Toeplitz-circulant Preconditioners for Toeplitz Systems and Their Applications to Queueing Networks with Batch Arrivals

Raymond H. Chan* Wai-Ki Ching†

November 14, 1994

Abstract

The preconditioned conjugate gradient method is employed to solve Toeplitz systems $T_n\mathbf{x} = \mathbf{b}$ where the generating functions of the n-by-n Toeplitz matrices T_n are functions with zeros. In this case, circulant preconditioners are known to give poor convergence, whereas band-Toeplitz preconditioners only offer linear convergence and can only handle real-valued functions with zeros of even orders. We here propose preconditioners which are products of band-Toeplitz matrices and circulant matrices. The band-Toeplitz matrices are used to cope with the zeros of the given generating function and the circulant matrices are to speed up the convergence rate of the algorithm. Our preconditioner can handle complex-valued functions with zeros of arbitrary orders. We prove that the preconditioned Toeplitz matrices have singular values clustered around 1 for large n. We apply our preconditioners to solve the stationary probability distribution vectors of Markovian queueing models with batch arrivals. We show that if the number of servers is fixed independent of the queue size n, then the preconditioners are invertible and the preconditioned matrices have singular values clustered around 1 for large n. Numerical results are given to illustrate the fast convergence of our methods.

Abbreviated Title: Toeplitz-circulant Preconditioners for Toeplitz Systems.

*Department of Mathematics, Chinese University of Hong Kong, Shatin, Hong Kong. Research supported by HKRGC grant CUHK 178/93E and CUHK-DAG 220600680.

†Department of Systems Engineering & Engineering Management, Chinese University of Hong Kong, Shatin, Hong Kong

Key Words: Preconditioning, Toeplitz matrix, circulant matrix, queueing network.

AMS(MOS) Subject Classifications: 65F10, 65F15, 65N20, 60K25.

1 Introduction

In this paper, we discuss the solutions of linear systems $T\mathbf{x} = \mathbf{b}$ where T is Toeplitz matrix. Direct methods that are based on the Levinson recursion formula are in constant use; see for instance, Trench [25]. For an n-by-n Toeplitz matrix T_n, these methods require $O(n^2)$ operations. Faster algorithms that require $O(n \log^2 n)$ operations have been developed, see Ammar and Gragg [1] for instance. The stability properties of these direct methods for symmetric positive definite matrices are discussed in Bunch [5].

Here we will consider solving Toeplitz systems by the preconditioned conjugate gradient squared (PCGS) method. There are many circulant preconditioning strategies for Toeplitz systems, see for instance [23, 11, 16, 15, 26]. The convergence results for these circulant preconditioners are all based on the regularity of the function $g(\theta)$ whose Fourier coefficients give the diagonals of T_n. The function $g(\theta)$ with $\theta \in [-\pi, \pi]$ is called the generating function of the sequence of Toeplitz matrices T_n. A general result is that if $g(\theta)$ is a positive function in the Wiener class, then for large enough n, the preconditioned matrix has eigenvalues values clustered around 1. In particular, the PCG method applied to the preconditioned system converges superlinearly and the n-by-n Toeplitz system can be solved in $O(n \log n)$ operations. However, we remark that if $g(\theta)$ has a zero, then the result fails to hold and the circulant preconditioned systems can converge at a very slow rate, see the numerical examples in §4 or in [9].

To this end, Chan in [6] has used trigonometric functions of the form $\sin^{\ell}(\theta - \theta_0)$ to approximate the function $g(\theta)$ around the zeros θ_0 of g. The power ℓ is the order of the zero θ_0 and is required to be an even number. The resulting preconditioner is a band-Toeplitz matrix which gives linear convergence. The band-width of the preconditioner is $(\ell + 1)$. To speed up the convergence rate, Chan and Tang [9] have considered using the Remez algorithm to find the best trigonometric polynomial that approximates g in the supremum norm and yet matches the order of the zeros of g in a neighborhood of the zeros. The resulting band-Toeplitz preconditioner can significantly reduce the condition number of the preconditioned systems at the expense of enlarging the band-width. We remark both methods work only for real-valued generating functions with zeros of even orders and fail for complex-valued functions or real-valued functions have zeros of odd order. A typical example is the Toeplitz matrix tridiag$[-1, 1, 0]$. Its generating function is given by $g(\theta) = 1 - e^{-i\theta}$ and it has a zero of order 1 at $\theta = 0$. We note that if we write $z = e^{i\theta}$, then g as a function of z has a zero of order 1 at $z = 1$.

In this paper, we will design preconditioners that give superlinear convergence and work for generating functions that are complex-valued and have zeros with arbitrary orders. Our idea is to approximate g, as a function of the complex variable z, around its zeros z_0 by functions of the form $(z-z_0)^\ell$ where ℓ is the order of the zero z_0. Then we approximate the quotient $g(z)/(z-z_0)^\ell$ by using the usual circulant approach. This results in a preconditioner which is a product of a band-Toeplitz matrix with band-width $(\ell+1)$ and a circulant matrix. We will prove that if the quotient is a nonvanishing Wiener class function, then the preconditioner is invertible and the iterative method converges superlinearly for large n.

