Riv. Mat. Univ. Parma (6) **3** (2000), 259-291

LUCA GEMIGNANI (*)

A numerical approach to the solution of stable resultant linear systems (**)

dedicated to the memory of Giulio Di Cola

1 - Introduction

The problem we consider here is the one of solving $(m+n)\times(m+n)$ resultant linear systems of the form

$$[T_n[\boldsymbol{c}]\,|\,T_m[\widehat{\boldsymbol{a}}]]\,\boldsymbol{x}=\boldsymbol{b}\,, \tag{1}$$

where $\boldsymbol{c}=[1, c_1, \dots, c_m]^T$ and $\widehat{\boldsymbol{a}}=[a_n, a_{n-1}, \dots, 1]^T$ are the coefficient vectors of two given Laurent polynomials

$$c(z)=1+\sum_{i=1}^{m} c_i z^{-i}, \qquad \widehat{a}(z)=\sum_{i=0}^{n-1} a_{n-i} z^{-i}+z^{-n},$$

and, moreover, for any polynomial $p(z)=\sum_{i=0}^{r} p_i z^{-i}$, the associated $(j+r)\times j$ triangular Toeplitz matrix $T_j[\boldsymbol{p}]$ is given by:

$$T_j[\boldsymbol{p}]^T=\begin{bmatrix} p_0 & p_1 & \cdots & \cdots & p_r & 0 & \cdots & 0\\ 0 & p_0 & & \cdots & p_{r-1} & p_r & \ddots & \vdots\\ \vdots & \ddots & \ddots & & & & \ddots & 0\\ 0 & \cdots & 0 & p_0 & \cdots & \cdots & & p_r \end{bmatrix}.$$

(*) Dip. Mat. Univ. Pisa, Via Buonarroti 2, 56127 Pisa, Italy.
(**) Received July 17, 2000. AMS classification 65F05.

If we introduce the Laurent polynomials $b(z)$, $x_+(z)$ and $x_-(z)$ defined by the coefficients of $\boldsymbol{b} = [b_{n-1}, \ldots, b_0, \ldots, b_{-m}]^T$ and $\boldsymbol{x} = [x_{n-1}, \ldots, x_0, \ldots, x_{-m}]^T$ according to

$$b(z) = \sum_{i=-m}^{n-1} b_i z^{-i}, \qquad x_+(z) = \sum_{i=0}^{n-1} x_i z^{-i}, \qquad x_-(z^{-1}) = \sum_{i=-m}^{-1} x_i z^{-i},$$

then it can be easily seen that (1) reduces to the polynomial equation

$$c(z^{-1})\, x_+(z) + a(z)\, x_-(z^{-1}) = b(z)\,, \tag{2}$$

where $a(z) = z^{-n}\widehat{a}(z^{-1})$. Throughout this paper we will refer to (2) as to the polynomial counterpart of (1).

Systems of the type (1) together with their polynomial formulation (2) are often encountered in many diverse applicative fields. In particular, they are a basic computation in computer algebra where the sub-resultant theory [**17**], [**13**], provides a fundamental tool in the construction of polynomial *GCD*-algorithms. Based on this close connection between resultant matrices and the Euclidean scheme computation, over the years several different fast and superfast solvers had appeared (see [**3**], [**4**], [**21**], [**22**], [**40**] and [**15**]). However, due to the exponential growth of the coefficients of Euclidean remainders, these solution methods may suffer from ill-conditioning problems and numerical instabilities and, therefore, their applicability is usually confined within the framework of symbolic computation and/or multi-precision software packages.

On the other hand, the solution of (1) and (2) often arises also in many relevant applicative and industrial problems of data modeling, control theory and digital signal processing, where the primary focus is on the study of process dynamics by means of numerical procedures. These include time series analysis, Wiener filtering, noise variance estimation, covariance matrix computations and the study of multichannel systems (see [**19**], [**18**], [**20**], [**43**], [**1**] and [**33**]). In all these applications, the considered input-output models are generally represented by means of transfer functions given by the ratio of two polynomials in z^{-1}. In addition, the transfer functions are supposed to be stable, which means that all the roots of the denominator polynomials lie inside the unit circle in the complex plane. These observations thus motivate the search of effective computational methods solving (1) and (2) under the auxiliary conditions that both $a(z)$ and $c(z)$ have all their roots inside the unit circle. If this is true, then we will refer to the system (1) as to a stable resultant linear system (*SRLS*).

In this paper a new purely numerical approach to the solution of stable resultant linear systems is devised. It relies on a blend of ideas from structured numerical linear algebra, computational complex analysis and linear operator theory; moreover, it provides interesting links to the inverse problem of finding the coefficients of the spectral factors $a(z)$ and $c(z)$ from those of $a(z)c(z^{-1})$ (compare with [**42**], [**5**], [**26**] and [**23**]).

Specifically, the original problems (1) and (2) are reduced to that one of determining the central coefficients of the reciprocal of the Laurent polynomial $a(z)\,c(z^{-1})$ given in its factored form. This latter computation is then re-phrased in the matrix setting as the evaluation of the central entries in the central column of the inverse of a bi-infinite Toeplitz matrix defining an invertible bounded linear operator T acting on the $l^2(\mathbb{Z})$ Hilbert space.

In order to approximate such entries of T^{-1}, the cyclic reduction process, originally introduced in [**14**] for the solution of partial differential equations and, more recently, adjusted in [**8**], [**7**] and [**6**] for solving certain queueing problems, can be used. It generates a sequence of invertible linear operators $\{T^{(s)}\}_{s \in \mathbb{N}}$, $T^{(0)} = T$, approaching a block diagonal bi-infinite matrix from which the sought entries can be retrieved. We are able to provide a description of the cyclic reduction process where the Wiener-Hopf factorizations of $T^{(s)}$ are recursively updated starting from that one of T which is explicitly determined by the coefficients of the known factors of the given Laurent polynomial. In this way, the updating relations involve the powers of the Frobenius matrices associated with the spectral factors and they can be performed in a very efficient way by means of fast polynomial arithmetic using FFTs.

Based on these achievements, effective finite and iterative schemes for computing the sought coefficients of $1/a(z)\,c(z^{-1})$ are obtained. They are the core of our composite methods for the solution of the resultant linear systems (1) generated by a pair of stable polynomials. Such methods can be implemented at the cost of the better existing procedures and, in particular, they lead to fast and superfast solvers. Moreover, our numerical experience indicates that they generally have quite good stability properties and, therefore, they are efficient and numerically reliable.

The paper is organized in the following way. In section 2 we provide a basic description of our approach to the solution of (1) and (2) by means of reciprocation of Laurent polynomials. In section 3 a matrix algorithm for this latter computation is introduced and analyzed in the framework of structured numerical linear algebra. In section 4, it is reformulated to take into account the specific features of the present case and, in this way, simplified versions are found. In section 5 we report the results of the numerical experiments performed with *Mathematica*^{TM} implementations of our algorithms. Conclusions and possible developments are finally drawn in section 6.

2 - Solving stable resultant linear systems and its polynomial counterpart

Let $a(z)$ and $c(z)$ denote two real stable polynomials of degree n and m, respectively:

$$(3)\qquad \begin{aligned} a(z) &= \prod_{i=1}^{n}(1-\alpha_i z^{-1}) = 1 + \sum_{i=1}^{n} a_i z^{-i}, \qquad a_i \in \mathbb{R}, \qquad 0 < |\alpha_i| < 1\,, \\ c(z) &= \prod_{i=1}^{m}(1-\gamma_i z^{-1}) = 1 + \sum_{i=1}^{m} c_i z^{-i}, \qquad c_i \in \mathbb{R}, \qquad 0 < |\gamma_i| < 1\,. \end{aligned}$$

For the sake of notational convenience, assume that the zeros α_i and γ_i of $a(z)$ and $c(z)$ are ordered so that

$$(4)\qquad 0 < |\alpha_1| \leqslant |\alpha_2| \leqslant \ldots \leqslant |\alpha_n| < 1, \qquad 0 < |\gamma_1| \leqslant |\gamma_2| \leqslant \ldots \leqslant |\gamma_m| < 1$$

holds.

In this paper we address the problem of efficiently computing the solution $\boldsymbol{x} \in \mathbb{R}^{m+n}$ of the linear system

$$(5)\qquad [T_n[\boldsymbol{c}]\,|\,T_m[\widehat{\boldsymbol{a}}]]\,\boldsymbol{x} = \boldsymbol{b}\,, \qquad \boldsymbol{x} \in \mathbb{R}^{m+n}$$

where $\widehat{a}(z) = z^{-n} a(z^{-1})$ and, moreover, for any polynomial $p(z) = \sum_{i=0}^{r} p_i z^{-i}$ with coefficient vector $\boldsymbol{p} = [p_0, p_1, \ldots, p_r]^T$, we define the associated $(j+r) \times j$ triangular Toeplitz matrix $T_j[\boldsymbol{p}]$ as follows:

$$(6)\qquad T_j[\boldsymbol{p}]^T = \begin{bmatrix} p_0 & p_1 & \cdots & \cdots & p_r & 0 & \cdots & 0 \\ 0 & p_0 & & \cdots & p_{r-1} & p_r & \ddots & \vdots \\ \vdots & \ddots & \ddots & & & & \ddots & 0 \\ 0 & \cdots & 0 & p_0 & \cdots & \cdots & & p_r \end{bmatrix}.$$

The coefficient matrix $[T_n[\boldsymbol{c}]\,|\,T_m[\widehat{\boldsymbol{a}}]]$ of (5) is usually named the Sylvester resultant matrix generated by $\widehat{a}(z)$ and $c(z)$. It is well known that its determinant can be explicitly expressed in terms of the zeros of its polynomial generators [**1**], namely,

$$\det[T_n[\boldsymbol{c}]\,|\,T_m[\widehat{\boldsymbol{a}}]] = \prod_{i=1}^{n}\prod_{j=1}^{m}(1-\alpha_i\gamma_j)\,.$$

Hence, the stability assumption (4) immediately implies that the coefficient matrix of (5) is nonsingular and, therefore, for any fixed known vector $\boldsymbol{b}$, the solution $\boldsymbol{x}$ of (5) is uniquely determined.

We state below the precise description of the considered matrix problem of solving resultant linear systems generated by a pair of stable polynomials.

Problem 1 [Solving a stable resultant linear system]. *Let $a(z)$ and $c(z)$ be two real polynomials in z^{-1} of degree n and m, respectively, whose zeros lie inside the unit circle. For any given real vector $\boldsymbol{b} \in \mathbb{R}^{m+n}$, compute the entries x_i, $-m \leq i \leq n-1$, of the solution $\boldsymbol{x}$ of the $(m+n) \times (m+n)$ linear system*

$$[T_n[\boldsymbol{c}] \,|\, T_m[\widehat{\boldsymbol{a}}]]\, \boldsymbol{x} = \boldsymbol{b} \,,$$

where we set $\widehat{a}(z) = z^{-n} a(z^{-1})$.

