Effective methods for solving banded Toeplitz systems

Dario Andrea Bini and Beatrice Meini
Dipartimento di Matematica, Università di Pisa
bini@dm.unipi.it, meini@dm.unipi.it

July 24, 1997

Abstract

We propose new algorithms for solving $n \times n$ banded Toeplitz systems with bandwidth m. If the function associated with the Toeplitz matrix has no zero in the unit circle, then $O(n \log m + m \log^2 m \log \log \epsilon^{-1})$ arithmetic operations (ops) are sufficient to approximate the solution of the system up to within the error ϵ, otherwise the cost becomes $O(n \log m + m \log^2 m \log \frac{n}{m})$ ops. Here $m = o(n)$, and $n > \log \epsilon^{-1}$. Some applications are presented. The methods can be applied to infinite and bi-infinite systems and to block matrices.

Key Words. Banded matrices, Toeplitz matrices, displacement rank, cyclic reduction, Graeffe's method.

AMS Subject Classification. 65F05, 65F10, 15A23

1 Introduction

Let $\{a_k\}_{k \in \boldsymbol{Z}}$ be a sequence of real numbers such that $a_k = 0$ for $k < m_2$, and for $k > m_1$, $a_{m_1} a_{m_2} \neq 0$, where $\boldsymbol{Z}$ is the ring of integers, and $m_1, m_2 \in \boldsymbol{Z}$, $m_2 < 0 < m_1$. We may associate with $\{a_k\}_{k \in \boldsymbol{Z}}$ the function $\phi(z) = \sum_{i=m_2}^{m_1} z^i a_i$ and a matrix $T = (t_{i,j})$, having entries $t_{i,j} = a_{i-j}$. Let us use different notations for T according to its size. That is, we denote with $T_{\pm\infty} = (t_{i,j})_{i,j \in \boldsymbol{Z}}$ the bi-infinite matrix, with $T_\infty = (t_{i,j})_{i,j \geq 0}$ the infinite matrix, and with $T_n = (t_{i,j})_{i,j=1,\ldots,n}$ the matrix having size n. We leave the notation T for a matrix having an unspecified size.

Indeed, T is a banded matrix with bandwidth $m = \max\{m_1, -m_2\}$, and has the Toeplitz structure, i.e., its entries are constant along the diagonals.

The problem of solving banded Toeplitz systems as $T\boldsymbol{x} = \boldsymbol{b}$, plays an important role in many applications. Example of these systems are frequently encountered in the numerical solution of ordinary differential equations with boundary values, where the banded Toeplitz matrix provides a finite difference discretization of the differential operator [31], and in the numerical solution of many Markov chains coming from the modeling of queueing problems, where the involved matrix defines the probabilities of transition from one state to the other [26], [25], [18]. Another indirect and more recent application is related to the solution of general Toeplitz systems, like the ones encountered in the problems in signal processing [16], where the banded Toeplitz matrix is used as a preconditioner for speeding up the convergence of preconditioned conjugate gradient techniques [17], [30], [29].

In all the above problems the dimension of the matrices involved is typically huge or, sometimes, may be even infinite [26]. Moreover, in multi dimensional problems the banded Toeplitz structure is expressed in terms of blocks, i.e., the scalar entries a_{j-i} are replaced by square matrices.

This makes particularly important devising efficient and reliable algorithms for the solution of (block) banded (block) Toeplitz systems that maintain their efficiency in the case of matrices having a very large, or even infinite, size.

A certain attention to this problem has been devoted in the literature specially concerning the analysis of the complexity in parallel models of computation, like the PRAM model. In the PRAM model it is assumed that a certain number of processors work concurrently, more specifically, at each time unit (parallel step) each processor performs a single arithmetic operation. In this model the cost of a computation is expressed in terms of the number of parallel steps and of the number of processors needed to carry out the computation.

Different direct parallel algorithms have been developed [20], [4], [6], [5]. They are based mainly on two different techniques: embedding the matrix into a slightly larger matrix that can be diagonalized by means of trigonometric transforms (FFT); correcting the matrix by means of a low rank additive correction in order to find a new matrix that can be diagonalized by a trigonometric transform and applying the Morrison-Sherman-Woodbury formula.

The algorithms obtained in this way have a parallel cost of $O(\log n + \log^2 m)$ steps with nm^2 processors, and their sequential cost of $O(n \log n + m^2 n)$ arithmetic operations (ops) is not inferior to the cost of Gaussian elimination, that is, $O(m^2 n)$ ops. Moreover most of these algorithms have problems of numerical stability.

Recently, a direct and fully parallelizable algorithm for inverting block Toeplitz matrices in block Hessenberg form, based on a doubling technique and on the use of the Morrison-Sherman-Woodbury formula, has been proposed in [11]. The algorithm, applied to solve the system $T_n \boldsymbol{x} = \boldsymbol{b}$, has a sequential cost of $O(m^2 \log \frac{n}{m} + n \log n)$ ops. However, this algorithm does not fully exploit the band structure of the system and requires a large storage.

Iterative algorithms, based on the preconditioned conjugate gradient (pcg) techniques, are fully parallelizable and have a sequential cost per iteration of $O(n \log n)$. There exists a wide literature concerning these techniques, see for instance [16], [15], [29], [32]. The number of pcg iterations needed in the worst case depends on the bandwidth m (compare [30] for the dependency of the number of iterations on the specific features of the function associated with the matrix). Due to the Axelsson-Lindskog theorem [2], the number of iterations sufficient to reach a super-linear convergence is proportional to m if a circulant [14], [32] or τ-class [6], [7], [28] preconditioner is used. Moreover, $O(mn \log n)$ ops are needed for the presently known preconditioners [30].

In this paper we propose a different approach to solving $n \times n$ (block) banded Toeplitz systems with bandwidth m, leading to an algorithm that can be used either as a direct algorithm or as an iterative one. The direct method, which can be applied with almost no restriction, has a sequential cost of $O(n \log m + m \log^2 m \log \frac{n}{m})$ ops. Under suitable conditions generally satisfied in most part of the applications, the direct algorithm turns into an iterative one, and its convergence speed is quadratic. In this way the number of ops needed to approximate n components of the solution up to within an $O(\epsilon)$ error is $O(n \log m + m \log^2 m \log \log \epsilon^{-1})$.

The cost $O(m \log^2 m \log \log \epsilon^{-1})$ is required for arriving at an approximate LU factorization of the (infinite) matrix obtained by means of a suitable permutation of rows and columns of T. The $O(n \log m)$ term is the cost of the back substitution stage of the algorithm. This nice convergence property allows us to solve infinite and bi-infinite systems with the same computational cost as well. In the PRAM model the algorithm costs $O(t_s(m) \log \frac{n}{m})$ steps with $\max\{t_p(m), n\}$ processors, where $t_s(m)$, $t_p(m)$ are the number of steps and of processors, respectively, needed to invert an $m \times m$ Toeplitz matrix.

The algorithm is based on the following technique. The matrix T is partitioned into $m \times m$ blocks, thus obtaining a block Toeplitz block tridiagonal matrix having Toeplitz blocks. The algorithm of cyclic reduction is applied to the reblocked matrix T, that is, an odd-even permutation of the block rows and the block columns of T is followed by a step of block Gaussian elimination.

We prove that the block tridiagonal block Toeplitz structure is kept almost everywhere by the cyclic reduction step. In other words, the Schur complement that we obtain after applying the single step of block Gaussian elimination to the matrix with permuted block rows and block columns is still a block tridiagonal block Toeplitz matrix except for the block in position (1,1).

On the other hand a sort of deterioration occurs in the inner Toeplitz structure of the blocks of T. In fact, the blocks of the Schur complement obtained just at the first step of cyclic reduction are not Toeplitz any more. In principle, this fact would imply that each block must be stored in m^2 memory locations and that managing with blocks cost $O(m^3)$ ops. In this way cyclic reduction would require $O(mn + m^3 \log \frac{n}{m})$ ops in order to solve an $n \times n$ banded Toeplitz system with bandwidth m.

Even though the Toeplitz structure of the blocks is lost, a more general structure is maintained by the blocks of the Schur complement at each level of cyclic reduction. More precisely, the main result that we prove in section 3 is that each block defining the Schur complement is a Toeplitz-like matrix of displacement rank at most four. Each block can be represented as the sum of at most four products of pairs of lower and upper triangular Toeplitz matrices. Therefore it can be stored in $O(m)$ memory locations by representing the block by means of few vectors. Furthermore, performing multiplications between the blocks costs $O(m \log m)$ ops., and inverting a block costs $t(m) \in \{m^2, m \log^2 m, m \log m\}$ ops, according to the feature of the system to be solved.

Moreover we prove explicit relations which hold among the vectors defining the blocks of the Schur complement at two consecutive steps. These relations allow us to implement the cyclic reduction technique in $O(n \log m + t(m) \log \frac{n}{m})$ ops, where $t(m)$ is the cost of solving an $m \times m$ Toeplitz-like system, i.e., $t(m) \in \{m^2, m \log^2 m, m \log m\}$ (compare [23], [1], [15]).

The algorithm can be easily extended to banded block Toeplitz systems.

The cyclic reduction algorithm obtained in this way has a further nice feature. We prove that the lower (upper) diagonal blocks of the Schur complement obtained at each step of cyclic reduction, tend to zero double exponentially if the polynomial $a(z) = z^{-m_2}\phi(z) = \sum_{i=m_2}^{m_1} a_i z^{i-m_2}$ has m zeros inside (outside) the unit circle and if the blocks have bounded norm.

For stochastic matrices, like the ones modeling Markov chains arising in QBD problems [26], the above conditions are satisfied [9], [10]. In the case where T is symmetric and the polynomial $a(z)$ has no zeros of modulus 1, then the above property is satisfied simultaneously inside and outside the unit circle, and the boundness of the blocks follows from the Cauchy interlace theorem. In this case we may prove that the off diagonal blocks of the Schur complement obtained at the j-th step of cyclic reduction, tend to zero as $O(\gamma^{2^j})$ where $\gamma < 1$ is the modulus of the zero of $a(z)$ in the unit disc, closest to the unit circle.

