

## SPECTRAL PROPERTIES OF PRECONDITIONED RATIONAL TOEPLITZ MATRICES: THE NONSYMMETRIC CASE\*

TA-KANG KU<sup>†</sup> AND C.-C. JAY KUO<sup>†</sup>

**Abstract.** Various preconditioners for symmetric positive-definite (SPD) Toeplitz matrices in circulant matrix form have recently been proposed. The spectral properties of the preconditioned SPD Toeplitz matrices have also been studied. In this research, Strang's preconditioner  $S_N$  and our preconditioner  $K_N$  are applied to an  $N \times N$  nonsymmetric (or nonhermitian) Toeplitz system  $T_N \mathbf{x} = \mathbf{b}$ . For a large class of Toeplitz matrices, it is proved that the singular values of  $S_N^{-1} T_N$  and  $K_N^{-1} T_N$  are clustered around unity except for a fixed number independent of  $N$ . If  $T_N$  is additionally generated by a rational function, the eigenvalues of  $S_N^{-1} T_N$  and  $K_N^{-1} T_N$  can be characterized directly. Let the eigenvalues of  $S_N^{-1} T_N$  and  $K_N^{-1} T_N$  be classified into the outliers and the clustered eigenvalues depending on whether they converge to 1 asymptotically. Then, the number of outliers depends on the order of the rational generating function, and the clustering radius is proportional to the magnitude of the last elements in the generating sequence used to construct the preconditioner. Numerical experiments are provided to illustrate our theoretical study.

**Key words.** Toeplitz matrix, preconditioned iterative method, rational generating function, nonsymmetric matrices

**AMS(MOS) subject classifications.** 65F10, 65F15

**1. Introduction.** Research on preconditioning symmetric positive-definite (SPD) Toeplitz matrices with circulant matrices has been active recently [1], [6], [8], [9], [17]. In this research, we generalize Strang's preconditioner  $S_N$  [17] and our preconditioner  $K_N$  [9] to nonsymmetric (or nonhermitian) Toeplitz matrices. Let  $T_N$  be an  $N \times N$  nonsymmetric Toeplitz matrix with elements  $t_{i,j} = t_{i-j}$ . The generalized Strang's preconditioner  $S_N$  is obtained by preserving  $N$  consecutive diagonals in  $T_N$ , i.e., diagonals with elements  $t_n, 1 - M \leq n \leq N - M$ , and using them to form a circulant matrix. One simple rule to determine  $M$  is to choose its value such that  $|t_{N-M}| \approx |t_{1-M}|$ . Note that half of the elements in  $T_N$  are not used in constructing  $S_N$ . The generalized preconditioner  $K_N$  is obtained from a  $2N \times 2N$  circulant matrix in such a way that all elements in  $T_N$  are used, and is a circulant matrix itself (see §2). Since  $S_N$  and  $K_N$  are circulant, the matrix-vector products  $S_N^{-1} \mathbf{v}$  and  $K_N^{-1} \mathbf{v}$  can be conveniently computed via fast Fourier transform (FFT) with  $O(N \log N)$  operations. The system of equations associated with the preconditioned Toeplitz matrix is then solved by iterative methods such as CGN (the conjugate gradient iteration applied to the normal equations) [7], GMRES (the generalized minimal residual) [15], and CGS (the conjugate gradient squared) [16].

The convergence rate of preconditioned iterative methods depends on the singular value or eigenvalue distribution of the preconditioned matrices [13]. The spectral properties of preconditioned SPD Toeplitz matrices have been widely studied. Chan and Strang [1], [5] proved that, for a symmetric Toeplitz with a positive generating function in the Wiener class, the preconditioned matrix has eigenvalues clustered around unity except a fixed number independent of  $N$ . If the Toeplitz is additionally

\* Received by the editors April 16, 1991; accepted for publication (in revised form) September 4, 1991. This work was supported by the University of Southern California Faculty Research and Innovation Fund and by a National Science Foundation Research Initiation Award (ASC-9009323).

<sup>†</sup> Signal and Image Processing Institute and Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, California 90089-2564 (tkku@sipi.usc.edu and cckuo@sipi.usc.edu).

generated by a rational function, even stronger results were proved by Trefethen [19] and the authors [12]. In contrast, relatively few results for preconditioned nonsymmetric Toeplitz have been obtained so far [4], [11], [14], [21]. One such preconditioner was studied by R. Chan and Yeung [4], where they generalized T. Chan's preconditioner  $C_N$  to nonsymmetric Toeplitz matrices.

In this research, we examine the spectral properties of  $S_N^{-1}T_N$  and  $K_N^{-1}T_N$  for nonsymmetric  $T_N$  in general, and nonsymmetric rational  $T_N$  in particular. The main results of our study are stated as follows. For a large class of general Toeplitz matrices, we prove that the singular values of  $S_N^{-1}T_N$  and  $K_N^{-1}T_N$ , or equivalently, the eigenvalues of  $(S_N^{-1}T_N)^H(S_N^{-1}T_N)$  and  $(K_N^{-1}T_N)^H(K_N^{-1}T_N)$ , are clustered around unity except for a fixed number independent of  $N$ . If  $T_N$  is additionally generated by a rational function of order  $(\alpha, \beta, \gamma, \delta)$ , we are able to characterize the eigenvalues of  $S_N^{-1}T_N$  and  $K_N^{-1}T_N$  directly. We classify the eigenvalues of  $S_N^{-1}T_N$  and  $K_N^{-1}T_N$  into two classes, i.e. the outliers and the clustered eigenvalues, depending on whether they converge to 1 asymptotically. Then, (i) the number of outliers is at most  $\eta = 2 \min(r, s)$  where  $r = \max(\alpha, \beta)$  and  $s = \max(\gamma, \delta)$ ; and (ii) the clustered eigenvalues are confined in a disk centered at 1 with radius  $\epsilon$ , where the clustering radius  $\epsilon$  is proportional to the magnitude of the last elements in the generating sequence used to construct the preconditioner.

With these spectral regularities, we can find appropriate preconditioned iterative methods to solve a nonsymmetric Toeplitz system efficiently. In particular, an  $N \times N$  rational Toeplitz system  $T_N \mathbf{x} = \mathbf{b}$  can be solved with  $O(N \log N)$  operations since the number of iterations required for convergence is independent of the problem size  $N$ . To compare the performance of  $S_N$  and  $K_N$ , the  $S_N^{-1}T_N$  and  $K_N^{-1}T_N$  have the same number of outliers so that they converge in the same number of iterations asymptotically. However, the performances of  $S_N$  and  $K_N$  for finite  $N$  are determined by the clustering radii of the clustered eigenvalues as well. The magnitudes of the last elements used to construct  $S_N$  and  $K_N$  are  $O(|t_{N-M}| + |t_{1-M}|)$  and  $O(|t_N| + |t_{-N}|)$ , respectively. Since  $O(|t_N| + |t_{-N}|) \leq O(|t_{N-M}| + |t_{1-M}|)$  for large  $N$ , iterative methods with preconditioner  $K_N$  converges faster than with preconditioner  $S_N$  for solving rational Toeplitz systems. This is confirmed by numerical experiments. By the parallelism provided by FFT, the iterative methods with preconditioners in circulant matrix form is highly parallelizable, and the time complexity of the method can be reduced to  $O(\log N)$  if  $O(N)$  processors are used.

When  $T_N$  is a symmetric rational Toeplitz, we have  $r = s$  and  $t_N = t_{-N}$ . Consequently, the number of outliers of  $K_N^{-1}T_N$  is  $\eta = 2r = 2 \max(\alpha, \beta)$  and the clustering radius is  $O(|t_N|)$ . They reduce to the case given in [12]. Although the results derived in this paper can be viewed as a generalization of the results in [12], we want to point out that the approach adopted in this research is very different from that in [12] and the proof techniques are much more involved. For example, in characterizing the clustering radius of clustered eigenvalues of  $K_N^{-1}T_N$  (or  $S_N^{-1}T_N$ ) for symmetric  $T_N$ , the intertwining theorem of eigenvalues was exploited in [12]. However, such a theorem does not exist for nonsymmetric matrices so that we use perturbation theory for eigenvalues instead.

It is worthwhile to mention that there exists a preconditioner based on the minimum-phase LU factorization (MPLU) technique [11] which has a faster or comparable convergence rate than preconditioners  $S_N$  and  $K_N$ . However, Toeplitz preconditioners in circulant matrix form have two advantages over the MPLU preconditioner. First, the circulant preconditioning technique can be easily generalized to multidimensional

mensional Toeplitz systems. See [2], [3], and [10] for the two-dimensional case (block Toeplitz matrices). Second, the resulting preconditioned iterative method with preconditioners in circulant form is highly parallelizable while the MPLU preconditioner has to be implemented sequentially.

This paper is organized as follows. The construction of preconditioners  $S_N$  and  $K_N$  for nonsymmetric Toeplitz  $T_N$  is discussed in §2. We describe the singular value distribution of  $K_N^{-1}T_N$  and  $S_N^{-1}T_N$  for general Toeplitz in §3, and characterize the eigenvalue distribution of  $K_N^{-1}T_N$  and  $S_N^{-1}T_N$  for rational Toeplitz in §4 and §5, respectively. Numerical experiments are given in §6 to illustrate the theoretical study.

**2. Constructions of Toeplitz preconditioners.** Let  $T_m$  be a sequence of  $m \times m$  nonsymmetric Toeplitz matrices with generating sequence  $t_n$ . Then,

$$T_N = \begin{bmatrix} t_0 & t_{-1} & \cdot & t_{-(N-2)} & t_{-(N-1)} \\ t_1 & t_0 & t_{-1} & \cdot & t_{-(N-2)} \\ \cdot & t_1 & t_0 & \cdot & \cdot \\ t_{N-2} & \cdot & \cdot & \cdot & t_{-1} \\ t_{N-1} & t_{N-2} & \cdot & t_1 & t_0 \end{bmatrix}.$$

Following the idea proposed by Strang [17], we construct the preconditioner  $S_N$  by preserving  $N$  consecutive diagonals in  $T_N$  and bringing them around to form a circulant matrix,

$$S_N = \begin{bmatrix} t_0 & t_{-1} & \cdot & t_{2-M} & t_{1-M} & t_{N-M} & \cdot & t_2 & t_1 \\ t_1 & t_0 & t_{-1} & \cdot & t_{2-M} & t_{1-M} & t_{N-M} & \cdot & t_2 \\ \cdot & \cdot \\ \cdot & \cdot & t_1 & t_0 & t_{-1} & \cdot & t_{2-M} & t_{1-M} & t_{N-M} \\ t_{N-M} & \cdot & \cdot & t_1 & t_0 & t_{-1} & \cdot & t_{2-M} & t_{1-M} \\ t_{1-M} & t_{N-M} & \cdot & \cdot & t_1 & t_0 & t_{-1} & \cdot & t_{2-M} \\ \cdot & \cdot \\ t_{-2} & \cdot & t_{1-M} & t_{N-M} & \cdot & \cdot & t_1 & t_0 & t_{-1} \\ t_{-1} & t_{-2} & \cdot & t_{1-M} & t_{N-M} & \cdot & \cdot & t_1 & t_0 \end{bmatrix}.$$

A simple rule of thumb to decide the value of  $M$  is to require  $|t_{N-M}| \approx |t_{1-M}|$ .

