Efficient computation of the extreme solutions of $X + A^*X^{-1}A = Q$ and $X - A^*X^{-1}A = Q$

Beatrice Meini[*]

January 18, 2000

Abstract

We propose a new quadratically convergent algorithm, having a low computational cost per step and good numerical stability properties, that allows the simultaneous approximation of the extreme solutions of the matrix equations $X + A^*X^{-1}A = Q$ and $X - A^*X^{-1}A = Q$. The algorithm is based on the cyclic reduction method.

Key words: Matrix equation, cyclic reduction, block Toeplitz matrix.
AMS Classification: 15A24, 65F10, 65H05.

1 Introduction

In this paper we are concerned with the design and analysis of a quadratically convergent algorithm for computing the extremal solutions of the matrix equations

$$X + A^*X^{-1}A = Q, \tag{1}$$
$$X - A^*X^{-1}A = Q, \tag{2}$$

where Q is an $m \times m$ Hermitian positive definite matrix, A is an $m \times m$ matrix, and A^* denotes the conjugate transposed of A. An algorithm is quadratically convergent (or, equivalently, has a double exponential convergence) if it generates a sequence X_n of approximations to the solution X such that $||X - X_n|| \le \gamma\sigma^{2^n}$, for positive constants $\sigma < 1$ and γ, where $||\cdot||$ denotes any operator norm; if the bound $||X - X_n|| \le \gamma\sigma^n$ holds, then we say that the convergence is linear.

In [8] it is shown that, if (1) has a positive definite solution X, then there exist minimal and maximal solutions X_- and X_+, respectively, such that $0 < X_- \le X \le X_+$ for any positive definite solution X. Here and hereafter, if X and Y are Hermitian matrices, $X \le Y$ ($X < Y$) means that $Y - X$ is positive semidefinite (definite).

*Dipartimento di Matematica, Università di Pisa, via Buonarroti 2, 56127 Pisa, Italy, meini@dm.unipi.it

Concerning equation (2), in [9] it is proved that there always exists a unique positive definite solution X_+, that is the largest one, and, if A is nonsingular, there exists a unique negative definite solution X_-, that is the smallest one. Here maximal and minimal must be intended according to the ordering for Hermitian matrices that we have introduced above.

We will refer to X_- and X_+ as to the extreme solutions of (1), or of (2).

Equations (1) and (2) arise in a wide variety of research areas, that include control theory, ladder networks, dynamic programming, stochastic filtering and statistics (see [2, 23] for a list of references concerning (1), and [9] for references about equation (2)). Recently, theoretical properties of the extreme solutions and numerical methods for their computation have been investigated by several authors (see the papers [13, 12, 9, 23, 24, 7, 8, 2]). The available numerical methods are mainly based on fixed point iterations, or on applications of Newton's algorithm.

In the recent paper [13] the authors analyze convergence properties of fixed point iterations and of Newton's method for computing the maximal solutions of (1) and (2). They show that the rate of convergence of both the algorithms is related to the spectral radius $\rho(X_+^{-1}A)$ of the matrix $X_+^{-1}A$. More precisely, $\rho(X_+^{-1}A)$, that is less than or equal to one, expresses the asymptotic average reduction per step of the error, so that fixed point iterations, having a linear convergence, slow down if $\rho(X_+^{-1}A)$ is close to 1, and it may be necessary to apply Newton's method, having a double exponential convergence. On the other hand, the drawback of the application of Newton's method is the large computational cost per iteration; in fact, at each step the solution X of the matrix equation $X + BXC = D$, where X, B, C, D are $m \times m$ matrices, must be computed.

In this paper we devise a new algorithm that provides the simultaneous approximation of the extreme solutions of (1), or of (2), that has a double exponential convergence, like Newton's method, and a low computational cost for each step, like fixed point iterations. Moreover, the algorithm shows good numerical stability properties.

The idea on which this algorithm is based arises from the closeness of equations (1) and (2) to certain problems related to Markov chains. Indeed, a similar (even though more complicated) matrix equation, is encountered in the numerical solution of Markov chains described by Quasi Birth and Death (QBD) processes (see [20]), where the fundamental problem is the computation of the minimal nonnegative solution G of

$$G = A_0 + A_1 G + A_2 G^2, \tag{3}$$

with A_0, A_1, A_2 nonnegative matrices such that $A_0 + A_1 + A_2$ is stochastic (i.e. $(A_0 + A_1 + A_2)\boldsymbol{e} = \boldsymbol{e}$ and $\boldsymbol{e}$ is the vector having all the entries equal to one). Here, minimal means that any other solution is component-wise larger than the minimal one. The matrix G is related to the minimal nonnegative solution U of the matrix equation

$$I - U - A_2(I - U)^{-1}A_0 = I - A_1, \tag{4}$$

in fact, $G = (I - U)^{-1}A_0$ (see [14, 16]). In the field of Markov chains, several algorithms have been developed for the numerical solution of (3), see [15, 17, 4, 5, 1, 18]. One of the most efficient methods, based on cyclic reduction [11], is proposed in [4, 5] and consists in generating a sequence of nonlinear matrix equations of the form (3), such that the matrix coefficients quadratically converge to the coefficients of a linear equation, that can be explicitly solved, and that provides the seeked solution G and the corresponding matrix U.

In this paper we show that the cyclic reduction algorithm can be efficiently applied for solving also the matrix equations (1) and (2). We observe that the matrix $X_+^{-1}A$ solves a quadratic matrix equation of the form (3) and, according to the ideas developed in [4, 5], we rewrite the matrix equation in terms of an infinite block tridiagonal block Toeplitz system, and apply the cyclic reduction algorithm to this system, generating a sequence of infinite block tridiagonal block systems, that are block Toeplitz except for the block entry in position $(1,1)$. Due to the nice spectral properties of $X_+^{-1}A$, we show that the sequence of matrices defining the infinite systems quadratically converges to a block diagonal matrix, and the first block diagonal entry of these matrices quadratically converges to X_+. More precisely, if we denote by X_n the block entry in position $(1,1)$ of the infinite matrix obtained at the n-th step of cyclic reduction, we prove that for equation (1), in the case where $\rho(X_+^{-1}A) < 1$, for any $\epsilon > 0$ and for any operator norm $||\cdot||$ we have $||I - X_nX_+^{-1}|| = O((\sigma + \epsilon)^{2\cdot 2^n})$, where $\sigma = \rho(X_+^{-1}A)$. For equation (2), it always holds $\rho(X_+^{-1}A) < 1$, thus the double exponential convergence is always guaranteed. If A is nonsingular, the algorithm provides the simultaneous approximation of X_-, for both the equations. Each step of the algorithm requires the solution of two linear systems and the computation of three matrix products; moreover, the conditioning of the matrices defining the linear systems is bounded from above by a constant.

