed that circulant preconditioners can be derived in general by convolving the generating function f with some kernels. For instance, convolving f with the Dirichlet kernel $\mathcal{D}_{\lfloor n/2 \rfloor}$ gives the Strang preconditioner. They proved that for any positive 2π-periodic continuous function f, if $\mathcal{C}_n$ is a kernel such that the convolution product $\mathcal{C}_n * f$ tends to f uniformly on $[-\pi, \pi]$, then the corresponding circulant preconditioned matrix $C_n^{-1} A_n$ will have clustered spectrum. In particular, the conjugate gradient method will converge superlinearly when solving the preconditioned system. This result turns the problem of finding a good preconditioner to the problem of approximating f with $\mathcal{C}_n * f$. Notice that $\mathcal{D}_{\lfloor n/2 \rfloor} * f$, being the partial sum of f, depends solely on the first $\lfloor n/2 \rfloor$ Fourier coefficients $\{a_j\}_{|j|<\lfloor n/2 \rfloor}$ of f. Thus the Strang preconditioner, and similarly for other circulant preconditioners constructed through kernels, does not require the explicitly knowledge of f.

In [**10**], Chan, Tso and Sun constructed circulant preconditioners through approximate convolution identities such as B-spline preconditioners. These are preconditioners obtained by convolution the Fourier transform of the B-splines with the generating function f. Well-known circulant preconditioners such as the Strang and the T. Chan circulant preconditioners fall into this class. They correspond to the B-spline preconditioners of order 0 and 1 respectively. The intuitive reason for using B-spline is that we can approximate f as accurately as possible by using higher order B-splines as they form a sequence of converging to the Dirac delta function (approximate convolution identity). The convolution is done in the Fourier space so that we need only the Fourier coefficients of f. Thus the method is applicable if we are only given the Toeplitz matrix but not f. B-splines also have an advantage that the higher order B-splines can be constructed recursively from the lower order B-splines by self-convolution. For all mildly ill-conditioned Toeplitz systems considered in [**10**], the numbers of iterations required for convergence after circulant preconditioning do not grow with n, provided that a B-spline of high enough order is used.

Recently, Potts and Steidl [**21**] have used convolution products that match the zeros of f to construct the circulant-type preconditioners. In particular, they have used the generalized Jackson kernels and B-spline kernels in their construction. They have proved that both the convolution products of Jackson and B-spline kernels with f can match the zeros of f. However, in order to guarantee that the preconditioners are positive definite, the position of the zeros of f is required which in general may not be readily available. In [**12**], Chan, Yip and Ng have also used the generalized Jackson kernel functions to construct circulant preconditioners for ill-conditioned Toeplitz systems. In contrast, our circulant preconditioners can be constructed without any explicit knowledge of the zeros of f.

The generalized Jackson kernel functions is defined as follows

$$\mathcal{K}_{m,2r}(\theta) \equiv \frac{k_{m,2r}}{m^{2r-1}} \left(\frac{\sin(\frac{m\theta}{2})}{\sin(\frac{\theta}{2})} \right)^{2r}, \quad r = 1, 2, \dots, \tag{5.1}$$

where $k_{m,2r}$ is a normalization constant such that $\int_{-\pi}^{\pi} \mathcal{K}_{m,2r}(\theta)d\theta = 1$. It is known that $k_{m,2r}$ is bounded above and below by constants independent of m, see [**18**, p.57] or [**12**, Lemma 3.2]. We note that $\mathcal{K}_{m,2}(\theta)$ is the Fejér kernel and $\mathcal{K}_{m,4}(\theta)$ is the Jackson kernel [**18**, p.57]. For any m, the Fejér kernel $\mathcal{K}_{m,2}(\theta)$ can be expressed as

$$\mathcal{K}_{m,2}(\theta) = \sum_{k=-m+1}^{m-1} b_k^{(m,2)} e^{ik\theta},$$

where

$$b_k^{(m,2)} = \frac{m-|k|}{2\pi m}, \quad k = 0, \pm 1, \pm 2, \cdots, \pm(m-1),$$

see for instance [**11**]. Note that $\int_{-\pi}^{\pi} \mathcal{K}_{m,2}(\theta)d\theta = 2\pi b_0^{(m,2)} = 1$.

