The cyclic reduction algorithm

Article *in* Numerical Algorithms · May 2009

DOI: 10.1007/s11075-008-9253-0 · Source: DBLP

2 authors:

Dario A. Bini
Università di Pisa

Beatrice Meini
Università di Pisa

Numer Algor
DOI 10.1007/s11075-008-9253-0

ORIGINAL PAPER

The cyclic reduction algorithm: from Poisson equation to stochastic processes and beyond

In memoriam of Gene H. Golub

Dario A. Bini · Beatrice Meini

Received: 11 July 2008 / Accepted: 3 November 2008

Abstract Cyclic reduction is an algorithm invented by G. H. Golub and R. W. Hockney in the mid 1960s for solving linear systems related to the finite differences discretization of the Poisson equation over a rectangle. Among the algorithms of Gene Golub, it is one of the most versatile and powerful ever created. Recently, it has been applied to solve different problems from different applicative areas. In this paper we survey the main features of cyclic reduction, relate it to properties of analytic functions, recall its extension to solving more general finite and infinite linear systems, and different kinds of nonlinear matrix equations, including algebraic Riccati equations, with applications to Markov chains, queueing models and transport theory. Some new results concerning the convergence properties of cyclic reduction and its applicability are proved under very weak assumptions. New formulae for overcoming breakdown are provided.

Keywords Cyclic reduction · Toeplitz systems · Hessenberg systems · Markov chains · Matrix equations

1 Introduction

In the mid 1960s, G. H. Golub and R. W. Hockney designed an efficient algorithm for solving a block tridiagonal linear system of the kind $T\boldsymbol{u} = \boldsymbol{b}$

D. A. Bini (✉) · B. Meini
Dipartimento di Matematica, Università di Pisa, Pisa, Italy
e-mail: bini@dm.unipi.it

B. Meini
e-mail: meini@dm.unipi.it

Published online: 21 November 2008

originated from the discretization of the Poisson equation over a rectangle. In this system, T is an $n \times n$ block tridiagonal matrix with $m \times m$ blocks, the blocks are constant along the diagonals, i.e., the matrix is block Toeplitz, the diagonal blocks are tridiagonal Toeplitz matrices having diagonal entries $(-1, 4, -1)$ and the remaining nonzero blocks coincide with the identity matrix.

The algorithm appeared the first time in 1965 in the paper by Hockney [58] where the author writes "*These equations form a tridiagonal system with periodic boundary conditions and a particularly efficient method of solution has been devised in collaboration with Dr. G. Golub. This involves the recursive application of the process of cyclic reduction which follows.*" The algorithm, called Cyclic Reduction (CR), was described and analyzed with more details some years later in the papers by Buzbee, Golub and Nielson [35], and by Buneman [34], who provided a more stable version. Since then it received much attention for its very nice computational features and had a great development.

In particular, the algorithm is nicely described in terms of Schur complements in the paper by Gander and Golub [43] where a historical survey is also given. The underlying idea on which CR is based is to re-order the block rows and columns of the matrix T by means of an odd-even permutation and to eliminate the odd indexed unknowns by means of a Schur complementation. The nice feature is that the new system of half the size obtained in this way has the same block tridiagonal block Toeplitz structure as the original system so that the same procedure can be cyclically repeated until one arrives at a single $m \times m$ system.

A surprising feature which makes CR very versatile is that it can be viewed both as a direct method for solving a linear system and as an iterative technique which makes it a reliable approximate solver. In fact, under suitable conditions, some matrices computed by CR have the property to converge quadratically to zero.

The wide literature generated in the 1970's and 1980's provides a lot of variations, improvements and extensions. Here we recall some of the most important contributions without pretending to provide a complete list of references. Many contributions have been given by Swarztrauber [86–89], and by Sweet [92, 93], with implementation [90] and analysis in the parallel model of computation [91, 94]. It is worth citing also the work of Diamond and Ferreira [37], Heller [54], and Reichel [77], the analysis of CR implemented with an interval arithmetic by Schwandt [83, 84], the comparisons of Temperton [95, 96], the use of CR as preconditioner by Rodrigue and Wolitzer [78], and the analysis of Rosmond and Faulkner [80].

The interest for the implementation of CR on parallel architectures, present in the works of Hockney and Jesshope [59], and Gallopoulos and Saad [41, 42], has had a great expansion in the 1990's with the contributions of many authors among whom Amodio, Briggs, Brugnano, Ho, Johnsson, Mastronardi, Politi, Rossi, Toivanen and Turnbull [1–3, 33, 57, 82].

Convergence properties initially investigated by Hockney [58] and proved by Heller [54], have been further analyzed by Bondeli and Gander [32], and by Yalamov and Pavlov [99]. Numerical stability of CR has been analyzed by Amodio and Mazzia [4] and by Yalamov and Pavlov [74, 99, 100]. Extensions and applications to other problems can be found in the papers [5, 7, 31, 38, 81, 82], and [85], by Amodio, Bialecki, Boisvert, Dodson, Levin, Paprzycki, Rossi, Sun, and Toivanen.

It is not surprising that CR has been rediscovered several times. In the framework of stochastic processes, Ye and Li in [101, 102] have provided a version of CR called *folding algorithm* for the solution of certain block tridiagonal finite systems whose solution is the probability invariant vector of a Markov chain. In the same framework Latouche and Ramaswami [64] have designed the *Logarithmic Reduction* algorithm for solving certain quadratic matrix equations related to Quasi–Birth-Death processes [65] which relies on the same idea of CR.

A substantial advance in the understanding of CR has been given in the 1990s in a series of papers by Bini and Meini [8, 19, 20, 63] with the contributions of Gemignani [10–13], Latouche [17, 18], Ramaswami [26, 27] and Spitkovsky [28]. The fundamental idea in these papers is that CR can be described in functional form by means of a sequence of analytic functions, and its convergence properties can be easily deduced from standard results concerning the theory of analytic functions of complex variable. In fact, it is proved that CR is the matrix counterpart of the Graeffe-Lobachevski-Dandelin iteration [72, 73] extended to matrix polynomials and to suitable matrix power series. For a polynomial $p(x)$, this iteration generates the sequence $\{p_k(x)\}_k$ where $p_{k+1}(x^2) = p_k(x)p_k(-x)$ and $p_0(x) = p(x)$, so that the roots of $p_{k+1}(x)$ are the square of the roots of $p_k(x)$.

The analysis of CR performed in terms of analytic functions was originated by the study of Markov chains of M/G/1 and G/M/1-type which model a large part of queueing problems from engineering and telecommunications [65, 70, 71].

This advance enabled the authors to extend and generalize CR to wider classes of problems like finite and infinite linear systems with the block Hessenberg structure [8, 19, 20], and banded Toeplitz linear systems [21].

By relying on this new functional formulation, some nonlinear problems of apparently different nature were solved through the application of CR starting from the mid 90's. Here, the main idea was to linearize the problems by transforming them into special infinite linear systems to be solved by means of CR [69]. To this regard it is worth citing the results of Bini, Böttcher, Fiorentino, Gemignani, Meini and Spitkovsky [9–12, 28] on the factorization of (matrix) polynomials and (matrix) power series, and the application of CR to the solution of quadratic matrix equations and nonlinear matrix equations defined by a matrix power series developed by Bini, Latouche, Meini, Ramaswami in [8, 18, 20, 26]. The role of CR in solving infinite systems and matrix equations encountered in queueing models is pointed out in the book [18].

More recently, CR has been applied by Guo, Higham, Lancaster and Tisseur in [49, 51], for solving the quadratic hyperbolic eigenvalue problem, and by Bini, Guo, Higham, Iannazzo, Latouche, Meini, Poloni, Ramaswami [15, 16, 24, 47, 48, 50, 76] for solving algebraic Riccati equations.

Concerning applications, CR has become the method of choice in the solution of queuing problems from engineering and telecommunications where some effective implementations in Matlab and Fortran 95 have been designed by Bini, Meini, Steffé and Van Houdt [29, 30] and are publically available.

Acceleration techniques, introduced by He, Meini and Rhee [53] and generalized and used in [13, 18, 46, 50], allow one to speed-up the convergence of CR and to keep the quadratic convergence even in the critical cases, like in the null recurrent stochastic processes, where the customary methods have linear or even sublinear convergence as shown by Guo in [45].

Nowadays, CR has become a powerful and versatile algorithm currently used for the efficient solution of diverse linear and nonlinear problems from different areas and applications. From the theoretical point of view, CR is a rich and variegated concept where tools from linear algebra, like the Schur complement, tools from complex analysis like the Cauchy integral theorem and the properties of analytic functions, tools from operator theory like the Wiener-Hopf factorization, and tools from polynomial computations like the Graeffe iteration, play important roles. The combination of these tools generate a synergetic action which enables one to prove rich and unexpected properties of CR.

In this paper we wish to illustrate this important idea of Gene Golub in all its theoretical and algorithmic facets by pointing out its richness and its relationships with the different areas of Mathematics. Indeed, since the subjects involved are so many, we cannot provide the detailed description of all the topics related to cyclic reduction, but we give pointers to the current literature where the reader can find more insights, details, and technical information.

Besides an overview on classical and modern cyclic reduction we present some new advances which concern convergence properties and breakdown conditions. More specifically, in Sections 3.2 and 4.2 we prove new convergence theorems for CR where the assumptions needed to prove convergence are much weakened with respect to the existing literature.

Concerning the cases of breakdown, we provide in Section 3.3 new different formulae for performing the CR step which avoid the matrix inversion in the case of singularity and ill conditioning. These formulae express the iteration in terms of the block entries in the four corners of the inverse of a suitable block tridiagonal (nonsingular) matrix. These new equations enable one to numerically optimize the implementation of CR by constructing a subsequence of the matrix sequence generated by CR which avoid the inversion of ill-conditioned matrices. An example is also given where it is shown that singularity and therefore breakdown, can be encountered at any step of CR and if the singularity is moved "to the infinity" then the conditions of convergence of CR do not hold anymore.

The paper is organized as follows. In Section 2 we recall the original formulation presented in [35] by Buzbee, Golub, Nielson and in [58] by Hockney for solving the Poisson equation over a rectangle, and the formulation given in [43] by Gander and Golub in terms of Schur complements. In Section 2.3 we discuss some computational issues of CR together with the roles of the even-odd and the odd-even permutations. The classical convergence properties are recalled in Section 2.4.

Section 3 is devoted to the description of more recent properties of CR including its functional formulation and its relationship with Graeffe iteration (Section 3.1). In particular, in Sections 3.2 and 4.2 we revisit convergence properties and give new results which hold under very weak assumptions. In Section 3.3 we prove the new formulae for performing the CR step by avoiding breakdown.

Extention of CR to block Hessenberg block Toeplitz matrices is reported in Section 4 both for finite and for infinite systems, together with the evaluation/interpolation technique for implementing CR. Convergence properties for this case are proved in Section 4.2.

Section 5 concerns applications of CR. Block banded Toeplitz systems, quadratic matrix equations, polynomial and power series matrix equations, matrix square root and algebraic Riccati equations are the main subjects of this section. Section 5.7 describes an effective technique which provides a substantial acceleration of CR which is fundamental in critical cases encountered in the applications.

The paper is closed by Section 6 with conclusions.

1.1 Notation

Throughout the paper we denote by $\mathrm{Trid}_n(B, A, C)$, the block tridiagonal block Toeplitz matrix with block size n having A on the main diagonal, B on the subdiagonal and C on the superdiagonal, where A, B, C are $m \times m$ matrices, that is,

$$\mathrm{Trid}_n(B, A, C) = \begin{bmatrix} A & C & & 0 \\ B & A & \ddots & \\ & \ddots & \ddots & C \\ 0 & & B & A \end{bmatrix}.$$

If the size is not specified, we denote the latter matrix by $\mathrm{Trid}(B, A, C)$. We denote by $\mathcal{A}(r, R) = \{z \in \mathbb{C} : \ r < |z| < R\}$ the open annulus made up by the complex numbers having modulus between r and R where $0 \le r < R$. The cost of an algorithm is defined as the number of arithmetic operations sufficient to carry out the algorithm.

2 Original formulation of CR

In this section we present the original formulation of CR in the context of solving the discrete Poisson equation over a rectangle, together with its formulation in terms of Schur complements and its main computational features.

2.1 CR and the Poisson equation

We report the algorithm of cyclic reduction as described by Buzbee, Golub and Nielson in [35] and by Hockney in [58] for solving the Poisson equation over a rectangle $\Omega \subset \mathbb{R}^2$

$$\begin{aligned} \frac{\partial^2 u(x,y)}{\partial x^2} + \frac{\partial^2 u(x,y)}{\partial x^2} &= f(x,y), (x,y) \in \Omega \\ u(x,y) &= g(x,y), (x,y) \in \partial\Omega \end{aligned} \tag{1}$$

where $u(x, y)$ is the unknown real valued function defined on Ω, $f(x, y)$ is a given function defined on Ω and $g(x, y)$ is a given real valued function defined on the boundary $\partial\Omega$ of Ω.

The finite-differences discretization of (1) where Ω is discretized by means of a grid of $m \times n$ interior equispaced points, leads to an $n \times n$ block tridiagonal system with $m \times m$ blocks

$$\mathrm{Trid}_n(-I, A, -I) \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_n \end{bmatrix}, \quad A = \mathrm{Trid}_m(-1, 4, -1), \tag{2}$$

where $\boldsymbol{u}_i, \boldsymbol{b}_i \in \mathbb{R}^m$, I is the $m \times m$ identity matrix.

Assume $n = 2^q - 1$, for q positive integer, and consider three consecutive block equations for $1 < j < n$, where we assume $\boldsymbol{u}_0 = \boldsymbol{u}_{2^q} = 0$, and j even:

$$\begin{aligned} -\boldsymbol{u}_{j-2} + A\boldsymbol{u}_{j-1} - \boldsymbol{u}_j &= \boldsymbol{b}_{j-1} \\ -\boldsymbol{u}_{j-1} + A\boldsymbol{u}_j - \boldsymbol{u}_{j+1} &= \boldsymbol{b}_j \\ -\boldsymbol{u}_j + A\boldsymbol{u}_{j+1} - \boldsymbol{u}_{j+2} &= \boldsymbol{b}_{j+1} \end{aligned} \tag{3}$$

for $j = 2, 4, 6 \ldots, 2^q - 2$. Multiplying the second equation by A and summing up the three equations yields the system involving the even numbered unknowns:

$$\begin{aligned} -\boldsymbol{u}_{j-2} + A^{(1)}\boldsymbol{u}_j - \boldsymbol{u}_{j+2} &= \boldsymbol{b}^{(1)}_{j/2} \\ \boldsymbol{b}^{(1)}_{j/2} &= \boldsymbol{b}_{j-1} + A\boldsymbol{b}_j + \boldsymbol{b}_{j+1} \end{aligned} \tag{4}$$

where $j = 2, 4, 6, \ldots, 2^q - 2$ and $A^{(1)} = A^2 - 2I$. Once the even numbered unknowns have been computed, the remaining unknowns can be obtained from the original system (2) by solving 2^{q-1} systems with the same matrix A, i.e.,

$$A\boldsymbol{u}_{2j-1} = \boldsymbol{b}_{2j-1} + \boldsymbol{u}_{2j-2} + \boldsymbol{u}_{2j}, \quad j = 1, 2, \ldots, 2^{q-1}. \tag{5}$$

Since the system (4) has the same structure as (2), the same process can be cyclically repeated until a single $m \times m$ system is obtained. The solution of the latter system and back substitution, by means of (5), complete the procedure.

