

# A PRECONDITIONED MINRES METHOD FOR NONSYMMETRIC TOEPLITZ MATRICES

J. PESTANA<sup>†</sup> AND A. J. WATHEN<sup>‡</sup>

**Abstract.** Circulant preconditioning for symmetric Toeplitz linear systems is well-established; theoretical guarantees of fast convergence for the conjugate gradient method are descriptive of the convergence seen in computations. This has led to robust and highly efficient solvers based on use of the fast Fourier transform exactly as originally envisaged in Gil Strang’s ‘Proposal for Toeplitz Matrix Calculations’ (Studies in Applied Mathematics, 74, pp. 171–176, 1986.). For nonsymmetric systems, the lack of generally descriptive convergence theory for most iterative methods of Krylov type has provided a barrier to such a comprehensive guarantee, though several methods have been proposed and some analysis of performance with the normal equations is available.

In this paper, by the simple device of reordering, we rigorously establish a circulant preconditioned short recurrence Krylov subspace iterative method of minimum residual type for nonsymmetric (and possibly highly nonnormal) Toeplitz systems. Convergence estimates similar to those in the symmetric case are established.

**Key words.** circulant preconditioner, MINRES, nonsymmetric matrix, Toeplitz matrix

**AMS subject classifications.** 65F08, 65F10, 15B05

**1. Introduction.** Krylov subspace methods are well suited to linear systems

$$(1.1) \quad A_n x_n = b_n$$

involving large nonsingular Toeplitz matrices  $A_n \in \mathbb{R}^{n \times n}$ , since matrix vector products can be cheaply performed via a circulant embedding and the fast Fourier transform [55, Section 4.2.4]. Additionally, preconditioned Krylov subspace methods can be more robust than tailored direct methods and have better complexity [8, Section 1.1]. Circulant preconditioners  $C_n \in \mathbb{R}^{n \times n}$  for  $A_n$  are attractive, since these can also be applied efficiently using the fast Fourier transform. However, theoretical guarantees of effectiveness of these preconditioners have mostly been restricted to symmetric (or Hermitian) positive definite Toeplitz matrices [4, 8, 11, 12, 14, 15, 32, 35, 38, 40, 49, 52]. This is largely due to the fact that when  $C_n$  and  $A_n$  are symmetric positive definite matrices, the preconditioned conjugate gradient method (CG) [29] can be applied with its favourable properties: namely, minimization of the error in the  $A_n$ -norm, short-term recurrences and convergence bounds that depend only on the eigenvalues of the preconditioned matrix. These eigenvalues can be analysed in many cases, so that convergence within a certain number of iterations of CG can be guaranteed.

If  $A_n$  is symmetric but indefinite we can apply preconditioned MINRES [42] as long as the preconditioner is positive definite. MINRES has similar properties to CG, namely minimization of the preconditioned residual, short-term recurrences and convergence bounds in terms of the eigenvalues of the preconditioned matrix. However, the spectrum of the preconditioned matrix may be more difficult to analyse than in case of positive definite  $A_n$  [9, 39].

Often considered hardest to solve by preconditioned iterative methods are systems with nonsymmetric (nonnormal) matrices. It may be more challenging to obtain

---

<sup>†</sup>School of Mathematics, The University of Manchester, Manchester, M13 9PL, United Kingdom (jennifer.pestana@manchester.ac.uk)

<sup>‡</sup>Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, United Kingdom (wathen@maths.ox.ac.uk)

a suitable preconditioner, although effective preconditioners for certain nonsymmetric Toeplitz matrices have been proposed [13, 33, 36, 41]. However, in general Krylov subspace methods for nonnormal matrices, such as GMRES [47], QMR [24], SQMR [25] (which requires less work than the original QMR method when applied to Toeplitz matrices) and BiCG-STAB [54], do not simultaneously minimize a relevant quantity and have short-term recurrences for general matrices [22]. Additionally, the convergence of these methods cannot be bounded using only the eigenvalues (or singular values). This is in stark contrast to methods for symmetric matrices. The exception is the application of CG to the normal equations (CGNE), or MINRES to an augmented system. The former does not preserve the Toeplitz structure, although  $A_n$  and  $A_n^T$  can still be cheaply applied, while decay properties of the Toeplitz matrix may be lost in the augmented matrix. Additionally, CGNE and MINRES on an augmented system may be more costly, may result in worse conditioning (and slower convergence), and may be more difficult to precondition effectively.

In this note we show that nonsymmetry and nonnormality of Toeplitz matrices pose no difficulty in terms of obtaining theoretical guarantees of convergence of a minimum residual method with short-term recurrences, since a simple permutation renders the Toeplitz matrix symmetric (but possibly indefinite). MINRES can, therefore, be applied to suitably preconditioned Toeplitz problems. The preconditioner is a circulant matrix that is as cheap and easy to apply as other proposed preconditioners for Toeplitz systems. Significantly, since the convergence of MINRES can be bounded using only the eigenvalues of the preconditioned matrix, we can obtain descriptive convergence bounds.

We restrict our attention to real Toeplitz matrices here. When applied to complex Toeplitz matrices, our method produces complex symmetric matrices and the associated linear system can be solved by SQMR [25] or by the recently proposed CS-MINRES [17], a minimal residual method for complex symmetric matrices that can be used with a complex symmetric preconditioner.

**2. Preliminaries.** Let  $A_n \in \mathbb{R}^{n \times n}$  be the nonsingular Toeplitz matrix

$$A_n = \begin{bmatrix} a_0 & a_{-1} & \dots & a_{-n+2} & a_{-n+1} \\ a_1 & a_0 & a_{-1} & & a_{-n+2} \\ \vdots & a_1 & a_0 & \ddots & \vdots \\ a_{n-2} & & \ddots & \ddots & a_{-1} \\ a_{n-1} & a_{n-2} & \dots & a_1 & a_0 \end{bmatrix}$$

and  $C_n \in \mathbb{R}^{n \times n}$  be the nonsingular circulant preconditioner

$$C_n = \begin{bmatrix} c_0 & c_{n-1} & \dots & c_2 & c_1 \\ c_1 & c_0 & c_{n-1} & & c_2 \\ \vdots & c_1 & c_0 & \ddots & \vdots \\ c_{n-2} & & \ddots & \ddots & c_{n-1} \\ c_{n-1} & c_{n-2} & \dots & c_1 & c_0 \end{bmatrix}.$$

If the Fourier matrix is denoted  $F_n = (f_{jk})$  with  $f_{jk} = e^{-2(j-1)(k-1)\pi i/n}$ ,  $j, k = 1, \dots, n$ , then the circulant  $C_n$  has diagonalization

$$(2.1) \quad C_n = U_n^* \Lambda_n U_n$$

where  $U_n = F_n/\sqrt{n}$  and  $\Lambda = \text{diag}(F_n c_n)$ , with  $c_n$  the first column of  $C_n$ . A consequence of the relationship between  $C_n$  and the Fourier matrix is that a matrix-vector product, or the solution of a linear system, with a circulant matrix of dimension  $n$  can be performed in  $O(n \log(n))$  operations using the fast Fourier transform (FFT) [55, Section 4.2.3]. Note that although  $C_n$  is real,  $F_n$  and  $U_n$  are generally complex.

Superfast direct solvers with  $O(n \log^2(n))$  complexity have been developed for Toeplitz systems but these require that a certain submatrix of  $A_n$  of dimension  $\lfloor n/2 \rfloor$  is nonsingular. The algorithms can, therefore, break down if a principal submatrix of  $A_n$  is ill-conditioned or singular, although look-ahead strategies have been proposed [14]. Iterative methods, in contrast, do not require that such a submatrix is nonsingular. Moreover, since each matrix-vector multiply with  $A_n$  can be performed in  $O(n \log(n))$  time, if the number of iterations is small the complexity of iterative methods is better than that of direct methods for solving Toeplitz systems. Typically, preconditioning is required to keep the number of iterations small.

