SIAM J. MATRIX ANAL. APPL.
Vol. , No. , pp. ,

ON THE SOLUTION OF A NONLINEAR MATRIX EQUATION ARISING IN QUEUEING PROBLEMS *

DARIO BINI† AND BEATRICE MEINI†

Abstract. By extending the cyclic reduction technique to infinite block matrices we devise a new algorithm for computing the solution G_0 of the matrix equation $G = \sum_{i=0}^{+\infty} G^i A_i$ arising in a wide class of queueing problems. Here A_i, $i = 0, 1, \ldots$, are $k \times k$ nonnegative matrices such that $\sum_{i=0}^{+\infty} A_i$ is column-stochastic. Our algorithm, which under mild conditions generates a sequence of matrices converging quadratically to G_0, can be fully described in terms of simple operations between matrix power series, i.e., power series in z having matrix coefficients. Such operations, like multiplication and reciprocation modulo z^m, can be quickly computed by means of FFT-based fast polynomial arithmetic; here m is the degree where the power series are numerically cut off in order to reduce them to polynomials. These facts lead to a dramatic reduction of the complexity of solving the given matrix equation, in fact, $O(k^3 m + k^2 m \, log\, m)$ arithmetic operations are sufficient to carry out each iteration of the algorithm. Numerical experiments and comparisons performed with the customary techniques, show the effectiveness of our algorithm. For a problem arising from the modelling of metropolitan networks, our algorithm resulted to be about 30 times faster than the algorithms customarily used in the applications. Cyclic reduction applied to QBD problems leads to an algorithm similar to the one of [11] but having a lower computational cost.

Key words. Queueing problems, M/G/1 type matrices, cyclic reduction, Toeplitz matrices.

AMS(MOS) subject classifications. 15A51, 15A24, 60J10, 60K25, 65U05.

1. Introduction. Let A_i, $\widetilde{A}_{i+1}$, $i = 0, 1, 2, \ldots$, be two sequences of $k \times k$ nonnegative matrices such that $\sum_{i=0}^{+\infty} A_i$ and $\sum_{i=1}^{+\infty} \widetilde{A}_i$ are column-stochastic. A matrix $A = \{a_{ij}\}_{ij}$ is column-stochastic if $0 \le a_{i,j} \le 1$, $\sum_i a_{i,j} = 1$ for any j; here and hereafter a stochastic matrix will be a column-stochastic matrix. Consider the infinite stochastic matrix P defined by

$$P = \begin{pmatrix} \widetilde{A}_1 & A_0 & & \bigcirc & \\ \widetilde{A}_2 & A_1 & A_0 & & \\ \widetilde{A}_3 & A_2 & A_1 & A_0 & \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}. \tag{1.1}$$

Matrices of the structure (1.1) are called M/G/1 type matrices [15], and arise in the mathematical modelling of a wide and important class of queueing problems occurring in many applications, where the matrix P^T is the transition matrix associated with a Markov chain [7]. Matrices of this kind have the block Toeplitz structure, except on the first block column, and are in lower block Hessenberg form. We recall that a block Toeplitz matrix is a block matrix having (i, j)-blocks which depend on $i - j$, and that a matrix is in lower block Hessenberg form if its blocks are null for $j > i + 1$.

The computation of the probability invariant vector associated with P, that is, a nonnegative vector $\boldsymbol{\pi}$ such that

$$P\boldsymbol{\pi} = \boldsymbol{\pi}, \ \ ||\boldsymbol{\pi}|| = 1, \tag{1.2}$$

*Received by the editors .....; accepted for publication (in revised form).......

†Dipartimento di Matematica, Università di Pisa, Italy (bini@dm.unipi.it, meini@dm.unipi.it).

plays an important role in the analysis of the queue. Here, for a vector $\mathbf{v} = (v_i)$ we denote $||\mathbf{v}|| = \sum_i |v_i|$.

If the matrix (1.1) is irreducible, the existence of a solution of (1.2) is equivalent to the positive recurrence of P [7]. In this case the solution $\boldsymbol{\pi}$ is unique and positive. For a complete and detailed study of matrices of the M/G/1 type we refer the reader to [15].

Properties of the solution $\boldsymbol{\pi}$ and numerical techniques for its computations are strongly related to the minimal nonnegative solution of the matrix equation

$$G = \sum_{i=0}^{+\infty} G^i A_i. \tag{1.3}$$

In particular, it can be proved that if P is irreducible and positive recurrent then the only nonnegative solution G_0 of (1.3) is a stochastic matrix [15]. The matrix G_0 has very important properties, in fact it is possible to express the solution $\boldsymbol{\pi} = \{\boldsymbol{\pi}_i\}_{i\geq 0}$ of (1.2) in terms of G_0 by means of the Ramaswami formula [15]

$$\boldsymbol{\pi}_i = (I - A_1^*)^{-1}\left[\widetilde{A}_{i+1}^*\boldsymbol{\pi}_0 + \sum_{j=1}^{i-1} A_{i+1-j}^*\boldsymbol{\pi}_j\right],$$

where $\boldsymbol{\pi}_i$ is the k-dimensional vector obtained by partitioning the vector $\boldsymbol{\pi}$ according to the structure (1.1) of P, and

$$A_i^* = \sum_{j=i}^{+\infty} G_0^{j-i} A_j,\ \ \widetilde{A}_{i+1}^* = \sum_{j=i}^{+\infty} G_0^{j-i}\widetilde{A}_{j+1},\ \ i \geq 1. \tag{1.4}$$

The Ramaswami formula, due also to its numerical stability, is an effective tool for solving the problem (1.2) once the solution G_0 of (1.3) has been computed. Thus, the availability of efficient numerical methods for computing the matrix G_0 is crucial in order to solve (1.2).

Different algorithms for computing the solution of (1.3) have been proposed and analyzed by several authors. The most of them are based on functional iteration techniques, obtained by manipulating the matrix equation (1.3). For instance, in [16] the iteration

$$X_{j+1} = \sum_{i=0}^{+\infty} X_j^i A_i, \quad X_0 \geq O \tag{1.5a}$$

is considered. Similar techniques, based on the recurrences

$$X_{j+1} = A_0(I - A_1)^{-1} + \sum_{i=2}^{+\infty} X_j^i A_i (I - A_1)^{-1} \tag{1.5b}$$

or

$$X_{j+1} = A_0(I - \sum_{i=1}^{+\infty} X_j^{i-1} A_i)^{-1} \tag{1.5c}$$

are introduced and analyzed in [15], [9], [14] in order to speed up the convergence. However, the convergence of these numerical schemes still remains linear. In [10] a sort of Newton's iteration is introduced in order to arrive at a quadratic convergence with an extreme increase of the computational cost. In [12] the approximation of G_0 is reduced to solving nested finite systems of linear equations associated with the matrix P by means of a doubling technique. In this way the solution for the matrix P is approximated with the solution of the problem obtained by cutting the infinite block matrix P to a suitable finite block size n.

In this paper we propose a new method for the numerical solution of (1.3) which is quadratically convergent, numerically stable and avoids the need of cutting the infinite matrix (1.1) to a finite size. Moreover this method, based on the use of FFT, requires a small number of arithmetic operations.

Our method relies on the technique of the successive state reduction (cyclic reduction [6]), introduced and analyzed in [2] for the computation of the vector $\boldsymbol{\pi}$, here extended to the computation of the solution G_0 of (1.3).

In the case where the matrix P of (1.1) is block tridiagonal, i.e., $A_i = \widetilde{A}_i = O$, for $i > 2$, our method is very similar to that of Latouche and Ramaswami [11], where a modified cyclic reduction scheme is applied to compute the matrix G_0 associated with a Quasi-Birth-Death process, and, moreover, is slightly faster than the latter algorithm.

Our method generates a sequence $P^{(j)}$ of matrices of the M/G/1 type, defined by the blocks $\{A_i^{(j)}\}_i$ and $\{\widetilde{A}_i^{(j)}\}_i$, such that the solution of the equation $G = \sum_{i=0}^{+\infty} G^i A_i^{(j)}$ is $G_0^{2^j}$. We associate with the sequence $\{P^{(j)}\}_j$ a sequence of $k \times k$ stochastic matrices $\{R^{(j)}\}_j$ that, under mild conditions, due to the quadratic convergence of the matrices $\{G_0^{2^j}\}_j$ to the limit $G' = lim_j\, G_0^{2^j}$, tends quadratically to the matrix G' as j tends to $+\infty$. This allows us to devise a fast algorithm for the approximation of G'. The matrix G_0 is then approximated in a back-substitution stage, by means of a suitable relation which involves G' and the matrices $A_i^{(j)}$. Moreover, we show that under further mild conditions the blocks $A_i^{(j)}$, $i > 1$, tend to the null matrix as j tends to $+\infty$.

The formulae which relate the sequence $\{A_i^{(j)}\}_i$ with the sequence $\{A_i^{(j+1)}\}_i$ are expressed in functional form in terms of the matrix power series $\varphi^{(j)}(z) = \sum_{i=0}^{+\infty} A_i^{(j)} z^i$ in the following way (compare [2])

$$\varphi^{(j+1)}(z) = z\varphi_{odd}^{(j)}(z) + \varphi_{even}^{(j)}(z)\big(I - \varphi_{odd}^{(j)}(z)\big)^{-1}\varphi_{even}^{(j)}(z),$$

where $\varphi_{even}^{(j)}(z) = \sum_{i=0}^{+\infty} A_{2i}^{(j)} z^i$, $\varphi_{odd}^{(j)}(z) = \sum_{i=0}^{+\infty} A_{2i+1}^{(j)} z^i$. The computation of the coefficients of $\varphi^{(j+1)}(z)$, given $\varphi^{(j)}(z)$, is performed by applying the above functional relation modulo z^m, where m (cutting level) is an integer which can be dynamically computed and is such that the matrix $\sum_{i=0}^{m} A_i^{(j)}$ is "numerically stochastic", i.e., $\mathbf{e}^T(I - \sum_{i=0}^{m} A_i^{(j)}) < \epsilon \mathbf{e}^T$, where ϵ is the machine precision. In this way the computation is reduced to compute products and reciprocals of matrix polynomials modulo z^m. This can be efficiently obtained by extending to matrix polynomials the FFT-based techniques which have been devised for scalar polynomial arithmetic [3], [2].

The computational cost of each step of our algorithm is $O(k^3 m + k^2 m \log m)$ arithmetic operations versus $O(k^3 m^2)$ operations needed if the customary polynomial arithmetic were used. The cost of the algorithm presented in [12] is $O(k^3 n \log^2 n)$, where n is the size at which the infinite matrix P is cut in order to reduce the infinite problem to a finite one (typically $n \geq m$). The cost of each step of the linearly convergent algorithms based on (1.5a), (1.5b) and (1.5c) is $O(k^3 m)$ operations.

