NORTH-HOLLAND

Circulant Preconditioners With Unbounded Inverses

E. E. Tyrtyshnikov

Institute of Numerical Mathematics
Russian Academy of Sciences
Leninskij Prospekt, 32-A
117334, Moscow, Russia

Submitted by Richard A. Brualdi

ABSTRACT

The eigenvalue and singular-value distributions for matrices $S_n^{-1}A_n$ and $C_n^{-1}A_n$ are examined, where A_n, S_n, and C_n are Toeplitz matrices, simple circulants, and optimal circulants generated by the Fourier expansion of some function f. Recently it has been proven that a cluster at 1 exists whenever f is from the Wiener class and strictly positive. Both restrictions are now weakened. A proof is given for the case when f may take the zero value, and hence the circulants are to have unbounded inverses. The main requirements on f are that it belong to L_2 and be in some sense, sparsely vanishing. Specifically, if f is nonnegative and circulants S_n (or C_n) are positive definite, then the eigenvalues of $S_n^{-1}A_n$ (or $C_n^{-1}A_n$) are clustered at 1. If f is complex-valued and S_n (or C_n) are nonsingular, then the singular values of $S_n^{-1}A_n$ (or $C_n^{-1}A_n$) are clustered at 1 as well. Also proposed and studied are the improved circulants. It is shown that (improved) simple circulants can be much more advantageous than optimal circulants. This depends crucially on the smoothness properties of f. Further, clustering-on theorems are given that pertain to multilevel Toeplitz matrices preconditioned by multilevel simple and optimal circulants.

1. INTRODUCTION

Let $A_n = [a_{i-j}]_{i,j=0}^{n}$ be a Hermitian positive definite Toeplitz matrix, and suppose that a linear algebraic system with A_n is given, and attacked with the method of conjugate gradients. Then some preconditioners of choice are the optimal (Cesaro) [7, 13, 14] circulant

$$C_n = \left[c_{i-j\,(\mathrm{mod}\, n+1)}\right]_{i,j=0}^{n},$$

LINEAR ALGEBRA AND ITS APPLICATIONS 216:1–23 (1995)

0024-3795/95/$9.50
SSDI 0024-3795(93)00092-E

where

$$c_0 = a_0, \qquad c_k = \frac{1}{n+1}\left[(n+1-k)a_k + ka_{-(n+1-k)}\right], \quad k = 1, \ldots, n, \quad (1.1)$$

and, alternatively, the simple (in fact, of Strang's type [12]) circulant

$$S_n = \left[s_{i-j\,(\mathrm{mod}\, n+1)}\right]_{i,j=0}^{n},$$

where

$$s_k = \begin{cases} a_k, & 0 \le k < n/2, \\ a_{-(n+1-k)}, & n/2 < k \le n, \\ 0, & k = n/2. \end{cases} \quad (1.2)$$

In this paper we are going to further substantiate the usefulness of these circulants. According to [2] the problem may be viewed as that of proving that the eigenvalues of $C_n^{-1}A_n$ or $S_n^{-1}A_n$ are condensed into a cluster.

We assume that matrices A_n originate from the Fourier expansion of

$$f(x) \sim \sum_{k=-\infty}^{\infty} a_k \exp(ikx), \qquad x \in \mathbf{R}. \quad (1.3)$$

If f belongs to the Wiener class, that is

$$\sum_{k=-\infty}^{\infty} |a_k| < +\infty, \quad (1.4)$$

and, in addition,

$$\inf_x f(x) \equiv \delta > 0, \quad (1.5)$$

then, as is shown in [3, 4], the eigenvalues of $C_n^{-1}A_n$ and $S_n^{-1}A_n$ are clustered at 1. In this case, for all n

$$\|C_n^{-1}\|_2 \le \delta^{-1}, \quad (1.6)$$

and for any $g > 1$, for all sufficiently large n

$$\|S_n^{-1}\|_2 \le g\delta^{-1}. \quad (1.7)$$

In the proof given in [3, 4], both (1.4) and (1.5) are essential. However, there are some cases of practical interest where $f(x)$ is nonnegative, but takes the zero value at some x. So (1.5) is no longer true. Consequently, even with S_n and C_n positive definite (as established in [13], C_n is positive definite whenever A_n is), the norms of S_n^{-1} and C_n^{-1} grow infinitely as n increases. Notwithstanding the lack of theoretical background, we often observe that the eigenvalues of $C_n^{-1}A_n$ and $S_n^{-1}A_n$ cluster at 1 even in this more general case.

For instance, this is so for Hermitian Toeplitz matrices

$$A_n = \left[\frac{-1}{(i-j)^2 - \frac{1}{4}}\right], \tag{1.8}$$

which are generated by

$$f(x) = 2\pi \left|\sin \frac{x}{2}\right|. \tag{1.9}$$

In this case (1.4) holds, whereas (1.5) does not, because $f(0) = 0$. Just the same, the convergence of the preconditioned conjugate-gradient method is rapid, and most of the eigenvalues of $C_n^{-1}A_n$ are amassed at 1:

Matrix order $(n+1)$	Number of iterations to reach relative error about 10^{-5}	Number of eigenvalues of $C_n^{-1}A_n$ which lie outside (0.9, 1.1)
64	8	6
128	9	6
256	10	9
512	10	9
1024	12	20

In this paper, we will furnish a proof that a cluster exists when $f(x)$ is allowed to take the zero value. In addition, the condition (1.4) will be greatly weakened, being replaced by the requirement that $f \in L_2$.

As a practical matter, our most important results have to do with twofold (doubly) Toeplitz matrices, threefold ones, and the like. These are the cases where iterative solvers meet no efficient direct method to vie with. Clustering-on theorems for multifold (multilevel) Toeplitz matrices are given in Section 5. In preceding sections, we take up the one-level case in such full detail that it enables us to treat the multilevel case rather in brief.

In Section 2, we present some preliminary results. After the notions of a general and proper cluster, we convey the definition of an asymptotic distribution, which is a generalization of a definition by H. Weyl (see [10]). One immediate advantage of the new definition is that it allows us to handle generating functions which belong to L_2 rather than L_∞. A few statements are given that are somewhat modified versions of those from our work [14], where one can find a full treatment.

In Section 3, the general clustering-on theorems concerning $S_n^{-1}A_n$ and $C_n^{-1}A_n$ are proposed. Instead of (1.5) we impose on f the demand that it should be "sparsely vanishing." Roughly speaking, this means that the set of those $x \in [-\pi, \pi]$ for which $f(x) = 0$ is not too large. Instead of (1.4) we require that $f \in L_2$. One more restriction we need is that S_n (or C_n) should be nonsingular. If so, then Theorem 3.3 states that the singular values of $S_n^{-1}A_n$ (or $C_n^{-1}A_n$) cluster at 1. If f is nonnegative, then A_n, S_n, and C_n are Hermitian, and Theorem 3.1 says that the eigenvalues of $S_n^{-1}A_n$ (or $C_n^{-1}A_n$) cluster at 1, provided that S_n (or C_n) is positive definite.

The restriction that S_n and C_n are nonsingular (positive definite) is not very significant, and can be easily obviated if we "improve" these circulants by changing their zero (nonpositive) eigenvalues, if any, so as to have them nonzero (positive). Such "improved" circulants, considered in Theorems 3.2 and 3.4, are always nonsingular (positive definite) and appear to maintain the clustering.