In this paper we develop both finite and iterative procedures for solving Problem 1. Our solution methods rely upon an equivalent polynomial formulation of Problem 1 obtained by expressing (5) as a polynomial equation. To do this, let us first introduce the Laurent polynomials $x(z)$ and $b(z)$ defined, respectively, by the coefficients of the solution vector $\boldsymbol{x} = [x_{n-1}, \ldots, x_0, \ldots, x_{-m}]^T$ and by the coefficients of the known vector $\boldsymbol{b} = [b_{n-1}, \ldots, b_0, \ldots, b_{-m}]^T$ according to the following rules:

$$b(z) = \sum_{i=-m}^{n-1} b_i z^{-i},$$

$$x_+(z) = \sum_{i=0}^{n-1} x_i z^{-i}, \qquad x_-(z^{-1}) = \sum_{i=-m}^{-1} x_i z^{-i}, \qquad x(z) = x_+(z) + x_-(z^{-1}).$$

Then it can be easily seen that (5) reduces to the polynomial equation

$$c(z^{-1})\, x_+(z) + a(z)\, x_-(z^{-1}) = b(z)\,. \tag{7}$$

The next statement gives the equivalent polynomial version of Problem 1. Here and hereafter we will denote by $\mathcal{L}_s^r$ the vector space of real Laurent polynomials of the form $\sum_{i=-s}^{r} p_i z^i$.

Problem 2 [Solving an equivalent polynomial equation]. *Given two stable polynomials $a(z) \in \mathcal{L}_n^0$ and $c(z) \in \mathcal{L}_m^0$, then, for any $b(z) \in \mathcal{L}_{n-1}^m$, determine the coefficients of $y_1(z) \in \mathcal{L}_{n-1}^0$ and $y_2(z) \in \mathcal{L}_{m-1}^0$ such that*

$$c(z^{-1})\, y_1(z) + a(z) z y_2(z^{-1}) = b(z)\,.$$

Our approach to the solution of Problem 2 is based on the observation that (7) can be equivalently rewritten as

$$\frac{x_+(z)}{a(z)} + \frac{x_-(z^{-1})}{c(z^{-1})} = \frac{b(z)}{a(z)c(z^{-1})} . \tag{8}$$

It is worth pointing out that the rational function $g_+(z) = x_+(z)/a(z)$ is analytic in the domain

$$G_+ = \{z \in \mathbb{C} : |z| > |\alpha_n|\}$$

and, therefore, in G_+ it admits a convergent Taylor expansion in powers of z^{-1}, namely,

$$g_+(z) = \sum_{i=0}^{\infty} g_i^{(+)} z^{-i}, \qquad \forall z \in G_+ . \tag{9}$$

Conversely, the rational function $g_-(z) = x_-(z^{-1})/c(z^{-1})$, where $g_-(0) = 0$, is analytic in the disk

$$G_- = \{z \in \mathbb{C} : |z| < |1/\gamma_m|\} .$$

Hence, in G_- it has a convergent Taylor expansion of the form,

$$g_-(z) = \sum_{i=1}^{\infty} g_i^{(-)} z^{i}, \qquad \forall z \in G_- . \tag{10}$$

By replacing (9) and (10) into the equation (8), we find that the function on the right hand side of (8) can be represented in the annulus $G = G_+ \cap G_-$ by the convergent Laurent series

$$g_+(z) + g_-(z) = \sum_{i=0}^{\infty} g_i^{(+)} z^{-i} + \sum_{i=1}^{\infty} g_i^{(-)} z^{i}, \qquad \forall z \in G .$$

It is now clear that the function $1/(a(z)c(z^{-1}))$ also possesses a Laurent expansion in G,

$$\frac{1}{a(z)\, c(z^{-1})} = \sum_{i \in \mathbb{Z}} h_i z^i, \qquad \forall z \in G ,$$

and, then, the same holds for $b(z)/(a(z)\, c(z^{-1}))$. From the uniqueness of the Laurent series of an analytic function in a given annulus [**30**], we may therefore conclude that

$$(11)\qquad \begin{bmatrix} h_0 & h_{-1} & \cdots & \cdots & \cdots & h_{1-n-m} \\ h_1 & h_0 & h_{-1} & & & h_{2-n-m} \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ h_{n+m-2} & \cdots & \cdots & h_1 & h_0 & h_{-1} \\ h_{n+m-1} & \cdots & \cdots & \cdots & h_1 & h_0 \end{bmatrix} \begin{bmatrix} b_{n-1} \\ \vdots \\ b_0 \\ b_{-1} \\ \vdots \\ b_{-m} \end{bmatrix} = \begin{bmatrix} g_{n-1}^{(+)} \\ \vdots \\ g_0^{(+)} \\ g_1^{(-)} \\ \vdots \\ g_m^{(-)} \end{bmatrix}.$$

The observation that the coefficients of $x_+(z)$ and $x_-(z)$ can be retrieved from $g_0^{(+)}, \ldots, g_{n-1}^{(+)}$ and from $g_1^{(-)}, \ldots, g_m^{(-)}$, respectively, finally leads to the following procedure for the solution of Problem 2.

Procedure **SolveSRLS**

1. Evaluate the central coefficients $h_{-m-n+1}, \ldots, h_{m+n-1}$ of the Laurent expansion of the reciprocal of $a(z)c(z^{-1})$.

2. Compute the first n coefficients of $g_+(z)$ and the first m coefficients of $g_-(z)$ by means of relations (11). Set

$$\widehat{g}_+(z) = g_+(z) (\bmod z^{-n}), \qquad \widehat{g}_-(z) = g_-(z) (\bmod z^{m+1}).$$

3. Determine the coefficients of $x_+(z) = y_1(z)$ such that

$$x_+(z) = a(z)\, \widehat{g}_+(z) \qquad (\bmod z^{-n}),$$

and, analogously, find the coefficients of $x_-(z^{-1}) = zy_2(z^{-1})$ by

$$x_-(z^{-1}) = c(z^{-1})\, \widehat{g}_-(z) \qquad (\bmod z^{m+1}).$$

Since the steps 2 and 3 of **SolveSRLS** essentially amount to perform polynomial multiplications, for which fast schemes based on FFTs can be applied at the cost of $O((m+n)\log(m+n))$ arithmetic operations, it is quite obvious that the most expensive computation of **SolveSRLS** is to be carried out at the step 1. Roughly speaking, this means that, from a computational point of view, the previous procedure reduces the solution of Problem 1 and Problem 2 to the evaluation of certain central coefficients of the Laurent series of the reciprocal of $a(z)\, c(z^{-1})$. We state below the precise formulation of this latter computational problem:

Problem 3 [Reciprocation of Laurent polynomials in factored form]. *Given an odd integer k and two stable polynomials* $a(z) \in \mathcal{L}_n^0$ *and* $c(z) \in \mathcal{L}_m^0$ *whose zeros*

α_i, $1 \leq i \leq n$, and γ_i, $1 \leq i \leq m$, satisfy (4), *compute the k central coefficients $h_{-(k-1)/2}, \ldots, h_{-(k-1)/2}$ of the Laurent series of $1/(a(z)c(z^{-1}))$ in the annulus $G = \{z \in \mathbb{C} : |\alpha_n| < |z| < |\gamma_m|^{-1}\}$.*

This problem is a specific instance of the more general issue of finding the central coefficients of the Laurent expansion of the reciprocal of a Laurent polynomial having no zeros on the unit circle in the complex plane. In the next section we will first introduce a matrix analogue of this latter polynomial computation involving manipulations with bi-infinite Toeplitz matrices (operators). Then, we will show that the knowledge of the spectral factorization of the given Laurent polynomial can be exploited to produce effective computational schemes for solving Problem 3. By complementing Procedure **SolveSRLS** with these algorithms, we thus obtain a family of composite procedures for the efficient solution of stable resultant linear systems.

3 - Reciprocation of Laurent polynomials in the framework of structured numerical linear algebra

This section is concerned with the problem of the reciprocation of Laurent polynomials. We first provide a solution of the general problem based on the cyclic reduction process and, then, we specialize it for the more specific case, treated in Problem 3, where the spectral factorization of the input polynomial is assumed to be known.

By virtue of the Cauchy integral representation of the Laurent coefficients of a meromorphic function [**30**], it follows that such coefficients can be computed by sampling the function in sufficiently many equidistant points on a circle and then by applying a discrete Fourier transform. This approach was considered in [**38**] and applied in [**36**] for the fast evaluation of contour integrals of rational functions. An implementation of this scheme needs the preliminary selection of the number of points and of the radius of the integration circle. Both of them are crucial parameters for the convergence and for the computational performance of the resulting quadrature procedure. A large number of points slows down the computation whereas big and small radii can lead to numerical instabilities.

The approach taken here proceeds in a very different way without requiring any critical initialization. A numerical comparison between the diverse methods for computing Laurent coefficients of meromorphic functions is planned in a subsequent work.

Let $p(z) = \sum_{i=-n}^{m} p_i z^i \in \mathcal{L}_n^m$ be a real Laurent polynomial of degree

$r = \max\{n, m\}$. Assume that $p(z)$ can be factored as

$$p(z) = \xi a(z)\, c(z^{-1}), \tag{12}$$

where the factors $a(z)$ and $c(z)$ satisfy (3) and (4) and $\xi = p_m/c_m$. Without loss of generality, throughout this paper $\xi = 1$ will be always assumed. If is so, then $a(z)$ and $c(z)$ are stable Laurent polynomials and the considered factorization of $p(z)$ is known as its spectral factorization.

The application that associates the Laurent polynomial $p(z)$ with the bi-infinite band Toeplitz matrix $T[p]$,

$$T[p] = \left[\begin{array}{cccc|ccccc} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & p_m & \cdots & p_0 & \cdots & p_{-n} & 0 & \cdots & \cdots \\ \hline \cdots & 0 & p_m & \cdots & p_0 & \cdots & p_{-n} & 0 & \cdots \\ \cdots & \cdots & 0 & p_m & \cdots & p_0 & \cdots & p_{-n} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{array}\right], \tag{13}$$

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. The key role played by the spectral factorization (12) is made clear by the following remarkable fact: it induces a very special triangular factorization — named Wiener-Hopf factorization — of the corresponding bi-infinite banded Toeplitz matrix $T[p]$.

The description of such an intimate connection between matrix and polynomial factorizations requires a preliminary introduction to linear algebra manipulations with infinite Toeplitz matrices ([**35**] and [**12**] give a detailed exposition of these subjects). The appropriate framework for studying infinite matrices is the theory of linear operators acting on Banach spaces of sequences. In general, we are concerned with the Hilbert space $l^2(\mathbb{Z})$ of real square summable sequences $\boldsymbol{w} = \{w_k\}_{k\in\mathbb{Z}}$ with the norm

$$\|\boldsymbol{w}\|^2 = \sum_{k\in\mathbb{Z}} w_k^2.$$

A bounded linear operator on $l^2(\mathbb{Z})$ can always be represented by a bi-infinite matrix $T = (t_{i,j})$, $i, j \in \mathbb{Z}$, in such a way that

$$T\boldsymbol{w} = \left\{\sum_{j\in\mathbb{Z}} t_{k,j} w_j\right\}_{k\in\mathbb{Z}}, \qquad \forall \boldsymbol{w} \in l^2(\mathbb{Z}),$$

with

$$\|T\| = \sup_{\|\boldsymbol{w}\| \leq 1} \|T\boldsymbol{w}\| .$$

In this way, the algebra of infinite matrix representations of bounded linear operators acting on $l^2(\mathbb{Z})$ provides a natural extension of the usual matrix algebra.

