A shifted cyclic reduction algorithm for QBDs

C. He[*] B. Meini[†] N.H. Rhee[‡]

May 9, 2000

Abstract

The problem of the computation of the rate matrix G associated with discrete-time QBD Markov chains is analyzed. We present a shifted cyclic reduction algorithm and show that the speed of convergence of the latter modified algorithm is always faster than that of the original cyclic reduction.

1 Introduction

The block tridiagonal and block Toeplitz structure of the probability transition matrix P associated with QBD problems, i.e.,

$$P = \begin{bmatrix} B_0 & B_1 & & 0 \\ A_0 & A_1 & A_2 & \\ & A_0 & A_1 & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix},$$

allowed the design of fast and reliable methods for computation of the rate matrix G [16, 11, 2, 3, 4, 12, 5]. Here, A_0, A_1, A_2, B_0, B_1 are $k \times k$ nonnegative matrices such that $A_0 + A_1 + A_2$ is irreducible, and P is irreducible, stochastic and positive recurrent. These fast methods rely on the property that the matrix G, which solves the nonlinear matrix equation

$$G = A_0 + A_1 G + A_2 G^2, \qquad (1)$$

[*]Sprint Corporation, Network Planning and Design, 7171 W. 95th Street Overland Park, KS 66212, charlie.he@mail.sprint.com

[†]Dipartimento di Matematica, Università di Pisa, Italy, meini@dm.unipi.it

[‡]Department of Mathematics and Statistics, University of Missouri at Kansas City, Kansas City, MO 64110, rhee@cctr.umkc.edu

can be computed by solving the infinite block tridiagonal, block Toeplitz system

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

Recently, in [2, 3, 4], Bini and Meini have devised a new quadratically convergent and numerically stable algorithm for the computation of G, based on a functional representation of cyclic reduction, which applies to general M/G/1 type Markov chains [16] and which extends the method of Latouche and Ramaswami [11].

The aim of this paper is to introduce a shifted cyclic reduction algorithm for QBDs and to show that, under the nonrestrictive assumption that G has only one eigenvalue of modulus one, the speed of convergence of a shifted cyclic reduction algorithm is always faster than that of the original cyclic reduction algorithm.

More precisely, since G has a known eigenvalue (equal to one) and a known corresponding right eigenvector, we may apply a shifting technique, that consists in moving the eigenvalue 1 to 0, and in maintaining the remaining ones. This trick leads to a new quadratic matrix equation, having as solution a singular matrix H, such that $G = H + \boldsymbol{e}\boldsymbol{u}^T$, where $\boldsymbol{e}$ is the vector having all the entries equal to 1, and $\boldsymbol{u}$ is a known arbitrary vector, with positive entries, such that $\boldsymbol{u}^T\boldsymbol{e} = 1$. Thus the problem of the computation of G is reduced to the problem of the computation of H, that shares with G all the eigenvalues, except for the eigenvalue 1, that is moved to 0.

In order to compute H we apply the cyclic reduction algorithm, and we show that the shifting leads to a faster convergence. More precisely, in standard cyclic reduction, the error of the approximation to G, for $j \to \infty$, goes to zero as $O\left((1/\sigma)^{2^j}\right)$, where $\sigma = \min\{|\lambda| : |\lambda| > 1, \ \phi(\lambda) = 0\}$, and $\phi(\lambda) = \det(-A_0 + (I - A_1)\lambda - A_2\lambda^2)$. With the shifting technique, the error goes to zero as $O\left((\theta/\sigma)^{2^j}\right)$, where θ is any real number such that $\max\{|\lambda| : |\lambda| < 1, \ \phi(\lambda) = 0\} < \theta < 1$. Thus the convergence speed can be much increased; in particular the improvement is more appreciated if $\sigma \approx 1$ and $\theta << 1$.

Finally, we introduce a means to measure the conditioning of the quadratic matrix equations, and we prove that the shifted equation is better conditioned than the original one. Indeed, the shifting technique leads to a better rate of convergence, but destroys the nonnegativity and the M-matrix properties of the blocks generated at each step of standard cyclic reduction, and in principle this fact could lead to a loss of accuracy of the results.

We have performed several numerical experiments, which show that the shifted cyclic reduction algorithm is fast and numerically accurate.

The paper is organized as follows. In Section 2 we recall the cyclic reduction algorithm for QBDs. In Section 3 we apply the shifting technique, and we show how the solutions of the shifted matrix equation are related to the solution of the original one. In Section 4 we analyze the convergence properties of the cyclic reduction algorithm applied to the shifted matrix equation. In Section 5 we study the conditioning of the two matrix equations. In Section 6 we present some numerical results.

2 The cyclic reduction algorithm

In this section we recall the cyclic reduction algorithm for QBDs, described in [3].

Let us consider the system (2). By recursively applying block cyclic reduction, i.e., an odd-even permutation of block rows and block columns, followed by one step of Gaussian elimination, the following sequence of infinite block tridiagonal systems is generated:

$$\begin{bmatrix} I-\widehat{A}_1^{(j)} & -A_2^{(j)} & & 0 \\ -A_0^{(j)} & I-A_1^{(j)} & -A_2^{(j)} & \\ & -A_0^{(j)} & I-A_1^{(j)} & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} G \\ G^{2^j+1} \\ G^{2\cdot 2^j+1} \\ \vdots \end{bmatrix} = \begin{bmatrix} A_0 \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \quad j \geq 0, \tag{3}$$

where $A_0^{(0)} = A_0$, $\widehat{A}_1^{(0)} = A_1^{(0)} = A_1$, $A_2^{(0)} = A_2$, and the blocks $\widehat{A}_1^{(j)}$, $A_i^{(j)}$, $i = 0,1,2$, $j \geq 1$, are defined by the recurrences:

$$\begin{aligned} A_0^{(j+1)} &= A_0^{(j)}\left(I - A_1^{(j)}\right)^{-1} A_0^{(j)} \\ A_1^{(j+1)} &= A_1^{(j)} + A_0^{(j)}\left(I - A_1^{(j)}\right)^{-1} A_2^{(j)} + A_2^{(j)}\left(I - A_1^{(j)}\right)^{-1} A_0^{(j)} \\ A_2^{(j+1)} &= A_2^{(j)}\left(I - A_1^{(j)}\right)^{-1} A_2^{(j)} \\ \widehat{A}_1^{(j+1)} &= \widehat{A}_1^{(j)} + A_2^{(j)}\left(I - A_1^{(j)}\right)^{-1} A_0^{(j)}. \end{aligned} \tag{4}$$

The sequences of matrices generated by the above relations allow the fast computation of the matrix G. Indeed, from (3) it follows that

$$(I - \widehat{A}_1^{(j)})G - A_2^{(j)}G^{2^j+1} = A_0. \tag{5}$$

On the other hand, if we denote by R the minimal nonnegative solution of the matrix equation $R = A_2 + RA_1 + R^2 A_0$, then the following equation is verified (see [3, 11])

$$-A_2^{(j)} + R^{2^j}(I - A_1^{(j)}) - R^{2\cdot 2^j} A_0^{(j)} = 0.$$

Moreover, the maximum modulus eigenvalue of R is real, simple and unique, and it is equal to $1/\sigma$ (see [8]), where $\sigma = \min\{|\lambda| :\ |\lambda| > 1,\ \phi(\lambda) = 0\}$, and $\phi(\lambda) = \det(-A_0 + (I - A_1)\lambda - A_2\lambda^2)$. Thus, there exists an operator norm $||\cdot||$ such that $||R|| = \rho(R)$, where $\rho(R)$ denotes the spectral radius of R; hence, we obtain

$$||A_2^{(j)}|| \leq \rho(R)^{2^j}(||I - A_1^{(j)}|| + \rho(R)^{2^j}||A_0^{(j)}||).$$

Since the matrices $A_1^{(j)}$ and $A_0^{(j)}$ are bounded in norm [3], it follows that $||A_2^{(j)}|| = O\left(\rho(R)^{2^j}\right)$ and from (5), since G^{2^j+1} is bounded, that

$$(I - \widehat{A}_1^{(j)})G - A_0 = O\left(\left(\frac{1}{\sigma}\right)^{2^j}\right). \tag{6}$$

Hence an approximation of the matrix G is given by $\left(I - \widehat{A}_1^{(j)}\right)^{-1} A_0$, for a sufficiently large value of j.

Due to the double exponential convergence to zero of the sequence $A_2^{(j)}$, just a small number of steps can be sufficient to reach a good approximation of G.

The rate of convergence is given by $1/\sigma$, and it is therefore related to the closeness to the unit circle of the smallest zero of $\phi(\lambda)$ of modulus larger than one.

In the next section we present a trick to improve the rate of convergence; indeed, we show that by applying a deflating technique, that consists in removing the zero $\lambda = 1$ of $\phi(\lambda)$, the rate of convergence can be reduced to θ/σ, where θ is any real number such that $\max\{|\lambda| :\ |\lambda| < 1,\ \phi(\lambda) = 0\} < \theta < 1$.

3 A shifted matrix equation

In the sequel we assume, without loss of generality, that $\lambda = 1$ is the only zero of the function $\det(-A_0 + (I - A_1)\lambda - A_2\lambda^2) = 0$ on the unit circle [7].

Since G has a known eigenvalue 1 and a known eigenvector $\boldsymbol{e}$ where $\boldsymbol{e}$ is the k-dimensional vector having all the entries equal to 1, we can modify

G so that the eigenvalue 1 is shifted to 0. This shifting technique improves the convergence speed of the cyclic reduction algorithm.