We then apply our preconditioner to solve the stationary probability distribution vectors for Markovian queueing networks with batch arrivals. We note that the generator matrices A_n for these queueing networks are singular matrices with a Toeplitz-like structure. In fact, when there is only one server in the system, A_n differs from a lower Hessenberg Toeplitz matrix by a rank one matrix. The preconditioner P_n is constructed by exploiting the near-Toeplitz structure of A_n and will also be a product of a band-Toeplitz matrix and a circulant matrix. We prove that if the number of servers is independent of the queue size n, then for all sufficiently large n, P_n are invertible and that the preconditioned matrices have singular values clustered around 1.

The outline of the paper is as follows. In §2, we define our preconditioners P_n for general Toeplitz matrices with generating functions that have zeros. We then prove that the preconditioned systems have singular values clustered around 1. In §3, we consider solving Markovian queueing networks with batch arrivals by using our preconditioners. In §4, numerical results are given to illustrate the fast convergence of our methods when compared to other methods and other preconditioners in solving Toeplitz systems and queueing networks. Finally, concluding remarks are given in §5.

2 Construction and Analysis of Preconditioners

In this section, we discuss how to construct preconditioners for Toeplitz systems T_n whose generating functions are functions having zeros. Then we analyze the convergence rate of the resulting preconditioned systems.

Let us first recall the definitions of Toeplitz and circulant matrices. An n-by-n matrix $T_n = (t_{i,j})$ is said to be Toeplitz if $t_{i,j} = t_{i-j}$, i.e. if T_n is constant along its diagonals. It is said to be circulant if its diagonals t_k further satisfies $t_{n-k} = t_{-k}$ for $0 < k \le n-1$. The idea of using circulant matrices as preconditioners for Toeplitz matrices has been studied extensively in recent years, see for instance [23, 11, 16, 26, 15]. In this paper, we will concentrate ourselves in the T. Chan circulant preconditioners. The results for the other circulant preconditioners can be obtained similarly, see §5.

For a given Toeplitz matrix T_n with diagonals $\{t_j\}_{j=-(n-1)}^{n-1}$, the T. Chan circulant preconditioner to T_n is defined to be the circulant matrix C_n which minimizes the Frobenius norm $||T_n - C_n||_F$ amongst all circulant matrices. The (i,j)th entry of C_n is given by c_{i-j} where

$$c_k = \begin{cases} \dfrac{(n-k)t_k + kt_{k-n}}{n}, & 0 \le k < n, \\ c_{n+k}, & 0 < -k < n, \end{cases} \tag{1}$$

see T. Chan [11]. We note that the diagonals $\{c_j\}_{j=-(n-1)}^{n-1}$ and hence the matrix C_n can be obtained in $O(n)$ operations. The eigenvalues of C_n, which are required in the inversion of C_n, can be computed in $O(n \log n)$ operations by using Fast Fourier Transforms, see Strang [23] for instance. Because of the good approximating properties of the T. Chan circulant preconditioners, they have been used in solving numerical elliptic partial differential equations [7] and signal processing problems [8].

We are going to analyze the convergence rate of the preconditioned systems $C_n^{-1}T_n$ in the limit $n \to \infty$, assuming that a fixed sequence of entries $\{t_j\}_{j=-\infty}^{\infty}$ has been prescribed. As usual in the study of Toeplitz matrices and operators, see for instance Grenander and Szegö [14], we consider the Laurent series

$$g(z) = \sum_{j=-\infty}^{\infty} t_j z^j$$

whose coefficients $\{t_j\}$ are the entries of T_n, with $t_{j,k} = t_{j-k}$ for $0 \le j, k < n$. We will call $g(z)$ the generating function of the sequence of Toeplitz matrices T_n and for clarity, we will denote such T_n by $\mathcal{T}_n[g]$ and the corresponding T. Chan's circulant preconditioner by $\mathcal{C}_n[g]$.

We note that if

$$\sum_{j=-\infty}^{\infty} |t_j| < \infty,$$

i.e., if $g(z)$ belongs to the Wiener class of functions $\mathcal{W}$ defined on the unit circle $|z| = 1$ and if $g(z)$ has no zeros on $|z| = 1$ then $\mathcal{C}_n[g]$ is a good approximation to $\mathcal{T}_n[g]$ as far as PCG methods are concerned.

