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 $n \times n$ Toeplitz matrix in block Hessenberg form for sufficiently large values of n. This matrix is reduced to a band matrix if $f(z)$ is a polynomial. We prove that the factorization problem can be also reduced to solving a matrix equation $\phi(X) = 0$ for an $h \times h$ matrix X, where ϕ is a matrix power series whose coefficients are Toeplitz matrices. The function ϕ is reduced to a matrix polynomial of degree 2 if $f(z)$ is a polynomial of degree N and $h \geq N/2$.

These reductions allow us to 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.

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.

Applications to solving spectral factorization problems and Markov chains are also shown.

Keywords. Koenig's theorem, Toeplitz matrices, polynomial zeros, zeros of analytic functions, polynomial factorization, spectral factorization, Markov chains.

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 ξ_i, $i = 1, 2, \ldots$, be the zeros of $f(z)$ such that $|\xi_{i+1}| \geq |\xi_i|$

*Supported by the European project FRISCO LTR 21.024 and by the CNR coordinated project TLI

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

for any i and $|\xi_{h+1}| > |\xi_h|$ for a given $h > 0$. Define $r(z) = \prod_{i=1}^{h}(z - \xi_i)$ 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

$$A = \begin{bmatrix} A_1 & A_0 & & & \bigcirc \\ A_2 & A_1 & A_0 & & \\ \vdots & \ddots & \ddots & \ddots & \end{bmatrix}. \tag{1}$$

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

$$\mathbf{v}^{(j)} = \widehat{\mathbf{v}} + O(\sigma^{jh}), \quad |\xi_h/\xi_{h+1}| < \sigma < 1. \tag{2}$$

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 also prove that the factorization problem is closely related to the solution of the matrix equation

$$\sum_{i=0}^{+\infty} X^i A_i = 0 \tag{3}$$

where X denotes the $m \times m$ unknown matrix and the coefficients A_i are Toeplitz matrices. In fact, we prove that the equation (3) has a solution whose first column is the coefficient vector $\widehat{\mathbf{v}}$ of the seeked polynomial factor. This allows us to adjust any algorithm for the solution of (3), like the functional iterations of [12],[23], [25], to factoring power series. Moreover, if $f(z)$ is a polynomial of degree N and if $m \geq N/2$, then the function in the left-hand side of (3) reduces to a matrix polynomial of degree 2.

We provide an adaptation of the cyclic reduction algorithm —originally introduced for the numerical solution of the Poisson equation [9] and adjusted in [2], [3] to solve banded Toeplitz linear systems and for the numerical solution of certain problems in queueing theory— for computing the sequence of vectors $\mathbf{v}^{(2^k)}$, $k = 1, 2, \ldots$.

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 $\mathbf{v}^{(2^k)}$ given $\mathbf{v}^{(2^{k-1})}$, 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 to the numerical degrees of the power series involved in the computation of $\mathbf{v}^{(2^k)}$. 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 is related to the numerical conditioning 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, our algorithm results to be numerically stable if A is diagonally dominant or positive definite. The latter property is satisfied in the problems of spectral factorization [33],[35], [17], where we have to compute the coefficients of a polynomial $p(z)$ of degree h given the coefficients of the polynomial $f(z) = p(z)p(z^{-1})z^h$ of degree $N = 2h$, where $p(z)$ has all its roots of modulus less than 1. The property of diagonal dominance is satisfied in the wide variety of problems in queueing theory where the numerical solution of Markov chains is requested [26]. In this case we have to compute the coefficients of the polynomial $\prod_{i=1}^{h}(z-\xi_i)$ where ξ_i are the zeros of $f(z)$ of modulus at most 1 and $a_i \leq 0$, $i \neq h$, $a_h = -\sum_{i=0,\ i\neq h}^{+\infty} a_i$.

The proof of (2) is based on some closely related extensions of the Koenig theorem developed by Householder [20]. The algorithm for solving the matrix problem is based, besides on the technique of cyclic reduction, on the concept of displacement rank [21], [18]. 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 $2m$.

The algorithm in its implementation requires the efficient inversion of Toeplitz-like matrices. This task can be accomplished by means of the (preconditioned) conjugate gradient techniques if the matrix A is positive definite (as in the spectral factorization problems [33], [35]) or diagonally dominant (as in the case of Markov chains). In the remaining cases the super fast algorithm designed and implemented by Kravanja and Van Barel [32] or the more recent techniques of [5] can be adopted.

Besides applications in Markov chains and for the spectral factorization, our algorithm can be used as a new tool for the recursive splitting of a polynomial into linear factors. 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 [34].

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 [20] and [31] 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 [13], [14], [15], 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 [30] and his "splitting circle method", the results of Kirrinnis [22] and the most recent use of mixed algebraic geometric and analytic techniques made by V. Pan and leading to the best complexity estimates [27], [28], where the basic tool for recursive splitting is the Graeffe iteration. Recursive splitting has been also considered in [10] and [7] 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 a Toeplitz system, and, equivalently, to solving the matrix equation (3). In section 3 we describe our algorithm, based on cyclic reduction, analyze its structural and computational properties and reformulate it in functional form. In section 4 we analyze the breakdown situations where the cyclic reduction algorithm cannot be carried out and propose suitable techniques to overcome this problem. In section 5, we report the results of the numerical experiments performed with a Fortran 90 implementation of our algorithm. As test suite we considered some problems of spectral factorization (among which the one presented in [24]), of polynomial factorization and solution of Markov chains. Conclusions and possible development are drawn in section 6.

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

$$f(z) = 1 + \sum_{i=1}^{+\infty} a_i z^i \tag{4}$$

with radius of convergence R so that [19] $a_i = \frac{1}{2\pi \underline{i}} \int_{|z|=\rho} f(z) z^{-i-1} dz$, 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

$$|a_j| \leq \mu \rho^{-j}, \quad j \geq 0. \tag{5}$$

We associate with the power series (4) the infinite lower triangular Toeplitz matrix

$$T = \begin{bmatrix} 1 & & & 0 \\ a_1 & 1 & & \\ a_2 & a_1 & 1 & \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix} \tag{6}$$

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 a polynomial factor by solving a Toeplitz system in block Hessenberg form.

For this purpose, let us denote by ξ_1, ξ_2, ξ_3, ..., the zeros of $f(z)$ (if they exist), where zeros of multiplicity j appear as j consecutive entries in the list, and assume that the zeros have been ordered so that $0 < |\xi_1| \le |\xi_2| \le |\xi_3| \le \ldots < R$. Further, if $|\xi_h| < |\xi_{h+1}|$, define $r_h(z) = \prod_{i=1}^{h}(1-\xi_i^{-1}z)$, and partition the matrix T of (6) into $h \times h$ Toeplitz blocks A_i, $i = 0, 1, \ldots$, so that

$$T = \begin{bmatrix} A_0 & & & 0 \\ A_1 & A_0 & & \\ A_2 & A_1 & A_0 & \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{7}$$

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 (4) and zeros $\xi_1, \xi_2, \ldots$, such that $0 < |\xi_1| \le |\xi_2| \le \ldots \le |\xi_h| < |\xi_{h+1}| \le \ldots < R$. Denote $r_h(z) = \prod_{i=1}^{h}(1-\xi_i^{-1}z) = \sum_{i=0}^{h} u_i z^i$ a factor of $f(z)$, and for any $j > 1$ consider the block linear system*

$$\left[X_1^{(j)}, X_2^{(j)}, \ldots, X_j^{(j)}\right] \begin{bmatrix} A_1 & A_0 & & 0 \\ A_2 & A_1 & \ddots & \\ \vdots & \ddots & \ddots & A_0 \\ A_j & \ldots & A_2 & A_1 \end{bmatrix} = [-A_0, 0, \ldots, 0] \tag{8}$$

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, $i \ge 0$. Then there exists an integer k such that $A^{(j)}$ is nonsingular for any $j \ge k$ and, moreover,

$$X_1^{(j)} \mathbf{e}_1 = \begin{bmatrix} -u_0/u_h \\ \vdots \\ -u_{h-1}/u_h \end{bmatrix} + O(\sigma^{jh}), \tag{9}$$

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

Notice that, for $0 \le j \le h$, u_j/u_h is the coefficient of z^j in the monic polynomial $r_h(z)/u_h = \prod_{i=1}^{h}(z-\xi_i)$. Further, observe that for $h = 1$ the first entry of the solution of (8) 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 theorem [20]. 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.

Another reduction of the problem of power series factorization is given by the following result.

Theorem 2 *In the hypotheses of Theorem 1 the matrix equation*

$$\sum_{i=0}^{+\infty} X^i A_i = 0$$

has the solution $X = JF^{-h}J$, where J is the permutation matrix with unit anti-diagonal entries and F is the Frobenius matrix associated with the polynomial $z^h r_h(z^{-1})$, that is,

$$F = \begin{bmatrix} 0 & & & -u_h \\ 1 & \ddots & & -u_{h-1} \\ & \ddots & 0 & \vdots \\ & & 1 & -u_1 \end{bmatrix}. \tag{10}$$

Moreover, $\lim_j X_i^{(j)} = JF^{-hi}J$, $i \geq 1$.

