Analyzing M/G/1 paradigms through QBDs: the role of the block structure in computing the matrix G

Dario A. Bini* Beatrice Meini* V. Ramaswami†

Abstract

Recently, it has been proved by V. Ramaswami that an M/G/1 type Markov chain can be rewritten in terms of a suitable QBD process, with infinite and strongly structured block entries [11]. This reformulation allows to derive new algorithms for the computation of the matrix G. We show that by applying the cyclic reduction algorithm to the QBD process we generate a sequence of QBDs, having constant displacement rank. More precisely, we show that these blocks belong to the algebra generated by a suitable block Frobenius matrix. This fact allows us to devise a new quadratically convergent algorithm, that makes use of Fast Fourier Transforms and of Toeplitz matrix computations, for approximating the matrix G. A complexity analysis is performed.

1 Introduction

In a recent paper [11] it has been proved by V.Ramaswami that an M/G/1 type Markov chain can be embedded into a suitable QBD process with an infinite number of phases. This fact allows one to apply, for the solution of an M/G/1 type Markov chain, any algorithm designed for solving QBD problems. Indeed, this fact provides new algorithmic possibilities and new approaches for solving general M/G/1 Markov chains and to reduce the solution of the matrix equation

$$G = \sum_{i=0}^{+\infty} D_i G^i$$

*Dipartimento di Matematica, Università di Pisa, 56127 Pisa, Italy.
†AT&T Labs, Holmdel, NJ, USA

associated to an M/G/1 type Markov chain, to the solution of

$$H = C + AH + BH^2 \tag{1}$$

associated with the embedding QBD.

An important feature of this reduction is that the special QBD problem derived from the general one, has a very strong structure; therefore, despite the new infinite size of the blocks, one might expect a reduction of the complexity if the underlying structure is suitably exploited when applying the QBD solver.

In this paper we provide a structure analysis of the QBD problem obtained by this reduction and show how to exploit this structure in order to speed up the logarithmic reduction algorithm (also known as cyclic reduction and hereafter referred to as CR) introduced in [8] and modified in [4, 3].

In particular, we prove that the sequence of QBD's generated by cyclic reduction are defined by blocks that are related to the matrix algebra defined by a suitable block Frobenius matrix. This fact allows us to prove that the displacement rank of each block is finite, constant and, consequently, that a single step of cyclic reduction can be performed with a low computational costs by means of a small number of FFT's.

More precisely, in order to solve the matrix equation (1), the arithmetic cost of one step of CR is $O(m^3N + m^2N \log N)$ arithmetic operations (ops), where m is the size of D_i (number of phases), N is the numerical degree of the matrix power series $\sum_{i=0}^{+\infty} D_i z^i$, that is, N is such that $\sum_{i=0}^{N} D_i \boldsymbol{e} < (1-\epsilon)\boldsymbol{e}$ where $\boldsymbol{e} = (1, \ldots, 1)^T$ and ϵ is the machine precision used in the computation.

The algorithm obtained in this way has been implemented in fortran 90 and the numerical experiments, performed so far, show very good convergence properties even for problems having a certain complexity [1]. Situations where the new approach can be more favorable with respect to directly applying CR to M/G/1 type Markov chains are investigated.

Similarly to the recent applications of CR to factoring polynomials and power series [2], the new approach seems to provide new and more efficient tools for computing factors of small degree of a given polynomial or power series.

The paper is organized as follows. In section 2 we recall the reduction of [11] from general QBD to M/G/1 problems. In section 3 we synthesize the algorithm of cyclic reduction with its convergence and basic structural properties. In section 4 we prove the main displacement structures of the blocks generated by cyclic reduction that are used in section 5 for the explicit description of an algorithm for the computation of the matrix G. In section 6 we briefly comment on some numerical experiments made with a prototype implementation of our algorithm and on possible applications of the newly presented approach.