More specifically, CR generates a matrix sequence and a vector sequence defined by

$$
\begin{aligned}
&A^{(k)} = (A^{(k-1)})^2 - 2I, \quad A^{(0)} = A, \\
&\boldsymbol{b}_j^{(k)} = \boldsymbol{b}_{2j-1}^{(k-1)} + A^{(k-1)}\boldsymbol{b}_{2j}^{(k-1)} + \boldsymbol{b}_{2j+1}^{(k-1)}, \quad j = 1, \ldots, 2^{q-k} - 1
\end{aligned}
\tag{6}
$$

for $k = 1, \ldots, q-1$, with $A^{(0)} = A$, $\boldsymbol{b}_i^{(0)} = \boldsymbol{b}_i$, $i = 1, \ldots, n$.

In the back substitution stage, 2^{q-k-1} systems with matrices $A^{(k)}$ must be solved for $k = 0, 1, \ldots, q-1$, i.e.,

$$
A^{(k)}\boldsymbol{u}_{(2i-1)2^k} = \boldsymbol{b}_{2i-1}^{(k)} + \boldsymbol{u}_{(i-1)2^{k+1}} + \boldsymbol{u}_{i2^{k+1}}, \quad i = 1, 2, \ldots, 2^{q-k-1}. \tag{7}
$$

Due to the recurrence relation (6), the block $A^{(k)}$ can be expressed in terms of a Chebyshev polynomial $P_{2^k}(x)$, of the first kind, of degree 2^k as $A^{(k)} = P_{2^k}(A)$. This property holds since Chebyshev polynomials of the first kind [44] satisfy the same formal equation $P_{2^k}(x) = P_{2^{k-1}}(x)^2 - 2$, $P_1(x) = x$, as the matrices $A^{(k)}$ in (6).

In this way, denoting by $\theta_i^{(2^k)}$, $i = 1, \ldots, 2^k$, the zeros of the Chebyshev polynomial $P_{2^k}(x)$ so that $P_{2^k}(x) = \prod_{i=1}^{2^k}\left(x - \theta_i^{(2^k)}\right)$, one can write $A^{(k)}$ as the product of 2^k tridiagonal matrices $A^{(k)} = \prod_{i=1}^{2^k}\left(A - \theta_i^{(2^k)} I\right)$. Therefore, each one of the 2^{q-k-1} systems in (7) is reduced to solving 2^k tridiagonal systems of size m, for the cost of $O(m2^{q-1})$. Thus, the cost of back substitution is just $O(nm\log_2 n)$ which asymptotically coincides with the overall computational cost of CR.

In this formulation, CR suffers of numerical instability encountered in the computation of the right-hand side $\boldsymbol{b}_j^{(k)}$. This drawback was overcome by Buneman [34] who provided a stable version for updating the right-hand side at each recursive step of CR.

2.2 Formulation in terms of Schur complements

Gander and Golub [43] have provided a nice formulation of CR given in terms of Schur complements. This formulation slightly differs from the original one given in [58] in the way the elimination of the odd numbered components is performed, moreover it can be applied to the more general block tridiagonal block Toeplitz matrix $\mathrm{Trid}(B, A, C)$.

Assume $n = 2^q - 1$ and consider the more general system

$$\begin{bmatrix} A & C & & 0 \\ B & A & \ddots & \\ & \ddots & \ddots & C \\ 0 & & B & A \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_n \end{bmatrix} \tag{8}$$

for $A, B, C \in \mathbb{R}^{m\times m}$. Apply an odd-even permutation to both block-columns and block-rows in (8) and get

$$\left[\begin{array}{ccc|ccc} A & & 0 & C & & 0 \\ & \ddots & & B & \ddots & \\ & & \ddots & & \ddots & C \\ 0 & & A & 0 & & B \\ \hline B & C & 0 & A & & 0 \\ & \ddots & \ddots & & \ddots & \\ 0 & B & C & 0 & & A \end{array}\right] \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_3 \\ \vdots \\ \boldsymbol{u}_{2\cdot 2^{q-1}-1} \\ \hline \boldsymbol{u}_2 \\ \boldsymbol{u}_4 \\ \vdots \\ \boldsymbol{u}_{2\cdot 2^{q-1}-2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_3 \\ \vdots \\ \boldsymbol{b}_{2\cdot 2^{q-1}-1} \\ \hline \boldsymbol{b}_2 \\ \boldsymbol{b}_4 \\ \vdots \\ \boldsymbol{b}_{2\cdot 2^{q-1}-2} \end{bmatrix}.$$

Now, rewrite the above system as

$$\begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_{\text{odd}} \\ \boldsymbol{u}_{\text{even}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_{\text{odd}} \\ \boldsymbol{b}_{\text{even}} \end{bmatrix},$$

assume A nonsingular, eliminate the odd block components by means of block Gaussian elimination, i.e., compute the Schur complement of $H_{1,1}$ and obtain the smaller system of block size $2^{q-1} - 1$:

$$(H_{22} - H_{21}H_{11}^{-1}H_{12})\boldsymbol{u}_{\text{even}} = \boldsymbol{b}^{(1)}, \quad \boldsymbol{b}^{(1)} = \boldsymbol{b}_{\text{even}} - H_{21}H_{11}^{-1}\boldsymbol{b}_{\text{odd}}.$$

Magically, the Schur complement has the same structure as the original matrix and the above system takes the form

$$\begin{bmatrix} A^{(1)} & C^{(1)} & & 0 \\ B^{(1)} & A^{(1)} & \ddots & \\ & \ddots & \ddots & C^{(1)} \\ 0 & & B^{(1)} & A^{(1)} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_2 \\ \boldsymbol{u}_4 \\ \vdots \\ \boldsymbol{u}_{2^q-2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1^{(1)} \\ \boldsymbol{b}_2^{(1)} \\ \vdots \\ \boldsymbol{b}_{2^{q-1}-1}^{(1)} \end{bmatrix},$$

where

$$\boldsymbol{b}_i^{(1)} = \boldsymbol{b}_{2i} - BA^{-1}\boldsymbol{b}_{2i-1} - CA^{-1}\boldsymbol{b}_{2i+1}, \quad i = 1, \ldots, 2^{q-1} - 1$$

and

$$\begin{aligned} A^{(1)} &= A - BA^{-1}C - CA^{-1}B \\ B^{(1)} &= -BA^{-1}B \\ C^{(1)} &= -CA^{-1}C \end{aligned} \tag{9}$$

while for the odd indexed block components one has

$$A\boldsymbol{u}_{2i-1} = \boldsymbol{b}_{2i-1} - B\boldsymbol{u}_{2i-2} - C\boldsymbol{u}_{2i+2}, \quad i = 1, 2, \ldots, 2^{q-1},$$

where we set $\boldsymbol{u}_0 = \boldsymbol{u}_{n+1} = 0$.

This process, can be cyclically repeated and generates the sequence of systems of block size $2^{q-k} - 1$:

$$\begin{bmatrix} A^{(k)} & C^{(k)} & & 0 \\ B^{(k)} & A^{(k)} & \ddots & \\ & \ddots & \ddots & C^{(k)} \\ 0 & & B^{(k)} & A^{(k)} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_{1\cdot 2^k} \\ \boldsymbol{u}_{2\cdot 2^k} \\ \vdots \\ \boldsymbol{u}_{(2^{q-k}-1)\cdot 2^k} \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1^{(k)} \\ \boldsymbol{b}_2^{(k)} \\ \vdots \\ \boldsymbol{b}_{2^{q-k}-1}^{(k)} \end{bmatrix}, \tag{10}$$

for $k = 0, 1, \ldots, q-1$. The matrices $A^{(k)}$, $B^{(k)}$, $C^{(k)}$, and the vectors $\boldsymbol{b}_i^{(k)}$ are defined by means of the following recursions

$$\boldsymbol{b}_i^{(k+1)} = \boldsymbol{b}_{2i}^{(k)} - B^{(k)}(A^{(k)})^{-1}\boldsymbol{b}_{2i-1}^{(k)} - C^{(k)}(A^{(k)})^{-1}\boldsymbol{b}_{2i+1}^{(k)}, \quad i = 1, \ldots, 2^{q-k} - 1$$

$$\begin{aligned} A^{(k+1)} &= A^{(k)} - B^{(k)}(A^{(k)})^{-1}C^{(k)} - C^{(k)}(A^{(k)})^{-1}B^{(k)} \\ B^{(k+1)} &= -B^{(k)}(A^{(k)})^{-1}B^{(k)} \\ C^{(k+1)} &= -C^{(k)}(A^{(k)})^{-1}C^{(k)} \end{aligned} \tag{11}$$

for $k = 0, 1, \ldots, q-2$, and $A^{(0)} = A$, $B^{(0)} = B$, $C^{(0)} = C$, provided that $\det A^{(k)} \neq 0$ for any k.

At the end of the process one gets the $m \times m$ system $A^{(q-1)}\boldsymbol{u}_{2^{q-1}} = \boldsymbol{b}_1^{(q-1)}$, from which one recovers $\boldsymbol{u}_{2^{q-1}}$. Back substitution, performed by solving the systems

$$A^{(k)}\boldsymbol{u}_{(2i-1)2^k} = \boldsymbol{b}_{2i-1}^{(k)} - B^{(k)}\boldsymbol{u}_{(2i-2)2^k} - C^{(k)}\boldsymbol{u}_{(2i)2^k},$$

for $i = 1, 2, \ldots, 2^{q-k-1}$, $k = q-2, \ldots, 0$, starting with $\boldsymbol{u}_{2^{q-1}}$ allows one to compute the remaining unknowns.

Observe that, this version of CR is slightly different from that of Section 2.1. In fact, here the elimination stage is performed by means of Schur complementation, while in Section 2.1, since $B = C = -I$, the elimination is performed by computing the matrix $H_{22}^2 - H_{21}H_{12}$ in place of the Schur complement.

The applicability of CR depends on the nonsingularity of the blocks $A^{(k)}$, that is on the nonsingularity of suitable leading principal submatrices of $\mathrm{Trid}(B, A, C)$. If k_0 is the first integer for which $\det A^{(k_0)} = 0$, then the matrix sequences $A^{(k)}$, $B^{(k)}$, $C^{(k)}$ generated by CR are not computable for $k > k_0$ by means of (11). We refer to this situation as a *breakdown* of CR.

Breakdown of CR will be discussed in more detail in Section 3.3. Here we point out that breakdown is not encountered if, say, $\mathrm{Trid}(B, A, C)$ is strongly diagonally dominant or irreducibly diagonal dominant, like the matrix of (2), or Hermitian and positive definite, or a nonsingular M-matrix or an irreducible singular M-matrix.

2.3 Some computational issues

In [92, 93] Sweet extended CR to systems of any size n. The same author in [94] provided a version of CR for the Poisson equation where the intrinsic sequential nature of the back substitution stage based on the successive solution of the systems with matrices $\left(A - \theta_i^{(2^k)} I\right)$ is completely removed. The idea is nice and elegant: the system $A^{(k)}\boldsymbol{u} = \boldsymbol{b}$ is rewritten as $\boldsymbol{u} = (P_{2^k}(A))^{-1}\boldsymbol{b}$, and the rational function $P_{2^k}(x)^{-1}$ is expressed by means of its partial fraction expansion as $\sum_{i=1}^{2^k} c_i/(x - d_i)$ for suitable constants c_i and d_i. In this way one has $\boldsymbol{u} = \sum_{i=2}^{2^k} c_i(A - d_i I)^{-1}\boldsymbol{b}$. This amounts to solving 2^k systems in parallel and to summing up their solutions with weights c_i. A different approach is followed by Swarztrauber in [89] where the function $P_{2^k}(x)^{-1}$ is approximated as a short sum of terms of the kind $c_i/(x - d_i)$.

Observe also that, by applying (3) with j odd and performing a similar linear combinations of the three equations yields

$$
\begin{aligned}
(A^2 - I)\boldsymbol{u}_1 - \boldsymbol{u}_3 &= A\boldsymbol{b}_1 + \boldsymbol{b}_2 \\
-\boldsymbol{u}_{j-2} + A^{(1)}\boldsymbol{u}_j - \boldsymbol{u}_{j+2} &= \boldsymbol{c}^{(1)}_{(j+1)/2} \\
\boldsymbol{c}^{(1)}_{(j+1)/2} &= \boldsymbol{b}_{j-1} + A\boldsymbol{b}_j + \boldsymbol{b}_{j+1}
\end{aligned}
\tag{12}
$$

with $j = 3, 5, \ldots, 2^q - 1$. This enables one to arrive at the two disjoint block tridiagonal systems (4) and (12) involving separately the odd and the even numbered unknowns. These two systems are independent of each other and can be solved in parallel. This observation provides an effective parallel version of CR in the divide-and-conquer style [43]. This property holds also for the version based on Schur complements. Analysis of CR in parallel models of computation have been carried out by several authors including Amodio, Briggs, Brugnano, Gallopoulos, Ho, Hockney, Jesshope, Johnsson, Mastronardi, Politi, Rossi, Saad, Swarztrauber, Sweet, Toivanen and Turnbull [1–3, 33, 41, 42, 57, 59, 82, 91, 94].

In the case of the system (2) discretizing the Poisson equation, the cost $O(mn\log_2 n)$ can be reduced to $O(mn\log_2\log_2 n)$ if the recursion is halted at step k when the size of the system is $O(n/\log_2 n)$ and the smaller system is solved by means of the fast sine transform. This variant of CR, known as FACR (Fourier Analysis and Cyclic Reduction), has been designed by Heller in [54]. Effective implementations of CR for elliptic equations are given in [90] by Swarztrauber and Sweet.

For general (dense) matrices A, B, C, the computation of $A^{(k)}, B^{(k)}, C^{(k)}$ costs $O(m^3\log_2 n)$, while back substitution costs $O(nm^2)$. If the blocks A, B, C belong to some matrix algebra (say, circulant, Hartley, triangular) then the blocks generated by CR belong to the same algebra as well. This property can be exploited for reducing the computational cost of the method. Other structures like that of band or Toeplitz matrices are not generally preserved by CR. Therefore the cost of the algorithm is the same as that of general dense blocks. Displacement [63] or quasi-separable structures [97], even though not preserved in general, are numerically maintained under certain assumptions on the numerical values of the blocks [22].

It is interesting to point out that for n even, one step of CR applied to $\mathrm{Trid}_n(B, A, C)$ provides the $(n/2) \times (n/2)$ block matrix

$$\begin{bmatrix} A^{(1)} & C^{(1)} & & & 0 \\ B^{(1)} & A^{(1)} & C^{(1)} & & \\ & \ddots & \ddots & \ddots & \\ & & B^{(1)} & A^{(1)} & C^{(1)} \\ 0 & & & B^{(1)} & \widetilde{A}^{(1)} \end{bmatrix}$$

where the blocks $A^{(1)}, B^{(1)}$ and $C^{(1)}$ are defined by (9) while $\widetilde{A}^{(1)} = A - BA^{-1}C$.

If one replaces the odd-even permutation by the even-odd permutation, then also the first diagonal block of the block tridiagonal matrix obtained after one step of CR is different from $A^{(1)}$ and is given by $\widehat{A}^{(1)} = A - CA^{-1}B$.

More generally, applying one step of CR to the $n_k \times n_k$ block matrix

$$T^{(k)} = \begin{bmatrix} \widehat{A}^{(k)} & C^{(k)} & & & 0 \\ B^{(k)} & A^{(k)} & C^{(k)} & & \\ & \ddots & \ddots & \ddots & \\ & & B^{(k)} & A^{(k)} & C^{(k)} \\ 0 & & & B^{(k)} & \widetilde{A}^{(k)} \end{bmatrix}$$

yields the matrix $T^{(k+1)}$ where the blocks $A^{(k+1)}$, $B^{(k+1)}$ and $C^{(k+1)}$ are defined as in (11), whereas the blocks $\widehat{A}^{(k+1)}$ and $\widetilde{A}^{(k+1)}$ are defined differently according to the parity of n_k and according to the use of the even-odd or odd-even permutation. More precisely, one has

$$\widehat{A}^{(k+1)} = \begin{cases} A^{(k)} - C^{(k)} \left(A^{(k)}\right)^{-1} B^{(k)} - B^{(k)} \left(\widehat{A}^{(k)}\right)^{-1} C^{(k)} & \text{odd-even permutation} \\ \widehat{A}^{(k)} - C^{(k)} \left(A^{(k)}\right)^{-1} B^{(k)} & \text{even-odd permutation} \end{cases}$$

$$\widetilde{A}^{(k+1)} = \begin{cases} A^{(k)} - C^{(k)} \left(\widetilde{A}^{(k)}\right)^{-1} B^{(k)} - B^{(k)} \left(A^{(k)}\right)^{-1} C^{(k)} & \text{case a)} \\ \widetilde{A}^{(k)} - B^{(k)} \left(A^{(k)}\right)^{-1} C^{(k)} & \text{case b)} \end{cases}$$

where in the case a) n_k is odd and the permutation is odd-even, or n_k is even and the permutation is even-odd; in the case b) n_k is odd and the permutation is even-odd, or n_k is even and the permutation is odd-even.

