SIAM J. MATRIX ANAL. APPL.
Vol. 24, No. 3, pp. 645–659

REAL-VALUED, LOW RANK, CIRCULANT APPROXIMATION*

MOODY T. CHU† AND ROBERT J. PLEMMONS‡

Abstract. Partially due to the fact that the empirical data collected by devices with finite bandwidth often neither preserves the specified structure nor induces a certain desired rank, retrieving the nearest structured low rank approximation from a given data matrix becomes an imperative task in many applications. This paper investigates the case of approximating a given target matrix by a real-valued circulant matrix of a specified, fixed, and low rank. A fast Fourier transform (FFT)-based numerical procedure is proposed to speed up the computation. However, since a conjugate-even set of eigenvalues must be maintained to guarantee a real-valued matrix, it is shown by numerical examples that the nearest real-valued, low rank, and circulant approximation is sometimes surprisingly counterintuitive.

Key words. real-valued circulant matrix, lower rank, nearest approximation, conjugate-even, fast Fourier transform, truncated singular value decomposition

AMS subject classifications. 41A29, 41A50, 15A18, 65F35, 15A60

PII. S0895479801383166

1. Introduction.

Finding a low rank approximation of a general data matrix is a critical task in many aspects. The list of applications includes image compression, noise reduction, seismic inversion, latent semantic indexing, principal component analysis, regularization for ill-posed problems, and so on. Practical means to tackle this low rank approximation problem include the truncated singular value decomposition (TSVD) method [9], the Lanczos bidiagonalization process [14], and the Monte Carlo algorithm [11]. When the underlying matrix is also required to retain a certain structure, however, few techniques are available. Some preliminary discussion on structured low rank approximation regarding its mathematical properties, interesting applications, and an outline of some possible numerical procedures can be found in [5]. This paper concerns the special case of real-valued low rank approximation with circulant structure.

By an $n \times n$ circulant matrix, we mean a matrix C of the form

$$C = \begin{bmatrix} c_0 & c_1 & & \cdots & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \cdots & c_{n-2} \\ c_{n-2} & c_{n-1} & c_0 & \cdots & c_{n-3} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ c_1 & c_2 & & c_{n-1} & c_0 \end{bmatrix},$$

where each of its rows is just the previous row cycled forward one step. A circulant matrix is uniquely determined by the entries of its first row. We shall denote

*Received by the editors January 3, 2001; accepted for publication (in revised form) by P. C. Hansen May 23, 2002; published electronically January 23, 2003.

http://www.siam.org/journals/simax/24-3/38316.html

†Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205 (chu@math.ncsu.edu). The research of this author was supported in part by the National Science Foundation under grant DMS-9803759.

‡Department of Computer Science and Mathematics, Wake Forest University, Winston-Salem, NC 27109 (plemmons@mthcsc.wfu.edu). The research of this author was supported in part by the Air Force Office of Scientific Research under grant AFOSR-F49620-00-1-0155, the Army Research Office under grant DAAD-19-00-1-0540, and the National Science Foundation under grant CCR-9732070.

a circulant matrix by *Circul*($\mathbf{c}$) if its first row is $\mathbf{c}$. In this paper, we are mainly concerned with the case when $\mathbf{c} \in \mathbb{R}^n$.

Let Π ($= \Pi_n$) denote the specific permutation matrix of order n,

$$\Pi := \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & & & & 1 \\ 1 & 0 & & \dots & 0 \end{bmatrix}. \tag{1.1}$$

It is easy to see that

$$C = \sum_{k=0}^{n-1} c_k \Pi^k \tag{1.2}$$

if and only if $C = \mathit{Circul}(\mathbf{c})$ with $\mathbf{c} := [c_0, \dots, c_{n-1}]$. It is convenient to represent this relationship as

$$\mathit{Circul}(\mathbf{c}) = P_{\mathbf{c}}(\Pi), \tag{1.3}$$

where

$$P_c(x) = \sum_{k=0}^{n-1} c_k x^k \tag{1.4}$$

is called the characteristic polynomial of *Circul*($\mathbf{c}$). Because of this representation, it follows that circulant matrices are closed under multiplication. It is also clear that circulant matrices commute under multiplication. Many important properties of circulant matrices can be traced back mainly to those of the matrix Π. The circulant structure often makes it possible to resolve many matrix-theoretic questions by "closed form" answers. The book by Davis [6] is generally considered the most complete reference on circulant matrices. It is also well known that circulant matrices are closely related to Fourier analysis [15]. That relationship will be used to develop a fast algorithm in this paper.

Circulant matrices have received much attention because the circulant form arises from areas such as acoustics, electrodynamics, image processing, mathematical statistics, number theory, numerical analysis, and stationary time series. To mention a few specific examples, circulant matrices often are used as preconditioners for ill-posed problems [2, 13]. In a recent book by Kailath and Sayed [10], circulant matrices are related to important applications of linear estimation theory. Circulant matrices even find applications to multiconjugate adaptive optics, as was discussed in [7, 8, 12].

Our goal in this paper is to retrieve as much information as possible from a given real-valued matrix A while enforcing a circulant structure and a rank condition; that is, we want to best approximate A with a real-valued circulant matrix C with a certain desired rank. Before moving on, we first point out the following three limitations imposed upon our approximation:

- We are emphasizing real-valued approximation. If there is no constraint requiring C to have real coefficients, then the nearest circulant approximation can easily be achieved via the notion of TSVD. (See Algorithm 3.1.)

- We are fixing the rank to a specific value and not to a certain range. In the latter, say, under the circumstances where singular values of the data matrix decay gradually to zero, the *precise* number of singular values included in a TSVD solution might not be very important. Such a flexibility on the rank condition, as was discussed in [5], is much easier to handle than the fixed rank condition.
- If the Frobenius matrix norm is used as the measurement of nearness, we may assume without loss of generality that the original matrix A is the Chan circulant matrix [3] to begin with. This can easily be seen from the fact that circulant matrices form a linear subspace, and thus the square of the distance from a target matrix to its nearest low rank circulant approximation is the sum of the squares of the distance from the target matrix to the linear subspace of circulant matrices (and hence its Chan circulant approximation) and the distance from the Chan circulant approximation to its nearest low rank circulant approximation.