In Section 4, some neater estimates of the clustering are provided. Specifically, we examine $\gamma_n(\varepsilon)$, the number of eigenvalues of $S_n^{-1}A_n$ (or $C_n^{-1}A_n$) which lie beyond ε-distance from 1. One may anticipate that the growth of $\gamma_n(\varepsilon)$ as n increases should be crucially dependent on the smoothness of f. We discover that this is true only as far as simple circulants are concerned. For optimal (Cesaro) circulants, enhancing the smoothness has no bearing on that growth. This discrepancy between simple and optimal circulants is in good agreement with the different approximation properties of partial Fourier sums and Cesaro sums.

An interesting extreme case is when f is analytic. In this case, we show $\gamma_n(\varepsilon) = (O \log n)$. Moreover, if f is a trigonometric polynomial, then $\gamma_n(\varepsilon) = O(1)$ (which means that the cluster is proper). We would like to stress that these nice estimates are valid only for simple circulants (properly "improved," if needed). For example, if

$$f(x) = 2 - 2\cos x \tag{1.10}$$

then

$$A_n = \begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & & \mathbf{0} \\ & -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & & \\ \mathbf{0} & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix} \in \mathbf{R}^{(n+1)\times(n+1)}. \tag{1.11}$$

Here, $f(0) = f'(0) = 0$ and $f''(0) \neq 0$. Hence by Theorem 4.2 we have $\gamma_n(\varepsilon) = O(n^{2/3})$ when using optimal circulants. At the same time, for improved simple circulants $\gamma_n(\varepsilon) = O(1)$, which is much better. The theory is entirely confirmed numerically:

In Section 5, the clustering-on theorems obtained in Section 3 are generalized to multilevel Toeplitz matrices. In Section 6 we give concluding remarks.

Matrix order $(n+1)$	$\gamma_n(\varepsilon)$ where $\varepsilon = 0.1$ for optimal circulants	$\gamma_n(\varepsilon)$ where $\varepsilon = 0.1$ for improved simple circulants
64	12	3
128	16	4
256	26	3
512	40	2
1024	77	3

2. PRELIMINARY RESULTS

Let us be given a sequence of real numbers $\{\lambda_k^{(n)}\}_{k=1}^n$, and let μ be real and fixed. Denote by $\gamma_n(\varepsilon)$, $\varepsilon > 0$, the number of those $k \in \{1, \ldots, n\}$ for which $\lambda_k^{(n)} \notin (\mu - \varepsilon,\ \mu + \varepsilon)$. We will say that $\{\lambda_k^{(n)}\}$ has a cluster at μ if $\gamma_n(\varepsilon) = o(n)$ [as usual, $o(n)$ designates a function of n such that $o(n)/n \to 0$ as $n \to \infty$]. A cluster is called proper if $\gamma_n(\varepsilon) \leq c(\varepsilon)$, where $c(\varepsilon)$ does not depend on n.

The notion of clusters is bound up with the notion of an asymptotic distribution for $\{\lambda_n^{(n)}\}$. Let f be a real-valued 2π-periodic Lebesgue-integrable function. We will say that $\{\lambda_k^{(n)}\}$ is distributed as $f(x)$ if for any continuous function $F(x)$ with bounded support,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} F(\lambda_k^{(n)}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F(f(x))\, dx. \tag{2.1}$$

This definition, which is a generalization of that given by H. Weyl, is proposed in [14]. Obviously, if $f(x) \equiv \mu$, then there is a cluster at μ.

LEMMA 2.1. *Given a sequence $\{A_n\}$ of Hermitian matrices $A_n \in \mathbf{C}^{n\times n}$, suppose that there are Hermitian matrices Δ_n such that*

$$\begin{aligned} \|A_n - I_n + \Delta_n\|_F^2 &= o(n), & (2.2)\\ \operatorname{rank} \Delta_n &= o(n). & (2.3) \end{aligned}$$

Then the eigenvalues of A_n are clustered at 1.

PROOF. Denote by $\lambda_i(A_n + \Delta_n)$ the eigenvalues of the Hermitian matrix $A_n + \Delta_n$, taken in nondecreasing order. Then, by the Hoffman-Wielandt theorem (see [16, pp. 104–108]), due to (2.2),

$$\sum_{k=1}^{n} |\lambda_k(A_n + \Delta_n) - 1|^2 = o(n). \tag{2.4}$$

Let $\gamma_n(\varepsilon; A_n + \Delta_n)$ be the number of $k \in \{1, \dots, n\}$ such that $\lambda_k(A_n + \Delta_n) \notin (1 - \varepsilon,\ 1 + \varepsilon)$. It readily follows from (2.4) that

$$\gamma_n(\varepsilon;\ A_n + \Delta_n) = o(n). \tag{2.5}$$

Now, remember the interlacing property which holds when a Hermitian matrix is perturbed by a Hermitian rank-1 matrix ([16, pp. 94–97]; see also [9, p. 412]). Then we immediately conclude that

$$\gamma_n(\varepsilon;\ A_n) \le \gamma_n(\varepsilon;\ A_n + \Delta_n) + \operatorname{rank} \Delta_n = O(n). \tag{2.6}$$

The proof is thus completed. ■

Note that this lemma is a consequence (in fact, a modified representation) of the results from Section 2 of our work [14]. We will also make use of the following

LEMMA 2.2. *If A_n, $\Delta_n \in \mathbf{C}^{n \times n}$ satisfy* (2.2) *and* (2.3), *then A_n's singular values have a cluster at* 1.

To this end, it is sufficient to construct Hermitian matrices

$$\widetilde{A}_n = \begin{bmatrix} 0 & A_n \\ A_n^* & 0 \end{bmatrix}, \qquad \widetilde{\Delta}_n = \begin{bmatrix} 0 & \Delta_n \\ \Delta_n^* & 0 \end{bmatrix}$$

and take into account that $\widetilde{A}_n$'s eigenvalues equal $\pm\sigma_k(A_n)$, where σ_k are the singular values of A_n. Applying Lemma 2.1 to $\widetilde{A}_n$ and $\widetilde{\Delta}_n$, we obtain the result.

LEMMA 2.3. *Suppose (complex) Teoplitz matrices A_n are generated by the Fourier extension of a (complex-valued) function $f \in L_2$, and optimal circulants C_n and simple circulants S_n are defined by* (1.1) *and* (1.2). *Then*

$$\|A_n - C_n\|_F^2 = o(n), \tag{2.7}$$

$$\|A_n - S_n\|_F^2 = o(n), \tag{2.8}$$

LEMMA 2.4. *Let $f \in L_2$. Then the singular values of C_n and S_n are distributed as $|f(x)|$. If f is real-valued, then C_n and S_n are Hermitian, and their eigenvalues are distributed as $f(x)$.*

REMARK. As far as C_n alone is concerned, all the statements stand even if $f \in L_1$. For proofs see [14].

3. CLUSTERING-ON THEOREMS

The results here will rest on the notion of a sparsely vanishing function. Let $f(x)$ be a 2π-periodic Lebesgue-integrable function such that

$$\lim_{\varepsilon \to +0} \int_{-\pi}^{\pi} \varphi_\varepsilon(|f(x)|)\, dx = 0, \tag{3.1}$$

where φ_ε is the characteristic function for the interval $[0, \varepsilon]$, i.e., $\varphi_\varepsilon(x) = 1$ if $x \in [0, \varepsilon]$, and $\varphi_\varepsilon(x) = 0$ otherwise. Such a function f will be called sparsely vanishing.