Let  $A_n$  have as its coefficients the Fourier coefficients of a function  $f$  that is  $L^2$ -integrable over  $[-\pi, \pi]$ , i.e.,

$$a_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx,$$

$k = 0, \pm 1, \pm 2, \dots$  with  $a_k$  the entries of  $A_n$  above. Then  $f$  is the generating function associated with  $A_n$ . If the real Toeplitz matrix  $A_n$  is associated with a real, non-negative generating function that is not identically zero then it is symmetric positive definite. In this case a number of circulant preconditioners, including the Strang [49] and optimal [15] preconditioners, are such that for large enough  $n$ ,  $C_n^{-1}A_n$  is a perturbation of the identity of the form  $R+E$ , where the rank of  $R$  is bounded independently of  $n$  and  $E$  has small norm. Consequently, the eigenvalues of  $C_n^{-1}A_n$  are clustered near one with the exception of a fixed number of outliers and CG for the preconditioned system is guaranteed to converge within a certain number of steps. Similar results can be obtained for certain circulant preconditioners for MINRES, with the eigenvalues clustered near 1 and -1 for large enough  $n$  [9, 39].

Circulant [13, 36, 41],  $\omega$ -circulant [10] and banded Toeplitz [5, 33, 44] preconditioners have been proposed for nonsymmetric  $A_n$ . For some of these preconditioners  $P_n$ , the preconditioned matrix  $P_n^{-1}A_n = I_n + R + E$ , where  $I_n$  is the identity matrix, and again the rank of  $R$  is bounded independently of  $n$  and  $E$  has small norm. However, as previously mentioned, this is insufficient to guarantee rapid convergence of most Krylov solvers, including GMRES, QMR, SQMR and BiCGSTAB. Despite the lack of rigorous convergence proofs, circulant preconditioners for non-Hermitian and nonsymmetric Toeplitz matrices have been used successfully with methods like GMRES [33, 41], which is perhaps favoured because it minimizes the preconditioned residual.

**3. MINRES for preconditioned Toeplitz matrices.** We would like to solve the Toeplitz system (1.1) by a minimal residual method with short-term recurrences and descriptive convergence bounds. It turns out that this can be achieved without any additional cost, since  $A_n$  is transformed to a symmetric (Hankel) matrix by swapping its rows or, equivalently, by forming  $Y_n A_n$ , where  $Y_n \in \mathbb{R}^{n \times n}$  is the exchange matrix

$$Y_n = \begin{bmatrix} & & & 1 \\ & & \cdot & \\ & \cdot & & \\ 1 & & & \end{bmatrix}.$$

An equivalent statement is that  $A_n$  is self-adjoint with respect to the bilinear form defined by  $Y_n$  [26, 45]. The matrix  $Y_n$  is symmetric and orthogonal, and is thus involutory. Moreover, since  $Y_n$  is orthogonal, both  $A_n$  and  $Y_n A_n$  have the same 2-norm condition number.

The permutation  $Y_n A_n$  converts a Toeplitz matrix that may be highly nonnormal to a matrix that is not only normal but symmetric, so that the system

$$(3.1) \quad Y_n A_n x = Y_n b$$

can be solved by MINRES. Within the MINRES algorithm, multiplications by  $Y_n$  are replaced by permutations. Convergence of MINRES for (3.1) depends on the eigenvalues of  $Y_n A_n$  and may be slow if these are not nicely distributed. However, if an effective (invertible) circulant preconditioner  $C_n$  for  $A_n$  is known, the unitary diagonalization of  $C_n$  allows us to easily replace  $C_n$  by its absolute value. Recall that the absolute value of a general matrix  $B \in \mathbb{R}^{n \times n}$  is given by  $|B| = (B^T B)^{\frac{1}{2}}$  [30, Chapter 8]. Since  $C_n$  is normal, we find that

$$(3.2) \quad |C_n| = (C_n^T C_n)^{\frac{1}{2}} = (C_n C_n^T)^{\frac{1}{2}} = U_n^* \Lambda_n U_n,$$

where  $\Lambda_n$  is the diagonal matrix in (2.1) with all entries replaced by their magnitudes. Additionally,

$$(3.3) \quad |C_n|^{-1} Y_n = Y_n |C_n|^{-1}$$

because  $|C_n|^{-1}$  is also symmetric and circulant, hence certainly Toeplitz, and

$$(3.4) \quad |C_n| = C_n \tilde{C}_n, \quad \tilde{C}_n = U_n^* \text{sgn}(\Lambda_n) U_n,$$

where  $\text{sgn}(\Lambda_n) = \text{diag}(\lambda_i/|\lambda_i|)$  and  $\tilde{C}_n \in \mathbb{R}^{n \times n}$  is orthogonal and involutory. Since  $|C_n|$  is a circulant, matrix-vector products, or linear solves, with  $|C_n|$  can also be applied in  $O(n \log(n))$  operations, so that  $|C_n|$  is no more difficult to apply than  $C_n$ . Note that if  $C_n$  is positive definite then  $\Lambda = |\Lambda|$ , i.e., the absolute value leaves  $C_n$  unchanged.

Accordingly, the solution  $x$  of (1.1) can be found by applying preconditioned MINRES, described in, for example, Elman, Silvester and Wathen [21, Chapter 6] to (3.1) with preconditioner  $|C_n|$ . The preconditioned MINRES method produces iterates  $\{x_k\}$ , with residual  $r_k = b - A_n x_k$  so that norm of the preconditioned residual

$$\| |C_n|^{-1} Y_n r_k \|_2 = \| Y_n |C_n|^{-1} r_k \|_2 = \| |C_n|^{-1} r_k \|_2$$

is minimized. Since  $|C_n|$  has unitary eigenvectors,  $\| |C_n|^{-1} r_k \|_2 = \| \Lambda_n^{-1} v_k \|_2$  with  $v_k = U_n r_k$ , so that  $\| v_k \|_2 = \| r_k \|_2$ . Relative preconditioned residuals are bounded as follows:

$$\frac{\| |C_n|^{-1} r_k \|_2}{\| |C_n|^{-1} r_0 \|_2} \leq \min_{\substack{p \in \Pi_k \\ p(0)=1}} \max_{\lambda \in \sigma(|C_n|^{-1} Y_n A_n)} |p(\lambda)|,$$

where  $\Pi_k$  is the set of polynomials of at most degree  $k$  and  $\sigma(|C_n|^{-1} Y_n A_n)$  is the spectrum of  $|C_n|^{-1} Y_n A_n$ . This shows that knowledge of the eigenvalues of  $|C_n|^{-1} Y_n A_n$ , or of inclusion regions for these eigenvalues, is sufficient to guarantee that preconditioned MINRES will converge to a given tolerance within a certain number of iterations. In

particular, if the eigenvalues of  $|C_n|^{-1}Y_nA_n$  lie in two intervals  $[-a, -b] \cup [c, d]$ , of equal length, so that  $d - c = a - b$ , then [21, pp. 306–307]

$$(3.5) \quad \frac{\| |C_n|^{-1}r_{2k} \|_2}{\| |C_n|^{-1}r_0 \|_2} \leq 2 \left( \frac{\sqrt{ad} - \sqrt{bc}}{\sqrt{ad} + \sqrt{bc}} \right)^k.$$

To this point we have used a row permutation to symmetrize  $A_n$ . However, we could instead have applied a column permutation, since  $A_nY_n$  is also symmetric. In general,  $A_nY_n \neq Y_nA_n$  but both approaches give the same iterates and residuals as the next lemma shows.

LEMMA 1. *When the same starting vector  $x_0 \in \mathbb{R}^n$  is used, the  $k$ th iterate  $x_k^L$  of preconditioned MINRES applied to (3.1) with preconditioner  $|C_n|$  is equal to the  $k$ th iterate  $x_k^R$  of preconditioned MINRES applied to*

$$(A_nY_n)z = b, \quad x = Y_nz,$$

with preconditioner  $|C_n|$  for  $k = 1, 2, \dots, n$ .

