Computations with infinite Toeplitz matrices and polynomials

D. A. Bini, L. Gemignani, B. Meini

Dipartimento di Matematica, Università di Pisa, Italy

Abstract

We relate polynomial computations with operations involving infinite band Toeplitz matrices and show applications to the numerical solution of Markov chains, of nonlinear matrix equations, to spectral factorizations and to the solution of finite Toeplitz systems. In particular two matrix versions of Graeffe's iteration are introduced and their convergence properties are analyzed. Correlations between Graeffe's iteration for matrix polynomials and cyclic reduction for block Toeplitz matrices are pointed out. The paper contains a systematic treatment of known topics and presentation of new results, improvements and extensions.

Key words: Toeplitz matrices, polynomial computations, Graeffe's method, nonlinear matrix equations, matrix polynomials, cyclic reduction, spectral factorization.

1 Introduction

Let $\boldsymbol{a} = [a_i]_{i\in\mathbb{Z}}$ and $\boldsymbol{A} = [A_i]_{i\in\mathbb{Z}}$ be bi-infinite (block) column vectors where $a_i \in \mathbb{R}$ and $A_i \in \mathbb{R}^{m\times m}$. The (block) Toeplitz matrices generated by $\boldsymbol{a}$ and $\boldsymbol{A}$ are defined according to:

$$\begin{aligned}
&T_n[\boldsymbol{a}] = (a_{i-j})_{i,j=0,n-1}, \quad T_\infty[\boldsymbol{a}] = (a_{i-j})_{i,j\in\mathbb{N}}, \quad T[\boldsymbol{a}] = (a_{i-j})_{i,j\in\mathbb{Z}},\\
&T_n[\boldsymbol{A}] = (A_{i-j})_{i,j=0,n-1}, T_\infty[\boldsymbol{A}] = (A_{i-j})_{i,j\in\mathbb{N}}, T[\boldsymbol{A}] = (A_{i-j})_{i,j\in\mathbb{Z}}.
\end{aligned} \tag{1}$$

Observe that $T_n[\boldsymbol{a}] \in \mathbb{R}^{n\times n}$, $T_\infty[\boldsymbol{a}]$ is an infinite matrix while $T[\boldsymbol{a}]$ is bi-infinite.

Let J represent the reversion operator mapping the vector $\boldsymbol{a}$ onto the vector $J(\boldsymbol{a}) = [a_{-i}]_{i\in\mathbb{Z}}$ obtained by reverting the order of the entries of $\boldsymbol{a}$. For the sake of notational simplicity, also denote with J the reversion operator acting on bi-infinite block vectors so that $J(\boldsymbol{A}) = [A_{-i}]_{i\in\mathbb{Z}}$. The Toeplitz matrices associated with $J(\boldsymbol{a})$ and $J(\boldsymbol{A})$ are defined similarly with (1) by replacing the entries a_i and A_i with the corresponding entries a_{-i} and A_{-i}, respectively. In particular, with this notation, $T[\boldsymbol{u}]$, $T[\boldsymbol{l}]$ are upper and lower triangular matrices, respectively, where $u_i = l_{-i} = 0$ for $i > 0$; moreover it holds $T[J(\boldsymbol{a})] = T[\boldsymbol{a}]^T$ where A^T denotes the transpose of A. Finally, if r, s are nonnegative integers and $\boldsymbol{a} = [a_i]_{-r\le i\le s}$, $\boldsymbol{A} = [A_i]_{-r\le i\le s}$, then the Toeplitz matrices generated by $\boldsymbol{a}$, $J(\boldsymbol{a})$, $\boldsymbol{A}$ and $J(\boldsymbol{A})$ are still defined according to (1) where we implicitly extend $\boldsymbol{a}$ and $\boldsymbol{A}$ to bi-infinite (block) vectors by assuming $a_i = 0$ and $A_i = 0$ for $i < -r$ or $i > s$; in this case the associated (block) Toeplitz matrices are (block) banded.

Solving a system of the kind $T\boldsymbol{x} = \boldsymbol{b}$, where T is any of the matrices $T_n[\boldsymbol{a}]$, $T_\infty[\boldsymbol{a}]$, $T[\boldsymbol{a}]$, is encountered in many problems in pure and applied mathematics: in particular, in the numerical solution of difference and differential equations [31], in polynomial computations [25], in the numerical solution of Markov chains that arise in queueing models [20, 82, 81, 74], in the numerical solution of nonlinear matrix equations [78, 15, 16], in the analysis of cellular automata [9].

The interplay between Toeplitz matrices and polynomial computations, already pointed out in [25, 79, 40, 24, 11, 89], is a key feature that relates and integrates two fields like numerical linear algebra and computer algebra. In fact, almost all the fundamental polynomial computations like computing gcd, lcm, Euclidean scheme, polynomial division, polynomial multiplication, modular computations, Padé approximation, polynomial interpolation etc, have their own counterparts expressed in terms of structured (Toeplitz) matrix computations. This fact allows one to map algorithms for matrix computations into algorithms for polynomial computations and vice-versa, leading to synergies in both the polynomial and matrix frameworks. Moreover, techniques and tools from a well consolidated area like numerical linear algebra can be used for the design and analysis of algorithms for polynomial computations. A profound impact of these connections has been in the applicative area of systems and control, where the interplay between structured numerical linear algebra and polynomial computations has strongly influenced the ways in which problems are now being solved [46, 39].

Recently, based on such correlations, new effective algorithms have been designed for solving different important problems: tools for solving certain problems in queueing theory are presented in [15, 16, 17, 23, 20], while techniques for polynomial factorization based on Toeplitz computations are introduced in [12], nonlinear matrix equations are treated in [13, 15, 17, 16, 78], applications to image restoration are shown in [21, 10, 22, 18], optimal algorithms for banded Toeplitz systems are designed in [19]. Many results concerning spectral properties of band Toeplitz matrices together with a wide literature can be found in [27, 28, 29]. As a result of these achievements, it is worth pointing out that algorithms based on the Toeplitz matrix technology have lead to a reduction of the cpu time needed for the numerical solution of Markov chains modeling certain metropolitan networks problems by a factor of several hundreds with respect to the customarily used algorithms [2, 75, 77].

In this paper, some of these recent results are revisited in a unifying framework with a systematic treatment of theory and algorithms. We provide new theoretical results, improvement of algorithms and extensions of the set of applications.

In section 2 we state the main computational problems analyzed in the paper, described both in the polynomial and in the matrix framework. In particular we consider: the solution of (bi-)infinite banded Toeplitz systems; the computation of the "central" coefficients of the Laurent series $x(z)$ which solves the equation $a(z)x(z) = b(z)$, for given Laurent polynomials $a(z), b(z)$; the computation of the UL factorization of (bi-)infinite banded Toeplitz matrices and the solution of the equivalent problem of factoring a polynomial. Finally we consider the problem of solving certain nonlinear matrix equations and its structured matrix version, which consists in solving suitable (bi-)infinite block tridiagonal or block Hessenberg block Toeplitz systems.

The main technique that we introduce for the analysis of these computational problems is the Graeffe iteration [65, 83]. In section 3 we recall the main properties of this iteration, describe its matrix formulation expressed in terms of bi-infinite banded Toeplitz matrices and apply it for the design of efficient bi-infinite banded Toeplitz solvers. Since the extension of Graeffe's iteration to the case of infinite and of finite (block) banded Toeplitz systems is not straightforward, in section 4 we introduce a modification of this iteration which has a much wider set of applications, and which applies in particular to matrix polynomials of degree 2. We show that the matrix counterpart of this modified iteration is the cyclic reduction (CR) algorithm of [62, 14, 15, 16] and, based on this equivalence, we finally deduce applicability conditions and convergence properties of CR by using numerical linear algebra tools.

In section 5 the modified Graeffe iteration and its matrix version are extended to matrix power series and to (bi-)infinite (or to finite) block Hessenberg block Toeplitz matrices. In section 6 we consider the case where banded Toeplitz matrices are reblocked into block tridiagonal block Toeplitz matrices, by showing that the scalar "Toeplitzness" of the initial coefficient matrix leads to many relevant improvements of CR from both the theoretical and compational point of view. Specifically, by means of the displacement theory, we are able to provide a complete description of CR in terms of a small set of generator vectors, in such a way that each step can be performed by means of few FFT's.

In section 7 we apply the generalized Graeffe iteration and its matrix version CR for the design of effective algorithms for the solution of banded Toeplitz systems, nonlinear matrix equations, polynomial factorization problems, solution of resultant-like systems. In particular, concerning polynomial factorization, we improve the result of [12] by providing a simpler iteration which allows us to approximate the factors of a given polynomial with a lower computational cost.

2 Band Toeplitz matrices and Laurent polynomials

The application that associates $T[\boldsymbol{a}]$, $\boldsymbol{a} = [a_i]_{i=-r,s}$, with the Laurent polynomial

$$a(z) = \sum_{i=-r}^{s} a_i z^i \tag{2}$$

is clearly an isomorphism between the ring of infinite band Toeplitz matrices, with the operations of addition and row by column multiplication, and the ring of Laurent polynomials. Here we consider different computational problems which are interesting in themselves and have also a great importance in certain applications like queueing theory, computer algebra, system and control theory, signal processing and data modeling [81, 82, 74, 25, 38, 57, 91]. Almost each problem has its own polynomial formulation and its matrix version.

Throughout the paper, if not differently specified, $a(z)$ denotes the Laurent polynomial (2) of degree $n = \max\{r, s\}$, $\boldsymbol{a} = [a_i]_{i\in\mathbb{Z}}$ the associated bi-infinite vector where $a_i = 0$ if $i < -r$ or $i > s$, $b(z) = \sum_{i=-\widehat{r}}^{\widehat{s}} b_i z^i$ a Laurent polynomial of degree $\widehat{n} = \max\{\widehat{r}, \widehat{s}\}$, and $\boldsymbol{b} = [b_i]_{i\in\mathbb{Z}}$ the associated bi-infinite vector where $b_i = 0$ if $i < -\widehat{r}$ or $i > \widehat{s}$. Moreover, for any integer m we call $a_{-m}, \ldots, a_m$ the $2m+1$ *central coefficients* of the Laurent polynomial (2) or more generally of the Laurent power series $a(z) = \sum_{i=-\infty}^{+\infty} a_i z^i$.

Problem 1 [Computing the Laurent expansion of a rational function]. *Given a positive integer m, the coefficients of the Laurent polynomial $a(z)$, such that $a(z) \neq 0$ for $|z| = 1$, and the coefficients of $b(z)$, compute the $2m+1$ central coefficients of the Laurent series $x(z) = \sum_{i=-\infty}^{+\infty} x_i z^i$ such that $a(z)x(z) = b(z)$.*

The equivalent matrix version of Problem 1 is given below; it involves the solution of a linear system whose coefficient matrix is the bi-infinite Toeplitz matrix $T[\boldsymbol{a}]$. It is worth pointing out that the appropriate framework for studying infinite matrices is generally the theory of linear operators acting on Banach spaces of sequences. In particular, in the Toeplitz case we are mainly concerned with the Hilbert space $\ell^2(\mathbb{Z})$ of real square summable sequences $\boldsymbol{x} = [x_i]_{i\in\mathbb{Z}}$ with the norm

$$\| \boldsymbol{x} \|^2 = \sum_{i\in\mathbb{Z}} x_i^2.$$

The algebra of infinite matrix representations of bounded linear operators acting on $\ell^2(\mathbb{Z})$ provides a natural extension of the usual matrix algebra [69].

Problem 2 [Solving a bi-infinite band Toeplitz system]. *Compute the components $x_{-m}, \ldots, x_m$ of the vector $\boldsymbol{x} = [x_i]_{i\in\mathbb{Z}}$ solution in $\ell^2(\mathbb{Z})$ of the system $T[\boldsymbol{a}]\boldsymbol{x} = \boldsymbol{b}$, where $a(z) \neq 0$ for $|z| = 1$.*

A special case of Problems 1 and 2 is obtained with $b(z) = 1$ where the goal is the computation of the $2m+1$ central coefficients of the reciprocal of $a(z)$.

Observe that the condition $a(z) \neq 0$ for $|z| \neq 1$ insures that $T[\boldsymbol{a}]$ is an invertible linear operator on $\ell^2(\mathbb{Z})$ [29] and therefore $\boldsymbol{x}$, defined in Problem 2, belongs to $\ell^2(\mathbb{Z})$. Observe also that any vector $\boldsymbol{x}$ solving Problem 2, not necessarily in $\ell^2(\mathbb{Z})$, can be viewed as the solution of a linear difference equation with constant coefficients. Hence, $\boldsymbol{x}$ can be written as a specific solution of the non-homogeneous equation plus a solution of the homogeneous equation. Further, this latter component is a linear combination of the elementary solutions $\boldsymbol{x}_j = [\xi_j^i]_{i\in\mathbb{Z}}$, where ξ_j, $j = 1, \ldots, r+s+1$ are the zeros of $a(z^{-1})$ which, for the sake of simplicity, here are assumed to be distinct.

In this way, an interesting link between Problem 2 and the root-finding problem for $a(z)$ is easily established. Its deep manifestation is the theoretical and computational equivalence between the solution of a sequence of nested Toeplitz systems and the *qd* algorithm of Rutishauser (a description of such a relationship, aimed at numerical applications to factorization problems, can be found in [32]). Indeed the theoretical properties of solutions of large banded Toeplitz linear systems have been used by many authors in order to develop iterative schemes for the computation of zeros and poles of polynomials and rational functions (see, for instance, [6, 71, 45], and [70]). More recently, in [12] the authors make use of a similar approach to numerically approximate the coefficients of a polynomial factor of an analytic function.

Problem 3 [Factorization of a Laurent polynomial]. *Given the Laurent polynomial $a(z)$ compute two polynomials $u(z) = \sum_{i=0}^{r} u_{-i}z^i$ and $l(z) = \sum_{i=0}^{s} l_i z^i$ of degree r and s respectively, such that $a(z) = u(z^{-1})l(z)$.*

Observe that the factorization of the above problem is equivalent to factoring the polynomial $z^r a(z)$ as $u_R(z)l(z)$, where $u_R(z) = z^r u(z^{-1})$ is obtained by reverting the coefficients of $u(z)$. If the zeros of $l(z)$ and $u(z)$ have modulus greater than 1, then the factorization of Problem 3 is called *spectral factoriza-*

tion [57]. Factorizations where $u(z)$ and $l(z)$ have zeros with modulus greater than 1 are particularly meaningful in the solution of Markov chains of the M/G/1 type that model queueing problems [82, 48]. Moreover, the spectral factors play a key role in many diverse problems of data modeling, control theory and digital signal processing, where the primary focus is on the study of process dynamics. These problems also include time series analysis, Wiener filtering, noise variance estimation, covariance matrix computations and the study of multichannel systems (see [38, 37, 39, 91, 5, 4]).

The key role played by the spectral factorization of $a(z)$ is made clear by the following remarkable fact: the spectral factorization induces a very special triangular factorization of the corresponding bi-infinite banded Toeplitz matrix $T[\boldsymbol{a}]$. This allows us to provide a very interesting counterpart of Problem 3 in the framework of Toeplitz computations.

Problem 4 [UL factorization of a (bi-)infinite band Toeplitz matrix]. *Given* $T[\boldsymbol{a}]$ *compute two vectors* $\boldsymbol{u} = [u_{-r}, \ldots, u_0]^T$, $\boldsymbol{l} = [l_0, \ldots, l_s]^T$ *such that* $T[\boldsymbol{a}] = T[\boldsymbol{u}]T[\boldsymbol{l}]$*; equivalently, such that* $T_\infty[\boldsymbol{a}] = T_\infty[\boldsymbol{u}]T_\infty[\boldsymbol{l}]$.

Note that the solution of Problems 3 and 4 exists but is not unique. If $a(z)$ has r zeros inside and s zeros outside the unit disk, then the solution where $l(z)$ and $u(z)$ have roots of modulus greater than one is unique. In this case the solution of Problem 4 is usually referred to as the *Wiener-Hopf factorization* [29] and the triangular factors $T[\boldsymbol{l}]$ and $T[\boldsymbol{u}]$ are themselves invertible operators in $\ell^2(\mathbb{Z})$.

Observe that Problem 2 can be formulated also for finite matrices or for infinite matrices. In the first case it corresponds to solving a linear difference equation with boundary conditions. In the second case it corresponds to solving a linear difference equation with initial conditions. However, there is no evidence of a polynomial version of these two formulations of Problem 2. Similarly, Problem 4 may have no solution if formulated in terms of finite matrices since the product of an upper triangular and a lower triangular Toeplitz matrix is not generally Toeplitz.

