Using Displacement Structure for Solving Non-Skip-Free M/G/1 Type Markov Chains

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

Abstract: We propose a new quadratically convergent algorithm for solving Non-Skip-Free M/G/1 type Markov chains. This algorithm consists in applying the cyclic reduction technique of [3], [5] to the reblocked M/G/1 type Markov chain and in exploiting the Toeplitz structure by means of the concept of displacement rank. The complexity of our algorithm, applied to a Markov chain with k boundary levels and with the number of phases equal to p, is $O(np^3k\log^2 k + p^2kn\log(kn))$ arithmetic operations, where n is the numerical truncation level of the power series involved. This complexity bound is better than the $O(m^3n + m^2n\log n)$ bound of the cyclic reduction applied to the reblocked matrix without using the inner structure of the problem, where $m = kp$. Structural properties of the corresponding G matrix are proved in terms of displacement operators. For QBD problems the cost is reduced to $O(p^3k\log^2 k)$ operations, moreover, convergence results are proved in terms of the location of the zeros of a suitable polynomial associated with the Markov chain. Numerical results are presented.

1 INTRODUCTION

Numerical methods for solving M/G/1 type Markov chains [20] have been widely investigated in the literature [8], [15], [16], [17], [18], [1]. Markov chains of this kind are defined by transition matrices P of the form

$$P^T = \begin{bmatrix} B_1 & A_0 & & & 0 \\ B_2 & A_1 & A_0 & & \\ B_3 & A_2 & A_1 & A_0 & \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix} \tag{1}$$

where A_i, B_{i+1}, $i \geq 0$, are $m \times m$ nonnegative matrices such that $\sum_{i=0}^{+\infty} A_i$ and $\sum_{i=1}^{+\infty} B_i$ are column stochastic. The fundamental problem of computing the probability invariant vector $\boldsymbol{\pi}$ such that $(I - P^T)\boldsymbol{\pi} = \mathbf{0}$, can be reduced to compute the minimal nonnegative solution G of the nonlinear matrix equation

$$X = \sum_{i=0}^{+\infty} X^i A_i \tag{2}$$

and to apply the celebrated Ramaswami formula [22].

Conversely, the equation (2) can be equivalently rewritten in terms of block Hessenberg block Toeplitz matrices in the following form

$$[G, G^2, G^3, \ldots] \begin{bmatrix} I - A_1 & -A_0 & & & 0 \\ -A_2 & I - A_1 & -A_0 & & \\ -A_3 & -A_2 & I - A_1 & -A_0 & \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} = [A_0, 0, 0, \ldots]. \tag{3}$$

A specific feature of the matrices in (1) and (3) is the block Hessenberg form. Moreover, except for the first block column of P^T, both the matrices in (1) and (3) share the block Toeplitz structure. Here we recall that a block matrix has a block Toeplitz structure if its blocks are constant along each block diagonal.

The matrix in (3) is uniquely determined by the formal matrix power series $A(z) = \sum_{i=0}^{+\infty} A_i z^i$ which is convergent for $|z| \leq 1$, due to the stochasticity of P^T.

Several numerical methods for solving (2) have been introduced and analyzed in the literature. Besides the more classical ones, based on functional iterations [15], [19], [21], and the ones based on the factorization of the matrix P^T [11], new quadratically convergent methods, that fully exploit the block Toeplitz and the block Hessenberg structure of (3), have been proposed in [6], [2], [3], [5]. The latter methods are based on specific tools for manipulating Toeplitz matrices [7] and are ultimately reduced to perform operations among matrix power series by means of Fast Fourier Trasforms (FFT).

More general M/G/1 type Markov chains are defined by transition matrices of the form

$$P^T = \begin{bmatrix} \beta_{1,1} & \ldots & \beta_{1,k} & \alpha_0 & & & 0 \\ \beta_{2,1} & \ldots & \beta_{2,k} & \alpha_1 & \alpha_0 & & \\ \beta_{3,1} & \ldots & \beta_{3,k} & \alpha_2 & \alpha_1 & \alpha_0 & \\ \vdots & & \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}, \tag{4}$$

where α_i, $\beta_{i+1,j}$, $i \geq 0$, $j = 1, \ldots, k$, are $p \times p$ nonnegative matrices such that $\sum_{i=0}^{+\infty} \alpha_i$ and $\sum_{i=1}^{+\infty} \beta_{i,j}$, $j = 1, \ldots, k$, are column stochastic. This kind of Markov chains are called *Non-Skip-Free* in [9]. It is convenient to associate with the matrix P of (4) the $p \times p$ matrix power series $\alpha(z) = \sum_{i=0}^{+\infty} \alpha_i z^i$.

The latter class of M/G/1 type Markov chains can be reduced to the block Hessenberg form (1) by partitioning the matrix P^T into blocks A_i, B_i, of dimension $m = kp$ such that

$$B_i = \begin{bmatrix} \beta_{(i-1)k+1,1} & \beta_{(i-1)k+1,2} & \dots & \beta_{(i-1)k+1,k} \\ \beta_{(i-1)k+2,1} & \beta_{(i-1)k+2,2} & \dots & \beta_{(i-1)k+2,k} \\ \vdots & \vdots & & \vdots \\ \beta_{ik,1} & \beta_{ik,2} & \dots & \beta_{ik,k} \end{bmatrix}, \ i \geq 1,$$
$$A_0 = \begin{bmatrix} \alpha_0 & 0 & \dots & 0 \\ \alpha_1 & \alpha_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \alpha_{k-1} & \dots & \alpha_1 & \alpha_0 \end{bmatrix}, \ A_i = \begin{bmatrix} \alpha_{ik} & \alpha_{ik-1} & \dots & \alpha_{ik-k+1} \\ \alpha_{ik+1} & \alpha_{ik} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \alpha_{ik-1} \\ \alpha_{ik+k-1} & \dots & \alpha_{ik+1} & \alpha_{ik} \end{bmatrix}, \ i \geq 1. \tag{5}$$

Thus, the algorithms for solving matrices of the Hessenberg structure (1) can still be applied to solve matrices of the more general form (4). However, in this way the more specific structure of the blocks A_i of the reblocked matrix is not exploited at all. Moreover, in the case where the number k of boundary levels is large, manipulating matrices of dimension $m = kp$ can be impractical.

Recently, in [9], a specific algorithm for computing the $m \times m$ matrix G associated with the reblocked Hessenberg matrix has been proposed. The authors observe that the matrix G can be expressed in terms of a $k \times k$ block Frobenius matrix with blocks of size p, that is uniquely defined by a nonnegative block vector $\boldsymbol{g}$ having k blocks of dimension p. In order to determine the block vector $\boldsymbol{g}$ it is proposed an iterative technique, for computing a sequence $\boldsymbol{g}_n$ of block vectors that monotonically converges to $\boldsymbol{g}$. This approach allows one to drastically reduce the computational cost, with respect to the customary functional iteration methods. However, the convergence speed of this scheme is still linear.

In this paper we adapt the cyclic reduction method, introduced and analyzed in [3], [2], [5], for the computation of the $m \times m$ matrix G associated with the M/G/1 type block Hessenberg matrix obtained by reblocking (4), and arrive at a quadratically convergent method for solving (3), having almost the same cost per iteration and the same memory storage requirement of the algorithm of [9].

The cyclic reduction algorithm generates two sequences of matrix power series

$$\begin{aligned} A^{(j)}(z) &= \textstyle\sum_{i=0}^{+\infty} A_i^{(j)} z^i, \ j \geq 0, \quad A^{(0)}(z) = \sum_{i=0}^{+\infty} A_i z^i, \\ \widehat{A}^{(j)}(z) &= \textstyle\sum_{i=0}^{+\infty} \widehat{A}_{i+1}^{(j)} z^i, \ j \geq 0, \quad \widehat{A}^{(0)}(z) = \sum_{i=0}^{+\infty} A_{i+1} z^i, \end{aligned}$$

such that G^{2^j} is the nonnegative solution of the matrix equation $X = \sum_{i=0}^{+\infty} X^i A_i^{(j)}$, and $G = A_0 \left(I - \sum_{i=0}^{+\infty} G^{i \cdot 2^j} \widehat{A}_{i+1}^{(j)} \right)^{-1}$. The two sequences of matrix power series $\left\{ A^{(j)}(z) \right\}_j$, $\left\{ \widehat{A}^{(j)}(z) \right\}_j$, quadratically converge to two matrix polynomials of degree one and zero, respectively. This allows us to recover the matrix G by means of a quadratically convergent scheme. For more details we refer the reader to [3], [5].