Lemma 1 *Let $g \in \mathcal{W}$ and has no zeros on $|z| = 1$. Then for large n, $\mathcal{C}_n[g]$ will be invertible and the sequence of matrices $\mathcal{C}_n[g]^{-1}\mathcal{T}_n[g]$ will have singular values clustered around 1. More precisely, for any fixed $\epsilon > 0$, there exist integers $M, N > 0$ such that for all $n > N$, $\mathcal{C}_n[g]$ is invertible and the matrix $\mathcal{C}_n[g]^{-1}\mathcal{T}_n[g]$ have no more than M singular values lying outside the interval $(1-\epsilon, 1+\epsilon)$.*

Proof: We just note that by the Weierstrass M-test, $g(z)$ is a 2π-periodic complex-valued function defined on the unit circle $|z| = 1$ with respect to the angle θ, see for instance Conway [12, p.29]. The Lemma now follows from Lemma 3 and Theorem 2 of Chan and Yeung [10]. $\square$

In this paper, we are going to relax the requirement that $g(z)$ has no zeros on $|z| = 1$. In particular, we consider $g(z)$ that are of the form

$$g(z) = \left\{ \prod_j (z - z_j)^{\ell_j} \right\} h(z),$$

where z_j are the roots of $g(z)$ on $|z| = 1$ with order ℓ_j and $h(z)$ is a non-vanishing function in $\mathcal{W}$. Our Toeplitz-circulant preconditioner P_n is defined to be

$$P_n = \mathcal{T}_n \left[\prod_j (z - z_j)^{\ell_j} \right] \mathcal{C}_n[h].$$

By expanding the product $\prod_j (z - z_j)^{\ell_j}$ we see that the Toeplitz matrix $\mathcal{T}_n[\prod_j (z - z_j)^{\ell_j}]$ is a lower triangular matrix with band-width equal to $(\ell + 1)$ where

$$\ell = \sum_j \ell_j.$$

Moreover, its main diagonal entry is 1 and therefore it is invertible for all n.

In each iteration of the PCG method, we have to solve a linear system of the form $P_n \mathbf{y} = \mathbf{r}$. We first claim that P_n is invertible for large n. As mentioned above, the Toeplitz matrix $\mathcal{T}_n[\prod_j (z - z_j)^{\ell_j}]$ is invertible for all n. Since $h \in \mathcal{W}$ and has no zeros, the invertibility of $\mathcal{C}_n[h]$ for large n is guaranteed by Lemma 1. Hence P_n is invertible for large n. Let us consider the cost of solving the system

$$P_n \mathbf{y} = \mathcal{T}_n \left[\prod_j (z - z_j)^{\ell_j} \right] \mathcal{C}_n[h] \mathbf{y} = \mathbf{r}.$$

As the matrix $\mathcal{T}_n[\prod_j (z - z_j)^{\ell_j}]$ is a lower triangular matrix with band-width $(\ell + 1)$, the system involving $\mathcal{T}_n[\prod_j (z - z_j)^{\ell_j}]$ can be solved by forward substitution and the cost is $O(\ell n)$ operations. Given any vector $\mathbf{x}$, the matrix-vector product $\mathcal{C}_n[h]^{-1}\mathbf{x}$ can be done by using Fast Fourier Transforms in $O(n \log n)$ operations, see Strang [23] and O'Leary [20]. Thus the system $P_n \mathbf{y} = \mathbf{r}$ can be solved in $O(n \log n) + O(\ell n)$ operations. In comparison, the systems involving the preconditioners proposed by Chan [6] and Chan and Tang [9] require $O(n \log n) + O(\ell^2 n)$ operations to be solved.

We now investigate the convergence rate of the preconditioned systems.

Theorem 1 *The sequence of matrices $P_n^{-1}\mathcal{T}_n[g]$ has singular values clustered around* 1 *for all sufficiently large n.*

Proof: Since $\mathcal{T}_n[\prod_j (z-z_j)^{\ell_j}]$ is a lower triangular Toeplitz matrix of band-width $(\ell+1)$, we see that the matrix

$$\left\{ \mathcal{T}_n[g] - \mathcal{T}_n \left[\prod_j (z-z_j)^{\ell_j} \right] \mathcal{T}_n[h] \right\}$$

only has non-zero entries in the first $l+1$ rows. Hence it is clear that

$$\mathcal{T}_n[g] = \mathcal{T}_n \left[\prod_j (z-z_j)^{\ell_j} \right] \mathcal{T}_n[h] + L_1,$$

where rank $L_1 \le \ell + 1$. Therefore

$$\begin{aligned} P_n^{-1}\mathcal{T}_n[g] &= \mathcal{C}_n[h]^{-1}\mathcal{T}_n \left[\prod_j (z-z_j)^{\ell_j} \right]^{-1} \mathcal{T}_n[g] \\ &= \mathcal{C}_n[h]^{-1}\mathcal{T}_n \left[\prod_j (z-z_j)^{\ell_j} \right]^{-1} (\mathcal{T}_n[\prod_j (z-z_j)^{\ell_j}]\mathcal{T}_n[h] + L_1) \\ &= \mathcal{C}_n[h]^{-1}\mathcal{T}_n[h] + L_2, \end{aligned} \qquad (2)$$

where rank $L_2 \le \ell$. Since h has no zeros, by Lemma 1, $\mathcal{C}_n[h]^{-1}\mathcal{T}_n[h]$ has clustered singular values. In particular, we can write $\mathcal{C}_n[h]^{-1}\mathcal{T}_n[h] = I + L_3 + U$ where U is a small norm matrix and rank L_3 is fixed independent of n, see [10, Corollary 1]. Hence (2) becomes

$$P_n^{-1}\mathcal{T}_n[g] = I + L_4 + U$$

where the rank of L_4 is again fixed independent of n. By using Cauchy interlace theorem [28, p.103] on

$$(P_n^{-1}\mathcal{T}_n[g])^*(P_n^{-1}\mathcal{T}_n[g]) = (I + L_4 + U)^*(I + L_4 + U),$$

it is straightforward to show that $P_n^{-1}\mathcal{T}_n[g]$ has singular values clustered around 1, see [10, Theorem 2] for details. $\square$

Accordingly, the PCG methods will converge fast when applied to solving the preconditioned systems, see Axelsson and Barker [2, p.26] for instance. Numerical examples are given in §4 to illustrate this fast convergence.

