2ND INTERNATIONAL WORKSHOP ON THE NUMERICAL SOLUTION OF MARKOV CHAINS

2ND INTERNATIONAL WORKSHOP ON THE NUMERICAL SOLUTION OF MARKOV CHAINS

EDITED BY

William J. Stewart
North Carolina State University
Raleigh, NC, USA

KLUWER ACADEMIC PUBLISHERS
Boston/London/Dordrecht

CONTENTS

CONTRIBUTORS

1

ON CYCLIC REDUCTION APPLIED TO A CLASS OF TOEPLITZ-LIKE MATRICES ARISING IN QUEUEING PROBLEMS

Dario Bini and Beatrice Meini

Dipartimento di Matematica, Università di Pisa

56100 Pisa, Italy

ABSTRACT

We observe that the cyclic reduction algorithm leaves unchanged the structure of a block Toeplitz matrix in block Hessenberg form. Based on this fact, we devise a fast algorithm for computing the probability invariant vector of stochastic matrices of a wide class of Toeplitz-like matrices arising in queueing problems. We prove that for any block Toeplitz matrix H in block Hessenberg form it is possible to carry out the cyclic reduction algorithm with $O(k^3 n + k^2 n \log n)$ arithmetic operations, where k is the size of the blocks and n is the number of blocks in each row and column of H. The probability invariant vector is computed within the same cost. This substantially improves the $O(k^3 n^2)$ arithmetic cost of the known methods based on Gaussian elimination. The algorithm, based on the FFT, is numerically weakly stable. In the case of semi-infinite matrices the cyclic reduction algorithm is rephrased in functional form by means of the concept of generating function and a convergence result is proved. An implementation of the algorithm for the infinite case is presented together with the results of the numerical experiments which confirm the effectiveness of our approach.

1 INTRODUCTION.

Given two sequences of $k \times k$ matrices $\{A_i\}_{i\in N}$, $\{\widetilde{A}_i\}_{i\in N}$, consider the semi-infinite matrix with the following block structure

$$\begin{pmatrix} \widetilde{A}_1 & A_0 & & \bigcirc & \\ \widetilde{A}_2 & A_1 & A_0 & & \\ \vdots & \vdots & \ddots & \ddots & \\ \widetilde{A}_{n-1} & A_{n-2} & \ddots & A_1 & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}. \tag{1.1}$$

Moreover, given the sequence $\{\widehat{A}_i\}_{i\in N}$, for any integer n we may define the finite $m \times m$ matrix, $m = nk$:

$$\begin{pmatrix} \widetilde{A}_1 & A_0 & & \bigcirc & \\ \widetilde{A}_2 & A_1 & A_0 & & \\ \vdots & \vdots & \ddots & \ddots & \\ \widetilde{A}_{n-1} & A_{n-2} & \dots & A_1 & A_0 \\ \widehat{A}_n & \widehat{A}_{n-1} & \dots & \widehat{A}_2 & \widehat{A}_1 \end{pmatrix}. \tag{1.2}$$

Here and hereafter A^T will denote a matrix of (1.1) or (1.2). We consider the problem of solving the homogeneous system $(I - A^T)\boldsymbol{\pi} = \mathbf{0}$, $\boldsymbol{\pi} = (\pi_i)$, subject to the constraint $\sum_i \pi_i = 1$, $\pi \geq 0$, in the special cases where the matrix A^T of (1.1) or of (1.2), is a row stochastic matrix, i.e., a matrix having nonnegative entries such that $A\mathbf{e} = \mathbf{e}$, $\mathbf{e} = (1, 1, \dots)^T$. Matrices of this kind arise in the analysis of queueing problems where $\boldsymbol{\pi}$ is the stationary probability vector and A is the transition matrix of the Markov chain associated with the problem.

Matrices sharing the structure (1.1) are called "matrices of the M/G/1 type" [19] and are special instances of a wider class of matrices arising in many applications, better known as the class of Toeplitz-like matrices [3], [12] [13]. A matrix A is (block) Toeplitz if its (blocks) entries are shift-invariant, i.e., they just depend on the difference of their subscripts. A matrix A is called (block) Toeplitz-like if the rank of $A - SAS^T$ is independent of n, here $S = (s_{i,j})$, $s_{i,i-1} = 1$, $s_{i,j} = 0$ elsewhere ($S = (S_{i,j})$, $S_{i,i-1} = I$, $S_{i,j} = 0$ elsewhere), denotes the down-shift (block) matrix. Observe that the matrices (1.1) and (1.2) have also the block Hessenberg form, i.e., the blocks in position (i, j) are null matrices if $j > i + 1$. We will refer to the matrices (1.1) as to matrices of the M/G/1 type and to the matrices (1.2) as to Toeplitz-like matrices in block Hessenberg form.

In this paper, we exploit the special structure of the matrices (1.1) and (1.2) in order to devise numerically stable algorithms for the computation of $\boldsymbol{\pi}$, which have a low computational cost.

Concerning the finite case (1.2), we prove that a single step of the cyclic reduction algorithm applied to the system $(I - A^T)\boldsymbol{\pi} = \mathbf{0}$, i.e., an odd-even reordering of block rows and block columns followed by a step of block Gaussian elimination, leaves unchanged the Toeplitz-like and the Hessenberg structures of (1.2) and leads to a problem of half the size. In this way, by taking advantage from the fast (FFT-based) algorithms for Toeplitz matrix computations, we devise fast and numerically stable algorithms for the computation of the vector $\boldsymbol{\pi}$ requiring $O(k^2 n \log n + k^3 n)$ arithmetic operations. This complexity bound substantially decreases the $O(k^3 n^2)$ bound obtained by applying the customary Gaussian elimination algorithm and slightly improves the $O(k^3 n \log^2 n)$ asymptotic bound of the algorithm of G.W. Stewart [21]. The storage required by our algorithms is kept to its minimum value, more specifically, only $k^3(3n-1)$ floating point numbers must be stored in all the computation, that is, the same number of input values of the matrix. Due to the nice computational properties of the FFT, our algorithms can be efficiently implemented in the PRAM model of parallel computation.

In the infinite case we express the cyclic reduction method in functional form. More precisely, we associate with the matrix (1.1) the matrix functions $\varphi(z) = \sum_{i=0}^{+\infty} A_i z^i$ and $\tilde{\varphi}(z) = \sum_{i=0}^{+\infty} \widetilde{A}_{i+1} z^i$ (generating functions) and determine a functional expression which relates the generating functions $\varphi^{(j)}(z)$ and $\tilde{\varphi}^{(j)}(z)$ of the matrix $A^{(j)}$ involved at the j-th recursive step of the cyclic reduction algorithm to the generating functions $\varphi^{(j-1)}(z)$ and $\tilde{\varphi}^{(j-1)}(z)$ of the matrix $A^{(j-1)}$ involved at the $(j-1)$-st step. By using this formulation we prove that there exists $\lim_{j\to+\infty} \widetilde{A}_1^{(j)} = W$, where $\tilde{\varphi}^{(j)}(z) = \sum_{i=0}^{+\infty} \widetilde{A}_{i+1}^{(j)} z^i$; estimates of the convergence speed are also given. Moreover, W^T is a row stochastic matrix such that $W\boldsymbol{\pi}_0 = \boldsymbol{\pi}_0$, where $\boldsymbol{\pi}^T = (\boldsymbol{\pi}_0^T, \boldsymbol{\pi}_1^T, \ldots)$. In this way, by means of cyclic reduction we may approximate the matrix W with $\widetilde{A}_1^{(h)}$, for a suitable h. Then, by computing the dominant eigenvector of $\widetilde{A}_1^{(h)}$, we may approximate the vector $\boldsymbol{\pi}_0$. Finally, a backward substitution allows us to recover the block components $\boldsymbol{\pi}_1, \boldsymbol{\pi}_2, \ldots, \boldsymbol{\pi}_{2^h-1}$. Moreover, we observe that only the blocks $A_0, \ldots, A_{2^h-1}$, $\widetilde{A}_1, \ldots, \widetilde{A}_{2^h}$, are involved in the computation; hence, the described functional relations involving $\varphi^{(j)}(z)$ and $\widetilde{\varphi}^{(j)}(z)$, can be applied modulo z^{2^h} and the computations can be reduced to operations between matrix polynomials as in the finite case.