Under these constraints, we follow the notion of the TSVD to propose a fast Fourier transform (FFT)-based fast algorithm. In order to keep the final low rank approximation a real-valued matrix, we recast the approximation as a data matching problem. As it turns out, we discover a situation where sometimes one may have to delete the largest eigenvalue in order to obtain a real-valued matrix. This surprising and somewhat counterintuitive case might not be significant in applications since, when the precise rank is not critically important, one may slightly relax the rank condition (say, from holding the given rank exactly to being no greater than the given rank), as we have indicated in the second bulleted item above. However, this discussion still might be worth noting in that it clearly demonstrates the disparity between fixed rank and variable rank and real-valued and complex-valued approximations.

2. Basic spectral properties. In this section, we briefly review some of the basic spectral properties relevant to our study. Most of the proofs can be found in [6, 15].

Let $i := \sqrt{-1}$. For a fixed integer $n \geq 1$, let ω $(= \omega_n)$ denote the primitive nth root of unity

$$\omega := \exp\left(\frac{2\pi i}{n}\right). \tag{2.1}$$

Let Ω $(= \Omega_n)$ denote the diagonal matrix

$$\Omega := \operatorname{diag}(1, \omega, \omega^2, \ldots, \omega^{n-1}), \tag{2.2}$$

and let F $(= F_n)$ denote the so-called discrete Fourier matrix whose Hermitian adjoint F^* is defined by

$$F^* := \frac{1}{\sqrt{n}} \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & \omega & \omega^2 & \ldots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \ldots & \omega^{2n-2} \\ \vdots & & & & \vdots \\ 1 & \omega^{n-1} & \omega^{n-2} & \ldots & \omega \end{bmatrix}. \tag{2.3}$$

Note that $\sqrt{n}F^*$ is the Vandermonde matrix generated by the row vector $[1, \omega, \omega^2, \ldots, \omega^{n-1}]$ and that F is a unitary matrix. The following spectral decomposition is a key to our discussion [6].

THEOREM 2.1. *The forward shift matrix* Π *is unitarily diagonalizable. Indeed,*

$$\Pi = F^* \Omega F. \tag{2.4}$$

The circulant matrix Circul$(\mathbf{c})$ *with any given row vector* $\mathbf{c}$ *has a spectral decomposition*

$$Circul(\mathbf{c}) = F^* P_{\mathbf{c}}(\Omega) F. \tag{2.5}$$

Observe that the vector of eigenvalues $\boldsymbol{\lambda} = [P_{\mathbf{c}}(1), \ldots, P_{\mathbf{c}}(\omega^{n-1})]$ of a circulant matrix *Circul*$(\mathbf{c})$ can quickly be calculated from

$$\boldsymbol{\lambda}^T = \sqrt{n} F^* \mathbf{c}^T. \tag{2.6}$$

From (2.6), the inverse eigenvalue problem of finding a circulant matrix with a prescribed spectrum can also be answered easily: Given any vector $\boldsymbol{\lambda} := [\lambda_0, \ldots, \lambda_{n-1}]$, the circulant matrix *Circul*$(\mathbf{c})$ with $\mathbf{c}$ defined by

$$\mathbf{c}^T = \frac{1}{\sqrt{n}} F \boldsymbol{\lambda}^T \tag{2.7}$$

will have eigenvalues $\{\lambda_0, \ldots, \lambda_{n-1}\}$. It is important to note that the matrix-vector multiplication involved in either (2.6) or (2.7) is precisely that involved in the FFT. Thus both the eigenvalue problem and the inverse problem for circulant matrices can be answered in $O(n \log_2 n)$ floating point operations [15]. Observe also that, if all of the eigenvalues are distinct, then there are precisely $n!$ many distinct circulant matrices with the prescribed spectrum.

For real circulant matrices, every complex-valued eigenvalue has the corresponding complex conjugate as another eigenvalue. Indeed, the spectrum of any real circulant matrix necessarily appears in a more special order, called *conjugate-even* in [15]. In order to obtain a *real-valued* circulant matrix by using the FFT in (2.7) for the inverse eigenvalue problem, the vector $\boldsymbol{\lambda}$ of the prescribed eigenvalues must also be arranged in a conjugate-even order. More precisely, the following arrangement of eigenvalues allows for efficient FFT calculation for real data [15].

THEOREM 2.2. *If the eigenvalues are arranged in the order that*

1. $\boldsymbol{\lambda} := [\lambda_0, \lambda_1, \ldots, \lambda_{m-1}, \lambda_m, \overline{\lambda_{m-1}}, \ldots, \overline{\lambda_1}]$, *where* $\lambda_0, \lambda_m \in \mathbb{R}$ *and* $n = 2m$, *or*
2. $\boldsymbol{\lambda} := [\lambda_0, \lambda_1, \ldots, \lambda_m, \overline{\lambda_m}, \ldots, \overline{\lambda_1}]$, *where* $\lambda_0 \in \mathbb{R}$ *and* $n = 2m + 1$,

then the circulant matrix Circul$(\mathbf{c})$ *with* $\mathbf{c}$ *obtained from* (2.7) *is real-valued and has entries in the prescribed vector* $\boldsymbol{\lambda}$ *as its spectrum.*

For later reference, we shall refer to λ_0 and λ_m, if $n = 2m$, and λ_0, if $n = 2m + 1$, in the above theorem as the *absolutely real* elements in $\boldsymbol{\lambda}$.

The singular value decomposition of *Circul*$(\mathbf{c})$ is also easy to establish. It follows from rewriting the expression (2.5) as

$$Circul(\mathbf{c}) = \left(F^* P_{\mathbf{c}}(\Omega) |P_{\mathbf{c}}(\Omega)|^{-1}\right) |P_{\mathbf{c}}(\Omega)| F, \tag{2.8}$$

where $|X|$ denotes the matrix of absolute values of the elements of X. The singular values of *Circul*$(\mathbf{c})$ are $|P_{\mathbf{c}}(\omega^k)|$, $k = 0, 1, \ldots, n-1$. Observe the following necessary characteristic for singular values of a real circulant matrix.