Observe that, due to Theorem 2 any algorithm for solving matrix equations, like the functional iteration techniques, can be adjusted to solve the polynomial factorization problem or to refine a computed approximation to a given factor.

It is easy to check that, under the hypothesis of Theorem 1, the super-diagonal entry A_0 is nonsingular; this fact will be repeatedly used in the proof of Theorem 1 itself. However, the nonsingularity of A_0 can be relaxed. In particular, if $m > h$ and $jh = mr$, we may repartition the matrix $A^{(j)}$ in (8) into $m \times m$ blocks $\widetilde{A}_i$, $i = 0, \ldots, r$, where $\widetilde{A}_0$ is singular. 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 $h \times h$ submatrix in the lower leftmost corner that coincides with A_0. Therefore the system

$$\left[\begin{array}{cccc} \widetilde{X}_1^{(j)}, & \widetilde{X}_2^{(j)}, & \ldots, & \widetilde{X}_r^{(j)} \end{array} \right] A^{(j)} = \left[\begin{array}{cccc} -\widetilde{A}_0, & 0, & \ldots, & 0 \end{array} \right] \tag{11}$$

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

Corollary 3 *Let us assume the hypotheses of Theorem 1 and consider the $jh \times jh$ 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 (11) we have $(\widetilde{X}_1^{(j)} \mathbf{e}_1)_{m-h+i} = -u_{i-1}/u_h + O(\sigma^{jh})$, $i = 1, \ldots, h$.*

We may look at the above corollary as to Theorem 1 extended to a power series having $m - h$ null zeros. The extension of Theorem 1 obtained in this way is extremely interesting also from the computational point of view since it allows us to choose the size of the blocks to be a power of 2 and to apply FFT-based algorithms.

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 [19] the function $t(z) = r_h(z)/f(z)$ can be analytically continued into ξ_i, $1 \le i \le h$. Thus, we can write $t(z) \in \mathcal{G}(\rho)$ for any ρ such that $0 < \rho < |\xi_{h+1}|$. If we denote $q(z)$ the polynomial of degree at most $h-1$ that interpolates $t(z)$ at ξ_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

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

In this way, we obtain that the decomposition

$$\frac{1}{f(z)} = \frac{q(z)}{r_h(z)} + s(z), \quad s(z) \in \mathcal{G}(\rho),\ 0 < \rho < |\xi_{h+1}|, \tag{12}$$

holds whenever $|z| \le \rho$ and, moreover, $z \ne \xi_i$, $1 \le i \le h$.

Now, observe that the formal power series expansion $1/f(z) = \sum_{i=0}^{+\infty} b_i z^i$, 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 $|z| < |\xi_1|$. Hence, we conclude that

$$b_i = k_i + e_i, \quad i \ge 0, \tag{13}$$

where

$$\frac{q(z)}{r_h(z)} = \sum_{i=0}^{+\infty} k_i z^i, \tag{14}$$

and, moreover,

$$s(z) = \sum_{i=0}^{+\infty} e_i z^i \tag{15}$$

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

Now we rephrase equations (12–15) 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 $i \ne j+1$, 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, $1 \le i \le h$, 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 $B_0, B_1, \ldots$, where B_i are Toeplitz matrices.

Since B is made up by the first h columns of T^{-1}, in the light of (7) it follows that

$$\begin{bmatrix} A_0 & & & 0 \\ A_1 & A_0 & & \\ A_2 & A_1 & A_0 & \\ \vdots & \ddots & \ddots & \ddots \end{bmatrix} \begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} I \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \tag{16}$$

where I is the identity matrix of order h.

Equation (13) 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 J_i the reversion matrix of size ih. If A is a Toeplitz matrix of order h, then JA and AJ are Hankel matrices, that is, their entries are constant along the anti-diagonals; 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 (14) we see that KJ is the 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} [18]. 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 (14) by $r_h(z)$ and comparing coefficients in z^h, z^{h+1}, ..., we obtain that $K_1 J = KJF$, where F is the Frobenius matrix (10) associated with the monic polynomial $z^h r_h(z^{-1})$, i.e., $r_h(z) = \sum_{i=0}^h u_i z^i$, $u_0 = 1$. Inductively, we find that $K_i = KJF^{h(i-1)+1}J$, $\; i \geq 1$, which leads to the following characterization of the matrices B_i [20].

Lemma 4 *For the blocks defined in (16) it holds*

$$B_i = (K + E_i JF^{-h(i-1)-1}J)JF^{h(i-1)+1}J. \tag{17}$$

Moreover, for any σ such that $\frac{|\xi_h|}{|\xi_{h+1}|} < \sigma < 1$ there exists a constant β such that

$$\| E_i JF^{-h(i-1)-1}J \|_\infty \leq \beta\sigma^{ih}. \tag{18}$$

Proof. Equation (17) holds from $B_i = K_i + E_i$ since $K_i = KJF^{h(i-1)+1}J$. Equation (18) holds from the Cauchy estimate (5) since the series (15) is convergent in the disk of radius $|\xi_{h+1}|$ (compare (12)) and since $\rho(F^{-h(i-1)}) = |\xi_h|^{h(i-1)}$. ■

Based on this result, Householder [20] and Stewart [31] developed numerical methods to approximate 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 (17) (18) imply that $(B_{j-1}B_j^{-1})^T$ tends to F^{-h}, since, as we will see later on, $KF^{-h}K^{-1} = J(F^h)^T J$. In this way, the first column of $JB_j^{-T}B_{j-1}^T J$ 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 $JB_j^{-T}B_{j-i}^T J$ coincides with $X_i^{(j)}$ in (8). Finally we prove the bound (9) by using Lemma 4 in order to estimate the convergence of $X_1^{(j)}$ to $JF^{-h}J$.

From (16) it follows that

$$\begin{bmatrix} A_1 & A_0 & & 0 \\ A_2 & A_1 & \ddots & \\ \vdots & \ddots & \ddots & A_0 \\ A_j & \dots & A_2 & A_1 \end{bmatrix} \begin{bmatrix} B_0 \\ B_1 \\ \vdots \\ B_{j-1} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -A_0 B_j \end{bmatrix}, \tag{19}$$

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 (19) with the nonsingularity of B_j. It provides an extension of the Wronski formula (see Theorem 1.3 of [19]) for the inversion of formal power series with coefficients over a field.

Lemma 5 *The coefficient matrix $A^{(j)}$ of the linear system (19) 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 submatrix T_{j+1} of T,

$$T_{j+1} = \begin{bmatrix} A_0 & & & 0 \\ A_1 & A_0 & & \\ \vdots & \ddots & \ddots & \\ A_j & \dots & A_1 & A_0 \end{bmatrix} = \begin{bmatrix} D^{(j)} & 0 \\ A^{(j)} & C^{(j)} \end{bmatrix}. \tag{20}$$

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 (20) we have

$$T_{j+1} \begin{bmatrix} W_0^{(j)} \\ \vdots \\ W_{j-1}^{(j)} \\ 0 \end{bmatrix} = \begin{bmatrix} A_0 W_0^{(j)} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

where $A_0 W_0^{(j)}$ is a non-zero matrix. This implies that the block vector

$$\begin{bmatrix} B_0 A_0 W_0^{(j)} - W_0^{(j)} \\ \vdots \\ B_{j-1} A_0 W_0^{(j)} - W_{j-1}^{(j)} \\ B_j A_0 W_0^{(j)} \end{bmatrix}$$

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

Conversely, assume that $A^{(j)}$ is nonsingular. In view of (20), if we consider the complementary persymmetric partitioning of T_{j+1}^{-1} we find

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

This shows that the Schur complement of $A^{(j)}$ in T_{j+1} 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 5 with the characterization of the B_i provided by (17) we see that $A^{(j)}$ is nonsingular for all sufficiently large indices j.

To prove equation (8) of Theorem 1, we provide a more suitable representation of the system (19) by reverting the rows and the columns of the matrix in (19). This operation corresponds 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_1 on its anti-diagonal, then we have

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

that for j large enough so that B_j is nonsingular, is reduced to (8) with

$$X_i^{(j)} = J B_j^{-T} B_{j-i}^T J, \quad i = 1, \ldots, j. \tag{21}$$

In order to obtain equation (9) of Theorem 1 we apply Lemma 4. In fact, from (17) and (21), we deduce that

$$X_1^{(j)} = J(K + E_j J F^{-h(j-1)-1} J)^{-T} J (F^{-h})^T J (K + E_{j-1} J F^{-h(j-2)-1} J)^T J. \tag{22}$$

We recall the following [16]:

Lemma 6 *If A and Δ are $n \times n$ matrices such that A is nonsingular and $\| A^{-1}\Delta \|_\infty < 1$, then $A + \Delta$ is nonsingular and*

$$\| (A+\Delta)^{-1} - A^{-1} \|_\infty \leq \| \Delta \|_\infty \| A^{-1} \|_\infty^2 / (1 - \| A^{-1}\Delta \|_\infty).$$

Applying this result to (22), since KJ is symmetric, we find that

$$X_1^{(j)} = K^{-1}(F^{-h})^T K + \Delta^{(j)}, \quad \| \Delta^{(j)} \|_\infty \leq \vartheta \sigma^{jh} \tag{23}$$

holds for sufficiently large j satisfying $\beta \parallel K^{-1} \parallel_\infty \sigma^{jh} < 1$, where β is the constant in (18).

