Università di Pisa

Dipartimento di Informatica

Technical Report : TR-98-08

Factorization of Analytic Functions by means of Koenig's Theorem and Toeplitz Computations

D. A. BINI, L. GEMIGNANI, B. MEINI

July 16, 1998

ADDR: Corso Italia 40,56125 Pisa,Italy. TEL: +39-50-887111. FAX: +39-50-887226

Factorization of Analytic Functions by means of Koenig's Theorem and Toeplitz Computations*

Dario Andrea Bini† Luca Gemignani‡ Beatrice Meini†

Abstract

By providing a matrix version of Koenig's theorem we reduce the problem of evaluating the coefficients of a monic factor $r(z)$ of degree h of a power series $f(z)$ to that of approximating the first h entries in the first column of the inverse of an infinite Toeplitz matrix in block Hessenberg form. This matrix is reduced to a band matrix if $f(z)$ is a polynomial.

We devise a suitable algorithm, based on cyclic reduction and on the concept of displacement rank, for generating a sequence of vectors $\mathbf{v}^{(2^j)}$ that quadratically converges to the vector $\mathbf{v}$ having as components the coefficients of the factor $r(z)$. In the case of a polynomial $f(z)$ of degree N, the cost of computing the entries of $\mathbf{v}^{(2^j)}$ given $\mathbf{v}^{(2^{j-1})}$ is $O(N \log N + \theta(N))$ arithmetic operations, where $\theta(N) = O(N \log^2 N)$ is the cost of solving an $N \times N$ Toeplitz-like system. In the case of analytic functions the cost depends on the numerical degree of the power series involved in the computation.

The algorithm is strictly related to the Graeffe method for lifting the roots of a polynomial.

From the numerical experiments performed with several test polynomials and power series, the algorithm has shown good numerical properties and promises to be a good candidate for implementing polynomial root-finders based on recursive splitting strategies.

Keywords. Koenig's theorem, Toeplitz matrices, polynomial roots, zeros of analytic functions, Graeffe's algorithm.

AMS Classification 65H05, 65F05.

1 INTRODUCTION

Let $f(z)$ be an analytic function such that the series $f(z) = \sum_{i=0}^{+\infty} a_i z^i$ is convergent for $|z| < R$. Let r_i, $i = 1, 2, \ldots$, be the zeros of $f(z)$ such that $|r_{i+1}| \geq |r_i|$ for any i and $|r_{h+1}| > |r_h|$ for a given $h > 0$. Define $r(z) = \prod_{i=1}^{h}(z - r_i)$ and denote by $\mathbf{v}$ the vector of the first h coefficients of $r(z)$.

Consider the infinite matrix $A = (a_{i-j+h})_{i,j=1,2,\ldots}$, defined by the coefficients $a_0, a_1, \ldots$, and by the integer h, where we assume $a_i = 0$ for $i < 0$. Observe that A has constant entries along its diagonals, i.e., A is a Toeplitz matrix; moreover, if we partition A into $m \times m$ blocks A_i, $i = 0, 1, \ldots$, for $m \geq h$, we obtain a block Toeplitz matrix in block Hessenberg form

$$A = \begin{bmatrix} A_1 & A_0 & & & \bigcirc \\ A_2 & A_1 & A_0 & & \\ \vdots & \ddots & \ddots & \ddots & \end{bmatrix}.$$

*Supported by the European project FRISCO LTR 21.024

†Dipartimento di Matematica, Università di Pisa, 56127 Pisa, Italy.

‡Dipartimento di Informatica, Università di Pisa, 56125 Pisa, Italy.

We prove that, for k sufficiently large, the leading principal sub-matrix $A^{(k)}$ of A of size $km \times km$ is nonsingular. Moreover, denoting $\mathbf{w}^{(k)}$ the vector made up by the first m components of the first column of $(A^{(k)})^{-1}$, we show that the sequence $\mathbf{v}^{(k)} = -A_0\mathbf{w}^{(k)}$ converges to a vector obtained by filling up $\mathbf{v}$ with $m-h$ zeros, more precisely

$$\mathbf{v}^{(k)} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \mathbf{v} \end{bmatrix} + O(\sigma^k), \quad |r_h/r_{h+1}| < \sigma < 1. \tag{1}$$

For a polynomial $f(z)$ of degree N the matrix A is banded; moreover, if the block partitioning is performed with size $m = \max\{h, N-h\}$ then A is block tridiagonal.

In this way we reduce the problem of approximating a factor of $f(z)$ to that of computing m entries of the first column of the inverse of a Toeplitz matrix in block Hessenberg form, that becomes a block tridiagonal matrix for a polynomial $f(z)$.

We provide an adaptation of the cyclic reduction algorithm —originally introduced for the numerical solution of the Poisson equation [7] and adjusted in [1], [2] to solve banded Toeplitz linear systems and for the numerical solution of certain problems in queueing theory— for computing the sequence of vectors $\mathbf{v}^{(2^j)}$, $j = 1, 2, \ldots$.

The algorithm that we obtain in this way has a quadratic convergence and a low computational cost. In fact, we prove that the computation of $\mathbf{v}^{(2^j)}$ given $\mathbf{v}^{(2^{j-1})}$, costs $O(m \log m + \theta(m))$ arithmetic operations (ops) for $f(z)$ polynomial of degree N, where $m = \max\{h, N-h\}$, $\theta(m) = O(m \log^2 m)$ is the cost of solving an $m \times m$ Toeplitz-like system.

For power series the cost estimate $O(mN(\log N + \log m) + N\theta(m))$ holds where N is an upper bound on the numerical degrees of the power series involved in the computation of $\mathbf{v}^{(2^j)}$. In special cases of particular interest in the applications, like the problems in queueing theory and Markov chains, the numerical degree of the series generated in the computation can be dynamically controlled and is always inferior to the numerical degree of the original series. Further, there is some numerical evidence that this property still holds for many important functions of mathematical physics such as the Bessel functions.

The numerical stability of our scheme depends on the numerical stability of the LU factorization, this subject is well investigated and understood in the literature. In particular, the stability can be controlled by means of known techniques like the look-ahead strategies. Moreover, as a consequence of classical results from numerical linear algebra, the scheme is stable where A is diagonally dominant. Numerical experiments show very good numerical performances even though the latter condition is not verified.

The proof of (1) is based on some closely related extensions of the Koenig theorem developed by Householder [14]. The algorithm for solving the matrix problem is based, besides on the technique of cyclic reduction, on the concept of displacement rank [15], [12]. This concept allows us to represent all the $m \times m$ matrices generated in the computation, by means of a few vectors, and to perform all the computations by means of FFT's of length m.

Approximating the roots of polynomials and of power series has received a great interest in the last decade not only for the challenging theoretical problems but also for its importance in the many applications of industry and technology [23]. Our algorithm adds one more tool for implementing recursive factorization techniques for polynomial and power series factorization.

As a matter of interest, a great effort has been spent in the European project FRISCO (see the web site www.nag.co.uk/projects/FRISCO.html) which aims to provide advanced and reliable software, based on the integration of numerical and symbolic tools, for solving polynomial systems arising in scientific and industrial problems.

Numerical methods for computing the zeros of polynomials and analytic functions have been analyzed in the years, see for instance [14] and [22] for the methods based on Padé approximation, on the Koenig theorem, and on the Sebastião-e-Silva and Graeffe iteration.

More recently, factorization techniques have been analyzed both as a practical means for extracting a factor of small degree [10], [11], and in terms of computational complexity where they are used as basic tools for the recursive reduction of a polynomial into factors. In this latter respect, it is interesting to cite the pioneering work of A. Schoenhage [21] and his "splitting circle method", the results of Kirrinnis [16] and the most recent use of mixed algebraic geometric and analytic techniques made by V. Pan and leading to the best complexity estimates [18], [19], where the basic tool for recursive splitting is the Graeffe iteration. Recursive splitting has been also considered in [8] and [5] where generalizations of Sebastião-e-Silva method are provided.

The paper is organized in the following way. In section 2 we provide a generalized matrix version of the Koenig theorem and we reduce the problem of factorizing a power series to that of solving an infinite Toeplitz system. In section 3 we provide our algorithm, based on cyclic reduction, and we describe it in functional form. In section 4 we analyze the break-down situations where the cyclic reduction algorithm cannot be carried out and propose suitable techniques to overcome this problem. In section 5, based on the concept of displacement rank, we recall some structural properties of the matrices generated by our algorithm. Moreover, we use these properties for obtaining low complexity bounds. Section 6 deals with implementative issues and, finally, in section 7 we report the results of some numerical experiments.

2 The basic reduction to Toeplitz computations

Let $R > 0$ be a given number and let

$$f(z) \in \mathcal{G}(R) = \{g(z) : \mathbf{C} \to \mathbf{C} \ : \ g(z) \text{ analytic for } |z| < R\}$$

have the series expansion

$$f(z) = 1 + \sum_{i=1}^{+\infty} a_i z^i \tag{2}$$

with radius of convergence R so that [13] $a_i = \frac{1}{2\pi \underline{i}} \int_{|z|=\rho} f(z) z^{-i-1} dz$, where $\underline{i}$ denotes the imaginary unit, $\underline{i}^2 = -1$, and ρ is any number such that $0 < \rho < R$. From this integral representation it follows that there exists a constant μ such that $|a_j| \le \mu \rho^{-j}$, $j \ge 0$.

We associate with the power series (2) the infinite lower triangular Toeplitz matrix

$$T = \begin{bmatrix} 1 & & & 0 \\ a_1 & 1 & & \\ a_2 & a_1 & 1 & \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix} \tag{3}$$

and relate the entries of its inverse T^{-1} with the zeros of the function $f(z)$. More precisely, we prove a sort of generalization of Koenig's theorem in matrix form which allows us to recover an approximation of the polynomial factor by solving a Toeplitz system in block Hessenberg form.

For this purpose, let us denote by r_1, r_2, r_3, ..., the zeros of $f(z)$ (if they exist), where zeros of multiplicity m appear as m consecutive entries in the list, and assume that the zeros have been ordered so that $0 < |r_1| \leq |r_2| \leq |r_3| \leq \ldots < R$. Further, if $|r_h| < |r_{h+1}|$, define $r_h(z) = \prod_{i=1}^{h}(1 - r_i^{-1}z)$, and partition the matrix T of (3) into $h \times h$ Toeplitz blocks A_i, $i = 0, 1, \ldots$, so that

$$T = \begin{bmatrix} A_0 & & & 0 \\ A_1 & A_0 & & \\ A_2 & A_1 & A_0 & \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{4}$$

The main goal of this section is to prove the following:

Theorem 1 *Let h be a positive integer and $f(z) \in \mathcal{G}(R)$ be an analytic function having the series expansion (2) and zeros $r_1, r_2, \ldots$, such that $0 < |r_1| \leq |r_2| \leq \ldots \leq |r_h| < |r_{h+1}| \leq \ldots < R$. Denote $r_h(z) = \prod_{i=1}^{h}(1 - r_i^{-1}z) = \sum_{i=0}^{h} u_i z^i$ a factor of $f(z)$, and consider the block linear systems*

$$\left[X_1^{(j)}, X_2^{(j)}, \ldots, X_j^{(j)}\right] \begin{bmatrix} A_1 & A_0 & & 0 \\ A_2 & A_1 & \ddots & \\ \vdots & \ddots & \ddots & A_0 \\ A_j & \ldots & \ldots & A_1 \end{bmatrix} = [-A_0, 0, \ldots, 0] \tag{5}$$

where the coefficient matrix, denoted by $A^{(j)}$, such that $A^{(j)} = (a_{r-s+h})$, $a_r = 0$ for $r < 0$, is a Toeplitz matrix in block Hessenberg form with the $h \times h$ Toeplitz blocks A_i, $i \geq 0$. Then there exists an integer k such that for any $j \geq k$ $A^{(j)}$ is nonsingular and, moreover,

$$X_1^{(j)}\mathbf{e}_1 = \begin{bmatrix} -u_0/u_h \\ \vdots \\ -u_{h-1}/u_h \end{bmatrix} + O(\sigma^j), \tag{6}$$

for any σ such that $\left|\frac{r_h}{r_{h+1}}\right| < \sigma < 1$ and $\mathbf{e}_1 = [1, 0, \ldots, 0]^T$.

Notice that, for $0 \leq j \leq h$, u_j/u_h is the coefficient of z^j in the monic polynomial $r_h(z)/u_h = \prod_{i=1}^{h}(z - r_i)$. Further, observe that for $h = 1$ the first entry of the solution of (5) can be expressed in terms of the coefficients of the formal power series expansion of $1/f(z)$ in such a way that the previous result reduces to the celebrated Koenig's theorem [14]. Therefore, in the sequel we will often refer to the Theorem 1 as to our matrix generalization of Koenig's theorem.

The above theorem provides a means for approximating a factor of the power series $f(z)$ by reducing this approximation to computing the first block entry in the north-western corner of the inverse of a Toeplitz matrix in block Hessenberg form. This computation can be performed with several algorithms having different efficiencies. In the next section we will introduce some advanced techniques, based on the Toeplitz matrix technology, for performing this computation in a very efficient way.