Generalizing the idea in [9], the preconditioner  $K_N$  is constructed based on a  $2N \times 2N$  circulant matrix  $R_{2N}$ ,

$$R_{2N} = \begin{bmatrix} T_N & \Delta T_N \\ \Delta T_N & T_N \end{bmatrix},$$

where  $\Delta T_N$  is determined by the elements of  $T_N$  to make  $R_{2N}$  circulant, i.e.,

$$\Delta T_N = \begin{bmatrix} 0 & t_{N-1} & \cdot & t_2 & t_1 \\ t_{-(N-1)} & 0 & t_{N-1} & \cdot & t_2 \\ \cdot & t_{-(N-1)} & 0 & \cdot & \cdot \\ t_{-2} & \cdot & \cdot & \cdot & t_{N-1} \\ t_{-1} & t_{-2} & \cdot & t_{-(N-1)} & 0 \end{bmatrix}.$$

This construction is motivated by the observation that the augmented circulant system,

$$\begin{bmatrix} T_N & \Delta T_N \\ \Delta T_N & T_N \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{b} \end{bmatrix},$$

is equivalent to  $(T_N + \Delta T_N)\mathbf{x} = \mathbf{b}$  so that  $(T_N + \Delta T_N)^{-1}\mathbf{b}$  can be computed efficiently via FFT and

$$(2.1) \quad K_N = T_N + \Delta T_N$$

can be used as a preconditioner for  $T_N$ . Note, however, that  $K_N$  itself is also circulant and can be inverted directly via  $N$ -point FFT rather than  $2N$ -point FFT.

**3. Spectral properties of preconditioned Toeplitz.** We assume that the generating sequence  $t_n$  satisfies the following two conditions:

$$(3.1) \quad \sum_{-\infty}^{\infty} |t_n| \leq B_T < \infty,$$

$$(3.2) \quad |T(e^{i\theta})| = \left| \sum_{-\infty}^{\infty} t_n e^{-in\theta} \right| \geq \mu_T > 0 \quad \forall \theta.$$

According to the construction of preconditioners  $K_N$  and  $S_N$ , the function  $T(e^{i\theta})$  describes the asymptotic eigenvalue distribution of preconditioners  $K_N$  and  $S_N$  but not necessarily the asymptotic eigenvalue distribution of Toeplitz matrix  $T_N$  (see Test Problem 5 in §6). Thus, the matrix-vector product  $K_N^{-1}\mathbf{v}$  (or  $S_N^{-1}\mathbf{v}$ ) can be completed successfully via FFT since magnitudes of eigenvalues of  $K_N^{-1}$  (or  $S_N^{-1}$ ) are bounded due to (3.2). With conditions (3.1) and (3.2), the condition numbers of preconditioners  $K_N$  and  $S_N$  are bounded independent of  $N$  due to the following theorem.

**THEOREM 1.** *Let  $T_N$  be an  $N \times N$  Toeplitz matrix with the corresponding generating sequence satisfying (3.1) and (3.2). The  $\|(K_N K_N^H)^{-1}\|_2$  and  $\|(S_N S_N^H)^{-1}\|_2$  are bounded for sufficiently large  $N$ .*

*Proof.* Since  $K_N$  is circulant, we have

$$K_N = F_N^H D_N F_N \quad \text{and} \quad K_N^H = F_N^H D_N^H F_N,$$

where  $F_N$  is the  $N \times N$  unitary Fourier matrix with  $N^{-1/2} e^{-i2\pi(m-1)(n-1)/N}$  as the  $(m, n)$  element and  $D_N$  a diagonal matrix formed by the eigenvalues of  $K_N$ . Thus,  $K_N$ ,  $K_N^H$ , and  $K_N K_N^H$  share the same eigenvectors, and the eigenvalues of  $K_N K_N^H$  are

$$\lambda(K_N K_N^H) = \lambda(K_N) \lambda^*(K_N) = |\lambda(K_N)|^2.$$

Any eigenvalue of  $K_N$  belongs to the set of eigenvalues of  $R_{2N}$ , which are

$$\rho_n = \lambda_n(R_{2N}) = \sum_{k=-(N-1)}^{N-1} t_k e^{i2\pi kn/2N}, \quad 1 \leq n \leq 2N.$$

It is clear that  $\rho_n$  is a partial sum of the infinite series  $\sum_{-\infty}^{\infty} t_k e^{-ik\theta}$  with  $\theta = -n\pi/N$ . With (3.2),  $|\rho_n| \geq \mu_T - \mu$ , where  $\mu$  can be made arbitrarily small by choosing sufficiently large  $N$  so that

$$\|(K_N K_N^H)^{-1}\|_2 \leq \frac{1}{(\mu_T - \mu)^2} < \infty.$$

Similar arguments can be used to prove the boundness of  $\|(S_N S_N^H)^{-1}\|_2$ , and the proof is completed.  $\square$

The next theorem describes the clustering property of the singular values of  $K_N^{-1}T_N$  and  $S_N^{-1}T_N$ .

**THEOREM 2.** *Let  $T_N$  be an  $N \times N$  Toeplitz matrix with the generating sequence satisfying (3.1) and (3.2). For sufficiently large  $N$ , the singular values of the preconditioned matrices  $K_N^{-1}T_N$  and  $S_N^{-1}T_N$  are clustered around unity except for a fixed number independent of  $N$ .*

*Proof.* Note that the singular value of  $K_N^{-1}T_N$  is equal to the square root of the corresponding eigenvalue of  $(K_N^{-1}T_N)^H(K_N^{-1}T_N)$ . Since  $(K_N^{-1}T_N)^H(K_N^{-1}T_N)$  and  $(K_N K_N^H)^{-1}(T_N T_N^H)$  are similar, the eigenvalues of  $(K_N K_N^H)^{-1}(T_N T_N^H)$  are examined to understand the singular values of  $K_N^{-1}T_N$ . With the relation  $K_N = T_N + \Delta T_N$ , we have

$$\lambda[(K_N K_N^H)^{-1}(T_N T_N^H)] = 1 - \lambda[(K_N K_N^H)^{-1}(K_N \Delta T_N^H + \Delta T_N K_N^H - \Delta T_N \Delta T_N^H)].$$

Let us define

$$W_N = K_N \Delta T_N^H + \Delta T_N K_N^H - \Delta T_N \Delta T_N^H.$$

For an arbitrary positive integer  $q$ , we denote the corresponding central  $(N - 2q) \times (N - 2q)$  diagonal block of  $(K_N K_N^H)^{-1}$  and  $W_N$  by  $\mathcal{K}_{N-2q}^{-1}$  and  $\mathcal{W}_{N-2q}$ , respectively. By the separation theorem (or intertwining theorem) of generalized eigenvalues [18], [20], there are at least  $N - 4q$  eigenvalues of  $(K_N K_N^H)^{-1}W_N$  bounded by the minimum and the maximum eigenvalues of  $\mathcal{K}_{N-2q}^{-1}\mathcal{W}_{N-2q}$ .

Since  $\mathcal{K}_{N-2q}^{-1}$  is a submatrix of the symmetric circulant matrix  $(K_N K_N^H)^{-1}$ ,

$$\|\mathcal{K}_{N-2q}^{-1}\|_2 \leq \|(K_N K_N^H)^{-1}\|_2.$$

According to the definition of  $\mathcal{W}_{N-2q}$ ,

$$\mathcal{W}_{N-2q} = \mathcal{K} \Delta T^H + \Delta T \mathcal{K}^H - \Delta T \Delta T^H,$$

where  $\mathcal{K}$  and  $\Delta T$  are  $(N - 2q) \times N$  matrices formed by the central  $(N - 2q)$  rows of  $K_N$  and  $\Delta T_N$ , respectively. It is easy to verify that, for  $p = 1, \infty$ ,

$$\|\mathcal{K}\|_p \leq 2 \sum_{n=-(N-1)}^{N-1} |t_n| \leq 2 \sum_{n=-\infty}^{\infty} |t_n| \leq 2B_T < \infty,$$

and

$$\|\Delta T\|_p \leq \sum_{n=q+1}^{N-1} (|t_n| + |t_{-n}|) \leq \sum_{n=q+1}^{\infty} (|t_n| + |t_{-n}|) = \sigma(q).$$

Since  $\|A\|_2 \leq (\|A\|_1 \|A\|_\infty)^{1/2}$  for an arbitrary matrix  $A$ , the above bounds also hold for  $p = 2$ . Similarly, we can argue that  $\|\mathcal{K}^H\|_2 \leq 2B_T < \infty$  and  $\|\Delta T^H\|_2 \leq \sigma(q)$ . Thus,

$$\begin{aligned} \|\mathcal{W}_{N-2q}\|_2 &\leq \|\mathcal{K}\|_2 \|\Delta T^H\|_2 + \|\Delta T\|_2 \|\mathcal{K}^H\|_2 + \|\Delta T\|_2 \|\Delta T^H\|_2 \\ &\leq 4B_T \sigma(q) + \sigma^2(q). \end{aligned}$$

By using Theorem 1 and the fact that  $\sigma(q)$  is smaller as  $q$  becomes larger due to (3.1), we conclude that for given  $\epsilon$  there exist  $q$  and  $\tilde{N}$  such that for all  $N \geq \tilde{N}$ ,

$$\|\mathcal{K}_{N-2q}^{-1}\|_2 \|\mathcal{W}_{N-2q}\|_2 \leq \|(K_N K_N^H)^{-1}\|_2 \|\mathcal{W}_{N-2q}\|_2 \leq \epsilon.$$

Hence, the eigenvalues of  $(K_N K_N^H)^{-1} (T_N T_N^H)$  are confined in the interval  $(1 - \epsilon, 1 + \epsilon)$  except at most  $4q$  outlying eigenvalues. Similar arguments can be used to prove the spectral clustering property of the singular values of  $S_N^{-1} T_N$ .  $\square$

The solution of a Toeplitz system  $T_N \mathbf{x} = \mathbf{b}$  can be determined by applying the CGN method to the preconditioned system  $K_N^{-1} T_N \mathbf{x} = K_N^{-1} \mathbf{b}$  (or  $S_N^{-1} T_N \mathbf{x} = S_N^{-1} \mathbf{b}$ ). With the clustering property on singular values of  $K_N^{-1} T_N$  (or  $S_N^{-1} T_N$ ), the CGN method converges superlinearly. A more detailed discussion about the convergence rate of the CGN method for solving Toeplitz systems were studied by Chan and Yeung [4]. When the generating function is additionally rational, we characterize the eigenvalues of the preconditioned matrices  $K_N^{-1} T_N$  and  $S_N^{-1} T_N$  directly. It will be detailed in the following sections.

**4. Spectral properties of preconditioned rational Toeplitz  $K_N^{-1} T_N$ .** The generating function of a sequence of Toeplitz matrices  $T_m$  is defined as

$$T(z) = \sum_{n=-\infty}^{\infty} t_n z^{-n}.$$

Let the generating function of  $T_N$  be of the form

$$(4.1) \quad T(z) = \frac{A(z^{-1})}{B(z^{-1})} + \frac{C(z)}{D(z)},$$

where

$$\frac{A(z^{-1})}{B(z^{-1})} = \frac{a_0 + a_1 z^{-1} + \dots + a_\alpha z^{-\alpha}}{1 + b_1 z^{-1} + \dots + b_\beta z^{-\beta}}, \quad \frac{C(z)}{D(z)} = \frac{c_0 + c_1 z + \dots + c_\gamma z^\gamma}{1 + d_1 z + \dots + d_\delta z^\delta}.$$

Note that  $a_\alpha b_\beta c_\gamma d_\delta \neq 0$  and polynomials  $A(z^{-1})$  and  $B(z^{-1})$  (or  $C(z)$  and  $D(z)$ ) have no common factor. We call  $T(z)$  a rational function of order  $(\alpha, \beta, \gamma, \delta)$  and  $T_N$  a rational Toeplitz matrix. To simplify the notation, we define  $r = \max(\alpha, \beta)$  and  $s = \max(\gamma, \delta)$ .