Since KJ is the Hankel matrix associated with $z^h r_h(z^{-1})$, it is well known that [29] $F^T KJ = KJF$. Using this commuting property, we obtain $K^{-1}(F^{-h})^T K = JF^{-h}J$, that, together with (23), implies $X_1^{(j)} = JF^{-h}J + \Delta^{(j)}$, which proves equation (9) 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

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

and, moreover,

$$F^{-h}\mathbf{e}_h = F^{-1}\mathbf{e}_1,$$

where, throughout this paper, $\mathbf{e}_j$ denotes the j-th column of the identity matrix of appropriate size. Thus, the first column of $JF^{-h}J$ provides the coefficients $-u_i/u_h$, $i = 0, \dots, h-1$.

By recursively computing the first entry of the solutions of (8) 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.

2.2 Proof of Theorem 2

Without loss of generality we may assume that $|\xi_h| < 1 \leq |\xi_{h+1}|$.

In fact, if this condition is not satisfied, one can consider the solution of the modified matrix equation $\sum_{i=0}^{+\infty} X^i \widehat{A}_i$, where the blocks $\widehat{A}_i$ are generated by the power series expansion of the function $g(z) = f(\lambda z)$ and λ is such that $|\xi_h| < \lambda \leq |\xi_{h+1}|$. Then, the statement for $f(z)$ easily follows by observing that

$$\widehat{A}_i = \lambda^{ih} D A_i D^{-1},$$

and, therefore,

$$\widehat{X}_i^{(j)} = \lambda^{-hi} D X_i^{(j)} D^{-1},$$

where D denotes the scaling diagonal matrix with diagonal entries equal to λ^{i-1}, $i = 1, \dots, h$.

Since $R > 1$, then the Cauchy estimate (5) for the coefficients of the series expansion of $f(z)$ implies that $\sum_{i=0}^{\infty} ||A_i||_\infty = M < +\infty$. Moreover, from Lemma 4 and (21) one finds that, for sufficiently large values of j and for any σ, with $|\xi_h| < \sigma < 1$ there exists $\theta > 0$ such that

$$X_i^{(j)} = J(F^{-hi})^T J + \Delta_i^{(j)}, \quad \| \Delta_i^{(j)} \|_\infty \leq \theta \sigma^{jh}.$$

Furthermore, since the spectral radius of $\widehat{F} = JF^{-1}J$ is $|\xi_h|$, we find that for any $\epsilon > 0$ there exists a norm $|| \cdot ||$ such that $||\widehat{F}|| \leq |\xi_h| + \epsilon$. This implies that, for any σ, $|\xi_h| < \sigma < 1$, there exists a constant α such that $||\widehat{F}^{ih}||_\infty \leq \alpha \sigma^{ih}$.

By combining all these facts, since $A_0+\sum_{i=1}^{j}X_i^{(j)}A_i=0$ we find that

$$\begin{array}{c}\|\sum_{i=0}^{\infty}\widehat{F}^{ih}A_i\|_\infty=\|\sum_{i=1}^{\infty}\widehat{F}^{ih}A_i-\sum_{i=1}^{j}X_i^{(j)}A_i\|_\infty\le\\ \le\|\sum_{i=j+1}^{\infty}\widehat{F}^{ih}A_i\|_\infty+\|\sum_{i=1}^{j}\widehat{F}^{ih}A_i-\sum_{i=1}^{j}X_i^{(j)}A_i\|_\infty\le\\ \le M(\alpha\sigma^{(j+1)h}+\theta\sigma^{jh}),\end{array}$$

holds for j sufficiently large. This completes the proof since $\sigma<1$ and j is arbitrarily large.

3 Solving Toeplitz systems in generalized 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 (8), for $j=2^k$. For a suitable sequence of permutation matrices $P^{(n)}$, this procedure will compute block LU factorizations of the matrices $P^{(n)}A^{(2^n)}P^{(n)T}$ for $n=0,1,2,\ldots,k$, provided they exist, and at the same time computes also the matrix $X_1^{(2^n)}$ of (8) 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. These assumptions are automatically satisfied if the matrix A is diagonally dominant or if it is positive definite like in the solution of Markov chains or in the spectral factorization problems; for the general case, the nonsigularity assumptions will be removed in the next section.

The algorithm that we show can be adjusted to solving the matrix equation of Theorem 2.

Now, for a given integer $k>0$ set $\ell_0=2^k$ and consider the following system

$$\begin{gathered}\left[\begin{array}{cccc}X_1, & X_2, & \ldots, & X_{\ell_0}\end{array}\right]H=\left[\begin{array}{cccc}C, & 0, & \ldots, & 0\end{array}\right]\\ H=H^{(0)}=\left[\begin{array}{ccccc}\widehat{H}_1 & H_0 & & & 0\\ \widehat{H}_2 & H_1 & H_0 & & \\ \vdots & \vdots & \ddots & \ddots & \\ \widehat{H}_{\ell_0-1} & H_{\ell_0-2} & \ldots & H_1 & H_0\\ \widehat{H}_{\ell_0} & H_{\ell_0-1} & \ldots & H_2 & H_1\end{array}\right],\end{gathered}\tag{24}$$

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 (8) for $\ell_0=j$, $\widehat{H}_i=H_i=A_i$, $C=-A_0$, $X_i=X_i^{(j)}$ and $m=h$, or to (11) for $m>h$ and $\ell_0=r$, $H_i=\widehat{H}_i=\widetilde{A}_i$, $C=-\widetilde{A}_0$, $X_i=\widetilde{X}_i^{(j)}$.

Here our goal is simply to compute the first block component X_1 of the solution of (24). Let us introduce the even-odd permutation $q_k:\{1,2,\ldots,\ell_0\}\to\{1,2,\ldots,\ell_0\}$ such that $q_k(2i)=i$, $q_k(2i-1)=2^{k-1}+i$, together with the associated block permutation matrix Q_k having null blocks except for the blocks in position $(q_k(i),i)$ which are the $m\times m$ identity matrix I. Consider the matrix

$$Q_kH^{(0)}Q_k^T=\left[\begin{array}{cc}T^{(0)} & W^{(0)}\\ Z^{(0)} & R^{(0)}\end{array}\right],\tag{25}$$

obtained by applying the even-odd permutation to the block rows and block columns of $H^{(0)}$. It holds

$$
\begin{gathered}
\left[\begin{array}{ll} X_+, & X_- \end{array}\right] Q_k H^{(0)} Q_k^T = \left[\begin{array}{lllllll} 0, & 0, & \ldots, & 0, & C, & 0, & \ldots, & 0 \end{array}\right] \\
X_+ = \left[\begin{array}{llll} X_2, & X_4, & \ldots, & X_{\ell_0} \end{array}\right], \quad X_- = \left[\begin{array}{llll} X_1, & X_3, & \ldots, & X_{\ell_0-1} \end{array}\right].
\end{gathered} \tag{26}
$$

We observe that $T^{(0)}$, $R^{(0)}$, $Z^{(0)}$ and $W^{(0)}$ are $(\ell_0/2)\times(\ell_0/2)$ block Toeplitz-like matrices, defined by

$$
W^{(0)} = \begin{bmatrix} \widehat{H}_2 & H_0 & & 0 \\ \widehat{H}_4 & H_2 & \ddots & \\ \vdots & \vdots & \ddots & H_0 \\ \widehat{H}_{\ell_0} & H_{\ell_0-2} & \ldots & H_2 \end{bmatrix}, \quad Z^{(0)} = \begin{bmatrix} H_0 & & & 0 \\ H_2 & H_0 & & \\ \vdots & \ddots & \ddots & \\ H_{\ell_0-2} & \ldots & H_2 & H_0 \end{bmatrix},
$$

$$
R^{(0)} = \begin{bmatrix} \widehat{H}_1 & & & & 0 \\ \widehat{H}_3 & H_1 & & & \\ \vdots & H_3 & \ddots & & \\ \widehat{H}_{\ell_0-3} & \vdots & \ddots & H_1 & \\ \widehat{H}_{\ell_0-1} & H_{\ell_0-3} & \ldots & H_3 & H_1 \end{bmatrix}, \quad T^{(0)} = \begin{bmatrix} H_1 & & & 0 \\ H_3 & H_1 & & \\ \vdots & \ddots & \ddots & \\ H_{\ell_0-1} & \ldots & H_3 & H_1 \end{bmatrix}.
$$

By applying one step of block Gaussian elimination to the 2×2 block matrix (25) we find that