Let $jh = mr$, $m \geq h$, and repartition the matrix $A^{(j)}$ in (5) into $m \times m$ blocks $\widetilde{A}_i$, $i = 0, \ldots, r$. Observe that the matrix $A^{(j)}$, partitioned in this way, is still block Toeplitz in block Hessenberg form. Moreover, the block $\widetilde{A}_0$ is zero everywhere except for its $h \times h$ sub-matrix in the lower leftmost corner that coincides with A_0. Therefore the system

$$\left[\begin{array}{cccc} \widetilde{X}_1^{(j)}, & \widetilde{X}_2^{(j)}, & \ldots, & \widetilde{X}_r^{(j)} \end{array}\right] A^{(j)} = \left[\begin{array}{cccc} -\widetilde{A}_0, & O, & \ldots, & O \end{array}\right] \tag{7}$$

is such that the $h \times h$ sub-matrix in the lower leftmost corner of $\widetilde{X}_1^{(j)}$ coincides with $X_1^{(j)}$. This implies the following corollary of Theorem 1.

Corollary 2 *Let us assume the hypotheses of Theorem 1 and consider the $jm \times jm$ Toeplitz matrix $A^{(j)} = (a_{i-s+h})$, where $jh = mr$, $m \geq h$ and $a_i = 0$ for $i < 0$. Let us partition $A^{(j)}$ into $m \times m$ blocks $\widetilde{A}_i$, $i = 0, \ldots, r$. Then, for the first component $\widetilde{X}_1^{(j)}$ of the solution of (7) we have $(\widetilde{X}_1^{(j)}\mathbf{e}_1)_{m-h+i} = -u_{i-1}/u_h + O(\sigma^j)$, $i = 1, \ldots, h$.*

We may look at the above corollary as to Theorem 1 extended to a power series having $m - h$ null zeros.

2.1 Proof of Theorem 1

In order to prove Theorem 1 we look for an asymptotic characterization of $X_1^{(j)}$ in terms of the zeros of $f(z)$. For this purpose observe that, by virtue of Theorem 4.4d of [13] the function $t(z) = r_h(z)/f(z)$ can be analytically continued into r_i, $1 \leq i \leq h$. Thus, we can write $t(z) \in \mathcal{G}(\rho)$ for any ρ such that $0 < \rho < |r_{h+1}|$. If we denote $q(z)$ the polynomial of degree at most $h-1$ that interpolates $t(z)$ at r_i, $i = 1, \ldots, h$ (where interpolation concerns also derivatives up to the order $p-1$ if a zero has multiplicity p), then the function $t(z) - q(z)$ is divisible by $r_h(z)$ so that

$$t(z) - q(z) = r_h(z)s(z), \quad s(z) \in \mathcal{G}(\rho),\ 0 < \rho < |r_{h+1}|.$$

In this way, we obtain that the decomposition

$$\frac{1}{f(z)} = \frac{q(z)}{r_h(z)} + s(z), \quad s(z) \in \mathcal{G}(\rho),\ 0 < \rho < |r_{h+1}|, \tag{8}$$

holds whenever $|z| \leq \rho$ and, moreover, $z \neq r_i$, $1 \leq i \leq h$.

Now, observe that the formal power series expansion $\frac{1}{f(z)} = \sum_{i=0}^{+\infty} b_i z^i$, whose coefficients uniquely determine the first column of the matrix T^{-1}, coincides with the Taylor expansion of the reciprocal function $1/f(z)$ over the domain $|z| < |r_1|$. Hence, we conclude that

$$b_i = k_i + e_i, \quad i \geq 0, \tag{9}$$

where

$$\frac{q(z)}{r_h(z)} = \sum_{i=0}^{+\infty} k_i z^i, \tag{10}$$

and, moreover,

$$s(z) = \sum_{i=0}^{+\infty} e_i z^i \tag{11}$$

represents the series expansion of $s(z)$ with radius of convergence ρ, $0 < \rho < |r_{h+1}|$.

Now we rephrase equations (8–11) in matrix form by means of a block Toeplitz system.

Firstly, let us introduce the infinite down-shift matrix $Z = (z_{i,j})$ defined by $z_{i+1,i} = 1$, $z_{i,j} = 0$ for $i \neq j+1$, denote with $\mathbf{b} = T^{-1}\mathbf{e}_1$ the infinite vector having entries b_i, where $\mathbf{e}_1 = (1, 0, \ldots)^T$, and with B the Toeplitz matrix whose i–th column, $1 \leq i \leq h$, is $Z^{i-1}\mathbf{b}$, i.e., $B = \begin{bmatrix} \mathbf{b} & Z\mathbf{b} & \ldots & Z^{h-1}\mathbf{b} \end{bmatrix}$. This matrix can be partitioned into $h \times h$ blocks B_0, B_1, ..., where B_i are Toeplitz matrices.

Since B is made up by the first h columns of T^{-1}, in the light of (4) it follows that

$$\begin{bmatrix} A_0 & & & 0 \\ A_1 & A_0 & & \\ A_2 & A_1 & A_0 & \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix} \begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \tag{12}$$

where I is the identity matrix of order h.

Equation (9) provides the following decomposition of B_i

$$B_i = \begin{bmatrix} b_{ih} & \dots & b_{(i-1)h+1} \\ \vdots & \ddots & \vdots \\ b_{(i+1)h-1} & \dots & b_{ih} \end{bmatrix} = K_i + E_i, \quad i \geq 0,$$

where the Toeplitz matrices K_i and E_i are defined by

$$K_i = \begin{bmatrix} k_{ih} & \dots & k_{(i-1)h+1} \\ \vdots & \ddots & \vdots \\ k_{(i+1)h-1} & \dots & k_{ih} \end{bmatrix}, \; E_i = \begin{bmatrix} e_{ih} & \dots & e_{(i-1)h+1} \\ \vdots & \ddots & \vdots \\ e_{(i+1)h-1} & \dots & e_{ih} \end{bmatrix},$$

with $b_i, k_i, e_i = 0$ whenever $i < 0$.

Let J be the $h \times h$ permutation matrix having unit anti-diagonal entries (reversion matrix). If the size is not clear from the context we denote with denote with J_r a reversion matrix of size rh. If A is a Toeplitz matrix of order h, then JA and AJ are Hankel matrices; further, $JAJ = A^T$ is again a Toeplitz matrix. Consider the Hankel matrix KJ where

$$K = \begin{bmatrix} k_{h-1} & \dots & k_0 \\ \vdots & \ddots & \vdots \\ k_{2h-2} & \dots & k_{h-1} \end{bmatrix}.$$

From (10) we see that KJ is a Hankel matrix generated by the Markov parameters of the transfer function $q(z^{-1})/(zr_h(z^{-1}))$ which is a series in z^{-1} [12]. We refer to KJ as to the $h \times h$ Hankel matrix associated with the two reversed polynomials $z^{h-1}q(z^{-1})$ and $z^h r_h(z^{-1})$. Since these polynomials are relatively prime, it follows that KJ is nonsingular.

Furthermore, after multiplying both sides of (10) by $r_h(z)$ and comparing coefficients in z^h, z^{h+1}, ..., we obtain that $K_1 J = KJF$, where F denotes the Frobenius matrix associated with the monic polynomials $z^h r_h(z^{-1})$, i.e.,

$$F = \begin{bmatrix} 0 & & & -u_h \\ 1 & \ddots & & -u_{h-1} \\ & \ddots & 0 & \vdots \\ & & 1 & -u_1 \end{bmatrix}, \quad r_h(z) = \sum_{i=0}^{h} u_i z^i, \; u_0 = 1.$$

Inductively, we find that $K_i = KJF^{h(i-1)+1}J, \quad i \geq 1$, which leads to the following characterization of the matrices B_i [14].

Lemma 3 *For the blocks defined in (12) it holds*

$$B_i = (K + E_i JF^{-h(i-1)-1}J)JF^{h(i-1)+1}J, \tag{13}$$

moreover, for any σ such that $\frac{|r_h|}{|r_{h+1}|} < \sigma < 1$ there exists a constant β such that

$$\| E_i JF^{-h(i-1)-1}J \|_\infty \leq \beta\sigma^i. \tag{14}$$

Based on this result, Householder [14] and Stewart [22] developed numerical methods to approximating the coefficients of a factor of $f(z)$. Approximations are constructed by means of determinantal relations involving Hankel matrices whose entries are defined by the series expansion of $1/f(z)$.

Here we follow a different line, that is based on the following observation. If B_j is asymptotically nonsingular then equations (13) (14) imply that $(B_{j-1}B_j^{-1})^T$ tends to F^{-h}, so that the first column of $JB_j^{-T}B_{j-1}^TJ$ provides accurate approximations to the coefficients of the factor $r_h(z)$ for j large enough. Therefore, we first prove that B_j is asymptotically nonsingular, then we show that the matrix $JB_j^{-T}B_{j-i}^TJ$ coincides with $X_i^{(j)}$ in (5). Finally we prove the bound (6) by using Lemma 3 in order to estimate the convergence of $X_1^{(j)}$ to $JF^{-h}J$.

From (12) it follows that

$$\begin{bmatrix} A_1 & A_0 & & \\ A_2 & A_1 & \ddots & \\ \vdots & \ddots & \ddots & A_0 \\ A_j & \dots & \dots & A_1 \end{bmatrix} \begin{bmatrix} B_0 \\ B_1 \\ \vdots \\ B_{j-1} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -A_0B_j \end{bmatrix}, \tag{15}$$

for any $j \geq 2$. Observe that the coefficient matrix —denoted by $A^{(j)}$— is a block Toeplitz matrix in lower Hessenberg form. Moreover, the super-diagonal blocks A_0 are nonsingular. The following result relates the solvability of (15) with the nonsingularity of B_j. It provides an extension of the Wronski formula (see Theorem 1.3 of [13]) for the inversion of formal power series with coefficients over a field.

Lemma 4 *The coefficient matrix $A^{(j)}$ of the linear system (15) is nonsingular if and only if B_j is nonsingular.*

Proof. Consider the following persymmetric partitioning of the $(j+1)\times(j+1)$ block leading principal sub-matrix T_{j+1} of T,

$$T_{j+1} = \begin{bmatrix} A_0 & & & 0 \\ A_1 & A_0 & & \\ \vdots & \ddots & \ddots & \\ A_j & \dots & \dots & A_0 \end{bmatrix} = \begin{bmatrix} D^{(j)} & 0 \\ A^{(j)} & C^{(j)} \end{bmatrix}. \tag{16}$$

If $A^{(j)}$ is singular then there exists a non-zero block vector satisfying

$$A^{(j)} \begin{bmatrix} W_0^{(j)} \\ \vdots \\ W_{j-1}^{(j)} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

From (16) we have

$$T_{j+1} \begin{bmatrix} W_0^{(j)} \\ \vdots \\ W_{j-1}^{(j)} \\ 0 \end{bmatrix} = \begin{bmatrix} A_0W_0^{(j)} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

where $A_0W_0^{(j)}$ is a non-zero matrix. This implies that the block vector

$$\begin{bmatrix} B_0A_0W_0^{(j)} - W_0^{(j)} \\ \vdots \\ B_{j-1}A_0W_0^{(j)} - W_{j-1}^{(j)} \\ B_jA_0W_0^{(j)} \end{bmatrix}$$

belongs to the kernel of T_{j+1} and, therefore, B_j has to be singular.

Conversely, assume that $A^{(j)}$ is nonsingular. In the view of (16), if we consider the complementary persymmetric partitioning of T_{j+1}^{-1} we find

$$\det T_{j+1} = \pm \det A^{(j)} \det(D^{(j)} {A^{(j)}}^{-1} C^{(j)}).$$

This shows that the Schur complement of $A^{(j)}$ is nonsingular. In addition, we obtain the following block representation of the inverse of T_{j+1},

$$T_{j+1}^{-1} = \begin{bmatrix} 0 & {A^{(j)}}^{-1} \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} -{A^{(j)}}^{-1} C^{(j)} \\ I \end{bmatrix} (D^{(j)} {A^{(j)}}^{-1} C^{(j)})^{-1} \begin{bmatrix} I, & -D^{(j)} {A^{(j)}}^{-1} \end{bmatrix},$$

from which it finally follows that $B_j = -(D^{(j)} {A^{(j)}}^{-1} C^{(j)})^{-1}$. ■

By combining Lemma 4 with the characterization of the B_i provided by (13) we see that $A^{(j)}$ is nonsingular for all sufficiently large indices j.