In the case where $T = (t_{i-j})$, $i, j \in \mathbb{Z}$, exhibits the Toeplitz structure, then many properties of T can be expressed in terms of corresponding properties of the associated symbol

$$t(z) = \sum_{k \in \mathbb{Z}} t_k z^k .$$

In particular, since $p(z)$ is a continuous function with no zeros on the unit circle, it follows that $T[p]$ defines an invertible bounded linear operator acting on $l^2(\mathbb{Z})$ with norm

$$\|T[p]\| = \max_{|z|=1} |p(z)| .$$

The inverse of $T[p]$ is the bi-infinite Toeplitz matrix $T[1/p] = (h_{i-j})$, $i, j \in \mathbb{Z}$, where

$$\frac{1}{p(z)} = \sum_{k \in \mathbb{Z}} h_k z^k , \qquad \forall z : |\alpha_n| < |z| < |\gamma_m|^{-1},$$

and, moreover,

$$\|(T[p])^{-1}\| = \left\| T\left[\frac{1}{p}\right] \right\| = \max_{|z|=1} \frac{1}{|p(z)|} = \frac{1}{\min_{|z|=1} |p(z)|} .$$

Further, the spectral factorization (12) induces the following Wiener-Hopf factorization of $T[p]$,

$$T[p] = T[\hat{c}]\, T[a] = T[a]\, T[\hat{c}] , \tag{14}$$

where $\hat{c}(z) = c(z^{-1})$. Note that $T[\hat{c}]$ is a bi-infinite lower triangular Toeplitz matrix whereas $T[a]$ is a bi-infinite upper triangular Toeplitz matrix; therefore, (14) represents a factorization of $T[p]$ into the product of triangular factors. Obviously, other different triangular factorizations of $T[p]$ are formally possible but (14) is the only one where the triangular factors are themselves invertible operators on $l^2(\mathbb{Z})$.

As an application of the preceding theory, one may consider the block solution of the linear system

$$T[p]\,X = E\,, \tag{15}$$

where X and E are bi-infinite vectors with $n+m$ columns $\in l^2(\mathbb{Z})$, $X, E: \mathbb{Z} \to \mathbb{R}^{(n+m)\times(n+m)}$, with $E(0) = I$ and, otherwise, $E(j) = 0$ for $j \neq 0$. Then, it can be easily seen that $X(0) = (h_{i-j})$ with $1 \leq i, j \leq n+m$. Hence, the first and the last column of $X(0)$ provide the sought coefficients of the reciprocal of the Laurent polynomial $p(z)$.

Remark 1. *By using the equivalence between Problem* 1 *and its polynomial counterpart, one may deduce the invertibility of the matrix* $X(0)$. *We proceed by contradiction. If* $X(0)$ *has a non-trivial kernel, then, in the view of relations* (11), *one finds that there exists a nonzero polynomial* $b(z)$ *such that* $x_+(z) = 0$ *and* $x_-(z) = 0$ *is a solution of the corresponding equation* (7).

Our approach to the solution of (15) is based upon the use of the cyclic reduction scheme, originally introduced in [**14**] for the solution of partial differential equations and, more recently, adjusted in [**8**], [**7**] and [**6**] for solving certain queueing problems. By using a block partitioning of $T[p]$ into blocks P_k, $k = -1, 0, 1$, $P_k = P_k^{(0)} = (p_{i-j+k(n+m)})$, $1 \leq i, j \leq n+m$, $p_k = 0$ if $k > m$ or $k < -n$, we are able to reduce the system (15) to a block tridiagonal form with a block Toeplitz structure, namely,

$$\begin{cases} P_1^{(0)} X(h-2) + P_0^{(0)} X(h-1) + P_{-1}^{(0)} X(h) = E(h-1) \\ P_1^{(0)} X(h-1) + P_0^{(0)} X(h) + P_{-1}^{(0)} X(h+1) = E(h), \qquad h \in \mathbb{Z}\,. \\ P_1^{(0)} X(h) + P_0^{(0)} X(h+1) + P_{-1}^{(0)} X(h+2) = E(h+1) \end{cases}$$

Assume now that $P_0^{(0)}$ is nonsingular; then, by multiplying the first equation by $P_1^{(0)}(P_0^{(0)})^{-1}$ and the last equation by $P_{-1}^{(0)}(P_0^{(0)})^{-1}$ and by subtracting them from the second one, we obtain

$$P_1^{(1)} X(h-2) + P_0^{(1)} X(h) + P_{-1}^{(1)} X(h+2) = E(h)\,, \qquad h = 2l,\ l \in \mathbb{Z}\,,$$

where we set

$$\begin{aligned} P_1^{(1)} &= -P_1^{(0)}(P_0^{(0)})^{-1} P_1^{(0)} \\ P_0^{(1)} &= P_0^{(0)} - P_1^{(0)}(P_0^{(0)})^{-1} P_{-1}^{(0)} - P_{-1}^{(0)}(P_0^{(0)})^{-1} P_1^{(0)}. \\ P_{-1}^{(1)} &= -P_{-1}^{(0)}(P_0^{(0)})^{-1} P_{-1}^{(0)} \end{aligned} \tag{16}$$

These formulae allow us to define a new bounded block tridiagonal operator $T^{(1)}$, with $T^{(0)} = T[p]$, given by

$$T^{(1)} = (P^{(1)}_{i-j})_{i,j \in \mathbb{Z}}, \qquad P^{(1)}_{i-j} = 0 \text{ if } |i-j| > 1 ,$$

such that

$$T^{(1)} X^{(1)} = E , \tag{17}$$

where $X^{(1)}(k) = X^{(0)}(2k) = X(2k)$, $k \in \mathbb{Z}$.

The iterative application of formulae (16) defines the cyclic reduction process for the approximation of $X(0)$. The structure properties of the matrices $P_i^{(j)}$ generated by this process as well as its convergence behavior have been widely investigated in the papers [8] and [7]. In what follows, we will consider a different look at the cyclic reduction algorithm which is here used as a means for the iterative construction of the Wiener-Hopf factorization of $T^{(i)}$ rather than of its block entries $P_k^{(i)}$, $k = -1, 0, 1$. This has the important advantage of replacing (16) with a set of formulae involving terms explicitly related to the given spectral factors $a(z)$ and $c(z^{-1})$ of $p(z)$.

Suppose that $T[\hat{c}]$ and $T[a]$ are partitioned commensurably with $T[p]$, that is, we set $T[\hat{c}] = (L_i^{(0)})$, $L_i^{(0)} \in \mathbb{R}^{(n+m)\times(n+m)}$, where $L_i^{(0)} = 0$ if $i < 0$ or $i > 1$, and $T[a] = (U_i^{(0)})$, $U_i^{(0)} \in \mathbb{R}^{(n+m)\times(n+m)}$, where $U_i^{(0)} = 0$ if $i < -1$ or $i > 0$. In addition, both the upper triangular matrix $U_0^{(0)}$ and the lower triangular matrix $L_0^{(0)}$ are nonsingular with unit diagonal entries.

In this way, the Wiener-Hopf factorization (14) of $T^{(0)}$ can be rewritten into a block form as

$$\begin{bmatrix} \ddots & & & \\ \ddots & I & & \\ & F_1^{(0)} & \ddots & \\ & & \ddots & \end{bmatrix} \begin{bmatrix} \ddots & & \\ & D_0^{(0)} & \\ & & \ddots \end{bmatrix} \begin{bmatrix} \ddots & \ddots & & \\ & I & G_{-1}^{(0)} & \\ & & \ddots & \ddots \end{bmatrix} = F^{(0)} D^{(0)} G^{(0)},$$

where $F_1^{(0)} = L_1^{(0)} (L_0^{(0)})^{-1}$, $D_0^{(0)} = L_0^{(0)} U_0^{(0)}$ and $G_{-1}^{(0)} = (U_0^{(0)})^{-1} U_{-1}^{(0)}$. Furthermore, the commuting property of (14) translates into the block relations:

$$U_0^{(0)} L_0^{(0)} + U_{-1}^{(0)} L_1^{(0)} = L_0^{(0)} U_0^{(0)} + L_1^{(0)} U_{-1}^{(0)}, \tag{18}$$

and

$$U_0^{(0)} L_1^{(0)} = L_1^{(0)} U_0^{(0)}, \quad U_{-1}^{(0)} L_0^{(0)} = L_0^{(0)} U_{-1}^{(0)}. \tag{19}$$

Observe that the equalities (19) also follow from the fact that upper (lower) triangular Toeplitz matrices commute.

The matrices $F_1^{(0)}$ and $G_{-1}^{(0)}$ show very interesting properties related to the zeros of $\hat{c}(z)$ and $a(z)$. In particular, the next result says that they coincide with the $(m+n)$-th power of a permuted version of the Frobenius matrix associated with $z^{n+m}c(z)$ and $z^{n+m}a(z)$, respectively. Since these polynomials have all their zeros inside the unit circle, it follows that both $G^{(0)}$ and $F^{(0)}$ define invertible operators.

Theorem 2. *Let $C(z^{n+m}a(z))$ denote the Frobenius matrix of order $n+m$ associated with the polynomial $z^{n+m}a(z)$, that is,*

$$C(z^{n+m}a(z)) = \begin{bmatrix} 0 & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ 0 & \cdots & -a_n & \cdots & -a_1 \end{bmatrix}.$$

Analogously, let $C(z^{n+m}c(z))$ be the Frobenius matrix of order $n+m$ associated with the polynomial $z^{n+m}c(z)$. Finally, let J be the $(n+m)\times(n+m)$ permutation matrix with unit anti-diagonal entries. Then we have that

$$F_1^{(0)} = -(JC(z^{n+m}c(z))^T J)^{n+m}, \qquad G_{-1}^{(0)} = -(JC(z^{n+m}a(z))J)^{n+m}. \tag{20}$$

Hence, the bi-infinite triangular matrices $F^{(0)}$ and $G^{(0)}$ are invertible and their inverses are given by:

$$((F^{(0)})^{-1})_{i,j} = ((JC(z^{n+m}c(z))^T J)^{m+n})^{i-j}, \qquad i \geq j,\ i, j \in \mathbb{Z},$$

and

$$((G^{(0)})^{-1})_{i,j} = ((JC(z^{n+m}a(z))J)^{m+n})^{j-i}, \qquad j \geq i,\ j, i \in \mathbb{Z}.$$

Proof. We shall consider the matrix $G^{(0)}$ only, since the results for $F^{(0)}$ can be proven in exactly the same way. Let $a_\varepsilon(z) = a(z) + \varepsilon z^{-n-m}$ be a Laurent polynomial of degree $n+m$, where ε is chosen in a neighborhood of the origin of the complex plane in such a way to guarantee that $a_\varepsilon(z)$ has $n+m$ distinct zeros denoted by $\alpha_{\varepsilon,1}, \ldots, \alpha_{\varepsilon,n+m}$. These zeros can be grouped so that, as ε approaches 0, $\alpha_{\varepsilon,i}$ converges to α_i for $1 \leq i \leq n$, whereas $\alpha_{\varepsilon,i}$ converges to 0 for $i > n$. The $(n+m) \times 2(n+m)$ Toeplitz matrix $T_{n+m}[\boldsymbol{a}_\varepsilon]^T$ of (6) satisfies

$$T_{n+m}[\boldsymbol{a}_\varepsilon]^T = [\, U_0^{(0)} \;\; U_{-1}^{(0)} + \varepsilon I \,],$$

from which it follows that

$$(\alpha_{\varepsilon,j}^{n+m} U_0^{(0)} + (U_{-1}^{(0)} + \varepsilon I))[\alpha_{\varepsilon,j}^{n+m-1}, \ldots, 1]^T = \mathbf{0}$$

holds for $1 \leq j \leq n+m$. This implies that $(U_0^{(0)})^{-1}(U_{-1}^{(0)} + \varepsilon I)$ coincides with $-J(C(z^{n+m} a_\varepsilon(z)))^{n+m} J$ and this matrix converges to $-J(C(z^{n+m} a(z)))^{n+m} J$ as ε goes to 0.