CR can be used in the more general case of block tridiagonal systems with nonconstant blocks along the diagonals, i.e.,

$$\begin{bmatrix} A_1 & C_1 & & 0 \\ B_1 & A_2 & \ddots & \\ & \ddots & \ddots & C_{n-1} \\ 0 & & B_{n-1} & A_n \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1 \\ \boldsymbol{u}_2 \\ \vdots \\ \boldsymbol{u}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_n \end{bmatrix}.$$

In this case, the cost of the kth CR step is $O(m^3 n/2^k)$ so that the overall cost is $O(m^3 n)$. For banded non constant blocks, the banded structure is not preserved. It is an open issue to analyze under which conditions (say diagonal dominance) the quasi-separable structure of the blocks is numerically preserved. In this case the cost of CR could be strongly reduced.

Numerical stability results are given by Amodio and Mazzia in [4] and by Yalamov and Pavlov in [74, 99, 100].

2.4 Convergence properties

It was observed by Hockney [58] that often the infinite sequences $\{B^{(k)}\}_k$ and $\{C^{(k)}\}_k$ generated by (11) converge to zero quadratically. It was proved by Heller [54] that under diagonal dominance of $\mathrm{Trid}(B, A, C)$ the matrices $B^{(k)}$ and $C^{(k)}$ converge to zero quadratically. Yalamov and Pavlov [99] have proved a similar convergence property under certain conditions on the blocks. These properties hold also for block tridiagonal matrices with nonconstant blocks.

Convergence to zero of the off-diagonal blocks $B^{(k)}$ and $C^{(k)}$ of (10) enables one to design an incomplete version of CR [43] where the iteration step is halted when a numerical block diagonal matrix is encountered. The same technique is fundamental in solving infinite systems [19] as we will see in Section 3.4.

3 A modern analysis of CR

This section is devoted to the description of some properties of CR which were found at the end of the 1990's in a series of papers by Bini and Meini motivated by the solution of structured Markov chains encountered in queueing models [65, 70, 71]. For a detailed analysis of these properties in the context of Markov chains and for the related references we refer the reader to the book by Bini, Latouche and Meini [18]. For the relations of CR with structured matrices and the Graeffe iteration we refer to the papers [12] and [13] by Bini, Gemignani and Meini. Here we formulate and prove such properties in a more general context not necessarily related to Markov processes.

In this section we also report some new results concerning the convergence and the applicability of CR, and some new formulae for its implementation which allow one to better manage breakdown due to singularity or ill-conditioning of some matrices encountered in the computation.

3.1 Functional interpretation

We consider the CR algorithm as defined in Section 2.2, applied to the block tridiagonal matrix $\mathrm{Trid}(B, A, C)$. The algorithm generates three matrix sequences $A^{(k)}$, $B^{(k)}$ and $C^{(k)}$, according to (11). Here, for simplicity, we assume that these sequences are defined for any $k \geq 0$.

Associate with the triple (B, A, C) the quadratic matrix polynomial

$$\varphi(z) = B + zA + z^2 C.$$

Observe that $\det \varphi(z)$ is a polynomial of degree at most $2m$. If C is nonsingular, the degree of $\det \varphi(z)$ is exactly $2m$, if C has rank ℓ then the degree is $m + \ell$. We denote by ξ_i, $i = 1, \ldots, 2m$ the roots of $\det \varphi(z)$ ordered so that

$$|\xi_1| \leq |\xi_2| \leq \ldots \leq |\xi_{2m}|,$$

where, for notational simplicity, we assume that $m - \ell$ roots are at the infinity if C has rank $\ell < m$.

Similarly, at the generic kth step of CR we associate with the triple $\left(B^{(k)}, A^{(k)}, C^{(k)}\right)$ the quadratic matrix polynomial

$$\varphi^{(k)}(z) = B^{(k)} + zA^{(k)} + z^2 C^{(k)}, \quad k \geq 0. \tag{13}$$

Then it is immediate to see by direct inspection that Eq. (11) can be equivalently written in functional form as

$$\varphi^{(k+1)}\left(z^2\right) = -\varphi^{(k)}(z)\left(A^{(k)}\right)^{-1}\varphi^{(k)}(-z), \quad k \geq 0. \tag{14}$$

This enable us to describe CR in terms of a sequence of matrix polynomials defined by means of a functional iteration.

It is a nice surprise to discover that the latter equation is nothing else but the natural extension to matrix polynomials of the Graeffe-Dandelin-Lobachevsky root-squaring iteration [72, 73] which, for a general scalar polynomial $p(z)$ of degree h, is simply defined by

$$p_{k+1}\left(z^2\right) = p_k(z)p_k(-z), \quad k \geq 0, \quad p_0(z) = p(z).$$

This iteration generates the sequence $\{p_k(z)\}_k$ of polynomials of degree h such that the roots of $p_{k+1}(z)$ are the square of the roots of $p_k(z)$.

Taking determinants on both sides of (14), one finds that this property holds for the polynomials $\det \varphi^{(k)}(z)$ as well. That is, the roots of $\det \varphi^{(k+1)}(z)$ are the square of the roots of $\det \varphi^{(k)}(z)$. Therefore, $\det \varphi^{(k)}(z)$ has roots $\xi_i^{2^k}$, $i = 1, \ldots, 2m$. Moreover, if $\det \varphi^{(0)}(z)$ has a balanced splitting of the roots with respect to the unit circle

$$|\xi_1| \leq \ldots \leq |\xi_m| < 1 < |\xi_{m+1}| \leq \ldots \leq |\xi_{2m}|, \tag{15}$$

i.e., m roots have modulus less than 1 and m roots have modulus greater than 1, then $\det \varphi^{(k)}(z)$ has m roots which converge to zero and m roots which converge to infinity for $k \to \infty$. Intuitively, the matrix polynomial $\varphi^{(k)}(z)$ should converge to a matrix polynomial of the kind zA^* with $\det A^* \neq 0$, which has m roots at zero and m roots at infinity. Consequently, the blocks $B^{(k)}$ and $C^{(k)}$ should converge to zero. This can be formally proved and we provide this proof in the next section.

A simpler functional formulation of CR can be given in terms of the inverses of the matrix functions $\varphi^{(k)}(z)$. To this end, we introduce the matrix Laurent polynomial $z^{-1}\varphi(z) = z^{-1}B + A + zC$ and

$$\psi^{(0)}(z) = (z^{-1}\varphi(z))^{-1} = (z^{-1}B + A + zC)^{-1}, \tag{16}$$

where the latter matrix function is defined and invertible for all $z \neq \xi_i$, $i = 1, \dots, 2m$, $z \neq 0$. In particular, if $|\xi_i| \neq 1$, for $i = 1, \dots, 2m$, then there exist r, R such that $r < 1 < R$ and $\psi^{(0)}(z)$ is analytic and invertible in the annulus $\mathcal{A} = \mathcal{A}(r, R)$. A possible choice for r and R is the largest modulus root $|\xi_i|$ inside the unit disk and the smallest modulus root $|\xi_i|$ outside the unit disk, respectively.

Since the matrix function $\psi^{(0)}(z)$ is analytic in $\mathcal{A}$, it can be represented by means of the Laurent matrix power series

$$\psi^{(0)}(z) = \sum_{i=-\infty}^{+\infty} z^i H_i \tag{17}$$

where H_i are suitable $m \times m$ matrices such that $\sum_{i=-\infty}^{+\infty} ||H_i||$ is finite for some norm $||\cdot||$. In fact, for the Cauchy integral theorem [55, Theorem 4.4c] any analytic function over the annulus $\mathcal{A}$ has coefficients which decay to zero exponentially, so that, for any matrix norm and for any $\epsilon > 0$ one has

$$\begin{aligned} ||H_i|| &\leq \theta(r+\epsilon)^i \quad \text{for } i > 0, \\ ||H_i|| &\leq \theta(R-\epsilon)^i \text{ for } i < 0, \end{aligned} \tag{18}$$

where θ is a suitable constant.

Consider the sequence $\{\psi^{(k)}(z)\}$ recursively formed by the even part of $\psi^{(0)}(z)$ as

$$\psi^{(k+1)}(z^2) = \frac{1}{2}\left(\psi^{(k)}(z) + \psi^{(k)}(-z)\right). \tag{19}$$

Observe that

$$\psi^{(k)}(z) = \sum_{i=-\infty}^{+\infty} z^i H_{i\cdot 2^k}. \tag{20}$$

Indeed, the function $\psi^{(k)}(z)$ is defined on $\mathcal{A}$, but we can easily prove that the analyticity domain is much wider. We have the following

Theorem 1 *The matrix function $\psi^{(k)}(z)$ is analytic in the annulus*

$$\mathcal{A}_k = \mathcal{A}\left(r^{2^k}, R^{2^k}\right). \tag{21}$$

Moreover, if $\det H_0 \neq 0$*, there exists $k_0 > 0$ such that for any $k \geq k_0$ the matrix function $\psi^{(k)}(z)$ is nonsingular in $\mathcal{A}$ and the sequence $z^{-1}\varphi^{(k)}(z)$ converges to H_0^{-1} uniformly over all the compact sets $\mathcal{K} \subset \mathcal{A}$.*

Proof Since $\psi(z)$ is analytic in $\mathcal{A}(r, R)$, its block coefficients H_i satisfy (18). Therefore, $||H_{i\cdot 2^k}|| \leq \theta(r+\epsilon)^{2^k\cdot i}$ for $i > 0$, and $||H_{i\cdot 2^k}|| \leq \theta(R-\epsilon)^{i\cdot 2^k}$ for $i < 0$.

Whence, since ϵ is arbitrary, one deduces that $\psi^{(k)}(z)$ is analytic in $\mathcal{A}_k$. Let $\mathcal{K}$ be a compact set in $\mathcal{A}$. Then

$$\sup_{z\in\mathcal{K}} \left|\psi^{(k)}(z) - H_0\right| = \sup_{z\in\mathcal{K}} \left|\sum_{i\neq 0} z^{i\cdot 2^k} H_{i\cdot 2^k}\right|.$$

Since $z \in \mathcal{K} \subset \mathcal{A}$, then there exists $\delta > 0$ such that $r + \delta < |z| < R - \delta$ so that

$$\sup_{z\in\mathcal{K}} |\psi^{(k)}(z) - H_0| \leq \sum_{i>0} (R-\delta)^{i\cdot 2^k} \cdot |H_{i\cdot 2^k}| + \sum_{i<0} (r+\delta)^{i\cdot 2^k} \cdot |H_{i\cdot 2^k}|.$$

By choosing $\epsilon < \delta$ in (18) one obtains that $\sup_{z\in\mathcal{K}} |\psi^{(k)}(z) - H_0|$ converges to zero, i.e., the sequence $\psi^{(k)}(z)$ uniformly converges to H_0 over $\mathcal{K}$. In particular, for the continuity of the function determinant there exists $k_0 > 0$ such that $\det \psi^{(k)}(z) \neq 0$ for any $z \in \mathcal{A}$ and $k \geq k_0$. $\square$

Notice that we can rewrite (14) as

$$\begin{aligned}\varphi^{(k+1)}(z^2) &= -\varphi^{(k)}(z) \left(\frac{\varphi^{(k)}(z) - \varphi^{(k)}(-z)}{2z}\right)^{-1} \varphi^{(k)}(-z)\\ &= \left(\frac{\varphi^{(k)}(z)^{-1} - \varphi^{(k)}(-z)^{-1}}{2z}\right)^{-1}. \end{aligned} \tag{22}$$

This equation is the basis to prove the following

Theorem 2 *Let the matrix polynomial $\varphi(z)$ be nonsingular for $z \in \mathcal{A}(r, R)$ for $r < 1 < R$. If $\det A^{(i)} \neq 0$, $i = 0, 1, \ldots, k-1$, then $\varphi^{(i)}(z)$ is well defined by (13) and (14), for $i = 0, 1, \ldots, k$, $\det \varphi^{(i)}(z) \neq 0$ for $z \in \mathcal{A}_i$, $i = 1, \ldots, k$, and it holds*

$$\varphi^{(i)}(z)^{-1} = z^{-1}(\psi^{(i)}(z)), \quad i = 0, \ldots, k, \tag{23}$$

and $\det \psi^{(i)}(z) \neq 0$ for $z \in \mathcal{A}_i$, $i = 0, \ldots, k$. Moreover, if $\psi^{(i)}(z)$, defined by (19), and (14), is nonsingular for $z \in \mathcal{A}_i$, for $i = 1, \ldots, k$, then $\varphi^{(i)}(z)$, defined by (13), (14), exists for $i = 1, \ldots, k$ and it holds $\varphi^{(i)}(z) = z(\psi^{(i)}(z))^{-1}$, for $i = 1, \ldots, k$.

Proof Since $\det A^{(i)} \neq 0$, $i = 0, \ldots, k-1$, then $\varphi^{(i)}(z)$ is well defined, for $i = 1, \ldots, k$, in view of (14). Moreover, from (14) the roots of $\det \varphi^{(i)}(z)$ are the 2^i powers of the roots of $\det \varphi(z)$, so that $\varphi^{(i)}(z)$ is nonsingular in $\mathcal{A}_i$, and (23) follows from (22) by using an induction argument on k. Concerning the second part, we proceed by induction on k. Assume $k = 1$. Under the hypothesis of the theorem, since $\varphi(z)$ is nonsingular in $\mathcal{A}$, then $\psi^{(0)}(z) = z\varphi(z)^{-1}$ is nonsingular in $\mathcal{A}$, and therefore from (19) one has

$$\psi^{(1)}(z^2) = \frac{1}{2}\psi^{(0)}(z) \left(\psi^{(0)}(z)^{-1} + \psi^{(0)}(-z)^{-1}\right) \psi^{(0)}(-z). \tag{24}$$

Since $\det \psi^{(1)}(z) \neq 0$ for $z \in \mathcal{A}_1$, in view of (16), the above equation implies that $A^{(0)}$ is nonsingular so that $\varphi^{(1)}(z)$ is well defined. Inverting both sides of (24) one finds that

$$\psi^{(1)}(z^2)^{-1} = -(1/z^2)\varphi^{(0)}(z)(A^{(0)})^{-1}\varphi^{(0)}(-z).$$

Therefore, $\varphi^{(1)}(z) = z(\psi^{(1)}(z))^{-1}$. By using the same arguments, the proof can completed by induction on k, relying on the equation

$$\psi^{(i+1)}(z^2) = (1/2)\psi^{(i)}(z)\left(\psi^{(i)}(z)^{-1} + \psi^{(i)}(-z)^{-1}\right)\psi^{(i)}(-z),$$

which is derived from (19). □

Observe that, in principle, $\varphi^{(k)}(z)$ might not be defined because of the singularity of some $A^{(h)}$, for $h < k$ whereas the matrix functions $\psi^{(h)}(z)$ are well defined and analytic over $\mathcal{A}$, for any h. This property enables us to define the matrix polynomials $\varphi^{(k)}(z)$ independently of $\varphi^{(h)}(z)$, $h < k$, by means of the matrix functions $\psi^{(k)}(z)$ which always exist, provided that $\det \psi^{(k)}(z) \neq 0$ for $z \in \mathcal{A}$.