This property generalizes and extends the results of [13] proved for tridiagonal Toeplitz matrices, where $m = 1$. The conditions under which the convergence occurs are much weaker than the conditions given in [22], [33].

For matrices associated with polynomials having some zeros of modulus 1, we provide a simple trick that enable us to move the zeros of modulus 1 slightly in/out the unit disc. Then constructing the normal equations leads us to an asymptotically well conditioned system having a banded matrix with a nearly Toeplitz structure, that satisfies the convergence theorem. This fact allows us to improve the robustness and the stability of our algorithm which, however, needs to be applied as a direct algorithm to solve the linear system in $O(n \log m + t(m) \log \frac{n}{m})$ ops.

The algorithm proposed in this paper, unlike the ones based on matrix embedding and on the Morrison-Sherman-Woodbury formula is numerically stable if applied to symmetric matrices. In fact, all the $m \times m$ linear systems involved at each step of cyclic reduction have condition numbers not greater than the condition number of the original matrix T.

Another interesting feature of the algorithm is its close relationship with polynomial computations. In fact, we show that applying the cyclic reduction algorithm to a banded Toeplitz matrix corresponds to applying the Graeffe iteration [27] to the polynomial $\alpha(z)$ having as zeros the m-th powers of the zeros of $a(z)$. Moreover, this similarity between polynomial and Toeplitz computations enables us to extend the Graeffe iteration to matrix polynomials.

We have implemented in Fortran 90 different versions of our algorithm which include the cases of matrices with scalar/block entries, symmetric/unsymmetric matrices, and stochastic matrices.

The algorithms have been compared with the preconditioned conjugate gradient method with circulant preconditioner. From the many numerical experiments performed over an alpha workstation, our algorithms have proved to be robust, numerically stable and fast. For matrices of size $n = 524288$ and bandwidth $m = 32$ our programs have been 16 times faster than the preconditioned conjugate gradient method. For larger dimensions the conjugate gradient method could not be applied for lack of memory.

The programs can be taken by anonymous ftp at the following URL address: morse.dm.unipi.it.

2 Preliminaries

Without loss of generality let us assume that $-m_2 \geq m_1$, moreover, let n be an integer multiple of m, i.e., $n = mq$, for an integer q. We refer to $\phi(z)$ as to the generating function of the matrix T and we call T_n an *n-section of* T_∞.

The matrix T can be partitioned into $m \times m$ blocks yielding a block tridiagonal Toeplitz matrix. In particular T_n has the following structure

$$T_n = \begin{bmatrix} A_0 & A_{-1} & & & \bigcirc \\ A_1 & A_0 & A_{-1} & & \\ & \ddots & \ddots & \ddots & \\ & & A_1 & A_0 & A_{-1} \\ \bigcirc & & & A_1 & A_0 \end{bmatrix}$$

where

$$A_1 = \begin{bmatrix} a_m & a_{m-1} & \dots & a_1 \\ 0 & a_m & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{m-1} \\ 0 & \dots & 0 & a_m \end{bmatrix},$$

$$A_0 = \begin{bmatrix} a_0 & a_{-1} & \dots & a_{-m+1} \\ a_1 & a_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{-1} \\ a_{m-1} & \dots & a_1 & a_0 \end{bmatrix} \qquad (1)$$

$$A_{-1} = \begin{bmatrix} a_{-m} & 0 & \dots & 0 \\ a_{-m+1} & a_{-m} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{-1} & \dots & a_{-m+1} & a_{-m} \end{bmatrix}$$

and similarly T_∞ and $T_{\pm\infty}$.

In the light of the above block structure we may associate with T the matrix polynomial $A(z) = A_{-1} + zA_0 + z^2 A_1$.

In this section we first recall the cyclic reduction algorithm, introduced in [8] for computing the probability invariant vector of stochastic matrices, and here adjusted in order to solve the block tridiagonal block Toeplitz system $T\boldsymbol{x} = \boldsymbol{b}$. Then we introduce the concept of displacement rank and we recall some properties that are needed for the design and analysis of our algorithm.

2.1 Cyclic reduction.

Consider the infinite system

$$T_\infty \boldsymbol{x} = \boldsymbol{b} \qquad (2)$$

and partition the infinite vectors $\boldsymbol{x}$ and $\boldsymbol{b}$ into blocks $\boldsymbol{x}_k$, $\boldsymbol{b}_k$ of dimension m, respectively, that is, $\boldsymbol{x} = (\boldsymbol{x}_k)_{k\geq 1}$, $\boldsymbol{b} = (\boldsymbol{b}_k)_{k\geq 1}$. By performing an odd-even permutation of the block rows and block columns in (2) we find that

$$\begin{bmatrix} D_1^{(0)} & U^{(0)} \\ L^{(0)} & D_2^{(0)} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_+^{(0)} \\ \boldsymbol{x}_-^{(0)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_+^{(0)} \\ \boldsymbol{b}_-^{(0)} \end{bmatrix} \qquad (3)$$

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

$$D_1^{(0)} = D_2^{(0)} = \begin{bmatrix} A_0 & & \bigcirc \\ & A_0 & \\ \bigcirc & & \ddots \end{bmatrix}, \quad L^{(0)} = \begin{bmatrix} A_{-1} & & & \bigcirc \\ A_1 & A_{-1} & & \\ & \ddots & \ddots & \\ \bigcirc & & \ddots & \ddots \end{bmatrix},$$

$$U^{(0)} = \begin{bmatrix} A_1 & A_{-1} & & \bigcirc \\ & A_1 & A_{-1} & \\ & & \ddots & \ddots \\ \bigcirc & & & \ddots \end{bmatrix}.$$

Applying one step of block-Gaussian elimination to the 2×2 block system (3) yields the equivalent system

$$\begin{cases} (D_2^{(0)} - L^{(0)} D_1^{(0)^{-1}} U^{(0)}) x_+^{(0)} = \boldsymbol{b}_-^{(0)} - L^{(0)} D_1^{(0)^{-1}} \boldsymbol{b}_+^{(0)} \\ \boldsymbol{x}_+^{(0)} = D_1^{(0)^{-1}} (\boldsymbol{b}_+^{(0)} - U^{(0)} \boldsymbol{x}_-^{(0)}) \end{cases} \qquad (4)$$

Denoting

$$T^{(1)} = D_2^{(0)} - L^{(0)} D_1^{(0)^{-1}} U^{(0)} \tag{5}$$

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

$$T^{(1)} \boldsymbol{x}^{(1)} = \boldsymbol{b}^{(1)}. \tag{6}$$

It is interesting to observe that $T^{(1)}$ is a block tridiagonal matrix which, except for the north west corner block, has the block Toeplitz structure, i.e.,

$$T^{(1)} = \begin{bmatrix} \widehat{A}^{(1)} & A_{-1}^{(1)} & & \bigcirc \\ A_1^{(1)} & A_0^{(1)} & A_{-1}^{(1)} & \\ & A_1^{(1)} & A_0^{(1)} & \ddots \\ \bigcirc & & \ddots & \ddots \end{bmatrix}.$$

In other words, $T^{(1)}$ is uniquely determined by the blocks $A_{-1}^{(1)}$, $A_0^{(1)}$, $A_1^{(1)}$, $\widehat{A}^{(1)}$.

We may recursively apply the same reduction (cyclic reduction) to the matrix equation (6) thus obtaining a sequence of systems $\{T^{(j)} \boldsymbol{x}^{(j)} = \boldsymbol{b}^{(j)}\}$, where $T^{(0)} = T_\infty$, such that $T^{(j)}$ is a block tridiagonal matrix having, except for the north west corner, the block Toeplitz structure. Each matrix $T^{(j)}$ is uniquely determined by the blocks $\widehat{A}^{(j)}$, $A_{-1}^{(j)}$, $A_0^{(j)}$, $A_1^{(j)}$.

The block entries $A_i^{(j+1)}$, $i = -1, 0, 1$, $\widehat{A}^{(j+1)}$, obtained at step $j+1$, are related to the block entries $A_i^{(j)}$, $i = -1, 0, 1$, $\widehat{A}^{(j)}$, obtained at step j, by means of simple relations. For this purpose let us associate with $T^{(j)}$ the following matrix polynomial

$$A^{(j)}(z) = A_{-1}^{(j)} + z A_0^{(j)} + z^2 A_1^{(j)} \tag{7}$$

and express in compact form the block odd-even permutation and the Gaussian elimination of one step of cyclic reduction in terms of $A^{(j)}(z)$.

In fact from (5) we may easily deduce that

$$\left\{ \begin{array}{l} A^{(j+1)}(z) = z A_0^{(j)} - (A_{-1}^{(j)} + z A_1^{(j)}) A_0^{(j)^{-1}} (A_{-1}^{(j)} + z A_1^{(j)}) \\ \widehat{A}^{(j+1)} = \widehat{A}^{(j)} - A_{-1}^{(j)} A_0^{(j)^{-1}} A_1^{(j)}. \end{array} \right. \tag{8}$$

The above relations yield the following matrix equations

$$\left\{ \begin{array}{ll} A_{-1}^{(j+1)} & = -A_{-1}^{(j)} A_0^{(j)^{-1}} A_{-1}^{(j)} \\ A_0^{(j+1)} & = A_0^{(j)} - A_1^{(j)} A_0^{(j)^{-1}} A_{-1}^{(j)} - A_{-1}^{(j)} A_0^{(j)^{-1}} A_1^{(j)} \\ A_1^{(j+1)} & = -A_1^{(j)} A_0^{(j)^{-1}} A_1^{(j)} \\ \widehat{A}^{(j+1)} & = \widehat{A}^{(j)} - A_{-1}^{(j)} A_0^{(j)^{-1}} A_1^{(j)} \end{array} \right. \tag{9}$$

which allow us to compute the sequence of blocks defining the matrices $T^{(j)}$, $j = 1, 2, \ldots$.

For a finite banded Toeplitz system $T_n \boldsymbol{x} = \boldsymbol{b}$, where $q = 2^p$, p positive integer, the cyclic reduction generates a sequence of systems $T_{m\frac{q}{2^j}}^{(j)} \boldsymbol{x}^{(j)} = \boldsymbol{b}^{(j)}$, $j = 1, 2, \ldots, p$, of dimension $n_j = m 2^{p-j}$. Moreover, the matrices $T_{m\frac{q}{2^j}}^{(j)}$ are n_j-sections of the infinite matrices $T^{(j)}$ whose blocks are defined by (9).

