Solving block banded block Toeplitz systems with banded Toeplitz blocks

Dario Andrea Bini[a] and Beatrice Meini[a]

[a]Dipartimento di Matematica, Università di Pisa,
via Buonarroti 2, 56127 Pisa, Italy
bini,meini@dm.unipi.it

ABSTRACT

We introduce the concept of ϵ-displacement rank, that allows us to devise a fast algorithm for the approximate solution of BBBT/BTB (Block Banded Block Toeplitz with Banded Toeplitz Blocks) systems by means of cyclic reduction. We also introduce the concept of incomplete displacement block LU factorization of a Toeplitz-like matrix, where the displacement structure is imposed to the blocks of the factors L and U. The role of the matrix LU as preconditioner is discussed. Finally we propose another preconditioner obtained by extending a BBBT/BTB matrix to a banded Toeplitz matrix. Some open problems are addressed.

Keywords: Block Toeplitz matrices, displacement rank, cyclic reduction.

1. INTRODUCTION

Many important problems in the field of queueing theory, image restoration, computer algebra, etc. require the solution of linear systems of large (or even infinite) size, with matrices strongly structured. In most of these applications, the locality and the shift invariance properties of certain functions associated with the problem lead to block banded block Toeplitz matrices, accompanied with a banded Toeplitz structure of the blocks. We refer to such class of matrices as to BBBT/BTB (Block Banded Block Toeplitz with Banded Toeplitz Blocks).

A particularly interesting problem arises from image restoration, where the Point Spread Function (PSF) that defines the blur transformation is space invariant and has a local action. In this way, the restoration process is reduced to solving a BBBT/BTB system.

Due to the large block size of the matrices and to the large size of the blocks (for image processing the product of the two sizes is the number of the pixels of the image), it is mandatory to exploit both the outer (BBBT) and the inner (BTB) structures to devise efficient algorithms for the solution of these systems. Indeed, in many situations, algorithms that exploit the outer banded block Toeplitz structure can destroy the inner structure of the blocks, and vice-versa. Several techniques for the solution of BBBT systems have been introduced in the literature, in particular we recall the PCG methods,[1–6] the multigrid techniques,[7] and the algorithms based on cyclic reduction.[8–13]

Concerning the latter class, in Ref. 8–12 the cyclic reduction technique (CR) has been successfully applied to solve systems arising in the modeling of Markov chains with block Hessenberg block Toeplitz matrices and with banded Toeplitz matrices; the speed-up with respect to the available algorithms for the solution of these systems has reached a factor of several hundreds.[9]

In the banded Toeplitz case the matrix is reblocked into a block Toeplitz block tridiagonal matrix, with Toeplitz blocks, and the cyclic reduction is applied to solve the reblocked system[10]; at each step of CR we obtain a block tridiagonal matrix, that is block Toeplitz except for the block entry in position (1,1). An interesting feature of this algorithm is that, even if the Toeplitz structure of the block entries is in general lost just after the first step, their Toeplitz-like structure is kept in all the computation. This is the main reason of the large speed of the algorithm based on CR.

The conservation of the Toeplitz-like structure is expressed in terms of the conservation of the displacement rank of certain matrix polynomials that provide a functional description of cyclic reduction. The uniform bound of the displacement rank relies on suitable algebraic properties of these matrix polynomials.

Among the different displacement operators, defining the concept of displacement rank, we considered $\Delta(A) = AZ - ZA$ for $A \in \boldsymbol{R}^{n\times n}$, where $Z = (z_{i,j})$, $z_{i+1,i} = 1$, $z_{i,j} = 0$, is the down-shift matrix. We say that A has displacement rank r if $\mathrm{Rank}(\Delta(A)) = r$ and we write $\mathrm{Drk}(A) = r$. A matrix with displacement rank r can be easily represented in terms of at most $2r+1$ triangular Toeplitz matrices. We refer the reader to section 2.1 for more details. Informally we call *Toeplitz-like* a matrix A such that $\mathrm{Drk}(A)$ is much smaller than n.

A BBBT/BTB matrix may be reblocked into a block tridiagonal block Toeplitz matrix, where the blocks are Toeplitz-like matrices. In this way we can apply CR for solving the reblocked system. Unfortunately, the matrix polynomials generated by CR do not satisfy the suitable algebraic properties that guarantee the conservation of the displacement rank, as in the case of banded Toeplitz matrices. Indeed, the displacement rank of the blocks can generally grow exponentially at each step of CR, until the block size is reached. For this reason the CR approach might seem unsuitable for the solution of BBBT/BTB systems.

In this paper we show how to overcome the above drawback by introducing different tools that allow us to apply CR still keeping the displacement structure of the blocks. In general we are concerned with block tridiagonal block Toeplitz matrices that may arise from different problems, say, by reblocking BBBT/BTB matrices, where the blocks have a Toeplitz-like structure. To be more specific, we denote with Q_n the block tridiagonal block Toeplitz matrix

$$Q_n = \begin{bmatrix} A & C & & \bigcirc \\ B & A & \ddots & \\ & \ddots & \ddots & C \\ \bigcirc & & B & A \end{bmatrix} \tag{1}$$

of block size n, with blocks of size m.

We first introduce the concept of ϵ- *and relative* ϵ*-displacement rank*

$$\mathrm{Drk}_\epsilon(A) = \min_{||E||\le\epsilon} \mathrm{Rank}(\Delta(A)+E), \quad \mathrm{Rdrk}_\epsilon(A) = \min_{||E||\le\epsilon||A||} \mathrm{Rank}(\Delta(A)+E)$$

where $||\cdot||$ is the Euclidean norm, and relate it to the SVD of $\Delta(A)$. Then we prove that under suitable properties of the block tridiagonal block Toeplitz matrix Q_n the (relative) ϵ-displacement rank of the blocks generated by CR is bounded from above along all the computation.

This fact is used for designing an algorithm that approximates the solution of the linear system $Q_n\boldsymbol{x} = \boldsymbol{b}$ in $O(h^2 m \log^2 m \log n + hnm\log m)$ arithmetic operations (ops), where h is the largest ϵ-displacement rank encountered in the overall computation.

Moreover, we introduce the concept of *incomplete displacement block LU factorization* of Q_n that is a sort of extension of the well-known incomplete Cholesky factorization[14] (for sparse matrices) to dense structured matrices. In the latter factorization the sparsity pattern is imposed in the lower and upper triangular factors L and U; in the former, the Toeplitz-like structure, given by the displacement operator, is imposed to the blocks of the matrices L and U.

The computation of the factors L and U costs $O(K^2 m \log^2 m \log n)$ ops, where K is the displacement rank of the blocks, and the solution of a linear system with the matrix L or U, costs $O(Knm\log m)$ ops. Moreover, the matrix LU seems to keep the main spectral features of Q_n, thus it can be used as a preconditioner for the solution of a BBBT/BTB system by means of PCG methods.

In particular, if Q_n is obtained by reblocking a BBBT/BTB matrix, the spectral properties of the matrix LU seem to be strictly related to the spectral properties of the band Toeplitz matrix T defined by the first row and column of Q_n. The matrix T has very good computational properties since CR can be applied at the cost $O(m\log^2 m \log n)$ with full conservation of the displacement rank.[10] Thus it can be used as preconditioner for the implementation of PCG methods.

Interesting open questions are the analysis of the spectral properties of T and of the matrix LU obtained through an incomplete displacement factorization of Q_n.

As a byproduct of this analysis we provide also a simpler and more general proof of the conservation of the displacement rank of the blocks generated by CR, that gives more insights about possible extension of this kind of algorithms.