In the case where the blocks A_i, $i \geq 0$, have the structure (5), at the first stage, for $j = 0$, the matrices $A^{(j)}(z)$, $\widehat{A}^{(j)}(z)$ are both block Toeplitz for any z. But after just the first step of cyclic reduction the Toeplitz structure of the matrices $A^{(j)}(z)$, $\widehat{A}^{(j)}(z)$ is destroyed.

The main result of this paper, on which the effectiveness of our algorithm is based, is that the matrices $A^{(j)}(z)$, $\widehat{A}^{(j)}(z)$ still keep the Toeplitz-like structure, for any $j \geq 1$, for any z. In other words, we prove that, despite the inner block Toeplitz structure of $A^{(j)}(z)$ and $\widehat{A}^{(j)}(z)$ is destroyed after few steps of cyclic reduction, the matrix power series $A^{(j)}(z)$ and $\widehat{A}^{(j)}(z)$ can be represented as sum of at most four products between block lower triangular and block upper triangular block Toeplitz matrices.

As a consequence, we may take advantage of this property both in terms of memory storage and in terms of computational cost. In fact, all the matrix power series generated by cyclic reduction, as well as the matrix G, can still be represented in terms of at most four products of block lower and block upper triangular block Toeplitz matrices defined by few block vectors of k components, each component being a $p \times p$ matrix; moreover we provide an explicit functional relation between the block vectors obtained at consecutive stages of cyclic reduction.

By using the properties of FFT and of Toeplitz matrix computations we prove that each step of cyclic reduction can be performed with $O(np^3 k \log^2 k + p^2 kn \log(kn))$ arithmetic operations, where n is the numerical truncation level such that A_i are negligible for $i > n$ with respect to the unit round-off. This complexity bound is better than the $O(m^3 n + m^2 n \log n)$ bound of the cyclic reduction applied to the reblocked matrix (4) without exploiting the inner Toeplitz structure of the blocks, where $m = kp$.

In the case of QBD problems, where the matrix obtained by reblocking (4) is block tridiagonal, our algorithm is much simplified and reduces to the algorithm proposed in [4]. In particular, the computational cost per step is reduced to $O(p^3 k \log^2 k)$ operations. Moreover, concerning the convergence speed, we prove that the sequences of blocks $A_2^{(j)}$ and $\widehat{A}_2^{(j)}$, generated by the cyclic reduction, tends to zero as $O(\theta^{-k2^j})$ for $j \to +\infty$, where

$$\theta = \min\{|z| : |z| > 1, \det(z^k I - \alpha(z)) = 0\}.$$

The information provided by the latter result is very useful in order to compare the speed of this method with the speed of the methods based on functional iterations, which depends on the location of the solutions of the equation $\det(z^k I - \alpha(z)) = 0$ inside the unit disk (compare [19]).

The paper is organized as follows. In Section 2 we recall the main features of cyclic reduction adjusted to deal with infinite stochastic matrices. Section 3 is devoted to the main tools needed for devising our algorithm. More specifically, Section 3.1 is concerned with power series arithmetic and FFT, while Section 3.2 deals with Toeplitz computations and with the concept of displacement rank. The main results of the paper are contained in Sections 4 and 5 where we prove the invariance of the displacement rank under cyclic reduction (Section 4) and relate the convergence of our technique to the location of the zeros of the polynomial $\det(z^k I - \alpha(z))$ in the case of QBD processes. In Section 6 we report the results of some numerical experiments and compare the performances of our algorithm and the algorithms of [9] and [3]. From the numerical experiments obtained with classes of randomly generated stochastic matrices it follows that our algorithm is much faster than the ones of [9], [3], for moderately large values of k.

2 CYCLIC REDUCTION

In this section we briefly recall the main features of cyclic reduction applied to compute the solution of the matrix equation (2) by solving the system (3). For more details we refer the reader to [2], [3], [5]. Hereafter we suppose that the matrix P is irreducible and positive recurrent.

By performing an odd-even permutation of the block rows and block columns in (3) we find that

$$\left[G^2, G^4, \ldots \mid G, G^3, \ldots\right] \begin{bmatrix} T_1 & W \\ V & T_2 \end{bmatrix} = [0, 0, \ldots \mid A_0, 0, \ldots] \tag{6}$$

where T_1, T_2 and V are infinite block lower triangular block Toeplitz matrices, W is an infinite block Hessenberg block Toeplitz matrix. Applying one step of block-Gaussian elimination to the 2×2 block system (6) yields

$$\left[G, G^3, G^5, \ldots\right] S^{(1)} = [A_0, 0, 0, \ldots] \tag{7}$$

where $S^{(1)} = T_2 - VT_1^{-1}W$ is the Schur complement of T_2 (the matrix T_1 is nonsingular, due to the irreducibility of P). It is interesting to observe that $Q^{(1)} = I - S^{(1)}$ is a nonnegative block Hessenberg matrix which, except for the first block column, has the block Toeplitz structure of the matrix of (3).

In other words, $S^{(1)}$ is uniquely determined by the blocks $\widehat{A}_i^{(1)}$, $i \geq 1$, $A_i^{(1)}$, $i \geq 0$, defining the first two block columns of $Q^{(1)}$.

We may recursively apply the same reduction (cyclic reduction) to the matrix equation (7) thus obtaining a sequence of matrices $\{S^{(j)}\}_{j\geq 0}$ such that $S^{(0)}$ is the matrix of (3), $Q^{(j)} = I - S^{(j)}$ is a nonnegative matrix having, except for the first block column, the same block Hessenberg and Toeplitz structure of the matrix in (3). Each matrix $Q^{(j)}$ is determined by the blocks $\widehat{A}_i^{(j)}$, $i \geq 1$, $A_i^{(j)}$, $i \geq 0$, defining its first two block columns.

At the general step j the following equation is satisfied

$$[G, G^{2^j+1}, G^{2\cdot 2^j+1}, \ldots]S^{(j)} = [A_0, 0, 0, \ldots]. \tag{8}$$

The solution of (8) provides the matrix G that solves (2). Moreover, as j grows, the equation (8) quadratically converges to a linear system that can be directly solved (see [3] and the next Theorem 2).

The block entries $\{A_i^{(j+1)}\}_{i\geq 0}$, $\{\widehat{A}_i^{(j+1)}\}_{i\geq 1}$, obtained at step $j+1$, are related to the block entries $\{A_i^{(j)}\}_{i\geq 0}$, $\{\widehat{A}_i^{(j)}\}_{i\geq 1}$, obtained at step j, by means of functional relations, that express in compact form the block odd-even permutation and the Gaussian elimination of one step of cyclic reduction. More precisely, for any $j \geq 0$, let us associate with the matrix sequences $\{A_i^{(j)}\}_{i\geq 0}$, $\{\widehat{A}_i^{(j)}\}_{i\geq 1}$, the formal matrix power series $A^{(j)}(z) = \sum_{i=0}^{+\infty} A_i^{(j)} z^i$, $\widehat{A}^{(j)}(z) = \sum_{i=0}^{+\infty} \widehat{A}_{i+1}^{(j)} z^i$, respectively, where $A_i^{(0)} = \widehat{A}_i^{(0)} = A_i$, $i \geq 1$, $A_0^{(0)} = A_0$. Then the following functional relations hold:

$$\left\{ \begin{array}{l} A^{(j+1)}(z) = zA_{odd}^{(j)}(z) + A_{even}^{(j)}(z)(I - A_{odd}^{(j)}(z))^{-1}A_{even}^{(j)}(z) \\ \widehat{A}^{(j+1)}(z) = \widehat{A}_{even}^{(j)}(z) + A_{even}^{(j)}(z)(I - A_{odd}^{(j)}(z))^{-1}\widehat{A}_{odd}^{(j)}(z) \end{array} \right. \tag{9}$$

where, for any matrix power series $B(z) = \sum_{i=0}^{+\infty} B_i z^i$, we define

$$\begin{array}{l} [B(z)]_{even} = B_{even}(z) = \sum_{i=0}^{+\infty} B_{2i} z^i, \\ [B(z)]_{odd} = B_{odd}(z) = \sum_{i=0}^{+\infty} B_{2i+1} z^i. \end{array}$$

The relations (9), besides expressing in compact form the recursions which the cyclic reduction method is based on, provide a basic tool for the efficient computation of the matrices $\{A_i^{(j)}\}_{i\geq 0}$, $\{\widehat{A}_i^{(j)}\}_{i\geq 1}$, by using FFT's.