*Proof.* If we solve  $Y_nA_nx = Y_nb$  by preconditioned MINRES, we effectively solve

$$(|C_n|^{-\frac{1}{2}}Y_nA_n|C_n|^{-\frac{1}{2}})|C_n|^{\frac{1}{2}}x = |C_n|^{-\frac{1}{2}}Y_nb$$

by MINRES. At the  $k$ th step MINRES selects an iterate  $x_k^L$  of the form

$$x_k^L = x_0 + |C_n|^{-\frac{1}{2}}p_{k-1}^L(|C_n|^{-\frac{1}{2}}Y_nA_n|C_n|^{-\frac{1}{2}})|C_n|^{-\frac{1}{2}}Y_nr_0,$$

where  $p_{k-1}^L$  is a polynomial of degree at most  $k-1$ . Since  $|C_n|$  is real symmetric and circulant, so is  $|C_n|^{-\frac{1}{2}}$ , and it follows that  $Y_n|C_n|^{-\frac{1}{2}} = |C_n|^{-\frac{1}{2}}Y_n$ . It is then straightforward to show that

$$(3.6) \quad x_k^L = x_0 + Y_n|C_n|^{-1}p_{k-1}^L(A_nY_n|C_n|^{-1})r_0.$$

The polynomial  $p_{k-1}^L$  is selected so that the 2-norm of the preconditioned residual

$$|C_n|^{-\frac{1}{2}}Y_nr_k = q_k^L(|C_n|^{-\frac{1}{2}}Y_nA_n|C_n|^{-\frac{1}{2}})|C_n|^{-\frac{1}{2}}Y_nr_0 = Y_n|C_n|^{-\frac{1}{2}}q_k^L(A_nY_n|C_n|^{-1})r_0$$

is minimized, where  $q_k^L(z) = 1 - zp_{k-1}^L(z)$ . Thus, the residual polynomial is

$$(3.7) \quad q_k^L = \arg \min_{\substack{q \in \Pi_k \\ q(0)=1}} \| |C_n|^{-\frac{1}{2}}q(A_nY_n|C_n|^{-1})r_0 \|_2.$$

On the other hand, solving  $(A_nY_n)z = b$ ,  $x = Y_nz$  by preconditioned MINRES is equivalent to solving

$$(|C_n|^{-\frac{1}{2}}A_nY_n|C_n|^{-\frac{1}{2}})|C_n|^{\frac{1}{2}}Y_nx = |C_n|^{-\frac{1}{2}}b.$$

The  $k$ th iterate  $x_k^R$  is of the form

$$x_k^R = x_0 + Y_n|C_n|^{-\frac{1}{2}}p_{k-1}^R(|C_n|^{-\frac{1}{2}}A_nY_n|C_n|^{-\frac{1}{2}})|C_n|^{-\frac{1}{2}}r_0$$

or

$$(3.8) \quad x_k^R = x_0 + Y_n|C_n|^{-1}p_{k-1}^R(A_nY_n|C_n|^{-1})r_0,$$

where  $p_{k-1}^R$  is also a polynomial of degree at most  $k-1$ . The 2-norm of the preconditioned residual

$$|C_n|^{-\frac{1}{2}} r_k = q_k^R (|C_n|^{-\frac{1}{2}} A_n Y_n |C_n|^{-\frac{1}{2}}) |C_n|^{-\frac{1}{2}} r_0 = |C_n|^{-\frac{1}{2}} q_k^R (A_n Y_n |C_n|^{-1}) r_0$$

is minimized by MINRES, with  $q_k^R(z) = 1 - zp_{k-1}^R(z)$ , and so

$$q_k^R = \arg \min_{\substack{q \in \Pi_k \\ q(0)=1}} \| |C_n|^{-\frac{1}{2}} q(A_n Y_n |C_n|^{-1}) r_0 \|_2.$$

Comparison with (3.7) shows that  $q_k^L = q_k^R$ . From the definitions of  $q_k^L$  and  $q_k^R$  we see that  $p_{k-1}^L = p_{k-1}^R$  and, from (3.6) and (3.8), that  $x_k^L = x_k^R$  at each iteration.  $\square$

Let us contrast this proposed MINRES method with two alternatives, CG for the normal equations and SQMR. When the eigenvalues of  $|C_n|^{-1} Y_n A_n$  are completely symmetric about the origin, CGNE converges in half the number of iterations of MINRES, but each CGNE iteration requires two matrix-vector products with  $A_n$  and two solves with  $C_n$ , so that the work of the two methods is roughly the same [23]. When this exact symmetry does not apply MINRES converges faster, i.e. requires fewer matrix-vector multiplies and preconditioner solves. Both methods minimize the residual and use short-term recurrences. However, it may be more difficult to find suitable preconditioners for the normal equations than it is to find preconditioners for MINRES. This issue is not restricted to Toeplitz matrices, but more generally a good preconditioner  $P$  for  $A$  does not imply that  $PP^T$  is a good preconditioner for  $AA^T$ .

Now consider SQMR, which uses  $Y_n$  to reduce the work of QMR. Unlike MINRES, SQMR can be used with an indefinite preconditioner but it does not minimize the residual or the error with respect to a standard norm and breakdown may still occur. These properties make guarantees on convergence rates difficult, if not impossible, to obtain. A similar analysis to that in the proof of Lemma 1 shows that SQMR applied to  $(Y_n C_n)^{-1} (Y_n A_n) x = (Y_n C_n)^{-1} (Y_n b)$  is equivalent to SQMR applied to  $(C_n Y_n)^{-1} (A_n Y_n) (Y_n x) = (C_n Y_n)^{-1} b$  and that SQMR applied to the system  $(C_n Y_n)^{-1} (Y_n A_n) x = (C_n Y_n)^{-1} (Y_n b)$  is equivalent to SQMR applied to  $(Y_n C_n)^{-1} (A_n Y_n) (Y_n x) = (Y_n C_n)^{-1} b$ .

**4. Using existing circulant preconditioners in MINRES.** The previous section showed that the proposed MINRES method can be used to solve systems (1.1) with nonsymmetric Toeplitz coefficient matrices, provided we can find suitable positive definite preconditioners  $|C_n|$  in (3.2). Here, we show that certain existing circulant preconditioners can be used to build  $|C_n|$ .

As discussed in Section 2, for certain nonsymmetric Toeplitz matrices and circulant preconditioners the preconditioned matrix  $C_n^{-1} A_n = I_n + R + E$ , where  $I_n$  is the  $n \times n$  identity matrix,  $R \in \mathbb{R}^{n \times n}$  has low rank and  $E \in \mathbb{R}^{n \times n}$  has small norm. Fortunately, a similar relationship also holds for the transformed matrix  $|C_n|^{-1} Y_n A_n$ .

PROPOSITION 2. *Let*

$$(4.1) \quad C_n^{-1} A_n = I + R + E$$

where  $\text{rank}(R) = K$  and  $\|E\| \leq \epsilon$ , where  $\|\cdot\|$  is a unitarily invariant norm. Then,

$$(4.2) \quad |C_n|^{-1} Y_n A_n = Q_n + \widehat{R} + \widehat{E},$$

where  $\text{rank}(\widehat{R}) = K$  and  $\|\widehat{E}\| = \|E\|$ . Moreover, the eigenvalues of  $Q_n$  are 1 or  $-1$ .

*Proof.* If (4.1) holds then, using (3.3) and (3.4),

$$|C_n|^{-1}Y_nA_n = Y_n|C_n|^{-1}A_n = Y_n\tilde{C}_nC_n^{-1}A_n = Q_n + \hat{R} + \hat{E},$$

where  $Q_n = Y_n\tilde{C}_n$ ,  $\hat{R} = Q_nR$  and  $\hat{E} = Q_nE$ . Since  $\tilde{C}_n$  is circulant,  $Q_n$  is not only orthogonal but symmetric and so has eigenvalues  $\pm 1$ , while the orthogonality of  $Q_n$  means that  $\text{rank}(\hat{R}) = \text{rank}(R) = K$  and  $\|\hat{E}\| = \|E\|$  with respect to any unitarily invariant norm.  $\square$

REMARK 1. A similar relationship between  $C_n$  and  $|C_n|$  was proposed by Ng and Potts for symmetric indefinite problems [39, Theorem 4.3].