The paper is organized in the following way. In section 2 we recall the concept and the properties of displacement rank and introduce the concept of approximate displacement rank. We provide also a characterization of the latter concept in terms of SVD. In section 3 we summarize the main features of the CR algorithm and provide conditions for the conservation of displacement and ϵ-displacement rank of the blocks generated at each step. In section 4 we describe an algorithm for the approximate solution of BBBT/BTB systems based on the ϵ-displacement rank, an algorithm for the incomplete displacement factorization of Q_n and discuss a preconditioner based on the Toeplitz extension of a BBBT/BTB matrix.

2. TOOLS BASED ON DISPLACEMENT PROPERTIES

In this section we recall some basic tools for the design of efficient algorithms for solving certain structured systems.

2.1. Displacement rank

The concept of displacement operators and displacement rank, introduced in Ref. 15,16, and elaborated in Ref. 17,18, is a powerful tool for dealing with structured matrices. Here we recall the main ideas and results.

Let us consider the operator

$$\Delta(A) = AZ - ZA, \quad Z = (z_{i,j}), \quad z_{i+1,i} = 1, \quad z_{i,j} = 0, \text{ otherwise}, \tag{2}$$

applied to $n \times n$ matrices A, and define the displacement rank of A, associated with Δ, as $\mathrm{Drk}(A) = \mathrm{Rank}(\Delta(A))$. A Toeplitz matrix has displacement rank at most 2. We call a matrix Toeplitz-like if its displacement rank is independent of the size n.

Define $L(\boldsymbol{a})$ the $n \times n$ lower triangular Toeplitz matrix defined by its first column $\boldsymbol{a}$. It holds

$$\Delta(A) = \sum_{i=1}^{r} \boldsymbol{u}_i \boldsymbol{v}_i^T \quad \Leftrightarrow A = L(A\boldsymbol{e}_1) + \sum_{i=1}^{r} L(\boldsymbol{u}_i) L^T(Z\boldsymbol{v}_i), \tag{3}$$

where $\boldsymbol{u}_i$, $\boldsymbol{v}_i$, $i = 1, \ldots, r$, are n-dimensional vectors, and $\boldsymbol{e}_1$ denotes the first column of the identity matrix. The relation in the right-hand side of (3) is called the displacement representation of A. Equivalent representations can be given in terms of different operators.

A simple consequence of the relation

$$\Delta(A^{-1}) = -A^{-1}\Delta(A)A^{-1}, \tag{4}$$

that can be directly derived from (2), is that $\mathrm{Drk}(A^{-1}) = \mathrm{Drk}(A)$ and the following displacement representation of A^{-1} holds

$$A^{-1} = L(A^{-1}\boldsymbol{e}_1) - \sum_{i=1}^{r} L(A^{-1}\boldsymbol{u}_i) L^T(ZA^{-T}\boldsymbol{v}_i).$$

The above representations allow us to make elementary operations among Toeplitz-like matrices at a low cost by means of FFT. In particular, computing the product of matrices with displacement rank r costs $O(rn \log n)$ arithmetic operations (ops), inverting a strongly nonsingular matrix with displacement rank r costs $O(r^2 n \log^2 n)$ ops by using a superfast algorithm like the Bitmead Anderson algorithm.[19,20]

It is a simple matter to show that

$$\Delta(AB) = A\Delta(B) + \Delta(A)B, \tag{5}$$

therefore it holds $\mathrm{Drk}(AB) \leq \mathrm{Drk}(A) + \mathrm{Drk}(B)$. Similarly we have $\mathrm{Drk}(A + B) \leq \mathrm{Drk}(A) + \mathrm{Drk}(B)$.

2.2. Approximate displacement rank

The concept of displacement rank provides a means to measure, to a certain extent, the distance of a given matrix from the class of lower triangular Toeplitz matrices; in fact $\mathrm{Drk}(T) = 0$ for any lower triangular Toeplitz matrix T. In certain situations it may be useful to measure this distance in an approximate way.

Definition 2.1. *Let $||\cdot||$ denote the Euclidean norm. For a given $\epsilon > 0$ define the ϵ-displacement rank of a matrix A as*

$$\mathrm{Drk}_\epsilon(A) = \min_{||E|| \leq \epsilon} \mathrm{Rank}(\Delta(A) + E),$$

and the relative ϵ-displacement rank as

$$\mathrm{Rdrk}_\epsilon(A) = \min_{||E|| \le \epsilon ||A||} \mathrm{Rank}(\Delta(A) + E).$$

Thus, from the above definition, if $r = \mathrm{Drk}_\epsilon(A)$ ($r = \mathrm{Rdrk}_\epsilon(A)$), and if $\Delta(A) = \sum_{i=1}^r \boldsymbol{u}_i \boldsymbol{v}_i^T + E$, with $||E|| \le \epsilon$ ($||E|| \le \epsilon ||A||$), we say that the matrix

$$A_\epsilon = L(A\boldsymbol{e}_1) + \sum_{i=1}^r L(\boldsymbol{u}_i) L^T(Z\boldsymbol{v}_i)$$

is an ϵ-displacement (relative ϵ-displacement) representation of A, and that the vectors $\boldsymbol{u}_i$, $\boldsymbol{v}_i$, $i = 1, \ldots, r$, are ϵ-displacement (relative ϵ-displacement) generators for A.

The following properties hold.

THEOREM 2.2. *Let $\Delta(A) = U\Sigma V^T$ be the singular values decomposition of $\Delta(A)$, where $\Sigma = \mathrm{Diag}(\sigma_1, \ldots, \sigma_n)$ and $\sigma_1 \ge \ldots \ge \sigma_n \ge 0$ are the singular values of $\Delta(A)$. Then $\mathrm{Drk}_\epsilon(A) = r$ if and only if $\sigma_r > \epsilon \ge \sigma_{r+1}$. Moreover $\mathrm{Rdrk}_\epsilon(A) = r$ if and only if $\sigma_r > \epsilon ||A|| \ge \sigma_{r+1}$.*

Proof. We prove that $\sigma_{r+1} \le \epsilon$ if and only if $\mathrm{Drk}_\epsilon(A) \le r$. This is sufficient to prove the theorem. Let $\boldsymbol{u}_i$, $\boldsymbol{v}_i$, $i = 1, \ldots, n$, be the i-th column of U and V, respectively, that is

$$\Delta(A) = \sum_{i=1}^r \boldsymbol{u}_i \sigma_i \boldsymbol{v}_i^T + \sum_{i=r+1}^n \boldsymbol{u}_i \sigma_i \boldsymbol{v}_i^T. \tag{6}$$

Define $E = \sum_{i=r+1}^n \boldsymbol{u}_i \sigma_i \boldsymbol{v}_i^T$, then $||E|| = \sigma_{r+1}$. Thus, if $\sigma_{r+1} \le \epsilon$, we have $\mathrm{Drk}_\epsilon(A) \le \mathrm{Drk}(\Delta(A) - E) = r$. Vice-versa, if $\mathrm{Drk}_\epsilon(A) = r$ then there exists a matrix E, with $||E|| \le \epsilon$, such that $\mathrm{Rank}(\Delta(A) + E) = r$. Define $\theta_r = \min_{\{X:\ \mathrm{Rank}(X) = r\}} ||\Delta(A) - X||$; since $\theta_r = \sigma_{r+1}$ and $\sigma_{r+1} = ||\Delta(A) - \sum_{i=1}^r \boldsymbol{u}_i \sigma_i \boldsymbol{v}_i^T||$, it follows that $\sigma_{r+1} = \theta_r \le ||E||$. □

The concept of ϵ-displacement rank has been implicitly used in Ref. 21 for implementing the Newton iteration for inverting Toeplitz matrices (compare also Ref. 18, chapter 4, section 8).