The first equation of (9) can be rewritten in a more convenient form that will be used in the next section to prove some useful structural properties of the cyclic reduction method applied to non-skip-free M/G/1 type Markov chains. For this purpose we introduce the matrix power series

$$H^{(j)}(z) = zI - A^{(j)}(z) = \sum_{i=0}^{+\infty} H_i^{(j)} z^i, \quad \widehat{H}^{(j)}(z) = I - \widehat{A}^{(j)}(z) = \sum_{i=0}^{+\infty} \widehat{H}_{i+1}^{(j)} z^i, \quad j \geq 0.$$

With the latter notations, since

$$B_{even}(z^2) = \frac{B(z)+B(-z)}{2}, \quad B_{odd}(z^2) = \frac{B(z)-B(-z)}{2z},$$

the equation (9) can be rewritten in the following way

$$\begin{array}{l} H^{(j+1)}(z^2) = \left(\frac{H^{(j)}(z)^{-1} - H^{(j)}(-z)^{-1}}{2z}\right)^{-1} = -H^{(j)}(z) H^{(j)}_{odd}(z^2)^{-1} H^{(j)}(-z), \\ \widehat{H}^{(j+1)}(z) = \widehat{H}^{(j)}_{even}(z) - H^{(j)}_{even}(z) H^{(j)}_{odd}(z)^{-1} \widehat{H}^{(j)}_{odd}(z). \end{array} \tag{10}$$

The matrix power series $A^{(j)}(z)$, $\widehat{A}^{(j)}(z)$ generated by the cyclic reduction algorithm satisfy useful relations that, together with suitable convergence properties, enable us to compute the solution G of the matrix equation (2). In fact, we have the following result [3]:

Theorem 1. *The blocks $\{A_i^{(j)}\}_{i\geq 0}$, $\{\widehat{A}_i^{(j)}\}_{i\geq 1}$ are nonnegative matrices such that $\sum_{i=0}^{+\infty} A_i^{(j)}$, $A_0 + \sum_{i=1}^{+\infty} \widehat{A}_i^{(j)}$ are column stochastic. Moreover, if the matrix*

$$I - \sum_{i=0}^{+\infty} G^{i\cdot 2^j} \widehat{A}_{i+1}^{(j)}$$

is nonsingular, then $G = A_0 \left(I - \sum_{i=0}^{+\infty} G^{i\cdot 2^j} \widehat{A}_{i+1}^{(j)}\right)^{-1}$.

Under mild conditions, usually satisfied in the applications, the block entries $A_i^{(j)}$ and $\widehat{A}_i^{(j)}$ quadratically converge to zero, for $i \geq 2$, for $j \to \infty$, as stated by the following theorem [3]:

Theorem 2. *Let $G' = \lim_{j\to\infty} G^j$. Then the following convergence properties hold:*

1. $\lim_{j\to\infty} \widehat{A}_i^{(j)} = 0$*, for $i \geq 2$;*
2. *if the entries of the matrix $(I - \sum_{i=1}^{+\infty} A_i^{(j)})^{-1}$ are bounded above by a constant, then the sequence of matrices $R^{(j)} = A_0^{(j)} \left(I - \sum_{i=1}^{+\infty} A_i^{(j)}\right)^{-1}$ converges quadratically to the matrix G';*
3. *if the solution G of (2) is irreducible then $\lim_{j\to\infty} A_0^{(j)}(I - A_1^{(j)})^{-1} = G'$ and $\lim_{j\to\infty} A_i^{(j)} = 0$, $i \geq 2$.*

The above theorems enable us to compute G. In fact, since the sequences $\left\{\widehat{A}_i^{(j)}\right\}_j$, $i \geq 2$, converge to null matrices, then

$$G = \lim_j A_0 (I - \widehat{A}_1^{(j)})^{-1}, \tag{11}$$

so that the matrix $A_0(I-\widehat{A}_1^{(j)})^{-1}$, for j large enough, provides an approximation of G. Otherwise part 2 or part 3 of Theorem 2 provide computable sequences converging to G'. The j-th term of each sequence can be used for numerically approximating G' (if j is large enough), and thus to replace G^{2^j} in Theorem 1, in order to obtain the approximation of G.

3 COMPUTATIONAL TOOLS AND STRUCTURES

We now introduce some structural properties and computational tools that are fundamental to prove the main results of this paper and for devising fast and effective algorithms for the solution of equation (3) for non-skip-free M/G/1 type Markov chains.

3.1 Power series arithmetic

Let us first recall some computational properties based on the Fast Fourier Transform (FFT).

Given a matrix power series $A(z) = \sum_{i=0}^{+\infty} A_i z^i$ which is convergent for $|z| = 1$ and a real positive ϵ, define the *ϵ-degree* of $A(z)$ the minimum integer q such that $\boldsymbol{e}^T \sum_{i=q+1}^{+\infty} |A_i| \leq \epsilon \boldsymbol{e}^T$, for $\boldsymbol{e}^T = (1, \ldots, 1)$. The concept of ϵ-degree is particularly meaningful when working with a floating point representation of real numbers where the entries of A_i are numerically negligible for $i > q$, if ϵ represents the unit roundoff of the floating point arithmetic. In this case we prefer to use the term *numerical degree* instead of ϵ-degree.

Throughout this section we assume that all the series are convergent in the closed unit disk.

We call by the name *p-dimensional $k \times k$ block column vector* any $kp \times p$ matrix. Similarly the transposed of a block column vector is called *block row vector.* The same notation is used for power series, i.e., we call block column (respectively row) vector power series any $kp \times p$ (respectively $p \times kp$) matrix power series.

The following results enable us to perform operations among block Toeplitz matrices by means of FFT with a low computational cost. Here we recall that an FFT of length n can be computed with $O(n \log n)$ arithmetic operations whenever n is an integer power of 2. For more details we refer the reader to [5], [3], [7].

Theorem 3. *The product of a $k \times k$ block Toeplitz matrix having blocks of size p and a block column vector can be performed by means of $O(p^2)$ FFT's of length $O(k)$ in $O(p^3 k + p^2 k \log k)$ arithmetic operations.*

A similar complexity bound can be obtained for computations with matrix power series.

Theorem 4. *The first n coefficients of the product $C(z)$ of an $m \times m$ matrix power series $A(z)$ and an $m \times p$ matrix power series $B(z)$, can be approximated within the error bound ϵ by means of $2mp+m^2$ FFT's of length n in $O(m^2pn+(mp+m^2)n\log n)$ arithmetic operations, where n is the ϵ-degree of $C(z)$.*

Proof. The product is computed by means of the evaluation/interpolation technique at the roots of 1. More precisely, first observe that for a given complex $\xi \in \boldsymbol{C}$ the values that $C(z)$ takes on at $z = \xi$ can be computed by means of the matrix product $C(\xi) = A(\xi)B(\xi)$ once $A(\xi)$ and $B(\xi)$ have been computed. In order to recover the first n coefficients of $C(z)$ within the error ϵ, it is sufficient to interpolate $C(z)$ at the n-th roots of 1 (compare proposition 4 of [5]). The evaluation of $C(z)$ is performed by first evaluating $A(z)$ and $B(z)$ at the roots of 1 by means of $m^2 + mp$ FFT's of length n for the arithmetic cost $O((m^2 + mp)n \log n)$, and then computing n matrix products of size $m \times m$ and $m \times p$ for the cost $O(nm^2p)$ arithmetic operations. The interpolation of $C(z)$ at the roots of 1 is performed by means of mp FFT's of length n for the cost $O(mpn \log n)$. $\square$

A similar result can be proved for the inversion of matrix power series.