We implemented our algorithm in Fortran 77 and performed comparisons with the iterative formula based on (1.5c), for $X_0 = 0$ and $X_0 = I$, which is the fastest among (1.5) (compare [14]), and with the algorithm of [12], on a problem arising from the mathematical modelling of a metropolitan network [1]. Our method resulted to be about 30 times faster than the method based on (1.5c), with $X_0 = I$, and about 170 times faster in the case $X_0 = 0$. The ratio between the number of iterations required by (1.5c) and our method has been 167 for $X_0 = I$, and 945 for $X_0 = 0$. From the numerical computations that we have performed, our method resulted to be much superior specially in the cases of "long queues" where either the initial cutting level m of the queue is very large or the number of the nonnegligible components of the vector $\boldsymbol{\pi}$ is very large. In fact, in this cases the advantages of the quadratic convergence and of the use of FFT are strongly evident. Another interesting feature of our algorithm that emerged by the numerical experiments, is that the cutting level m_j at the j-th recursive step of cyclic reduction has an almost decreasing behaviour. This contribuites to speed up the computation.

The paper is organized as follows. In section 2 we describe the cyclic reduction technique applied to matrices of the M/G/1 type, and prove the main convergence result. In section 3 we give further convergence properties. In section 4 we present two different algorithms for the computation of G_0 based on cyclic reduction: the first one, which is slightly faster, requires to store a certain number of blocks that are used in the back-substitution stage; the second one performs the back-substitution implicitely in the first stage and does not need to store auxiliary matrices. In section 5 we describe our implementations, by giving a detailed description of our algorithm together with the FFT-based techniques for manipulating matrix power series. Finally, in section 6 we report the results of the numerical experiments and the comparisons with the known algorithms.

2. Cyclic reduction and its convergence properties. In this section we recall the method of cyclic reduction introduced and analyzed in [2] for the computation of the invariant vector $\boldsymbol{\pi}$, and prove some convergence results which are needed to extend the method to the computation of the matrix G_0.

Here and hereafter we assume that the stochastic matrix P is irreducible and positive recurrent, i.e.,

$$\rho = \mathbf{e}^T \sum_{i=1}^{+\infty} i A_i \mathbf{a} < 1, \tag{2.1}$$

where $\sum_{i=0}^{+\infty} A_i \mathbf{a} = \mathbf{a}$, $||\mathbf{a}|| = 1$, so that the matrix equation (1.3) has only one non-negative solution, which is stochastic [15], and the vector $\boldsymbol{\pi}$ is positive. Under these hypotheses it is possible to prove that [15]

$$r\left(\sum_{i=1}^{+\infty} A_i^*\right) < 1, \tag{2.2}$$

where $r(A)$ denotes the spectral radius of the matrix A and A_i^*, $i = 1, 2, \ldots$, are defined in (1.4).

Let G_0 be the solution of (1.3) and set $H_0 = -A_0$, $H_1 = I - A_1$, $H_i = -A_i$, $i \geq 2$, $\widetilde{H}_1 = I - \widetilde{A}_1$, $\widetilde{H}_i = -\widetilde{A}_i$, $i \geq 2$, $W = \sum_{i=0}^{+\infty} G_0^i \widetilde{A}_{i+1}$. Then from (1.3), we may easily obtain the matrix equation

$$(I, G_0, G_0^2, \ldots) \begin{pmatrix} \widetilde{H}_1 & H_0 & & \bigcirc & \\ \widetilde{H}_2 & H_1 & H_0 & & \\ \widetilde{H}_3 & H_2 & H_1 & H_0 & \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix} = (I - W, O, O, \ldots). \tag{2.3}$$

Performing an odd-even permutation of the block rows and block columns in the matrix equation (2.3) we find that

$$(G_0, G_0^3, \ldots \mid I, G_0^2, \ldots) \begin{pmatrix} T_1^{(0)} & W^{(0)} \\ Z^{(0)} & T_2^{(0)} \end{pmatrix} = (O, O, \ldots \mid I - W, O, \ldots) \tag{2.4}$$

where

$$T_1^{(0)} = \begin{pmatrix} H_1 & & & \bigcirc \\ H_3 & H_1 & & \\ H_5 & H_3 & H_1 & \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}, \quad T_2^{(0)} = \begin{pmatrix} \widetilde{H}_1 & & & \bigcirc \\ \widetilde{H}_3 & H_1 & & \\ \widetilde{H}_5 & H_3 & H_1 & \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix},$$

$$Z^{(0)} = \begin{pmatrix} H_0 & & & \bigcirc \\ H_2 & H_0 & & \\ H_4 & H_2 & H_0 & \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}, \quad W^{(0)} = \begin{pmatrix} \widetilde{H}_2 & H_0 & & & \bigcirc \\ \widetilde{H}_4 & H_2 & H_0 & & \\ \widetilde{H}_6 & H_4 & H_2 & H_0 & \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}.$$

Applying one step of block-Gaussian elimination to the 2×2 block matrix defined in (2.4), yields

$$(I, G_0^2, G_0^4, \ldots)(T_2^{(0)} - Z^{(0)} T_1^{(0)^{-1}} W^{(0)}) = (I - W, O, O \ldots). \tag{2.5}$$

It is interesting to point out (compare [2]) that the Schur complement $Q^{(1)} = T_2^{(0)} - Z^{(0)} T_1^{(0)^{-1}} W^{(0)}$ of $T_2^{(0)}$, which appears in (2.5), is such that $P^{(1)} = I - Q^{(1)}$ is a lower block Hessenberg matrix which, except for the first block column, has the block Toeplitz structure of the matrix $P = P^{(0)} = I - Q^{(0)}$ of (1.1) (Toeplitz-like structure). In other words, $Q^{(1)}$ is univocally determined by the blocks $\widetilde{H}_i^{(1)}$, $i \geq 1$, $H_i^{(1)}$, $i \geq 0$, defining the first two block columns of $Q^{(1)}$; in this way (2.5) can be rewritten as

$$(I, G_0^2, G_0^4, \ldots) \begin{pmatrix} \widetilde{H}_1^{(1)} & H_0^{(1)} & & \bigcirc & \\ \widetilde{H}_2^{(1)} & H_1^{(1)} & H_0^{(1)} & & \\ \widetilde{H}_3^{(1)} & H_2^{(1)} & H_1^{(1)} & H_0^{(1)} & \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix} = (I - W, O, O, \ldots). \tag{2.6}$$

We may recursively apply the same reduction (cyclic reduction) to the matrix equation (2.6) thus obtaining a sequence of matrices $\{Q^{(j)}\}$ such that $P^{(j)} = I - Q^{(j)}$ has the same lower block Hessenberg and Toeplitz-like structure of (1.1). Each matrix $Q^{(j)}$ is univocally determined by the blocks $\widetilde{H}_i^{(j)}$, $i \geq 1$, $H_i^{(j)}$, $i \geq 0$, defining its first two block columns. The matrices $P^{(j)}$, $j = 0, 1, \ldots$, have further properties as shown by the following

LEMMA 2.1. *All the matrices $P^{(j)} = I - Q^{(j)}$ recursively generated by applying cyclic reduction to the stochastic, irreducible and positive recurrent matrix $P^{(0)}$ of (1.1) are stochastic, irreducible and positive recurrent.*

Proof. It is sufficient to prove the lemma for $P^{(1)}$, then the proof can be completed by induction. Since $\mathbf{e}^T P^{(0)} = \mathbf{e}^T$, i.e., $\mathbf{e}^T T_1^{(0)} + \mathbf{e}^T Z^{(0)} = \mathbf{0}$ and $\mathbf{e}^T W^{(0)} + \mathbf{e}^T T_2^{(0)} = \mathbf{0}$, from (2.5) we have $\mathbf{e}^T(I - P^{(1)}) = \mathbf{e}^T(T_2^{(0)} - Z^{(0)} T_1^{(0)^{-1}} W^{(0)}) = \mathbf{e}^T T_2^{(0)} + \mathbf{e}^T T_1^{(0)} T_1^{(0)^{-1}} W^{(0)} = \mathbf{e}^T T_2^{(0)} + \mathbf{e}^T W^{(0)} = \mathbf{0}$, that is, $P^{(1)}$ is stochastic. Now, in order to prove the irreducibility of $P^{(1)}$, we first show that a reducible stochastic matrix M having a positive probability invariant vector $\mathbf{p}$ such that $M\mathbf{p} = \mathbf{p}$, $||\mathbf{p}|| = 1$, has at least a probability invariant vector having some null components. Indeed, without loss of generality, we may assume that $M = \begin{pmatrix} M_{1,1} & O \\ M_{2,1} & M_{2,2} \end{pmatrix}$, where $M_{1,1}$ and $M_{2,2}$ are square (possibly infinite) matrices and $M_{1,1}$ is irreducible. Let us partition the vector $\mathbf{p}$ as $\mathbf{p}^T = (\mathbf{p}_1^T, \mathbf{p}_2^T)^T$, according to the block structure of M. From the condition $M\mathbf{p} = \mathbf{p}$ we deduce that $M_{1,1}\mathbf{p}_1 = \mathbf{p}_1$, whence

$$\mathbf{e}^T M_{1,1} \mathbf{p}_1 = \mathbf{e}^T \mathbf{p}_1. \tag{2.7}$$

Since $\mathbf{p}_1 > \mathbf{0}$ and $\mathbf{e}^T M_{1,1} \leq \mathbf{e}^T$ (due to the stochasticity of M), we find that if a component of $\mathbf{e}^T M_{1,1}$ were less than 1, we would have $\mathbf{e}^T M_{1,1}\mathbf{p}_1 < \mathbf{e}^T \mathbf{p}_1$, which would contraddict (2.7). Therefore $\mathbf{e}^T M_{1,1} = \mathbf{e}^T$, i.e., $M_{1,1}$ is stochastic and $M_{2,1} = O$. Whence it follows that the vector $\mathbf{y} = (\mathbf{p}_1^T, \mathbf{0})^T / ||\mathbf{p}_1||$ is such that $M\mathbf{y} = \mathbf{y}$, $||\mathbf{y}|| = 1$. Now, since the vector made up by the even block components of $\boldsymbol{\pi}$ is a positive probability invariant vector of $P^{(1)}$, we deduce that, if $P^{(1)}$ were reducible, then there would exist a vector $\mathbf{x} \geq \mathbf{0}$ having some null components such that $P^{(1)}\mathbf{x} = \mathbf{x}$, $||\mathbf{x}|| = 1$. In this way the vector $\mathbf{y} = (\mathbf{y}_1^T, \mathbf{y}_2^T)^T$ defined by $\mathbf{y}_1 = -T_1^{(0)^{-1}} W^{(0)}\mathbf{x}$, $\mathbf{y}_2 = \mathbf{x}$, is such that $P^{(0)}\Pi^T\mathbf{y} = \Pi^T\mathbf{y}$, where Π is the odd-even permutation matrix. This follows from the relation $\Pi Q^{(0)} \Pi^T = \begin{pmatrix} T_1^{(0)} & W^{(0)} \\ Z^{(0)} & T_2^{(0)} \end{pmatrix} = \begin{pmatrix} I & O \\ Z^{(0)} T_1^{(0)^{-1}} & I \end{pmatrix} \begin{pmatrix} T_1^{(0)} & W^{(0)} \\ O & Q^{(1)} \end{pmatrix}$, $Q^{(1)} = T_2^{(0)} - Z^{(0)} T_1^{(0)^{-1}} W^{(0)}$ (compare (2.4)). This contraddicts the assumptions, since the positive recurrent matrix $P^{(0)}$ cannot have a probability invariant vector with some null components. In order to prove that $P^{(1)}$ is positive recurrent it is sufficient to show that there exists a positive invariant vector of $P^{(1)}$ with finite norm (compare [7], page 152). This vector is given by the even blocks of $\boldsymbol{\pi}$.