THEOREM 2.3. *Any* $n \times n$ *real-valued circulant matrix can have at most* $\lceil \frac{n+1}{2} \rceil$ *distinct singular values. More precisely, the singular values must appear in the following way:*

1. $\sigma_{n_0}, \sigma_{n_1}, \sigma_{n_1}, \ldots, \sigma_{n_{m-1}}, \sigma_{n_{m-1}}, \sigma_{n_m}$ *if* $n = 2m$ *or*
2. $\sigma_{n_0}, \sigma_{n_1}, \sigma_{n_1}, \ldots, \sigma_{n_m}, \sigma_{n_m}$ *if* $n = 2m + 1$.

3. Low rank approximation. Given a general matrix $A \in \mathbb{R}^{n\times n}$, its nearest circulant matrix approximation measured in the Frobenius norm is simply the Chan circulant matrix *Circul*($\mathbf{c}$) obtained by averaging over diagonals of A, as shown in [3]. Indeed, if $\mathbf{c} = [c_0, \ldots, c_{n-1}]$, then

$$c_k := \frac{1}{n}\langle A, \Pi^k\rangle, \quad k = 0, \ldots, n-1, \tag{3.1}$$

where

$$\langle X, Y\rangle = \operatorname{trace}(XY^T)$$

stands for the Frobenius inner product. This projection *Circul*($\mathbf{c}$) is generally of full rank even if A has lower rank to begin with. Recall that the TSVD gives rise to the nearest low rank approximation in the Frobenius norm. Observe further that the low rank approximation *Circul*($\hat{\mathbf{c}}$) of a circulant matrix *Circul*($\mathbf{c}$) by the TSVD is automatically circulant. We thus have the following algorithm for low rank circulant approximation.

ALGORITHM 3.1. *Given a general $n\times n$ matrix A, the matrix Circul($\hat{\mathbf{c}}$) computed below is a nearest circulant matrix to A with rank no higher than $\kappa \le n$.*

1. *Use the projection* (3.1) *to find the nearest circulant matrix approximation Circul($\mathbf{c}$) of A.*
2. *Use the inverse FFT* (2.6) *to calculate the spectrum $\boldsymbol{\lambda}$ of the matrix Circul($\mathbf{c}$).*
3. *Let $\hat{\boldsymbol{\lambda}}$ be the vector consisting of elements of $\boldsymbol{\lambda}$, but those corresponding to the $n-\kappa$ smallest (in modulus) singular values are set to zero.*
4. *Apply the FFT* (2.7) *to $\hat{\boldsymbol{\lambda}}$ to compute a nearest circulant matrix Circul($\hat{\mathbf{c}}$) of rank κ to A.*

The above algorithm is fast due to the employment of efficient FFT calculation. The resulting matrix *Circul*($\hat{\mathbf{c}}$), however, is complex-valued in general. To construct real-valued low rank approximation, the truncated singular values must be specifically selected so that the resulting vector $\hat{\boldsymbol{\lambda}}$ of *truncated* eigenvalues is conjugate-even. Recall from Theorem 2.3 that most of the singular values are paired. Thus, to preserve the conjugate-even property, the deletion of one complex eigenvalue necessitates the deletion of its complex conjugate as well. To achieve the desired rank, the criteria for truncation must be modified in a special way, as we shall now describe.

It is clear from Theorem 2.1 that all circulant matrices of the same size have the same set of unitary eigenvectors. The real-valued low rank circulant approximation problem, therefore, is equivalent to the following data matching problem (DMP):

(DMP) *Given a conjugate-even vector $\boldsymbol{\lambda} \in \mathbb{C}^n$, find its nearest conjugate-even approximation $\hat{\boldsymbol{\lambda}} \in \mathbb{C}^n$ in the 2-norm subject to the constraint that $\hat{\boldsymbol{\lambda}}$ has exactly $n-\kappa$ zeros.*

Note that finding the closest vector approximation $\hat{\boldsymbol{\lambda}}$ in the 2-norm produces the closest matrix approximation in the Frobenius norm. If there were no conjugate-even constraint, the DMP could easily be answered. See, for example, [1, 4]. With the conjugate-even constraint, we claim that the DMP could be solved according to the following sorting scheme.

THEOREM 3.1. *The optimal solution $\hat{\boldsymbol{\lambda}}$ to the DMP must be such that its nonzero entries match precisely with the first κ conjugate-even components of $\boldsymbol{\lambda}$ according to the descending order of their moduli.*

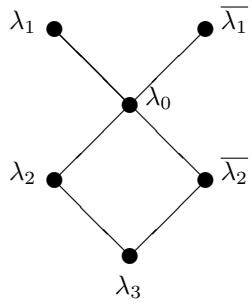

FIG. 1. *Tree graph of* $\lambda_1, \overline{\lambda_1}, \lambda_0, \lambda_2, \overline{\lambda_2}, \lambda_3$ *with* $|\lambda_1| \geq |\lambda_0| > |\lambda_2| \geq |\lambda_3|$.

Proof. Without loss of generality, we may write $\hat{\boldsymbol{\lambda}} = [\hat{\boldsymbol{\lambda}}_1, 0] \in \mathbb{C}^n$ with $\hat{\boldsymbol{\lambda}}_1 \in \mathbb{C}^\kappa$ to be determined and consider the problem of minimizing

$$F(P, \hat{\boldsymbol{\lambda}}) = \|P\hat{\boldsymbol{\lambda}}^T - \boldsymbol{\lambda}^T\|_2^2,$$

where the permutation matrix P is used to search for the match. Partition the permutation matrix into $P = [P_1, P_2]$ with $P_1 \in R^{n\times\kappa}$. The objective function in the least squares problem is reduced to

$$F(P, \hat{\boldsymbol{\lambda}}) = \|P_1\hat{\boldsymbol{\lambda}}_1^T - \boldsymbol{\lambda}^T\|_2^2,$$

which obviously has its optimal solution with

$$\hat{\boldsymbol{\lambda}}_1 = \boldsymbol{\lambda}P_1.$$

This proves the important fact that the entries of $\hat{\boldsymbol{\lambda}}_1$ must come from the rearrangement of κ components of $\boldsymbol{\lambda}$. Indeed, the objective function becomes

$$F(P, \hat{\boldsymbol{\lambda}}) = \|(P_1P_1^T - I)\boldsymbol{\lambda}\|_2^2,$$

where $P_1P_1^T - I$ is but a projection. To minimize $F(P, \boldsymbol{\lambda}P_1)$, the optimal permutation P should be such that $P_1P_1^T$ projects $\boldsymbol{\lambda}$ onto its first κ components with as large a modulus as possible while maintaining the conjugate-even condition. □