2 M/G/1 paradigms through QBDs: structure analysis

Let

$$P = \begin{bmatrix} E_0 & E_1 & E_2 & \dots \\ D_0 & D_1 & D_2 & \dots \\ 0 & D_0 & D_1 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix} \tag{2}$$

be the transition matrix of an M/G/1 type Markov chain, where E_i, D_i, $i \geq 0$, are $m \times m$ matrices, and let G be the minimal nonnegative solution of $G = \sum_{i=0}^{\infty} D_i G^i$. Consider the QBD chain obtained by embedding the M/G/1 type Markov chain, as in [11]:

$$Q = \begin{bmatrix} \widetilde{A} & B & & \\ C & A & B & \\ & C & A & \ddots \\ & & \ddots & \ddots \end{bmatrix} \tag{3}$$

where

$$\begin{aligned} \widetilde{A} &= \begin{bmatrix} E_0 & E_1 & E_2 & \dots \\ 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \end{bmatrix}, \quad A = \begin{bmatrix} D_1 & D_2 & D_3 & \dots \\ 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \end{bmatrix}, \\ C &= \begin{bmatrix} D_0 & 0 & 0 & \dots \\ 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 0 & \dots \\ I & 0 & 0 & \dots \\ 0 & I & 0 & \dots \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}. \end{aligned} \tag{4}$$

We recall the following result of [11] that allows us to reduce a general M/G/1-type Markov chain to a QBD process with infinite phases.

Theorem 1 *Let G be the minimal nonnegative solution of $G = \sum_{i=0}^{+\infty} D_i G^i$, and denote H the minimal nonnegative solution of the matrix equation $H =$*

$C + AH + BH^2$, where A, B, C are defined in (4). Then it holds

$$H = \begin{bmatrix} G & 0 & 0 & \dots \\ G^2 & 0 & 0 & \dots \\ G^3 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \end{bmatrix}.$$

Let us denote A_N, B_N, C_N and H_N the block $N \times N$ matrices obtained by truncating A, B, C and H at the block size N. Then it is easy to verify that, if $D_i = 0$, for $i > N$ then $H_N = C_N + A_N H_N + B_N H_N^2$.

Besides its theoretical interest, the above results allow us to apply algorithms designed for the QBD case in order to solve general M/G/1-type Markov chains. In particular we may consider quadratically convergent methods like the logarithmic reduction technique of [8] or the slightly modified version of [3] based on the cyclic reduction algorithm.

The main issue in the application of these methods is the exploitation of the structures of the blocks (4) and of the solution H that must be taken into account if we want to arrive at an efficient design of the solution algorithm.

Since $\sum_{i=0}^{+\infty} D_i$ is stochastic, it holds that D_i is negligible for i large enough. In practice, if the entries of D_i are less than the machine precision of the floating point arithmetic used in the computation for $i > N$, then we may assume that $D_i = 0$ for $i > N$. In this way without loss of generality, for practical applications, we may assume that the blocks A, B, C of (4) have finite size N. Here and hereafter we make this assumption.

3 Properties of cyclic reduction

In this section we briefly recall the cyclic reduction technique [3] for the solution of the matrix equation $H = B + AH + CH^2$ and analyze its behavior in the case (4).

The cyclic reduction method generates the sequences of matrices A_n, B_n, C_n, $\widehat{A}_n$, $n \geq 0$, such that

$$\begin{array}{l} \widehat{A}_0 = A_0 = A,\ B_0 = B,\ C_0 = C, \\ A_{n+1} = A_n + B_n(I - A_n)^{-1}C_n + C_n(I - A_n)^{-1}B_n, \\ B_{n+1} = B_n(I - A_n)^{-1}B_n, \\ C_{n+1} = C_n(I - A_n)^{-1}C_n, \\ \widehat{A}_{n+1} = \widehat{A}_n + C_n(I - A_n)^{-1}B_n, \quad n \geq 0. \end{array} \tag{5}$$

It holds

$$H = C + \widehat{A}_n H + B_n H^{2^n+1}. \tag{6}$$

Moreover, we have the following convergence result [9, 5, 4, 3].

Theorem 2 *If the QBD process (3) is positive recurrent then the polynomial* $f(z) = \det(C + zA + z^2B - zI)$ *has* m *zeros in the closed unit disk. Moreover, if the polynomial* $f(z)$ *has* m *zeros in the closed unit disk then* $||B_n|| = O(\lambda^{2^n})$*, where* $||\cdot||$ *is any induced matrix norm,* $\lambda = 1/\mu < 1$ *and* $\mu = \min\{|z| : \ \ f(z) = 0, \ \ |z| > 1\}$.

From this property and from equation (6) we find that $(I - \widehat{A}_n)^{-1}C$ provides accurate approximations of H for n large enough.

3.1 Structure of the blocks

The aim of this section is to describe the structural properties of the blocks A_n, B_n, C_n and $\widehat{A}_n$ generated by the cyclic reduction (5). We recall that we are assuming that A, B, C are defined by (4) and have finite size N.