By (5.1), we see that $\mathcal{K}_{m,2r}(\theta)$ is the r-th power of $\mathcal{K}_{m,2}(\theta)$ up to a scaling. Hence we have

$$\mathcal{K}_{m,2r}(\theta) = \sum_{k=-r(m-1)}^{r(m-1)} b_k^{(m,2r)} e^{ik\theta}, \tag{5.2}$$

where the coefficients $b_k^{(m,2r)}$ can be obtained by convolving the vector $(b_{-m+1}^{(m,2)}, \cdots, b_0^{(m,2)}, \cdots, b_{m-1}^{(m,2)})$ with itself for $r-1$ times and this can be done by fast Fourier transforms, see [**24**, pp.294–296]. Thus the cost of computing the coefficients $\{b_k^{(m,2r)}\}$ for all $|k| \le r(m-1)$ is of order $O(rm \log m)$ operations. In order to guarantee that $\int_{-\pi}^{\pi} \mathcal{K}_{m,2r}(\theta)d\theta = 1$, we can normalize $b_0^{(m,2r)}$ to $1/(2\pi)$ by dividing all coefficients $b_k^{(m,2r)}$ by $2\pi b_0^{(m,2r)}$.

The convolution product of two arbitrary functions $g = \sum_{k=-\infty}^{\infty} b_k e^{ik\theta}$ and $h = \sum_{k=-\infty}^{\infty} c_k e^{ik\theta}$ in $\mathbf{C}_{2\pi}$ is defined as

$$(g * h)(\theta) \equiv \int_{-\pi}^{\pi} g(t)h(\theta - t)dt = 2\pi \sum_{k=-\infty}^{\infty} b_k c_k e^{ik\theta}. \tag{5.3}$$

When we are given an n-by-n Toeplitz matrix $A_n[f]$, our proposed circulant preconditioner is $C_n[\mathcal{K}_{m,2r} * f]$, where $m = \lceil n/r \rceil$, i.e.,

$$r(m-1) < n \le rm. \tag{5.4}$$

By (5.2) and (5.3), since $f = \sum_{k=-\infty}^{\infty} a_k e^{ik\theta}$, the convolution product of $\mathcal{K}_{m,2r} * f$ is given by

$$(\mathcal{K}_{m,2r} * f)(\theta) = 2\pi \sum_{k=-r(m-1)}^{r(m-1)} a_k b_k^{(m,2r)} e^{ik\theta} = \sum_{k=-n+1}^{n-1} d_k e^{ik\theta}, \tag{5.5}$$

where

$$d_k = \begin{cases} 2\pi a_k b_k^{(m,2r)}, & |k| \le r(m-1), \\ 0, & \text{otherwise.} \end{cases}$$

Clearly, $\mathcal{K}_{m,2r} * f$ depends only on a_k for $|k| < n$, i.e., only on the entries of the given n-by-n Toeplitz matrix $A_n[f]$. For a given function g, we define the circulant preconditioner $C_n[g]$ to be the n-by-n circulant matrix with its j-th eigenvalue given by

$$\lambda_j(C_n[g]) = g\left(\frac{2\pi j}{n}\right), \quad 0 \le j < n. \tag{5.6}$$

Notice that by (5.6), to construct our preconditioner $C_n[\mathcal{K}_{m,2r} * f]$, we only need the values of $\mathcal{K}_{m,2r} * f$ at $2\pi j/n$ for $0 \le j < n$. By (5.5), these values can be obtained by taking one fast Fourier transform of length n. Thus the cost of constructing $C_n[\mathcal{K}_{m,2r} * f]$ is of $O(n \log n)$ operations.

THEOREM 5.1. *Let $f \in \mathbf{C}_{2\pi}^+$. The preconditioner $C_n[\mathcal{K}_{m,2r} * f]$ is positive definite for all positive integers m, n and r.*

PROOF. By (5.1), $\mathcal{K}_{m,2r}(\theta)$ is positive except at $\theta = 2k\pi/m$, $k = \pm 1, \pm 2, \ldots, \pm(n-1)$. Since $f \in \mathbf{C}_{2\pi}^+$ is nonnegative and not identically zero, the function

$$(\mathcal{K}_{m,2r} * f)(\theta) \equiv \int_{-\pi}^{\pi} \mathcal{K}_{m,2r}(t) f(\theta - t) dt$$

is clearly positive for all $\theta \in [-\pi, \pi]$. Hence by (5.6), the preconditioners $C_n[\mathcal{K}_{m,2r} * f]$ are positive definite. $\square$

In the following, we analyze the spectra of the circulant preconditioned matrices when the generating function has a zero. The case where the generating function has multiple zeros can be found in [**9**].