The spectral properties of  $K_N^{-1} T_N$  can be determined from that of  $T_N^{-1} \Delta T_N$  via

$$(4.2) \quad [\lambda(K_N^{-1} T_N)]^{-1} = \lambda(T_N^{-1} (T_N + \Delta T_N)) = 1 + \lambda(T_N^{-1} \Delta T_N).$$

The eigenvalues of  $K_N^{-1} T_N$  clustered around 1 correspond to those of  $T_N^{-1} \Delta T_N$  clustered around 0. We summarize the procedures in examining the spectral properties of  $T_N^{-1} \Delta T_N$  as follows:

*Step 1.* Show that the  $\Delta T_N$  is asymptotically equivalent to a low-rank Toeplitz matrix  $\Delta F_N$  (Lemma 2).

*Step 2.* Study the rank of  $\Delta F_N$  by transforming it to a matrix  $Q_F$  which has at most  $d = r + s$  nonzero columns (Lemma 3).

*Step 3.* Show that the  $Q_F$  is asymptotically equivalent to a matrix  $\overline{Q}_F$  which has at most  $2 \min(r, s)$  nonzero eigenvalues (Lemma 4).

*Step 4.* Use perturbation theory to determine the radius of the clustered eigenvalues of  $T_N^{-1} \Delta T_N$  and  $K_N^{-1} T_N$  (Lemmas 5 and 6 and Theorem 3).

The number of outliers of  $K_N^{-1} T_N$ , i.e.,  $2 \min(r, s)$ , is determined from Steps 1–3, and the clustering radius is determined from Step 4.

**4.1. The number of outliers of  $K_N^{-1}T_N$ .** Note that the sequence  $t_n$  can be recursively calculated for large  $|n|$ . This is stated as follows.

LEMMA 1. *The sequence  $t_n$  generated by (4.1) follows the recursions,*

$$(4.3) \quad \begin{aligned} t_{n+1} &= -(b_1 t_n + b_2 t_{n-1} + \dots + b_\beta t_{n-\beta+1}), & n \geq r = \max(\alpha, \beta), \\ t_{n-1} &= -(d_1 t_n + d_2 t_{n+1} + \dots + d_\delta t_{n+\delta-1}), & n \leq -s = -\max(\gamma, \delta). \end{aligned}$$

*Proof.* The proof is similar to the proof of Lemma 1 in [12].  $\square$

Since elements  $t_n$  satisfy the recursion given in Lemma 1, we construct a low-rank Toeplitz matrix  $\Delta F_N$  as

$$(4.4) \quad \Delta F_N = F_{1,N} + F_{2,N},$$

where

$$F_{1,N} = \begin{bmatrix} t_N & t_{N-1} & \cdot & t_2 & t_1 \\ t_{N+1} & t_N & t_{N-1} & \cdot & t_2 \\ \cdot & t_{N+1} & t_N & \cdot & \cdot \\ t_{2N-2} & \cdot & \cdot & \cdot & t_{N-1} \\ t_{2N-1} & t_{2N-2} & \cdot & t_{N+1} & t_N \end{bmatrix}$$

and

$$F_{2,N} = \begin{bmatrix} t_{-N} & t_{-(N+1)} & \cdot & t_{-(2N-2)} & t_{-(2N-1)} \\ t_{-(N-1)} & t_{-N} & t_{-(N+1)} & \cdot & t_{-(2N-2)} \\ \cdot & t_{-(N-1)} & t_{-N} & \cdot & \cdot \\ t_{-2} & \cdot & \cdot & \cdot & t_{-(N+1)} \\ t_{-1} & t_{-2} & \cdot & t_{-(N-1)} & t_{-N} \end{bmatrix},$$

and where  $t_n, n \geq r$  or  $n \leq -s$ , are recursively defined by (4.3). Due to the recursion given by (4.3), the ranks of  $F_{1,N}$  and  $F_{2,N}$  are bounded by  $r$  and  $s$ , respectively. Thus, the rank of  $\Delta F$  is bounded by  $d = r + s$ . The following lemma shows that  $\Delta T_N$  and  $\Delta F_N$  are, in fact, asymptotically equivalent.

LEMMA 2. *Let  $T_N$  be an  $N \times N$  Toeplitz matrix generated by  $T(z)$  in (4.1) with the corresponding generating sequence satisfying (3.1) and (3.2). The  $\Delta T_N$  and  $\Delta F_N$  are asymptotically equivalent.*

*Proof.* Let us denote the difference between  $\Delta F_N$  and  $\Delta T_N$  by

$$(4.5) \quad \Delta E_N = \Delta F_N - \Delta T_N = \begin{bmatrix} t_N + t_{-N} & t_{-(N+1)} & \cdot & t_{-(2N-2)} & t_{-(2N-1)} \\ t_{N+1} & t_N + t_{-N} & t_{-(N+1)} & \cdot & t_{-(2N-2)} \\ \cdot & t_{N+1} & t_N + t_{-N} & \cdot & \cdot \\ t_{2N-2} & \cdot & \cdot & \cdot & t_{-(N+1)} \\ t_{2N-1} & t_{2N-2} & \cdot & t_{N+1} & t_N + t_{-N} \end{bmatrix}.$$

It can be easily verified that the  $l_1$  and  $l_\infty$  norms of  $\Delta E_N$  are both bounded by

$$(4.6) \quad \tau_E = \sum_{n=N}^{2N-1} |t_n| + \sum_{n=-N}^{-(2N-1)} |t_n|.$$

Consequently, we have

$$\|\Delta E_N\|_2 \leq (\|\Delta E_N\|_1 \|\Delta E_N\|_\infty)^{1/2} \leq \tau_E.$$

Since  $\tau_E$  goes to zero as  $N$  goes to infinity due to (3.1), the proof is completed.  $\square$

Since  $\Delta T_N$  is asymptotically equivalent to  $\Delta F_N$  and the rank of  $\Delta F_N$  is bounded by  $d$ , the number of outliers of  $T_N^{-1}\Delta T_N$  (or  $K_N^{-1}T_N$ ) is bounded by  $d$ , which is, however, not tight. We are able to determine a tighter bound by introducing another asymptotically equivalent matrix of  $\Delta T_N$  (or  $\Delta F_N$ ), which has only  $2 \min(r, s)$  nonzero eigenvalues in the following. This turns out to be the exact number of outliers actually observed in all our numerical experiments. To exploit the low-rank structure of  $\Delta F_N$ , we transform  $\Delta F_N$  to

$$(4.7) \quad Q_F = \Delta F_N U_D L_B,$$

where  $U_D$  is an  $N \times N$  upper triangular Toeplitz matrix with the first  $N$  coefficients in  $D(z)$  as the first row, and  $L_B$  is an  $N \times N$  lower triangular Toeplitz matrix with the first  $N$  coefficients in  $B(z^{-1})$  as the first column. Note that since  $U_D$  and  $L_B$  are full-rank matrices, the  $Q_F$  and  $\Delta F_N$  have the same rank. The structure of  $Q_F$  is described in the following lemma.

**LEMMA 3.** *Let  $T_N$  be an  $N \times N$  Toeplitz matrix generated by  $T(z)$  in (4.1) with the corresponding generating sequence satisfying (3.1) and (3.2). The elements of  $Q_F$  are zeros except for the first  $s$  and the last  $r$  columns.*

*Proof.* Note that  $F_{1,N}$  and  $F_{2,N}$  are Toeplitz matrices with elements

$$(F_{1,N})_{i,j} = t_{N+i-j} \quad \text{and} \quad (F_{2,N})_{i,j} = t_{-N+i-j}.$$

The  $(i, j)$  elements of  $F_{1,N}U_D L_B$  and  $F_{2,N}U_D L_B$  are

$$\sum_{n=1}^N \sum_{m=1}^N t_{N+i-m} d_{n-m} b_{n-j} \quad \text{and} \quad \sum_{n=1}^N \sum_{m=1}^N t_{-N+i-m} d_{n-m} b_{n-j},$$

where  $b_0 = 1$  ( $d_0 = 1$ ) and  $b_i = 0$  ( $d_i = 0$ ) if the subscript  $i$  is not in the range  $0 \leq i \leq \beta$  ( $0 \leq i \leq \delta$ ). If  $s < j \leq N - r$ , we can simplify the above summations as

$$\sum_{m'=0}^{\delta} \left( \sum_{n'=0}^{\beta} t_{N+i+m'-n'-j} b_{n'} \right) d_{m'} = 0$$

and

$$\sum_{n'=0}^{\beta} \left( \sum_{m'=0}^{\delta} t_{-N+i+m'-n'-j} d_{m'} \right) b_{n'} = 0,$$

where  $m' = n - m$ ,  $n' = n - j$ , and the equalities are due to the recursion defined in (4.3). Thus, the elements of

$$Q_F = \Delta F_N U_D L_B = (F_{1,N} + F_{2,N}) U_D L_B$$

are zeros except for the first  $s$  and the last  $r$  columns.  $\square$

Consequently, we decompose the complex  $N$ -tuple space  $C^N$  into two orthogonal complement subspaces,

$$(4.8) \quad \begin{aligned} \mathcal{R}(Q_F) &= \{ \mathbf{v} \in C^N \mid \mathbf{v}_i = 0, s < i \leq N - r \}, \\ \mathcal{N}(Q_F) &= \{ \mathbf{v} \in C^N \mid \mathbf{v}_i = 0, 1 \leq i \leq s \text{ or } N - r < i \leq N \}, \end{aligned}$$

with dimensions

$$\dim \mathcal{R}(Q_F) = d \quad \text{and} \quad \dim \mathcal{N}(Q_F) = N - d.$$

The subspace  $\mathcal{N}(Q_F)$  is contained in the null space of  $Q_F$ . Let  $Q_{NW}$  denote the northwest  $s \times s$  block in  $Q_F$ , and  $Q_{NE}$ ,  $Q_{SW}$ , and  $Q_{SE}$  the corresponding corner blocks in  $Q_F$  with sizes  $s \times r$ ,  $r \times s$ , and  $r \times r$ , respectively. By using the subspace decomposition (4.8), it is easy to see that the nonzero eigenvalues of  $Q_F$  only depend on the corresponding four corner blocks of  $Q_F$ , and are also the eigenvalues of the  $d \times d$  matrix,

$$P_F = \begin{bmatrix} Q_{NW} & Q_{NE} \\ Q_{SW} & Q_{SE} \end{bmatrix}.$$

In other words, the rank of  $Q_F$  is the same as that of  $P_F$ .