Set $H = G - \boldsymbol{e}\boldsymbol{u}^T$, where $\boldsymbol{u}$ is any vector whose elements are positive and such that $\boldsymbol{u}^T\boldsymbol{e} = 1$. Then the eigenvalues of H are those of G except that in H the eigenvalue 1 of G is replaced by 0. Moreover, $\boldsymbol{e}$ is an eigenvector of H corresponding to the eigenvalue 0, and hence $H\boldsymbol{e} = 0$. So we have

$$G = H + \boldsymbol{e}\boldsymbol{u}^T \quad \text{and} \quad G^2 = (H + \boldsymbol{e}\boldsymbol{u}^T)^2 = H^2 + \boldsymbol{e}\boldsymbol{u}^T H + \boldsymbol{e}\boldsymbol{u}^T.$$

By replacing G by H in (1) we obtain that H solves the following shifted equation

$$B_0 + B_1 H + B_2 H^2 = H, \tag{7}$$

where

$$\begin{aligned} B_0 &= A_0 + (A_1 + A_2 - I)\boldsymbol{e}\boldsymbol{u}^T = A_0(I - \boldsymbol{e}\boldsymbol{u}^T) \\ B_1 &= A_1 + A_2\boldsymbol{e}\boldsymbol{u}^T \\ B_2 &= A_2. \end{aligned} \tag{8}$$

So the computation of the matrix G can be reduced to the computation of the matrix $H = G - \boldsymbol{e}\boldsymbol{u}^T$, that solves the nonlinear matrix equation (7).

3.1 Spectral properties of the shifted matrix polynomial

The rank one corrections of the matrices A_0 and A_1 have the effect to keep unchanged all the zeros of the polynomial $\phi(\lambda) = \det A(\lambda)$, $A(\lambda) = -A_0 + (I - A_1)\lambda - A_2\lambda^2$, except for the zero $\lambda = 1$, that is moved to $\lambda = 0$. Indeed, the following result holds:

Theorem 1 *The zeros of the polynomial $\psi(\lambda) = \det B(\lambda)$, $B(\lambda) = -B_0 + (I - B_1)\lambda - B_2\lambda^2$, are*

$$\{\lambda \ : \ \det A(\lambda) = 0, \ \lambda \neq 1\} \cup \{0\}.$$

To prove the above theorem we need to introduce some notations and some results on pairs of matrices (see, for example, [9]).

Definition 1 *Let A and B be $n \times n$ complex matrices. We define $\Lambda(A, B)$ by*

$$\Lambda(A, B) = \{\lambda \in \mathbf{C} : \det(A - \lambda B) = 0\}.$$

If $\lambda \in \Lambda(A, B)$, λ is called an eigenvalue of the pair (A, B). If λ is an eigenvalue of the pair (A, B), then there exists a nonzero vector $\boldsymbol{x}$ such that $A\boldsymbol{x} = \lambda B\boldsymbol{x}$, and such an $\boldsymbol{x}$ is called an eigenvector *of the pair (A, B) corresponding to the eigenvalue λ.*

Theorem 2 *If A and B are $n \times n$ complex matrices, then there exist unitary matrices U and V such that $U^H AV = R$ and $U^H BV = \tilde{R}$ where R and $\tilde{R}$ are upper triangular. If for some i, r_{ii} and $\tilde{r}_{ii}$ are both zero, then $\Lambda(A, B) = \mathbf{C}$. Otherwise*

$$\Lambda(A, B) = \{r_{ii}/\tilde{r}_{ii} : \tilde{r}_{ii} \neq 0, \ i = 1, \ldots, n\}.$$

Note that $\boldsymbol{v}_1$, the first column of V, is an eigenvector of the pair (A, B) corresponding to the eigenvalue $r_{11}/\tilde{r}_{11}$.

Definition 2 *Two pairs of matrices (A, B) and (C, D) are said to be* equivalent *if there exist nonsingular matrices L and M such that $C = LAM$ and $D = LBM$. It is easy to verify that if (A, B) and (C, D) are equivalent pairs, then $\Lambda(A, B) = \Lambda(C, D)$.*

Let us now prove Theorem 1:
Proof of Theorem 1. Let

$$\Lambda(A) = \{\lambda \in \mathbf{C} : \det A(\lambda) = 0\}, \quad \Lambda(B) = \{\lambda \in \mathbf{C} : \det B(\lambda) = 0\}.$$

If we set

$$S = \begin{bmatrix} I - A_1 & -A_0 \\ I & 0 \end{bmatrix} \qquad \text{and} \qquad T = \begin{bmatrix} A_2 & 0 \\ 0 & I \end{bmatrix}$$

it is easy to verify that $\Lambda(A) = \Lambda(S, T)$. Similarly, if we set

$$\tilde{S} = \begin{bmatrix} I - B_1 & -B_0 \\ I & 0 \end{bmatrix}$$

then $\Lambda(B) = \Lambda(\tilde{S}, T)$.

Now if we let

$$L = \begin{bmatrix} I & -A_2 \boldsymbol{e}\boldsymbol{u}^T \\ 0 & I \end{bmatrix} \qquad \text{and} \qquad M = \begin{bmatrix} I & \boldsymbol{e}\boldsymbol{u}^T \\ 0 & I \end{bmatrix}$$

then by direct computation one can verify that

$$\tilde{S} = L(S - T \begin{bmatrix} \boldsymbol{e} \\ \boldsymbol{e} \end{bmatrix} [\mathbf{0}, \boldsymbol{u}^T])M \qquad \text{and} \qquad T = LTM.$$

So

$$\Lambda(B) = \Lambda(\tilde{S}, T) = \Lambda(S - T \begin{bmatrix} \boldsymbol{e} \\ \boldsymbol{e} \end{bmatrix} [\mathbf{0}, \boldsymbol{u}^T], T).$$

Note that the vector $[\boldsymbol{e}^T, \boldsymbol{e}^T]^T$ is an eigenvector of the pair (S, T) corresponding to the eigenvalue 1. So there exist unitary matrices U and V such that

$$U^H S V = R \qquad \text{and} \qquad U^H T V = \tilde{R}$$

with

$$\boldsymbol{v}_1 = \frac{1}{\sqrt{2k}} \begin{bmatrix} \boldsymbol{e} \\ \boldsymbol{e} \end{bmatrix}.$$

Since $\boldsymbol{v}_1$ is an eigenvector of the pair (S, T) corresponding to the eigenvalue 1, we have $r_{11}/\tilde{r}_{11} = 1$. By Theorem 2 we know that

$$\begin{aligned} \Lambda(A) &= \Lambda(S, T) \\ &= \{\tfrac{r_{11}}{\tilde{r}_{11}}, \tfrac{r_{22}}{\tilde{r}_{22}}, \ldots, \tfrac{r_{l,l}}{\tilde{r}_{l,l}}\} \\ &= \{1, \tfrac{r_{22}}{\tilde{r}_{22}}, \ldots, \tfrac{r_{l,l}}{\tilde{r}_{l,l}}\}. \end{aligned} \tag{9}$$

Here we assumed that $\tilde{r}_{i,i} = 0$ for $i = l+1, \ldots, 2k$.

Since $U^H T [\boldsymbol{e}^T, \boldsymbol{e}^T]^T = \tilde{R} V^H (\sqrt{2k} \boldsymbol{v}_1) = \sqrt{2k} \tilde{r}_{11} \boldsymbol{e}_1$, where $\boldsymbol{e}_1$ is the first standard unit vector of length $2k$, and the first component of the row vector $[\boldsymbol{0}, \boldsymbol{u}^T] V$ is $1/\sqrt{2k}$, it follows that the matrix

$$U^H \left\{ S - T \begin{bmatrix} \boldsymbol{e} \\ \boldsymbol{e} \end{bmatrix} [\boldsymbol{0}, \boldsymbol{u}^T] \right\} V = R - U^H T \begin{bmatrix} \boldsymbol{e} \\ \boldsymbol{e} \end{bmatrix} [\boldsymbol{0}, \boldsymbol{u}^T] V$$

is upper triangular and has $r_{11} - \tilde{r}_{11}$, r_{22}, $\ldots$, $r_{2k,2k}$ as its diagonal elements. Hence by Theorem 2 we have

$$\begin{aligned} \Lambda(B) &= \Lambda(\tilde{S}, T) \\ &= \Lambda(S - T \begin{bmatrix} \boldsymbol{e} \\ \boldsymbol{e} \end{bmatrix} [\boldsymbol{0}, \boldsymbol{u}^T], T) \\ &= \{\tfrac{r_{11} - \tilde{r}_{11}}{\tilde{r}_{11}}, \tfrac{r_{22}}{\tilde{r}_{22}}, \ldots, \tfrac{r_{l,l}}{\tilde{r}_{l,l}}\} \\ &= \{0, \tfrac{r_{22}}{\tilde{r}_{22}}, \ldots, \tfrac{r_{l,l}}{\tilde{r}_{l,l}}\}. \end{aligned} \tag{10}$$

From (9) and (10) we see that $\Lambda(B)$ is the same as $\Lambda(A)$ except that in $\Lambda(B)$ $1 \in \Lambda(A)$ is replaced by 0. $\square$

3.2 Solutions of the shifted matrix equation

The blocks B_0, B_1, B_2 are such that $H = G - \boldsymbol{e}\boldsymbol{u}^T$ solves the nonlinear matrix equation (7). The following theorem shows how the minimal nonnegative solutions of the nonlinear matrix equations (11) are transformed by the shifting:

Theorem 3 *Let $\boldsymbol{u}$ be positive, and let G, R, S and F be the minimal non-negative solutions of the matrix equations*

$$\begin{array}{l} G = A_0 + A_1 G + A_2 G^2 \\ S = A_0 S^2 + A_1 S + A_2 \\ R = R^2 A_0 + R A_1 + A_2 \\ F = A_0 + F A_1 + F^2 A_2, \end{array} \tag{11}$$

respectively. Then $H = G - \boldsymbol{e}\boldsymbol{u}^T$, $K = (I - \boldsymbol{e}\boldsymbol{u}^T S)S(I - \boldsymbol{e}\boldsymbol{u}^T S)^{-1}$, $T = R$ and $V = F - \boldsymbol{w}\boldsymbol{u}^T (I - A_1 - A_0 S)^{-1}$, where $\boldsymbol{w} = (1 - \boldsymbol{u}^T S\boldsymbol{e})^{-1} A_0 (I - S)\boldsymbol{e}$, solve the matrix equations:

$$\begin{array}{l} H = B_0 + B_1 H + B_2 H^2 \\ K = B_0 K^2 + B_1 K + B_2 \\ T = T^2 B_0 + T B_1 + B_2 \\ V = B_0 + V B_1 + V^2 B_2. \end{array} \tag{12}$$

Proof. First observe that if $\boldsymbol{u}$ is positive then $\boldsymbol{u}^T S\boldsymbol{e} < 1$, since $S\boldsymbol{e} \le \boldsymbol{e}$ and S is substochastic [10, 14], hence $I - \boldsymbol{e}\boldsymbol{u}^T S$ is nonsingular.