The following result provides a bound to the error in the ϵ-displacement representation of A, where $||A||_1 = \max_j \sum_i |a_{i,j}|$ denotes the 1-norm of A.

THEOREM 2.3. *Let $r = \mathrm{Drk}_\epsilon(A)$, and let A_ϵ be an ϵ-displacement representation of A. Then it holds*

$$||A - A_\epsilon||_1 \le \sum_{i=r+1}^n \sigma_i \le (n - r)\epsilon.$$

Similarly, if $r = \mathrm{Rdrk}_\epsilon(A)$ and if A_ϵ is a relative ϵ-displacement representation of A, then it holds

$$||A_\epsilon - A||_1 \le \sum_{i=r+1}^n \sigma_i \le (n - r)\epsilon ||A||.$$

Proof. Let $\Delta(A) = U\Sigma V^T$ be the singular values decomposition of $\Delta(A)$. From theorem 2.2 and (6) it holds $A - A_\epsilon = \sum_{i=r+1}^n L(\boldsymbol{u}_i) L^T(\sigma_i Z \boldsymbol{v}_i)$, where $||\boldsymbol{u}_i| = ||\boldsymbol{v}_i|| = 1$. Since $||L(\boldsymbol{u})||_1 = ||L^T(\boldsymbol{u})||_1 = ||\boldsymbol{u}||_1 \le ||\boldsymbol{u}||$, and $||Z\boldsymbol{u}|| \le ||\boldsymbol{u}||$ for any vector $\boldsymbol{u}$, we have $||A - A_\epsilon||_1 \le \sum_{i=r+1}^n \sigma_i$. The proof is completed since $|\sigma_i| \le \epsilon$ for $r = \mathrm{Drk}_\epsilon(A)$ and $|\sigma_i| \le \epsilon ||A||$ for $r = \mathrm{Rdrk}_\epsilon(A)$, $i = r+1, \ldots, n$. □

The following inequalities are an immediate consequence of the definition:

$$\begin{array}{ll}
\mathrm{Drk}_\delta(A + B) \le \mathrm{Drk}_\epsilon(A) + \mathrm{Drk}_\epsilon(B), & \delta = 2\epsilon \\
\mathrm{Drk}_\delta(AB) \le \mathrm{Drk}_\epsilon(A) + \mathrm{Drk}_\epsilon(B), & \delta = \epsilon(||A|| + ||B||) \\
\mathrm{Rdrk}_\delta(AB) \le \mathrm{Rdrk}_\epsilon(A) + \mathrm{Rdrk}_\epsilon(B), & \delta = \epsilon \frac{||A|| \cdot ||B||}{||AB||} \\
\mathrm{Rdrk}_\delta(A^{-1}) \le \mathrm{Rdrk}_\epsilon(A), & \delta = \epsilon \mathrm{Cond}(A),
\end{array} \tag{7}$$

where $\mathrm{Cond}(A) = ||A|| \cdot ||A^{-1}||$.

3. CYCLIC REDUCTION

Consider the system

$$Q_n \boldsymbol{x} = \boldsymbol{b}, \tag{8}$$

where Q_n, that is defined in (1), has block size $n = 2^q$. Partition the vectors $\boldsymbol{x}$ and $\boldsymbol{b}$ into blocks $\boldsymbol{x}_i$, $\boldsymbol{b}_i$ of dimension m, respectively, that is, $\boldsymbol{x} = (\boldsymbol{x}_i)_{i=1,\ldots,2^q}$, $\boldsymbol{b} = (\boldsymbol{b}_i)_{i=1,\ldots,2^q}$. By performing an even-odd permutation of the block rows and block columns in (8) we find that

$$\begin{bmatrix} D_1^{(0)} & U^{(0)} \\ L^{(0)} & D_2^{(0)} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_+^{(0)} \\ \boldsymbol{x}_-^{(0)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_+^{(0)} \\ \boldsymbol{b}_-^{(0)} \end{bmatrix} \tag{9}$$

where $\boldsymbol{x}_+^{(0)} = (\boldsymbol{x}_{2k})$, $\boldsymbol{x}_-^{(0)} = (\boldsymbol{x}_{2k-1})$, $\boldsymbol{b}_+^{(0)} = (\boldsymbol{b}_{2k})$, $\boldsymbol{b}_-^{(0)} = (\boldsymbol{b}_{2k-1})$;

$$D_1^{(0)} = D_2^{(0)} = \begin{bmatrix} A & & \bigcirc \\ & \ddots & \\ \bigcirc & & A \end{bmatrix}, \quad L^{(0)} = \begin{bmatrix} C & & & \bigcirc \\ B & C & & \\ & \ddots & \ddots & \\ \bigcirc & & B & C \end{bmatrix}, \quad U^{(0)} = \begin{bmatrix} B & C & & \bigcirc \\ & B & \ddots & \\ & & \ddots & C \\ \bigcirc & & & B \end{bmatrix}.$$

Applying one step of block-Gaussian elimination to the 2×2 block system (9) yields the equivalent system

$$\begin{cases} (D_2^{(0)} - L^{(0)} D_1^{(0)^{-1}} U^{(0)}) \boldsymbol{x}_-^{(0)} = \boldsymbol{b}_-^{(0)} - L^{(0)} D_1^{(0)^{-1}} \boldsymbol{b}_+^{(0)} \\ \boldsymbol{x}_+^{(0)} = D_1^{(0)^{-1}} (\boldsymbol{b}_+^{(0)} - U^{(0)} \boldsymbol{x}_-^{(0)}). \end{cases} \tag{10}$$

Denoting

$$Q_n^{(1)} = D_2^{(0)} - L^{(0)} D_1^{(0)^{-1}} U^{(0)} \tag{11}$$

the Schur complement of $D_2^{(0)}$, $\boldsymbol{x}^{(1)} = \boldsymbol{x}_-^{(0)}$, $\boldsymbol{b}^{(1)} = \boldsymbol{b}_-^{(0)} - L^{(0)} D_1^{(0)^{-1}} \boldsymbol{b}_+^{(0)}$, the above system is ultimately reduced to solving

$$Q_n^{(1)} \boldsymbol{x}^{(1)} = \boldsymbol{b}^{(1)}. \tag{12}$$

Once $\boldsymbol{x}^{(1)}$ has been computed, the solution of the original system (8) can be recovered by means of back substitution through (10).

It is easy to observe that $Q_n^{(1)}$ is a block tridiagonal matrix of block size 2^{q-1} which, except for the north western corner block, has the block Toeplitz structure. If we apply to system (12) the same block even-odd permutation, followed by one step of Gaussian elimination (i.e., one step of cyclic reduction), we obtain a new system, whose matrix $Q_n^{(2)}$ is a block tridiagonal matrix which, except for the north western corner block, has the block Toeplitz structure. In this way we obtain a sequence of systems $\{Q_n^{(j)} \boldsymbol{x}^{(j)} = \boldsymbol{b}^{(j)}\}$, of block size 2^{q-j}, $j = 0, \ldots, q$, where $Q_n^{(0)} = Q_n$, and

$$Q_n^{(j)} = \begin{bmatrix} F^{(j)} & C^{(j)} & & \bigcirc \\ B^{(j)} & A^{(j)} & \ddots & \\ & \ddots & \ddots & C^{(j)} \\ \bigcirc & & B^{(j)} & A^{(j)} \end{bmatrix}.$$

The block entries $A^{(j+1)}$, $B^{(j+1)}$, $C^{(j+1)}$, $F^{(j+1)}$, obtained at step $j+1$, that define the matrix $Q_n^{(j+1)}$, are related to the block entries $A^{(j)}$, $B^{(j)}$, $C^{(j)}$, $F^{(j)}$, obtained at step j, by means of the following simple relations:

$$\begin{cases} B^{(j+1)} = -B^{(j)} A^{(j)^{-1}} B^{(j)} \\ A^{(j+1)} = A^{(j)} - B^{(j)} A^{(j)^{-1}} C^{(j)} - C^{(j)} A^{(j)^{-1}} B^{(j)} \\ C^{(j+1)} = -C^{(j)} A^{(j)^{-1}} C^{(j)} \\ F^{(j+1)} = F^{(j)} - C^{(j)} A^{(j)^{-1}} B^{(j)}. \end{cases} \tag{13}$$

The above relations can be carried out provided that the block $A^{(j)}$ is nonsingular. This property is guaranteed if the leading principal submatrix of Q_n of block size 2^j is nonsingular. In particular this nonsingularity holds if Q_n is diagonally dominant, or symmetric and positive (negative) definite. In the latter case, for the Cauchy interlace theorem, the spectral condition number of the diagonal blocks $A^{(j)}$ does not exceed the spectral condition number of Q_n. Moreover, in the symmetric case, even though the matrix is not definite, the blocks $A^{(j)}$, $F^{(j)}$ are symmetric and $B^{(j)T} = C^{(j)}$ for any j.