Let us denote by $\boldsymbol{e}_1$ the $Nm \times m$ matrix made up by the first m columns of the identity matrix of size Nm and by Z the block down-shift matrix, i.e., the $N \times N$ finite section of the matrix B in (4). Here and hereafter we call *block vector* any $Nm \times m$ matrix.

Theorem 3 *The matrices* A_n, B_n, C_n *and* $\widehat{A}_n$ *satisfy the following relations*

$$\begin{array}{l} \widehat{A}_n = \boldsymbol{e}_1 \boldsymbol{v}_0^T + Z\boldsymbol{u}_n \boldsymbol{e}_1^T, \\ A_n = \boldsymbol{e}_1 \boldsymbol{v}_n^T + Z\boldsymbol{u}_n \boldsymbol{e}_1^T, \\ B_{n+1} = B_n(I + \hat{\boldsymbol{u}}_n S_n \boldsymbol{v}_n^T)B_n, \\ C_n = \boldsymbol{e}_1 D_0^{(n)} \boldsymbol{e}_1^T, \\ (I - A_n)^{-1} = I + \left[\begin{array}{c|c} \boldsymbol{e}_1 & Z\boldsymbol{u}_n \end{array}\right] \left[\begin{array}{cc} S_n & T_n \\ S_n & I + T_n \end{array}\right] \left[\begin{array}{c} \boldsymbol{v}_n^T \\ \boldsymbol{e}_1^T \end{array}\right], \quad n \geq 0, \end{array} \tag{7}$$

where

$$\begin{array}{l} B_0 = Z, \ D_0^{(0)} = D_0, \ \hat{\boldsymbol{u}}_0 = \boldsymbol{e}_1, \ \boldsymbol{v}_0^T = [D_1, \ldots, D_N], \\ D_0^{(n+1)} = D_0^{(n)}(S_n v_{n,1} + T_n + I)D_0^{(n)}, \\ \hat{\boldsymbol{u}}_{n+1} = \hat{\boldsymbol{u}}_n + B_n \hat{\boldsymbol{u}}_n (S_n v_{n,1} + T_n + I) D_0^{(n)}, \\ \boldsymbol{v}_{n+1}^T = \boldsymbol{v}_n^T + D_0^{(n)} S_n \boldsymbol{v}_n^T B_n, \ n \geq 0, \end{array} \tag{8}$$

and

$$S_n = (I - \boldsymbol{v}_n^T \hat{\boldsymbol{u}}_n)^{-1}, \quad T_n = S_n \boldsymbol{v}_n^T Z \boldsymbol{u}_n, \quad \boldsymbol{u}_n = Z^T \hat{\boldsymbol{u}}_n,$$

and $v_{n,1}$ *is the first block entry of the block vector* $\boldsymbol{v}_n^T$.

Proof. Proceed by induction on n. For $n = 0$ the theorem trivially holds. In the general case, by applying (5) we immediately obtain the representation of B_n, C_n and A_n. The representation of $(I - A_n)^{-1}$ follows from the Sherman-Morrison-Woodbury formula [6] applied to $(I - A_n)$. □

From the above theorem it follows that the blocks $\widehat{A}_n$, A_n and C_n still keep the sparsity properties of $\widehat{A}_0$, A_0 and C_0, respectively. More precisely, the matrices $\widehat{A}_n$ and A_n have null block entries except for the blocks in the first block column and in the first block row; and the matrices C_n have null block entries except for the one in position (1,1). At a first glance, the matrices B_n have no sparsity nor structural properties; indeed, it is easy to verify that if $n < \log_2 N$, then B_n has the following structure

$$B_n = \begin{bmatrix} 0 & 0 & \dots & 0 \\ * & * & \dots & * \\ \vdots & \vdots & & \vdots \\ * & * & \dots & * \\ I & & & \\ & \ddots & & \\ & & I & \end{bmatrix}$$

where $*$ denotes a possibly nonzero block entry and where the number of nonzero block rows is $2^n - 1$; if $n \geq \log_2 N$, then B_n seems to be a general $N \times N$ matrix without any structural properties. The lack of structure of B_n would lead to an inefficient implementation of the cyclic reduction algorithm, in fact each step of CR would cost $O(N^3 m^3)$ ops.

Fortunately, it is possible to prove that the matrix B_n has a very strong structure (the displacement structure) that allows us to implement one step of CR with $O(m^2 N \log N + m^3 N)$ ops. This property is not trivial and needs some sophisticated technique to be proved. We will do this in section 4.