THEOREM 3.1. *Let f be a nonnegative sparsely vanishing function from L_2, and A_n, C_n, S_n be Toeplitz, optimal, and simple circulant matrices, respectively, generated by f. Suppose in addition that the matrices C_n (or S_n) are positive definite. Then eigenvalues of $C_n^{-1}A_n$ (or $S_n^{-1}A_n$) are real, and have a cluster at* 1.

PROOF. Take $\varepsilon > 0$, and define the continuous function F_ε as follows:

$$F_\varepsilon(x) = \begin{cases} 0, & x \le -\varepsilon, \\ 1 + x/\varepsilon, & -\varepsilon \le x \le 0, \\ 1, & 0 \le x \le \varepsilon, \\ 2 - x/\varepsilon, & \varepsilon \le x \le 2\varepsilon, \\ 0, & 2\varepsilon \le x. \end{cases} \tag{3.2}$$

Denote by $\beta_n(\varepsilon)$ the number of those eigenvalues of C_n which fall inside the interval $[0, \varepsilon]$. Then, by Lemma 2.4, for n sufficiently large we obtain

$$\begin{aligned} \frac{\beta_n(\varepsilon)}{n+1} &= \frac{1}{n+1} \sum_{k=0}^{n} \varphi_\varepsilon\left(\lambda_k(C_n)\right) \\ &\le \frac{1}{n+1} \sum_{k=0}^{n} F_\varepsilon\left(\lambda_k(C_n)\right) \\ &\le \frac{1}{2\pi} \int_{-\pi}^{\pi} F_\varepsilon(f(x))\, dx + \varepsilon \\ &\le \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi_{2\varepsilon}(f(x))\, dx + \varepsilon \equiv \alpha(\varepsilon). \end{aligned} \tag{3.3}$$

From the well-known properties of circulants, we can write

$$C_n = Q_n^* \operatorname{diag}(\lambda_0, \ldots, \lambda_n)\, Q_n, \tag{3.4}$$

where Q_n is unitary and shared by all circulants of the same order. Starting from (3.4), we set

$$C_{n;\varepsilon} = Q_n^* \operatorname{diag}\left(\lambda_{0;\varepsilon}, \ldots, \lambda_{n;\varepsilon}\right) Q_n, \tag{3.5}$$

where

$$\lambda_{k;\varepsilon} = \begin{cases} \lambda_k & \text{if} \quad |\lambda_k| > \varepsilon, \\ \varepsilon & \text{if} \quad |\lambda_k| \le \varepsilon, \end{cases} \tag{3.6}$$

According to the theory of circulants, $C_{n;\varepsilon}$ is a circulant. Further, it is easy to verify that

$$C_{n;\varepsilon}^{-1/2}(A_n - C_n)\, C_{n;\varepsilon}^{-1/2} = C_n^{-1/2} A_n C_n^{-1/2} - I + \Delta_{n;\varepsilon}, \tag{3.7}$$

where

$$\begin{aligned} \Delta_{n;\varepsilon} \;=\; & C_{n;\varepsilon}^{-1/2}\big(C_{n;\varepsilon} - C_n\big) C_{n;\varepsilon}^{-1/2} + \big(C_{n;\varepsilon}^{-1/2} - C_n^{-1/2}\big) A_n C_{n;\varepsilon}^{-1/2} \\ & + C_n^{-1/2} A_n \big(C_{n;\varepsilon}^{-1/2} - C_n^{-1/2}\big). \end{aligned} \tag{3.8}$$

By virtue of (3.3)–(3.6) we thus find

$$\operatorname{rank} \Delta_{n;\varepsilon} \le 3\beta_n(\varepsilon) \le 3\alpha(\varepsilon)(n+1). \tag{3.9}$$

At the same time, according to (3.5), (3.6),

$$\|C_{n;\varepsilon}^{-1/2}\|_2 \le \varepsilon^{-1/2},$$

and hence, allowing for (3.7) and Lemma 2.3, we obtain

$$\|C_n^{-1/2} A_n C_n^{-1/2} - I + \Delta_{n;\varepsilon}\|_F^2 \le \varepsilon^{-2} \|A_n - C_n\|_F^2 = \varepsilon^{-2} o(n). \tag{3.10}$$

Therefore, there exists $N(\varepsilon)$ such that for $n \ge N(\varepsilon)$

$$\frac{1}{n+1} \|C_n^{-1/2} A_n C_n^{-1/2} - I + \Delta_{n;\varepsilon}\|_F^2 \le \varepsilon, \tag{3.11}$$

with

$$\frac{1}{n+1} \operatorname{rank} \Delta_{n;\varepsilon} \le 3\alpha(\varepsilon). \tag{3.12}$$

Since f is assumed to be sparsely vanishing, $\alpha(\varepsilon) \to 0$ as $\varepsilon \to 0$. Hence, the relationships (3.11), (3.12) mean that the sequence of Hermitian matrices $\{C_n^{-1/2} A_n C_n^{-1/2}\}$ satisfies all the hypotheses of Lemma 2.1. So the eigenvalues of $C_n^{-1/2} A_n C_n^{-1/2}$, which coincide with those of $C_n^{-1} A_n$, are clustered at 1. The above reasoning is still available if C_n is replaced by S_n. The proof is thus completed. ■

REMARK. If we require A_n to be positive definite, then this will entail the positive definiteness of C_n [13].

Usually, the main restriction on f—that it is sparsely vanishing—is easy to check. The function (1.9) from the introduction is evidently such. Since Toeplitz matrices (1.8) are diagonally dominant, they are positive definite. Using the above remark, we see that in this case all hypotheses of Theorem 3.1 are fulfilled.

The other restriction—that circulants C_n (or S_n) are positive definite—can be, in some sense, abandoned, if we replace C_n or S_n by other close by rank circulants which are positve definite. Specifically, if $C_n = Q_n^* \operatorname{diag}(\lambda_0, \dots, \lambda_n) Q_n$ as in (3.4), we put

$$\widehat{C}_n = Q_n^* \operatorname{diag}(\widehat{\lambda}_0, \dots, \widehat{\lambda}_n) Q_n, \tag{3.13}$$

where

$$\widehat{\lambda}_k = \begin{cases} \lambda_k & \text{if} \quad \lambda_k \neq 0, \\ \delta & \text{otherwise;} \end{cases} \tag{3.14}$$

here δ is an arbitrary positive number. Similarly, if

$$S_n = Q_n^* \operatorname{diag}(\mu_0, \dots, \mu_n) Q_n, \tag{3.15}$$

then we set

$$\widehat{S}_n = Q_n^* \operatorname{diag}(\widehat{\mu}_0, \dots, \widehat{\mu}_n) Q_n, \tag{3.16}$$

where

$$\widehat{\mu}_k = \begin{cases} \mu_k & \text{if} \quad \mu_k \neq 0, \\ \delta & \text{otherwise.} \end{cases} \tag{3.17}$$

In addition to $\widehat{S}_n$, we also define

$$\widetilde{S}_n = Q_n^* \operatorname{diag}(\widetilde{\mu}_0, \dots, \widetilde{\mu}_n) Q_n, \tag{3.18}$$

where

$$\widetilde{\mu}_k = \begin{cases} \mu_k & \text{if } \mu_k > 0, \\ \delta & \text{otherwise.} \end{cases} \tag{3.19}$$

With these improved circulants we state the following

THEOREM 3.2. *Let $f \in L_2$ be a nonnegative sparsely vanishing function, A_n be Hermitian Toeplitz matrices generated by f, and $\widehat{C}_n$ and $\widetilde{S}_n$ be improved circulants defined by* (3.13), (3.14) *and* (3.18), (3.19). *Then the matrices $\widehat{C}_n$ and $\widetilde{S}_n$ are Hermitian positive definite, and the eigenvalues of $\widehat{C}_n^{-1} A_n$ and $\widetilde{S}_n^{-1} A_n$ are clustered at* 1.