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

where I_s denotes the $sm \times sm$ identity matrix. Moreover the system (26) turns into

$$
X_- \; H^{(1)} = \left[\begin{array}{llll} C, & 0, & \ldots, & 0 \end{array}\right] \tag{27}
$$

together with $X_+ = -X_- Z^{(0)}(T^{(0)})^{-1}$. Since our goal is to compute X_1 that is the first block component of X_-, we may ignore the latter equation and reduce the original problem (24) to the problem (27) of half the size. Moreover, it can be easily verified (compare [3]) that the matrix $H^{(1)}$ of (27) has exactly the same structure as the original matrix $H^{(0)}$ in (24), that 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 (27) thus obtaining the sequence of $\ell_n \times \ell_n$ matrices $H^{(n)}$, for $\ell_n = 2^{k-n}$ where

$$
H^{(n)} = \begin{bmatrix} \widehat{H}_1^{(n)} & H_0^{(n)} & & & 0 \\ \widehat{H}_2^{(n)} & H_1^{(n)} & H_0^{(n)} & & \\ \vdots & \vdots & \ddots & \ddots & \\ \widehat{H}_{\ell_n-1}^{(n)} & H_{\ell_n-2}^{(n)} & \ldots & H_1^{(n)} & H_0^{(n)} \\ \widehat{H}_{\ell_n}^{(n)} & H_{\ell_n-1}^{(n)} & \ldots & H_2^{(n)} & H_1^{(n)} \end{bmatrix}. \tag{28}
$$

and

$$H^{(n)} = R^{(n-1)} - Z^{(n-1)}(T^{(n-1)})^{-1}W^{(n-1)}$$

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

Similarly, the sequence of systems

$$X^{(n)}H^{(n)} = \begin{bmatrix} C, & 0, & \ldots, & 0 \end{bmatrix}$$

$$X^{(n)} = \begin{bmatrix} X_1, & X_{1+2^n}, & X_{1+2\cdot 2^n}, & \ldots, & X_{1+(2^{k-n}-1)\cdot 2^n} \end{bmatrix}$$

is generated, for $n = 0, 1, \ldots, k$ 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 7 Observe that if we denote with $P^{(n)}$ the $\ell_0 \times \ell_0$ block cumulative permutation matrix such that

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

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 (28).

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

$$H_\infty^{(0)} = \begin{bmatrix} \widehat{H}_1 & H_0 & & & 0 \\ \widehat{H}_2 & H_1 & H_0 & & \\ \widehat{H}_3 & H_2 & H_1 & H_0 & \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{29}$$

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

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

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

One step of cyclic reduction applied to the infinite matrix (29) 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 [4]) that $H_\infty^{(1)}$ has the structure (29), 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

$$H_\infty^{(n)} = \begin{bmatrix} \widehat{H}_1^{(n)} & H_0^{(n)} & & & 0 \\ \widehat{H}_2^{(n)} & H_1^{(n)} & H_0^{(n)} & & \\ \widehat{H}_3^{(n)} & H_2^{(n)} & H_1^{(n)} & H_0^{(n)} & \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}. \tag{30}$$

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:

Remark 8 The $\ell_n \times \ell_n$ block matrix (28) obtained at the n-th step of cyclic reduction applied to the $2^k \times 2^k$ block matrix (24), coincides with the $2^{k-n} \times 2^{k-n}$ block leading principal submatrix of the infinite matrix (30) obtained at the n-th step of cyclic reduction, applied to the infinite matrix of (29). In particular, the block entry $\widehat{H}_1^{(n)}$ in position $(1,1)$ of (30) coincides with the one generated at the last step of the cyclic reduction algorithm applied to its leading principal submatrix of block size 2^n. Hence, $\widehat{H}_1^{(n)}$ is nonsingular if and only if $A^{(2^n)}$ is nonsingular.

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

Theorem 9 *Assume that the cyclic reduction process applied to $H^{(0)} = A^{(\ell)}$, where $A^{(\ell_0)}$ is the coefficient matrix of (8) and $\ell_0 = 2^k$, 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 (8) it holds*

$$X_1^{(2^n)} = -A_0(\widehat{H}_1^{(n)})^{-1}, \ n = 0, \dots, k.$$

Since we double the size of the linear system (8) 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 (9)). In practice we apply a finite number of steps of cyclic reduction to the infinite matrix (29). This provides the desired sequence $X_1^{(2^n)}$.

An interesting issue is the analysis of the condition number of the matrices $H_1^{(n)}$. It is easy to verify that these matrices are principal submatrices of certain Schur complements in the initial matrix $H^{(0)}$. Indeed, if $H^{(0)}$ is diagonally dominant then also all the Schur complements in $H^{(0)}$ are diagonally dominant [16] and the condition numbers of the Schur complements are not worse than the condition number of the initial matrix. This implies that, under strong diagonal dominance of $H^{(0)}$, the condition number of the matrices that are inverted at each step of the algorithm is bounded from above by a constant. A similar favorable situation holds for positive definite matrices, where, due to the Cauchy interlace theorem, the condition number of the Schur complements is bounded from above by the condition number of the initial matrix $H^{(0)}$.

In section 4 we will discuss the applicability of the cyclic reduction in the presence of breakdowns where $T_\infty^{(n)}$ is singular for certain values of n, $1 \leq n \leq k$;

in particular we will show that this situation occurs if and only if the leading principal submatrix of $H_\infty^{(0)}$ of block size $2^{n+1}-1$ is singular. In such singular cases, we suitably modify the triangularization scheme in order to compute the subsequence $\{X_1^{(2^s)}\}$, $1 \le s \le k$, corresponding to the values of s for which the leading principal submatrices of $H_\infty^{(0)}$ of order 2^s-1 and 2^s are nonsingular.

3.1 Functional formulation and computational issues

Let us associate with the infinite matrix $H_\infty^{(n)}$, obtained at the n-th step of cyclic reduction applied to the matrix (29), 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 $H_\infty^{(n)}$. Moreover, given a vector $\mathbf{u}=(u_i)$ of m components, denote with

$$L(\mathbf{u})=\begin{bmatrix} u_1 & & & 0 \\ u_2 & u_1 & & \\ \vdots & \ddots & \ddots & \\ u_m & \dots & u_2 & u_1 \end{bmatrix}$$

the $m \times m$ lower triangular Toeplitz matrix defined by its first column $\mathbf{u}$ and set $Z=L([0,1,0,\dots,0]^T)$.

Then, the first and the second block column of the matrix $H_\infty^{(n+1)}$ for $n \ge 0$ are determined by the formal matrix power series [3]

$$\begin{cases} \widehat{H}^{(n+1)}(z)=\widehat{H}_{even}^{(n)}(z)-H_{even}^{(n)}(z)\left(H_{odd}^{(n)}(z)\right)^{-1}\widehat{H}_{odd}^{(n)}(z) \\ H^{(n+1)}(z)=zH_{odd}^{(n)}(z)-H_{even}^{(n)}(z)\left(H_{odd}^{(n)}(z)\right)^{-1}H_{even}^{(n)}(z) \end{cases} \tag{31}$$

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

$$[R(z)]_{even}=R_{even}(z)=\textstyle\sum_{i=0}^{+\infty}R_{2i}z^i,$$

$$[R(z)]_{odd}=R_{odd}(z)=\textstyle\sum_{i=0}^{+\infty}R_{2i+1}z^i.$$

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.

Due to the Toeplitz structure of the blocks A_i, it is possible to provide a functional formulation of cyclic reduction in terms of triangular Toeplitz matrix power series [4].

Theorem 10 *For the matrix power series $H^{(n)}(z)$, $\widehat{H}^{(n)}(z)$ generated at the n-th step of cyclic reduction we have*

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

where J is the permutation matrix having 1 on the anti-diagonal, $\mathbf{v}^{(n)}(z) = \sum_{i=0}^{+\infty} \mathbf{v}_i^{(n)} z^i$ *is a suitable m-dimensional vector power series,* $\mathbf{c}^{(n)}(z) = H^{(n)}(z)\mathbf{e}_1$ *is the first column of $H^{(n)}(z)$ and* $\widehat{\mathbf{r}}^{(n)T}(z) = \mathbf{e}_m^T \widehat{H}^{(n)}(z)$, $\widehat{\mathbf{c}}^{(n)}(z) = \widehat{H}^{(n)}(z)\mathbf{e}_1$.

From the above result it follows that, at each step n, only the four vector power series $\mathbf{v}^{(n)}(z)$, $\mathbf{c}^{(n)}(z)$, $\widehat{\mathbf{c}}^{(n)}(z)$, $\widehat{\mathbf{r}}^{(n)}(z)$, are sufficient to represent $H^{(n)}(z)$ and $\widehat{H}^{(n)}(z)$. We call this set of vector power series the *generators* of the matrix power series $H^{(n)}(z)$ and $\widehat{H}^{(n)}(z)$.

Explicit relations between the generators at two subsequent steps of cyclic reduction are given in [4] and reported below.