The following lemma relates the matrices $A_i^{(j)}$ generated by the cyclic reduction to the matrix $G_0^{2^j}$.

LEMMA 2.2. *(1) For every positive integer j, the matrix $G_j = G_0^{2^j}$ is the minimal nonnegative solution of the matrix equation*

$$G = \sum_{i=0}^{+\infty} G^i A_i^{(j)} \tag{2.8}$$

where $A_0^{(j)} = -H_0^{(j)}$, $A_1^{(j)} = I - H_1^{(j)}$, $A_i^{(j)} = -H_i^{(j)}$, $i \geq 2$, and $A_i^{(0)} = A_i$, $i \geq 0$.

(2) There exists $\lim_j G_j = G'$. Moreover $G' = \mathbf{g}\,\mathbf{e}^T$, where $\mathbf{g} = (g_1, \ldots, g_k)^T$ is the nonnegative vector such that $\mathbf{e}^T\mathbf{g} = 1$, $G_0\mathbf{g} = \mathbf{g}$.

(3) The sequence of matrices $\{E_j\}_j$, where E_j is such that $G_j = G' + E_j$, for $j \geq 0$, converges quadratically to zero.

Proof. By extending (2.3)–(2.6) to the j-th step of cyclic reduction it follows that $G_j = G_0^{2^j}$ is a nonnegative solution of (2.8). From the irreducibility and positive recurrence of the stochastic matrix $P^{(j)} = I - Q^{(j)}$, obtained at each step j of cyclic reduction, the nonnegative solution G_j of (2.8) is unique. The convergence results stated by the lemma follow since G_0 is a stochastic matrix and the only eigenvalue having the largest modulus is 1 [15]. Indeed, let $G_0 = \mathbf{g}\ \mathbf{e}^T + E_0$ so that $\mathbf{e}^T E_0 = \mathbf{0}$, $E_0 \mathbf{g} = \mathbf{0}$, and the spectral radius of E_0 is strictly less than 1. Then $G_j = \mathbf{g}\ \mathbf{e}^T + E_j$, where $E_j = E_0^{2^j}$; whence E_j tends quadratically to zero and $lim_j\, G_j = \mathbf{g}\ \mathbf{e}^T$.

In [2] a functional formulation of the cyclic reduction method has been given. More precisely, the matrix power series

$$\begin{aligned} \varphi^{(j)}(z) &= \sum_{i=0}^{+\infty} A_i^{(j)} z^i, \\ \widetilde{\varphi}^{(j)}(z) &= \sum_{i=0}^{+\infty} \widetilde{A}_{i+1}^{(j)} z^i, \end{aligned} \tag{2.9}$$

which converge for $|z| \le 1$, have been introduced. It has been observed that the equations which relate the blocks $A_i^{(j+1)}$ with the blocks $A_i^{(j)}$, $i \ge 0$, and the blocks $\widetilde{A}_i^{(j+1)}$ with the blocks $\widetilde{A}_i^{(j)}$, $i \ge 1$, can be formally expressed in functional form as

$$\varphi^{(j+1)}(z) = z\varphi_{odd}^{(j)}(z) + \varphi_{even}^{(j)}(z)\big(I - \varphi_{odd}^{(j)}(z)\big)^{-1}\varphi_{even}^{(j)}(z) \tag{2.10a}$$

$$\widetilde{\varphi}^{(j+1)}(z) = \widetilde{\varphi}_{odd}^{(j)}(z) + \varphi_{even}^{(j)}(z)(I - \varphi_{odd}^{(j)}(z))^{-1}\widetilde{\varphi}_{even}^{(j)}(z) \tag{2.10b}$$

where

$$\begin{cases} \varphi_{even}^{(j)}(z) = \sum_{i=0}^{+\infty} A_{2i}^{(j)} z^i, & \varphi_{odd}^{(j)}(z) = \sum_{i=0}^{+\infty} A_{2i+1}^{(j)} z^i \\ \widetilde{\varphi}_{even}^{(j)}(z) = \sum_{i=0}^{+\infty} \widetilde{A}_{2(i+1)}^{(j)} z^i, & \widetilde{\varphi}_{odd}^{(j)}(z) = \sum_{i=0}^{+\infty} \widetilde{A}_{2i+1}^{(j)} z^i. \end{cases} \tag{2.11}$$

These functional relations not only allow us to express in compact form the recurrence which this method is based on, but also provide the basic tool on which the efficient computation of the matrices $A_i^{(j)}$ relies. For the computational and numeric issues we refer the reader to the next sections 4 and 5.

From the functional relations it is immediate to obtain the explicit expression for $A_0^{(j+1)}$, i.e.,

$$A_0^{(j+1)} = A_0^{(j)}(I - A_1^{(j)})^{-1} A_0^{(j)}. \tag{2.12}$$

THEOREM 2.3. *We have*

$$A_0^{(j)}\left(I - \sum_{i=1}^{+\infty} A_i^{(j)}\right)^{-1} = G' - \left(E_{1j} - \sum_{i=1}^{+\infty} E_{ij} A_i^{(j)}\right)\left(I - \sum_{i=1}^{+\infty} A_i^{(j)}\right)^{-1} \tag{2.13}$$

where $E_{ij} = G_j^i - G' = G_0^{i2^j} - G'$. *Moreover, 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'.

Proof. Since the matrices $P^{(j)}$ are irreducible and positive recurrent, the matrix

$$S^{(j)} = \sum_{l=1}^{+\infty}\sum_{i=l}^{+\infty} G_j^{i-l} A_i^{(j)} \tag{2.14}$$

has spectral radius less than 1 (compare (2.2), (1.4) and [15]). Therefore, since $\sum_{i=1}^{+\infty} A_i^{(j)} \leq S^{(j)}$, for the Perron-Frobenius theorem [18] it holds that $r(\sum_{i=1}^{+\infty} A_i^{(j)}) \leq r(S^{(j)}) < 1$, that is, the matrix $I - \sum_{i=1}^{+\infty} A_i^{(j)}$ is nonsingular. By replacing G_j^i with $G' + E_{ij}$ in (2.8) we arrive at (2.13). The quadratic convergence holds for Lemma 2.2, part (3).

The matrix $(I - \sum_{i=1}^{+\infty} A_i^{(j)})^{-1}$ is bounded above in particular if there exists $P' = lim_j\, P^{(j)}$ and the matrix P' is positive recurrent. In fact, if $P' = lim_j\, P^{(j)}$ is positive recurrent, then there exists $S' = lim_j\, S^{(j)}$ and $1 > r(S') \geq r(\sum_{i=1}^{+\infty} A_i')$, since $S' \geq \sum_{i=1}^{+\infty} A_i'$, where $A_i' = lim_j\, A_i^{(j)}$.

It is intereseting to point out that, in the case where the matrix P is block tridiagonal (as in the case of QBD processes), all the matrices $P^{(j)}$ generated by the cyclic reduction are block tridiagonal and the functions $\varphi^{(j)}(z)$ and $\widetilde{\varphi}^{(j)}(z)$ are matrix polynomials of degree 2 and 1, respectively. Thus the functional relations (2.10) become

$$\begin{aligned} \varphi^{(j+1)}(z) &= zA_1^{(j)} + (A_0^{(j)} + zA_2^{(j)})(I - A_1^{(j)})^{-1}(A_0^{(j)} + zA_2^{(j)}) \\ \widetilde{\varphi}^{(j+1)}(z) &= \widetilde{A}_1^{(j)} + (A_0^{(j)} + zA_2^{(j)})(I - A_1^{(j)})^{-1}\widetilde{A}_2^{(j)} \end{aligned} \tag{2.15}$$

Similar relations have been obtained by Latouche and Ramaswami in [11].

3. Further convergence properties. Under suitable assumptions, further useful convergence results can be proved. From (2.10a) we easily arrive at the following results

LEMMA 3.1. *For every integer $j \geq 0$, the matrix $I - A_1^{(j)}$ is nonsingular and the nonnegative matrix sequence $\{A_1^{(j)}\}_j$ is nondecreasing. Moreover, there exists $lim_j\, A_1^{(j)} = A_1'$.*

Proof. By following the same argument used in the proof of Theorem 2.3, since $r(S^{(j)}) < 1$ (compare (2.14)), it follows that for any j, the matrix $I - A_1^{(j)}$ is nonsingular. Let us prove that the sequence of matrices $\{A_1^{(j)}\}_j$ is monotonically convergent. Comparing the coefficients of z in both the sides of (2.10a), it follows that, for every $j \geq 0$, $A_1^{(j+1)} = A_1^{(j)} + C^{(j)}$, where $C^{(j)}$ is the coefficient of z in the series $\varphi_{even}^{(j)}(z)\big(I - \varphi_{odd}^{(j)}(z)\big)^{-1}\varphi_{even}^{(j)}(z)$. Since the coefficients of $\varphi_{even}^{(j)}(z)$ and $(I - \varphi_{odd}^{(j)}(z))^{-1}$ are nonnegative, we find that $C^{(j)} \geq 0$, whence $A_1^{(j+1)} \geq A_1^{(j)}$. Moreover, since $A_1^{(j)}$ has nonnegative entries not greater than 1, it follows that there exists $lim_j\, A_1^{(j)} = A_1'$.

LEMMA 3.2. *For every convergent subsequence $\{A_0^{(\sigma_j)}\}_j$ of the sequence of matrices $\{A_0^{(j)}\}_j$, it holds $lim_j\, A_0^{(\sigma_j)} \neq O$.*

Proof. Let $\{A_0^{(\sigma_j)}\}_j$ be a convergent subsequence of $\{A_0^{(j)}\}_j$. We prove that, if $lim_j\, A_0^{(\sigma_j)} = O$, then $lim_j\, G_{\sigma_j} = O$, which would contraddict part 2 of Lemma 2.2. In order to arrive at the condition $lim_j\, G_{\sigma_j} = O$, we prove that for any $\epsilon > 0$ there exists a sequence of nonnegative integers $m(j)$ such that for any j it holds

$$\mathbf{e}^T G_j - \epsilon \mathbf{e}^T \leq \mathbf{e}^T - \mathbf{e}^T (M^{(j)})^{m(j)}, \tag{3.1}$$

where $M^{(j)} = \sum_{h=1}^{+\infty} A_h^{(j)}$. First we introduce the matrix sequence defined for $i, j \geq 0$,

$$\begin{cases} X_{0,j} = O \\ X_{i+1,j} = \sum_{n=0}^{+\infty} X_{i,j}^n A_n^{(j)}, \ i \geq 0 \end{cases} \tag{3.2}$$

such that $G_j = lim_i\, X_{i,j}$. Then we observe that

$$\mathbf{e}^T X_{i,j} \leq \mathbf{e}^T - \mathbf{e}^T M^{(j)^i}, \tag{3.3}$$

as it can be easily proved by induction on i (we leave it to the reader). Since $0 \leq \mathbf{e}^T M^{(j)^i} \leq \mathbf{e}^T$, there exists an increasing sequence of nonnegative numbers $i_n = i_n(j)$ such that the subsequence $\{M^{(j)^{i_n}}\}_n$ is convergent. Therefore, from (3.3), we find that the matrix $G_j = lim_n\, X_{i_n,j}$ satisfies the inequality

$$\mathbf{e}^T G_j \leq \mathbf{e}^T - \mathbf{e}^T \lim_n M^{(j)^{i_n}}.$$

Hence, fixed $\epsilon > 0$, there exists a positive integer $\nu = \nu(j)$ such that (3.1) holds with $m(j) = i_{\nu(j)}(j)$. If the sequence $\{A_0^{(\sigma_j)}\}_j$ converges to the null matrix, then the sequence of matrices $\left\{M^{(\sigma_j)}\right\}_j$ converges to a stochastic matrix, since $\left\{\sum_{h=0}^{+\infty} A_h^{(\sigma_j)}\right\}_j$ is stochastic. Whence, by replacing j with σ_j in (3.1) and by taking limits for $j \longrightarrow +\infty$, we conclude that $lim_j\, G_{\sigma_j} = O$.