The paper is organized as follows, in section 2 we recall some useful results concerning the computation of $\boldsymbol{\pi}$ by means of matrix factorizations in the finite case. In sections 3 we analyze the cyclic reduction algorithm in the finite case, leading to the block LU factorization of a suitable permuted matrix. In section 4 we describe our algorithms for the computation of the vector $\boldsymbol{\pi}$ in the finite case, based on cyclic reduction and analyze their computational costs. In section 5 we give the functional formulation of the cyclic reduction algorithm in the infinite case, present the convergence properties and discuss the results of the numerical experiments.

2 PRELIMINARIES.

The problem of computing the steady state distribution of a finite Markov chain has been investigated by several authors. Let us recall some of the main results on direct approaches based on LU and UL matrix factorizations of an M-matrix; we refer the reader to [1], [2], [6], [8], [9], [10] and [11] for more details and further references. We recall that an M-matrix is any matrix of the form $\alpha I - A$ where A has nonnegative entries and the spectral radius $\rho(A)$ of A is such that $\rho(A) \geq \alpha$. Thus, for any row-stochastic matrix A, both $I - A$ and $I - A^T$ are singular M-matrices since $\rho(A) = \rho(A^T) = 1$.

The next two propositions, which are a special case of more general results given in [22], [7], are at the basis of the methods given in [9], [6], [7], [10], [11] and [2] for the computation of $\boldsymbol{\pi}$.

Proposition 1 *For any $m \times m$ row stochastic matrix A there exist matrices L and U, such that $I - A^T = LU$, where L is lower triangular with unit diagonal entries, and U is upper triangular. Moreover, L and U are M-matrices, therefore the off-diagonal entries of L and U are nonpositive, the diagonal entries of U are nonnegative. If in addition A is irreducible then $u_{i,i} \neq 0$, $i = 1, \ldots, m-1$, $u_{m,m} = 0$.*

Observe that proposition 1 still holds for block matrices if the LU factorization is replaced with the block LU factorization and the condition $u_{m,m} = 0$ is replaced with $\det U_{m,m} = 0$.

A similar result can be obtained in terms of (block) UL factorization, by applying the following:

Proposition 2 *For any $m \times m$ row stochastic irreducible matrix A there exist matrices L and U such that $I - A = LU$, where L is lower triangular with unit diagonal entries, and U is upper triangular. Moreover, L and U are M-matrices, therefore the off-diagonal entries of L and U are nonpositive, the diagonal entries of U are $u_{i,i} > 0$, $i = 1, \ldots, m-1$, $u_{m,m} = 0$.*

By applying proposition 2 to the row stochastic matrix JAJ, where J is the permutation (reversion) matrix having unit entries along the secondary diagonal, i.e., for $i = j - n + 1$, we have the factorization $I - JAJ = \hat{L}\hat{U}$. In this way we obtain the factorization $I - A^T = \tilde{U}\tilde{L}$, where $\tilde{U} = (J\hat{U}J)^T$, $\tilde{L} = (J\hat{L}J)^T$. It is a simple matter to check that the first entry (block) on the main diagonal of $\tilde{U}$ is zero (singular).

If the factors L, U, or $\tilde{L}$, $\tilde{U}$ are computed by means of Gaussian elimination with the diagonal adjustment technique of [9], the computation is *strongly numerically stable* [5]. That is, the relative rounding errors in the output can be bounded componentwise. We say that an algorithm is *weakly numerically stable* if the relative rounding errors in the output can be only bounded in norm. Once the (block) triangular factors L and U, $\tilde{L}$ and $\tilde{U}$, respectively, have been computed, the relations $I - A^T = LU$, $I - A^T = \tilde{U}\tilde{L}$ allow us to compute the components of the probability invariant vector $\boldsymbol{\pi}$ in a numerically stable way. In fact, the following three conditions are equivalent

$$(I - A^T)\boldsymbol{\pi} = \mathbf{0}, \tag{1.3}$$

$$U\boldsymbol{\pi} = 0, \tag{1.4a}$$

$$\begin{cases} \tilde{L}\boldsymbol{\pi} = 0 \\ \tilde{U}_{1,1}\boldsymbol{\pi}_0 = \mathbf{0}. \end{cases} \tag{1.4b}$$

The method proposed in [6], [7] consists of solving equation (1.4a): we may arbitrarily fix the last component of $\boldsymbol{\pi}$ and use back substitution in order to recover all the remaining components. Normalization allows us to compute the desired values. This procedure is numerically stable since it involves only additions of positive numbers. If $I - A^T = LU$ is a block factorization, we have just to solve the problem $U_{n,n}\boldsymbol{\pi}_{n-1} = 0$ (say, by means of the computation of the LU or $\tilde{U}\tilde{L}$ factorization of $U_{n,n}$) and then recover the remaining block components by means of back substitution. In this last stage we have to solve $n-1$ linear systems having matrix $U_{i,i}$, $i = 1, \ldots, n-1$. Again, this computation can be performed by means of the LU, or $\tilde{U}\tilde{L}$ factorizations of the matrices $U_{i,i}$, $i = 1, \ldots, n-1$. Also in the block case, the back substitution stage is numerically stable since it is reduced to performing additions of positive numbers.

A similar computation holds if (1.4b) is used instead of (1.4a). In this case we have the method proposed in [9]: the back substitution stage is replaced by forward substitution, once the first component (block) of $\boldsymbol{\pi}$ has been fixed (computed).

In the approach of [10], [11], [1] the matrix $I - A^T$ is partitioned as follows

$$I - A^T = \begin{pmatrix} B & \mathbf{y} \\ \mathbf{z}^T & \beta \end{pmatrix}, \tag{1.5}$$

where B is an $(n-1)\times(n-1)$ matrix, $\mathbf{x}$, $\mathbf{z}$ are vectors having $n-1$ components and β is a real number. Thus, the problem is reduced to solving the system $B\mathbf{x} = -\mathbf{y}$ and to setting $\boldsymbol{\pi}^T = (\mathbf{x}^T, 1)/(1 + \sum x_i)$. In [2] Barlow performs an accurate error analysis and proves that this approach is numerically stable provided that the system $B\mathbf{x} = -\mathbf{y}$ is solved by means of the LU, or QR factorization, with the diagonal adjustment technique of Grassman, Taksar and Heyman [9].

Now observe that if A is stochastic then P^TAP is stochastic for any permutation matrix P and even irreducibility is preserved under this similarity transformation. Therefore propositions 1 and 2 still hold if A is replaced with $I - P^TAP$, for any permutation matrix P. This fact is used in the next section where the LU factorization of the matrix P^TAP is computed, for a suitable matrix P, by means of cyclic reduction.