3.2 Convergence results

Before giving the convergence properties of the cyclic reduction algorithm, we recall that the $N \times N$ block Frobenius matrix, associated with the monic matrix polynomial $P(z) = \sum_{i=0}^{N-1} P_i z^i + z^N I$, is the matrix

$$F = \begin{bmatrix} -P_{N-1} & -P_{N-2} & \dots & -P_0 \\ I & 0 & \dots & 0 \\ & \ddots & \ddots & \vdots \\ \bigcirc & & I & 0 \end{bmatrix};$$

moreover, it holds $\det P(z) = \det(zI - F)$ (compare [7]).

Concerning the convergence results we have the following

Theorem 4 *Let* $f(z) = \det(C + zA + z^2 B - zI)$, $g(z) = \det(\sum_{i=0}^{N} D_i z^i - zI)$. *Then we have* $f(z) = z^{m(N-1)} g(z)$.

Proof. Assume that D_0 is nonsingular. It holds

$$f(z) = \det\left(\begin{bmatrix} D_0 & & & \bigcirc \\ & -zI & & \\ & & \ddots & \\ \bigcirc & & & -zI \end{bmatrix}(I - zF)\right)$$

where F is the block Frobenius matrix defined by the blocks $-D_0^{-1}(D_1 - I)$, $-D_0^{-1}D_i$, $i = 2, \ldots, N$. Therefore, $f(z) = z^{m(N-1)} \det D_0 \det(z(z^{-1}I - F))$. In the light of the properties of the determinants of block Frobenius matrices we have $f(z) = z^{m(N-1)} \det(zI - \sum_{i=0}^{N} D_i z^i) = z^{m(N-1)} g(z)$. If the matrix D_0 is singular we may apply a continuity argument in order to complete the proof. □

The result expressed in the above theorem is fundamental for proving the convergence of cyclic reduction. In fact, if the M/G/1 Markov chain (2) is positive recurrent then the polynomial $g(z)$ has m roots inside the unit disk [10] (compare also theorem 2), therefore the polynomial $f(z)$ has mN roots of modulus at most 1 of which $m(N-1)$ are zero. In the light of theorem 2 the latter property implies the following convergence result:

Theorem 5 *If the M/G/1 Markov chain (2) is positive recurrent then* $||B_n|| = O(\lambda^{2^n})$*, where* $||\cdot||$ *is any induced matrix norm,* $\lambda = 1/\mu$ *and* $\mu = \min\{|z| : g(z) = 0, \quad |z| > 1\}$*. Moreover, the sequences* $\{A_n\}$, $\{C_n\}$, $\{\widehat{A}_n\}$ *are convergent.*

From the above theorem we deduce the following property

Theorem 6 *If the matrix* G *is nonsingular then* $u_{*,i} = G^i$, $i = 1, \ldots, N-1$, *where* $u_{*,i} = \lim_n u_{n,i}$ *and* $u_{n,i}$ *is the* i*-th block component of the vector* $\boldsymbol{u}_n$.

Proof. From (6) and from theorem 5 we obtain $(I - \widehat{A}_*)H = B$ where $\widehat{A}_* = \lim_n \widehat{A}_n$. Since $\widehat{A}_*$ has the same structure of $\widehat{A}_n$ we deduce that $u_{*,i}G = G^{i+1}$, $i = 1, \ldots, N-1$, whence, from the nonsingularity of G we obtain that $u_{*,i} = G^i$. □

Observe that if the matrix G is nonsingular, then it can be approximated by the matrix $u_{n,1}$, for sufficiently large n; on the other hand, if G is singular, then it can be approximated by the matrix $D_0(I - D_1 - \sum_{i=2}^{N} D_i u_{n,i-1})^{-1}$ that can be easily derived from the first block column of (6).