In order to prove equation (5) of Theorem 1, we provide a more suitable representation of the system (15) by reverting the rows and the columns of the matrix in (15). This operation correspond to multiplying $A^{(j)}$ on the left and on the right by the $jh \times jh$ permutation matrix J_j having 1 on the anti-diagonal. Since it holds $(J_j A^{(j)} J_j)^T = A^{(j)}$ and moreover J_j can be viewed as a block matrix having the $h \times h$ block J on its anti-diagonal, then we have

$$\left[B_{j-1}^T J, B_{j-2}^T J, \ldots, B_0^T J\right] \begin{bmatrix} A_1 & A_0 & & 0 \\ A_2 & A_1 & \ddots & \\ \vdots & \ddots & \ddots & A_0 \\ A_j & \ldots & \ldots & A_1 \end{bmatrix} = \left[-B_j^T A_0^T J, 0, \ldots, 0\right],$$

that for j large enough so that B_j is nonsingular, is reduced to (12) with

$$X_i^{(j)} = J B_j^{-T} B_{j-i}^T J, \quad i = 1, \ldots, j. \tag{17}$$

In order to obtain equation (6) of Theorem 1 we apply Lemma 3. In fact, from (13) and (17), we deduce that

$$X_1^{(j)} = J(K + E_j J F^{-h(j-1)-1} J)^{-T} J (F^{-h})^T J (K + E_{j-1} J F^{-h(j-2)-1} J)^T J \tag{18}$$

We recall the following [9]:

Lemma 5 *If A and Δ are an $n \times n$ matrices such that A is nonsingular and $\| A^{-1}\Delta \|_\infty = r < 1$, then $A + \Delta$ is nonsingular and $\| (A+\Delta)^{-1} - A^{-1} \|_\infty \leq \frac{\|\Delta\|_\infty \|A^{-1}\|_\infty^2}{1-r}$.*

Applying this result to (18), we find that

$$X_1^{(j)} = K^{-1}(F^{-h})^T K + \Delta^{(j)}, \quad \| \Delta^{(j)} \|_\infty \leq \vartheta \sigma^j \tag{19}$$

holds for sufficiently large j satisfying $\beta \| K^{-1} \|_\infty \sigma^j < 1$. Since KJ is the Hankel matrix associated with $z^h r_h(z^{-1})$, it is well known that [20] $F^T KJ = KJF$. Using this commuting property, we obtain $K^{-1}(F^{-h})^T K = JF^{-h}J$, that, together with (19), implies $X_1^{(j)} = JF^{-h}J + \Delta^{(j)}$, which proves equation (6) of Theorem 1. In fact, in order to complete the proof of Theorem 1 observe that the inverse of the Frobenius matrix F is given by

$$F^{-1} = \begin{bmatrix} -u_{h-1}/u_h & 1 & & O \\ -u_{h-2}/u_h & 0 & \ddots & \\ \vdots & \vdots & \ddots & 1 \\ -u_0 & 0 & \ldots & 0 \end{bmatrix}.$$

Thus, the first column of F^{-h} provides the coefficients $-u_i/u_h$, $i = 0, \ldots, h-1$.

By recursively computing the first entry of the solutions of (5) for increasing values of the size jh, we obtain a method for the numerical approximation of the coefficients of $r_h(z)/u_h$. This method will be outlined in the next section.

3 Solving block Toeplitz systems in block Hessenberg form by means of cyclic reduction

In this section we devise a recursive triangularization procedure, based on the cyclic reduction technique, that we apply to the coefficient matrix $A^{(j)}$ of (5). This procedure will compute a block LU factorization of the matrices $A^{(2^n)}$ for $n = 0, 1, 2, \ldots, k$ provided they exist, where $k > 0$ is a given integer, and at the same time computes also the matrix $X_1^{(2^n)}$ of (5) for $n = 0, 1, 2, \ldots, k$.

In its basic form, the proposed method requires the nonsingularity of certain submatrices of $A^{(2^k)}$. For the sake of simplicity, throughout this section we assume that all the matrices to be inverted are nonsingular. This assumption will be removed in the next section.

Let us introduce some notations. Given an $\ell \times \ell$ block matrix R with $m \times m$ blocks $R_{i,j}$, $i, j = 1, \ldots, \ell$, where $m \geq h$, and given two vectors $\mathbf{m} = (m_i) \in \{1, \ldots, \ell\}^r$, $\mathbf{n} = (n_i) \in \{1, \ldots, \ell\}^s$, such that $m_i \neq m_j$ and $n_i \neq n_j$ for $i \neq j$, we denote $R(\mathbf{m}, \mathbf{n})$ the $r \times s$ block sub-matrix of R with entries R_{m_i, n_j}, $i = 1, \ldots, r$, $j = 1, \ldots, s$. For simplicity we denote $R(\mathbf{m}) = R(\mathbf{m}, \mathbf{m})$.

Furthermore we associate with a given permutation $p : \{1, \ldots, \ell\} \to \{1, \ldots, \ell\}$ the $\ell \times \ell$ block permutation matrix $P = (P_{i,j})$ such that $P_{i,j} = I$ for $i = p(j)$, $P_{i,j} = O$ elsewhere. Here I denotes the $m \times m$ identity matrix, moreover if the size is not clear from the context we write I_r for the identity matrix of order rm.

Now, for a given integer $k > 0$ set $\ell = 2^k$ and consider the following system

$$
\begin{gathered}
\begin{bmatrix} X_1, & X_2, & \ldots, & X_\ell \end{bmatrix} H = \begin{bmatrix} C, & O, & \ldots, & O \end{bmatrix} \\
H = H^{(0)} = \begin{bmatrix} \widehat{H}_1 & H_0 & & & 0 \\ \widehat{H}_2 & H_1 & H_0 & & \\ \vdots & \vdots & \ddots & \ddots & \\ \widehat{H}_{\ell-1} & H_{\ell-2} & \ldots & H_1 & H_0 \\ \widehat{H}_\ell & H_{\ell-1} & \ldots & H_2 & H_1 \end{bmatrix},
\end{gathered}
\tag{20}
$$

where H_i, $\widehat{H}_i$ have dimension m. Hereafter we assume that $H^{(0)} = A^{(2^k)}$. In this way the above system reduces to (5) for $\ell = j$, $\widehat{H}_i = H_i = A_i$, $C = -A_0$, $X_i = X_i^{(j)}$ and $m = h$, or to (7) for $m > h$ and $\ell = r$, $H_i = \widehat{H}_i = \widetilde{A}_i$, $C = -\widetilde{A}_0$, $X_i = \widetilde{X}_i^{(j)}$.

Here our goal is simply to compute the first component X_1 of the solution of (20). Let us introduce the even-odd permutation $q_k : \{1, 2, \ldots, \ell\} \to \{1, 2, \ldots, \ell\}$ such that $q_k(2i) = i$, $q_k(2i-1) = 2^{k-1} + i$, together with the associated block permutation matrix Q_k and the vector $\mathbf{m} = (q_k^{-1}(i)) = (2, 4, \ldots, \ell, 1, 3, \ldots, \ell - 1)$. Then we have

$$
H^{(0)}(\mathbf{m}) = Q_k H^{(0)} Q_k^T = \begin{bmatrix} T^{(0)} & W^{(0)} \\ Z^{(0)} & R^{(0)} \end{bmatrix},
\tag{21}
$$

moreover

$$
\begin{gathered}
\begin{bmatrix} X_+, & X_- \end{bmatrix} H^{(0)}(\mathbf{m}) = \begin{bmatrix} O, & O, & \ldots, & O, & C, & O, & \ldots, & O \end{bmatrix} \\
X_+ = \begin{bmatrix} X_2, & X_4, & \ldots, & X_\ell \end{bmatrix}, \quad X_- = \begin{bmatrix} X_1, & X_3, & \ldots, & X_{\ell-1} \end{bmatrix}.
\end{gathered}
\tag{22}
$$

We observe that $T^{(0)}$, $R^{(0)}$, $Z^{(0)}$ and $W^{(0)}$ are $(\ell/2)\times(\ell/2)$ block Toeplitz-like matrices, defined by

$$W^{(0)} = \begin{bmatrix} \widehat{H}_2 & H_0 & & 0 \\ \widehat{H}_4 & H_2 & \ddots & \\ \vdots & \vdots & \ddots & H_0 \\ \widehat{H}_\ell & H_{\ell-2} & \dots & H_2 \end{bmatrix}, \; Z^{(0)} = \begin{bmatrix} H_0 & & & 0 \\ H_2 & H_0 & & \\ \vdots & \ddots & \ddots & \\ H_{\ell-2} & \dots & H_2 & H_0 \end{bmatrix},$$

$$R^{(0)} = \begin{bmatrix} \widehat{H}_1 & & & & 0 \\ \widehat{H}_3 & H_1 & & & \\ \vdots & \ddots & \ddots & & \\ \widehat{H}_{\ell-3} & \dots & H_3 & H_1 & \\ \widehat{H}_{\ell-1} & H_{\ell-3} & \dots & H_3 & H_1 \end{bmatrix}, \; T^{(0)} = \begin{bmatrix} H_1 & & & 0 \\ H_3 & H_1 & & \\ \vdots & \ddots & \ddots & \\ H_{\ell-1} & \dots & H_3 & H_1 \end{bmatrix}.$$

By applying one step of block Gaussian elimination to the 2×2 block matrix (21) we find that

$$\begin{gathered} Q_k H^{(0)} Q_k^T = \begin{bmatrix} I_{2^{k-1}} & O \\ V^{(0)} & I_{2^{k-1}} \end{bmatrix} \begin{bmatrix} T^{(0)} & W^{(0)} \\ O & H^{(1)} \end{bmatrix}, \\ V^{(0)} = Z^{(0)}(T^{(0)})^{-1}, \; H^{(1)} = R^{(0)} - V^{(0)}W^{(0)}, \end{gathered} \tag{23}$$

moreover the system (22) turns into

$$X_- H^{(1)} = [\; C, \quad O, \quad \dots, \quad O \;] \tag{24}$$

together with $X_+ = -X_- Z^{(0)}(T^{(0)})^{-1}$. Since our goal is to compute X_1 that is the first component of X_-, we may ignore the latter equation and reduce the original problem (20) to the problem (24) of half the size. Moreover, it can be easily verified (compare [2]) that the matrix $H^{(1)}$ of (24) has exactly the same structure of the original matrix $H^{(0)}$ in (20), that is, is a block Toeplitz matrix, except for its first block column, in block Hessenberg form. In this way the same scheme can be cyclically applied to (24) thus obtaining the sequence of $\ell_n \times \ell_n$ matrices $H^{(n)}$, for $\ell_n = 2^{k-n}$ where

$$H^{(n)} = \begin{bmatrix} \widehat{H}_1^{(n)} & H_0^{(n)} & & & 0 \\ \widehat{H}_2^{(n)} & H_1^{(n)} & H_0^{(n)} & & \\ \vdots & \vdots & \ddots & \ddots & \\ \widehat{H}_{\ell_n-1}^{(n)} & H_{\ell_n-2}^{(n)} & \dots & H_1^{(n)} & H_0^{(n)} \\ \widehat{H}_{\ell_n}^{(n)} & H_{\ell_n-1}^{(n)} & \dots & H_2^{(n)} & H_1^{(n)} \end{bmatrix}. \tag{25}$$

and

$$\begin{aligned} & H^{(n)} = R^{(n-1)} - Z^{(n-1)}(T^{(n-1)})^{-1}W^{(n-1)} \\ & Q_{k-n+1} H^{(n-1)} Q_{k-n+1}^T = \begin{bmatrix} T^{(n-1)} & W^{(n-1)} \\ Z^{(n-1)} & R^{(n-1)} \end{bmatrix}. \end{aligned} \tag{26}$$

Similarly, the sequence of systems

$$\begin{aligned} & X^{(n)} H^{(n)} = [\; C, \quad O, \quad \dots, \quad O \;] \\ & X^{(n)} = [\; X_1, \quad X_{1+2^n}, \quad X_{1+2\cdot 2^n}, \quad \dots, \quad X_{1+(2^{k-n}-1)\cdot 2^n} \;] \end{aligned}$$

is generated, for $n = 0, 1, \dots, k$ and for $n = k$ we have $H^{(k)} = \widehat{H}_1^{(k)}$, $X_1 \widehat{H}_1^{(k)} = C$, that allows us to recover the matrix $X_1 = C {\widehat{H}_1^{(k)}}^{-1}$.

Remark 6 Observe that if we denote with $P^{(n)}$ the $\ell \times \ell$ block cumulative permutation matrix such that

$$P^{(n)} = \begin{bmatrix} I_{2^k-2^{k-n+1}} & 0 \\ 0 & Q_{k-n+1} \end{bmatrix} P^{(n-1)}, \quad P^{(0)} = Q_k,$$

then the cyclic reduction procedure at each step n provides an $(n+1) \times (n+1)$ block triangular factorization of $P^{(n)} H^{(0)} P^{(n)T}$, where the block entry in position $(n+1, n+1)$ of the upper triangular factor is the $\ell_n \times \ell_n$ block matrix $H^{(n)}$ of (25).