Remark 5 [On the reduction to finite systems]. It is possible to reduce a certain instance of Problem 2 to a finite dimension still keeping valid the equivalent polynomial formulation of Problem 1. In fact, it is sufficient to reformulate Problem 1 modulo an assigned monic polynomial $\phi(z) = \sum_{i=0}^{n} \phi_i z^i$, $\phi_n = 1$, such that $\phi(0) \neq 0$ so that $z^{-1} \bmod \phi(z)$ exists and, moreover, $x(z)$ is defined on the zeros of $\phi(z)$. In this case, by setting $a_\phi(z) = a(z) \bmod \phi(z)$, $b_\phi(z) = b(z) \bmod \phi(z)$ and $x_\phi(z) = x(z) \bmod \phi(z)$, we find that $a_\phi(z)x_\phi(z) = b_\phi(z) \bmod \phi(z)$. In this way, if F_ϕ denotes the $n \times n$ Frobenius matrix associated with $\phi(z)$, i.e., the matrix $F_\phi = (f_{i,j})_{i,j=1,n}$ such that $f_{i+1,i} = 1$, $f_{i,n} = -\phi_{i-1}$, $f_{i,j} = 0$ elsewhere, then the evaluation of the coefficients of $x_\phi(z)$ turns into solving the equation $\sum_{i=0}^{n-1} a_{\phi,i} F_\phi^i \sum_{i=0}^{n-1} x_{\phi,i} F_\phi^i = \sum_{i=0}^{n-1} b_{\phi,i} F_\phi^i$. Without loss of generality we may restrict the above equation to the first column of both members and obtain $(\sum_{i=0}^{n-1} a_{\phi,i} F_\phi^i)\boldsymbol{x}_\phi = \boldsymbol{b}_\phi$. In particular, if $\phi(z) = z^n - 1$ the matrix in the left hand side of the above equation is a band circulant matrix [36] with band width n. Hence, fast schemes based on FFT's can be used in order to calculate $\boldsymbol{x}_\phi$.

Remark 6 [Extension to bivariate polynomials]. Observe that Problem 1 can be generalized in terms of bivariate Laurent polynomials of the kind $a(z, w) = \sum_{i=-r}^{s} \sum_{j=-r}^{s} a_{i,j} z^i w^j$. The matrix counterpart of this generalization involves a two-level band Toeplitz matrix A, where A is a bi-infinite block band block Toeplitz matrix and where the blocks are bi-infinite band Toeplitz matrices. By operating modulo $\phi(z)$ and modulo $\psi(w)$, where $\phi(z)$ and $\psi(w)$ are monic polynomials of degree n such that $\psi(0) \neq 0$ and $\phi(0) \neq 0$, we obtain a matrix version of the problem with the linear system defined by a matrix of the form $\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \tilde{a}_{i,j} F_\phi^i \otimes F_\psi^j$, where F_ϕ and F_ψ are the Frobenius matrices associated with the polynomials ϕ, ψ, respectively, and $\otimes$ denotes the Kronecker product. In particular for $\phi(z) = \psi(z) = z^n - 1$ we obtain a 2-level circulant matrix.

Another kind of generalization can be obtained by assuming that the components of the vectors $\boldsymbol{a}$, $\boldsymbol{x}$ and $\boldsymbol{b}$ are matrices. In this case the polynomials $a(z)$, $b(z)$, $x(z)$ turn into *matrix polynomials* [55]. A problem involving matrix polynomials that can be viewed as an extension of Problems 1 and 2 is formulated below. Here, for the sake of simplicity, matrix operations are to be considered by a formal point of view and we omit to give conditions under which the considered infinite and bi-infinite matrices define bounded invertible operators on a certain Banach space. This is because these assumptions are generally tailored for the special applications where diverse instances of Problem 7 below arise and, therefore, we prefer to postpone their description in the section devoted to the applicability of the proposed solution methods.

Problem 7 [Solving block tridiagonal block Toeplitz systems]. *Given* $\boldsymbol{A} = [A_i]_{i\in\mathbb{Z}}$*,* $A_i \in \mathbb{R}^{m\times m}$ *and a block vector* $\boldsymbol{b} = (\boldsymbol{b}_i)_{i\in\mathbb{Z}}$*,* $\boldsymbol{b}_i \in \mathbb{R}^m$*, solve the block tridiagonal system* $T[\boldsymbol{A}]\boldsymbol{x} = \boldsymbol{b}$*, where* $A_i = 0$ *if* $i < -1$ *or* $i > 1$*,* $\boldsymbol{x} = (\boldsymbol{x}_i)_{i\in\mathbb{Z}}$*,* $\boldsymbol{x}_i \in \mathbb{R}^m$*. More generally, solve the upper block Hessenberg system* $T[\boldsymbol{A}]\boldsymbol{x} = \boldsymbol{b}$*, where* $A_i = 0$ *for* $i > 1$*.*

We associate with the above problem the following one:

Problem 8 [Solving nonlinear matrix equations]. *Given the matrices* A_{-1}*,* A_0*,* A_1*, compute an* $m \times m$ *matrix* G *which solves the equation*

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

More generally, given the infinite sequence of $m \times m$ matrices $\{A_i\}_{i \geq -1}$ compute an $m \times m$ matrix G which solves the equation

$$\sum_{i=-1}^{+\infty} A_i X^{i+1} = 0. \tag{4}$$

Conditions under which matrix equations (3) and (4) have a solution are provided in section 7.2.

Observe that Problem 7 is strictly related to Problem 2. In fact, by re-blocking the infinite matrix $T[\boldsymbol{a}]$ of Problem 2 into $m \times m$ blocks, where $m = \max\{r, s\}$, we obtain a block tridiagonal block Toeplitz matrix $T[\boldsymbol{A}]$, $\boldsymbol{A} = [A_i]_{i \in \mathbb{Z}}$, where the blocks $A_{-1} = [a_{i-j-m}]_{i,j=1,m}$, $A_0 = [a_{i-j}]_{i,j=1,m}$, $A_1 = [a_{m+i-j}]_{i,j=1,m}$ are themselves Toeplitz matrices and $A_i = 0$ for $i < -1$ or $i > 1$. In this way, Problem 2 can be reformulated as a specific instance of Problem 7.

Problems 7 and 8 are related each other since if G is any nonsingular solution of (3) or of (4) then, for any vector $\boldsymbol{u} \in \mathbb{R}^m$, the block vector $\boldsymbol{x} = (\boldsymbol{x}_i)_{i \in \mathbb{Z}}$, $\boldsymbol{x}_i = G^i \boldsymbol{u}$ solves Problem 7 in the form $T[J(\boldsymbol{A})]\boldsymbol{x} = \boldsymbol{0}$, where $J(\boldsymbol{A}) = [A_{-i}]_{i \in \mathbb{Z}}$. Problem 8 is particularly interesting in the numerical solution of Markov chains modeling QBD processes, since from the solution G of the matrix equation (3) we may easily recover the solution of the infinite block tridiagonal Toeplitz-like system $\boldsymbol{\pi}^T(T_\infty[J(\boldsymbol{A})] + E) = \boldsymbol{0}^T$, where $\boldsymbol{A} = [A_i]_{i \in \mathbb{Z}}$, $A_i = 0$ for $i < -1$ or $i > 1$, and E is a suitable infinite matrix having null entries except in its $m \times m$ leading principal submatrix (see [81, 74]). Indeed, for the solution $\boldsymbol{\pi}^T = [\boldsymbol{\pi}_i^T]_{i \geq 0}$, where $\boldsymbol{\pi}_i$ are m-dimensional vectors, that represents the stationary probability distribution, we have $\boldsymbol{\pi}_i^T = \boldsymbol{\pi}_0^T R^i$, $i \geq 1$, where R is a suitable matrix that can be computed by means of G [81]. A similar property also holds for M/G/1 type Markov chains [82], where the stationary probability vector $\boldsymbol{\pi}^T = [\boldsymbol{\pi}_i^T]_{i \geq 0}$ solves the block Hessenberg system $\boldsymbol{\pi}^T(T_\infty[J(\boldsymbol{A})] + E) = \boldsymbol{0}^T$, where E is an infinite matrix with null entries except for its first m rows and $\boldsymbol{A} = [A_i]_{i \in \mathbb{Z}}$, $A_i = 0$ for $i < -1$. In this case the solution $\boldsymbol{\pi}$ can be easily expressed in terms of the solution G of the equation (4) by means of Ramaswami's formula [82, 85].

Problem 8 is also intimately related to Problems 3 and 4. Indeed, it is easy to show by a direct inspection (see also [82, 76]) that if G solves (3) then the matrix $T_\infty[A_1, A_0, A_{-1}]$ can be factorized as

$$\begin{bmatrix} A_0 & A_1 & & \\ A_{-1} & A_0 & A_1 & \\ & \ddots & \ddots & \ddots \end{bmatrix} = \begin{bmatrix} A_0 + A_1 G & A_1 & & 0 \\ & A_0 + A_1 G & A_1 & \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} I & & 0 \\ -G & I & \\ 0 & \ddots & \ddots \end{bmatrix}$$

and the matrix polynomial factorization holds

$$A_{-1} + A_0 z + A_1 z^2 = (A_0 + A_1 G + A_1 z)(Iz - G).$$

Similarly, if G solves (4) then, at least formally, we have

$$T_\infty[J(\boldsymbol{A})] = \begin{bmatrix} B_0 & B_1 & B_2 & \dots \\ & B_0 & B_1 & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} I & & 0 \\ -G & I & \\ 0 & \ddots & \ddots \end{bmatrix}$$

where $B_i = \sum_{j=i}^{\infty} A_j G^{j-i}$, $i = 0, 1, \ldots$, $\boldsymbol{A} = [A_i]_{i \in \mathbb{Z}}$, $A_i = 0$ for $i < -1$. In this way it can be easily seen that the following matrix power series factorization holds

$$\sum_{i=-1}^{+\infty} A_i z^{i+1} = \left(\sum_{i=0}^{+\infty} B_i z^i\right)(Iz - G).$$

3 Graeffe's iteration and its matrix version

Let $p(z) = \sum_{i=0}^{m} a_i z^i$ be a polynomial of degree m, with zeros ξ_i, $i = 1, \ldots, m$, ordered such that $|\xi_1| \le \ldots \le |\xi_m|$. Consider the sequence of polynomials $p_j(z)$ generated by

$$p_{j+1}(z^2) = p_j(z) p_j(-z), \quad j = 0, 1, \ldots \tag{5}$$

with $p_0(z) = p(z)$. Observe that $\xi_i^{2^j}$, $i = 1, \ldots, m$, are the zeros of $p_j(z)$. This is the fundamental property that makes the sequence $p_j(z)$ useful in many polynomial computations [86, 34, 26, 84]. The iteration (5), known as *Graeffe's iteration*, has been introduced by Dandelin, Lobachevski and Graeffe [83, 65], and has been widely exploited in the design of algorithms for the approximation of polynomial zeros [44, 84, 86].

Observe that, if $p(z)$ has h zeros of modulus less than 1 and $m - h$ zeros of modulus greater than 1 then h zeros of $p_j(z)$ tend to zero and $m - h$ zeros tend to infinity so that the normalized polynomial $p_j(z)/a_h^{(j)}$ tends to z^h, where $a_i^{(j)}$ is the coefficient of z^i in $p_j(z)$. Moreover, since the coefficients of $p_j(z)$ are the symmetric functions of the zeros of $p_j(z)$, it follows that $|a_i^{(j)}/a_h^{(j)}|$ tends to zero as $O(1/|\xi_{h+1}|^{2^j})$ if $i > h$, and as $O(|\xi_h|^{2^j})$ if $i < h$.

Graeffe's iteration provides a useful tool for the efficient solution of Problems 1 and 2. In this section we describe the application of this iteration for solving

Problem 1, then provide its matrix version (for solving Problem 2) which, to our knowledge, has never been pointed out in the literature except for [9] where it has been applied in a modular form; finally we discuss on its application to finite matrices.

Let $n = \max\{r, s\}$, $\hat{n} = \max\{\hat{r}, \hat{s}\}$ be the degrees of the Laurent polynomials $a(z)$ and $b(z)$, respectively. Let $m > n$, $\widehat{m} = m + n$ and observe that, once the Laurent polynomial $u_{\widehat{m}}(z)$ made up by the $2\widehat{m} + 1$ central coefficients of the Laurent series $a(z)^{-1}$ has been computed, the Laurent polynomial $x_m(z)$ made up by the $2m + 1$ central coefficients of the solution $x(z)$ of Problem 1 is obtained from $u_{\widehat{m}}(z)b(z)$. Therefore, we may restrict our attention to the computation of $u_{\widehat{m}}(z)$ and consider the equation

$$a(z)u(z) = 1. \tag{6}$$

For any polynomial or power series $a(z)$ denote $a_+(z)$ and $a_-(z)$ the even and the odd part of $a(z)$, respectively, such that $a_+(z^2) = \frac{1}{2}(a(z) + a(-z))$, and $a_-(z^2) = \frac{1}{2z}(a(z) - a(-z))$. Multiplying both members of (6) by $a(-z)$ and rewriting $u(z) = u_+(z^2) + zu_-(z^2)$, $a(z) = a_+(z^2) + za_-(z^2)$, we reduce (6) to $a_1(z)u_+(z) = a_+(z)$, $a_1(z)u_-(z) = -a_-(z)$, where $a_1(z^2) = a(z)a(-z)$, $a_-(z)$ and $a_+(z)$ are Laurent polynomials of degree at most $n/2$, and $a_1(z)$ has degree n. That is, the problem $\mathcal{P}(\widehat{m}, n)$ of computing the $2\widehat{m} + 1$ central coefficients of the inverse of a Laurent polynomial of degree $n = \max\{r, s\}$ is reduced to $\mathcal{P}((\widehat{m} + n)/2, n)$ and to computing the products of two polynomials of degree $\widehat{m} + n/2$ and a polynomial of degree $n/2$. Denoting $C(m, n)$ the arithmetic cost of $\mathcal{P}(m, n)$ and recalling that the product of a polynomial of degree m and a polynomial of degree n, $m > n$, can be computed in $O(m \log n)$ ops by means of FFT [25], we have the recurrence $C(m, n) = C((m + n)/2, n) + O(m \log n)$. Observe that after j steps of this iteration the moduli of the coefficients of $a_j(z)/a_h^{(j)} - z^h$ are $O(|\xi_h|^{2^j} + 1/|\xi_{h+1}|^{2^j})$ therefore $a_j(z)/a_h^{(j)}$ can be easily approximated with the polynomial z^h. The overall cost needed to solve problem $\mathcal{P}(m, n)$ with error at most $O(\epsilon)$ is thus $O(m \log n \log\log \epsilon^{-1})$.

A modular version of the above algorithm has been implicitly used in [9] where Problem 1 is considered modulo $z^k - 1$ and k is an integer power of 2. In this case, the equation (6) modulo $z^k - 1$ is reduced to $a_1(z)x_+(z) = a_+(z) \bmod z^{k/2} - 1$, $a_1(z)x_-(z) = -a_-(z) \bmod z^{k/2} - 1$. Thus, after $\log_2 k$ steps we obtain equations modulo $z - 1$ where polynomials turn into constants. Therefore the algorithm enables the computation with no approximation errors. In [87], Graeffe's iteration is used for reducing power series inversion to polynomial multiplication.

The matrix version of Graeffe's iteration for the solution of Problem 2 can be easily obtained from the isomorphism between bi-infinite band Toeplitz matrices and Laurent polynomials. More precisely, if we denote $\hat{\boldsymbol{a}} = [(-1)^i a_i]$, we find that $T[\hat{\boldsymbol{a}}]T[\boldsymbol{a}]$ is a band Toeplitz matrix with null entries in position (i,j) if $i-j$ is odd. Hence, by permuting its rows and columns according to the odd-even permutation, we obtain a 2×2 block diagonal matrix with diagonal blocks $T[\boldsymbol{a}_1]$ where $\boldsymbol{a}_1$ is the coefficient vector of the Laurent polynomial $a_1(z)$, $a_1(z^2)=a(-z)a(z)$. That is, in matrix form we have

$$\Pi\ T[\hat{\boldsymbol{a}}]T[\boldsymbol{a}]\ \Pi^T = \begin{bmatrix} T[\boldsymbol{a}_1] & 0 \\ 0 & T[\boldsymbol{a}_1] \end{bmatrix},$$

where Π denotes the permutation matrix associated with the odd-even permutation. The algorithm for computing the inverse of a bi-infinite band Toeplitz matrix can be easily deduced from the Graeffe iteration and the above arguments.

Except for the modular case where the computation is performed modulo z^k-1, the reduction of the above technique to infinite and to finite band Toeplitz matrices does not work. In fact, due to the finiteness of the size, multiplying $T_n[\boldsymbol{a}]$ by $T_n[\hat{\boldsymbol{a}}]$, or $T_\infty[\boldsymbol{a}]$ by $T_\infty[\hat{\boldsymbol{a}}]$ introduces some "garbage" in the corner(s) of the matrix product that destroys the striped diagonal structure. Another drawback of this approach is that it relies on the commutative property of the product so that it cannot be applied to block matrices, or to matrix polynomials, where the entries of the matrices and the coefficients of the involved polynomials are matrices. A way for overcoming these limitations is described in the next section.