In other words, without the conjugate-even constraints, the answer to the DMP corresponds precisely to the usual selection criterion mentioned in Algorithm 3.1, i.e., $\hat{\boldsymbol{\lambda}}$ is obtained by setting the $n - \kappa$ elements of $\boldsymbol{\lambda}$ with smallest modulus to zeros. With the conjugate-even constraint, the above criterion remains effective, but the truncation also depends on the conjugate-even structure inside $\boldsymbol{\lambda}$, as we explain next.

Consider the case $n = 6$ as an example. We shall first assume that neither λ_1 nor λ_2 is a real number. There are six possible conjugate-even structures. For convenience, we shall denote each structure by a tree graph. Each node in the tree represents an element of $\boldsymbol{\lambda}$. Arrange the nodes from top to bottom according to the descending order of their moduli. In case of a tie, arrange the complex conjugate nodes at the same level, and place the real node below the complex nodes. Thus the conjugate-even structure $\lambda_1, \overline{\lambda_1}, \lambda_0, \lambda_2, \overline{\lambda_2}, \lambda_3$, arranged in the descending order of their moduli, will be denoted by the tree in Figure 1.

The nearest conjugate-even vectors to $\boldsymbol{\lambda}$ of rank 5, 3, and 2, respectively, are easy to determine. Their trees are given in Figure 2, where ∘ and • at each node denotes,

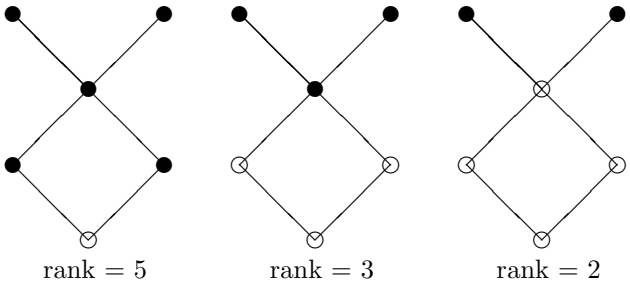

FIG. 2. *Tree graphs of* $\hat{\boldsymbol{\lambda}}$ *with rank* 5, 3, *and* 2.

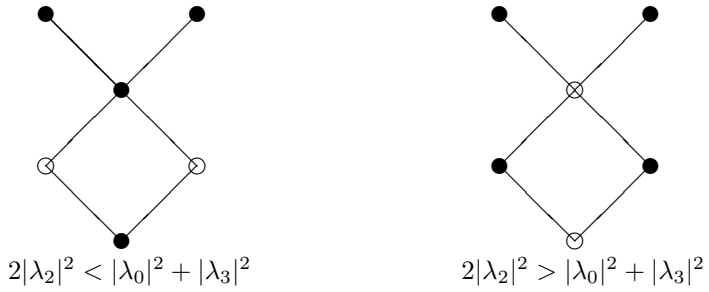

FIG. 3. *Tree graphs of* $\hat{\boldsymbol{\lambda}}$ *with rank* 4.

respectively, that the particular node is being replaced by zero or remains unchanged from the original tree. For these ranks and for this specific tree structure depicted in Figure 1, the conjugate-even requirement has no effect.

However, depending upon whether $2|\lambda_2|^2 > |\lambda_0|^2 + |\lambda_3|^2$, there are two choices for $\hat{\lambda}$ as the nearest conjugate-even approximation of rank 4. See Figure 3. Finally, the nearest rank-1 conjugate-even approximation for the tree of $\boldsymbol{\lambda}$ given by Figure 1 is depicted in Figure 4.

It should be noted that we have implicitly assumed that, if $n = 2m$, then the two absolutely real elements in a conjugate-even $\boldsymbol{\lambda}$ are λ_0 and λ_m and that $|\lambda_0| \geq |\lambda_m|$. We have also assumed that the remaining $2m - 2$ elements are "potentially" complex-valued (some of them could in fact turn out to be real-valued), that they are paired up (necessarily), and that they are arranged in descending order, i.e., $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_{m-1}|$. A similar assumption can be made for the case in which $n = 2m + 1$. See the ordering stated in Theorem 2.2. Nevertheless, we will *never* assume any ordering relationship between the absolutely real element(s) and the potential complex elements. Indeed, it is precisely such an ordering relationship that will determine the truncation criteria as we have demonstrated above for the case in which $n = 6$. In other words, assuming that there are exactly $m + 1$ distinct absolute values of elements in $\boldsymbol{\lambda}$, there are exactly $m(m + 1)/2$ many possible conjugate-even structures for the case in which $n = 2m$, depending upon where the moduli of the absolutely real elements fit into the moduli of the potentially complex elements when a *total* ordering is taken.

Again, under the assumption that neither λ_1 or λ_2 is a real number, we further illustrate our point by considering other cases for $n = 6$ in Figure 5. The leftmost column in Figure 5 represents the six possible conjugate-even structures of $\boldsymbol{\lambda}$ when

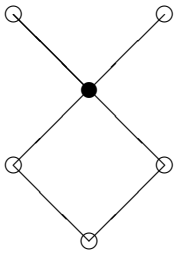

FIG. 4. *Tree graph of* $\hat{\boldsymbol{\lambda}}$ *with rank* 1.

elements are arranged in descending order of their moduli. For each fixed structure, moving from left to right, Figure 5 demonstrates the plan of how the nodes on the original tree should be "pruned" to solve the DMP for various lower rank conditions. There are four cases, A4, A2, B2, and F4, in which additional comparisons are needed to further discern which plan should be used. This situation happens when an even number of nodes from a "loop" are to be dropped. We have already discussed the case F4 in Figure 3. Other cases can easily be identified.