4 Displacement structure

We recall the following properties concerning the displacement rank of a matrix. Let $\Delta(A) = AZ - ZA$ and consider the following decomposition

$$\Delta(A) = \sum_{i=1}^{k} \boldsymbol{r}_i \boldsymbol{s}_i^T \tag{9}$$

where $\boldsymbol{r}_i$ and $\boldsymbol{s}_i$ are block vectors. We define the *block displacement rank* of A as the minimum integer k such that (9) holds. If (9) holds then (compare [5])

$$A = L(A\boldsymbol{e}_1) - \sum_{i=1}^{k} L(\boldsymbol{r}_i)L(Z\boldsymbol{s}_i)^T$$

where $L(\boldsymbol{w})$ denotes the block lower triangular block Toeplitz matrix defined by the block column $\boldsymbol{w}$ having block components w_i, $i = 1, \ldots, N$,

$$L(\boldsymbol{w}) = \begin{bmatrix} w_1 & & & O \\ w_2 & \ddots & & \\ \vdots & \ddots & \ddots & \\ w_N & \ldots & w_2 & w_1 \end{bmatrix}.$$

We recall that a block Toeplitz matrix A is a matrix having block entries $(A)_{i,j} = A_{i-j}$. It is immediate to check that a block Toeplitz matrix has block displacement rank at most 2 and a block lower triangular block Toeplitz matrix has block displacement rank 0.

We recall also that the product $L(\boldsymbol{w})\boldsymbol{v}$, for a pair of block vectors $\boldsymbol{w}$ and $\boldsymbol{v}$ can be computed with $O(m^2 N \log_2 N + m^3 N)$ ops by using the Fast Fourier Transforms (FFT) [5].

4.1 Frobenius matrices and displacement rank

Let F be an $N \times N$ block Frobenius matrix. Then the powers of F can be expressed in terms of block Toeplitz matrices by means of the following result:

Theorem 7 *For any integer $k \geq 0$ it holds*

$$\Delta(F^k) = \boldsymbol{e}_1 \boldsymbol{u}^T F^k - F^k \boldsymbol{e}_1 \boldsymbol{u}^T \tag{10}$$

where $\boldsymbol{u}^T = \boldsymbol{e}_1^T F$ is the first block row of F.

Proof. Since $Z = F - \boldsymbol{e}_1\boldsymbol{u}^T$ it holds $\Delta(F^k) = F^k Z - Z F^k = F^k(F - \boldsymbol{e}_1\boldsymbol{u}^T) - (F - \boldsymbol{e}_1\boldsymbol{u}^T)F^k = \boldsymbol{e}_1\boldsymbol{u}^T F^k - F^k \boldsymbol{e}_1 \boldsymbol{u}^T$. □

From the above theorem it follows that any power of a block Frobenius matrix has block displacement rank at most 2.

4.2 Displacement properties of the blocks

Now, for the sake of simplicity we assume that $\det D_0 \neq 0$. We recall a suitable relation that expresses the blocks A_n, B_n, C_n in functional form. By using this relation we may prove the displacement structure of the blocks B_n at each step of the cyclic reduction.

Theorem 8 *Let $\phi_n(z) = zI - B_n - zA_n - z^2C_n$. If D_0 is nonsingular, then at each step n of cyclic reduction the matrix $\boldsymbol{e}_1^T F^{2^n-1}\boldsymbol{e}_1$ is nonsingular and it holds:*

$$\phi_n(z) = (I - \boldsymbol{w}_n(z)H_n\boldsymbol{e}_1^T)(zI - F^{2^n}) \tag{11}$$

where $H_n = (\boldsymbol{e}_1^T F^{2^n-1}\boldsymbol{e}_1)^{-1}$, $\boldsymbol{w}_n(z) = F^{2^n-1}\boldsymbol{e}_1 + z\boldsymbol{e}_1 D_0$ and

$$F = \begin{bmatrix} -D_0^{-1}(D_1 - I) & -D_0^{-1}D_2 & \dots & -D_0^{-1}D_N \\ I & 0 & \dots & 0 \\ & \ddots & \ddots & \vdots \\ \bigcirc & & I & 0 \end{bmatrix}.$$

In particular, for the block B_n we have

$$B_n = \phi_n(0) = -\left(F^{2^n-1} - F^{2^n-1}\boldsymbol{e}_1 H_n \boldsymbol{e}_1^T F^{2^n-1}\right) Z. \tag{12}$$

