Applied Mathematics and Computation 468 (2024) 128473

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Full Length Article

Circulant decomposition of a matrix and the eigenvalues of Toeplitz type matrices

Hariprasad M., Murugesan Venkatapathi *

Department of Computational & Data Sciences, Indian Institute of Science, Bangalore, India

ARTICLE INFO

MSC:
15A04
15B05
65F08
65F15

Keywords:
Sparse approximation
Similarity transformation
Periodic entries
Fast Fourier transform
Preconditioners

ABSTRACT

We begin by showing that any $n \times n$ matrix can be decomposed into a sum of n circulant matrices with periodic relaxations on the unit circle. This decomposition is orthogonal with respect to a Frobenius inner product, allowing recursive iterations for these circulant components. It is also shown that the dominance of a few circulant components in the matrix allows sparse similarity transformations using Fast-Fourier-transform (FFT) operations. This enables the evaluation of all eigenvalues of *dense* Toeplitz, block-Toeplitz, and other periodic or quasi-periodic matrices, to a reasonable approximation in $\mathcal{O}(n^2)$ arithmetic operations. The utility of the approximate similarity transformation in preconditioning linear solvers is also demonstrated.

1. Introduction

Matrix decompositions, low-rank approximations, and projections on low complexity classes are some of the approaches useful in reducing the computation incurred with dense matrices [1]. Other approaches are also available in the case of sparse and effectively-sparse matrices, for example, in efficiently and accurately evaluating eigenvalues [2–7]. Dense matrices that occur frequently in the numerical solution of eigenvalue problems such as Toeplitz, block-Toeplitz, and other matrices with periodicity in the diagonals are a class where further reduced computing may be possible. An efficient decomposition of the given matrix into circulant components i.e. circulant matrices with periodic relaxations on the unit circle, is shown to allow drastic reduction in such computations at a relatively small cost in the accuracy.

$\mathcal{O}(n^2)$ algorithms to evaluate the characteristic polynomial and its derivative were proposed a few decades ago for Toeplitz [8], block-Toeplitz [9,10] and Hankel matrices [11]. These algorithms were amalgamated with the Newton methods to evaluate only the eigenvalues of interest. On the other hand, the asymptotic behavior of spectra of Toeplitz and block-Toeplitz operators in the limit of large dimensions are well studied [12–16] and they can also be approximated by appropriate circulant matrices. Matrix-less methods scaling as $\mathcal{O}(n)$ have been proposed for evaluating eigenvalues of certain classes of Toeplitz matrices using asymptotic expansions [17].

An approximation of a finite-dimensional Toeplitz matrix using a circulant matrix for speed up of operations was suggested decades ago [18,19]. Similarly, circulant preconditioners were constructed for Toeplitz matrices by minimizing the Frobenius norm of the residue [20]. Spectral preserving properties of this preconditioner were shown, and block circulant versions were also proposed

* Corresponding author.
E-mail addresses: mhariprasadkansur@gmail.com (M. Hariprasad), murugesh@iisc.ac.in (M. Venkatapathi).

https://doi.org/10.1016/j.amc.2023.128473
Received 28 September 2023; Accepted 21 November 2023

Available online 28 December 2023
0096-3003/

and analyzed [21]. They are expected to speed up iterative methods in regularization and optimization even though they can fail in special cases of poor approximation, and they may not be sufficiently accurate as a similarity transformation of the given Toeplitz matrix. This preconditioner based on minimizing the Frobenius norm of residue, is shown to be the first term in the circulant decomposition of the matrix discussed here. Using a single circulant component of a Toeplitz matrix for estimating all its eigenvalues results in a $\mathcal{O}(n \log n)$ algorithm. The relation between this single-term circulant approximation of a Toeplitz matrix, and its *symbol*, is highlighted in the appendix. The approach of approximating the Toeplitz matrix by a circulant matrix minimizing the norm of the residue, was also extended to a generalized circulant preconditioner where only the absolute values of the entries preserve the periodicity [22]. The full circulant decomposition of a matrix presented here enhances and broadens the scope of such preconditioners and approximate similarity transformations, to matrices with any periodicity along the diagonals. Such matrices are relevant to signal and image processing, solving differential and integral equations, applications of queuing theory, and polynomial computations [23,24].

We first recall in Remark 1.1 that any matrix can be decomposed into n cycles that generate its $2n$-1 diagonals. We later show in Lemma 1.3 and Theorem 1.4 that the decomposition of a matrix into n circulant matrices with periodic relaxations on the unit circle, is equivalent to a decomposition of a similar matrix into such cycles. By including only the dominant cycles of the similar matrix, we include the dominant circulant components of the given matrix. The sparse similar matrix can be operated by a non-symmetric Lanczos algorithm for (block) tridiagonalization, and one can evaluate all eigenvalues of such dense matrices in $\mathcal{O}(n^2)$ arithmetic operations. Other relevant approaches for eigenvalues of sparse matrices have been reported as well [2].

Let I_n be the identity matrix of dimension n, and C be the permutation matrix corresponding to a full cycle.

$$C = \begin{bmatrix} 0 & 1 \\ I_{n-1} & 0 \end{bmatrix}_{n\times n}.$$

Remark 1.1. Any matrix A can be decomposed into n cycles given by a power series in C such that $A = \sum_{k=0}^{n-1} \Lambda_k C^k$, where the Hadamard product $A \circ C^k = \Lambda_k C^k$, and Λ_k are diagonal matrices. Entries supported on C^k i.e. diagonal entries of Λ_k in the above decomposition, are referred as the k^{th} cycle of the matrix A.

Example 1.1. The decomposition of a matrix into cycles, $A = \sum_{j=0}^{n-1} \Lambda_j C^j$ for an order 3 magic square.

$$\begin{bmatrix} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{bmatrix} = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix} C^0 + \begin{bmatrix} 6 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 9 \end{bmatrix} C^1 + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 4 \end{bmatrix} C^2$$

Let the permutation matrix given by a flipped identity matrix be

$$J = \begin{bmatrix} 0 & \cdots & 0 & 0 & 1 \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & 1 & 0 & 0 \\ \vdots & \iddots & 0 & 0 & 0 \\ 1 & \cdots & 0 & 0 & 0 \end{bmatrix}_{n\times n}.$$

Let i denote $\sqrt{-1}$. For a sequence $\{x(p)\}_0^{n-1}$, its Discrete-Fourier-Transform (DFT) is given by $\{X(k)\}_0^{n-1}$, with $X(k) = \sum_{p=0}^{n-1} x(p) e^{\mathrm{i}\frac{2\pi pk}{n}}$. Eigenvalues of a circulant matrix R are given by the DFT of the first row $\{R(0,j)\}_{j=0}^{n-1}$. The corresponding eigenvectors are given by the columns of the matrix W with $W(p,q) = \frac{1}{\sqrt{n}} e^{-\mathrm{i}\frac{2\pi pq}{n}}$, with $p,q = 0,1,2\cdots n-1$.

Remark 1.2. $W^2 = CJ$, and any circulant matrix has an eigen decomposition $R = W\Lambda W^*$. Here W^* is the conjugate transpose of the matrix W. R is also given by $R = W^*\tilde{\Lambda}W$ where $\tilde{\Lambda} = CJ\Lambda J^T C^T$. Thus for $1 \le k \le n-1$, we have $\tilde{\Lambda}(k,k) = \Lambda(n-k, n-k)$ and $\Lambda(0,0) = \tilde{\Lambda}(0,0)$.

Lemma 1.3. *Given diagonal matrices D_k with $D_k(q,q) = e^{\mathrm{i}\frac{2\pi kq}{n}}$ (for $0 \le q \le n-1$) and any circulant matrix R with eigenvalues given by a diagonal matrix $\tilde{\Lambda}$, the matrices RD_k and $\tilde{\Lambda}C^k$ are similar.*

Proof. Using W for a linear transformation of RD_k,

$$\begin{aligned} WRD_kW^* &= WW^*\tilde{\Lambda}WD_kW^* \\ &= \tilde{\Lambda}WD_kW^* \end{aligned}$$

$$= \tilde{\Lambda} C^k.$$

The substitution $WD_kW^* = C^k$, can be deduced by evaluating its $(p,l)^{\text{th}}$ entry.

$$\begin{aligned}(p,l)^{\text{th}} \text{ entry of } WD_kW^* &= \frac{1}{n}\sum_{q=0}^{n-1} e^{-\mathrm{i}\frac{2\pi pq}{n}} e^{\mathrm{i}\frac{2\pi kq}{n}} e^{\mathrm{i}\frac{2\pi ql}{n}} \\ &= \frac{1}{n}\sum_{q=0}^{n-1} e^{\mathrm{i}\frac{2\pi(-p+k+l)q}{n}} \\ &= 1 \text{ when } p \equiv l+k \mod n, \text{ and 0 otherwise.}\end{aligned}$$

Hence, the eigenvalues of RD_k are the eigenvalues of the matrix $\tilde{\Lambda} C^k$.