Now let us consider the infinite block Toeplitz-like matrix in block Hessenberg form

$$H_\infty^{(0)} = \begin{bmatrix} \widehat{H}_1 & H_0 & & & 0 \\ \widehat{H}_2 & H_1 & H_0 & & \\ \widehat{H}_3 & H_2 & H_1 & H_0 & \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{27}$$

and, similarly to the finite case, we construct the infinite block sub-matrices $T_\infty^{(0)}$, $R_\infty^{(0)}$, $Z_\infty^{(0)}$, $W_\infty^{(0)}$ of $H_\infty^{(0)}$ obtained by considering separately block rows and block columns having odd, even position, respectively, i.e.,

$$T_\infty^{(0)} = \begin{bmatrix} H_1 & & & 0 \\ H_3 & H_1 & & \\ H_5 & H_3 & H_1 & \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}, \quad R_\infty^{(0)} = \begin{bmatrix} \widehat{H}_1 & & & 0 \\ \widehat{H}_3 & H_1 & & \\ \widehat{H}_5 & H_3 & H_1 & \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix},$$

$$Z_\infty^{(0)} = \begin{bmatrix} H_0 & & & 0 \\ H_2 & H_0 & & \\ H_4 & H_2 & H_0 & \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}, \quad W_\infty^{(0)} = \begin{bmatrix} \widehat{H}_2 & H_0 & & & 0 \\ \widehat{H}_4 & H_2 & H_0 & & \\ \widehat{H}_6 & H_4 & H_2 & H_0 & \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}.$$

One step of cyclic reduction applied to the infinite matrix (27) leads to the matrix $H_\infty^{(1)} = R_\infty^{(0)} - Z_\infty^{(0)} T_\infty^{(0)^{-1}} W_\infty^{(0)}$. It is easy to verify (compare [3]) that $H_\infty^{(1)}$ has the structure (27), i.e., it is an infinite block Hessenberg matrix, that, except for the first block column, has the block Toeplitz structure. We may recursively apply the same reduction (cyclic reduction), thus obtaining the sequence of matrices

$$H_\infty^{(n)} = \begin{bmatrix} \widehat{H}_1^{(n)} & H_0^{(n)} & & & 0 \\ \widehat{H}_2^{(n)} & H_1^{(n)} & H_0^{(n)} & & \\ \widehat{H}_3^{(n)} & H_2^{(n)} & H_1^{(n)} & H_0^{(n)} & \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{28}$$

Each matrix $H_\infty^{(n)}$ is uniquely determined by the blocks $\widehat{H}_i^{(n)}$, $i \geq 1$, $H_i^{(n)}$, $i \geq 0$, defining its first two block columns. Moreover, due to the triangular and Toeplitz structure of $Z_\infty^{(0)}$ and $T_\infty^{(0)}$, it can be easily verified that the $\ell_n \times \ell_n$ block matrix (25) obtained at the n-th step of cyclic reduction applied to the $2^k \times 2^k$ finite matrix (20), coincides with the $2^{k-n} \times 2^{k-n}$ block leading principal sub-matrix of the infinite matrix (28) obtained at the n-th step of cyclic reduction, applied to the infinite matrix of (27). In particular, the block entry $\widehat{H}_1^{(n)}$ in position $(1,1)$ of (28) coincides with the one generated at the last step of the cyclic reduction algorithm applied to its leading principal sub-matrix of order 2^n. Hence, $\widehat{H}_1^{(n)}$ is nonsingular if $A^{(2^n)}$ is nonsingular.

This fact leads to the following result that, in the view of Theorem 1, provides a means for computing a sequence of approximations of a factor of an analytic function.

Theorem 7 *Assume that the cyclic reduction process applied to $H^{(0)} = A^{(\ell)}$, where $A^{(\ell)}$ is the coefficient matrix of (5) and $\ell = 2^k$, can be carried out without breakdowns and that the matrices $A^{(2^n)}$, $1 \leq n \leq k$, are nonsingular. Then, for the first block entry $X_1^{(2^n)}$ of the solution of (5) it holds*

$$X_1^{(2^n)} = -A_0(\widehat{H}_1^{(n)})^{-1}, \ n = 1, \ldots, k.$$

Since we double the size of the linear system (5) at each step, the computed sequence of approximations $X_1^{(2^n)}\mathbf{e}_1$ converges quadratically to the vector of the coefficients of the desired factor (compare (6)). In practice we apply cyclic reduction to the infinite matrix (27). This provides the desired sequence $X_1^{(2^n)}$. In section 4 we will discuss the applicability of the cyclic reduction in the presence of breakdowns where some blocks $T_\infty^{(n)}$ may be singular. Specifically, we suitably modify the triangularization scheme in order to compute the subsequence $\{X_1^{(2^s)}\}$, $1 \leq s \leq k$, corresponding to the values of s for which the leading principal sub-matrices of $A^{(\ell)}$ of order $2^s - 1$ and 2^s are nonsingular.

3.1 Functional formulation of cyclic reduction

For computational purposes, it is useful to associate with the infinite matrix $H_\infty^{(n)}$, obtained at the n-th step of cyclic reduction applied to the matrix (27), the formal matrix power series $\widehat{H}^{(n)}(z) = \sum_{i=0}^{+\infty} \widehat{H}_{i+1}^{(n)} z^i$, $H^{(n)}(z) = \sum_{i=0}^{+\infty} H_i^{(n)} z^i$, which represent the first and the second block column of the matrix $H_\infty^{(n)}$. Then, the first and the second block column of the matrix $H_\infty^{(n+1)}$ are determined by the formal matrix power series [2]

$$\begin{cases} \widehat{H}^{(n+1)}(z) = \widehat{H}_{even}^{(n)}(z) - H_{even}^{(n)}(z)\left(H_{odd}^{(n)}(z)\right)^{-1} \widehat{H}_{odd}^{(n)}(z) \\ H^{(n+1)}(z) = zH_{odd}^{(n)}(z) - H_{even}^{(n)}(z)\left(H_{odd}^{(n)}(z)\right)^{-1} H_{even}^{(n)}(z) \end{cases} \tag{29}$$

with $H^{(0)}(z) = \sum_{i=0}^{+\infty} A_i z^i$, $\widehat{H}^{(0)}(z) = \sum_{i=0}^{+\infty} A_{i+1} z^i$, where for any matrix power series $R(z) = \sum_{i=0}^{+\infty} R_i z^i$ we define

$$\begin{array}{l} [R(z)]_{even} = R_{even}(z) = \sum_{i=0}^{+\infty} R_{2i} z^i, \\ [R(z)]_{odd} = R_{odd}(z) = \sum_{i=0}^{+\infty} R_{2i+1} z^i. \end{array}$$

Hence, we have a recursive formula expressed in terms of formal matrix power series, that allows the computation of the coefficients $H_i^{(n)}$, $\widehat{H}_i^{(n)}$, at each step of cyclic reduction. The relation (29) can also be rewritten in the equivalent form

$$\begin{array}{c} H^{(n+1)}(z^2) = \left((H^{(n)}(z)^{-1} - H^{(n)}(-z)^{-1})/(2z)\right)^{-1} = \\ -H^{(n)}(z)H_{odd}^{(n)}(z^2)^{-1}H^{(n)}(-z). \end{array} \tag{30}$$

In the case where $H^{(0)}$ is a block Toeplitz matrix in a block tridiagonal form, then the matrices generated by cyclic reduction are still block tridiagonal and, except for the north western corner block, have the block Toeplitz structure. The block entries $H_i^{(n+1)}$, $i = 0, 1, 2$, $\widehat{H}_1^{(n+1)}$, are explicitly given by the relations:

$$\begin{array}{l} H_0^{(n+1)} = -H_0^{(n)}(H_1^{(n)})^{-1}H_0^{(n)} \\ H_1^{(n+1)} = H_1^{(n)} - H_0^{(n)}(H_1^{(n)})^{-1}H_2^{(n)} - H_2^{(n)}(H_1^{(n)})^{-1}H_0^{(n)} \\ H_2^{(n+1)} = -H_2^{(n)}(H_1^{(n)})^{-1}H_2^{(n)} \\ \widehat{H}_1^{(n+1)} = \widehat{H}_1^{(n)} - H_0^{(n)}(H_1^{(n)})^{-1}H_2^{(n)}. \end{array} \tag{31}$$

Moreover the second equation in (30) reduces to

$$H^{(n+1)}(z^2) = -H^{(n)}(z)\left(H_1^{(n)}\right)^{-1} H^{(n)}(-z).$$

By taking determinants of both the sides of the above equation, and by setting $\gamma^{(n)}(z) = \det(H^{(n)}(z))$, we obtain that $\gamma^{(n+1)}(z^2) = -\gamma^{(n)}(z)\gamma^{(n)}(-z)/\det(H_1^{(n)})$. This means that the cyclic reduction step is equivalent to the squaring step in the Graeffe algorithm [14] for factoring polynomials. Such a connection was already pointed out in [1] where the authors used it in order to prove certain convergence properties of the cyclic reduction process applied for the solution of infinite banded Toeplitz linear systems.

The evaluation of the formulae (29) involves arithmetic operations among the formal matrix power series $H^{(n)}(z)$ and $\widehat{H}^{(n)}(z)$. In section 5, using the displacement rank theory, it will be shown that, for any $n \geq 0$ and any z, these matrices are Toeplitz-like. Moreover, in the tridiagonal case, where the equations (29) reduce to (31), the matrices $H_i^{(n)}$, $i = 0, 1, 2$ and $\widehat{H}_1^{(n)}$ themselves are Toeplitz-like. These facts allow us to dramatically reduce the computational cost of a single step of cyclic reduction. This will be discussed in section 6.

4 Applicability of cyclic reduction

In this section we investigate the conditions which ensure that the cyclic reduction technique applied to the linear system (5) can be carried out without the occurrence of a breakdown at some intermediate step due to the singularity of the matrices to be inverted. Our analysis leads to a suitable generalization of the basic triangularization scheme which can be applied to any arbitrary block Toeplitz matrix in a block lower Hessenberg form. Similar developments can also be used in order to improve the stability properties of the factorization process.

As already observed in the previous section, we recall that the cyclic reduction algorithm produces a block triangular factorization of a matrix obtained from $A^{(2^k)}$ by suitably permuting its rows and columns. Specifically, the n-th step, for $n \leq k$, provides a block triangular factorization of $P^{(n)}A^{(2^k)}P^{(n)^T}$, where $P^{(n)}$ denotes the block cumulative permutation matrix defined in Section 3. Let us denote with $p_n : \mathcal{I} \to \mathcal{I}$, $\mathcal{I} = \{1, \ldots, 2^k\}$, the permutation associated with the matrix $P^{(n)}$. In the light of Remark 6, for the properties of the Schur complement, the matrix $H^{(n)}$ is such that $H^{(n)^{-1}}$ coincides with the $2^{k-n} \times 2^{k-n}$ trailing principal sub-matrix of $(P^{(n)}A^{(2^k)}P^{(n)^T})^{-1}$, that is the principal sub-matrix of $A^{(2^k)^{-1}}$ corresponding to block row indices $p_n^{-1}(i)$ with $2^{k-n}(2^n - 1) + 1 \leq i \leq 2^k$. Therefore, any different permutation strategy which leads to a block triangular factorization of $\tilde{P}^{(n)}A^{(2^k)}\tilde{P}^{(n)T}$ generates the same inverse matrix whenever $p_n^{-1}(i) = \tilde{p}_n^{-1}(i)$ for all i such that $2^k - 2^{k-n} + 1 \leq i \leq 2^k$.

The possibility of reaching the same block factorization by means of different permutations enables us to choose permutation strategies that do not lead to breakdown or that improve the numerical stability of the algorithm. Let us begin by comparing two different permutation strategies for computing a block triangular factorization of a nonsingular block matrix A. Similar comparisons are generally used for proving the equivalence of some variants of Gaussian elimination with pivoting [9]. We assume that A is a generic block nonsingular matrix of order 2^k with $m \times m$ block entries $A_{i,j}$, $1 \leq i, j \leq 2^k$.

Let n be an integer such that $0 < n \leq k$, and let $w_1 : \mathcal{I} \to \mathcal{I}$, $w_2 : \mathcal{I} \to \mathcal{I}$ and $w_3 : \{1, \ldots, 2^{k-n+1}\} \to \{1, \ldots, 2^{k-n+1}\}$ be three permutations associated with the permutation matrices P_1, P_2 and P_3, respectively. Let us assume that

$$w_1^{-1}(2^k - 2^{k-n} + j) = w_2^{-1}(2^k - 2^{k-n} + w_3^{-1}(j)), \quad 1 \le j \le 2^{k-n}, \qquad (32)$$

or, equivalently in matrix form, that

$$P_1(\mathbf{m},\mathbf{n}) = (\begin{bmatrix} I_{2^k-2^{k-n+1}} & 0 \\ 0 & P_3 \end{bmatrix} P_2)(\mathbf{m},\mathbf{n}), \qquad (33)$$

If $A([\ w_1^{-1}(1) \quad \dots \quad w_1^{-1}(2^k - 2^{k-n})\ ])$ is nonsingular, then we can write

$$P_1AP_1^T = \begin{bmatrix} I_{2^k-2^{k-n}} & 0 \\ L^{(1)} & I_{2^{k-n}} \end{bmatrix} \begin{bmatrix} T^{(1)} & W^{(1)} \\ 0 & S^{(1)} \end{bmatrix}, \qquad (34)$$

where $T^{(1)} = A([\ w_1^{-1}(1) \quad \dots \quad w_1^{-1}(2^k - 2^{k-n})\ ])$ and $S^{(1)}$ is the Schur complement of $T^{(1)}$ in $P_1AP_1^T$. In the view of the nonsingularity of A, we find that $S^{(1)}$ is nonsingular and, moreover, ${S^{(1)}}^{-1} = A^{-1}([\ w_1^{-1}(2^k - 2^{k-n} + 1) \quad \dots \quad w_1^{-1}(2^k)\ ])$.