PROOF. The eigenvalues of C_n are shown in [14] to equal

$$\lambda_k(C_n) = \sigma_n \left(\frac{2\pi k}{n+1} \right), \qquad k = 0, 1, \ldots, n, \tag{3.20}$$

where $\sigma_n(x)$ is the Cesaro sum, possessing the Fejer representation

$$\sigma_n(x) = \int_{-\pi}^{\pi} K_n(x, t) f(t)\, dt, \tag{3.21}$$

where

$$K_n(x, t) = \frac{1}{2\pi(n+1)} \frac{\sin^2 (n+1)(x-t)/2}{\sin^2 (x-t)/2}, \tag{3.22}$$

and, as is well known,

$$\int_{-\pi}^{\pi} K_n(x, t)\, dt = 1. \tag{3.23}$$

We thus see that if $f(t) \geq 0$ then $\sigma_n(x) \geq 0$, and hence $\lambda_k(C_n) \geq 0$ for all k. It follows that the matrices $\widehat{C}_n$ are positive definite. The matrices $\widetilde{S}_n$ are positive definite by their very definition.

The remainder of the proof is quite analogous to the proof of the Theorem 3.1. Instead of (3.7) and (3.8) we write

$$C_{n;\varepsilon}^{-1/2} (A_n - C_n) C_{n;\varepsilon}^{-1/2} = \widehat{C}_n^{-1/2} A_n \widehat{C}_n^{-1/2} - I + \widehat{\Delta}_{n;\varepsilon}, \tag{3.24}$$

where

$$\begin{aligned} \widehat{\Delta}_{n;\varepsilon} &= C_{n;\varepsilon}^{-1/2} (C_{n;\varepsilon} - C_n) C_{n;\varepsilon}^{-1/2} + (C_{n;\varepsilon}^{-1/2} - \widehat{C}_n^{-1/2}) A_n C_{n;\varepsilon}^{-1/2} \\ &\quad + \widehat{C}_n^{-1/2} A_n (C_{n;\varepsilon}^{-1/2} - \widehat{C}_n^{-1/2}). \end{aligned} \tag{3.25}$$

This time it can be guaranteed that

$$\operatorname{rank} \widehat{\Delta}_{n;\varepsilon} \leq 5\beta_n(\varepsilon) \leq 5\alpha(\varepsilon)(n+1), \tag{3.26}$$

and we have as well

$$\| \widehat{C}_n^{-1/2} A_n \widehat{C}_n^{-1/2} - I + \widehat{\Delta}_{n;\varepsilon} \|_F^2 \leq \varepsilon^{-2} o(n). \tag{3.27}$$

Again, the sequence $\widehat{C}_n^{-1/2} A_n \widehat{C}_n^{-1/2}$ satisfies the hypotheses of Lemma 2.1, and we thus achieve the desired result. Everything stands if $\widetilde{S}_n$ replaces $\widehat{C}_n$. ■

Thus, while use of C_n and S_n is accompanied by some hazard (C_n and S_n may turn out to be singular), the improved circulants $\widehat{C}_n$ and $\widetilde{S}_n$ are always positive definite. Nothing ever prevents us from using the improved circulants. It is worth noting that getting improved circulants needs no additional work, because circulants are usually diagonalized via the FFT prior to further iterative calculations (see [13, 15]).

Now consider a more general situation, when $f(x)$ may change sign, and perhaps be complex-valued. The previous circulants can be still regarded as preconditioners, the only difference being that the conjugate gradient method should be used with some symmetrization technique, e.g. the transition to $(C_n^{-1}A_n)^*(C_n^{-1}A_n)$ or $(\widehat{C}_n^{-1}A_n)^*(\widehat{C}_n^{-1}A_n)$. Consequently, the effect of such preconditioning depends on the distribution of singular values of $C_n^{-1}A_n$ or $\widehat{C}_n^{-1}A_n$. We explore this in the following theorems.

THEOREM 3.3. *Suppose (complex-valued) $f \in L_2$ is sparsely vanishing, and let A_n and C_n (or S_n) be Toeplitz matrices and optimal (or simple) circulants allied with f. If C_n (or S_n) are nonsingular, then the singular values of $C_n^{-1}A_n$ (or $S_n^{-1}A_n$) are clustered at* 1.

THEOREM 3.4. *Let $f \in L_2$ be sparsely vanishing. Then the improved optimal and simple circulants, $\widehat{C}_n$ and $\widehat{S}_n$, are nonsingular, and the singular values of $\widehat{C}_n$ and $\widehat{S}_n$ have a cluster at* 1.

The proofs of these theorems are almost identical with those of Theorems 3.1 and 3.2, differing only in a final reference to Lemma 2.2 instead of Lemma 2.1. In fact, the proofs turn out to be simpler, because we need no similarity transition to Hermitian matrices.

4. SOME ESTIMATES

More precise information about clusters can be gained by specifying $\gamma_n(\varepsilon)$ introduced at the beginning of Section 2. Here we have a look at the behavior of $\gamma_n(\varepsilon)$ when ε is fixed and n is growing.

Suppose a 2π-periodic function $f(x)$ is such that its mth derivative $f^{(m)}(x)$ is piecewise continuous and has a bounded derivative on each continuity interval. Let K_m denote the set of all such functions.

It is an easy matter to verify that if $f \in K_m$ then

$$a_k = O\left(\frac{1}{k^{m+1}}\right). \tag{4.1}$$

We thence infer that the nth Fourier sums

$$f_n(x) = \sum_{k=-n}^{n} a_k e^{ikx} \tag{4.2}$$

approximate f, so that

$$\max_{-\pi \leq x \leq \pi} |f(x) - f_n(x)| = O\left(\frac{1}{n^m}\right), \tag{4.3}$$

while the nth Cesaro sums

$$\sigma_n(x) = \sum_{k=-n}^{n} \left(1 - \frac{|k|}{n+1}\right) a_k e^{ikx} \tag{4.4}$$

satisfy

$$\max_{-\pi \leq x \leq \pi} |f(x) - \sigma_n(x)| = \begin{cases} O\left(\frac{1}{n}\right), & m \geq 2, \\ O\left(\frac{\ln n}{n}\right), & m = 1. \end{cases} \tag{4.5}$$

Further, suppose that there are a finite number of points $x_j \in [-\pi, \pi]$, $j = 1, \ldots, t$, such that

$$f(x_j) = 0, \qquad j = 1, \ldots, t. \tag{4.6}$$

Let $f^{(p)}(x+0)$ and $f^{(p)}(x-0)$ denote the pth right- and left-hand derivatives of f at x, respectively. Assume that for every x_j, there exists p_j^+ and p_j^- such that

$$\begin{aligned} f^{(1)}(x_j+0) = \cdots = f^{(p_j^+-1)}(x_j+0) = 0, &\qquad f^{(p_j^+)}(x_j+0) \neq 0, \\ f^{(1)}(x_j-0) = \cdots = f^{(p_j^--1)}(x_j-0) = 0, &\qquad f^{(p_j^-)}(x_j-0) \neq 0, \end{aligned} \tag{4.7}$$

and set

$$p = \max\{p_j^{\pm} : j = 1, \ldots, t\}. \tag{4.8}$$

Let $K_m^{(p)}$ signify the set of those $f \in K_m$ which are characterized by relationships (4.6)–(4.8) where p is fixed, and t is arbitrary but finite.