The matrix H solves the first equation of (12) by construction. For the matrix K we have:

$$\begin{array}{l} B_0 K^2 + B_1 K + B_2 = \\ B_0 (I - \boldsymbol{e}\boldsymbol{u}^T S) S^2 (I - \boldsymbol{e}\boldsymbol{u}^T S)^{-1} + B_1 (I - \boldsymbol{e}\boldsymbol{u}^T S) S (I - \boldsymbol{e}\boldsymbol{u}^T S)^{-1} + B_2 = \\ \Big(A_0 (I - \boldsymbol{e}\boldsymbol{u}^T) S^2 + \Big(A_1 + (A_0 \boldsymbol{e} - \boldsymbol{e}) \boldsymbol{u}^T S + A_2 \boldsymbol{e}\boldsymbol{u}^T \Big) S + \\ + A_2 (I - \boldsymbol{e}\boldsymbol{u}^T S) \Big) (I - \boldsymbol{e}\boldsymbol{u}^T S)^{-1} = \\ (S - \boldsymbol{e}\boldsymbol{u}^T S^2)(I - \boldsymbol{e}\boldsymbol{u}^T S)^{-1} = K. \end{array}$$

For the matrix $T = R$, since [10] $R = A_2 (I - A_1 - A_2 G)^{-1}$ and

$$\begin{array}{rcl} A_2 \boldsymbol{e} & = & R(I - A_1 - R A_0)\boldsymbol{e} = R\Big(I - A_1 - A_2 (I - A_1 - A_2 G)^{-1} A_0 \Big) \boldsymbol{e} = \\ & & R(I - A_1 - A_2)\boldsymbol{e} = R A_0 \boldsymbol{e}, \end{array}$$

we obtain

$$\begin{array}{rcl} B_2 + R B_1 + R^2 B_0 & = & A_2 + R(A_1 + A_2 \boldsymbol{e}\boldsymbol{u}^T) + R^2 A_0 (I - \boldsymbol{e}\boldsymbol{u}^T) = \\ & & A_2 + R A_1 + R^2 A_0 + R A_2 \boldsymbol{e}\boldsymbol{u}^T - R^2 A_0 \boldsymbol{e}\boldsymbol{u}^T = \\ & & R + R(A_2 \boldsymbol{e} - R A_0 \boldsymbol{e})\boldsymbol{u}^T = R. \end{array}$$

Concerning the matrix V, observe that $I - B_1 - B_0 K$ is nonsingular, since $I - A_1 - A_0 S - B_2 G$ is nonsingular [10, 14], and $I - A_1 - A_0 S - B_2 G =$

$(I - B_1 - B_0K - B_2H)(I - \boldsymbol{e}\boldsymbol{u}^T S) = (I - B_1 - B_0K)(I - KH)(I - \boldsymbol{e}\boldsymbol{u}^T S)$. It can be shown by direct substitution that $V = B_0(I - B_1 - B_0K)^{-1}$ solves the last matrix equation in (12). By rewriting the matrix V in terms of the blocks A_0, A_1, A_2 we obtain that:

$$V = A_0(I - \boldsymbol{e}\boldsymbol{u}^T)(I - A_1 - A_0S - A_2\boldsymbol{e}\boldsymbol{u}^T)^{-1}.$$

On the other hand it holds:

$$\begin{aligned}
&(I - A_1 - A_0S - A_2\boldsymbol{e}\boldsymbol{u}^T)^{-1} = \\
&(I - A_1 - A_0S)^{-1} + \\
&(1 - \boldsymbol{u}^T(I - A_1 - A_0S)^{-1}A_2\boldsymbol{e})^{-1}(I - A_1 - A_0S)^{-1}A_2\boldsymbol{e}\boldsymbol{u}^T(I - A_1 - A_0S)^{-1} = \\
&(I - A_1 - A_0S)^{-1} + (1 - \boldsymbol{u}^T S\boldsymbol{e})^{-1}S\boldsymbol{e}\boldsymbol{u}^T(I - A_1 - A_0S)^{-1} = \\
&\left(I - (1 - \boldsymbol{u}^T S\boldsymbol{e})^{-1}S\boldsymbol{e}\boldsymbol{u}^T\right)(I - A_1 - A_0S)^{-1}.
\end{aligned}$$

From the latter equation and from the relation $F = A_0(I - A_1 - A_0S)^{-1}$ [10], we obtain:

$$\begin{aligned}
V &= F + (1 - \boldsymbol{u}^T S\boldsymbol{e})^{-1}A_0S\boldsymbol{e}\boldsymbol{u}^T(I - A_1 - A_0S)^{-1} - \\
&\quad - \left(A_0\boldsymbol{e} + (\boldsymbol{u}^T S\boldsymbol{e})(1 - \boldsymbol{u}^T S\boldsymbol{e})^{-1}A_0\boldsymbol{e}\right)\boldsymbol{u}^T(I - A_1 - A_0S)^{-1} = \\
&\quad F - \boldsymbol{w}\boldsymbol{u}^T(I - A_1 - A_0S)^{-1},
\end{aligned}$$

where $\boldsymbol{w} = (1 - \boldsymbol{u}^T S\boldsymbol{e})^{-1}A_0(I - S)\boldsymbol{e}$. $\square$

The spectral properties of the solutions of the shifted matrix equations allow us to prove that the rate of convergence is improved, with respect to standard cyclic reduction. Let us define

$$\begin{aligned}
\bar{\theta} &= \max\{|\lambda| : \ |\lambda| < 1, \ \phi(\lambda) = 0\} \\
\sigma &= \min\{|\lambda| : \ |\lambda| > 1, \ \phi(\lambda) = 0\}.
\end{aligned} \tag{13}$$

The convergence properties that will be proved in the next section rely on the fact that $\bar{\theta} < 1$, $\sigma = 1/\rho(R) > 1$, S and R have the same eigenvalues (see [14]), and thus also K and R, and $\rho(K) = 1/\sigma < 1$.

4 A shifted cyclic reduction algorithm

In this section we apply the cyclic reduction algorithm to solve the shifted matrix equation (7), and we show that the rate of convergence is improved, with respect to the same algorithm applied to the original matrix equation (1).

By following the approach described in section 2, we generate by means of cyclic reduction the sequence of infinite block tridiagonal systems

$$\begin{bmatrix} I-\widehat{B}_1^{(j)} & -B_2^{(j)} & & 0 \\ -B_0^{(j)} & I-B_1^{(j)} & -B_2^{(j)} & \\ & -B_0^{(j)} & I-B_1^{(j)} & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} H \\ H^{2^j+1} \\ H^{2\cdot 2^j+1} \\ \vdots \end{bmatrix} = \begin{bmatrix} B_0 \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \quad j \geq 0, \tag{14}$$

where $B_0^{(0)} = B_0$, $\widehat{B}_1^{(0)} = B_1^{(0)} = B_1$, $B_2^{(0)} = B_2$, and the blocks $\widehat{B}_1^{(j)}$, $B_i^{(j)}$, $i = 0, 1, 2$, $j \geq 1$, are defined by the recurrences:

$$\begin{aligned} B_0^{(j+1)} &= B_0^{(j)} \left(I - B_1^{(j)}\right)^{-1} B_0^{(j)} \\ B_1^{(j+1)} &= B_1^{(j)} + B_0^{(j)} \left(I - B_1^{(j)}\right)^{-1} B_2^{(j)} + B_2^{(j)} \left(I - B_1^{(j)}\right)^{-1} B_0^{(j)} \\ B_2^{(j+1)} &= B_2^{(j)} \left(I - B_1^{(j)}\right)^{-1} B_2^{(j)} \\ \widehat{B}_1^{(j+1)} &= \widehat{B}_1^{(j)} + B_2^{(j)} \left(I - B_1^{(j)}\right)^{-1} B_0^{(j)}. \end{aligned} \tag{15}$$

In the shifted case, the problem of the nonsingularity of the blocks $I - B_1^{(j)}$ must be considered. In fact, in the original case, the matrices $I - A_1^{(j)}$ are nonsingular M-matrices for any j, since $I - A_1^{(j)}$ can be viewed as a Schur complement of the block $(2^{j+1}-1) \times (2^{j+1}-1)$ matrix obtained by truncating the infinite matrix (2) at the block size $2^{j+1}-1$ (see [1, 6]), and this finite matrix is a nonsingular M-matrix. Analogously, the matrix $I - B_1^{(j)}$ can be viewed as a Schur complement of the matrix $Q_{2^{j+1}-2}$, where Q_n is the $(n+1) \times (n+1)$ block matrix

$$Q_n = \begin{bmatrix} I-B_1 & -B_2 & & 0 \\ -B_0 & \ddots & \ddots & \\ & \ddots & \ddots & -B_2 \\ 0 & & -B_0 & I-B_1 \end{bmatrix}. \tag{16}$$