On the other hand, if also $A([\ w_2^{-1}(1) \quad \dots \quad w_2^{-1}(2^k - 2^{k-n+1})\ ])$ is nonsingular, then we analogously find that

$$P_2AP_2^T = \begin{bmatrix} I_{2^k-2^{k-n+1}} & 0 \\ L^{(2)} & I_{2^{k-n+1}} \end{bmatrix} \begin{bmatrix} T^{(2)} & W^{(2)} \\ 0 & S^{(2)} \end{bmatrix}. \qquad (35)$$

By virtue of (33), the leading principal sub-matrix of order $2^k - 2^{k-n}$ of

$$\begin{bmatrix} I_{2^k-2^{k-n+1}} & 0 \\ 0 & P_3 \end{bmatrix} P_2AP_2^T \begin{bmatrix} I_{2^k-2^{k-n+1}} & 0 \\ 0 & P_3 \end{bmatrix}^T$$

coincides up to a suitable reordering of rows and columns with $T^{(1)}$ and, hence, is nonsingular. This means that the leading principal sub-matrix of order $2^{k-n+1} - 2^{k-n}$ of $P_3S^{(2)}P_3^T$ is nonsingular, too. In this way, by applying again the triangularization process to $P_3S^{(2)}P_3^T$, we obtain that

$$P_3S^{(2)}P_3^T = \begin{bmatrix} I_{2^{k-n+1}-2^{k-n}} & 0 \\ L^{(3)} & I_{2^{k-n}} \end{bmatrix} \begin{bmatrix} T^{(3)} & W^{(3)} \\ 0 & S^{(3)} \end{bmatrix}. \qquad (36)$$

By combining (35) with (36), we arrive at a block triangular factorization of

$$\begin{bmatrix} I_{2^k-2^{k-n+1}} & 0 \\ 0 & P_3 \end{bmatrix} P_2AP_2^T \begin{bmatrix} I_{2^k-2^{k-n+1}} & 0 \\ 0 & P_3 \end{bmatrix}^T$$

which shows that $S^{(3)} = S^{(1)}$.

Summing up, the following result holds.

Theorem 8 *Let n be an integer such that $0 < n \le k$, and let $w_1 : \mathcal{I} \to \mathcal{I}$, $w_2 : \mathcal{I} \to \mathcal{I}$ and $w_3 : \{1,\dots,2^{k-n+1}\} \to \{1,\dots,2^{k-n+1}\}$ be three permutations such that (32) holds. If $A([\ w_2^{-1}(1) \quad \dots \quad w_2^{-1}(2^k - 2^{k-n+1})\ ])$ is nonsingular then $S^{(1)}$, as defined by (34), is computable if and only if $S^{(3)}$, given by (35) and (36), is computable. Moreover, if also A is nonsingular, then $S^{(3)} = S^{(1)}$.*

In the sequel of this section we use this result for comparing two computationally different triangularization processes applied to the nonsingular matrix $A = A^{(2^k)}$. More precisely we show that the Schur complement of size 2^{k-n} obtained after n steps of cyclic reduction coincides with the one obtained in a single step of block Gaussian elimination applied to a suitably permuted matrix $C_{n,k}AC_{n,k}^T$.

In the light of Theorem 8, the block permutation matrix $C_{n,k}$ of order 2^k is such that (33) holds with $P_1(\mathbf{m},\mathbf{n}) = C_{n,k}(\mathbf{m},\mathbf{n})$ and with $P_2 = P^{(n-1)}$, $P_3 = Q_{k-n+1}$, as defined in the previous section. In order to satisfy (33) the matrix $C_{n,k}$ must verify the condition $(C_{n,k})_{i,j} = I_h$ if $j = 1+(i-1-2^k+2^{k-n})2^n$, for $2^k-2^{k-n}+1 \le i \le 2^k$.

The remaining entries of $C_{n,k}$ are chosen in such a way that the block leading principal sub-matrix of order $2^k - 2^{k-n}$ of $C_{n,k}AC_{n,k}^T$, partitioned into $(2^n-1) \times (2^n-1)$ blocks, results a $2^{k-n} \times 2^{k-n}$ block lower triangular matrix, where the diagonal blocks are equal to the block leading principal sub-matrix of A of order $2^n - 1$.

For this purpose, we partition $\mathcal{I}$ into disjoint sets $\{i_1^{(j)}, \ldots, i_{2^{k-n}}^{(j)}\}$, $j = 1, \ldots, 2^n$, such that $i_r^{(j)} = j + (r-1)2^n$, $1 \le r \le 2^{k-n}$, $1 \le j \le 2^n$. Then, for $1 \le i \le 2^k - 2^{k-n}$, we define $(C_{n,k})_{i,j} = I$ for $j = i_s^{(r)}$, where, $s = \lceil i/(2^n-1) \rceil$ and $r = 2^n$ if $i = 0 \bmod 2^n - 1$, $r = 1 + (i \bmod 2^n - 1)$ if $i \neq 0 \bmod 2^n - 1$. Here $\lceil x \rceil$ denotes the minimum integer greater than or equal to x.

For instance, in the case where $k = 3$ and $n = 2$ we obtain that

$$C_{2,3} = \begin{bmatrix} 0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & I \\ I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 & 0 \end{bmatrix}.$$

In the view of Theorem 8, collecting all these observations leads to the following result that relates the applicability of the cyclic reduction algorithm with the nonsingularity of certain block leading principal sub-matrices of $A^{(2^k)}$.

Theorem 9 *Assume that the first $n-1$ steps of the cyclic reduction algorithm applied to the nonsingular matrix $A^{(2^k)}$ are well-defined. Then, the n-th step is well-defined if and only if the block leading principal sub-matrix of order $2^n - 1$ of $A^{(2^k)}$ is nonsingular.*

This immediately gives the following necessary and sufficient condition for the applicability of the cyclic reduction algorithm applied to $A^{(2^k)}$.

Corollary 10 *The cyclic reduction algorithm applied to the nonsingular matrix $A^{(2^k)}$ can be carried out without the occurrence of a breakdown at some intermediate step if and only if all the block leading principal sub-matrices of order 2^n-1 of $A^{(2^k)}$ are nonsingular for $1 \le n \le k$.*

This result implies that the cyclic reduction algorithm is well-defined when applied to almost any matrix. However, if any of the considered leading principal sub-matrices results to be singular, then Theorem 8 can be used again for avoiding the singularity.

We say that there is a jump of order $r > 1$ between the leading principal sub-matrices of order $2^n - 1$ and $2^{n+r} - 1$ of $A^{(2^k)}$ whenever $A^{(2^n-1)}$ and $A^{(2^{n+r}-1)}$ are nonsingular whereas $\det(A^{(2^{n+j}-1)}) = 0$ for $1 \le j < r$. For the sake of notational simplicity in the sequel we will consider the problem of overcoming jumps of order 2 only. However, the proposed techniques can be easily generalized in order to cure also breakdowns generated by jumps of order greater than 2.

By virtue of Theorem 9, the occurrence of a jump of order 2 means that the $(n+1)$-st step of the cyclic reduction algorithm applied to $A^{(2^k)}$ can not be performed.

On the other hand, if $A^{(2^k)}(\epsilon) = A^{(2^k)} + \epsilon I_{2^k}$ where ϵ is chosen in a neighborhood of 0 in such a way to guarantee that $\det(A^{(2^j-1)}(\epsilon)) \neq 0$ for $1 \le j \le k$, then the cyclic reduction algorithm applied to $A^{(2^k)}(\epsilon)$ goes on to termination without breakdowns. In particular, the Schur complement at the $(n+2)$-nd step is well-defined and, when ϵ converges to zero, it will tend to the inverse of the trailing principal sub-matrix of order 2^{k-n-2} of $(P^{(n+2)}A^{(2^k)}P^{(n+2)T})^{-1}$. In the view of Theorem 8, the same Schur complement can be obtained by one step of block Gaussian elimination applied to the matrix $C_{2,k-n}\{(P^{(n)}A^{(2^k)}P^{(n)^T})^{-1}([\ 2^k - 2^{k-n}+1,\ \ \ldots,\ \ 2^k\ ])\}^{-1}C_{2,k-n}^T$, where $((P^{(n)}A^{(2^k)}P^{(n)T})^{-1}([\ 2^k - 2^{k-n}+1,\ \ \ldots,\ \ 2^k\ ]))^{-1}$ is generated at the $n-$th step of the cyclic reduction algorithm. This Schur complement is obtained with respect to the leading principal sub-matrix of order $3\ 2^{k-n-2}$. The latter sub-matrix has a useful block structure. In fact, it is a block lower triangular matrix having diagonal blocks of block dimension 3, with a block Hessenberg structure.

In the case where the matrix $A^{(2^k)} = H^{(0)}$ is block tridiagonal, then the double-step technique leads to a simple generalization of the relations (31). Specifically, assume that $H_1^{(n)}$ is singular but, on the contrary, the 3×3 block tridiagonal matrix

$$K^{(n)} = \begin{bmatrix} H_1^{(n)} & H_0^{(n)} & 0 \\ H_2^{(n)} & H_1^{(n)} & H_0^{(n)} \\ 0 & H_2^{(n)} & H_1^{(n)} \end{bmatrix}$$

is invertible with

$$(K^{(n)})^{-1} = \begin{bmatrix} K_{11}^{(n)} & K_{12}^{(n)} & K_{13}^{(n)} \\ K_{21}^{(n)} & K_{22}^{(n)} & K_{23}^{(n)} \\ K_{31}^{(n)} & K_{32}^{(n)} & K_{33}^{(n)} \end{bmatrix}.$$

Moreover assume for the sake of simplicity that $H^{(n)}$ has block dimension 8, then

$$C_{2,3}H^{(n)}C_{2,3}^T = \left[\begin{array}{ccc|ccc|cc} H_1^{(n)} & H_0^{(n)} & 0 & 0 & 0 & 0 & H_2^{(n)} & 0 \\ H_2^{(n)} & H_1^{(n)} & H_0^{(n)} & 0 & 0 & 0 & 0 & 0 \\ 0 & H_2^{(n)} & H_1^{(n)} & 0 & 0 & 0 & 0 & H_0^{(n)} \\ \hline 0 & 0 & 0 & H_1^{(n)} & H_0^{(n)} & 0 & 0 & H_2^{(n)} \\ 0 & 0 & 0 & H_2^{(n)} & H_1^{(n)} & H_0^{(n)} & 0 & 0 \\ 0 & 0 & 0 & 0 & H_2^{(n)} & H_1^{(n)} & 0 & 0 \\ \hline H_0^{(n)} & 0 & 0 & 0 & 0 & 0 & \widehat{H}_1^{(n)} & 0 \\ 0 & 0 & H_2^{(n)} & H_0^{(n)} & 0 & 0 & 0 & H_1^{(n)} \end{array}\right]. \tag{37}$$

In this way, by taking the Schur complement of the 2×2 block leading principal sub-matrix in (37), we may define

$$\begin{aligned} H_0^{(n+2)} &= -H_0^{(n)}K_{13}^{(n)}H_0^{(n)} \\ H_1^{(n+2)} &= H_1^{(n)} - H_0^{(n)}K_{11}^{(n)}H_2^{(n)} - H_2^{(n)}K_{33}^{(n)}H_0^{(n)} \\ H_2^{(n+2)} &= -H_2^{(n)}K_{31}^{(n)}H_2^{(n)} \\ \widehat{H}_1^{(n+2)} &= \widehat{H}_1^{(n)} - H_0^{(n)}K_{11}^{(n)}H_2^{(n)}. \end{aligned} \tag{38}$$

Analogous extensions can be obtained both for jumps of order greater than 2 and in the Hessenberg case. These variants of the cyclic reduction algorithm yield a generalization of Theorem 7 for input matrices with arbitrary rank profile. Indeed, if they are applied to the matrix $H^{(0)} = A^{(2^k)}$, then they perform the recursive computation of a subsequence of matrices $H^{(n_i)}$ of order 2^{k-n_i} such that

$$X_1^{(2^{n_i})}\widehat{H}_1^{(n_i)} = -A_0.$$

We finally conclude by observing that the proposed approach for skipping singular sub-matrices can be also used in a numerical setting in order to avoid manipulations involving ill-conditioned sub-matrices.

5 Structural properties of cyclic reduction

The main goal of this section is to point out the invariance of certain structures for the matrices recursively computed by the cyclic reduction algorithm and exploit them for computational purposes. A direct inspection shows that the Toeplitz structure of $H^{(0)}(z) = \sum_{i=0}^{+\infty} A_i z^i$ and $\widehat{H}^{(0)}(z) = \sum_{i=0}^{+\infty} A_{i+1} z^i$ is generally lost by the matrix power series $H^{(n)}(z)$ and $\widehat{H}^{(n)}(z)$, for $n \geq 1$. However, we may show that both $H^{(n)}(z)$ and $\widehat{H}^{(n)}(z)$, have the Toeplitz-like structure, for any z, and for any $n \geq 0$. We recall that a matrix A is said to be *Toeplitz-like of displacement rank at most* r if it can be written as sum of r terms, where each term is the product of a lower triangular and an upper triangular Toeplitz matrix, that is

$$A = \sum_{i=1}^{r} L(\mathbf{u}_i) L(\mathbf{v}_i)^T, \tag{39}$$

where $L(\mathbf{u})$ denotes the lower triangular Toeplitz matrix defined by its first row $\mathbf{u}$. The set of $2r$ vectors $\mathbf{u}_i$, $\mathbf{v}_i$, $i = 1, \ldots, r$ is called a *generator of rank* r of A and the formula (39) is called a *displacement representation of length* r of A. For more details on Toeplitz-like matrices and their correlations with the displacement operators, we refer to [4], [12], [15].