3 LU FACTORIZATION BY MEANS OF CYCLIC REDUCTION.

Let $p \geq 2$ be an integer and let $n = 2^p$. Consider the Toeplitz-like matrix in block Hessenberg form

$$H = H^{(0)} = I - A^T = \begin{pmatrix} \widetilde{H}_1 & H_0 & & & \bigcirc \\ \widetilde{H}_2 & H_1 & H_0 & & \\ \vdots & \vdots & \ddots & \ddots & \\ \widetilde{H}_{n-1} & H_{n-2} & \dots & H_1 & H_0 \\ \widehat{H}_n & \widehat{H}_{n-1} & \dots & \widehat{H}_2 & \widehat{H}_1 \end{pmatrix}, \tag{1.6}$$

uniquely defined by the blocks $H_0 = -A_0$, $H_1 = I - A_1$, $\widetilde{H}_1 = I - \widetilde{A}_1$, $\widehat{H}_1 = I - \widehat{A}_1$, $H_i = -A_i$, $\widetilde{H}_i = -\widetilde{A}_i$, $\widehat{H}_i = -\widehat{A}_i$, $i = 2, \dots, n-2$, $\widetilde{H}_{n-1} = -\widetilde{A}_{n-1}$, $\widehat{H}_n = -\widehat{A}_n$, $\widehat{H}_{n-1} = -\widehat{A}_{n-1}$.

Interchanging rows and columns of $H^{(0)} = H$ according to the odd-even permutation $(1, 3, 5, \ldots, n-1, 0, 2, 4, 6, \ldots, n-2)$ we find the matrix

$$P^{(0)} H^{(0)} P^{(0)T} = \begin{pmatrix} T_1^{(0)} & W^{(0)} \\ Z^{(0)} & T_2^{(0)} \end{pmatrix}, \tag{1.7}$$

where $P^{(0)}$ is the odd-even permutation block matrix, $T_1^{(0)}$, $T_2^{(0)}$, $Z^{(0)}$ and $W^{(0)}$ are $(n/2) \times (n/2)$ block Toeplitz-like matrices, defined by

$$\begin{aligned}
W^{(0)} &= \begin{pmatrix} \widetilde{H}_2 & H_0 & & \bigcirc \\ \widetilde{H}_4 & H_2 & \ddots & \\ \vdots & \vdots & \ddots & H_0 \\ \widehat{H}_n & \widehat{H}_{n-2} & \ldots & \widehat{H}_2 \end{pmatrix}, \; Z^{(0)} = \begin{pmatrix} H_0 & & & \bigcirc \\ H_2 & H_0 & & \\ \vdots & \ddots & \ddots & \\ H_{n-2} & \ldots & H_2 & H_0 \end{pmatrix}, \\
T_1^{(0)} &= \begin{pmatrix} H_1 & & & & \bigcirc \\ H_3 & H_1 & & & \\ \vdots & \vdots & \ddots & & \\ H_{n-3} & H_{n-5} & \ldots & H_1 & \\ \widehat{H}_{n-1} & \widehat{H}_{n-3} & \ldots & \widehat{H}_3 & \widehat{H}_1 \end{pmatrix}, \\
T_2^{(0)} &= \begin{pmatrix} \widetilde{H}_1 & & & \bigcirc & \\ \widetilde{H}_3 & H_1 & & & \\ \vdots & \ddots & \ddots & & \\ \widetilde{H}_{n-3} & \ldots & H_3 & H_1 & \\ \widetilde{H}_{n-1} & H_{n-3} & \ldots & H_3 & H_1 \end{pmatrix}.
\end{aligned} \tag{1.8}$$

By applying one step of block Gaussian elimination to the 2×2 block matrix (1.7) we find that

$$\begin{aligned}
P^{(0)} H^{(0)} P^{(0)T} &= \begin{pmatrix} I & 0 \\ L^{(0)} & I \end{pmatrix} \begin{pmatrix} T_1^{(0)} & W^{(0)} \\ 0 & H^{(1)} \end{pmatrix}, \\
L^{(0)} &= Z^{(0)} T_1^{(0)^{-1}}, \; H^{(1)} = T_2^{(0)} - L^{(0)} W^{(0)}.
\end{aligned} \tag{1.9}$$

Now, denote

$$T(B_0, B_1, \ldots, B_{q-1}) = \begin{pmatrix} B_0 & & & \bigcirc \\ B_1 & B_0 & & \\ \vdots & \ddots & \ddots & \\ B_{q-1} & \ldots & B_1 & B_0 \end{pmatrix}$$

the $q \times q$ block lower triangular block Toeplitz matrix uniquely defined by its first block column $(B_0^T, B_1^T, \ldots, B_{q-1}^T)^T$, and define the $(n/2-1) \times (n/2-1)$

block matrices

$$F_0^{(0)} = T(O, H_1, H_3, \ldots, H_{n-5}),\ \ F_1^{(0)} = T(H_0, H_2, \ldots, H_{n-4}), \\ F_2^{(0)} = T(H_1, H_3, \ldots, H_{n-3}). \tag{1.10}$$

Observe that from (1.9) and (1.8) it follows that

$$H^{(1)} = \begin{pmatrix} U^{(1)} & R^{(1)} \\ S^{(1)} & V^{(1)} \end{pmatrix} \tag{1.11}$$

where

$$\begin{aligned} R^{(1)} &= F_0^{(0)} - F_1^{(0)} F_2^{(0)^{-1}} F_1^{(0)} \\ U^{(1)} &= \begin{pmatrix} \widetilde{H}_1 \\ \widetilde{H}_3 \\ \vdots \\ \widetilde{H}_{n-3} \end{pmatrix} - F_1^{(0)} F_2^{(0)^{-1}} \begin{pmatrix} \widetilde{H}_2 \\ \widetilde{H}_4 \\ \vdots \\ \widetilde{H}_{n-2} \end{pmatrix}, \\ (S^{(1)}, V^{(1)}) &= (\widetilde{H}_{n-1}, H_{n-3}, \ldots, H_1) - (H_{n-2}, H_{n-4}, \ldots, H_0) T_1^{(0)^{-1}} W^{(0)}. \end{aligned} \tag{1.12}$$

Since the linear space of block lower triangular block Toeplitz matrices is closed under multiplication and inversion, i.e., constitutes a matrix algebra, we deduce that $R^{(1)}$ is a block triangular block Toeplitz matrix. Whence, the Schur complement $H^{(1)}$ has the same structure of (1.6), i.e., $A^{(1)T} = I - H^{(1)}$ is still a block Toeplitz-like matrix in block Hessenberg form. $H^{(1)}$ is uniquely defined by the matrices $H_0^{(1)}, \ldots, H_{n/2-2}^{(1)}$, $\widetilde{H}_1^{(1)}, \ldots, \widetilde{H}_{n/2-1}^{(1)}$, and $\widehat{H}_1^{(1)}, \ldots, \widehat{H}_{n/2}^{(1)}$ such that

$$\begin{aligned} R^{(1)} &= T(H_0^{(1)}, H_1^{(1)}, \ldots, H_{n/2-2}^{(1)}), \\ U^{(1)} &= (\widetilde{H}_1^{(1)T}, \widetilde{H}_2^{(1)T}, \ldots, \widetilde{H}_{n/2-1}^{(1)T})^T, \\ (S^{(1)}, V^{(1)}) &= (\widehat{H}_{n/2}^{(1)}, \widehat{H}_{n/2-1}^{(1)}, \ldots, \widehat{H}_1^{(1)}). \end{aligned} \tag{1.13}$$

Moreover, it is easy to check that $A^{(1)T}$ is a row stochastic matrix. Therefore the cyclic reduction process can be recursively applied $p-1$ times until the $2k \times 2k$ matrix $H^{(p-1)}$ is computed.