4 Graeffe's iteration for matrix polynomials and its matrix version: the case of polynomials of degree 2.

In this section we provide a different formulation of Graeffe's iteration, introduced in [15, 16] and in [73], that does not use commutativity, can be extended to matrix polynomials and can be applied to infinite and finite matrices. This iteration is strictly related to the cyclic reduction algorithm of [33, 62], to LU factorization and to the Schur complementation, from which it has inherited the numerical stability features that have made this technique an effective tool for the solution of problems in queueing theory [14, 2, 8, 23].

Let us consider the case of a matrix polynomial $\phi(z)=A_{-1}+A_0z+A_1z^2$ of degree 2. A way of getting rid of the commutativity property is to "normalize" $\phi(z)$ by multiplying it on the left by A_0^{-1} (where we assume A_0 nonsingular). In this way the Graeffe iteration (5) takes the following form $\widehat{\phi}_j(z) = B_0^{(j)^{-1}}\widetilde{\phi}_j(z)$, $\widetilde{\phi}_{j+1}(z^2)=\widehat{\phi}_j(-z)\widehat{\phi}_j(z)$, $j=0,1,\ldots,$ where the sequence of matrix polynomials $\widetilde{\phi}_j(z) = B_{-1}^{(j)} + B_0^{(j)} z + B_1^{(j)} z^2$ is generated by starting with $\widetilde{\phi}_0(z) = A_{-1} + A_0 z + A_1 z^2$ and where we assume that $B_0^{(j)}$ is nonsingular [73, 15, 80]. In fact, it can be easily verified that the terms of odd degree cancel since the coefficient of z in $\widehat{\phi}_j(z)$ is the identity matrix.

This iteration is the basis of the Logarithmic Reduction algorithm derived by G. Latouche and V. Ramaswami in [73] with probabilistic arguments, for the numerical solution of QBD Markov chains [81, 74].

A simpler version of the non-commutative formulation of Graeffe's iteration can be obtained by performing the normalization step in a slightly different way:

$$\phi_{j+1}(z^2) = -\phi_j(-z) A_0^{(j)^{-1}} \phi_j(z), \quad j = 0, 1, \ldots \tag{7}$$

where the sequence of matrix polynomials $\phi_j(z) = A_{-1}^{(j)} + A_0^{(j)} z + A_1^{(j)} z^2$ is generated starting with $\phi_0(z) = A_{-1} + A_0 z + A_1 z^2$ and where we assume that $A_0^{(j)}$ is nonsingular. It can be easily verified that also in this case the cancellation of the odd terms holds, thus obtaining the recurrences

$$\begin{cases} A_{-1}^{(j+1)} = -A_{-1}^{(j)} A_0^{(j)^{-1}} A_{-1}^{(j)} \\ A_0^{(j+1)} = A_0^{(j)} - A_{-1}^{(j)} A_0^{(j)^{-1}} A_1^{(j)} - A_1^{(j)} A_0^{(j)^{-1}} A_{-1}^{(j)} \\ A_1^{(j+1)} = -A_1^{(j)} A_0^{(j)^{-1}} A_1^{(j)}, \quad j \geq 0, \end{cases} \tag{8}$$

which define the coefficients of the matrix polynomials obtained by the Graeffe iteration (7).

By using an inductive argument we may easily relate the matrix coefficients $A_i^{(j)}$ with $B_i^{(j)}$, $i = -1, 0, 1$, by means of $B_i^{(j)} = -A_0^{(j-1)^{-1}} A_i^{(j)}, \;\; i = -1, 0, 1, \;\; j \geq 1$. Observe also that if $A_{-1} = A_1^T$, $A_0 = A_0^T$, then for any j it holds $A_{-1}^{(j)} = A_1^{(j)T}$, $A_0^{(j)} = A_0^{(j)T}$.

An immediate consequence of (7) is that the polynomials $p_j(z) = \det \phi_j(z)$ satisfy the normalized Graeffe recurrence $p_{j+1}(z^2) = p_j(z) p_j(-z) \alpha_j, \quad \alpha_j = -1/\det A_0^{(j)}$, from which we deduce the following:

Theorem 9 *If the matrices $A_0^{(j)}$ are nonsingular for any $j = 0, 1, \ldots$, i.e., if the matrices $A_i^{(j)}$, $i = -1, 0, 1$, can be computed without any break-down, then the zeros of the polynomial $p_j(z) = \det \phi_j(z)$ are given by $\xi_1^{2^j}, \ldots, \xi_{2m}^{2^j}$, where $\xi_1, \ldots, \xi_{2m}$ are the zeros of $p(z) = \det \phi(z)$, completed with $2m - k$ zeros at the infinity if $p(z)$ has degree $k < 2m$ (i.e., if A_1 is singular).*

Another interesting relation that we obtain from (7) is that for any complex z for which the matrices $\phi_j(z)$ and $\phi_j(-z)$ are nonsingular, it holds

$$\phi_{j+1}(z^2) = \left(\frac{\phi_j(z)^{-1} - \phi_j(-z)^{-1}}{2z} \right)^{-1}. \tag{9}$$

Thus, if we introduce the matrix functions $\psi_j(z) = \phi_j(z)^{-1}$, defined for the values of z such that $\det \phi_j(z) \neq 0$, we have

$$\psi_{j+1}(z^2) = \frac{\psi_j(z) - \psi_j(-z)}{2z}. \tag{10}$$

This relation will be used in section 4 for proving applicability conditions of Graeffe iteration and in section 6 to show structural properties of the blocks $A_i^{(j)}$ [21].

The Graeffe iteration in its non-commutative version (7) can be reformulated in matrix form in terms of block LU factorization by means of a cyclic reduction step. In order to see this, let $\boldsymbol{A} = [A_i]_{i\in\mathbb{Z}}$, where $A_i \in \mathbb{R}^{m\times m}$, $A_i = 0$ for $i < -1$ or $i > 1$, and consider the block tridiagonal Toeplitz matrix $T[J(\boldsymbol{A})] = T[A_1, A_0, A_{-1}]$ (throughout this section, in order to be consistent with the polynomial notations of (7), we will deal with the matrix $T[J(\boldsymbol{A})]$ instead of $T[\boldsymbol{A}]$). Apply the even-odd permutation to block rows and block columns of $T[J(\boldsymbol{A})]$ by means of the even-odd permutation matrix Π and obtain

$$\Pi T[J(\boldsymbol{A})]\Pi^T = \begin{bmatrix} I\otimes A_0 & I\otimes A_1 + Z\otimes A_{-1} \\ I\otimes A_{-1} + Z^T\otimes A_1 & I\otimes A_0 \end{bmatrix}$$

where I is the bi-infinite identity matrix and Z is the bi-infinite matrix having unit entries in position $(i+1, i)$ and zeros elsewhere.

Applying one step of Gaussian elimination to the above matrix in order to reduce to zero the south-west block, yields the Schur complement

$$I\otimes A_0 - (I\otimes A_{-1} + Z^T\otimes A_1)(I\otimes A_0^{-1})(I\otimes A_1 + Z\otimes A_{-1})$$

which coincides with $T[A_1^{(1)}, A_0^{(1)}, A_{-1}^{(1)}]$ where $A_i^{(1)}$, $i = -1, 0, 1$, are defined by the Graeffe iteration (7), (8). In other words, the blocks $A_i^{(1)}$, $i = -1, 0, 1$, coincide with suitable submatrices of the Schur complement of a submatrix of an infinite block tridiagonal block Toeplitz matrix. Applying again the same technique to the new matrix $T[A_1^{(1)}, A_0^{(1)}, A_{-1}^{(1)}]$ yields the sequence $T[A_1^{(j)}, A_0^{(j)}, A_{-1}^{(j)}]$

of block tridiagonal block Toeplitz matrices whose blocks are the coefficients of the matrix polynomials obtained by means of the Graeffe iteration (7), (8).

The technique of applying the even-odd permutation followed by one step of block Gaussian elimination, called Cyclic Reduction (CR), was introduced at the end of the sixties for solving certain linear systems which discretize elliptic equations [33, 62]. We will refer to the Graeffe iteration for quadratic matrix polynomials as to CR.

Remark 10 [On CR applied to infinite and finite matrices]. CR can be applied also to the infinite matrix $T_\infty[A_1, A_0, A_{-1}]$. If we use the even-odd permutation the equations (8) which relate $A_i^{(j)}$, $i = -1, 0, 1$, at two subsequent steps of CR are unchanged. On the other hand if we apply CR with the odd-even permutation, we obtain a sequence of block tridiagonal matrices of the form

$$\begin{bmatrix} \widehat{A}^{(j)} & A_1^{(j)} & \\ A_{-1}^{(j)} & A_0^{(j)} & \ddots \\ & \ddots & \ddots \end{bmatrix}$$

where the block Toeplitz structure is kept almost everywhere except for the block $\widehat{A}^{(j)}$ in position $(1, 1)$. Moreover, the equations (8) still apply, complemented with the following one

$$\widehat{A}^{(j+1)} = \widehat{A}^{(j)} - A_1^{(j)} A_0^{(j)^{-1}} A_{-1}^{(j)}. \tag{11}$$

The CR technique can be also applied to the finite matrix $T_n[A_1, A_0, A_{-1}]$. Depending on the parity of n and on the kind of permutation (odd-even or even-odd), the sequence of block tridiagonal matrices generated by CR may loose the Toeplitz structure also for the block $\widetilde{A}^{(j)}$ in position (n, n) where we have to introduce a new updating equation. For instance, if $n = 2^k$ and we apply the even-odd permutation, then the equations (8) must be complemented with $\widetilde{A}^{(j+1)} = \widetilde{A}^{(j)} - A_{-1}^{(j)} A_0^{(j)^{-1}} A_1^{(j)}$. If $n = 2^k$ and we apply the odd-even permutation, then (8), complemented with (11) fully characterize CR. If $n = 2^k - 1$ and we apply the even-odd permutation, then (8) are sufficient and the block Toeplitz structure is preserved at all the CR steps.

Let us now start with stating some conditions under which the matrices $A_0^{(j)}$ are nonsingular, so that CR can be carried out with no break-down.

In the view of formulae (8) and (11), the study of the applicability of CR (Graeffe's iteration) is reduced to determine conditions under which the matrices $A_0^{(j)}$, $j \ge 0$, are nonsingular. Moreover there are no differences between applying the even-odd and the odd-even permutation since the blocks $A_i^{(j)}$, $i = -1, 0, 1$, generated with the two permutations are the same (compare Remark 10). In addition there are no differences among the finite, infinite and bi-infinite case. In fact, it is easily observed that the block matrix obtained at the j-th step of CR applied to $T_{2^k}[A_1, A_0, A_{-1}]$ coincides with the $2^{k-j} \times 2^{k-j}$ block leading principal submatrix of the infinite matrix $T[A_1^{(j)}, A_0^{(j)}, A_{-1}^{(j)}]$ obtained at the j-th step of cyclic reduction, applied to the infinite matrix $T[A_1, A_0, A_{-1}]$. In the light of Remark 10, since $T[A_1^{(j+1)}, A_0^{(j+1)}, A_{-1}^{(j+1)}]$ can be viewed as the Schur complement of a suitable principal submatrix of $T[A_1^{(j)}, A_0^{(j)}, A_{-1}^{(j)}]$, we may apply well known tools of numerical linear algebra and easily deduce applicability conditions of CR. For instance, it is well known that if an $n \times n$ matrix $V = (v_{i,j})$ is diagonally dominant, i.e., $|v_{i,i}| > \sum_{j=1,n,\ j\neq i} |v_{i,j}|$ for $i = 1, 2, \ldots, n$, then any Schur complement in V is diagonally dominant, and thus nonsingular [56]. This fact, together with the property that $T_n[A_1, A_0, A_{-1}]$ is diagonally dominant for any $n > 2$ if and only if $T_3[A_1, A_0, A_{-1}]$ is diagonally dominant, implies the following result:

Theorem 11 *If A_{-1}, A_0, A_1 are such that $T_3[A_1, A_0, A_{-1}]$ is diagonally dominant, then* $\det A_0^{(j)} \neq 0$ *for any j.*

A different applicability condition, expressed in terms of matrix norm and derived as a particular case of a general result of [92], is expressed by the following:

Theorem 12 *If* $\det A_0 \neq 0$ *and there exist a matrix norm $|| \cdot ||$ and a real number $0 < \alpha < 1$ such that $||A_0^{-1}A_{-1}|| + ||A_0^{-1}A_1|| \le \alpha$ then* $\det A_0^{(j)} \neq 0$ *for any j and, moreover, it holds $||A_0^{(j)^{-1}} A_{-1}^{(j)}|| + ||A_0^{(j)^{-1}} A_1^{(j)}|| \le \alpha^{2^j}$.*

The following result is slightly less immediate to prove.

Theorem 13 *Let the matrices $A_0^{(i)}$, for $i = 0, \ldots, j-1$, be nonsingular. Then the $(j+1)$-st step of cyclic reduction can be performed, i.e., $A_0^{(j)}$ is nonsingular, if and only if $T_{2^{j+1}-1}[A_1, A_0, A_{-1}]$ is nonsingular.*

PROOF. Proceed by induction. If $j = 1$, apply the even-odd permutation of block rows and columns of $T_3[A_1, A_0, A_{-1}]$ and obtain the matrix

$$\left[\begin{array}{cc|c} A_0 & 0 & A_1 \\ 0 & A_0 & A_{-1} \\ \hline A_{-1} & A_1 & A_0 \end{array}\right] = \begin{bmatrix} D_2[A_0] & L_1[A_1, A_{-1}] \\ U_1[A_{-1}, A_1] & D_1[A_0] \end{bmatrix}$$

where $D_k[B]$ is the $k \times k$ block diagonal matrix with diagonal blocks equal to B, $L_k[C, A]$ is the $(k+1) \times k$ block lower bidiagonal matrix having C on the main diagonal and A on the lower diagonal, $U_k[C, A]$ is the $k \times (k+1)$ block upper bidiagonal matrices having A on the main diagonal and C on the upper diagonal. By applying one step of block Gaussian elimination to the above permuted matrix we obtain that $A_0^{(1)} = D_1[A_0] - U_1[A_{-1}, A_1]D_2[A_0^{-1}]L_1[A_1, A_{-1}]$. From the properties of the Schur complement it follows that if A_0 is nonsingular, then $A_0^{(1)}$ is nonsingular if and only if $T_3[A_1, A_0, A_{-1}]$ is nonsingular. Now, let us suppose that the theorem holds for $j - 1$, and show it for j; applying the even-odd permutation of block rows and columns to the matrix $T_{2^{j+1}-1}[A_1, A_0, A_{-1}]$, yields

$$\begin{bmatrix} D_{2^j}[A_0] & L_{2^j-1}[A_1, A_{-1}] \\ U_{2^j-1}[A_{-1}, A_1] & D_{2^j-1}[A_0] \end{bmatrix}.$$

After performing one step of Gaussian elimination we obtain the matrix

$$T_{2^j-1}[A_1^{(1)}, A_0^{(1)}, A_{-1}^{(1)}] = D_{2^j-1}[A_0] - U_{2^j-1}[A_{-1}, A_1]D_{2^j}[A_0^{-1}]L_{2^j-1}[A_1, A_{-1}].$$

Since A_0 is nonsingular, $T_{2^j-1}[A_1^{(1)}, A_0^{(1)}, A_{-1}^{(1)}]$ is nonsingular if and only if $T_{2^{j+1}-1}[A_1, A_0, A_{-1}]$ is nonsingular. From the inductive hypothesis, assuming $A_0^{(i)}$, $i = 1, \ldots, j-1$, nonsingular, then $T_{2^j-1}[A_1^{(1)}, A_0^{(1)}, A_{-1}^{(1)}]$ is nonsingular if and only if the j-th step of cyclic reduction can be performed, starting with blocks $A_{-1}^{(1)}$, $A_0^{(1)}$, $A_1^{(1)}$, i.e., if and only if $A_0^{(j)}$ is nonsingular. $\square$

If A_{-1}, A_0, A_1 are such that $T_{2^i-1}[A_1, A_0, A_{-1}]$, $i = 1, \ldots, j$, is symmetric positive definite then from the above theorem it follows in particular that $\det A_0^{(i)} \neq 0$ for $i = 0, \ldots, j-1$. Moreover, the spectral condition number of $A_0^{(i)}$ does not exceed the spectral condition number of $T_{2^j-1}[A_1, A_0, A_{-1}]$.

