ON A CLASS OF MATRIX PENCILS AND ℓ-IFICATIONS EQUIVALENT TO A GIVEN MATRIX POLYNOMIAL*

DARIO A. BINI † AND LEONARDO ROBOL ‡

Abstract. A new class of linearizations and ℓ-ifications for $m \times m$ matrix polynomials $P(x)$ of degree n is proposed. The ℓ-ifications in this class have the form $A(x) = D(x)+(e\otimes I_m)W(x)$ where D is a block diagonal matrix polynomial with blocks $B_i(x)$ of size m, W is an $m\times qm$ matrix polynomial and $e = (1, \dots , 1)^t \in \mathbb{C}^q$, for a suitable integer q. The blocks $B_i(x)$ can be chosen a priori, subjected to some restrictions. Under additional assumptions on the blocks $B_i(x)$ the matrix polynomial $A(x)$ is a strong ℓ-ification, i.e., the reversed polynomial of $A(x)$ defined by $A^{\#}(x) := x^{\deg A(x)}A(x^{-1})$ is an ℓ-ification of $P^{\#}(x)$. The eigenvectors of the matrix polynomials $P(x)$ and $A(x)$ are related by means of explicit formulas. Some practical examples of ℓ-ifications are provided. A strategy for choosing $B_i(x)$ in such a way that $A(x)$ is a well conditioned linearization of $P(x)$ is proposed. Some numerical experiments that validate the theoretical results are reported.

AMS classification: 65F15, 15A21, 15A03
Keywords: Matrix polynomials, matrix pencils, linearizations, companion matrix, tropical roots.

1. Introduction. A standard way to deal with an $m \times m$ matrix polynomial $P(x) = \sum_{i=0}^{n} P_i x^i$ is to convert it to a linear pencil, that is to a linear matrix polynomial of the form $L(x) = Ax - B$ where A and B are $mn \times mn$ matrices such that $\det P(x) = \det L(x)$. This process, known as *linearization*, has been considered in [?].

In certain cases, like for matrix polynomials modeling Non-Skip-Free stochastic processes [?], it is more convenient to reduce the matrix polynomial to a quadratic polynomial of the form $Ax^2 + Bx + C$, where A, B, C are matrices of suitable size [?]. The process that we obtain this way is referred to as *quadratization*. If $P(x)$ is a matrix power series, like in M/G/1 Markov chains [?, ?], the quadratization of $P(x)$ can be obtained with block coefficients of infinite size [?]. In this framework, the quadratic form is desirable since it is better suited for an effective solution of the stochastic model; in fact it corresponds to a QBD process for which there exist efficient solution algorithms [?], [?]. In other situations it is preferable to reduce the matrix polynomial $P(x)$ of degree n to a matrix polynomial of lower degree ℓ. This process is called *ℓ-ification* in [?].

Techniques for linearizing a matrix polynomial have been widely investigated. Different companion forms of a matrix polynomial have been introduced and analyzed, see for instance [?, ?, ?] and the literature cited therein. A wide literature exists on matrix polynomials with contribution of many authors [?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?], motivated both by the theoretical interest of this subject and by the many applications that matrix polynomials have [?, ?, ?, ?, ?, ?]. Techniques for reducing a matrix polynomial, or a matrix power series into quadratic form, possibly with coefficients of infinite size, have been investigated in [?, ?]. Reducing a matrix polynomial to a polynomial of degree ℓ is analyzed in [?].

Denote by $\mathbb{C}[x]^{m\times m}$ the set of $m \times m$ matrix polynomials over the complex field $\mathbb{C}$. If $P(x) = \sum_{i=0}^{n} P_i x^i \in \mathbb{C}[x]^{m\times m}$ and $P_n \neq 0$ we say that $P(x)$ has *degree n*. If $\det P(x)$ is not identically zero we say that $P(x)$ is *regular*. Throughout the paper

*Work supported by Gruppo Nazionale di Calcolo Scientifico (GNCS) of INdAM

†Dipartimento di Matematica, Università di Pisa, Largo Bruno Pontecorvo 5, 56127 Pisa (`bini@dm.unipi.it`)

‡Scuola Normale Superiore, Piazza dei Cavalieri, Pisa (`leonardo.robol@sns.it`)

we assume that $P(x)$ is a regular polynomial of degree n. The following definition is useful in our framework.