The resulting algorithm, which computes the block LU factorization of a matrix obtained by suitably permuting rows and columns of H, is described by the formulae (1.6)–(1.13) adjusted to the generic j-th step. This algorithm outputs the blocks $H_i^{(j+1)}$, $i = 0, \ldots, 2^{p-j-1}-2$, $\widehat{H}_i^{(j+1)}$, $i = 1, \ldots, 2^{p-j-1}$, and $\widetilde{H}_i^{(j+1)}$, $i = 1, \ldots, 2^{p-j-1}-1$, defining the block matrices $T_1^{(j)}$, $W^{(j)}$, $L^{(j)}$, $H^{(j+1)}$,

$j = 0, 1, \ldots, p-2$. The analysis of this algorithm and its use for computing the probability invariant vector, is postponed until the next section where we apply the FFT-based algorithms for Toeplitz matrices in order to speed up the computation.

We observe that in the case where the matrix H is block tridiagonal the whole computation can be carried out in a simpler form and we arrive at the finite-dimensional analogue of the algorithm of Latouche and Ramaswami [17], (see also [23]). In fact, let

$$H = \begin{pmatrix} \widetilde{A} & B & & & \bigcirc \\ \widetilde{C} & A & B & & \\ & C & \ddots & \ddots & \\ & & \ddots & A & B \\ \bigcirc & & & \widehat{C} & \widehat{A} \end{pmatrix}; \tag{1.14}$$

then the block tridiagonal matrix $H^{(j)}$ obtained at the j-th recursive step of cyclic reduction, defined by the blocks $A^{(j)}$, $B^{(j)}$, $C^{(j)}$, $\widetilde{A}^{(j)}$, $\widetilde{C}^{(j)}$, $\widehat{A}^{(j)}$, $\widehat{C}^{(j)}$, $j = 0, \ldots, \log_2 n - 1$, is such that

$$\begin{aligned} A^{(j+1)} &= A^{(j)} - (C^{(j)} A^{(j)^{-1}} B^{(j)} + B^{(j)} A^{(j)^{-1}} C^{(j)}), \\ B^{(j+1)} &= -B^{(j)} A^{(j)^{-1}} B^{(j)}, \quad C^{(j+1)} = -C^{(j)} A^{(j)^{-1}} C^{(j)} \\ \widetilde{A}^{(j+1)} &= \widetilde{A}^{(j)} - B^{(j)} A^{(j)^{-1}} \widetilde{C}^{(j)}, \\ \widehat{A}^{(j+1)} &= A^{(j)} - (B^{(j)} \widehat{A}^{(j)^{-1}} \widehat{C}^{(j)} + C^{(j)} A^{(j)^{-1}} B^{(j)}), \\ \widehat{C}^{(j+1)} &= C^{(j+1)}, \quad \widetilde{C}^{(j)} = -C^{(j)} A^{(j)^{-1}} \widetilde{C}^{(j)} \end{aligned} \tag{1.15}$$

where $A_0 = A, B_0 = B, C_0 = C$, $\widetilde{A}_0 = \widetilde{A}, \widehat{A}_0 = \widehat{A}, \widetilde{C}_0 = \widetilde{C}$ and $\widehat{C}_0 = \widehat{C}$.

If we apply the technique of Harrod and Plemmons in [11], [10] based on (1.5), we have to solve a nonsingular linear system whose matrix has the same structure as that of (1.6). The only difference is that the blocks $\widehat{H}_i$, $i = 2, \ldots, n$, now have dimension $(k-1) \times k$, the last block $\widehat{H}_1$ on the main diagonal has dimension $(k-1) \times (k-1)$ and the block H_0 in the last block column has dimension $k \times (k-1)$.

By following the same lines shown in this section, it is easy to prove that one step of the cyclic reduction method leaves unchanged this specific structure of the matrix, and it is possible to devise an analogous method for solving the linear system in this slightly different case. We leave the details to the reader.

4 THE ALGORITHM AND ITS ANALYSIS.

The cyclic reduction method of section 3 can be used to compute the probability invariant vector $\boldsymbol{\pi}$, which solves the system

$$(I - A^T)\boldsymbol{\pi} = \mathbf{0}. \tag{1.16}$$

A first approach consists in applying this method directly to the homogeneous system (1.16) (compare (1.3),(1.4a)). A second approach, of [10], [11], [1], is that described at the end of section 2, and consists in reducing the homogeneous system (1.16) into a nonhomogeneous system obtained by removing the last equation of (1.16) (compare (1.5)). Here we describe only the algorithm derived by the first approach and we leave to the reader the analysis of the algorithm based on (1.5).

Let us partition the vector $\boldsymbol{\pi}$ into blocks of length k according to the block structure of H, i.e., set $\boldsymbol{\pi}^T = (\boldsymbol{\pi}_0^T, \boldsymbol{\pi}_1^T, \ldots, \boldsymbol{\pi}_{n-1}^T)$. Moreover, denote by $\boldsymbol{\pi}_-^{(j)}$ the vector having blocks $\boldsymbol{\pi}_{i2^j}$, $i = 0, \ldots, 2^{p-j} - 1$, $j = 0, 1, \ldots, p-1$, where $n = 2^p$, $p \geq 2$. Similarly, denote by $\boldsymbol{\pi}_+^{(j)}$ the vector having blocks $\boldsymbol{\pi}_{(2i+1)2^{j-1}}$, $i = 0, \ldots, 2^{p-j} - 1$, for $j = 1, \ldots, p-1$. In this way, at the j-th recursive step of the cyclic reduction algorithm applied to (1.1) we have the system $H^{(j)}\boldsymbol{\pi}_-^{(j)} = \mathbf{0}$, $j = 0, \ldots, p-1$. After the odd-even permutation on the rows and the columns and after performing one step of Gaussian elimination we obtain the system

$$\begin{gathered} \begin{pmatrix} T_1^{(j)} & W^{(j)} \\ O & H^{(j+1)} \end{pmatrix} \begin{pmatrix} \boldsymbol{\pi}_+^{(j+1)} \\ \boldsymbol{\pi}_-^{(j+1)} \end{pmatrix} = \mathbf{0}, \\ H^{(j+1)} = T_2^{(j)} - Z^{(j)} {T_1^{(j)}}^{-1} W^{(j)}, \end{gathered} \tag{1.17}$$

(compare (1.9)). At the end of the cyclic reduction algorithm we have to solve the $2k \times 2k$ homogeneous system $H^{(p-1)}\boldsymbol{\pi}_-^{(p-1)} = \mathbf{0}$, $\boldsymbol{\pi}_-^{(p-1)} = \begin{pmatrix} \boldsymbol{\pi}_0 \\ \boldsymbol{\pi}_{2^{(p-1)}} \end{pmatrix}$, and we may recover the remaining components of the vector $\boldsymbol{\pi}$ by means of back substitution by using the following relation derived by (1.17):

$$\boldsymbol{\pi}_+^{(j+1)} = -{T_1^{(j)}}^{-1} W^{(j)} \boldsymbol{\pi}_-^{(j+1)}, \quad j = p-2, \ldots, 0$$

The most expensive part of each recursive step of the cyclic reduction is to compute the blocks of the matrix $H^{(j)}$. For this computation we make use of the formulae (1.11), (1.12), developed in section 3 in the case $j = 0$ and easily extendable to the general case. These formulae allow us to reduce the computation of the blocks of the matrix $H^{(j)}$ to computations with block triangular block Toeplitz matrices. In fact, observe that the most expensive part of (1.12) is the computation of the matrix $F_1^{(0)} F_2^{(0)^{-1}} F_1^{(0)}$, where $F_1^{(0)}$, $F_2^{(0)}$ as well as $F_2^{(0)^{-1}}$ are block triangular block Toeplitz matrices.