For the matrices $\widehat{A}^{(j)}$ generated as in (11) by CR applied to both finite and infinite matrices, by using the same arguments of the proof of Theorem 13, we can give the following condition:

Theorem 14 *Let the matrices $A_0^{(i)}$, for $i = 0, \dots, j$, be nonsingular. Then the matrix $\widehat{A}^{(j)}$, defined in (11), is nonsingular if and only if $T_{2^{j+1}}[A_1, A_0, A_{-1}]$ is nonsingular.*

Observe that the subset of $(\mathbb{R}^{m\times m})^3$ made up by the triples A_{-1}, A_0, A_1 such that CR breaks down for some j is the numerable union of algebraic varieties of dimension lower than m^6 and has null measure in $(\mathbb{R}^{m\times m})^3$. This means that a random choice of A_{-1}, A_0 and A_1 leads to break down with probability zero. Moreover, for any "singular" triple A_{-1}, A_0, A_1 there exists a neighborhood $\mathcal{U}$ of zero such that the triple A_{-1}, $A_0 + \epsilon I$, A_1 does not lead to break down in CR for any $\epsilon \in \mathcal{U} \setminus \{0\}$.

In the symmetric case, where $A_0 = A_0^T$ and $A_{-1} = A_1^T$, the following sufficient condition for the applicability can be given:

Theorem 15 [On the applicability of CR for symmetric matrices] *Assume that $A_0 = A_0^T$, $A_{-1} = A_1^T$ and define, for any complex number z with $|z| = 1$, the Hermitian matrix $\gamma(z) = z^{-1}\phi(z) = A_{-1}z^{-1} + A_0 + A_1 z$. If $\gamma(z)$ is positive definite for any z, then for any $j \geq 1$ the matrices $A_0^{(j-1)}$ and $\gamma_j(z) = z^{-1}\phi_j(z)$, $|z| = 1$, are Hermitian positive definite. Moreover, we have $\mathrm{cond}_2(A_0^{(j)}) \leq \mu_{\max}/\mu_{\min}$, where $\mathrm{cond}_2(\cdot)$ denotes the spectral condition number, $\mu_{\max} = \max_{|z|=1} \max \lambda(\gamma(z))$, $\mu_{\min} = \min_{|z|=1} \min \lambda(\gamma(z))$, and $\lambda(\gamma(z))$ is the set of the eigenvalues of $\gamma(z)$.*

PROOF. Let us proceed by induction. Since $\gamma(z)$, $|z| = 1$, is positive definite, then $A_0 = (\gamma(z) + \gamma(-z))/2$ is positive definite. Now, from (7), since $\phi_0(z) = z\gamma(z)$ is nonsingular for $|z| = 1$ and A_0 is nonsingular, we deduce that $\phi_1(z)$ is nonsingular for $|z| = 1$, and hence $\gamma_1(z) = z^{-1}\phi_1(z)$ is nonsingular for $|z| = 1$. Moreover $\gamma_1(z)^{-1} = z\psi_1(z)$ and we may apply (10), thus obtaining that $\gamma_1(z^2)^{-1} = z^2\psi_1(z^2) = (\gamma(z)^{-1} + \gamma(-z)^{-1})/2$. Hence, from the positive definiteness of $\gamma(z)^{-1}$ we conclude that $\gamma_1(z)^{-1}$, and thus $\gamma_1(z)$ is positive definite. For the inductive step, by using the same argument as before, from the positive definiteness of $\gamma_j(z)$, deduce that $A_0^{(j)}$ is positive definite and $\phi_j(z)$ is nonsingular for $|z| = 1$, and thus, from (7), that $\gamma_{j+1}(z)$ is nonsingular for $|z| = 1$. Therefore we may apply (10) and deduce that $\gamma_{j+1}(z^2)^{-1} = (\gamma_j(z)^{-1} + \gamma_j(-z)^{-1})/2$, and thus $\gamma_{j+1}(z)^{-1}$ is positive definite. Concerning the condition number, observe that from the relation between $\gamma_{j+1}(z^2)^{-1}$ and $\gamma_j(z)^{-1}$ we obtain that $\max_{|z|=1} \max \lambda(\gamma_{j+1}(z)^{-1}) \leq \max_{|z|=1} \max \lambda(\gamma_j(z)^{-1})$ and similarly $\min_{|z|=1} \min \lambda(\gamma_{j+1}(z)^{-1}) \geq \min_{|z|=1} \min \lambda(\gamma_j(z)^{-1})$. Therefore, from $A_0^{(j)} = (\gamma_j(z) + \gamma_j(-z))/2$ we deduce that $\max \lambda(A_0^{(j)}) \leq \mu_{\max}$ and $\min \lambda(A_0^{(j)}) \geq \mu_{\min}$, whence $\mathrm{cond}_2(A_0^{(j)}) \leq \mu_{\max}/\mu_{\min}$. $\square$

As by-product of the previous proof, we also obtain that CR can be carried out under the weaker conditions where $\gamma(z)$ is Hermitian nonnegative definite for $|z| = 1$ and there exists ξ, $|\xi| = 1$, such that $\gamma(\xi)$ is positive definite. Under this weaker hypothesis the boundness of the condition number of $A_0^{(j)}$ is not guaranteed.

4.2 Convergence properties of Graeffe's iteration

Under suitable conditions the cyclic reduction technique has very nice convergence properties. In fact, as we will see later on, it is possible to prove that for a wide class of problems it holds $\lim_j A_{-1}^{(j)} = 0$ and $\lim_j A_1^{(j)} = 0$. Convergence properties of this kind have been investigated in the literature under diverse assumptions [92, 15, 19, 22, 18]. Here we report the main results [18, 22].

Theorem 16 [On the convergence of CR]. *Let $A_i^{(j)}$, $i = -1, 0, 1$, $j \geq 0$, be the matrices generated by the CR algorithm (8) applied to $T[A_1, A_0, A_{-1}]$, where $\det A_0^{(j)} \neq 0$ so that CR can be carried out. Assume that the quadratic matrix equations $A_{-1} + A_0X + A_1X^2 = 0$ and $A_{-1}Y^2 + A_0Y + A_1 = 0$ have solutions X and Y with spectral radius $\rho(X) < 1$, $\rho(Y) < 1$, respectively. Then for any matrix norm $||\cdot||$, the sequences $||A_0^{(j)}||$, $||A_0^{(j)^{-1}}||$ are bounded from above by a constant and for any $\epsilon > 0$ it holds $||A_{-1}^{(j)}|| = O((\epsilon + \rho(X))^{2^j})$, $||A_1^{(j)}|| = O((\epsilon + \rho(Y))^{2^j})$.*

PROOF. By using an induction argument, it can be shown that at each step j of CR the following equations hold:

$$\begin{aligned} &A_{-1}^{(j)} + A_0^{(j)}X^{2^j} + A_1^{(j)}X^{2\cdot 2^j} = 0, \\ &A_{-1}^{(j)}Y^{2\cdot 2^j} + A_0^{(j)}Y^{2^j} + A_1^{(j)} = 0, \end{aligned} \tag{12}$$

and therefore

$$(A_0^{(j)})^{-1}A_{-1}^{(j)} + X^{2^j} + (A_0^{(j)})^{-1}A_1^{(j)}X^{2\cdot 2^j} = 0, \tag{13}$$

$$(A_0^{(j)})^{-1}A_1^{(j)} + Y^{2^j} + (A_0^{(j)})^{-1}A_{-1}^{(j)}Y^{2\cdot 2^j} = 0. \tag{14}$$

We first prove that the matrices $C_{-1}^{(j)} = (A_0^{(j)})^{-1}A_{-1}^{(j)}$ and $C_1^{(j)} = (A_0^{(j)})^{-1}A_1^{(j)}$ are bounded in norm. Given $\epsilon > 0$ such that $\rho(X)+\epsilon < 1$ and $\rho(Y)+\epsilon < 1$, let $||\cdot||_X$, $||\cdot||_Y$ be matrix norms such that $||X||_X \leq \rho(X)+\epsilon$ and $||Y||_Y \leq \rho(Y)+\epsilon$ (for the exisstence of such norms see [56]). Let $\alpha_j = ||C_{-1}^{(j)}||_X$. If α_j is not bounded, then there exists a subsequence α_{j_h} which diverges to infinity. From (13) we have

$$\alpha_{j_h} \le ||C_1^{(j_h)}||_X ||X^{2\cdot 2^{j_h}}||_X + ||X^{2^{j_h}}||_X \le$$
$$(\rho(X)+\epsilon)^{2^{j_h}} \left((\rho(X)+\epsilon)^{2^{j_h}} ||C_1^{(j_h)}||_X + 1 \right).$$

Thus,

$$||C_1^{(j_h)}||_X \ge \left(\frac{\alpha_{j_h}}{(\rho(X)+\epsilon)^{2^{j_h}}} - 1 \right) \frac{1}{(\rho(X)+\epsilon)^{2^{j_h}}}$$

and, since α_{j_h} diverges to infinity and $\rho(X)+\epsilon$, there exists a constant $c > 0$ such that

$$||C_1^{(j_h)}||_X \ge c \frac{\alpha_{j_h}}{(\rho(X)+\epsilon)^{2^{j_h}}}.$$

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

$$||C_1^{(j_h)}||_Y \ge c' \frac{\alpha_{j_h}}{(\rho(X)+\epsilon)^{2^{j_h}}}, \tag{15}$$
$$||C_{-1}^{(j)}||_Y \le c'' ||C_{-1}^{(j)}||_X = c'' \alpha_j. \tag{16}$$

On the other hand, from (14) we have $||C_1^{(j_h)}||_Y \le ||C_{-1}^{(j_h)}||_Y ||Y^{2\cdot 2^{j_h}}||_Y + ||Y^{2^{j_h}}||_Y$. Thus, from the latter inequality and (16) we have $||C_1^{(j_h)}||_Y \le c''\alpha_{j_h}(\rho(Y)+\epsilon)^{2\cdot 2^{j_h}} + (\rho(Y)+\epsilon)^{2^{j_h}}$. Hence, from (15), we obtain

$$c' \frac{\alpha_{j_h}}{(\rho(X)+\epsilon)^{2^{j_h}}} \le c''\alpha_{j_h}(\rho(Y)+\epsilon)^{2\cdot 2^{j_h}} + (\rho(Y)+\epsilon)^{2^{j_h}},$$

that contradicts the assumption that α_{j_h} goes to infinity. By using a similar argument we can prove that also $C_1^{(j)}$ is bounded in norm. From (13), we obtain $||C_{-1}^{(j)}||_X \le (\rho(X)+\epsilon)^{2^j} + ||C_1^{(j)}||_X(\rho(X)+\epsilon)^{2\cdot 2^j}$. Hence, since $||C_1^{(j)}||_X$ is bounded, then there exists a constant $\gamma > 0$ such that $||C_{-1}^{(j)}||_X \le \gamma(\rho(X)+\epsilon)^{2^j}$. Thus, for the equivalence of the matrix norms, for any matrix norm $||\cdot||$ there exists a constant θ such that

$$||C_{-1}^{(j)}|| \le \theta(\rho(X)+\epsilon)^{2^j}. \tag{17}$$

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

$$||C_1^{(j)}|| \le \gamma'(\rho(Y)+\epsilon)^{2^j}. \tag{18}$$

From (17) and (18), since $A_0^{(j+1)} = A_0^{(j)}(I - C_{-1}^{(j)}C_1^{(j)} - C_1^{(j)}C_{-1}^{(j)})$, we have $||A_0^{(j+1)}|| \le ||A_0^{(j)}||(1+\sigma_j)$, where $\sigma_j = O\left((\rho(X)+\epsilon)^{2^j}(\rho(Y)+\epsilon)^{2^j}\right)$. Thus the matrices $A_0^{(j)}$, $j \ge 0$ are bounded in norm. Similarly, it holds

$$||(A_0^{(j+1)})^{-1}|| \le \frac{1}{1-\sigma_j}||(A_0^{(j)})^{-1}||,$$

thus $||(A_0^{(j)})^{-1}||$ is bounded. Now, from the boundness of $||A_0^{(j)}||$, and from the relations (12) we can show that $||A_1^{(j)}||$ and $||A_{-1}^{(j)}||$ are bounded, and thus that $||A_{-1}^{(j)}|| = O\left((\rho(X)+\epsilon)^{2^j}\right)$, $||A_1^{(j)}|| = O\left((\rho(Y)+\epsilon)^{2^j}\right)$. □

Conditions for the existence of the solutions X, Y of the matrix equations in the above theorem are given in section 7. Observe that the eigenvalues of X and the reciprocal of the eigenvalues of Y are zeros of the polynomial $p(z) = \det\phi(z)$. Denoting ξ_i, $i = 1,\dots,2m$ these zeros ordered such that $|\xi_1| \le |\xi_2| \le \dots \le |\xi_{2m}|$, and completed with zeros at infinity if $p(z)$ has degree less than $2m$, if $\rho(X) < 1$ and $\rho(Y) < 1$ then we have $|\xi_m| < 1 < |\xi_{m+1}|$ and $\rho(X) = |\xi_m|$, $\rho(Y) = |\xi_{m+1}|^{-1}$.

Remark 17 It is worth pointing out that the convergence to zero of the non-diagonal blocks holds even if the matrix $T[A_1, A_0, A_{-1}]$ is not diagonally dominant. Moreover, the necessary condition $|\xi_m| < 1 < |\xi_{m+1}|$ can be relaxed into $|\xi_m| < |\xi_{m+1}|$. In fact, if $|\xi_m| < |\xi_{m+1}|$, and θ is such that $\theta|\xi_m| < 1 < \theta|\xi_{m+1}|$, we may scale the variable z by θ so that the matrix equations $A_{-1} + \theta^{-1}A_0X + \theta^{-2}A_1X^2 = 0$, $\theta^2A_{-1}Y^2 + \theta A_0Y + A_1 = 0$, have solutions θX and $\theta^{-1}Y$, respectively, that satisfy the conditions of Theorem 16. Furthermore, if $|\xi_m| < |\xi_{m+1}|$ and $|\xi_m| = 1$ ($|\xi_{m+1}| = 1$), then the spectral radius of X (Y, respectively) is equal to one, and the convergence to zero of $A_1^{(j)}$ ($A_{-1}^{(j)}$) can be established.

If the matrix $T[A_1, A_0, A_{-1}]$ is real symmetric we have the following result [78, 22]:

Theorem 18 [On the convergence of CR for symmetric matrices]. *Assume that $A_0 = A_0^T$, $A_{-1} = A_1^T$ and that $\gamma(z) = z^{-1}\phi(z) = A_{-1}z^{-1} + A_0 + A_1z$ is positive definite for any complex number z with $|z| = 1$. Then the polynomial $p(z) = \det\phi(z)$ has zeros ξ_i, $i = 1,\dots,2m$ (completed with zeros at the infinity if A_1 is singular), such that $|\xi_1| \le \dots \le |\xi_m| < 1$, $\xi_{m+i} = 1/\xi_{m-i+1}$. Moreover, the blocks $A_1^{(j)}$, $A_0^{(j)}$, $A_0^{(j)^{-1}}$ have norm bounded from above by a constant and $||A_1^{(j)}|| = ||A_{-1}^{(j)T}|| = O((\epsilon + |\xi_m|)^{2^j})$ for any $\epsilon > 0$ and for any matrix norm $||\cdot||$.* □

Remark 19 The previous result holds also for the matrices generated by means of (11). In fact, since $\widehat{A}^{(j)}$ is a Schur complement of a suitable principal submatrix of $T_\infty[A_1, A_0, A_{-1}]$, it is invertible and uniformly bounded. The blocks $\widehat{A}^{(j)}$ intervene in the CR applied to infinite systems with the even-odd permutation (see Remark 10).

5 Graeffe's iteration for matrix polynomials and its matrix version: the general case

In this section we show that Graeffe's iteration, that we have described in the previous section for matrix polynomials of degree 2, can be formulated also for matrix polynomials of larger degree, or for matrix power series.

Consider the matrix power series $\phi(z) = \sum_{i=-1}^{\infty} A_i z^{i+1}$, where A_i, $i = -1, 0, \ldots$ are $m \times m$ matrices, and $A_i = 0$ for $i > n$ if $\phi(z)$ is a matrix polynomial of degree n.

Recurrence (7), that expresses Graeffe's iteration for matrix polynomials of degree 2, can be extended in the following way:

$$\phi_{j+1}(z^2) = -\phi_j(z)\left(\frac{\phi_j(z) - \phi_j(-z)}{2z}\right)^{-1}\phi_j(-z), \quad j \geq 0. \tag{19}$$

Indeed (19) reduces to (7) in the case where $A_i = 0$ for $i > 1$. Moreover, it can be easily checked that relations (9) and (10) still hold in the more general case where $\phi(z) = \sum_{i=-1}^{\infty} A_i z^{i+1}$.