Observe that CR provides the block LU factorization of $\Pi Q_n \Pi^T$, where Π is the permutation matrix resulting from the composition of the even-odd permutations performed at all the levels of CR.

3.1. Convergence properties

Convergence properties of the blocks generated at each step of cyclic reduction have been investigated by several authors[13,22] in terms of diagonal dominance of the matrix Q_n. Here we give some convergence properties, related to the location of the zeros of the polynomial $p(z) = \det(C + zA + z^2B)$.

THEOREM 3.1. *Let ξ_i, $i = 1, \ldots, 2m$, be the zeros of $p(z)$ ordered such that $|\xi_1| \le |\xi_2| \le \ldots \le |\xi_{2m}|$, where we assume that $p(z)$ has zeros equal to infinity, if its degree is less than $2m$. Assume that $|\xi_m| < 1 < |\xi_{m+1}|$ and the matrix polynomials $\phi_1(z) = C + zA + z^2B$ and $\phi_2(z) = z^2C + zA + B$ have two nonsingular Jordan chains of dimension m corresponding to $\xi_1, \ldots, \xi_m$, and to $1/\xi_{m+1}, \ldots, 1/\xi_{2m}$, respectively. Then the sequences $||A^{(j)}||$, $||{A^{(j)}}^{-1}||$ are bounded and $||C^{(j)}|| = O(|\xi_m|^{2^j})$, $||B^{(j)}|| = O((1/|\xi_{m+1}|)^{2^j})$.*

Proof. For the Jordan chains properties[23,24] there exist $m \times m$ matrices H and K that solve the matrix equations $C + HA + H^2B = 0$ and $K^2C + KA + B = 0$ whose eigenvalues, counting multiplicities, are $\xi_1, \ldots, \xi_m$ and $1/\xi_{m+1}, \ldots, 1/\xi_{2m}$, respectively. It can be proved that[11] at each step j of cyclic reduction, it holds $C^{(j)}{A^{(j)}}^{-1} + H^{2^j} + H^{2 \cdot 2^j} B^{(j)}{A^{(j)}}^{-1} = 0$, $K^{2 \cdot 2^j} C^{(j)}{A^{(j)}}^{-1} + K^{2^j} + B^{(j)}{A^{(j)}}^{-1} = 0$. From the latter equations it follows that $||C^{(j)}{A^{(j)}}^{-1}|| = O(|\xi_m|^{2^j})$ and $||B^{(j)}{A^{(j)}}^{-1}|| = O((1/|\xi_{m+1}|)^{2^j})$. Hence, the sequences $||A^{(j)}||$, $||{A^{(j)}}^{-1}||$ are bounded[22] and therefore $||C^{(j)}|| = O(|\xi_m|^{2^j})$, $||B^{(j)}|| = O((1/|\xi_{m+1}|)^{2^j})$. □

It is important to point out that, for the latter theorem, the non diagonal blocks converge to zero, even if the starting matrix Q_n is not diagonally dominant. Moreover, the condition $|\xi_m| < 1 < |\xi_{m+1}|$ is not restrictive, since, if $|\xi_m| < |\xi_{m+1}|$, we can find a block diagonal matrix $D = \mathrm{Diag}(\theta^i I, i = 0, \ldots, n-1)$, such that the zeros associated with the block tridiagonal matrix $\widehat{Q}_n = DQ_nD^{-1}$ satisfy the condition $|\hat{\xi}_m| < 1 < |\hat{\xi}_{m+1}|$, and we can apply cyclic reduction to the matrix $\widehat{Q}_n$.

The following theorem gives conditions under which the blocks $A^{(j)}$, $B^{(j)}$, $C^{(j)}$ have bounded norm, in the case where $C = B^T$ and $A = A^T$, i.e., Q_n is symmetric.

THEOREM 3.2. *Assume that $A = A^T$ and $C = B^T$. Consider the set $\mathcal{S} = \{\lambda : \ \det(z^{-1}B + A + zC - \lambda I) = 0, |z| = 1\}$. If $0 \notin \mathcal{S}$ then $||A^{(j)}||$, $||B^{(j)}||$, $||F^{(j)}||$, $||{A^{(j)}}^{-1}||$, $||{F^{(j)}}^{-1}||$ are bounded from above by a constant. Moreover the polynomial $p(z) = \det(C + zA + z^2B)$ has m roots inside and m roots outside the unit disk.*

Proof. Let $\theta_1 = \inf \mathcal{S}$, $\theta_2 = \sup \mathcal{S}$. From Ref. 25 it follows that $\theta_1 < \mu < \theta_2$ for any eigenvalue μ of Q_n, for any n. For the Cauchy interlace theorem and from the properties of the Schur complements[14] it follows that the Schur complement of any principal submatrix of Q_n has eigenvalues in (θ_1, θ_2) for any n. Since the block tridiagonal matrices defined by the blocks $A^{(j)}$, $B^{(j)}$, $B^{(j)T}$, $F^{(j)}$ are Schur complements, also their eigenvalues belong to (θ_1, θ_2). This implies the boundness of the blocks $A^{(j)}$, $B^{(j)}$, $F^{(j)}$; analogously, we may prove the boundness of ${A^{(j)}}^{-1}$ and ${F^{(j)}}^{-1}$. Since $0 \notin \mathcal{S}$ we have $p(z) \neq 0$ for $|z| = 1$. Moreover from the symmetricity of Q_n it follows that if $p(\xi) = 0$ then $p(\xi^{-1}) = 0$. This fact completes the proof. □

In the symmetric case the condition $0 \notin \mathcal{S}$ guarantees the boundness of the blocks and the convergence to zero, if it holds, depends on the modulus of the zero ξ closest to the unit circle.

Another important situation where the blocks generated by cyclic reduction have bounded norm is the case where A is a nonsingular M-matrix, $B \ge 0$, $C \ge 0$, and $A + B + C$ is an irreducible singular M-matrix; this case occurs for example in the analysis of Markov chains arising in queueing problems.[11,9]

A nice consequence of theorem 3.1 is that the block tridiagonal matrix $Q_n^{(j)}$ generated after $j = O(\log\log\epsilon^{-1})$ steps can be approximated, within the error $O(\epsilon)$, by a block (bi-)diagonal matrix. In particular, if the cyclic reduction algorithm is applied to solve the finite system $Q_n \boldsymbol{x} = \boldsymbol{b}$, where Q has block size $n = 2^p$ and $p >> \log\log \epsilon^{-1}$, it is sufficient to perform $j = O(\log\log \epsilon^{-1})$ steps of cyclic reduction, and then to compute the solution of $Q_n^{(j)} \boldsymbol{x}^{(j)} = \boldsymbol{b}^{(j)}$, within the error ϵ, by solving a block (bi-)diagonal system, and then to recover by back substitution the remaining components.