Theorem 5. *Let $A(z)$ be an $m \times m$ matrix power series such that $\det A(0) \neq 0$, $A(z)^{-1}$ exists and has ϵ-degree n. Then the first n coefficients of the matrix power series $A(z)^{-1}$ can be approximated within the error ϵ in $O(m^3n+m^2n\log n)$ operations by means of the evaluation/interpolation technique at the roots of 1.*

Observe that, by combining Theorems 4 and 3 we may easily arrive at the following

Theorem 6. *Let $A(z)$ be a $kp \times kp$ matrix power series having the $k \times k$ block Toeplitz structure with $p \times p$ blocks. Let $B(z)$ be a $kp \times p$ matrix power series. Then the first n matrix coefficients of the product $C(z) = A(z)B(z)$ can be approximated within the error ϵ in $O(np^3k + p^2kn \log kn)$ arithmetic operations, where n is the ϵ-degree of $C(z)$.*

3.2 Toeplitz matrices and displacement rank

The concept of displacement rank, introduced by Kailath et al. in [13], is fundamental in devising and analyzing algorithms related to Toeplitz matrices [12], [7].

Define the $k \times k$ *block down-shift* matrix

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

where the blocks have dimension p, and consider the following *block displacement operator*

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

defined for any $k \times k$ block matrix H. For simplicity the matrix $\Delta(H)$ is called the displacement of H.

Let us also introduce the following notations: the block vectors $\boldsymbol{e}_1 = [I, 0, \dots, 0]^T$, $\boldsymbol{e}_2 = [0, \dots, 0, I]^T$ denote $kp \times p$ matrices made up by the first and the last p columns of the identity matrix of order kp.

We say that the block matrix H has *block displacement rank* r if r is the minimum integer such that there exist block column vectors $\boldsymbol{u}^{(i)}$, and block row vectors $\boldsymbol{v}^{(i)}$, $i = 1, \dots, r$, satisfying the equation $\Delta(H) = \sum_{i=1}^{r} \boldsymbol{u}^{(i)} \boldsymbol{v}^{(i)}$.

It can be easily observed that, if H is any $k \times k$ block Toeplitz matrix, then H has block displacement rank at most 2. A sequence $\{H_k\}_k$ of $k \times k$ block matrices, having $p \times p$ blocks, is called a sequence of *block Toeplitz-like matrices* if the block displacement rank of H_k is independent of the block dimension k.

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

$$\Delta(H^{-1}) = -H^{-1}\Delta(H)H^{-1}. \tag{12}$$

From this property it follows in particular that the inverse of a block Toeplitz, or Toeplitz-like, matrix is a block Toeplitz-like matrix.

The following result can be easily proved by extending to block matrices the same proof given in [7].

Theorem 7. *Let K be a $k \times k$ block matrix such that $\Delta(K) = \sum_{i=1}^{r} \boldsymbol{u}^{(i)} \boldsymbol{v}^{(i)}$, where $\boldsymbol{u}^{(i)}$ and $\boldsymbol{v}^{(i)}$ are k-dimensional block column and block row vectors, respectively. Then we have*

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

where $L(\boldsymbol{w})$ denotes the $k \times k$ block lower triangular block Toeplitz matrix defined by its first block column $\boldsymbol{w}$, and $L^T(\boldsymbol{w})$ is its transposed matrix.

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

If the matrix K is nonsingular then the above representation theorem can be applied to K^{-1} in the light of (12) yielding the following formula

$$K^{-1} = L(K^{-1}\boldsymbol{e}_1) + \sum_{i=1}^{r} L(K^{-1}\boldsymbol{u}^{(i)}) L^T(ZK^{-T}\boldsymbol{v}^{(i)T}). \tag{13}$$

From the above results it follows that a block Toeplitz-like matrix, and its inverse, can be expressed by means of a sum of few products between block lower triangular and block upper triangular block Toeplitz matrices. This fact allows us to perform block Toeplitz-like matrix computations in terms of a few multiplications between matrix polynomials, in the light of the results of Section 3.1, with a substantial reduction of the computational cost.

The known algorithms for the inversion of block Toeplitz matrices can be extended to the inversion of block Toeplitz-like matrices. In particular the "fast" and the "superfast" algorithms for Toeplitz inversion [7] turn into algorithms for block Toeplitz-like matrix inversion having a cost $O(p^3k^2)$ and $O(p^3k\log^2 k)$ operations, respectively. The conjugate gradient method can also be applied to solve the few systems that provide the block vectors defining the inverse matrix. This algorithm has a cost of $O(p^3k\log k)$ operations per step, and requires at most k iterations. In many cases, with a suitable preconditioning technique, the number of iterations can be reduced to $O(1)$.

The above remarks and the results of Section 3.1 allow us to state the following theorem that constitutes a useful tool for implementing the cyclic reduction in the point-wise version [5]:

Theorem 8. *Let $F(X_1,\ldots,X_r)$ be a rational function in the variables $X_1,\ldots,X_r$. Assume that the $k\times k$ block matrix power series $X_i(z)$, $i=1,\ldots,r$, with blocks of size p, are such that $F(X_1,\ldots,X_r)$ exists and has ϵ-degree n, $X_i(z)$ as well as $F(X_1,\ldots,X_r)$ are block Toeplitz-like. Then the first n coefficients of $F(X_1,\ldots,X_r)$ can be approximated in $O(np^3k\log^2 k+p^2kn\log(kn))$ operations, within the error ϵ, where we assume r constant with respect to p and k.*

4 STRUCTURAL PROPERTIES OF CYCLIC REDUCTION

The main goal of this section is to prove that the matrix power series $H^{(j)}(z)$, $\widehat{H}^{(j)}(z)$ generated by the cyclic reduction process applied to a non-skip-free M/G/1 type Markov chain have block displacement rank at most three and to provide explicit formulae relating the displacement of $H^{(j)}(z)$ to the displacement of $H^{(j+1)}(z)$.

These properties allow us to devise an algorithm, based on FFT, that performs a single cyclic reduction step in $O(np^3k\log^2 k+p^2kn\log(kn))$ arithmetic operations. This result generalizes the algorithm devised in [4] for block banded block Toeplitz matrices.

Let $\boldsymbol{r}^{(j)}(z)=\boldsymbol{e}_2^TH^{(j)}(z)$, $\boldsymbol{c}^{(j)}(z)=H^{(j)}(z)\boldsymbol{e}_1$, be the last block row and the first block column of $H^{(j)}(z)$, respectively. Similarly, denote $\widehat{\boldsymbol{r}}^{(j)}(z)=\boldsymbol{e}_2^T\widehat{H}^{(j)}(z)$, $\widehat{\boldsymbol{c}}^{(j)}(z)=\widehat{H}^{(j)}(z)\boldsymbol{e}_1$.

We prove the following result.