We shall need the following auxiliary lemma.

LEMMA 4.1. *Suppose the points $y_k \in \mathbf{R}$ form a uniform mesh with the step size h, and let M be a finite union of nonintersecting intervals with total length d. Let* $\operatorname{in} M$ *denote the number of indices k such that $y_k \in M$. Then*

$$(\operatorname{in} M)h \leq d + ch,$$

where c depends on the number of intervals but not on h.

The proof is evident.

THEOREM 4.1. *Suppose $f \in K_m^{(p)}$ is nonnegative, and A_n and S_n are Toeplitz matrices and simple circulants generated by f. Assume that the matrices S_n are positive definite. Then the eigenvalues of $S_n^{-1}A_n$ are real, and clustered at* 1 *so that*

$$\gamma_n(\varepsilon) = O\big(n^{p/(p+m)}\big). \tag{4.9}$$

PROOF. Take arbitrary but sufficiently small $\delta > 0$, and set

$$M(\delta) \equiv \{x \in [-\pi, \pi] \ : \ |f(x)| \leq \delta\} \tag{4.10}$$

Using the theorem's hypotheses, we can deduce that $M(\delta)$ is embedded in some $M'(\delta)$ which is a union of t nonintersecting intervals with total length

$$d(\delta) = O(\delta^{1/p}), \tag{4.11}$$

provided that δ is sufficiently small.

The eigenvalues of S_n are expressed as (see [14])

$$\lambda_k(S_n) = f_{[n/2]}\left(\frac{2\pi k}{n+1}\right), \qquad k = 0, 1, \dots, n. \tag{4.12}$$

Denote by $\beta_n(\delta)$ the number of those $k \in \{0, 1, \dots, n\}$ such that $|\lambda_k(S_n)| \leq \delta$. Let $\varphi_\delta(x) = 1$ if $x \in [0, \delta]$, and $\varphi_\delta(x) = 0$ elsewhere. Next, let $\operatorname{in} M$ be the number of those $k \in \{0, 1, \dots, n\}$ for which $2\pi k/(n+1) \in M$. Then by (4.3), (4.11) and on the basis of Lemma 4.1 we find

$$\begin{aligned} \frac{\beta_n(\delta)}{n+1} &\leq \frac{1}{n+1}\sum_{k=0}^{n} \varphi_\delta\left(f_{[n/2]}\left(\frac{2\pi k}{n+1}\right)\right) \\ &\leq \frac{1}{n+1}\sum_{k=0}^{n} \varphi_{\delta+O(n^{-m})}\left(f\left(\frac{2\pi k}{n+1}\right)\right) \end{aligned}$$

$$
\begin{aligned}
&= \frac{\operatorname{in} M(\delta + O(n^{-m}))}{n+1} \leq \frac{\operatorname{in} M'(\delta + O(n^{-m}))}{n+1} \\
&= O((\delta + O(n^{-m}))^{1/p}) + O\left(\frac{1}{n}\right).
\end{aligned}
$$

Therefore,

$$
\beta_n(\delta) \leq c_1 \delta^{1/p} n + c_2 n^{1-m/p} + c_3, \tag{4.13}
$$

where $c_1,\ c_2,\ c_3 > 0$ are independent of n and δ.

As in [3, 4], we bring in matrices

$$
\Delta_{n;N} = \begin{bmatrix}
 & & a_{[N]} & \cdots & a_1 \\
 & & & \ddots & \vdots \\
a_{[N]} & & 0 & & a_{[N]} \\
\vdots & \ddots & & & \\
a_{-1} & \cdots & a_{-[N]} & &
\end{bmatrix} \in \mathbf{C}^{(n+1)\times(n+1)}, \tag{4.14}
$$

and allowing for (1.2) and (4.1) obtain

$$
\|A_n - S_n + \Delta_{n;N}\|_\infty \leq \frac{c_0}{N^m}, \tag{4.15}
$$

where $c_0 > 0$ is independent of n and N. As previously, we consider the unitary diagonalization of S_n:

$$
S_n = Q_n^* \operatorname{diag}(\lambda_k(S_n))\, Q_n, \tag{4.16}
$$

and set

$$
S_{n;\,\delta} = Q_n^* \operatorname{diag}\left(\lambda_{k;\,\delta}\right) Q_n, \tag{4.17}
$$

where

$$
\lambda_{k;\,\delta} = \begin{cases} \lambda_k(S_n) & \text{if} \quad \lambda_k(S_n) > \delta, \\ \delta & \text{otherwise.} \end{cases} \tag{4.18}
$$

The theory of circulants tells that $S_{n;\,\delta}$ is a circulant, and because of (4.18)

$$
\left\|S_{n;\,\delta}^{-1/2}\right\|_\infty \leq \delta^{-1/2}. \tag{4.19}
$$

Being interested in the eigenvalues of $S_n^{-1}A_n$, which coincide with those of the Hermitian matrix $S_n^{-1/2}A_nS_n^{-1/2}$, we proceed as follows:

$$
S_{n;\,\delta}^{-1/2}\left(A_n - S_n + \Delta_{n;N}\right) S_{n;\,\delta}^{-1/2} = S_n^{-1/2}(A_n - S_n)\, S_n^{-1/2} + \Delta_{n;N}^{(\delta)}, \tag{4.20}
$$

where

$$\begin{aligned}\Delta_{n;N}^{(\delta)} &= (S_{n;\delta}^{-1/2} - S_n^{-1/2})(A_n - S_n)S_n^{-1/2} \\ &\quad + S_{n;\delta}^{-1/2}(A_n - S_n)(S_{n;\delta}^{-1/2} - S_n^{-1/2}) \\ &\quad + S_{n;\delta}^{-1/2}\Delta_{n;N}S_{n;\delta}^{-1/2},\end{aligned} \tag{4.21}$$

and the relationships (4.13), (4.14), and (4.16)–(4.18) imply

$$\operatorname{rank}\Delta_{n;N}^{(\delta)} \le 2(c_1\delta^{1/p}n + c_2 n^{1-m/p} + c_3 + N). \tag{4.22}$$

At the same time, by virtue of (4.15) and (4.19),

$$\left\|S_{n;\delta}^{-1/2}\left(A_n - S_n + \Delta_{n;N}\right)S_{n;\delta}^{-1/2}\right\|_\infty \le \frac{c_0}{\delta N^m}. \tag{4.23}$$

Let us take $\varepsilon > 0$ and set $c_0/\delta N^m = \varepsilon$, that is,

$$N = \left(\frac{c_0}{\delta\varepsilon}\right)^{1/m}. \tag{4.24}$$

Then all the eigenvalues of the matrix in the left-hand side of (4.20) are not greater in modulus than ε, and the number of those eigenvalues of $S_n^{-1/2}A_nS_n^{-1/2} - I$ which are greater in modulus than ε does not exceed

$$\operatorname{rank}\Delta_{n;N}^{(\delta)} \le 2\left[c_1\delta^{\frac{1}{p}}n + c_2 n^{1-m/p} + c_3 + \left(\frac{c_0}{\delta\varepsilon}\right)^{1/m}\right]. \tag{4.25}$$