Proof. We prove, by induction on n, that $\psi_n(z) = \phi_n(z)^{-1}$ verifies

$$(zI - F^{2^n})\psi_n(z) = \left(I - \boldsymbol{v}_n(z)\boldsymbol{e}_1^T\right) \tag{13}$$

for all the values z for which $\psi_n(z)$ exists, where $\boldsymbol{v}_n(z) = \boldsymbol{e}_1 + z^{-1}F^{2^n-1}\boldsymbol{e}_1 D_0^{-1}$. For $n = 0$ the equation trivially holds. Assuming the equation true for n, evaluate (13) for the values z and $-z$, compute the sums and the differences of both members of the equations obtained this way. Recover $\psi_n(z) + \psi_n(-z)$ from one equation and substitute this value in the other one, and recover $(z^2I - F^{2^{n+1}})(\psi_n(z) - \psi_n(-z))/(2z)$. The induction step is completed since $\psi_{n+1}(z^2) = (\psi_n(z) - \psi_n(-z))/(2z)$ (compare [5]). Applying the Sherman-Woodbury-Morrison formula to the matrix $(I - \boldsymbol{v}_n(z)\boldsymbol{e}_1^T)$ we obtain that this matrix is nonsingular if and only if $z^{-1}\boldsymbol{e}_1^T F^{2^n-1}\boldsymbol{e}_1$ is nonsingular. Since $\phi_n(z)$ is singular for a finite number of values of z, the matrix $\psi_n(z) = \phi_n(z)^{-1}$ exists and is nonsingular for an infinite number of values of z. Therefore from (13) it follows that there exists at least a value $z_0 \neq 0$ of z such that the matrix $I - \boldsymbol{v}_n(z_0)\boldsymbol{e}_1^T$ is nonsingular. Whence we obtain that $\boldsymbol{e}_1^T F^{2^n-1}\boldsymbol{e}_1$ is nonsingular and it holds that $(I - \boldsymbol{v}_n(z)\boldsymbol{e}_1^T)^{-1} = I - \boldsymbol{w}_n(z)H_n\boldsymbol{e}_1^T$ for any $z \neq 0$. Therefore from (13) we obtain $\phi_n(z) = \left(I - \boldsymbol{w}_n(z)H_n\boldsymbol{e}_1^T\right)(zI - F^{2^n})$.

By applying a continuity argument we deduce that the above relation holds for any z. In particular for $z = 0$ we obtain (12). $\square$

An immediate consequence of the above theorem is that B_n has displacement rank at most 2 whenever D_0 is nonsingular.

Theorem 9 *If D_0 is nonsingular, the matrices B_n, $n \geq 0$, generated by (5), are such that*

$$\Delta(B_n) = \left((\boldsymbol{e}_1 - F^{2^n-1}\boldsymbol{e}_1 H_n)\boldsymbol{u}^T + ZF^{2^n-1}\boldsymbol{e}_1 H_n \boldsymbol{e}_1^T\right) F^{2^n-1} Z \tag{14}$$

where $H_n = (\boldsymbol{e}_1^T F^{2^n-1}\boldsymbol{e}_1)^{-1}$, thus they have block displacement rank at most 2.

Proof. For the sake of notational simplicity set $k = 2^n - 1$. We have

$$\begin{aligned} \Delta(B_n) = \Delta(F^k - F^k\boldsymbol{e}_1 H_n \boldsymbol{e}_1^T F^k)Z = \\ \Delta(F^k)Z - (F^k\boldsymbol{e}_1 H_n \boldsymbol{e}_1^T F^k Z - ZF^k\boldsymbol{e}_1 H_n \boldsymbol{e}_1^T F^k)Z. \end{aligned} \tag{15}$$

Since $Z = F - \boldsymbol{e}_1\boldsymbol{u}^T$ we obtain $F^k\boldsymbol{e}_1 H_n \boldsymbol{e}_1^T F^k Z = F^k\boldsymbol{e}_1 H_n \boldsymbol{u}^T F^k - F^k\boldsymbol{e}_1\boldsymbol{u}^T$. Plugging the latter equation into (15) we arrive at (14). $\square$

If $\det D_0 = 0$, since B_n is well defined through (5), by using a continuity argument we arrive at the following:

Corollary 9.1 *The matrices B_n have block displacement rank at most 2 for any $n \geq 0$.*

Observe that the matrix B_n can be represented by means of five block vectors: the first block column $\boldsymbol{b}_n$ of B_n and the 4 block vectors, $\boldsymbol{c}_1^{(n)}$, $\boldsymbol{c}_2^{(n)}$, $\boldsymbol{r}_1^{(n)}$, $\boldsymbol{r}_2^{(n)}$, such that

$$\Delta(B_n) = \boldsymbol{c}_1^{(n)}\boldsymbol{r}_1^{(n)T} + \boldsymbol{c}_2^{(n)}\boldsymbol{r}_2^{(n)T},$$

defining the block displacement rank of B_n. In the next section we describe an algorithm for their computation.