These feature is fundamental for the computation of a finite number of components of a (bi-)infinite block tridiagonal block Toeplitz system.[8,11,10]

The convergence properties will be used in section 3.3 to prove the invariance of the approximate displacement rank of the blocks.

3.2. Conditions for structure invariance

It is possible to describe the cyclic reduction process in functional form[9,10] by means of the matrix polynomial $\phi^{(j)}(z) = C^{(j)} + zA^{(j)} + z^2 B^{(j)}$, $j \geq 0$. From (11) we may easily deduce that $\phi^{(j+1)}(z) = zA^{(j)} - (C^{(j)} + zB^{(j)}) A^{(j)^{-1}} (C^{(j)} + zB^{(j)})$ and that $\phi^{(j+1)}(z^2) = \left(\frac{\phi^{(j)}(z)^{-1} - \phi^{(j)}(-z)^{-1}}{2z} \right)^{-1} = -\phi^{(j)}(z) A^{(j)^{-1}} \phi^{(j)}(-z)$. Thus, if we introduce the formal matrix power series $\psi^{(j)}(z) = \phi^{(j)}(z)^{-1}$, we have

$$\psi^{(j+1)}(z^2) = \frac{\psi^{(j)}(z) - \psi^{(j)}(-z)}{2z}, \tag{14}$$

that is, the matrix coefficients of $\psi^{(j+1)}(z)$ are the odd matrix coefficients of $\psi^{(j)}(z)$. In other words, the matrix power series $\psi^{(j)}(z)$ are determined by matrix coefficients that are suitable coefficients of the matrix power series $\psi(z) = \psi^{(0)}(z)$.

Relation (14) is useful to prove structural properties of the block entries obtained at each step of cyclic reduction. Indeed, from (4) it holds that for any z, $\psi^{(j)}(z)$ and $\phi^{(j)}(z)$ have the same displacement rank. Unfortunately this does not imply that $\mathrm{Drk}(\psi^{(j)}(z)) = \mathrm{Drk}(\psi(z))$ and thus that $\mathrm{Drk}(\phi^{(j)}(z)) = \mathrm{Drk}(\psi(z))$ for any j. However, under suitable conditions on $\phi(z)$ the former property is satisfied. This occurs for instance if $\phi(z)$ belongs to the same matrix algebra for any z, or even under weaker conditions.

It is important to point out that if the matrix polynomials $\phi^{(j)}(z)$ have a small displacement rank for any j and z, then the process of cyclic reduction can be implemented at a very low cost with the help of FFT, and with a low memory storage.

For banded Toeplitz matrices reblocked into block tridiagonal form the matrix polynomial $\phi^{(0)}(z)$ is z-circulant for any z.[10] The class of z-circulant matrices is the algebra generated by the matrix $Z + \boldsymbol{e}_1 z \boldsymbol{e}_n^T$, where $\boldsymbol{e}_n$ is the last column of the identity matrix. This class is closed under inversion, thus $\psi^{(0)}(z) = \phi^{(0)}(z)^{-1}$ is z-circulant. Since z-circulant matrices are also Toeplitz, and the latter class is a linear space, it follows that $\psi^{(j)}(z)$ is Toeplitz for any z and j, whence $\phi^{(j)}(z) = \psi^{(j)}(z)^{-1}$ has displacement rank at most 2. This property has been proved in Ref. 10 by using an induction argument. The proof given above is easier and provides more insights for possible generalizations.

We point out that if $\psi^{(0)}(z)$ has a linear space structure for any z except possibly for a small rank corrections having a kernel independent of z, then the matrix polynomials $\phi^{(j)}(z)$ have the bilinear structure (say the displacement structure) associated with the linear space, except for a small rank correction. This makes the structural properties an important tool for the design of efficient algorithms. An example of such an analysis can be found in Ref. 26 where the Frobenius algebra plays an important role.

3.3. Conditions for approximated structure invariance

Suppose that the blocks A, B and C, defining Q_n in (1) are Toeplitz-like matrices. Without specific assumptions on their displacement structure, from formulae (13) and (7) it follows that the approximate displacement rank of the blocks $A^{(j)}$, $B^{(j)}$, $C^{(j)}$, $F^{(j)}$ should roughly triplicate at each step of cyclic reduction.

However, the concept of approximate displacement rank is very useful if the blocks have some convergence properties. More precisely, if the blocks $A^{(j)}$, $B^{(j)}$, $C^{(j)}$ have a double exponential convergence, then after $O(\log\log \epsilon^{-1})$ iterations of CR, the blocks obtained at two subsequent steps have the same (relative) ϵ-displacement rank. In this way, according to theorem 2.3, if the size m is such that $m >> \log\log \epsilon^{-1}$, then after $O(\log\log \epsilon^{-1})$ steps the blocks can be approximated, within the error $O(\epsilon)$, by matrices with small displacement rank with respect to their size m.

In the particular case of symmetric matrices, if the hypothesis of theorem 3.2 are satisfied, then after $O(\log\log \epsilon^{-1})$ steps of cyclic reduction it holds $||B^{(j)}|| < \epsilon$, thus $\mathrm{Drk}_\epsilon(B^{(j)}) = 0$.

4. ALGORITHMS, APPLICATIONS AND OPEN PROBLEMS

In this section we show how the cyclic reduction algorithm, together with the concepts of displacement and approximate displacement rank, can be used to derive algorithms to compute approximate solutions, or to compute suitable preconditioners, for BBBT/BTB problems.

4.1. ϵ-rank algorithm: approximate solution

Here we describe an algorithm based on the ϵ-displacement rank for the approximate solution of the system $Q_n\boldsymbol{x} = \boldsymbol{b}$, where Q_n is a block tridiagonal matrix, with blocks having small displacement rank. This case occurs for example in the solution of BBBT/BTB systems, where the matrix of the system is reblocked into a block tridiagonal matrix.

The algorithm consists in applying the cyclic reduction algorithm, described in section 3, where at each step j, the matrices $A^{(j)}$, $B^{(j)}$, $C^{(j)}$, $F^{(j)}$, are approximated with matrices $\widetilde{A}^{(j)}$, $\widetilde{B}^{(j)}$, $\widetilde{C}^{(j)}$, $\widetilde{F}^{(j)}$, such that $\mathrm{Rdrk}_{\epsilon_j}(A^{(j)}) \leq \mathrm{Drk}(\widetilde{A}^{(j)})$, $\mathrm{Rdrk}_{\epsilon_j}(B^{(j)}) \leq \mathrm{Drk}(\widetilde{B}^{(j)})$, $\mathrm{Rdrk}_{\epsilon_j}(C^{(j)}) \leq \mathrm{Drk}(\widetilde{C}^{(j)})$, $\mathrm{Rdrk}_{\epsilon_j}(F^{(j)}) \leq \mathrm{Drk}(\widetilde{F}^{(j)})$, for a suitable ϵ_j. More precisely, at the j-th step of cyclic reduction, instead of dealing with the exact blocks $A^{(j)}$, $B^{(j)}$, $C^{(j)}$, $F^{(j)}$, we perform computations with the corresponding relative approximate displacement representations and, in the light of theorem 2.2, we reduce the approximate displacement rank representation of the new blocks by means of their relative ϵ-displacement representation, where ϵ is a fixed error bound. That is, at the $(j+1)$-th step we compute the matrices $\widetilde{A}^{(j+1)}$, $\widetilde{B}^{(j+1)}$, $\widetilde{C}^{(j+1)}$, $\widetilde{F}^{(j+1)}$ such that