REMARK 2. An analogous result can be proved for right preconditioning.

Proposition 2 shows that if  $C_n^{-1}A_n$  satisfies (4.1), the eigenvalues of  $|C_n|^{-1}Y_nA_n$  are clustered near  $\pm 1$  with a certain number of outliers. Additionally, the eigenvalues of  $Y_n|C_n|^{-1}A_n = |C_n|^{-1}Y_nA_n$  are real. This clustering of eigenvalues, in conjunction with the results of Campbell et al. [3] allow us to provide a more precise bound than (3.5) in this case.

PROPOSITION 3. *Let  $|C_n|^{-1}Y_nA_n$  have  $p$  eigenvalues  $\lambda \in \{z : |1 - z| \leq \epsilon\}$  and  $q$  eigenvalues  $\lambda \in \{z : |1 + z| \leq \epsilon\}$  with the remaining  $d = n - p - q$  outlying eigenvalues satisfying  $\lambda \in \{z : |1 - z| > \epsilon \text{ and } |-1 - z| > \epsilon\}$ . If we apply preconditioned MINRES to (3.1) with preconditioner  $|C_n|$  then, after the  $d$ th iteration, the preconditioned residuals satisfy*

$$(4.3) \quad \frac{\||C_n|^{-1}r_{d+2k}\|_2}{\||C_n|^{-1}r_0\|_2} \leq C(2\epsilon + \epsilon^2)^k,$$

where  $C$  is independent of  $k$  but depends on  $\epsilon$  and the distance from the outliers to the clusters at 1 and  $-1$ .

*Proof.* Recognising that MINRES and GMRES are mathematically equivalent, the result is obtained directly from Proposition 5.1 in Campbell et al. [3] with  $\sigma = 2 + \epsilon$ ,  $\rho = \epsilon$  and  $P = 2$  clusters.  $\square$

Combining the results of Propositions 2 and 3 we see that if  $C_n$  is such that  $C_n^{-1}A_n = I + R + E$  where  $\text{rank}(R) = K$  and  $\|E\| \leq \epsilon$ , and  $\|\cdot\|$  is a unitarily invariant norm, we can expect convergence of our MINRES method in a number of iterations that depends on  $K$  and on  $\epsilon$ . Moreover, if  $K$  and  $\epsilon$  are independent of the dimension  $n$  of the problem then MINRES will converge in a number of iterations that is independent of  $n$ .

Proposition 3 also leads to a 2-step superlinear convergence result that is a natural extension of the results for circulant preconditioners for positive definite Toeplitz matrices [8, page 431]. To achieve this superlinear convergence, as in the positive definite case, we require the notion of a clustered spectrum. For this purpose, we generalize the definition for positive definite matrices given in the survey paper by Chan and Ng [8, Definition 1.1] to our problem.

DEFINITION 4. *A sequence of matrices  $\{Z_n\}_{n=1}^{\infty}$  is said to have clustered spectrum around 1 and  $-1$  if for any  $\epsilon > 0$  there exist positive integers  $n_1$  and  $n_2$  such that for all  $n > n_1$ , at most  $n_2$  eigenvalues  $\lambda$  of  $Z_n$  are such that  $|\lambda - 1| > \epsilon$  and  $|\lambda + 1| > \epsilon$ . Examining (4.3) in Proposition 3 we see that if  $C_n^{-1}A_n$  has clustered spectra then we can expect 2-step superlinear convergence (c.f. equation (1.8) in Chan and Ng [8]).*

Although it seems more common in the literature to prove that (4.1) holds, there are cases in which eigenvalue bounds are given for the normal equations coefficient matrix  $(C_n^{-1}A_n)^T(C_n^{-1}A_n)$ . Proposition 5 below shows that any preconditioner that

works well for the normal equations will also work well for MINRES provided the condition number of  $C_n$  is not too large.

PROPOSITION 5. *Let the eigenvalues of  $(C_n^{-1}A_n)^T(C_n^{-1}A_n)$  be contained in the interval  $[1-\epsilon, 1+\epsilon]$  with the exception of  $K$  outliers. Then the eigenvalues of  $|C_n|^{-1}Y_nA_n$  are contained in*

$$\left[ \sqrt{\kappa}(-1-\epsilon), \frac{-1+\epsilon}{\sqrt{\kappa}} \right] \cup \left[ \frac{1-\epsilon}{\sqrt{\kappa}}, \sqrt{\kappa}(1+\epsilon) \right],$$

where  $\kappa = \kappa_2(C_n)$  is the 2-norm condition number of  $C_n$ , with the exception of  $K$  outliers.

*Proof.* Since

$$(C_n^{-1}A_n)^T(C_n^{-1}A_n) = (Y_n|C_n|^{-1}A_n)^T(Y_n|C_n|^{-1}A_n) = (|C_n|^{-1}Y_nA_n)^T(|C_n|^{-1}Y_nA_n),$$

the singular values of  $|C_n|^{-1}Y_nA_n$  are clustered in  $[1-\epsilon, 1+\epsilon]$  with the exception of  $K$  outliers. These singular values are related to those of  $|C_n|^{-\frac{1}{2}}Y_nA_n|C_n|^{-\frac{1}{2}}$  by the bound [20, Theorem 3.1]

$$\frac{\sigma_i(|C_n|^{-1}Y_nA_n)}{\sqrt{\kappa_2(|C_n|)}} \leq \sigma_i(|C_n|^{-\frac{1}{2}}Y_nA_n|C_n|^{-\frac{1}{2}}) \leq \sqrt{\kappa_2(|C_n|)}\sigma_i(|C_n|^{-1}Y_nA_n),$$

where  $\sigma_i$  is the  $i$ th singular value. The matrix  $|C_n|^{-\frac{1}{2}}Y_nA_n|C_n|^{-\frac{1}{2}}$  is symmetric and so its singular values are the absolute values of its eigenvalues. The condition number of  $|C_n|$  is equal to that of  $C_n$  (see (3.2)), which gives the result.  $\square$

Note that even when  $\kappa$  is large, the eigenvalues of  $|C_n|^{-1}Y_nA_n$  may be clustered. As an example, consider the Toeplitz matrix of dimension 1000 generated by the function  $f(x) = |x|e^{ix}$ ,  $x \in [-\pi, \pi]$ , and the Strang preconditioner. All but five eigenvalues  $\lambda$  of  $(C_n^{-1}A_n)^T(C_n^{-1}A_n)$  satisfy  $|1-\lambda| \leq 0.1$ , with the outliers equal to  $5.5 \times 10^{-8}$ , 0.27, 0.63, 0.88 and 978. Since the Strang preconditioner has condition number  $\kappa = 2.5 \times 10^3$  (see Table 6.6), we might expect the eigenvalues of  $|C_n|^{-1}Y_nA_n$  to be spread along the real line. However, all but five eigenvalues  $\lambda$  of  $|C_n|^{-1}Y_nA_n$  satisfy  $|1-\lambda| \leq 0.1$  or  $|1+\lambda| \leq 0.1$ , with the outliers at  $-1.22$ ,  $-0.70$ ,  $-5 \times 10^{-4}$ , 1.3 and 5.3. Note also that these outliers do not vary as much in magnitude as the outliers of  $(C_n^{-1}A_n)^T(C_n^{-1}A_n)$ , although they lie on both sides of the imaginary axis. Clearly then, Proposition 5 may be pessimistic. Additionally, in our experiments the performance of our MINRES method is typically better than that of LSQR (an implementation of CGNE) even when  $C_n$  is ill-conditioned.