DEFINITION 1.1. *Let $P(x) \in \mathbb{C}[x]^{m\times m}$ be a matrix polynomial of degree n. Let q be an integer such that $0 < q \leqslant n$. We say that a matrix polynomial $Q(x) \in \mathbb{C}[x]^{qm\times qm}$ is* equivalent *to $P(x)$, and we write $P(x) \approx Q(x)$ if there exist two matrix polynomials $E(x), F(x) \in \mathbb{C}[x]^{qm\times qm}$ such that $\det E(x)$ and $\det F(x)$ are nonzero constants, that is $E(x)$ and $F(x)$ are* unimodular*, and*

$$E(x)Q(x)F(x) = \begin{bmatrix} I_{m(q-1)} & 0 \\ 0 & P(x) \end{bmatrix} =: I_{m(q-1)} \oplus P(x).$$

Denote $P^{\#}(x) = x^n P(x^{-1})$ the reversed polynomial *obtained by reverting the order of the coefficients. We say that the polynomials $P(x)$ and $Q(x)$ are* strongly equivalent *if $P(x) \approx Q(x)$ and $P^{\#}(x) \approx Q^{\#}(x)$. If the degree of $Q(x)$ is 1 and $P(x) \approx Q(x)$ we say that $Q(x)$ is a* linearization *of $P(x)$. Similarly, we say that $Q(x)$ is a* strong linearization *if $Q(x)$ is strongly equivalent to $P(x)$ and $\deg Q(x) = 1$. If $Q(x)$ has degree ℓ we use the terms* ℓ-ification *and* strong ℓ-ification.

It is clear from the definition that $P(x) \approx Q(x)$ implies $\det P(x) = \kappa \det Q(x)$ where κ is some nonzero constant, but the converse is not generally true. The equivalence property is actually stronger because it preserves also the eigenstructure of the matrix polynomial, and not only the eigenvalues. For a more in-depth view of this subject see [**?**].

In the literature, a number of different linearizations have been proposed. The most known are probably the Frobenius and the Fiedler linearizations [**?**]. One of them is, for example,

$$xA - B = x \begin{bmatrix} I_m & & & \\ & \ddots & & \\ & & I_m & \\ & & & P_n \end{bmatrix} - \begin{bmatrix} & & & -P_0 \\ I_m & & & -P_1 \\ & \ddots & & \vdots \\ & & I_m & -P_{n-1} \end{bmatrix}, \qquad (1.1)$$

where I_m denotes the identity matrix of size m.

1.1. New contribution. In this paper we provide a general way to transform a given $m \times m$ matrix polynomial $P(x) = \sum_{i=0}^{n} P_i x^i$ of degree n into a strongly equivalent matrix polynomial $A(x)$ of lower degree ℓ and larger size endowed with a strong structure. The technique relies on representing $P(x)$ with respect to a basis of matrix polynomials of the form $C_i(x) = \prod_{j=1,\, j\neq i}^{q} B_j(x)$, $i = 1, \ldots, q$, where $B_i(x)$ such that $\deg B_i(x) = d_i$ satisfy the following requirements:

1. For every i, j, $B_i(x)B_j(x) - B_j(x)B_i(x) = 0$, i.e., the $B_i(x)$ commute;
2. $B_i(x)$ and $B_j(x)$ are right coprime for every $i \neq j$. This implies that there exist $\alpha_{i,j}(x)$, $\beta_{i,j}(x)$ appropriate matrix polynomials such that $B_i(x)\alpha_{i,j}(x) + B_j(x)\beta_{i,j}(x) = I$.

The above conditions are sufficient to obtain an ℓ-ification. In order to provide a strong ℓ-ification, we need the following additional assumptions

1. $\deg B_i(x) = d$ for $i = 1, \ldots, q$;
2. $B_i^{\#}(x)$ and $B_j^{\#}(x)$ are right coprime for every $i \neq j$.

According to the choice of the basis we arrive at different ℓ-ifications $A(x)$, where $\ell \geqslant \lceil n/q \rceil$ is determined by the degree of the $B_i(x)$, represented as a $q \times q$ block diagonal matrix with $m \times m$ blocks plus a matrix of rank at most m.

Moreover, we provide an explicit version of right and left eigenvectors of $A(x)$ in the general case.

An example of ℓ-ification $A(x)$ is given by

$$\begin{aligned}
&A(x) = D(x) + (e \otimes I_m)[W_1, \dots, W_q],\\
&D(x) = \operatorname{diag}(B_1(x), \dots, B_q(x)), \quad e = (1, \dots, 1)^t \in \mathbb{C}^q,\\
&B_i(x) = b_i(x) I_m, \quad \text{for } i = 1, \dots, q-1,\\
&B_q(x) = b_q(x) P_n + s I_m, \quad d_q = \deg b_q(x),\\
&W_i(x) \in \mathbb{C}[x]^{m \times m}, \quad \deg W_i(x) < \deg B_i(x),
\end{aligned}$$

where $b_1(x), \dots, b_q(x)$ are pairwise co-prime monic polynomials of degree $d_1, \dots, d_q$, respectively, such that $n = d_1 + \cdots + d_q$, and s is such that $\lambda b_q(\xi) + s \neq 0$ for any eigenvalue λ of P_n and for any root ξ of $b_i(x)$ for $i = 1, \dots, q-1$. The matrix polynomial $A(x)$ has degree $\ell = \max\{d_1, \dots, d_q\} \geqslant \lceil \frac{n}{q} \rceil$ and size $mq \times mq$.