$$\begin{array}{ll} \mathrm{Drk}(\widetilde{A}^{(j+1)}) = \mathrm{Rdrk}_\epsilon(\widehat{A}^{(j+1)}), & \mathrm{Drk}(\widetilde{B}^{(j+1)}) = \mathrm{Rdrk}_\epsilon(\widehat{B}^{(j+1)}), \\ \mathrm{Drk}(\widetilde{C}^{(j+1)}) = \mathrm{Rdrk}_\epsilon(\widehat{C}^{(j+1)}), & \mathrm{Drk}(\widetilde{F}^{(j+1)}) = \mathrm{Rdrk}_\epsilon(\widehat{F}^{(j+1)}), \end{array} \tag{15}$$

where

$$\begin{array}{ll} \widehat{A}^{(j+1)} = \widetilde{A}^{(j)} - \widetilde{B}^{(j)}\widetilde{A}^{(j)^{-1}}\widetilde{C}^{(j)} - \widetilde{C}^{(j)}\widetilde{A}^{(j)^{-1}}\widetilde{B}^{(j)}, & \widehat{B}^{(j+1)} = -\widetilde{B}^{(j)}\widetilde{A}^{(j)^{-1}}\widetilde{B}^{(j)} \\ \widehat{C}^{(j+1)} = -\widetilde{C}^{(j)}\widetilde{A}^{(j)^{-1}}\widetilde{C}^{(j)}, & \widehat{F}^{(j+1)} = \widetilde{F}^{(j)} - \widetilde{C}^{(j)}\widetilde{A}^{(j)^{-1}}\widetilde{B}^{(j)} \end{array} \tag{16}$$

and where $\widetilde{A}^{(0)} = A^{(0)}$, $\widetilde{B}^{(0)} = A^{(0)}$, $\widetilde{C}^{(0)} = C^{(0)}$, $\widetilde{F}^{(0)} = F^{(0)}$. For simplicity we give an outline of the algorithm that describes the generic j-th step of cyclic reduction with the ϵ-displacement approximation.

Algorithm.

INPUT: An error bound ϵ and a displacement representation of $\widetilde{A}^{(j)}$, $\widetilde{B}^{(j)}$, $\widetilde{C}^{(j)}$, $\widetilde{F}^{(j)}$, that approximate (in the sense of relative approximate displacement rank) $A^{(j)}$, $B^{(j)}$, $C^{(j)}$, $F^{(j)}$, respectively.

OUTPUT: A displacement representation of the matrices $\widetilde{A}^{(j+1)}$, $\widetilde{B}^{(j+1)}$, $\widetilde{C}^{(j+1)}$, $\widetilde{F}^{(j+1)}$, such that (15) holds.

COMPUTATION:

1. By using the displacement representation of $\widetilde{A}^{(j)}$, compute the displacement representation of $\widetilde{A}^{(j)^{-1}}$, by means of (4).
2. From the latter, and from the displacement representation of $\widetilde{B}^{(j)}$ compute a displacement representation of $\widehat{B}^{(j+1)} = -\widetilde{B}^{(j)}\widetilde{A}^{(j)^{-1}}\widetilde{B}^{(j)}$, according to formula (5).

   Let h be the rank of such representation and S_1 and S_2 be the $m \times h$ matrices whose columns define a relative ϵ-representation, i.e., such that $\Delta(\widehat{B}^{(j+1)}) = S_1 S_2^T + E$, $||E|| \leq \epsilon||\widehat{B}^{(j+1)}||$.
3. Compute the QR factorizations $S_1 = Q_1R_1$, $S_2 = Q_2R_2$ and the $h \times h$ matrix $W = R_1R_2^T$. Compute the singular value decomposition $W = U\Sigma V^T$ of W and denote with $\sigma_1 \geq \cdots \geq \sigma_h$ the singular values of W. Let r be such that $\sigma_r > \epsilon||\widehat{B}^{(j+1)}|| \geq \sigma_{r+1}$ and denote with $\widehat{U}$, $\widehat{V}$, the matrices made up by the first r columns of U and V, respectively. Similarly denote with $\widehat{\Sigma}$ the leading $r \times r$ principal submatrix of Σ.
4. Define the matrix $\widetilde{B}^{(j+1)}$ as the one represented by the displacement representation $Q_1\widehat{U}\widehat{\Sigma}$, $Q_2\widehat{V}$, i.e., $\widetilde{B}^{(j+1)}$ is the matrix such that $\Delta(\widetilde{B}^{(j+1)}) = (Q_1\widehat{U}\widehat{\Sigma})(Q_2\widehat{V})^T$. This matrix provides a relative ϵ-displacement representation of $\widehat{B}^{(j+1)}$.
5. Perform similar computations for the matrices $\widehat{A}^{(j+1)}$, $\widehat{C}^{(j+1)}$, $\widehat{F}^{(j+1)}$.

	$\epsilon = 10^{-2}$		$\epsilon = 10^{-4}$		$\epsilon = 10^{-6}$	
m	Drk_ϵ	Rdrk_ϵ	Drk_ϵ	Rdrk_ϵ	Drk_ϵ	Rdrk_ϵ
64	12	12	20	22	28	34
128	12	12	20	22	28	34
256	12	12	20	22	28	34

Table 1. $\alpha = 8$

m	Drk_ϵ	Rdrk_ϵ
64	52	22
128	88	42
256	150	62

Table 2. $\alpha = 2, \epsilon = 10^{-2}$

The cost of the above algorithm is $O(h^2 m \log^2 m)$ ops, where h is the largest rank encountered in the computation. Once the computation of the blocks $\widetilde{A}^{(j)}$, $\widetilde{B}^{(j)}$, $\widetilde{C}^{(j)}$, $\widetilde{F}^{(j)}$, $j = 1, \ldots, \log_2 n$, has been carried out, that is, once the approximate block LU factorization of $\Pi Q_n \Pi^T$ has been computed, the approximate solution of the system $Q_n \boldsymbol{x} = \boldsymbol{b}$ can be obtained by means of back substitution. The latter computation costs $O(hnm \log m)$ ops, where h is the largest rank encountered in the overall computation, while the approximate factorization costs $O(h^2 m \log^2 m \log n)$ ops. Observe that in the solution of many linear systems $Q_n \boldsymbol{x} = \boldsymbol{b}$, obtained with different values of $\boldsymbol{b}$, the approximate factorization can be computed once for all, and the displacement representation of the blocks $\widetilde{A}^{(j)}$, $\widetilde{B}^{(j)}$, $\widetilde{C}^{(j)}$, $\widetilde{F}^{(j)}$ can be stored; this requires a memory space of $O(hm \log n)$, that is negligible with respect to n.

The algorithm has been implemented and tested for a problem of image restoration of $N \times N$ pixels images, with the PSF defined by the mask

$$\sigma \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & \alpha & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}, \quad \sigma = 1/(24 + \alpha).$$

This mask generates a 5-block diagonal block Toeplitz matrix of block size N, with blocks of size N; the off diagonal blocks are 5-diagonal Toeplitz matrices with all the nonzero entries equal to σ, the diagonal blocks are 5-diagonal Toeplitz matrices having $\sigma\alpha$ on the main diagonal and σ on the remaining nonzero diagonals. The matrix obtained in this way is reblocked into a block tridiagonal form, thus obtaining (1).

We have considered images of size $N = 64$, that generate a block tridiagonal matrix with block size $n = 32$, with blocks of size $m = 128$. We have reconstructed the image by solving the system with our modified CR algorithm. For $\alpha = 5$ the conditions of theorem 3.2 are verified and we have reconstructed the image with values of ϵ in the range $[10^{-6}, 10^{-3}]$, keeping a small value of the ϵ-displacement rank and an accurate precision in the approximation. Larger values of ϵ did not provide adequate approximation of the solution. The original, blurred and restored images are reported in figures 1, 2.