3 Markovian Queueing Networks

In this section, we consider using the PCG method with our Toeplitz-circulant preconditioners for solving the stationary probability distribution vectors for Markovian queueing models with batch arrivals. This kind of queueing system occurs in many applications, such as telecommunication networks [19] and loading dock models [21]. We will see that the generator matrices of these systems have a near-Toeplitz structure and our preconditioners are constructed by exploiting this fact.

Let us first introduce the following queueing parameters. Definitions of queueing theory terminologies used below can be found in Cooper [13]. The input of the queueing system will be an exogenous Poisson batch arrival process with mean batch inter-arrival time λ^{-1}. For $k \geq 1$, denote λ_k to be the batch arrival rate for batches with size k. We note that

$$\lambda_k = \lambda p_k \tag{3}$$

where p_k is the probability that the arrival batch size is k. Clearly we have

$$\sum_{k=1}^{\infty} \lambda_k = \lambda. \tag{4}$$

The number of servers in the queueing system will be denoted by s. The service time of each server is independent of the others and is exponentially distributed with mean μ^{-1}. The waiting room is of size $(n-s-1)$ and the queueing discipline is blocked customers cleared. If the arrival batch size is larger than the number of waiting places left, then only part of the arrival batch will be accepted; the other customers will be treated as overflows and will be cleared from the system.

By ordering the state-space lexicographically, i.e. the i-th variable corresponds to the state where there are $i-1$ customers in the system, the queueing model can be characterized by the infinitesimal generator matrix

$$A_n = \begin{pmatrix} \lambda & -\mu & 0 & 0 & 0 & \dots & 0 \\ -\lambda_1 & \lambda+\mu & -2\mu & 0 & 0 & \dots & 0 \\ -\lambda_2 & -\lambda_1 & \lambda+2\mu & \ddots & \ddots & & \vdots \\ \vdots & -\lambda_2 & \ddots & \ddots & -s\mu & \ddots & \\ & \vdots & \ddots & \ddots & \lambda+s\mu & \ddots & 0 \\ -\lambda_{n-2} & -\lambda_{n-3} & & & \vdots & \lambda+s\mu & -s\mu \\ -r_1 & -r_2 & -r_3 & \cdots & -r_{s+1} & \cdots & s\mu \end{pmatrix}, \tag{5}$$

where r_i are such that each column sum of A_n is zero, see Seila [21].

Clearly A_n has zero column sum, positive diagonal entries and non-positive off diagonal entries. Moreover the matrix A_n is irreducible. In fact, if $\lambda_i = 0$ for all $i = 1, \ldots, n-2$, then $r_1 = \lambda$ and the matrix is irreducible. If the λ_i's are not all zero, say λ_j is the first nonzero λ_i, then $r_{n-j} = \lambda$, and hence A_n is also irreducible. From Perron and Frobenius theory [27, p.30], A_n has a 1-dimensional null-space with a positive null vector.

The stationary probability distribution vector $\mathbf{p}$ of the queueing system is the normalized null-vector of the generator matrix A_n given in (5). Many useful information about the queueing system, such as the blocking probability and the expected waiting time of customers can be obtained from $\mathbf{p}$. Since A_n has a 1-dimensional null-space, $\mathbf{p}$ can be found by deleting the last column and the last row of A_n and solving the $(n-1)$-by-$(n-1)$ reduced linear system $Q_{n-1}\mathbf{y} = (0, \ldots, 0, s\mu)^t$. After getting $\mathbf{y}$, the distribution vector $\mathbf{p}$ can then be obtained by normalizing the vector $(\mathbf{y}^t, 1)^t$.

Thus let us concentrate ourselves in solving nonhomogeneous systems of the form

$$Q_n \mathbf{x} = \mathbf{b} \tag{6}$$

where

$$Q_n = \begin{pmatrix} \lambda & -\mu & 0 & 0 & 0 & \ldots & 0 \\ -\lambda_1 & \lambda+\mu & -2\mu & 0 & 0 & \ldots & 0 \\ -\lambda_2 & -\lambda_1 & \lambda+2\mu & \ddots & \ddots & & \vdots \\ \vdots & -\lambda_2 & \ddots & \ddots & -s\mu & \ddots & \\ & \vdots & \ddots & \ddots & \lambda+s\mu & \ddots & 0 \\ -\lambda_{n-2} & -\lambda_{n-3} & & \ddots & \ddots & \ddots & -s\mu \\ -\lambda_{n-1} & -\lambda_{n-2} & & \cdots & -\lambda_2 & -\lambda_1 & \lambda+s\mu \end{pmatrix}. \tag{7}$$

Notice that if all of the λ_i, $i = 1, \ldots, n-1$ are zeros, then Q_n will be a bidiagonal matrix and can easily be inverted. Therefore in the following, we assume that at least one of the λ_i is non-zero. Then clearly, Q_n^t is an irreducibly diagonally dominant matrix. In particular, if the system (6) is solved by classical iterative methods such as the Jacobi or the Gauss-Seidel methods, both methods will converge for arbitrary initial guesses, see for instance Varga [27, Theorem 3.4].