A generalization of formulae (8) can be expressed in terms of operations between matrix power series. Indeed, we have [15]

$$\begin{cases} \phi_{j+1}(z) = z\phi_{j,-}(z) - \phi_{j,+}(z)\phi_{j,-}(z)^{-1}\phi_{j,+}(z), \\ \phi_j(z) = \sum_{i=-1}^{\infty} A_i^{(j)} z^{i+1}, \end{cases} \tag{20}$$

where, for a power series $f(z)$, we define $f_-(z)$, $f_+(z)$ as

$$f_-(z^2) = \frac{f(z) - f(-z)}{2z}, \quad f_+(z^2) = \frac{f(z) + f(-z)}{2}.$$

Functional relation (20), that reduces to (8) in the case of matrix polynomials of degree 2, allows the efficient computation of the coefficients of $\phi_{j+1}(z)$ by means of evaluation/interpolation at suitable Fourier points. More specifically, assuming that $\phi_{j+1}(z)$, $\phi_j(z)$ are convergent in the closed unit disk, the first

n block coefficients of $\phi_{j+1}(z)$ can be approximated within the error bound ϵ in $O(m^3n + m^2n\log n)$ ops, where n is such that $\sum_{i=n+1}^{+\infty} ||A_i^{(j+1)}||_\infty < \epsilon$. We refer to [16] for details on this subject.

Recurrences (19) and (20) can be reformulated in terms of operations between block Toeplitz matrices. Indeed, let $\boldsymbol{A} = [A_i]_{i\in\mathbb{Z}}$, $A_i = 0$ for $i < -1$, consider the block Toeplitz matrix $T[J(\boldsymbol{A})]$, and apply the even-odd permutation of block rows and columns, thus obtaining

$$\Pi T[J(\boldsymbol{A})]\Pi^T = \begin{bmatrix} U & V \\ W & U \end{bmatrix},$$

where $U = (A_{2(j-i)})_{i,j\in\mathbb{Z}}$, $V = (A_{2(j-i)+1})_{i,j\in\mathbb{Z}}$, $W = (A_{2(j-i)-1})_{i,j\in\mathbb{Z}}$, assuming $A_i = 0$ if $i < -1$. By performing a Schur complementation, i.e., eliminating the block in position $(2,1)$ in the above matrix, we obtain the matrix $T[J(\boldsymbol{A}^{(1)})] = U - WU^{-1}V$, $\boldsymbol{A}^{(1)} = [A_i^{(1)}]_{i\in\mathbb{Z}}$, $A_i^{(1)} = 0$ for $i < -1$. It can be easily verified that $\phi_1(z)$, defined by means of (20), is such that $\phi_1(z) = \sum_{i=-1}^{\infty} A_i^{(1)} z^{i+1}$.

If we apply the even-odd permutation followed by the Schur complementation to the infinite matrix $T_\infty[J(\boldsymbol{A})]$, the matrix that we obtain is the Toeplitz matrix $T_\infty[J(\boldsymbol{A}^{(1)})]$, defined by the same block entries of the bi-infinite matrix. In certain applications, like the solution of nonlinear matrix equations, it is more useful to perform an odd-even permutation, instead of the even-odd one; in this case the matrix that we obtain after the Schur complementation is Toeplitz except for its first block row, i.e., it has the structure (compare Remark 10):

$$\begin{bmatrix} \widehat{A}_0^{(1)} & \widehat{A}_1^{(1)} & \widehat{A}_2^{(1)} & \dots \\ A_{-1}^{(1)} & A_0^{(1)} & A_1^{(1)} & \dots \\ & A_{-1}^{(1)} & A_0^{(1)} & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix}.$$

This structure is kept at each step of CR, and the blocks $\widehat{A}_i^{(j)}$ are defined by means of the functional relation

$$\begin{cases} \widehat{\phi}_{j+1}(z) = \widehat{\phi}_{j,+}(z) - \widehat{\phi}_{j,-}(z)\phi_{j,-}(z)^{-1}\phi_{j,+}(z), \\ \widehat{\phi}_j(z) = \sum_{i=0}^{\infty} \widehat{A}_i^{(j)} z^i, \end{cases} \tag{21}$$

where $\widehat{\phi}_0(z) = \sum_{i=0}^{\infty} A_i z^i$. Again, by using evaluation/interpolation at suitable Fourier points, the first n coefficients of $\widehat{\phi}_{j+1}(z)$, given $\widehat{\phi}_j(z)$, can be approximated within the error ϵ in $O(m^3 n + m^2 n \log n)$ ops, where n is such that $\sum_{i=n+1}^{+\infty} ||\widehat{A}_i^{(j+1)}||_\infty < \epsilon$ [15, 16].

In the finite case, CR generates matrices that are block Toeplitz except at most the first block row and the last block column, depending on the kind of permutation, and on the block size n of the matrix that can be chosen, say, $n = 2^p$ or $n = 2^p - 1$ (see Remark 10 and [15]). In any case, the computation of the j-th step of CR requires $O(m^3 2^j + m^2 j 2^j)$ ops.

Concerning applicability, Theorem 13 can be easily extended in the following way:

Theorem 20 *Let the matrices $A_0^{(i)}$, for $i = 0, \ldots, j-1$, be nonsingular. Then the $(j+1)$-st step of cyclic reduction can be performed, i.e., $A_0^{(j)}$ is nonsingular, if and only if $T_{2^{j+1}-1}[J(\boldsymbol{A})]$ is nonsingular.*

Convergence properties have been proved under suitable conditions which are generally satisfied in queueing theory problems [13].

Theorem 21 *Let the matrix power series $\phi(z) = \sum_{i=-1}^{+\infty} A_i z^{i+1}$ be such that $A_0 = I - B_1$, $A_i = -B_i$, $i \neq 0$, $B_i \geq 0$, $B = \sum_{i=-1}^{+\infty} B_i$ is stochastic irreducible, $\phi(z)$ is meromorphic in the complex plane, $\sum_{i=0}^{+\infty}(i+1)A_i < +\infty$, and assume that the dominating left eigenvector $\boldsymbol{b}$ of B satisfies $\boldsymbol{b}^T \boldsymbol{e} = 1$, $\boldsymbol{b}^T \sum_{i=0}^{+\infty}(i+1)B_i \boldsymbol{e} < 1$, where $\boldsymbol{e} = (1, \ldots, 1)^T$. Then the equation (3) has one nonnegative solution G with spectral radius 1 and $G = -(\widehat{A}_0^{(j)})^{-1} A_{-1} + O(\sigma^{2^j})$, for any σ such that $1/\xi \leq \sigma < 1$, where $\xi = \min\{|z| : \ z \in \mathbb{C}, \ |z| > 1, \ \det \phi(z) = 1\}$.*

6 Graeffe's iteration for polynomials with structured matrix coefficients

In this section we consider CR applied to block Hessenberg block Toeplitz matrices, in the case where the blocks are obtained by reblocking into block Hessenberg form a generalized Hessenberg Toeplitz matrix. More specifically, consider the Laurent power series $a(z) = \sum_{i=-r}^{+\infty} a_i z^i$, and the generalized Hessenberg matrix $T[J(\boldsymbol{a})]$ where $\boldsymbol{a} = (a_j)_{j \in \mathbb{Z}}$, $a_j = 0$ if $j < -r$. We may easily verify that, if we define the $m \times m$ blocks, $m \geq r$,

$$A_h = (a_{j-i+mh})_{i,j=1,\ldots,m}, \quad , h = -1, 0, 1, \ldots \tag{22}$$

then $T[J(\boldsymbol{A})] = T[J(\boldsymbol{a})]$, $\boldsymbol{A} = [A_i]_{i \in \mathbb{Z}}$. If $a_j = 0$ for $j > s$, i.e., if $a(z)$ is a Laurent polynomial, and $m = \max\{r, s\}$ then $A_i = 0$ for $i > 1$, i.e., the reblocked matrix is block tridiagonal.

The nice feature of the reblocking technique is that it allows us to apply the Graeffe iteration, in the matrix form, also to infinite and finite banded Toeplitz matrices. In fact, the techniques introduced in section 3 enabled us to deal with the bi-infinite case only.

Due to the structure of the blocks A_i, $i \geq -1$, the matrix power series $\phi(z) = \sum_{i=-1}^{+\infty} A_i z^{i+1}$ is a z-circulant matrix (i.e., it can be expressed as a polynomial in the matrix $C = (c_{i,j})_{i,j=1,m}$, $c_{i,j} = 1$ if $i - j = 1$, $c_{1,m} = z$, $c_{i,j} = 0$ elsewhere), and therefore

$$\det \phi(z) = z^m \prod_{i=0}^{m-1} a(\omega_m^i z) \tag{23}$$

where ω_m is a primitive m-th root of 1 (see [47, 19]).

The property of $\phi(z)$ of being a z-circulant matrix, together with the functional relation (10), allows us to prove that at each step of CR the matrices $\psi_j(z) = \phi_j(z)^{-1}$ are Toeplitz for any value of z. More specifically, since $\phi(z)$ is a z-circulant matrix, and the set of z-circulant matrices is an algebra, then also $\phi(z) = \psi(z)^{-1}$ is z-circulant, for any z such that $\phi(z)$ and $\psi(z)^{-1}$ exist. On the other hand, from relation (10) and from the property that Toeplitz matrices form a linear space, it follows that $\psi_j(z)$ is a Toeplitz matrix for any z and for any j, and therefore $\phi_j(z) = \psi_j(z)^{-1}$ is the inverse of a Toeplitz matrix. It is well known that inverses of Toeplitz matrices can be represented by means of few vectors, the (displacement) generators, as a sum of products of triangular Toeplitz matrices [61, 25]. Matrices that can be represented in this way are called Toeplitz-like.

This structural property is very important since it allows the fast computation of $\phi_j(z)$ at each step j. In particular, it is possible to give explicit functional relations, involving operations between Toeplitz matrices and vectors, for the generators representing $\phi_{j+1}(z)$, given the generators representing $\phi_j(z)$ (we refer to [17, 19] for details on this subject). In the special case where $A_i = 0$ for $i > 1$, the blocks $A_i^{(j)}$, $i = -1, 0, 1$, can be computed in $O(m \log m + t(m))$ ops, where $t(m)$ is the arithmetic cost of solving an $m \times m$ Toeplitz-like system, i.e., $t(m) \in \{m^2, m \log^2 m, m \log m\}$ according to the adopted algorithm (see [68, 1, 35, 22, 18]).

In the case of Laurent polynomials, from relation (23) and from Theorem 16, the convergence properties of CR are related to the location of the zeros of $z^m a(z)$. In particular, we show that the condition $|\xi_m| < 1 < |\xi_{m+1}|$ is also sufficient to guarantee the convergence of CR.

Theorem 22 *Let $a(z) = \sum_{i=-r}^{s} a_i z^i$, be a Laurent polynomial and assume for simplicity that $m = \max\{r, s\} = r$. If the zeros ξ_i of $z^m a(z)$ are such that $|\xi_1| \leq$*

$\ldots \leq |\xi_{m+s}|$, $|\xi_m| < |\xi_{m+1}|$, then the matrix equations $A_{-1} + A_0X + A_1X^2 = 0$, $A_{-1}Y^2 + A_0Y + A_1 = 0$ have solutions X, Y, with spectral radius $\rho(X) = |\xi_m|$, $\rho(Y) = |\xi_{m+1}|^{-1}$. Moreover, if $|\xi_m| < 1 < |\xi_{m+1}|$ then Theorem 16 applies.

PROOF. Assume for simplicity that $\xi_i \neq \xi_j$ for $i \neq j$. From the relation $A_{-1}V + A_0VD^m + A_1VD^{2m} = 0$, where $V = (\xi_j^{i-1})_{i,j=1,m}$, $D = \mathrm{diag}(\xi_1, \ldots, \xi_m)$, it follows that $X = VD^mV^{-1}$ solves the equation $A_{-1} + A_0X + A_1X^2 = 0$ and $\rho(X) = \xi_m$. If the zeros ξ_i are not simple, then it is sufficient to chose as V the generalized Vandermonde matrix and to adjust consequently D. Similarly we do for Y. $\square$

7 Applications of Graeffe's iteration

CR has many applications to the solution of different problems. In this section we present applications to the solution of the problems of section 2, in particular to the solution of Toeplitz linear systems, to the solution of certain nonlinear matrix equations, to polynomial factorization and to the solution of resultant systems.

7.1 Solving Toeplitz linear systems

Consider the system

$$T\boldsymbol{x} = \boldsymbol{b} \tag{24}$$

where T can be finite, infinite, or bi-infinite (block) Toeplitz matrix in block Hessenberg form. If we apply the even-odd permutation of (block) rows and columns in the above system we obtain the 2×2 block system

$$\Pi T \Pi^T \Pi \boldsymbol{x} = \Pi \boldsymbol{b} = \begin{bmatrix} U & V \\ W & Y \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_+ \\ \boldsymbol{x}_- \end{bmatrix} = \begin{bmatrix} \boldsymbol{b}_+ \\ \boldsymbol{b}_- \end{bmatrix},$$

where Π is the permutation matrix, $\boldsymbol{x}_+$, $\boldsymbol{x}_-$, $\boldsymbol{b}_+$, $\boldsymbol{b}_-$ are the even and odd (block) components of $\boldsymbol{x}$ and $\boldsymbol{b}$. After performing the Schur complementation we obtain that $\boldsymbol{x}_1 = \boldsymbol{x}_-$ solves the system

$$T_1\boldsymbol{x}_1 = \boldsymbol{b}_1 \tag{25}$$

where T_1 is the matrix obtained by applying one step of CR to T and $\boldsymbol{b}_1 = \boldsymbol{b}_- - WU^{-1}\boldsymbol{b}_+$. In this way the solution of system (24) can be computed by solving (25) and then by recovering $\boldsymbol{x}_+$ by means of the relation $\boldsymbol{x}_+ = U^{-1}(\boldsymbol{b}_+ - V\boldsymbol{x}_-)$. We can recursively apply the same technique, thus generating the sequence of linear systems $T_j\boldsymbol{x}_j = \boldsymbol{b}_j$, where the matrix T_j is obtained by applying j steps of CR to T.

In the case of finite systems where $T = T_n[J(\boldsymbol{a})]$ and $a_i = 0$ for $i < -1$ (or $T = T_n[J(\boldsymbol{A})]$ and $A_i = 0$ for $i < -1$) and where for simplicity we assume $n = 2^p$, after $\log_2 n$ steps of CR we get a 1×1 (block) system, that can be immediately solved. Moreover if some of the hypotheses that guarantee the convergence of CR are verified, we can stop the process before performing $\log_2 n$ steps, when the matrix T_j generated by CR is "close" to a (block) diagonal matrix, and approximate $\boldsymbol{x}_j$ by solving a (block) diagonal system. The computational cost of the solution of the system can be divided into two parts: the cost of the computations of T_j, $j = 1, \ldots, \log_2 n$, and the cost of updating $\boldsymbol{b}_j$ and of back substitution. For block tridiagonal systems, the former is $O(m^3 \log n)$, the latter is $O(m^2 n)$. For systems in block Hessenberg form, the former is $O(m^3 n \log n)$, the latter is $O(m^2 n \log n)$. Moreover, if convergence of CR occurs then the $\log n$ factors in the above complexity bounds can be replaced by $\log\log\epsilon^{-1}$, where ϵ is the approximation error. The cost of solving $n \times n$ banded Toeplitz systems with $2m + 1$ diagonals, reblocked into $q \times q$ block tridiagonal block Toeplitz systems with $m \times m$ blocks (we assume that $n = mq$, for q integer), is $O(n \log m + m \log^2 m \log n)$ ops, where the $\log n$ factor can be replaced with $\log\log\epsilon^{-1}$ if convergence of CR occurs.

In order to apply CR for the solution of infinite and bi-infinite systems, the hypotheses that guarantee the convergence of $\{T_j\}_j$ to a (block) diagonal (or lower bi-diagonal in the infinite case) must be verified. Thus, for a sufficiently large value of j we can approximate a suitable number of components of $\boldsymbol{x}_j$, and then we may recover other components of $\boldsymbol{x}$ by means of back substitution. For block tridiagonal systems the computational cost for approximating n central block components of $\boldsymbol{x}$ within the error ϵ is $O(m^3 \log k + m^2 k)$ ops, where $k = \max\{n, \log_2 \epsilon^{-1}\}$.

7.2 Nonlinear matrix equations

Cyclic reduction and the generalized Graeffe iteration are closely related to the solution of the quadratic matrix equation (3). In fact, we easily find that the equation $A_{-1}^{(j)} + A_0^{(j)}X + A_1^{(j)}X^2 = 0$ is solved by $X = G^{2^j}$, where G solves (3); moreover, the matrix G solves the equation

$$A_{-1} + \widehat{A}^{(j)}X + A_1^{(j)}X^{2^j+1} = 0. \tag{26}$$

This property can be proved by applying CR with odd-even permutation to the infinite system $T_\infty[A_1, A_0, A_{-1}]\boldsymbol{X} = [-A_{-1}^T, 0, 0, \ldots]^T$, where $\boldsymbol{X}$ is the infinite block column vector having block components $X, X^2, X^3, \ldots$, in the light of Remark 10. Under the hypotheses of Theorem 16 we have $\lim_j A_1^{(j)} = 0$, thus $G = -\lim_j \widehat{A}^{(j)-1} A_{-1}$, so that CR provides a fast approximation to the solution G of (3).