For simplicity, we will use θ to denote the function θ defined on the whole real line $\mathbb{R}$ in the following discussion. For clarity, we will use $\theta_{2\pi}$ to denote the periodic extension of θ on $[-\pi, \pi]$, i.e. $\theta_{2\pi}(\theta) = \tilde{\theta}$ if $\theta = \tilde{\theta} \pmod{2\pi}$ and $\tilde{\theta} \in [-\pi, \pi]$. It is clear that $\theta_{2\pi} \in \mathbf{C}_{2\pi}^+$. In order to estimate $(\mathcal{K}_{m,2r} * \theta_{2\pi}^{2p})(\phi)$ for $\phi \neq 0$, we need to replace the function $\theta_{2\pi}^{2p}$ in the convolution product by θ^{2p} defined on $\mathbb{R}$.

LEMMA 5.2. *Let p be a positive integer. Then*

$$1 \le \frac{\left(\mathcal{K}_{m,2r} * \theta^{2p}\right)(\phi)}{\left(\mathcal{K}_{m,2r} * \theta_{2\pi}^{2p}\right)(\phi)} \le 3^{2p}, \quad \forall \phi \in [-\frac{\pi}{2}, \frac{\pi}{2}]. \tag{5.7}$$

The next theorem states that $\mathcal{K}_{m,2r} * \theta_{2\pi}^{2p}$ and $\theta_{2\pi}^{2p}$ are essentially the same away from the zero of $\theta_{2\pi}^{2p}$. We remark the proof of the following theorems can be found in [**12**]. Here we give a proof by using the fact that $\mathcal{K}_{m,2r}$ is a summation kernel [**18**] with the following property

$$\lim_{m \to \infty} \|\mathcal{K}_{m,2r} * f - f\|_\infty = 0 \quad \forall f \in \mathbf{C}_{2\pi}, \tag{5.8}$$

where $\|\cdot\|_\infty$ denotes the supremum norm.

THEOREM 5.3. *Let $f \in \mathbf{C}_{2\pi}^+$ and have a zero of order $2p$ at $\theta_0 \in [-\pi, \pi]$. Let $r > p$ be any integer and $m = \lceil n/r \rceil$. Then there exist positive numbers α and β, independent of n, such that for all sufficiently large n,*

$$\alpha \le \frac{\left(\mathcal{K}_{m,2r} * f\right)(\phi)}{f(\phi)} \le \beta, \qquad \forall\ \frac{\pi}{n} \le |\phi - \theta_0| \le \pi.$$

PROOF. By the definition of zeros (see §2), $f(\theta) = (\theta - \theta_0)_{2\pi}^{2p} g(\theta)$ for some positive continuous function $g(\theta)$ on $[-\pi, \pi]$. Write

$$\frac{\left(\mathcal{K}_{m,2r} * f\right)(\phi)}{f(\phi)} = \frac{\left[\mathcal{K}_{m,2r} * (\theta - \theta_0)_{2\pi}^{2p} g(\theta)\right](\phi)}{\left[\mathcal{K}_{m,2r} * (\theta - \theta_0)_{2\pi}^{2p}\right](\phi)} \cdot \frac{\left[\mathcal{K}_{m,2r} * (\theta - \theta_0)_{2\pi}^{2p}\right](\phi)}{(\phi - \theta_0)_{2\pi}^{2p}} \cdot \frac{1}{g(\phi)}.$$

Clearly the last factor is uniformly bounded above and below by positive constants. As for the first factor, by the Mean Value Theorem for integrals, there exists a $\zeta \in [-\pi, \pi]$ such that

$$[\mathcal{K}_{m,2r} * (\theta-\theta_0)_{2\pi}^{2p} g(\theta)](\phi) = g(\zeta)[\mathcal{K}_{m,2r} * (\theta-\theta_0)_{2\pi}^{2p}](\phi).$$

Hence

$$0 < g_{\min} \leq \frac{\left[\mathcal{K}_{m,2r} * (\theta-\theta_0)_{2\pi}^{2p} g(\theta)\right](\phi)}{\left[\mathcal{K}_{m,2r} * (\theta-\theta_0)_{2\pi}^{2p}\right](\phi)} \leq g_{\max}, \quad \forall \phi \in [-\pi, \pi],$$

where $g_{\min}$ and $g_{\max}$ are the minimum and maximum of g respectively.