We will see in §3.2 that the costs per iteration of the Jacobi and the Gauss-Seidel methods are $O(n \log n)$ and $O(n^2)$ respectively. The memory requirement is $O(n)$ for both methods. We remark that the system (6) can also be solved by Gaussian elimination in $O(n^2)$ operations with $O(n^2)$ memory. In the remaining of this section, we are interested in solving (6) by the PCG method. We will see that the cost per iteration of the method is $O(n \log n)$ and memory requirement is $O(n)$, the same as those of the Jacobi method. However, we are able to show that if s is independent of n, then with our Toeplitz-circulant preconditioner, the PCG method converges superlinearly for all sufficiently large n. In particular, the method converges in finite number of steps independent of the queue size n. Therefore the total cost of finding the steady-state probability distribution is $O(n \log n)$ operations.

3.1 Construction of the Preconditioner

We observe that in the single server case, i.e. when $s = 1$, the matrix Q_n given in (7) differs from a lower Hessenberg Toeplitz matrix by only its $(1,1)$ entry. In general, Q_n can be written as

$$Q_n = T_n + R_n, \tag{8}$$

where T_n is a Toeplitz matrix:

$$T_n = \begin{pmatrix} \lambda + s\mu & -s\mu & 0 & 0 & 0 & \dots & 0 \\ -\lambda_1 & \lambda + s\mu & -s\mu & 0 & 0 & \dots & 0 \\ -\lambda_2 & -\lambda_1 & \lambda + s\mu & \ddots & \ddots & & \vdots \\ \vdots & -\lambda_2 & \ddots & \ddots & -s\mu & \ddots & \\ & \vdots & \ddots & \ddots & \lambda + s\mu & \ddots & 0 \\ -\lambda_{n-2} & & & \ddots & \ddots & \ddots & -s\mu \\ -\lambda_{n-1} & -\lambda_{n-2} & & \dots & -\lambda_2 & -\lambda_1 & \lambda + s\mu \end{pmatrix}, \tag{9}$$

and R_n is a matrix of rank s. From (9), we see that $T_n = \mathcal{T}_n[g]$, where the generating function $g(z)$ of T_n is given by

$$g(z) = -s\mu\frac{1}{z} + \lambda + s\mu - \sum_{k=1}^{\infty} \lambda_k z^k. \tag{10}$$

We note that by (4), $g \in \mathcal{W}$.

Unfortunately, it is also clear from (10) and (4) that $g(z)$ has a zero at $z = 1$ and therefore Lemma 1 is not applicable. However, if we look at the real part of $g(z)$ on the unit circle $|z| = 1$, we see that

$$\mathrm{Re}\{g(z)\} = -s\mu\cos\theta + \lambda + s\mu - \sum_{k=1}^{\infty} \lambda_k \cos(k\theta) \geq s\mu - s\mu\cos\theta.$$

Hence the zeros of $g(z)$ can only occur at $z = 1$. In particular, we can write

$$g(z) = (z-1)^{\ell} b(z), \tag{11}$$

where ℓ is the order of the zero of $g(z)$ at $z = 1$ and $b(z)$ will have no zeros on the unit circle. According to the discussion in §2, we define our preconditioner for Q_n as

$$P_n = \mathcal{T}_n[(z-1)^\ell]\mathcal{C}_n[b]. \tag{12}$$

Let us consider cases where the quotient function $b(z)$ will be in $\mathcal{W}$. We first note that if the radius of convergence ρ of the power series $\sum_{k=1}^{\infty} \lambda_k z^k$ in (10) is greater than 1, then $g(z)$ and hence $b(z)$ are analytic functions in a neighborhood of $|z| = 1$, see Conway [12, p.31]. In particular, $h(z)$ will be in $\mathcal{W}$. A formula for computing ρ is given by

$$\frac{1}{\rho} = \limsup |\lambda_j|^{1/j}, \tag{13}$$

see Conway [12, p.31].