We have also computed the maximum values of Drk_ϵ and Rdrk_ϵ of the blocks generated by CR for different values of α, ϵ and the original block size N. The results are summarized in tables 1, 2, where we report the size m of the blocks of the reblocked tridiagonal matrix and the largest values of Drk_ϵ and Rdrk_ϵ of the blocks generated by CR, for different values of ϵ. We observe that for $\alpha = 2$ the hypotheses of theorem 3.2 are not satisfied and even though the Rdrk_ϵ depends on N, its growth is rather moderate.

4.2. Fixed rank algorithm: incomplete displacement block LU factorization

The cyclic reduction technique can be implemented in a different way where we fix from the beginning the maximum value of the displacement rank at which truncating the intermediate blocks generated by the computation. More precisely, at the $(j+1)$-th step we compute the matrices $\widetilde{A}^{(j+1)}$, $\widetilde{B}^{(j+1)}$, $\widetilde{C}^{(j+1)}$, $\widetilde{F}^{(j+1)}$ having displacement rank

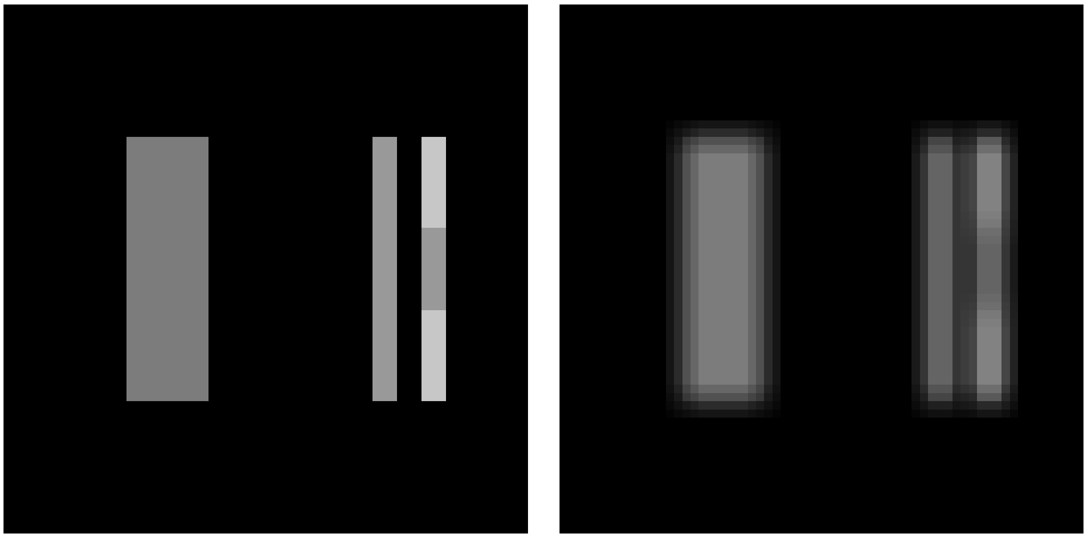

Figure 1. Original and blurred images

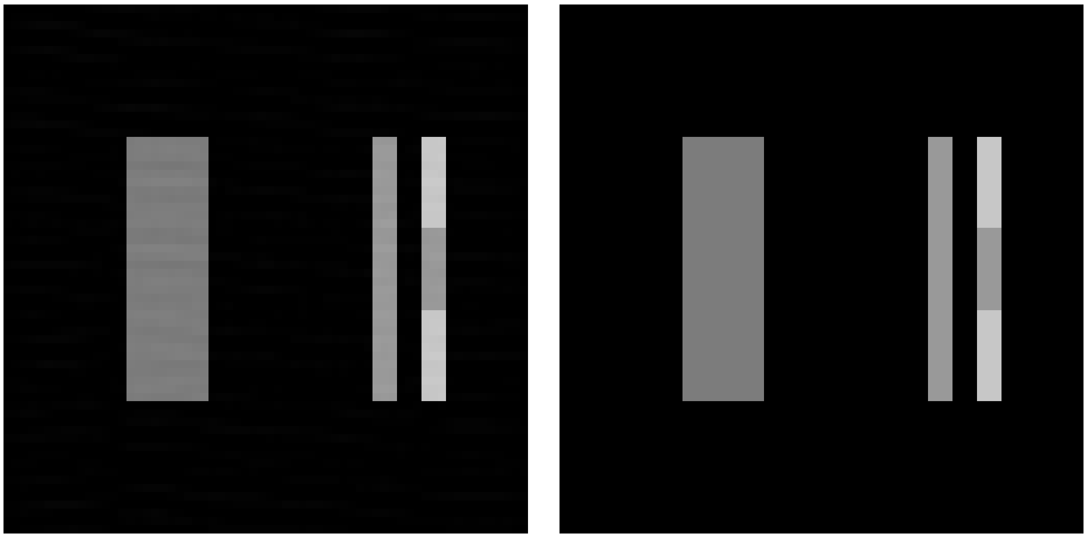

Figure 2. Restored images, with $\epsilon = 10^{-3}$ and $\epsilon = 10^{-6}$

at most K, obtained by selecting at most the first K largest singular values of the SVD of $\Delta(\widehat{A}^{(j+1)})$, $\Delta(\widehat{B}^{(j+1)})$, $\Delta(\widehat{C}^{(j+1)})$, $\Delta(\widehat{F}^{(j+1)})$, respectively, where $\widehat{A}^{(j+1)}$, $\widehat{B}^{(j+1)}$, $\widehat{C}^{(j+1)}$, $\widehat{F}^{(j+1)}$ are defined by means of formula (16). The algorithm for performing one step of CR works in the following way:

Algorithm.

INPUT: An integer $K > 1$, and a displacement representation of displacement rank at most K of the matrices $\widetilde{A}^{(j)}$, $\widetilde{B}^{(j)}$, $\widetilde{C}^{(j)}$, $\widetilde{F}^{(j)}$.

OUTPUT: A displacement representation of displacement rank at most K of the matrices $\widetilde{A}^{(j+1)}$, $\widetilde{B}^{(j+1)}$, $\widetilde{C}^{(j+1)}$, $\widetilde{F}^{(j+1)}$.

COMPUTATION:

1. By using the displacement representation of $\widetilde{A}^{(j)}$, compute the displacement representation of ${\widetilde{A}^{(j)}}^{-1}$, by means of (4).
2. From the latter, and from the displacement representation of $\widetilde{B}^{(j)}$ compute a displacement representation of $\widehat{B}^{(j+1)} = -\widetilde{B}^{(j)}{\widetilde{A}^{(j)}}^{-1}\widetilde{B}^{(j)}$, according to formula (5).

   Let h be the rank of such representation and S_1 and S_2 be the $m \times h$ matrices such that $\Delta(\widehat{B}^{(j+1)}) = S_1 S_2^T$.
3. Compute the QR factorizations $S_1 = Q_1 R_1$, $S_2 = Q_2 R_2$ and the $h \times h$ matrix $W = R_1 R_2^T$. Compute the singular value decomposition $W = U\Sigma V^T$ of W and denote with $\sigma_1 \geq \cdots \geq \sigma_h$ the singular values of W. If $h > K$, denote with $\widehat{U}$, $\widehat{V}$, the matrices made up by the first K columns of U and V, respectively. Similarly denote with $\widehat{\Sigma}$ the leading $K \times K$ principal submatrix of Σ. Otherwise, if $h \leq K$, set $\widehat{U} = U$, $\widehat{V} = V$, $\widehat{\Sigma} = \Sigma$.
4. Define the matrix $\widetilde{B}^{(j+1)}$ as the one represented by $Q_1\widehat{U}\widehat{\Sigma}$, $Q_2\widehat{V}$, i.e., $\widetilde{B}^{(j+1)}$ is the matrix such that $\Delta(\widetilde{B}^{(j+1)}) = (Q_1\widehat{U}\widehat{\Sigma})(Q_2\widehat{V})^T$.
5. Perform similar computations for the matrices $\widehat{A}^{(j+1)}$, $\widehat{C}^{(j+1)}$, $\widehat{F}^{(j+1)}$.