The existence of the linear inverse of $(G^{(0)})^{-1}$ now follows from Banach's theorem [**35**]. In fact, in view of the matrix theory for the finite dimensional case, one obtains that, for any σ satisfying $|\alpha_n|^{n+m} < \sigma < 1$, there exists a suitable matrix norm $\|\cdot\|_*$ such that $\|G_{-1}^{(0)}\|_* \leq \sigma$. Moreover, for any matrix norm $\|\cdot\|$ on $\mathbb{R}^{(n+m)\times(n+m)}$ there exists a positive constant C satisfying $\|B\| \leq C\|B\|_*$, for any $B \in \mathbb{R}^{(n+m)\times(n+m)}$. Thus, by setting $G^{(0)} = I - \widehat{G}^{(0)}$, we find that, for any integer h, it holds

$$\|(\widehat{G}^{(0)})^h\| \leq C\sigma^h. \tag{21}$$

This finally implies that $G^{(0)}$ is invertible and its inverse is given by

$$(G^{(0)})^{-1} = \sum_{i=0}^{\infty} (\widehat{G}^{(0)})^i. \qquad \blacksquare$$

Theorem 2 describes the Wiener-Hopf factorization of $T^{(0)} = T[p]$ in terms of the Frobenius matrices associated with the spectral factors of $p(z)$. The next result shows that this decomposition can be used as the initial guess of an iterative scheme which produces the sequence of the triangular factorizations of the linear operators $T^{(s)}$ generated by the cyclic reduction process.

Theorem 3. *Let* $\{T^{(s)}\}_{s\in\mathbb{N}}$, $T^{(s)} = (P_{i,j}^{(s)})$, *be the sequence of linear operators generated by the cyclic reduction process starting from* $T^{(0)} = T[p]$ *by means of relations* (16), *where we assume that all the matrices to be inverted are nonsingular and, therefore, the process does not break down at any step. We have that* $T^{(s)}$, $s \in \mathbb{N}$, *is a bi-infinite block Toeplitz matrix in block tridiagonal* *form and, moreover, it has a block triangular factorization of the form*

$$T^{(s)} = F^{(s)} D^{(s)} G^{(s)}, \tag{22}$$

where $F^{(s)}$ *is block lower bidiagonal,* $G^{(s)}$ *is block upper bidiagonal and* $D^{(s)}$ *is block diagonal. In addition, for* $s = 1, 2, \ldots,$ *the factorization of* $T^{(s)}$ *can be constructed recursively from the one of* $T^{(s-1)}$ *according to the following rules. It holds:*

$$F^{(s)} = (F^{(s)}_{i-j})_{i,j\in\mathbb{Z}}, \qquad F^{(s)}_0 = I\,, \qquad F^{(s)}_1 = -(F^{(s-1)}_1)^2,$$

$$G^{(s)} = (G^{(s)}_{i-j})_{i,j\in\mathbb{Z}}, \qquad G^{(s)}_0 = I\,, \qquad G^{(s)}_{-1} = -(G^{(s-1)}_{-1})^2,$$

and $D^{(s)} = (D^{(s)}_{i-j})_{i,j\in\mathbb{Z}}$ *with*

$$D^{(s)}_0 = D^{(s-1)}_0 - D^{(s-1)}_0 G^{(s-1)}_{-1} (D^{(s-1)}_0 + F^{(s-1)}_1 D^{(s-1)}_0 G^{(s-1)}_{-1})^{-1} F^{(s-1)}_1 D^{(s-1)}_0 .$$

Proof. The proof follows from some straightforward calculations and, without loss of generality, we may restrict ourselves to the case where $s = 1$. From the Wiener-Hopf factorization of $T^{(0)} = T[p]$, it follows that

$$\begin{aligned} &P^{(0)}_0 = D^{(0)}_0 + F^{(0)}_1 D^{(0)}_0 G^{(0)}_{-1};\\ &P^{(0)}_{-1} = D^{(0)}_0 G^{(0)}_{-1}, \qquad P^{(0)}_1 = F^{(0)}_1 D^{(0)}_0; \end{aligned} \tag{23}$$

therefore, $D^{(1)}_0$ is well defined if and only if $P^{(0)}_0$ is nonsingular or, equivalently, the first step of cyclic reduction can be performed. Hence, we are able to introduce the block Toeplitz matrix in block tridiagonal form $\widehat{T}$ defined by

$$\widehat{T} = (\widehat{T}_{i-j}) = F^{(1)} D^{(1)} G^{(1)}.$$

The theorem is so established by first replacing (23) into the formulae (16) and, then, by showing that $T^{(1)} = \widehat{T}$. For the sake of notational simplicity we omit both the superscripts and the subscripts when it is possible. In this way, we set $F^{(0)}_1 = F$, $G^{(0)}_{-1} = G$ and $D^{(0)}_0 = D$. Then, we find that

$$\begin{aligned} \widehat{T}_0 - P^{(1)}_0 &= F^2\{D - DG(D + FDG)^{-1} FD\} G^2 + -FDG + FD(D + FDG)^{-1} DG \\ &= F\{FDG - FDG(D + FDG)^{-1} FDG - D + D(D + FDG)^{-1} D\}\, G\,. \end{aligned}$$

Since we have

$$FDG - FDG(D + FDG)^{-1} FDG = D(D + FDG)^{-1} FDG\,,$$

one gets that

$$\widehat{T}_0 - P_0^{(1)} = F\{D(D+FDG)^{-1}FDG - D + D(D+FDG)^{-1}D\}\, G$$

$$= FD(D+FDG)^{-1}\{FDG - (D+FDG) + D\}\, G = 0\,.$$

Analogously, by comparing the subdiagonal block entries, we obtain that

$$\widehat{T}_1 - P_1^{(1)} = -F^2\{D - DG(D+FDG)^{-1}FD\} + FD(D+FDG)^{-1}FD$$

$$= \{-F + F^2DG(D+FDG)^{-1} + FD(D+FDG)^{-1}\}FD$$

$$= \{-F(D+FDG) + F^2DG - FD\}(D+FDG)^{-1}FD = 0\,.$$

The proof for the superdiagonal block entry follows in exactly the same way. ■

Based on this proposition, next we shall prove that the cyclic reduction process applied to $T^{(0)} = T[p]$ generates a sequence of bounded invertible linear operators on $l^2(\mathbb{Z})$ converging in norm to a suitable bi-infinite block diagonal matrix.

Theorem 4. *Let us assume that the cyclic reduction algorithm applied to $T^{(0)} = T[p]$ does not break down at any step. Then, it generates a sequence of linear operators $\{T^{(s)}\}_{s\in\mathbb{N}}$ acting on $l^2(\mathbb{Z})$ for which the following statements hold.*

1. *Each $T^{(s)}$ is a bounded linear invertible operator, and, therefore, the bi-infinite block vector $X^{(s)}$, $s\in\mathbb{N}$, defined recursively by $X^{(s)}(k) = X^{(s-1)}(2k)$, $k\in\mathbb{Z}$, with $X^{(0)} = X$ solution of* (15), *is the unique solution of the linear systems*

$$T^{(s)}X^{(s)} = E\,, \qquad s\in\mathbb{N}\,.$$

2. *Denote by I_∞ the bi-infinite identity matrix. For any σ with*

$$\max\{|\alpha_n|^{n+m},\ |\gamma_m|^{n+m}\} < \sigma < 1\,,$$

there exists a positive constant C such that

$$\|T^{(s)} - D^{(\infty)}\| \leq C\sigma^{2^s},$$

where $D^{(\infty)} = I_\infty \otimes D$, $D = (X^{(0)}(0))^{-1}$, $X^{(0)}(0) = (h_{i-j})$, $1 \leq i, j \leq n+m$.

Proof. The first part of the theorem is established by showing that the inverse of $D^{(s)}$ can be explicitly constructed starting from the one of $D^{(s-1)}$. Again, as in the proof of Theorem 3, for the sake of notational simplicity, we consider the case $s = 1$ and we omit superscripts and subscripts when it is possible. Note that $D_0^{(0)}$ is invertible since it is given by the product of invertible triangular Toeplitz matrices. As the cyclic reduction process is applicable, $D_0^{(1)} = D - DG(D + DFG)^{-1} FD$ is well defined. The matrix $W = D^{-1} + GD^{-1} F$ is our candidate to be the inverse of $D_0^{(1)}$. We have that:

$$D_0^{(1)} W = I + DGD^{-1}F - DG(D + FDG)^{-1} F - DG(D + FDG)^{-1} FDGD^{-1} F$$

$$= I + DG(D + FDG)^{-1} \{I + DDGD^{-1} - I - FDGD^{-1}\} F = I .$$

In this way, by induction, we find that

$$(D_0^{(s)})^{-1} = (D_0^{(s-1)})^{-1} + G_{-1}^{(s-1)} (D_0^{(s-1)})^{-1} F_1^{(s-1)}, \qquad s \in \mathbb{N}^+ . \tag{24}$$

This relation also implies that the matrices $(D_0^{(s)})^{-1}$ are of uniformly bounded norm. In fact, from (24) it follows that

$$\|(D_0^{(s)})^{-1}\| \leq (1 + C_1 \sigma^{2^s}) \|(D_0^{(s-1)})^{-1}\| , \qquad s \in \mathbb{N}^+ ,$$

holds for a suitable positive constant C_1. Let $h > 1$ a positive integer such that $C_1 \sigma^{2s} \leq 1$ for $s \geq h$. The increasing sequence $\{d_k\}$ defined by

$$d_k = \prod_{i=0}^{k} (1 + C_1 \sigma^{2^{h+i}}), \qquad k = 0, 1, \dots ,$$

can be bounded from above as follows:

$$d_k = EXP(\log (d_k)) = EXP\left(\sum_{i=0}^{k} \log (1 + C_1 \sigma^{2^{h+i}})\right) \leq EXP\left(C_1 \sum_{i=0}^{k} \sigma^{2^{h+i}}\right) \leq L .$$

Hence, one easily gets that

$$\|(D_0^{(s)})^{-1}\| \leq L \prod_{i=1}^{h-1} (1 + C_1 \sigma^{2^i}) \|(D_0^{(0)})^{-1}\| , \qquad s \geq 1 .$$

In the view of Theorem 3, one therefore finds that,

$$\|(T^{(s)})^{-1} - (D^{(s)})^{-1}\| \leq C_2 \sigma^{2^s},$$

where C_2 is a given positive constant. On the other hand, for any $s \in \mathbb{N}$, the matrices $(T^{(s)})^{-1}$ have the same entries in the positions i, j with $1 \leq i, j \leq n + m$. Under our notations, $(X^{(0)}(0))$ is the $(n + m) \times (n + m)$ Toeplitz matrix made up by these entries, where we recall that $X^{(0)}(0) = X^{(s)}(0) = (h_{i-j})$, for $s \geq 0$, and, moreover, $X^{(0)}(0)$ is nonsingular by virtue of Remark 1. Thus we obtain that the diagonal block $(D_0^{(s)})^{-1}$ of $(D^{(s)})^{-1}$ approaches $X^{(0)}(0)$ so that

$$\|(D_0^{(s)})^{-1} - X^{(0)}(0)\| \leq C_3 \sigma^{2^s}, \qquad s \in \mathbb{N}.$$

This means that $D^{(s)}$ are themselves uniformly bounded and, therefore, by using Theorem 3 again, we may conclude that

$$\|T^{(s)} - D^{(\infty)}\| \leq C\sigma^{2^s}, \qquad s \in \mathbb{N},$$

holds for a suitable positive constant C. ■

This theorem says that the cyclic reduction scheme applied to $T[p]$ is eligible for the task of evaluating the sought coefficients of the Laurent series of $1/p(z) = 1/(a(z)c(z^{-1}))$. In the next section, computationally effective procedures for performing this task will be developed that are based on the given description of cyclic reduction in terms of the block entries of the Wiener-Hopf factorization of $T[p]$.