For a bi-infinite block tridiagonal block Toeplitz system $T_{\pm\infty} \boldsymbol{x} = \boldsymbol{b}$, where $\boldsymbol{x} = (\boldsymbol{x}_i)_{i \in \boldsymbol{Z}}$, $\boldsymbol{b} = (\boldsymbol{b}_i)_{i \in \boldsymbol{Z}}$, the cyclic reduction generates a sequence of bi-infinite block tridiagonal block Toeplitz systems $T_{\pm\infty}^{(j)}\boldsymbol{x}^{(j)} = \boldsymbol{b}^{(j)}$ where $T_{\pm\infty}^{(j)}$ is the bi-infinite block tridiagonal block Toeplitz matrix defined by the $m \times m$ blocks $A_i^{(j)}$, $i = -1, 0, 1$.

A direct inspection shows that the Toeplitz structure of the blocks $A_i^{(0)}$, $i = -1, 0, 1$, is not kept any more by the blocks $A_i^{(j)}$, $i = -1, 0, 1$, and $\widehat{A}^{(j)}$. For this property one might deduce that the initial structure cannot be exploited in order to increase the efficiency of the cyclic reduction algorithm.

However, we may prove that a more general structure, i.e., the displacement structure is maintained by the blocks $A_i^{(j)}$, $i = -1, 0, 1$, and $\widehat{A}^{(j)}$. This fact allows us to devise a suitable FFT-based implementation of the cyclic reduction algorithm thus reaching a dramatic reduction of the complexity bound.

For this purpose we need to rewrite the functional relation (8) for the matrix power series $A^{(j)}(z)$ in the following equivalent form:

$$A^{(j+1)}(z^2) = 2z(A^{(j)}(z)^{-1} - A^{(j)}(-z)^{-1})^{-1} = -2zA^{(j)}(z){A_0^{(j)}}^{-1}A^{(j)}(-z). \quad (10)$$

2.2 Displacement rank.

The concept of displacement rank, introduced by Kailath et al. in [24], is fundamental in devising and analyzing algorithms related to Toeplitz matrices. Several displacement operators have been introduced and analyzed in the literature; here we adopt an operator that seems to be particularly suitable for dealing with our problems.

Define the $m \times m$ *down-shift* matrix

$$Z = \begin{bmatrix} 0 & & & \\ 1 & 0 & & \\ & \ddots & \ddots & \\ & & 1 & 0 \end{bmatrix},$$

and consider the following *displacement operator*

$$\Delta(H) = ZH - HZ$$

defined for any $m \times m$ matrix H. Moreover, denote by $L(\boldsymbol{w})$ the $m \times m$ lower triangular Toeplitz matrix defined by its first column $\boldsymbol{w}$.

Let us also introduce the following notations: the vectors $\boldsymbol{e}_1 = [1, 0, \ldots, 0]^T$, $\boldsymbol{e}_m = [0, \ldots, 0, 1]^T$ denote the first and the last column of the identity matrix of order m, respectively.

Observe that for a general Toeplitz matrix A the displacement $\Delta(A)$ is zero except for the entries in the first row and in the last column, whence the rank of $\Delta(A)$ is at most 2.

We say that the matrix H has *displacement rank r* if r is the rank of $\Delta(H)$, i.e., if there exist vectors $\boldsymbol{u}_i, \boldsymbol{v}_i$, $i = 1, \ldots, r$, satisfying the equation $\Delta(H) = \sum_{i=1}^{r} \boldsymbol{u}_i \boldsymbol{v}_i^T$.

According to this definition an $m \times m$ Toeplitz matrix A has displacement rank at most 2, more precisely it holds

$$\Delta(A) = -\boldsymbol{e}_1\boldsymbol{e}_m^T AZ + ZA\boldsymbol{e}_m\boldsymbol{e}_m^T.$$

A nice property of $\Delta(H)$ is that the displacement of the inverse matrix can be explicitly related to the displacement of the matrix itself, in fact, if H is nonsingular, then

$$\Delta(H^{-1}) = -H^{-1}\Delta(H)H^{-1}. \quad (11)$$

The following result has been proved in [3], [12].

Theorem 1 *Let K be an $m \times m$ matrix such that $\Delta(K) = \sum_{i=1}^{r} \boldsymbol{u}_i \boldsymbol{v}_i^T$ where $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are m-dimensional vectors. Then we have*

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

The above result allows one to represent any matrix K as a sum of products of lower and upper triangular Toeplitz matrices defined by the first column of K and by the vectors $\boldsymbol{u}_i$, $\boldsymbol{v}_i$ associated with the displacement of K.

If the matrix K is nonsingular then the above representation theorem can be applied to K^{-1} in the light of (11) yielding the following result

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

We recall that multiplying an $m \times m$ Toeplitz matrix and a vector costs $O(m \log m)$ ops if the computation is performed by means of FFT [12]. In the PRAM model, the cost is $O(\log m)$ steps with m processors.

3 Structural properties of cyclic reduction.

In this section we prove that the matrix polynomials $A^{(j)}(z)$, $j = 0, 1, \ldots$, defined in (7), (10), have displacement rank at most 2, and provide explicit formulae relating the displacement of $A^{(j)}(z)$ to the displacement of $A^{(j+1)}(z)$. As a consequence we find that $A_{-1}^{(j)}$ and $A_1^{(j)}$ have displacement rank at most 2, $A_0^{(j)}$ has displacement rank at most 4. Moreover, we show that the blocks $\widehat{A}^{(j)}$ have displacement rank at most 3.

These properties allow us to devise an algorithm, based on FFT, that performs a single cyclic reduction step in $O(m \log^2 m)$ arithmetic operations if super-fast algorithms for the inversion of $m \times m$ Toeplitz-like matrices are used.

We need the following result.

Theorem 2 *For the matrix polynomial $A^{(j)}(z)$ generated by the cyclic reduction (10) we have*

$$\Delta(A^{(j)}(z)) = \boldsymbol{a}^{(j)}(z)\boldsymbol{u}^{(j)T}(z) - \boldsymbol{v}^{(j)}(z)\boldsymbol{c}^{(j)T}(z), \quad j \geq 0, \tag{12}$$

where the vectors $\boldsymbol{a}^{(j)}(z)$, $\boldsymbol{c}^{(j)}(z)$, $\boldsymbol{u}^{(j)}(z)$, $\boldsymbol{v}^{(j)}(z)$ have entries which are polynomials in z, and are defined by the following equations:

$$\begin{aligned}
&\boldsymbol{a}^{(j)}(z) = A^{(j)}(z)\boldsymbol{e}_1, \ \boldsymbol{c}^{(j)T}(z) = \boldsymbol{e}_m^T A^{(j)}(z), \ j \geq 0\\
&\boldsymbol{u}^{(j)}(z^2)^T = (-\boldsymbol{u}^{(j-1)}(z)^T (A_0^{(j-1)})^{-1} A^{(j-1)}(-z) +\\
&\qquad\qquad \boldsymbol{u}^{(j-1)}(-z)^T (A_0^{(j-1)})^{-1} A^{(j-1)}(z))/(2z), \quad j \geq 1;\\
&\boldsymbol{v}^{(j)}(z^2) = (-A^{(j-1)}(-z)(A_0^{(j-1)})^{-1}\boldsymbol{v}^{(j-1)}(z) +\\
&\qquad\qquad A^{(j-1)}(z)(A_0^{(j-1)})^{-1}\boldsymbol{v}^{(j-1)}(-z))/(2z), \quad j \geq 1;\\
&\boldsymbol{u}^{(0)}(z)^T = z\boldsymbol{e}_m^T, \ \boldsymbol{v}^{(0)}(z) = z\boldsymbol{e}_1.
\end{aligned} \tag{13}$$

Proof. We prove relation (12) by induction on j. For $j = 0$ a simple calculation shows that

$$\begin{aligned}\Delta(A^{(0)}(z)) &= z(A^{(0)}(z)\boldsymbol{e}_1\boldsymbol{e}_m^T - \boldsymbol{e}_1\boldsymbol{e}_m^T A^{(0)}(z)) =\\ &A^{(0)}(z)\boldsymbol{e}_1\boldsymbol{u}^{(0)T}(z) - \boldsymbol{v}^{(0)}(z)\boldsymbol{e}_m^T A^{(0)}(z).\end{aligned}$$

Suppose that relation (12) holds for a fixed $j \geq 0$; we show that it holds for $j+1$. For simplicity let us denote $B = A^{(j)}(z)^{-1} - A^{(j)}(-z)^{-1}$. For the properties of the displacement operator (11) and for (10) we have:

$$\begin{array}{l}\Delta(A^{(j+1)}(z^2)) = 2z\Delta(B^{-1}) = -2zB^{-1}\Delta(B)B^{-1} = \\ -2zB^{-1}(-A^{(j)}(z)^{-1}\Delta(A^{(j)}(z))A^{(j)}(z)^{-1} + A^{(j)}(-z)^{-1}\Delta(A^{(j)}(-z))A^{(j)}(-z)^{-1})B^{-1};\end{array}$$

by substituting relation (12) in the above equation we obtain

$$\begin{array}{l}\Delta(A^{(j+1)}(z^2)) = 2zB^{-1}\boldsymbol{e}_1\boldsymbol{u}^{(j)}(z)^T A^{(j)}(z)^{-1}B^{-1} - 2zB^{-1}A^{(j)}(z)^{-1}\boldsymbol{v}^{(j)}(z)\boldsymbol{e}_m^T B^{-1} - \\ 2zB^{-1}\boldsymbol{e}_1\boldsymbol{u}^{(j)}(-z)^T A^{(j)}(-z)^{-1}B^{-1} + 2zB^{-1}A^{(j)}(-z)^{-1}\boldsymbol{v}^{(j)}(-z)\boldsymbol{e}_m^T B^{-1} = \\ A^{(j+1)}(z^2)\boldsymbol{e}_1\boldsymbol{u}^{(j+1)}(z^2)^T - \boldsymbol{v}^{(j+1)}(z^2)\boldsymbol{e}_m^T A^{(j+1)}(z^2).\end{array}$$