The performed experiments show the effectiveness of our algorithm, that is about 60 times faster than Newton's method, and 1300 times faster than fixed point iterations.

The paper is organized as follows. In Section 2 we recall theoretical results about the existence and properties of the solutions. In Sections 3 and 4 we present the new algorithm and perform some comparisons with the known ones for the solution of (1) and (2), respectively. In Section 5 we report some numerical results.

2 Solutions of the matrix equations $X+A^*X^{-1}A = Q$ and $X - A^*X^{-1}A = Q$

In this section we recall conditions about the existence of the extreme solutions of (1) and (2), and some spectral properties of the matrix $X_+^{-1}A$, that will be used in the subsequent sections to show the convergence of our algorithm.

2.1 The equation $X + A^*X^{-1}A = Q$

Necessary and sufficient conditions for the existence of a positive definite solution of (1) are provided in [8]. More specifically, let us introduce the rational matrix function

$$\psi(\lambda) = \lambda A + Q + \lambda^{-1}A^*, \tag{5}$$

defined on the unit circle $\mathcal{C}$ of the complex plane, that is Hermitian for any $\lambda \in \mathcal{C}$. This function is said regular if there exists at least a $\lambda \in \mathcal{C}$ such that $\det \psi(\lambda) \neq 0$.

The following fundamental results hold [8]:

Theorem 2.1 *Equation (1) has a positive definite solution X if and only if $\psi(\lambda)$ is regular and $\psi(\lambda) \geq 0$ for all $\lambda \in \mathcal{C}$. In that case $\psi(\lambda)$ factors as*

$$\psi(\lambda) = (C_0^* + \lambda^{-1}C_1^*)(C_0 + \lambda C_1)$$

with $\det C_0 \neq 0$, and $X = C_0^*C_0$ is a solution of (1). Every positive definite solution is obtained in this way.

Theorem 2.2 *Assume that equation (1) has a positive definite solution. Then (1) has a largest and smallest solution X_+ and X_-, respectively. Moreover, X_+ is the unique solution such that $X + \lambda A$ is nonsingular for $|\lambda| < 1$, while X_- is the unique solution such that $X + \lambda A^*$ is nonsingular for $|\lambda| > 1$.*

Theorem 2.3 *Suppose A nonsingular. Then X solves*

$$X + A^*X^{-1}A = I \tag{6}$$

if and only if $Y = I - X$ solves

$$Y + AY^{-1}A^* = I. \tag{7}$$

In particular, if Y_+ is the largest solution of (7), then $X_- = I - Y_+$ is the smallest solution of (6).

The latter theorem can be easily generalized to the matrix equation (1) by multiplying (6) and (7) on the right and on the left by $Q^{1/2}$. Thus, if A is nonsingular, the smallest solution of (1) is $X_- = Q - Y_+$, where Y_+ is the largest solution of the equation

$$Y + AY^{-1}A^* = Q. \tag{8}$$

Let $r(T)$ be the numerical radius of an $m \times m$ matrix T, defined as

$$r(T) = \max\{|\boldsymbol{x}^*T\boldsymbol{x}| : \boldsymbol{x} \in \boldsymbol{C}^m, \quad \boldsymbol{x}^*\boldsymbol{x} = 1\}.$$

Concerning the spectral properties of $X_+^{-1}A$, it is proved in [23] that $\rho(X_+^{-1}A) \leq 1$. In [13] the authors characterize the eigenvalues of $X_+^{-1}A$ and give necessary and sufficient conditions so that $\rho(X_+^{-1}A) < 1$:

Theorem 2.4 *The eigenvalues of $X_+^{-1}A$ are precisely the eigenvalues of the matrix pencil $\lambda F - H$, inside the closed unit disk, with half of the partial multiplicities for each eigenvalue on the unit circle, where*

$$F = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -I & 0 \end{bmatrix}, \quad H = \begin{bmatrix} 0 & 0 & -I \\ Q & -I & A^* \\ -A & 0 & 0 \end{bmatrix}.$$

From the above theorem it follows that the eigenvalues of $X_+^{-1}A$ are the zeros in the closed unit disk of the polynomial $\det(-A + \lambda Q - \lambda^2 A^*)$.

Theorem 2.5 *It holds $\rho(X_+^{-1}A) < 1$ if and only if $r(Q^{-1/2}AQ^{-1/2}) < 1/2$.*

Moreover, in [8] the following necessary and sufficient condition is given:

Theorem 2.6 *It holds $\psi(\lambda) > 0$ for all $\lambda \in \mathcal{C}$ if and only if $r(Q^{-1/2}AQ^{-1/2}) < 1/2$.*

From the above theorems it follows that $X_+^{-1}A$ has spectral radius strictly less than one if and only if $\psi(\lambda)$ is positive definite on the unit circle.

2.2 The equation $X - A^*X^{-1}A = Q$

Concerning equation (2), in [9] the authors prove the following interesting results:

Theorem 2.7 *The rational matrix function*

$$\psi(\lambda) = \lambda A + Q - \lambda^{-1}A^* \tag{9}$$

is nonsingular for any $\lambda \in \mathcal{C}$.