Theorem 9. *For the matrix power series $H^{(j)}(z) = zI - A^{(j)}(z)$, $\widehat{H}^{(j)}(z) = I - \widehat{A}^{(j)}(z)$ generated after j steps of cyclic reduction we have*

$$\begin{aligned}\Delta(H^{(j)}(z)) &= \boldsymbol{c}^{(j)}(z)\boldsymbol{u}^{(j)}(z) - \boldsymbol{v}^{(j)}(z)\boldsymbol{r}^{(j)}(z)\\ \Delta(\widehat{H}^{(j)}(z)) &= \boldsymbol{e}_1\boldsymbol{r}^{(0)}(0) - \boldsymbol{c}^{(j)}(z)\widehat{\boldsymbol{u}}^{(j)}(z) - \boldsymbol{v}^{(j)}(z)\widehat{\boldsymbol{r}}^{(j)}(z)\end{aligned} \tag{14}$$

where $\boldsymbol{u}^{(j)}(z) = \sum_{i=0}^{+\infty} \boldsymbol{u}_i^{(j)} z^i$, $\widehat{\boldsymbol{u}}^{(j)}(z) = \sum_{i=0}^{+\infty} \widehat{\boldsymbol{u}}_i^{(j)} z^i$ are block row vector power series, $\boldsymbol{v}^{(j)}(z) = \sum_{i=0}^{+\infty} \boldsymbol{v}_i^{(j)} z^i$ is a block column vector power series, defined by

$$\begin{aligned}\boldsymbol{u}^{(j+1)}(z) &= -\left[\boldsymbol{u}^{(j)}(z)\left(H_{odd}^{(j)}(z^2)\right)^{-1} H^{(j)}(-z)\right]_{even}\\ \boldsymbol{v}^{(j+1)}(z) &= -\left[H^{(j)}(-z)\left(H_{odd}^{(j)}(z^2)\right)^{-1} \boldsymbol{v}^{(j)}(z)\right]_{even}\\ \widehat{\boldsymbol{u}}^{(j+1)}(z) &= \widehat{\boldsymbol{u}}_{odd}^{(j)}(z) + \boldsymbol{u}_{odd}^{(j)}(z)\left(H_{odd}^{(j)}(z)\right)^{-1} \widehat{H}_{even}^{(j)}(z)\end{aligned} \tag{15}$$

for $j \geq 0$, and $\boldsymbol{u}^{(0)}(z) = z\boldsymbol{e}_2^T$, $\widehat{\boldsymbol{u}}^{(0)}(z) = \boldsymbol{e}_2^T$, $\boldsymbol{v}^{(0)}(z) = z\boldsymbol{e}_1$.

Proof. We prove relation (14) by induction on j. For $j = 0$ a simple calculation shows that

$$\begin{aligned}\Delta(H^{(0)}(z)) = z(\boldsymbol{c}^{(0)}(z)\boldsymbol{e}_2^T - \boldsymbol{e}_1\boldsymbol{r}^{(0)}(z)) =\\ \boldsymbol{c}^{(0)}(z)\boldsymbol{u}^{(0)}(z) - \boldsymbol{v}^{(0)}(z)\boldsymbol{r}^{(0)}(z).\end{aligned}$$

Suppose that relation (14) holds for a fixed $j \geq 0$; we show that it holds for $j+1$. For simplicity let us denote $B = H^{(j)}(z)^{-1} - H^{(j)}(-z)^{-1}$. For the properties of the displacement operator (12) and for (10) we have:

$$\begin{aligned}&\Delta(H^{(j+1)}(z^2)) = 2z\Delta(B^{-1}) = -2zB^{-1}\Delta(B)B^{-1} =\\ &-2zB^{-1}(-H^{(j)}(z)^{-1}\Delta(H^{(j)}(z))H^{(j)}(z)^{-1}+\\ &H^{(j)}(-z)^{-1}\Delta(H^{(j)}(-z))H^{(j)}(-z)^{-1})B^{-1};\end{aligned}$$

by substituting relation (14) in the above equation we obtain

$$\begin{aligned}&\Delta(H^{(j+1)}(z^2)) = 2zB^{-1}\boldsymbol{e}_1\boldsymbol{u}^{(j)}(z)H^{(j)}(z)^{-1}B^{-1} - 2zB^{-1}H^{(j)}(z)^{-1}\boldsymbol{v}^{(j)}(z)\boldsymbol{e}_2^TB^{-1} -\\ &2zB^{-1}\boldsymbol{e}_1\boldsymbol{u}^{(j)}(-z)H^{(j)}(-z)^{-1}B^{-1} + 2zB^{-1}H^{(j)}(-z)^{-1}\boldsymbol{v}^{(j)}(-z)\boldsymbol{e}_2^TB^{-1} =\\ &\boldsymbol{c}^{(j+1)}(z^2)\boldsymbol{u}^{(j+1)}(z^2) - \boldsymbol{v}^{(j+1)}(z^2)\boldsymbol{r}^{(j+1)}(z^2),\end{aligned}$$

where $\boldsymbol{u}^{(j+1)}(z)$ and $\boldsymbol{v}^{(j+1)}(z)$ satisfy (15). Similarly, we proceed for proving the second equation in (14) and the third in (15). □

From the above theorem, the following representations hold for the matrix power series $H^{(j)}(z)$ and $\widehat{H}^{(j)}(z)$:

Theorem 10. *At each step j of cyclic reduction it holds that*

$$\begin{aligned} H^{(j)}(z) &= L(\boldsymbol{c}^{(j)}(z))(I - L^T(Z\boldsymbol{u}^{(j)}(z)^T)) + L(\boldsymbol{v}^{(j)}(z))L^T(Z\boldsymbol{r}^{(j)}(z)^T) \\ \widehat{H}^{(j)}(z) &= L(\widehat{\boldsymbol{c}}^{(j)}(z)) + L^T(Z\boldsymbol{r}^{(0)}(0)^T) + \\ & \quad L(\boldsymbol{c}^{(j)}(z))L^T(Z\widehat{\boldsymbol{u}}^{(j)}(z)^T) + L(\boldsymbol{v}^{(j)}(z))L^T(Z\widehat{\boldsymbol{r}}^{(j)}(z)^T). \end{aligned}$$

Proof. The proof simply follows from Theorems 7 and 9. □

From the above results it follows that, at each step j, only the seven block vector power series $\boldsymbol{u}^{(j)}(z)$, $\widehat{\boldsymbol{u}}^{(j)}(z)$, $\boldsymbol{v}^{(j)}(z)$, $\boldsymbol{c}^{(j)}(z)$, $\boldsymbol{r}^{(j)}(z)$, $\widehat{\boldsymbol{c}}^{(j)}(z)$, $\widehat{\boldsymbol{r}}^{(j)}(z)$, need to be computed in order to represent the matrix power series $H^{(j)}(z)$ and $\widehat{H}^{(j)}(z)$.

Explicit relations between the block vector power series $\boldsymbol{c}^{(j)}(z) = \sum_{i=0}^{+\infty} \boldsymbol{c}_i^{(j)} z^i$, $\boldsymbol{r}^{(j)}(z) = \sum_{i=0}^{+\infty} \boldsymbol{r}_i^{(j)} z^i$, $\widehat{\boldsymbol{c}}^{(j)}(z) = \sum_{i=0}^{+\infty} \widehat{\boldsymbol{c}}_i^{(j)} z^i$, $\widehat{\boldsymbol{r}}^{(j)}(z) = \sum_{i=0}^{+\infty} \widehat{\boldsymbol{r}}_i^{(j)} z^i$ at two subsequent steps of cyclic reduction can be obtained by multiplying the functional relations (9) on the left by $\boldsymbol{e}_2^T$, and on the right by $\boldsymbol{e}_1$.

Moreover, the block vectors defining $H^{(j)}(z)$ and $\widehat{H}^{(j)}(z)$ are related to the block vectors obtained at step $j+1$, which define $H^{(j+1)}(z)$ and $\widehat{H}^{(j+1)}(z)$, by means of functional relations that only involve sums of products between block triangular block Toeplitz matrices. Hence, in the light of the results of Section 3 and of Theorem 8, the computational cost needed to perform the j-th step of cyclic reduction by using FFT reduces to $O(n_j p^3 k \log^2 k + p^2 k n_j \log k n_j)$, where n_j is an upper bound to the numerical degree of the matrix power series $H^{(j)}(z)$ and $\widehat{H}^{(j)}(z)$.