We now discuss a selection of preconditioners for nonsymmetric Toeplitz matrices for which our MINRES method works well. Oseledets and Tyrtysnikov [41] show that (4.1) holds with  $\|E\| = 0$  and  $K$  independent of  $n$  for a wide class of Toeplitz matrices, namely matrices with generating function

$$f = P(z) + \frac{Q(z)}{L(z)}, \quad z = e^{ix},$$

where  $P$ ,  $Q$  and  $L$  are polynomials,  $L$  has no roots on the unit circle, the degree of  $Q$  is not greater than the degree of  $L$ , and  $L$  and  $Q$  have no common roots. For these problems, Proposition 3 shows that MINRES terminates after at most  $K$  iterations and, since Definition 4 holds, we can expect 2-step superlinear convergence. For Toeplitz matrices with generating function

$$f = g + \sum_{\alpha=0}^{\ell} \sum_{k=0}^m A_{k\alpha} (z - \psi_k)^\alpha \log(z - \psi_k), \quad z = e^{ix}, \quad |\psi_k| = 1,$$



where  $g$  is analytic in a disc containing  $|z| = 1$ , Proposition 2 holds for any  $\epsilon > 0$  with  $K = \log \epsilon^{-1}[c_0 + c_1 \log \epsilon^{-1} + c_2 \log n] + c_3$ , where  $c_0, c_1, c_2$  and  $c_3$  are independent of  $n$  and  $\epsilon$ . Thus, for this class of Toeplitz matrix, which includes all examples in papers on superlinear preconditioners, we can also expect 2-step superlinear convergence with our MINRES method [41, 56]. Additionally, Oseledets and Tyrtshnikov propose a method for constructing circulant preconditioners that satisfy (4.1) using only the elements of  $A_n$ , i.e., without reference to a generating function. GMRES is used to solve their preconditioned systems, but MINRES could be applied to the transformed systems instead.

For large enough dimension  $n$ , and for Toeplitz matrices with  $2\pi$ -periodic, continuous, complex-valued generating function  $f$  having no zeros, the optimal preconditioner proposed by Chan [15] satisfies (4.1) and has clustered spectrum [13, Corollary 1]. Thus, we can expect 2-step superlinear convergence. Additional properties of the optimal preconditioner can be found in the paper by Chan, Jin and Yeung [7]. Tyrtshnikov, Yeregin and Zamarashkin [53] consider improved circulants for Toeplitz matrices with a  $2\pi$ -Lebesgue-integrable complex-valued generating function  $f$  that is sparsely vanishing, i.e., for which

$$\int_{-\pi}^{\pi} \phi_{\epsilon}(|f(x)|) dx = 0,$$

where  $\phi_{\epsilon}$  is the characteristic function of the interval  $[0, \epsilon]$ . For these improved circulant preconditioners, the preconditioned Toeplitz matrix satisfies  $A_n C_n^{-1} = I + E$  and an analogous result to Proposition 2 holds. Additionally,  $\|E\|_F = o(n)$ , where  $\|\cdot\|_F$  is the Frobenius norm, so that  $A_n C_n^{-1}$  has a clustered spectrum and 2-step superlinear convergence is achieved.

Let  $C_n$  be the Strang preconditioner and let the entries of  $A_n$  be such that

$$\sum_{k=-\infty}^{\infty} |a_k| \leq \beta < \infty, \quad \left| \sum_{k=inf ty}^{\infty} a_k e^{-ik\theta} \right| \geq \gamma > 0,$$

for all  $\theta$  and for constants  $\alpha$  and  $\beta$ . Then  $(C_n^{-1} A_n)^T (C_n^{-1} A_n)$  has all eigenvalues clustered at 1 with the exception of a fixed number, independent of  $n$ . Thus, by Proposition 5, the modified Strang preconditioner  $|C_n|$  also gives clustered eigenvalues as long as  $|C_n|$  is well conditioned. Similar results were proved for a second circulant preconditioner that utilizes all elements of  $A_n$ .

**5. Extension to block matrices.** Multilevel Toeplitz matrices are a natural extension of Toeplitz matrices, with the simplest, separable matrices, formed from Kronecker products of Toeplitz matrices. Separable matrices arise in, for example, partial differential equations [16, 28], two-dimensional signal processing [6] and image deblurring [18]. Preconditioners for Toeplitz matrices have been used in Kronecker products to solve multilevel problems (see, e.g., [6, 18, 34]). Note, however, that unlike in the previous section we cannot expect the number of outlying eigenvalues, i.e., the eigenvalues not belonging to a cluster, to be independent of the dimension of the matrix [48] since for multilevel Toeplitz problems only multigrid methods show convergence rates independent of the dimension [1].

To get a feel for how the results of previous sections can be applied to separable matrices, and the differences between the Toeplitz and multilevel Toeplitz cases, let  $B_{nm} = S_n \otimes T_m$  where  $S_n \in \mathbb{R}^{n \times n}$  and  $T_m \in \mathbb{R}^{m \times m}$  are Toeplitz matrices and

$\otimes$  represents a Kronecker product. (Pertinent properties of Kronecker products are given in, for example, the book by Horn and Johnson [31, Section 4.2].) Then if  $W_{nm} = Y_n \otimes Y_m$ , the product

$$W_{nm}B_{nm} = (Y_n \otimes Y_m)(S_n \otimes T_m) = (Y_n S_n) \otimes (Y_m T_m).$$

Since the Kronecker product of symmetric matrices is symmetric,  $W_{nm}B_{nm}$  is symmetric. We can define a preconditioner  $G_{nm} = J_n \otimes K_m$  for  $B_{nm}$ , where  $J_n \in \mathbb{R}^{n \times n}$  is a circulant preconditioner for  $S_n$  and  $K_m \in \mathbb{R}^{m \times m}$  is a circulant preconditioner for  $T_m$ . Then, if  $|J_n|$  is the absolute value of  $J_n$  (defined analogously to  $|C_n|$ ) and  $|K_m|$  is the absolute value of  $K_m$ , we can define the symmetric positive definite preconditioner

$$|G_{nm}| = |J_n| \otimes |K_m|$$

for which

$$|G_{nm}|^{-1}W_{nm}B_{nm} = |J_n|^{-1}Y_n S_n \otimes |K_m|^{-1}Y_m T_m.$$

The eigenvalues of  $|G_{nm}|^{-1}W_{nm}B_{nm}$  are products of the eigenvalues of  $|J_n|^{-1}Y_n S_n$  and  $|K_m|^{-1}Y_m T_m$ . Specifically, if  $\mu_1, \dots, \mu_n$  are the eigenvalues of  $|J_n|^{-1}Y_n S_n$  and  $\nu_1, \dots, \nu_m$  are the eigenvalues of  $|K_m|^{-1}Y_m T_m$  then the eigenvalues  $\lambda_{1,1}, \dots, \lambda_{m,n}$  of  $|G_{nm}|^{-1}W_{nm}B_{nm}$  are  $\lambda_{i,j} = \mu_i \nu_j$ ,  $i = 1, \dots, n$ ,  $j = 1, \dots, m$ .

The fact that the spectrum of  $|G_{nm}|^{-1}W_{nm}B_{nm}$  comprises all products  $\mu_i \nu_j$  is an indicator that the eigenvalues of this preconditioned separable matrix may not be as ‘‘nice’’ as those of a preconditioned Toeplitz matrix. This is substantiated by the block version of Proposition 2 given below.

**PROPOSITION 6.** *Let  $S_n \in \mathbb{R}^{n \times n}$  and  $T_m \in \mathbb{R}^{m \times m}$  be nonsingular Toeplitz matrices and let  $J_n \in \mathbb{R}^{n \times n}$  and  $K_m \in \mathbb{R}^{m \times m}$  be circulants. Additionally, let*

$$J_n^{-1}S_n = I + R_1 + E_1 \text{ and } K_m^{-1}T_m = I + R_2 + E_2$$

where  $\text{rank}(R_1) = K_1$ ,  $\text{rank}(R_2) = K_2$ ,  $\|E_1\| \leq \epsilon_1$  and  $\|E_2\| \leq \epsilon_2$ , where  $\|\cdot\|$  is a unitarily invariant norm. Then,

$$(5.1) \quad G_{nm}^{-1}B_{nm} = I + R + E + F,$$

where  $\text{rank}(R) \leq K_1 + K_2 + K_1 K_2$ ,  $\|E\| \leq \epsilon_1 + \epsilon_2 + \epsilon_1 \epsilon_2$  and  $F = R_1 \otimes E_2 + R_2 \otimes E_1$ . Moreover,

$$(5.2) \quad |G_{nm}|^{-1}W_{nm}B_{nm} = Q_{mn} + \hat{R} + \hat{E} + \hat{F},$$

where  $Q_{mn}$  is orthogonal with eigenvalues  $\pm 1$ ,  $\text{rank}(\hat{R}) = \text{rank}(R)$ ,  $\|\hat{E}\| = \|E\|$  and  $\hat{F} = Q_{mn}F$ .