Let us pick

$$\delta = n^{-pm/(p+m)}. \tag{4.26}$$

Then, because $n^{1-m/p} \le n^{p/(p+m)}$, we have

$$\operatorname{rank}\Delta_{n;N}^{(\delta)} = O(n^{p/(p+m)}), \tag{4.27}$$

and that completes the proof. ■

THEOREM 4.2. *Suppose $f \in K_m^{(p)}$ is nonnegative, and A_n and C_n are Toeplitz matrices and optimal circulants generated by f. Assume that the matrices C_n are positive definite. Then the eigenvalues of $C_n^{-1}A_n$ are real, and clustered at 1 so that*

$$\gamma_n(\varepsilon) = \begin{cases} O(n^{p/(p+1)}), & m > 1, \\ O(n^{p/(p+1)}\ \ln n), & m = 1. \end{cases} \tag{4.28}$$

PROOF. We will use the notation from the preceding proof, throughout replacing S_n by C_n. Thus, $\beta_n(\delta)$ now means the number of $k \in \{0, 1, \dots, n\}$ such that $|\lambda_k(C_n)| \le \delta$. Since the eigenvalues of C_n, by (3.20), are the values of $\sigma_n(x)$, and (4.5) holds, we have, for $m > 1$,

$$\begin{aligned}
\frac{\beta_n(\delta)}{n+1} &\le \frac{1}{n+1}\sum_{k=0}^{n} \varphi_\delta\left(\sigma_n\left(\frac{2\pi k}{n+1}\right)\right) \\
&\le \frac{1}{n+1}\sum_{k=0}^{n} \varphi_{\delta+O(n^{-1})}\left(f\left(\frac{2\pi k}{n+1}\right)\right) \\
&= \frac{\operatorname{in} M(\delta + O(n^{-1}))}{n+1} \le \frac{\operatorname{in} M'(\delta + O(n^{-1}))}{n+1} \\
&= O\left(\left[\delta + O\left(n^{-1}\right)\right]^{1/p}\right) + O\left(\frac{1}{n}\right),
\end{aligned} \tag{4.29}$$

and hence

$$\beta_n(\delta) \le c_1\delta^{1/p}n + c_2 n^{1-1/p} + c_3, \tag{4.30}$$

where c_1, c_2, $c_3 > 0$ do not depend on n and δ. Further,

$$\|A_n - C_n + \Delta_{n;N}\|_\infty \le \frac{c_0}{N}, \tag{4.31}$$

and, instead of (4.25), we this time obtain

$$\operatorname{rank} \Delta_{n;N}^{(\delta)} \le 2\left(c_1\delta^{1/p}n + c_2 n^{1-1/p} + c_3 + \frac{c_0}{\delta\varepsilon}\right). \tag{4.32}$$

To achieve what we are after, it is sufficient to choose $\delta = n^{-p/(p+1)}$.

If $m = 1$, then, instead of (4.30)–(4.31), we arrive at

$$\beta_n(\delta) \le c_1\delta^{1/p}n + c_2 n^{1-1/p} \ln^{1/p} n + c_3, \tag{4.33}$$

$$\|A_n - C_n + \Delta_{n;N}\|_\infty \le c_0\frac{\ln N}{N}. \tag{4.34}$$

If we take

$$N = \frac{c_0}{\delta\varepsilon} \ln \frac{c_0}{\delta\varepsilon},$$

then, as is easily seen,

$$\frac{c_0 \ln N}{\delta N} \le \varepsilon,$$

and, instead of (4.32),

$$\operatorname{rank} \Delta_{n;N}^{(\delta)} \le 2\left(c_1\delta^{1/p}n + c_2 n^{1-1/p} \ln^{1/p} n + c_3 + \frac{c_0}{\delta\varepsilon} \ln \frac{c_0}{\delta\varepsilon}\right). \tag{4.35}$$

Trying $\delta = n^{-p/(p+1)} \ln n$, we obviously arrive at (4.28) for $m = 1$, and that completes the proof. ■

Next, we turn to the case when f may be complex-valued. If such functions are included in K_m, the definition remains unaltered. Suppose f has only a finite number of zeros $x_j \in [-\pi, \pi]$, $j = 1, \ldots, t$, and, analogously to (4.7), assume that at each x_j, $|f(x)|$ has nonzero left- and right-hand derivatives of some order. Denote by p the maximum order of such first nonzero derivatives, and by $\widehat{K}_m^{(p)}$ the corresponding class of complex-valued functions $f \in K_m$.

THEOREM 4.3. *If $f \in \widehat{K}_m^{(p)}$ and simple circulants S_n are nonsingular, then the singular values of $S_n^{-1}A_n$ are clustered at* 1, *so that $\gamma_n(\varepsilon)$ is of the form* (4.9).

PROOF. Almost everything from the proof of Theorem 4.1 is available. Allowing for

$$\max_{-\pi \le x \le \pi} \left| |f(x)| - |f_n(x)| \right| \le \max_{-\pi \le x \le \pi} |f(x) - f_n(x)|,$$

we again arrive at (4.13). In parallel with (4.15), we also have

$$\|A_n - S_n + \Delta_{n;N}\|_1 \le \frac{c_0}{N^m}. \tag{4.36}$$

Moreover,

$$\|S_{n;\delta}^{-1}\|_\infty = \|S_{n;\delta}^{-1}\|_1 \le \delta^{-1}, \tag{4.37}$$

and consequently

$$\begin{aligned} \|S_{n;\delta}^{-1}(A_n - S_n + \Delta_{n;N})\|_\infty &\le \frac{c_0}{\delta N^m}, \\ \|S_{n;\delta}^{-1}(A_n - S_n + \Delta_{n;N})\|_1 &\le \frac{c_0}{\delta N^m}. \end{aligned} \tag{4.38}$$

It immediately follows that all singular values of $S_{n;\delta}^{-1}(A_n - S_n + \Delta_{n;N})$ are no greater in modulus than $c_0/\delta N^m$. Furthermore,

$$S_{n;\delta}^{-1}(A_n - S_n + \Delta_{n;N}) = S_n^{-1}(A_n - S_n) + \widehat{\Delta}_{n;N}^{(\delta)}, \tag{4.39}$$

where

$$\widehat{\Delta}_{n;N}^{(\delta)} = (S_{n;\delta}^{-1} - S_n^{-1})(A_n - S_n) + S_{n;\delta}^{-1}\Delta_{n;N}. \tag{4.40}$$

Clearly, rank $\widehat{\Delta}_{n;N}^{(\delta)}$ can be upper-bounded by the right-hand side of (4.22), and the remainder of the proof is similar to that of Theorem 4.1. ■

THEOREM 4.4. *If $f \in \widehat{K}_m^{(p)}$ and optimal circulants C_n are nonsingular, then the singular values of $C_n^{-1}A_n$ are clustered at* 1 *so that $\gamma_n(\varepsilon)$ satisfies* (4.28).

We omit the proof, because it is wholly analogous to those of Theorems 4.1 and Theorem 4.3.