□

By induction on j, from (13), it readily follows that $\boldsymbol{u}^{(j)}(z)$ and $\boldsymbol{v}^{(j)}(z)$ have components which are polynomials of degree at most 1. Therefore, since $\boldsymbol{a}^{(j)}(z)$ and $\boldsymbol{c}^{(j)}(z)$ have degree at most 2, we may denote

$$\begin{array}{l}\boldsymbol{u}^{(j)}(z) = \boldsymbol{u}_{-1}^{(j)} + z\boldsymbol{u}_0^{(j)}, \\ \boldsymbol{v}^{(j)}(z) = \boldsymbol{v}_{-1}^{(j)} + z\boldsymbol{v}_0^{(j)}, \\ \boldsymbol{a}^{(j)}(z) = \boldsymbol{a}_{-1}^{(j)} + z\boldsymbol{a}_0^{(j)} + z^2\boldsymbol{a}_1^{(j)}, \\ \boldsymbol{c}^{(j)}(z) = \boldsymbol{c}_{-1}^{(j)} + z\boldsymbol{c}_0^{(j)} + z^2\boldsymbol{c}_1^{(j)}.\end{array}$$

Moreover, by comparing the term of degree 3 in (12) we may deduce that, for any $j \geq 0$, it holds

$$\boldsymbol{a}_1^{(j)}\boldsymbol{u}_0^{(j)T} = \boldsymbol{v}_0^{(j)}\boldsymbol{c}_1^{(j)T}. \tag{14}$$

In this way, from the above properties and from Theorem 2 we deduce explicit relations for the vectors $\boldsymbol{u}_i^{(j)}$, $\boldsymbol{v}_i^{(j)}$, $i = -1, 0$, $\boldsymbol{a}_i^{(j)}$, $\boldsymbol{c}_i^{(j)}$, $i = -1, 0, 1$, $j \geq 0$, defining $\Delta(A^{(j)}(z))$:

$$\begin{array}{ll}\boldsymbol{u}_{-1}^{(j+1)T} & = \boldsymbol{u}_{-1}^{(j)T} - \boldsymbol{u}_0^{(j)T}A_0^{(j)^{-1}}A_{-1}^{(j)} \\ \boldsymbol{u}_0^{(j+1)T} & = -\boldsymbol{u}_0^{(j)T}A_0^{(j)^{-1}}A_1^{(j)} \\ \boldsymbol{v}_{-1}^{(j+1)} & = \boldsymbol{v}_{-1}^{(j)} - A_{-1}^{(j)}A_0^{(j)^{-1}}\boldsymbol{v}_0^{(j)} \\ \boldsymbol{v}_0^{(j+1)} & = -A_1^{(j)}A_0^{(j)^{-1}}\boldsymbol{v}_0^{(j)} \\ \boldsymbol{a}_{-1}^{(j+1)} & = -A_{-1}^{(j)}A_0^{(j)^{-1}}\boldsymbol{a}_{-1}^{(j)} \\ \boldsymbol{a}_0^{(j+1)} & = \boldsymbol{a}_0^{(j)} - A_{-1}^{(j)}A_0^{(j)^{-1}}\boldsymbol{a}_1^{(j)} - A_1^{(j)}A_0^{(j)^{-1}}\boldsymbol{a}_{-1}^{(j)} \\ \boldsymbol{a}_1^{(j+1)} & = -A_1^{(j)}A_0^{(j)^{-1}}\boldsymbol{a}_1^{(j)} \\ \boldsymbol{c}_{-1}^{(j+1)T} & = -\boldsymbol{c}_{-1}^{(j)^T}A_0^{(j)^{-1}}A_{-1}^{(j)} \\ \boldsymbol{c}_0^{(j+1)T} & = \boldsymbol{c}_0^{(j)T} - \boldsymbol{c}_{-1}^{(j)T}A_0^{(j)^{-1}}A_1^{(j)} - \boldsymbol{c}_1^{(j)T}A_0^{(j)^{-1}}A_{-1}^{(j)} \\ \boldsymbol{c}_1^{(j+1)T} & = -\boldsymbol{c}_1^{(j)T}A_0^{(j)^{-1}}A_1^{(j)}\end{array} \tag{15}$$

where $\boldsymbol{u}_{-1}^{(0)} = \boldsymbol{v}_{-1}^{(0)} = \boldsymbol{0}$, $\boldsymbol{u}_0^{(0)} = \boldsymbol{e}_m$, $\boldsymbol{v}_0^{(0)} = \boldsymbol{e}_1$.

The above relations, together with Theorem 2, yield the following result:

Theorem 3 *For the matrices $A_{-1}^{(j)}$, $A_0^{(j)}$, $A_1^{(j)}$, $j \geq 0$, generated by the cyclic reduction (8) we have*

$$\begin{array}{l}\Delta(A_{-1}^{(j)}) = \boldsymbol{a}_{-1}^{(j)}\boldsymbol{u}_{-1}^{(j)T} - \boldsymbol{v}_{-1}^{(j)}\boldsymbol{c}_{-1}^{(j)T} \\ \Delta(A_0^{(j)}) = \boldsymbol{a}_{-1}^{(j)}\boldsymbol{u}_0^{(j)T} + \boldsymbol{a}_0^{(j)}\boldsymbol{u}_{-1}^{(j)T} - \boldsymbol{v}_{-1}^{(j)}\boldsymbol{c}_0^{(j)T} - \boldsymbol{v}_0^{(j)}\boldsymbol{c}_{-1}^{(j)T} \\ \Delta(A_1^{(j)}) = \boldsymbol{r}^{(j)}\boldsymbol{u}_0^{(j)T} - \boldsymbol{v}_0^{(j)}\widehat{\boldsymbol{c}}^{(j)T},\end{array} \tag{16}$$

where

$$\begin{aligned} \boldsymbol{r}^{(0)} &= \boldsymbol{a}_0^{(0)}, \\ \boldsymbol{r}^{(j+1)} &= \boldsymbol{r}^{(j)} - A_1^{(j)} A_0^{(j)^{-1}} \boldsymbol{a}_{-1}^{(j)}, \quad j \geq 0, \end{aligned} \tag{17}$$

and $\widehat{\boldsymbol{c}}^{(j)T} = \boldsymbol{e}_m^T \widehat{A}^{(j)}$.

Proof. By comparing the terms in z of degree 0 and 1 in relation (12) we obtain the expression for $\Delta(A_{-1}^{(j)})$ and $\Delta(A_0^{(j)})$, respectively. Similarly, by comparing the terms of degree 2 we have

$$\Delta(A_1^{(j)}) = \boldsymbol{a}_0^{(j)} \boldsymbol{u}_0^{(j)T} + \boldsymbol{a}_1^{(j)} \boldsymbol{u}_{-1}^{(j)T} - \boldsymbol{v}_{-1}^{(j)} \boldsymbol{c}_1^{(j)T} - \boldsymbol{v}_0^{(j)} \boldsymbol{c}_0^{(j)T}. \tag{18}$$

For $j = 0$ relation (16) immediately follows. For $j \geq 1$, from an inductive application of (15) we may deduce that $\boldsymbol{u}_{-1}^{(j)T} = -\sum_{i=0}^{j-1} \boldsymbol{u}_0^{(i)T} A_0^{(i)^{-1}} A_{-1}^{(i)}$ and $\boldsymbol{v}_{-1}^{(j)} = -\sum_{i=0}^{j-1} A_{-1}^{(i)} A_0^{(i)^{-1}} \boldsymbol{v}_0^{(i)}$, whence

$$\boldsymbol{a}_1^{(j)} \boldsymbol{u}_{-1}^{(j)T} = -\sum_{i=0}^{j-1} \boldsymbol{a}_1^{(j)} \boldsymbol{u}_0^{(i)T} A_0^{(i)^{-1}} A_{-1}^{(i)}. \tag{19}$$

On the other hand, from (9), it can be easily proved by induction on j that, for any $j \geq 1$ and $0 \leq i < j$, it holds that

$$A_1^{(j)} = (-1)^{j-i} \left(\prod_{h=i}^{j-1} A_1^{(j-h+i-1)} A_0^{(j-h+i-1)^{-1}} \right) A_1^{(i)}, \tag{20}$$

where $\prod_{h=h_1}^{h_2} X_{\sigma(h)} = X_{\sigma(h_1)} \cdots X_{\sigma(h_2)}$. Whence, from (19), (20) and (14), we have

$$\begin{aligned} &\boldsymbol{a}_1^{(j)} \boldsymbol{u}_{-1}^{(j)T} = -\sum_{i=0}^{j-1} (-1)^{j-i} \left(\prod_{h=i}^{j-1} A_1^{(j-h+i-1)} A_0^{(j-h+i-1)^{-1}} \right) \boldsymbol{a}_1^{(i)} \boldsymbol{u}_0^{(i)T} A_0^{(i)^{-1}} A_{-1}^{(i)} = \\ &-\sum_{i=0}^{j-1} (-1)^{j-i} \left(\prod_{h=i}^{j-1} A_1^{(j-h+i-1)} A_0^{(j-h+i-1)^{-1}} \right) \boldsymbol{v}_0^{(i)} \boldsymbol{c}_1^{(i)T} A_0^{(i)^{-1}} A_{-1}^{(i)}. \end{aligned} \tag{21}$$

By applying the recursive relation for $\boldsymbol{v}_0^{(j)}$, given in formula (15), it may be proved by induction on j that, for any $j \geq 1$, $0 \leq i < j$, it holds that

$$\boldsymbol{v}_0^{(j)} = (-1)^{j-i} \left(\prod_{h=i}^{j-1} A_1^{(j-h+i-1)} A_0^{(j-h+i-1)^{-1}} \right) \boldsymbol{v}_0^{(i)} \tag{22}$$