Theorem 2.8 *The set of solutions of (2) is non-empty, and admits a maximal element X_+ and a minimal element X_-. X_+ is the unique positive definite solution, and if A is nonsingular, X_- is the unique negative definite solution. Moreover, if A is nonsingular, $X_- = Q - Y_+$, where Y_+ is the maximal solution of the equation*

$$Y - AY^{-1}A^* = Q. \tag{10}$$

Concerning the matrix $X_+^{-1}A$ in [9] it is proved that its spectral radius is strictly less than one, and in [13] the following characterization of the eigenvalues is given:

Theorem 2.9 *The eigenvalues of $X_+^{-1}A$ are precisely the eigenvalues of the matrix pencil $\lambda F - H$ inside the closed unit disk, where*

$$F = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -I & 0 \end{bmatrix}, \quad H = \begin{bmatrix} 0 & 0 & -I \\ Q & -I & -A^* \\ -A & 0 & 0 \end{bmatrix}.$$

3 Computation of the extreme solutions of $X + A^*X^{-1}A = Q$

In this section we describe the new algorithm, based on cyclic reduction, that has a computational cost roughly larger of a factor two with respect to fixed point iterations, and a double exponential convergence, like Newton's method; moreover, if A is nonsingular it allows the simultaneous computation of X_+ and X_-. Finally we compare our algorithm with the known methods.

Throughout this section we suppose $\psi(\lambda)$ regular and positive semidefinite for any $\lambda \in \mathcal{C}$, where $\psi(\lambda)$ is given in (5), so that the conditions of Theorem 2.1 are satisfied.

Let X be a solution of (1). Then, by multiplying on the right both sides of (1) by X^{-1}, we find that

$$-I + QX^{-1} - A^*X^{-1}AX^{-1} = 0. \tag{11}$$

Thus, the matrix $G = X^{-1}A$ solves the quadratic matrix equation

$$-A + QG - A^*G^2 = 0. \tag{12}$$

In addition, if Y is a solution of the matrix equation (8) then we have

$$-I + QY^{-1} - AY^{-1}A^*Y^{-1} = 0, \tag{13}$$

and thus $H = Y^{-1}A^*$ solves the quadratic matrix equation

$$-A^* + QH - AH^2 = 0. \tag{14}$$

In particular, if X_+ is the maximal solution of (1) and Y_+ is the maximal solution of (8), since $\rho(X_+^{-1}A) \le 1$ and $\rho(Y_+^{-1}A^*) \le 1$, then the matrix equations (12), (14) have a solution $G_+ = X_+^{-1}A$, $H_+ = Y_+^{-1}A^*$, respectively, with spectral radius at most 1. Moreover, for Theorem 2.4, the eigenvalues of G_+ and H_+ are

$$\lambda(G_+) = \{\lambda \ : \ g(\lambda) = 0, \ |\lambda| \le 1\},$$
$$\lambda(H_+) = \{\lambda \ : \ h(\lambda) = 0, \ |\lambda| \le 1\},$$

respectively, where $g(\lambda) = \det(-A + Q\lambda - A^*\lambda^2)$, $h(\lambda) = \det(-A^* + Q\lambda - A\lambda^2)$. Since $g(\lambda) = 0$ if and only if $g(\lambda)^* = 0$, and since

$$g(\lambda)^* = \det(-A^* + Q\bar{\lambda} - A\bar{\lambda}^2) = f(\bar{\lambda}),$$

we conclude that λ is an eigenvalue of G_+ if and only if $\bar{\lambda}$ is an eigenvalue of H_+. In particular it holds $\rho(G_+) = \rho(H_+)$.

The nice relation between the matrix equations (1), (8) and the quadratic matrix equations (12), (14), and the spectral properties of the solutions of the latter equations, allow us to derive a fast algorithm for the simultaneous computation of X_+ and Y_+.

If A is nonsingular, the matrix Y_+ enables us to recover the minimal solution X_- of (1) by means of the relation $X_- = Q - Y_+$.

The matrices X_+ and Y_+ can be simultaneously and efficiently computed by rewriting the matrix equations (11), (12), (13), (14) in terms of linear systems, and by applying the cyclic reduction algorithm, according to the ideas developed in [4, 5, 6]. In fact, we observe that the following equations are satisfied:

$$\begin{bmatrix} Q & -A^* & & 0 \\ -A & Q & -A^* & \\ & -A & Q & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} I \\ G_+ \\ G_+^2 \\ \vdots \end{bmatrix} X_+^{-1} = \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \tag{15}$$

$$\begin{bmatrix} Q & -A & & 0 \\ -A^* & Q & -A & \\ & -A^* & Q & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} I \\ H_+ \\ H_+^2 \\ \vdots \end{bmatrix} Y_+^{-1} = \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \end{bmatrix}. \tag{16}$$

By following the strategy successfully devised in [4, 5, 6] for solving nonlinear matrix equations arising in Markov chains, we apply the cyclic reduction algorithm to the above systems. This consists in performing an even-odd permutation of the block rows and columns, followed by one step of Gaussian elimination, thus generating the sequence of systems:

$$\begin{aligned} &\begin{bmatrix} X_n & -A_n^* & & 0 \\ -A_n & Q_n & -A_n^* & \\ & -A_n & Q_n & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} I \\ G_+^{2^n} \\ G_+^{2\cdot 2^n} \\ \vdots \end{bmatrix} X_+^{-1} = \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \\ &\begin{bmatrix} Y_n & -A_n & & 0 \\ -A_n^* & Q_n & -A_n & \\ & -A_n^* & Q_n & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} I \\ H_+^{2^n} \\ H_+^{2\cdot 2^n} \\ \vdots \end{bmatrix} Y_+^{-1} = \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \quad n \geq 0. \end{aligned} \tag{17}$$

The block entries of each system are defined by the following recursions:

$$\begin{aligned} &A_0 = A, \quad Q_0 = X_0 = Y_0 = Q, \\ &A_{n+1} = A_n Q_n^{-1} A_n, \\ &Q_{n+1} = Q_n - A_n Q_n^{-1} A_n^* - A_n^* Q_n^{-1} A_n, \\ &X_{n+1} = X_n - A_n^* Q_n^{-1} A_n, \\ &Y_{n+1} = Y_n - A_n Q_n^{-1} A_n^*, \quad n \geq 0. \end{aligned} \tag{18}$$

Observe that the matrices Q_n, X_n and Y_n are Hermitian, and thus the matrices in (17) are Hermitian.