Theorem 11 *For the generators* $\mathbf{c}^{(n)}(z)$, $\widehat{\mathbf{c}}^{(n)}(z)$, $\widehat{\mathbf{r}}^{(n)}(z)$, $\mathbf{v}^{(n)}(z)$, *we have*

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

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

In the block tridiagonal case, where the power series $f(z)$ reduces to a polynomial, i.e., $H_i = 0$ for $i > 2$, the above results are further simplified. We can give explicit expressions of the blocks $\widehat{H}_1^{(n)}$, $H_i^{(n)}$, $i = 0, 1, 2$. In fact the matrices generated by cyclic reduction are still block tridiagonal and, except for the north western corner block $\widehat{H}_1^{(n)}$, 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:

$$
\begin{aligned}
H_0^{(n+1)} &= -H_0^{(n)}(H_1^{(n)})^{-1}H_0^{(n)}\\
H_1^{(n+1)} &= H_1^{(n)} - H_0^{(n)}(H_1^{(n)})^{-1}H_2^{(n)} - H_2^{(n)}(H_1^{(n)})^{-1}H_0^{(n)}\\
H_2^{(n+1)} &= -H_2^{(n)}(H_1^{(n)})^{-1}H_2^{(n)}\\
\widehat{H}_1^{(n+1)} &= \widehat{H}_1^{(n)} - H_0^{(n)}(H_1^{(n)})^{-1}H_2^{(n)}.
\end{aligned}
\tag{32}
$$

Moreover, we have the following result [2].

Theorem 12 *If $H_i = 0$ 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 satisfy the following equations*

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

$$
\begin{aligned}
H_1^{(n)^{-1}} = \;& L(H_1^{(n)^{-1}}\mathbf{e}_1) + \\
& L^T(ZH_1^{(n)^{-T}}J\mathbf{v}_0^{(n)}) + L(H_1^{(n)^{-1}}\mathbf{c}_0^{(n)})L^T(ZH_1^{(n)^{-T}}J\mathbf{v}_1^{(n)}) - \\
& L(H_1^{(n)^{-1}}\mathbf{v}_1^{(n)})L^T(ZH_1^{(n)^{-T}}J\mathbf{c}_0^{(n)T}), \\
\widehat{H}_1^{(n)} = \;& L(\widehat{\mathbf{c}}^{(n)}) + L^T(ZJ\mathbf{c}_0^{(0)T}) - L(\mathbf{c}_0^{(n)})L^T(ZJ\mathbf{v}_1^{(n)T}) + L(\mathbf{v}_0^{(n)})L^T(Z\widehat{\mathbf{r}}^{(n)T}),
\end{aligned}
$$

where the generators $\mathbf{c}_i^{(n)}$, $i = 0,1,2$, $\mathbf{v}_i^{(n)}$, $i = 0,1$, $\widehat{\mathbf{c}}^{(n)}$, $\widehat{\mathbf{r}}^{(n)}$ *and* $\mathbf{s}^{(n)}$ *satisfy the following recurrences*

$$
\begin{aligned}
&\mathbf{c}_0^{(n+1)} = -H_0^{(n)}H_1^{(n)^{-1}}\mathbf{c}_0^{(n)}, \\
&\mathbf{c}_1^{(n+1)} = \mathbf{c}_1^{(n)} - H_0^{(n)}H_1^{(n)^{-1}}\mathbf{c}_2^{(n)} - H_2^{(n)}H_1^{(n)^{-1}}\mathbf{c}_0^{(n)}, \\
&\mathbf{c}_2^{(n+1)} = -H_2^{(n)}H_1^{(n)^{-1}}\mathbf{c}_2^{(n)}, \\
&\mathbf{v}_0^{(n+1)} = \mathbf{v}_0^{(n)} - H_0^{(n)}H_1^{(n)^{-1}}\mathbf{v}_1^{(n)}, \\
&\mathbf{v}_1^{(n+1)} = -H_2^{(n)}H_1^{(n)^{-1}}\mathbf{v}_1^{(n)}, \\
&\widehat{\mathbf{c}}^{(n+1)} = \widehat{\mathbf{c}}^{(n)} - H_0^{(n)}H_1^{(n)^{-1}}\mathbf{c}_2^{(n)}, \\
&\widehat{\mathbf{r}}^{(n+1)T} = \widehat{\mathbf{r}}^{(n)T} - \mathbf{c}_0^{(n)T}JH_1^{(n)^{-1}}H_2^{(n)}, \\
&\mathbf{s}^{(n+1)} = \mathbf{s}^{(n)} - H_2^{(n)}H_1^{(n)^{-1}}\mathbf{c}_0^{(n)}, \; n \geq 1.
\end{aligned}
\tag{33}
$$

with $\mathbf{v}_0^{(0)} = \mathbf{0}$, $\mathbf{v}_1^{(0)} = \mathbf{e}_1$, $\mathbf{s}^{(0)} = \mathbf{c}_1^{(0)}$, $\mathbf{c}_i^{(0)} = H_i^{(0)}\mathbf{e}_1$, $i = 0,1,2$, $\widehat{\mathbf{c}}^{(0)} = \mathbf{c}^{(0)}$, $\widehat{\mathbf{r}}^{(0)} = J\mathbf{c}_1^{(0)}$.

Theorems 10, 11 and 12 have very important computational consequences. We recall that the product of an $m \times m$ (triangular) Toeplitz matrix and a vector can be computed by means of FFT, in $O(m \log m)$ arithmetic operations [6], and that the inverse of a Toeplitz-like matrix, i.e., a matrix of the kind $\sum_{i=1}^{h} L(\mathbf{u}_i)L^T(\mathbf{v}_i)$, is still Toeplitz-like and can be computed by means of superfast algorithms in $O(m \log^2 m)$ ops. We refer the reader to [32] for a recent effective implementation of a superfast algorithm. Alternatively, if the Toeplitz-like matrix is positive definite, its inverse can be computed by means of (preconditioned) conjugate gradient techniques where the cost per iteration is $O(m \log m)$ and the number of iterations is typically very low.

It is easy to verify that, if $H_i = 0$ for $i > 2$ (as in the case of a polynomial), the cyclic reduction step (32) can be performed by means of Theorem 12 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.

We display below an outline of the algorithm for polynomial factorization.

Algorithm *Polynomial factorization.*

INPUT. Integers h and k, the degree $N > h$ and the coefficients of the polynomial $p(z) = \sum_{i=0}^{N} a_i z^i = a_N \prod_{i=1}^{N}(z - \xi_i)$, such that $|\xi_i| \leq |\xi_j|$ for $i < j$ and $|\xi_h| < |\xi_{h+1}|$. Without loss of generality, assume $h \geq N/2$ so that $h = m = \max\{h, N-h\}$.

OUTPUT. The sequence $\mathbf{w}^{(n)} = X_1^{(2^n)}\mathbf{e}_1$ of approximations to the coefficients of the factor $\prod_{i=1}^{h}(z - \xi_i)$ provided by the first column of the matrix $X_1^{(2^n)}$ of Theorem 1.

Computation.

1. Consider the $h \times h$ matrices $A_s = (a_{i-j+sh})_{i,j}$, $s = 0, 1, 2$, where we assume $a_i = 0$ if $i < 0$ or $i > N$. Consider the system (24) with $H_0 = A_0$, $H_1 = \widehat{H}_1 = A_1$, $H_2 = A_2$, $C = -A_0$. Initialize the vectors $\mathbf{v}_0^{(0)} = \mathbf{0}$, $\mathbf{v}_1^{(0)} = \mathbf{e}_1$, $\mathbf{c}_i^{(0)} = A_i\mathbf{e}_1$, $i = 0, 1, 2$, $\hat{\mathbf{c}}^{(0)} = \mathbf{c}_1^{(0)}$, $\hat{\mathbf{r}} = J\mathbf{c}_1^{(0)}$, $\mathbf{s}^{(0)} = \mathbf{c}_1^{(0)}$ so that the equations of Theorem 12 hold for $n = 0$.
2. For $n = 1, 2, \ldots, k$ compute the vectors $\mathbf{c}_i^{(n)}$, $i = 0, 1, 2$, $\mathbf{v}_i^{(n)}$, $i = 0, 1$, $\mathbf{s}^{(n)}, \widehat{\mathbf{c}}^{(n)}$ and $\widehat{\mathbf{r}}^{(n)}$ by means of (33) where the matrices $H_i^{(n-1)}$, $i = 0, 1, 2$, $\widehat{H}_1^{(n-1)}$ and $H_1^{(n-1)^{-1}}$ are represented according to Theorem 12 in terms of lower/upper triangular Toeplitz matrices. Compute and output $\mathbf{w}^{(n)} = X_1^{(2^n)}\mathbf{e}_1 = -A_0\widehat{H}_1^{(n)^{-1}}\mathbf{e}_1$ (compare with Theorem 9).