*Proof.* Using standard properties of Kronecker products, we find that

$$G_{nm}^{-1}B_{nm} = (I + R_1 + E_1) \otimes (I + R_2 + E_2) = I + R + E + F,$$

where  $R = R_1 \otimes I + I \otimes R_2 + R_1 \otimes R_2$ ,  $E = E_1 \otimes I + I \otimes E_2 + E_1 \otimes E_2$  and  $F$  is as defined in the proposition. The subadditivity property of both the rank and norm give that  $\text{rank}(R) \leq \text{rank}(R_1 \otimes I) + \text{rank}(I \otimes R_2) + \text{rank}(R_1 \otimes R_2)$  and  $\|E\| \leq \|E_1 \otimes I\| + \|I \otimes E_2\| + \|E_1 \otimes E_2\|$ . For any two matrices  $X$  and  $Y$ ,  $\text{rank}(X \otimes Y) = \text{rank}(X)\text{rank}(Y)$  and  $\|X \otimes Y\| = \|X\|\|Y\|$  which proves the first part.

Also, since  $|J_n|$  and  $|K_m|$  are symmetric and Toeplitz,  $Y_n|J_n|^{-1} = |J_n|^{-1}Y_n$  and  $Y_m|K_m|^{-1} = |K_m|^{-1}Y_m$ . It follows that

$$|G_{nm}|^{-1}W_{nm}B_{nm} = (Y_n\tilde{J}_n \otimes Y_m\tilde{K}_m)(J_n^{-1}S_n \otimes K_m^{-1}T_m),$$

where  $\tilde{J}_n$  and  $\tilde{K}_m$  are defined analogously to  $\tilde{C}_n$  in (3.4). The matrix  $Q_{mn} = (Y_n\tilde{J}_n \otimes Y_m\tilde{K}_m)$  is orthogonal and symmetric, so its eigenvalues are  $\pm 1$ . Thus (5.2) holds with  $\hat{R} = Q_{mn}R$ ,  $\hat{E} = Q_{mn}E$  and  $\hat{F} = Q_{mn}F$ . The rank and norm of a matrix are not affected by multiplication by a unitary matrix and so we have the second result.  $\square$

We see from this result that the rank of  $R$  may be larger than that of  $R_1$  or  $R_2$  and  $\|E\|$  may be larger than  $\|E_1\|$  or  $\|E_2\|$ . However, more problematic is the matrix  $F$  which need not be low-rank or have small norm. This is not surprising in light of the results of Serra Capizzano and Tyrtyshnikov [48]. Given these results we do not consider block Toeplitz matrices further, although there may be examples for which the approach described in this section is appropriate.

**6. Numerical results.** In this section we compare the proposed preconditioned MINRES method, with preconditioner  $|C_n|$ , with other solvers for linear systems with nonsymmetric Toeplitz matrices. Specifically, we apply right-preconditioned GMRES and left-preconditioned LSQR [43] (which can be considered as a good implementation of CGNE) with preconditioner  $C_n$  to (1.1) and SQMR to (3.1) with (possibly indefinite) preconditioner  $Y_nC_n$ . Note that without preconditioning the SQMR and MINRES methods are equivalent.

The Toeplitz matrices themselves have been chosen for their nonnormality. Our first example is a Jordan block with eigenvalue 1.1 while our second is the Grcar matrix

$$A_n = \begin{bmatrix} 1 & 1 & 1 & 1 & & & & \\ -1 & 1 & 1 & 1 & 1 & & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ & & -1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix},$$

which is well-conditioned but very nonnormal [51, Chapter 7]. The third is the tridiagonal Toeplitz matrix

$$A_n = \begin{bmatrix} 1 & 0.01 & & & & & \\ 1 & 1 & 0.01 & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & & 1 & 1 & 0.01 & \\ & & & & 1 & 1 & \end{bmatrix},$$

which was used by Liesen and Strakoš [37] as an example of a nonnormal tridiagonal Toeplitz matrix. Our final example is the dense matrix defined by the generating function  $f(x) = |x|e^{ix}$ ,  $x \in [-\pi, \pi]$ . The condition numbers of these matrices are given in Table 6.1.

In all cases, the right-hand side vector  $b$  is a random vector scaled to have unit length. Our stopping criteria are that the relative residual satisfies  $\|r_k\|_2/\|r_0\|_2 < 10^{-8}$  or that this residual reduction is not achieved in 1000 iterations. We denote the latter condition by ‘—’. We use the Strang [49] and optimal [15] preconditioners and additionally give numerical results also for so-called superoptimal preconditioning [52], although our theory does not (obviously) cover this last case. Note that

TABLE 6.1  
*Condition numbers of unpreconditioned matrices.*

$n$	10	100	1000
Jordan block	9.2	20.1	21.0
Grcar	2.9	3.6	3.6
Tridiagonal	14	207	$2.6 \times 10^6$
Dense	83	$1.3 \times 10^3$	$1.8 \times 10^4$

TABLE 6.2  
*Matrix-vector products for the Jordan block with eigenvalue 1.1.*

	n	GMRES	LSQR	MINRES	SQMR
Unpreconditioned	10	10	20	10	10
	100	98	200	100	100
	1000	183	372	368	368
Strang	10	3	6	4	2
	100	3	6	4	2
	1000	3	6	4	2
Optimal	10	9	14	10	8
	100	8	18	13	8
	1000	5	12	10	5
Superoptimal	10	10	18	10	9
	100	12	30	22	12
	1000	7	18	16	8

the superoptimal preconditioner can be expressed in terms of the optimal preconditioner [7].

Tables 6.2–6.5 show the numbers of matrix-vector products for the different preconditioners and iterative methods. (All methods require one matrix-vector product per iteration except LSQR, which requires two.) We see that for each of these problems the preconditioners are effective for all methods tested. However, theoretical bounds on convergence exist only for LSQR and MINRES, both of which minimize the preconditioned residual. LSQR requires more matrix-vector products than the proposed MINRES method for all problems and so it appears that the latter may be preferable in practice. It is clear that GMRES converges faster than MINRES for these examples; however, there are no theoretical guarantees of fast convergence and it is well known that convergence is not necessarily governed by eigenvalues [2, 27]. Moreover, GMRES requires long recurrences, in contrast to the proposed MINRES method, so that each iteration becomes more expensive. From Tables 6.4 and 6.5 we see that the superoptimal preconditioner is not as effective as the optimal or Strang preconditioners for the two ill-conditioned problems. This may be related to the lack of satisfaction of our theory, but it is also known in the symmetric case that when  $A_n$  is ill conditioned the superoptimal preconditioner may preserve small eigenvalues that hamper convergence [19, 50] (c.f. the eigenvalue distribution for the example presented after Proposition 5).