The spectral theory of Hermitian block Toeplitz matrices guarantees the positive definitiveness, and thus the nonsingularity, of the blocks Q_n. To prove this, let us introduce the function $f : (-\pi, \pi) \to \boldsymbol{H}^m$, defined as $f(\theta) = -e^{\boldsymbol{i}\theta} A + Q - e^{-\boldsymbol{i}\theta} A^*$, where $\boldsymbol{i}$ is the imaginary unit, and $\boldsymbol{H}^m$ is the set of $m \times m$ Hermitian matrices. Let us denote by

$$\begin{aligned} \mu_1 &= \inf_{\theta \in (-\pi,\pi)} \lambda(f(\theta)), \\ \mu_2 &= \sup_{\theta \in (-\pi,\pi)} \lambda(f(\theta)), \end{aligned}$$

where $\lambda(f(\theta))$ is the set of eigenvalues of $f(\theta)$.

Since $\psi(\lambda) \geq 0$ for any $\lambda \in \mathcal{C}$, it holds $f(\theta) \geq 0$ and $\mu_1 \geq 0$. Moreover, since $\psi(\lambda)$ is regular and $f(\theta)$ is a continuous function, there exists an interval $(\alpha, \beta) \subset (-\pi, \pi)$ such that $\psi(\lambda) > 0$ for any $\theta \in (\alpha, \beta)$, hence the set where $f(\theta)$ is positive definite has strictly positive Lebesgue measure.

From these properties, the following result, that guarantees the applicability of the cyclic reduction algorithm and the boundness in norm of Q_n, X_n and Y_n, can be proved:

Theorem 3.1 *The matrices Q_n, X_n, Y_n, $n \geq 0$, are positive definite, and their eigenvalues belong to the interval $[\mu_1, \mu_2]$. Moreover, it holds $0 < Q_{n+1} \leq Q_n$, $0 < X_{n+1} \leq X_n$, $0 < Y_{n+1} \leq Y_n$, for $n \geq 0$.*

Proof. For the properties of cyclic reduction [6, 3], the matrices Q_n can be viewed as a Schur complement in $T_{2^{n+1}-1}$, where T_n is the $n \times n$ block matrix obtained by truncating the matrix in (15) at the block size n; similarly X_n and Y_n can be viewed as Schur complements in T_{2^n} and U_{2^n}, respectively, where U_n is the $n \times n$ block matrix obtained by truncating the matrix in (16) at the block size n. On the other hand, since T_n and U_n are Hermitian block Toeplitz matrices, their eigenvalues belong to the interval $[\mu_1, \mu_2]$ (see [22, 19, 21]); moreover, since the set where $f(\theta)$ is positive definite has strictly positive Lebesgue measure, then the matrices T_n and U_n are positive definite for any n [22, 19]. In particular, also the the Schur complements Q_n, X_n, Y_n are positive definite and their eigenvalues belong to the interval $[\mu_1, \mu_2]$, thus they have bounded norm. From (18) it follows that the sequences Q_n, X_n and Y_n are monotone. $\square$

If $\psi(\lambda) > 0$ for any $\lambda \in \mathcal{C}$ then $\mu_1 > 0$, thus also Q_n^{-1} is bounded in norm, and the condition number of Q_n is bounded. In this case the sequences X_n and Y_n quadratically converge to X_+ and Y_+, respectively:

Theorem 3.2 *If $\psi(\lambda) > 0$ for any $\lambda \in \mathcal{C}$, then for any $\epsilon > 0$ and for any operator norm $|| \cdot ||$ it holds $||I - X_n X_+^{-1}|| = O\left((\sigma + \epsilon)^{2 \cdot 2^n}\right)$, $||I - Y_n Y_+^{-1}|| = O\left((\sigma + \epsilon)^{2 \cdot 2^n}\right)$, where $\sigma = \rho(X_+^{-1} A) < 1$. Moreover, it holds $||A_n|| = O\left((\sigma + \epsilon)^{2^n}\right)$.*

Proof. From (17) we obtain that, for any n, the following relations hold:

$$-I + X_n X_+^{-1} - A_n^* G_+^{2^n} X_+^{-1} = 0, \tag{19}$$

$$-A_n + Q_n G_+^{2^n} - A_n^* G_+^{2 \cdot 2^n} = 0, \tag{20}$$

$$-I + Y_n Y_+^{-1} - A_n H_+^{2^n} Y_+^{-1} = 0, \tag{21}$$

$$-A_n^* + Q_n H_+^{2^n} - A_n H_+^{2 \cdot 2^n} = 0. \tag{22}$$

From (20) and (22) we have

$$A_n(I - H_+^{2\cdot 2^n} G_+^{2\cdot 2^n}) = Q_n(I - H_+^{2^n} G_+^{2^n}) G_+^{2^n}. \tag{23}$$

Since $\rho(G_+) = \rho(H_+) = \sigma < 1$ for Theorems 2.5, 2.6, for any $\epsilon > 0$ there exists an operator norm $||\cdot||$ such that $||G_+|| \leq \sigma + \epsilon$. Thus, since Q_n is bounded in norm, from (23) it follows that $||A_n|| = O\left((\sigma + \epsilon)^{2^n}\right)$. From (19) and (21) we obtain that $||I - X_n X_+^{-1}|| = O\left((\sigma + \epsilon)^{2\cdot 2^n}\right)$. Similarly, we prove that $||I - Y_n Y_+^{-1}|| = O\left((\sigma + \epsilon)^{2\cdot 2^n}\right)$. □

The algorithm for the computation of X_+ and X_- is synthesized by the following scheme:

Algorithm 3.1 (Cyclic reduction for (1))

1 *Set* $X_0 = Y_0 = Q_0 = Q$, $A_0 = A$.

2 *For* $n = 0, 1, \ldots$ *compute*

$$\begin{aligned} A_{n+1} &= A_n Q_n^{-1} A_n, \\ Q_{n+1} &= Q_n - A_n Q_n^{-1} A_n^* - A_n^* Q_n^{-1} A_n, \\ X_{n+1} &= X_n - A_n^* Q_n^{-1} A_n, \\ Y_{n+1} &= Y_n - A_n Q_n^{-1} A_n^*, \end{aligned}$$

until $||X_{n+1} - X_n||_\infty < \epsilon$, $||Y_{n+1} - Y_n||_\infty < \epsilon$, *for a fixed error bound* $\epsilon > 0$.

3 X_{n+1} *provides an approximation to* X_+, *and, if* A *is nonsingular,* $Q - Y_{n+1}$ *provides an approximation to* X_-.

If we are interested only in the computation of the maximal solution X_+ it is sufficient to skip the computation of the sequence Y_n.