Next we consider the case $\ell = 1$ in more depth. By straightforward division of $g(z)$ in (10) by $(z-1)$, we have

$$b(z) = s\mu\frac{1}{z} - \lambda - \sum_{k=1}^{\infty}(\lambda - \sum_{j=1}^{k}\lambda_j)z^k. \tag{14}$$

Therefore, by (3) and (4),

$$b(1) = s\mu - \sum_{k=0}^{\infty}\sum_{j=k+1}^{\infty}\lambda_j = s\mu - \lambda\sum_{k=0}^{\infty}kp_k = s\mu - \lambda E(B) \tag{15}$$

where $E(B)$ is the expected value of the arrival batch size. Thus if $s\mu \neq \lambda E(B)$ then $b(1) \neq 0$ and hence $\ell = 1$. Moreover, if $E(B) < \infty$, then $b \in \mathcal{W}$. Clearly from (14), the first n Laurent coefficients of $b(z)$, i.e. $\sum_{j=1}^{k}\lambda_j - \lambda$, $k = 1, 2, \ldots, n$, can be computed recursively in $O(n)$ operations. Hence by using (1), $\mathcal{C}_n[h]$ and also P_n can be constructed in $O(n)$ operations.

3.2 Convergence Analysis and Computation Cost

In this section, we prove the fast convergence of the PCG method and discuss its computational cost.

Theorem 2 *Let $b(z)$ defined in (11) be in $\mathcal{W}$ and the number of servers s in the queue be independent of the queue size n. Then the sequence of preconditioned matrices $P_n^{-1}Q_n$ has singular values clustered around 1 for large n.*

Proof: By (8) and (12),

$$P_n^{-1}Q_n = \mathcal{C}_n[b]^{-1}\mathcal{T}_n[(z-1)^\ell]^{-1}(\mathcal{T}_n[g]+R_n) = \mathcal{C}_n[b]^{-1}\mathcal{T}_n[(z-1)^\ell]^{-1}\mathcal{T}_n[g]+L_5$$

where rank $L_5 \le s$. By Theorem 1 and Cauchy interlace theorem, we see that $P_n^{-1}Q_n$ has singular values clustered around 1 for sufficiently large n. $\square$

It follows from standard convergence theory of the PCG method that the method will converge superlinearly and in particular in finite number of steps independent of n.

In each iteration of the PCG method, the main computational cost consists of solving a linear system $P_n\mathbf{y} = \mathbf{r}$ and multiplying Q_n to some vector $\mathbf{r}$. We first recall from §2 that the cost of solving $P_n\mathbf{y} = \mathbf{r}$ is of $O(n\log n) + O(\ell n)$ operations. To compute $Q_n\mathbf{r}$, we make use of the partitioning (8). Note that R_n in (8) is a matrix containing only $2s-1$ nonzero entries, we therefore need $O(s)$ operations for computing $R_n\mathbf{r}$. Since T_n is a Toeplitz matrix, $T_n\mathbf{r}$ can be computed in $O(n\log n)$ operations by embedding T_n into a $2n$-by-$2n$ circulant matrix, see Strang [23]. Hence $Q_n\mathbf{r}$ can be obtained in $O(n\log n)$ operations. Thus the number of operations required for each iteration of the PCG method is of order $O(n\log n)$.

Finally, we consider the memory requirement. We note that besides some n-vectors, we only have to store the first column (or eigenvalues) of the matrices T_n, $\mathcal{T}_n[(z-1)^\ell]$ and $\mathcal{C}_n[b]$ but not the whole matrices. Thus we need $O(n)$ memory for the PCG method.

4 Numerical Results

In this section, we test the performance of our Toeplitz-circulant preconditioners P_n on solving Toeplitz systems and the queueing problems discussed in §3. All computations were done by Matlab on an HP 715 workstation.

For the tests on Toeplitz systems, we tried the following generating functions:

(i) $$g_1(z) = \frac{(z^4-1)}{(z-\frac{3}{2})(z-\frac{1}{2})} = \frac{15}{8}\sum_{k=1}^{\infty}\frac{1}{(2z)^k} + \frac{13}{24} + \frac{7}{36}z - \frac{11}{54}z^2 - \frac{65}{24}\sum_{k=3}^{\infty}(\frac{2z}{3})^k.$$

(ii) $$g_2(z) = \frac{(z+1)^2(z-1)^2}{(z-\frac{3}{2})(z-\frac{1}{2})} = -\frac{9}{8}\sum_{k=1}^{\infty}\frac{1}{(2z)^k} + \frac{5}{24} + \frac{47}{36}z + \frac{29}{54}z^2 - \frac{25}{24}\sum_{k=3}^{\infty}(\frac{2z}{3})^k.$$

(iii) $$g_3(z) = \frac{(z+1)^2(z-1)}{(z-\frac{3}{2})(z-\frac{1}{2})} = \frac{9}{4}\sum_{k=1}^{\infty}\frac{1}{(2z)^k} + \frac{11}{12} - \frac{7}{18}z - \frac{25}{12}\sum_{k=2}^{\infty}(\frac{2z}{3})^k.$$

Clearly the functions g_i, $i = 1,2,3$, all have zeros on $|z| = 1$. We remark that the preconditioners proposed in Chan [6] and Chan and Tang [9] are not applicable here because g_i are complex-valued functions on $|z| = 1$.