An interesting case is that of infinite smoothness. In particular, assume that

$$f(x) = \Phi(e^{ix}), \tag{4.41}$$

where $\Phi(z)$ is a function of the complex variable z which is analytic in a ring $r_1 < |z| < r_2$, where $r_1 < 1 < r_2$. Then for some $q > 1$

$$|a_k| = O\left(\frac{1}{q^k}\right). \tag{4.42}$$

Denote by K_∞ the set of all such functions f, and let $K_\infty^{(p)}$ and $\widehat{K}_\infty^{(p)}$ be defined in a similar way to $K_m^{(p)}$ and $\widehat{K}_m^{(p)}$. Theorem 4.1 and 4.3 are naturally complemented by the following

THEOREM 4.5. *If $f \in K_\infty^{(p)}$ is nonnegative and S_n are positive definite, the eigenvalues of $S_n^{-1}A_n$ are distributed so that*

$$\gamma_n(\varepsilon) = O(\ln n). \tag{4.43}$$

If $f \in \widehat{K}_\infty^{(p)}$ and S_n are nonsingualr, then the singular values of $S_n^{-1}A_n$ are clustered at 1 *so that* (4.43) *holds as well.*

The proof can follow the same scheme as in the above proofs. One need only note that, instead of (4.23), we have the right-hand side of the form $c_0/\delta q^N$. Setting this equal to ε, we find $N = \log_2(c_0/\delta\varepsilon)$. Instead of (4.23), rank $\Delta_{n;N}^{(\delta)}$ is now estimated as $O(\delta^{1/P}n + nq^{-p/n} + \ln\, \delta^{-1})$. Picking $\delta = n^{-1/p}$, we thus arrive at (4.43). The proof for the singular value case is very much the same.

It is curious, though rather foreseen, that the infinite smoothness has no effect on the estimates for optimal circulants. We should also caution against a possible misunderstanding by stressing that infinite smoothness as we mean it here is not the same as the existence of the infinite number of derivatives. The latter may not provide (4.42).

One more limiting case seems to deserve some attention. That is the one when

$$f(x) = \sum_{k=-\nu}^{\mu} a_k e^{ikx}. \tag{4.44}$$

In this case, the Toeplitz matrices A_n become banded.

THEOREM 4.6. *Suppose f is of the form* (4.44), *and not everywhere zero. If f is nonnegative and the circulants S_n are positive definite, then the eigenvalues of $S_n^{-1}A_n$ are clustered so that*

$$\gamma_n(\varepsilon) = O(1). \tag{4.45}$$

In a wider case, if the circulants S_n are nonsingular, then the singular values of $S_n^{-1}A_n$ are clustered so that (4.45) *stands as well.*

PROOF. If we take $N \geq \mu + \nu$, then for sufficiently large n, $A_n - S_n + \Delta_{n\,;\,N}$ approaches zero. Let $f \geq 0$ and $S_n > 0$. Then $S_n^{-1/2}(A_n - S_n)S_n^{-1/2}$ differs from the zero matrix by $S_n^{-1/2}\Delta_{n;\,N}S_n^{-1/2}$, whose rank is upper-bounded by $2N$. Hence at most $2N$ eigenvalues of $S_n^{-1}A_n - I$ can be distinct from zero. Similar arguments are needed concerning the singular-value case. ■

We are now ready to present six more theorems. These are, in effect, the above six theorems of this section adjusted to deal with improved circulants defined by (3.13)–(3.19). When f is nonnegative, we replace S_n by $\widetilde{S}_n$, which must always be positive definite. When f is complex-valued, we use $\widehat{S}_n$, which must always be nonsingular. Further, instead of C_n we may always use $\widehat{C}_n$. The necessary adjustment of formulations will be clear.

5. MULTILEVEL CIRCULANT PRECONDITIONERS

Here we take up a block Toeplitz matrix composed of $(n_1 + 1) \times (n_1 + 1)$ blocks, and suppose that every block is again a block Toeplitz matrix composed of $(n_2 + 1) \times (n_2 + 1)$ blocks, and so on. If there are L levels of such nested partitionings, then the corresponding matrix will be called an L-level Toeplitz matrix. The order of such a matrix is obviously equal to

$$n = (n_1 + 1) \cdots (n_L + 1), \tag{5.1}$$

and for its (i, j) entry, $0 \leq i, j \leq n$, it is convenient to write

$$a_{ij} = a_{i_1 - j_1;\, \ldots;\, i_L - j_L}, \tag{5.2}$$

where

$$\begin{aligned} i &= i_1 \prod_{\nu=2}^{L} (n_\nu + 1) + i_2 \prod_{\nu=3}^{L} (n_\nu + 1) + \cdots + i_{L-1}(n_L + 1) + i_L, \\ j &= j_1 \prod_{\nu=2}^{L} (n_\nu + 1) + j_2 \prod_{\nu=3}^{L} (n_\nu + 1) + \cdots + j_{L-1}(n_L + 1) + j_L, \\ & 0 \le i_1, j_1 \le n_1, \ldots, \qquad 0 \le i_L, j_L \le n_L. \end{aligned} \tag{5.3}$$

We shall assume that L-level Toeplitz matrices

$$A_{\bar{n}}^{\nu} = [a_{i_1 - j_1; \ldots; i_L - j_L}], \qquad \bar{n} = (n_1, \ldots, n_L), \tag{5.4}$$

are generated by the L-dimensional Fourier series

$$f(x_1, \ldots, x_L) = \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_L=-\infty}^{\infty} a_{k_1; \ldots; k_L} \exp[i(k_1 x_1 + \cdots + k_L x_L)]. \tag{5.5}$$

In order to precondition A_n, it is natural to try L-level circulants, i.e. matrices of the form

$$C = [c_{i_1 - j_1 (\mathrm{mod}\, n_1 + 1); \ldots; i_L - j_L (\mathrm{mod}\, n_L + 1)}].$$

Here we focus on the optimal (Cesaro) L-level circulants [13, 14],

$$C_{\bar{n}} = \left[c^{(\bar{n})}_{i_1 - j_1 (\mathrm{mod}\, n_1 + 1); \ldots; i_L - j_L (\mathrm{mod}\, n_L + 1)} \right]. \tag{5.6}$$

where

$$c^{(\bar{n})}_{k_1; \ldots; k_L} = \frac{1}{n} \sum_{\substack{i_1, j_1 = 0 \\ i_1 - j_1 = k_1 (\mathrm{mod}\, n_1 + 1)}}^{n_1} \cdots \sum_{\substack{i_L, j_L = 0 \\ i_L - j_L = k_L (\mathrm{mod}\, n_L + 1)}}^{n_L} a_{i_1 - j_1; \ldots; i_L - j_L}, \tag{5.7}$$

and also on the simple L-level circulants [14]

$$S_{\bar{n}} = [A_{i_1 - j_1 (\mathrm{mod}\, n_1 + 1); \ldots; i_L - j_L (\mathrm{mod}\, n_L + 1)}], \tag{5.8}$$

where

$$A_{k_1; \ldots; k_L} = \begin{cases} 0 & \text{if there exists } i \text{ such that} \\ & \quad k_i = (n_i + 1)/2, \\ a_{\varphi_1(k_1); \ldots; \varphi_L(k_L)} & \text{otherwise,} \end{cases} \tag{5.9}$$

$$\varphi_i(k_i) = \begin{cases} k_i & \text{if} \quad k_i < (n_i + 1)/2, \\ -(n_i + 1 - k_i) & \text{otherwise.} \end{cases} \tag{5.10}$$

The contents of Section 2 can be naturally extended to the case of multilevel matrices [14], and this permits us to produce the analogs of the previous results.

By a sparsely vanishing function of L variables, we mean a Lebesgue-integrable function $f(x_1, \ldots, x_L)$ which is 2π-periodic with respect to each argument, and enjoys the relationship

$$\lim_{\varepsilon \to +0} \int_{-\pi}^{\pi} \cdots \int_{-\pi}^{\pi} \varphi_\varepsilon(|f(x_1, \ldots, x_L)|)\, dx_1 \cdots dx_L = 0, \tag{5.11}$$

where $\varphi_\varepsilon(x) = 1$ if $x \in [0, \varepsilon]$, and $\varphi_\varepsilon(x) = 0$ elsewhere.