One step of the above algorithm requires the solution of two Hermitian linear systems (i.e., the computation of $A_n Q_n^{-1}$ and $A_n^* Q_n^{-1}$), with bounded condition number, and the computation of three matrix products.

If the hypothesis of Theorem 3.2 are not satisfied, then $\rho(X_+^{-1} A) = 1$. In this case X_n, Y_n, Q_n are still Hermitian positive definite and bounded in norm. In general, the sequence Q_n^{-1} may be not bounded. However, if the sequence A_n converges to zero, and the sequence $(X_+^{-1} A)^{2^n}$ is bounded, the sequence X_n still converges to X_+.

Our algorithm can be compared with classical algorithms, that are based on fixed point iterations or Newton's method. The basic fixed point iteration consists in the following:

Algorithm 3.2 (Fixed point iteration for (1))

1 *Set* $X_0 = Q$.

2 *For* $n = 0, 1, 2, \ldots$, *compute* $X_{n+1} = Q - A^* X_n^{-1} A$ *until* $||X_{n+1} - X_n||_\infty < \epsilon$.

3 X_{n+1} *provides an approximation to* X_+.

In [13] it is proved that X_n converges to X_+ and $\limsup_{n\to\infty} \sqrt[n]{||X_n - X_+||} \leq \rho(X_+^{-1}A)^2$; thus the sequence linearly converges to X_+ and the mean asymptotic rate of convergence is $\rho(X_+^{-1}A)^2$, that is the same quantity that appears in the convergence Theorem 3.2. The main difference is that the sequence generated by Algorithm 3.2 converges only linearly. Each step of the above algorithm requires the solution of an Hermitian linear system and a matrix product; hence the cost of the cyclic reduction algorithm is roughly twice that of fixed point iterations. On the other hand, if we wish to compute also the minimal solution X_- with fixed point iterations, we must apply from scratch the algorithm to the matrix equation (8), or apply an "ad hoc" fixed point iteration, described in [8]; thus the computational cost per step is the same as cyclic reduction.

To avoid possible numerical instability problems due to matrix inversions, the following inversion free variant is proposed in [13]:

Algorithm 3.3 (Inversion free fixed point iteration for (1))

1 *Set* $X_0 = Q$, $Y_0 = I/||Q||_\infty$.

2 *For* $n = 0, 1, 2, \ldots$, *compute*

$$\begin{aligned} Y_{n+1} &= Y_n(2I - X_n Y_n), \\ X_{n+1} &= Q - A^* Y_{n+1} A, \end{aligned}$$

until $||X_{n+1} - X_n||_\infty < \epsilon$.

3 X_{n+1} *provides an approximation to* X_+.

The latter algorithm requires four matrix products at each step; in [13] it is shown that the inversion free variant can be faster than Algorithm 3.2 by a factor 2.

Finally, Newton's method consists in the following:

Algorithm 3.4 (Newton's method for (1))

1 *Set* $X_0 = Q$.

2 *For* $n = 1, 2, \ldots$, *compute* $L_n = X_{n-1}^{-1} A$ *and solve*

$$X_n - L_n^* X_n L_n = Q - 2L_n^* A, \tag{24}$$

until $||X_n - X_{n-1}||_\infty < \epsilon$.

3 X_n *provides an approximation to* X_+.

Each step of Newton's method requires the solution of the Hermitian system $X_{n-1}L_n = A$, the solution of the Stein equation (24) and one matrix product; thus the computational cost of each step if much larger than the cost of fixed point iterations and cyclic reduction. If equation (24) is solved with the algorithm described in [10] the computational time for iteration is roughly 10-15 times that for Algorithms 3.2, 3.3 (see [13]). Moreover, if X_+A^{-1} has eigenvalues very close to the unit circle, equation (24) can be nearly singular [13]. In [13] it is proved that the convergence of X_n to X_+ is quadratic if $\rho(X_+^{-1}A) < 1$, either quadratic or linear with rate $1/2$ if $\rho(X_+^{-1}A) = 1$ and all the eigenvalues of $X_+^{-1}A$ on the unit circle are semi-simple.

Hence Newton's method presents the same nice features of cyclic reduction in terms of convergence, but it is much more expensive in terms of computational cost, and more difficult to be implemented.

4 Computation of the extreme solution of $X - A^*X^{-1}A = Q$

For the computation of the maximal and minimal solutions X_+ and X_-, respectively, of $X - A^*X^{-1}A = Q$ we can apply a technique similar to the one used the previous section. Specifically, we introduce the maximal solution Y_+ of (10), and observe that it holds:

$$\begin{aligned} -I + QX_+^{-1} + A^*X_+^{-1}AX_+^{-1} &= 0 \\ -I + QY_+^{-1} + AY_+^{-1}A^*Y_+^{-1} &= 0. \end{aligned}$$

Thus, by setting $G_+ = X_+^{-1}A$ and $H_+ = Y_+^{-1}A^*$, the following linear systems are verified:

$$\begin{gathered} \begin{bmatrix} Q & A^* & & 0 \\ -A & Q & A^* & \\ & -A & Q & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} I \\ G_+ \\ G_+^2 \\ \vdots \end{bmatrix} X_+^{-1} = \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \\ \begin{bmatrix} Q & A & & 0 \\ -A^* & Q & A & \\ & -A^* & Q & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} I \\ H_+ \\ H_+^2 \\ \vdots \end{bmatrix} Y_+^{-1} = \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \end{bmatrix}. \end{gathered} \tag{25}$$

The infinite matrices in the above systems are block Toeplitz, but are not Hermitian. However, if we apply one step of cyclic reduction, we obtain the following systems:

$$
\begin{array}{c}
\begin{bmatrix}
X_1 & -A_1^* & & 0 \\
-A_1 & Q_1 & -A_1^* & \\
 & -A_1 & Q_1 & \ddots \\
0 & & \ddots & \ddots
\end{bmatrix}
\begin{bmatrix} I \\ G_+^2 \\ G_+^4 \\ \vdots \end{bmatrix} X_+^{-1} = \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \\
\begin{bmatrix}
Y_1 & -A_1 & & 0 \\
-A_1^* & Q_1 & -A_1 & \\
 & -A_1^* & Q_1 & \ddots \\
0 & & \ddots & \ddots
\end{bmatrix}
\begin{bmatrix} I \\ H_+^2 \\ H_+^4 \\ \vdots \end{bmatrix} Y_+^{-1} = \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \end{bmatrix},
\end{array}
\tag{26}
$$

where

$$
\begin{array}{l}
A_1 = AQ^{-1}A, \\
Q_1 = Q + AQ^{-1}A^* + A^*Q^{-1}A, \\
X_1 = Q + A^*Q^{-1}A, \quad Y_1 = Q + AQ^{-1}A^*.
\end{array}
\tag{27}
$$