For the second factor, we see from Lemma 5.2 that $(\mathcal{K}_{m,2r} * \theta_{2\pi}^{2p})(\phi)$ can be replaced by an another function for $\phi \in [-\pi/2, \pi/2]$. Hence, we proceed the proof by considering ϕ in $[-\pi/2, \pi/2]$. We first consider $\phi \in [\pi/n, \pi/2]$. By the binomial expansion,

$$\begin{aligned} \left(\mathcal{K}_{m,2r} * \theta^{2p}\right)(\phi) &\equiv \int_{-\pi}^{\pi} \mathcal{K}_{m,2r}(t)(\phi-t)^{2p} dt \\ &= \int_{-\pi}^{\pi} \mathcal{K}_{m,2r}(t) \sum_{k=0}^{2p} \binom{2p}{k} \phi^{2p-k}(-t)^k dt. \end{aligned} \tag{5.9}$$

For odd k, $\int_{-\pi}^{\pi} \mathcal{K}_{m,2r}(t) t^k dt = 0$. Thus

$$\frac{\left(\mathcal{K}_{m,2r} * \theta^{2p}\right)(\phi)}{\phi_{2\pi}^{2p}} = \frac{\left(\mathcal{K}_{m,2r} * \theta^{2p}\right)(\phi)}{\phi^{2p}} = \sum_{k=0}^{p} \binom{2p}{2k} \phi^{-2k} \int_{-\pi}^{\pi} \mathcal{K}_{m,2r}(t) t^{2k} dt.$$

By using $\int_{-\pi}^{\pi} \mathcal{K}_{m,2r}(t) t^{2k} dt = c_{k,2r}/m^{2k}$ in [**12**, Lemma 3.2], we then have

$$\frac{\left(\mathcal{K}_{m,2r} * \theta^{2p}\right)(\phi)}{\phi_{2\pi}^{2p}} = \sum_{k=0}^{p} \binom{2p}{2k} \frac{c_{k,2r}}{\phi^{2k} m^{2k}}, \tag{5.10}$$

where $c_{k,2r}$ are bounded above and below by positive constants independent of m for $k = 0, \ldots p$, see [**12**, Lemma 3.2]. Since by (5.4), $\pi/r \leq \pi m/n \leq \phi m$, we have

$$c_{0,2r} \leq \sum_{k=0}^{p} \binom{2p}{2k} \frac{c_{k,2r}}{\phi^{2k} m^{2k}} \leq \sum_{k=0}^{p} \left(\frac{r}{\pi}\right)^{2k} \binom{2p}{2k} c_{k,2r}.$$

Thus by (5.10),

$$c_{0,2r} \leq \frac{\left(\mathcal{K}_{m,2r} * \theta^{2p}\right)(\phi)}{\phi_{2\pi}^{2p}} \leq \sum_{k=0}^{p} \left(\frac{r}{\pi}\right)^{2k} \binom{2p}{2k} c_{k,2r}.$$

The case with $\phi \in [-\pi/2, -\pi/n]$ is similar to the case where $\phi \in [\pi/n, \pi/2]$. Hence by (5.7), the second factor is uniformly bounded for $\pi/n \leq |\phi - \theta_0| \leq \pi/2$.

Next we consider the case ϕ in $[-\pi, -\pi/2]$ or $[\pi/2, \pi]$. By using (5.8), it implies that there exists a positive integer m_0 (or a positive integer n_0, cf. (5.4)) such that for $m \geq m_0$

$$-\alpha_1 \leq \left(\mathcal{K}_{m,2r} * |\theta|^p\right)(\phi) - |\phi|_{2\pi}^p \leq \alpha_1, \quad \forall \phi \in [-\pi, \pi],$$

where α_1 is a constant less than π^p. It follows that

$$1 - \frac{\alpha_1}{\pi^{2p}} \leq \frac{\left(\mathcal{K}_{m,2r} * \theta^{2p}\right)(\phi)}{\phi_{2\pi}^{2p}} \leq 1 + \frac{\beta_1}{\left(\frac{\pi}{2}\right)^{2p}}, \quad \phi \in [-\pi, -\frac{\pi}{2}] \text{ or } [\frac{\pi}{2}, \pi].$$

By combining the above results, the second factor is uniformly bounded above and below. Hence the result follows. □

So far we have considered only the interval $\pi/n \le |\phi - \theta_0| \le \pi$. For $|\phi - \theta_0| \le \pi/n$, we now show that the convolution product $\mathcal{K}_{m,2r} * f$ matches the order of the zero of f at the zero of f.