We note that the Toeplitz matrices formed by g_i's are nonsymmetric, therefore the systems $\mathcal{T}_n[g_i]\mathbf{x} = \mathbf{b}$ are solved by the preconditioned conjugate gradient squared (PCGS) method, see Sonneveld [22]. The stopping criterion we used is

$$\frac{||\mathbf{r}_k||_2}{||\mathbf{r}_0||_2} < 10^{-6}, \tag{16}$$

where $\mathbf{r}_k$ is the residual at the kth iteration. The right hand side vector is $(1, 1, \ldots, 1)^t$ and the initial guess is the zero vector. Table 1 gives the numbers of iterations required for convergence by using preconditioners I, P_n and $\mathcal{C}_n[g]$. The symbol $**$ there denotes that the method does not converge in 5000 iterations. We see that the circulant preconditioner does not work well when the generating function has zeros on $|z| = 1$ and that the number of iterations required for convergence actually grows with n. However, our preconditioner P_n gives very fast convergence in all cases and the rate is actually improving with increasing n.

	g_1			g_2			g_3		
n	I	P_n	$\mathcal{C}_n[g_1]$	I	P_n	$\mathcal{C}_n[g_2]$	I	P_n	$\mathcal{C}_n[g_3]$
8	8	7	8	8	8	7	8	9	7
16	22	6	9	26	7	9	26	5	12
32	68	5	9	132	6	11	68	6	12
64	315	4	9	**	6	14	202	5	13
128	3417	4	10	**	5	15	573	5	17
256	**	4	10	**	5	18	**	5	22
512	**	4	10	**	5	25	**	5	28

Table 1. Numbers of Iterations for Different Preconditioners.

Next we test our preconditioner for queueing networks mentioned in §3. Since Q_n in (7) is irreducibly diagonally dominant, both Jacobi and Gauss-Seidel methods converge when applied to solving the system (6). However, by using the partitioning of Q_n as in (8) and taking advantage of the Toeplitz matrix-vector multiplication (see §3.2), we see that each iteration of the Jacobi method can be done in $O(n \log n)$ operations, the same count as that of the PCGS method. This special property is not enjoyed by the Gauss-Seidel method which will still require $O(n^2)$ operations per iteration. Thus in our comparisons, we used only the Jacobi method.

We tried two sets of queueing parameters:

(i) $\lambda_j = \dfrac{1}{2^j}$, $j = 1, 2, \ldots$, and

(ii) $\lambda_j = \dfrac{90}{(\pi j)^4}$, $j = 1, 2, \ldots$.

We note that, in both cases, $\lambda = \sum_{k=1}^{\infty} \lambda_k = 1$. The service rate μ is set to $\mu = \lambda/s$. By (15), we see that $b(1) \neq 0$ and hence $\ell = 1$. Clearly, in both cases the mean arrival batch size $E(B)$ is finite because $\sum_j j\lambda_j < \infty$. Therefore, $b(z) \in \mathcal{W}$ and is given by (14). We remark that by using (13), the radius of convergence ρ for the first set of queueing parameters is 2. Hence regardless of the values of μ, its $b(z)$ will always be in $\mathcal{W}$.

The initial guess for both methods is $(1, 1, \ldots, 1)/n$. The stopping criteria for the PCGS method is again given by (16), whereas for the Jacobi method, it is $||\mathbf{x}_k - \mathbf{x}_{k-1}||_2 < 10^{-6}$, where $\mathbf{x}_k$ is the solution obtained at the kth iteration. Tables 2–3 give the numbers of iterations required for convergence for $s = 1, 4$ and $n-1$. The symbol J there means the Jacobi method is used. Again $**$ signifies that the method does not converge in 5000 iterations. The symbol "kflop" means 1000 floating point operations. Note that the case $s = n-1$ is not covered by our Theorem 2. However, we note that in all the cases we tested, our preconditioner P_n is clearly the best choice.

s	1				4				$n-1$			
n	I	P_n	$\mathcal{C}_n[g]$	J	I	P_n	$\mathcal{C}_n[g]$	J	I	P_n	$\mathcal{C}_n[g]$	J
8	8	5	6	95	8	5	7	78	8	6	6	45
16	15	4	6	115	15	5	7	86	13	7	8	70
32	28	4	7	213	27	5	8	209	21	7	9	114
64	**	4	7	307	**	5	8	306	48	7	10	173
128	**	3	7	470	**	5	8	469	**	7	10	267
256	**	3	8	768	**	5	8	768	**	7	10	434
512	**	3	8	1331	**	5	8	1331	**	6	10	746

Table 2: Numbers of iterations for $\lambda_j = 1/2^j$.

s	1				4				$n-1$			
n	I	P_n	$\mathcal{C}_n[g]$	J	I	P_n	$\mathcal{C}_n[g]$	J	I	P_n	$\mathcal{C}_n[g]$	J
8	8	5	6	389	8	5	7	264	8	6	8	112
16	16	4	7	1050	16	6	9	899	16	8	12	212
32	32	4	9	2253	32	6	10	2139	32	12	15	372
64	64	4	11	3431	64	5	12	3398	79	15	23	577
128	125	4	13	3874	124	5	15	3842	**	18	29	1005
256	365	4	17	**	388	5	18	**	**	21	34	1841
512	**	3	21	**	**	5	21	**	**	17	38	3163

Table 3: Numbers of iterations for $\lambda_j = 90/(\pi j)^4$.

n	I	P_n	$\mathcal{C}_n[g]$	J
8	21	20	23	413
16	82	37	56	2410
32	340	80	150	11213
64	1437	172	388	37042
128	5997	373	977	90536
256	37423	807	2696	**
512	**	1421	7060	**

Table 4: Numbers of kflops for $\lambda_j = 90/(\pi j)^4$ and $s = 1$.