Whence, from (21) and (22), it follows that

$$\boldsymbol{a}_1^{(j)} \boldsymbol{u}_{-1}^{(j)T} = -\sum_{i=0}^{j-1} \boldsymbol{v}_0^{(j)} \boldsymbol{c}_1^{(i)T} A_0^{(i)^{-1}} A_{-1}^{(i)},$$

and, in equation (18), we obtain

$$\boldsymbol{a}_1^{(j)} \boldsymbol{u}_{-1}^{(j)T} - \boldsymbol{v}_0^{(j)} \boldsymbol{c}_0^{(j)T} = -\boldsymbol{v}_0^{(j)} \left(\boldsymbol{c}_0^{(j)T} + \sum_{i=0}^{j-1} \boldsymbol{c}_1^{(i)T} A_0^{(i)^{-1}} A_{-1}^{(i)} \right) = -\boldsymbol{v}_0^{(j)} \widehat{\boldsymbol{c}}^{(j)T}.$$

Similarly, it may be proved that

$$\boldsymbol{a}_0^{(j)} \boldsymbol{u}_0^{(j)T} - \boldsymbol{v}_{-1}^{(j)} \boldsymbol{c}_1^{(j)T} = \left(\boldsymbol{a}_0^{(j)} + \sum_{i=0}^{j-1} A_{-1}^{(i)} A_0^{(i)^{-1}} \boldsymbol{a}_1^{(i)} \right) \boldsymbol{u}_0^{(j)T} = \boldsymbol{r}^{(j)} \boldsymbol{u}_0^{(j)T}$$

□

The above theorem, in the light of the results of Section 2.2, yields the following representations for the matrices $A_{-1}^{(j)}$, $A_0^{(j)}$, $A_1^{(j)}$:

Theorem 4 *At each step j of cyclic reduction, the matrices $A_{-1}^{(j)}$, $A_0^{(j)}$, $A_1^{(j)}$, $A_0^{(j)^{-1}}$ can be rewritten as*

$$\begin{aligned}
A_{-1}^{(j)} &= L(\boldsymbol{a}_{-1}^{(j)})(I - L^T(Z\boldsymbol{u}_{-1}^{(j)}) + L(\boldsymbol{v}_{-1}^{(j)})L^T(Z\boldsymbol{c}_{-1}^{(j)})\\
A_0^{(j)} &= L(\boldsymbol{a}_0^{(j)})(I - L^T(Z\boldsymbol{u}_{-1}^{(j)})) - L(\boldsymbol{a}_{-1}^{(j)})L^T(Z\boldsymbol{u}_0^{(j)}) + L(\boldsymbol{v}_{-1}^{(j)})L^T(Z\boldsymbol{c}_0^{(j)}) +\\
&\qquad L(\boldsymbol{v}_0^{(j)})L^T(Z\boldsymbol{c}_{-1}^{(j)})\\
A_1^{(j)} &= L(\boldsymbol{a}_1^{(j)}) - L(\boldsymbol{r}^{(j)})L^T(Z\boldsymbol{u}_0^{(j)}) + L(\boldsymbol{v}_0^{(j)})L^T(Z\widehat{\boldsymbol{c}}^{(j)})\\
A_0^{(j)^{-1}} &= L(A_0^{(j)^{-1}}\boldsymbol{e}_1) + L^T(ZA_0^{(j)^{-T}}\boldsymbol{u}_{-1}^{(j)}) + L(A_0^{(j)^{-1}}\boldsymbol{a}_{-1}^{(j)})L^T(ZA_0^{(j)^{-T}}\boldsymbol{u}_0^{(j)}) -\\
&\qquad L(A_0^{(j)^{-1}}\boldsymbol{v}_0^{(j)})L^T(ZA_0^{(j)^{-T}}\boldsymbol{c}_{-1}^{(j)}).
\end{aligned}$$

Proof. The proof readily follows from Theorem 3 and from Theorem 1. □

Similar properties hold for the matrices $\widehat{A}^{(j)}$:

Theorem 5 *For the matrices $\widehat{A}^{(j)}$, $j \geq 0$, generated by the cyclic reduction (8) we have*

$$\Delta(\widehat{A}^{(j)}) = -\boldsymbol{e}_1\boldsymbol{c}_{-1}^{(0)T} + \boldsymbol{a}_{-1}^{(j)}\boldsymbol{u}_0^{(j)T} - \boldsymbol{v}_{-1}^{(j)}\widehat{\boldsymbol{c}}^{(j)T}, \tag{23}$$

whence the matrices $\widehat{A}^{(j)}$ can be rewritten as

$$\widehat{A}^{(j)} = L(\widehat{\boldsymbol{a}}^{(j)}) + L^T(Z\boldsymbol{c}_{-1}^{(0)}) - L(\boldsymbol{a}_{-1}^{(j)})L^T(Z\boldsymbol{u}_0^{(j)}) + L(\boldsymbol{v}_{-1}^{(j)})L^T(Z\widehat{\boldsymbol{c}}^{(j)}),$$

where $\widehat{\boldsymbol{a}}^{(j)} = \widehat{A}^{(j)}\boldsymbol{e}_1$, $\widehat{\boldsymbol{c}}^{(j)T} = \boldsymbol{e}_m^T\widehat{A}^{(j)}$.

Proof. We prove (23) by induction on j. For $j = 0$ the equation is obviously verified. Let us suppose that (23) holds for j and let us prove it for $j+1$. From (8) and from the properties of displacement rank we have

$$\Delta(\widehat{A}^{(j+1)}) = \Delta(\widehat{A}^{(j)}) - \Delta(A_{-1}^{(j)}A_0^{(j)^{-1}}A_1^{(j)}). \tag{24}$$

From (11) it follows that

$$\begin{aligned}
&\Delta(A_{-1}^{(j)}A_0^{(j)^{-1}}A_1^{(j)}) =\\
&\Delta(A_{-1}^{(j)})A_0^{(j)^{-1}}A_1^{(j)} - A_{-1}^{(j)}A_0^{(j)^{-1}}\Delta(A_0^{(j)})A_0^{(j)^{-1}}A_1^{(j)} + A_{-1}^{(j)}A_0^{(j)^{-1}}\Delta(A_1^{(j)}).
\end{aligned}$$

By substituting in the above equation relations (16), it may be easily proved, by applying (15), that

$$\begin{aligned}
&\Delta(A_{-1}^{(j)}A_0^{(j)^{-1}}A_1^{(j)}) =\\
&-\boldsymbol{v}_{-1}^{(j+1)}\boldsymbol{c}_{-1}^{(j)T}A_0^{(j)^{-1}}A_1^{(j)} - \boldsymbol{a}_{-1}^{(j+1)}\boldsymbol{u}_0^{(j+1)T} + A_{-1}^{(j)}A_0^{(j)^{-1}}\boldsymbol{v}_{-1}^{(j)}\boldsymbol{c}_1^{(j)} + A_{-1}^{(j)}A_0^{(j)^{-1}}\boldsymbol{r}^{(j)}\boldsymbol{u}_0^{(j)}.
\end{aligned} \tag{25}$$

Whence, from (24), (25) and from the inductive hypothesis, it follows that

$$\begin{aligned}
&\Delta(\widehat{A}^{(j+1)}) =\\
&-\boldsymbol{e}_1\boldsymbol{c}_{-1}^{(0)T} + \boldsymbol{a}_{-1}^{(j+1)}\boldsymbol{u}_0^{(j+1)T} - \boldsymbol{v}_{-1}^{(j+1)}\boldsymbol{e}_m^T(\widehat{A}^{(j)} - A_{-1}^{(j)}A_0^{(j)^{-1}}A_1^{(j)}) +\\
&(\boldsymbol{a}_{-1}^{(j)} - A_{-1}^{(j)}A_0^{(j)^{-1}}\boldsymbol{r}^{(j)})\boldsymbol{u}_0^{(j)T} - A_{-1}^{(j)}A_0^{(j)^{-1}}\boldsymbol{v}_{-1}^{(j)}\boldsymbol{c}_1^{(j)T}.
\end{aligned}$$

On the other hand, by following the the same lines of the proof of Theorem 3 and from (17), it may be verified that

$$(\boldsymbol{a}_{-1}^{(j)} - A_{-1}^{(j)}A_0^{(j)^{-1}}\boldsymbol{r}^{(j)})\boldsymbol{u}_0^{(j)T} = A_{-1}^{(j)}A_0^{(j)^{-1}}\boldsymbol{v}_{-1}^{(j)}\boldsymbol{c}_1^{(j)T}.$$

□

The vectors $\widehat{\boldsymbol{a}}^{(j)}$, $\widehat{\boldsymbol{c}}^{(j)}$ obtained at two subsequent stages of cyclic reduction are related by means of the equations:

$$\begin{aligned}\widehat{\boldsymbol{a}}^{(j+1)} &= \widehat{\boldsymbol{a}}^{(j)} - A_{-1}^{(j)} A_0^{(j)^{-1}} \boldsymbol{a}_1^{(j)} \\ \widehat{\boldsymbol{c}}^{(j+1)T} &= \widehat{\boldsymbol{c}}^{(j)T} - \boldsymbol{c}_{-1}^{(j)T} A_0^{(j)^{-1}} A_1^{(j)}, \quad j \geq 0.\end{aligned} \tag{26}$$

Observe that all the above displacement properties still hold when the entries a_i of the blocks A_i, $i = -1, 0, 1$, of (1) are square matrices of dimension $h > 1$. In this case $\Delta(\cdot)$ is the block displacement operator, Z is the $m \times m$ block down-shift matrix, and the vectors $\boldsymbol{u}_i^{(j)}$, $\boldsymbol{v}_i^{(j)}$, $\boldsymbol{a}_i^{(j)}$, $\boldsymbol{c}_i^{(j)}$, $\boldsymbol{r}^{(j)}$, $\widehat{\boldsymbol{a}}_i^{(j)}$, $\widehat{\boldsymbol{c}}_i^{(j)}$ have m components, each component being a $h \times h$ matrix.

In the case of scalar entries a_i, the formulae displayed in Theorems 2–5 can be substantially simplified. In fact, from the equation (10), we deduce the following:

Proposition 6 *The matrix polynomials $A^{(j)}(z)$ and the vectors $\boldsymbol{r}^{(j)}(z)$, $\widehat{\boldsymbol{c}}^{(j)}(z)$ satisfy the following relations*

$$A^{(j)}(z)^T = JA^{(j)}(z)J, \quad \widehat{\boldsymbol{c}}^{(j)}(z) = J\boldsymbol{r}^{(j)}(z),$$

where J denotes the permutation (reversion) matrix having unit entries along the anti-diagonal.