More details on the issues concerning the applicability of CR will be discussed in Section 3.3.

3.2 Convergence properties

In the light of the Cauchy integral theorem [55, Theorem 4.4c], the analyticity properties of the functions $\varphi^{(k)}(z)$ and $\psi^{(k)}(z)$, imply interesting convergence properties of CR. In particular, from (18) it follows that the Laurent series $\psi^{(k)}(z)$, except for its constant block, has block coefficients which tend to zero double exponentially with k. The nonsingularity of H_0 is needed to export this convergence property to the functions $\varphi^{(k)}(z) = z(\psi^{(k)}(z))^{-1} = z^{-1}B^{(k)} + A^{(k)} + zC^{(k)}$. This property is formalized in the following

Theorem 3 *Let the matrix polynomial $\varphi(z)$ be nonsingular in $\mathcal{A} = \mathcal{A}(r, R)$ where $r < 1 < R$. Assume that H_0 of (16), (17) is nonsingular. If CR can be carried out with no breakdown, then for any $\epsilon > 0$ and for any matrix norm $||\cdot||$, there exists $\theta > 0$ such that*

$$||B^{(k)}|| \leq \theta(r+\epsilon)^{2^k}, \quad ||C^{(k)}|| \leq \theta/(R-\epsilon)^{2^k}$$
$$||A^{(k)} - H_0^{-1}|| \leq \theta\left(\tfrac{r+\epsilon}{R-\epsilon}\right)^{2^k}.$$

Moreover, the roots of $\det\varphi(z)$ *satisfy the splitting property* (15).

Proof Equating the constant terms and the coefficients of z and z^{-1} in the identity

$$\left(\sum_{i=-\infty}^{+\infty} z^i H_{i\cdot 2^k}\right)\left(z^{-1}B^{(k)} + A^{(k)} + zC^{(k)}\right) = I$$

yields

$$\begin{cases} H_0 B^{(k)} + H_{-2^k} A^{(k)} + H_{-2^{k+1}} C^{(k)} = 0 \\ H_{2^k} B^{(k)} + H_0 A^{(k)} + H_{-2^k} C^{(k)} = I \\ H_{2^{k+1}} B^{(k)} + H_{2^k} A^{(k)} + H_0 C^{(k)} = 0 \end{cases}$$

whence

$$\begin{bmatrix} I & H_0^{-1} H_{-2^k} & H_0^{-1} H_{-2^{k+1}} \\ H_0^{-1} H_{2^k} & I & H_0^{-1} H_{-2^k} \\ H_0^{-1} H_{2^{k+1}} & H_0^{-1} H_{2^k} & I \end{bmatrix} \begin{bmatrix} B^{(k)} \\ A^{(k)} - H_0^{-1} \\ C^{(k)} \end{bmatrix} = \begin{bmatrix} -H_0^{-1} H_{-2^k} H_0^{-1} \\ 0 \\ -H_0^{-1} H_{2^k} H_0^{-1} \end{bmatrix}.$$

Since the inverse of the matrix in the above system can be written as

$$\begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} + \sum_{i=1}^{+\infty} (-1)^i \begin{bmatrix} 0 & H_0^{-1} H_{-2^k} & H_0^{-1} H_{-2^{k+1}} \\ H_0^{-1} H_{2^k} & 0 & H_0^{-1} H_{-2^k} \\ H_0^{-1} H_{2^{k+1}} & H_0^{-1} H_{2^k} & 0 \end{bmatrix}^i$$

one obtains

$$\begin{bmatrix} B^{(k)} \\ A^{(k)} - H_0^{-1} \\ C^{(k)} \end{bmatrix} \doteq \begin{bmatrix} -H_0^{-1} H_{-2^k} H_0^{-1} \\ H_0^{-1} H_{-2^k} H_0^{-1} H_{2^k} H_0^{-1} + H_0^{-1} H_{2^k} H_0^{-1} H_{-2^k} H_0^{-1} \\ -H_0^{-1} H_{2^k} H_0^{-1} \end{bmatrix}$$

where $\doteq$ denotes equality up to lower order terms. The above equation together with (18) implies the convergence properties. Concerning the splitting of the roots of $\det \varphi(z)$, since $B^{(k)}$ and $C^{(k)}$ converge to zero and $A^{(k)}$ converges to a nonsingular matrix, and for the continuity of the roots of polynomials, the polynomial $\det \varphi^{(k)}$ has m roots which converge to zero and m roots which converge to the infinity. Since these roots are the 2^k powers of the roots of $\det \varphi(z)$, then the latter polynomial has m roots of modulus less than 1 and m roots of modulus greater than 1. □

The above theorem gives general conditions under which the convergence of CR is quadratic and the convergence speed is related to the location of the roots of $\det \varphi(z)$.

The condition $\det H_0 \neq 0$ and the splitting property (15) are related to the existence of the solution of two matrix equations with spectral radius less than 1 [18].

Theorem 4 *If the two matrix equations*

$$B + AX + CX^2 = 0$$

$$BY^2 + AY + C = 0$$

have solutions X and Y such that $\rho(X) < 1$ and $\rho(Y) < 1$ then $\det H_0 \neq 0$, *the roots ξ_i, $i = 1, \ldots, 2m$ of* $\det \varphi(z)$ *satisfy* (15), *moreover $\rho(X) = |\xi_m|$, $\rho(Y) = 1/|\xi_{m+1}|$.*

It is interesting to point out that there exist examples where only one of the above matrix equations has a solution with spectral radius less than 1, and H_0 is singular, see for instance Example 1 in the next Section 3.3.

In [13] it has been proved that under the condition (15) if the solution X to $B + AX + CX^2 = 0$ with spectral radius $|\xi_m|$ exists, then the solution W to $W^2B + WA + C = 0$ exists with spectral radius $1/|\xi_{m+1}|$. Similarly, if the solution Y to $BY^2 + AY + C = 0$ with spectral radius $1/|\xi_{m+1}|$ exists, then the solution Z to $B + ZA + Z^2C = 0$ exists with spectral radius $|\xi_m|$.

The assumptions of the results in this section can be weakened by allowing that either $r = 1$ or $R = 1$. In this case we still have convergence to zero of at least one of the two sequences $B^{(k)}$ or $C^{(k)}$. If $r = R = 1$, then the annulus $\mathcal{A}(r, R)$ is empty and convergence is not guaranteed in general. For problems encountered in Markov chains, where $A + B + C$ is a singular M-matrix, one may encounter the different cases $r = 1 < R$, $r < R = 1$, or $r = R = 1$. In the latter, known as *null recurrent* case, convergence still holds under additional conditions even though it turns to linear [45].

3.3 Applicability

In principle, CR can have a breakdown if some matrix $A^{(k)}$ is singular. If k_0 is the first integer for which $\det A^{(k_0)} = 0$, then the matrix sequences $A^{(k)}$, $B^{(k)}$, $C^{(k)}$ generated by CR are not computable for $k > k_0$ by means of (11).

On the other hand, the sequence $\psi^{(k)}(z)$ of (19) is defined also in the case of breakdown, moreover, if $\det \psi^{(k)}(z) \neq 0$ for some $k > k_0$ and $z \in \mathcal{A}$, we may still define $\varphi^{(k)}(z) = z\psi^{(k)}(z)^{-1}$.

In the case where breakdown is not encountered, the functions $\varphi^{(k)}(z)$ defined in this way coincide with the functions (13) in view of Theorem 2.

Therefore, with an abuse of notation, in the case of breakdown we keep denoting by $\varphi^{(k)}(z) = z\psi^{(k)}(z)^{-1}$ even though $\varphi^{(k)}(z)$ cannot defined by means of (11).

Observe also that, if breakdown is not encountered, the matrix functions $\varphi^{(k)}(z)$ are quadratic matrix polynomials. A continuity argument applied to the triple B, $A + \epsilon I$, C, for ϵ in a neighborhood of 0, enables us to prove that $\varphi^{(k)}(z) = z\psi^{(k)}(z)^{-1}$ are still quadratic matrix polynomials even in case of breakdown.

The fact that the quadratic matrix polynomials $\varphi^{(k)}(z)$ can be defined for $k > k_0$, in the case of breakdown at k_0, has mainly a theoretical relevance. In order to use this property computationally it is more convenient to interpret CR in terms of Schur complements.

Given an $n \times n$ matrix H with indices in the set $\mathcal{N}$ of cardinality n, partition $\mathcal{N}$ into two disjoint subsets $\mathcal{I}_1, \mathcal{I}_2$ and denote by $H_{\mathcal{I}_i,\mathcal{I}_j}$, the submatrix of H with indices in $\mathcal{I}_i \times \mathcal{I}_j$. Assume that $\det H_{\mathcal{I}_1,\mathcal{I}_1} \neq 0$ and denote by $S(H, \mathcal{N}, \mathcal{I}_1) = H_{\mathcal{I}_2,\mathcal{I}_2} - H_{\mathcal{I}_2,\mathcal{I}_1} H^{-1}_{\mathcal{I}_1,\mathcal{I}_1} H_{\mathcal{I}_1,\mathcal{I}_2}$ the Schur complement in H of $H_{\mathcal{I}_1,\mathcal{I}_1}$.

Let us recall the following property of Schur complements known as the quotient property [36].

Lemma 1 *Let H be an $n \times n$ matrix with indices in the set $\mathcal{N}$ and partition $\mathcal{N}$ into three disjoint sets $\mathcal{I}_1$, $\mathcal{I}_2$, $\mathcal{I}_3$. If $\det H_{\mathcal{I}_1,\mathcal{I}_1} \neq 0$ and $\det H_{\mathcal{I}_1\cup\mathcal{I}_2,\mathcal{I}_1\cup\mathcal{I}_2} \neq 0$, then the principal submatrix of $S(H, \mathcal{N}, \mathcal{I}_1)$ with indices in $\mathcal{I}_2$ is nonsingular and*

$$S(H, \mathcal{N}, \mathcal{I}_1 \cup \mathcal{I}_2) = S(S(H, \mathcal{N}, \mathcal{I}_1), \mathcal{I}_2 \cup \mathcal{I}_3, \mathcal{I}_2).$$

By using this lemma we may prove the following

Theorem 5 *Assume that the matrices $A^{(k)}$, $B^{(k)}$, and $C^{(k)}$ can be constructed with no breakdown by means of (11) for $k = 1, \dots, q$ and let $\bar{\mathcal{I}}_k$ be the complement of the set $\mathcal{I}_k = \{i \cdot 2^k,\ \ i = 1, \dots, 2^{q-k} - 1\}$ in $\mathcal{N} = \{1, 2, \dots, 2^q - 1\}$. Then the matrix $\mathrm{Trid}_{2^{q-k}-1}(B^{(k)}, A^{(k)}, C^{(k)})$ is the Schur complement in $T_{2^q-1} = \mathrm{Trid}_{2^q-1}(B, A, C)$ of the principal submatrix having block indices in $\bar{\mathcal{I}}_k$.*

Proof We proceed by induction on k. For $k = 1$ the property clearly holds since $\mathrm{Trid}_{2^{q-1}-1}(B^{(1)}, A^{(1)}, C^{(1)})$ is the Schur complement in $\mathrm{Trid}_{2^q-1}(B, A, C)$ of the principal submatrix having block indices in the set of odd integers. Let us consider the inductive step. By assuming that the property holds for $k - 1$, one finds that

$$\mathrm{Trid}_{2^{q-k+1}-1}(B^{(k-1)}, A^{(k-1)}, C^{(k-1)}) = S(T_{2^q-1}, \mathcal{N}, \bar{\mathcal{I}}_{k-1}).$$

Applying the cyclic reduction step to $\mathrm{Trid}_{2^{q-k+1}-1}(B^{(k-1)}, A^{(k-1)}, C^{(k-1)})$, one obtains

$$\mathrm{Trid}_{2^{q-k}-1}(B^{(k)}, A^{(k)}, C^{(k)}) = S(\mathrm{Trid}_{2^{q-k+1}-1}(B^{(k-1)}, A^{(k-1)}, C^{(k-1)}), \mathcal{I}_{k-1}, \bar{\mathcal{J}}),$$

where $\mathcal{J}$ is the subset of $\mathcal{I}_{k-1} = \{i2^{k-1},\ \ i = 1, 2, \dots, 2^{q-k+1} - 1\}$ obtained with even values for i. That is, $\mathcal{J} = \mathcal{I}_k$. The proof is completed in view of Lemma 1. $\square$

Observe that, in view of the above theorem, the existence of $B^{(k)}$, $A^{(k)}$, and $C^{(k)}$ does not require the nonsingularity of the blocks $A^{(h)}$ for $h = 1, \dots, k - 1$. In fact it is sufficient that the submatrix of $\mathrm{Trid}_{2^q-1}(B, A, C)$ with block indices in $\bar{\mathcal{I}}_{k-1}$ is nonsingular. The following result provides a better understanding of this fact.

Theorem 6 *The kth step of CR applied to the matrix $\mathrm{Trid}_{2^q-1}(B, A, C)$, $q > k$, can be performed if and only if the matrix $T_{2^k-1} = \mathrm{Trid}_{2^k-1}(B, A, C)$ is nonsingular.*

Proof By Theorem 5, the matrix $\mathrm{Trid}_{2^{q-k}-1}(B^{(k)}, A^{(k)}, C^{(k)})$ generated after k steps of CR is the Schur complement in $\mathrm{Trid}_{2^q-1}(B, A, C)$ of the principal submatrix having block indices in the complement of the set $\{i \cdot 2^k, i = 1, \dots, 2^{q-k} - 1\}$. Let us permute block rows and block columns of $\mathrm{Trid}_{2^q-1}(B,$

A, C) so that the ones with indices $i \cdot 2^k$ are at the bottom. In this way the matrix that we obtain has the following structure

$$\left[\begin{array}{cccc|cccc} T & & & 0 & \mathcal{C} & & & 0 \\ & T & & & \mathcal{B} & \ddots & & \\ & & \ddots & & & & \ddots & \mathcal{C} \\ 0 & & & T & 0 & & & \mathcal{B} \\ \hline \widehat{\mathcal{B}} & \widehat{\mathcal{C}} & & 0 & A & & & 0 \\ & \ddots & \ddots & & & & \ddots & \\ 0 & & \widehat{\mathcal{B}} & \widehat{\mathcal{C}} & 0 & & & A \end{array}\right] \tag{25}$$

where $T = \mathrm{Trid}_{2^k-1}(B, A, C)$, the number of diagonal blocks of the kind T is $2^q - 2^{q-k}$ and the number of diagonal blocks of the kind A is $2^{q-k} - 1$. Moreover,

$$\mathcal{B} = \begin{bmatrix} B \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathcal{C} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ C \end{bmatrix}, \quad \widehat{\mathcal{B}} = \begin{bmatrix} 0 \ldots 0 \; B \end{bmatrix}, \quad \widehat{\mathcal{C}} = \begin{bmatrix} C \; 0 \ldots 0 \end{bmatrix}.$$

Therefore the Schur complement can be computed if and only if $\det T \neq 0$. □

Observe that, in view of the above theorem, the applicability of CR is guaranteed if for instance $\mathrm{Trid}(B, A, C)$ is strongly diagonally dominant or irreducibly diagonally dominant, or if it is a non-singular M-matrix or a singular irreducible M-matrix, or symmetric and positive definite. There are important applications where these hypotheses are verified. For instance, in the analysis of Quasi–Birth-Death processes [65] one has $A = I - S$, $S \geq 0$ $B, C \leq 0$ and $S - B - C$ is an irreducible stochastic matrix. Under the assumption of irreducibility of $\mathrm{Trid}(B, A, C)$, the Gerschgorin theorem guarantees nonsingularity of all its principal submatrices.