Note that the matrix $\tilde{\Lambda} C^k$ is sparse and represents a single cycle, while the similar matrix RD_k is dense. Remark 1.1 recalls that any matrix has a decomposition into such cycles. The cycle decomposition of a matrix A is equivalent to a circulant decomposition of a transformed similar matrix WAW^*. Conversely, a circulant decomposition of the given matrix A is equivalent to a cycle decomposition of the transformed similar matrix WAW^*, as presented in the theorem below.

Theorem 1.4. *Any $n \times n$ square matrix A can be represented as a sum of n circulant matrices with periodic relaxations taking values from the n^{th} root of unity. It is of the form $A = \sum_{k=0}^{n-1} R_k D_k$ where R_k is a circulant matrix and D_k is a diagonal matrix with $D_k(q,q) = e^{\mathrm{i}\frac{2\pi kq}{n}}$ (with $0 \le q \le n-1$).*

Proof. Consider a matrix $B = WAW^*$. Recalling its decomposition into cycles, and using the substitution $C^k = WD_kW^*$ from the proof of Lemma 1.3, we have $B = \sum_{k=0}^{n-1} \tilde{\Lambda}_k C^k = \sum_{k=0}^{n-1} \tilde{\Lambda}_k W D_k W^*$, where $\tilde{\Lambda}_k$ are diagonal matrices. Thus, the original matrix A is given by:

$$A = W^* B W = \sum_{k=0}^{n-1} W^* \tilde{\Lambda}_k W D_k = \sum_{k=0}^{n-1} R_k D_k. \tag{1.1}$$

Example 1.2. The decomposition of a matrix into circulant components, $A = \sum_{k=0}^{n-1} R_k D_k$ for an order 3 magic square.

$$\begin{aligned}\begin{bmatrix} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{bmatrix} &= \begin{bmatrix} 5 & 4 & 6 \\ 6 & 5 & 4 \\ 4 & 6 & 5 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &+ \begin{bmatrix} 1.5-0.86\mathrm{i} & 1.73\mathrm{i} & -1.5-0.86\mathrm{i} \\ -1.5-0.86\mathrm{i} & 1.5-0.86\mathrm{i} & 1.73\mathrm{i} \\ 1.73\mathrm{i} & -1.5-0.86\mathrm{i} & 1.5-0.86\mathrm{i} \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{\mathrm{i}\frac{2\pi}{3}} & 0 \\ 0 & 0 & e^{\mathrm{i}\frac{4\pi}{3}} \end{bmatrix} \\ &+ \begin{bmatrix} 1.5+0.86\mathrm{i} & -1.73\mathrm{i} & -1.5+0.86\mathrm{i} \\ -1.5+0.86\mathrm{i} & 1.5+0.86\mathrm{i} & -1.73\mathrm{i} \\ -1.73\mathrm{i} & -1.5+0.86\mathrm{i} & 1.5+0.86\mathrm{i} \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{\mathrm{i}\frac{4\pi}{3}} & 0 \\ 0 & 0 & e^{\mathrm{i}\frac{2\pi}{3}} \end{bmatrix}.\end{aligned}$$

Frobenius inner product: Let $\langle A, B\rangle_F = \sum_{i,j} \overline{A(i,j)} B(i,j)$ denote the Frobenius inner product of matrices A and B. Note that any unitary transformation of the matrices preserves this inner product. It is also shown below that the circulant decomposition in Theorem 1.4 is an orthogonal decomposition with respect to $\langle .,. \rangle_F$.

Lemma 1.5. *For any unitary matrices U, V and matrices A, B, $\langle A, B\rangle_F = \langle UAV, UBV\rangle_F$.*

Proof. Let $\widehat{\ }$ denote a vector formed by concatenating the columns of a matrix such as

$$\widehat{A} = \begin{bmatrix} A(:,1) \\ A(:,2) \\ \vdots \\ A(:,n) \end{bmatrix}.$$

Then $\langle A, B\rangle_F = \widehat{A}^* \widehat{B}$. Let [] denote the expansion into a matrix of diagonal blocks, such as

$$[U] = \begin{bmatrix} U & 0 & \cdots & 0 \\ 0 & U & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & U \end{bmatrix}_{n^2 \times n^2}.$$

Let P be the permutation such that $P\widehat{A} = \widehat{A^T}$, and $G = UAV,\ H = UBV$. Then

$$\widehat{G} = [V]P[U]\widehat{A} = T\widehat{A},$$

where $T = [V]P[U]$ is also a unitary matrix. Thus we have

$$\begin{aligned} \langle UAV, UBV \rangle_F = \langle G, H \rangle_F &= \widehat{A}^* T^* T \widehat{B}, \\ &= \widehat{A}^* \widehat{B}, \\ &= \langle A, B \rangle_F. \end{aligned}$$

Theorem 1.6. *In the decomposition,* $A = \sum_{k=0}^{n-1} R_k D_k$, *we have* $\langle D_i, D_j \rangle_F = n\delta_{i,j}$, *(with* $\delta_{i,j} = 1$ *if* $i = j$ *and zero otherwise). Also, for* $i \neq j$, $\langle R_i D_i, R_j D_j \rangle_F = 0$.

Proof. It is easy to verify $\langle D_i, D_j \rangle_F = n\delta_{i,j}$. From the transformation $A \to WAW^*$, and by Lemma 1.5 we can show that,

$$\begin{aligned} \langle R_i D_i, R_j D_j \rangle_F &= \langle W R_i D_i W^*, W R_j D_j W^* \rangle_F, \\ &= \langle C^i \Lambda_i, C^j \Lambda_j \rangle_F, \\ &= 0 \text{ for } i \neq j. \end{aligned}$$

The Theorem 1.6 implies a Gram-Schmidt orthogonalization type of procedure to obtain the individual circulant matrices in the decomposition of Theorem 1.4.

Remark 1.7. Recursive iterations for circulant components: The following procedure gives the individual circulant components for a given matrix A.

- Initialize $A_0 = A$,
- for $k = 0, 1, 2, \cdots, n-2$
  - $R_k(j, 0) = \frac{1}{n}\left(\mathbf{1}^T (A_k \circ C^j)\mathbf{1}\right)$ for $j = 0, 1, \cdots n-1$, and $\mathbf{1}$ is a n-vector with entries all ones i.e. Average the entries of A_k in the corresponding cycles to find the n unknown entries of circulant R_k.
  - $A_{k+1} = (A_k - R_k)D_{-1}$.
- $R_{n-1} = A_{n-1}$

The circulant matrix component R_k of A can also be evaluated using an inverse transformation of a cycle of WAW^*, given by $W^*(WAW^* \circ C^k)W$. Note that R_0 minimizes $\|A - R\|_F$ for any circulant matrix R.

Lemma 1.8. *Given* Λ *with diagonal entries* $\{\lambda_j\}_{j=0}^{n-1}$, *let* $R_1 = W\Lambda W^*$ *and* $R_2 = W^*\Lambda W$. *Then* $|R_1(0,0)| = |R_2(0,0)|$ *and* $|R_1(0,j)| = |R_2(0, n-j)|$ *for* $1 \leq j \leq n-1$.

Proof. $R_1(0, j)$ is the jth coefficient of the inverse discrete Fourier transform of the diagonals of Λ. We have,

$$|R_1(0,j)| = \left| \sum_{k=0}^{n-1} \lambda_k e^{\mathrm{i}2\pi \frac{kj}{n}} \right| = \left| \sum_{k=0}^{n-1} \lambda_k e^{-\mathrm{i}2\pi \frac{k(-j)}{n}} \right| \tag{1.2}$$

$$= \left| \sum_{k=0}^{n-1} \lambda_k e^{-\mathrm{i}2\pi \frac{k(n-j)}{n}} \right|. \tag{1.3}$$

Note that the first row and first column of the matrices correspond to index zero of the discrete Fourier transform. So the claim follows.