4 - Efficient direct and iterative methods for reciprocation of factored Laurent polynomials

In this section the theory developed above is used in order to derive two different computational schemes for the efficient evaluation of the central coefficients $h_{-n-m+1}, \ldots, h_{n+m-1}$ of the Laurent series of the reciprocal of $a(z)c(z^{-1})$.

The first method comes from the observation that

$$(T^{(0)})^{-1} = (X_{i-j}^{(0)}) = (G^{(0)})^{-1}(D^{(0)})^{-1}(F^{(0)})^{-1}.$$

This implies that the Toeplitz matrix $X = X^{(0)}(0) = (h_{i-j})$, $1 \leq i, j \leq n+m$, can be expressed as

$$X = \sum_{i=0}^{\infty} (JC(z^{n+m}a(z))\,J)^{i(n+m)}(D_0^{(0)})^{-1}(JC(z^{n+m}c(z))^T J)^{i(n+m)}.$$

Hence, it follows that X is the unique solution of the discrete Lyapunov matrix equation

$$X - (JC(z^{n+m}a(z))\,J)^{n+m} X (JC(z^{n+m}c(z))^T J)^{n+m} = (D_0^{(0)})^{-1}. \tag{25}$$

The subsequent investigation of (25) largely follows from the results of [**29**]. By virtue of Theorem 2, this equation can equivalently be rewritten as

$$U_0^{(0)} X L_0^{(0)} - U_{-1}^{(0)} (U_0^{(0)})^{-1} U_0^{(0)} X L_0^{(0)} (L_0^{(0)})^{-1} L_1^{(0)} = I ,$$

from which, by setting $Y = U_0^{(0)} X L_0^{(0)}$, one has that Y is the unique solution of

$$Y = I + U_{-1}^{(0)} (U_0^{(0)})^{-1} Y (L_0^{(0)})^{-1} L_1^{(0)}.$$

By means of a straightforward calculation we find that

$$Y^{-1} = I - W = (I + U_{-1}^{(0)} (U_0^{(0)})^{-1} Y (L_0^{(0)})^{-1} L_1^{(0)})^{-1}$$

$$= I - U_{-1}^{(0)} (U_0^{(0)})^{-1} (Y^{-1} + (L_0^{(0)})^{-1} L_1^{(0)} U_{-1}^{(0)} (U_0^{(0)})^{-1})^{-1} (L_0^{(0)})^{-1} L_1^{(0)}.$$

This implies that the matrix W is the unique solution of the matrix equation

$$W = U_{-1}^{(0)} (L_0^{(0)} U_0^{(0)} - L_0^{(0)} W U_0^{(0)} + L_1^{(0)} U_{-1}^{(0)})^{-1} U_1^{(0)}.$$

In this way, by using the commuting properties (18) and (19), it is easily found that

$$W = (L_0^{(0)})^{-1} U_{-1}^{(0)} L_1^{(0)} (U_0^{(0)})^{-1},$$

from which we finally conclude that

$$D = X^{-1} = L_0^{(0)} U_0^{(0)} - U_{-1}^{(0)} L_1^{(0)}.$$

Summing up, we have the following:

Theorem 5. *For the* $(n+m)\times(n+m)$ *Toeplitz-like matrix* $D = (X^{(0)}(0))^{-1}$, $X^{(0)}(0) = (h_{i-j})$, $1 \leqslant i, j \leqslant n+m$, *the following representation holds*

$$D = X^{-1} = L_0^{(0)} U_0^{(0)} - U_{-1}^{(0)} L_1^{(0)}. \tag{26}$$

By means of this result, the problem of computing the entries in the first and in the last column of X^{-1} is then reduced to the problem of solving a linear system whose coefficient matrix is of the form of (26). Systems of this type are well studied and investigated in the structured numerical linear algebra framework and over the years many diverse fast and superfast algorithms for their solution have appeared [34]. By incorporating any of these efficient solvers in the procedure **SolveSRLS**, we obtain an effective computational scheme for the solution of both Problem 1 and Problem 2.

In particular, the importance of the present reduction is evident in the case where $a(z) = c(z)$ that is relevant for applications in signal and control theory

[**18**]. In fact, in this situation the matrix D results to be positive definite and, therefore, superfast algorithms can be applied in a stable way. Moreover, in this case the matrix D is also related to the Schur-Cohn test for determining the stability of a given polynomial. A numerical comparison of the performance of superfast algorithms for solving the considered definite linear system as well as the exploitation of the Schur-Cohn algorithm in such a context are presently ongoing works.

A completely different approach to the computation of the sought coefficients of $1/(a(z)c(z^{-1}))$ is based on the iterative approximation of X itself, instead of its inverse matrix D, by using relation (24). Indeed, (24) provides an iterative process quadratically converging towards the matrix X. The problem we are interested now is the efficient implementation of this process at a low computational cost.

To do this, let us introduce the displacement operator

$$\mathcal{F}_1 : \mathbb{R}^{(n+m)\times(n+m)} \longrightarrow \mathbb{R}^{(n+m)\times(n+m)},$$

defined by

$$\mathcal{F}_1(A) = A - (JC(z^{n+m}a(z))\,J)\,A(JC(z^{n+m}c(z))^T J)\,.$$

This operator can immediately be related to the more classical displacement operator

$$\mathcal{F}_2(A) = A - ZAZ^T,$$

where Z denotes the down-shift matrix of order $n+m$ given by

$$Z = [\boldsymbol{e}_2 \,|\, \cdots \,|\boldsymbol{e}_{n+m}\,|\,\boldsymbol{0}\,],$$

and $\boldsymbol{e}_i$ is the i-th column of the identity matrix I of order $n+m$. In fact, we have

$$JC(z^{n+m}a(z))\,J = Z - \boldsymbol{e}_1\boldsymbol{a}^T Z, \qquad JC(z^{n+m}c(z))\,J = Z - \boldsymbol{e}_1\boldsymbol{c}^T Z\,, \tag{27}$$

where $\boldsymbol{a} = [1, a_1, \ldots, a_n, 0, \ldots, 0]^T$ and $\boldsymbol{c} = [1, c_1, \ldots, c_m, 0, \ldots, 0]^T$.

The rank of $\mathcal{F}_j(A)$, $j = 1, 2$, is called the j-displacement rank of A. The 2-displacement rank of a matrix A can also be regarded as the smallest integer l such that A can be written as

$$A = \sum_{i=1}^{l} L_i U_i,$$

where L_i and U_i are lower and upper triangular Toeplitz matrices, respectively. More precisely, the following result holds [**11**].

Theorem 6. *Let us assume that the* $(m+n)\times(m+n)$ *matrix* A *has 2-displacement rank* l, *that is,*

$$\mathcal{F}_2(A) = A - ZAZ^T = \sum_{i=1}^{l} \boldsymbol{x}_i \boldsymbol{y}_i^T .$$

Then, we have that

$$A = \sum_{i=1}^{l} L(\boldsymbol{x}_i)\, U(\boldsymbol{y}_i),$$

where $L(\boldsymbol{x})$ *denotes the lower triangular Toeplitz matrix whose first column is* $\boldsymbol{x}$ *and* $U(\boldsymbol{y}) = L(\boldsymbol{y})^T$.

The powers of a Frobenius matrix also inherit a special displacement structure **[1]**, **[37]**.

Theorem 7. *Let* $p(z) = \sum_{i=0}^{n+m} p_i z^i$, *be a polynomial in* z *of degree* $n+m$. *For any integer* k, *let us denote by* $q^{(k)}(z)$ *and* $r^{(k)}(z)$, *respectively, the quotient and the remainder in the Euclidean division of* z^k *by* $p(z)$, *that is,*

$$z^k = q^{(k)}(z)p(z) + r^{(k)}(z), \tag{28}$$

where the degree $l(k)$ *of* $r^{(k)}(z) = \sum_{i=0}^{l(k)} r_i^{(k)} z^i$ *is smaller than the degree of* $p(z)$. *For the k-th power* $(C(p(z)))^k$ *of the Frobenius matrix* $C(p(z))$ *of order* $n+m$ *associated with* $p(z)$ *we have*

$$(C(p(z)))^k = r^{(k)}(C(p(z))) = J(U(\widehat{\boldsymbol{p}}))^{-1} J\{L(\widehat{\boldsymbol{p}})\, U(\boldsymbol{r}^{(k)}) - L(\widehat{\boldsymbol{r}}^{(k)})\, U(\boldsymbol{p})\} \tag{29}$$

where $\boldsymbol{p} = [p_0, \ldots, p_{n+m-1}]^T$, $\boldsymbol{r}^{(k)} = [r_0^{(k)}, \ldots, r_{l(k)}^{(k)}, 0, \ldots, 0]^T$, $\widehat{\boldsymbol{p}} = [p_{n+m}, \ldots, p_1]^T$, *and* $\widehat{\boldsymbol{r}}^{(k)} = [0, \ldots, 0, r_{l(k)}^{(k)}, \ldots, r_1^{(k)}]^T$.

Next two results show that each matrix $(D_0^{(s)})^{-1}$, $s \geqslant 0$, has 1-displacement rank bounded from above by 2, that is, the rank of $\mathcal{F}_1((D_0^{(s)})^{-1})$ is at most 2. Let us start by considering the inverse of the initial matrix $D_0^{(0)}$. Recall that, from Theorem 7 it follows that

$$C(z^{n+m}a(z)) = J(U_0^{(0)})^{-1}B_a, \qquad C(z^{n+m}c(z)) = J(L_0^{(0)})^{-1}B_c,$$

for suitable symmetric matrices B_a and B_c.