5 Concluding Remarks

We remark that although we concentrate ourselves in the T. Chan circulant preconditioners here, the convergence results in Theorems 1 and 2 can easily be extended to include other circulant preconditioners. For instance, results for Strang's circulant preconditioners can be obtained if we replace Lemma 1 by theorems in [24]. In particular, using Theorem 6 there, we can show that if the quotient function $h(z)$ is a rational function of type (μ, ν), then our method converges in at most $(1 + 2\max\{\mu, \nu\} + \ell)$ steps for large n.

References

[1] G. Ammar and W. Gragg, *Superfast Solution of Real Positive Definite Toeplitz Systems*, SIAM J. Matrix Anal. Appl., 9 (1988), 61–76.

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

[3] F. Benedetto, *Analysis of Preconditioning Techniques for Ill-Conditioned Toeplitz Matrices*, to appear on SIAM J. Sci. Statist. Comput..

[4] F. Benedetto, G. Fiorentino and S. Serra, *C.G. Preconditioning for Toeplitz Matrices*, Computers Math. Appl., 25 (1993) 35–45.

[5] J. Bunch, *Stability of Methods for Solving Toeplitz Systems of Equations*, SIAM J. Sci. Statist. Comput., 6 (1985), 349–364.

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

[7] R. Chan and T. Chan, *Circulant Preconditioners for Elliptic Problems*, J. of Num. Lin. Alg. Appls., 1 (1992), 77–101.

[8] R. Chan, J. Nagy and R. Plemmons, *FFT-Based Preconditioners for Toeplitz-Block Least Squares Problems*, SIAM J. Numer. Anal., 30 (1993), 1740–1768.

[9] R. Chan and P. Tang, *Fast Band-Toeplitz Preconditioners for Hermitian Toeplitz Systems*, SIAM J. Sci. Statist. Comput., 15 (1994), 164–171.

[10] R. Chan and M. Yeung, *Circulant Preconditioners for Complex Toeplitz Matrices*, SIAM J. Numer. Anal., 30 (1993), 1193–1207.

[11] T. Chan, *An Optimal Circulant Preconditioner for Toeplitz Systems*, SIAM J. Sci. Statist. Comput., 9 (1988), 767–771.

[12] J. Conway, *Functions of One Complex Variable*, Springer-Verlag, Berlin, 1973.

[13] R. Cooper, Introduction to queueing theory, 2nd ed., London, Arnold, 1981.

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

[15] T. Huckle, *Circulant and Skew-circulant Matrices for Solving Toeplitz Matrix Problems*, SIAM J. Matrix Anal. Appl., 13 (1992), 767–777.

[16] T. Ku and C. Kuo, *Design and Analysis of Toeplitz Preconditioners*, IEEE Trans. Signal Processing, 40 (1991), 129–141.

[17] T. Ku and C. Kuo, *A Minimum-Phase LU Factorization Preconditioner for Toeplitz Matrices*, SIAM J. Sci. Statist. Comput., 13 (1992), 1470–1480.

[18] T. Ku and C. Kua, *Spectral Properties of Preconditioned Rational Toeplitz Matrices: the Nonsymmetric Case*, SIMAX 14 (1993), 521–544.

[19] T. Oda, *Moment Analysis for Traffic Associated with Markovian Queueing Systems,* IEEE Trans. Commun., 39 (1991), 737–745.

[20] D. O'Leary and J. Simmons, *A Bidiagonalization-Regularization Procedure for Large Scale Discretizations of Ill-Posed Problems*, SIAM J. Sci. Statist. Comput., 2 (1981), 474–488.

[21] A. Seila, *Multivariate Estimation of Conditional Performance Measure in Regenerative Simulation,* American Journal of Mathematical and Management Sciences, 10 (1990), 17–45.

[22] P. Sonneveld, *A Fast Lanczos-type Solver for Non-symmetric Linear Systems,* SIAM J. Sci. Statist. Comput., 10 (1989), 36–52.

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

[24] L. Trefethen, *Approximation Theory and Numerical Linear Algebra*, Algorithms for Approximation II, J. Mason and M. Cox, eds., Chapman and Hall, London, 1990.

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

[26] E. Tyrtyshnikov, *Optimal and Super-optimal Circulant Preconditioners,* SIAM J. Matrix Anal. Appl., 13 (1992), 459–473.

[27] R. Varga, *Matrix Iterative Analysis*, Prentice-Hall, New Jersey, 1963.

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