The bounds for the elements of  $Q_{NW}$ ,  $Q_{NE}$ ,  $Q_{SW}$ , and  $Q_{SE}$  are summarized as follows:

$$(4.9) \quad \begin{aligned} |(Q_{NW})_{i,j}| &\leq \tau_{NW}, & \tau_{NW} &= O(|t_N| + |t_{-N}|), \\ |(Q_{SE})_{i,j}| &\leq \tau_{SE}, & \tau_{SE} &= O(|t_N| + |t_{-N}|), \\ |(Q_{NE})_{i,j}| &\leq (F_{1,N} U_D L_B)_{i,N-r+j} + \tau_{NE}, & \tau_{NE} &= O(|t_{-2N}|), \\ |(Q_{SW})_{i,j}| &\leq (F_{2,N} U_D L_B)_{N-s+i,j} + \tau_{SW}, & \tau_{SW} &= O(|t_{2N}|). \end{aligned}$$

To derive (4.9), recall that the  $(i, j)$  element of  $Q_F$  is

$$\sum_{n=1}^N \sum_{m=1}^N t_{N+i-m} d_{n-m} b_{n-j} + \sum_{n=1}^N \sum_{m=1}^N t_{-N+i-m} d_{n-m} b_{n-j},$$

which is bounded by

$$\sum_{m'=0}^{\delta} \sum_{n'=0}^{\beta} |t_{N+i+m'-n'-j}| |d_{m'}| |b_{n'}| + \sum_{m'=0}^{\delta} \sum_{n'=0}^{\beta} |t_{-N+i+m'-n'-j}| |d_{m'}| |b_{n'}|.$$

Since the elements of  $Q_{NW}$  are the same as those of  $Q_F$  with subscript  $(i, j)$ ,  $i, j \leq s$ , they are bounded by

$$\tau_{NW} = \sum_{i=0}^{\beta} |b_i| \sum_{j=0}^{\delta} |d_j| \left( \max_{-(s+\beta) < n < s+\delta} |t_{N+n}| + \max_{-(s+\beta) < n < s+\delta} |t_{-N+n}| \right).$$

To determine the bound for  $\sum_{i=0}^{\beta} |b_i|$ , we factorize  $B(z^{-1})$  as

$$B(z^{-1}) = (1 - r_1 z^{-1})(1 - r_2 z^{-1}) \cdots (1 - r_{\beta} z^{-1}).$$

A direct consequence of (3.1) is that all poles of  $A(z^{-1})/B(z^{-1})$  should lie inside the unit circle, i.e.,  $|r_i| < 1$ ,  $1 \leq i \leq \beta$ , so that

$$|b_k| \leq \binom{\beta}{k} (\max |r_i|)^k \leq \binom{\beta}{k}, \quad \text{where} \quad \binom{\beta}{k} \equiv \frac{\beta!}{(\beta - k)! k!}.$$

Therefore, we obtain

$$\sum_{k=0}^{\beta} |b_k| \leq \sum_{k=0}^{\beta} \binom{\beta}{k} \leq 2^{\beta}.$$

Similarly,  $\sum_{k=0}^{\delta} |d_k| \leq 2^{\delta}$  and thus, the elements of  $Q_{NW}$  are bounded by

$$\tau_{NW} = 2^{(\beta+\delta)} (|t_{N-s-\beta}| + |t_{-N+s+\delta}|) = O(|t_N| + |t_{-N}|),$$

where the last equality is due to the fact that, for large  $n$ ,  $t_n$  can be approximated by

$$(4.10) \quad t_n \approx cr_j^n, \quad \text{where} \quad |r_j| = \max_i |r_i|,$$

and where  $c$  is a constant. Similarly, we can prove that the elements of  $Q_{SE}$  are bounded by

$$\tau_{SE} = 2^{(\beta+\delta)} (|t_{N-r-\beta}| + |t_{-N+r+\delta}|) = O(|t_N| + |t_{-N}|).$$

The  $(i, j)$ ,  $1 \leq i \leq s, 1 \leq j \leq r$ , element of  $Q_{NE}$  is the sum of the  $(i, N-r+j)$  elements of  $F_{1,N}U_D L_B$  and  $F_{2,N}U_D L_B$ . It is straightforward to verify that the  $(i, N-r+j)$  element of  $F_{1,N}U_D L_B$  remains unchanged while that of  $F_{2,N}U_D L_B$  is bounded by  $\tau_{NE} = 2^{(\beta+\delta)} |t_{-2N+d+\delta}| = O(|t_{-2N}|)$  for sufficiently large  $N$ . Similarly, we can derive the bound for the elements in  $Q_{SW}$  as given by (4.9).

Thus, when  $N$  becomes asymptotically large, the  $P_F$  converges to

$$\bar{P}_F = \begin{bmatrix} 0 & \bar{Q}_{NE} \\ \bar{Q}_{SW} & 0 \end{bmatrix},$$

where  $\bar{Q}_{NE}$  is the converged northeast  $s \times r$  block in  $F_{1,N}U_D L_B$  and  $\bar{Q}_{SW}$  is the converged southwest  $r \times s$  block in  $F_{2,N}U_D L_B$ . Since the ranks of  $\bar{Q}_{NE}$  and  $\bar{Q}_{SW}$  are both bounded by  $\min(r, s)$ , the rank of  $\bar{P}_F$  is bounded by  $\eta = 2 \min(r, s)$ .

Let us define a matrix  $\bar{Q}_F$  by replacing the four corner blocks in  $Q_F$  with the corresponding blocks in  $\bar{P}_F$ . Then, we have

$$\begin{aligned} \tau_Q &= \|Q_F - \bar{Q}_F\|_p = \|P_F - \bar{P}_F\|_p \\ &\leq s\tau_{NW} + r\tau_{SE} + \max(r, s)(\tau_{NE} + \tau_{SW}) \\ &= O(|t_N| + |t_{-N}|), \end{aligned}$$

for  $p = 1$  and  $\infty$ . The above bounds also hold for  $p = 2$  because

$$\|A\|_2 \leq (\|A\|_1 \|A\|_{\infty})^{1/2}$$

for an arbitrary matrix  $A$ . Since  $\tau_Q$  goes to zero as  $N$  goes to infinity due to (3.1), the asymptotic equivalence between  $Q_F$  and  $\bar{Q}_F$  is established. This result is summarized in the following lemma.

LEMMA 4. Let  $T_N$  be an  $N \times N$  Toeplitz matrix generated by  $T(z)$  in (4.1) with the corresponding generating sequence satisfying (3.1) and (3.2). The  $Q_F$  and  $\bar{Q}_F$  are asymptotically equivalent.

Based on Lemmas 2–4, (4.2) and (4.7),  $T_N^{-1} \Delta T_N$  is asymptotically equivalent to  $T_N^{-1} \bar{Q}_F L_B^{-1} U_D^{-1}$  whose rank is bounded by  $\eta = 2 \min(r, s)$  and  $K_N^{-1} T_N$  has at most  $\eta$  asymptotic eigenvalues not converging to one (outliers).

**4.2. The clustering radius of  $K_N^{-1} T_N$ .** We use perturbation theory to estimate the clustering radius of the  $N - \eta$  clustered eigenvalues. Instead of examining the eigenvalues of  $T_N^{-1} \Delta T_N$  directly, we study those of the similar matrix

$$G_N = L_B^{-1} U_D^{-1} T_N^{-1} \Delta T_N U_D L_B = L_B^{-1} U_D^{-1} T_N^{-1} Q_T,$$

where  $Q_T = \Delta T_N U_D L_B$ . Let us define

$$H_N = L_B^{-1} U_D^{-1} T_N^{-1} \overline{Q}_F.$$

It is clear that  $H_N$  has only  $d$  nonzero columns as  $\overline{Q}_F$  (or  $Q_F$ ). The  $G_N$  can be viewed as a matrix obtained from  $H_N$  by adding the perturbation matrix

$$(4.11) \quad \Delta G_N = G_N - H_N = L_B^{-1} U_D^{-1} T_N^{-1} (Q_T - \overline{Q}_F).$$

A bound of  $\|\Delta G_N\|_2$  is given below so that we can estimate the clustering radius of the clustered eigenvalues by using perturbation theory for eigenvalues.

LEMMA 5. *Let  $T_N$  be an  $N \times N$  Toeplitz matrix generated by  $T(z)$  in (4.1) with the corresponding generating sequence satisfying (3.1) and (3.2). Then, for sufficiently large  $N$ , the  $\|\Delta G_N\|_2$  is bounded by  $\epsilon = O(|t_N| + |t_{-N}|)$ .*

*Proof.* We first study the 2-norm of  $Q_T - \overline{Q}_F$ , which is bounded by

$$\|Q_T - \overline{Q}_F\|_2 \leq \|Q_T - Q_F\|_2 + \|Q_F - \overline{Q}_F\|_2.$$

As shown in the proof of Lemma 4, the second term  $\|Q_F - \overline{Q}_F\|_2$  is bounded by  $\tau_Q = O(|t_N| + |t_{-N}|)$  while the first term  $\|Q_T - Q_F\|_2$  is bounded by

$$\|Q_T - Q_F\|_2 \leq \|\Delta T_N - \Delta F_N\|_2 \|U_D\|_2 \|L_B\|_2 = \|\Delta E_N\|_2 \|U_D\|_2 \|L_B\|_2.$$

Recall from (4.6) that  $\|\Delta E_N\|_2 \leq \sum_{n=N}^{2N-1} (|t_n| + |t_{-n}|)$ . By using (4.10), we have

$$\sum_{n=N}^{2N-1} |t_n| \leq \sum_{n=N}^{\infty} |q_j r_j^n| = \frac{|t_N|}{1 - |r_j|} = M_B |t_N|, \quad \text{where } M_B = \frac{1}{1 - |r_j|}.$$

Similarly,  $\sum_{n=N}^{2N-1} |t_{-n}| \leq M_D |t_{-N}|$ . Besides,  $\|L_B\|_2 \leq \sum_{k=0}^{\beta} |b_k| \leq 2^{\beta}$  and  $\|U_D\|_2 \leq \sum_{k=0}^{\delta} |d_k| \leq 2^{\delta}$ . Thus, we obtain a bound for the first term, i.e.,

$$\|Q_T - Q_F\|_2 \leq 2^{(\beta+\delta)} (M_B |t_N| + M_D |t_{-N}|) = O(|t_N| + |t_{-N}|),$$

and conclude that

$$\|Q_T - \overline{Q}_F\|_2 \leq O(|t_N| + |t_{-N}|).$$

With (4.11), we have

$$\|\Delta G_N\|_2 \leq \|L_B^{-1}\|_2 \|U_D^{-1}\|_2 \|T_N^{-1}\|_2 \|Q_T - \overline{Q}_F\|_2.$$