Triangular Toeplitz matrices and polynomials are intimately connected [3]. Let $A = T(A_0, \ldots, A_{n-1})$, $B = T(B_0, \ldots, B_{n-1})$ be block triangular block Toeplitz matrices and let us associate with A and B the matrix polynomials $A(z) = \sum_{i=0}^{n-1} A_i z^i$, $B(z) = \sum_{i=0}^{n-1} B_i z^i$, respectively. Let $C(z) = \sum_{i=0}^{n-1} C_i z^i$ be such that $C(z) = A(z)B(z) \bmod z^n$. Then we have $AB = T(C_0, \ldots, C_{n-1})$. Moreover the matrices C_i can be computed by means of the evaluation/interpolation technique in $O(k^2 n \log n + k^3 n)$ arithmetic operations. More precisely, let q be a positive integer such that $m = 2^q > 2n$, let ω be a primitive m-th root of 1, and set $A_j = B_j = O$, for $j = n+1, \ldots, m$. If $P_i = \sum_{j=0}^{m-1} \omega^{ij} A_j$, $Q_i = \sum_{j=0}^{m-1} \omega^{ij} B_j$, $i = 0, \ldots, m-1$, are the values that the matrix polynomials $A(z)$ and $B(z)$ take at the m-th roots ω^i, $i = 0, \ldots, m-1$, of 1, then the coefficients of $C(z)$ can be expressed as $C_i = \frac{1}{m} \sum_{j=0}^{m-1} \omega^{-ij} P_j Q_j$.

Observe that the three summations involving ω correspond to $3k^2$ discrete Fourier transforms of vectors of m real components, i.e., a transform for each entry of the three $k \times k$ matrices involved. The cost of their computation is $3k^2 \phi(m)$, where we denote with $\phi(m)$ the number of real arithmetic operations needed for performing a discrete Fourier transform of a vector having m real components. We recall that, if m is an integer power of 2, then $\phi(m) = (5/2) m \log_2 m$, up to within constant terms. Since the complex matrices P_j and Q_j are obtained as discrete Fourier transforms of real matrices, the computation of the complex matrix products $P_j Q_j$, $j = 0, \ldots, m-1$ can be reduced to $m/2$ complex matrix multiplications and two real matrix multiplications. Each complex matrix product can be computed with three real matrix multiplication and with 5 matrix additions, so that the computation of the matrix products $P_j Q_j$, $j = 0, \ldots, m-1$, can be carried out with about $m(3k^3 + k^2)$ real operations. Thus, the overall cost of computing the product of the matrices A and B (block convolution) is about $(15/2) k^2 n \log_2 n + 3nk^3$ real arithmetic operations. Since the first block column of C is the product of A and the first block column of B, the same cost bound applies to the computation of the product of a block triangular block Toeplitz matrix and a block vector.

A similar cost bound holds for matrix inversion, in fact, it is easy to prove that the inverse of a nonsingular block triangular block Toeplitz matrix A is still a block triangular block Toeplitz matrix uniquely defined by its first block column. Moreover, partitioning A into four square blocks $A = \begin{pmatrix} M_1 & O \\ M_2 & M_1 \end{pmatrix}$, we find that $A^{-1} = \begin{pmatrix} M_1^{-1} & O \\ -M_1^{-1}M_2M_1^{-1} & M_1^{-1} \end{pmatrix}$ i.e., the first block column of A^{-1} can be obtained by inverting a block triangular block Toeplitz matrix of half the size and computing the first block column of the product of three block Toeplitz matrices. This leads to a recursive algorithm that have been discovered by Lafon for scalar matrices, and which corresponds to the Sieveking-Kung algorithm (see [15], [3]) in the polynomial setting. The cost γ_n of this algorithm, which can be easily recovered by the recurrence $\gamma_n = \gamma_{\frac{n}{2}} + 5k^2\phi(n) + 3k^3n$, is given by $\gamma_n \le 10k^2\phi(n) + 6k^3n$. It is easy to rephrase triangular Toeplitz matrix inversion in terms of polynomials. In fact, if $A^{-1} = T(B_0, \dots, B_{n-1})$ then $B(z) = A(z)^{-1} \bmod z^n$.

The algorithm thus obtained for the computation of the probability invariant vector $\boldsymbol{\pi}$ can be carried out in the following way.

- **Stage 1.** Let $H^{(0)} = H$ and, for $j = 1, \dots, p-1$, recursively apply the cyclic reduction algorithm and compute the $k \times k$ matrices $H_i^{(j)}$, $\widetilde{H}_{i+1}^{(j)}$, $i = 0, 1, \dots, n_j - 2$, and $\widehat{H}_{i+1}^{(j)}$, $i = 0, 1, \dots, n_j - 1$, $n_j = n/2^j$, defining the $n_j \times n_j$ block matrices $H^{(j)} = T_2^{(j-1)} - Z^{(j-1)}T_1^{(j-1)^{-1}}W^{(j-1)}$, by means of the formulae (1.17)–(1.12) of section 3, there represented for $j = 0$ and here adjusted for the recursive j-th step of the cyclic reduction algorithm as follows

$$\begin{aligned} R^{(j+1)} &= F_0^{(j)} - F_1^{(j)}F_2^{(j)^{-1}}F_1^{(j)} \\ U^{(j+1)} &= \begin{pmatrix} \widetilde{H}_1^{(j)} \\ \widetilde{H}_3^{(j)} \\ \vdots \\ \widetilde{H}_{n_j-3}^{(j)} \end{pmatrix} - F_1^{(j)}F_2^{(j)^{-1}} \begin{pmatrix} \widetilde{H}_2^{(j)} \\ \widetilde{H}_4^{(j)} \\ \vdots \\ \widetilde{H}_{n_j-2}^{(j)} \end{pmatrix}, \qquad (1.18) \\ (S^{(j+1)}, V^{(j+1)}) &= (\widetilde{H}_{n_j-1}^{(j)}, H_{n_j-3}^{(j)}, \dots, H_1^{(j)}) - \\ &\qquad (H_{n_j-2}^{(j)}, H_{n_j-4}^{(j)}, \dots, H_0^{(j)})T_1^{(j)^{-1}}W^{(j)} \end{aligned}$$

  where $F_0^{(j)} = T(O, H_1^{(j)}, H_3^{(j)}, \dots, H_{n_j-5}^{(j)})$, $F_1^{(j)} = T(H_0^{(j)}, H_2^{(j)}, \dots, H_{n_j-4}^{(j)})$, $F_2^{(j)} = T(H_1^{(j)}, H_3^{(j)}, \dots, H_{n_j-3}^{(j)})$.

- **Stage 2.** Solve, by means of LU or $\tilde{U}\tilde{L}$ factorization, the $2k \times 2k$ system $H^{(p-1)}\boldsymbol{\pi}_-^{(p-1)} = \mathbf{0}$.

- **Stage 3.** Compute $\boldsymbol{\pi}_+^{(j+1)} = -T_1^{(j)^{-1}} W^{(j)} \boldsymbol{\pi}_-^{(j+1)}$ for $j = p-2, \ldots, 0$.
- **Stage 4.** Normalize the vector $\boldsymbol{\pi}$.