The condition numbers of the preconditioned matrices are given in Table 6.6. We see that the condition number of  $|C_n|^{-1/2} Y_n A_n |C_n|^{-1/2} = |C_n|^{-1/2} A_n |C_n|^{-1/2}$  is typically smaller than that of  $C_n^{-1} A_n$ . Moreover, the conditioning of  $|C_n|^{-1/2} A_n |C_n|^{-1/2}$  is relevant to the speed of convergence of the proposed MINRES method, whereas

TABLE 6.3  
Matrix-vector products for the Grcar matrix.

	n	GMRES	LSQR	MINRES	SQMR
Unpreconditioned	10	10	20	10	10
	100	79	64	60	60
	1000	300	66	66	66
Strang	10	6	18	10	5
	100	5	18	10	4
	1000	5	18	10	4
Optimal	10	11	20	10	10
	100	9	22	16	8
	1000	7	20	14	6
Superoptimal	10	11	20	10	10
	100	9	24	18	9
	1000	7	20	14	7

TABLE 6.4  
Matrix-vector products for the tridiagonal matrix.

	n	GMRES	LSQR	MINRES	SQMR
Unpreconditioned	10	10	20	10	10
	100	97	200	100	100
	1000	985	—	—	—
Strang	10	4	12	6	3
	100	4	12	6	3
	1000	4	12	6	3
Optimal	10	9	14	9	9
	100	10	20	13	11
	1000	11	32	18	11
Superoptimal	10	10	20	10	10
	100	18	48	29	17
	1000	30	106	57	30

TABLE 6.5  
Matrix-vector products for the dense matrix.

	n	GMRES	LSQR	MINRES	SQMR
Unpreconditioned	10	—	20	10	10
	100	—	—	—	—
	1000	923	—	—	—
Strang	10	8	20	9	7
	100	11	26	16	9
	1000	13	34	18	12
Optimal	10	10	20	10	10
	100	12	28	17	11
	1000	15	44	24	15
Superoptimal	10	10	20	10	10
	100	19	54	31	19
	1000	41	138	66	35

TABLE 6.6  
*Condition numbers of the Strang (S), optimal (O) and superoptimal (SO) circulant preconditioned matrices.*

		$\kappa(C_n)$			$\kappa(C_n^{-1}A_n)$			$\kappa( C_n ^{-\frac{1}{2}}A_n C_n ^{-\frac{1}{2}})$		
		10	100	1000	10	100	1000	10	100	1000
Jordan	S	21	21	21	8.5	6.6	6.6	3.6	3.5	3.5
	O	10	19	21	4.6	6.3	6.6	2.4	3.4	3.5
	SO	3.1	10	19	4.0	5.1	6.3	3.6	3.2	3.4
Grcar	S	3.2	2.6	3.6	2.5	2.7	2.7	2.2	2.2	2.2
	O	2.4	3.5	3.6	2.3	2.6	2.7	2.0	2.2	2.2
	SO	2.2	3.4	3.6	2.3	2.6	2.7	2.0	2.1	2.2
Tridiagonal	S	201	201	201	104	187	$1.3 \times 10^6$	10	9.0	$1.2 \times 10^5$
	O	21	$2.0 \times 10^4$	223	11	$1.7 \times 10^4$	$1.4 \times 10^6$	2.4	36	$1.1 \times 10^5$
	SO	2.6	486	31	9.0	$1.6 \times 10^4$	$6.3 \times 10^5$	8.6	$1.0 \times 10^4$	$2.5 \times 10^5$
Dense	S	23	246	$2.5 \times 10^3$	96	$3.7 \times 10^3$	$1.3 \times 10^5$	28	584	$1.1 \times 10^4$
	O	20	110	734	86	$1.2 \times 10^3$	$5.3 \times 10^4$	26	397	$6.4 \times 10^3$
	SO	3.4	13	43	46	734	$1.4 \times 10^4$	39	335	$4.1 \times 10^3$

convergence of GMRES or SQMR cannot be directly related to the conditioning of  $C_n^{-1}A_n$ . Since the condition number of  $C_n$  is small for the Jordan block and the Grcar matrix, Proposition 5 guarantees that the eigenvalues of  $|C_n|^{-1/2}A_n|C_n|^{-1/2}$  of the Strang preconditioned matrix are clustered, similarly to the eigenvalues of  $(C_n^{-1}A_n)^T(C_n^{-1}A_n)$  for the normal equations. The same comparison cannot be made in all cases for the tridiagonal or dense matrices, since  $C_n$  may be ill-conditioned, yet we find that convergence is still better faster than that of LSQR for these problems, as measured by the number of matrix-vector products needed to achieve the desired residual norm reduction. Additionally, the eigenvalue results following Proposition 5 indicate that the eigenvalue bounds may not reflect the actual eigenvalue distribution.

**7. Conclusions.** We have shown that linear systems with nonnormal Toeplitz coefficient matrices can be efficiently solved by a preconditioned MINRES method by a simple reordering of rows or columns and the application of a suitable circulant preconditioner. The resulting method minimizes the preconditioned residual in the Euclidean norm, uses short-term recurrences and its rate of convergence is bounded in terms of the eigenvalues of the preconditioned coefficient matrix. The preconditioner itself is no more difficult to apply than existing circulant preconditioners for Toeplitz matrices, and our theory shows that the preconditioned MINRES method typically has fast convergence.

Our numerical results demonstrate that when modified Strang, optimal and superoptimal preconditioners are used, the preconditioned MINRES method converges rapidly. Although convergence can be achieved with fewer matrix-vector products when GMRES and SQMR are applied, each of these methods has its drawbacks: GMRES requires long recurrences while SQMR does not minimize the residual. Moreover, for neither method do there exist nice convergence bounds like those for MINRES.

Since every real matrix is self-adjoint with respect to some bilinear form, in theory this idea of symmetrizing a matrix and applying MINRES can always be applied [45, 46]. In practice, the bilinear form is often difficult to obtain but there are certain matrices, such as the persymmetric matrices described above, or Hamiltonian matrices [26, page 215], for which the bilinear form is known and is easy to use, and in these cases the proposed MINRES method could be used in an analogous manner.

**Acknowledgements** The authors would like to thank the referees for their helpful

suggestions and references. We also thank David Titley-Peloquin for fruitful discussions and comments on an earlier version of this manuscript.