Remark 3.3. We observe that from Lemma 3.2 it follows that, if $\sum_{i=1}^{+\infty} A_i^{(j)}$ is irreducible and if no accumulation point of $\sum_{i=1}^{+\infty} A_i^{(j)}$ is reducible, then the matrix $(I - \sum_{i=1}^{+\infty} A_i^{(j)})^{-1}$ is bounded above for any j. In order to prove this we show that, under these assumptions, $r(\sum_{i=1}^{+\infty} A_i^{(j)}) \leq \gamma < 1$ for any j. Indeed, let $\gamma = lim\, sup_j\, r(\sum_{i=1}^{+\infty} A_i^{(j)})$ and let σ_j be an increasing sequence of nonnegative numbers such that $\gamma = lim_j\, r(\sum_{i=1}^{+\infty} A_i^{(\sigma_j)}) = r(lim_j \sum_{i=1}^{+\infty} A_i^{(\sigma_j)})$ and $lim_j\, A_0^{(\sigma_j)} = A_0'$ (observe that σ_j exists since $\sum_{i=1}^{+\infty} A_i^{(j)}$ and $A_0^{(j)}$ belong to a compact set for any j). In this way, since $\sum_{i=0}^{+\infty} A_i^{(j)}$ is stochastic for any j, we have that $A' + A_0'$ is stochastic, where $A' = lim_j \sum_{i=1}^{+\infty} A_i^{(\sigma_j)}$. Since A' is irreducible and, for Lemma 3.2, $A_0' \geq O$, $A_0' \neq O$, from the Perron-Frobenius theorem [18] we deduce that $1 = r(A' + A_0') > r(A') = \gamma$.

In the light of Lemma 3.1, if $A_1^{(j)}$ is irreducible for $j = j_0$, then the conditions of Remark 3.3 on the irreducibility of $\sum_{i=1}^{+\infty} A_i^{(j)}$ and of the accumulation points of the latter sequence are satisfied.

LEMMA 3.4. *If the matrix G_0 is irreducible, then the sequence*

$$\theta_j = \mathbf{e}^T A_0^{(j)} (I - A_1^{(j)})^{-1} \mathbf{g} \tag{3.4}$$

is such that $lim_j\, \theta_j = 1$.

Proof. Since $0 \leq \theta_j \leq 1$ for every integer j, it follows that the sequence θ_j has convergent subsequences, whose limits belong to the set $[0, 1]$. We first prove that 0 cannot be a limit. Let $\{\theta_{\sigma_j}\}_j$ be a convergent subsequence. Assume without loss of generality that the sequence $\{A_0^{(\sigma_j)}\}_j$ is also convergent (the matrices $A_0^{(j)}$ belong to a compact set) and set $A_0' = lim_j\, A_0^{(\sigma_j)}$. If $lim_j\, \theta_{\sigma_j} = 0$, then from (3.4) it holds that $lim_j\, A_0^{(\sigma_j)} = 0$ since the vector $\mathbf{g}$ is strictly positive, due to the irreducibility of G_0, and the nonnegative matrix $(I - A_1^{(j)})^{-1}$ has diagonal entries greater than 1. This contraddicts Lemma 3.2. Now we show that any number θ such that $0 < \theta < 1$ cannot be the limit of any subsequence $\{\theta_{\sigma_j}\}_j$. From (2.8) it follows that

$$A_0^{(\sigma_j)} = G_{\sigma_j} - G_{\sigma_j} \sum_{i=1}^{+\infty} G_{\sigma_j}^{i-1} A_i^{(\sigma_j)};$$

by substituting $G_{\sigma_j} = \mathbf{g}\ \mathbf{e}^T + E_{\sigma_j}$ (compare Lemma 2.2) we obtain

$$A_0^{(\sigma_j)} = \mathbf{g}\ \mathbf{e}^T A_0^{(\sigma_j)} + E_{\sigma_j} \sum_{i=1}^{+\infty} G_{\sigma_j}^{i-1} A_i^{(\sigma_j)}.$$

Hence, by taking limits for $j \to \infty$, we find that $A_0' = \mathbf{g}\ \mathbf{e}^T A_0' = \mathbf{g}\ \mathbf{v}^T$, where $\mathbf{v}^T = \mathbf{e}^T A_0'$. From the recursive equation (2.12) it follows that the sequence $\mathbf{e}^T A_0^{(\sigma_j+1)}$ is convergent and

$$\lim_j \mathbf{e}^T A_0^{(\sigma_j+1)} = \theta \mathbf{v}^T.$$

It can be easily shown by induction on h that for every integer $h \geq 0$ there exists

$$\lim_j \mathbf{e}^T A_0^{(\sigma_j+h)} = \theta^{2^h-1} \mathbf{v}^T.$$

Hence, for every nonnegative integer h there exists a nonnegative integer j_h such that

$$\forall j \geq j_h \qquad ||\mathbf{e}^T A_0^{(\sigma_j+h)} - \theta^{2^h-1}\mathbf{v}^T|| < \theta^{2^h-1}. \tag{3.5}$$

Consider the strictly increasing sequence of nonnegative integers

$$\begin{cases} \nu_0 = \sigma_{j_0} \\ \nu_h = h + min\{\sigma_j : j \geq j_h, \sigma_j > \nu_{h-1}\}, \quad h \geq 1. \end{cases}$$

It is readly seen, from (3.5), that the sequence $\mathbf{e}^T A_0^{(\nu_h)}$ is such that $\mathbf{e}^T A_0^{(\nu_h)} \leq \mathbf{u}^T \theta^{2^h-1}$ for every nonnegative integer h, where $\mathbf{u}^T$ is a constant nonnegative vector. Whence the sequence of matrices $A_0^{(\nu_h)}$ converges to the null matrix, giving a contraddiction, again for Lemma 3.2. Therefore, the unique accumulation point of the sequence $\{\theta_j\}_j$ is 1; hence the sequence $\{\theta_j\}_j$ converges to 1.

THEOREM 3.5. *If the matrix G_0 is irreducible, then the sequence of matrices $P^{(j)}$ converges to a stochastic matrix P' having the M/G/1 structure (1.1). Moreover,*

$$\lim_j A_0^{(j)}(I - A_1^{(j)})^{-1} = G'$$

and the matrix P', defined by the blocks $\widetilde{A}_i'$, $i \geq 1$, of its first block column and by the block entries A_i', $i \geq 0$, of its second block column, is such that

$$\widetilde{A}_i' = A_i' = O, \ i = 2, 3, \ldots$$

Proof. In order to prove the convergence of the block entries $\widetilde{A}_i^{(j)}$, $i \geq 1$, we follow [2]. For this purpose, denote $\boldsymbol{\pi}^{(j)} = (\boldsymbol{\pi}_0^T, \boldsymbol{\pi}_{2^j}^T, \boldsymbol{\pi}_{2\cdot 2^j}^T, \ldots)^T$, where $\boldsymbol{\pi} = \boldsymbol{\pi}^{(0)}$ is the probability invariant vector defined in (1.2). In this way, due to the cyclic odd-even permutation, $\boldsymbol{\pi}^{(j)}$ is the probability invariant vector associated with $P^{(j)}$. For the positive recurrence of $P = P^{(0)}$, it follows that $\boldsymbol{\pi}_0 > 0$. Moreover, since $||\boldsymbol{\pi}|| = 1$, we have $lim_j \boldsymbol{\pi}_{i\cdot 2^j} = \mathbf{0}$, for $i \geq 1$. Whence, from the condition $(I - P^{(j)})\boldsymbol{\pi}^{(j)} = \mathbf{0}$, we deduce that $lim_j \widetilde{A}_i^{(j)} = O$, for $i \geq 2$. Let us now analyze the convergence of the block entries $A_i^{(j)}$, $i \geq 0$. Let $\{A_0^{(\sigma_j)}\}_j$ be a convergent subsequence of $\{A_0^{(j)}\}_j$. Then for Lemma 3.4 we have

$$\lim_j \mathbf{e}^T A_0^{(\sigma_j)} (I - A_1^{(\sigma_j)})^{-1} \mathbf{g} = 1.$$

Hence, for the stochasticity of the matrix $\sum_{i=0}^{+\infty} A_i^{(\sigma_j)}$, it holds

$$\lim_j \mathbf{e}^T \sum_{i=2}^{+\infty} A_i^{(\sigma_j)}(I - A_1^{(\sigma_j)})^{-1}\mathbf{g} = 0.$$

Since the vector $\mathbf{g}$ is positive and the diagonal entries of the matrices $(I - A_1^{(\sigma_j)})^{-1}$ are greater than or equal to 1, then it holds

$$\lim_j \mathbf{e}^T \sum_{i=2}^{+\infty} A_i^{(\sigma_j)} = \mathbf{0}.$$

Therefore, the unique accumulation points of the sequence $\{P^{(j)}\}_j$ are stochastic matrices P' such that $A_i' = O$, $i = 2, 3, \ldots$. Moreover, for the convergence of the sequence G_j and from (2.8), we obtain that $A_0' = G' - G'A_1'$ and that $G' = \lim_j A_0^{(j)}(I - A_1^{(j)})^{-1}$.

4. Computing the matrix G. The convergence results given by Theorem 2.3 can be used to derive two efficient and stable algorithms for computing the solution G_0 of the nonlinear matrix equation (1.3). In the following, the matrix $B = |A|$, where $A = \{a_{ij}\}_{ij}$, is the matrix defined by the entries $b_{ij} = |a_{ij}|$.