We recall that the product of an $m \times m$ Toeplitz matrix and a vector can be computed by means of FFT in $O(m \log m)$ ops, therefore the product of a Toeplitz-like matrix of displacement rank at most r and a vector can be computed in $O(rm \log m)$ ops once a generator of rank r is known. Moreover, given a generator of rank r of a nonsingular matrix A, it is possible to compute a generator of rank r of A^{-1} by means of the Bitmead-Anderson algorithm [6] in $O(r^2 m \log^2 m)$ ops.

Here we provide displacement representation formulae of lengths 2 and 3 for the sequences of matrix power series $H^{(n)}(z)$ and $\widehat{H}^{(z)}(z)$, respectively, and relate their displacement generators at two subsequent steps n and $n-1$. For more details we refer to [1], [3], where these properties have been proved in the context of the numerical solution of Markov chains.

To be more specific, let $\mathbf{r}^{(n)T}(z) = \mathbf{e}_m^T H^{(n)}(z)$, $\mathbf{c}^{(n)}(z) = H^{(n)}(z)\mathbf{e}_1$, be the last row and the first column of $H^{(n)}(z)$, respectively. Similarly, denote $\widehat{\mathbf{r}}^{(n)T}(z) = \mathbf{e}_m^T \widehat{H}^{(n)}(z)$, $\widehat{\mathbf{c}}^{(n)}(z) = \widehat{H}^{(n)}(z)\mathbf{e}_1$. Then the following result holds [3].

Theorem 11 *For the matrix power series $H^{(n)}(z)$, $\widehat{H}^{(n)}(z)$ generated at the n-th step of cyclic reduction we have*

$$\begin{aligned} H^{(n)}(z) &= L(\mathbf{c}^{(n)}(z))(I - L^T(Z\mathbf{u}^{(n)}(z))) + L(\mathbf{v}^{(n)}(z))L^T(Z\mathbf{r}^{(n)}(z)) \\ \widehat{H}^{(n)}(z) &= L(\widehat{\mathbf{c}}^{(n)}(z)) + L^T(Z\mathbf{r}^{(0)}(0)) + \\ &\quad L(\mathbf{c}^{(n)}(z))L^T(Z\widehat{\mathbf{u}}^{(n)}(z)) + L(\mathbf{v}^{(n)}(z))L^T(Z\widehat{\mathbf{r}}^{(n)}(z)), \end{aligned}$$

where $\mathbf{u}^{(n)}(z) = \sum_{i=0}^{+\infty} \mathbf{u}_i^{(n)} z^i$, $\widehat{\mathbf{u}}^{(n)}(z) = \sum_{i=0}^{+\infty} \widehat{\mathbf{u}}_i^{(n)} z^i$, $\mathbf{v}^{(n)}(z) = \sum_{i=0}^{+\infty} \mathbf{v}_i^{(n)} z^i$ are m-dimensional vector power series, and $Z = (z_{i,j})$ is the $m \times m$ down-shift matrix such that $z_{i+1,i} = 1$, $z_{i,j} = 0$ for $i - 1 \neq j$.

From the above result it follows that, at each step n, only the seven vector power series $\mathbf{u}^{(n)}(z)$, $\widehat{\mathbf{u}}^{(n)}(z)$, $\mathbf{v}^{(n)}(z)$, $\mathbf{c}^{(n)}(z)$, $\mathbf{r}^{(n)}(z)$, $\widehat{\mathbf{c}}^{(n)}(z)$, $\widehat{\mathbf{r}}^{(n)}(z)$, are sufficient to represent the matrix power series $H^{(n)}(z)$ and $\widehat{H}^{(n)}(z)$.

Moreover, since it holds that $H^{(n)}(z)^T = JH^{(n)}(z)J$, where J denotes the $m\times m$ permutation (reversion) matrix having unit entries along the anti-diagonal, then the vectors $\mathbf{u}^{(n)}(z)$, $\mathbf{v}^{(n)}(z)$, $\mathbf{c}^{(n)}(z)$, $\mathbf{r}^{(n)}(z)$ are related, as stated by the following [3]

Proposition 12 *For the vector power series $\mathbf{u}^{(n)}(z)$, $\mathbf{v}^{(n)}(z)$, $\mathbf{r}^{(n)}(z)$, $\mathbf{c}^{(n)}(z)$ we have $\mathbf{v}^{(n)}(z) = J\mathbf{u}^{(n)}(z)$, $\mathbf{c}^{(n)}(z) = J\mathbf{r}^{(n)}(z)$.*

In this way the number of essential vectors is reduced to 5, say, $\mathbf{c}^{(n)}(z)$, $\widehat{\mathbf{c}}^{(n)}(z)$, $\widehat{\mathbf{r}}^{(n)}(z)$, $\mathbf{v}^{(n)}(z)$, $\widehat{\mathbf{u}}^{(n)}(z)$.

Explicit relations between the above set of vector power series (generators) at two subsequent steps of cyclic reduction are given in [3].

Proposition 13 *For the generators $\mathbf{c}^{(n)}(z)$, $\widehat{\mathbf{c}}^{(n)}(z)$, $\widehat{\mathbf{r}}^{(n)}(z)$, $\mathbf{v}^{(n)}(z)$, $\widehat{\mathbf{u}}^{(n)}(z)$, we have*

$$
\begin{aligned}
&\mathbf{v}^{(n)}(z) = -\left[H^{(n-1)}(-z)\left(H_{odd}^{(n-1)}(z^2)\right)^{-1}\mathbf{v}^{(n-1)}(z)\right]_{even},\\
&\widehat{\mathbf{u}}^{(n)T}(z) = \widehat{\mathbf{u}}_{odd}^{(n-1)T}(z) + \mathbf{u}_{odd}^{(n-1)T}(z)\left(H_{odd}^{(n-1)}(z)\right)^{-1}\widehat{H}_{even}^{(n-1)}(z),\\
&\mathbf{c}^{(n)}(z) = z\mathbf{c}_{odd}^{(n-1)}(z) - H_{even}^{(n-1)}(z)(H_{odd}^{(n-1)}(z))^{-1}\mathbf{c}_{even}^{(n-1)}(z),\\
&\widehat{\mathbf{c}}^{(n)}(z) = \widehat{\mathbf{c}}_{even}^{(n-1)}(z) - H_{even}^{(n-1)}(z)(H_{odd}^{(n-1)}(z))^{-1}\widehat{\mathbf{c}}_{odd}^{(n-1)}(z),\\
&\widehat{\mathbf{r}}^{(n)T}(z) = \widehat{\mathbf{r}}_{even}^{(n-1)T}(z) - \mathbf{c}_{even}^{(n-1)T}(z)J(H_{odd}^{(n-1)}(z))^{-1}\widehat{H}_{odd}^{(n-1)}(z),
\end{aligned}
$$

for $n\geq 1$, and $\widehat{\mathbf{u}}^{(0)T}(z) = \mathbf{e}_m^T$, $\mathbf{v}^{(0)}(z) = z\mathbf{e}_1$.

In the case where $A_i = 0$ for $i > 2$ the above results are further simplified. We can give explicit expressions of the blocks $H_i^{(n)}$, $i = 0, 1, 2$, $\widehat{H}_1^{(n)}$. In fact we have the following theorem [1].

Theorem 14 *If $A_i = O$ for $i > 2$, then the matrices $H_0^{(n)}$, $H_1^{(n)}$, $H_2^{(n)}$, $H_1^{(n)^{-1}}$, $\widehat{H}_1^{(n)}$ generated by the cyclic reduction (29) can be rewritten as*

$$
\begin{aligned}
H_0^{(n)} &= L(\mathbf{c}_0^{(n)})(I - L^T(Z\mathbf{u}_0^{(n)T})) + L(\mathbf{v}_0^{(n)})L^T(Z\mathbf{r}_0^{(n)T}),\\
H_1^{(n)} &= L(\mathbf{c}_1^{(n)})(I - L^T(Z\mathbf{u}_0^{(n)T})) - L(\mathbf{c}_0^{(n)})L^T(Z\mathbf{u}_1^{(n)T}) +\\
&\quad L(\mathbf{v}_0^{(n)})L^T(Z\mathbf{r}_1^{(n)T}) + L(\mathbf{v}_1^{(n)})L^T(Z\mathbf{r}_0^{(n)T}),\\
H_2^{(n)} &= L(\mathbf{c}_2^{(n)}) - L(\mathbf{s}^{(n)})L^T(Z\mathbf{u}_1^{(n)T}) + L(\mathbf{v}_1^{(n)})L^T(Z\widehat{\mathbf{r}}_1^{(n)T}),\\
H_1^{(n)^{-1}} &= L(H_1^{(n)^{-1}}\mathbf{e}_1) + L^T(ZH_1^{(n)^{-T}}\mathbf{u}_0^{(n)T}) +\\
&\quad L(H_1^{(n)^{-1}}\mathbf{c}_0^{(n)})L^T(ZH_1^{(n)^{-T}}\mathbf{u}_1^{(n)T}) - L(H_1^{(n)^{-1}}\mathbf{v}_1^{(n)})L^T(ZH_1^{(n)^{-T}}\mathbf{r}_0^{(n)T}),\\
\widehat{H}_1^{(n)} &= L(\widehat{\mathbf{c}}^{(n)}) + L^T(Z\mathbf{r}_0^{(0)T}) - L(\mathbf{c}_0^{(n)})L^T(Z\mathbf{u}_1^{(n)T}) + L(\mathbf{v}_0^{(n)})L^T(Z\widehat{\mathbf{r}}^{(n)T}),
\end{aligned}
$$

where $\mathbf{s}^{(0)} = \mathbf{c}_1^{(0)}$, $\mathbf{s}^{(n)} = \mathbf{s}^{(n-1)} - H_2^{(n-1)}H_1^{(n-1)^{-1}}\mathbf{c}_0^{(n-1)}$, for $n\geq 1$.

In this way the blocks $H_0^{(n)}$, $H_1^{(n)}$, $H_2^{(n)}$, $H_1^{(n)^{-1}}$, $\widehat{H}_1^{(n)}$ are uniquely determined by the vectors $\mathbf{c}_i^{(n)}$, $i = 0, 1, 2$, $\mathbf{v}_i^{(n)}$, $i = 0, 1$, $\widehat{\mathbf{c}}^{(n)}$, $\widehat{\mathbf{r}}^{(n)}$, $\mathbf{s}^{(n)}$. Moreover, the relations between such vectors with the corresponding ones of the previous iterations are defined by the following equations

$$
\begin{aligned}
\mathbf{v}_0^{(n)} &= \mathbf{v}_0^{(n-1)} - H_0^{(n-1)} H_1^{(n-1)^{-1}} \mathbf{v}_1^{(n-1)},\\
\mathbf{v}_1^{(n)} &= -H_2^{(n-1)} H_1^{(n-1)^{-1}} \mathbf{v}_1^{(n-1)},\\
\mathbf{c}_0^{(n)} &= -H_0^{(n-1)} H_1^{(n-1)^{-1}} \mathbf{c}_0^{(n-1)},\\
\mathbf{c}_1^{(n)} &= \mathbf{c}_1^{(n-1)} - H_0^{(n-1)} H_1^{(n-1)^{-1}} \mathbf{c}_2^{(n-1)} - H_2^{(n-1)} H_1^{(n-1)^{-1}} \mathbf{c}_0^{(n-1)},\\
\mathbf{c}_2^{(n)} &= -H_2^{(n-1)} H_1^{(n-1)^{-1}} \mathbf{c}_2^{(n-1)},\\
\widehat{\mathbf{c}}^{(n)} &= \widehat{\mathbf{c}}^{(n-1)} - H_0^{(n-1)} H_1^{(n-1)^{-1}} \mathbf{c}_2^{(n-1)},\\
\widehat{\mathbf{r}}^{(n)T} &= \widehat{\mathbf{r}}^{(n-1)T} - \mathbf{c}_0^{(n-1)T} J H_1^{(n-1)^{-1}} H_2^{(n-1)},\ n \geq 1,
\end{aligned} \tag{40}
$$

where $\mathbf{v}_0^{(0)} = \mathbf{0}$, $\mathbf{v}_1^{(0)} = \mathbf{e}_1$.

Similar representation formulae can be derived for the blocks obtained with a jump of arbitrary order.

It is a simple matter to verify that, if $A_i = O$ for $i > 2$, the cyclic reduction step can be performed by means of (40) in $O(m \log^2 m)$ ops. In fact only one Toeplitz-like matrix inversion and a constant number of multiplications of Toeplitz-like matrices and vectors must be computed.

A slightly more complicated situation occurs in the general Hessenberg case where the Toeplitz-like matrices involved, and therefore their generators, have coefficients in the algebra $\mathcal{F}$ of power series. In this case it is possible to use the evaluation/interpolation technique at the roots of 1 (Fourier points) in order to reduce multiplication and inversion of power series to the corresponding operations among complex numbers. The number of Fourier points that must be used in order to have a good accuracy in the results depends on the numerical degree of the power series involved.

We say that a matrix power series $\Gamma(z) \in \mathcal{F}^{m\times n}$, $\Gamma(z) = \sum_{i=0}^{+\infty} \Gamma_i z^i$, $\Gamma_i \in \mathbf{C}^{m\times n}$, convergent in the closed unit disk, has ϵ-degree N if N is the minimum integer such that $\mathbf{e}^T \sum_{i=N+1}^{+\infty} |\Gamma_i| \leq \epsilon \mathbf{e}^T$, where $\mathbf{e}$ denotes the $m-$dimensional vector having entries equal to 1. If ϵ is the machine precision of the floating arithmetic used, then N is called the numerical degree of the power series. In the case where $\Gamma(z)$ reduces to a matrix polynomial of degree N then, for any $\epsilon \geq 0$, its ϵ-degree is bounded from above by N. Further, whenever we know that the result of an arithmetic operation over $\mathcal{F}$ has ϵ-degree N, then it can be approximated within the error bound ϵ by the result of the corresponding operation between polynomials of degree less than N. In this way, the following complexity estimates can be established [2], [3].