5 Computing the matrix G

Now we are ready to describe the main algorithm for performing one step of cyclic reduction.

Algorithm: *Computing one step of cyclic reduction.*

INPUT:

– The block vectors $\boldsymbol{u}$, $\boldsymbol{v}$, such that $A = \boldsymbol{e}_1\boldsymbol{v}^T + Z\boldsymbol{u}\boldsymbol{e}_1^T$, $\widehat{A} = \boldsymbol{e}_1\boldsymbol{v}_0^T + Z\boldsymbol{u}\boldsymbol{e}_1^T$;

– the block entry D such that $C = \boldsymbol{e}_1 D\boldsymbol{e}_1^T$;

– the block vectors $\boldsymbol{b} = B\boldsymbol{e}_1$, and $\boldsymbol{c}_1$, $\boldsymbol{c}_2$, $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, such that $\Delta(B) = \boldsymbol{c}_1\boldsymbol{r}_1^T + \boldsymbol{c}_2\boldsymbol{r}_2^T$.

OUTPUT:

– The block vectors $\boldsymbol{u}'$, $\boldsymbol{v}'$, such that $A' = \boldsymbol{e}_1\boldsymbol{v}'^T + Z\boldsymbol{u}'\boldsymbol{e}_1^T$, $\widehat{A}' = \boldsymbol{e}_1\boldsymbol{v}_0^T + Z\boldsymbol{u}'\boldsymbol{e}_1^T$;

– the block entry D' such that $C' = \boldsymbol{e}_1 D'\boldsymbol{e}_1^T$;

– the block vectors $\boldsymbol{b}' = B'\boldsymbol{e}_1$, and $\boldsymbol{c}_1'$, $\boldsymbol{c}_2'$, $\boldsymbol{r}_1'$, $\boldsymbol{r}_2'$, such that $\Delta(B') = \boldsymbol{c}_1'\boldsymbol{r}_1'^T + \boldsymbol{c}_2'\boldsymbol{r}_2'^T$,

where A', B', C', $\widehat{A}'$ are the matrices obtained after one step of cyclic reduction applied to $A, B, C, \widehat{A}$.

COMPUTATION:

1. Compute $S = (I - \boldsymbol{v}^T\hat{\boldsymbol{u}})^{-1}$, $T = S\boldsymbol{v}^T Z\boldsymbol{u}$, that define $(I-A)^{-1}$ by means of formula (7).
2. Compute the block entry D' by means of (8).
3. Compute the vectors $\hat{\boldsymbol{u}}'$, $\boldsymbol{v}'$, by means of (8), where the matrix B is written in the form
$$B = L(\boldsymbol{b}) - L(\boldsymbol{r}_1)L(Z\boldsymbol{c}_1)^T - L(\boldsymbol{r}_2)L(Z\boldsymbol{c}_2)^T. \qquad (16)$$
4. Compute the vector $\boldsymbol{b}' = B'\boldsymbol{e}_1 = -B(I-A)^{-1}\boldsymbol{b}$, where the matrix B is written in the form (16).
5. Compute the $(Nm)\times(7m)$ matrices W and Y such that
$$WY^T = -\Delta(B)(I-A)^{-1}B - B\Delta((I-A)^{-1})B - B(I-A)^{-1}\Delta(B).$$
6. Compute a QR factorization of W and Y, i.e., $W = Q_1R_1$, $Y = Q_2R_2$, where Q_1, Q_2 are $(Nm)\times(7m)$ matrices, and R_1, R_2 are $(7m)\times(7m)$ matrices; in this way $WY^T = Q_1(R_1R_2^T)Q_2^T$.
7. Compute the singular value decomposition of $R_1R_2^T$, i.e., $R_1R_2^T = U\Sigma V^T$, where $\Sigma = \mathrm{diag}(\sigma_1,\ldots,\sigma_{7m})$, $\sigma_1 \geq \cdots \geq \sigma_{7m}$.

8. Set

$$\left[\, \boldsymbol{c}_1' \mid \boldsymbol{c}_2' \,\right] = Q_1 U_{2m} \mathrm{diag}(\sigma_1, \ldots, \sigma_{2m}), \quad \begin{bmatrix} \boldsymbol{r}_1'^T \\ \boldsymbol{r}_2'^T \end{bmatrix} = V_{2m}^T Q_2^T,$$

where U_{2m}, V_{2m} are the $(Nm) \times (2m)$ matrices made up by the first $2m$ columns of U and V, respectively.