Observe that the entries of the matrices $T_1^{(j)^{-1}}$, $-W^{(j)}$, $j = 0, \ldots, p-2$, are nonnegative. Therefore the computation of the vector $\boldsymbol{\pi}$, given the entries of the matrices $T_1^{(j)^{-1}}$, $-W^{(j)}$, is numerically stable since it requires only additions of nonnegative numbers. Performing convolutions and deconvolutions by means of the FFT is a weakly stable numerical computation (compare [18]). This fact, together with the nonnegativeness of the entries of the matrices $-W^{(j)}$, $T_1^{(j)^{-1}}$ and of the vector $\boldsymbol{\pi}$, leads to the weak numerical stability of the above algorithm when applied together with the FFT-based techniques for Toeplitz computation. If the FFT is not used for convolutions then the resulting algorithm is strongly stable.

Let us analyze the computational cost of the algorithm. By formulae (1.18), at each step of the cyclic reduction scheme, stage 1 involves one inversion of a triangular Toeplitz matrix and five products between a block triangular block Toeplitz matrix and a block vector. Stage 2 has a cost independent of n, stage 3 involves two products of block triangular block Toeplitz matrices and a vector. Thus, the overall cost of the j-th step is $\theta_{n_j} = O(k^2 n_j \log n_j + k^3 n_j)$ arithmetic operations and the overal cost of the algorithm to compute $\boldsymbol{\pi}$ by means of cyclic reduction is $\theta_n + \theta_{\frac{n}{2}} + \ldots + \theta_2 = O(\theta_{2n}) = O(k^2 n \log n + k^3 n)$ arithmetic operations. On the other hand, it can be easily verified that the computational cost obtained by applying the customary block Gaussian elimination is $O(k^3 n^2)$. Moreover, Gaussian elimination with the diagonal adjustment of [9] is strongly numerically stable ([5]). It is worth pointing out that for moderate values of n below a certain threshold value $\tilde{n}$, the customary algorithms for matrix-vector product are faster. The value $\tilde{n}$ depends on the block size k, the larger is k the smaller is $\tilde{n}$. In the case where the matrix H has the block tridiagonal structure (1.14) the computations are simplified: by formulae (1.15) it is easily seen that each step of cyclic reduction involves two inversions of $k \times k$ matrices and ten products between $k \times k$ matrices. Hence, the overall cost for the computation of $\boldsymbol{\pi}$ is $O(k^3 \log n)$, against the $O(k^3 n)$ cost of Gaussian elimination.

5 THE INFINITE CASE.

The cyclic reduction tecnique for successive state reduction can also be applied to stochastic matrices A of $M/G/1$ type, i.e., to infinite matrices of the form

(1.1). Here we assume that A is irreducible so that the associated Markov chain is nonnull recurrent. We also introduce the vector norm $||\mathbf{v}|| = \sum_i |v_i|$, both for finite and infinite vectors $\mathbf{v} = (v_i)$, and the induced matrix norm $||B|| = \sup_j \sum_i |b_{i,j}|$, both for finite and infinite matrices B. Indeed, analogously to what we have done in section 3, we may consider the infinite block submatrices $T_1^{(0)}$, $T_2^{(0)}$, $Z^{(0)}$, $W^{(0)}$ of

$$H = I - A^T = \begin{pmatrix} \widetilde{H}_1 & H_0 & & & \bigcirc \\ \widetilde{H}_2 & H_1 & H_0 & & \\ \widetilde{H}_3 & H_2 & H_1 & H_0 & \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

obtained by considering separately block rows and block columns having odd, even position, respectively, i.e.,

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

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

Similarly, define the vectors $\boldsymbol{\pi}_+^{(1)T} = (\boldsymbol{\pi}_1^T, \boldsymbol{\pi}_3^T, \boldsymbol{\pi}_5^T, \ldots)$ and $\boldsymbol{\pi}_-^{(1)T} = (\boldsymbol{\pi}_0^T, \boldsymbol{\pi}_2^T, \boldsymbol{\pi}_4^T, \ldots)$, where $\boldsymbol{\pi}^T = (\boldsymbol{\pi}_0^T, \boldsymbol{\pi}_1^T, \boldsymbol{\pi}_2^T, \ldots)$ is the probability invariant vector associated with A, i.e., $H\boldsymbol{\pi} = \mathbf{0}$.

One step of cyclic reduction yields the matrix $H^{(1)} = T_2^{(0)} - Z^{(0)} T_1^{(0)^{-1}} W^{(0)}$ such that $A^{(1)T} = I - H^{(1)}$ is a stochastic matrix of the $M/G/1$ type (1.1). Moreover, $H^{(1)}\boldsymbol{\pi}_-^{(1)} = \mathbf{0}$. More generally, the matrix

$$H^{(j+1)} = T_2^{(j)} - Z^{(j)} T_1^{(j)^{-1}} W^{(j)}, \tag{1.19}$$

generated by applying the same technique to $H^{(j)}$, is such that $A^{(j+1)T} = I - H^{(j+1)}$ is a stochastic matrix of the $M/G/1$ type (1.1) and $H^{(j+1)}\boldsymbol{\pi}_-^{(j+1)} = \mathbf{0}$. Explicit recursive formulae for the block entries $\left(\widetilde{A}_i^{(j)}\right)_{i \in \mathbf{N}^+}$, $\left(A_i^{(j)}\right)_{i \in \mathbf{N}}$ defining the matrices $A^{(j)^T} = I - H^{(j)}$ are rather complicate. However, we can give a simple matrix functional representation of such stochastic matrices, by using the natural isomorphism between formal power series and triangular Toeplitz matrices. More precisely, we associate with the infinite block triangular block Toeplitz matrix T, defined by its first block column $[S_0^T, S_1^T, S_2^T, \ldots]^T$, the formal matrix power series $\varphi_T(z) = \sum_{j=0}^{\infty} S_j z^j$, called the *generating function* of the matrix T. Generating functions allow us to express, in terms of formal power series, the product between block lower triangular block Toeplitz matrices and the inverse (if it exists) of block lower triangular block Toeplitz matrices. In fact, the following straightforward result holds:

Proposition 3 *Let T_1 and T_2 be two infinite block lower triangular, block Toeplitz matrices, with blocks of the same dimension. Let $\varphi_{T_1}(z)$ and $\varphi_{T_2}(z)$ their generating functions, respectively. Then the matrix $T = T_1T_2$ is a block lower triangular, block Toeplitz matrix whose generating function $\varphi_T(z)$ is such that $\varphi_T(z) = \varphi_{T_1}(z)\varphi_{T_2}(z)$, i.e., the row by column product of $\varphi_{T_1}(z)$ and $\varphi_{T_2}(z)$. Moreover, if T_1 is nonsingular, then the inverse matrix T_1^{-1} is a block lower triangular, block Toeplitz matrix and its generating function $\varphi_{T_1^{-1}}(z)$ satisfies $\varphi_{T_1}(z)\varphi_{T_1^{-1}}(z) = I$.*

We associate with the stochastic matrix $A^{(j)^T} = I - H^{(j)}$, obtained at the j-th step of cyclic reduction, the formal power series $\widetilde{\varphi}^{(j)}(z) = \sum_{i=0}^{\infty} \widetilde{A}_{i+1}^{(j)} z^i$, $\varphi^{(j)}(z) = \sum_{i=0}^{\infty} A_i^{(j)} z^i$, which represent the first and the second block column of $A^{(j)^T}$. Denote with

$$\widetilde{\varphi}_{odd}^{(j)}(z) = \sum_{i=0}^{\infty} \widetilde{A}_{2i+1}^{(j)} z^i, \qquad \widetilde{\varphi}_{even}^{(j)}(z) = \sum_{i=0}^{\infty} \widetilde{A}_{2(i+1)}^{(j)} z^i,$$

$$\varphi_{even}^{(j)}(z) = \sum_{i=0}^{\infty} A_{2i}^{(j)} z^i, \qquad \varphi_{odd}^{(j)}(z) = \sum_{i=0}^{\infty} A_{2i+1}^{(j)} z^i,$$

and observe that $\varphi_{even}^{(j)}(z^2) = (\varphi^{(j)}(z) + \varphi^{(j)}(-z))/2$, $\varphi_{odd}^{(j)}(z^2) = (\varphi^{(j)}(z) - \varphi^{(j)}(-z))/(2z)$.