Thus, after one step of cyclic reduction, we obtain two systems with Hermitian matrices, with the structure of (17), where the diagonal blocks are positive definite matrices. Hence, one step of cyclic reduction leads to an Hermitian system, like in the case of equation (1). Moreover, observe that the function

$$
\psi_1(\lambda) = \lambda A_1 + Q_1 + \lambda^{-1}A_1^*,
$$

is such that $\psi_1(\lambda) > 0$ for any $\lambda \in \mathcal{C}$, since $\psi_1(\lambda) = Q + (A - \bar{\lambda}A^*)Q^{-1}(A^* - \lambda A)$. If we apply cyclic reduction to systems (26), we generate the sequence (17), for $n \geq 1$, where

$$
\begin{array}{l}
A_{n+1} = A_n Q_n^{-1} A_n, \\
Q_{n+1} = Q_n - A_n Q_n^{-1} A_n^* - A_n^* Q_n^{-1} A_n, \\
X_{n+1} = X_n - A_n^* Q_n^{-1} A_n, \\
Y_{n+1} = Y_n - A_n Q_n^{-1} A_n^*, \quad n \geq 1,
\end{array}
\tag{28}
$$

and A_1, Q_1, X_1 and Y_1 are defined in (27).

Without any assumption, the following convergence result can be proved:

Theorem 4.1 *For the sequences of matrices X_n, Y_n, Q_n, $n \geq 1$, defined in (28), it holds:*

1. $0 < X_{n+1} \leq X_n$, $0 < Y_{n+1} \leq Y_n$, $0 < Q_{n+1} \leq Q_n$, $n = 1, 2, \ldots$*;*
2. Q_n *and* Q_n^{-1} *are bounded in norm.*

Moreover, for any $\epsilon > 0$ and for any operator norm $||\cdot||$ it holds

$$
\begin{array}{l}
||I - X_n X_+^{-1}|| = O\left((\sigma + \epsilon)^{2\cdot 2^n}\right), \\
||I - Y_n Y_+^{-1}|| = O\left((\sigma + \epsilon)^{2\cdot 2^n}\right), \\
||A_n|| = O\left((\sigma + \epsilon)^{2^n}\right),
\end{array}
$$

where $\sigma = \rho(X_+^{-1}A) < 1$.

Proof. For the properties of cyclic reduction, the matrices Q_n, for $n \geq 2$, can be viewed as Schur complements of a suitable finite block tridiagonal block Toeplitz matrix, with Q_1 on the main block diagonal, $-A_1^*$ on the block lower diagonal, and $-A_1$ on the block upper diagonal. Since $\psi_1(-\lambda)$, that is the matrix function associated with the latter block Toeplitz matrix, is positive definite for any λ on the unit circle, the eigenvalues of Q_n belong to the interval $[\mu_1, \mu_2]$, where $0 < \mu_1 = \inf\{|\mu| \ : \ \mu \text{ is an eigenvalue of } \psi_1(-\lambda),\ |\lambda| = 1\}$, $\lambda_2 = \sup\{|\mu| \ : \ \mu \text{ is an eigenvalue of } \psi_1(-\lambda),\ |\lambda| = 1\}$ (see [19, 21, 22]); hence Q_n is positive definite and the sequences Q_n and Q_n^{-1} are bounded in norm. Let us now prove that the matrices X_n, Y_n are positive definite; let us denote by $X_1^{(1)}$, $Y_1^{(1)}$, $Q_1^{(1)}$, $A_1^{(1)}$ the blocks obtained by applying one step of cyclic reduction to systems (15) and (16); then we have $0 < X_1^{(1)} \leq X_1$, $0 < Y_1^{(1)} \leq Y_1$, $0 < Q_1^{(1)} \leq Q_1$, $A_1^{(1)} = A_1$; hence, since all the finite sections of the matrices in (17), for $n = 1$, are positive definite, then also all the finite sections of the matrices in (26) are positive definite; in particular, X_n, Y_n, being Schur complements of suitable finite sections of the matrices in (26), are positive definite. The monotonicity trivially holds. The convergence properties can be proved by using the same argument of Theorem 3.2. □

From the above theorem, the quadratic convergence is always guaranteed, and the condition number of Q_n is always bounded.

If A is nonsingular, the matrix X_- can be recovered from Y_+ by means of the relation $X_- = Q - Y_+$ (see [9]).

The resulting algorithm is resumed below:

Algorithm 4.1 (Cyclic reduction for (2))

1 *Set*

$$A_1 = AQ^{-1}A, \quad Q_1 = Q + AQ^{-1}A^* + A^*Q^{-1}A,$$
$$X_1 = Q + A^*Q^{-1}A, \quad Y_1 = Q + AQ^{-1}A^*.$$

2 *For* $n = 1, 2, \ldots$ *compute*

$$\begin{aligned} A_{n+1} &= A_n Q_n^{-1} A_n, \\ Q_{n+1} &= Q_n - A_n Q_n^{-1} A_n^* - A_n^* Q_n^{-1} A_n, \\ X_{n+1} &= X_n - A_n^* Q_n^{-1} A_n, \\ Y_{n+1} &= Y_n - A_n Q_n^{-1} A_n^*, \end{aligned}$$

until $||X_{n+1} - X_n|| < \epsilon$, $||Y_{n+1} - Y_n|| < \epsilon$.

3 X_{n+1} *provides an approximation to* X_+, *and, if* A *is nonsingular,* $Q - Y_{n+1}$ *provides an approximation to* X_-.

If we are interested only in X_+, it is sufficient to skip the computation of the sequence Y_n.

The above algorithm has the same nice features, in terms of computational cost and convergence properties of Algorithm 3.1.

We may compare our algorithm with the ones customarily used in the literature, that we recall below:

Algorithm 4.2 (Fixed point iteration for (2))

1 *Set* $X_0 = Q$.