If $b_i(x) = x - \beta_i$ are linear polynomials then $\deg A(x) = 1$ and the above equivalence turns into a strong linearization, moreover the eigenvalues of $P(x)$ can be viewed as the generalized eigenvalues of the matrix pencil $A(x)$

$$A(x) = x \begin{bmatrix} I & & & \\ & \ddots & & \\ & & I & \\ & & & P_n \end{bmatrix} - \begin{bmatrix} \beta_1 I & & & \\ & \ddots & & \\ & & \beta_{n-1} I & \\ & & & \beta_n P_n - sI \end{bmatrix} + \begin{bmatrix} I \\ \vdots \\ \vdots \\ I \end{bmatrix} [W_1 \dots W_n]$$

where

$$W_i = \begin{cases} \frac{P(\beta_i)}{\prod_{j=1,\, j\neq i}^{n-1} (\beta_i - \beta_j)} ((\beta_i - \beta_n) P_n + s I_m)^{-1} & \text{for } i < n, \\ \frac{P(\beta_n)}{\prod_{j=1}^{n-1} (\beta_n - \beta_j)} - s I_m - s \sum_{j=1}^{n-1} \frac{W_i}{\beta_n - \beta_j} & \text{otherwise.} \end{cases} \tag{1.2}$$

If $P(x)$ is a scalar polynomial then $\det A(x) = \prod_{i=1}^n (x - \beta_i)(\sum_{j=1}^n \frac{W_j}{x - \beta_j} + 1)$ so that the eigenvalue problem can be rephrased in terms of the secular equation $\sum_{j=1}^n \frac{W_j}{x - \beta_j} + 1 = 0$. Motivated by this fact, we will refer to this linearization as *secular linearization*.

Observe that this kind of linearization relies on the representation of $P(x) - \prod_{i=1}^n B_i(x)$ in the Lagrange basis formed by $C_i(x) = \prod_{j=1,\, j\neq i}^n B_j(x)$, $i = 1, \dots, n$ which is different from the linearization given in [**?**] where the pencil $A(x)$ has an arrowhead structure. Unlike the linearization of [**?**], our linearization does not introduce eigenvalues at infinity.

This secular linearization has some advantages with respect to the Frobenius linearization (1.1). For a monic matrix polynomial we show that with the linearization obtained by choosing $\beta_i = \omega_n^i$, where ω_n is a principal nth root of 1, our linearization is unitarily similar to the block Frobenius pencil associated with $P(x)$. By choosing $\beta_i = \alpha \omega_n^i$, we obtain a pencil unitarily similar to the scaled Frobenius one. With these choices, the eigenvalues of the secular linearization have the same condition number as the eigenvalues of the (scaled) Frobenius matrix.

This observation leads to better choices of the nodes β_i performed according to the magnitude of the eigenvalues of $P(x)$. In fact, by using the information provided by the tropical roots in the sense of [**?**], we may compute at a low cost particular values of the nodes β_i which greatly improve the condition number of the eigenvalues. From an experimental analysis we find that in most cases the conditioning of the eigenvalues of the linearization obtained this way is lower by several orders of magnitude with respect to the conditioning of the eigenvalues of the Frobenius matrix even if it is scaled with the optimal parameter.

Our experiments, reported in Section 6 are based on some randomly generated polynomials and on some problems taken from the repository NLEVP [**?**].

We believe that the information about the tropical roots, used in [**?**] for providing better numerically conditioned problems, can be more effectively used with our ℓ-ification. This analysis is part of our future work.

Another advantage of this representation is that any matrix in the form "diagonal plus low-rank" can be reduced to Hessenberg form H by means of Givens rotation with a low number of arithmetic operations provided that the diagonal is real. Moreover, the function $p(x) = \det(xI - H)$ as well as the Newton correction $p(x)/p'(x)$ can be computed in $O(nm^2)$ operations [**?**]. This fact can be used to implement the Aberth iteration in $O(n^2m^3)$ ops instead of $O(nm^4 + n^2m^3)$ of [**?**]. This complexity bound seems optimal in the sense that for each one of the mn eigenvalues all the $m^2(n+1)$ data are used at least once.