Thus, the $(j+1)$-st step of cyclic reduction can be performed, i.e., $I - B_1^{(j)}$ is nonsingular, if and only if $Q_{2^{j+1}-2}$ is nonsingular [1]. Based on this property, we prove the following result:

Theorem 4 *Let $\boldsymbol{u}$ be any positive vector such that $\boldsymbol{u}^T\boldsymbol{e} = 1$ and let $H = G - \boldsymbol{u}\boldsymbol{e}^T$, $K = (I - \boldsymbol{e}\boldsymbol{u}^T S) S (I - \boldsymbol{e}\boldsymbol{u}^T S)^{-1}$, where G and S are the minimal nonnegative solutions of the matrix equations $G = A_0 + A_1 G + A_2 G^2$, $S =$*

$A_0S^2 + A_1S + A_2$. *Let* $B_0^{(j)}$, $B_1^{(j)}$, $B_2^{(j)}$, $\widehat{B}_1^{(j)}$ *be the blocks generated at the* j*-th step of cyclic reduction,* $j \geq 1$. *Then, the following conditions are equivalent:*

1. $I - (HK^{2^{j+1}-1})(KH^{2^{j+1}-1})$ *is nonsingular;*
2. $I - (KH^{2^{j+1}-1})(HK^{2^{j+1}-1})$ *is nonsingular;*
3. $I - B_1^{(j)}$ *is nonsingular.*

Proof. The matrix $I - B_1^{(j)}$ is nonsingular if and only if $Q_{2^{j+1}-2}$ is singular, where Q_n is defined in (16). For simplicity of notations, set $n = 2^{j+1} - 2$. Observe that the matrix $W = I - B_1 - B_0K - B_2H$ is nonsingular, since $W = (I - A_1 - A_0S - A_2G)(I - \boldsymbol{e}\boldsymbol{u}^TS)$ and $I - A_1 - A_0S - B_2G$ is nonsingular (see [10, 14]). Let $\boldsymbol{x} = (\boldsymbol{x}_i)_{i=0,\ldots,n}$ such that $Q_n\boldsymbol{x} = 0$. Therefore, since $I - B_1 - B_0K - B_2H$ is nonsingular, then

$$\boldsymbol{x}_i = H^i\boldsymbol{r} + K^{n-i}\boldsymbol{s}, \quad i = 0, \ldots, n$$

where $\boldsymbol{r}$ and $\boldsymbol{s}$ are suitable vectors, such that the boundary conditions

$$\begin{array}{l} (I - B_1)\boldsymbol{x}_0 - B_2\boldsymbol{x}_1 = \boldsymbol{0} \\ -B_0\boldsymbol{x}_{n-1} + (I - B_1)\boldsymbol{x}_n = \boldsymbol{0} \end{array} \tag{17}$$

are satisfied (see [15]). By imposing the above equalities, we obtain that $\boldsymbol{r}$ and $\boldsymbol{s}$ must solve the following homogeneous linear system:

$$\left[\begin{array}{cc} I - B_1 - B_2H & ((I - B_1)K - B_2)\,K^{n-1} \\ ((I - B_1)H - B_0)\,H^{n-1} & I - B_1 - B_0K \end{array}\right]\left[\begin{array}{c} \boldsymbol{r} \\ \boldsymbol{s} \end{array}\right] = 0,$$

that can be written as

$$\left[\begin{array}{cc} I - B_1 - B_2H & B_0K^{n+1} \\ B_2H^{n+1} & I - B_1 - B_0K \end{array}\right]\left[\begin{array}{c} \boldsymbol{r} \\ \boldsymbol{s} \end{array}\right] = 0, \tag{18}$$

since H and K solve the matrix equations (1), (7). Hence, $\boldsymbol{x} = \boldsymbol{0}$ if and only if $\boldsymbol{r} = \boldsymbol{s} = \boldsymbol{0}$. Now, the block diagonal entries in the matrix of (18) are nonsingular since $I - B_1 - B_0K - B_2H = (I - B_1 - B_2H)(I - HK) = (I - B_1 - B_0K)(I - KH)$, and $I - B_1 - B_0K - B_2H$ is nonsingular. By computing the Schur complement of $I - B_1 - B_2H$ in $I - B_1 - B_0K$ we obtain the matrix

$$\begin{array}{rcl} S_1 & = & I - B_1 - B_0K - B_2H^{n+1}(I - B_1 - B_2H)^{-1}B_0K^{n+1} = \\ & & I - B_1 - B_0K - B_2H^{n+2}K^{n+1}. \end{array}$$

Thus, S_1 (and hence Q_n) is singular if and only if $(I - B_1 - B_0K)^{-1}S_1 = I - KH^{n+2}K^{n+1}$ is nonsingular. By taking the Schur complement of $I - B_1 - B_0K$ in $I - B_1 - B_0K$ we complete the proof of the theorem. □

The above theorem gives necessary and sufficient conditions for the applicability of cyclic reduction. Since $\rho(H) < 1$ and $\rho(K) < 1$, the matrix $Z_j = I - (HK^{2^{j+1}-1})(KH^{2^{j+1}-1})$ is nonsingular for sufficiently large values of j. From the numerical experiments that we have performed, and from the fact that $I - (GS^{2^{j+1}-1})(SG^{2^{j+1}-1})$ is nonsingular for any value of j, we conjecture that also Z_j is nonsingular for any value of j. However, if it were not, the cyclic reduction algorithm can be still applied, by performing a different permutation of block rows and columns, that allows one to skip the steps that cannot be performed for the singularity of the blocks $I - B_1^{(j)}$ (we refer to the paper [1] for details of this subject).

Here and hereafter we assume that the matrices $I - B_1^{(j)}$ are nonsingular for any j.

The nice convergence properties stated by the next theorem enable us to efficiently compute the matrix H:

Theorem 5 *Let $B_0^{(j)}$, $B_1^{(j)}$, $B_2^{(j)}$ be the blocks generated at the j-th step of cyclic reduction. Then $B_1^{(j)}$, $(I - B_1^{(j)})^{-1}$ are bounded, and for any operator norm it holds $||B_2^{(j)}|| = O\left((1/\sigma)^{2^j}\right)$ and $||B_0^{(j)}|| = O\left(\theta^{2^j}\right)$, for any $\bar\theta < \theta < 1$, where $\bar\theta$ and σ are defined in (13).*

Proof. From (14), at each step j the following equations hold:

$$H^{2^j} = B_0^{(j)} + B_1^{(j)} H^{2^j} + B_2^{(j)} H^{2\cdot 2^j}, \tag{19}$$
$$K^{2^j} = B_0^{(j)} K^{2\cdot 2^j} + B_1^{(j)} K^{2^j} + B_2^{(j)}, \tag{20}$$

and therefore

$$(I - B_1^{(j)})^{-1} B_0^{(j)} + H^{2^j} + (I - B_1^{(j)})^{-1} B_2^{(j)} H^{2\cdot 2^j} = 0, \tag{21}$$
$$(I - B_1^{(j)})^{-1} B_2^{(j)} + K^{2^j} + (I - B_1^{(j)})^{-1} B_0^{(j)} K^{2\cdot 2^j} = 0. \tag{22}$$

We first prove that the matrices $C_0^{(j)} = (I - B_1^{(j)})^{-1} B_0^{(j)}$ and $C_2^{(j)} = (I - B_1^{(j)})^{-1} B_2^{(j)}$ are bounded in norm.

Since the maximum modulus eigenvalue of R, and thus of K, is real, simple and unique [8] then there exists an operator norm $|| \cdot ||_K$ such that $||K||_K = \rho(K)$. Moreover, let $\epsilon > 0$ such that $\rho(H) + \epsilon < 1$, and let $|| \cdot ||_{H,\epsilon}$ be an operator norm such that $||H||_{H,\epsilon} \le \rho(H) + \epsilon$ (see [9]).