We proceed further to derive the relationship between the frequencies in the varying entries along the matrix diagonals, and the cycles of WAW^*. The following definitions are relevant for this exercise.

Weight of a cycle - w_i**:** Let $B = \sum_{i=0}^{n-1} R_i D_i$. The relative weight of circulant component R_i in B is $w_i = \frac{\|R_i\|_F^2}{\sum_j \|R_j\|_F^2}$. When $B = WAW^*$, note that w_i also represents the relative weight of the cycle i in A, and $\sum_i w_i = 1$, $0 \le w_i \le 1$.

Partial energy of a set of frequencies - E_i**:** Let $A = \sum_{i=0}^{n-1} \Lambda_i C^i$, and the discrete Fourier transform of the diagonal entries $\{\Lambda_i(j,j)\}_{j=0}^{n-1}$ be γ_j^i. Let $S_k = \{a_1, a_2, \cdots a_k\}$ be a set of indices, with $\frac{\sum_{j=1}^{k}(|\gamma_{a_j}^i|)^2}{\sum_{j=1}^{n}(|\gamma_j^i|)^2} = E_i$, where $0 \le E_i \le 1$. We say that the frequencies given by S_k have a partial energy E_i on the cycle i, where energy refers to the square of the magnitude.

Also note that δ_j^i, the inverse discrete Fourier transform of the diagonal entries, have the corresponding index set $T_k = \{b_1, b_2, \cdots b_k\}$ given by Lemma 1.8 ($b_j = a_j$ if $a_j = 0$, else $b_j = n - a_j$), such that $\frac{\sum_{j=1}^{k}(|\delta_{b_j}^i|)^2}{\sum_{j=1}^{n}(|\delta_j^i|)^2} = E_i$.

Relative magnitude of a set of cycles - s**:** For a matrix $B = \sum_{i=0}^{n-1} \Lambda_i C^i$, when the cycles corresponding to indices $J_k = \{j_1, j_2, \cdots j_k\}$ have $\frac{\sum_{j \in J_k} \|\Lambda_j\|_F^2}{\|B\|_F^2} = s$, we say that the cycles J_k have a relative magnitude s.

Theorem 1.9. *When E_i is the partial energy in a set of frequencies indexed by S_k on the cycle i of matrix A, the cycles of matrix $B = WAW^*$ indexed by the corresponding T_k have a relative magnitude $s = \sum w_i E_i$ where $\sum w_i = 1$ and $0 \le w_i \le 1$.*

Proof. Consider the decomposition,

$$A = \sum_{i=0}^{n-1} \Lambda_i C^i \text{ and } B = W(\sum_{i=0}^{n-1} \Lambda_i C^i) W^* = \sum_{i=0}^{n-1} R_i D_i.$$

By Parsevals theorem, $\|B\|_F^2 = \sum_i \|R_i\|_F^2$, and in the decomposition $B = \sum_{j=0}^{n-1} \tilde{\Lambda}_j C^j$, cyclic diagonal entries $\tilde{\Lambda}_j$ are given by the discrete Fourier transform of the sequence $\{R_i(0,j)\}_{i=0}^{n-1}$ (and scaled by $\sqrt{n}$). Considering the ratio of entries corresponding to the cycles indexed by T_k,

$$\frac{\sum_{j\in T_k} \|\tilde{\Lambda}_j\|_F^2}{\|B\|_F^2} = n \frac{\sum_{j\in T_k} \sum_{i=0}^{n-1} |R_i(0,j)|^2}{\sum_i \|R_i\|_F^2}, \qquad \text{using Parseval's theorem} \tag{1.4}$$

$$= n \frac{\sum_{i=0}^{n-1} \sum_{j\in T_k} |R_i(0,j)|^2}{\sum_i \|R_i\|_F^2}, \qquad \text{changing the order of summation} \tag{1.5}$$

$$= n \frac{\sum_{i=0}^{n-1} \frac{\|R_i\|_F^2}{n} E_i}{\sum_i \|R_i\|_F^2}, \qquad \text{by definition of } E_i \tag{1.6}$$

$$= \sum_{i=0}^{n-1} \left(\frac{\|R_i\|_F^2}{\sum_i \|R_i\|_F^2} \right) E_i. \tag{1.7}$$

Using $w_i = \frac{\|R_i\|_F^2}{\sum_j \|R_j\|_F^2}$, we get $s = \sum_{i=0}^{n-1} w_i E_i$.

Note that $s \to 1$ as $E_i \to 1$. Thus, when a matrix A has only $k \ll n$ dominant frequencies in the variation of entries along its diagonals, the similar matrix $B = WAW^*$ is effectively sparse and has only k dominant cycles, with the Frobenius norm of the other $n - k$ cycles being negligible.

Corollary 1.10. *For a Toeplitz matrix*

$$A=\begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{-1} & a_0 & a_1 & a_2 & \ddots \\ a_{-2} & a_{-1} & a_0 & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ a_{-(n-1)} & a_{-(n-2)} & \cdots & \cdots & a_0 \end{bmatrix},$$

by Theorem 1.9, the relative magnitude of the diagonal of $B = WAW^*$ is $s^0 = \sum_i w_i E_i^0$, where E_i^0 includes only the zero-frequency or the average value of entries in the cycle i.

$$s^0 = \sum_i \frac{(n-i)|a_{-i}|^2 + i|a_{n-i}|^2}{\|A\|_F^2} \frac{|(n-i)a_{-i} + ia_{n-i}|^2}{n((n-i)|a_{-i}|^2 + i|a_{n-i}|^2)}, \tag{1.8}$$

$$= \frac{\sum_i |(n-i)a_{-i} + ia_{n-i}|^2}{n\|A\|_F^2}. \tag{1.9}$$

In the above (1.9) note that for the cases $a_{-i} = a_{n-i}$, representing the circulant matrices, $s^0 = 1$ showing that only the diagonal of the corresponding B has non-zero entries. Minimizing the first circulant component s^0 for a Toeplitz matrix using special pathological cases such as $a_{-i} = -ia_i/(n-i)$ where $E_i^0 = 0$ for all $i > 0$, the lower bound on s^0 is

$$s^0 \geq \frac{n|a_0|^2}{\|A\|_F^2}.$$

The above lower bound shows that the circulant matrix R directly minimizing $\|A - R\|_F$ need not be an effective approximation for a Toeplitz matrix in general, even in the limit of large n. Using expressions such as (1.9) for the other partial energies, one can show that when the entries of the Toeplitz matrix $\{a_{-(n-1)}, a_{-(n-2)}, \cdots a_0, \cdots a_{(n-2)}, a_{(n-1)}\}$ are randomly chosen, partial energies E_i^k and E_i^{n-k} for frequencies $k \ll n$ are dominant corresponding to a small set of cycles in B. In other cases, depending on the set of given $2n-1$ entries, only the dominant cycles (representing the dominant partial energies) can be included in a sparse approximation of B with a negligible residue.

Corollary 1.11. *Given the set of entries $\{a_{-(n-1)}, a_{-(n-2)}, \cdots a_0, \cdots a_{(n-2)}, a_{(n-1)}\}$ of a Toeplitz matrix, the distribution of the partial energies E_i^k in the different frequencies indexed by $k = 1, 2, \dots n-1$, for a cycle i is given by:*

$$E_i^k = \frac{|a_{-i} - a_{n-i}|^2}{n((n-i)|a_{-i}|^2 + i|a_{n-i}|^2)} \left| \frac{\sin \frac{\pi(n-i+1)k}{n}}{\sin \frac{\pi k}{n}} \right|^2. \tag{1.10}$$

Corollary 1.12. *For a block Toeplitz matrix, we have constant matrices of a block size m replacing the scalar entries a_i and a_{-i} in the Toeplitz matrix. Here, we have m frequencies of interest given by the indices $S_m = \{n/m, 2n/m, 3n/m, \cdots n\}$. Correspondingly, we have the cycles of interest indexed by $T_m = \{n(m-1)/m, n(m-2)/m, \cdots 0\}$ in $B = WAW^*$.*

2. Approximating eigenvalues using the circulant decomposition

When a matrix A has k dominant frequencies in the variation of entries along its diagonals, the similar matrix $B = WAW^*$ is effectively sparse and has k dominant cycles, with the Frobenius norm of the other $n-k$ cycles being negligible (see Theorem 1.9). We can approximate B by a sparse $\tilde{B} = \sum_{i=1}^{k} \Lambda_{a_i} C^{a_i}$ by choosing the k dominant cycles in $a_i \in \{0, 1, 2 \cdots n-1\}$. This is followed by the application of an algorithm suited for the eigenvalues of a sparse matrix.