THEOREM 5.1. *Let $f(x_1, \ldots, x_L) \in L_2$ be sparsely vanishing, and $A_{\bar{n}}$, $C_{\bar{n}}$, $S_{\bar{n}}$ be L-level Toeplitz matrices and optimal and simple L-level circulants, respectively. Assume that $C_{\bar{n}}(S_{\bar{n}})$ are positive definite. Then the eigenvalues of $C_{\bar{n}}^{-1} A_{\bar{n}}$ (or $S_{\bar{n}}^{-1} A_{\bar{n}}$) are real and clustered at* 1.

As previously, we can bring in improved circulants. Since all L-level circulants with common $\bar{n}$ are unitarily diagonalized via the same transforming matrix (see [13]), we can write

$$C_{\bar{n}} = Q_{\bar{n}}^* \operatorname{diag}(\lambda_k^{(\bar{n})})\, Q_{\bar{n}}, \qquad S_{\bar{n}} = Q_{\bar{n}}^* \operatorname{diag}(\mu_k^{(\bar{n})})\, Q_{\bar{n}},$$

and then set

$$\begin{aligned} \widehat{C}_{\bar{n}} &= Q_{\bar{n}}^* \operatorname{diag}(\widehat{\lambda}_k^{(\bar{n})})\, Q_{\bar{n}}, \\ \widehat{S}_{\bar{n}} &= Q_{\bar{n}}^* \operatorname{diag}(\widehat{\mu}_k^{(\bar{n})})\, Q_{\bar{n}}, \\ \widetilde{S}_{\bar{n}} &= Q_{\bar{n}}^* \operatorname{diag}(\widetilde{\mu}_k^{(\bar{n})})\, Q_{\bar{n}}, \end{aligned}$$

where

$$\begin{aligned} \widehat{\lambda}_k^{(\bar{n})} &= \begin{cases} \lambda_k^{(\bar{n})} & \text{if} \quad \lambda_k^{(\bar{n})} \neq 0, \\ \delta & \text{otherwise}, \end{cases} \\ \widehat{\mu}_k^{(\bar{n})} &= \begin{cases} \mu_k^{(\bar{n})} & \text{if} \quad \mu_k^{(\bar{n})} \neq 0, \\ \delta & \text{otherwise}, \end{cases} \\ \widetilde{\mu}_k^{(\bar{n})} &= \begin{cases} \mu_k^{(\bar{n})} & \text{if} \quad \mu_k^{(\bar{n})} > 0, \\ \delta & \text{otherwise}; \end{cases} \end{aligned}$$

here δ is arbitrary but positive.

THEOREM 5.2. *If $f(x_1, \ldots, x_L) \in L_2$ is nonegative and sparsely vanishing, then the eigenvalues of $\widehat{C}_{\bar{n}}^{-1} A_{\bar{n}}$ (or $\tilde{S}_n^{-1} A_{\bar{n}}$) are real and clustered at* 1.

THEOREM 5.3. *Suppose that $f(x_1, \ldots, x_L) \in L_2$ is sparsely vanishing and the matrices $C_{\bar{n}}(S_{\bar{n}})$ are nonsingular. Then the singular value of $C_{\bar{n}}^{-1} A_{\bar{n}}$ (or $S_{\bar{n}}^{-1} A_{\bar{n}}$) are clustered at* 1.

THEOREM 5.4. *If $f(x_1, \ldots, x_L) \in L_2$ is sparsely vanishing, then the improved circulants $\widehat{C}_{\bar{n}}$ and $\widehat{S}_{\bar{n}}$ are nonsingualr, and the singular values of $\widehat{C}_{\bar{n}}^{-1} A_{\bar{n}}$ and $\widehat{S}_{\bar{n}}^{-1} A_{\bar{n}}$ are clustered at* 1.

The proofs echo those of theorems 3.1–3.4 almost word by word.

6. CONCLUDING REMARKS

It should be said that we do not know how sharp are the estimates on $\gamma_n(\varepsilon)$ obtained in Section 4.

All the same, even if those estimates were sharp it would not immediately follow that simple circulants are much better than optimal ones. First of all, the estimates are of an asymptotic nature, and the corresponding constants are not yet known. We guess that gleaning them is not a simple task. In practice, we often observe that $\gamma_n(\varepsilon)$ grows in a very moderate fashion. As a consequence, circulant preconditioners prove to be nicely efficient even when their inverses have unbounded norm.

For "bad" functions f, optimal circulants are sometimes better than simple ones, but they seem to never be much better. At the same time, the simple circulants, in practice, do not very often outperform the optimal ones. However, there are some cases when they distinctly do (see the example of banded matrices in the Introduction).

Finally, when using circulants we recommend always using the improved circulants. Perhaps these should be improved simple circulants.

I gratefully acknowledge the assistance by V. Strela in obtaining the above-discussed numerical information. I am also indebted to the referee, who kindly suggested many corrections to improve the language of this paper.

REFERENCES

1 F. Avram, On bilinear forms on Gaussian variables and Toeplitz matrices, *Probab. Theory Related Fields* 79:37–45 (1988).

2 O. Axelsson and G. Lindskog, On the rate of convergence of the preconditioned conjugate method, *Number. Math.* 48:499–523 (1986).
3 R. Chan, Circulant preconditioners for Hermitian Toeplitz systems, *SIAM J. Matrix Anal. Appl.* 10:542–550 (1989).
4 R. Chan, The spectrum of a family of circulant preconditioned Toeplitz systems, *SIAM J. Matrix Anal. Appl.* 10:542–550 (1989).
5 R. Chan and X. Jin, A family of block preconditioners for block systems *SIAM J. Sci. Statist. Comput.* 13(5):1218–1235 (1992).
6 R. Chang and M. Yeung, *Circulant Preconditioners for Complex Toeplitz Matrices,* Research Rep. HKUM-91-6, Dept. of Mathematics, Univ. of Hong Kong, 1991.
7 T. F. Chan, An optimal circulant preconditioner for Toeplitz systems, *SIAM J. Sci. Statist Comput.* 9:766–771 (1988).
8 T. Huckle, *Circulant and skewcirculant matrices for solving Toeplitz matrix problems, SIAM J. Sci. Statist. Comput.* 13(5):767–780 (1992).
9 G. H. Golub and C.F. Van Loan, *Matrix Computations*, 2nd ed., John Hopkins U.P., Baltimore, 1989.
10 U. Grenander and G. Szego, *Toeplitz Forms and Their Applications*, Univ. of California Press, Berkeley, 1958.
11 S.V. Parter, On the distribution of the singular values of Toeplitz matrices, *Linear Algebra Appl.* 80:115–130 (1986).
12 G. Strang, A proposal for Toeplitz matrix calculations, *Stud. Appl. Math* 74:171–176 (1986).
13 E. E. Tyrtyshnikov, Optimal and superoptimal circulant preconditioners, *SIAM J. Matrix Anal. Appl.* 13(2):459–473 (1992).
14 E. E. Tyrtyshnikov, A unifying approach to some old and new theorems on distribution and clustering, *Linear Algebra Appl.* submitted for publication.
15 E. E. Tyrtyshnikov, *Toeplitz Matrices, Some of their Analogues and Applications* (in Russian), Dept. of Numerical Mathematics, USSR Acad. of Sciences, Moscow, 1989.
16 J. W. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon, Oxford, 1965.

Received 20 July 1992; final manuscript accepted 15 April 1993