Under the assumption that  $\|T_N^{-1}\|_2$  is bounded by a constant  $c_T$  independent of  $N$ , we want to show that  $\|L_B^{-1}\|_2$  and  $\|U_D^{-1}\|_2$  are also bounded. Let us factorize  $B(z^{-1})$  as

$$B(z^{-1}) = (1 - r_1 z^{-1})(1 - r_2 z^{-1}) \cdots (1 - r_{\beta} z^{-1}),$$

where we assume that all roots  $r_i$  are distinct for simplicity. By applying the isomorphism between the ring of the power series and the ring of semi-infinite lower (or upper) triangular Toeplitz matrices, the  $L_B$  and  $L_B^{-1}$  can be decomposed into the products

$$L_B = L_{r_1} L_{r_2} \cdots L_{r_{\beta}}, \quad L_B^{-1} = L_{r_{\beta}}^{-1} \cdots L_{r_2}^{-1} L_{r_1}^{-1},$$

where  $L_{r_i}, 1 \leq i \leq \beta$  is an  $N \times N$  lower triangular Toeplitz matrix with  $[1, -r_i, 0, \dots, 0]^T$  as the first column. It can be easily verified that  $L_{r_i}^{-1}$  is a lower triangular Toeplitz matrix with  $[1, r_i, r_i^2, \dots, r_i^{N-1}]^T$  as the first column. Therefore,

$$\|L_{r_i}^{-1}\|_p \leq \sum_{k=0}^{N-1} |r_i^k| \leq \sum_{k=0}^{\infty} |r_i^k| = \frac{1}{1 - |r_i|}, \quad p = 1, 2, \infty,$$

and

$$\|L_B^{-1}\|_2 \leq \prod_{i=1}^{\beta} \|L_{r_i}^{-1}\|_2 \leq \prod_{i=1}^{\beta} \frac{1}{1 - |r_i|} = c_B.$$

Similar arguments can be used to prove that  $\|U_D^{-1}\|_2 \leq c_D$ . Finally, we have

$$(4.12) \quad \|\Delta G_N\|_2 \leq \epsilon \equiv c_B c_D c_T \|(Q_T - \bar{Q}_F)\|_2 = O(|t_N| + |t_{-N}|).$$

The proof is completed.  $\square$

Let us denote the rank of  $H_N = L_B^{-1} U_D^{-1} T_N^{-1} \bar{Q}_F$  by  $\tilde{\eta}$ . Clearly,  $\tilde{\eta} \leq \eta = 2 \min(r, s)$ . We arrange the eigenvalues of  $H_N$  in a descending order so that  $|\lambda_n| \geq |\lambda_{n+1}|$  ( $\lambda_n = 0$  for  $\tilde{\eta} < n \leq N$ ), and denote the corresponding normalized right-hand and left-hand eigenvectors by  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$  and  $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N$ , respectively. Besides, vectors  $\mathbf{x}_n$  with  $\tilde{\eta} < n \leq N$  are chosen to be othogonal. The complex  $N$ -tuple space is decomposed into the row and the null spaces of  $H_N$ ,

$$\text{Row}(H_N) = \text{span}\{\mathbf{x}_n, n \leq \tilde{\eta}\}, \quad \text{Null}(H_N) = \text{span}\{\mathbf{x}_n, \tilde{\eta} < n \leq N\}.$$

Since  $G_N = H_N + \Delta G_N$  and  $\|\Delta G_N\|_2 \leq \epsilon$ , the eigenvalues and the right-hand eigenvectors of  $G_N$  are denoted by  $\lambda_n(\epsilon)$  and  $\mathbf{x}_n(\epsilon)$ , respectively. By using results from perturbation theory for repeated eigenvalues [20], the eigenvectors  $\mathbf{x}_n(\epsilon)$  with  $\tilde{\eta} < n \leq N$  must take the form

$$(4.13) \quad \mathbf{x}_n(\epsilon) = \sum_{m=1}^{\tilde{\eta}} \frac{\xi_{mn}}{(\lambda_n - \lambda_m)s_m} \mathbf{x}_m + \sum_{m=\tilde{\eta}+1}^N g_{mn} \mathbf{x}_m + O(\epsilon^2),$$

where  $\xi_{mn} = \mathbf{y}_m^H \Delta G_N \mathbf{x}_n$ ,  $\lambda_n = 0$ ,  $s_m = \mathbf{y}_m^H \mathbf{x}_m$  and  $g_{nn} = 1$ . Due to the construction, we know that

$$(4.14) \quad \|\mathbf{x}_n(\epsilon)\|_2 \geq \|\mathbf{x}_n\|_2 = 1.$$

The factor  $|\xi_{mn}|$  is bounded by

$$|\xi_{mn}| = |\mathbf{y}_m^H \Delta G_N \mathbf{x}_n| \leq \|\mathbf{y}_m\|_2 \|\Delta G_N\|_2 \|\mathbf{x}_n\|_2 \leq \epsilon.$$

The  $|s_m^{-1}|, 1 \leq m \leq \tilde{\eta}$ , is also bounded as given in the following lemma.

**LEMMA 6.** *Let  $T_N$  be an  $N \times N$  Toeplitz matrix generated by  $T(z)$  in (4.1) with the corresponding generating sequence satisfying (3.1) and (3.2). Then, the  $|s_m^{-1}|, 1 \leq m \leq \tilde{\eta}$ , of  $H_N$  is bounded by a constant independent of  $N$ .*

*Proof.* The eigenvalues  $\lambda$  and the right-hand eigenvectors  $\mathbf{x}$  of  $H_N$  satisfy

$$L_B \bar{Q}_F \mathbf{x} = \lambda L_B T_N U_D L_B \mathbf{x}.$$

Since the elements of  $\overline{Q}_F$  are zeros except the first  $s$  and the last  $r$  columns, so are the elements of  $L_B \overline{Q}_F$ . Thus, the nonzero eigenvalues of  $H_N$  only depend on the northwest  $s \times s$ , northeast  $s \times r$ , southwest  $r \times s$ , and southeast  $r \times r$  blocks of  $L_B \overline{Q}_F$  and  $L_B T_N U_D L_B$ . The boundness of  $|s_m^{-1}|, 1 \leq m \leq \tilde{\eta}$ , is guaranteed if the elements of the four corner blocks of  $L_B \overline{Q}_F$  and  $L_B T_N U_D L_B$  remain unchanged for sufficiently large  $N$ .

By using the band structure of  $L_B$  and the special structure of  $\overline{Q}_F$ , it is straightforward to verify that the four blocks of  $L_B \overline{Q}_F$  remain unchanged for large  $N$ . Next, we examine the matrix  $L_B T_N U_D L_B$ . By using (4.1) and the isomorphism between the ring of the power series and the ring of the semi-infinite lower (or upper) triangular Toeplitz matrices, we can express  $T_N$  as

$$T_N = L_A L_B^{-1} + U_C U_D^{-1},$$

where  $L_A$  is an  $N \times N$  lower triangular Toeplitz matrix with the first  $N$  coefficients in  $A(z^{-1})$  as the first column, and  $U_C$  is an  $N \times N$  upper triangular Toeplitz matrix with the first  $N$  coefficients in  $C(z)$  as the first row. Then, we have

$$L_B T_N U_D L_B = L_A U_D L_B + L_B U_C L_B,$$

whose four corner blocks remain unchanged for large  $N$ . Thus,  $\lambda_m$  and  $s_m = \mathbf{y}_m^H \mathbf{x}_m$  with  $1 \leq m \leq \tilde{\eta}$  do not change with  $N$ , when  $N$  becomes sufficiently large.  $\square$

Let  $\mathbf{v}_n(\epsilon)$  be the normalized vector of  $\mathbf{x}_n(\epsilon)$ ,

$$\mathbf{v}_n(\epsilon) = \frac{\mathbf{x}_n(\epsilon)}{\|\mathbf{x}_n(\epsilon)\|_2},$$

which can be decomposed as

$$\mathbf{v}_n(\epsilon) = \mathbf{v}_N(\epsilon) + \mathbf{v}_R(\epsilon),$$

where  $\mathbf{v}_N(\epsilon) \in \text{Null}(H_N)$  and  $\mathbf{v}_R(\epsilon) \in \text{Row}(H_N)$ . The magnitude of  $\lambda_n(\epsilon), \tilde{\eta} < n \leq N$ , of  $G_N$  is approximated by

$$|\lambda_n(\epsilon)| = \|G_N \mathbf{v}_n(\epsilon)\|_2 = \|H_N \mathbf{v}_R(\epsilon) + \Delta G_N \mathbf{v}_n(\epsilon)\|_2.$$

By using (4.12)–(4.14), we obtain that

$$\begin{aligned} \max_{\tilde{\eta} < n \leq N} |\lambda_n(\epsilon)| &\leq \max_{\tilde{\eta} < n \leq N} \|H_N \mathbf{v}_R(\epsilon)\|_2 + \max_{\tilde{\eta} < n \leq N} \|\Delta G_N \mathbf{v}_n(\epsilon)\|_2 \\ &\leq \sum_{m=1}^{\tilde{\eta}} \frac{\|\xi_{mn} H_N \mathbf{x}_m\|_2}{|\lambda_m s_m| \|\mathbf{x}_n(\epsilon)\|_2} + \|\Delta G_N\|_2 \\ &\leq \sum_{m=1}^{\tilde{\eta}} \frac{\epsilon}{|s_m|} + \epsilon = \epsilon_K \\ &= O(|t_N| + |t_{-N}|) \end{aligned}$$

for sufficiently large  $N$ . The above analysis is concluded in the following theorem.

**THEOREM 3.** *Let  $T_N$  be an  $N \times N$  Toeplitz matrix generated by  $T(z)$  in (4.1) with the corresponding generating sequence satisfying (3.1) and (3.2). For sufficiently large  $N$ , the preconditioned Toeplitz matrix  $K_N^{-1} T_N$  has the following two properties:*

P1. *The number of outliers is at most  $\eta = 2 \min(r, s)$ .*

P2. There are at least  $N - \eta$  eigenvalues confined in the disk centered at 1 with radius  $\epsilon_K$ , where

$$\epsilon_K = O(|t_N| + |t_{-N}|).$$

**5. Spectral properties of preconditioned rational Toeplitz  $S_N^{-1}T_N$ .** The preconditioned Toeplitz matrix  $S_N^{-1}T_N$  has similar spectral properties as  $K_N^{-1}T_N$ . The number of outliers of  $S_N^{-1}T_N$  can be obtained by proving that  $\Delta S_N = S_N - T_N$  and  $\Delta F_N$  given by (4.4) are asymptotically equivalent.

LEMMA 7. Let  $T_N$  be an  $N \times N$  Toeplitz matrix generated by  $T(z)$  in (4.1) with the corresponding generating sequence satisfying (3.1) and (3.2).  $S_N^{-1}T_N$  has asymptotically at most  $\eta = 2 \min(r, s)$  eigenvalues not converging to 1.

Proof. Let us define  $\Delta S_N = S_N - T_N$ , and express the difference between  $\Delta F_N$  in (4.4) and  $\Delta S_N$  as

$$\Delta F_N - \Delta S_N = E_{1,N} + E_{2,N},$$

where  $E_{1,N}$  and  $E_{2,N}$  are  $N \times N$  Toeplitz matrices with elements

$$(E_{1,N})_{i,j} = \begin{cases} t_{N+i-j}, & -(M-1) \leq i-j \leq N-1, \\ t_{i-j}, & -(N-1) \leq i-j \leq -M, \end{cases}$$

and

$$(E_{2,N})_{i,j} = \begin{cases} t_{i-j}, & N-(M-1) \leq i-j \leq N-1, \\ t_{i-j-N}, & -(N-1) \leq i-j \leq N-M, \end{cases}$$

respectively. By using similar arguments in deriving Lemma 2, we can prove that  $\Delta S_N$  and  $\Delta F_N$  are asymptotically equivalent. Since  $\Delta F_N$  is asymptotically equivalent to the matrix  $\bar{Q}_F L_B^{-1} U_D^{-1}$  with rank  $\tilde{\eta} \leq \eta = 2 \min(r, s)$  as described in Lemma 4, the proof is completed.  $\square$

Similar arguments used in §4.2 can be applied to derive the following theorem.