Let $\alpha_j = ||C_0^{(j)}||_{H,\epsilon}$. If α_j is not bounded, then there exists a subsequence α_{j_h} such that α_{j_h} diverges to infinity. From (21) we have

$$\begin{aligned}\alpha_{j_h} \quad \leq \quad & ||C_2^{(j_h)}||_{H,\epsilon}||H^{2\cdot 2^{j_h}}||_{H,\epsilon} + ||H^{2^{j_h}}||_{H,\epsilon} \leq \\ & (\rho(H)+\epsilon)^{2^{j_h}}\left((\rho(H)+\epsilon)^{2^{j_h}}||C_2^{(j_h)}||_{H,\epsilon} + 1\right).\end{aligned}$$

Thus,

$$||C_2^{(j_h)}||_{H,\epsilon} \geq \left(\frac{\alpha_{j_h}}{(\rho(H)+\epsilon)^{2^{j_h}}} - 1\right)\frac{1}{(\rho(H)+\epsilon)^{2^{j_h}}}$$

and, since α_{j_h} diverges to infinity, there exists a constant $c > 0$ such that

$$||C_2^{(j_h)}||_{H,\epsilon} \geq c\frac{\alpha_{j_h}}{(\rho(H)+\epsilon)^{2^{j_h}}}.$$

For the equivalence of the operator norms, there exist constants $c' > 0$ and $c'' > 0$ such that

$$||C_2^{(j_h)}||_K \geq c'\frac{\alpha_{j_h}}{(\rho(H)+\epsilon)^{2^{j_h}}}, \tag{23}$$

$$||C_0^{(j)}||_K \leq c''||C_0^{(j)}||_{H,\epsilon} = c''\alpha_j. \tag{24}$$

On the other hand, from (22) we have

$$||C_2^{(j_h)}||_K \leq ||C_0^{(j_h)}||_K||K^{2\cdot 2^{j_h}}||_K + ||K^{2^{j_h}}||_K.$$

Thus, from the latter inequality and (24) we have

$$||C_2^{(j_h)}||_K \leq c''\alpha_{j_h}\rho(K)^{2\cdot 2^{j_h}} + \rho(K)^{2^{j_h}}.$$

Hence, from (23), we obtain

$$c'\frac{\alpha_{j_h}}{(\rho(H)+\epsilon)^{2^{j_h}}} \leq c''\alpha_{j_h}\rho(K)^{2\cdot 2^{j_h}} + \rho(K)^{2^{j_h}},$$

that contradicts the assumption that α_{j_h} goes to infinity. By using a similar argument we can prove that also $C_2^{(j)}$ is bounded in norm. From (21), we obtain

$$||C_0^{(j)}||_{H,\epsilon} \leq (\rho(H)+\epsilon)^{2^j} + ||C_2^{(j)}||_{H,\epsilon}(\rho(H)+\epsilon)^{2\cdot 2^j}.$$

Hence, since $||C_2^{(j)}||_{H,\epsilon}$ is bounded, then there exists a constant $\gamma > 0$ such that

$$||C_0^{(j)}||_{H,\epsilon} \leq \gamma(\rho(H)+\epsilon)^{2^j}.$$

Thus, for the equivalence of the operator norms, for any operator norm $||\cdot||$ there exists a constant γ such that

$$||C_0^{(j)}|| \leq \gamma(\rho(H)+\epsilon)^{2^j}. \tag{25}$$

Similarly, for any operator norm $||\cdot||$ there exists a constant γ' such that

$$||C_2^{(j)}|| \leq \gamma'\rho(K)^{2^j}. \tag{26}$$

From (25) and (26), since

$$I - B_1^{(j+1)} = (I - B_1^{(j)})(I - C_0^{(j)}C_2^{(j)} + C_2^{(j)}C_0^{(j)}),$$

we have

$$||I - B_1^{(j+1)}|| \leq ||I - B_1^{(j)}||(1+\sigma_j)$$

where $\sigma_j = O\left((\rho(H)+\epsilon)^{2^j}\rho(K)^{2^j}\right)$. Thus the matrices $I - B_1^{(j)}$, and hence the matrices $B_1^{(j)}$, are bounded in norm. Similarly, it holds

$$||(I - B_1^{(j+1)})^{-1}|| \leq \frac{1}{1-\sigma_j}||(I - B_1^{(j)})^{-1}||,$$

thus $||(I - B_1^{(j)})^{-1}||$ is bounded.

Now, from the boundness of $||I - B_1^{(j)}||$, and from the relations

$$\begin{aligned} -B_0^{(j)} + (I - B_1^{(j)})H^{2^j} - B_2^{(j)}H^{2\cdot 2^j} &= 0, \\ -B_2^{(j)} + (I - B_1^{(j)})K^{2^j} - B_0^{(j)}K^{2\cdot 2^j} &= 0, \end{aligned}$$

derived from (19) and (20), by applying the same argument used to derive (25) and (26), we can show that $||B_2^{(j)}||$ and $||B_0^{(j)}||$ are bounded, and thus that $||B_0^{(j)}|| = O\left((\rho(H)+\epsilon)^{2^j}\right)$, $||B_2^{(j)}|| = O\left(\rho(K)^{2^j}\right)$. $\square$

The matrix G can be directly recovered, without computing the matrix H, according to the following result:

Theorem 6 *For any operator norm $||\cdot||$ and for any $\bar{\theta} < \theta < 1$ it holds*

$$A_0 - (I - \widehat{B}_1^{(j)})G = O\left(\left(\frac{\theta}{\sigma}\right)^{2^j}\right). \tag{27}$$

Proof. From the relation

$$(I - \widehat{B}_1^{(j)})H - B_2^{(j)}H^{2^j+1} - B_0 = 0$$

and from Theorem 5 it follows that for any operator norm $||\cdot||$ and for any $\bar{\theta} < \theta < 1$ it holds

$$B_0 - (I - \widehat{B}_1^{(j)})H = O\left(\left(\frac{\theta}{\sigma}\right)^{2^j}\right).$$

On the other hand, it can be easily proved by induction that $(I - \widehat{B}_1^{(j)})\boldsymbol{e} = A_0\boldsymbol{e}$ for any $j \geq 0$. Thus, by replacing H with $G - \boldsymbol{e}\boldsymbol{u}^T$ and B_0 with $A_0 - A_0\boldsymbol{e}\boldsymbol{u}^T$, we arrive at (27). □

From the above result, compared with (6), it follows that the shifted cyclic reduction can be much faster than the original one. Indeed, if the second largest modulus eigenvalue $\bar{\theta}$ of G is far from the unit circle, then the rate of convergence is much improved.

Thus, the deflating technique leads to a better rate of convergence, but destroys the nonnegativity and M-matrix properties of the blocks generated at each step. Indeed, in general neither $I - B_1^{(j)}$ is an M-matrix, nor $B_0^{(j)}$, $B_2^{(j)}$ are nonnegative matrices (from the numerical experiments it seems that $I - \widehat{B}_1^{(j)}$ is an M-matrix for $j = 0, 1, \ldots$). In principle this fact could lead to a loss of accuracy of the results obtained with the shifting technique. In practice, we have not observed any differences, in terms of accuracy, between the results obtained with the two algorithms. Furthermore, as we will prove in the next section, the shifted equation is better conditioned than the original one.

5 Conditioning of the shifted matrix equation

In this section we introduce a measure of the conditioning of the matrix equation, and we show that the shifted equation is better conditioned than the original one.

Consider the matrix equation (1) and a perturbed equation

$$G+\triangle G = (A_0+\triangle A_0)+(A_1+\triangle A_1)(G+\triangle G)+(A_2+\triangle A_2)(G+\triangle G)^2. \quad (28)$$

Using (1) this perturbed equation simplifies up to the first oder in $\triangle G$

$$(I - A_1 - A_2G)\triangle G - A_2(\triangle G)G = \triangle A, \quad (29)$$

where

$$\triangle A = \triangle A_0 + (\triangle A_1)G + (\triangle A_2)G^2.$$

Note that (29) can be written

$$W\mathbf{vec}(\triangle G) = \mathbf{vec}(\triangle A), \tag{30}$$

where

$$W = I \otimes (I - A_1 - A_2 G) + G^T \otimes (-A_2)$$

and $\mathbf{vec}(A)$ is the k^2-dimensional vector obtained by column-wise arranging the entries of the matrix A.

Let $\Lambda(M)$ denote the set of all eigenvalues of any square matrix M.

Theorem 7 *The $k^2 \times k^2$ matrix W is nonsingular. Furthermore*

$$\min\{|\lambda| : \lambda \in \Lambda(W)\} = 1 - \rho(A_1 + A_2 G + A_2).$$

Proof. Let S be the Schur canonical form of G^T. Then the matrix W is similar to

$$\tilde{W} = I \otimes (I - A_1 - A_2 G) + S \otimes (-A_2).$$

Due to the upper triangular structure of S and the fact that $\Lambda(S) = \Lambda(G)$ it is easy to see that

$$\Lambda(W) = \Lambda(\tilde{W}) = \bigcup_{\lambda \in \Lambda(G)} \Lambda(I - A_1 - A_2 G - \lambda A_2).$$

Note that whenever $\lambda \in \Lambda(G)$

$$\begin{aligned} \rho(A_1 + A_2 G + \lambda A_2) &\leq \rho(|A_1 + A_2 G + \lambda A_2|) \\ &\leq \rho(A_1 + A_2 G + |\lambda| A_2) \\ &\leq \rho(A_1 + A_2 G + A_2) \\ &< 1. \end{aligned}$$

In the last step we used the fact that since $A_0 + A_1 + A_2$ is irreducible and positive recurrent the matrix $A_1 + A_2 G + A_2$ has spectral radius less than 1. It follows that 0 does not belong to $\Lambda(W)$, and hence W is nonsingular. Clearly

$$\min\{|\lambda| : \lambda \in \Lambda(W)\} = 1 - \rho(A_1 + A_2 G + A_2)$$

□

Since W is nonsingular

$$\mathbf{vec}(\triangle G) = W^{-1}\mathbf{vec}(\triangle A).$$

Hence

$$\begin{aligned}
\|\triangle G\|_F &= \|\mathbf{vec}(\triangle G)\|_2 \\
&\leq \|W^{-1}\|_2\|\mathbf{vec}(\triangle A)\|_2 \\
&= \frac{1}{\sigma_{min}(W)}\|\triangle A_0 + (\triangle A_1)G + (\triangle A_2)G^2\|_F \\
&\leq \frac{\sqrt{k}}{\sigma_{min}(W)}(\|\triangle A_0\|_F + \|\triangle A_1\|_F + \|\triangle A_2\|_F).
\end{aligned}$$

Here $\sigma_{min}(W)$ is the minimum singular value of the matrix W. So we may view $1/\sigma_{min}(W)$ as a condition number of the equation (1). Even though $1/\min\{|\lambda| : \lambda \in \Lambda(W)\} \leq 1/\sigma_{min}(W)$, we may consider

$$\frac{1}{\min\{|\lambda| : \lambda \in \Lambda(W)\}} = \frac{1}{1 - \rho(A_1 + A_2 G + A_2)}$$

as a number which reflects the conditioning of the equation (1).