It is entirely possible that there are real-valued elements other than the two (when n is even) absolutely real elements in a conjugate-even λ. The eigenvalues of a symmetric circulant matrix, for instance, are conjugate-even and are all real. When this happens, these conjugate-even real-valued elements must appear in pairs, and the truncation criteria are further complicated. Using the example discussed in Figure 1 but assuming further that $\lambda_2 = \overline{\lambda_2}$, we illustrate our point below. First, we use a dashed link in Figure 6 and larger dots to indicate the occurrence of $\lambda_2 = \overline{\lambda_2}$. It is important to note that, in contrast to the two drawings in Figure 3, the tree graph of the nearest conjugate-even approximation $\hat{\lambda}$ with rank 4 changes its structure in this case. See Figure 7.

4. Algorithm. While the aforementioned graph-theoretic concept should be quite easy to follow, a general purpose code is not as straightforward. To facilitate the discussion, we now present an algorithm for computing the real-valued low rank circulant approximation. In order to highlight the notion on how the singular values of *Circul*($\mathbf{c}$) should be truncated, we simplify many computational operations by adopting a pseudo-MATLAB syntax. The commands for these abridged operations are denoted in boldface (whereas, to avoid distraction, the vectors $\mathbf{c}$ and $\boldsymbol{\lambda}$ are denoted as ordinary c and λ) in the steps of the following algorithm.

ALGORITHM 4.1. *Given* $\mathbf{c} \in \mathbb{R}^n$ *and a positive integer* $1 \le \kappa < n$, *let* $m = \lfloor \frac{n}{2} \rfloor$. *Define* $tol = n\epsilon\|\mathbf{c}\|$ *where* ϵ *is the machine accuracy as the threshold of system zero. The matrix Circul*($\hat{\mathbf{c}}$) *with* $\hat{\mathbf{c}}$ *computed at the end of the following steps has eigenvalues* $\hat{\boldsymbol{\lambda}}$ *containing exactly* $n - \kappa$ *zeros and is the nearest approximation to Circul*($\mathbf{c}$).

1. $\lambda = n * \mathbf{ifft}(c)$; *(Indices of* λ *start with* 1.*)*
   $\hat{\lambda} = \lambda(1 : m + 1)$;
2. *if* $n = 2m$
   $I_r = \mathbf{find}(\mathbf{abs}(\mathbf{imag}(\lambda(2 : m))) < tol) + 1$;
   $I_c = \mathbf{find}(\tilde{}\mathbf{ismember}(2 : m, I_r)) + 1$;
   else
   $I_r = \mathbf{find}(\mathbf{abs}(\mathbf{imag}(\lambda(2 : m + 1))) < tol) + 1$;
   $I_c = \mathbf{find}(\tilde{}\mathbf{ismember}(2 : m + 1, I_r)) + 1$;

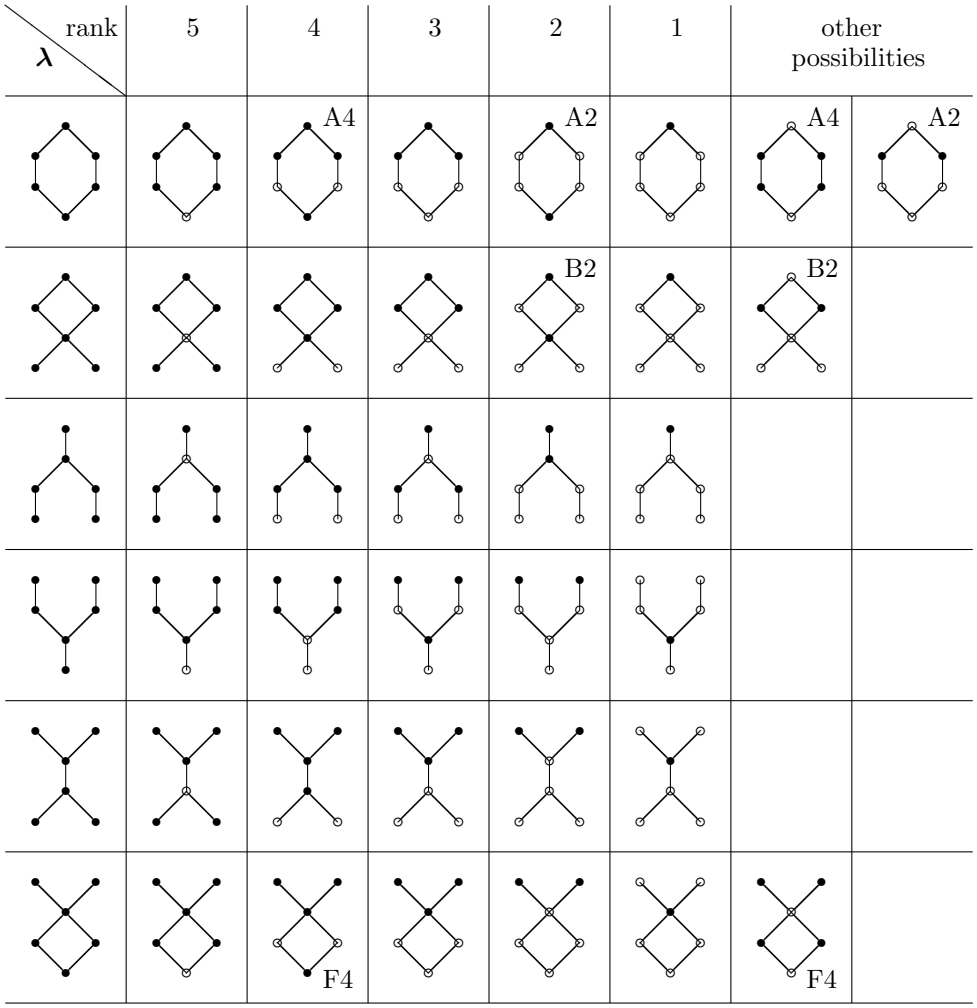

FIG. 5. *Possible solutions to the DMP when* $n = 6$.