It is readily seen that, for every $j \geq 0$, it holds

$$G_j = \left(\sum_{i=0}^{+\infty} G_j^{2i} A_{2i}^{(j)}\right)\left(I - \sum_{i=0}^{+\infty} G_j^{2i} A_{2i+1}^{(j)}\right)^{-1}, \tag{4.1}$$

where G_j is the unique nonnegative solution of the matrix equation (2.8). Observe that, for every j, the matrix $I - \sum_{i=0}^{+\infty} G_j^{2i} A_{2i+1}^{(j)}$ is nonsingular, since $\sum_{i=0}^{+\infty} G_j^{2i} A_{2i+1}^{(j)} \leq S^{(j)}$ (compare (2.14), (2.2), (1.4) and [15]). Moreover, by replacing G_j^2 with G_{j+1} (compare Lemma 2.2) in (4.1), we arrive at the following recursive relation for the sequence of matrices G_j

$$G_j = \left(\sum_{i=0}^{+\infty} G_{j+1}^{i} A_{2i}^{(j)}\right)\left(I - \sum_{i=0}^{+\infty} G_{j+1}^{i} A_{2i+1}^{(j)}\right)^{-1}, \quad j \geq 0. \tag{4.2}$$

Equations (4.2), together with the convergence properties stated by Theorem 2.3, allow us to derive the following algorithm for computing the matrix G_0, either in the hypotheses of Theorem 3.5 or in the case where the entries of the matrices $(I - \sum_{i=1}^{+\infty} A_i^{(j)})^{-1}$ are bounded by a constant (compare Remark 3.3):

Algorithm 4.1.

1. Apply cyclic reduction by means of (2.10a) and compute the matrices $A_i^{(j)}$, $j = 1, 2, \ldots, q$, until one of the following conditions is satisfied

$$|R^{(q)} - R^{(q-1)}| < \epsilon E, \ \mathbf{e}^T(I - A_0^{(q)}(I - A_1^{(q)})^{-1}) < \epsilon \mathbf{e}^T,$$

where, at each step j, $R^{(j)} = A_0^{(j)}(I - \sum_{i=1}^{+\infty} A_i^{(j)})^{-1}$, $\epsilon > 0$ is fixed and E is the $k \times k$ matrix having all the entries equal to 1.

2. Compute an approximation of G_0 by replacing G_q with $R^{(q)}$ in (4.2) for $j = q-1$, and then apply back substitution in (4.2) for $j = q-1, q-2, \ldots, 0$, until the approximation of G_0 is obtained.

In order to apply the back substitution stage of Algorithm 4.1, it is needed to store the blocks $A_i^{(j)}$ for $j \le q$ computed at stage 1. This may need a large amount of memory in the case where the matrices $A_i^{(j)}$ are negligible only for large values of i. This problem is overcome by the next algorithm where back substitution is implicitely performed in the cyclic reduction stage and no storage is required.

Observe that the solution G_0 of (1.3) satisfies the matrix equation

$$G_0(I, G_0, G_0^2, \ldots)\begin{pmatrix} \widehat{H}_1 & H_0 & & \bigcirc \\ \widehat{H}_2 & H_1 & H_0 & \\ \widehat{H}_3 & H_2 & H_1 & H_0 \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix} = (A_0, O, O, \ldots) \tag{4.3}$$

where $H_0 = -A_0$, $H_1 = I - A_1$, $H_i = -A_i$, $i \ge 2$, $\widehat{H}_1 = I - A_1$, $\widehat{H}_i = -A_i$, $i \ge 2$. By recursively applying cyclic reduction we obtain (as it can be easily seen by extending (2.3)–(2.6) to equation (4.3)) the sequence of matrix equations

$$G_0(I, G_j, G_j^2, \ldots)\begin{pmatrix} \widehat{H}_1^{(j)} & H_0^{(j)} & & \bigcirc \\ \widehat{H}_2^{(j)} & H_1^{(j)} & H_0^{(j)} & \\ \widehat{H}_3^{(j)} & H_2^{(j)} & H_1^{(j)} & H_0^{(j)} \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix} = (A_0, O, O, \ldots), \quad j \ge 1, \tag{4.4}$$

where $G_j = G_0^{2^j}$ and the block entries $H_i^{(j)}$ coincide with the block entries of the second block column of the matrix $Q^{(j)}$ of section 2. The matrices $\widehat{A}_1^{(j)} = I - \widehat{H}_1^{(j)}$, $\widehat{A}_i^{(j)} = -\widehat{H}_i^{(j)}$, $i \ge 2$, $A_0^{(j)} = -H_0^{(j)}$, $A_1^{(j)} = I - H_1^{(j)}$, $A_i^{(j)} = -H_i^{(j)}$, $i \ge 2$, can still be represented by the functional equations (2.10b) and (2.10a), respectively, where in (2.10b) the functions $\widetilde{\varphi}^{(j)}(z)$ must be replaced by $\widehat{\varphi}^{(j)}(z) = \sum_{i=0}^{+\infty} \widehat{A}_{i+1}^{(j)} z^i$, with $\widehat{\varphi}^{(0)}(z) = \sum_{i=0}^{+\infty} A_{i+1} z^i$, so that (2.10b) is replaced by

$$\widehat{\varphi}^{(j+1)}(z) = \widehat{\varphi}_{odd}^{(j)}(z) + \varphi_{even}^{(j)}(z)\left(I - \varphi_{odd}^{(j)}(z)\right)^{-1} \widehat{\varphi}_{even}^{(j)}(z), \tag{4.5}$$

for $\widehat{\varphi}_{odd}^{(j)}(z)) = \sum_{i=0}^{+\infty} \widehat{A}_{2i+1}^{(j)} z^i$ and $\widehat{\varphi}_{even}^{(j)}(z) = \sum_{i=0}^{+\infty} \widehat{A}_{2(i+1)}^{(j)} z^i$. Since, from (4.5) it holds $\mathbf{e}^T \widehat{\varphi}^{(j)}(1) = \mathbf{e}^T \widehat{\varphi}^{(0)}(1) = \mathbf{e}^T - \mathbf{e}^T A_0$, then we have $\mathbf{e}^T \sum_{i=1}^{+\infty} \widehat{A}_i^{(j)} = \mathbf{e}^T - \mathbf{e}^T A_0$, for every $j \ge 0$. Moreover, if the matrix $\sum_{i=0}^{+\infty} G_j^i \widehat{H}_{i+1}^{(j)} = I - \sum_{i=0}^{+\infty} G_j^i \widehat{A}_{i+1}^{(j)}$ is nonsingular, then it holds

$$G_0 = A_0 \left(I - \sum_{i=0}^{+\infty} G_j^i \widehat{A}_{i+1}^{(j)} \right)^{-1}. \tag{4.6}$$

Suppose that, for every $j \ge 0$, the matrix $I - \sum_{i=0}^{+\infty} G_j^i \widehat{A}_{i+1}^{(j)}$ is nonsingular. Observe that, if the matrix A_0 has no null columns, then the matrix $I - \sum_{i=0}^{+\infty} G_j^i \widehat{A}_{i+1}^{(j)}$ is nonsingular for any j. Then, the following algorithm for the computation of the matrix G_0 can be carried out:

Algorithm 4.2.

1. Apply cyclic reduction to (4.3) by means of (2.10a) and (4.5), for $\varphi^{(0)}(z) = \sum_{i=0}^{+\infty} A_i z^i$, $\widehat{\varphi}^{(0)}(z) = \sum_{i=0}^{+\infty} A_{i+1} z^i$, and compute the matrices $A_i^{(j)}$,$\widehat{A}_i^{(j)}$, $j = 1, 2, \ldots, q$, until one of the following conditions is satisfied

$$|R^{(q)} - R^{(q-1)}| < \epsilon E, \ \mathbf{e}^T (I - A_0^{(q)}(I - A_1^{(q)})^{-1}) < \epsilon \mathbf{e}^T, \tag{4.7}$$

where, at each step j, $R^{(j)} = A_0^{(j)}(I - \sum_{i=1}^{+\infty} A_i^{(j)})^{-1}$, $\epsilon > 0$ is fixed and E is the $k \times k$ matrix having all the entries equal to 1.

2. Compute an approximation of G_0 by replacing G_q with $R^{(q)}$ in (4.6) for $j = q$.

We observe that both Algorithms 4.1 and 4.2 only involve multiplications and additions of nonnegative matrices and inversions of M-matrices [18], i.e., matrices having nonpositive off-diagonal entries and nonnegative inverse. These computations, if the diagonal adjustment technique [8] is used, can be reduced to performing additions of positive numbers, multiplications and divisions. This makes such computations strongly numerically stable, i.e., the relative rounding errors in the results can be bounded componentwise.

By following the proof of Theorem 3.5 to show the convergence to zero of the blocks $\widetilde{A}_i^{(j)}$, $i \geq 2$, for $j \to +\infty$, we now prove the following

THEOREM 4.1. *For the matrices $\widehat{A}_i^{(j)} = -\widehat{H}_i^{(j)}$, $i \geq 2$, of (4.4) it holds $lim_j \widehat{A}_i^{(j)} = O$, for any $i \geq 2$.*

Proof. The probability invariant vector $\boldsymbol{\pi}$ of (1.2) satisfies the following equation

$$\begin{pmatrix} I - A_1 & -A_0 & & \bigcirc & \\ -A_2 & I - A_1 & -A_0 & & \\ -A_3 & -A_2 & I - A_1 & -A_0 & \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} \boldsymbol{\pi}_1 \\ \boldsymbol{\pi}_2 \\ \boldsymbol{\pi}_3 \\ \vdots \end{pmatrix} = \begin{pmatrix} \widetilde{A}_2 \\ \widetilde{A}_3 \\ \widetilde{A}_4 \\ \vdots \end{pmatrix} \boldsymbol{\pi}_0,$$

where the block matrix in the left-hand size is the same block matrix of (4.3). Thus, applying j steps of cyclic reduction to the above system we find that

$$\begin{pmatrix} \widehat{H}_1^{(j)} & H_0^{(j)} & & \bigcirc & \\ \widehat{H}_2^{(j)} & H_1^{(j)} & H_0^{(j)} & & \\ \widehat{H}_3^{(j)} & H_2^{(j)} & H_1^{(j)} & H_0^{(j)} & \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} \boldsymbol{\pi}_1 \\ \boldsymbol{\pi}_{2^j+1} \\ \boldsymbol{\pi}_{2 \cdot 2^j+1} \\ \vdots \end{pmatrix} = \begin{pmatrix} B_1^{(j)} \\ B_2^{(j)} \\ B_3^{(j)} \\ \vdots \end{pmatrix} \boldsymbol{\pi}_0,$$

where $B_i^{(j)}$ is such that $B_i^{(j)} \geq O$, for every $i \geq 1$. Since $lim_j \boldsymbol{\pi}_{i \cdot 2^j+1} = \mathbf{0}$, for $i \geq 1$, we have

$$\lim_j \left(\begin{pmatrix} B_1^{(j)} \\ B_2^{(j)} \\ B_3^{(j)} \\ \vdots \end{pmatrix} \boldsymbol{\pi}_0 - \begin{pmatrix} \widehat{H}_1^{(j)} \\ \widehat{H}_2^{(j)} \\ \widehat{H}_3^{(j)} \\ \vdots \end{pmatrix} \boldsymbol{\pi}_1 \right) = \lim_j \left(\begin{pmatrix} B_1^{(j)} \\ B_2^{(j)} \\ B_3^{(j)} \\ \vdots \end{pmatrix} \boldsymbol{\pi}_0 + \begin{pmatrix} \widehat{A}_1^{(j)} - I \\ \widehat{A}_2^{(j)} \\ \widehat{A}_3^{(j)} \\ \vdots \end{pmatrix} \boldsymbol{\pi}_1 \right) = \mathbf{0}.$$

Therefore, for the conditions $\boldsymbol{\pi}_0 > 0$, $\boldsymbol{\pi}_1 > 0$, $B_i^{(j)} \geq O$, $\widehat{A}_i^{(j)} \geq O$, we deduce that $lim_j B_i^{(j)} = lim_j \widehat{A}_i^{(j)} = O$, for $i \geq 2$.