Theorem 15 *Assume we are given a vector $\mathbf{u}(z) \in \mathcal{F}^{m\times 1}$ and an $m \times m$ Toeplitz-like matrix power series $\Gamma(z) \in \mathcal{F}^{m\times m}$, both convergent in the closed unit disk. Then the first N coefficients of the product $\Gamma(z)\mathbf{u}(z)$ can be approximated within the error bound ϵ at the cost of $O(mN \log(mN))$ arithmetic operations, where N is the ϵ-degree of $\Gamma(z)\mathbf{u}(z)$. Moreover, if $\Gamma(z)^{-1}$ exists, is convergent in the closed unit disk, and has ϵ-degree equal to N, then its first N coefficients can be approximated within the error bound ϵ at the overall cost of $O(mN(\log^2 m + \log(mN)))$ ops.*

In the light of this result we may prove that the cost of one step of cyclic reduction applied to the Hessenberg case amounts to $O(mN(\log(mN) + \log^2 m))$ ops, if N is an upper bound on the numerical degrees of the power series involved in the computation.

6 Implementation of the algorithm

In this section we discuss some issues related to the implementation of the cyclic reduction scheme for determining the coefficients of the polynomial $r(z)$ having the

h zeros of minimum modulus of an analytic function $f(z)$. We assume that $f(z)$ is assigned in terms of the coefficients a_i of its Taylor expansion.

The algorithm proceeds as follows. Assuming that $|r_h| < |r_{h+1}|$, the cyclic reduction process is applied to the infinite Toeplitz matrix in block Hessenberg form $H_\infty^{(0)}$ of (27) partitioned into $m \times m$ blocks H_i. In general, for power series, in the light of Theorem 7 we set $m = h$ and then check the convergence of the sequence $-A_0(\widehat{H}_1^{(n)})^{-1}\mathbf{e}_1$, $n \geq 1$, to the vector of the coefficients of $r(z)$.

For a polynomial $f(z)$ of degree N, we apply Corollary 2 and choose $m = \max\{h, N-h\}$. In fact in this way, the block Hessenberg matrix A reduces to the block tridiagonal form so that the algorithm is much simplified.

Using computations with infinite matrices allows us to estimate dynamically the depth of the recursion phase needed in order to reach the desired accuracy. Nevertheless, in the analytic case where $H^{(0)}$ is block Hessenberg, this approach implies some restrictions on the input functions which our method can be applied to. Roughly speaking, matrix operations among infinite block Hessenberg matrices can be carried out within an arbitrary high precision whenever they can numerically approximated by infinite block band matrices. This always happens in the polynomial case without additional conditions on $p(z)$ since $H^{(i)}$ are block tridiagonal for any $i \geq 0$.

6.1 The polynomial case

In the case of polynomials of degree N, the matrix (27) has a block tridiagonal structure. If $H_1^{(n)}$ is nonsingular, then the cyclic reduction step from $H^{(n)}$ to $H^{(n+1)}$ is described by means of the relations (31), defining the four $m \times m$ block entries $H_i^{(n+1)}$, $i = 0, 1, 2$, and $\widehat{H}_1^{(n+1)}$ of the block tridiagonal matrix $H^{(n+1)}$. Otherwise, if $H_1^{(n)}$ is singular, we consider the matrix obtained from $H^{(n)}$ by removing the first block row and the first block column. Then we check the invertibility of its block leading principal sub-matrices of order $2^s - 1$, $s = 1, 2, \ldots$, until we find a nonsingular matrix of order $2^r - 1$. We compute the block tridiagonal matrix $H^{(n+r)}$ from $H^{(n)}$ by means of a generalization of (38) that holds for $r = 2$.

It is clear that this look-ahead strategy can be also applied in order to skip computations involving ill-conditioned matrices by replacing the singularity test with a suitable ill-conditioning test. If we impose no restriction on the maximum step size, the accuracy of the algorithm is guaranteed but, at the same time, the computational cost increases. In order to control the cost of the algorithm, one may perform the step size test for s ranging from 1 to a fixed value L which is assumed to be independent of N. It cannot be excluded that the algorithm has reached the maximum step size L without satisfying the test on the conditioning. In this case, we increase the upper bound on the condition number in such a way to satisfy our test in the maximal look-ahead range.

6.2 The analytic case

In the case of general analytic functions, the matrix (27) has a block Hessenberg structure and we choose $m = h$. Thus the cyclic reduction algorithm is described by means of the functional relations (29), defining the first and the second block column of the block Hessenberg matrix obtained at each step. In this way the matrix operations are recasted in terms of manipulations among power series with complex coefficients.

Suppose now to know the coefficients of the matrix power series $\widehat{H}^{(n)}(z)$, $H^{(n)}(z)$, and assume that $\widehat{H}^{(n+1)}(z)$, $H^{(n+1)}(z)$ are convergent in the closed unit circle, and have numerical degree at most d_{n+1}. Then, in order to approximate the coefficients of the matrix power series $\widehat{H}^{(n+1)}(z)$ and $H^{(n+1)}(z)$, it is sufficient to evaluate/interpolate the functional relations (29) at the $(d_{n+1}+1)$-st roots of 1.

Since $\widehat{H}^{(n+1)}(z)$ and $H^{(n+1)}(z)$ are Toeplitz-like and, therefore, are completely specified by their generators (see Theorem 11), then, in the light of Theorem 15, the computational cost of the above computation is $O(hd_{n+1}(\log^2 h + \log(hd_{n+1})))$.

The practical feasibility of our approach for the numerical approximation of a factor is therefore restricted to the input set of analytic functions such that the numerical degree of the power series involved in the computations can be either bounded apriori or estimated dynamically. Fortunately, as we will show in the sequel of this section, many important functions arising in problems of probability and statistics fall in this set. Moreover, it should be emphasized that there is some numerical evidence that other important functions of mathematical physics, such as the Bessel functions, can be dealt with our techniques. This point deserves further investigations.

For certain analytic functions, it is possible to devise a simple criterium that allows to dynamically estimate the numerical degree of the matrix power series generated at each step of cyclic reduction. The analytic functions we are referring to are for example those of the form $a(z) = z^h - \sum_{i=0}^{\infty} \alpha_i z^i$, where $\alpha_i \geq 0$, $i = 0, 1, \ldots$, $\sum_{i=0}^{\infty} \alpha_i = 1$. These functions arise in the modeling of Markov chains of M/G/1 type [17], where $\alpha(z) = \sum_{i=0}^{\infty} \alpha_i z^i$ is a probability distribution. By an easy application of the Rouché theorem [13] there follows that, if the Markov chain is irreducible and positive recurrent then $a(z)$ has h zeros in closed unit circle. Here, the fundamental problem is that of computing the coefficients of the factor corresponding to these zeros in the closed unit circle.

It can be easily verified that the cyclic reduction algorithm applied to the matrix $H^{(0)}$ formed from the coefficients of $a(z)$ never stops for the occurrence of a breakdown. In addition, at each step n of the cyclic reduction process the $h \times h$ blocks $H_0^{(n)}$, $H_i^{(n)}$, $\widehat{H}_i^{(n)}$, $i > 1$ are nonpositive, the off-diagonal entries of $H_1^{(n)}$, $\widehat{H}_1^{(n)}$ are nonpositive, the matrix power series $H^{(n)}(z)$, $\widehat{H}^{(n)}(z)$ are convergent in the closed unit circle, and

$$\mathbf{e}^T H^{(n)}(1) = \mathbf{0}^T, \quad \mathbf{e}^T \widehat{H}^{(n)}(1) + \mathbf{e}^T H_0 = \mathbf{0}^T.$$

Moreover, it is possible to estimate the numerical degree of the power series that will be computed at the subsequent step of cyclic reduction. More precisely, if $P(z)$ and $\widehat{P}(z)$ are matrix polynomials obtained by evaluating/interpolating functional relations (29) at the N-th roots of 1, then the following results provide a test to check if $P(z)$ and $\widehat{P}(z)$ are good approximations of $H^{(n+1)}(z)$ and $\widehat{H}^{(n+1)}(z)$, respectively. Hereafter we set, for any $n \geq 0$,

$$\Gamma^{(n)}(z) = zI - H^{(n)}(z) = \sum_{i=0}^{+\infty} \Gamma_i^{(n)} z^i, \quad \Gamma_i^{(n)} \in \mathbf{C}^{h \times h},$$

and

$$\widehat{\Gamma}^{(n)}(z) = I - \widehat{H}^{(n)}(z) = \sum_{i=0}^{+\infty} \widehat{\Gamma}_{i+1}^{(n)} z^i, \quad \widehat{\Gamma}_i^{(n)} \in \mathbf{C}^{h \times h}.$$

Recursive relations for the series $\Gamma^{(n)}(z)$ and $\widehat{\Gamma}^{(n)}(z)$ can be easily obtained from the following modifications of (29)

$$\begin{cases} \widehat{\Gamma}^{(n+1)}(z) = \widehat{\Gamma}_{even}^{(n)}(z) + \Gamma_{even}^{(n)}(z) \left(I - \Gamma_{odd}^{(n)}(z) \right)^{-1} \widehat{\Gamma}_{odd}^{(n)}(z) \\ \Gamma^{(n+1)}(z) = z\Gamma_{odd}^{(n)}(z) + \Gamma_{even}^{(n)}(z) \left(I - \Gamma_{odd}^{(n)}(z) \right)^{-1} \Gamma_{even}^{(n)}(z). \end{cases} \tag{41}$$

The following theorems from [2] provide an efficient tool to check the accuracy of the approximations $P(z)$, $\widehat{P}(z)$ to $\Gamma^{(n+1)}(z)$, $\widehat{\Gamma}^{(n+1)}(z)$, respectively.

Theorem 16 *Let, for any $n \geq 0$,*

$$\boldsymbol{\alpha}^{(n)T} = \mathbf{e}^T \sum_{i=1}^{+\infty} i\Gamma_i^{(n)}, \; \widehat{\boldsymbol{\alpha}}^{(n)T} = \mathbf{e}^T \sum_{i=1}^{+\infty} i\widehat{\Gamma}_{i+1}^{(n)}.$$

Then, taking derivatives of (41) yields the following recursive relations :

$$\begin{aligned} \boldsymbol{\alpha}^{(n+1)T} &= \left(\mathbf{e}^T + \boldsymbol{\alpha}^{(n)T} - (\mathbf{e}^T - \boldsymbol{\alpha}^{(n)T})(I - \Gamma_{odd}^{(n)}(1))^{-1}\Gamma_{even}^{(n)}(1)\right)/2 \, , \\ \widehat{\boldsymbol{\alpha}}^{(n+1)T} &= \left(\widehat{\boldsymbol{\alpha}}^{(n)T} - (\mathbf{e}^T - \boldsymbol{\alpha}^{(n)T})(I - \Gamma_{odd}^{(n)}(1))^{-1}\widehat{\Gamma}_{odd}^{(n)}(1)\right)/2. \end{aligned} \tag{42}$$

Theorem 17 *Let, at step n, $P(z) = \sum_{i=0}^{N-1} P_i z^i$ and $\widehat{P}(z) = \sum_{i=0}^{N-1} \widehat{P}_i z^i$ be the matrix polynomials of degree $N-1$ interpolating the series $\Gamma^{(n)}(z)$, $\widehat{\Gamma}^{(n)}(z)$ at the N-th roots of unity. Then the following inequalities hold:*

$$\mathbf{e}^T \textstyle\sum_{i=N}^{+\infty} \Gamma_i^{(n)} \leq \boldsymbol{\alpha}^{(n)T} - \mathbf{e}^T \sum_{i=1}^{N-1} iP_i \, ,$$

$$\mathbf{e}^T \textstyle\sum_{i=N+1}^{+\infty} \widehat{\Gamma}_i^{(N)} \leq \widehat{\boldsymbol{\alpha}}^{(n)T} - \mathbf{e}^T \sum_{i=1}^{N-1} i\widehat{P}_i.$$

Now let us apply the above results. Suppose we know the coefficients of the series $\Gamma^{(n)}(z)$, $\widehat{\Gamma}^{(n)}(z)$, and compute the coefficients P_i, $\widehat{P}_i$, $i = 0, \ldots, N-1$, of the approximations $P(z) = \sum_{i=0}^{N-1} P_i z^i$, $\widehat{P}(z) = \sum_{i=0}^{N-1} \widehat{P}_i z^i$ of the series $\Gamma^{(n+1)}(z)$, $\widehat{\Gamma}^{(n+1)}(z)$ by interpolating the functional relations (41) at the N-th roots of unity. From Theorem 17 it follows that, if

$$\boldsymbol{\alpha}^{(n+1)T} - \mathbf{e}^T \sum_{i=1}^{N-1} iP_i \leq \epsilon \mathbf{e}^T \, , \tag{43}$$

and

$$\widehat{\boldsymbol{\alpha}}^{(n+1)T} - \mathbf{e}^T \sum_{i=1}^{N-1} i\widehat{P}_i \leq \epsilon \mathbf{e}^T, \tag{44}$$

then the series $\Gamma^{(n+1)}(z)$, $\widehat{\Gamma}^{(n+1)}(z)$ have ϵ-degree $N-1$ and the matrix coefficients of $P(z)$ and $\widehat{P}(z)$ are approximations of the corresponding coefficients of $\Gamma^{(n+1)}(z)$, $\widehat{\Gamma}^{(n+1)}(z)$ within the error ϵ. Otherwise, if (43) and (44) are not satisfied, then we determine $P(z)$ and $\widehat{P}(z)$ by interpolation of (41) at the $2N$-th roots of unity and repeat the tests until they hold.