end

3. $[t, J] = \mathbf{sort}(\mathbf{abs}(\lambda))$*;*
   $J = \mathbf{fliplr}(J)$*;* *(J is the index set sorting λ in descending order.)*
   for $i = 1 : m + 1$

$$I(:,:,i) = \begin{cases} [2,0]\,, & \textit{if } \mathbf{ismember}(J(i), I_c); \\ [2,1]\,, & \textit{if } \mathbf{ismember}(J(i), I_r); \\ [1,1]\,, & \textit{otherwise}; \end{cases}$$

   end

4. $\sigma = 0$*;*
   $s = m + 1$*;*
   while $\sigma < n - \kappa$
   $\sigma = \sigma + I(1, 1, s)$*;*
   $s = s - 1$*;*
   end
   $idx = s + 1$*;* *(idx indicates the place where λ is to be cut.)*

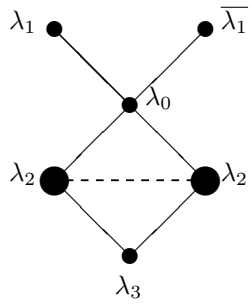

FIG. 6. *Tree graph of* $\lambda_1, \overline{\lambda_1}, \lambda_0, \lambda_2, \lambda_2, \lambda_3$ *with* $|\lambda_1| \geq |\lambda_0| > |\lambda_2| \geq |\lambda_3|$.

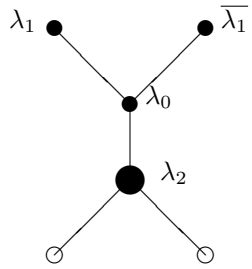

FIG. 7. *Tree graph of* $\hat{\boldsymbol{\lambda}}$ *with rank* 4 *when* $\lambda_2 = \overline{\lambda_2}$.

5. *if* $\sigma = n - \kappa$
   $\hat{\lambda}(J(idx : m+1)) = \mathbf{zeros}(1, m - idx + 2)$*;*
   go to 7
6. $k_\kappa = \mathbf{min}(\mathbf{find}(I(1,1,idx+1:m+1) == 1)) + idx$*;*
   $k_u = \mathbf{max}(\mathbf{find}(I(1,1,1:idx) == 1))$*;*
   if $I(:,:,idx) == [2,1]$
      if ~**isempty**(k_κ)
         $\hat{\lambda}(J(idx : m+1)) = \mathbf{zeros}(1, m - idx + 2)$*;*
         $\hat{\lambda}(J(k_\kappa)) = \lambda(J(idx))$*;*
      else
         $\hat{\lambda}(J(k_u)) = 0$*;*
         $\hat{\lambda}(J(idx+1 : m+1)) = \mathbf{zeros}(1, m - idx + 1)$*;*
      end
   else
      if ~**isempty**(k_κ)
         if **isempty**(k_u)
            $\hat{\lambda}(J(idx : m+1)) = \mathbf{zeros}(1, m - idx + 2)$*;*
            $\hat{\lambda}(J(k_\kappa)) = \lambda(J(k_\kappa))$*;*
         else
            $t_1 = 2 * \mathbf{abs}(\lambda(J(idx)))^2$*;*
            $t_2 = \mathbf{abs}(\lambda(J(k_u)))^2 + \mathbf{abs}(\lambda(J(k_\kappa)))^2$*;*
            if $t_1 \leq t_2$
               $\hat{\lambda}(J(idx : m+1)) = \mathbf{zeros}(1, m - idx + 2)$*;*
               $\hat{\lambda}(J(k_\kappa)) = \lambda(J(k_\kappa))$*;*

else
$\hat{\lambda}(J(idx+1:m+1)) = \mathbf{zeros}(1, m-idx+1)$*;*
$\hat{\lambda}(J(k_u)) = 0$*;*
end
end
else
$\hat{\lambda}(J(k_u)) = 0$*;*
$\hat{\lambda}(J(idx+1:m+1)) = \mathbf{zeros}(1, m-idx+1)$*;*
end
end

7. $\hat{\lambda} = \begin{cases} \left[\hat{\lambda}, \mathbf{fliplr}(\mathbf{conj}(\hat{\lambda}(2:m)))\right], & \text{if } n = 2m; \\ \left[\hat{\lambda}, \mathbf{fliplr}(\mathbf{conj}(\hat{\lambda}(2:m+1)))\right], & \text{if } n = 2m+1; \end{cases}$
$\hat{c} = \mathbf{real}(\mathbf{fft}(\hat{\lambda}))/n$*;*

5. Numerical examples. In this section, we illustrate our algorithm with some numerical examples. We report all numerics using only four significant digits, although entries of all matrices in consideration and the corresponding eigenvalues are originally the full length of the double precision. All calculations are done using MATLAB.

Example 1. Consider the 8×8 symmetric circulant matrix whose first row is given by a randomly generated vector

$$\mathbf{c} = [0.5404, 0.2794, 0.1801, -0.0253, -0.2178, -0.0253, 0.1801, 0.2794].$$

The corresponding eigenvalues, arranged in descending order of their moduli, are

$$\{1.1909, 1.1891, 1.1891, 0.3273, 0.3273, 0.1746, -0.0376, -0.0376\}.$$

The singular values clearly are given by the moduli of these eigenvalues. Observe the parity caused by the conjugate-evenness, whereas 1.1909 and 0.1746 are what we call absolutely real eigenvalues.

The nearest circulant approximation of rank 7 would be simply to set the last eigenvalue, i.e., -0.0376, to zero by using Algorithm 3.1, but such a TSVD approach would result in a complex matrix. To obtain the nearest real-valued circulant approximation of rank 7, we have to keep the pair of -0.0376 and zero out the value 0.1746. Using the conjugate-even eigenvalues

$$\hat{\boldsymbol{\lambda}} = [1.1909, 1.1891, -0.0376, 0.3273, 0, 0.3273, -0.0376, 1.1892],$$

we can construct the nearest real-valued rank-7 circulant approximation to $Circul(\mathbf{c})$ via the FFT and obtain the first row vector

$$\hat{\mathbf{c}} = [0.5186, 0.3012, 0.1583, -0.0035, -0.2396, -0.0035, 0.1583, 0.3012].$$

In yet another scenario, the first row of the nearest rank 4 circulant approximation is given by the row vector

$$\hat{\mathbf{c}} = [0.4871, 0.3182, 0.1898, -0.1023, -0.1075, -0.1023, 0.1898, 0.3182]$$

with eigenvalues $\hat{\boldsymbol{\lambda}}$

$$\hat{\boldsymbol{\lambda}} = [1.1909, 1.1892, 0, 0, 0.3273, 0, 0, 1.1892],$$

where we see that the last pair of eigenvalues in $\boldsymbol{\lambda}$ are set to zero while the value 0.1746 together with the value 0.3273 cause a topology change in the graph tree the same as in Figures 6 and 7.