The result of Theorem 4.1 suggests to add to the above stop conditions (4.7) the following one

$$\mathbf{e}^T(I - A_0(I - \widehat{A}_1^{(q)})^{-1}) < \epsilon \mathbf{e}^T. \tag{4.8}$$

If this condition is verified then the matrix G_0 can be readily approximated by $A_0(I - \widehat{A}_1^{(q)})^{-1}$.

In the case of QBD processes, where P is a block tridiagonal matrix, the matrix equation (1.3) becomes $G = A_0 + GA_1 + G^2 A_2$ and Algorithm 4.2 can be rewritten as follows (compare (2.15))

Algorithm 4.3. *QBD processes.*

1. Apply cyclic reduction to (4.3) by means of the following formulae

$$\begin{aligned}
A_0^{(j+1)} &= A_0^{(j)}(I - A_1^{(j)})^{-1}A_0^{(j)}\\
A_1^{(j+1)} &= A_1^{(j)} + A_0^{(j)}(I - A_1^{(j)})^{-1}A_2^{(j)} + A_2^{(j)}(I - A_1^{(j)})^{-1}A_0^{(j)}\\
A_2^{(j+1)} &= A_2^{(j)}(I - A_1^{(j)})^{-1}A_2^{(j)}\\
\widehat{A}_1^{(j+1)} &= \widehat{A}_1^{(j)} + A_0^{(j)}(I - A_1^{(j)})^{-1}\widehat{A}_2^{(j)}
\end{aligned}$$

with $A_0^{(0)} = A_0$, $A_1^{(0)} = A_1$, $A_2^{(0)} = A_2, \widehat{A}_1^{(0)} = A_1$, for $j = 1, 2, \ldots, q$, until one of the following two conditions is satisfied $\mathbf{e}^T(I - A_0^{(q)}(I - A_1^{(q)})^{-1}) < \epsilon\mathbf{e}^T$, $\mathbf{e}^T(I - A_0^{(q)}(I - \widehat{A}_1^{(q)})^{-1}) < \epsilon\mathbf{e}^T$.

2. Compute an approximation $\widetilde{G}$ of G_0 by means of $\widetilde{G} = A_0(I - \widehat{A}_1^{(q)})^{-1}$.

Observe that in the QBD case it holds $\widehat{A}_2^{(j)} = A_2^{(j)}$, this leads to a saving of the computational cost.

Each step of Algorithm 4.3 requires 6 matrix multiplications and one matrix inversion. The algorithm of [11] is similar to Algorithm 4.3 but it requires 8 matrix multiplications and one matrix inversion per step.

5. Implementation of Algorithm 4.2 by means of FFT-based techniques. In this section we give a detailed description of the implementation of Algorithm 4.2 that we sketched in section 4.

Accordingly to the customary way of dealing with infinite sequences of blocks ([15], [9]), we truncate the sequences $\{A_i^{(j)}\}_i$ $\{\widehat{A}_i^{(j)}\}_i$ to the index m_j and $\widehat{m}_j$, respectively, such that $\mathbf{e}^T(I - \sum_{i=0}^{m_j} A_i^{(j)}) < \epsilon\mathbf{e}^T$ and $\mathbf{e}^T(I - A_0 - \sum_{i=1}^{\widehat{m}_j}\widehat{A}_i^{(j)}) < \epsilon\mathbf{e}^T$, where ϵ is an upper bound to the machine precision of the floating point arithmetic used in the computation, and is related to the accuracy of the computed approximation of G_0. In other words, we consider only those blocks which make the matrices $\sum_{i=0}^{m_j} A_i^{(j)}$ and $A_0 + \sum_{i=1}^{\widehat{m}_j}\widehat{A}_i^{(j)}$ numerically stochastic with respect to ϵ. A nonnegative matrix A is said *numerically stochastic with respect to* ϵ, or ϵ*-stochastic*, if $\mathbf{e}^T(I - A) < \epsilon\mathbf{e}^T$. In this way relations (2.10a) and (4.5) can be rewritten in terms of matrix polynomials, in the variable z, instead of matrix power series and all the operations involved in (2.10a) and (4.5) are thus reduced to multiplying matrix polynomials and inverting a matrix polynomial modulo z^m, for a suitable m.

This kind of truncation seems to be quite natural since it fits in with the floating point arithmetic used in the computation in the sense that only the terms which cannot add information to the problem are neglected. A different meaning has the truncation used in [12] where the user must replace the infinite matrix (1.1) with a finite one. The determination of the cutting level in the latter case constitutes a nontrivial problem. It is not easy to recover an a-priori estimate of this cutting level. Even though this value could be dynamically determined by a suitable modification of the algorithm, its value should be at least greater than or equal to m_0 otherwise some significant input data would be lost.

We now introduce the basic tools on which the fast version of our algorithm is based, by extending to matrix power series the known FFT-based techniques for dealing with scalar power series (see [3]).

The use of FFT techniques allows us to reduce the cost of the j-th step of cyclic reduction from $O(k^3 n_j^2)$ operations (needed if the customary way of dealing with matrix power series is used) to $O(k^3 n_j + k^2 n_j \, log \, n_j)$ operations, where $n_j = max\ \{m_j, \widehat{m}_j\}$. This computational advantage is paid by a slight deterioration of the numerical features of the algorithm. In fact, customary techniques for manipulating power series related to stochastic matrices are strongly numerically stable, i.e., the relative errors generated by the floating point arithmetic can be bounded entry-wise. Whereas, with the use of FFT it is possible to give very good relative error bounds but only in terms of norm (weak stability [3], [13]). However, from the many numerical experiments that we have performed we appreciated no substantial difference in the numerical behaviour of the two diverse implementations.

The description of the FFT-based techniques for matrix polynomials is performed according to the following lines. We first introduce Algorithms 5.1 for matrix polynomial multiplication modulo $z^m - 1$, and then use it for the computation of the product of a block Toeplitz matrix and a block vector. Then we reduce inversion of a matrix polynomial modulo z^m to inverting a lower block triangular block Toeplitz matrix, and we introduce Algorithm 5.2 for this computation. Finally we describe Algorithm 5.3 which expands, with full computational details, Algorithm 4.2 for the numerical solution of the matrix equation (1.3).

Throughout this section we will denote $\mathbf{y} = (y_0, \ldots, y_{d-1}) = \mathrm{DFT}(\mathbf{x})$ the Discrete Fourier Transform of the vector $\mathbf{x} = (x_0, \ldots, x_{d-1})$ defined by $y_i = \sum_{j=0}^{d-1} \omega^{ij} x_j$, $i = 0, \ldots, d-1$, where $\omega = \cos\frac{2\pi}{d} + \mathbf{i} \sin\frac{2\pi}{d}$, and $\mathbf{i}$ is the imaginary unit such that $\mathbf{i}^2 = -1$. Analogously $\mathbf{x} = \mathrm{IDFT}(\mathbf{y})$ denotes the Inverse Discrete Fourier Transform such that $\mathbf{x} = \frac{1}{d}\sum_{j=0}^{d-1} \omega^{-ij} y_j$, $i = 0, \ldots, d-1$. It is well known that if $d = 2^M$, where M is a positive integer, the computation of DFT and IDFT can be performed with $O(d \, log \, d)$ real arithmetic operations (hereafter denoted by ops) by means of the base-2 FFT algorithms. More precisely, in the case of real input the cost of DFT and IDFT is $\frac{5}{2} d \, log \, d + O(1)$ ops and $\frac{5}{2} d \, log \, d + d + O(1)$ ops, respectively, if we do not count the cost of computing the d-th roots of 1.

Similarly, we extend the above notation to block vectors. Given the block vector $X = (X_0, \ldots, X_{d-1})$, where $X_s = (x_{i,j}^{(s)})$ are $k \times k$ matrices, we define $Y = (Y_0, \ldots, Y_{d-1}) = \mathrm{DFT}(X)$ the block vector such that $Y_s = (y^{(s)})_{i,j}$ are $k \times k$ matrices and $(y_{i,j}^{(0)}, \ldots, y_{i,j}^{(d-1)}) = \mathrm{DFT}(x_{i,j}^{(0)}, \ldots, x_{i,j}^{(d-1)})$, $i, j = 1, \ldots, k$. Analogously we set $X = \mathrm{IDFT}(Y)$ if $(x_{i,j}^{(0)}, \ldots, x_{i,j}^{(d-1)}) = \mathrm{IDFT}(y_{i,j}^{(0)}, \ldots, y_{i,j}^{(d-1)})$. In this way the computation of $\mathrm{DFT}(X)$ and $\mathrm{IDFT}(Y)$ is reduced to the computation of k^2 DFT's and IDFT's, respectively, for the cost of $O(k^2 d \, log \, d)$ ops.

Let $P(z) = \sum_{i=0}^{m-1} P_i z^i$, $Q(z) = \sum_{i=0}^{m-1} Q_i z^i$, $R(z) = \sum_{i=0}^{m-1} R_i z^i$ be matrix polynomials such that $R(z) = P(z)Q(z) \bmod (z^m - 1)$. Then the matrix coefficients of $R(z)$ can be computed, given the matrix coefficients of $P(z)$ and $Q(z)$, by means of the following algorithm.

Algorithm 5.1. *Computation of matrix polynomial product modulo $z^m - 1$.*

Input. Positive integers M, k and the $k \times k$ matrices $P_0, \ldots, P_{m-1}$, $Q_0, \ldots, Q_{m-1}$, coefficients of the matrix polynomials $P(z) = \sum_{i=0}^{m-1} P_i z^i$, $Q(z) = \sum_{i=0}^{m-1} Q_i z^i$, where $m = 2^M$.

Output. The $k \times k$ matrices $R_0, \ldots, R_{m-1}$, coefficients of the matrix polynomial $R(z) = \sum_{i=0}^{m-1} R_i z^i$ such that $R(z) = P(z)Q(z) \bmod (z^m - 1)$.

Computation.

1. (Evaluation) Compute the entries of the $2m$ matrices $U_s = P(\omega^s)$, $V_s = Q(\omega^s)$, $s = 0, \ldots, m-1$, $\omega = \cos\frac{2\pi}{m} + \mathbf{i}\sin\frac{2\pi}{m}$, in the following way
$$(U_0, \ldots, U_{m-1}) = \mathrm{DFT}(P_0, \ldots, P_{m-1}),\ (V_0, \ldots, V_{m-1}) = \mathrm{DFT}(Q_0, \ldots, Q_{m-1}).$$
2. Compute the m matrix products $W_s = U_s V_s$, $s = 0, \ldots, m-1$, such that $W_s = P(\omega^s)Q(\omega^s) = R(\omega^s)$.
3. (Interpolation) Compute the entries of the matrices $R_0, \ldots, R_{m-1}$ by means of the equation $(R_0, \ldots, R_{m-1}) = \mathrm{IDFT}(W_0, \ldots, W_{m-1})$.