It is important to point out that (43) and (44) can be easily applied without knowing the coefficients of the series $\Gamma^{(n+1)}(z)$, $\widehat{\Gamma}^{(n+1)}(z)$. In fact, the mean values $\boldsymbol{\alpha}^{(n+1)}$ and $\widehat{\boldsymbol{\alpha}}^{(n+1)}$ can be explicitly obtained by means of (42) at the cost of $O(\theta(h))$ arithmetic operations.

7 Numerical experiments

We implemented our algorithm in Mathematica^{TM} both for the full factorization of a polynomial into linear factors and for the computation of only one factor of a power series.

In this set of preliminary experiments, we did not exploit the Toeplitz-like structure of the blocks generated by the cyclic reduction process. In fact, our aim was to check the convergence properties and the numerical stability of the basic cyclic reduction scheme. A complete implementation, based on FFT and the Toeplitz matrix technology, is planned in a subsequent stage in the C language. For this reason we did not report the cpu time needed by the execution of this implementation that, however, in most cases is slightly higher than the time needed by the built-in function MSolve, based on the Jenkins-Traub algorithm. Instead, we report the number of iterations required by our algorithm, the number of digits of working precision used in the computation, and the number of digits of output precision.

Our implementation recursively applies cyclic reduction for factoring the given polynomial into linear factors, in this way a polynomial of degree N, say $N = 16$, is factorized into two polynomials of degree $n/2$, the two polynomials are factorized into four polynomials of degree $/4$ and so forth, until we arrive at polynomials of degree 1 or 2 which are explicitly solved. In the tables we report the values N of the degrees and the maximum number of iterations needed to factorize the polynomials of degree N.

As test suite for our experiments we considered the following polynomials and power series.

- Wilkinson polynomials $p(z) = \prod_{i=1}^{N}(z - i)$ of degree $N \in \{20, 40, 80\}$. They are characterized by a large condition number (about 10^{14} for $N = 20$).

- Mignotte-like polynomial with the shifted variable $p(z) = q(z + \underline{i})$, where $q(z) = z^{20} + (1000z - 1)^2$ and $\underline{i}^2 = -1$. The polynomial $q(z)$ has two very close zeros near $1/1000$ with a large condition number (about 10^{27}). The remaining zeros are displaced roughly along a circle.

- The exponential polynomials $p(z) = \sum_{i=0}^{N} \frac{(z-N/4)^i}{i!}$, $N \in \{16, 32, 64, 128\}$, obtained by truncating the power series of the exponential function at the term of degree N, where the exponential function is evaluated at $z - N/4$. The zeros of these polynomials are displaced along a curve and have condition numbers that grow with N.

- The power series

  $$e(z) = z^8 - e^{-\lambda} \sum_{i=0}^{+\infty} \frac{\lambda^i}{i!} z^i, \quad 0 < \lambda < 8,$$

  that is frequently encountered in queueing problems where the goal is to compute the factor of $e(z)$ of degree 8 having zeros inside the unit disk.

- The power series of the Bessel function

  $$p(z) = J_0(2\sqrt{150z}) = \sum_{i=0}^{+\infty} \frac{(-150z)^i}{i!i!},$$

  where the scaling of the variable is such that $p(z)$ has 8 zeros in $[0, 1]$.

The following tables report the number of iterations required at all the intermediate steps with a working precision of wp digits. We also report the number d of correct output digits and $\log_{10}$ cond, the logarithm of the condition number of the worst conditioned zero ξ estimated by evaluating cond $= \sum_{i=0}^{N} |a_i||\xi|^i/|\xi p'(\xi)|$, where $p(z) = \sum_{i=0}^{N} a_i z^i$, and ξ is a simple zero. According to the conditioning theorems, from a numerical algorithm we cannot pretend more than $d \approx wp - \log_{10}$ cond correct digits.

For the Wilkinson polynomial of degree 20, for the Mignotte-like polynomial, and for the exponential polynomial of degree 20, the number of correct output

N	20	10	5	3
it	7	9	10	10

Table 1: Wilkinson polynomial, $N = 20$, $wp = 20$, $d = 8$, $\log_{10} \mathrm{cond} = 13$.

N	40	20	10	5	3
it	8	10	11	12	13

Table 2: Wilkinson polynomial, $N = 40$, $wp = 50$, $d = 11$, $\log_{10} \mathrm{cond} = 28$.

N	80	40	20	10	5	3
it	9	11	12	13	14	15

Table 3: Wilkinson polynomial, $N = 80$, $wp = 100$, $d = 8$, $\log \mathrm{cond} = 59$.

digits respected the above relation. For instance, the algorithm computed at least 8 digits of the zeros of the Wilkinson polynomial of degree 20 with $wp = 20$.

For the Wilkinson polynomial of degree 40 with $wp = 50$ we obtained at least $d = 11$ correct digits, while with $wp = 100$ we obtained at least $d = 61$ correct digits. Since in this case $\log_{10} \mathrm{cond}$ is about 28, we have a loss of about 11 digits.

A similar situation holds for the Wilkinson polynomial of degree 80 and for the exponential polynomial of degree 64 and 128. In the former case we obtain a number of correct digits in the zeros that ranges from 100 to 8, where $wp = 100$. Since the condition number in the worst-case is about 10^{59}, we have a loss of about 33 digits in the approximation of the worstly conditioned zero.

The extra amount of working precision, as it results from the numerical experiments, can be estimated as an additive $O(N)$ term. This need of a higher working precision is typical of all the algorithms based on recursive splitting of a polynomial into factors and can be even proved theoretically [18].

N	20	10	5	3
it	8	21	21	23

Table 4: Mignotte-like polynomial, $N = 20$, $wp = 80$, $d = 45$, $\log \mathrm{cond} = 27$.

One more consideration related to this stability issues, is that the algorithm is not self-correcting. This fact may suggest that a substantial improvement could be obtained if any factor that is approximated in the intermediate steps of our algorithm is refined by means of functional iterations, using the uncorrupted information stored in the coefficients of the initial polynomial, before the factor itself is given as input to the recursive step of cyclic reduction.

We approximated the 8 zeros of minimum modulus of the two power series $e(z)$ (with $\lambda = 1/2$) and $J_0(2\sqrt{150z})$ by means of the 8 zeros of minimum modulus of the corresponding polynomials obtained by truncating both the power series at the term of degree 256,. The zeros of $e(z)$ are fairly well-conditioned whereas the maximum of the condition numbers of the zeros of the Bessel function is about $3.0\ 10^9$. We applied the cyclic reduction algorithm to $A^{(32)}$ with $h = 8$ and $wp = 16$ in order to approximate a factor of degree 8. Then we approximated the zeros of these factors by means of the Mathematica^{TM} function NSolve. The number of correct digits in the approximations obtained in this way has been $d = 15$ for the power series $e(z)$ and $d = 6$ for the power series $J_0(z)$ in accordance with the condition numbers of the corresponding zeros.

N	16	8	4
it	6	7	8

Table 5: Exponential polynomial, $N = 16$, $wp = 20$, $d = 20$.

N	32	16	8	4
it	6	7	8	9

Table 6: Exponential polynomial, $N = 32$, $wp = 20$, $d = 14$.

N	64	32	16	8	4
it	7	8	9	11	11

Table 7: Exponential polynomial, $N = 64$, $wp = 40$, $d = 17$.

N	128	64	32	16	8	4
it	8	9	10	11	12	13

Table 8: Exponential polynomial, $N = 128$, $wp = 80$, $d = 21$.

8 Conclusions and further extensions

We have reduced the problem of factoring a polynomial $f(z)$ of degree N to that of computing the first N entries in the first column of the inverse of an infinite banded Toeplitz matrix with bandwidth N. If $f(z)$ is a power series, then the Toeplitz matrix is in block Hessenberg form. An algorithm for performing this computation, based on the cyclic reduction technique, has been provided. This algorithm computes part of the LU factorization of the infinite band matrix with rows and columns suitably permuted.

If the banded matrix is diagonally dominant, then well known results in numerical linear algebra guarantee the numerical stability of the LU factorization.

It would be interesting to give conditions on the distribution of the zeros of the polynomial under which the banded matrix, suitably scaled, is diagonally dominant.

It is rather immediate to give necessary conditions for the diagonal dominance. In fact, if the matrix is diagonally dominant, i.e., if $|a_m| > \sum_{i=0, i\neq m}^{N} |a_i|$, where a_m is the diagonal entry, by applying Rouché theorem to the polynomials $\sum_{i=0}^{N} a_i z^i$ and $a_m z^m$, we deduce that m zeros are inside the unit disk and $N - m$ zeros are outside. It is even possible to express the isolation ratio of these two sets of zeros (i.e., the ratio R/r, where $r < 1 < R$ and both the disks of center 0 and radius r, R contain m zeros), in terms of the ratio $|a_m| / \sum_{i=0, i\neq m}^{N} |a_i|$ that measures the diagonal dominance of the matrix.

It would be more interesting to give conditions on the zeros of a polynomial sufficient to guarantee the diagonal dominance of the associated matrix.

References

[1] D. Bini, B. Meini, *Effective Methods for Solving Banded Toeplitz Systems*, to appear in SIAM J. Matrix Anal. Appl..

[2] D. Bini, B. Meini, *Improved Cyclic Reduction for Solving Queueing Problems*, Numerical Algorithms, **15** (1997), pp. 57-74.

[3] D. Bini, B. Meini, *Using Displacement Structure for Solving Non-Skip-Free M/G/1 Type Markov Chains* , Advances in Matrix Analytic Methods for Stochastic Models, A. Alfa and S. Chakravarthy Eds., 1998, Notable Publications Inc., NJ, pp. 17–37.

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

[5] D. Bini, V. Y. Pan, *Graeffe's, Chebyshev-like and Cardinal's Processes for Splitting a Polynomial into Factors*, J. of Complexity, **12** (1996), pp. 492-511.

[6] R. R. Bitmead and B. D. O. Anderson, *Asymptotically Fast Solution of Toeplitz and Related Systems of Linear Equation*, Linear Algebra Appl. **34**, (1980), pp. 103–116.

[7] B. L. Buzbee, G. H. Golub, C. W. Nielson, *On Direct Methods for Solving Poisson's Equations*, SIAM J. Num. Anal., **7** (1970), pp. 627-656.

[8] J. P. Cardinal, *On Two Iterative Methods for Approximating the Roots of a Polynomial*, Proc. of Workshop: Mathematics of Numerical Analysis: Real Numbers Algorithms, Park City, Utah, (1995) (J. Renegar, M. Shub and S. Smale Eds.), Lect. in Appl. Math., Amer. Math. soc., Providence, RI.

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

[10] L. Gemignani, *Computing a Factor of a Polynomial by means of Multishift LR Algorithms*, SIAM J. Matrix Anal. Appl., **19** (1998), pp. 161-181.

[11] L. Gemignani, *Polynomial Root Computation by means of the LR Algorithm*, BIT, **37** (1997), pp. 333-345.

[12] G. Heinig, K. Rost, *Algebraic Methods for Toeplitz-like Matrices and Operators*, Oper. Theory 13, Birkhäuser, Boston, MA, 1984.

[13] P. Henrici, *Applied and Computational Complex Analysis*, Vol. 1, John Wiley and Sons, New York, NY, 1974.

[14] 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.

[15] T. Kailath, A. H. Sayed, *Displacement Structure: Theory and Applications*, SIAM Review, **37** (1995), pp. 297-386.

[16] P. Kirrinnis, *Partial Fraction Decomposition in* $\mathbf{C}(z)$ *and Simultaneous Newton Iteration for Factorization in* $\mathbf{C}[z]$, J. of Complexity, to appear.

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

[18] V. Y. Pan, *Optimal and Nearly Optimal Algorithms for Approximating Polynomial Zeros*, Comput. Math. Appl., **31** (1996), pp. 97-138.

[19] V. Y. Pan, *Deterministic Improvement of Complex Polynomial Factorization Based on the Properties of Associated Resultant*, Comput. Math. Appl., **30** (1995), pp. 237-250.

[20] V. Pták, *Lyapunov, Bézout and Hankel*, Linear Algebra Appl., **58** (1984), pp. 363-390.

[21] A. Schoenhage, *The Fundamental Theorem of Algebra in Terms of Computational Complexity*, manuscript, 1982.

[22] G. W. Stewart, *On a Companion Operator for Analytic Functions*, Numer. Math., **18** (1971), pp. 26-43.

[23] J. Verschelde, *Homotopy Continuation Methods for Solving Polynomial Systems*, PhD Dissertation, Katholieke Universiteit Leuven, 1996.