2.1. Identifying the dominant circulant components

The similarity transformation by the matrix W is shown to restrict the larger magnitudes to certain cycles for the Toeplitz and block Toeplitz matrices (Fig. 1). However, the similarity transform fails to show such behavior for random matrices with no restriction on the entries (Fig. 2), as expected.

Toeplitz matrices A, for example, have constant entries along the diagonals and the similarity transformation typically produces a very large magnitude for the cycle that represents a zero frequency (the main diagonal of the similar matrix B), and a few other cycles representing a low frequency of variations (see Corollary 1.11). On the other hand, a block Toeplitz matrix A with a block size m is m-periodic. Its similar matrix B has dominant cycles given by integers n/m and its multiples (see Corollary 1.12). Similarly, in the case of a quasi-periodic matrix where we have a random variation of k frequencies along diagonals, the dominant cycles resemble a mixture of block-Toeplitz matrices of k different block sizes. This is illustrated with examples in Section 3.

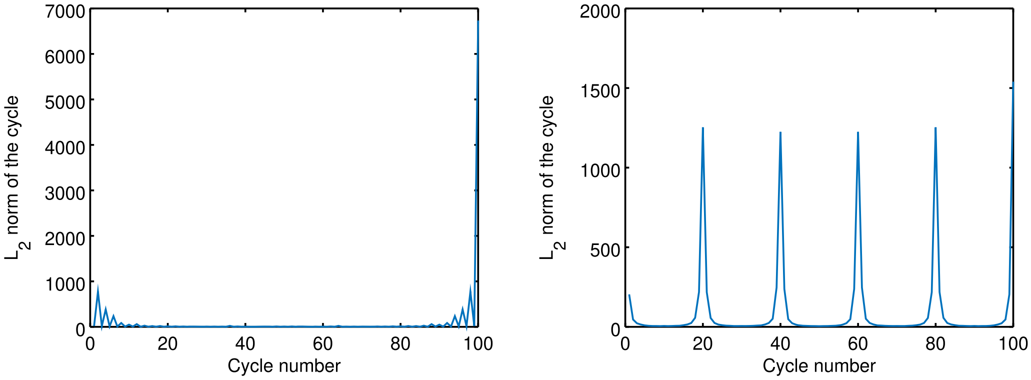

Fig. 1. L_2 norm of cycles of WAW^* numbered 0 to $n-1$ where $n = 100$. The dominant cycle numbering zero given by the diagonal is folded as number 100 in the plot. *Left*: A is a Toeplitz matrix with randomly chosen entries *Right*: A is a random block Toeplitz matrix of block-size 5.

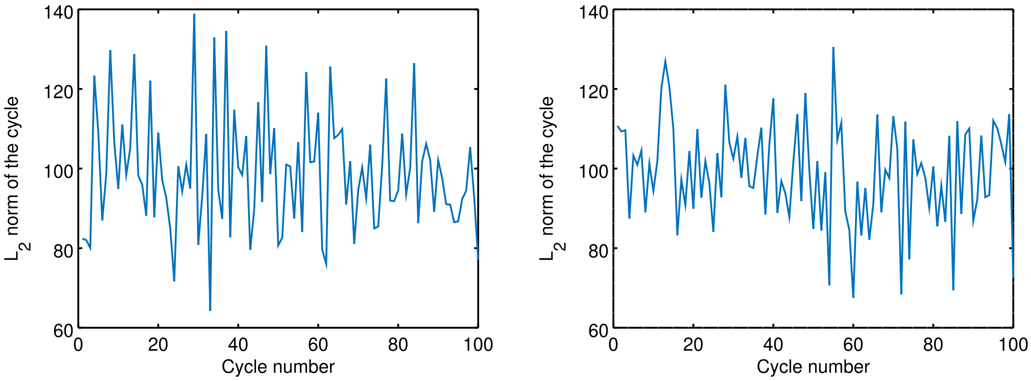

Fig. 2. The L_2 norm of cycles numbered 0 to $n-1$ where $n = 100$. The cycle numbered zero is folded as number 100 on the X-axis. *Left*: Cycles of a random matrix A with entries from $\mathcal{N}(0,1)$. *Right*: Cycles of WAW^* where entries of A are given by $\mathcal{N}(0,1)$.

2.2. Computing WAW^*

The (p,q) entry of the matrix $B = WAW^*$ is given by

$$B(p,q) = \frac{1}{n}\sum_{k=1}^{n}\sum_{j=1}^{n} e^{-\mathrm{i}\frac{2\pi pk}{n}} A(k,j) e^{\mathrm{i}\frac{2\pi jq}{n}}. \tag{2.1}$$

Relation with fast Fourier transform: If $\mathcal{F}(A)$ represents the discrete Fourier transform (DFT) of the columns of matrix A, we have

$$B = \frac{1}{n}\mathcal{F}(\mathcal{F}(A)^*)^*. \tag{2.2}$$

If we denote the two dimensional DFT of the matrix A by $\mathcal{F}^2(A)$, then we have $B = [\mathcal{F}^2(A)]/n$. Thus WAW^* is computed in $2n^2 \log n$ operations when the full linear transformation is required.

Evaluating only k cycles of WAW^*: In the n-point fast-Fourier-transform (FFT) of a vector there are $\log_2 n$ stages in its butterfly structure i.e. decomposition of a larger size DFT recursively into smaller sizes. When we require only k frequency components of the given matrix at the final stage, the number of required points double every previous stage until $2^m k = n$ for some $m = \log_2\left(\frac{n}{k}\right)$. So the number of arithmetic operations O_k for including k frequency components is:

$$O_k = \sum_{j=0}^{m-1} 2^j k + n\left(\log_2 n - \log_2\left(\frac{n}{k}\right)\right),$$

$$O_k = k(2^m - 1) + n\log_2 k,$$

$$O_k = (n-k) + n\log_2 k.$$

The total number of required operations for n vectors is nO_k, which is $\mathcal{O}(n^2)$ for a given k. In the special case of $k = 1$ for a Toeplitz matrix, the averaging of entries on the cycles to evaluate the n entries of the first circulant R_0 requires $2n$ arithmetic operations (see Remark 1.7), followed by the evaluation of $WR_0D_0W^*$ requires $n\log n$ arithmetic operations, where $D_0 = I$. This produces the diagonal of the similar matrix B and the corresponding approximation of the eigenvalues.

Algorithm 2.1 Approximation of eigenvalues using a circulant decomposition.

Output : $\Lambda \longleftarrow [\ ]$; a set of eigenvalues of given matrix.
Initialization: $W(p,q) \longleftarrow \frac{1}{\sqrt{n}} e^{-i\frac{2\pi}{n}pq}$; Construct transformation matrix.
Sparsification : $\tilde{B} \longleftarrow WAW^* - \Delta$; Construct $\tilde{B}$ using the dominant cycles.
Reduced evaluation : $\Lambda \longleftarrow \lambda_i\{\tilde{B}\}$; Eigenvalue algorithm for sparse matrices.

2.3. Error in approximations

The similar matrix B is reduced to a sparse matrix $\tilde{B}$ by selecting dominant cycles with the largest Frobenius norms. Let $\Delta = B - \tilde{B}$, and λ_i be the eigenvalues of B, and $\tilde{\lambda}_i$ be the eigenvalues of $\tilde{B}$. Let $B = X\Lambda X^{-1}$. From Bauer-Fike theorem,

$$|\lambda_i - \tilde{\lambda}_i| \le \kappa(X)\|\Delta\|_2. \tag{2.3}$$

Here $\kappa(X) = \|X\|\|X^{-1}\|$ is the condition number of the eigenvector matrix X. So the eigenvalues are better approximated when the eigenvector matrix X is well conditioned, and $\|\Delta\|_2$ is minimized. Similarly, the relative error can be bound when B is non singular and diagonalizable.

$$\frac{|\lambda_i - \tilde{\lambda}_i|}{|\lambda_i|} \le \kappa(X)\|B^{-1}\Delta\|_2. \tag{2.4}$$

2.4. Positive definiteness of $\tilde{B}$

It can be shown that for a positive definite A, the diagonal matrix with the diagonal entries of B is positive definite [25]. But the matrix $\tilde{B}$ with other cycles included, need not be positive definite. Here we provide a sufficient condition for the matrix $\tilde{B}$ to be positive definite.