Consider now the shifted matrix equation (7) and the perturbed equation

$$H + \triangle H = (B_0 + \triangle B_0) + (B_1 + \triangle B_1)(H + \triangle H) + (B_2 + \triangle B_2)(H + \triangle H)^2.$$

As before this perturbed equation simplifies up to the first oder in $\triangle H$

$$(I - B_1 - B_2 H)\triangle H - B_2(\triangle H)H = \triangle B, \tag{31}$$

where

$$\triangle B = \triangle B_0 + (\triangle B_1)H + (\triangle B_2)H^2.$$

Note that (31) can be written

$$Q\mathbf{vec}(\triangle H) = \mathbf{vec}(\triangle B),$$

where

$$Q = I \otimes (I - B_1 - B_2 H) + H^T \otimes (-B_2).$$

Note that

$$B_1 + B_2 H = A_1 + A_2 \boldsymbol{e}\boldsymbol{u}^T + A_2(G - \boldsymbol{e}\boldsymbol{u}^T) = A_1 + A_2 G.$$

Hence

$$Q = I \otimes (I - A_1 - A_2 G) + H^T \otimes (-A_2).$$

Note that if we denote $\Lambda(G)$ by

$$\Lambda(G) = \{\lambda_1, \lambda_2, \ldots, \lambda_k\}$$

with

$$1 = \lambda_1 > |\lambda_2| \geq \ldots \geq |\lambda_k|$$

then

$$\Lambda(H) = \{\lambda_2, \lambda_3, \ldots, \lambda_k, 0\}$$

and $|\lambda_2| = \bar{\theta}$.

Theorem 8 *The $k^2 \times k^2$ matrix Q is nonsingular. Furthermore*

$$\min\{|\lambda| : \lambda \in \Lambda(Q)\} \geq 1 - \rho(A_1 + A_2G + \bar{\theta}A_2)$$

and the equality holds if λ_2 is a positive real number.

Proof. As in the proof of Theorem 7 it is easy to see that

$$\Lambda(Q) = \Lambda(\tilde{Q}) = \bigcup_{\lambda \in \Lambda(H)} \Lambda(I - A_1 - A_2G - \lambda A_2).$$

Note that since $\rho(H) = \bar{\theta} = |\lambda_2|$ whenever $\lambda \in \Lambda(H)$

$$\begin{aligned} \rho(A_1 + A_2G + \lambda A_2) &\leq \rho(|A_1 + A_2G + \lambda A_2|) \\ &\leq \rho(A_1 + A_2G + |\lambda| A_2) \\ &\leq \rho(A_1 + A_2G + |\lambda_2| A_2) \\ &< 1. \end{aligned}$$

In the last step we used the fact that

$$\rho(A_1 + A_2G + |\lambda_2|A_2) < \rho(A_1 + A_2G + A_2) < 1.$$

It follows that 0 does not belong to $\Lambda(Q)$, and hence Q is nonsingular. Clearly

$$\min\{|\lambda| : \lambda \in \Lambda(Q)\} \geq 1 - \rho(A_1 + A_2G + |\lambda_2|A_2)$$

and the equality holds if λ_2 is a positive real number. $\square$

Since Q is nonsingular

$$\mathbf{vec}(\triangle H) = Q^{-1}\mathbf{vec}(\triangle B).$$

Hence

$$\begin{aligned}
\|\triangle H\|_F &= \|\mathbf{vec}(\triangle H)\|_2 \\
&\le \|Q^{-1}\|_2 \|\mathbf{vec}(\triangle B)\|_2 \\
&= \frac{1}{\sigma_{min}(Q)} \|\triangle B_0 + (\triangle B_1)H + (\triangle B_2)H^2\|_F \\
&\le \frac{2\sqrt{k}}{\sigma_{min}(Q)} (\|\triangle B_0\|_F + \|\triangle B_1\|_F + \|\triangle B_2\|_F).
\end{aligned}$$

Here $\sigma_{min}(Q)$ is the minimum singular value of the matrix Q. So we may view $1/\sigma_{min}(Q)$ as a condition number of the shifted equation (7). Even though $1/\min\{|\lambda| : \lambda \in \Lambda(Q)\} \le 1/\sigma_{min}(Q)$, we may consider $1/\min\{|\lambda| : \lambda \in \Lambda(Q)\}$ as a number which reflects the conditioning of the shifted equation (7). Since $1/\min\{|\lambda| : \lambda \in \Lambda(Q)\} \le 1/(1 - \rho(A_1 + A_2 G + \bar{\theta} A_2))$,

$$\frac{1}{1 - \rho(A_1 + A_2 G + \bar{\theta} A_2)}$$

is a number which reflects the conditioning of the shifted equation (7). The inequality

$$\frac{1}{1 - \rho(A_1 + A_2 G + \bar{\theta} A_2)} < \frac{1}{1 - \rho(A_1 + A_2 G + A_2)},$$

suggests that the shifted equation (7) has a better conditioning than the original equation (1).

6 Numerical results

We have tested the cyclic reduction algorithm and the shifted cyclic reduction algorithm on the examples in [13] using Matlab. In the case of the cyclic reduction algorithm we stopped when

$$\|\widehat{A}_1^{(j)} - \widehat{A}_1^{(j-1)}\|_\infty \le 10^{-12}$$

and we accepted

$$\tilde{G} = \left(I - \widehat{A}_1^{(j)}\right)^{-1} A_0$$

as an approximation of G. Similarly, in the case of a shifted cyclic reduction we stopped when

$$\|\widehat{B}_1^{(j)} - \widehat{B}_1^{(j-1)}\|_\infty \le 10^{-12}$$

	Cyclic Reduction			Shifted Cyclic Reduction		
p	Iter.	Res.	Stoc.	Iter.	Res.	Stoc.
10^{-1}	9	0.0	$6.7 \cdot 10^{-16}$	2	0.0	$2.2 \cdot 10^{-16}$
10^{-2}	13	0.0	$7.9 \cdot 10^{-15}$	2	0.0	$1.1 \cdot 10^{-16}$
10^{-3}	16	0.0	$6.0 \cdot 10^{-15}$	2	0.0	0.0
10^{-4}	20	$1.1 \cdot 10^{-16}$	$3.4 \cdot 10^{-13}$	2	0.0	0.0
10^{-5}	23	$2.2 \cdot 10^{-16}$	$1.3 \cdot 10^{-11}$	2	$2.2 \cdot 10^{-16}$	0.0
10^{-6}	26	$2.2 \cdot 10^{-16}$	$4.9 \cdot 10^{-11}$	2	$2.2 \cdot 10^{-16}$	$1.1 \cdot 10^{-16}$
10^{-7}	30	$2.2 \cdot 10^{-16}$	$3.1 \cdot 10^{-10}$	2	$5.6 \cdot 10^{-17}$	$2.2 \cdot 10^{-16}$
10^{-8}	33	0.0	$7.9 \cdot 10^{-9}$	2	0.0	0.0
10^{-9}	36	$1.1 \cdot 10^{-16}$	$3.0 \cdot 10^{-8}$	2	$5.6 \cdot 10^{-17}$	$1.1 \cdot 10^{-16}$
10^{-10}	39	0.0	$2.8 \cdot 10^{-7}$	2	0.0	0.0

Table 1: Example 1

and we accepted

$$\tilde{G} = \left(I - \widehat{B}_1^{(j)}\right)^{-1} A_0$$

as an approximation of G. For each example we have reported tables with the number of iterations, the residual error, Res., which is defined by $\|\tilde{G} - A_0 - A_1\tilde{G} - A_2\tilde{G}^2\|_\infty$, and the closeness to stochasticity, Stoc., which is defined by $\|\tilde{G}\boldsymbol{e} - \boldsymbol{e}\|_\infty$, for standard and shifted cyclic reduction. For particular examples we have also reported tables with the values of $\bar{\theta}$, of σ, $1/\sigma$ and $\bar{\theta}/\sigma$, to show the reduction of the rate of convergence, and tables with $1-\rho(A_1+A_2G+A_2)$, $1-\rho(A_1+A_2G+\bar{\theta}A_2)$, $\sigma_{min}(W)$, $\sigma_{min}(Q)$, that show the conditioning of the original and shifted matrix equations.

Example 1 Here the blocks, having dimension 2, depend on a parameter $p > 0$, and are defined as

$$A_0 = \begin{bmatrix} 1-p & 0 \\ 0 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & p \\ 2p & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1-2p \end{bmatrix}$$

We have tested with ten different p values and obtained the following results summarized in Tables 1, 2, and 3.

Example 2 In this example, A_0, A_1, and A_2 are defined as follows. First we define 24×24 matrices A'_0, A'_1, and A'_2 as follows:

$$(A'_0) = \begin{cases} 192(1-i/24), & i=j \\ 0, & i \neq j \end{cases} \qquad (A'_2) = \begin{cases} 192\rho_d, & i=j \\ 0, & i \neq j \end{cases}$$

p	θ	σ	$1/\sigma$	θ/σ
10^{-1}	0	1.125	0.8889	0
10^{-2}	0	1.0102	0.9899	0
10^{-3}	0	1.001	0.9990	0
10^{-4}	0	1.0001	0.9999	0
10^{-5}	0	1.00001	0.99999	0

Table 2: Example 1

p	$1-\rho(A_1+A_2G+A_2)$	$1-\rho(A_1+A_2G+\theta A_2)$	$\sigma_{min}(W)$	$\sigma_{min}(Q)$
10^{-1}	0.09	0.68	0.05	0.35
10^{-6}	$1.0 \cdot 10^{-6}$	0.999	$5.0 \cdot 10^{-7}$	0.30
10^{-10}	$1.0 \cdot 10^{-10}$	1.0	$5.0 \cdot 10^{-11}$	0.44

Table 3: Example 1

and

$$(A'_1) = \begin{cases} ar(M-i)/M, & i-j=-1 \\ ir, & i-j=1 \\ 0, & \text{elsewhere} \end{cases}$$

Here i and j are integers between 0 and 23, and a, r, M, and ρ_d are parameters. Now A_0, A_1, and A_2 are defined such that $A_0 = sA'_0$, $A_1 = sA'_1$, and $A_2 = sA'_2$ where s is a scalar which makes $A_0 + A_1 + A_2$ stochastic. Tables 4, 5, and 6 report the results obtained by choosing different values of M while we fix $r = 1/300$, $a = 18.244$, and $\rho_d = 0.280$. Tables 7, 8, and 9 report the results obtained by choosing different values of ρ_d while we fix $r = 1/100$, $a = 18.244$, and $M = 512$.