THEOREM 4. Let  $T_N$  be an  $N \times N$  Toeplitz matrix generated by  $T(z)$  in (4.1) with the corresponding generating sequence satisfying (3.1) and (3.2). For sufficiently large  $N$ , the preconditioned Toeplitz matrix  $S_N^{-1}T_N$  has the following two properties:

P1. The number of outliers is at most  $\eta = 2 \min(r, s)$ .

P2. There are at least  $N - \eta$  eigenvalues confined in the disk centered at 1 with radius  $\epsilon_S$ , where

$$\epsilon_S = O(|t_{N-M}| + |t_{1-M}|).$$

**6. Numerical results.** Five test problems, including both rational and nonrational  $T_N$ , are used to illustrate the above analysis. For the nonsymmetric Toeplitz system  $T_N \mathbf{x} = \mathbf{b}$  to be solved, we choose  $\mathbf{b} = (1, \dots, 1)^T$  and zero initial guesses in the first four experiments. Without further specification,  $M$  is chosen such that  $|t_{N-M}| \approx |t_{1-M}|$  to construct preconditioner  $S_N$ . We use the first test problem, which is generated by a nonrational function, to examine the clustering effect of singular values. Test Problems 2–4 are generated by rational functions so that the number of outliers and the clustering radius can be observed, which confirm the theoretical

TABLE 1  
Number of iterations required for the CGN method.

N	$T_N$	$S_N$	$K_N$
32	24	12	9
64	33	15	11
128	49	17	13

TABLE 2  
Number of iterations required for the CGS method.

N	$T_N$	$S_N$	$K_N$
32	15	7	9
64	21	8	10
128	26	9	10

results developed in §4 and §5. One underdetermined Toeplitz system (singular  $T_N$ ) is also given in Test Problem 5 with appropriate vector  $\mathbf{b}$ .

TEST PROBLEM 1. Nonrational  $T_N$ . Let  $T_N$  be a Toeplitz matrix with generating sequence

$$t_n = \begin{cases} 1/\log(2-n), & n \leq -1, \\ 1/\log(2-n) + 1/(1+n), & n = 0, \\ 1/(1+n), & n \geq 1. \end{cases}$$

The singular values of  $S_N^{-1}T_N$  and  $K_N^{-1}T_N$  are plotted in Fig. 1(a) for  $N = 32, 64,$  and  $128$ . Both  $S_N^{-1}T_N$  and  $K_N^{-1}T_N$  have clustered singular values. The eigenvalues of  $K_N^{-1}T_N$  with  $N = 32$  are plotted in Fig. 1(b). It is clear that the eigenvalues possess a certain clustering property. We apply both the CGN and CGS methods to solve the preconditioned Toeplitz system  $P_N^{-1}T_N\mathbf{x} = P_N^{-1}\mathbf{b}$ . The numbers of iterations required for the CGN and CGS methods to achieve  $\|\mathbf{b} - T_N\mathbf{x}\|_2 < 10^{-12}$  are summarized in Tables 1 and 2, respectively. The case without preconditioning, denoted by  $T_N$ , is also included for comparison. The use of preconditioners does accelerate the convergence rate of these iterative methods. The number of iterations required for  $S_N$  and  $K_N$  increase slightly as  $N$  becomes large. The  $K_N$  performs better than  $S_N$  in the CGN method. However, their performances are comparable for the CGS method. Note also that it requires more iterations for the CGN method to converge than the CGS method. Since the operation counts inside each iteration of the CGN and CGS methods are approximately equal, the CGS method performs better for this test problem. However, the CGN method may have a more robust convergence behavior for general problems. We refer to [13] for a more detailed comparison of these two nonsymmetric iterative algorithms.

Although the necessary conditions for convergence of the CGS method are not clear yet, there exists a direct relation between the convergence rate and the eigenvalue distribution of the iteration matrix [11], [13]. We will only present results of the CGS method for remaining test problems to verify the theoretic results about the eigenvalue distributions of the preconditioned rational Toeplitz matrices.

TEST PROBLEM 2. Rational  $T_N$  with  $(r, s) = (1, 1)$ . The generating function of  $T_N$  is chosen to be

$$T(z) = \frac{1 + 0.7z^{-1}}{1 - 0.9z^{-1}} + \frac{1 - 0.8z}{1 + 0.7z}.$$

To show that the simple rule for choosing  $M$ , i.e.,  $|t_{N-M}| \approx |t_{1-M}|$ , does provide a better spectral clustering property and a better convergence rate for  $S_N^{-1}T_N$ , two

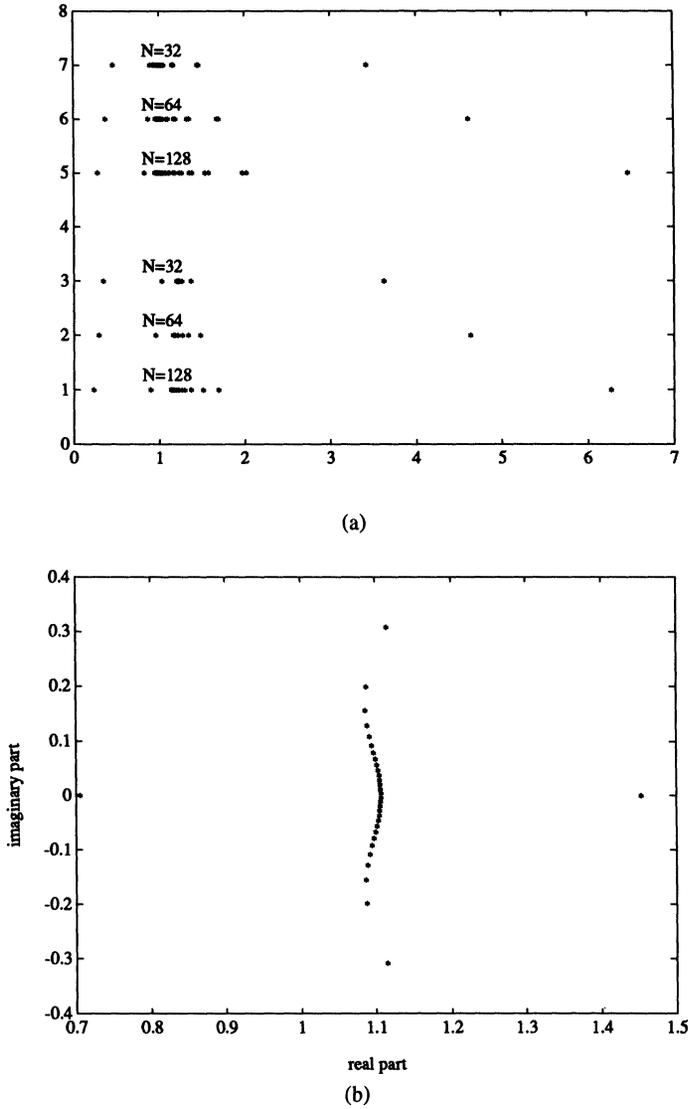


FIG. 1. (a) The singular value distribution of  $S_N^{-1}T_N$  and  $K_N^{-1}T_N$ , and (b) the eigenvalue distribution of  $K_N^{-1}T_N$  for Test Problem 1.

preconditioners  $S_N$  and  $\tilde{S}_N$  are constructed. The  $S_N$  is constructed with  $M$  such that  $|t_{N-M}| \approx |t_{1-M}|$  while the  $\tilde{S}_N$  is constructed with  $M = \lceil N/2 \rceil$ . The eigenvalues of  $T_N$ ,  $\tilde{S}_N^{-1}T_N$ ,  $S_N^{-1}T_N$  and  $K_N^{-1}T_N$  with  $N = 32$  are plotted in Figs. 2(a)–(d). All preconditioned Toeplitz matrices have eigenvalues clustered around 1 except  $2 = 2 \min(r, s)$  outliers. The  $K_N^{-1}T_N$  has the best clustering effect, and the eigenvalues of  $S_N^{-1}T_N$  are more closely clustered than those of  $\tilde{S}_N^{-1}T_N$ . The sums of magnitudes of the last elements in constructing  $S_N$  and  $K_N$  and the corresponding clustering radii are listed in Table 3. They are of approximately the same order, as stated in Theorems 3 and 4.

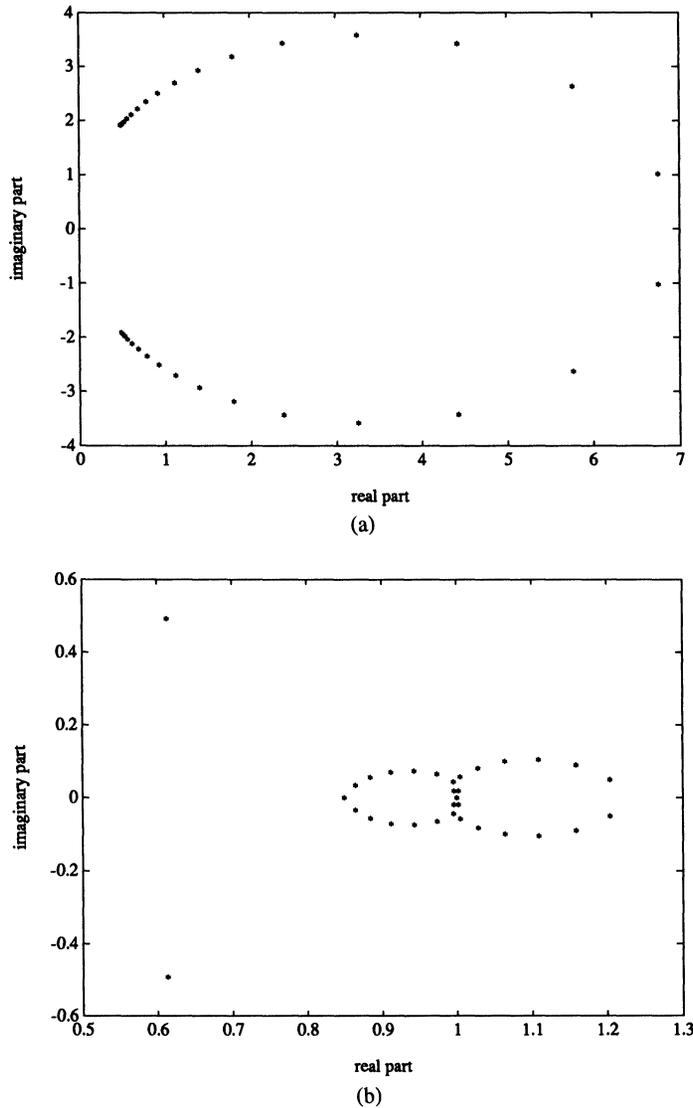


FIG. 2. The eigenvalue distribution of (a)  $T_N$  and (b)  $\tilde{S}_N^{-1}T_N$  for Test Problem 2.

The convergence history of the CGS method with various preconditioners is plotted in Fig. 3 with  $N = 32$ . The convergence rate of the CGS method without preconditioning (the curve denoted by  $T_N$ ) is very slow. This phenomenon is not surprising by examining the eigenvalue distribution given in Fig. 2(a). Preconditioning improves the convergence behavior dramatically. It is clear that  $K_N$  performs the best, while  $S_N$  performs better than  $\tilde{S}_N$ .