Theorem 2.1. *When the approximation $\tilde{B}$ includes the diagonal of B along with any k symmetric cycles indexed by a set of some whole numbers $a_i \in \mathcal{T}$, the approximated matrix is positive definite if $\frac{\sqrt{B(i,i)B(a_j,a_j)}}{k} \ge B(i,a_j)$ for all i, a_j.*

Proof. By using the expansion of $u^*\tilde{B}u$ with constraint $\|u\| = 1$, it can be divided into two parts

$$u^*\tilde{B}u = \sum_i |u_i|^2 B(i,i) + \sum_{p,q\in\mathcal{T}} u_p\overline{u_q}B(p,q) + u_q\overline{u_p}\overline{B(p,q)}, \tag{2.5}$$

$$u^*\tilde{B}u \ge \sum_i |u_i|^2 B(i,i) - \sum_{p,q\in\mathcal{T}} 2|u_p||u_q||\mathrm{Re}(B(p,q))|, \tag{2.6}$$

$$u^*\tilde{B}u \ge \sum_{p,q\in\mathcal{T}} \left(\frac{(|u_p|^2 B(p,p) + |u_q|^2 B(q,q))}{|\mathcal{T}|} - 2|u_p||u_q||\mathrm{Re}(B(p,q))| \right). \tag{2.7}$$

From the last equality, for every $p, q \in \mathcal{T}$ if

$$\frac{(|u_p|^2 B(p,p) + |u_q|^2 B(q,q))}{|\mathcal{T}|} - 2|u_p||u_q||\mathrm{Re}(B(p,q))| \ge 0, \tag{2.8}$$

then it satisfies a sufficient condition for the matrix $\tilde{B}$ to be positive definite. With $\theta = \left|\frac{u_p}{u_q}\right|$, this implies

$$\frac{\theta B(p,p) + \frac{1}{\theta}B(q,q)}{|\mathcal{T}|} \ge 2|\mathrm{Re}(B(p,q))|. \tag{2.9}$$

Now minimizing $f(\theta) = \theta B(p,p) + \frac{1}{\theta}B(q,q)$ w.r.t. θ, we find the minimum value to be $2\sqrt{B(p,p)B(q,q)}$, giving us the sufficient condition,

$$\frac{\sqrt{B(p,p)B(q,q)}}{|\mathcal{T}|} \ge |\mathrm{Re}(B(p,q))|. \tag{2.10}$$

On the other hand, $\frac{B(i,i)+B(j,j)}{2} < |\mathrm{Re}(\tilde{B}(i,j))|$ for any i,j implies a vector $u = ae_i + be_j$ with suitable a, b such that $u^*Bu < 0$. So we note that $\frac{B(i,i)+B(j,j)}{2} > |\mathrm{Re}(\tilde{B}(i,j))|$ for all i,j when B is positive definite.

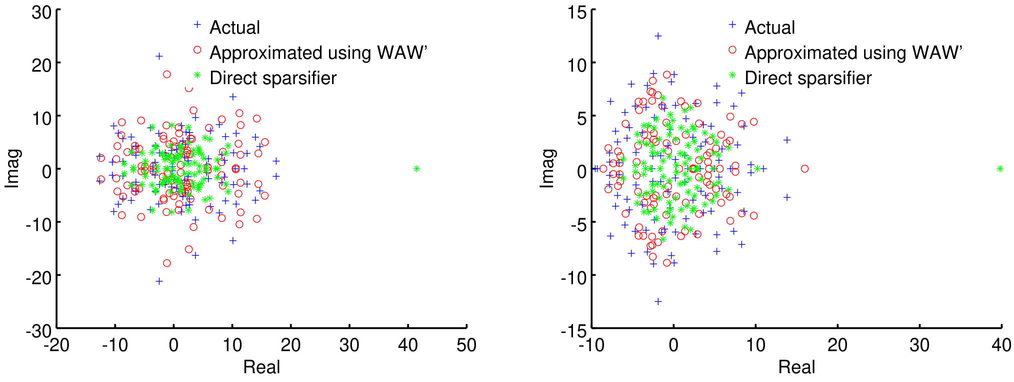

Fig. 3. A qualitative comparison of eigenvalue approximations using the direct [26] and the circulant sparsifiers for identical number of non-zero entries in the resulting matrix. The smaller magnitudes of the former show that the information lost by the direct sparsifier is significantly larger compared to that of the suggested computationally efficient sparsification using the dominant circulant components. *Left*: A random Toeplitz matrix. *Right*: A random block Toeplitz matrix of block size 5.

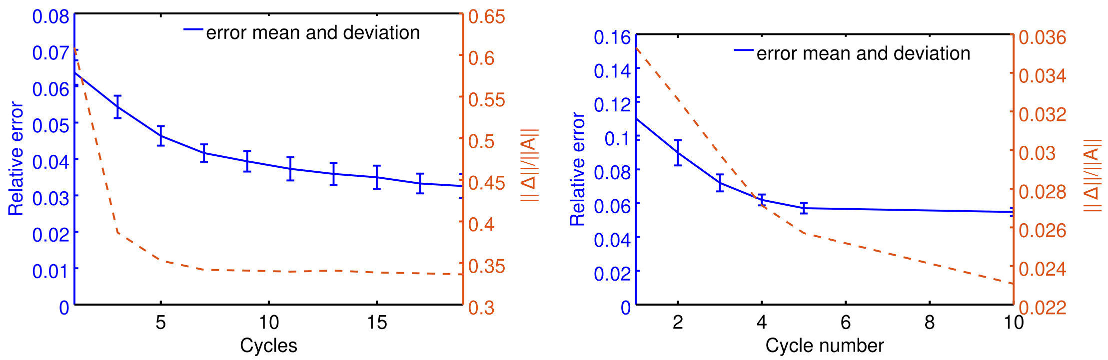

Fig. 4. Average and deviation of relative error in the approximated eigenvalues with increasing number of cycles considered. Corresponding relative errors in the similarity transformation are given by the dashed line and Y-axis on the right. *Left*: Toeplitz matrix of dimension $n = 1000$. *Right*: Block Toeplitz matrix of block size 5 and dimension $n = 1000$.

3. Numerical results

3.1. *Eigenvalue evaluations*

We begin by qualitatively highlighting the advantages of the suggested sparsification over a direct sparsifier, for matrices with periodic properties. Later, we present quantitative results of the errors in the similarity transformation and eigenvalue approximations of Toeplitz, block Toeplitz and quasi-periodic matrices. Random matrices with $\mathcal{N}(0,1)$ as entries of the matrix or its blocks were used for corresponding numerical experiments on Toeplitz and block Toeplitz matrices. It should be noted that in applications where the matrices are not random, one can expect even better results. The results show that including a few dominant cycles provides us lower relative errors for eigenvalues of such matrices, compared to the single-term circulant approximation. Also, while the relative errors may be notable for the smaller eigenvalues, their absolute errors are very small, as expected.

3.1.1. *Comparison with direct sparsifiers*

The eigenvalues of the approximated Toeplitz and block Toeplitz matrices using WAW^* and a direct sparsifier [26], along with the actual eigenvalues are shown in Fig. 3. In general, the eigenvalues with the direct sparsifier are smaller in magnitude i.e. more centered in the plot, compared to the eigenvalues with the proposed sparsifier in the frequency domain. Intuitively, the information lost by the direct sparsifier by thresholding some entries in the matrix to zero, is larger compared to the information lost in the frequency domain by removing the same number of entries as cycles of the matrix WAW^*.