Example 3 In this example we construct a QBD problem defined by the $k \times k$ matrices $A_0 = R + \delta I$, $A_1 = A_2 = R$, where R is a matrix having null diagonal entries and constant off-diagonal entries, and $0 < \delta < 1$. As was observed in [13], the rate $\rho = p^T(A_1 + 2A_2)e$, where $p^T(A_0 + A_1 + A_2) = p^T$, $p^Te = 1$, is exactly $1-\delta$. We have tested with eight different δ values. Tables 10, 11, and 12 report the results obtained with $k = 16$ and Tables 13 and 14 report the results obtained with sizes $k = 32$ and $k = 64$, respectively.

	Cyclic Reduction			Shifted Cyclic Reduction		
M	Iter.	Res.	Stoc.	Iter.	Res.	Stoc.
64	19	$1.6 \cdot 10^{-16}$	$3.2 \cdot 10^{-12}$	18	$5.1 \cdot 10^{-16}$	$4.4 \cdot 10^{-16}$
128	20	$2.2 \cdot 10^{-16}$	$8.4 \cdot 10^{-13}$	19	$6.9 \cdot 10^{-16}$	$4.4 \cdot 10^{-16}$
256	21	$3.0 \cdot 10^{-16}$	$5.9 \cdot 10^{-12}$	19	$4.6 \cdot 10^{-16}$	$4.4 \cdot 10^{-16}$
512	22	$2.2 \cdot 10^{-16}$	$1.4 \cdot 10^{-12}$	19	$4.7 \cdot 10^{-16}$	$4.4 \cdot 10^{-16}$
1024	23	$2.4 \cdot 10^{-16}$	$2.1 \cdot 10^{-11}$	19	$5.3 \cdot 10^{-16}$	$5.6 \cdot 10^{-16}$
2048	24	$3.0 \cdot 10^{-16}$	$6.1 \cdot 10^{-11}$	19	$5.0 \cdot 10^{-16}$	$4.4 \cdot 10^{-16}$
4096	25	$2.6 \cdot 10^{-16}$	$5.4 \cdot 10^{-11}$	19	$5.0 \cdot 10^{-17}$	$6.7 \cdot 10^{-16}$
8192	26	$3.1 \cdot 10^{-16}$	$3.9 \cdot 10^{-10}$	19	$4.8 \cdot 10^{-16}$	$6.7 \cdot 10^{-16}$
16384	27	$2.2 \cdot 10^{-16}$	$2.0 \cdot 10^{-11}$	19	$5.9 \cdot 10^{-17}$	$6.7 \cdot 10^{-16}$
32768	29	$1.3 \cdot 10^{-16}$	$1.1 \cdot 10^{-9}$	19	$3.5 \cdot 10^{-16}$	$4.4 \cdot 10^{-16}$
65536	34	$2.2 \cdot 10^{-16}$	$5.5 \cdot 10^{-8}$	19	$5.0 \cdot 10^{-16}$	$6.7 \cdot 10^{-16}$

Table 4: Example 2: $r = 1/300$, $a = 18.244$, $\rho_d = 0.280$

M	$\bar{\theta}$	σ	$1/\sigma$	$\bar{\theta}/\sigma$
64	0.9999	1.0002	0.9998	0.9997
128	0.9999	1.0001	0.9999	0.9998
256	0.9999	1.000039	0.999961	0.9998
512	0.9999	1.000019	0.999981	0.9998
1024	0.9999	1.000009	0.999991	0.9998
2048	0.9999	1.000004	0.999996	0.9998
4096	0.9999	1.000002	0.999998	0.9999
8192	0.9999	1.000001	0.999999	0.9999
16384	0.9999	1.0000004	0.9999996	0.9999
32768	0.9999	1.0000001	0.9999999	0.9999
65536	0.9999	1.000000002	0.999999998	0.9999

Table 5: Example 2: $r = 1/300$, $a = 18.244$, $\rho_d = 0.280$

M	$1 - \rho(A_1 + A_2G + A_2)$	$1 - \rho(A_1 + A_2G + \theta A_2)$	$\sigma_{min}(W)$	$\sigma_{min}(Q)$
64	$1.4 \cdot 10^{-4}$	$2.2 \cdot 10^{-4}$	$2.7 \cdot 10^{-5}$	$3.8 \cdot 10^{-5}$
8192	$7.1 \cdot 10^{-7}$	$1.1 \cdot 10^{-4}$	$1.5 \cdot 10^{-7}$	$1.5 \cdot 10^{-5}$
65536	$1.6 \cdot 10^{-9}$	$1.0 \cdot 10^{-4}$	$3.5 \cdot 10^{-10}$	$1.5 \cdot 10^{-5}$

Table 6: Example 2: $r = 1/300$, $a = 18.244$, $\rho_d = 0.280$

	Cyclic Reduction			Shifted Cyclic Reduction		
ρ_d	Iter.	Res.	Stoc.	Iter.	Res.	Stoc.
0.01	6	$2.2 \cdot 10^{-16}$	$4.4 \cdot 10^{-16}$	6	$3.9 \cdot 10^{-16}$	$2.2 \cdot 10^{-16}$
0.025	7	$3.3 \cdot 10^{-16}$	$6.6 \cdot 10^{-16}$	7	$5.7 \cdot 10^{-16}$	$4.4 \cdot 10^{-16}$
0.05	10	$1.5 \cdot 10^{-16}$	$8.9 \cdot 10^{-16}$	10	$3.0 \cdot 10^{-16}$	$4.4 \cdot 10^{-16}$
0.075	12	$2.2 \cdot 10^{-16}$	$9.0 \cdot 10^{-15}$	12	$6.1 \cdot 10^{-16}$	$4.4 \cdot 10^{-16}$
0.1	14	$2.2 \cdot 10^{-16}$	$2.2 \cdot 10^{-15}$	14	$4.2 \cdot 10^{-16}$	$4.4 \cdot 10^{-16}$
0.12	14	$2.2 \cdot 10^{-16}$	$2.8 \cdot 10^{-14}$	14	$3.9 \cdot 10^{-16}$	$4.4 \cdot 10^{-16}$
0.14	15	$2.2 \cdot 10^{-16}$	$2.1 \cdot 10^{-13}$	15	$5.1 \cdot 10^{-17}$	$4.4 \cdot 10^{-16}$
0.16	16	$2.2 \cdot 10^{-16}$	$1.5 \cdot 10^{-14}$	16	$5.3 \cdot 10^{-16}$	$4.4 \cdot 10^{-16}$
0.18	16	$2.2 \cdot 10^{-16}$	$5.9 \cdot 10^{-14}$	16	$4.7 \cdot 10^{-17}$	$3.3 \cdot 10^{-16}$
0.2	17	$2.2 \cdot 10^{-16}$	$1.0 \cdot 10^{-14}$	16	$4.5 \cdot 10^{-16}$	$6.7 \cdot 10^{-16}$
0.22	18	$2.2 \cdot 10^{-16}$	$4.3 \cdot 10^{-13}$	17	$6.0 \cdot 10^{-16}$	$3.3 \cdot 10^{-16}$
0.24	18	$2.3 \cdot 10^{-16}$	$7.8 \cdot 10^{-13}$	17	$3.7 \cdot 10^{-16}$	$6.7 \cdot 10^{-16}$
0.26	19	$1.9 \cdot 10^{-16}$	$8.6 \cdot 10^{-13}$	17	$5.0 \cdot 10^{-16}$	$4.4 \cdot 10^{-16}$
0.28	20	$2.3 \cdot 10^{-16}$	$3.5 \cdot 10^{-12}$	17	$5.1 \cdot 10^{-16}$	$3.3 \cdot 10^{-16}$
0.29	22	$2.6 \cdot 10^{-16}$	$1.8 \cdot 10^{-11}$	17	$3.7 \cdot 10^{-16}$	$2.2 \cdot 10^{-16}$
.29568	31	$2.6 \cdot 10^{-16}$	$1.9 \cdot 10^{-8}$	17	$6.0 \cdot 10^{-16}$	$4.4 \cdot 10^{-16}$

Table 7: Example 2: $r = 1/100$, $a = 18.244$, $M = 512$

ρ_d	$\bar{\theta}$	σ	$1/\sigma$	$\bar{\theta}/\sigma$
0.01	0.9998	4.3182	0.2316	0.2315
0.025	0.9998	1.7714	0.5645	0.5644
0.05	0.9998	1.0814	0.9248	0.9246
0.075	0.9998	1.0170	0.9833	0.9831
0.1	0.9998	1.0063	0.9938	0.9936
0.12	0.9998	1.0034	0.9966	0.9964
0.14	0.9997	1.0021	0.9979	0.9977
0.16	0.9997	1.0013	0.9987	0.9984
0.18	0.9997	1.0009	0.9991	0.9988
0.2	0.9997	1.0006	0.9994	0.9991
0.22	0.9997	1.0004	0.9996	0.9993
0.24	0.9996	1.0002	0.9998	0.9994
0.26	0.9996	1.0001	0.9999	0.9995
0.28	0.9996	1.0001	0.9999	0.9995
0.29	0.9996	1.000019	0.99998	0.9995
0.29568	0.9995	1.00000003	0.99999997	0.9995