TEST PROBLEM 3. Rational  $T_N$  with  $(r, s) = (3, 1)$ . The generating function of

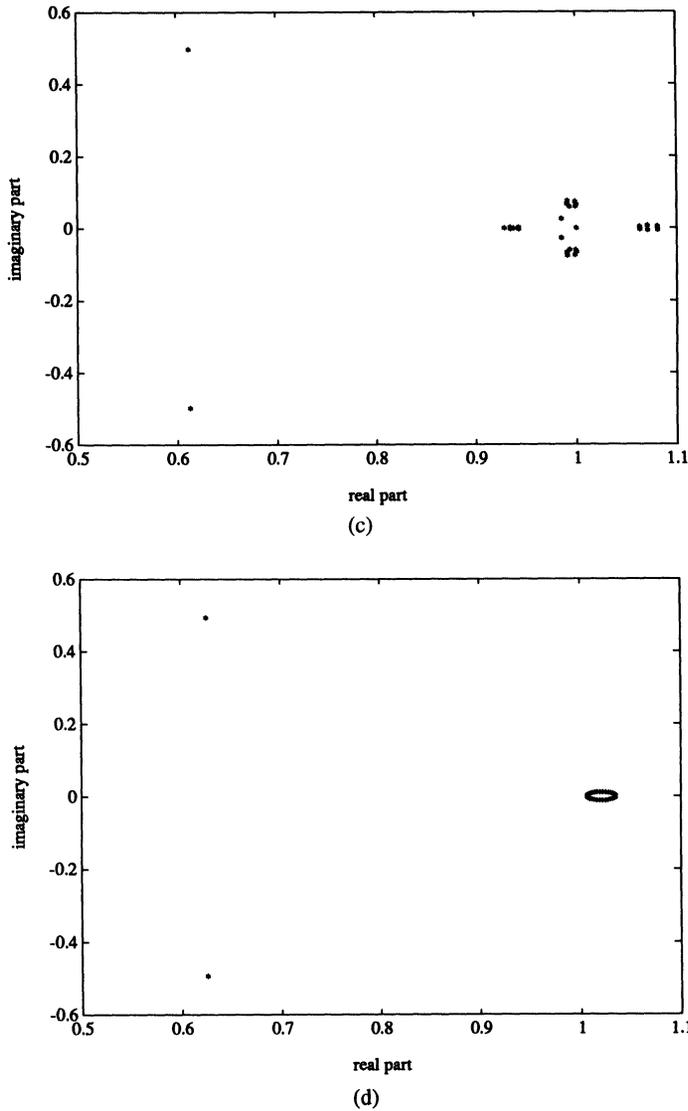


FIG. 2. The eigenvalue distribution of (c)  $S_N^{-1}T_N$  and (d)  $K_N^{-1}T_N$  for Test Problem 2.

$T_N$  is chosen to be

$$T(z) = \frac{(1 + 0.5z^{-1})(1 + 0.7z^{-1})}{(1 - 0.4z^{-1})(1 - 0.6z^{-1})(1 - 0.8z^{-1})} + \frac{1 + 0.8z}{1 + 0.9z}.$$

The eigenvalues of  $T_N$ ,  $S_N^{-1}T_N$ , and  $K_N^{-1}T_N$  with  $N = 64$  are plotted in Figs. 4(a)–(c). It is clear that  $K_N^{-1}T_N$  has  $2 = 2 \min(r, s)$  outliers. The outliers of  $S_N^{-1}T_N$  are not easy to identify for this case. However, two outliers can be observed more easily for larger  $N$ . Besides, the eigenvalues of  $K_N^{-1}T_N$  are more closely clustered than those of  $S_N^{-1}T_N$ . We list in Table 4 the sums of magnitudes of the last elements

TABLE 3  
Clustering radii  $\epsilon$  of preconditioners  $S_N$  and  $K_N$  for Test Problem 2.

$N$	$\epsilon_S$	$ t_{N-M}  +  t_{1-M} $	$\epsilon_K$	$ t_{N-1}  +  t_{1-N} $
32	$8.2 \times 10^{-2}$	$2.8 \times 10^{-1}$	$3.5 \times 10^{-2}$	$6.8 \times 10^{-2}$
64	$4.6 \times 10^{-2}$	$2.1 \times 10^{-2}$	$1.2 \times 10^{-3}$	$2.3 \times 10^{-3}$
128	$3.3 \times 10^{-5}$	$1.1 \times 10^{-4}$	$1.4 \times 10^{-6}$	$2.7 \times 10^{-6}$

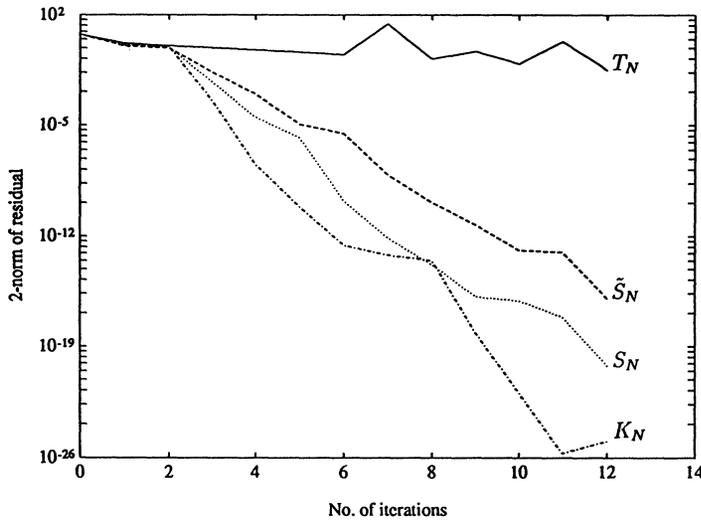


FIG. 3. Convergence history of the CGS method for Test Problem 2.

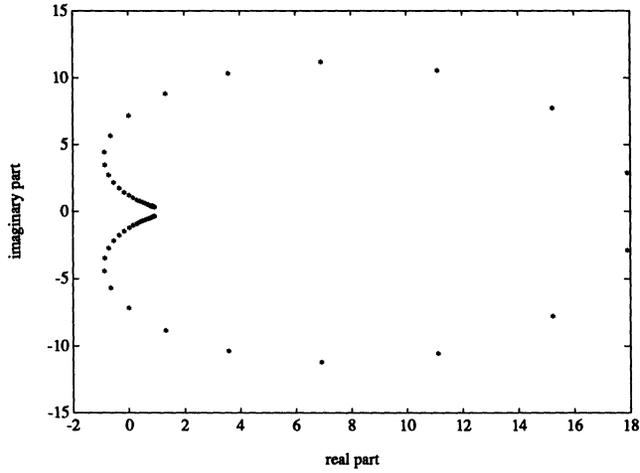
in constructing  $S_N$  and  $K_N$  and the corresponding clustering radii. The convergence history of the CGS method with  $N = 64$  is plotted in Fig. 5. We observe that the CGS method without preconditioning does not converge and that the CGS method with preconditioners  $K_N$  and  $S_N$  converges in 4 and 6 iterations, respectively. This seems to suggest that the use of preconditioners does not only accelerate the convergence rate by providing better spectral properties, but also improves the convergence of nonsymmetric iterative algorithms by making the preconditioned matrix more close to normal.

TEST PROBLEM 4. Rational triangular  $T_N$  with  $(r, s) = (1, 0)$ . The generating function of  $T_N$  is chosen to be

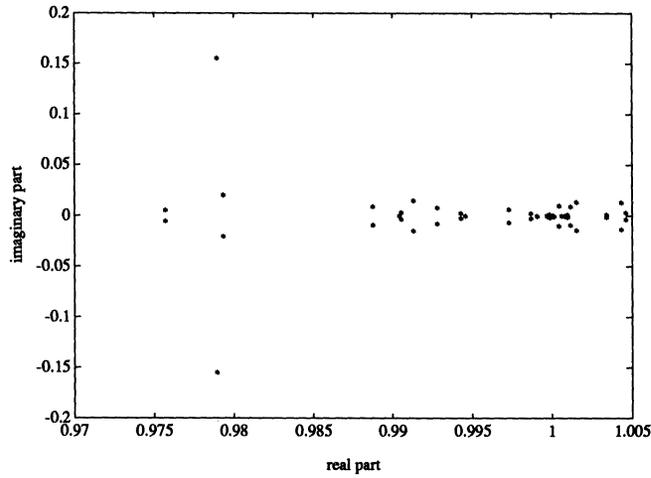
$$T(z) = \frac{1 - 0.7z^{-1}}{1 + 0.5z^{-1}}.$$

Since there are only  $N$  nonzero elements in  $T_N$ , we can make  $S_N$  the same as  $K_N$ . The eigenvalues of  $K_N^{-1}T_N$  with  $N = 32$  are plotted in Fig. 6(a). We see that all eigenvalues are clustered around 1 with radius  $\epsilon_K = O(|t_N|) = 10^{-9}$ . This is consistent with Theorem 3, which predicts that  $K_N^{-1}T_N$  has  $0 = 2 \min(r, s)$  outliers. The convergence history of the CGS method with  $N = 32$  is plotted in Fig. 6(b). The CGS method with preconditioner  $K_N$  converges in two iterations while the CGS method without preconditioning does not converge.

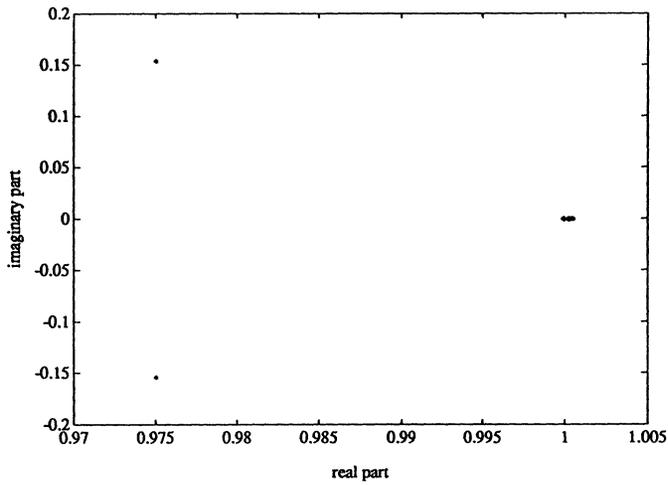
TEST PROBLEM 5. Ill-conditioned  $T_N$ . Let  $T_N$  be the  $N \times N$  Toeplitz matrix



(a)



(b)



(c)

FIG. 4. The eigenvalue distribution of (a)  $T_N$ , (b)  $S_N^{-1}T_N$ , and (c)  $K_N^{-1}T_N$  for Test Problem 3.

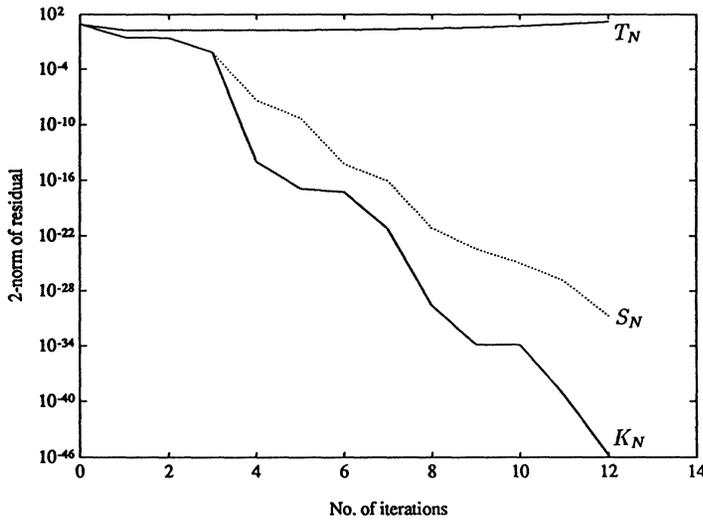


FIG. 5. Convergence history of the CGS method for Test Problem 3.