2 *For* $n = 0, 1, 2, \ldots$, *compute* $X_{n+1} = Q + A^* X_n^{-1} A$, *until* $||X_{n+1} - X_n||_\infty < \epsilon$.

3 X_{n+1} *provides an approximation to* X_+.

Algorithm 4.3 (Newton's method for (2))

1 *Set* X_0 *sufficiently close to* X_+.

2 *For* $n = 1, 2, \ldots$, *compute* $L_n = X_{n-1}^{-1} A$ *and solve*

$$X_n + L_n^* X_n L_n = Q + 2L_n^* A, \tag{29}$$

3 X_n *provides an approximation to* X_+.

The above algorithms present the same characteristics of Algorithms 3.2, 3.4, in terms of computational cost and convergence rate. Concerning Newton's method, the initial guess X_0 must be chosen close to the solution X_+, in order to guarantee the convergence [13]; in practice, the starting point X_0 is chosen as the approximation provided after few steps of fixed point iterations (Algorithm 4.2).

The cyclic reduction algorithm has always quadratic convergence, low computational cost for iteration, and good numerical stability properties; moreover, with respect to Newton's method, we do not need to have an initial guess X_0 sufficiently close to the solution X_+.

5 Numerical results

Concerning (1), we have compared the proposed Algorithm 3.1 with the fixed point iterations (Algorithms 3.2 and 3.3), and with Newton's method (Algorithm 3.4), for the computation of X_+; for equation (2) we have compared Algorithm 4.1 with the fixed point iterations (Algorithm 4.2). For all the algorithms the value of ϵ for the stopping condition is $\epsilon = 10^{-11}$.

We implemented the algorithms in Fortran 90, and run the programs on a Pentium II. For the solution of systems (24), (29) we have used the Fortran implementation of [10].

We considered the following problems:

Example 5.1 This is an example concerning (1), presented in [13]. The matrices A and Q are

$$A = \begin{bmatrix} 0.37 & 0.13 & 0.12 \\ -0.30 & 0.34 & 0.12 \\ 0.11 & -0.17 & 0.29 \end{bmatrix}, \quad Q = \begin{bmatrix} 1.20 & -0.30 & 0.10 \\ -0.30 & 2.10 & 0.20 \\ 0.10 & 0.20 & 0.65 \end{bmatrix}.$$

To reach the requested accuracy, the cyclic reduction (Algorithm 3.1) needs 7 iterations, fixed point iterations (Algorithm 3.2) 49 iterations, the inversion free variant (Algorithm 3.3) 52 iterations, and Newton's method (Algorithm 3.4) 6 iterations. The CPU time needed by all the algorithms is negligible, since it is less than 0.01 seconds.

Example 5.2 We consider again equation (1), with $Q = I$ and A symmetric, whose entries are defined by the following scheme:

1. Fix a real $0 \leq \alpha < 1/2$.
2. For $i = 1, \ldots, m$, for $j = i, \ldots, m$:
   (a) set $a_{i,j} = i^2 + j$;
   (b) compute $s_1 = \sum_{j=1}^{i-1} a_{i,j}$, $s_2 = \sum_{j=i}^{m} a_{i,j}$;
   (c) set
   $$a_{i,j} = a_{i,j}\frac{(1/2 - \alpha - s_1)}{s_2}, \quad a_{j,i} = a_{i,j}.$$

The matrix A obtained in this way is symmetric, nonnegative and such that $A\boldsymbol{e} = (1/2-\alpha)\boldsymbol{e}$, where $\boldsymbol{e}$ is the vector having all the entries equal to 1. Thus the spectral radius of A is $1/2 - \alpha$, and $r(Q^{-1/2}AQ^{-1/2}) = r(A) = \rho(A) = 1/2 - \alpha$. If $\alpha = 0$ it holds $r(A) = 1/2$ and $\rho(X_+^{-1}A) = 1$; thus, the hypothesis of Theorem 3.2 are not satisfied. Small values of α provide very critical test problems.

For this example, since $Q = I$ and A is normal with $||A||_2 \leq 1/2$, the maximal solution is given by the matrix [24]

$$X_+ = \frac{1}{2}\left(I + (I - 4A^*A)^{1/2}\right).$$

Table 1 reports, for different values of the size m and of the parameter α, the number of iterations needed to satisfy the stopping conditions and the error $||\widetilde{X}_+ - X_+||_\infty / ||X_+||_\infty$, where $\widetilde{X}_+$ is the approximation provided by Algorithms 3.1, 3.2, 3.3, 3.4 (denoted by CR, FPI, IF-FPI, NM, respectively). Table 2 reports the CPU time (in seconds) needed by the same algorithms.

It is worth point out how the advantage, in terms of CPU time and accuracy, of the cyclic reduction gets more consistent as the size m increases and the spectral radius of $X_+^{-1}A$ approaches to 1. In particular, Algorithm 3.1 shows a good convergence also in the case $\alpha = 0$, i.e. $\rho(X_+^{-1}A) = 1$, when the hypothesis of Theorem 3.2 are not satisfied.

	CR		FPI		IF-FPI		NM	
α	Iter.	Err.	Iter.	Err.	Iter.	Err.	Iter.	Err.
	$m = 20$							
0.4	4	1.4E-15	6	1.2E-14	6	2.3E-14	3	4.0E-15
0.2	5	1.4E-15	11	2.6E-12	11	5.2E-12	4	3.7E-15
0.1	6	1.5E-15	16	2.9E-11	17	1.4E-11	5	3.9E-15
0.01	7	1.3E-15	50	1.7E-10	51	1.5E-10	6	3.1E-15
0.001	9	2.3E-15	141	6.9E-10	142	6.8E-10	8	3.2E-15
0.0001	10	5.7E-15	388	2.4E-09	389	2.4E-09	9	3.7E-15
0	30	1.4E-09	58723	5.9E-06	58723	5.9E-06	67	6.2E-09
	$m = 40$							
0.4	4	2.9E-15	6	1.2E-14	6	2.3E-14	3	6.9E-15
0.2	5	2.8E-15	11	2.6E-12	11	5.2E-12	4	6.6E-15
0.1	6	2.3E-15	16	2.9E-11	17	1.4E-11	5	5.9E-15
0.01	7	1.8E-15	50	1.7E-10	51	1.5E-10	6	5.3E-15
0.001	9	1.9E-15	141	6.8E-10	142	6.7E-10	8	5.3E-15
0.0001	10	4.5E-15	388	2.4E-09	389	2.4E-09	9	1.2E-14
0	31	6.3E-09	58222	5.8E-06	58223	5.8E-06	148	6.6E-09
	$m = 80$							
0.4	4	5.1E-15	6	1.31E-14	6	1.45E-14	5	4.3E-15
0.2	5	4.4E-15	11	2.6E-12	11	5.2E-12	4	6.6E-15
0.1	6	4.4E-15	16	2.9E-11	17	1.4E-11	5	5.9E-15
0.01	7	3.7E-15	50	1.7E-10	51	1.5E-10	6	5.3E-15
0.001	9	4.9E-15	141	6.8E-10	142	6.7E-10	8	5.3E-15
0.0001	10	5.4E-15	388	2.4E-09	389	2.4E-09	9	1.2E-14
0	28	6.7E-09	58222	5.8E-06	58223	5.8E-06	148	6.6E-09