The paper is organized as follows. In Section 2 we provide the reduction of any matrix polynomial $P(x)$ to the equivalent form

$$A(x) = D(x) + (e \otimes I_m)[W_1(x), \ldots, W_q(x)],$$

that is, the ℓ-ification of $P(x)$. In Section 2.1 we show that $P(x)$ is strongly equivalent to $A(x)$ in the sense of Definition 1.1. In Section 3 we provide the explicit form of left and right eigenvectors of $A(x)$ relating them to the corresponding eigenvectors of $P(x)$. In Section 4 we analyze the case where $B_i(x) = b_i(x)I$ for $i = 1, \ldots, q-1$ and $B_q(x) = b_q(x)P_n + sI$ for scalar polynomials $b_i(x)$. Section 5 outlines an algorithm for computing the ℓ-ifications. Finally, in Section 6 we present the results of some numerical experiments.

2. A diagonal plus low rank ℓ-ification. Here we recall a known companion-like matrix for scalar polynomials represented as a diagonal plus a rank-one matrix, provide a more general formulation and then extend it to the case of matrix polynomials. This form was introduced by B.T. Smith in [**?**], as a tool for providing inclusion regions for the zeros of a polynomial, and used by G.H. Golub in [**?**] in the analysis of modified eigenvalue problems.

Let $p(x) = \sum_{i=0}^{n} p_i x^i$ be a polynomial of degree n with complex coefficients, assume $p(x)$ monic, i.e., $p_n = 1$, consider a set of pairwise different complex numbers $\beta_1, \ldots, \beta_n$ and set $e = (1, \ldots, 1)^t$. Then it holds that [**?**], [**?**]

$$\begin{aligned} &p(x) = \det(xI - D + ew^t), \\ &D = \operatorname{diag}(\beta_1, \ldots, \beta_n), \quad w = (w_i),\ w_i = \frac{p(\beta_i)}{\prod_{j \neq i}(\beta_i - \beta_j)}. \end{aligned} \tag{2.1}$$

Now consider a monic polynomial $b(x)$ of degree n factored as $b(x) = \prod_{i=1}^{q} b_i(x)$, where $b_i(x)$, $i = 1, \ldots, q$ are monic polynomials of degree d_i which are co-prime, that is, $\gcd(b_i(x), b_j(x)) = 1$ for $i \neq j$, where gcd denotes the monic greatest common divisor. Recall that given a pair $u(x), v(x)$ of polynomials there exist unique polynomials $s(x), r(x)$ such that $\deg s(x) < \deg v(x)$, $\deg r(x) < \deg u(x)$, and $u(x)s(x) + v(x)r(x) = \gcd(u(x), v(x))$. From this property it follows that if $u(x)$ and $v(x)$ are co-prime, there exists $s(x)$ such that $s(x)u(x) \equiv 1 \mod v(x)$. This polynomial can be viewed as the reciprocal of $u(x)$ modulo $v(x)$. Here and hereafter we denote $u(x) \mod v(x)$ the remainder of the division of $u(x)$ by $v(x)$.

This way, we may uniquely represent any polynomial of degree n in terms of the generalized Lagrange polynomials $c_i(x) = b(x)/b_i(x)$, $i = 1, \ldots, q$ as follows.

LEMMA 2.1. *Let $b_i(x)$, $i = 1, \ldots, q$ be co-prime monic polynomials such that $\deg b_i(x) = d_i$ and $b(x) = \prod_{i=1}^q b_i(x)$ has degree n. Define $c_i(x) = b(x)/b_i(x)$. Then there exist polynomials $s_i(x)$ such that $s_i(x)c_i(x) \equiv 1 \mod b_i(x)$, moreover, any monic polynomial $p(x)$ of degree n can be uniquely written as*

$$\begin{aligned} p(x) &= b(x) + \sum_{i=1}^q w_i(x)c_i(x), \\ w_i(x) &\equiv p(x)s_i(x) \mod b_i(x), \quad i = 1, \ldots, q, \end{aligned} \tag{2.2}$$

where $\deg w_i(x) < d_i$.