THEOREM 5.4. *Let $f \in \mathbf{C}_{2\pi}^{+}$ and have a zero of order $2p$ at $\theta_0 \in [-\pi, \pi]$. Let $r > p$ be any integer and $m = \lceil n/r \rceil$. Then for any $|\phi - \theta_0| \le \pi/n$, we have*

$$(\mathcal{K}_{m,2r} * f)(\phi) = O\left(\frac{1}{n^{2p}}\right).$$

PROOF. By (5.10), we can obtain $\left(\mathcal{K}_{m,2r} * \theta_{2\pi}^{2p}\right)(\phi) \le O(1/n^{2p})$ for $|\phi - \theta_0| \le \pi/n$. On the other hand, from (5.9), we have

$$(\mathcal{K}_{m,2r} * \theta^{2p})(\phi) \ge \int_{-\pi}^{\pi} \mathcal{K}_{m,2r}(t) t^{2p} dt = \frac{c_{p,2r}}{m^{2p}} = O\left(\frac{1}{n^{2p}}\right).$$

Hence the theorem follows. □

With Theorems 5.3 and 5.4, we have our main theorem which states the spectra of the preconditioned matrices are essentially bounded.

THEOREM 5.5. *Let $f \in \mathbf{C}_{2\pi}^{+}$ and have a zero of order $2p$ at θ_0. Let $r > p$ and $m = \lceil n/r \rceil$. Then there exist positive numbers $\alpha < \beta$, independent of n, such that for all sufficiently large n, at most $2p + 1$ eigenvalues of $C_n^{-1}[\mathcal{K}_{m,2r} * f]A_n[f]$ are outside the interval $[\alpha, \beta]$.*

PROOF. For any function $g \in \mathbf{C}_{2\pi}$, we let $\tilde{C}_n[g]$ to be the n-by-n circulant matrix with the j-th eigenvalue given by

$$\lambda_j(\tilde{C}_n[g]) = \begin{cases} \dfrac{1}{n^{2p}}, & \text{if } \left|\dfrac{2\pi j}{n} - \gamma\right| < \dfrac{\pi}{n} \\ g\left(\dfrac{2\pi j}{n}\right), & \text{otherwise,} \end{cases} \tag{5.11}$$

for $j = 0, \ldots, n-1$. Since there is at most one j such that $|2\pi j/n - \gamma| < \pi/n$, by (5.6), $\tilde{C}_n[g] - C_n[g]$ is a matrix of rank at most 1.

By assumption, $f(\theta) = \sin^{2p}((\theta - \gamma)/2) g(\theta)$ for some positive function g in $\mathbf{C}_{2\pi}$. We use the following decomposition of the Rayleigh quotient to prove the theorem:

$$\begin{aligned} \frac{\mathbf{x}^* A_n[f]\mathbf{x}}{\mathbf{x}^* C_n[\mathcal{K}_{m,2r} * f]\mathbf{x}} \quad = \quad & \frac{\mathbf{x}^* A_n[f]\mathbf{x}}{\mathbf{x}^* A_n\left[\sin^{2p}\left(\frac{\theta-\gamma}{2}\right)\right]\mathbf{x}} \cdot \frac{\mathbf{x}^* A_n\left[\sin^{2p}\left(\frac{\theta-\gamma}{2}\right)\right]\mathbf{x}}{\mathbf{x}^* \tilde{C}_n\left[\sin^{2p}\left(\frac{\theta-\gamma}{2}\right)\right]\mathbf{x}} \cdot \\ & \cdot \frac{\mathbf{x}^* \tilde{C}_n\left[\sin^{2p}\left(\frac{\theta-\gamma}{2}\right)\right]\mathbf{x}}{\mathbf{x}^* \tilde{C}_n[f]\mathbf{x}} \cdot \frac{\mathbf{x}^* \tilde{C}_n[f]\mathbf{x}}{\mathbf{x}^* \tilde{C}_n[\mathcal{K}_{m,2r} * f]\mathbf{x}} \cdot \\ & \cdot \frac{\mathbf{x}^* \tilde{C}_n[\mathcal{K}_{m,2r} * f]\mathbf{x}}{\mathbf{x}^* C_n[\mathcal{K}_{m,2r} * f]\mathbf{x}}. \end{aligned} \tag{5.12}$$

We remark that by Theorem 3.1 and the definitions of C_n and $\tilde{C}_n$, all matrices in the factors in the right hand side of (5.12) are positive definite.