The matrix G has interesting structural properties; in particular it can be factorized into the product of a block lower triangular, block Toeplitz matrix and a block upper triangular, block Toeplitz matrix. In fact, in [9] it is proved that G is the k-th power of a $k \times k$ block Frobenius matrix. We recall that that the k-th power of a scalar $k \times k$ Frobenius matrix can be factored as the product of a lower triangular Toeplitz matrix and an upper triangular Toeplitz matrix [7], and deduce that for $p = 1$ the matrix $G = \widehat{F}^k$ can be factored in the same way. This property can be proved even for $p > 1$ in terms of block triangular block Toeplitz matrices. This decomposition of the matrix G can also be obtained by using the concept of displacement rank, extended to infinite matrices, and applied to the system (3). In fact, from (3) we deduce that $G = A_0 E$, where E is the $m \times m$ leading principal submatrix of $S^{(0)^{-1}}$. Moreover, A_0 is a block lower triangular block Toeplitz matrix and, due to the properties of the displacement operator, the matrix $S^{(0)^{-1}}$ can be factored into the product of a block lower triangular block Toeplitz matrix and a block upper triangular block Toeplitz matrix. Thus, E and consequently G, admit the block LU factorization with L and U block Toeplitz matrices. The existence of this factorization can be also proved by means of the relation (11); in fact, from the general converge properties presented in Section 2 and from the displacement structure of $\widehat{H}_1^{(j)} = I - \widehat{A}_1^{(j)}$, it follows that

$\widehat{H}_1^{(j)^{-1}}$ converges to a matrix that can be factored into the product of a block lower triangular block Toeplitz matrix and a block upper triangular block Toeplitz matrix.

In the case where $H_i = 0$, for $i > 2$, the results concerning the structural properties of the blocks generated by cyclic reduction are further simplified. We can give explicit expressions of the blocks $H_i^{(j)}$, $i = 0, 1, 2$, $\widehat{H}_1^{(j)}$. Moreover, since $H^{(j)}(z)$ is a matrix polynomial of degree 2 having displacement rank 2, then it follows that its matrix coefficients have displacement rank at most 3. In fact we have the following theorems:

Theorem 11. *If $A_i = 0$ for $i > 2$, then for the matrices $H_i^{(j)}$, $i = 0, 1, 2$, $\widehat{H}_1^{(j)}$, $j \geq 0$, generated by the cyclic reduction (9) we have*

$$\begin{aligned}
\Delta(H_0^{(j)}) &= \boldsymbol{c}_0^{(j)}\boldsymbol{u}_0^{(j)} - \boldsymbol{v}_0^{(j)}\boldsymbol{r}_0^{(j)}\\
\Delta(H_1^{(j)}) &= \boldsymbol{c}_0^{(j)}\boldsymbol{u}_1^{(j)} + \boldsymbol{c}_1^{(j)}\boldsymbol{u}_0^{(j)} - \boldsymbol{v}_0^{(j)}\boldsymbol{r}_1^{(j)} - \boldsymbol{v}_1^{(j)}\boldsymbol{r}_0^{(j)}\\
\Delta(H_2^{(j)}) &= \boldsymbol{s}^{(j)}\boldsymbol{u}_1^{(j)} - \boldsymbol{v}_1^{(j)}\widehat{\boldsymbol{r}}_1^{(j)}\\
\Delta(\widehat{H}_1^{(j)}) &= -\boldsymbol{e}_1\boldsymbol{r}_0^{(0)} - \boldsymbol{c}_0^{(j)}\boldsymbol{u}_1^{(j)} - \boldsymbol{v}_0^{(j)}\widehat{\boldsymbol{r}}_1^{(j)},
\end{aligned}$$

where $\boldsymbol{s}^{(0)} = \boldsymbol{c}_1^{(0)}$, $\boldsymbol{s}^{(j+1)} = \boldsymbol{s}^{(j)} - H_2^{(j)}H_1^{(j)^{-1}}\boldsymbol{c}_0^{(j)}$, for $j \geq 0$.

Proof. The proof follows by comparing the terms of the same degree in formula (14). □

Theorem 12. *At each step j of cyclic reduction, the matrices $H_0^{(j)}$, $H_1^{(j)}$, $H_2^{(j)}$, $H_1^{(j)^{-1}}$, $\widehat{H}_1^{(j)}$ can be rewritten as*

$$\begin{aligned}
H_0^{(j)} =\ & L(\boldsymbol{c}_0^{(j)})(I - L^T(Z\boldsymbol{u}_0^{(j)T})) + L(\boldsymbol{v}_0^{(j)})L^T(Z\boldsymbol{r}_0^{(j)T})\\
H_1^{(j)} =\ & L(\boldsymbol{c}_1^{(j)})(I - L^T(Z\boldsymbol{u}_0^{(j)T})) - L(\boldsymbol{c}_0^{(j)})L^T(Z\boldsymbol{u}_1^{(j)T}) + L(\boldsymbol{v}_0^{(j)})L^T(Z\boldsymbol{r}_1^{(j)T}) +\\
& L(\boldsymbol{v}_1^{(j)})L^T(Z\boldsymbol{r}_0^{(j)T})\\
H_2^{(j)} =\ & L(\boldsymbol{c}_2^{(j)}) - L(\boldsymbol{s}^{(j)})L^T(Z\boldsymbol{u}_1^{(j)T}) + L(\boldsymbol{v}_1^{(j)})L^T(Z\widehat{\boldsymbol{r}}_1^{(j)T})\\
H_1^{(j)^{-1}} =\ & L(H_1^{(j)^{-1}}\boldsymbol{e}_1) + L^T(ZH_1^{(j)^{-T}}\boldsymbol{u}_0^{(j)T}) + L(H_1^{(j)^{-1}}\boldsymbol{c}_0^{(j)})L^T(ZH_1^{(j)^{-T}}\boldsymbol{u}_1^{(j)T}) -\\
& L(H_1^{(j)^{-1}}\boldsymbol{v}_1^{(j)})L^T(ZH_1^{(j)^{-T}}\boldsymbol{r}_0^{(j)T})\\
\widehat{H}_1^{(j)} =\ & L(\widehat{\boldsymbol{c}}_1^{(j)}) + L^T(Z\boldsymbol{r}_0^{(0)T}) - L(\boldsymbol{c}_0^{(j)})L^T(Z\boldsymbol{u}_1^{(j)T}) + L(\boldsymbol{v}_0^{(j)})L^T(Z\widehat{\boldsymbol{r}}_1^{(j)T}).
\end{aligned}$$

Proof. The proof follows from Theorem 11 and by the results of Section 3.2. □

Explicit relations for the vectors $\boldsymbol{u}_i^{(j)}$, $\boldsymbol{v}_i^{(j)}$ $\boldsymbol{c}_i^{(j)}$, $\boldsymbol{r}_i^{(j)}$, $\widehat{\boldsymbol{c}}_i^{(j)}$, $\widehat{\boldsymbol{r}}_i^{(j)}$, can be provided:

$$
\begin{aligned}
\boldsymbol{u}_0^{(j)} &= \boldsymbol{u}_0^{(j-1)} - \boldsymbol{u}_1^{(j-1)} H_1^{(j-1)^{-1}} H_0^{(j-1)}\\
\boldsymbol{u}_1^{(j)} &= -\boldsymbol{u}_1^{(j-1)} H_1^{(j-1)^{-1}} H_2^{(j-1)}\\
\boldsymbol{v}_0^{(j)} &= \boldsymbol{v}_0^{(j-1)} - H_0^{(j-1)} H_1^{(j-1)^{-1}} \boldsymbol{v}_1^{(j-1)}\\
\boldsymbol{v}_1^{(j)} &= -H_2^{(j-1)} H_1^{(j-1)^{-1}} \boldsymbol{v}_1^{(j-1)}\\
\boldsymbol{c}_0^{(j)} &= -H_0^{(j-1)} H_1^{(j-1)^{-1}} \boldsymbol{c}_0^{(j-1)}\\
\boldsymbol{c}_1^{(j)} &= \boldsymbol{c}_1^{(j-1)} - H_0^{(j-1)} H_1^{(j-1)^{-1}} \boldsymbol{c}_2^{(j-1)} - H_2^{(j-1)} H_1^{(j-1)^{-1}} \boldsymbol{c}_0^{(j-1)}\\
\boldsymbol{c}_2^{(j)} &= -H_2^{(j-1)} H_1^{(j-1)^{-1}} \boldsymbol{c}_2^{(j-1)}\\
\boldsymbol{r}_0^{(j)} &= -\boldsymbol{r}_0^{(j-1)} H_1^{(j-1)^{-1}} H_0^{(j-1)}\\
\boldsymbol{r}_1^{(j)} &= \boldsymbol{r}_1^{(j-1)} - \boldsymbol{r}_0^{(j-1)} H_1^{(j-1)^{-1}} H_2^{(j-1)} - \boldsymbol{r}_2^{(j-1)} H_1^{(j-1)^{-1}} H_0^{(j-1)}\\
\boldsymbol{r}_2^{(j)} &= -\boldsymbol{r}_2^{(j-1)} H_1^{(j-1)^{-1}} H_2^{(j-1)}\\
\widehat{\boldsymbol{r}}_1^{(j)} &= \widehat{\boldsymbol{r}}_1^{(j-1)} - H_0^{(j-1)} H_1^{(j-1)^{-1}} \boldsymbol{r}_2^{(j-1)}\\
\widehat{\boldsymbol{c}}_1^{(j)} &= \widehat{\boldsymbol{c}}_1^{(j-1)} - \boldsymbol{c}_0^{(j-1)} H_1^{(j-1)^{-1}} H_2^{(j-1)}
\end{aligned}
$$