Table 1: Example 5.2: number of iterations and errors

α	CR	FPI	IF-FPI	NM
	$m = 20$			
0.4	0.01	0.01	0.01	0.04
0.2	0.01	0.01	0.01	0.03
0.1	0.01	0.01	0.01	0.05
0.01	0.01	0.03	0.05	0.06
0.001	0.01	0.09	0.10	0.07
0.0001	0.01	0.25	0.25	0.09
0	0.01	37.11	37.22	0.62
	$m = 40$			
0.4	0.04	0.02	0.04	0.19
0.2	0.05	0.04	0.05	0.26
0.1	0.05	0.06	0.09	0.32
0.01	0.05	0.20	0.27	0.38
0.001	0.08	0.54	0.75	0.51
0.0001	0.08	1.45	2.03	0.57
0	0.26	218.08	306.98	9.29
	$m = 80$			
0.4	0.24	0.15	0.02	1.51
0.2	0.29	0.30	0.40	2.03
0.1	0.34	0.42	0.64	2.53
0.01	0.41	1.35	1.86	3.00
0.001	0.53	3.70	5.18	4.03
0.0001	0.61	10.38	14.14	4.64
0	1.64	1548.12	2112.56	85.20

Table 2: Example 5.2: CPU time

Example 5.3 This example is taken from [13] and concerns equation (2). Here

$$A = \begin{bmatrix} 50 & 20 \\ 10 & 60 \end{bmatrix}, \quad Q = \begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix}.$$

The cyclic reduction (Algorithm 4.1) needs 9 iterations to reach the required accuracy, while Algorithm 4.2 needs 426 iterations. The residual error of the approximation provided by Algorithm 4.1 is $2.3 \cdot 10^{-15}$, while residual error of the approximation provided by Algorithm 4.2 is $9.1 \cdot 10^{-11}$.

References

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

[2] W. N. Anderson Jr., T. D. Morley, and G. E. Trapp. Positive solutions to $X = A - BX^{-1}B^*$. *Linear Algebra Appl.*, 134:53–62, 1990.

[3] D. A. Bini, L. Gemignani, and B. Meini. Factorization of Analytic Functions by means of Koenig's Theorem and Toeplitz Computations. Submitted for publication, 1998.

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

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

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

[7] J. C. Engwerda. On the existence of a positive definite solution of the matrix equation $X + A^T X^{-1} A = I$. *Linear Algebra Appl.*, 194:91–108, 1993.

[8] J. C. Engwerda, A. C. M. Ran, and A. L. Rijkeboer. Necessary and sufficient conditions for the existence of a positive definite solution of the matrix equation $X + A^* X^{-1} A = Q$. *Linear Algebra Appl.*, 186:255–275, 1993.

[9] A. Ferrante and B. C. Levy. Hermitian solutions of the equation $X = Q + NX^{-1}N^*$. *Linear Algebra Appl.*, 247:359–373, 1996.

[10] J. D. Gardiner, A. L. Laub, J. J. Amato, and C. B. Moler. Solution of the Sylvester matrix equation $AXB^T + CXD^T = E$. *ACM Trans. Math. Software*, 18:223–231, 1992.

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

[12] C.-H. Guo. Newtons's method for discrete algebraic Riccati equations when the closed-loop matrix has eigenvalues on the unit circle. *SIAM J. Matrix Anal. Appl.*, 20:279–294, 1999.

[13] C.-H. Guo and P. Lancaster. Iterative solution of two matrix equations. *Math. Comp.*, 68:1589–1603, 1999.

[14] G. Latouche. A note on two matrices occurring in the solution of quasi-birth-and-death processes. *Commun. Statist. Stochastic Models*, 3:251–257, 1987.

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

[16] G. Latouche and V. Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modeling*. ASA-SIAM Series on Statistics and Applied Probability 5. SIAM, 1999.

[17] G. Latouche and G.W. Stewart. Numerical methods for M/G/1 type queues. In W. J. Stewart, editor, *Computations with Markov Chains*, pages 571–581. Kluwer Academic Publishers, 1995.

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

[19] M. Miranda and P. Tilli. Block Toeplitz matrices and preconditioning. *Calcolo*, 33(1-2):79–86 (1998), 1996. Toeplitz matrices: structures, algorithms and applications (Cortona, 1996).

[20] M. F. Neuts. *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. The Johns Hopkins University Press, Baltimore, MD, 1981.

[21] P. Tilli. Asymptotic spectral distribution of Toeplitz-related matrices. In T. Kailath and A. H. Sayed, editors, *Fast reliable algorithms for matrices with structure*, chapter 6, pages 153–187. SIAM, Philadelphia, 1999.

[22] P. Tilli and M. Miranda. Asymptotic spectrum of Hermitian and block Toeplitz matrices and preconditioning results. *SIAM J. Matrix Anal. Appl.*, 1999.

[23] X. Zhan. Computing the extremal positive definite solutions of a matrix equation. *SIAM J. Sci. Comput.*, 17:1167–1174, 1996.

[24] X. Zhan and J. Xie. On the matrix equation $X + A^T X^{-1} A = I$. *Linear Algebra Appl.*, 247:337–345, 1996.