Observe that the matrix-vector products at stage 2 of the above algorithm are reduced to multiplying lower/upper triangular Toeplitz matrices and vectors, in the light of the representation formulae provided by the Theorem 12. This computation can be performed by means of FFT's at the cost of $O(N \log N)$ ops. The inversion of the Toeplitz-like matrices $H_1^{(n)}$ that appears in Theorem 12, and the computation of the first column of $\widehat{H}_1^{(n)^{-1}}$, can be performed by means of any superfast algorithm for Toeplitz-like matrix inversion at the cost of $O(N \log^2 N)$ ops (see for instance [32]). Alternatively, with a more practical approach, the inversion can be performed by means of iterative techniques, like preconditioned conjugate gradient [11] or Newton's iteration [5], almost with the same asymptotical cost. Thus the overall cost of the above algorithm is $O(kN \log^2 N)$ ops.

Remark 13 If $f(z)$ is a polynomial of degree N then the banded matrix defined by the reversed polynomial $f(z^{-1})z^N$ is the transposed of A of (1). Moreover the blocks generated by the above algorithm applied to $f(z^{-1})z^N$ coincide with $H_0^{(n)T}$, $H_1^{(n)T}$, $H_2^{(n)T}$ and $\widehat{H}_1^{(n)T}$. From Theorem 9 applied to $f(z^{-1})z^N$ the first column of $-H_2^T(\widehat{H}_1^{(n)T})^{-1}$ provides approximations to the coefficients of the complementary factor of $f(z)$. Therefore, the above algorithm can be slightly modified in order to compute both factors of $f(z)$ within almost the same computational cost.

A slightly more complicated situation occurs in the general Hessenberg case where the Toeplitz-like matrices involved have coefficients in the algebra $\mathcal{F}$ of power series. In this case the basic computations to be performed are: computing the product $\mathbf{w}(z) = L(\mathbf{u}(z))\mathbf{v}(z)$ of a lower triangular matrix power series $L(\mathbf{u}(z))$ and a vector power series $\mathbf{v}(z)$, given vector power series $\mathbf{u}(z), \mathbf{v}(z) \in \mathcal{F}^m$; the solution of an $m \times m$ system of the kind $\Gamma(z)\mathbf{x}(z) = \mathbf{b}(z)$, where $\Gamma(z) = \sum_{i=1}^{r} L(\mathbf{u}^{(i)}(z))L^T(\mathbf{v}^{(i)}(z)) \in \mathcal{F}^{m\times m}$, given the vector power series $\mathbf{u}^{(i)}(z), \mathbf{v}^{(i)}(z), \mathbf{b}(z) \in \mathcal{F}^m$. Both the above computations can be reduced to customary computations with Toeplitz-like matrices in the complex field by means of the evaluation/interpolation technique at the roots of 1 (Fourier points). We describe this procedure in the case of the computation of $\mathbf{w}(z) = L(\mathbf{u}(z))\mathbf{v}(z)$; the computation of $\mathbf{x}(z)$ can be carried out similarly.

For simplicity assume that $\mathbf{u}(z)$, $\mathbf{v}(z)$ and $\mathbf{w}(z)$ are vector polynomials of degree less than $K = 2^k$. Compute the values of $\mathbf{u}(z)$, $\mathbf{v}(z)$ at the K-th roots of unity, $\omega_1, \ldots, \omega_K$, by means of $2m$ FFT's, compute the vectors $\mathbf{w}(\omega_i) = L(\mathbf{u}(\omega_i))\mathbf{v}(\omega_i)$, recover the vector coefficients of $\mathbf{w}(z)$ by interpolating to the values $\mathbf{w}(\omega_i)$ at the K-th roots of 1 by applying m inverse FFT's. The overall cost of this computation is K multiplications between lower triangular (scalar) Toeplitz matrices and vectors, plus $3m$ FFT's of length K, that is $O(Km\log m + mK\log K) = O((Km)\log(Km))$ ops.

For matrix power series the number K 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 $\Gamma(z) \in \mathcal{F}^{m\times m}$, $\Gamma(z) = \sum_{i=0}^{+\infty} \Gamma_i z^i$, $\Gamma_i \in \mathbf{C}^{m\times m}$, convergent in the closed unit disk, has ϵ-degree N if N is the minimum integer such that $\mathbf{e}^T \sum_{i=N+1}^{+\infty} |\Gamma_i| \le \epsilon \mathbf{e}^T$, where $\mathbf{e}$ denotes the $m-$dimensional vector having entries equal to 1. If ϵ is the machine precision of the floating point 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 $\epsilon \ge 0$, 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 at most N.

In this way, the following complexity estimates can be established [3], [4].

Theorem 14 *Assume we are given a vector $\mathbf{u}(z) \in \mathcal{F}^m$ and an $m \times m$ Toeplitz-like matrix power series $\Gamma(z) \in \mathcal{F}^{m\times m}$ convergent in the closed unit disk. Then the first N coefficients of the product $\mathbf{w}(z) = \Gamma(z)\mathbf{u}(z)$ 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 $\mathbf{w}(z)$. Moreover, if also $\Gamma(z)^{-1}$ 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 (31) applied to the Hessenberg case (i.e., for factoring a power series) and implemented by means of Theorems 10 and 11, 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.

For wide classes of analytic functions $f(z)$ it is possible to estimate the numerical degree by means of simple criteria that are dynamically applied in the computation of the cyclic reduction step [3].

4 Applicability of cyclic reduction

As already observed in the previous section, the polynomial factorization algorithm is numerically reliable if applied to positive definite or diagonally dominant matrices. In the general case, breakdown situations or instability problems may interrupt the algorithm or spoil the accuracy of the computed output.

In this section we investigate the conditions which ensure that the cyclic reduction technique applied to the linear system (8) 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, stated in the previous section, 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.

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 by Remark 7. The possibility of reaching the same Schur complement 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.

Necessary and sufficient conditions for the applicability of cyclic reduction are given by the following results:

Theorem 15 *Assume that the first $n-1$ steps of the cyclic reduction algorithm applied to $A^{(2^k)}$, $n < k$, can be performed. Then, the n-th step can be performed if and only if the block leading principal submatrix of order $2^n - 1$ of $A^{(2^k)}$ is nonsingular.*

Proof. The Schur complement obtained by applying $n-1$ steps of cyclic reduction with the odd-even permutation to the matrix $A^{(2^n-1)}$ coincides with $H_1^{(n-1)}$. Therefore $A^{(2^n-1)}$ is nonsingular if and only if $H_1^{(n-1)}$ is nonsingular. ■

This immediately implies the following:

Corollary 16 *The cyclic reduction algorithm applied to $A^{(2^k)}$ can be carried out without the occurrence of a breakdown at some intermediate step if and only if all the block leading principal submatrices of order $2^n - 1$ of $A^{(2^k)}$ are nonsingular for $1 \leq n \leq k$.*

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 submatrices results to be singular, then the next theorem can be used for avoiding the singularity.

Theorem 17 *Assume that A is a generic nonsingular matrix of order l and let P_1 and P_2 be two $l \times l$ permutation matrices with the same last h rows. Denote by T_1 and by T_2 the $(l-h) \times (l-h)$ leading principal submatrix of $P_1 A P_1^T$ and of $P_2 A P_2^T$, respectively. Then T_1 is nonsingular if and only if T_2 is nonsingular. Moreover, if T_1 is nonsingular, then $P_1 A P_1^T$ and $P_2 A P_2^T$ have the following block triangular factorization*

$$P_1 A P_1^T = \begin{bmatrix} I & 0 \\ L_1 & I \end{bmatrix} \begin{bmatrix} T_1 & W_1 \\ 0 & S_1 \end{bmatrix}, \quad P_2 A P_2^T = \begin{bmatrix} I & 0 \\ L_2 & I \end{bmatrix} \begin{bmatrix} T_2 & W_2 \\ 0 & S_2 \end{bmatrix},$$

where $S_1 = S_2$, I denotes an identity matrix of appropriate size and L_1, L_2 are $h \times (l-h)$ matrices.

Proof. The proof immediately follows from the observation that T_1 and T_2 are equal up to a suitable reordering of rows and columns and, moreover, both S_1^{-1} and S_2^{-1} coincide with the $h \times h$ trailing principal submatrix of the inverse of $P_1 A P_1^T$. ■

The above result allows to perform a triangularization processes also in the case of breakdowns, due to the singularity of certain leading principal submatrices of $A^{(2^k)}$. Indeed, based on the above theorem, if Π is a permutation matrix whose last 2^{k-n} block rows coincide with the last 2^{k-n} block rows of the cumulative permutation matrix $P^{(n)}$ of Remark 7, then the Schur complement $H^{(n)}$ 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 $\Pi H^{(0)}\Pi^T$. In this way it is possible to compute $H^{(n)}$ even though one of the intermediate diagonal blocks $H_1^{(i)}$, $i = 0, \ldots, n-1$, is singular.