Some conditions for the existence of the solution G of (3) are reported in [64]. More general conditions, that rely on the properties of Jordan chains [55], are stated in the following theorem [18, 22].

Theorem 23 [On the solution of the matrix equation (3)]. *Let $\phi(z) = A_{-1} + A_0 z + A_1 z^2$, $p(z) = \det \phi(z)$ and denote by ξ_i, $i = 1, \ldots, 2m$ the zeros of $p(z)$ where we set $\xi_{q+1} = \ldots = \xi_{2m} = +\infty$ if $\deg p(z) = q < 2m$. Let $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{2m}$ be a set of eigenvectors/Jordan chains for $\phi(z)$ corresponding to $\xi_1, \ldots, \xi_{2m}$. If there exists a subset of $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{2m}$, made up by m linearly independent vectors, say $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$, then the matrix equation* (3) *has a solution G whose eigenvalues are $\xi_1, \ldots, \xi_m$.*

PROOF. We prove the theorem in the simpler case where the linear independent vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ are such that $\phi(\xi_i)\boldsymbol{u}_i = 0$, $i = 1, \ldots, m$. The general case is left to the reader. Since the matrix U having columns $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m$ is nonsingular, denoting $G = U \mathrm{diag}(\xi_1, \ldots, \xi_m) U^{-1}$ we have, $G = \sum_{i=1}^m \boldsymbol{u}_i \xi_i \boldsymbol{v}_i^T$, where $\boldsymbol{v}_i^T$ are the rows of U^{-1}, moreover, $A_{-1} + A_0 G + A_1 G^2 = A_{-1} + A_0 \sum_{i=1}^m \boldsymbol{u}_i \xi_i \boldsymbol{v}_i^T + A_1 \sum_{i=1}^m \boldsymbol{u}_i \xi_i^2 \boldsymbol{v}_i^T = \sum_{i=1}^m (A_{-1} + A_0 \xi_i + A_1 \xi_i^2) \boldsymbol{u}_i \boldsymbol{v}_i^T = 0$. Thus G is solution of (3). $\square$

In the solution of M/G/1 type Markov chains the existence and uniqueness of a solution G with spectral radius equal to 1 hold under very mild conditions [82, 74]. Observe also that among the different solutions of (3), the one that is approximated by means of CR is the solution whose eigenvalues are made up by the zeros of minimum modulus of the polynomial $p(z)$.

A more general problem is the computation of a solution G of the matrix equation (4). In this case we can apply CR to the infinite system $T_\infty[J(\boldsymbol{A})]\boldsymbol{X} = [-A_{-1}^T, 0, 0, \ldots]^T$, where $\boldsymbol{A} = [A_i]_{i \in \mathbb{Z}}$, $A_i = 0$ for $i < -1$, and $\boldsymbol{X}$ is the infinite block column vector having block components $X, X^2, X^3, \ldots$. In this way we obtain the sequence of matrix equations $\sum_{i=-1}^\infty A_i^{(j)} X^{i+1} = 0$, having as solution $G_j = G^{2^j}$. Moreover we have $A_{-1} + \sum_{i=0}^\infty \widehat{A}_i^{(j)} G^{2^i+1} = 0$. Under the hypotheses of Theorem 21 we have $\lim_j \widehat{A}_i^{(j)} = 0$ for $i > 1$ so that $G = -\lim_j \widehat{A}_1^{(j)-1} A_{-1}$ and CR provides a powerful tool for the solution of (4).

Similar techniques can be designed for the computation of the extreme solutions of the matrix equations $X + A^T X^{-1} A = Q$ and $X - A^T X^{-1} A = Q$ where $Q, A \in \mathbb{R}^{m\times m}$ and Q is symmetric positive definite matrix. These equations occur in a wide variety of research areas, that include control theory, ladder networks, dynamic programming, stochastic filtering and statistics (see [3, 94, 43]). The algorithms based on the Toeplitz techniques, developed in [78], outperform the classical algorithms of [60, 59, 43, 94, 95, 41, 42, 3] mainly based on fixed point iterations, or on applications of Newton's algorithm.

7.3 Factoring polynomials

In recent years, the polynomial approach was successfully used for solving many problems of control and signal theory. For single-input, single-output systems the solution of Problem 3 in the symmetric case, i.e., $u(z) = l(z)$, is relevant and the problem is known as the spectral factorization problem. It plays a key role in problems of optimal control and robust control [39], in Wiener filtering and system identification [38], in wavelet construction and data modeling [57] and in time series analysis [91]. Moreover, it can be extended to consider matrix coefficients and matrix power series [93].

Besides applications for the spectral factorization, the problem of approximating a factor of an analytic function is also important in many other contexts. Special cases of particular interest are the problems in queueing theory, where the numerical solution of Markov chains is requested [82, 81, 74], and in physical sciences, where the zeros and turning points of special functions are to be computed [72].

CR provides a means for devising effective numerical methods for the approximation of a factor of a polynomial or, more generally, of an analytic function $f(z)$ given in terms of its power series expansion,

$$f(z) = \sum_{i=0}^{\infty} f_i z^i, \quad f_0 = 1. \tag{27}$$

Results along this line have been presented in [12] where CR is used to approximate the coefficients of the polynomial $w(z)$ of degree r made up by the first r zeros of $f(z)$. Specifically, it is shown that, under very mild assumption, the application of CR to the solution of a suitable infinite Toeplitz system in block Hessenberg form generates a sequence of vectors quadratically converging to that one formed by the coefficients of $w(z)$. The theoretical foundations of such an approach are summarized in the following theorem [12]. A generalization of this result to the case where Newton series are considered is presented in [52] while the connections with the solution of matrix difference equations of unbounded order are investigated in [49].

Theorem 24 *Let r be a positive integer and $f(z) : \mathcal{B}(R) \to \mathbb{C}$, $\mathcal{B}(R) = \{z \in \mathbb{C} : |z| < R\}$ with $R > 1$, be an analytic function having the series expansion (27) and zeros $\xi_1, \xi_2, \ldots$, such that $0 < |\xi_1| \leq \ldots \leq |\xi_r| < 1 < |\xi_{r+1}| \leq \ldots < R$. Denote by $w(z) = \prod_{i=1}^{r}(z - \xi_i) = \sum_{i=0}^{r} w_i z^i$, the factor of $f(z)$ with zeros inside the unit disk. Then, the block Toeplitz matrix in block Hessenberg form $T = T[A_{-1}, A_0, A_1, \ldots]$, where $A_k = (f_{i-j+r(k+1)})_{i,j=1,r}$ and $f_k = 0$ if $k < 0$, defines an invertible linear operator acting on $\ell^2(\mathbb{N})$. Therefore, it is uniquely determined the solution $\boldsymbol{X}^T = \left[X_1^T, X_2^T, \ldots\right]$ of the block linear system*

$$\boldsymbol{X}^T T = [-A_{-1}, 0, \ldots], \tag{28}$$

where $\boldsymbol{X}$ is a block vector with r columns $\in \ell^2(\mathbb{N})$. Moreover, the first row of X_1 provides the coefficients of $w(z)$, that is,

$$\boldsymbol{e}_1^T X_1 = \boldsymbol{v}^T = -[w_0, \ldots, w_{r-1}]. \tag{29}$$

Different approaches for the solution of the block Hessenberg system (28) are possible thus leading to many diverse methods for the approximation of the coefficients of the polynomial factor $w(z)$. Recently, new algorithms have also appeared in [90, 51] and [53]. However, in the case of block matrices obtained by partitioning a scalar Toeplitz matrix, like in the case of Theorem 24, all these algorithms are not competitive, both in terms of time and storage, with the CR scheme as described in section 6.

Assuming that break down does not occur, the CR process applied to the matrix T for the solution of the linear system (28) with the odd-even permutation generates a sequence of infinite block Hessenberg matrices $T^{(j)}$, $T^{(0)} = T$, which are block Toeplitz except for the entries $\widehat{A}_i^{(j)}$, $i = 0, 1, 2 \ldots$, in its first block column. Furthermore, it can be shown that (see [12] and Theorem 22 for the polynomial case) the vector sequence $\{\boldsymbol{v}^{(j)}\}$ given by

$$\boldsymbol{v}^{(j)} = -A_{-1}(\widehat{A}_0^{(j)})^{-1}\boldsymbol{e}_1, \tag{30}$$

converges quadratically to the vector $\boldsymbol{v}$ of (29), that is,

$$\| \boldsymbol{v}^{(j)} - \boldsymbol{v} \|_\infty = O(|\xi_r/\xi_{r+1}|^{2^j}).$$

In practice we apply a finite number of steps of CR to the infinite matrix T. To do this, we make use of the functional formulation of CR (compare with (20) and (21)) that reduces to the formulae (8), complemented with (11), in the polynomial/tridiagonal case where $f(z)$ reduces to a polynomial. In this latter situation, it follows that the computation of $\boldsymbol{v}^{(j+1)}$ of (30) essentially amounts to the solution of 2 Toeplitz-like linear systems of size r.

In this section we show that the reduction of the considered polynomial factorization problem to the solution of a bi-infinite band Toeplitz linear system provides a better approach. In fact, in this way we are able to prove that approximations to the coefficients of the factor $w(z)$ of $f(z)$ can directly be retrieved from the displacement generators of $A_0^{(j)}$ without computing the sequence $\{\widehat{A}_0^{(j)}\}$ through (11) and, thus, by roughly halving the overall computational cost.

Let us assume that $f(z)$ is a polynomial, $a(z) = z^{-r}f(z)$ is the Laurent polynomial (2) and $a(z) = l(z)u(z^{-1})$ is its spectral factorization, where $l(z) = \sum_{i=0}^{s} l_i z^i$, $w(z) = u_R(z) = z^r u(z^{-1})$ and $u(z) = \sum_{i=0}^{r} u_{-i} z^i$ with $u_0 = 1$. Moreover, for the sake of simplicity we also suppose that $r = s$; the more general case where $r \neq s$ can be reduced to the present one by using simple continuity arguments.

According to the results of section 2, the spectral factorization of $a(z)$ induces the Wiener-Hopf factorization of $T[\boldsymbol{a}]$, namely

$$T[\boldsymbol{a}] = T[\boldsymbol{u}]T[\boldsymbol{l}] = T[\boldsymbol{l}]T[\boldsymbol{u}], \tag{31}$$

where $\boldsymbol{u} = [u_{-r}, \ldots, u_0]^T$ and $\boldsymbol{l} = [l_0, \ldots, l_s]^T$. The matrix $T[\boldsymbol{a}]$ can be seen as a bi-infinite block tridiagonal matrix with $r \times r$ Toeplitz blocks $A_k^{(0)}$, $k = -1, 0, 1$, where $A_k^{(0)} = (a_{i-j+rk})$. Analogously, the upper triangular matrix $T[\boldsymbol{u}]$ and the lower triangular matrix $T[\boldsymbol{l}]$ can be partitioned into a block bi-diagonal form with blocks of size r. That is, we set $T[\boldsymbol{u}] = T[U_{-1}, U_0]$, $U_0 = (u_{i-j})_{i,j=1,r}$, $U_{-1} = (u_{i-j-r})_{i,j=1,r}$, and $T[\boldsymbol{l}] = T[L_0, L_1]$, $L_0 = (l_{i-j})_{i,j=1,r}$, $L_1 = (u_{i-j+r})_{i,j=1,r}$, where $u_i = 0$ if $i < -r$, $l_i = 0$ if $i > s$. Moreover, both the upper triangular matrix U_0 and the lower triangular matrix L_0 are nonsingular.

In this way, the Wiener-Hopf factorization (31) of $T[\boldsymbol{a}]$ can be rewritten into a block form as

$$T[\boldsymbol{a}] = \begin{bmatrix} \ddots & & & \\ \ddots & I & & \\ & M_1 & \ddots & \\ & & & \ddots \end{bmatrix} \begin{bmatrix} \ddots & & \\ & D_0 & \\ & & \ddots \end{bmatrix} \begin{bmatrix} \ddots & \ddots & & \\ & I & N_{-1} & \\ & & \ddots & \ddots \end{bmatrix} = MDN, \tag{32}$$

where $N_{-1} = (U_0)^{-1}U_{-1}$, $D_0 = L_0U_0$ and $M_1 = L_1(L_0)^{-1}$. In addition, the commuting property (31) turns into the block relations:

$$U_0L_0 + U_{-1}L_1 = L_0U_0 + L_1U_{-1}, \tag{33}$$

and

$$U_0L_1 = L_1U_0, \quad U_{-1}L_0 = L_0U_{-1}. \tag{34}$$

Observe that the equalities (34) merely follow also from the fact that upper (lower) triangular Toeplitz matrices commute.

The matrices M_1 and N_{-1} have very interesting properties related to the zeros ξ_i, $1 \le i \le 2r$, of $a(z)$. In particular, the next result says that they coincide with the r-th power of the Frobenius matrix associated with $l_R(z)$ and $u_R(z)$, respectively. Since these polynomials have all zeros inside the unit circle, it follows that both M and N define invertible operators.

Theorem 25 *Let F_{u_R} denote the Frobenius matrix of order r associated with the polynomial $u_R(z) = z^r u(z^{-1})$. Analogously, let F_{l_R} be the Frobenius matrix of order r associated with the polynomial $l_R(z) = z^r l(z^{-1})$. Finally, let J_r be the $r \times r$ permutation matrix with unit anti-diagonal entries. Then we have*

$$M_1 = -(J_rF_{l_R}J_r)^r, \quad N_{-1} = -(J_rF_{u_R}^TJ_r)^r. \tag{35}$$

Hence, the bi-infinite triangular matrices N and M are invertible and their inverses have block entries given by:

$$(M^{-1})_{i,j} = ((J_rF_{l_R}J_r)^r)^{i-j}, \quad i \ge j, \ i, j \in \mathbb{Z},$$

and

$$(N^{-1})_{i,j} = ((J_rF_{u_R}^TJ_r)^r)^{j-i}, \quad j \ge i, \ j, i \in \mathbb{Z}.$$

PROOF. We shall consider the matrix N only, since the results for M can be proven in exactly the same way. Let $u(z^{-1};\epsilon) = u(z^{-1}) + \epsilon z^{-r}$ be a Laurent polynomial of degree r, where ϵ is chosen in a neighborhood of the origin of the complex plane in such a way to guarantee that $u_R(z;\epsilon) = z^r u(z^{-1};\epsilon)$ has r distinct zeros denoted by $\xi_{\epsilon,1}, \ldots, \xi_{\epsilon,r}$. It is easy to check that, for $1 \le j \le r$,

$$\left[U_0 \middle| U_{-1} + \epsilon I_r \right] \left[\xi_{\epsilon,j}^{2r-1}, \xi_{\epsilon,j}^{2r-2}, \ldots, 1 \right]^T = \mathbf{0},$$

from which it follows that

$$\left(\xi_{\epsilon,j}^{r} U_0 + (U_{-1} + \epsilon I_r)\right) \left[\xi_{\epsilon,j}^{r-1}, \xi_{\epsilon,j}^{r-2}, \ldots, 1\right]^T = \mathbf{0}.$$

This implies that $U_0^{-1}(U_{-1} + \epsilon I_r) = -J_r V_\epsilon D_\epsilon^r V_\epsilon^{-1} J_r$, where $V_\epsilon = (\xi_{\epsilon,j}^{i-1})_{i,j=1,r}$ is the Vandermonde matrix and $D_\epsilon = \mathrm{diag}(\xi_{\epsilon,1}, \ldots, \xi_{\epsilon,r})$. Since $(V_\epsilon D_\epsilon V_\epsilon^{-1})^T$ coincides with the Frobenius matrix associated with the polynomial $u_R(z;\epsilon)$ [55], it follows that $U_0^{-1}(U_{-1} + \epsilon I_r)$ converges to $-(J_r F_{u_R}^T J_r)^r$ as ϵ goes to 0. The existence of the inverse of N^{-1} now follows from Banach's theorem [69] by observing that $u_R(z)$ has zeros ξ_i, $1 \le i \le r$, of modulus less than 1. In fact, in view of the matrix theory for the finite dimensional case, since the spectral radius of N_{-1} is $|\xi_r^r|$, one obtains that, for any σ satisfying $|\xi_r|^r < \sigma < 1$, there exists a suitable matrix norm $\| \cdot \|_*$ such that $\| N_{-1} \|_* \le \sigma$. Moreover, for any matrix norm $\| \cdot \|$ on $\mathbb{C}^{r\times r}$ there exists a positive constant γ satisfying $\| B \| \le \gamma \| B \|_*$, for any $B \in \mathbb{C}^{r\times r}$. Thus, by setting $N = I - \widehat{N}$, we find that, for any integer h, it holds

$$\| \widehat{N}^h \| \le \gamma \sigma^h. \tag{36}$$

This finally implies that N is invertible and its inverse is given by $N^{-1} = \sum_{i=0}^{\infty} \widehat{N}^i$. $\square$

This theorem provides a useful characterization of the blocks of the inverse of $T[\boldsymbol{a}]$. As already observed in section 2, the entries of this inverse are determined by the coefficients x_i of the Laurent expansion of $x(z) = 1/a(z)$. Hence, $X = (T[\boldsymbol{a}])^{-1}$ is a scalar Toeplitz matrix which can be partitioned into a block form with $r \times r$ Toeplitz blocks X_{i-j} given by $(X_{i-j})_{m,n} = x_{(i-j)r+m-n}$. By combining relation (32) with Theorem 25, we arrive at the following relation expressing the central block X_0 in terms of the Frobenius matrices associated with the factors of $a(z)$. It holds

$$X_0 = \sum_{i=0}^{\infty} (J_r F_{u_R}^T J_r)^{ir} (D_0)^{-1} (J_r F_{l_R} J_r)^{ir}, \tag{37}$$

from which it follows that X_0 is the solution of a discrete Lyapunov matrix equation of the form

$$X_0 - (J_r F_{u_R}^T J_r)^r X_0 (J_r F_{l_R} J_r)^r = (D_0)^{-1}. \tag{38}$$

This solution is unique since the spectral radii of F_{u_R} and F_{l_R} are less than 1.