By Theorem 3.1, the first factor in the right hand side of (5.12) is uniformly bounded above and below. Similarly, by (5.11), the third factor is also uniformly bounded. The eigenvalues of the two circulant matrices in the fourth factor differ only when $|2\pi j/n - \gamma| \geq \pi/n$. But by Theorem 5.3, the ratios of these eigenvalues are all uniformly bounded when n is large. The eigenvalues of the two circulant matrices in the last factor differ only when $|2\pi j/n - \gamma| < \pi/n$. But by Theorem 5.4, their ratios are also uniformly bounded.

It remains to handle the second factor. We note that when $n > 2p$,

$$C_n\left[s_{2p}(\theta)\right] = A_n\left[s_{2p}(\theta)\right] + \begin{bmatrix} 0 & 0 & R_p \\ 0 & 0 & 0 \\ R_p^* & 0 & 0 \end{bmatrix}, \tag{5.13}$$

where R_p is a p-by-p matrix, see (4.4). Thus $A_n[s_{2p}(\theta)] = \tilde{C}_n[s_{2p}(\theta)] + R_n$ where the n-by-n matrix R_n is of rank at most $2p+1$. By using a similar argument in Theorem 4.2, we can prove that at most $2p+1$ eigenvalues of $C_n[\mathcal{K}_{m,2r} * f]^{-1}A_n[f]$ are are outside the interval $[\alpha, \beta]$. Hence the theorem follows. ☐

Finally we use a trick proposed in [**3, 20**] to prove that all the eigenvalues of the preconditioned matrices are bounded from below by a constant independent of n. Hence the computational cost for solving this class of n-by-n Toeplitz systems will be of $O(n \log n)$ operations.

THEOREM 5.6. *Let $f \in \mathbf{C}_{2\pi}^+$ and have a zero of order $2p$ at θ_0. Let $r > p$ and $m = \lceil n/r \rceil$. Then there exists a constant c independent of n, such that for all n sufficiently large, all eigenvalues of the preconditioned matrix $C_n^{-1}[\mathcal{K}_{m,2r} * f]A_n[f]$ are larger than c.*

PROOF. In view of the proof of Theorem 5.5, it suffices to get a lower bound of the second Rayleigh quotient in the right hand side of (5.12). Equivalently, we have to get an upper bound of $\rho(A_n^{-1}\left[s_{2p}(\theta)\right]\tilde{C}_n\left[s_{2p}(\theta)\right])$, where $\rho(\cdot)$ denotes the spectral radius and $s_{2p}(\theta) = \sin^{2p}\left(\frac{\theta-\gamma}{2}\right)$.

We note that by the definition (5.11), $\tilde{C}_n\left[s_{2p}(\theta)\right] = C_n\left[s_{2p}(\theta)\right] + E_n$, where E_n is either the zero matrix or is given by

$$F_n^* \mathrm{diag}\left(\cdots, 0, \frac{1}{n^{2p}} - s_{2p}\left(\frac{2\pi j}{n}\right), 0, \cdots\right) F_n$$

for some j such that $|2\pi j/n - \gamma| < \pi/n$. Thus $\|E_n\|_2 = O(1/n^{2p})$.

By Theorem 2.1, $A_n^{-1}[s_{2p}(\theta)]$ is positive definite. Thus the matrix

$$A_n^{-1}\left[s_{2p}(\theta)\right]\tilde{C}_n\left[s_{2p}(\theta)\right]$$

is similar to the symmetric matrix

$$A_n^{-1/2}\left[s_{2p}(\theta)\right]\tilde{C}_n\left[s_{2p}(\theta)\right]A_n^{-1/2}\left[s_{2p}(\theta)\right].$$