Whence, the vectors $\boldsymbol{u}^{(j)}(z)$, $\boldsymbol{v}^{(j)}(z)$, $\boldsymbol{a}^{(j)}(z)$, $\boldsymbol{c}^{(j)}(z)$ are related, as stated by the following:

Proposition 7 *For the vectors $\boldsymbol{u}^{(j)}(z)$, $\boldsymbol{v}^{(j)}(z)$, $\boldsymbol{a}^{(j)}(z)$, $\boldsymbol{c}^{(j)}(z)$ we have $\boldsymbol{v}^{(j)}(z) = J\boldsymbol{u}^{(j)}(z)$, $\boldsymbol{c}^{(j)}(z) = J\boldsymbol{a}^{(j)}(z)$, thus*

$$\Delta(A^{(j)}(z)) = \boldsymbol{a}^{(j)}(z)\boldsymbol{v}^{(j)T}(z)J - \boldsymbol{v}^{(j)}(z)\boldsymbol{a}^{(j)T}(z)J.$$

In the case of symmetric matrices the formulae can be further simplified, since it can be easily verified that at each step j we have $A_1^{(j)} = A_{-1}^{(j)T}$, $A_0^{(j)} = A_0^{(j)T}$, $\widehat{A}^{(j)} = \widehat{A}^{(j)T}$. This simplification allows us to describe the algorithm in terms of the six vectors $\boldsymbol{a}_i^{(j)}$, $\boldsymbol{v}_i^{(j)}$, $i = -1, 0$, $\widehat{\boldsymbol{a}}^{(j)}$, $\boldsymbol{r}^{(j)}$.

The following algorithm enables us to perform a single cyclic reduction step by computing the blocks $\widehat{A}^{(j+1)}$, $A_i^{(j+1)}$, $i = -1, 0, 1$, given the blocks $\widehat{A}^{(j)}$, $A_i^{(j)}$, $i = -1, 0, 1$ by means of few FFT's of length $O(m)$. For the sake of simplicity we assume that $m = 2^s$, for a positive integer s, and that the entries a_i are scalar numbers.

Algorithm.

Input: the vectors $\boldsymbol{r}^{(j)}$, $\widehat{\boldsymbol{a}}^{(j)}$, $\boldsymbol{v}_i^{(j)}$, $i = -1, 0$, $\boldsymbol{a}_i^{(j)}$, $i = -1, 0, 1$, defining the matrices $\widehat{A}^{(j)}$, $A_i^{(j)}$, $i = -1, 0, 1$;

Output: the vectors $\boldsymbol{r}^{(j+1)}$, $\widehat{\boldsymbol{a}}^{(j+1)}$, $\boldsymbol{v}_i^{(j+1)}$, $i = -1, 0$, $\boldsymbol{a}_i^{(j+1)}$, $i = -1, 0, 1$, defining the matrices $\widehat{A}^{(j+1)}$, $A_i^{(j+1)}$, $i = -1, 0, 1$;

Computation:

1. Compute the vectors $A_0^{(j)^{-1}}\boldsymbol{e}_1$, $A_0^{(j)^{-1}}\boldsymbol{a}_{-1}^{(j)}$, $A_0^{(j)^{-1}}\boldsymbol{v}_0^{(j)}$ $ZJA_0^{(j)^{-1}}\boldsymbol{v}_{-1}^{(j)}$, that fully define the 6 vectors representing the matrix $A_0^{(j)^{-1}}$, by solving 4 Toeplitz-like $m \times m$ systems.

2. Apply relations (15), (17), (26) in order to compute the vectors $\boldsymbol{r}^{(j+1)}$, $\widehat{\boldsymbol{a}}^{(j+1)}$, $\boldsymbol{v}_i^{(j+1)}$, $i=-1,0$, $\boldsymbol{a}_i^{(j+1)}$, $i=-1,0,1$. By using the representation formula of $A_0^{(j)^{-1}}$ and of the matrices $A_1^{(j)}$, $A_{-1}^{(j)}$, $\widehat{A}^{(j)}$, together with the results of section 2.2, this computation can be performed with a finite number of real FFT's of length $2m$.

The above results, integrated with the update of the right hand side $\boldsymbol{b}^{(j)}$ by means of (4), and with a back substitution stage, leads to a direct algorithm for the solution of the finite system $T_n\boldsymbol{x}=\boldsymbol{b}$, where T_n is an $n\times n$ matrix, with $n=mq$, $q=2^p$, p positive integer.

The arithmetic cost of the latter algorithm can be easily estimated. In fact, denoting $t(m)$ the cost of solving an $m\times m$ Toeplitz-like system, it follows that $O(n\log m+(m\log m+t(m))\log\frac{n}{m})$ arithmetic operations are sufficient to carry out the algorithm. The solution of an $m\times m$ Toeplitz-like system can be computed by extending the known algorithms for Toeplitz systems. Therefore, according to the specific properties of the matrices involved (symmetry, positive definiteness, conditioning) algorithms like the fast, super fast, or iterative ones can be used for the computational cost $t(m)$ being $O(m^2)$, $O(m\log^2 m)$, $O(m\log m)$, respectively.

The cyclic reduction algorithm has also nice properties with respect to parallel computations. In fact, consider the PRAM model where we assume to have several processors working in parallel, at each step a processor performs an arithmetic operation, and the complexity of an algorithm is measured in terms of the number of parallel steps and of processors needed. Then the algorithm can be carried out in $O(t_s(m)\log\frac{n}{m})$ parallel steps, using at most $\max\{t_p(m),n\}$ processors, where $t_s(m)$, $t_p(m)$ denotes the number of parallel steps and of processors, respectively, needed to solve an $m\times m$ Toeplitz-like system. For well conditioned matrices, where the preconditioned conjugate gradient technique can be used, we have $t_s(m)=\log m$, $t_p(m)=m$.

4 Convergence properties.

In this section we relate the polynomial $a(z)=z^m\sum_{k=m_2}^{m_1}z^ka_k$ associated with the matrix T with the polynomial $\alpha(z)=\det A(z)$. The results that we obtain are used in the next section to prove some useful convergence properties of the cyclic reduction algorithm.

Let us first recall the following result of [18] that relates the zeros of $a(z)$ and $\alpha(z)$.

Theorem 8 *For the polynomial $\alpha(z)$ it holds*

$$\alpha(z^m)=\prod_{i=0}^{m-1}a(z\omega^i),\quad \omega=\cos\left(\frac{2\pi}{m}\right)+\boldsymbol{i}\sin\left(\frac{2\pi}{m}\right),\quad \boldsymbol{i}^2=-1.$$

A consequence of the above theorem is that, if ξ is zero of the polynomial $a(z)$ then ξ^m is zero of $\alpha(z)$. Conversely, if η is zero of $\alpha(z)$ then there exists an m-th root of η which is zero of $a(z)$.

Observe also that both the polynomials $a(z)$ and $\alpha(z)$ have degree $m+m_1$. This is trivial for $a(z)$; in order to prove it for $\alpha(z)$ it is sufficient to consider, in the expansion of $\det A(z)$, the term obtained by multiplying the entries on the m_1-th diagonal below the main diagonal and the entries in the $(m-m_1)$-th diagonal above the main one yielding the z^{m+m_1} term.

It is surprising to observe that the cyclic reduction step, is equivalent to the squaring step in the Graeffe algorithm which is customarily used for factoring polynomials [27], [12]. In fact by taking determinants in both sides of (10) we obtain

$$\alpha^{(j+1)}(z^2) = -\alpha^{(j)}(z)\alpha^{(j)}(-z)/\det A_0^{(j)}.$$

An immediate consequence of the above relation is that the polynomials $\alpha^{(j)}(z)$ have degree $m+m_1$ and their zeros are explicitly given by

$$\xi_i^{m2^j},\ i=1,\dots m+m_1,$$

where

$$a(\xi_i)=0,\quad i=1,\dots,m+m_1.$$

This property is very useful to prove some convergence results of the cyclic reduction algorithm.

First observe that the vector $\boldsymbol{\xi}^T=(1,\xi,\xi^2,\dots)$ is such that $\boldsymbol{\xi}^T T_\infty = (\boldsymbol{w}^T,0,\dots)$, where $\boldsymbol{w}^T$ is a suitable vector of m components, for ξ ranging in the set of zeros of $a(z)$. Moreover, assume for simplicity that $a(z)$ has at least m distinct non-null zeros $\xi_1,\dots,\xi_m$, and consider the $m\times m$ Vandermonde matrix $V_m=(\xi_i^{j-1})_{i,j=1,m}$ made up with this set of zeros. It holds

$$(V_m, DV_m, D^2V_m,\dots)T_\infty = (W,O,\dots),\quad D=\mathrm{Diag}(\xi_1^m,\dots,\xi_m^m). \tag{27}$$

Observe that, for a polynomial $a(z)$ having multiple zeros, the above relation still holds if the matrix V_m is replaced by a generalized Vandermonde matrix. The homogeneous part of the above relation can be rewritten as

$$V_mA_{-1}+DV_mA_0+D^2V_mA_1=O.$$

By applying the cyclic reduction to the system (27) we deduce that

$$V_mA_{-1}^{(j)}+D^{2^j}V_mA_0^{(j)}+D^{2^{j+1}}V_mA_1^{(j)}=O,\quad j=0,1,\dots$$

Multiplying the above relation on the left by V_m^{-1} we find

$$A_{-1}^{(j)}+F^{m2^j}A_0^{(j)}+F^{m2^{j+1}}A_1^{(j)}=O,\quad j=0,1,\dots$$

where $F=V_m^{-1}DV_m$ is the Frobenius matrix associated with the polynomial having zeros ξ_i, $i=1,\dots,m$.

The following result can be obtained from the above relations.