Theorem 8. *We have that*

$$\mathcal{F}_1((D_0^{(0)})^{-1}) = \boldsymbol{e}_1\boldsymbol{e}_1^T - (JC(z^{n+m}a(z))\,J)^{n+m}\boldsymbol{e}_1\boldsymbol{e}_1^T(JC(z^{n+m}c(z))^T J)^{n+m}.$$

Proof. By using the previous representation of the involved Frobenius matrices, we find that

$$\mathcal{F}_1((D_0^{(0)})^{-1}) = (U_0^{(0)})^{-1}(I - C(z^{n+m}a(z))^T C(z^{n+m}c(z)))(L_0^{(0)})^{-1}.$$

By replacing (27) into this formula, one finally gets that

$$\mathcal{F}_1((D_0^{(0)})^{-1}) = (U_0^{(0)})^{-1}(I - (Z - ZJ\boldsymbol{a}\boldsymbol{e}_{n+m}^T)(Z^T - \boldsymbol{e}_{n+m}\boldsymbol{c}^T JZ^T))(L_0^{(0)})^{-1}$$

$$= (U_0^{(0)})^{-1}(\boldsymbol{e}_1\boldsymbol{e}_1^T - ZJ\boldsymbol{a}\boldsymbol{c}^T JZ^T)(L_0^{(0)})^{-1} = \boldsymbol{e}_1\boldsymbol{e}_1^T - (U_0^{(0)})^{-1}U_{-1}^{(0)}\boldsymbol{e}_1\boldsymbol{e}_1^T L_1^{(0)}(L_0^{(0)})^{-1}$$

$$= \boldsymbol{e}_1\boldsymbol{e}_1^T - G_{-1}^{(0)}\boldsymbol{e}_1\boldsymbol{e}_1^T F_1^{(0)}. \qquad \blacksquare$$

Theorem 9. *For any $s \geqslant 0$, we find that*

$$\mathcal{F}_1((D_0^{(s)})^{-1}) = \boldsymbol{e}_1\boldsymbol{e}_1^T - (JC(z^{n+m}a(z))\,J)^{2^s(n+m)}\boldsymbol{e}_1\boldsymbol{e}_1^T(JC(z^{n+m}c(z))^T J)^{2^s(n+m)}.$$

Proof. The proof is by induction on s. The case $s = 0$ is established by Theorem 8. For $s > 0$, $l_{s-1} = 2^{s-1}(n+m)$, observe that

$$\mathcal{F}_1((D_0^{(s)})^{-1}) =$$

$$\mathcal{F}_1((D_0^{(s-1)})^{-1}) + (JC(z^{n+m}a(z))\,J)^{l_{s-1}}\mathcal{F}_1((D_0^{(s-1)})^{-1})(JC(z^{n+m}c(z))^T J)^{l_{s-1}} =$$

$$\boldsymbol{e}_1\boldsymbol{e}_1^T - (JC(z^{n+m}a(z))\,J)^{l_{s-1}}\boldsymbol{e}_1\boldsymbol{e}_1^T(JC(z^{n+m}c(z))^T J)^{l_{s-1}} +$$

$$(JC(z^{n+m}a(z))\,J)^{l_{s-1}}\boldsymbol{e}_1\boldsymbol{e}_1^T(JC(z^{n+m}c(z))^T J)^{l_{s-1}} +$$

$$-(JC(z^{n+m}a(z))\,J)^{l_s}\boldsymbol{e}_1\boldsymbol{e}_1^T(JC(z^{n+m}c(z))^T J)^{l_s} =$$

$$\boldsymbol{e}_1\boldsymbol{e}_1^T - (JC(z^{n+m}a(z))\,J)^{l_s}\boldsymbol{e}_1\boldsymbol{e}_1^T(JC(z^{n+m}c(z))^T J)^{l_s}. \qquad \blacksquare$$

As an immediate consequence of this result we find that the displacement rank of the matrices $(D_0^{(s)})^{-1}$, $s \geqslant 0$, w.r.t. the displacement operator $\mathcal{F}_2$ can also be bounded from above by a small constant integer.

Corollary 10. *For any $s \geqslant 0$, there are uniquely determined $n+m$-vectors $\boldsymbol{r}^{(s)}$ and $\boldsymbol{t}^{(s)}$ such that*

$$\mathcal{F}_2((D_0^{(s)})^{-1}) = \boldsymbol{e}_1\boldsymbol{r}^{(s)^T} + \boldsymbol{t}^{(s)}\boldsymbol{e}_1^T - (JC(z^{n+m}a(z))\,J)^{l_s}\boldsymbol{e}_1\boldsymbol{e}_1^T(JC(z^{n+m}c(z))^T J)^{l_s},$$

where $l_s = 2^s(n+m)$ and the first entry of $\boldsymbol{r}^{(s)}$ is zero.

Proof. The proof immediately follows by replacing (27) into the displacement equation provided by Theorem 9. The uniqueness of $\boldsymbol{r}^{(s)}$ and $\boldsymbol{t}^{(s)}$ is evident in the light of Theorem 6. In fact, it says that we can represent $(D_0^{(s)})^{-1}$ as

$$L(\boldsymbol{t}^{(s)}) + U(\boldsymbol{r}^{(s)}) - L((JC(z^{n+m}a(z))\,J)^{l_s}\boldsymbol{e}_1)\,U(JC(z^{n+m}c(z))\,J)^{l_s}\boldsymbol{e}_1). \tag{30}$$

Therefore, $\boldsymbol{t}^{(s)}$ is the first column of

$$(D_0^{(s)})^{-1} + L((JC(z^{n+m}a(z))\,J)^{l_s}\boldsymbol{e}_1)\,U(JC(z^{n+m}c(z))\,J)^{l_s}\boldsymbol{e}_1),$$

whereas $\boldsymbol{r}^{(s)^T}$ is obtained by its first row by replacing the entry in position 1 with zero. ■

In this way the problem of approximating the central coefficients $h_{-n-m+1}, \dots, h_{n+m-1}$ of the Laurent series of the reciprocal of $a(z)\,c(z^{-1})$ is reduced to that one of computing the sequences $\{\boldsymbol{r}^{(s)}\}$ and $\{\boldsymbol{t}^{(s)}\}$ that are quadratically convergent towards the vector $[0, h_{-1}, \dots, h_{-n-m+1}]^T$ and the vector $[h_0, h_1, \dots, h_{n+m-1}]^T$, respectively. By combining Theorem 7 with the inverse representation formula (30), we find that the iterative evaluation of the entries of these sequences can be accomplished at the cost of $O((n+m)\log(n+m))$ arithmetic operations thus leading at a superfast algorithm for the computation of the sought coefficients within an arbitrarily small precision ε. The resulting procedure **EvaluateCCRLS** is given below, where the acronym **CCRLS** stands for central coefficients of the reciprocal of a Laurent series. Here $r_a^{(k)}(z)$ and $r_c^{(k)}(z)$ denote the remainder of the Euclidean division of z^k by $p_a(z) = z^{n+m}a(z)$ and by $p_c(z) = z^{n+m}c(z)$, respectively. Moreover, powers of Frobenius matrices are always represented by means of (29) as a sum of products of triangular Toeplitz matrices.

Procedure **EvaluateCCRLS**

input: The coefficients of $a(z)$ and $c(z)$ and the value of the stop parameter s.

- Solve the linear systems

$$D_0^{(0)}\tilde{\boldsymbol{t}}^{(0)} = \boldsymbol{e}_1, \qquad D_0^{(0)^T}\tilde{\boldsymbol{r}}^{(0)} = \boldsymbol{e}_1.$$

- Set

$$r_a^{(n+m)}(z) = -(a_1 z^{n+m-1} + \ldots + a_n z^m),$$

$$r_c^{(n+m)}(z) = -(c_1 z^{n+m-1} + \ldots + c_m z^n).$$

- Evaluate the entries of the vectors

$$\boldsymbol{f} = (\boldsymbol{f}_i) = (JC(z^{n+m} a(z))\, J)^{n+m} \boldsymbol{e}_1, \qquad \boldsymbol{g} = (\boldsymbol{g}_i) = (JC(z^{n+m} c(z))\, J)^{n+m} \boldsymbol{e}_1.$$

- Compute

$$\boldsymbol{t}^{(0)} = \tilde{\boldsymbol{t}}^{(0)} + \boldsymbol{g}_0 \boldsymbol{f}, \qquad \boldsymbol{r}^{(0)} = \tilde{\boldsymbol{r}}^{(0)} + \boldsymbol{f}_0 \boldsymbol{g}\ ,$$

where the first entry of $\boldsymbol{r}^{(0)}$ is replaced by 0.

- **For** $i = 1, 2, \ldots, s$, **do**:
  1. Find the first column $\tilde{\boldsymbol{t}}^{(i)}$ and the first row $\tilde{\boldsymbol{r}}^{(i)^T}$ of $(D_0^{(i)})^{-1}$ by means of (24), where $(D_0^{(i-1)})^{-1}$ is expressed as in (30).
  2. Determine the coefficients of the remainders

$$r_a^{((n+m)2^i)}(z) = (r_a^{((n+m)2^{i-1})}(z))^2 (\bmod\ p_a(z)),$$

  and

$$r_c^{((n+m)2^i)}(z) = (r_c^{((n+m)2^{i-1})}(z))^2 (\bmod\, p_c(z)).$$

  3. Evaluate the entries of the vectors

$$\boldsymbol{f} = (\boldsymbol{f}_i) = (JC(z^{n+m} a(z))\, J)^{(n+m)2^i} \boldsymbol{e}_1$$

  and

$$\boldsymbol{g} = (\boldsymbol{g}_i) = (JC(z^{n+m} c(z))\, J)^{(n+m)2^i} \boldsymbol{e}_1.$$

  4. Compute

$$\boldsymbol{t}^{(i)} = \tilde{\boldsymbol{t}}^{(i)} + \boldsymbol{g}_0 \boldsymbol{f}\,, \qquad \boldsymbol{r}^{(i)} = \tilde{\boldsymbol{r}}^{(i)} + \boldsymbol{f}_0 \boldsymbol{g}\ ,$$

  where the first entry of $\boldsymbol{r}^{(i)}$ is replaced by 0.
- **endfor**

As the matrix $D_0^{(0)}$ is the product of two triangular Toeplitz matrices whose entries are explicitly given in terms of the coefficients of $a(z)$ and of $c(z)$, it is easily found that the initialization phase can be performed at the cost of $O((n+m)\log(n+m))$ arithmetic operations by means of the Sieveking-Kung algorithm [**9**]. Similarly, steps 1 and 3 require nothing but a small number of multiplications of a triangular Toeplitz matrix by a vector. This operation essentially amounts to a polynomial multiplication and, therefore, it can be carried out at the cost of $O((n+m)\log(n+m))$ arithmetic operations by using FFTs. Concerning the step 2, observe that these Euclidean divisions can also be performed in a stable way by means of convolutions [**16**] at the cost of $O((n+m)\log(n+m))$ arithmetic operations. In conclusion, Procedure **EvaluateCCRLS** can be implemented at the overall cost of $O(s(n+m)\log(n+m))$ arithmetic operations. Hence, in view of the quadratic convergence of the sequences $\{\boldsymbol{r}^{(s)}\}$ and $\{\boldsymbol{t}^{(s)}\}$, we find that, for a fixed precision ε, approximations $\boldsymbol{r}^{(s)}$ and $\boldsymbol{t}^{(s)}$ such that

$$\|\boldsymbol{t}^{(s)} - X\boldsymbol{e}_1\|_\infty \leq \varepsilon\,, \qquad \|\boldsymbol{r}^{(s)^T} - \boldsymbol{e}_1^T(X - h_0\boldsymbol{e}_1\boldsymbol{e}_1^T)\|_\infty \leq \varepsilon\,,$$

can be determined in $O(\log(\log\varepsilon^{-1}) + |\log(\log\sigma^{-1})|)$ steps at the total cost of $O((n+m)\log(n+m)(\log(\log\varepsilon^{-1}) + |\log(\log\sigma^{-1})|))$ arithmetic operations, where σ denotes the separation ratio as defined in Theorem 4.