Before proceeding to express in a convenient way the inverse of the solution of (38), we intend now to remark the relationships with the CR scheme. In order to establish these connections, we first observe that the CR algorithm can be applied to the system $T[\boldsymbol{a}]\boldsymbol{X} = \boldsymbol{B}$, $\boldsymbol{X} = [X_i]_{i\in\mathbb{Z}}$, $\boldsymbol{B} = [B_i]_{i\in\mathbb{Z}}$, $X_i, B_i \in \mathbb{R}^{r\times r}$, $B_0 = I$, $B_i = 0$ elsewhere, for the computation of the central block column $\boldsymbol{X}$ of $T[\boldsymbol{a}]^{-1}$. If no break down occurs, this process produces a sequence of linear systems $T^{(j)}\boldsymbol{X}^{(j)} = \boldsymbol{B}$, $\boldsymbol{X}^{(j)} = [X_{i2^j}]_{i\in\mathbb{Z}}$, where $T^{(j)}$ are invertible operators; each $T^{(j)}$ is a block Toeplitz matrix in block tridiagonal form with blocks $A_i^{(j)}$, $i = -1, 0, 1$ defined by (8). Thus, in the view of Theorems 22 and 16, the sequence $T^{(j)}$ quadratically converges towards a block diagonal operator of the form $I \otimes X_0^{-1}$ and, moreover,

$$\| A_0^{(j)} - X_0^{-1} \|_\infty = O(|\xi_r/\xi_{r+1}|^{2^j}).$$

The subsequent investigation of (38) largely follows from the results of [58] and it is here presented only for completeness. Recall that every block T of the matrices $T[\boldsymbol{u}]$ and $T[\boldsymbol{l}]$ is a triangular Toeplitz matrix for which it holds $J_r T J_r = T^T$. By means of straightforward calculations, from (38) we then find that $Y = U_0 X_0 L_0$ is the unique solution of

$$Y = I + U_{-1}(U_0)^{-1}Y(L_0)^{-1}L_1.$$

Hence, it follows that

$$\begin{aligned} &Y^{-1} = (I + U_{-1}(U_0)^{-1}Y(L_0)^{-1}L_1)^{-1} = \\ &I - U_{-1}(U_0)^{-1}(Y^{-1} + (L_0)^{-1}L_1U_{-1}(U_0)^{-1})^{-1}(L_0)^{-1}L_1. \end{aligned}$$

This implies that the matrix $W = I - Y^{-1}$ is the unique solution of the matrix equation

$$W = U_{-1}(L_0U_0 - L_0WU_0 + L_1U_{-1})^{-1}U_1.$$

In this way, by using the commuting properties (33) and (34), it is easily found that

$$W = (L_0)^{-1}U_{-1}L_1(U_0)^{-1},$$

from which we may finally conclude that

$$X_0^{-1} = L_0U_0 - U_{-1}L_1.$$

Summing up, we have the following:

Theorem 26 *Let us assume that the CR process applied to $T^{(0)} = T[\boldsymbol{a}]$ goes on without break-downs so that it generates a sequence of block tridiagonal block Toeplitz matrices $T^{(j)}$. Let $A_0^{(j)}$ be the $r \times r$ matrix located on the main diagonal of $T^{(j)}$. Moreover, suppose that $a(z)$ can be split as $a(z) = u(z^{-1})l(z)$ with $u(z) = \prod_{i=1}^{r}(1-\xi_i z) = \sum_{i=0}^{r} u_{-i}z^i$ and $l(z) = a_r \prod_{i=1}^{r}(z-\xi_{i+r}) = \sum_{i=0}^{r} l_i z^i$ with $0 < |\xi_1| \leq \ldots \leq |\xi_r| < 1 < |\xi_{r+1}| \leq \ldots < |\xi_{2r}|$. Finally, define the Toeplitz-like matrix $A_0^{(\infty)}$ as follows:*

$$A_0^{(\infty)} = L_0 U_0 - U_{-1} L_1,$$

where $L_0 = (l_{i-j})$, $L_1 = (l_{r+i-j})$, $U_0 = (u_{i-j})$, and $U_{-1} = (u_{-r+i-j})$, with $u_s = 0$ if $s > 0$ or $s < -r$ and $l_s = 0$ if $s < 0$ or $s > r$. Then, we have

$$\| A_0^{(j)} - A_0^{(\infty)} \|_\infty = O(|\xi_r/\xi_{r+1}|^{2^j}).$$

The above result provides a description of a Gohberg-Semencul type formula for the inverse of a nonsingular Toeplitz matrix in terms of the coefficients of the factors obtained by the spectral factorization of its symbol. The idea of relating suitable representations of the inverse of a Toeplitz matrix with the polynomials found by means of the solution of a factorization problem involving its symbol is not new. According to Iohvidov's book [66] it was apparently ascertained for the first time by G. Baxter and I. Hirschman [7]. Some years later, A. A. Semencul [88] also made use of similar developments to prove an inversion formula and, this result was fundamental in the corresponding section of the book of I. C. Gohberg and I. A. Fel'dman [54]. Unlike these theoretical contributions, we believe that the possibility of extending the derivation of Theorem 26 to the more general case where no restriction is imposed on the spectrum of the considered polynomials should also be investigated for computational purposes. In fact, we guess that approximate factorizations of a polynomial could be used in order to construct approximate representations of the inverse of a Toeplitz matrix with several applications to the preconditioning theory [35]. This investigation is presently an ongoing work.

Theorem 26 implies that a sequence of approximations to the factors of $a(z)$ can be constructed directly from the displacement of the diagonal blocks $A_0^{(j)}$ generated by CR without solving the auxiliary linear systems (30). Since the limit matrix $A_0^{(\infty)}$ is the sum of products of lower triangular by upper triangular Toeplitz matrices, it is quite natural to consider the displacement operator:

$$\mathcal{F} : \mathbb{C}^{r\times r} \to \mathbb{C}^{r\times r}, \quad \mathcal{F}(A) = A - Z_r A Z_r^T,$$

where Z_r is the $r \times r$ down-shift matrix. In particular, if we denote by $\boldsymbol{l}_0$ and $\boldsymbol{l}_1$, respectively, the first and the last column of L_0 and L_1, and by $\boldsymbol{u}_0$ and

$\boldsymbol{u}_{-1}$, respectively, the last and the first column of U_0 and U_{-1}, then we find that

$$\mathcal{F}(A_0^{(\infty)}) = \boldsymbol{u}_{-1}(J_r\boldsymbol{l}_1)^T - \boldsymbol{l}_0(J_r\boldsymbol{u}_0)^T. \tag{39}$$

This relation allows us to determine the coefficients of the polynomial factors $u_R(z)$ and $l(z)$ from the entries $t_{i,j}$ of the first and the second column of $\mathcal{F}(A_0^{(\infty)})$ denoted, respectively, by $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$. In fact, by using

$$u_{-r}l_1 - l_0u_{-r+1} = t_{r,1}, \quad u_{-r}l_1 + l_0u_{-r+1} = a_{-r+1},$$

and

$$u_{-r}l_2 - l_0u_{-r+2} = t_{r,2}, \quad u_{-r}l_2 + l_0u_{-r+2} + l_1u_{-r+1} = a_{-r+2},$$

the values of l_1, u_{-r+1}, l_2 and u_{-r+2} are immediately found. In this way, one obtains the following linear system

$$[\boldsymbol{u}_{-1}, \boldsymbol{l}_0] \begin{bmatrix} l_1 & l_2 \\ u_{-r+1} & u_{-r+2} \end{bmatrix} = [\boldsymbol{t}_1, \boldsymbol{t}_2],$$

which provides the coefficients of $u_R(z)$ and $l(z)$ at a linear computational cost. The same rational algorithm can be clearly applied to $\mathcal{F}(A_0^{(j)})$ by yielding the desired approximations to $\boldsymbol{u}_{-1}$, and $\boldsymbol{l}_0$ within the same accuracy as the one of $A_0^{(j)}$ to $A_0^{(\infty)}$.

7.4 Solution of resultant-like systems

The solution of resultant-like linear systems is often encountered in many relevant applicative and industrial problems of data modeling, system theory, control system design and digital signal processing, where the primary focus is on the study of process dynamics. In such applications a topic of considerable interest is the estimation of the transfer function of the considered input-output model from process records consisting of two times series: the input time series and the output time series. In the presence of appreciable noise, methods based on the choice of special inputs are usually not satisfactory and more involved statistical methods based on the properties of cross-covariance and cross-correlation functions should be used [30].

Let us consider, for instance, the basic problem of evaluating the sequence $\rho(n)$ of the cross-covariance between the output of two discrete time systems driven by the same white noise. From a mathematical point of view, this reduces to the problem of estimating integrals of the form

$$\rho(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{b(e^{i\theta})d(e^{i\theta})}{a(e^{i\theta})c(e^{i\theta})} e^{in\theta} d\theta,$$

where $a(z)$, $b(z)$, $c(z)$ and $d(z)$ are stable polynomials in z^{-1}, that is, all their zeros are inside the unit circle in the complex plane. In [30] it is shown that the evaluation of these integrals is equivalent to the solution of a polynomial equation of the form

$$c(z^{-1})y(z) + a(z)zx(z^{-1}) = b(z)d(z^{-1})$$

which represents a linear system whose coefficient matrix is a permuted version of the Sylvester resultant matrix associated with the polynomials $a(z)$ and $c(z)$. Since the determinant of this matrix can be explicitly expressed in terms of the zeros of its polynomial generators [5], then, from the stability assumption it immediately follows that this system is solvable and its solution is unique.

The auto-covariance case, where $b(z) = d(z)$ and $a(z) = c(z)$ has been also widely studied [37]. In particular, in this case the previous polynomial approach reduces to solving a structured linear system whose coefficient matrix is a Jury matrix generated by the coefficients of $a(z)$. Jury matrices are a clear example of structured matrices. Given the real polynomial $\gamma(z) = \sum_{i=0}^{m} \gamma_i z^i$, where $\gamma_0 > 0$ and $\gamma_m \neq 0$, having zeros ξ_i, $0 < |\xi_1| \leq \ldots \leq |\xi_m| < 1$, we define the $(m+1)$-th order Jury matrix $\mathcal{J}(\gamma)$ associated with $\gamma(z)$ as follows:

$$\mathcal{J}(\gamma) = T(\gamma) + H(\gamma) \tag{40}$$

where $T(\gamma)$ denotes the lower triangular Toeplitz matrix associated with $\gamma(z)$ and $H(\gamma)$ is the upper triangular Hankel matrix with respect to the main anti-diagonal, namely,

$$T(\gamma) = \begin{bmatrix} \gamma_0 & & \\ \vdots & \ddots & \\ \vdots & & \ddots & \\ \gamma_m & \cdots\cdots & \gamma_0 \end{bmatrix}, \quad H(\gamma) = \begin{bmatrix} \gamma_0 & \cdots\cdots & \gamma_m \\ \vdots & & \iddots \\ \vdots & \iddots & \\ \gamma_m & & \end{bmatrix}.$$

It is remarkable to observe that the solution of the spectral factorization problem can also lead to solving Jury linear systems. In fact, if we apply the

Newton-Raphson iteration to the quadratic equation $a(z) = u(z^{-1})l(z)$, the coefficients of the spectral factors being unknown, then we obtain a linear system whose (Jacobian) matrix is a Jury matrix (see [91]).

In [37] it was shown that the solution of Jury linear systems of the form $\mathcal{J}(\gamma)\boldsymbol{y} = \boldsymbol{b}$ and $\boldsymbol{y}^T\mathcal{J}(\gamma) = \boldsymbol{b}^T$ can be reduced to the solution of the following polynomial problem:

Problem 27 [Polynomial counterpart of Jury systems]. *Given $\gamma(z)$ and a polynomial $b(z)$ of degree $n \leq m$, determine the polynomial $y(z)$ of degree at most n such that*

$$\frac{b(z)b(z^{-1})}{\gamma(z)\gamma(z^{-1})} = \frac{y(z)}{\gamma(z)} + \frac{y(z^{-1})}{\gamma(z^{-1})}. \tag{41}$$

Our approach to the solution of (41) is based on the results of the previous sections concerning the evaluation of the coefficients of the Laurent series of the reciprocal of

$$\frac{1}{a(z)} = \frac{1}{\gamma(z)\gamma(z^{-1})}.$$

In fact, under our assumptions on the zeros ξ_i of $\gamma(z)$, it immediately follows that the function $g(z) = y(z^{-1})/\gamma(z^{-1})$ is analytic for $|z| < 1/|\xi_m|$, whereas the function $g(z^{-1})$ is analytic for $|z| > |\xi_m|$. Hence, we find that $g(z)+g(z^{-1})$ has a Laurent expansion for $z \in G$, $G = \{z \in \mathbb{C} : |\xi_m| < |z| < 1/|\xi_m|\}$. That is,

$$g(z) + g(z^{-1}) = \sum_{i\in\mathbb{Z}} c_i z^i = (c_0/2 + \sum_{i=1}^{\infty} c_i z^i) + (c_0/2 + \sum_{i=1}^{\infty} c_i z^{-i}), \quad \forall z \in G$$

where the first series is given by the Taylor expansion of $g(z^{-1})$. It is now clear that the function $1/a(z)$ also possesses a Laurent expansion in G,

$$\frac{1}{a(z)} = \sum_{i\in\mathbb{Z}} x_i z^i, \quad \forall z \in G,$$

and, then, the same holds for $b(z)b(z^{-1})/a(z)$. From the uniqueness of the Laurent series of an analytic function in a given annulus [63], we may therefore conclude that

$$T_{4m}[x_{1-2m}, \ldots, x_{2m-1}] \left[\widehat{b}_{-m+1}, \ldots, \widehat{b}_m\right]^T = [c_{-m+1}, \ldots, c_m]^T, \tag{42}$$

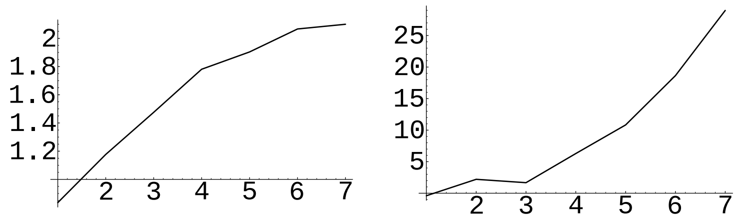

Fig. 1. Plot of $\mathrm{Log}(\mathrm{cond}(T)/\mathrm{cond}(\mathcal{J}(\gamma)))$ and of $\mathrm{Log}(\max_i\{\mathrm{cond}(\mathcal{J}(\gamma))_i)\}/\mathrm{cond}(T))$

where $b(z)b(z^{-1}) = \sum_{i=-m}^{m} \widehat{b}_i z^i$, $c_{-i} = c_i$, $i = 1, \ldots, m-1$. The observation that the coefficients of $y(z)$ can be retrieved from $c_0, \ldots, c_m$ finally leads to the following procedure for the solution of Problem 27.

Procedure **SolveJury**

(1) compute the central $4m-1$ coefficients $x_{-2m+1}, \ldots, x_0, \ldots, x_{2m-1}$ of the reciprocal of a Laurent polynomial $a(z) = \gamma(z)\gamma(z^{-1})$;
(2) determine the first $m+1$ coefficients $c_0, \ldots, c_m$ of the Laurent series of $g(z) + g(z^{-1})$ by means of (42);
(3) find the coefficients of $y(z^{-1})$ such that

$$\frac{z^m y(z^{-1})}{z^m \gamma(z^{-1})} = \frac{c_0}{2} + \sum_{i=1}^{m} c_i z^i \quad (\bmod z^{m+1}).$$

It is clear that the most expensive computation is to be performed at the step 1. In particular, the evaluation of the required central coefficients of $1/a(z)$ can be carried out by means of the super-fast algorithm of section 3 based on Graeffe's iteration. Alternatively, since in this case the spectral factorization of $a(z)$ is already known, one can use Theorem 26 in order to reduce this task to the solution of a definite symmetric Toeplitz-like system of order $2m$. Obviously, also in this case, well known super-fast procedures could be applied.