## REFERENCES

- [1] A. ARICÒ AND M. DONATELLI, *A V-cycle multigrid for multilevel matrix algebras: Proof of optimality*, Numer. Math., 105 (2007), pp. 511–547.
- [2] M. ARIOLI, V. PTÁK, AND Z. STRAKOŠ, *Krylov sequences of maximal length and convergence of GMRES*, BIT, 38 (1998), pp. 636–643.
- [3] STEPHEN L CAMPBELL, ILSE CF IPSEN, C TIM KELLEY, AND CARL D MEYER, *Gmres and the minimal polynomial*, BIT, 36 (1996), pp. 664–675.
- [4] R. H. CHAN, *Circulant preconditioners for Hermitian Toeplitz systems*, SIAM J. Matrix Anal. Appl., 10 (1989), pp. 542–550.
- [5] R. H. CHAN AND W.-K. CHING, *Toeplitz-circulant preconditioners for Toeplitz systems and their application to queuing networks with batch arrivals*, SIAM J. Sci. Comput., 17 (1996), pp. 762–772.
- [6] R. H. CHAN AND X.-Q. JIN, *A family of block preconditioners for block systems*, SIAM J. Sci. Stat. Comput., 13 (1992), pp. 1218–1235.
- [7] R. H. CHAN, X.-Q. JIN, AND M.-C. YEUNG, *The circulant operator in the Banach algebra of matrices*, Linear Algebra Appl., 149 (1991), pp. 41–53.
- [8] R. H. CHAN AND M. K. NG, *Conjugate gradient methods for Toeplitz systems*, SIAM Rev., 38 (1996), pp. 427–482.
- [9] R. H. CHAN, D. POTTS, AND G. STEIDL, *Preconditioners for nondefinite Hermitian Toeplitz systems*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 647–665.
- [10] ———, *Preconditioners for non-Hermitian Toeplitz systems*, Numer. Linear Algebra Appl., 8 (2001), pp. 83–98.
- [11] R. H. CHAN AND G. STRANG, *Toeplitz equations by conjugate gradients with circulant preconditioners*, SIAM J. Sci. and Stat. Comput., 10 (1989), pp. 104–119.
- [12] R. H. CHAN AND M.-C. YEUNG, *Circulant preconditioners constructed from kernels*, SIAM J. Numer. Anal., 29 (1992), pp. 1093–1103.
- [13] ———, *Circulant preconditioners for complex Toeplitz matrices*, SIAM J. Numer. Anal., 30 (1993), pp. 1193–1207.
- [14] R. H.-F. CHAN AND X.-Q. JIN, *An Introduction to Iterative Toeplitz Solvers*, SIAM, Philadelphia, PA, 2007.
- [15] T. F. CHAN, *An optimal circulant preconditioner for Toeplitz systems*, SIAM J. Sci. Stat. Comput., 9 (1988), pp. 766–771.
- [16] W. M. CHEUNG AND M. K. NG, *Block-circulant preconditioners for systems arising from discretization of the three-dimensional convection-diffusion equation*, J. Comput. Appl. Math., 140 (2002), pp. 143–158.
- [17] S.-C. T. CHOI, *Minimal residual methods for complex symmetric, skew symmetric, and skew Hermitian systems*, Tech. Report ANL/MCS-P3028-0812, Computation Institute, University of Chicago, 2013.
- [18] F. DI BENEDETTO, C. ESTATICO, AND S. SERRA-CAPIZZANO, *Superoptimal preconditioned conjugate gradient iteration for image deblurring*, SIAM J. Sci. Comput., 26 (2005), pp. 1012–1035.
- [19] F. DI BENEDETTO AND S. SERRA-CAPIZZANO, *Spectral analysis of the superoptimal matrix algebra operators*, Tech. Report 439, Dept. of Mathematics, Università di Genova, 2001.
- [20] S. C. EISENSTAT AND I. C. F. IPSEN, *Relative perturbation techniques for singular value problems*, SIAM J. Numer. Anal., 32 (1995), pp. 1972–1988.
- [21] H. ELMAN, D. SILVESTER, AND A. WATHEN, *Finite elements and fast iterative solvers with applications in incompressible fluid dynamics*, Oxford University Press, 2005.
- [22] V. FABER AND T. MANTEUFFEL, *Necessary and sufficient conditions for the existence of a conjugate gradient method*, SIAM J. Numer. Anal., 21 (1984), pp. 352–362.
- [23] B. FISCHER, A. RAMAGE, D. J. SILVESTER, AND A. J. WATHEN, *Minimum residual methods for augmented systems*, BIT, 38 (1998), pp. 527–543.
- [24] R. W. FREUND AND N. M. NACHTIGAL, *QMR: A quasi-minimal residual method for non-Hermitian linear systems*, Numer. Math., 60 (1991), pp. 315–339.
- [25] R. W. FREUND AND N. M. NACHTIGAL, *Software for simplified Lanczos and QMR algorithms*, Appl. Numer. Math., 19 (1995), pp. 319–341.
- [26] I. GOHBERG, P. LANCASTER, AND L. RODMAN, *Indefinite Linear Algebra and Applications*, Birkhäuser Verlag, Basel, 2005.

- [27] A. GREENBAUM AND Z. STRAKOŠ, *Matrices that generate the same Krylov residual spaces*, in Recent Advances in Iterative Methods, G. Golub, A. Greenbaum, and M. Luskin, eds., vol. 60 of The IMA Volumes in Mathematics and its Applications, Springer, New York, NY, 1994.
- [28] L. HEMMINGSSON, *A semi-circulant preconditioner for the convection-diffusion equation*, Numer. Math., 81 (1998), pp. 211–248.
- [29] M. R. HESTENES AND E. STIEFEL, *Methods of conjugate gradients for solving linear systems*, J. Res. Nat. Bur. Standards, 49 (1952), pp. 409–436.
- [30] N. J. HIGHAM, *Functions of Matrices*, SIAM, Philadelphia, PA, USA, 2008.
- [31] R. A. HORN AND C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, 1991.
- [32] T. HUCKLE, *Circulant and skewcirculant matrices for solving Toeplitz matrix problems*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 767–777.
- [33] T. HUCKLE, S. SERRA-CAPIZZANO, AND C. TABLINO-POSSIO, *Preconditioning strategies for non-Hermitian Toeplitz linear systems*, Numer. Linear Algebra Appl., 12 (2005), pp. 211–220.
- [34] T. KU AND C.-C. J. KUO, *On the spectrum of a family of preconditioned block Toeplitz matrices*, SIAM J. Sci. and Stat. Comput., 13 (1992), pp. 948–966.
- [35] T.-K. KU AND C.-C. J. KUO, *Design and analysis of Toeplitz preconditioners*, IEEE T. Signal Process., 40 (1992), pp. 129–141.
- [36] ———, *Spectral properties of preconditioned rational Toeplitz matrices: the nonsymmetric case*, SIAM J. Mat. Anal. Appl., 14 (1993), pp. 521–544.
- [37] J. LIESEN AND Z. STRAKOŠ, *Convergence of GMRES for tridiagonal Toeplitz matrices*, SIAM J. Matrix Anal. Appl., 26 (2004), pp. 233–251.
- [38] M. K. NG, *Iterative Methods for Toeplitz Systems*, Oxford University Press, 2004.
- [39] M. K. NG AND D. POTTS, *Circulant preconditioners for indefinite Toeplitz systems*, BIT, 41 (2001), pp. 1079–1088.
- [40] J. OLKIN, *Linear and Nonlinear Deconvolution Problems*, PhD thesis, Rice University, Houston, TX, 1986.
- [41] I. OSELEDETS AND E. TYRTYSHNIKOV, *A unifying approach to the construction of circulant preconditioners*, Linear Algebra Appl., 418 (2006), pp. 435–449.
- [42] C. C. PAIGE AND M. A. SAUNDERS, *Solution of sparse indefinite systems of linear equations*, SIAM J. Numer. Anal., 12 (1975), pp. 617–629.
- [43] ———, *LSQR: An algorithm for sparse linear equations and sparse least squares*, ACM T. Math. Software, 8 (1982), pp. 43–71.
- [44] H.-K. PANG, Y.-Y. ZHANG, AND X.-Q. JIN, *Tri-diagonal preconditioner for pricing options*, J. Comput. Appl. Math., 236 (2012), pp. 4365–4374.
- [45] J. PESTANA, *Nonstandard Inner Products and Preconditioned Iterative Methods*, PhD thesis, University of Oxford, 2011.
- [46] J. PESTANA AND A. J. WATHEN, *On choice of preconditioner for minimum residual methods for non-Hermitian matrices*, J. Comput. Appl. Math., 249 (2013), pp. 57–68.
- [47] Y. SAAD AND M. H. SCHULTZ, *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986), pp. 856–869.
- [48] S. SERRA CAPIZZANO AND E. TYRTYSHNIKOV, *Any circulant-like preconditioner for multilevel matrices is not superlinear*, SIAM J. Mat. Anal. Appl., 21 (1999), pp. 431–439.
- [49] G. STRANG, *A proposal for Toeplitz matrix calculations*, Stud. Appl. Math., 74 (1986), pp. 171–176.
- [50] V. V. STRELA AND E. E. TYRTYSHNIKOV, *Which circulant preconditioner is better?*, Math. Comput., 65 (1996), pp. 137–150.
- [51] L. N. TREFETHEN AND M. EMBREE, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, Princeton, 2005.
- [52] E. E. TYRTYSHNIKOV, *Optimal and superoptimal circulant preconditioners*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 459–473.
- [53] E. E. TYRTYSHNIKOV AND A. YU. YEREMIN AND N. L. ZAMARASHKIN, *Clusters, preconditioners, convergence*, Linear Algebra Appl., 263 (1997), pp. 25–48.
- [54] H. A. VAN DER VORST, *Bi-CGSTAB: A fast and smoothly converging variant of BiCG for the solution of nonsymmetric linear systems*, SIAM J. Sci. Statist. Comput., 13 (1992), pp. 631–644.
- [55] C. VAN LOAN, *Computational Frameworks for the Fast Fourier Transform*, SIAM, Philadelphia, PA, 1992.
- [56] N. L. ZAMARASHKIN, I. V. OSELEDETS, AND E. E. TYRTYSHNIKOV, *Approximation of Toeplitz matrices by sums of circulants and small-rank matrices*, Dokl. Math., 73 (2006), pp. 100–101.