TABLE 4  
Clustering radii  $\epsilon$  of preconditioners  $S_N$  and  $K_N$  for Test Problem 3.

$N$	$\epsilon_S$	$ t_{N-M}  +  t_{1-M} $	$\epsilon_K$	$ t_{N-1}  +  t_{1-N} $
32	$1.7 \times 10^{-1}$	$1.4 \times 10^{-1}$	$6.1 \times 10^{-2}$	$2.8 \times 10^{-2}$
64	$2.7 \times 10^{-2}$	$1.3 \times 10^{-2}$	$5.1 \times 10^{-4}$	$1.6 \times 10^{-4}$
128	$1.7 \times 10^{-3}$	$1.6 \times 10^{-4}$	$5.8 \times 10^{-7}$	$1.7 \times 10^{-7}$

generated by polynomial  $T(z) = \epsilon + z^{-1}$  with  $\epsilon < 1$ , i.e.,

$$T_N = \begin{bmatrix} \epsilon & 0 & 0 & 0 & \cdot & 0 & 0 \\ 1 & \epsilon & 0 & 0 & \cdot & \cdot & 0 \\ 0 & 1 & \epsilon & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \epsilon & 0 & 0 \\ 0 & \cdot & \cdot & 0 & 1 & \epsilon & 0 \\ 0 & 0 & \cdot & 0 & 0 & 1 & \epsilon \end{bmatrix}.$$

Matrix  $T_N$  is nearly singular for large  $N$  and  $\|T_N^{-1}\|$  increase with  $N$  at the rate  $O(\epsilon^{-N})$ .

The preconditioner  $K_N (= S_N)$  is

$$K_N = \begin{bmatrix} \epsilon & 0 & 0 & 0 & \cdot & 0 & 1 \\ 1 & \epsilon & 0 & 0 & \cdot & \cdot & 0 \\ 0 & 1 & \epsilon & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \epsilon & 0 & 0 \\ 0 & \cdot & \cdot & 0 & 1 & \epsilon & 0 \\ 0 & 0 & \cdot & 0 & 0 & 1 & \epsilon \end{bmatrix},$$

and  $\|K_N^{-1}\|$  is bounded for large  $N$ . Since the rank of the matrix  $K_N - T_N$  is one,  $K_N^{-1}T_N$  has  $N - 1$  eigenvalues repeated at 1 and our theoretic results still hold.

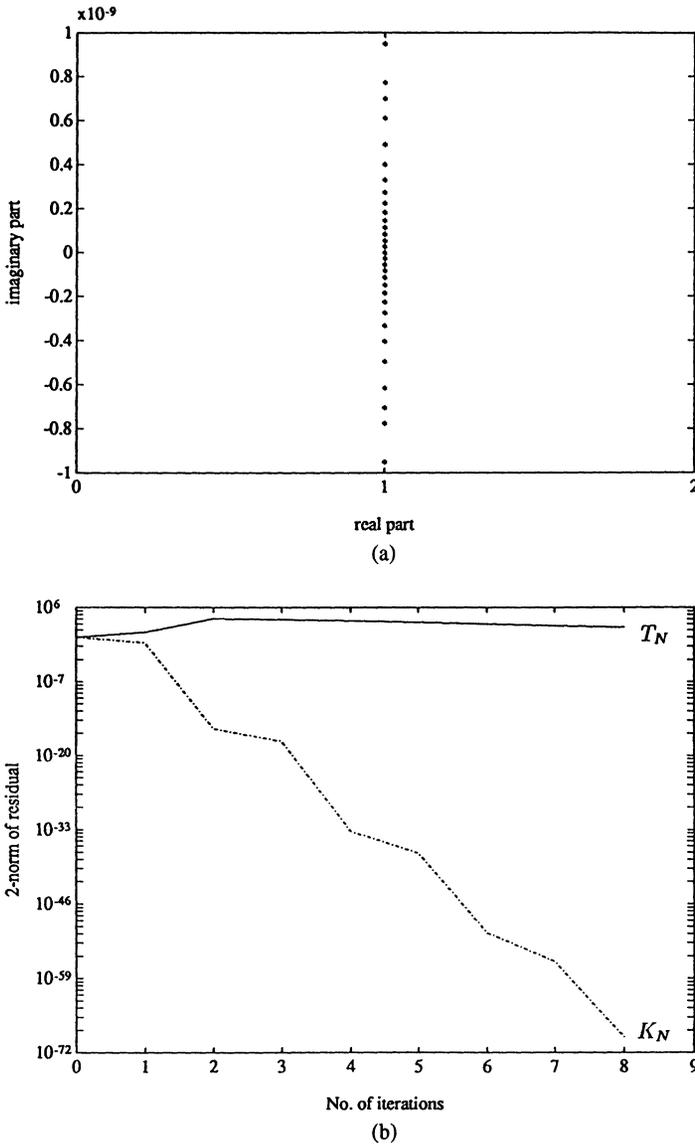


FIG. 6. (a) The eigenvalue distribution of  $K_N^{-1}T_N$  and (b) the convergence history of the CGS method for Test Problem 4.

However, preconditioning  $T_N$  by  $K_N$  cannot improve the ill-conditionedness of matrix  $T_N$ . The preconditioned matrix  $K_N^{-1}T_N$  is still nearly singular, and the solution is not accurate due to the ill-conditionedness.

For the extreme case  $\varepsilon = 0$ , matrix  $T_N$  is singular with eigenvalues repeated at 0, and  $T_N^{-1}$  does not exist. If  $\mathbf{b}$  is in the rank space of  $T_N$ ,  $T_N\mathbf{x} = \mathbf{b}$  has more than one

solution of the following form,

$$(6.1) \quad \mathbf{x} = \mathbf{x}^* + \mathbf{v},$$

where  $\mathbf{x}^*$  is in the row space of  $T_N$  with  $T_N \mathbf{x}^* = \mathbf{b}$ , and  $\mathbf{v}$  is in the null space of  $T_N$ . Note that  $K_N^{-1}$  exists even though  $T_N$  is singular. Since only the inversion of the preconditioner is required in each iteration of preconditioned iterative methods, the iteration can be performed without difficulty. It can be verified that  $K_N^{-1} T_N$  has one eigenvalue at 0 and all other  $N - 1$  eigenvalues repeated exactly at 1. As a consequence of the eigenvalue distribution, the CGS method converges to one solution in form (6.1) after one iteration.

**7. Conclusion.** In this paper, we generalized the circulant preconditioning technique from symmetric to nonsymmetric Toeplitz matrices. The resulting preconditioned Toeplitz systems are then solved by various iterative methods such as CGN and CGS. For a large class of Toeplitz matrices, we proved that the singular values of  $S_N^{-1} T_N$  and  $K_N^{-1} T_N$  are clustered around unity except for a fixed number independent of  $N$ . When the generating function is rational, the eigenvalues of  $K_N^{-1} T_N$  and  $S_N^{-1} T_N$  are classified into clustered eigenvalues and outliers. The number of outliers depends on the order of the rational generating function. The clustered eigenvalues are confined in the disk centered at 1 with the radii  $\epsilon_K = O(|t_N| + |t_{-N}|)$  and  $\epsilon_S = O(|t_{N-M}| + |t_{1-M}|)$  for  $K_N^{-1} T_N$  and  $S_N^{-1} T_N$ , respectively. Since the eigenvalues of  $K_N^{-1} T_N$  are more closely clustered than those of  $S_N^{-1} T_N$ , preconditioner  $K_N$  performs better than  $S_N$  for solving rational Toeplitz systems.

**Acknowledgments.** The authors thank Professor Robert J. Plemmons for helpful discussion. The authors are also grateful to the referees for valuable suggestions, including Test Problem 5 in §6.

#### REFERENCES

- [1] R. H. CHAN, *Circulant preconditioners for Hermitian Toeplitz systems*, SIAM J. Matrix Anal. Appl., 10 (1989), pp. 542–550.
- [2] R. H. CHAN AND T. F. CHAN, *Circulant preconditioners for elliptic problems*, Tech. Rep., Dept. of Mathematics, Univ. of California at Los Angeles, CA, Dec. 1990.
- [3] R. H. CHAN AND X. JIN, *A family of block preconditioners for block systems*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 1218–1235.
- [4] R. H. CHAN, X. JIN, AND M. YEUNG, *Circulant preconditioners for complex Toeplitz matrices*, Tech. Rep., Dept. of Mathematics, Univ. of Hong Kong, Apr. 1991.
- [5] R. H. CHAN AND G. STRANG, *Toeplitz equations by conjugate gradients with circulant preconditioner*, SIAM J. Sci. Statist. Comput., 10 (1989), pp. 104–119.
- [6] T. F. CHAN, *An optimal circulant preconditioner for Toeplitz systems*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 766–771.
- [7] M. R. HESTENES AND E. STIEFEL, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Standards, 49 (1952), pp. 409–436.
- [8] T. HUCKLE, *Circulant and skew-circulant matrices for solving Toeplitz matrices problems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 767–777.
- [9] T. K. KU AND C. J. KUO, *Design and analysis of Toeplitz preconditioners*, IEEE Trans. Signal Processing, 40 (1992), pp. 129–141.
- [10] ———, *On the spectrum of a family of preconditioned block Toeplitz matrices*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 948–966.
- [11] ———, *A minimum-phase LU factorization preconditioner for Toeplitz matrices*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 1470–1487.
- [12] ———, *Spectral properties of preconditioned rational Toeplitz matrices*, SIAM J. Matrix Anal. Appl., 14 (1993), to appear.
- [13] N. M. NACHTIGAL, S. C. REDDY, AND L. N. TREFETHEN, *How fast are nonsymmetric matrix iterations?*, SIAM J. Matrix Anal. Appl., 13 (1993), pp. 778–795.

- [14] J. G. NAGY AND R. J. PLEMMONS, *Some fast Toeplitz least squares algorithms*, in SPIE Internat. Symp. Optical Applied Sci. Engrg., San Diego, CA, 1991.
- [15] Y. SAAD AND M. H. SCHULTZ, *GMRES: A generalized minimum residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 856–869.
- [16] P. SONNEVELD, *CGS, a fast Lanczos-type solver for nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 10 (1989), pp. 36–52.
- [17] G. STRANG, *A proposal for Toeplitz matrix calculations*, Stud. Appl. Math., 74 (1986), pp. 171–176.
- [18] ———, *Linear Algebra and Its Applications*, 3rd ed., Harcourt Brace Jovanovich, Inc., Orlando, FL, 1988.
- [19] L. N. TREFETHEN, *Approximation theory and numerical linear algebra*, in Algorithms for Approximation II, M. Cox and J. C. Mason, eds., Chapman and Hall, London, 1988.
- [20] J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford, U.K., 1965.
- [21] A. YAMASAKI, *New preconditioners based on low-rank elimination*, Tech. Rep. Numerical Analysis 89-10, Dept. of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, Dec. 1989.