The computational cost of the algorithm is $O(K^2 m \log^2 m)$ ops.

Similarly to the incomplete Cholesky factorization for a sparse matrix A, where the sparsity is imposed in the L factor, here we obtain a sort of incomplete displacement block LU factorization (IDF) of the matrix $\Pi Q_n \Pi^T$, where the blocks of L and U have a fixed displacement rank and where the product LU does not coincide in general with $\Pi Q_n \Pi^T$. This kind of factorization can be used as a tool for preconditioning block banded block Toeplitz systems with Toeplitz-like blocks.

Observe once again that the memory storage required for allocating the IDF is just $O(Km \log n)$. The IDF can can be computed once for all, with the cost of $O(K^2 m \log^2 m \log n)$ ops, and provided to the PCG solver. One PCG iteration with the IDF preconditioner simply costs $O(Knm \log m)$ ops.

4.3. Extended Toeplitz preconditioner

In this section we propose a different preconditioner for block banded block Toeplitz systems with banded Toeplitz blocks, that seems particularly suitable for solving problems related to image restoration. The preconditioner is obtained by extending the matrix of the system to a banded Toeplitz matrix, by replacing some zero entries with suitable nonzero values; more precisely, the preconditioner is the banded Toeplitz matrix T defined by the first row and the first column of the BBBT/BTB matrix. The matrix T thus can be reblocked into a block Toeplitz block tridiagonal matrix, and the blocks defining the matrix obtained at each step of cyclic reduction applied to T have displacement rank at most three[10] (see section 3.2). In this way we can compute once for all the LU factorization of $\Pi T \Pi^T$, at a cost of $O(m \log^2 m \log n)$ ops, then, at each step of the PCG method, solve a system $T\boldsymbol{x} = \boldsymbol{b}$, at a cost of $O(nm \log m)$ ops.

4.4. Open problems

Concerning the algorithm for the approximate solution described in section 4.1, issues like the error analysis, the numerical stability and the control of the level ϵ where the singular values are truncated are still to be investigated; it should be also interesting to analyze the growth of the ϵ-displacement rank, and see how it may depend on displacement, or spectral properties of the initial matrix.

Concerning the preconditioners proposed in sections 4.3, issues like the spectral properties, or the problem of the non singularity of the blocks $\widetilde{A}^{(j)}$ have not yet been analyzed.

REFERENCES

1. R. H. Chan and M. K. Ng, "Conjugate gradient methods for Toeplitz systems," *SIAM Review* **38**, pp. 427–482, 1996.
2. S. Serra, "Superlinear PCG methods for symmetric Toeplitz systems," *Math. Comp.* **68**(226), pp. 793–803, 1999.
3. S. Serra, "Spectral and computational analysis of block Toeplitz matrices having nonnegative definite matrix-valued generating functions," *BIT* **39**(1), pp. 152–175, 1999.
4. S. Serra, "Asymptotic results on the spectra of block Toeplitz preconditioned matrices," *SIAM J. Matrix Anal. Appl.* **20**(1), pp. 31–44 (electronic), 1999.
5. S. Serra, "New PCG based algorithms for the solution of Hermitian Toeplitz systems," *Calcolo* **32**(3-4), pp. 153–176 (1997), 1995.
6. M. Miranda and P. Tilli, "Block Toeplitz matrices and preconditioning," *Calcolo* **33**(1-2), pp. 79–86 (1998), 1996. Toeplitz matrices: structures, algorithms and applications (Cortona, 1996).
7. G. Fiorentino and S. Serra, "Multigrid methods for indefinite Toeplitz matrices," *Calcolo* **33**(3-4), pp. 223–236 (1998), 1996. Toeplitz matrices: structures, algorithms and applications (Cortona, 1996).
8. D. A. Bini and B. Meini, "On cyclic reduction applied to a class of Toeplitz-like matrices arising in queueing problems," in *Computations with Markov Chains*, W. J. Stewart, ed., pp. 21–38, Kluwer Academic Publisher, 1995.
9. D. A. Bini and B. Meini, "Improved cyclic reduction for solving queueing problems," *Numerical Algorithms* **15**, pp. 57–74, 1997.
10. D. A. Bini and B. Meini, "Effective methods for solving banded Toeplitz systems," *SIAM J. Matrix Anal. Appl.* **20**, pp. 700–719, 1999.
11. D. A. Bini and B. Meini, "On the solution of a nonlinear matrix equation arising in queueing problems," *SIAM J. Matrix Anal. Appl.* **17**, pp. 906–926, 1996.
12. D. A. Bini and B. Meini, "Using displacement structure for solving non-skip-free M/G/1 type Markov chains," in *Advances in Matrix Analytic Methods for Stochastic Models - Proceedings of the 2nd international conference on matrix analytic methods*, A. Alfa and S. Chakravarthy, eds., pp. 17–37, Notable Publications Inc, NJ, 1998.
13. D. Heller, "Some aspects of the cyclic reduction algorithm for block tridiagonal linear systems," *SIAM J. Num. Anal.* **13**, pp. 484–496, 1976.
14. G. Golub and C. V. Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1989.
15. T. Kailath, A. Viera, and M. Morf, "Inverses of Toeplitz operators, innovations and orthogonal polynomials," *SIAM Review* **20**, pp. 106–119, 1978.
16. T. Kailath, S. Kung, and M. Morf, "Displacement rank of a matrix," *Bulletin of the American Mathematical Society* **1**, pp. 769–773, 1979.
17. G. Heinig and K. Rost, *Algebraic Methods for Toeplitz-like Matrices and Operators*, Akademie-Verlag, Berlin, and Birkhäuser, Boston, 1984.
18. D. A. Bini and V. Pan, *Matrix and Polynomial Computations, Vol. 1: Fundamental Algorithms*, Birkhäuser, Boston, 1994.
19. R. R. Bitmead and B. D. O. Anderson, "Asymptotically fast solution of Toeplitz and related systems of linear equations," *Linear Algebra Appl.* **34**, pp. 103–116, 1980.
20. M. Morf, *Fast algorithms for multivariable systems.* PhD thesis, Department of Electrical Engineering, Stanford University, 1974.
21. V. Pan, "Concurrent iterative algorithm for Toeplitz-like linear systems," *IEEE Trans. on Parallel and Distributive Systems* **4–5**, pp. 592–600, 1993.
22. P. Yalamov and V. Pavlov, "Stability of the block cyclic reduction," *Linear Algebra Appl.* **249**, pp. 341–358, 1996.
23. P. Lancaster and M. Tismenetski, *Matrix Polynomials*, Academic Press Inc., New York, 1982.
24. H. R. Gail, S. L. Hantler, and B. A. Taylor, "Solutions of the basic matrix equation for M/G/1 and G/M/1 type Markov chains," *Commun. Statist. Stochastic Models* **10**, pp. 1–43, 1994.
25. P. Tilli, "A note on the spectral distribution of Toeplitz matrices," *Linear and Multilinear Algebra* **45**, pp. 147–159, 1998.
26. D. A. Bini, B. Meini, and V. Ramaswami, "Analyzing M/G/1 paradigms through QBDs: the role of the block structure in computing the matrix G." in preparation.