for $j \geq 1$, where $\boldsymbol{u}_0^{(0)} = \boldsymbol{v}_0^{(0)} = \boldsymbol{0}$, $\boldsymbol{u}_1^{(0)} = \boldsymbol{e}_2^T$, $\boldsymbol{v}_0^{(0)} = \boldsymbol{e}_1$. For this particular case, the computational cost per step is reduced to $O(p^3 k \log^2 k)$ operations.

For the blocks $A_i^{(j)}$, $\widehat{A}_i^{(j)}$ the quadratically convergence properties of Section 2 still hold. Hence, due to the displacement structure of the blocks A_i of (5), we may devise a quadratically convergent algorithm, with low computational cost, for non-skip-free M/G/1 type Markov chains, just by applying the cyclic reduction technique where the computations of the blocks $H_i^{(j)}$, $\widehat{H}_i^{(j)}$ is performed by exploiting their Toeplitz-like structure according to the formulae displayed in this section. This new algorithm particularly shows its effectiveness, with respect to the linearly convergent method of [9], when the block dimension k is large; in fact, the computational cost of one step of the latter method is $O(p^3 k)$, and many iterations may need to be computed in order to reach a good approximation of G.

Numerical experiments which show the efficiency of our approach are presented in Section 6.

In the case of scalar entries α_i, the formulae displayed in Theorem 10 can be substantially simplified. In fact, from the equation (10), we deduce the following:

Proposition 13. *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 14. *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*

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

5 CONVERGENCE PROPERTIES IN THE QBD CASE

Consider the power series $\gamma(z)$ and $\Gamma(z)$ defined as $\gamma(z) = \det(z^k I - \alpha(z))$, $\alpha(z) = \sum_{i=0}^{+\infty} z^i \alpha_i$, $\Gamma(z) = \det(zI - A(z))$.

In this section we prove some convergence properties of the cyclic reduction for non-skip-free Markov chains of QBD processes that improve the known properties valid in the general case [3] and express in a different form the convergence results of [17]. More specifically we prove that the lower diagonal blocks $A_2^{(j)}$, generated by cyclic reduction, tend to zero as $O(\theta^{-2^j})$, where θ is the minimum modulus of the zeros of the function $\gamma(z)$ lying out of the unit disk.

Let us first recall the following result of [9] that relates the zeros of $\gamma(z)$ and $\Gamma(z)$.

Theorem 15. *For the power series $\Gamma(z)$ it holds*

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

A consequence of the above theorem is that, if ξ is zero of the power series $\gamma(z)$ then ξ^k is zero of $\Gamma(z)$. Conversely, if η is zero of $\Gamma(z)$ then there exists a k-th root of η which is zero of $\gamma(z)$.

Throughout this section we assume, without loss of generality, that $\alpha(z)$ is a matrix polynomial such that $\gamma(z)$ is a polynomial of degree $2m$. In fact, in the QBD case, $\gamma(z)$ is a polynomial of degree at most $2m$; moreover, if $\gamma(z)$ has degree less than $2m$ we may replace $\alpha(z)$ with $\alpha_\epsilon(z) = \alpha(z) + z^{2k}\epsilon I - \epsilon\beta(z)$, where $\beta(z)$ is chosen in such a way that $\alpha_\epsilon(z)$ has nonnegative coefficients and $\alpha_\epsilon(1)$ is column stochastic. The power series $\beta(z)$ exists since $\alpha(1)$ has not null columns. In this way, $\gamma_\epsilon(z) = \det(z^k I - \alpha_\epsilon(z))$ has degree $2m$, and the corresponding Markov chain is positive recurrent, for any $\epsilon \neq 0$ in a suitable neighborhood of 0 so that we may apply a homotopy argument based on the continuity of the eigenvalues/eigenvectors of a matrix polynomial with respect to its coefficients.

Let us denote $\Gamma^{(j)}(z) = \det(zI - A^{(j)}(z))$. By taking determinants in both sides of (10) we have

$$\Gamma^{(j+1)}(z^2) = -\Gamma^{(j)}(z)\Gamma^{(j)}(-z)/\det H_0^{(j)}.$$

An immediate consequence of the above relation is that the $2m$ zeros of the polynomial $\Gamma^{(j)}(z)$ are explicitly given by

$$\xi_i^{m2^j}, \ i = 1, \ldots, 2m,$$

where

$$\gamma(\xi_i) = 0, \quad i = 1, \ldots, 2m.$$

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

First observe that, for any ξ_i there exists a vector $\boldsymbol{x}_i$ such that $\boldsymbol{x}_i^T(\xi_i^k I - \alpha(\xi_i)) = 0$. Multiplying on the left the matrix $S^{(0)}$ of (3) by the vector $\boldsymbol{\xi}_i^T = [\boldsymbol{x}_i^T, \xi_i \boldsymbol{x}_i^T, \xi_i^2 \boldsymbol{x}_i^T, \ldots]$ yields $\boldsymbol{\xi}_i^T S^{(0)} = [\boldsymbol{w}_i^T, 0, 0, \ldots]$, where $\boldsymbol{w}_i$ is a suitable vector of m components.

As it is well known [20], for the positive recurrence of the Markov chain, the polynomial $\Gamma(z)$ has m zeros in the closed unit disk, and consequently, m zeros of modulus greater than 1. Let us arrange the zeros ξ_i in such a way that $|\xi_1| \geq |\xi_2| \geq \ldots > |\xi_{m+1}| = 1 \geq \ldots \geq |\xi_{2m}|$.

We consider the $m \times p$ matrix U whose i-th row is $\boldsymbol{x}_i^T$, $i = 1, \ldots, m$, and the $m \times m$ diagonal matrix $D = \mathrm{Diag}(\xi_1, \ldots, \xi_m)$. Moreover, we define the $k \times k$ block matrix

$$V = [U, DU, D^2U, \ldots, D^{k-1}U]. \tag{16}$$

Observe that if $p = 1$, then V is a Vandermonde matrix.

It holds

$$\left[V, D^kV, D^{2k}V, \ldots\right] S^{(0)} = [W, 0, 0, \ldots], \tag{17}$$

where $W = -D^{-k}VH_0$.