3.1.2. *Toeplitz and block-Toeplitz matrices*

The average relative error and standard deviation of the errors of all the n eigenvalues, decrease with an increase in the number of cycles in the approximation of eigenvalues of a Toeplitz matrix, as shown in Fig. 4. The cumulative relative error of the approximate similarity transformation in terms of the Frobenius norms, $\|\Delta\|/\|A\|$, is also plotted in these figures. Also, the mean of the average relative error and its deviation over 1000 matrices in Fig. 5 show reasonably small errors, and better results are expected for block Toeplitz and other periodic matrices.

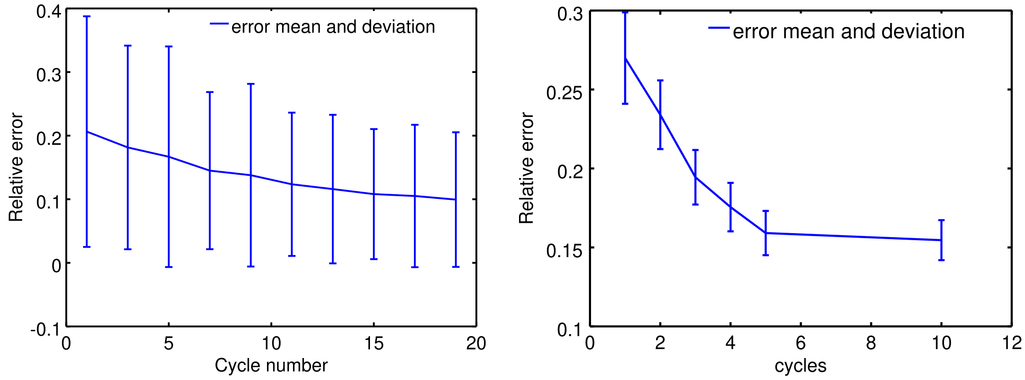

Fig. 5. Mean of the average relative error and its deviation in the approximated eigenvalues with the increasing number of included cycles; evaluated using 1000 random matrices of dimension $n = 100$. Note that the errors further reduce significantly with the increasing dimensions n as shown in Fig. 6. *Left*: Toeplitz matrices. *Right*: Block Toeplitz matrices of block size 5.

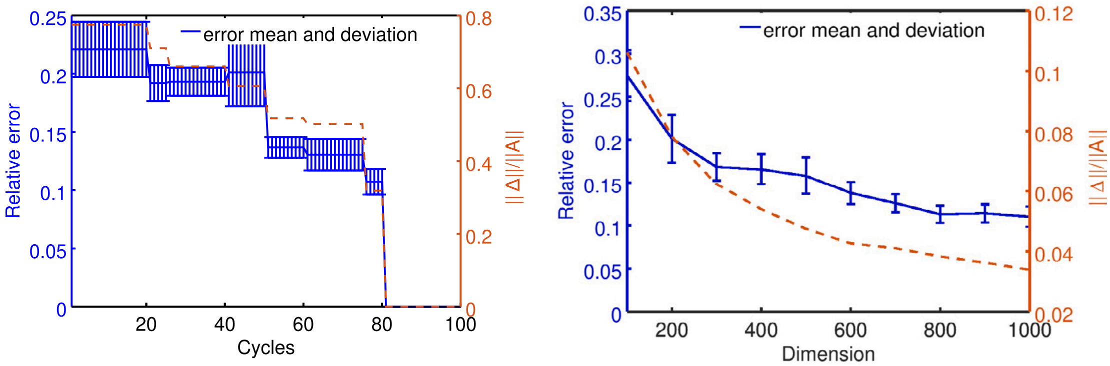

Fig. 6. *Left*: Average and deviation of relative error in the approximated eigenvalues of a random periodic matrix of dimension $n = 100$ with the included cycles. Periods 4, 5, 10 were chosen randomly along the diagonals with a uniform probability. In the suggested algorithm, cycles numbering around 20, 40, 50, 60 and 80 are preferentially included in the sparse similarity transformation. *Right*: Average and deviation of relative error in the approximated eigenvalues using only two cycles (C^0, $C^{\lfloor \frac{n}{2} \rfloor}$), of block Toeplitz matrices of block-size five, for dimensions n ranging from 100 to 1000. Corresponding relative errors in the similarity transformation are given by the dashed line and Y-axis on the right.

Block-Toeplitz matrices with random entries from $\mathcal{N}(0,1)$ were used for the numerical experiments. The average relative error and standard deviation of the errors decrease with an increase in the number of cycles in the approximation as shown in the example in Fig. 4. The mean of the average relative error and its deviation over 1000 matrices are shown in Fig. 5. Note that inclusion of a few cycles may be sufficient to achieve a reasonable accuracy in approximating all the eigenvalues for these matrices. The improvements in the accuracy of the eigenvalue estimation for a given k, with the increase in the size of the matrix, are also illustrated in the Fig. 6.

3.1.3. *Other periodic and quasi-periodic matrices*

The above properties of Toeplitz and block-Toeplitz matrices can describe the properties and the efficacy of sparsification of a quasi-periodic matrix as well. The dominant components for such matrices become the corresponding sizes of periodic blocks in the matrix; thus displaying the characteristics similar to that of a block-Toeplitz matrix. It can be used to include only the dominant cycles of the similar matrix B. Average relative error and deviation in the relative errors are plotted with the cycle number for a random periodic matrix in Fig. 6. Here, a uniform probability of periods $m = 4$, 5, and 10 were used to generate the entries along the diagonals for a matrix of dimension $n = 100$. These result in dominant cycles in the similar matrix B that are multiples of the corresponding integers n/m given by 25, 20, and 10 respectively, with their common multiples being even more significant. Note that the error in Fig. 6 reduces mostly for corresponding cycles, as in the case of a mixture of block Toeplitz matrices.

3.2. *Preconditioning linear systems*

Toeplitz linear systems appear, for example, in solving linear ordinary differential equations and delay differential equations. Here, T. Chan and generalized T. Chan's preconditioners are used in speeding up Conjugate Gradient (PCG) algorithms [21,25]. In Example-1 of the article [25], preconditioned CG is applied on a symmetric Toeplitz matrix with first row given by $A(1,:) = \begin{bmatrix} 2 & \frac{-1}{2} & \frac{-1}{2^2} & \frac{-1}{2^3} & \cdots & \frac{-1}{2^{n-1}} \end{bmatrix}$ with right hand side vector $b = \begin{bmatrix} 1 & 2 & 3 & \cdots & n \end{bmatrix}^T$. The preconditioner considered was a generalization of the T. Chan's preconditioner, given as $P(k) = W^*(B \circ Q(k))W$ with the matrix $Q(k)$ given as

Table 1
Toeplitz matrices of dimension 2000. The first case is Example-1 of [25], and others were generated using random entries. The preconditioner 'I' represents the plain solver. Both preconditioners are identical when only a single cycle of WAW^* is considered. The stopping criterion was a relative residue less than 10^{-6}.

Number of non-zero entries in preconditioner and iterations required ⟶	I	$P(n)$	$P(3n)$	$P(5n)$	$P(7n)$	$P(9n)$
Generalized T. Chan	683	30	23	23	23	23
WAW^* cycles		30	44	43	45	47
Generalized T. Chan	61	38	38	38	38	38
WAW^* cycles		38	27	24	21	21
Generalized T. Chan	103	61	61	61	61	61
WAW^* cycles		61	45	40	34	33
Generalized T. Chan	51	33	33	33	33	33
WAW^* cycles		33	23	21	21	19
Generalized T. Chan	56	36	36	36	36	34
WAW^* cycles		36	26	23	22	20

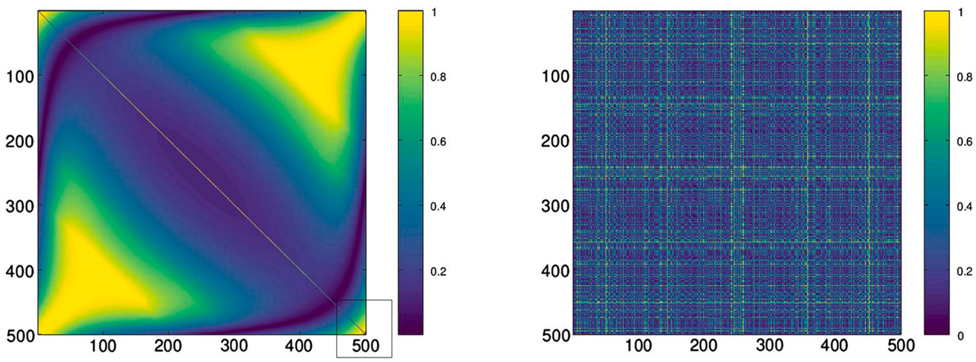

Fig. 7. Cycle-wise dominant entries of the matrix $B = WAW^*$. Heat map represents the magnitude of the matrix entry normalized by the largest magnitude among entries in the corresponding cycle. The matrices A are symmetric Toeplitz matrices of dimension 500. Left: When A has entries from Example-1 of [25], and note the higher relative magnitudes of the cycles at the bottom right indicated by a box. Right: When A has random Gaussian $\mathcal{N}(0,1)$ entries.