5 - Numerical experiments

We have produced a preliminary implementation of Procedure **SolveSRLS** by using *Mathematica* ^{TM} and, then, we have tested it by performing numerical experiments on a computer with 16 decimal digits of precision.

Our program approximates the coefficient matrix $X(0)$ of (11) by the matrix $(D_0^{(s)})^{-1}$ obtained after s-th steps of (24), where s is such that

$$\|(D_0^{(s)})^{-1} - (D_0^{(s-1)})^{-1}\|_\infty \leq \varepsilon\,,$$

and ε denotes here the machine precision. The inverse of $D_0^{(0)}$ is computed by finding firstly the first column and the first row of $(L_0^{(0)})^{-1}$ and $(U_0^{(0)})^{-1}$ by means of the customary back and forward substitution processes. Then, the inverse matrices $(U_0^{(0)})^{-1}$ and $(L_0^{(0)})^{-1}$ are formed and their multiplication is carried out by using the standard matrix-by-matrix algorithm. Once $(D_0^{(0)})^{-1}$ is available, the iterative process (24) is started. The computation of $(D_0^{(i)})^{-1}$ from $(D_0^{(i-1)})^{-1}$, $1 \leq i \leq s$, is performed in a straightforward way by using the classical algorithms for operations between general matrices. The total cost of our implementation is therefore $O((n+m)^3 s)$ arithmetic operations.

However, at the present time our main interest is on the study of the numerical stability properties of our approach. In respect of this point of view, we argue that our plain version of **SolveSRLS** should mimic closely the experimental performance of a more sophisticated version of **SolveSRLS** complemented with a fast implementation of Procedure **EvaluateCCRLS** using both FFTs and the displacement rank theory. In fact, the Sieveking-Kung algorithm used in the initialization phase of **EvaluateCCRLS** results to be backward stable when it is applied for the solution of a triangular Toeplitz system whose coefficient matrix is well conditioned [**10**]. This is the case of the linear systems defined by the matrices $L_0^{(0)}$ and $U_0^{(0)}$ whose inverses are exponentially decaying by virtue of the Cauchy estimates [**30**] for the coefficients of the Taylor series of $1/c(z^{-1})$ and $1/a(z^{-1})$. Regarding at the other steps of **EvaluateCCRLS**, the essential difference between the two versions consists of performing Toeplitz-by-vector multiplications

$\boldsymbol{Ty}$, $T = (t_{i-j})$, by means of the usual method instead of using fast techniques based on FFTs. Again, we may observe that these two algorithms generally have a quite comparable numerical behavior. More specifically, it can be shown that the vector $fl_s(\boldsymbol{Ty})$ computed by the standard matrix-by-vector algorithm satisfies

$$\|\boldsymbol{Ty} - fl_s(\boldsymbol{Ty})\|_2 \leqslant (n+m)^2 \|T\|_2 \|\boldsymbol{y}\|_2 \varepsilon + O(\varepsilon^2).$$

Moreover, a similar relation holds for the vector $fl_f(\boldsymbol{Ty})$ obtained at the cost of $O((n+m)\log(n+m))$ arithmetic operations by determining the first $n+m$ components of the convolution product of the two $2(n+m)$ −vectors:

$$\boldsymbol{t}^T = [t_0, \ldots, t_{n+m-1}, 0, t_{-n-m+1}, \ldots, t_{-1}], \qquad \tilde{\boldsymbol{y}}^T = [\boldsymbol{y}^T, \boldsymbol{0}^T].$$

Indeed we have [**10**]

$$\|\boldsymbol{Ty} - fl_f(\boldsymbol{Ty})\|_2 \leqslant 16(n+m)\sqrt{n+m}\,\log_2(n+m)\|T\|_2 \|\boldsymbol{y}\|_2 \varepsilon + O(\varepsilon^2).$$

In our numerical experiments, we generated stable polynomials $a(z)$ and $c(z)$ of degree n, $n = 32, 64, 128, 256, 512$ by using the Kakeya-Eneström theorem [**30**]. It says that all the zeros of the polynomial $p(z)$ of degree n,

$$p(z) = 1 + p_1 z^{-1} + p_2 z^{-2} + \ldots + p_n z^{-n},$$

lie inside the unit circle whenever its coefficients p_i, $1 \leqslant i \leqslant n$, satisfy

$$1 > p_1 > p_2 > \ldots > p_n.$$

For each pair $(a(z), c(z))$, we considered the solution of the associated linear system (1) of order $2n$ with $\boldsymbol{b} = \boldsymbol{e}_n + \boldsymbol{e}_{n+1}$. Firstly, this system was solved to high precision (32 decimal digits of precision) by means of the *Mathematica*^{TM} function *LinearSolve* which makes use of a suitable version of the Gaussian elimination algorithm with partial pivoting. Then, the so computed solution $\boldsymbol{x}_G$ was assumed to be the exact solution of (1) and we measured the final relative error by

$$err = \frac{\|\boldsymbol{x}_G - \boldsymbol{x}\|_\infty}{\|\boldsymbol{x}_G\|_\infty},$$

where $\boldsymbol{x}$ denotes the corresponding solution produced by our procedure. As test suite we considered the following set of polynomials.

1. Balanced case: $a(z) = \sum_{i=0}^{n} a_i z^{-1}$, $c(z) = \sum_{i=0}^{n} c_i z^{-1}$, where $a_0 = c_0 = 1$ and, for $i = 2, \ldots, n$, we set

$$a_i = \frac{a_{i-1}}{1 + (0.1 \text{ Random}[\,])^{\theta}}, \qquad c_i = \frac{c_{i-1}}{1 + (0.1 \text{ Random}[\,])^{\theta}}.$$

The intrinsic function Random[ ] returns a uniformly distributed random number in the interval $(0, 1)$. For integer values of the parameter θ, the coefficients of both $a(z)$ and $c(z)$ are quite close and, as θ increases, their zeros approach the unit circumference by slowing down the convergence of the iterative process (24).

2. Unbalanced case: $a(z) = \sum_{i=0}^{n} a_i z^{-1}$, $c(z) = \sum_{i=0}^{n} c_i z^{-1}$, where $a_0 = c_0 = 1$ and, for $i = 2, \ldots, n/2$, we set

$$a_i = \frac{a_{i-1}}{1 + (0.1 \text{ Random}[\,])^{\theta}}, \qquad c_i = \frac{c_{i-1}}{1 + (0.1 \text{ Random}[\,])^{\theta}}.$$

The values of the remaining coefficients are determined according to the following rules:

$$a_{n/2+1} = \frac{a_{n/2}}{100^{\theta}}, \qquad c_{n/2+1} = \frac{c_{n/2}}{100^{\theta}},$$

and, for $n/2 + 2 \leqslant i \leqslant n$,

$$a_i = \frac{a_{i-1}}{1 + (0.1 \text{ Random}[\,])^{\theta}}, \qquad c_i = \frac{c_{i-1}}{1 + (0.1 \text{ Random}[\,])^{\theta}}.$$

Differently from the previous set, in this case the distribution of the coefficients of $a(z)$ and $c(z)$ is quite unbalanced and this fact should affect the conditioning of the associated resultant linear system.

For any considered set of test polynomials, we generated 100 pairs $(a(z), c(z))$ and we evaluated the arithmetic means err_a and r_a of the estimated relative errors and of the computed residuals, respectively. Tables 1, 2, 3, 4, 5, 6, 7 and 8 report the degree n, the value of the parameter θ, the average value s_a of the stop parameter s, the maximum $\text{err}_{\max}$ and the minimum $\text{err}_{\min}$ of the estimated relative errors, the average relative error err_a and the average residual r_a. Tables 5, 6, 7 and 8 also report the average value cond_a of the spectral condition number of the considered coefficient matrices $[T_n[\boldsymbol{c}]\,|\,T_n[\widehat{\boldsymbol{a}}]]$. Our numerical experience confirms that an unbalanced distribution of the coefficients of $a(z)$ and $c(z)$ can usually lead to ill-conditioning problems. However, in each of the performed experiments, our method is shown to be as accurate as Gaussian elimination with partial pivoting.

TABLE 1. – *Balanced case for* $\theta = 1$.

n	s_a	$\mathrm{err}_{\min}$	$\mathrm{err}_{\max}$	err_a	r_a
32	5	3.5e-16	1.2e-15	6.6e-16	7.2e-15
64	4	3.1e-16	1.2e-15	7.8e-16	8.8e-15
128	4	2.9e-16	1.3e-15	7.0e-16	9.4e-15
256	3	3.3e-16	2.7e-15	1.2e-15	1.7e-14
512	2	7.6e-16	4.8e-15	2.8e-15	2.8e-14

TABLE 2. – *Balanced case for* $\theta = 2$.

n	s_a	$\mathrm{err}_{\min}$	$\mathrm{err}_{\max}$	err_a	r_a
32	9	5.1e-16	5.8e-14	1.6e-15	5.8e-14
64	8	7.0e-16	2.1e-15	1.4e-15	1.3e-13
128	7	0.8e-15	3.6e-15	2.0e-15	1.4e-13
256	6	1.3e-15	4.9e-15	2.6e-15	3.2e-13
512	5	9.0e-16	5.9e-15	2.7e-15	3.5-13

TABLE 3. – *Balanced case for* $\theta = 3$.

n	s_a	$\mathrm{err}_{\min}$	$\mathrm{err}_{\max}$	err_a	r_a
32	13	4.5e-15	4.1e-14	1.2e-14	1.7e-12
64	11	4.2e-15	9.5e-15	5.6e-15	1.4e-12
128	11	3.6e-15	7.0e-15	5.1e-15	2.3e-12
256	10	2.7e-15	7.6e-15	4.2e-15	3.1e-12
512	8	3.6e-15	7.3e-15	5.4e-15	4.3e-12

TABLE 4. – *Balanced case for* $\theta = 4$.

n	s_a	$\mathrm{err}_{\min}$	$\mathrm{err}_{\max}$	err_a	r_a
32	18	3.6e-14	1.8e-13	1.2e-13	1.7e-11
64	16	2.2e-14	2.0e-13	8.3e-14	2.3e-11
128	14	2.5e-14	7.8e-14	4.4e-14	2.2e-11
256	13	1.1e-14	3.0e-14	1.7e-14	4.9e-11
512	11	1.9e-14	2.7e-13	8.6e-13	6.6e-11

TABLE 5. – *Unbalanced case for* $\theta = 1$.

n	cond_a	s_a	$\mathrm{err}_{\min}$	$\mathrm{err}_{\max}$	err_a	r_a
32	1.1e2	5	4.6e-16	1.6e-15	8.0e-15	8.2e-15
64	1.4e2	5	5.6e-16	2.1e-15	1.1e-15	9.8e-15
128	1.5e2	4	4.2e-16	1.4e-15	7.7e-16	9.9e-15
256	1.6e2	3	7.6e-16	3.3e-15	1.7e-15	2.1e-14
512	2.1e2	3	8.1e-16	3.6e-15	2.0e-15	2.8e-14