Theorem 9 *Let ξ_i, $i=1,\dots,m+m_1$, be the zeros of $a(z)$ ordered such that $|\xi_1|\le|\xi_2|\le\dots\le|\xi_{m+m_1}|$. If $|\xi_m|<1$ and if the blocks $A_i^{(j)}$ have norm bounded from above by a constant, then $\lim A_{-1}^{(j)}=O$. Moreover, if ξ_m is simple then for any operator norm $||\cdot||$ it holds $||A_{-1}^{(j)}||=O(|\xi_m|^{m2^j})$, otherwise, for any $\epsilon>0$ there exists an operator norm $||\cdot||$ such that $||A_{-1}^{(j)}||=O((|\xi_m|+\epsilon)^{m2^j})$. Similarly, if $|\xi_{m_1+1}|>1$ and the blocks have bounded norm then $\lim A_1^{(j)}=O$. Moreover it holds $||A_1^{(j)}||=O(|\xi_{m_1+1}|^{-m2^j})$, for any operator norm, or $||A_1^{(j)}||=O((|\xi_{m_1+1}|-\epsilon)^{-m2^j})$ for a suitable norm, according to the multiplicity of ξ_{m_1+1} as root of $a(z)$.*

Proof. For any $\epsilon>0$ let $||\cdot||$ be an operator norm such that $||F||<\rho(F)+\epsilon$ (compare [19]), where $\rho(F)$ denotes the spectral radius of F.. Then we have

$$||A_{-1}^{(j)}||\le||F||^{m2^j}||A_0^{(j)}||+||F||^{m2^{j+1}}||A_1^{(j)}||,$$

whence, since $||A_0^{(j)}||,||A_1^{(j)}||$ are bounded from above by a constant, it holds

$$||A_{-1}^{(j)}||=O((\rho(F)+\epsilon)^{m2^j})=O((|\xi_m|+\epsilon)^{m2^j}).$$

If ξ_m is a simple root of $a(z)$ then there exists a norm $||\cdot||$ such that $||F|| = |\xi_m|$. Therefore, by applying the same argument as before the above relation turns into $||A_{-1}^{(j)}|| = O(|\xi_m|^{m2^j})$. In order to prove that $||A_1^{(j)}|| = O((|\xi_{m_1+1}| - \epsilon)^{-m2^j})$ we apply a similar argument to the relation $A_{-1}^{(j)} + \hat{F}^{m2^j} A_0^{(j)} + \hat{F}^{m2^{j+1}} A_1^{(j)} = O$, where $\hat{F}$ is the Frobenius matrix associated with the polynomial having zeros $\xi_{m_1+1}, \ldots, \xi_{m_1+m}$. $\square$

Conditions under which the blocks $A_i^{(j)}$ have bounded norm are given in [22], [33]. These condition can be substantially relaxed in the real symmetric case where we only require that the polynomial $a(z)$ has no zeros of modulus 1. In this case the convergence speed depends on the modulus of the zero ξ closest to the unit circle. These properties are investigated with more details in the next section.

A very important situation where the blocks $A_i^{(j)}$ have bounded norm is the case $T_n = I - P_n$, with P_n stochastic matrix. This case occurs in the analysis of Markov chains arising in queueing problems [9], [10].

4.1 The real symmetric positive definite case.

Consider the case where the matrix T_n is real symmetric, so that for the blocks $A_i^{(j)}$ it holds $A_{-1}^{(j)} = A_1^{(j)T}$ and $A_0^{(j)} = A_0^{(j)T}$, $j = 0, 1, \ldots$. Assume for simplicity that the polynomial $a(z)$ has no zero of modulus 1. Due to the symmetricity properties, the polynomial $a(z)$ is symmetric, that is, $a(z) = z^{2m} a(z^{-1})$, and its zeros occur in pairs $\xi_i = \xi_{2m-i+1}^{-1}$, $i = 1, \ldots, m$. Let us arrange the zeros ξ_i in such a way that $|\xi_1| \le |\xi_2| \le \ldots \le |\xi_m| < 1 < |\xi_{m+1}| \le \ldots \le |\xi_{2m}|$.

Due to the Szëgo theory on Toeplitz matrices [21], the matrices T_n, are either positive definite or negative definite for any n according to the sign that $a(z)$ takes over the unit circle. Throughout we assume, for simplicity T_n positive definite. Moreover, the eigenvalues of the matrix T_n belong to the interval $[\min_{|z|=1} |a(z)|, \max_{|z|=1} |a(z)|]$.

It is useful to observe that the block tridiagonal matrices $T^{(j)}$, generated at each step of the cyclic reduction, being Schur complements in T_n, have their eigenvalues belonging to the interval $[\min \lambda(T_n), \max \lambda(T_n)]$, where $\lambda(A)$ denotes the set of eigenvalues of the matrix A. Similarly the diagonal blocks $A_0^{(j)}$, $\widehat{A}^{(j)}$, $\begin{bmatrix} A_0^{(j)} & A_1^{(j)T} \\ A_1^{(j)} & A_0^{(j)} \end{bmatrix}$, share the same properties due to the Cauchy interlace theorem [19]. Therefore it follows that the all the blocks $A_0^{(j)}$, $A_1^{(j)}$, $\widehat{A}^{(j)}$, have norm bounded from above by a constant.

These observations allow us to conclude with the following theorem

Theorem 10 *Let T_n be a real symmetric matrix (Hermitian) matrix such that the associated polynomial $a(z)$ has no zeros in the unit circle. Then for the matrices $A_0^{(j)}$, $A_1^{(j)}$, $\widehat{A}^{(j)}$, generated by the cyclic reduction we have*

$$\lim A_1^{(j)} = O, \quad ||A_1^{(j)}|| \le \eta |\xi_m|^{m2^j}$$

The above theorem can be used in order to devise an algorithm for the approximation of k components of the solution of the system $T_n \boldsymbol{x} = \boldsymbol{b}$, $q \in \boldsymbol{N}_\infty$ in $O((m \log m + t(m)) \log\log \epsilon^{-1} + k \log m)$ arithmetic operations provided that $k \le mq$. More precisely, in the light of theorem 9, we may replace the term $\log\log \epsilon^{-1}$ with the expression $\log_2 \log \theta\epsilon^{-1} - \log_2 \log |\xi_{m-1}|^{-m}$ for a suitable constant θ depending on the specific norm adopted in the proof of theorem 9.

4.2 The unsymmetric/indefinite case: normal equations

If the polynomial $a(z)$ is not symmetric or if $a(\xi)=0$ for some ξ of unit modulus, the convergence results of the previous section do not generally hold and the cyclic reduction technique cannot be used for solving the linear system $T_n\boldsymbol{x}=\boldsymbol{b}$ in an iterative fashion. Moreover, we cannot bound with a constant the condition number of the $m\times m$ blocks to be inverted at each step of cyclic reduction unless $a(z)$ is nonnegative for $|z|=1$. Therefore, the stability of the cyclic reduction technique is not insured in this case.

However, with a simple trick we may reduce the general case to the positive definite one.

Let us first assume that $a(z)$ has zeros $\xi_1,\ldots\xi_{m+m_1}$ and that $|\xi_k|=1$ for a given k. We may replace the system $T_n\boldsymbol{x}=\boldsymbol{b}$ with the systems

$$\begin{aligned}&DT_nD^{-1}\boldsymbol{y}=D\boldsymbol{b}\\&\boldsymbol{y}=D\boldsymbol{x},\end{aligned}$$

where $D=\mathrm{Diag}(1,\gamma,\gamma^2,\ldots,\gamma^{n-1})$, for a suitable γ. In this way the new matrix DT_nD^{-1} is still Toeplitz and is associated with the polynomial $a_\gamma(z)=a(\gamma z)$. The parameter γ can be chosen in such a way that $a_\gamma(z)$ has no zero of modulus 1. This can be done, deterministically, if informations about the moduli of the zeros of $a(z)$ are known, or by means of randomization with a random choice of γ in the range $[1,1+\sigma]$ for a given positive σ, which guarantees with probability 1 that $a_\gamma(z)$ has no zero on the unit circle.

Indeed, if T_n is real symmetric, the new system is not anymore symmetric. In order to keep the symmetricity property, we switch to the normal equations

$$\begin{aligned}&K_n\boldsymbol{y}=\boldsymbol{c}\\&K_n=(DT_nD^{-1})^T(DT_nD^{-1})\\&\boldsymbol{c}=(DT_nD^{-1})^TD\boldsymbol{b}.\end{aligned}\tag{28}$$

It is easy to check that K_n has $4m+1$ diagonals and is still Toeplitz except for the entries in the leading and in the trailing $m\times m$ sub-matrices. Moreover, for the Toeplitz part, the matrix is associated with the polynomial $q_\gamma(z)=\phi_\gamma(z)\phi_\gamma(z^{-1})z^{(m_1+m)}$ having zeros $\xi_1,\ldots,\xi_{m+m_1},\xi_1^{-1},\ldots,\xi_{m_1+m}^{-1}$. More precisely $K_n=V_n-\mathrm{Diag}(R,O,S)$, where V_n is the banded Toeplitz matrix associated with $q_\gamma(z)$, $R=A_{-1}^TA_0^{-1}$, $S=A_1^TA_1$, and O denotes a null matrix of size $n-2m$.