$$Q(k) = \begin{bmatrix} I_{n-\lceil\sqrt{n-k}\rceil} & 0 \\ 0 & \mathbf{1}_{\lceil\sqrt{n-k}\rceil} \end{bmatrix}.$$

Here $\mathbf{1}_k$ is a $k \times k$ matrix of all ones. Note that this preconditioner includes a single circulant component i.e. the diagonal of WAW^* along with a square sub-matrix at the bottom-right corner in the matrix B. Table 1 shows the number of iterations required by a preconditioned CG for positive definite symmetric Toeplitz matrices. The generalized T. Chan's preconditioner performs marginally better than the proposed banded preconditioner with a few cycles of WAW^* for Example-1, and not in the other cases. This can be explained using Fig. 7. The matrix WAW^* for Example-1 has higher relative magnitudes of all cycles in the bottom right corner. As mentioned, the T. Chan's preconditioner can capture most of these entries as a bottom-right sub-matrix, in comparison with the corresponding banded preconditioner. Whereas other Toeplitz matrices may not show such a cluster of high magnitude entries near the diagonal. Thus, T. Chan's preconditioner does not reduce required iterations with small increments in the size of its non-zero sub-matrix, in general. Thus the banded preconditioner using the cycles of WAW^* performs better than the generalized T. Chan's preconditioner in the other positive definite Toeplitz matrices with arbitrary entries.

In the case of block-Toeplitz matrices we see that the banded preconditioner speeds up computations with inclusion of cycles of WAW^* as shown in Table 2, whereas the T. Chan's preconditioner fails to do so for the same number of non-zero entries in it. These results highlight the generality of the dominant cycles of matrix $B = WAW^*$, as an appropriate preconditioner for matrices with some periodicity in entries.

4. Conclusion

We began with a decomposition of any given matrix into circulant matrices with periodic relaxations on the unit circle. Exploiting the periodicity of entries along the diagonals of a matrix, and the dominance of a few circulant components, we can reduce the given matrix using appropriate fast-Fourier-transform operations for an approximate and sparse similarity transformation.

Using numerical results, we highlighted the efficacy of the sparsification of the periodic matrices using the circulant decomposition, in terms of errors both in approximation of entries of the matrix and also its eigenvalues. Examples of the relative errors in the eigenvalues were produced as a function of the number of circulant components included in the approximation. Results in preconditioning linear systems were presented where the generalized T. Chan's preconditioner was compared. These results show

Table 2
Block-Toeplitz matrices of dimension 1100 of block size 11. The matrices are symmetric positive definite with symmetric blocks, and condition number of the matrices is $\mathcal{O}(10^4)$. The stopping criterion was a relative residue less than 10^{-6}. The preconditioner 'I' represents the plain solver. Both preconditioners are identical when only a single cycle of WAW^* is considered.

Number of non-zero entries in preconditioner and iterations required ⟶	I	$P(n)$	$P(3n)$	$P(5n)$	$P(7n)$	$P(9n)$	$P(11n)$
Generalized T. Chan	168	162	162	162	162	162	162
WAW^* cycles		162	148	121	106	78	19
Generalized T. Chan	136	121	121	121	121	121	121
WAW^* cycles		121	119	106	94	66	17
Generalized T. Chan	172	172	172	172	172	172	172
WAW^* cycles		172	131	134	105	66	19
Generalized T. Chan	408	402	403	402	400	403	399
WAW^* cycles		402	375	327	284	232	29
Generalized T. Chan	125	124	124	124	124	124	124
WAW^* cycles		124	120	100	90	51	17

that the suggested sparse similarity transformation of a matrix is useful in efficiently approximating eigenvalues, and preconditioning linear systems, and may as well be exploited for other evaluations when a dense matrix has any periodicity along its diagonals.

Funding

Murugesan Venkatapathi acknowledges the support of the Science and Engineering Research Board (SERB) grant MTR/2019/000712 in performing this research.

Data availability

No data was used for the research described in the article.

Appendix A. Single-term circulant approximation of a Toeplitz matrix, and its symbol

A Toeplitz matrix is of the form,

$$A_n = \begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{-1} & a_0 & a_1 & a_2 & \ddots \\ a_{-2} & a_{-1} & a_0 & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ a_{-(n-1)} & a_{-(n-2)} & \cdots & \cdots & a_0 \end{bmatrix}.$$

The function $a(e^{i\theta}) = \sum_{k=-\infty}^{\infty} a_k e^{ik\theta}$ for $\theta \in [0, 2\pi)$ is called the *symbol* of the family of Toeplitz matrices A_n [15]. This section considers different types of symbols, the corresponding eigenvalues of the single-term circulant approximation, and the actual spectrum of the Toeplitz matrices. The single-term approximation is given by only the diagonal entries of WA_nW^*. We have the diagonal entry

$$(WA_nW^*)(p,p) = \frac{1}{n}\sum_{q=0}^{n-1}\sum_{k=0}^{n-1} A_n(q,k)e^{i2\pi p\frac{q-k}{n}} = \frac{1}{n}\left(\sum_{m=1}^{n}\sum_{s=-(m-1)}^{(m-1)} a_s e^{ip\frac{2\pi s}{n}}\right),$$

as the Cesáro sum. When a is continuous and 2π periodic, the Cesaro sum, i.e. the diagonal elements of $(WA_nW^*)(p,p)$, converge uniformly to the range of the symbol evaluated at $e^{i\frac{2\pi p}{n}}$, with $p = 0, \dots, n-1$. Both allow approximation of eigenvalues of Toeplitz matrices; the symbol in $\mathcal{O}(n^2)$ arithmetic operations and the single-circulant approximation in $\mathcal{O}(n\log n)$ arithmetic operations. While the symbol may be better in approximating the spectra of Toeplitz matrices, an approximation using one or more circulant components is applicable even to block-Toeplitz and other dense matrices with some periodicity in entries with a $\mathcal{O}(n^2)$ scaling in the computing effort. Here we present three cases where the relationship between the above two methods of approximating the eigenvalues of a Toeplitz matrix can be qualified, but note that these cases are not exhaustive.

A.1. Case 1: the symbol is of the Tilli class

Spectrum of A_n is said to be canonically distributed if the limiting spectrum approaches the range of the symbol. It was shown by Tilli, that if the complement of the range is a connected set, then the spectrum follows the symbol [27]. Several conditional theorems on the symbol such that the spectrum of A_n is canonically distributed are presented in [15]. Fig. 8 shows the spectrum, single-term circulant approximation, and the symbol for $a(\theta) = (1+\theta)e^{i\theta}$.

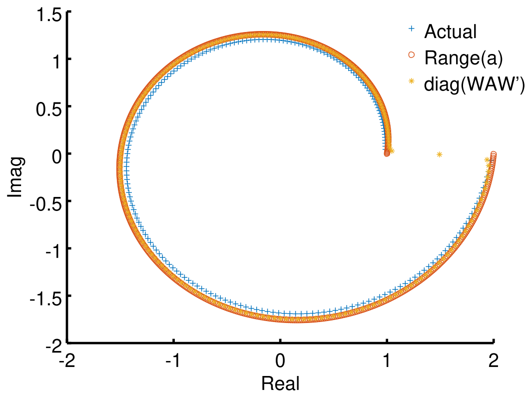

Fig. 8. Range of the symbol, and diagonal entries of WA_nW^* for a Toeplitz matrix of dimension 200 with a symbol $a(\theta) = (1+\theta)e^{i\theta}$.