Example 2. Consider the 9×9 circulant matrix whose first row is given by

$$\mathbf{c} = [1.6864, 1.7775, 1.9324, 2.9399, 1.9871, 1.7367, 4.0563, 1.2848, 2.5989].$$

The corresponding eigenvalues have conjugate-even structure given by

$$[20.0000, -2.8130 \pm 1.9106i, 3.0239 \pm 1.0554i, -1.3997 \pm 0.7715i, -1.2223 \pm 0.2185i].$$

Note the absolute real eigenvalue has modulus much larger than any other eigenvalues. To obtain a real-valued circulant approximation of rank 8, we have no choice but to select the vector (in its ordering)

$$\begin{aligned}\hat{\boldsymbol{\lambda}} = [&0, -2.8130 - 1.9106i, 3.0239 - 1.0554i,\\ &-1.3997 - 0.7715i, -1.2223 + 0.2185i, -1.2223 - 0.2185i,\\ &-1.3997 + 0.7715i, 3.0239 + 1.0554i, -2.8130 + 1.9106i]\end{aligned}$$

to produce

$$\hat{\mathbf{c}} = [-0.5358, -0.5872, -1.1736, -0.3212, 1.0198, 1.4013, -0.0761, -0.4115, 0.6844]$$

as the first row of its nearest real-valued circulant approximation. The fact that the *largest* eigenvalue (singular value) of *Circul*($\mathbf{c}$) must be set to zero to produce the nearest rank-8 approximation is quite counterintuitive to the usual sense of TSVD approximation.

On the other hand, it is worth noting that, if we slightly modify our approximation criteria by requesting only a nearest low rank approximation with rank *no greater than* 8, the answer could be completely different. In this particular example, such a nearest matrix turns out to be of rank 7 and is in agreement with the usual TSVD approximation by truncating the *pair* of eigenvalues with the smallest moduli.

Example 3. Let $C_\kappa \in \mathbb{R}^{n\times n}$ be a given circulant matrix of rank κ. With probability one, any random noise added to C_κ will destroy the circulant structure as well as the rank condition. To establish a comparison, we may assume, without loss of generality, that, after the projection step (3.1) mentioned in Algorithm 3.1, the added noise is a circulant matrix. Let $E \in \mathbb{R}^{n\times n}$ denote a random but fixed circulant matrix with unit Frobenius norm. Consider the perturbation of C_κ by an additive noise of magnitude (in Frobenius norm) 10^{-j}; i.e., consider the circulant matrices

$$W_j = C_\kappa + 10^{-j}E, \quad j = 1, \ldots, 12.$$

It is almost certain that, under such a random perturbation, the matrix W_j will be of full rank. Note that $\|W_j - C_\kappa\| = 10^{-j}$. It will be interesting to see if W_j has any closer circulant matrix approximation of rank κ than C_κ, especially when j is large.

Toward that end, we report a test case with $n = 100$, $\kappa = 73$, and a predetermined matrix C_{73}. In Figure 8, the (continuous) lines depict the distribution of singular values of the perturbed matrices W_j for $j = 1, \ldots, 12$, respectively, whereas the singular values of the original C_{73} are marked by $*$. Observe how the perturbation affects the last 27 (machine zero) singular values of C_{73} more significantly than the first 73 (larger) singular values according to the magnitude 10^{-j}.

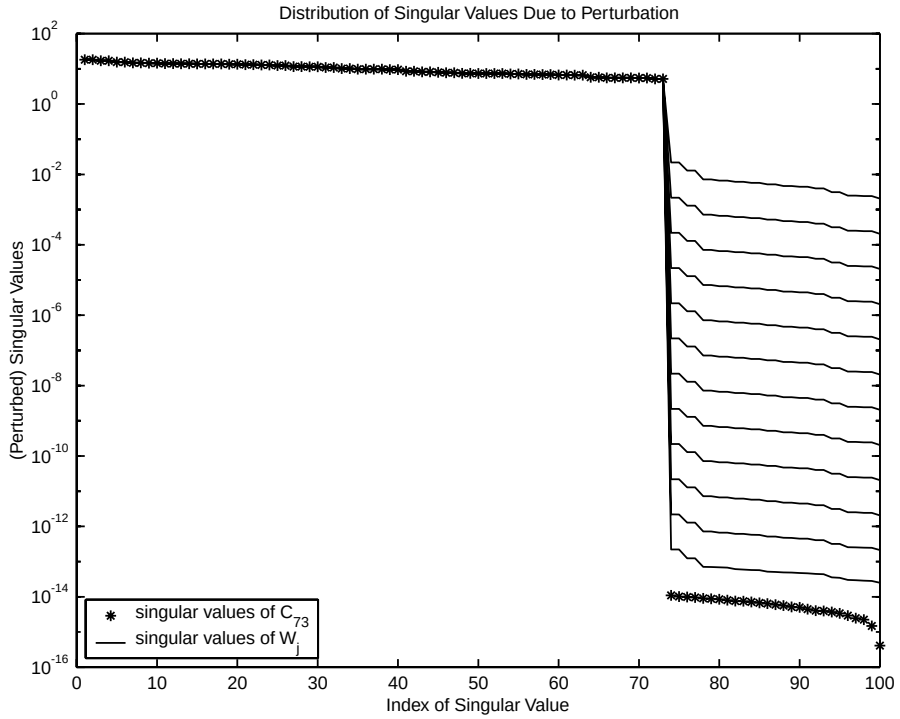

FIG. 8. *Distribution of singular values.*

Using our algorithm to find the best circulant approximation Z_j to W_j, we find that it is always the case that

$$\|W_j - Z_j\| < \|W_j - C_\kappa\|$$

for each j. This is indicated in Figure 9 by the fact that the circles $\circ$ are always below the diagonal line. Also marked in Figure 9 by $+$ signs is the difference between Z_j and C_κ.