Hence we have

$$\begin{aligned}
& \rho\left(A_n^{-1}\left[s_{2p}(\theta)\right]\tilde{C}_n\left[s_{2p}(\theta)\right]\right) \\
= \; & \rho\left(A_n^{-1/2}\left[s_{2p}(\theta)\right]\tilde{C}_n\left[s_{2p}(\theta)\right]A_n^{-1/2}\left[s_{2p}(\theta)\right]\right) \\
\leq \; & \rho\left(A_n^{-1/2}\left[s_{2p}(\theta)\right]C_n\left[s_{2p}(\theta)\right]A_n^{-1/2}\left[s_{2p}(\theta)\right]\right) \\
& \qquad + \rho\left(A_n^{-1/2}\left[s_{2p}(\theta)\right]E_n A_n^{-1/2}\left[s_{2p}(\theta)\right]\right) \\
\leq \; & \rho\left(A_n^{-1}\left[s_{2p}(\theta)\right]C_n\left[s_{2p}(\theta)\right]\right) + \left\|A_n^{-1}\left[s_{2p}(\theta)\right]\right\|_2 \|E_n\|_2. \qquad (5.14)
\end{aligned}$$

By [**7**, Theorem 1], we have $\left\|A_n^{-1}\left[s_{2p}(\theta)\right]\right\|_2 = O(n^{2p})$. Hence the last term in (5.14) is of $O(1)$.

It remains to estimate the first term in (5.14). According to (5.13), we partition $A_n^{-1}[s_{2p}(\theta)]$ as

$$A_n^{-1}\left[s_{2p}(\theta)\right] = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12}^* & B_{22} & B_{23} \\ B_{13}^* & B_{23}^* & B_{33} \end{bmatrix},$$

where B_{11} and B_{33} are p-by-p matrices. Then by (5.13),

$$\begin{aligned} \rho\left(A_n^{-1}\left[s_{2p}(\theta)\right] C_n\left[s_{2p}(\theta)\right]\right) &\leq 1+\rho\left(\begin{bmatrix} B_{13}R_p^* & 0 & B_{11}R_p \\ B_{23}R_p^* & 0 & B_{12}^*R_p \\ B_{33}R_p^* & 0 & B_{13}^*R_p \end{bmatrix}\right) \\ &= 1+\rho\left(\begin{bmatrix} B_{13}R_p^* & B_{11}R_p \\ B_{33}R_p^* & B_{13}^*R_p \end{bmatrix}\right), \end{aligned} \tag{5.15}$$

where the last equality follows because the 3-by-3 block matrix in the equation has vanishing central column blocks. In [**3**, Theorem 4.3], it has been shown that R_p, B_{11}, B_{13} and B_{33} all have bounded ℓ_1-norms and ℓ_∞-norms. Hence using the fact that $\rho(\cdot) \leq \|\cdot\|_2 \leq \{\|\cdot\|_1\|\cdot\|_\infty\}^{1/2}$, we see that (5.15) is bounded and the theorem follows. ☐

By combining Theorems 5.5 and 5.6, the number of preconditioned conjugate gradient (PCG) iterations required for convergence is of $O(1)$, see [**3**]. Since each PCG iteration requires $O(n\log n)$ operations (see [**8**]) and so is the construction of the preconditioner, the total complexity of the PCG method for solving Toeplitz systems generated by $f \in \mathbf{C}_{2\pi}^+$ is of $O(n\log n)$ operations.

6. Concluding Remarks

In 1986, Strang addressed the question of whether iterative methods can compete with direct methods for solving symmetric positive definite Toeplitz systems. The answer has turned out to be an unqualified yes. The conjugate gradient method coupled with a suitable circulant preconditioner can solve a large class of n-by-n well-conditioned and ill-conditioned Toeplitz systems in $O(n\log n)$ operations, as compared to the $O(n\log^2 n)$ operations required by fast direct Toeplitz solvers. In this paper, we summarize some of the developments of this iterative method for mildly ill-conditioned Toeplitz systems in the past few years.

We remark that even for mildly ill-conditioned matrices with condition number of order $O(n^p)$, if $p > 6$, then the matrix A_n will be very ill-conditioned already for moderate n, say $n = 100$. Thus regularization is also needed in this case. Once the system is regularized, our preconditioner $C_n[\mathcal{K}_{m,8} * f]$ will work even if $p > 6$, cf. the signal restoration example in []. Hence in general, the circulant preconditioner $C_n[\mathcal{K}_{m,8} * f]$ should be able to handle all cases, whether the matrix A_n is well-conditioned, mildly ill-conditioned, or very ill-conditioned but regularized. However, there are still many open problems which should be solved in particular to make Toeplitz solvers applicable for practical tasks. A possible direction of future work is the extension of our new circulant preconditioners to other areas where solution of indefinite Toeplitz or non-Hermitian Toeplitz systems are sought.

References

[1] G. Ammar and W. Gragg, *Superfast solution of real positive definite Toeplitz systems*, SIAM J. Matrix Anal. Appl., 9 (1988), pp. 61–67.