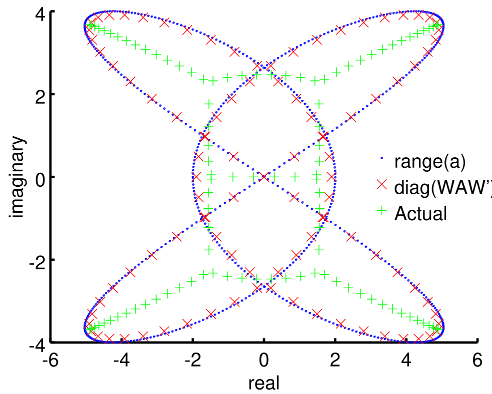

Fig. 9. Range of the symbol, and diagonal entries of WA_nW^* for a banded Toeplitz matrix of dimension 100, with 7 non-zero entries in each row. The work in this paper is not directed at such sparse matrices.

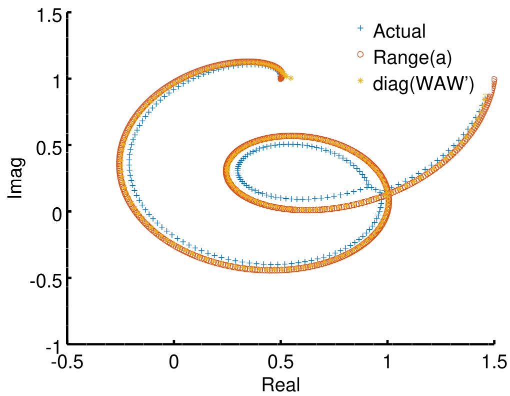

Fig. 10. Range of the symbol, and diagonal entries of WA_nW^* for a Toeplitz matrix of dimension 200 with a symbol $a(\theta) = \frac{\theta}{2\pi} + i\frac{1}{\pi^2}(\theta-\pi)^2 + e^{i2\theta}$.

A.2. Case 2: the symbol of the banded Toeplitz matrix is a trigonometric polynomial

For a banded Toeplitz matrix A_n, with first row $a_0, a_1, a_2, \cdots a_l$ and first column entries $a_0, a_{-1}, \cdots a_{-m}$, the symbol given by $a(e^{i\theta}) = \sum_{k=-m}^{l} a_k e^{ik\theta}$ is a curve in the complex plane. The eigenvalues of such matrices lie in the convex hull (denoted $\mathcal{H}(a)$) of the curve $a(e^{i\theta})$ for $\theta \in [0, 2\pi)$ [14,27]. On the other hand, we have the diagonal entries of WA_nW^* as FFT of the sequence

$$(a_0, \frac{n-1}{n}a_1, \frac{n-2}{n}a_2, \cdots \frac{n-l}{n}a_l, 0, 0, \cdots 0, \frac{n-m}{n}a_{-m}, \cdots \frac{n-1}{n}a_{-1}).$$

Note that these values lie in near proximity to the trace of the symbol in the complex plane when $l, m \ll n$, as shown in Fig. 9.

A.3. Case 3: the symbol is a sum of a polynomial and a trigonometric polynomial

This can be regarded as the combination of the previous two cases. Even though the spectra mimic the symbol on open arcs, they lie inside the closed arcs. Typically, when the symbol is a sum of a polynomial and a trigonometric polynomial we observe such a scenario. Here, Theorems 1 and Theorem 3 of Tilli [27] have to be applied. The spectrum of the matrix with symbol $a(e^{i\theta}) = \frac{\theta}{2\pi} + i\frac{1}{\pi^2}(\theta - \pi)^2 + e^{i2\theta}$ is shown in Fig. 10.

References

[1] Sivaram Ambikasaran, Eric Darve, An $O(N \log N)$ fast direct solver for partial hierarchically semi-separable matrices, J. Sci. Comput. 57 (3) (2013) 477–501.
[2] Yousef Saad, Numerical Methods for Large Eigenvalue Problems, vol. 66, revised edition, SIAM, 2011.
[3] Th Koslowski, W. Von Niessen, Linear combination of Lanczos vectors: a storage-efficient algorithm for sparse matrix eigenvector computations, J. Comput. Chem. 14 (7) (1993) 769–774.
[4] Christopher Conway Paige, The computation of eigenvalues and eigenvectors of very large sparse matrices, PhD thesis, University of London, 1971.
[5] James R. Bunch, Donald J. Rose, Sparse Matrix Computations, Academic Press, 2014.
[6] Sven-Erik Ekström, Carlo Garoni, Stefano Serra-Capizzano, Are the eigenvalues of banded symmetric Toeplitz matrices known in almost closed form?, Exp. Math. 27 (4) (2018) 478–487.
[7] M. Hariprasad, Murugesan Venkatapathi, Semi-analytical solutions for eigenvalue problems of chains and periodic graphs, Appl. Math. Comput. 411 (2021) 126512.
[8] F. William Trench, Numerical solution of the eigenvalue problem for Hermitian Toeplitz matrices, SIAM J. Matrix Anal. Appl. 10 (2) (1989) 135–146.
[9] William F. Trench, On the eigenvalue problem for Toeplitz band matrices, Linear Algebra Appl. 64 (1985) 199–214.
[10] D. Bini, V. Pan, Efficient algorithms for the evaluation of the eigenvalues of (block) banded Toeplitz matrices, Math. Comput. 50 (182) (1988) 431–448.
[11] Franklin T. Luk, Sanzheng Qiao, A fast eigenvalue algorithm for Hankel matrices, Linear Algebra Appl. 316 (1–3) (2000) 171–182.
[12] R. Gray, On the asymptotic eigenvalue distribution of Toeplitz matrices, IEEE Trans. Inf. Theory 18 (6) (1972) 725–730.
[13] N.K. Bose, K.J. Boo, Asymptotic eigenvalue distribution of block-Toeplitz matrices, IEEE Trans. Inf. Theory 44 (2) (1998) 858–861.
[14] Albrecht Böttcher, Sergei M. Grudsky, Spectral Properties of Banded Toeplitz Matrices, SIAM, 2005.
[15] Albrecht Böttcher, Bernd Silbermann, Introduction to Large Truncated Toeplitz Matrices, Springer Science & Business Media, 2012.
[16] Zhihui Zhu, Michael B. Wakin, On the asymptotic equivalence of circulant and Toeplitz matrices, IEEE Trans. Inf. Theory 63 (5) (2017) 2975–2992.
[17] Manuel Bogoya, Sven-Erik Ekström, Stefano Serra-Capizzano, Fast Toeplitz eigenvalue computations, joining interpolation-extrapolation matrix-less algorithms and simple-loop theory, Numer. Algorithms (2022) 1–24.
[18] Raymond Chan, Gilbert Strang, The asymptotic Toeplitz-circulant eigenvalue problem, 1987.
[19] Peter Arbenz, Computing eigenvalues of banded symmetric Toeplitz matrices, SIAM J. Sci. Stat. Comput. 12 (4) (1991) 743–754.
[20] Tony F. Chan, An optimal circulant preconditioner for Toeplitz systems, SIAM J. Sci. Stat. Comput. 9 (4) (1988) 766–771.
[21] Xiao-Qing Jin, Yi-Min Wei, A survey and some extensions of T. Chan's preconditioner, Linear Algebra Appl. 428 (2–3) (2008) 403–412.
[22] Silvia Noschese, Lothar Reichel, A note on superoptimal generalized circulant preconditioners, Appl. Numer. Math. 75 (2014) 188–195.
[23] T. Kailath, Remarks on the origin of the displacement-rank concept, Appl. Math. Comput. 45 (2) (1991) 193–206.
[24] Jesper Nielsen, Tobias Jensen, Jesper Rindom Jensen, Mads Christensen, Søren Jensen, Fast fundamental frequency estimation: making a statistically efficient estimator computationally efficient, Signal Process. 135 (01 2017).
[25] Ming-Chao Cai, Xiao-Qing Jin, Yi-Min Wei, A generalization of T. Chan's preconditioner, Linear Algebra Appl. 407 (2005) 11–18.
[26] Dimitris Achlioptas, Frank McSherry, Fast computation of low-rank matrix approximations, J. ACM 54 (2) (2007), 9–es.
[27] Paolo Tilli, Some results on complex Toeplitz eigenvalues, J. Math. Anal. Appl. 239 (2) (1999) 390–401.