For the properties of the generating functions and (1.19) it holds that the first and the second block column of the matrix $A^{(j+1)^T}$ are determined by the formal power series

$$\begin{cases} \widetilde{\varphi}^{(j+1)}(z) = \widetilde{\varphi}_{odd}^{(j)}(z) + \varphi_{even}^{(j)}(z)\Big(I - \varphi_{odd}^{(j)}(z)\Big)^{-1} \widetilde{\varphi}_{even}^{(j)}(z) \\ \varphi^{(j+1)}(z) = z\varphi_{odd}^{(j)}(z) + \varphi_{even}^{(j)}(z)\Big(I - \varphi_{odd}^{(j)}(z)\Big)^{-1} \varphi_{even}^{(j)}(z) \end{cases} \tag{1.20}$$

which allow us to represent in functional form the matrices $A^{(j)}$ generated by cyclic reduction. Hence, we have a recursive formula expressed in terms of formal power series.

We observe that $\boldsymbol{\pi}_-^{(j)}$ converges to the vector $(\boldsymbol{\pi}_0^T, 0, 0, \ldots)^T$, as j tends to infinity; in fact, $\boldsymbol{\pi}_-^{(j)} = (\boldsymbol{\pi}_0^T, \boldsymbol{\pi}_{2^j}^T, \boldsymbol{\pi}_{2\cdot 2^j}^T, \ldots)^T$ and $\lim_{j\to\infty} \boldsymbol{\pi}_{i\cdot 2^j} = 0$, for $i = 1, 2, \ldots$ because of $||\boldsymbol{\pi}|| = 1$.
Since $\boldsymbol{\pi}_0 > 0$, due to the irreducibility assumption (compare [4]), and $\widetilde{A}_i^{(j)}$, $A_i^{(j)} \geq 0$, from the condition $H^{(j)}\boldsymbol{\pi}_-^{(j)} = \mathbf{0}$ we deduce that $\lim_{j\to\infty} \widetilde{A}_i^{(j)} = O$, $i \geq 2$. On the other hand, since $\widetilde{\varphi}^{(j)}(0) = \widetilde{\varphi}_{odd}^{(j)}(0) = \widetilde{A}_1^{(j)}$, from (1.20) we have that the sequence $\{\widetilde{A}_1^{(j)}\}_j$ is monotonically increasing and, since $||\widetilde{A}_1^{(j)}|| \leq 1$ we have that there exists $\lim_{j\to\infty} \widetilde{A}_1^{(j)} = W$ and it is a stochastic matrix. We can also estimate the speed of convergence to zero of the blocks $\widetilde{A}_i^{(j)}$, $i \geq 2$, as j tends to the infinity. In fact, since $\sum_{i=0}^{+\infty} ||\boldsymbol{\pi}_i|| = 1$, we have $||\boldsymbol{\pi}_i|| \leq c/i$, for $i \geq 1$, and for a suitable positive constant c. Therefore, from the condition $H^{(j)}\boldsymbol{\pi}_-^{(j)} = \mathbf{0}$, it follows that $||\widetilde{A}_i^{(j)}\boldsymbol{\pi}_0|| \leq \frac{c}{2^j}\sum_{r=0}^{i-1}\frac{||H_r^{(j)}||}{i-r} \leq \frac{c'}{2^j}$, for $i \geq 2$ and for a constant c'. Hence $||\widetilde{A}_i^{(j)}||$ converges to zero, as j tends to the infinity, at least exponentially.

Moreover, if $||\boldsymbol{\pi}_i|| \leq \lambda^i$, for $i \geq 0$, $0 < \lambda < 1$, we have a double exponential convergence. In fact, by following the same argument used above, we arrive at the inequality $||\widetilde{A}_i^{(j)}\boldsymbol{\pi}_0|| \leq \lambda^{2^j}\sum_{r=0}^{i-1}||H_r^{(j)}||\lambda^{i-r} \leq c''\lambda^{2^j}, \quad i \geq 2$, for a suitable constant c''. In this way we may apply the technique of cyclic reduction in order to approximate the matrix W and the components of the vector $\boldsymbol{\pi}_0$. For instance, the iterations can be performed until the matrix $\widetilde{A}_1^{(j)}$ satisfies the condition $||\mathbf{e}^T\widetilde{A}_1^{(j)} - \mathbf{e}^T|| \leq \epsilon$ for a fixed positive ϵ. Then, a vector $\widetilde{\boldsymbol{\pi}}_0$ such that $\widetilde{A}_1^{(j)}\widetilde{\boldsymbol{\pi}}_0 = \lambda\widetilde{\boldsymbol{\pi}}_0$ can be computed, where λ is the eigenvalue of $\widetilde{A}_1^{(j)}$ having maximum modulus. Finally, setting $\widetilde{\boldsymbol{\pi}}_{i2^j} = \mathbf{0}$, $i \geq 1$, we may recover, by means of back substitution, the components $\widetilde{\boldsymbol{\pi}}_i$ through the formula $\widetilde{\boldsymbol{\pi}}_+^{(r+1)} = -{T_1^{(r)}}^{-1}W^{(r)}\widetilde{\boldsymbol{\pi}}_-^{(r+1)}, \quad r = j-1, \ldots, 0$.

Finally, we observe that in order to compute the entries of the matrix $\widetilde{A}_1^{(j)}$, only the blocks $A_0, \ldots, A_{2^j-1}$, $\widetilde{A}_1, \ldots, \widetilde{A}_{2^j}$ are involved. This allows us to apply the functional relation (1.20) modulo z^{2^j}, thus reducing the whole computation with power series to computation with polynomials. For this task we may apply the tools of section 3 for an overall computational cost of $O(k^2 2^j j + k^3 2^j)$ arithmetic operations. This cost strictly depends on the decay rate of $||\boldsymbol{\pi}_j||$, this problem, together with the comparison of our approach with the Ramaswami formula [19] is under investigation.

In the case of a block tridiagonal input matrix our method is very similar to that of Latouche and Ramaswami [17], where the matrix G solving the equation $G = A_0^T + A_1^T G + A_2^T G^2$ is found as a limit of a sequence of matrices obtained by applying a modified cyclic reduction scheme to the matrix A. It is possible to prove that in the case of block tridiagonal matrices, the matrices G and W are related by the following equation $W^T = \widetilde{A}_1^T + \widetilde{A}_2^T G$.

We have implemented, in fortran, the cyclic reduction algorithm for the computation of the probability invariant vector of an infinite stochastic matrix of M/G/1 type. We adopted the strategy of setting to zero the matrices $A_i^{(j)}$ for $i > i_0$ and $\widetilde{A}_i^{(j)}$ for $i > \tilde{i}_0$, if $\mathbf{e}^T(I - \sum_{s=0}^{i_0} A_s^{(j)}) < \epsilon \mathbf{e}^T$, $\mathbf{e}^T(I - \sum_{s=0}^{\tilde{i}_0} \widetilde{A}_s^{(j)}) < \epsilon \mathbf{e}^T$ for $\epsilon = 10^{-12}$, where the above inequalities hold componentwise. Moreover we adjusted the diagonal entry of the stochastic matrices obtained in this way by means of the GTH technique of [9]. With this choice, the Toeplitz-like matrices involved at the j-th step of the algorithm are banded and the computation is performed with FFT's of length depending on the bandwidth of the matrices. It is worth saying that with this choice we do not have to cut the infinite matrix to a preassigned size, and the number of steps of the computation is determined dynamically.