The following example shows that breakdown can occur at any arbitrary step of CR.

Example 1 Let

$$G = \begin{bmatrix} 1/2 & 0 \\ 1 & 1/2 \end{bmatrix}, \quad R = \begin{bmatrix} 1/2 & \alpha \\ 0 & 1/2 \end{bmatrix},$$

where α is a real parameter, and set $\varphi(z) = B + zA + z^2 C$ where

$$B = -G, \quad A = (I + RG), \quad C = -R.$$

The polynomial $\det \varphi(z)$ has zeros $\{1/2, 1/2, 2, 2\}$, moreover the matrix function $\psi(z) = \sum_{i=-\infty}^{+\infty} z^i H_i$ is such that

$$H_0 = \begin{bmatrix} 4/3 & 8\alpha/9 \\ 8/9 & 4/3 + 80\alpha/27 \end{bmatrix}.$$

Applying CR to Trid(B, A, C) generates matrices $A^{(k)}$ such that $\det A^{(k)}$ is a linear polynomial $p_k(\alpha)$ in α. Therefore for any k there exists α_k such that $p_k(\alpha_k) = 0$, i.e., $A^{(k)}$ is singular. In particular, $\alpha_0 = -25/16$ and $\alpha_1 = -7225/11024$. Moreover, it holds $\lim_k \alpha_k = -9/16$. For this value of α one has $\det H_0 = 0$.

Observe that, while the solutions with minimal spectral radius to the matrix equations

$$-RX^2 + (I + RG)X - G = 0, \quad -Y^2G + Y(I + RG) - R = 0$$

exist and coincide with G and R, respectively, the solutions to the dual matrix equations

$$-GX^2 + (I + RG)X - R = 0, \quad -Y^2R + Y(I + RG) - G = 0$$

exist only if $\det H_0 \neq 0$ and they are given by

$$X = \frac{1}{18 + 32\alpha}\begin{bmatrix} 9 + 4\alpha & 18\alpha \\ -8\alpha & 9 + 28\alpha \end{bmatrix}, \quad Y = \frac{1}{18 + 32\alpha}\begin{bmatrix} 9 + 4\alpha & -8\alpha^2 \\ 18 & 9 + 28\alpha \end{bmatrix}.$$

In fact, for $\alpha = -9/16$ even though CR has no breakdown, the conditions for the convergence of CR given in Theorem 3 are not satisfied since $\det H_0 = 0$. It can be easily verified that in fact CR does not converge in this case, even though the splitting (15) holds. □

Breakdown situations are not a serious drawback of cyclic reduction. In fact, relying on the Schur complement interpretation of Theorem 5 it is possible to compute the matrix sequences $A^{(k)}$, $B^{(k)}$, $C^{(k)}$ at least for all the values of k for which the matrix $\mathrm{Trid}_{2^k-1}(B, A, C)$ is nonsingular. The following two theorems provide explicit expressions for the computation of the blocks $A^{(k)}$, $B^{(k)}$, $C^{(k)}$. Similar formulas for the solution of the block tridiagonal system (8) in case of breakdown can be derived in the same way by relying on the block LU factorization of the matrix (25). Concerning the latter issue we leave the detail to the reader in order to make the presentation less technical.

Theorem 7 *Let* $\mathrm{Trid}_{2^k-1}(B, A, C)$ *be nonsingular and denote by* $S_{i,j}^{(k)}$ *the blocks of* $S^{(k)} = \mathrm{Trid}_{2^k-1}(B, A, C)^{-1}$. *Then with* $n = 2^k - 1$ *it holds*

$$\begin{aligned} A^{(k)} &= A - BS_{n,n}^{(k)}C - CS_{1,1}^{(k)}B, \\ B^{(k)} &= -BS_{n,1}^{(k)}B, \\ C^{(k)} &= -CS_{1,n}^{(k)}C. \end{aligned} \tag{26}$$

Proof In view of (25), computing the Schur complement of the matrix in the upper leftmost corner provides the blocks $A^{(k)}$, $B^{(k)}$ and $C^{(k)}$. For the structure of (25) one finds that $A^{(k)} = A - \widehat{\mathcal{C}}T^{-1}\mathcal{B} + \widehat{\mathcal{B}}T^{-1}\mathcal{C}$, and $B^{(k)} = -\widehat{\mathcal{B}}T^{-1}\mathcal{B}$, $C^{(k)} = -\widehat{\mathcal{C}}T^{-1}\mathcal{C}$. In view of the structure of the blocks $\mathcal{B}, \mathcal{C}, \widehat{\mathcal{B}}, \widehat{\mathcal{C}}$, the proof is completed. □

It is interesting to observe that (26) generalizes equation (11), obtained for $k = 1$ where $\mathrm{Trid}_1(B, A, C) = A$. Observe also that this formula is useful to avoid situations where $A^{(h)}$ is numerically ill conditioned for some $h < k$. Equation (26) can be also used to compute in a different way, suitable subsequences of the sequences generated by the standard CR. In fact, since $\mathrm{Trid}_{2^{q-h-k}-1}(B^{(k+h)}, A^{(k+h)}, C^{(k+h)})$ can be viewed as the matrix obtained by applying k steps of CR to the matrix $\mathrm{Trid}_{2^{q-h}-1}(B^{(h)}, A^{(h)}, C^{(h)})$, one has

$$\begin{aligned} A^{(h+k)} &= A^{(h)} - B^{(h)} S_{1,1}^{(h,k)} C^{(h)} - C^{(h)} S_{n,n}^{(h,k)} B^{(h)}, \\ B^{(h+k)} &= -B^{(h)} S_{1,n}^{(h,k)} B^{(h)}, \\ C^{(h+k)} &= -C^{(h)} S_{n,1}^{(h,k)} C^{(h)} \end{aligned} \tag{27}$$

where $S^{(h,k)} = \mathrm{Trid}_{2^k-1}\left(B^{(h)}, A^{(h)}, C^{(h)}\right)^{-1}$ and $n = 2^k - 1$.

By relying on the Sherman-Woodbury-Morrison (SWM) formula, it is possible to relate the blocks $S_{i,j}^{(n)}$ with the blocks $S_{i,j}^{(2n-1)}$.

Theorem 8 *Let $n = 2^k - 1$ and assume that $\mathrm{Trid}_n(B, A, C)$ is nonsingular, moreover denote by $S_{i,j}^{(k)}$ the blocks of $S^{(k)} = \mathrm{Trid}_n(B, A, C)^{-1}$. Then the matrix $\mathrm{Trid}_{2n+1}(B, A, C)$ is nonsingular if and only if $A - CS_{n,n}^{(k)} B - BS_{1,1}^{(k)} C$ is nonsingular. Moreover, denoting by $W^{(n)} = (A - CS_{n,n}^{(k)} B - BS_{1,1}^{(k)} C)^{-1}$, it holds*

$$\begin{aligned} S_{1,1}^{(k+1)} &= S_{1,1}^{(k)} - S_{1,n}^{(k)} C W^{(k)} B S_{n,1}^{(k)} \\ S_{p,p}^{(k+1)} &= S_{n,n}^{(k)} - S_{n,1}^{(k)} B W^{(k)} C S_{1,n}^{(k)} \\ S_{1,p}^{(k+1)} &= -S_{1,n}^{(k)} C W^{(k)} C S_{1,n}^{(k)} \\ S_{p,1}^{(k+1)} &= -S_{n,1}^{(k)} B W^{(k)} B S_{n,1}^{(k)} \end{aligned} \tag{28}$$

where $p = 2^{k+1} - 1$.

Proof It follows by applying the SWM formula to the partitioning

$$\mathrm{Trid}_p(B, A, C) = \left[\begin{array}{c|c|c} \mathrm{Trid}_n(B, A, C) & 0 & 0 \\ \hline \widehat{\mathcal{B}} & I & \widehat{\mathcal{C}} \\ \hline 0 & 0 & \mathrm{Trid}_n(B, A, C) \end{array}\right] + \begin{bmatrix} 0 \\ \\ C \\ A - I \\ B \\ \\ 0 \end{bmatrix} \left[0 \dots 0 \; I \; 0 \dots 0\right].$$

For the sake of brevity we omit the technical details. $\square$

It is interesting to observe that, by applying the SWM formula to a different partitioning where the two diagonal blocks $\mathrm{Trid}_n(B, A, C)$ are replaced by $\mathrm{Trid}_p(B, A, C)$ and $\mathrm{Trid}_q(B, A, C)$, respectively, one arrives at the more general formula where we set $r = p + q + 1$:

$$\begin{aligned} S_{1,1}^{(r)} &= S_{1,1}^{(p)} - S_{1,r}^{(p)} C W^{(p,q)} B S_{r,1}^{(p)} \\ S_{r,r}^{(r)} &= S_{r,r}^{(q)} - S_{r,1}^{(q)} B W^{(p,q)} C S_{1,r}^{(q)} \\ S_{1,r}^{(r)} &= -S_{1,r}^{(p)} C W^{(p,q)} C S_{1,r}^{(q)} \\ S_{r,1}^{(r)} &= -S_{r,1}^{(q)} B W^{(p,q)} B S_{r,1}^{(q)} \end{aligned} \tag{29}$$

for $W^{(p,q)} = (A - BS_{r,r}^{(p)} C - CS_{1,1}^{(q)} B)^{-1}$, where we assumed $\mathrm{Trid}_p(B, A, C)$, $\mathrm{Trid}_q(B, A, C)$ and $\mathrm{Trid}_r(B, A, C)$ nonsingular.

Equations (28) and (29) provide a means for differently implementing CR where for singularity or ill-conditioning of some principal minor it is convenient to skip some inversion. For instance, assume that the blocks $S_{i,j}^{(r_i)}$ are available for some given $r_1, r_2, \ldots, r_k$ obtained in the previous steps. We may choose the largest values of r_i and r_j, if any, such that the matrix $A - BS_{r_i,r_i}^{(r_i)} C - CS_{1,1}^{(r_j)} B$ is well conditioned, and then apply (29) with $p = r_i$, $q = r_j$ and $r = r_i + r_j + 1$.

It is important to point out that in the solution of matrix equations encountered in queuing models, where some acceleration techniques are applied, the applicability of CR is not guaranteed and breakdown situations can be encountered. Here, the goal is not solving the linear system (8) but is rather computing the limit of the sequence $A^{(k)}$. In these applications, the matrix H_0 is nonsingular so that the matrix $\psi^{(k)}(z)$ is nonsingular for sufficiently large k for $z \in \mathcal{A}$. This means that there exists k_0 such that CR has no breakdown for any $k \geq k_0$. Therefore, under the convergence condition of Theorem 3, formulae (26), (27), (28) enable one to apply CR even though some block $A^{(k)}$ is singular.

3.4 Infinite systems

The problem of solving infinite systems of the kind

$$\begin{bmatrix} \widehat{A} & C & & 0 \\ B & A & C & \\ 0 & \ddots & \ddots & \ddots \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \end{bmatrix},$$

where A, B, C and $\widehat{A}$ are $m \times m$ matrices and $\boldsymbol{x}_i$, $\boldsymbol{b}_i$, $i \geq 1$, are m-dimensional vectors, is encountered in several applications. For instance, in a wide variety of stochastic processes (Quasi-Birth-Death) [65], where $A = I - V$, $V \geq 0$, $B \leq 0$, $C \leq 0$ and $V - B - C$ is stochastic, the solution of the above systems provides the invariant probability (steady-state) vector.

Under the hypotheses of convergence of Theorem 3, CR is a suitable tool for computing an arbitrary number of components provided that the sequence

$\{\boldsymbol{x}_i\}_i$ is uniformly bounded in some norm. This in particular occurs when the solution is a probability vector with nonnegative components which sum up to 1.

The idea is to apply CR relying on the even/odd permutation rather than on the odd/even permutation so that at the general kth step the variables $\boldsymbol{x}_{i\cdot 2^k+1}$, for $i = 0, 1, \ldots$, are involved in the system

$$\begin{bmatrix} \widehat{A}^{(k)} & C^{(k)} & & 0 \\ B^{(k)} & A^{(k)} & C^{(k)} & \\ 0 & \ddots & \ddots & \ddots \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_{2^k+1} \\ \boldsymbol{x}_{2\cdot 2^k+1} \\ \vdots \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1^{(k)} \\ \boldsymbol{b}_2^{(k)} \\ \vdots \end{bmatrix}$$

where $A^{(k)}$, $B^{(k)}$ and $C^{(k)}$ are defined as in (11), while

$$\widehat{A}^{(k+1)} = \widehat{A}^{(k)} - B^{(k)}(A^{(k)})^{-1}C^{(k)}, \quad k \geq 0,$$

and $\widehat{A}^{(0)} = \widehat{A}$. The iteration is continued until $C^{(k)}$ is sufficiently small so that from the first block equation $\widehat{A}^{(k)}\boldsymbol{x}_1 + C^{(k)}\boldsymbol{x}_{2^k+1} = \boldsymbol{b}_1^{(k)}$ one can numerically compute $\boldsymbol{x}_1$ by neglecting $C^{(k)}\boldsymbol{x}_{2^k+1}$, and then compute as many components $\boldsymbol{x}_{i\cdot 2^k+1}$ as needed. Back substitution completes the recovery of all the previous components. More details of this technique are given in [19].

4 The Hessenberg case

As proved in [19] in the context of Markov chains, cyclic reduction can be easily extended to block Hessenberg block Toeplitz systems $H\boldsymbol{x} = \boldsymbol{b}$ where

$$H = \begin{bmatrix} A_0 & A_1 & \ldots & A_{n-1} \\ A_{-1} & A_0 & \ddots & \vdots \\ & \ddots & \ddots & A_1 \\ 0 & & A_{-1} & A_0 \end{bmatrix},$$

A_i, for $i = -1, \ldots, n-1$, are $m \times m$ matrices, and where $\boldsymbol{x}$ and $\boldsymbol{b}$ are partitioned into n vectors $\boldsymbol{x}_i$, $\boldsymbol{b}_i$, $i = 1, \ldots, n$ of dimension m.