The cost of Algorithm 5.1 is $O(k^3 m + k^2 m \, log\, m)$ ops. More precisely, for real input matrices the cost is less than $5k^2 m \, log\, m + m(3k^3 + 2k^2)$. Indeed, due to the real input we have $W_s = \bar{W}_{m-s}$, $s = 1, \ldots, m/2 - 1$, where $\bar{W}_{m-s}$ is the complex conjugate of W_{m-s}, moreover W_0, $W_{m/2}$ are real matrices. Thus the computation at stage 2 is reduced to $m/2 - 1$ complex matrix products and 2 real matrix products. Complex matrix products are computed by means of the algorithm which uses 3 multiplications and 5 additions ([5]).

Remark 5.1. Observe that if the input polynomials $P(z)$ and $Q(z)$ have degree at most $m/2 - 1$ then $R(z) = P(z)Q(z)$, thus, matrix polynomial product can be computed by means of Algorithm 5.1.

Remark 5.2. Algorithm 5.1 can be also used for efficiently computing the product of a block Toeplitz matrix and a vector. In fact, if $P(z)$ has degree $m-1$ and $Q(z)$ has degree $m/2 - 1$, then, comparing the terms of degree $m/2, \ldots, m-1$ in the equation $P(z)Q(z) = R(z) \bmod (z^m - 1)$, we arrive at the following equation involving a block Toeplitz matrix

$$\begin{pmatrix} R_{m/2} \\ \vdots \\ R_{m-1} \end{pmatrix} = \begin{pmatrix} P_{m/2} & \ldots & P_1 \\ \vdots & \ddots & \vdots \\ P_{m-1} & \ldots & P_{m/2} \end{pmatrix} \begin{pmatrix} Q_0 \\ \vdots \\ Q_{m/2-1} \end{pmatrix}.$$

Comparing the coefficients of the term of degree i in both sides of the equation $P(z)Q(z) = I \bmod z^m$ yields the matrix equation involving a lower block triangular block Toeplitz matrix

$$\begin{pmatrix} P_0 & & & \bigcirc \\ P_1 & P_0 & & \\ \vdots & \ddots & \ddots & \\ P_{m-1} & \ldots & P_1 & P_0 \end{pmatrix} \begin{pmatrix} Q_0 \\ Q_1 \\ \vdots \\ Q_{m-1} \end{pmatrix} = \begin{pmatrix} I \\ O \\ \vdots \\ O \end{pmatrix}. \tag{5.1}$$

Let us denote by T_m the matrix in (5.1), assume for simplicity $m = 2^M$, M positive integer, and partition T_m as follows

$$T_m = \begin{pmatrix} T_{m/2} & O \\ H_{m/2} & T_{m/2} \end{pmatrix}$$

where all the four blocks are $(m/2) \times (m/2)$ block Toeplitz matrices. Then the matrix T_m^{-1} can be easily written as

$$T_m^{-1} = \begin{pmatrix} T_{m/2}^{-1} & O \\ -T_{m/2}^{-1} H_{m/2} T_{m/2}^{-1} & T_{m/2}^{-1} \end{pmatrix} \tag{5.2}$$

and the first block column $(Q_0^T, \dots, Q_{m-1}^T)^T$ of T_m^{-1}, which solves (5.1), can be computed from the first block column $(Q_0^T, \dots, Q_{m/2-1}^T)^T$ of $T_{m/2}$ by means of two multiplications between an $(m/2) \times (m/2)$ block Toeplitz matrix and a block vector. In this way we arrive at the following algorithm which extends to the block case the algorithm of Lafon and Sieveking-Kung (compare [4]).

Algorithm 5.2. *Solution of the congruence* $P(z)Q(z) = I \bmod z^m$*, or, equivalently, of the system (5.1).*

Input. A positive integer M, the $k \times k$ matrices $P_0, P_1, \dots, P_{m-1}$, $m = 2^M$, defining the matrix polynomial $P(z) = \sum_{i=0}^{m-1} P_i z^i$, $det\, P_0 \neq 0$.

Output. The matrices $Q_0, Q_1, \dots, Q_{m-1}$ satisfying (5.1), or equivalently, such that the polynomial $Q(z) = \sum_{i=0}^{m-1} Q_i z^i$ solves the congruence $P(z)Q(z) = I \bmod z^m$.

Computation.

1. Compute $Q_0 = P_0^{-1}$.
2. For $i = 0, \dots, M-1$, given the first column $U = (Q_0^T, \dots, Q_{2^i-1}^T)^T$ of $T_{2^i}^{-1}$, compute the block vector $V = (Q_{2^i}^T, \dots, Q_{2^{i+1}-1}^T)^T$, which defines the remaining blocks of the first column of $T_{2^{i+1}}^{-1}$, by applying equation (5.2) with $m = 2^{i+1}$ in the following way: by applying Algorithm 5.1 and Remark 5.2 compute the products $W = H_{2^i} U$ and $V = T_{2^i}^{-1} W$.

Observe that at the i-th stage of Algorithm 5.2 the computation of 3 DFT's and 2 IDFT's of order 2^{i+1} must be performed, together with the multiplication of about 2^{i+1} complex $k \times k$ matrices. The overall cost is about $6mk^3 + \frac{25}{2} k^2 m \log m$ ops.

Remark 5.3. In our applications $P(z) = I - \varphi_{odd}^{(j)}(z)$ (compare (2.11)) is such that the power series $P(z)^{-1} = \sum_{i=0}^{+\infty} Q_i z^i$ is convergent for $|z| = 1$. In this way we have $lim_i\, Q_i = O$, thus $P(z)^{-1}$ can be numerically truncated to the polynomial $Q(z) = \sum_{i=0}^{m} Q_i z^i$, for a suitable m such that

$$\mathbf{e}^T (P(1)^{-1} - Q(1)) < \epsilon \mathbf{e}^T \tag{5.3}$$

The value of the integer m such that (5.3) holds, can be dynamically adjusted by modifying Algorithm 5.2 in such a way to test, at each step i, the inequality (5.3).

In the following, we will denote with the same symbols $\varphi^{(j)}(z)$ and $\widehat{\varphi}^{(j)}(z)$ the polynomials of degree m_j and $\widehat{m}_j$, respectively, obtained by numerically truncating the power series $\varphi^{(j)}(z)$ and $\widehat{\varphi}^{(j)}(z)$ of (2.10a) and (4.5). Now, we are ready to give a detailed description of Algorithm 4.2 by means of the following

Algorithm 5.3. *Computation of the solution G_0 of (1.3).*

Input. Positive integers m_0, k, an error bound $\epsilon > 0$, and the nonnegative $k \times k$ matrices A_i, $i = 0, 1, \ldots, m_0$, such that the matrix $\sum_{i=0}^{m_0} A_i$ is ϵ-stochastic and the associated Markov chain is positive recurrent.

Output. An approximation $\widetilde{G}$ of the solution G_0 of (1.3) together with an error bound δ such that $\mathbf{e}^T |\sum_{i=0}^{+\infty} \widetilde{G}^i A_i - \widetilde{G}| \le \delta \mathbf{e}^T$.

Computation.

1. (Inizialization) Let $\varphi^{(0)}(z) = \sum_{i=0}^{m_0} A_i z^i$, $\widehat{\varphi}^{(0)}(z) = \sum_{i=0}^{m_0-1} A_{i+1} z^i$, $R^{(0)} = A_0(I - \sum_{i=1}^{m_0} A_i)^{-1}$, $j = 0$.

2. (Computation of the coefficients of the polynomials $\varphi^{(j)}(z)$ and $\widehat{\varphi}^{(j)}(z)$, of degree m_j and $\widehat{m}_j$, respectively, for $j = 1, 2, \ldots$, where m_j, $\widehat{m}_j$ are such that the matrices $\varphi^{(j)}(1)$ and $A_0 + \widehat{\varphi}^{(j)}(1)$ are ϵ-stochastic.)

Repeat

2.1. By means of Algorithm 5.2 and Remark 5.3 compute an integer q and the coefficients of the polynomial $Q(z) = \sum_{i=0}^{q-1} Q_i z^i$ such that

$$(I - \varphi_{odd}^{(j)}(z))Q(z) = I \bmod z^q$$

and (5.3) is verified for $P(z) = I - \varphi_{odd}^{(j)}(z)$.

2.2. By means of Algorithm 5.1 and Remark 5.1 compute the coefficients of the polynomial

$$S(z) = \varphi_{even}^{(j)}(z)Q(z).$$

2.3. By means of Algorithm 5.1 and Remark 5.1 compute the coefficients of the polynomials

$$T(z) = S(z)\varphi_{even}^{(j)}(z), \quad \widehat{T}(z) = S(z)\widehat{\varphi}_{even}^{(j)}(z).$$

2.4. Set $j = j + 1$.

2.5. Compute the integers m_j, $\widehat{m}_j$ and the coefficients of the polynomials $\varphi^{(j)}(z)$ and $\widehat{\varphi}^{(j)}(z)$ of degree m_j and $\widehat{m}_j$, respectively, such that

$$\varphi^{(j)}(z) = z\varphi_{odd}^{(j-1)}(z) + T(z) \bmod z^{m_j}$$
$$\widehat{\varphi}^{(j)}(z) = \widehat{\varphi}_{odd}^{(j-1)}(z) + \widehat{T}(z) \bmod z^{\widehat{m}_j}$$

and the matrices $\varphi^{(j)}(1)$ and $A_0 + \widehat{\varphi}^{(j)}(1)$ are ϵ-stochastic.

2.6 Compute $R^{(j)} = A_0^{(j)}(I - \sum_{i=1}^{m_0} A_i^{(j)})^{-1}$.

Until one of the following conditions is verified.

$$(C1) \qquad |R^{(j)} - R^{(j-1)}| < \epsilon E$$
$$(C2) \qquad \mathbf{e}^T(I - A_0^{(j)}(I - A_1^{(j)})^{-1}) < \epsilon \mathbf{e}^T$$
$$(C3) \qquad \mathbf{e}^T(I - A_0(I - \widehat{A}_1^{(j)})^{-1}) < \epsilon \mathbf{e}^T$$

3. (Computation of the matrix $\widetilde{G}$)

3.1. If condition (C1) is verified then $\widetilde{G} = A_0(I - \sum_{i=0}^{\widehat{m}_j-1} R^{(j)^i} \widehat{A}_{i+1}^{(j)})^{-1}$.

3.2. If condition (C2) is verified then $\widetilde{G} = A_0(I - \widehat{A}_1^{(j)})^{-1}$.

3.3. If condition (C3) is verified then $\widetilde{G} = A_0(I - \widehat{A}_1^{(j)})^{-1}$.

4. (Computation of the error bound δ.) Compute $\delta = max_{j=1,\dots,k} \sum_{i=1}^{k} |w_{i,j}|$, for $W = (w_{i,j})_{i,j}$, $W = \sum_{i=0}^{m_0} \widetilde{G}^i A_i - \widetilde{G}$.