Observe that the matrices W and Y^T have size $(Nm) \times (7m)$ since $\Delta(B)$ has rank at most $2m$ and $\Delta((I-A)^{-1})$ has rank at most $3m$.

The computational cost of stage 1 is $O(m^3 N)$ ops and the cost of stage 2 is $O(m^3)$ ops; at stage 3 we have to perform products between matrices with block displacement rank at most 2, and block vectors, that can be computed in $O(m^3 N + m^2 N \log N)$ ops; similar computations are involved in stages 4 and 5, with the same cost; at stage 6 the computation of the QR factorizations costs $O(m^3 N)$ ops, and the computation of the singular values decomposition at stage 7 costs $O(m^3)$ ops. Thus, the overall computational cost of one step of cyclic reduction is $O(m^3 N + m^2 N \log N)$.

6 Numerical experiments.

The algorithm has been implemented, in a prototypal version, in fortran 90. The program has been run on a pentium II server with the Linux system for different input data. The experiments have been performed for testing the convergence properties of the technique and its numerical stability. In this preliminary version, the Toeplitz structure of the blocks B_n has not been exploited for the fast computation of the matrix-vector products. Therefore in the table below we report only the number of steps and the residual error of the computation. The problems that we have considered are the ones of [1] with different values of the parameter ρ. For $\rho = 1$ the chain is not positive recurrent. The size m of the blocks is 16, the value N depends on the parameter ρ. The results have been compared with the experiments reported in [4] (PWCR algorithm).

We observe that the number of steps and the residual errors of the method described in this paper are comparable with the correspondent values of the algorithm of [4].

By looking at the complexity bounds we may determine the cases where the new approach is more convenient than the point-wise cyclic reduction. In fact, the cost of the point-wise cyclic reduction of [4] depends on the numerical degree of the power series generated at the intermediate steps. At the first step this number is N and in the limit is 1, but in the intermediate

		New Algorithm		PWCR	
ρ	N	Steps	Residual	Steps	Residual
0.1	256	10	$6.8 \cdot 10^{-15}$	9	$1.8 \cdot 10^{-13}$
0.8	256	13	$1.2 \cdot 10^{-14}$	13	$5.4 \cdot 10^{-14}$
0.9	512	14	$3.6 \cdot 10^{-14}$	14	$1.6 \cdot 10^{-14}$
0.97	512	20	$3.2 \cdot 10^{-14}$	20	$4.2 \cdot 10^{-14}$

Table 1: Number of steps and residual errors

steps this number could be larger than N. On the other hand for the new method the cost is a function of only the numerical degree of the input matrix power series. Therefore the new approach should be more convenient in the cases where the numerical degrees of the intermediate power series generated by the point-wise cyclic reduction exceed N.

Another interesting application of the new approach concerns the problem of approximating a factor of a polynomial or of a power series. In [2] CR has been successfully applied to split a polynomial of degree n into two factors of degrees $n-k$ and k. The most expensive stage consists in inverting a Toeplitz like matrix of size $\max\{n-k, k\}$. With the new approach we have to invert a matrix of size $\min\{n-k, k\}$. This might be a great advantage if $k << n/2$.

References

[1] G. Anastasi, L. Lenzini, and B. Meini. Performance evaluation of a worst case model of the MetaRing MAC protocol with global fairness. *Performance Evaluation*, 29:127–151, 1997.

[2] D. Bini, L. Gemignani and B. Meini. Factorization of analytic functions by means of Koenig's theorem and Toeplitz computations. Technical Report, Dipartimento di Informatica, Università di Pisa, 1998.

[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. Improved cyclic reduction for solving queueing problems. *Numerical Algorithms*, 15:57–74, 1997.

[5] D. A. Bini and B. Meini. Effective methods for solving banded Toeplitz systems. *SIAM J. Matrix Anal. Appl.*, 20:700–719, 1999.

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

[7] P. Lancaster and M. Tismenetski. *Matrix Polynomials.* Academic Press Inc., New York, 1982.

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

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

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

[11] V. Ramaswami. The generality of Quasi Birth-and-Death processes. In A. Alfa and S. Chakravarthy, editors, *Advances in Matrix Analytic Methods for Stochastic Models - Proceedings of the 2nd international conference on matrix analytic methods*, pages 93–113. Notable Publications Inc, NJ, 1998.