TABLE 6. – *Unbalanced case for* $\theta = 2$.

n	cond_a	s_a	$\mathrm{err}_{\min}$	$\mathrm{err}_{\max}$	err_a	r_a
32	2.5e3	11	2.5e-15	1.1e-14	5.7e-15	1.2e-13
64	3.2e3	9	3.2e-15	8.0e-15	6.0e-15	1.9e-13
128	4.1e3	7	5.5e-15	1.7e-14	1.1e-14	4.4e-13
256	5.0e3	5	6.0e-15	5.6e-14	1.6e-14	5.1e-13
512	7.1e3	4	7.8e-15	9.1e-14	4.1e-14	8.6e-13

TABLE 7. – *Unbalanced case for* $\theta = 3$.

n	cond_a	s_a	$\mathrm{err}_{\min}$	$\mathrm{err}_{\max}$	err_a	r_a
32	2.0e4	12	6.1e-15	4.2e-14	9.9e-15	8.7e-13
64	4.8e4	12	8.1e-15	6.7e-14	3.2e-14	1.2e-12
128	7.6e4	11	9.6e-15	1.5e-12	8.7e-14	6.6e-12
256	7.8e4	11	4.2e-14	2.1e-12	1.6e-13	9.3e-12
512	8.0e4	9	4.4e-14	3.1e-11	8.2e-13	4.0e-11

TABLE 8. – *Unbalanced case for* $\theta = 4$.

n	cond_a	s_a	$\mathrm{err}_{\min}$	$\mathrm{err}_{\max}$	err_a	r_a
32	7.4e5	19	4.4e-14	3.2e-12	7.9e-13	8.7e-11
64	7.9e5	17	6.7e-14	7.0e-12	8.7e-13	1.9e-10
128	8.1e5	16	9.1e-14	2.1e-11	1.3e-12	2.8e-10
256	8.8e5	16	1.1e-13	2.8e-11	4.6e-12	8.9e-10
512	1.1e6	13	8.7e-13	7.1e-11	5.5e-11	1.5e-9

6 - Conclusions and further extensions

A novel numerical approach to the efficient solution of stable resultant linear systems has been presented. It is based upon the close connections between the matrix problem, its polynomial formulation and the problem of factoring polynomials with respect to the unit circle in the complex plane (spectral factorization problem). The experimental results of a preliminary implementation of the proposed solution algorithm are also reported by showing its effectiveness and robustness.

Many theoretical and experimental issues are however still open.

Firstly, an extensive numerical experience with a Fortran 90 implementation of Procedure **EvaluateCCRLS**, complemented with fast routines based on FFTs, is required to confirm its stability properties in the case where it is applied in a superfast way. Presently, this is an ongoing work.

Secondly, the exploitation of other diverse ways of solving the bi-infinite block tridiagonal system (15) would be very interesting. Observe that the tridiagonal form is a special case of the the more general Hessenberg structure and, recently, many efficient sequential and parallel algorithms for the solution of block Hessenberg linear systems have been developed (see [**41**], [**27**] and [**25**]). An analysis of the properties of such algorithms in the present context might be useful to devise alternative procedures for the solution of stable resultant linear systems.

Thirdly, in this paper we have shown that spectral factorization methods can lead to superfast 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 [**32**]. The study of similar results and relations between Hurwitz factorization methods [**24**] and the solution of structured linear systems should be welcome.

Finally, Theorem 5 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 [**31**], it was, apparently, ascertained for the first time by G. Baxter and I. Hirschman [**2**]. Some years later, A. A. Semencul [**39**] 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 [**28**]. Unlike of these theoretical contributions, we believe that the possibility of extending the derivation of Theorem 5 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 use in order to construct approximate representations of the inverse of a Toeplitz matrix with several applications to the preconditioning theory.

References

[1] S. BARNETT, *Polynomials and Linear Control Systems*, Marcel Dekker, New York 1983.

[2] G. BAXTER and I. I. HIRSCHMAN, *An explicit inversion formula for finite-section Wiener-Hopf operators*, Bull. Amer. Math. Soc. **70** (1964), 820-823.

[3] D. BINI and L. GEMIGNANI, *Fast parallel computation of the polynomial remainder sequences via Bezout and Hankel matrices*, SIAM J. Comput. **24** (1995), 63-77.

[4] D. A. BINI and L. GEMIGNANI, *Fast fraction-free triangularization of Bezoutians with applications to sub-resultant chain computation*, Linear Algebra Appl. **284** (1998), 19-39.

[5] D. A. BINI, L. GEMIGNANI and B. MEINI, *Factorization of analytic functions by means of Koenig's theorem and Toeplitz computations*, Numerische Mathematik (1998), to appear.

[6] D. A. BINI and B. MEINI, *Improved cyclic reduction for solving queueing problems*, Numerical Algorithms **15** (1997), 57-74.

[7] D. A. BINI and B. MEINI, *Effective methods for solving banded Toeplitz systems*, SIAM J. Matrix Anal. Appl. **20** (1999), 700-719.

[8] 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, SIAM, 1999, 211-244.

[9] D. A. BINI and V. PAN, *Matrix and Polynomial Computations*, **1**, Fundamental Algorithms, Birkhäuser, Boston 1994.

[10] D. A. BINI and V. PAN, *Polynomials and Matrix Computations*, **2**, Birkhäuser, Boston 2000, to appear.

[11] R. BITMEAD and B. ANDERSON, *Asympotically fast solution of Toeplitz and related systems of linear equations*, Linear Algebra Appl. **34** (1980), 103-116.

[12] A. BÖTTCHER and B. SILBERMANN, *Introduction to Large Truncated Toeplitz Matrices*, Springer-Verlag, 1999.

[13] W. S. BROWN and J. F. TRAUB, *On Euclid's algorithm and the theory of subresultants*, J. Assoc. Comput. Mach. **18** (1971), 505-514.

[14] B. L. BUZBEE, G. H. GOLUB and C. W. NIELSON, *On direct methods for solving Poisson's equation*, SIAM J. Num. Anal. **7** (1970), 627-656.

[15] S. CABAY and P. KOSSOWSKI, *Power series remainder sequences and Padé fractions over integral domain*, J. Symbolic Comp. **10** (1990), 139-163.

[16] J. P. CARDINAL, *On two iterative methods for approximating the roots of a polynomial*, in The mathematics of numerical Analysis, Lectures in Appl. Math., AMS, 1996, 165-188.

[17] G. E. COLLINS, *Sub-resultants and reduced polynomial remainder sequences*, J. Assoc. Comput. Mach. **14** (1967), 128-142.

[18] 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** (1989), 545-552.

[19] 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** (1990), 1697-1709.

[20] P. A. VAN DOOREN, *Some numerical challenges in control theory*, in Linear Algebra for Control Theory, vol. **62** of IMA Vol. Math. Appl. Springer, 1994.

[21] L. GEMIGNANI, *Solving Hankel systems over the integers*, J. Symbolic Comp. **18** (1994), 573-584.

[22] L. GEMIGNANI, *Schur complements of Bezoutians and the inversion of block Hankel and block Toeplitz matrices*, Linear Algebra Appl. **249** (1996), 79-91.

[23] 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 apppear.

[24] L. GEMIGNANI, *Computing a Hurwitz factorization of a polynomial*, J. Comput. Appl. Math. **126** (2000), 369-380.

[25] L. GEMIGNANI, *Efficient and stable solution of structured Hessenberg linear systems arising from difference equations*, Numer. Linear Algebra with Appl. **7** (2000), 39-335.

[26] L. GEMIGNANI, *Polynomial factors, lemniscates and structured matrix technology*, in D. Bini, P. Yalamov and E. Tyrtyshnikov, ed., *Structured matrices: recent developments in theory and computation*, NOVA Science, 2000.

[27] L. GEMIGNANI and G. LOTTI, *A recursive block Gaussian elimination algorithm for solving block Hessenberg systems*, submitted.

[28] I. C. GOHBERG and I. A. FEL'DMAN, *Convolution equations and projection methods for their solution*, Amer. Math. Soc. (1974).

[29] W. B. GRAGG and L. REICHEL, *On singular values of Hankel operator of finite rank*, Linear Algebra Appl. **121** (1989), 53-70.

[30] P. HENRICI, *Applied and Computational Complex Analysis*, **1**, Wiley, 1974.

[31] I. S. IOHVIDOV, *Hankel and Toeplitz Matrices and Forms*, Birkhäuser, Boston 1982.

[32] J. JEŽEK, *Conjugated and symmetric polynomial equations*, Kybernetika, **19** (1983).

[33] E. I. JURY, K. J. ÅSTRÖM and R. G. AGNIEL, *A numerical methods for the evaluation of complex integrals*, IEEE Trans. Automat. Contr., **15** (1970), 468-471.

[34] T. KAILATH and A. H. SAYED, *Displacement structure: Theory and applications*, SIAM Review **37** (1995), 297-386.

[35] L. V. KANTOROVICH and G. P. AKILOV, *Functional Analysis in Normed Spaces*, Pergamon Press, 1964.

[36] P. KIRRINNIS, *Fast computation of contour integral of rational functions*, Manuscript, 1998.

[37] P. LANCASTER and M. TISMENETSKY, *The Theory of Matrices*, Academic Press, New York 1985.

[38] A. SCHÖNHAGE, *Asymptotically fast algorithms for the numerical multiplication and division of polynomials with complex coefficients*, in J. Calmet, ed., Proc. EUROCAM 1982, vol. **144** Lecture Notes in Comput. Sci., Springer, Berlin 1982.

[39] 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.

[40] J. R. SENDRA and J. LLOVET, *GCD of polynomials and Hankel matrices*, J. Symbolic Comp. **13** (1992), 25-39.

[41] G. W. STEWART, *On the solution of block Hessenberg systems*, Numer. Linear Algebra Appl. **2** (1995), 287-296.

[42] G. RODRIGUEZ, T. N. T. GOODMAN, C. A. MICCHELLI and S. SEATZU, *Spectral factorization of Laurent polynomials*, Adv. Comput. Math. **7** (1997), 429-454.

[43] G. T. WILSON, *Factorization of the covariance generating function of a pure moving-average process*, SIAM J. Num. Anal. **6** (1969), 1-7.

Abstract

Devising efficient methods for the solution of resultant linear systems is a relevant issue in many diverse fields like computer algebra, control theory, signal processing and data modeling. Over the years, several fast and superfast algorithms have been proposed that are based either on purely numerical techniques or on mixed numeric-symbolic procedures. In this paper we present a new solution scheme falling in the former class that works under some auxiliary conditions on the separation of the spectrum of the polynomials associated with the initial coefficient matrix. Such assumptions are usually satisfied in the considered applications of control and signal theory and their exploitation allows us to reduce the original matrix problem to the equivalent one of finding the reciprocal of a Laurent polynomial. To carry out this computation we develop both finite and iterative processes employing a blend of ideas from structured numerical linear algebra, computational complex analysis and linear operator theory. The effectiveness and the robustness of the resulting composite solution methods is then confirmed by means of numerical experiments that are finally reported and discussed.

* * *