[2] O. Axelsson and V. Barker, *Finite Element Solution of Boundary Value Problems, Theory and Computation*, Academic Press, Orlando, 1984.

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

[4] F. Di Benedetto, *Solution of Toeplitz Normal Equations by Sine Transform Based Preconditioning*, Lin. Alg. Appl., 285 (1998), pp. 229-255.

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

[6] R. Chan, *Circulant Preconditioners for Hermitian Toeplitz Systems*, SIAM J. Matrix Anal. Appl., 10 (1989), pp. 542–550.

[7] R. Chan, *Toeplitz Preconditioners for Toeplitz Systems with Nonnegative Generating Functions*, IMA J. Numer. Anal., 11 (1991), pp. 333–345.

[8] R. Chan and M. Ng, *Conjugate Gradient Methods for Toeplitz Systems*, SIAM Review, 38 (1996), pp. 427–482.

[9] R. Chan, M. Ng and A. Yip, *The Best Circulant Preconditioners for Hermitian Toeplitz Systems II: The Case of Multiple Zeros*, preprint, 1999.

[10] R. Chan, T. Tso and H. Sun, *Circulant Preconditioners from B-Splines*, Proceedings to the SPIE Symposium on Advanced Signal Processing: Algorithms, Architectures, and Implementations, 3162, pp. 338–347, San Diego CA, July, 1997,

[11] R. Chan and M. Yeung, *Circulant Preconditioners Constructed from Kernels*, SIAM J. Numer. Anal., 29 (1992), pp. 1093–1103.

[12] R. Chan, A. Yip and M. Ng, *The Best Circulant Preconditioners for Hermitian Toeplitz Systems*, Math. Dept. Res. Rept. 99-11, The Chinese University of Hong Kong, 1999.

[13] T. Chan, *An Optimal Circulant Preconditioner for Toeplitz Systems*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 766–771.

[14] E. Cheney, *Introduction to Approximation Theory*, McGraw-Hill, New York, 1966.

[15] P. Davis, *Circulant Matrices*, John Wiley & Sons, New York, 1979.

[16] R. Horn and C. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

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

[18] G. Lorentz, *Approximation of Functions*, Holt, Rinehart and Winston, New York, 1966.

[19] D. Luenberger, *Linear and Nonlinear Programming*, 2nd ed., Addison-Wesley, Reading, MA, 1984.

[20] D. Potts and G. Steidl, *Preconditioners for Ill-Conditioned Toeplitz Matrices*, BIT, 39 (1999), pp. 513–533.

[21] D. Potts and G. Steidl, *Preconditioners for Ill-Conditioned Toeplitz Systems Constructed from Positive Kernels*, preprint, 1999.

[22] S. Serra, *Preconditioning Strategies for Hermitian Toeplitz Systems with Nondefinite Generating Functions*, SIAM J. Matrix Anal. Appl., **17** (1996), pp. 1007–1019.

[23] S. Serra, *On the extreme eigenvalues of Hermitian (block) Toeplitz matrices*, Lin. Alg. Appl., **270** (1998), pp. 109–129.

[24] G. Strang, *Introduction to Applied Mathematics*, Wellesley-Cambridge Press, Cambridge, 1986.

[25] G. Strang, *A Proposal for Toeplitz Matrix Calculations*, Stud. Appl. Math., **74** (1986), pp. 171–176.

[26] P. Tang, *A Fast Algorithm for Linear Complex Chebyshev Approximations*, Math. Comp., 51 (1988), pp. 721–739.

[27] W. Trench, *An Algorithm for the Inversion of Finite Toeplitz Matrices*, SIAM J. Appl. Math., **12** (1964), pp. 515–522.

[28] E. Tyrtyshnikov, *Circulant Preconditioners with Unbounded Inverses*, Lin. Alg. Appl., **216** (1995), pp. 1–24.

[29] E. Tyrtyshnikov and V. Strela, *Which Circulant Preconditioners are Better?*, Math. Comp., 65 (1996), pp. 137–150.

Department of Mathematics, The Chinese University of Hong Kong, Shatin, Hong Kong.
E-mail address: rchan@math.cuhk.edu.hk

Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong.
E-mail address: mng@maths.hku.hk

Department of Mathematics, The University of Hong Kong, Pokfulam Road, Hong Kong.
E-mail address: mhyipa@hkusua.hku.hk