For the sake of clarity, we introduce a computational example in order to illustrate our analysis. A formal derivation of the results given below can be found in [1]. We assume $k = 3$ and $H^{(0)}$ block tridiagonal. Suppose that $H_1^{(0)}$ is singular and, therefore, the first step of cyclic reduction applied to $H^{(0)}$ cannot be carried out. Observe that the last 2 block rows of the matrix Π, defined by

$$\Pi = \begin{bmatrix} 0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & I \\ I & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I & 0 & 0 & 0 \end{bmatrix},$$

coincide with the ones of the cumulative permutation matrix $P^{(1)}$ introduced in Remark 7. Also, note that

$$\Pi H^{(0)}\Pi^T = \left[\begin{array}{ccc|ccc|cc} H_1^{(0)} & H_0^{(0)} & 0 & 0 & 0 & 0 & H_2^{(0)} & 0 \\ H_2^{(0)} & H_1^{(0)} & H_0^{(0)} & 0 & 0 & 0 & 0 & 0 \\ 0 & H_2^{(0)} & H_1^{(0)} & 0 & 0 & 0 & 0 & H_0^{(0)} \\ \hline 0 & 0 & 0 & H_1^{(0)} & H_0^{(0)} & 0 & 0 & H_2^{(0)} \\ 0 & 0 & 0 & H_2^{(0)} & H_1^{(0)} & H_0^{(0)} & 0 & 0 \\ 0 & 0 & 0 & 0 & H_2^{(0)} & H_1^{(0)} & 0 & 0 \\ \hline H_0^{(0)} & 0 & 0 & 0 & 0 & 0 & \widehat{H}_1^{(0)} & 0 \\ 0 & 0 & H_2^{(0)} & H_0^{(0)} & 0 & 0 & 0 & H_1^{(0)} \end{array}\right]. \quad (34)$$

If $A^{(8)}(\epsilon) = H^{(0)}(\epsilon) = H^{(0)} + \epsilon I_8$ 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 $1 \leq j \leq 3$, then the cyclic reduction algorithm applied to $H^{(0)}(\epsilon)$ goes on to termination without breakdowns. In particular, $H^{(2)}(\epsilon)$ is well-defined and, when ϵ converges to zero, it will tend to the inverse of the trailing principal submatrix of order 2 of $(P^{(1)}A^{(8)}P^{(1)T})^{-1}$, where $P^{(1)}$ is the cumulative permutation matrix defined in Remark 7. In view of Theorem 17, the same Schur complement can be obtained by one step of block Gaussian elimination applied to the matrix $\Pi A^{(8)}\Pi^T$. Specifically, assume that the 3×3 block tridiagonal matrix

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

is invertible with

$$(K^{(0)})^{-1} = \begin{bmatrix} K_{11}^{(0)} & K_{12}^{(0)} & K_{13}^{(0)} \\ K_{21}^{(0)} & K_{22}^{(0)} & K_{23}^{(0)} \\ K_{31}^{(0)} & K_{32}^{(0)} & K_{33}^{(0)} \end{bmatrix}.$$

By taking the Schur complement of the 2×2 block leading principal submatrix in (34), we may define

$$\begin{aligned} H_0^{(2)} &= -H_0^{(0)} K_{13}^{(0)} H_0^{(0)} \\ H_1^{(2)} &= H_1^{(0)} - H_0^{(0)} K_{11}^{(0)} H_2^{(0)} - H_2^{(0)} K_{33}^{(0)} H_0^{(0)} \\ H_2^{(2)} &= -H_2^{(0)} K_{31}^{(0)} H_2^{(0)} \\ \widehat{H}_1^{(2)} &= \widehat{H}_1^{(0)} - H_0^{(0)} K_{11}^{(0)} H_2^{(0)}. \end{aligned}$$

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 9 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 submatrices can also be used in a numerical setting in order to avoid manipulations involving ill-conditioned submatrices.

5 Numerical experiments

We implemented in Fortran 90 the algorithm of section 3 for the computation of the factors $r_h(z)$ and $s(z) = f(z)/r_h(z)$ of the polynomial $f(z)$ of degree N. Here $r_h(z)$ has degree h and has zeros of modulus less than 1. The program relies on FFT for performing multiplications of Toeplitz matrices and vectors. The computation of the inverses of all the Toeplitz-like matrices involved is obtained by means of the preconditioned conjugate gradient method (PCG). As preconditioner we chose the circulant matrix that better approximates, in the sense of the Frobenius norm, the Toeplitz matrix defined by the first row and the first column of the matrix to be inverted.

As test suite we considered the following polynomials and power series (where the power series are truncated to polynomials according to their numerical degree).

1. Spectral factorization:
   - (a) $f(z) = p(z)p(z^{-1})z^h$, $p(z) = \sum_{i=0}^{h}(i+1)z^i$, $N = 2h$ [33]. This problem for $h = 512$ has been introduced in [24]. The matrix A is not diagonally dominant.
   - (b) $f(z) = \sum_{i=0}^{N} z^i + (2h-1)z^h$, $N = 2h$, where A is diagonal dominant.
   - (c) $f(z) = \sum_{i=0}^{N} z^i + (0.5h-1)z^h$, $N = 2h$, where A is not diagonal dominant.

2. Polynomial factorization:

   (a) $f(z) = \sum_{i=0}^{2h}(i+1)z^i + h^2 z^h$, where A is diagonally dominant.

   (b) $f(z) = \sum_{i=0}^{2h}(i+1)z^i + 50hz^h$, where A is not diagonally dominant.

   (c) $f(z) = (\sum_{i=0}^{h}(i+1)z^i)(\sum_{i=0}^{h}(2(h-i)+1)z^i)$ where A is not diagonally dominant.

3. Power series:

   (a) $f(z) = z^h - e^{-\lambda}\sum_{i=0}^{+\infty}\frac{\lambda^i}{i!}z^i, \quad 0 < \lambda < h$. This series is encountered in the numerical solution of certain Markov chains. Critical situations occur if $\lambda \approx h$. In fact, for $\lambda = h$ the Markov chain is not positive recurrent. We tested the cases $\lambda = h - 1/h$, $\lambda = h - 1/h^2$.

The experiments have been performed on a pentium 450 workstation with the Linux system, by using the standard IEEE 53-bit floating point arithmetic, where all the floating point variables in the program have been declared as double.

For certain computations we have applied one step of iterative refinement, based on Newton's iteration, for improving the approximation delivered by the cyclic reduction algorithm.

Tables 1, 2, 3, 4, 5, 6 and 7 report the degree N, the number of iterations of cyclic reduction and the maximum number of iterations of the PCG for the inversion of Toeplitz like matrices, the overall cpu time needed for the computation (excluding the refinement), the maximum residual error, the maximum relative approximation error of the coefficients (only for problems 1a and 2c where the output is known).

Observe that the number of steps of cyclic reduction is very small for all the tested polynomials (power series) and does not exceed 6 (22). For diagonally dominant matrices the number of preconditioned conjugate gradient steps needed for inverting the Toeplitz-like matrices is independent of N (see for instance table 2 where it does not exceed 12 even for very large values of N, or table 3 where it does not exceed 35). Moreover, we did not encounter numerical instability problems even for very large values of N.

In the positive definite case (problems of spectral factorization) our algorithm still has a good numerical behavior where we reach acceptable residual values even for large values of the degree (table 1) and where a single step of Newton's refinement is enough to get a residual of the order of the machine precision.

More critical situations are encountered for polynomials that do not correspond to symmetric or to diagonal dominant matrices, like in the case of table 5. In fact, in this case the residual error seems to have a faster growth as a function of the degree N of $f(z)$. However, in this situation our method still provides good approximations even for moderately large values of the degree N and accurate initial guesses to be refined in one step of Newton's iteration, even for large values of N. We observe also that the number of iteration of PCG is not constant and seems to grow like $O(\log N)$.

The estimated complexity bound that we may deduce from the numerical experiments (cpu time) grows as $O(N \log N)$ for diagonally dominant matrices and as $O(N \log^2 N)$ for general matrices.

	cyclic reduction				refinement	
N	cpu	iter	res	err	res	err
64	0.1	5/12	5.9e-14	6.9e-14	1.0e-17	6.7e-15
128	0.2	5/16	9.1e-12	1.6e-11	8.4e-16	2.3e-14
256	0.4	5/17	7.7e-11	2.3e-10	1.1e-15	5.1e-14
512	0.9	5/20	2.0e-9	1.1e-9	2.6e-15	1.9e-13
1024	2.0	5/20	2.0e-9	1.1e-9	2.6e-15	1.9e-13
2048	4.3	5/28	1.2e-7	4.9e-8	8.9e-15	6.9e-12
4096	13.0	5/22	4.6e-8	4.2e-8	3.2e-13	8.9e-12
8192	36.8	5/34	1.4e-6	2.5e-6	1.3e-12	2.2e-11

Table 1: Polynomial $f(z) = p(z)p(z^{-1})z^h$, $p(z) = \sum_{i=0}^{h}(i+1)z^i$, $N = 2h$.