6. Numerical Experiments. We have implemented Algorithm 4.2 in Fortran 77, by using the customary arithmetic for matrix polynomials (program CCR: Customary Cyclic Reduction) and by using the FFT based arithmetic for matrix polynomials, as described in details by Algorithm 5.3 (program FCR: Fast Cyclic Reduction). The programs have been run on a Spark Workstation using the standard double precision IEEE arithmetic. The value of ϵ defining the stopping condition has been choosen as $\epsilon = 10^{-13}$. We have compared our algorithms with the functional iteration (1.5c), which is the fastest among the iterations (1.5) (compare [14]), starting with $X_0 = O$ and $X_0 = I$, respectively, and with the method proposed in [12]. The implementation of the functional iteration (1.5c) has been performed in Fortran 77 (program FIF: Functional Iteration Formula), while for the algorithm of Latouche and Stewart [12] we have adapted to the specific problem the implementation in C given by Stewart [17] (program LS). In this adaptation we did not use FFT.

We have tested our programs on a problem analyzed in [1] arising from the mathematical modelling of a metropolitan network. The blocks A_i of (1.3) have dimension $k = 16$. We have tested the algorithms for different values of the parameter ρ of (2.1). For $\rho = 0.1$ the initial cutting level m_0 is 168; for $\rho = 0.8$ the initial cutting level m_0 is 240; for the remaining tested values of $\rho \geq 0.9$ the initial cutting level m_0 is 264. The problems with ρ close to 1 are characterized by a "long queue", i.e., the components $\boldsymbol{\pi}_i$ of the invariant probability vector are not negligible even for very large values of i. The programs FCR and CCR provided an approximation to the solution G_0 with a residual $\delta < 10^{-13}$ for $\rho = 0.1$, and $\delta < 10^{-12}$ for the tested values of $\rho \geq 0.8$.

Table 6.1 reports the CPU time and the number of iterations needed by algorithms CCR, FCR, together with the residual $||\mathbf{e}^T|\widetilde{G} - \sum_{i=0}^{m_0} \widetilde{G}^i A_i|\,||$. Observe that, despite the weak stability property of FCR, the residual values of FCR and CCR are equal. We may observe that the use of FFT leads to a substantial reduction of the CPU time.

	FCR			CCR		
	Time (s.)	Iterations	Residual	Time (s.)	Iterations	Residual
$\rho = 0.1$	27	9	$8.5 \cdot 10^{-14}$	185	9	$8.5 \cdot 10^{-14}$
$\rho = 0.8$	61	13	$2.7 \cdot 10^{-13}$	484	13	$2.7 \cdot 10^{-13}$
$\rho = 0.9$	62	14	$2.7 \cdot 10^{-13}$	545	14	$2.7 \cdot 10^{-13}$
$\rho = 0.95$	65	16	$1.8 \cdot 10^{-13}$	592	16	$1.8 \cdot 10^{-13}$
$\rho = 0.96$	65	17	$2.0 \cdot 10^{-13}$	592	17	$2.0 \cdot 10^{-13}$
$\rho = 0.97$	66	20	$2.3 \cdot 10^{-13}$	603	20	$2.3 \cdot 10^{-13}$

Table 6.1. Cyclic Reduction.

Table 6.2 reports the time and the number of iterations needed by program FIF, starting with $X_0 = O$ and $X_0 = I$, respectively, to compute an approximation of the matrix G_0 with a residual value less than 10^{-12}.

	FIF ($X_0 = I$)			FIF ($X_0 = O$)		
	Time (s.)	Iterations	Residual	Time (s.)	Iterations	Residual
$\rho = 0.1$	8.4	22	$2.3 \cdot 10^{-14}$	8.4	22	$2.6 \cdot 10^{-14}$
$\rho = 0.8$	79	148	$1.1 \cdot 10^{-13}$	116	211	$2.0 \cdot 10^{-13}$
$\rho = 0.9$	220	373	$1.2 \cdot 10^{-13}$	294	495	$2.3 \cdot 10^{-13}$
$\rho = 0.95$	902	1534	$1.2 \cdot 10^{-13}$	1104	1866	$1.2 \cdot 10^{-13}$
$\rho = 0.96$	1862	3157	$1.3 \cdot 10^{-13}$	2561	4343	$1.4 \cdot 10^{-13}$
$\rho = 0.97$	1967	3336	$1.2 \cdot 10^{-13}$	11148	18900	$1.5 \cdot 10^{-13}$

Table 6.2. Functional Iterations.

Table 6.3 reports the CPU time and the cut off level needed by algorithm LS, together with the residual $||\mathbf{e}^T|\widetilde{G} - \sum_{i=0}^{m_0} \widetilde{G}^i A_i|\ ||$. For $\rho = 0.1$ the cut off level sufficient to obtain a residual value less than 10^{-12} is 600. For $\rho \geq 0.8$ we have higher residual values since we were not able to increase the cut off level above 3500.

	LS		
	Time (s.)	Cut off level	Residual
$\rho = 0.1$	1019	600	$7.8 \cdot 10^{-13}$
$\rho = 0.8$	12332	3500	$5.3 \cdot 10^{-8}$
$\rho = 0.9$	12332	3500	$6.5 \cdot 10^{-6}$
$\rho = 0.95$	12332	3500	$2.2 \cdot 10^{-5}$
$\rho = 0.96$	12332	3500	$2.4 \cdot 10^{-5}$
$\rho = 0.97$	12332	3500	$2.3 \cdot 10^{-5}$

Table 6.3. Algorithm LS.

Table 6.4 reports the ratio between the CPU times needed by algorithms CCR, FIF (with $X_0 = O$ and $X_0 = I$), LS and the CPU time needed by algorithm FCR. The symbol "$*$" means that the corresponding residual value is higher than the residual value obtained with program FCR and we were not able to decrease it.

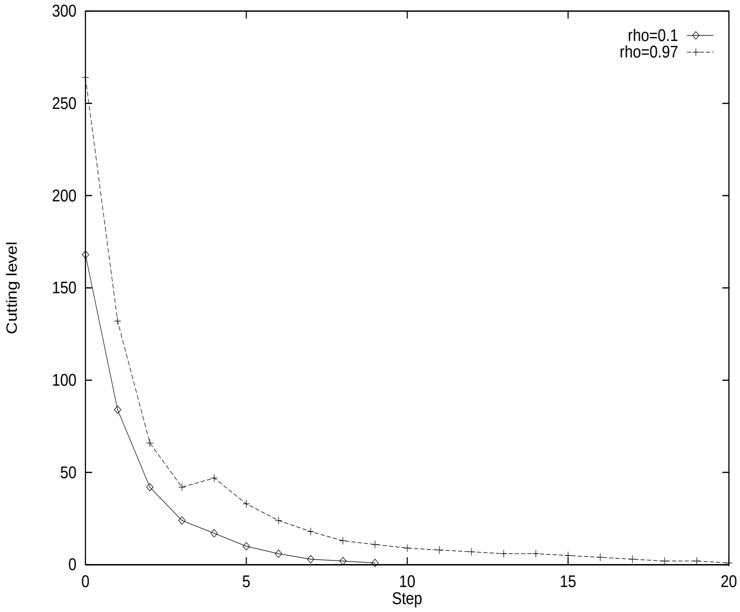
	CCR/FCR	FIF/FCR (X_0=I)	FIF/FCR (X_0=O)	LS/FCR
$\rho = 0.1$	6.8	0.3	0.3	37.7
$\rho = 0.8$	7.9	1.3	1.9	202.2 (*)
$\rho = 0.9$	8.8	3.5	4.7	198.9 (*)
$\rho = 0.95$	9.1	13.9	17.0	189.7 (*)
$\rho = 0.96$	9.1	28.6	39.4	189.7 (*)
$\rho = 0.97$	9.1	29.8	168.9	186.8 (*)

Table 6.4. Ratio of CPU times.

We observe that our method is almost unsensible with respect to the values of ρ. Whereas the performance of FIF strongly deteriorates when ρ approaches 1. It is interesting to point out the different rate of convergence of FIF in the case $X_0 = O$ and $X_0 = I$. This behaviour, already pointed out by Latouche [9], has received a full theoretical explanation in [14].

The number of steps needed by our algorithm seems to depend on $log_2 d$, where d is a positive integer such that the components $\boldsymbol{\pi}_i$ are negligible for $i > d$. It is worth pointing out that our algorithm is particularly suitable for "hard" problems, where the value of ρ is close to 1 and the classical functional iteration formulae (1.5) converge very slowly.

The following figure shows the size of the cutting level m_j at each step of cyclic reduction for $\rho = 0.1$ and $\rho = 0.97$. The decreasing behaviour leads to a further reduction of the complexity.



REFERENCES

[1] G. Anastasi, L. Lenzini, B. Meini, *Performance evaluation of a worst case model of the Metaring MAC Protocol with global fairness*, Submitted for publication.

[2] D. Bini, B. Meini, *On cyclic reduction applied to a class of Toeplitz-like matrices arising in queueing problems*, in Proceedings of the Second International Workshop on Numerical Solution of Markov Chains, Raleigh, North Carolina, 1995, pp. 21–38.

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

[4] ———, *Polynomial division and its computational complexity*, J. Complexity, 6 (1986), pp. 179–203.

[5] A. Borodin, I. Munro, *The computational complexity of algebraic and numeric problems*, American Elsevier, New York, 1975.

[6] B. L. Buzbee, G. H. Golub, C. W. Nielson, *On direct methods for solving Poisson's equation,* SIAM J. Numer. Anal., 7 (1970), pp. 627–656.

[7] E. Çinlar, *Introduction to Stochastic Processes,* Prentice-hall, Englewood Cliffs, N.J., 1975.

[8] W. K. Grassman, M. I. Taksar, D. P. Heyman, *Regenerative analysis and steady state distribution for Markov chains,* Oper. Res., 33 (1985), pp. 1107–1116.

[9] G. Latouche, *Algorithms for evaluating the matrix G in Markov chains of PH/G/1 type,* Bellcore Technical Report, 1992.

[10] ———, *Newton's iteration for non-linear equations in Markov chains,* IMA J. of Numerical Analysis, 14 (1994), pp. 583–598.

[11] G. Latouche, V. Ramaswami, *A logarithmic reduction algorithm for quasi-birth-death processes,* J. Appl. Prob., 30 (1993), pp. 650–674.

[12] G. Latouche, G.W. Stewart, *Numerical methods for M/G/1 type queues,* in Proceedings of the Second International Workshop on Numerical Solution of Markov Chains, Raleigh, North Carolina, 1995, pp. 571–581.

[13] E. Linzer, *On the stability of transform-based circular deconvolution,* SIAM J. Numer. Anal., 29 (1992), pp. 1482–1492.

[14] B. Meini, *New convergence results on functional iteration techniques for the numerical solution of M/G/1 type Markov chains,* Submitted for publication.

[15] M.F. Neuts, *Structured Stochastic Matrices of M/G/1 Type and Their Applications,* Dekker Inc., New York, 1989.

[16] V. Ramaswami, *Nonlinear matrix equations in applied probability - Solution techniques and open problems,* SIAM Review, 30 (1988), pp. 256-263.

[17] G.W. Stewart, *Implementing an algorithm for solving block Hessenberg systems,* Technical Report CS-TR-3295, Department of Computer Science, University of Maryland, College Park, 1993.

[18] R. Varga, *Matrix Iterative Analysis,* Prentice-Hall, Englewood Cliffs, N.J., 1963.