Here we show this extension in the general case. Assume $n = 2^q - 1$, applying an odd/even permutation to block columns and rows in the system above yields the equivalent system

$$\left[\begin{array}{cccc|ccc} A_0 & A_2 & \ldots & A_{n-1} & A_1 & \ldots & A_{n-2} \\ & A_0 & \ddots & \vdots & A_{-1} & \ddots & \vdots \\ & & \ddots & A_2 & & \ddots & A_1 \\ 0 & & & A_0 & 0 & & A_{-1} \\ \hline A_{-1} & A_1 & \ldots & A_{n-2} & A_0 & \ldots & A_{n-3} \\ & \ddots & \ddots & \vdots & & \ddots & \vdots \\ 0 & & A_{-1} & A_1 & 0 & & A_0 \end{array}\right] \left[\begin{array}{c} \boldsymbol{x}_1 \\ \boldsymbol{x}_3 \\ \vdots \\ \boldsymbol{x}_n \\ \hline \boldsymbol{x}_2 \\ \boldsymbol{x}_4 \\ \vdots \\ \boldsymbol{x}_{n-1} \end{array}\right] = \left[\begin{array}{c} \boldsymbol{b}_1 \\ \boldsymbol{b}_3 \\ \vdots \\ \boldsymbol{b}_n \\ \hline \boldsymbol{b}_2 \\ \boldsymbol{b}_4 \\ \vdots \\ \boldsymbol{b}_{n-1} \end{array}\right]$$

which we rewrite in compact form as

$$\begin{bmatrix} H_{1,1} & H_{1,2} \\ H_{2,1} & H_{2,2} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_{\text{odd}} \\ \hline \boldsymbol{x}_{\text{even}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_{\text{odd}} \\ \hline \boldsymbol{b}_{\text{even}} \end{bmatrix}.$$

Assume A_0 nonsingular, eliminate the odd block components by means of block Gaussian elimination, i.e., compute the Schur complement S of $H_{1,1}$ and obtain the smaller system of block size $2^{q-1} - 1$:

$$(H_{2,2} - H_{2,1}H_{1,1}^{-1}H_{1,2})\boldsymbol{x}_{\text{even}} = \boldsymbol{b}^{(1)}, \quad \boldsymbol{b}^{(1)} = \boldsymbol{b}_{\text{even}} - H_{2,1}H_{1,1}^{-1}\boldsymbol{b}_{\text{odd}}.$$

One can easily verify that the matrix $S = H_{2,2} - H_{2,1}H_{1,1}^{-1}H_{1,2}$ in the above system is the matrix obtained by deleting the first block column and the last block row in the following expression involving block upper triangular block Toeplitz matrices

$$\begin{bmatrix} 0 & A_0 & \dots & A_{n-3} \\ & 0 & \ddots & \vdots \\ & & \ddots & A_0 \\ 0 & & & 0 \end{bmatrix} - UH_{1,1}^{-1}U, \quad \text{where } U = \begin{bmatrix} A_{-1} & A_1 & \dots & A_{n-2} \\ & A_{-1} & \ddots & \vdots \\ & & \ddots & A_1 \\ 0 & & & A_{-1} \end{bmatrix}.$$

Due to the block Toeplitz structure, one finds that the Schur complement has the same form of H, i.e.,

$$S = \begin{bmatrix} A_0^{(1)} & A_1^{(1)} & \dots & A_{2^{q-1}-1}^{(1)} \\ A_{-1}^{(1)} & A_0^{(1)} & \ddots & \vdots \\ & \ddots & \ddots & A_1^{(1)} \\ 0 & & A_{-1}^{(1)} & A_0^{(1)} \end{bmatrix}$$

where

$$\begin{bmatrix} A_{2^{q-1}-1}^{(1)} \\ \vdots \\ A_0^{(1)} \\ A_{-1}^{(1)} \end{bmatrix} = \begin{bmatrix} A_{n-3} \\ \vdots \\ A_0 \\ 0 \end{bmatrix} - UH_{1,1}^{-1} \begin{bmatrix} A_{n-2} \\ \vdots \\ A_1 \\ A_{-1} \end{bmatrix}. \tag{30}$$

In fact, block triangular block Toeplitz matrices are closed under sum, product and inversion. The above formula provides a means for performing one step of CR. It requires the solution of a block triangular block Toeplitz system with matrix $H_{1,1}$ and the computation of the product of a block Toeplitz matrix and a block vector. All these operations can be performed in $O(m^3n + m^2n\log n)$ by means of FFT [18].

Once the even indexed components have been computed, one can compute the odd-indexed ones just by solving the block triangular block Toeplitz system $H_{1,1}\boldsymbol{x}_{\text{odd}} = \boldsymbol{b}_{\text{odd}} - H_{1,2}\boldsymbol{x}_{\text{even}}$. The cost of the latter computation is $O(m^2n\log n)$.

As in the block tridiagonal case, this technique can be cyclically repeated. Its overall cost is still $O(m^3n + m^2n\log n)$.

It is interesting to observe that the (30) which relates the blocks at two subsequent steps of CR can be written in terms of matrix polynomials. Let $p^{(k)}(z) = \sum_{i=-1}^{2q-k-1} z^{i+1} A_i^{(k)}$, then one has

$$p^{(k+1)}(z) = zp_-^{(k)}(z) - p_+^{(k)}(z)p_-^{(k)}(z)^{-1}p_+^{(k)}(z) \mod z^{2^{q-k-1}+1} \tag{31}$$

where $p_+^{(k)}(z)$ and $p_-^{(k)}(z)$ are the even and the odd parts of $p^{(k)}(z)$, respectively defined by $p_+^{(k)}(z^2) = \big(p^{(k)}(z) + p^{(k)}(-z)\big)/2$, $p_-^{(k)}(z^2) = \big(p^{(k)}(z) - p^{(k)}(-z)\big)/(2z)$.

4.1 Infinite systems

For infinite block Hessenberg systems of the kind

$$\begin{bmatrix} \widehat{A}_0 & \widehat{A}_1 & \widehat{A}_2 & \dots \\ A_{-1} & A_0 & A_1 & \ddots \\ & A_{-1} & A_0 & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \end{bmatrix},$$

CR with the even/odd permutation leads to the sequence of infinite systems

$$\begin{bmatrix} \widehat{A}_0^{(k)} & \widehat{A}_1^{(k)} & \widehat{A}_2^{(k)} & \dots \\ A_{-1}^{(k)} & A_0^{(k)} & A_1^{(k)} & \ddots \\ & A_{-1}^{(k)} & A_0^{(k)} & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_{2^k+1} \\ \boldsymbol{x}_{2\cdot 2^k+1} \\ \vdots \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_1^{(k)} \\ \boldsymbol{b}_2^{(k)} \\ \vdots \end{bmatrix}, \quad k = 1, 2, \dots.$$

Equation (31) still holds by replacing the matrix polynomials $p^{(k)}(z)$ with matrix power series $\varphi^{(k)}(z) = \sum_{i=-1}^{\infty} z^{i+1} A_i^{(k)}$ and by removing the modulo operation. Indeed, one can prove that CR is defined by the following equations

$$\varphi^{(k+1)}(z) = z\varphi_-^{(k)}(z) - \varphi_+^{(k)}(z)\varphi_-^{(k)}(z)^{-1}\varphi_+^{(k)}(z), \quad k \geq 0 \tag{32}$$

where $\varphi_+^{(k)}(z)$ and $\varphi_-^{(k)}(z)$ are the even and the odd parts of $\varphi^{(k)}(z)$, respectively defined by $\varphi_+^{(k)}(z^2) = \big(\varphi^{(k)}(z) + \varphi^{(k)}(-z)\big)/2$, $\varphi_-^{(k)}(z^2) = \big(\varphi^{(k)}(z) - \varphi^{(k)}(-z)\big)/(2z)$, and with $\varphi^{(0)}(z) = \sum_{i=-1}^{+\infty} z^{i+1} A_i$. The recursive formulae for the blocks $\widehat{A}_i^{(k)}$ can be expressed as

$$\widehat{\varphi}^{(k+1)}(z) = \widehat{\varphi}_+^{(k)}(z) - \widehat{\varphi}_-^{(k)}(z)\varphi_-^{(k)}(z)^{-1}\varphi_+^{(k)}(z), \quad k \geq 0, \tag{33}$$

where $\widehat{\varphi}^{(k)}(z) = \sum_{i=0}^{\infty} z^i \widehat{A}_i^{(k)}$ and $\widehat{\varphi}^{(0)}(z) = \sum_{i=0}^{+\infty} z^i A_i$, and $\widehat{\varphi}_+^{(k)}(z)$, $\widehat{\varphi}_-^{(k)}(z)$ are the even and odd part of $\widehat{\varphi}^{(k)}(z)$, respectively.

An equivalent functional formulation of CR can be given in the same form as in the block tridiagonal case as follows:

$$\varphi^{(k+1)}(z^2) = -\varphi^{(k)}(z)\left(\varphi_+^{(k)}(z^2)\right)^{-1}\varphi^{(k)}(-z), \quad k \geq 0.$$

This functional formulation enables us to provide convergence results and a simple implementation based on the technique of evaluation/interpolation at the Fourier points. These topics are examined in the next Sections 4.2 and 4.3, respectively.

Also for Hessenberg systems one can define the matrix function $\psi(z) = z\varphi(z)^{-1}$ and prove that $\psi^{(k)}(z) = z\varphi^{(k)}(z)^{-1}$ for any z for which $\varphi(z)$ is nonsingular, where

$$\psi^{(k+1)}(z^2) = \left(\psi^{(k)}(z) + \psi^{(k)}(-z)\right)/2, \quad k \geq 0,$$

$$\psi^{(0)}(z) = \psi(z) = \sum_{i=-\infty}^{+\infty} z^i H_i.$$

4.2 Convergence properties

Convergence properties of $\varphi^{(k)}(z)$ can be given under suitable additional assumptions [18, 28] for problems related to Markov chains. Here we provide more general results which extend the convergence properties obtained in the block tridiagonal case.

Theorem 9 *Assume that the matrix function $\varphi(z)$ is analytic and nonsingular for $z \in \mathcal{A}(r, R)$, where $r < 1 < R$, and that* $\det H_0 \neq 0$, *where $\psi(z) = z\varphi(z)^{-1} = \sum_{i=-\infty}^{+\infty} z^i H_i$. Then,*

1. *$\varphi^{(k)}(z)$ and $\psi^{(k)}(z)$ are analytic for $z \in \mathcal{A}_k = \mathcal{A}\left(r^{2^k}, R^{2^k}\right)$;*
2. *there exists $k_0 > 0$ such that for any $k \geq k_0$ the matrix function $\psi^{(k)}(z)$ is nonsingular in $\mathcal{A}$ and the sequence $z^{-1}\varphi^{(k)}(z)$ converges to H_0^{-1} uniformly over all the compact sets $\mathcal{K} \subset \mathcal{A}$;*
3. *for any matrix norm $||\cdot||$ and for any ϵ there exist $k_0 > 0$ and positive constants c_i, $i = -1, 1, 2\ldots$ such that*

$$||A_i^{(k)}|| \leq \begin{cases} c_i(R-\epsilon)^{-i\cdot 2^k} & \textit{for } i > 0 \\ c_i(r+\epsilon)^{2^k} & \textit{for } i = -1 \end{cases}$$

for any $k \geq k_0$.

Proof Since $\varphi(z)$ is analytic and invertible in $\mathcal{A} = \mathcal{A}(r, R)$ then $\psi(z) = z\varphi(z)^{-1}$ is analytic and invertible in $\mathcal{A}$. By following the same argument used in

Theorem 1 we deduce part 1 and part 2 of the theorem. Part 3 follows from the Cauchy integral theorem [55, Theorem 4.4c] by following the same arguments used in the proof of Theorem 8.2 of [18]. □

Conditions for the nonsingularity of H_0, needed in the assumptions of Theorem 9, are given in the following result of [18]:

Theorem 10 *Let the matrix function $\varphi(z)$ be analytic for $|z| < R$, $R > 1$, and nonsingular for $|z| = 1$. If there exist the factorizations*

$$z^{-1}\varphi(z) = \left(\sum_{j=0}^{+\infty} z^j U_j\right)\left(I - z^{-1}G\right)$$
$$z\varphi(z^{-1}) = (I - zV)\left(\sum_{j=0}^{+\infty} z^{-j} W_j\right)$$

valid for $|z| = 1$, where $\rho(G), \rho(V) < 1$, and the matrix functions $\sum_{j=0}^{+\infty} z^j U_j$ and $\sum_{j=0}^{+\infty} z^j W_j$ are nonsingular for $|z| < 1$, then $\det H_0 \neq 0$.

The two factorizations in the assumptions of the above theorem hold if there exist solutions G and V to the matrix equations

$$\sum_{i=-1}^{\infty} A_i X^{i+1} = 0, \quad \sum_{i=-1}^{\infty} Y^{i+1} A_i = 0, \tag{34}$$

respectively, with $\rho(G) < 1$, $\rho(V) < 1$, and the function $\det\varphi(z)$ has exactly m roots inside the open unit disk. This condition is generally satisfied in many applicative problems which model stochastic processes described by infinite Markov chains [18].

4.3 Implementation

Relying on the functional formulation (32), (33), CR can be easily implemented even though in principle an infinite amount of block coefficients are needed in order to deal with matrix power series.

Observe that under the assumption of analyticity, all the matrix power series have matrix coefficients which decay exponentially to zero. This way, from the numerical point of view, a matrix power series can be viewed as a matrix polynomial by ignoring the coefficients which have modulus less than a given tolerance ε. This fact makes it easy to approximate the matrix coefficients by means of an evaluation/interpolation scheme which allows one to apply (32) and (33) pointwise at a given set of interpolatory nodes, say the roots of the unity, and to perform interpolation to these computed values in order to recover the coefficients.

This general scheme provides very efficient algorithms which have been described in details in [18] and [20].

5 Applications

Here we report on some applications of CR to problems from different fields and on acceleration techniques. In some cases, in particular for problems concerning queuing models, the applicability of CR is guaranteed. In other cases, like solving general banded Toeplitz systems, breakdown can be encountered. In principle, breakdown is possible also for problems involving queuing models where the acceleration technique of Section 5.7 is applied to improve the convergence. In these cases the formulae of Section 3.3 can be successfully employed.

5.1 Banded Toeplitz systems and polynomial factorization

Consider matrices $E_i \in \mathbb{R}^{m\times m}$, $i = -p, \ldots, p$ and define the $n \times n$ block banded block Toeplitz matrix T having $(2p+1)$ diagonals and blocks entries $T_{i,j} = E_{j-i} \in \mathbb{R}^{m\times m}$, for $|j-i| \le p$ and $T_{i,j} = 0$, otherwise, for $i = 1, \ldots, n$.

Observe that if n is a multiple of p, the matrix T can be reblocked into $mp \times mp$ blocks so that it can be viewed as a block tridiagonal block Toeplitz matrix $\mathrm{Trid}_{n/p}(B, A, C)$ where the blocks A, B, C are $p \times p$ block Toeplitz matrices with $m \times m$ blocks.

In this way, CR applied to $\mathrm{Trid}_{n/p}(B, A, C)$ can be used for solving a block banded block Toeplitz system. It is interesting to observe that for the Toeplitz structure of the blocks A, B, C, the matrix polynomial $\varphi(z)$ is a block z-circulant matrix [21]. Since block z-circulant matrices form an algebra, then also $\psi(z)$ is block z-circulant so that the matrix functions $\psi^{(k)}(z)$ are block Toeplitz. This fact enables one to prove that even though the matrices $A^{(k)}, B^{(k)}, C^{(k)}$ do not maintain the Toeplitz structure, their displacement rank remains constant since $\varphi^{(k)}(z)$ are inverses of block Toeplitz matrices and inverses of block Toeplitz matrices have the displacement structure [63]. We refer the reader to the book [63] by Kailath and Sayed for the concept of displacement rank. The property of maintaining the displacement structure is at the basis of a fast algorithm for solving banded Toeplitz system developed in [21].

Solving banded Toeplitz systems is strictly related to computing polynomial factorization and the Wiener-Hopf factorization of a given matrix Laurent series. Analysis, algorithms and application of CR to these problems can be found in [9–12, 28].

Variations of cyclic reduction for the case of matrices with unbalanced band have been analyzed in [66].

5.2 Quadratic Matrix Equations

Consider the quadratic matrix equation

$$B + AX + CX^2 = 0$$

where the matrix coefficients A, B, C are such that the hypotheses of Theorem 4 are satisfied so that there exists the solution X with minimal spectral radius $\rho(X) = |\xi_m| < 1$. This solution X is such that

$$\begin{bmatrix} A & C & & 0 \\ B & A & C & \\ 0 & \ddots & \ddots & \ddots \end{bmatrix} \begin{bmatrix} X \\ X^2 \\ X^3 \\ \vdots \end{bmatrix} = \begin{bmatrix} -B \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

so that CR can be used to compute X. More precisely, applying CR to the above system by using the even-odd permutation in place of the odd-even yields the sequence of systems

$$\begin{bmatrix} \widehat{A}^{(k)} & C^{(k)} & & 0 \\ B^{(k)} & A^{(k)} & C^{(k)} & \\ 0 & \ddots & \ddots & \ddots \end{bmatrix} \begin{bmatrix} X \\ X^{2^k+1} \\ X^{2\cdot 2^k+1} \\ \vdots \end{bmatrix} = \begin{bmatrix} -B \\ 0 \\ 0 \\ \vdots \end{bmatrix}$$

where

$$\widehat{A}^{(k+1)} = \widehat{A}^{(k)} - B^{(k)} \left(A^{(k)}\right)^{-1} C^{(k)}, \quad k \geq 0$$
$$\widehat{A}^{(0)} = A.$$

Observe that the first block equation of the latter system yields

$$\widehat{A}^{(k)} X = -B - C^{(k)} X^{2^k+1}.$$

Therefore, for the convergence properties of Theorem 3, since the spectral radius of X is $\rho(X) = |\xi_m| < 1$, then $\lim_k \widehat{A}^{(k)} X = -B$. Therefore, if the matrices $\left(\widehat{A}^{(k)}\right)^{-1}$ are uniformly bounded, since for any norm and for any $\epsilon > 0$ there exists $\theta > 0$ such that $||X^{2^k+1}|| \leq \theta(|\xi_m| + \epsilon)^{2^k}$ and $||C^{(k)}|| \leq \theta/(|\xi_{m+1}| - \epsilon)^{2^k}$, one has

$$X = -\left(\widehat{A}^{(k)}\right)^{-1} B + O\left(\left(\frac{|\xi_m| + \epsilon}{|\xi_{m+1}| - \epsilon}\right)^{2^k}\right).$$

This algorithm is the method of choice in solving Markov chains associated with Quasi-Birth-Death processes [18, 65]. Effective implementations can be found in [29, 30].