6. Summary. Structured low rank approximation is an important and challenging task both theoretically and computationally. The special case of real-valued, low rank approximation with circulant structure is studied in this paper. For any given real data matrix, its nearest real circulant approximation can simply be determined from the average of its diagonal entries. Its nearest lower rank approximation can also be determined effectively from the TSVD and the FFT. However, such an approximation usually will be complex-valued. To simultaneously maintain the circulant structure, induce a specific lower rank, and keep the matrix real, the conjugate-even structure must be taken into account. These requirements, in turn, can substantially change the truncation criteria. Some counterintuitive examples illustrate the effect if the rank is fixed to a precise number and the approximation is required to be real-valued. We have proposed a fast algorithm to accomplish all of these approximation objectives.

Acknowledgments. The authors would like to thank Nikos Pitsianis and Xiaobai Sun for helpful comments that motivated much of this work.

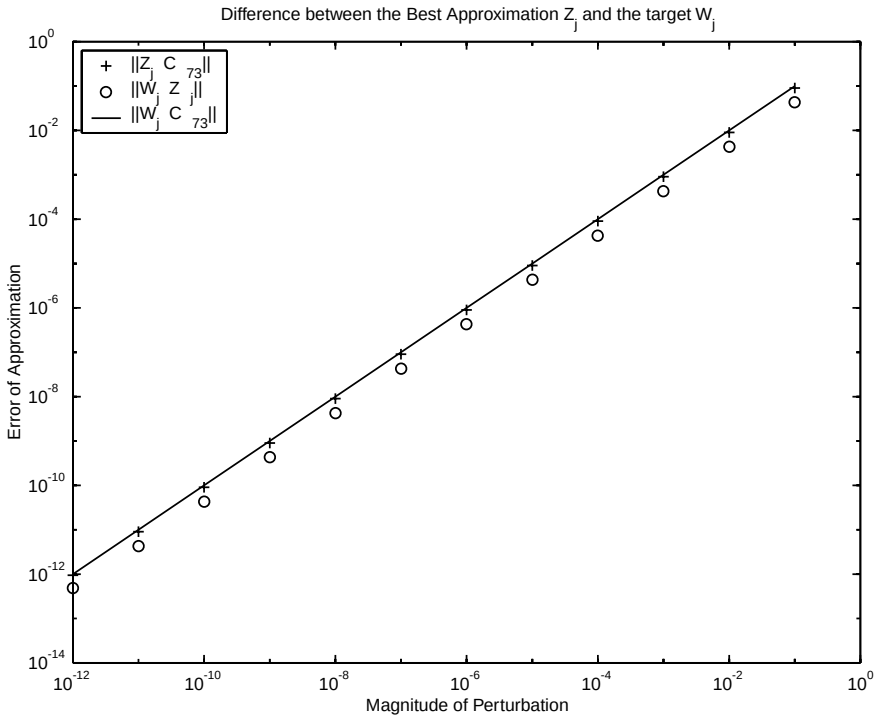

FIG. 9. *Errors in approximation.*

REFERENCES

[1] R. W. BROCKETT, *Least square matching problems*, Linear Algebra Appl., 122/123/124 (1989), pp. 761–777.

[2] R. H. CHAN, J. G. NAGY, AND R. J. PLEMMONS, *Circulant preconditioned Toeplitz least squares iterations*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 80–97.

[3] T. F. CHAN, *An optimal circulant preconditioner for Toeplitz systems*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 766–771.

[4] M. T. CHU AND K. R. DRIESSEL, *The projected gradient method for least squares matrix approximations with spectral constraints*, SIAM J. Numer. Anal., 27 (1990), pp. 1050–1060.

[5] M. T. CHU, R. E. FUNDERLIC, AND R. J. PLEMMONS, *Structured low rank approximation*, Linear Algebra Appl., to appear.

[6] P. J. DAVIS, *Circulant Matrices*, John Wiley and Sons, New York, 1979.

[7] B. L. ELLERBROEK AND F. J. RIGAUT, *Methods for correcting tilt anisoplanatism in laser-guide-star-based multi-conjugate adaptive optics*, J. Opt. Soc. Amer. A, 18 (2001), pp. 2539–2547.

[8] M. HANKE, J. G. NAGY, AND R. J. PLEMMONS, *Preconditioned iterative regularization for ill-posed problems*, in Numerical Linear Algebra and Scientific Computing, De Gruyter Press, Berlin, 1993, pp. 141–163.

[9] P. C. HANSEN, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*, SIAM Monogr. Math. Model. Comput. 4, SIAM, Philadelphia, PA, 1997.

[10] T. KAILATH, A. H. SAYED, AND B. HASSIBI, *Linear Estimation*, Prentice Hall, Upper Saddle River, NJ, 2000.

[11] A. FRIEZE AND R. KANNAA, *Fast Monte-Carlo Algorithm for Finding Low Rank Approximations*, available online from http://www.cs.yale.edu/~kannan/ (1998).

[12] J. G. NAGY, V. P. PAUCA, R. J. PLEMMONS, AND T. C. TORGERSEN, *Space-varying restoration of optical images*, J. Opt. Soc. Amer. A, 14 (1997), pp. 3162–3174.

[13] M. Ng, *Circulant preconditioners for convolution-like integral equations with higher order quadrature rules*, Electron. Trans. Numer. Anal., 6 (1997), pp. 18–28.
[14] H. D. Simon and H. Zha, *Low-rank matrix approximation using the Lanczos bidiagonalization process with applications*, SIAM J. Sci. Comput., 21 (2000), pp. 2257–2274.
[15] C. F. Van Loan, *Computational Frameworks for the Fast Fourier Transform*, Frontiers Appl. Math. 10, SIAM, Philadelphia, PA, 1992.