The stability properties of these many diverse approaches are currently under investigation and the results of an extensive numerical experience will be reported in subsequent works. Here, we only intend to discuss some preliminary observations on the conditioning of solving Jury systems. The efficient solution methods described in [37] are essentially equivalent to certain recursive variants of Gaussian elimination without pivoting applied to $\mathcal{J}(\gamma)$, and, therefore, their stability depends on the conditioning of the leading principal submatrices of $\mathcal{J}(\gamma)$. In our numerical experiments, we have observed that well conditioned Jury matrices can frequently admit many extremely ill-conditioned leading principal submatrices. This is clearly due to the fact that $\mathcal{J}(\gamma)$ is not positive definite. On the other hand, our approach reduces the solution of a Jury linear system to that one of a definite structured linear system of order $2m$. In this way, if the conditioning of this system —denoted by T— is comparable with the one of $\mathcal{J}(\gamma)$, then super-fast recursive procedure could be applied in a stable way. Table 1 reports the results of some numerical experiments performed with Mathematica^{TM}. We have generated real polynomials $\gamma(z) = \sum_{i=0}^{m} \gamma_i (1.2z)^i$ of degree $m = 2^k$, $k = 2, \ldots, 7$, according to the following rule

$$\gamma_0 = 1, \quad \gamma_i = \gamma_{i-1} + \ Random[\ ],$$

where *Random*[ ] provides a real random number uniformly distributed in the interval $[0, 1]$. Since we have $\gamma_0 < \gamma_1 < \ldots < \gamma_m$, in the view of the Kakeya-Eneström theorem [63], it follows that the generated polynomials are stable. Then, we have compared the condition number of $\mathcal{J}(\gamma)$ of order $m+1$ with the one of the matrix T of order $2m$ whose inverse is the symmetric Toepliz matrix having $[x_0, \ldots, x_{2m-1}]$ as its first row. Table 1 clearly illustrates that for a typical set of test polynomials these condition numbers are asymptotically within a constant whereas $\mathcal{J}(\gamma)$ has (many) arbitrarily ill-conditioned leading principal submatrices.

In conclusion, we have shown that spectral factorization methods can lead to super-fast algorithms for the numerical treatment of a certain class of structured linear systems. A continuous analogue of the spectral factorization problem is the problem of factoring polynomials with respect to the imaginary axis in the complex plane (Hurwitz factorization problem) which plays a key role in the synthesis of continuous quadratically optimal controllers [67]. The study of similar results and relations between Hurwitz factorization methods [50] and the solution of structured linear systems is an interesting research topic.

References

[1] G. S. Ammar and B. Gragg. Numerical experience with superfast real Toeplitz solver. *Linear Algebra Appl.*, 121:185–206, 1989.

[2] G. Anastasi, L. Lenzini, and B. Meini. Performance evaluation of a worst case model of the MetaRing MAC protocol with global fairness. *Performance Evaluation*, 29:127–151, 1997.

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

[4] K. J. Åström, E. I. Jury, and R. G. Agniel. A numerical method for the evaluation of complex integrals. *IEEE Trans. Automat. Contr.*, 15:468–471, 1970.

[5] S. Barnett. *Polynomials and Linear Control Systems.* Number 77 in Textbooks in Pure and Applied Mathematics. Marcel Dekker, Inc., New York, 1983.

[6] F. L. Bauer. Ein direktes iterations Verfahren zur Hurwitz -zerlegung eines Polynoms. *Arch. Elektr. Uebertragung*, 9:285–290, 1955.

[7] G. Baxter and I. I. Hirschman. An explicit inversion formula for finite-section Wiener-Hopf operators. *Bull. Amer. Math. Soc.*, 70:820–823, 1964.

[8] D. A. Bini, S. Chakravarthy, and B. Meini. A new algorithm for the design of finite capacity service units. In B. Plateau, W. J. Stewart, and M. Silva, editors, *Numerical Solution of Markov Chains, 3rd International Workshop, Saragoza (Spain), Sept. 1999*, pages 247–260. Prensas Unversitarias de Zaragoza, 1999.

[9] D. A. Bini, G. M. Del Corso, G. Manzini, and L. Margara. Inversion of circulant matrices over $\boldsymbol{Z}_m$. *Math. Comp.* To appear.

[10] D. A. Bini, A. Farusi, G. Fiorentino, and B. Meini. On the regularized solution of block banded block Toeplitz systems. Proceedings of SPIE, 2000.

[11] D. A. Bini and L. Gemignani. Fast parallel computation of the polynomial remainder sequence via Bezout and Hankel matrices. *SIAM J. Comput.*, 24:63–77, 1995.

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

[13] D. A. Bini, G. Latouche, and B. Meini. Quadratically convergent algorithms for solving matrix polynomial equations. Submitted for publication, 2000.

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

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

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

[17] D. A. Bini and B. Meini. Using displacement structure for solving non-skip-free M/G/1 type Markov chains. In A. Alfa and S. Chakravarthy, editors, *Advances in Matrix Analytic Methods for Stochastic Models - Proceedings of the 2nd international conference on matrix analytic methods*, pages 17–37. Notable Publications Inc, NJ, 1998.

[18] D. A. Bini and B. Meini. Approximate displacement rank and applications. In V. Olshevsky, editor, *Proceedings AMS Conference "Structured Matrices in Operator Theory, Numerical Analysis, Control, Signal and Image Processing", Boulder*. AMS, 1999. To appear.

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

[20] D. A. Bini and B. Meini. Fast algorithms for structured problems with applications to Markov chains and queueing models. In T. Kailath and A. H. Sayed, editors, *Fast reliable methods for matrices with structure*,

chapter 8, pages 211–243. SIAM, Philadelphia, 1999.

[21] D. A. Bini and B. Meini. Solving block banded block Toeplitz systems with banded Toeplitz blocks. In F. T. Luk, editor, *Advanced Signal Processing Algorithms, Architectures, and Implementations IX*, volume 3807, pages 300–311. Proceedings of SPIE, 1999.

[22] D. A. Bini and B. Meini. Solving block banded block Toeplitz systems with structured blocks: algorithms and applications. In D. A. Bini, E. Tyrtyshnikov, and P. Yalamov, editors, *Structured matrices: recent developments in theory and computation*. NOVA, 2000. In press.

[23] D. A. Bini, B. Meini, and V. Ramaswami. Analyzing M/G/1 paradigms through QBDs: the role of the block structure in computing the matrix G. In G. Latouche and P. Taylor, editors, *Advances in Algorithmic Methods for Stochastic Models*, pages 73–86. Notable Publications, New Jersey, USA, 2000. Proceedings of the Third Conference on Matrix Analytic Methods.

[24] D. A. Bini and V. Pan. Polynomial division and its computational complexity. *J. Complexity*, 2:179–203, 1986.

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

[26] D. A. Bini and V. Y. Pan. Graeffe's, Chebyshev-like, and Cardinal's processes for splitting a polynomial into factors. *J. Complexity*, 12(4):492–511, 1996. Special issue for the Foundations of Computational Mathematics Conference (Rio de Janeiro, 1997).

[27] A. Böttcher and S. Grudsky. Twenty old and new asymptotic results on Toeplitz band matrices. In D. A. Bini, E. Tyrtyshnikov, and P. Yalamov, editors, *Structured matrices: recent developments in theory and computation*. NOVA, 2000. In press.

[28] A. Böttcher, S. Grudsky, A. Kozak, and B. Silbermann. Convergence speed estimates for the norms of the inverses of large truncated Toeplitz matrices. *Calcolo*, 36(2):103–122, 1999.

[29] A. Böttcher and B. Silbermann. *Introduction to Large Truncated Toeplitz Matrices*. Springer-Verlag, New York, 1999.

[30] G. E. P. Box and G. M. Jenkins. *Time series analysis, forecasting and control*. Holden-Day, 1970.

[31] L. Brugnano and D. Trigiante. *Solving Differential Problems by Multistep Initial and Boundary Value Methods*. Gordon and Berach Science Publishers, Amsterdam, 1998.

[32] A. Bultheel. Zeros of a rational function defined by its Laurent expansion. In *Padé approximation and its applications*, volume 1071 of *Lecture Notes in Math.*, pages 34–48. Springer-Verlag, 1984.

[33] B. L. Buzbee, G. H. Golub, and C. W. Nielson. On direct methods for solving Poisson's equation. *SIAM J. Num. Anal.*, 7:627–656, 1970.

[34] Jean-Paul Cardinal. On two iterative methods for approximating the roots of a polynomial. In *The mathematics of numerical analysis (Park City, UT, 1995)*, pages 165–188. Amer. Math. Soc., Providence, RI, 1996.

[35] R. H. Chan and M. K. Ng. Conjugate gradient methods for Toeplitz systems. *SIAM Review*, 38:427–482, 1996.

[36] P. Davis. *Circulant Matrices.* Wiley, New York, 1974.

[37] C. J. Demeure and C. T. Mullis. The Euclid algorithm and the fast computation of cross-covariance and autocovariance sequences. *IEEE Trans. Acoust., Speech and Signal Processing*, 37:545–552, 1989.

[38] C. J. Demeure and C. T. Mullis. A Newton-Raphson method for moving-average spectral factorization using Euclid algorithm. *IEEE Trans. Acoust., Speech and Signal Processing*, 38:1697–1709, 1990.

[39] P. A. Van Dooren. Some numerical challenges in control theory. In *Linear Algebra for Control Theory*, volume 62 of *IMA Vol. Math. Appl.* Springer, 1994.

[40] I.Z. Emiris and V. Y. Pan. The structure of sparse resultant matrices. In *Proc. of the ACM International Symposium on Symbolic and Algebraic Computation (ISSAC97)*, pages 189–196. ACM Press, New York, 1997.

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

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

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

[44] Miroslav Fiedler. Über das Gräffesche Verfahren. *Czechoslovak Math. J.*, 5(80):506–516, 1955.

[45] L. Fox and L. Hayes. Polynomial factorization and the Q-D algorithm. *Linear Algebra Appl.*, 1:445–468, 1968.

[46] P. A. Fuhrmann. *A Polynomial Approach to Linear Algebra*. Springer-Verlag, New York, 1996.

[47] H. R. Gail, S. L. Hantler, and B. A. Taylor. Non-skip-free M/G/1 and G/M/1 type Markov chains. *Adv. Appl. Prob.*, 29:733–758, 1997.

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

[49] L. Gemignani. On a generalization of Poincaré's theorem for matrix difference equations arising from root-finding problems. In V. Olshevsky, editor, *Proceedings AMS Conference "Structured Matrices in Operator Theory, Numerical Analysis, Control, Signal and Image Processing", Boulder.* AMS, 1999. To appear.

[50] L. Gemignani. Computing a Hurwitz factorization of a polynomial. *J. Computational and Applied Mathematics*, 2000. To appear.

[51] L. Gemignani. Efficient and stable solution of structured Hessenberg linear systems arising from difference equations. *Numer. Lin. Algebra*

with Appl., 2000. To appear.

[52] L. Gemignani. Polynomial factors, lemniscates and structured matrix technology. In D. A. Bini, P. Yalamov, and E. Tyrtyshnikov, editors, *Structured matrices: recent developments in theory and computation.* NOVA Science, 2000.

[53] L. Gemignani and G. Lotti. A recursive block Gaussian elimination algorithm for solving block Hessenberg systems. Submitted.

[54] I. Gohberg and I. A. Fel'dman. *Convolution equations and projection methods for their solution.* American Mathematical Society, 1974.

[55] I. Gohberg, P. Lancaster, and L. Rodman. *Matrix Polynomials.* Academic Press Inc., New York, 1982.

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

[57] T. N. T. Goodman, C. A. Micchelli, G. Rodriguez, and S. Seatzu. Spectral factorization of Laurent polynomials. *Adv. Comput. Math.*, 7:429–454, 1997.

[58] W. B. Gragg and L. Reichel. On singular values of Hankel operator of finite rank. *Linear Algebra Appl.*, 121:53–70, 1989.

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

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

[61] G. Heinig and K. Rost. *Algebraic Methods for Toeplitz-like Matrices and Operators.* Akademie-Verlag, Berlin, and Birkhäuser, Boston, 1984.

[62] D. Heller. Some aspects of the cyclic reduction algorithm for block tridiagonal linear systems. *SIAM J. Num. Anal.*, 13:484–496, 1976.

[63] P. Henrici. *Applied and Computational Complex Analysis*, volume 1. Wiley, 1974.

[64] N. J. Higham and H.-M. Kim. Numerical analysis of a quadratic matrix equation. Numerical Analysis Report No. 347, Manchester Centre for Computational Mathematics, Manchester, England, August 1999. To appear in IMA J. of Numerical Analysis.

[65] A. S. Householder. *The Numerical Treatment of a Single Nonlinear Equation.* International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, NY, 1970.

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

[67] J. Ježek. Conjugated and symmetric polynomial equations. *Kybernetika*, 19, 1983.

[68] T. Kailath and A. H. Sayed. Displacement structure: theory and applications. *SIAM Review*, 37:297–386, 1995.

[69] L. V. Kantorovich and G. P. Akilov. *Functional Analysis in Normed Spaces.* Pergamon Press, 1964.

[70] D. Kershaw. An analysis of the method of L. Fox and L. Hayes for the

factorization of a polynomial. *Linear Algebra Appl.*, 86:179–187, 1987.

[71] V. E. Kondrashov. Extracting the roots of a polynomial in the order of increase of their moduli. *Soviet. Math. Dokl.*, 35:77–81, 1987.

[72] P. Kravanja and M. Van Barel. *Computing the Zeros of Analytic Functions.* Number 1727 in Lectures Notes in Mathematics. Springer, 2000.

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

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

[75] L. Lenzini, B. Meini, and E. Mingozzi. An efficient numerical method for performance analysis of contention MAC protocols: a case study (PRMA++). *IEEE J. Selected Areas Comm.*, 16:653–667, 1998.

[76] B. Meini. An improved FFT-based version of Ramaswami's formula. *Commun. Statist. Stochastic Models*, 13:223–238, 1997.

[77] B. Meini. Solving M/G/1 type Markov chains: recent advances and applications. *Commun. Statist. Stochastic Models*, 14:479–496, 1998.

[78] B. Meini. Efficient computation of the extreme solutions of $X + A^*X^{-1}A = Q$, $X - A^*X^{-1}A = Q$. Submitted for publication, 1999.

[79] B. Mourrain and V. Y. Pan. Multivariate polynomials, duality and structured matrices. *J. Complexity*, 16:110–180, 2000.

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

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

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

[83] M. A. Ostrowski. Recherches sur la methode de Graeffe et les zeros des polynomes et des series de Laurent. *Acta Math.*, 72:99–257, 1940.

[84] V. Y. Pan. Solving a polynomial equation: Some history and recent progress. *SIAM Review*, 39, 2:187–220, 1997.

[85] V. Ramaswami. A stable recursion for the steady state vector in Markov chains of M/G/1 type. *Commun. Statist. Stochastic Models*, 4:183–188, 1988.

[86] A. Schönhage. The fundamental theorem of algebra in terms of computational complexity. Math. Dept., University of Tübingen, Tübingen, Germany, 1982.

[87] A. Schönhage. Variations on computing reciprocals of power series. *Inform. Process. Lett.*, 74(1-2):41–46, 2000.

[88] A. A. Semencul. *Inversion of finite Toeplitz matrices and their continual analoga*, chapter Appendix II in: Projection methods in the solution of Wiener-Hopf equations, by I. C. Gohberg and I. A. Fel'dman. R. I. O.

Akad. Nauk. MSSR, 1967. In Russian.

[89] H.J. Stetter. Analysis of zero cluster in multivariate polynomial systems. In R. Caviness, editor, *Proceedings of ISSAC 96*, pages 127–136, 1996.

[90] G. W. Stewart. On the solution of block Hessenberg systems. *Numer. Lin. Algebra with Appl.*, 2:287–296, 1995.

[91] G. T. Wilson. Factorization of the covariance generating function of a pure moving-average process. *SIAM J. Num. Anal.*, 6:1–7, 1969.

[92] P. Yalamov and V. Pavlov. Stability of the block cyclic reduction. *Linear Algebra Appl.*, 249:341–358, 1996.

[93] D. C. Youla and N. N. Kazanjian. Bauer-type factorization of positive matrices and the theory of matrix polynomials orthogonal on the unit circle. *IEEE Trans. Circuits Syst.*, 25:57–69, 1978.

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

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