Applications of CR to the hyperbolic quadratic eigenvalue problem are given by Guo, Higham and Tisseur in [49] and by Guo and Lancaster in [51].

5.3 Polynomial and power series matrix equations

Given $m \times m$ matrices A_i, $i = -1, 0, 1, \ldots$, consider the matrix equation

$$A_{-1} + A_0 X + A_1 X^2 + A_2 X^3 + \cdots = 0 \tag{35}$$

where we assume that $\varphi(z) = \sum_{i=-1}^{+\infty} z^{i+1} A_i$ is analytic for $|z| < R$, $R > 1$, and $\det \varphi(z) \neq 0$ for $r < |z| < R$. If there exist solutions G and V to the matrix equations (34) such that $\rho(G), \rho(V) < 1$, and the function $\det \varphi(z)$ has exactly m roots of modulus less than 1, then the assumptions of Theorem 9 are satisfied and CR converges with r being the largest modulus root of $\det \varphi(z)$ in the open unit disk. Therefore, if CR has no breakdown, it can be applied for computing the solution G to the matrix equation (35) by computing the first block component of the solution of the following infinite linear system

$$\begin{bmatrix} A_0 & A_1 & A_2 & \dots \\ A_{-1} & A_0 & A_1 & \ddots \\ & A_{-1} & A_0 & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} X \\ X^2 \\ X^3 \\ \vdots \end{bmatrix} = \begin{bmatrix} -A_{-1} \\ 0 \\ \vdots \end{bmatrix}. \tag{36}$$

After k steps of CR with the even-odd permutation one obtains the system

$$\begin{bmatrix} \widehat{A}_0^{(k)} & \widehat{A}_1^{(k)} & \widehat{A}_2^{(k)} & \dots \\ A_{-1}^{(k)} & A_0^{(k)} & A_1^{(k)} & \ddots \\ & A_{-1}^{(k)} & A_0^{(k)} & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} X \\ X^{2^k+1} \\ X^{2\cdot 2^k+1} \\ \vdots \end{bmatrix} = \begin{bmatrix} -A_{-1} \\ 0 \\ \vdots \end{bmatrix}.$$

For the convergence properties of Theorem 9, since $\rho(G) = r$, and assuming that $\big(\widehat{A}_0^{(k)}\big)^{-1}$ is uniformly bounded, one has

$$X = -\big(\widehat{A}_0^{(k)}\big)^{-1} A_{-1} + O\left(\left(\frac{r+\epsilon}{R-\epsilon}\right)^{2^k}\right).$$

It is important to point out that this kind of matrix equations is fundamental in the solution of M/G/1-type and G/M/1-type Markov chains which model most part of queueing problems [70, 71]. Algorithms based on CR are the methods of choice for these problems [29, 30].

If the matrix equation is polynomial, i.e., if

$$A_{-1} + A_0 X + A_1 X^2 + A_2 X^3 + \dots + A_p X^p = 0$$

then the system (36) is banded and we may reblock it in order to find a block tridiagonal system as we did in Section 5.1. In this way applying CR in the original form, designed for block tridiagonal systems, provides a different solution algorithm. This approach is analyzed in [23].

A different approach which works even for infinite power series is due to Ramaswami [75]. It consists in transforming the system (36) into a block tridiagonal block Toeplitz system where the blocks have infinite size. In this way CR can be still applied in its original form by exploiting the specific structure of the infinite blocks. Details on the analysis of this approach can be found in [13, 26]. The analysis of CR for block tridiagonal systems with blocks of infinite size is performed in [13].

5.4 Matrix square root

The principal square root X of a matrix A can be effectively computed by means of CR as shown in [68]. The principal square root $A^{1/2}$ of a matrix A is the solution of the equation $X^2 - A = 0$ having eigenvalues with nonnegative real parts. The principal square root exists unique if A has no real negative eigenvalues [56]. This condition is assumed throughout this section.

Define $\varphi(z) = (I - A) + 2z(I + A) + z^2(I - A)$, then it has been proved in [68] that $z^{-1}\varphi(z)$ is invertible for any $z \in \mathbb{C}$ such that $r < |z| < 1/r$ where

$$r = \rho\left((A^{1/2} - I)(A^{1/2} + I)^{-1}\right) < 1.$$

Moreover, $H(z) = z\varphi(z)^{-1} = H_0 + \sum_{i=1}^{\infty} H_i(z^i + z^{-i})$ is such that $H_0 = \frac{1}{4}A^{-1/2}$. In this way, the desired solution $A^{1/2}$ can be directly obtained by computing the matrix H_0^{-1} by means of CR. More details in this regard can be found in [68].

A similar idea is used by Bini, Higham and Meini in [14] for computing the principal pth root of a matrix A.

Scaling techniques which improve the numerical stability of CR applied to the computation of the matrix square root have been designed and analyzed by Iannazzo in [60].

5.5 Algebraic Riccati equations

Consider the Nonsymmetric Algebraic Riccati Equation (NARE)

$$XCX - AX - XD + B = 0 \tag{37}$$

where the unknown X is an $m \times n$ matrix, and the coefficients A, B, C and D have sizes $m \times m$, $m \times n$, $n \times m$ and $n \times n$, respectively.

The matrix coefficients of the NARE (37) define the $(m + n) \times (m + n)$ matrix

$$M = \begin{bmatrix} D & -C \\ -B & A \end{bmatrix}, \tag{38}$$

which, throughout this section, we assume to be an M-matrix. Equations of this kind describe different models in the applications like fluid queues [76, 79, 98] and transport equations [61, 62]. The solution of interest is the matrix S with nonnegative entries which, among all the nonnegative solutions, is the one with component-wise minimal entries. A survey on this specific equation can be found in [16].

Efficient techniques for solving (37) have been recently designed in [25]. They are based on reducing the equation to a unilateral quadratic matrix equation of the kind $\mathcal{B}Y^2 + \mathcal{A}Y + \mathcal{C} = 0$ to which CR can be applied as described in Section 5.2.

Theorem 11 *X solves the NARE (37) iff*

$$Y = \begin{bmatrix} D - CX & 0 \\ X & 0 \end{bmatrix}$$

solves the matrix equation

$$\mathcal{B}Y^2 + \mathcal{A}Y + \mathcal{C} = 0 \tag{39}$$

where

$$\mathcal{A} = \begin{bmatrix} -I & -C \\ 0 & -A \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} 0 & 0 \\ 0 & -I \end{bmatrix}, \quad \mathcal{C} = \begin{bmatrix} D & 0 \\ B & 0 \end{bmatrix}.$$

The M-matrix property of (38) implies that the matrix coefficients $\mathcal{A}, \mathcal{B}, \mathcal{C}$ of (39) are such that the hypotheses of Theorem 4 are satisfied. Therefore CR applied to the matrix equation (39) converges.

The structure of the block coefficients of (39) allows one to implement cyclic reduction with just $\frac{68}{3}m^3$ operations per step, for $m = n$.

In [25] different reductions to a unilateral quadratic matrix equation are given. In one of these reductions, a function $\varphi(z) = \mathcal{B} + z\mathcal{A} + z^2\mathcal{C}$ is given, such that the constant coefficient H_0 of the matrix function $\psi(z) = z\varphi(z)^{-1}$ provides the sought solution to the Riccati equation. Therefore, CR provides a powerful tool for solving (37). Its cost is just $\frac{64}{3}m^3$ operations per step. Using this reduction it is possible to prove that the structure preserving doubling algorithm (SDA) of [6, 52] is just cyclic reduction applied to the unilateral quadratic matrix equation.

Variants of cyclic reduction have been applied to solve algebraic Riccati equations by Bini, Guo, Higham, Iannazzo, Latouche, Meini, Poloni, and Ramaswami in the papers [15, 47, 48, 50, 76].

5.6 Other matrix equations

Equations of the kind $X \pm A^* X^{-1} A - Q = 0$, studied in [39, 40] by Engwerda, Ferrante, Levy, Ran and Rijkeboer, can be efficiently solved by means of CR as shown by Meini in [67]. In fact, equations in this class can be reduced to unilateral quadratic matrix equations for which the convergence conditions of cyclic reduction are satisfied.

Other matrix functions like the geometric mean, the sign function or sector function, can be related to the coefficient H_0 of $z\varphi(z)^{-1}$ for a suitable quadratic matrix polynomial $\varphi(z)$. Therefore they can be computed by means of CR (Iannazzo, personal communication).

5.7 Acceleration

According to Theorems 3 and 9, the convergence of CR is quadratic and depends on the ratio r/R where $\mathcal{A}(r, R)$ is the domain where the function $\varphi(z)$ is analytic and invertible. The smaller is r/R the faster is convergence. We may improve convergence by enlarging the invertibility and analyticity domain of

$\varphi(z)$ by means of the shift technique introduced by He, Meini and Rhee in [53], extended to more general systems in [13, 28] and applied to algebraic Riccati equations in [15, 48], and [50]. The shift technique consists in moving one of the zeros of $\det\varphi(z)$ with modulus r or R to zero or to infinity, respectively, if the function is sufficiently regular. This goal can be obtained by replacing $\varphi(z)$ with $\widetilde{\varphi}(z) = \varphi(z)(I - z^{-1}\lambda \boldsymbol{u}\boldsymbol{v}^T)^{-1}$, where λ is a root of $\det\varphi(z)$ of modulus r, $\boldsymbol{v} \neq 0$ is a vector such that $\varphi(\lambda)\boldsymbol{v} = 0$, and $\boldsymbol{u}$ is any vector such that $\boldsymbol{v}^T\boldsymbol{u} = 1$. With this transformation, the roots of $\det\widetilde{\varphi}(z)$ are the roots of $\det\varphi(z)$, except that the root λ is replaced by 0. A similar formula is used for shifting the root of modulus R to infinity.

The effectiveness of this acceleration is more important in the case where $r = R$ where the quadratic convergence properties given in Theorems 3 and 9 do not hold anymore. In this difficult situation, typically encountered in *null recurrent* Markov chains, performing the shift of one of the two unwanted roots allows one to maintain the quadratic convergence of CR.

6 Conclusions

Cyclic Reduction, initially introduced to solving the discrete Poisson equation, has been extended to more general systems and to matrix equations. Its formulation given in terms of analytic functions provides a better understanding of its convergence properties.

New proofs of convergence have been given under weaker conditions. A complete characterization of breakdown situations has been given by providing alternative formulae for implementing CR in case of singularity or ill-conditioning.

Several applications to solving infinite systems and nonlinear matrix equations, including algebraic Riccati equations, have been given.

Acknowledgements The authors wish to thank the anonymous referees for their helpful comments and remarks which enabled the improvement of the presentation.

References

1. Amodio, P.: Optimized cyclic reduction for the solution of linear tridiagonal systems on parallel computers. Comput. Math. Appl. **26**(3), 45–53 (1993)
2. Amodio, P., Brugnano, L., Politi, T.: Parallel factorizations for tridiagonal matrices. SIAM J. Numer. Anal. **30**(3), 813–823 (1993)
3. Amodio, P., Mastronardi, N.: A parallel version of the cyclic reduction algorithm on a hypercube. Parallel Comput. **19**(11), 1273–1281 (1993)
4. Amodio, P., Mazzia, F.: Backward error analysis of cyclic reduction for the solution of tridiagonal systems. Math. Comp. **62**(206), 601–617 (1994)
5. Amodio, P., Paprzycki, M.: A cyclic reduction approach to the numerical solution of boundary value ODEs. SIAM J. Sci. Comput. **18**(1), 56–68 (1997) (Dedicated to C. William Gear on the occasion of his 60th birthday)
6. Anderson, B.D.O.: Second-order convergent algorithms for the steady-state Riccati equation. Int. J. Control **28**(2), 295–306 (1978)

7. Bialecki, B.: Cyclic reduction and FACR methods for piecewise Hermite bicubic orthogonal spline collocation. Numer. Algorithms **8**(2–4), 167–184 (1994)
8. Bini, D., Meini, B.: On the solution of a nonlinear matrix equation arising in queueing problems. SIAM J. Matrix Anal. Appl. **17**(4), 906–926 (1996)
9. Bini, D.A., Böttcher, A.: Polynomial factorization through Toeplitz matrix computations. Linear Algebra Appl. **366**, 25–37 (2003) (Special issue on structured matrices: analysis, algorithms and applications (Cortona, 2000))
10. Bini, D.A., Fiorentino, G., Gemignani, L., Meini, B.: Effective fast algorithms for polynomial spectral factorization. Numer. Algorithms **34**(2–4), 217–227 (2003) (International Conference on Numerical Algorithms, Vol. II (Marrakesh, 2001))
11. Bini, D.A., Gemignani, L., Meini, B.: Factorization of analytic functions by means of Koenig's theorem and Toeplitz computations. Numer. Math. **89**(1), 49–82 (2001)
12. Bini, D.A., Gemignani, L., Meini, B.: Computations with infinite Toeplitz matrices and polynomials. Linear Algebra Appl. **343/344**, 21–61 (2002) (Special issue on structured and infinite systems of linear equations)
13. Bini, D.A., Gemignani, L., Meini, B.: Solving certain matrix equations by means of Toeplitz computations: algorithms and applications. In: Fast Algorithms for Structured Matrices: Theory and Applications (South Hadley, MA, 2001), vol. 323 of Contemp. Math., pp. 151–167. American Mathematical Society, Providence (2003)
14. Bini, D.A., Higham, N.J., Meini, B.: Algorithms for the matrix pth root. Numer. Algorithms **39**(4), 349–378 (2005)
15. Bini, D.A., Iannazzo, B., Latouche, G., Meini, B.: On the solution of algebraic Riccati equations arising in fluid queues. Linear Algebra Appl. **413**(2–3), 474–494 (2006)
16. Bini, D.A., Iannazzo, B., Meini, B., Poloni, F.: Nonsymmetric algebraic Riccati equations associated with an M-matrix: recent advances and algorithms. In: Bini, D., Meini, B., Ramaswami, V., Remiche, M.-A., Taylor, P. (eds.) Numerical Methods for Structured Markov Chains, number 07461 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, (2008). Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl
17. Bini, D.A., Latouche, G., Meini, B.: Solving matrix polynomial equations arising in queueing problems. Linear Algebra Appl. **340**, 225–244 (2002)
18. Bini, D.A., Latouche, G., Meini, B.: Numerical Methods for Structured Markov Chains. Numerical Mathematics and Scientific Computation. Oxford University Press, New York (2005) (Oxford Science Publications)
19. Bini, D.A., Meini, B.: On cyclic reduction applied to a class of Toeplitz-like matrices arising in queueing problems. In: Stewart, W.J. (ed.) Computations with Markov Chains, pp. 21–38. Kluwer, Dordrecht (1995)
20. Bini, D.A., Meini, B.: Improved cyclic reduction for solving queueing problems. Numer. Algorithms **15**(1), 57–74 (1997)
21. Bini, D.A., Meini, B.: Effective methods for solving banded Toeplitz systems. SIAM J. Matrix Anal. Appl. **20**(3), 700–719 (1999) (electronic)
22. Bini, D.A., Meini, B.: Approximate displacement rank and applications. In: Structured Matrices in Mathematics, Computer Science, and Engineering, II (Boulder, CO, 1999), vol. 281 of Contemp. Math., pp. 215–232. American Mathematical Society, Providence (2001)
23. Bini, D.A., Meini, B.: Non-skip-free $M/G/1$-type Markov chains and Laurent matrix power series. Linear Algebra Appl. **386**, 187–206 (2004)
24. Bini, D.A., Meini B., Poloni, F.: Fast solution of a certain Riccati equation through Cauchy-like matrices. Electron. Trans. Numer. Anal. (2009). http://etna.mcs.kent.edu/
25. Bini, D.A., Meini, B., Poloni, F.: From algebraic Riccati equations to unilateral quadratic matrix equations: old and new algorithms. In: Bini, D., Meini, B., Ramaswami, V., Remiche, M.-A., Taylor, P. (eds.) Numerical Methods for Structured Markov Chains, number 07461 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, (2008). Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl
26. Bini, D.A., Meini, B., Ramaswami, V.: Analyzing M/G/1 paradigms through QBDs: the role of the block structure in computing the matrix G. In: Latouche, G., Taylor, P. (eds.) Advances in Algorithmic Methods for Stochastic Models, pp. 73–86. Notable Publications, Neshanic Station (2000) (Proceedings of the Third Conference on Matrix Analytic Methods)
27. Bini, D.A., Meini, B., Ramaswami, V.: A probabilistic interpretation of Cyclic Reduction and its relationships with Logarithmic Reduction. Calcolo **45**, 207–216 (2008)