	$\alpha = 2h$			$\alpha = h/2$		
N	cpu	iter	res	cpu	iter	res
64	0.08	5/7	1.1e-16	0.15	7/10	1.1 e-15
128	0.17	5/7	1.1e-16	0.32	7/10	4.4e-16
256	0.3	5/7	1.1e-16	0.56	7/10	3.3e-16
512	0.7	5/7	5.5e-16	1.34	7/11	7.8e-16
1024	1.5	5/7	1.2e-15	3.02	7/11	7.8e-16
2048	3.2	5/7	1.2e-15	6.6	7/11	4.4e-16
4096	8.48	5/7	4.4e-16	17	7/11	8.9e-16
8192	19.2	5/7	1.5e-15	38.6	7/11	8.9e-16
16384	42.1	5/7	1.9e-15	85.6	7/10	1.5e-15
32768	99.3	5/7	2.4e-15	199.7	7/12	2.2e-16
75536	264.8	5/7	5.9e-15	531.4	7/11	3.3e-16

Table 2: Polynomial $f(z) = \sum_{i=0}^{2h} z^i + (\alpha - 1)z^h$, $N = 2h$, $\alpha = 2h$, $\alpha = h/2$.

Concerning power series, we tested the situation where $\lambda \approx h$. This case is critical since the Markov chains is not positive recurrent for $\lambda = h$, i.e., $|\xi_h| = |\xi_{h+1}|$. Methods based on the functional iteration approach have a very slow convergence and require a huge number of steps that make the methods practically unusable. On the contrary, our algorithm still maintains very good performances in terms of speed and precision, even though a slight slow down (due to the closeness to 1 of $|\xi_h|/|\xi_{h+1}|$) is evident.

6 Conclusions and further extensions

We have reduced the problem of factoring a polynomial $f(z)$ of degree N to that of computing at most the first $N/2$ 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 banded matrix with rows and columns suitably permuted.

N	cpu	iter	res
64	0.2	6/28	3.1e-16
128	0.5	6/31	4.4e-16
256	1.1	6/35	2.2e-16
512	2.2	6/34	4.4e-16
1024	4.7	6/34	3.5e-15
2048	10.4	6/35	6.6e-16
4096	27.4	6/35	3.3e-15
8192	64.2	6/35	6.7e-16

Table 3: Polynomial $p(z) = \sum_{i=0}^{2h}(i+1)z^i + h^2 z^h$, $N = 2h$.

	cyclic reduction			refinement
N	cpu	iter	res	res
64	0.16	4/13	1.1e-15	1.2e-17
128	0.37	4/14	4.2e-15	1.4e-16
256	1.0	5/21	9.0e-15	1.4e-16
512	2.5	5/22	3.7e-14	3.9e-17
1024	6.5	5/27	2.3e-11	1.4e-15
2048	21.3	6/43	5.9e-12	2.9e-15

Table 4: Polynomial $f(z) = \sum_{i=0}^{2h}(i+1)z^i + 50hz^h$, $N = 2h$.

	cyclic reduction				refinement	
N	cpu	iter	res	err	res	err
64	0.3	5/20	4.2e-12	3.4e-12	4.1e-16	9.6e-15
128	0.7	5/22	5.3e-11	7.1e-11	1.0e-15	3.1e-14
256	1.5	5/25	1.5e-9	1.9e-9	2.2e-15	7.8e-14
512	3.4	5/26	1.1e-8	1.5e-8	2.5e-13	2.3e-13
1024	8.2	5/30	8.2e-7	1.0e-6	3.0e-13	8.9e-13
2048	21.4	5/45	4.4e-5	5.9e-5	4.7e-10	6.2e-10

Table 5: $f(z) = (\sum_{i=0}^{h}(i+1)z^i)(\sum_{i=0}^{h}(2(h-i)+1)z^i)$, $N = 2h$.

h	N	block size	cpu	iter	res
32	90	64	2.9	8/66	4.0e-16
64	140	128	9.5	9/100	1.8e-15
128	231	128	8.7	11/75	5.9e-16
256	397	256	29	12/115	4.2e-16
512	708	512	103	13/187	7.1e-16

Table 6: Power series $f(z) = z^h - e^{-\lambda}\sum_{i=0}^{+\infty}\frac{\lambda^i}{i!}z^i$, $\lambda = h - 1/h$.

h	N	block size	cpu	iter	res
32	89	64	5.1	13/78	6.4e-16
64	140	128	17.6	15/126	3.3e-15
128	232	128	15.5	18/91	3.7e-16
256	397	256	54.8	20/169	5.8e-16
512	707	512	201	22/293	8.3e-16

Table 7: Power series $f(z) = z^h - e^{-\lambda} \sum_{i=0}^{+\infty} \frac{\lambda^i}{i!} z^i$, $\lambda = h - 1/h^2$.

If the banded matrix is diagonally dominant or positive definite, then well known results in numerical linear algebra guarantee the numerical stability of the LU factorization. This is confirmed by the numerical experiments.

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 $|a_h| > \sum_{i=0, i\neq h}^{N} |a_i|$, where a_h is the diagonal entry, by applying Rouché's theorem to the polynomials $\sum_{i=0}^{N} a_i z^i$ and $a_h z^h$, we deduce that h zeros are inside the unit disk and $N - h$ zeros are outside. It is even possible to express the isolation ratio of these two sets of zeros (i.e., the ratio γ_1/γ_2, where $\gamma_1 < 1 < \gamma_2$ and both the disks of center 0 and radius γ_1, γ_2 contain h zeros), in terms of the ratio $|a_h| / \sum_{i=0, i\neq h}^{N} |a_i|$ that measures the diagonal dominance of the matrix.

A possible improvement of the numerical behavior of our algorithm relies on using better displacement representations of the Toeplitz-like matrices generated by cyclic reduction. In [5] we introduced the orthogonal displacement representation of Toeplitz-like matrices, based on the svd, that aims to reduce the cancellation problems arising from the unbalancing of displacement representations. This tool may also be used together with the concept of the ϵ-displacement rank, for the stable inversion of Toeplitz-like matrices by means of Newton's iteration [5]. The availability of stable Toeplitz-like solvers might greatly improve the numerical stability of our factorization method. A rough implementation of our algorithm, where the inversion of the Toeplitz-like matrices is performed by means of the netlib routines, allowed us to solve successfully the critical spectral factorization problems cited in [17].

Acknowledgements. We wish to thank the anonymous referees for their suggestions that greatly improved the presentation of our paper.

References

[1] D. Bini, L. Gemignani, B. Meini, Factorization of Analytic Functions by means of Koenig's Theorem and Toeplitz Computations, Tech. Rep. n.98-08 , Computer Science Dep., University of Pisa, Pisa, 1998.

[2] D. Bini, B. Meini, *Effective Methods for Solving Banded Toeplitz Systems*, SIAM J. Matrix Anal. Appl., **20** (1999), pp. 700–719.

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

[4] 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.

[5] D. Bini, B. Meini, *Approximate Displacement Rank and Applications*, Proceedings AMS Conference on Structured Matrices, Boulder, 1999.

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

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] R. H. Chan, M. K. Ng, *Conjugate Gradient Methods for Toeplitz Systems*, SIAM Review, **38** (1996), pp. 427-482.

[12] H. R. Gail, S.L. Hantler, B. A. Taylor, *Non-Skip-Free M/G/1 and G/M/1 type Markov Chains*, Advances in Applied Probability **29** (1997), pp. 733-758.

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

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

[15] L. Gemignani, *Computing a Hurwitz Factorization of a Polynomial*, J. Comput. Appl. Math., to appear.

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

[17] T.N.T. Goodman, C.A. Micchelli, G. Rodriguez, S. Seatzu, *Spectral Factorization of Laurent Polynomials*, Adv. in Comp. Math. **7** (1997) pp. 429–454.

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

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

[20] 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.

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

[22] P. Kirrinnis, *Partial Fraction Decomposition in* $\mathbf{C}(z)$ *and Simultaneous Newton Iteration for Factorization in* $\mathbf{C}[z]$, J. of Complexity, **14** (1998), pp. 378–444.

[23] G. Latouche, *Newton's Iteration for Nonlinear Equations in Markov Chains*, IMA J. on Numerical Analysis, **14** (1994), pp. 583-598.

[24] D. P. Laurie, *Efficient Implementation of Wilson's Algorithm for Factorizing a Self-Reciprocal Polynomial*, BIT **20** (1980), 257-259.

[25] B. Meini, *New Convergence Results on Functional Iteration Techniques for the Numerical Solution of M/G/1 Type Markov Chains*, Numer. Math., **78** (1997), pp. 39-58.

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

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

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

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

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

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

[32] M. Van Barel, P. Kravanja, *A Stabilized Superfast Solver for Indefinite Hankel Systems*, Linear Algebra Appl. **284** (1998), 275-295.

[33] P. Van Dooren, *Some Numerical Challenges in Control Theory*, Linear Algebra for Control Theory, IMA Vol. Math. Appl. 62, Springer, NY, 1994.

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

[35] M. Vidyasagar, *Control System Synthesis: A Factorization Approach*, MIT Press, Cambridge MA, 1985.