The homogeneous part of the above relation can be rewritten as

$$VH_0 + D^kVH_1 + D^{2k}VH_2 = 0.$$

By applying the cyclic reduction to the system (17) we deduce that

$$VH_0^{(j)} + D^{k2^j}VH_1^{(j)} + D^{k2^{j+1}}VH_2^{(j)} = 0, \quad j = 0, 1, \ldots$$

Let us consider the case $p = 1$. If the zeros ξ_i, $i = 1, \ldots, m$ are pair-wise distinct, then the matrix V is nonsingular. Otherwise, if the polynomial $\gamma(z)$ has multiple zeros out of the unit circle, the above equation holds by replacing V with the generalized Vandermonde matrix (compare [14], pages 69–70), that is invertible and satisfies relation (17). Multiplying the above relation on the left by V^{-1} we find

$$H_0^{(j)} + F^{k2^j}H_1^{(j)} + F^{k2^{j+1}}H_2^{(j)} = 0, \quad j = 0, 1, \ldots$$

where $F = V^{-1}DV$ is the Frobenius matrix associated with the polynomial having zeros ξ_i, $i = 1, \ldots, m$ (compare [23]). Whence we find that

$$H_2^{(j)} + F^{-k2^j} H_1^{(j)} + F^{-k2^{j+1}} H_0^{(j)} = 0, \quad j = 0, 1, \ldots \tag{18}$$

Analogously, it can be easily observed that

$$V\widehat{H}_1^{(j)} + D^{k2^j} V\widehat{H}_2^{(j)} = W, \quad j = 0, 1, \ldots,$$

whence, premultiplying by V^{-1}, we find that

$$\widehat{H}_2^{(j)} + F^{-k2^j} \left(\widehat{H}_1^{(j)} - V^{-1}W\right) = 0, \quad j = 0, 1, \ldots \tag{19}$$

In the case $p > 1$ the equation (18) still holds provided that V is nonsingular. Hereafter we assume this condition. It is easy to prove that in this case the matrix $V^{-1}DV$ is a block Frobenius matrix (compare [14], pages 280-281).

The following result can be obtained from the above relations.

Theorem 16. *Let ξ_i, $i = 1, \ldots, 2m$, be the zeros of the polynomial $\gamma(z)$, ordered such that $|\xi_1| \geq |\xi_2| \geq \ldots > |\xi_{m+1}| = 1 \geq |\xi_{m+2}| \geq \ldots \geq |\xi_{2m}|$, and assume that the matrix V in (16) is nonsingular. Then we have $\lim_j A_2^{(j)} = \lim_j \widehat{A}_2^{(j)} = 0$. Moreover, if ξ_m is simple, then for any operator norm $||\cdot||$, both $||A_2^{(j)}||$ and $||\widehat{A}_2^{(j)}||$ tend to zero as $O(|\xi_m|^{-k2^j})$. Otherwise, for any $\epsilon > 0$, there exists an operator norm $||\cdot||$ such that $||A_2^{(j)}||$ and $||\widehat{A}_2^{(j)}||$ tend to zero as $O((|\xi_m|^{-1} + \epsilon)^{k2^j})$.*

Proof. For any $\epsilon > 0$ let $||\cdot||$ be an operator norm such that $||F^{-1}|| < \rho(F^{-1}) + \epsilon$ (compare [10]), where $\rho(F^{-1})$ denotes the spectral radius of F^{-1}. Observe that for the (block) Frobenius matrix F it holds $\rho(F^{-1}) = |\xi_m|^{-1}$ ([14], pages 280–281). Then from (18) we have

$$||A_2^{(j)}|| \leq ||F^{-1}||^{k2^j} ||I - A_1^{(j)}|| + ||F^{-1}||^{k2^{j+1}} ||A_0^{(j)}||,$$

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

$$||A_2^{(j)}|| = O((\rho(F^{-1}) + \epsilon)^{k2^j}) = O((|\xi_m|^{-1} + \epsilon)^{k2^j}).$$

If ξ_m is a simple root of $\gamma(z)$ then there exists a norm $||\cdot||$ such that $||F^{-1}|| = |\xi_m|^{-1}$. Therefore, by applying the same argument as before, the above relation turns into $||A_2^{(j)}|| = O(|\xi_m|^{-k2^j})$. The convergence of the sequence $\widehat{A}_2^{(j)}$ follows by applying the same arguments to equation (19). $\square$

6 NUMERICAL EXPERIMENTS

We implemented our algorithm for reblocked QBD problems in Fortran 90 (Non-skip-free Cyclic Reduction, NCR Algorithm) and compared it with the cyclic reduction (CR Algorithm) for general QBD of [3], and with the algorithm of Gail, Hantler, Taylor [9] (GHT Algorithm). We generated random stochastic matrices with $p = 4$ for several values of k.

The following table reports the time (in seconds) needed by the programs to compute the matrix G, together with the residual error and the number of iterations sufficient to reach the requested accuracy. A "*" denotes failure of the program due to lack of memory.

The computations have been performed on an alpha workstation.

	NCR			CR			GHT		
k	Steps	Time	Residual	Steps	Time	Residual	Steps	Time	Residual
16	10	14	$6 \cdot 10^{-16}$	10	1	$1 \cdot 10^{-15}$	1534	6	$1 \cdot 10^{-10}$
32	9	19	$6 \cdot 10^{-15}$	9	10	$1 \cdot 10^{-15}$	848	13	$1 \cdot 10^{-10}$
64	9	38	$5 \cdot 10^{-13}$	9	136	$2 \cdot 10^{-14}$	980	60	$1 \cdot 10^{-10}$
128	10	89	$1 \cdot 10^{-11}$	*	*	*	2229	555	$1 \cdot 10^{-10}$
256	9	146	$2 \cdot 10^{-15}$	*	*	*	800	816	$1 \cdot 10^{-10}$
512	12	448	$3 \cdot 10^{-11}$	*	*	*	1380	5830	$2 \cdot 10^{-4}$

Table 1: Non-skip-free Markov chains

From the numerical experiments it results that our algorithm is much faster than CR and GHT if the number k of boundary levels is sufficiently large. Moreover, for the small memory storage, it is much robust than CR. Unlike GHT, the performance of our algorithm does not seem to be dependent on the values of the input data. This is likely due to the quadratic convergence of the cyclic reduction.

REFERENCES

[1] N. Akar and K. Sohraby. An invariant subspace approach in M/G/1 and G/M/1 type Markov chains. *Commun. Statist. Stochastic Models*, 13:381–416, 1997.

[2] 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, 1995.

[3] 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.

[4] D. Bini and B. Meini. Effective methods for solving banded Toeplitz systems. Annual SIAM Meeting, Stanford, CA, 1997. Submitted for publication, 1997.

[5] D. Bini and B. Meini. Improved cyclic reduction for solving queueing problems. *Numerical Algorithms*, 15:57–74, 1997.

[6] 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.*, 272:1–16, 1998.

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

[8] 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*, 10:1–43, 1994.

[9] 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.*, 29:733–758, 1997.

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

[11] W. K. Grassman and D. P. Heyman. Computation of steady state probabilities for infinite Markov chains with repeating rows. *ORSA J. on Computing*, 5:292–303, 1993.

[12] G. Heinig and K. Rost. *Algebraic Methods for Toeplitz-like Matrices and Operators.* Akademie-Verlag, Berlin, and Birkhäuser, Boston, 1984.

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

[14] P. Lancaster and M. Tismenetski. *The Theory of Matrices.* Academic Press, 1985.

[15] G. Latouche. Algorithms for infinite Markov chains with repeating columns. In C.D. Meyer and R.J. Plemmons, editors, *Linear Algebra, Queueing Models and Markov Chains*, pages 231–265. Springer-Verlag, New York, 1993.

[16] G. Latouche. Newton's iteration for non-linear equations in Markov chains. *IMA J. of Numerical Analysis*, 14:583–598, 1994.

[17] G. Latouche and V. Ramaswami. A logarithmic reduction algorithm for Quasi-Birth-Death processes. *J. Appl. Probability*, 30:650–674, 1993.

[18] G. Latouche and G.W. Stewart. Numerical methods for M/G/1 type queues. In W. J. Stewart, editor, *Computations with Markov Chains*, pages 571–581. Kluwer Academic Publishers, 1996.

[19] B. Meini. New convergence results on functional iteration techniques for the numerical solution of M/G/1 type Markov chains. *Numer. Math.*, 78:39–58, 1997.

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

[21] V. Ramaswami. Nonlinear matrix equations in applied probability - Solution techniques and open problems. *SIAM Review*, 30:256–263, 1988.

[22] V. Ramaswami. A stable recursion for the steady state vector in Markov chains of M/G/1 type. *Commun. Statist. Stochastic Models*, 4:183–188, 1988.

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