28. Bini, D.A., Meini, B., Spitkovsky, I.M.: Shift techniques and canonical factorizations in the solution of $M/G/1$-type Markov chains. Stoch. Models **21**(2–3), 279–302 (2005)
29. Bini, D.A., Meini, B., Steffé, S., Van Houdt, B.: Structured Markov chains solver: algorithms. In: Proceedings of ValueTools, Pisa. ACM, New York (2006)
30. Bini, D.A., Meini, B., Steffé, S., Van Houdt, B.: Structured Markov chains solver: software tools. In: Proceedings of ValueTools, Pisa. ACM, New York (2006)
31. Boisvert, R.F.: Algorithms for special tridiagonal systems. SIAM J. Sci. Statist. Comput. **12**(2), 423–442 (1991)
32. Bondeli, S., Gander, W.: Cyclic reduction for special tridiagonal systems. SIAM J. Matrix Anal. Appl. **15**(1), 321–330 (1994)
33. Briggs, W.L., Turnbull, T.: Fast Poisson solvers for MIMD computers. Parallel Comput. **6**(3), 265–274 (1988)
34. Buneman, O.: A compact non-iterative Poisson solver. Technical Report SUIPR Rep. 294, Institue for Plasma Research, Stanford University, Paolo Alto, California, (1969)
35. Buzbee, B.L., Golub, G.H., Nielson, C.W.: On direct methods for solving Poisson's equations. SIAM J. Numer. Anal. **7**, 627–656 (1970)
36. Cottle, R.W.: Manifestations of the Schur complement. Linear Algebra Appl. **8**, 189–211 (1974)
37. Diamond, M.A., Ferreira, D.L.V.: On a cyclic reduction method for the solution of Poisson's equations. SIAM J. Numer. Anal. **13**(1), 54–70 (1976)
38. Dodson, D.S., Levin, S.A.: A tricyclic tridiagonal equation solver. SIAM J. Matrix Anal. Appl. **13**(4), 1246–1254 (1992)
39. Engwerda, J.C., Ran, A.C.M., Rijkeboer, A.L.: 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)
40. Ferrante, A., Levy, B.C.: Hermitian solutions of the equation $X = Q + N X^{-1} N^*$. Linear Algebra Appl. **247**, 359–373 (1996)
41. Gallopoulos, E., Saad, Y.: A parallel block cyclic reduction algorithm for the fast solution of elliptic equations. In: Supercomputing (Athens, 1987), vol. 297 of Lecture Notes in Comput. Sci., pp. 563–575. Springer, Berlin (1988)
42. Gallopoulos, E., Saad, Y.: A parallel block cyclic reduction algorithm for the fast solution of elliptic equations. Parallel Comput. **10**(2), 143–159 (1989)
43. Gander, W., Golub, G.H.: Cyclic reduction—history and applications. In: Scientific computing (Hong Kong, 1997), pp. 73–85. Springer, Singapore (1997)
44. Gautschi, W.: Orthogonal Polynomials: Computation and Approximation. Numerical Mathematics and Scientific Computation. Oxford University Press, New York (2004) (Oxford Science Publications)
45. Guo, C.-H.: Convergence analysis of the Latouche-Ramaswami algorithm for null recurrent quasi-birth-death processes. SIAM J. Matrix Anal. Appl. **23**(3), 744–760 (2001/2002) (electronic)
46. Guo, C.-H.: Comments on a shifted cyclic reduction algorithm for quasi-birth-death problems. SIAM J. Matrix Anal. Appl. **24**(4), 1161–1166 (2003) (electronic)
47. Guo, C.-H.: Efficient methods for solving a nonsymmetric algebraic Riccati equation arising in stochastic fluid models. J. Comput. Appl. Math. **192**(2), 353–373 (2006)
48. Guo, C.-H., Higham, N.J.: Iterative solution of a nonsymmetric algebraic Riccati equation. SIAM J. Matrix Anal. Appl. **29**(2), 396–412 (2007)
49. Guo, C.-H., Higham, N.J., Tisseur, F.: Detecting and solving hyperbolic quadratic eigenvalue problems. Technical Report MIMS EPrint: 2007.117, Manchester Institute for Mathematical Sciences, School of Mathematics (2007)
50. Guo, C.-H., Iannazzo, B., Meini, B.: On the doubling algorithm for a (shifted) nonsymmetric algebraic Riccati equation. SIAM J. Matrix Anal. Appl. **29**(4), 1083–1100 (2007)
51. Guo, C.-H., Lancaster, P.: Algorithms for hyperbolic quadratic eigenvalue problems. Math. Comp. **74**(252), 1777–1791 (2005) (electronic)
52. Guo, X.-X., Lin, W.-W., Xu, S.-F.: A structure-preserving doubling algorithm for nonsymmetric algebraic Riccati equation. Numer. Math. **103**(3), 393–412 (2006)
53. He, C., Meini, B., Rhee, N.H., Sohraby, K.: A quadratically convergent Bernoulli-like algorithm for solving matrix polynomial equations in Markov chains. Electron. Trans. Numer. Anal. **17**, 151–167 (2004) (electronic)

54. Heller, D.: Some aspects of the cyclic reduction algorithm for block tridiagonal linear systems. SIAM J. Numer. Anal. **13**(4), 484–496 (1976)
55. Henrici, P.: Applied and Computational Complex Analysis, vol. 1. Wiley Classics Library. Wiley, New York (1988) (Power series—integration—conformal mapping—location of zeros, Reprint of the 1974 original, A Wiley-Interscience Publication)
56. Higham, N.J.: Functions of Matrices: Theory and Computation. Society for Industrial and Applied Mathematics, Philadelphia (2008)
57. Ho, C.-T., Johnsson, S.L.: Optimizing tridiagonal solvers for alternating direction methods on Boolean cube multiprocessors. SIAM J. Sci. Statist. Comput. **11**(3), 563–592 (1990)
58. Hockney, R.W.: A fast direct solution of Poisson's equation using Fourier analysis. J. Assoc. Comput. Mach. **12**, 95–113 (1965)
59. Hockney, R.W., Jesshope, C.R.: Parallel Computers: Architecture, Programming and algorithms. Hilger, Bristol (1981)
60. Iannazzo, B.: A note on computing the matrix square root. Calcolo **40**(4), 273–283 (2003)
61. Juang, J.: Global existence and stability of solutions of matrix Riccati equations. J. Math. Anal. Appl. **258**(1), 1–12 (2001)
62. Juang, J., Lin, W.-W.: Nonsymmetric algebraic Riccati equations and Hamiltonian-like matrices. SIAM J. Matrix Anal. Appl. **20**(1), 228–243 (1999) (electronic)
63. Kailath, T., Sayed, A.H. (eds.) Fast Reliable Algorithms for Matrices with Structure. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (1999)
64. Latouche, G., Ramaswami, V.: A logarithmic reduction algorithm for quasi-birth-death processes. J. Appl. Probab. **30**(3), 650–674 (1993)
65. Latouche, G., Ramaswami, V.: Introduction to Matrix Analytic Methods in Stochastic Modeling. SIAM, Philadelphia (1999)
66. Lotti, G.: A note on the solution of not balanced banded Toeplitz systems. Numer. Linear Algebra Appl. **14**(8), 645–657 (2007)
67. Meini, B.: Efficient computation of the extreme solutions of $X + A^*X^{-1}A = Q$ and $X - A^*X^{-1}A = Q$. Math. Comp. **71**(239), 1189–1204 (2002) (electronic)
68. Meini, B.: The matrix square root from a new functional perspective: theoretical results and computational issues. SIAM J. Matrix Anal. Appl. **26**(2), 362–376 (2004/2005) (electronic)
69. Meini, B.: Nonlinear matrix equations and structured linear algebra. Linear Algebra Appl. **413**(2–3), 440–457 (2006)
70. Neuts, M.F.: Matrix-geometric Solutions in Stochastic Models, vol. 2 of Johns Hopkins Series in the Mathematical Sciences. Johns Hopkins University Press, Baltimore (1981) (An algorithmic approach)
71. Neuts, M.F.: Structured Stochastic Matrices of $M/G/1$ Type and Their Applications, vol. 5 of Probability: Pure and Applied. Marcel Dekker, New York (1989)
72. Ostrowski, A.: Recherches sur la méthode de Graeffe et les zéros des polynomes et des séries de Laurent. Chapitres III et IV. Acta Math. **72**, 157–257 (1940)
73. Ostrowski, A.: Addition à notre mémoire: "Recherches sur la méthode de Graeffe et les zéros des polynômes et des séries de Laurent." Acta Math. **75**, 183–186 (1943)
74. Pavlov, V., Yalamov, P.: Stabilization by perturbation of ill-conditioned cyclic reduction. Int. J. Comput. Math. **68**(3–4), 273–283 (1998)
75. Ramaswami, V.: The generality of Quasi Birth-and-Death processes. In: Alfa, A.S., Chakravarthy, S.R. (eds.) Advances in Matrix Analytic Methods for Stochastic Models, pp. 93–113. Notable Publications, Neshanic Station (1998)
76. Ramaswami, V.: Matrix analytic methods for stochastic fluid flows. In: Smith, D., Hey, P. (eds.) Teletraffic Engineering in a Competitive World, Proceedings of the 16th International Teletraffic Congress, pp. 1019–1030. Elsevier Science B.V., Edimburgh (1999)
77. Reichel, L.: The ordering of tridiagonal matrices in the cyclic reduction method for Poisson's equation. Numer. Math. **56**(2–3), 215–227 (1989)
78. Rodrigue, G., Wolitzer, D.: Preconditioning by incomplete block cyclic reduction. Math. Comp. **42**(166), 549–565 (1984)
79. Rogers, L.C.G.: Fluid models in queueing theory and Wiener-Hopf factorization of Markov chains. Ann. Appl. Probab. **4**(2), 390–413 (1994)
80. Rosmond, T.E., Faulkner, F.D.: Direct solution of elliptic equations by block cyclic reduction and factorization. Mon. Weather Rev. **104**, 641–649 (1976)

81. Rossi, T., Toivanen, J.: A nonstandard cyclic reduction method, its variants and stability. SIAM J. Matrix Anal. Appl. **20**(3), 628–645 (1999) (electronic)
82. Rossi, T., Toivanen, J.: A parallel fast direct solver for block tridiagonal systems with separable matrices of arbitrary dimension. SIAM J. Sci. Comput. **20**(5), 1778–1796 (1999) (electronic)
83. Schwandt, H.: Cyclic reduction for tridiagonal systems of equations with interval coefficients on vector computers. SIAM J. Numer. Anal. **26**(3), 661–680 (1989)
84. Schwandt, H.: Truncated interval arithmetic block cyclic reduction. Appl. Numer. Math. **5**(6), 495–527 (1989)
85. Sun, W.: Cyclic reduction algorithm for solving collocation systems. Int. J. Comput. Math. **61**(3–4), 293–305 (1996)
86. Swarztrauber, P.N.: A direct method for the discrete solution of separable elliptic equations. SIAM J. Numer. Anal. **11**, 1136–1150 (1974)
87. Swarztrauber, P.N.: The methods of cyclic reduction, Fourier analysis and the FACR algorithm for the discrete solution of Poisson's equation on a rectangle. SIAM Rev. **19**(3), 490–501 (1977)
88. Swarztrauber, P.N.: Fast Poisson solvers. In: Studies in numerical analysis, vol. 24 of MAA Stud. Math., pp. 319–370. Mathematical Association of America, Washington, DC (1984)
89. Swarztrauber, P.N.: Approximate cyclic reduction for solving Poisson's equation. SIAM J. Sci. Statist. Comput. **8**(3), 199–209 (1987)
90. Swarztrauber, P.N., Sweet, R.A.: Algorithm 541—Efficient Fortran subprograms for the solution of separable elliptic partial differential equations. ACM Trans. Math. Softw. **5**, 325–364 (1979)
91. Swarztrauber, P.N., Sweet, R.A.: Vector and parallel methods for the direct solution of Poisson's equation. J. Comput. Appl. Math. **27**(1–2), 241–263 (1989) (Reprinted in *Parallel algorithms for numerical linear algebra*, 241–263, North-Holland, Amsterdam, 1990)
92. Sweet, R.A.: A generalized cyclic reduction algorithm. SIAM J. Numer. Anal. **11**, 506–520 (1974)
93. Sweet, R.A.: A cyclic reduction algorithm for solving block tridiagonal systems of arbitrary dimension. SIAM J. Numer. Anal. **14**(4), 706–720 (1977)
94. Sweet, R.A.: A parallel and vector variant of the cyclic reduction algorithm. SIAM J. Sci. Statist. Comput. **9**(4), 761–765 (1988)
95. Temperton, C.: Direct methods for the solution of the discrete Poisson equation: some comparisons. J. Comput. Phys. **31**(1), 1–20 (1979)
96. Temperton, C.: On the FACR(l) algorithm for the discrete Poisson equation. J. Comput. Phys. **34**(3), 314–329 (1980)
97. Vandebril, R., Van Barel, M., Mastronardi, N.: Matrix computations and semiseparable matrices, vol. 1. Johns Hopkins University Press, Baltimore (2008) (Linear systems)
98. Williams, D.: A "potential-theoretic" note on the quadratic Wiener-Hopf equation for Q-matrices. In: Seminar on Probability, XVI, vol. 920 of Lecture Notes in Math., pp. 91–94. Springer, Berlin (1982)
99. Yalamov, P., Pavlov, V.: Stability of the block cyclic reduction. Linear Algebra Appl. **249**, 341–358 (1996)
100. Yalamov, P.Y.: On the stability of the cyclic reduction without back substitution for tridiagonal systems. BIT **35**(3), 428–447 (1995)
101. Ye, J., Li, S.-q.: Analysis of multi-media traffic queues with finite buffer and overload control—part 1: Algorithm, pp. 1464–1474. INFOCOM, Paris (1991)
102. Ye, J., Li, S.-q.: Analysis of multi-media traffic queues with finite buffer and overload control—part ii: Applications, pp. 848–859. INFOCOM, Paris (1992)