Partitioning K_n into $2m\times 2m$ blocks yields

$$K_n=\begin{bmatrix}\widehat{C} & B_{-1} & & & \\ B_1 & B_0 & B_{-1} & & \\ & \ddots & \ddots & \ddots & \\ & & B_1 & B_0 & B_{-1}\\ & & & B_1 & \widetilde{C}\end{bmatrix}\tag{29}$$

Let us assume $n=2mq$, where $q=2^p-1$ and p is a positive integer. In this way the cyclic reduction algorithm applied to the system (28) generates a sequence of systems whose matrices, of dimension $(2^{p-j}-1)2m$, have the same structure of (29) and are defined by the $2m\times 2m$ blocks $\widehat{C}^{(j)}$, $\widetilde{C}^{(j)}$, $B_i^{(j)}$, $i=-1,0,1$, such that

$$\left\{\begin{aligned}B_{-1}^{(j+1)}&=-B_{-1}^{(j)}B_0^{(j)^{-1}}B_{-1}^{(j)}\\B_0^{(j+1)}&=B_0^{(j)}-B_1^{(j)}B_0^{(j)^{-1}}B_{-1}^{(j)}-B_{-1}^{(j)}B_0^{(j)^{-1}}B_1^{(j)}\\B_1^{(j+1)}&=-B_1^{(j)}B_0^{(j)^{-1}}B_1^{(j)}\\\widehat{B}^{(j+1)}&=\widehat{B}^{(j)}-B_{-1}^{(j)}B_0^{(j)^{-1}}B_1^{(j)}\\\widehat{C}^{(j+1)}&=\widehat{B}^{(j+1)}+\widehat{C}-B_0\\\widetilde{C}^{(j+1)}&=J\widehat{B}^{(j+1)}J+\widetilde{C}-B_0\end{aligned}\right.$$

	CR		PCG		
n	Time (s.)	Residual	Time (s.)	Residual	PCG/CR
8192	0.6	3. e -13	1.0	2. e -11	1.7
16384	1.0	3. e -13	2.6	2. e -11	2.6
32768	1.8	7. e -13	6.9	2. e -11	3.8
65536	3.5	7. e -13	16.0	2. e -11	4.6
131072	6.7	7. e -13	36.9	2. e -11	5.5
262144	13.3	7. e -13	132.0	2. e -11	9.9
524288	26.5	7. e -13	376.3	2. e -11	14.2
1048576	52.8	7. e -13	*	*	*
2097152	98.3	7. e -13	*	*	*

Table 1: $m = 16,\ a_0 = 8$

In the light of the results of the previous section the matrices $\widehat{C}^{(j)}$, $\widetilde{C}^{(j)}$ and $B_i^{(j)}$, $i = -1, 0, 1$, have displacement rank independent of j and an algorithm similar to the one of section 4.1 applies.

5 Numerical experiments and applications.

We have implemented several versions of our algorithms in Fortran 90. The most general one deals with banded (unsymmetric) block Toeplitz matrices. Two more specific versions deal with the cases of (unsymmetric) scalar Toeplitz matrices and with symmetric scalar Toeplitz matrices. For the inversion of the diagonal blocks we used the conjugate gradient method with no preconditioning. We performed several numerical experiments, here we report the ones concerning the scalar symmetric case.

The test matrices that we have considered are defined by the function $\phi(z) = \sum_{i=-m}^{m} z^i + (a_0 - 1)$, for $m = 16, 32, 64$, $n = 2^h$, $h = 14, 15, \ldots, 22$, where the value of a_0 has been chosen differently in such a way that the polynomial $z^m\phi(z)$ has some zeros more or less close to the unit circle. The right-hand vector of the system has been chosen so that the system has the solution $\boldsymbol{x} = (1, \ldots, 1)^T$.

More specifically, for $m = 16$, we tested the algorithm with $a_0 = 8.2$. In this case, the polynomial $\phi(z)z^m$ has no zeros in the unit circle, the zeros closest to the unit circle are two pairs of modulus 1.003052 and 0.99695731. Their ratio, 1.006, is far enough from 1 in order to allow the double exponential convergence of the cyclic reduction algorithm.

For $m = 32$ we have chosen $a_0 = 15.1315$. With this value, the polynomial $\phi(z)z^m$ has no zeros of modulus 1. The zeros closest to the unit circle have modulus 1.00009 and 0.9999099, and their ratio has modulus 1.00018.

For $m = 64$ we set $a_0 = 25$. With this value the cyclic reduction algorithm does not converge since the polynomial $\phi(z)z^m$ has two pairs of complex zeros of modulus 1.

The algorithm has been compared with the preconditioned conjugate gradient method obtained by using the circulant preconditioner of G. Strang [32], on an alpha workstation.

In the tables 1, 2, 3 we report the time in seconds needed by the two algorithms (a "*" denotes failure of the algorithm due to lack of memory), the maximum error in the solution and the ratios between the times needed.

The numerical experiments confirm that our algorithm is faster than the preconditioned conjugate gradient method, more efficient in terms of memory use and

	CR		PCG		
n	Time (s.)	Residual	Time (s.)	Residual	PCG/CR
8192	0.7	2. e -11	1.3	1. e -12	1.9
16384	1.1	1. e -10	3.3	1. e -12	3.0
32768	1.8	2. e -10	7.9	1. e -12	4.3
65536	3.4	2. e -10	17.2	4. e -11	5.1
131072	6.4	2. e -10	38.9	4. e -11	6.1
262144	12.3	2. e -10	139.3	4. e -11	11.3
524288	24.3	2. e -10	395.9	4. e -11	16.3
1048576	48.2	2. e -10	*	*	*
2097152	95.6	2. e -10	*	*	*

Table 2: $m = 32,\ a_0 = 15.1315$

	CR		PCG		
n	Time (s.)	Residual	Time (s.)	Residual	PCG/CR
8192	0.8	2. e -11	1.0	3. e -12	1.3
16384	1.2	4. e -11	2.5	1. e -11	2.1
32768	2.1	1. e -11	6.6	2. e -11	3.1
65536	3.6	6. e -12	15.4	6. e -12	4.3
131072	6.8	2. e -10	37.0	1. e -11	5.4
262144	13.0	4. e -11	123.0	4. e -11	9.5
524288	25.4	4. e -11	328.0	6. e -12	12.9
1048576	49.8	1. e -9	*	*	*
2097152	99.2	8. e -11	*	*	*

Table 3: $m = 64,\ a_0 = 25$

numerically stable.

References

[1] G. S. Ammar and B. Gragg. Numerical experience with superfast real Toeplitz solver. *Linear Algebra Appl.*, 121:185–206, 1989.

[2] O. Axelsson and G. Lindskög. The rate of convergence of the preconditioned conjugate gradient method. *Numer. Math.*, 52:499–523, 1986.

[3] D. Bini. On a class of matrices related to Toeplitz matrices. Technical Report 83-5, SUNY University, Albany, New York, 1983.

[4] D. Bini. Parallel solution of certain Toeplitz linear systems. *SIAM J. Comput.*, 13:268–276, 1984.

[5] D. Bini. Matrix structures in parallel matrix computations. *Calcolo*, 25:37–51, 1988.

[6] D. Bini and M. Capovani. Spectral and computational properties of band symmetric Toeplitz matrices. *Linear Algebra Appl.*, 52:99–126, 1983.

[7] D. Bini and F. Di Benedetto. A new preconditioner for the parallel solution of positive definite Toeplitz systems. In *Proceedings* 2^{nd} *Annual SPAA, Crete, Greece*, pages 220–223. ACM Press, July 1990.

[8] D. Bini and B. Meini. On cyclic reduction applied to a class of Toeplitz-like matrices arising in queueing problems. In W. J. Stewart, editor, *Computations with Markov Chains*, pages 21–38. Kluwer Academic Publisher, 1996.

[9] D. Bini and B. Meini. On the solution of a nonlinear matrix equation arising in queueing problems. *SIAM J. Matrix Anal. Appl.*, 17:906–926, 1996.

[10] D. Bini and B. Meini. Improved cyclic reduction for solving queueing problems. *Numerical Algorithms*, 1997. To appear.

[11] D. Bini and B. Meini. Inverting block Toeplitz matrices in block Hessenberg form by means of displacement operators: application to queueing problems. *Linear Algebra Appl.*, 1997. To appear.

[12] D. Bini and V. Pan. *Matrix and Polynomial Computations, Vol. 1: Fundamental Algorithms.* Birkhäuser, Boston, 1994.

[13] S. Bondeli and W. Gander. Cyclic reduction for special tridiagonal systems. *SIAM J. Matrix Anal. Appl.*, 15:321–330, 1994.

[14] R. H. Chan. Circulant preconditioners for Hermitian Toeplitz systems. *SIAM J. Matrix Anal. Appl.*, 10:542–550, 1989.

[15] R. H. Chan and M. K. Ng. Conjugate gradient methods for Toeplitz systems. *SIAM Review*, 38:427–482, 1996.

[16] R. H. Chan and M. K. Ng. Scientific applications of iterative Toeplitz solvers. *Calcolo*, 33:249–267, 1996.

[17] R. H. Chan and P. Tang. Fast band-Toeplitz preconditioners for Hermitian Toeplitz systems. *SIAM J. Sci. Comput.*, 15:164–171, 1994.

[18] H. R. Gail, S. L. Hantler, and B. A. Taylor. Non-skip-free M/G/1 and G/M/1 type Markov chains. *Adv. Appl. Prob.*, 1997. To appear.

[19] G.H. Golub and C.F. van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, 1989.

[20] J. Grcar and A. Sameh. On certain parallel Toeplitz linear system solvers. *SIAM J. Sci. Stat. Comput.*, 2:238–256, 1981.

[21] U. Grenander and G. Szegö. *Toeplitz Forms and Their Applications*. Chelsea Publishing, New York, 2nd edition, 1984.

[22] Don Heller. Some aspects of the cyclic reduction algorithm for block tridiagonal linear systems. *SIAM J. Num. Anal.*, 13:484–496, 1976.

[23] T. Kailath and A.H. Sayed. Displacement structure: theory and applications. *SIAM Review*, 37:297–386, 1995.

[24] T. Kailath, A. Viera, and M. Morf. Inverses of Toeplitz operators, innovations and orthogonal polynomials. *SIAM Review*, 20:106–119, 1978.

[25] M. F. Neuts. *Matrix-Geometric Solutions in Stochastic Models*. Johns Hopkins University Press, Baltimore, 1981.

[26] M. F. Neuts. *Structured Stochastic Matrices of M/G/1 Type and Their Applications*. Marcel Dekk., New York, 1989.

[27] M. A. Ostrowski. Recherches sur la methode de Graeffe et les zeros des polynomes et des series de Laurent. *Acta Math.*, 72:99–257, 1940.

[28] S. Serra. Preconditioning strategies for asymptotically ill-conditioned block Toeplitz systems. *BIT*, 34:579–594, 1994.

[29] S. Serra. Optimal, quasi-optimal and superlinear preconditioners for asymptotically ill-conditioned positive definite Toeplitz systems. *Math. Comp.*, 1997. to appear.

[30] S. Serra Capizzano. Toeplitz preconditioner constructed from linear approximation processes. submitted for publication, 1997.

[31] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*. Springer Verlag, New York, 1980.

[32] G. Strang. A proposal for Toeplitz matrix computations. *Studies in Applied Mathematics*, 74:171–176, 1986.

[33] P. Yalamov and V. Pavlov. Stability of the block cyclic reduction. *Linear Algebra Appl.*, 249:341–358, 1996.