Table 8: Example 2: $r = 1/100$, $a = 18.244$, $M = 512$

p	$1-\rho(A_1+A_2G+A_2)$	$1-\rho(A_1+A_2G+\theta A_2)$	$\sigma_{min}(W)$	$\sigma_{min}(Q)$
0.01	0.6265	0.6265	0.6259	0.6175
0.28	$4.0\cdot 10^{-5}$	$3.4\cdot 10^{-4}$	$8.7\cdot 10^{-6}$	$5.3\cdot 10^{-5}$
0.29568	$1.9\cdot 10^{-8}$	$3.2\cdot 10^{-4}$	$3.6\cdot 10^{-9}$	$4.4\cdot 10^{-5}$

Table 9: Example 2: $r=1/100$, $a=18.244$, $M=512$

	Cyclic Reduction			Shifted Cyclic Reduction		
δ	Iter.	Res.	Stoc.	Iter.	Res.	Stoc.
10^{-1}	8	$5.7\cdot 10^{-16}$	$1.3\cdot 10^{-15}$	5	$2.8\cdot 10^{-16}$	$3.3\cdot 10^{-16}$
10^{-2}	11	$6.5\cdot 10^{-16}$	$7.2\cdot 10^{-15}$	4	$3.7\cdot 10^{-16}$	0.0
10^{-3}	14	$6.7\cdot 10^{-16}$	$9.7\cdot 10^{-14}$	4	$1.9\cdot 10^{-16}$	$2.2\cdot 10^{-16}$
10^{-4}	17	$6.4\cdot 10^{-16}$	$8.8\cdot 10^{-13}$	4	$4.5\cdot 10^{-16}$	$2.2\cdot 10^{-16}$
10^{-5}	21	$1.1\cdot 10^{-15}$	$1.2\cdot 10^{-11}$	4	$3.5\cdot 10^{-16}$	$4.4\cdot 10^{-16}$
10^{-6}	24	$6.5\cdot 10^{-16}$	$5.8\cdot 10^{-11}$	4	$2.5\cdot 10^{-16}$	$8.9\cdot 10^{-16}$
10^{-7}	27	$6.3\cdot 10^{-16}$	$1.4\cdot 10^{-9}$	4	$2.5\cdot 10^{-16}$	$6.7\cdot 10^{-16}$
10^{-8}	29	$7.0\cdot 10^{-16}$	$3.5\cdot 10^{-9}$	4	$2.5\cdot 10^{-16}$	$4.4\cdot 10^{-16}$

Table 10: Example 3: $k=16$

δ	$\bar{\theta}$	σ	$1/\sigma$	$\bar{\theta}/\sigma$
10^{-1}	0.0783	1.3333	0.75	0.0587
10^{-2}	0.0174	1.0303	0.9706	0.0140
10^{-3}	0.0207	1.0030	0.9970	0.0207
10^{-4}	0.0216	1.0003	0.9997	0.0216
10^{-5}	0.0217	1.00003	0.99997	0.0217

Table 11: Example 3: $k=16$

δ	$1-\rho(A_1+A_2G+A_2)$	$1-\rho(A_1+A_2G+\bar{\theta}A_2)$	$\sigma_{min}(W)$	$\sigma_{min}(Q)$
10^{-1}	0.1	0.3765	0.1	0.3234
10^{-5}	$1.0\cdot 10^{-5}$	0.3261	$1.0\cdot 10^{-5}$	0.2310
10^{-8}	$1.0\cdot 10^{-8}$	0.3261	$1.0\cdot 10^{-8}$	0.2721

Table 12: Example 3: $k=16$

	Cyclic Reduction			Shifted Cyclic Reduction		
δ	Iter.	Res.	Stoc.	Iter.	Res.	Stoc.
10^{-1}	8	$1.3 \cdot 10^{-15}$	$1.1 \cdot 10^{-15}$	5	$1.2 \cdot 10^{-15}$	$8.9 \cdot 10^{-16}$
10^{-2}	11	$9.5 \cdot 10^{-16}$	$1.1 \cdot 10^{-14}$	4	$4.1 \cdot 10^{-16}$	$7.8 \cdot 10^{-16}$
10^{-3}	14	$1.4 \cdot 10^{-15}$	$8.0 \cdot 10^{-14}$	4	$5.4 \cdot 10^{-16}$	$1.1 \cdot 10^{-15}$
10^{-4}	17	$1.6 \cdot 10^{-15}$	$3.7 \cdot 10^{-12}$	4	$6.8 \cdot 10^{-16}$	$1.8 \cdot 10^{-15}$
10^{-5}	21	$1.1 \cdot 10^{-15}$	$1.0 \cdot 10^{-11}$	4	$5.1 \cdot 10^{-16}$	$6.6 \cdot 10^{-16}$
10^{-6}	24	$1.2 \cdot 10^{-15}$	$2.1 \cdot 10^{-10}$	4	$5.5 \cdot 10^{-16}$	$8.9 \cdot 10^{-16}$
10^{-7}	27	$1.1 \cdot 10^{-15}$	$7.9 \cdot 10^{-10}$	4	$6.5 \cdot 10^{-16}$	$1.0 \cdot 10^{-15}$
10^{-8}	29	$8.2 \cdot 10^{-16}$	$1.4 \cdot 10^{-8}$	4	$7.1 \cdot 10^{-16}$	$7.8 \cdot 10^{-16}$

Table 13: Example 3: $k = 32$

	Cyclic Reduction			Shifted Cyclic Reduction		
δ	Iter.	Res.	Stoc.	Iter.	Res.	Stoc.
10^{-1}	8	$2.8 \cdot 10^{-15}$	$4.9 \cdot 10^{-15}$	5	$2.0 \cdot 10^{-15}$	$1.8 \cdot 10^{-15}$
10^{-2}	11	$3.1 \cdot 10^{-15}$	$1.1 \cdot 10^{-14}$	4	$1.4 \cdot 10^{-15}$	$2.7 \cdot 10^{-15}$
10^{-3}	14	$3.1 \cdot 10^{-15}$	$1.8 \cdot 10^{-13}$	4	$1.3 \cdot 10^{-15}$	$1.1 \cdot 10^{-15}$
10^{-4}	17	$2.4 \cdot 10^{-15}$	$4.6 \cdot 10^{-12}$	4	$1.2 \cdot 10^{-15}$	$2.9 \cdot 10^{-15}$
10^{-5}	21	$2.3 \cdot 10^{-15}$	$2.9 \cdot 10^{-11}$	4	$1.6 \cdot 10^{-15}$	$4.4 \cdot 10^{-16}$
10^{-6}	24	$3.3 \cdot 10^{-15}$	$2.8 \cdot 10^{-10}$	4	$1.4 \cdot 10^{-15}$	$2.4 \cdot 10^{-15}$
10^{-7}	27	$2.4 \cdot 10^{-15}$	$1.9 \cdot 10^{-10}$	4	$1.3 \cdot 10^{-15}$	$3.8 \cdot 10^{-15}$
10^{-8}	29	$2.6 \cdot 10^{-15}$	$1.7 \cdot 10^{-8}$	4	$1.2 \cdot 10^{-15}$	$3.3 \cdot 10^{-15}$

Table 14: Example 3: $k = 64$

References

[1] D. A. Bini, L. Gemignani, and B. Meini. Factorization of analytic functions by means of Koenig's theorem and Toeplitz computations. *Numer. Math.*, 2000. To appear.

[2] D. A. Bini and B. Meini. On cyclic reduction applied to a class of Toeplitz-like matrices arising in queueing problems. In W. J. Stewart, editor, *Computations with Markov Chains*, pages 21–38. Kluwer Academic Publisher, 1995.

[3] D. 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.

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

[5] D. A. Bini and B. Meini. Inverting block Toeplitz matrices in block Hessenberg form by means of displacement operators: application to queueing problems. *Linear Algebra Appl.*, 272:1–16, 1998.

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

[7] H. R. Gail, S. L. Hantler, and B. A. Taylor. Spectral analysis of $M/G/1$ and $G/M/1$ type Markov chains. *Adv. in Appl. Probab.*, 28:114–165, 1996.

[8] H. R. Gail, S. L. Hantler, and B. A. Taylor. Use of characteristics roots for solving infinite state Markov chains. In W. K. Grassmann, editor, *Computational Probability*, pages 205–255. Kluwer Academic Publishers, 2000.

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

[10] 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.

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

[12] 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.

[13] B. Meini. Solving QBD problems: the cyclic reduction algorithm versus the invariant subspace method. *Adv. Perf. Anal.*, 1:215–225, 1998.

[14] V. Naoumov. Matrix-multiplicative approach to quasi-birth-and-death processes analysis. In A. S. Alfa and S. Chakravarthy, editors, *Matrix-analytic methods in stochastic models (Flint, MI)*, pages 85–106. Dekker, New York, 1997.

[15] V. Naoumov, U. R. Krieger, and D. Wagner. Analysis of a multiserver delay-loss system with a general Markovian arrival process. In A. S. Alfa and S. Chakravarthy, editors, *Matrix-analytic methods in stochastic models (Flint, MI)*, pages 43–66. Dekker, New York, 1997.

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