We have tested our program with a block matrix A arising from a queueing problem with Poisson arrivals. The $k \times k$ blocks of the matrix A are defined in the following way: $(A_h)_{i,j} = \alpha_{kh+i-j}$, $(\widetilde{A}_{h+1})_{i,j} = \alpha_{kh+i-1}$, $i, j = 1, \ldots, k$, $h = 1, 2, \ldots$, where $\alpha_i = 0$ for $i < 0$, $\alpha_i = e^{-\lambda}\frac{\lambda^i}{i!}$, $i \geq 0$, $0 < \lambda < k$, and $(A_h)_{i,j}$ denotes the (i, j) entry of the matrix A_h. We chosed $k = 50$ and $\lambda \in \{30, 35, 40, 45, 46, 47, 48, 49\}$. We observe that the value $\lambda = k$ is a critical point: if $\lambda < k$ then the Markov chain is positive recurrent, for $\lambda > k$ the chain is transient.

We implemented two versions of our algorithm. The version RCF which makes use of the FFT-based algorithms for Toeplitz computations and the RCC version which adopts the customary algorithms for computing the products of matrices. All the implementations have been performed in the double precision IEEE floating point arithmetic.

The following table reports the time t (seconds) needed for the computation of b blocks of the probability invariant vector by means of RCF, RCC and Ramaswami formula (denoted by RAM) together with the maximum component

r of the residual vector. Moreover we report also the number s of steps needed by RCF and RCC and the number S of steps that are needed to solve the equation $G = \sum_{i=0}^{+\infty} A_i^T G^i$ by means of the recurrence $G_{j+1} = \sum_{i=0}^{+\infty} A_i^T G_j^i$, $G_0 = A_0^T$, within the error 10^{-12}.

Table 1.

λ	RCC	RCF	RAM
30	$t = 2.87,\ r = 10^{-17}$	$t = 2.70,\ r = 10^{-17}$	$t = 2.80,\ r = 10^{-17}$
	$b = 2,\ s = 1$	$b = 2,\ s = 1$	$b = 2,\ S = 8$
35	$t = 2.80,\ r = 10^{-17}$	$t = 2.64,\ r = 10^{-17}$	$t = 4.20,\ r = 10^{-17}$
	$b = 2,\ s = 1$	$b = 2,\ s = 1$	$b = 2,\ S = 13$
40	$t = 4.50,\ r = 10^{-17}$	$t = 4.30,\ r = 10^{-17}$	$t = 8.30,\ r = 10^{-17}$
	$b = 4,\ s = 2$	$b = 4,\ s = 2$	$b = 4,\ S = 28$
45	$t = 5.80,\ r = 10^{-16}$	$t = 5.30,\ r = 10^{-17}$	$t = 26.5,\ r = 10^{-17}$
	$b = 4,\ s = 2$	$b = 4,\ s = 2$	$b = 4,\ S = 92$
46	$t = 7.50,\ r = 10^{-17}$	$t = 6.80,\ r = 10^{-17}$	$t = 36.5,\ r = 10^{-17}$
	$b = 8,\ s = 3$	$b = 8,\ s = 3$	$b = 8,\ S = 130$
47	$t = 7.50,\ r = 10^{-16}$	$t = 7.00,\ r = 10^{-16}$	$t = 56.0,\ r = 10^{-16}$
	$b = 8,\ s = 3$	$b = 8,\ s = 3$	$b = 8,\ S = 199$
48	$t = 7.30,\ r = 10^{-15}$	$t = 7.20,\ r = 10^{-15}$	$t = 94.3,\ r = 10^{-14}$
	$b = 8,\ s = 3$	$b = 8,\ s = 3$	$b = 8,\ S = 340$
49	$t = 10,\ r = 10^{-16}$	$t = 9.5,\ r = 10^{-16}$	$t = 210,\ r = 10^{-15}$
	$b = 16,\ s = 4$	$b = 16,\ s = 4$	$b = 16,\ S = 752$

Acknowledgements

The authors would like to thank the anonymous referees for valuable suggestions.

REFERENCES

[1] J. L. Barlow, *On the smallest positive singular value of a singular M-matrix with applications to ergodic Markov Chains,* SIAM J. Alg. Discr. Meth., 7 (1986), pp. 414-424.

[2] J. L. Barlow, *Error bounds for the computation of null vectors with application to Markov chains,* SIAM J. Matrix Anal. Appl., 14 (1993), pp. 798-812.

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

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

[5] C.A. O'Cinneide, *Entrywise perturbation theory and error analysis for Markov chains,* Numer. Math., 65, (1993), pp. 109-120.

[6] R. E. Funderlic, J. B. Mankin, *Solution of homegeneous systems of linear equations arising in compartmental models,* SIAM J. Sci. Stat. Comput., 2 (1981), pp. 375-383.

[7] R. E. Funderlic, R. J. Plemmons, *LU decomposition of M-matrices by elimination without pivoting,* Linear Algebra Appl., 41 (1981), pp. 99-110.

[8] R. E. Funderlic, R. J. Plemmons, *A combined direct-iterative method for certain M-matrix linear systems,* SIAM J. Alg. Discr. Meth., 5 (1984), pp. 33-42.

[9] 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.

[10] W. J. Harrod, *Rank modification methods for certain singular systems of linear equations,* Ph. D. thesis, Univ. of Tennessee, Knoxville. TN, Department of Mathematics, 1984.

[11] W. J. Harrod, R. J. Plemmons, *Comparison of some direct methods for computing stationary distributions of Markov chains,* J. Sci. Stat. Comput., 5 (1984), pp. 453-469.

[12] I.S. Iohvidov, *Hankel and Toeplitz Matrices and Forms, Algebraic Theory,* Birkhäuser, Boston 1982.

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

[14] L. Kaufman, *Matrix methods for queueing problems,* J. Sci. Stat. Comput., 4 (1983), pp. 525-552.

[15] D.E. Knuth, *The Art of Computer Programming: Seminumerical Algorithms, 2,* Addison Wesley, Reading, Mass., 1981.

[16] R. Kouri, D. F. McAllister, W. J. Stewart, *A numerical comparison of block iterative methods with aggregation for computing stationary probability vectors for nearly completely decomposable Markov Chains,* SIAM J. Alg. Discr. Meth., 5 (1984), pp. 164-186.

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

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

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

[20] D. J. Rose, *Convergent regular splittings for singular M-matrices,* SIAM J. Alg. Discr. Meth., 5 (1984), pp. 133-144.

[21] G.W. Stewart, *On the solution of block Hessenberg systems,* CS-TR-2973, Department of Computer Science, University of Maryland, 1992, To appear in Numerical Linear Algebra and Applications.

[22] R. S. Varga, D.-Y. Cai, *On the LU factorization of M-matrices,* Numer. Math., 38 (1981), pp. 179-192.

[23] J. Ye, San-qi Li, *Analisys of multi-media traffic queues with finite buffer and overload control—Part I: algorithm,* Proc. IEEE Infocom 91, Bal Harbour, 1464–1474.