Proof. Since $\gcd(b_i(x), b_j(x)) = 1$ for $i \neq j$ then $b_i(x)$ and $c_i(x) = \prod_{j \neq i} b_j(x)$ are co-prime. Therefore there exists $s_i(x) \equiv 1/c_i(x) \mod b_i(x)$. Moreover, setting $w_i(x) \equiv p(x)s_i(x) \mod b_i(x)$ for $i = 1, \ldots, q$, it turns out that the equation $p(x) = b(x) + \sum_{i=1}^q w_i(x)c_i(x)$ is satisfied modulo $b_i(x)$ for $i = 1, \ldots, q$. For the primality of $b_1(x), \ldots, b_q(x)$, this means that the polynomial $\psi(x) := p(x) - b(x) - \sum_{i=1}^q w_i(x)c_i(x)$ is a multiple of $\prod_{i=1}^q b_i(x)$ which has degree n. Since $\psi(x)$ has degree at most $n-1$ it follows that $\psi(x) = 0$. That is (2.2) provides a representation of $p(x)$. This representation is unique since another representation, say, given by $\tilde{w}_i(x)$, $i = 1, \ldots, q$, would be such that $\sum_{i=1}^q (\tilde{w}_i(x) - w_i(x))c_i(x) = 0$, whence $(\tilde{w}_i(x) - w_i(x))c_i(x) \equiv 0 \mod b_i(x)$. That is, for the co-primality of $b_i(x)$ and $c_i(x)$, the polynomial $\tilde{w}_i(x) - w_i(x)$ would be multiple of $b_i(x)$. The property $\deg(b_i(x)) < \deg(\tilde{w}_i(x) - w_i(x))$ implies that $\tilde{w}_i(x) - w_i(x) = 0$. □

The polynomial $p(x)$ in Lemma 2.1 can be represented by means of the determinant of a (not necessarily linear) matrix polynomial as expressed by the following result which provides a generalization of (2.1)

THEOREM 2.2. *Under the assumptions of Lemma 2.1 we have*

$$p(x) = \det A(x), \quad A(x) = D(x) + e[w_1(x), \ldots, w_q(x)]$$

for $D = \operatorname{diag}(b_1(x), \ldots, b_q(x))$ and $e = [1, \ldots, 1]^t$.

Proof. Formally, one has $A(x) = D(x)(I + D(x)^{-1}e[w_1(x), \ldots, w_q(x)])$ so that

$$\begin{aligned} \det A(x) &= \det D(x) \det(I_q + D(x)^{-1}e[w_1(x), \ldots, w_q(x)]) \\ &= b(x)(1 + [w_1(x), \ldots, w_q(x)]D(x)^{-1}e), \end{aligned}$$

where $b(x) = \prod_{i=1}^q b_i(x)$. Whence, we find that $\det A(x) = b(x) + \sum_{i=1}^q w_i(x)c_i(x) = p(x)$, where the latter equality holds in view of Lemma 2.1. □

Observe that for $d_i = 1$ the above result reduces to (2.1) where w_i are constant polynomials. From the computational point of view, the polynomials $w_i(x)$ are obtained by performing a polynomial division since $w_i(x)$ is the remainder of the division of $p(x)s_i(x)$ by $b_i(x)$.

2.1. Strong ℓ-ifications for matrix polynomials. The following technical Lemma is needed to prove the next Theorem 2.5.

LEMMA 2.3. *Let $B_1(x), B_2(x) \in \mathbb{C}[x]^{m \times m}$ be regular and such that $B_1(x)B_2(x) = B_2(x)B_1(x)$. Assume that $B_1(x)$ and $B_2(x)$ are right co-prime, that is, there exist*

$\alpha(x), \beta(x) \in \mathbb{C}[x]^{m\times m}$ *such that* $B_1(x)\alpha(x) + B_2(x)\beta(x) = I_m$. *Then the* 2×2 *block-matrix polynomial* $F(x) = \left[\begin{smallmatrix} \alpha(x) & B_2(x) \\ -\beta(x) & B_1(x)\end{smallmatrix}\right]$ *is unimodular.*

Proof. From the decomposition

$$\begin{bmatrix} I_m & 0 \\ B_1(x) & -B_2(x)\end{bmatrix}\begin{bmatrix} \alpha(x) & B_2(x) \\ -\beta(x) & B_1(x)\end{bmatrix} = \begin{bmatrix} \alpha(x) & B_2(x) \\ I_m & 0\end{bmatrix}$$

we have $-\det B_2(x)\det F(x) = -\det B_2(x)$. Since $B_2(x)$ is regular then $\det F(x) = 1$. □

LEMMA 2.4. *Let* $P(x), Q(x)$ *and* $T(x)$ *be pairwise commuting right co-prime matrix polynomials. Then* $P(x)Q(x)$ *and* $T(x)$ *are also right co-prime.*

Proof. We know that there exists $\alpha_P(x), \beta_P(x), \alpha_Q(x), \beta_Q(x)$, matrix polynomials such that

$$P(x)\alpha_P(x) + T(x)\beta_P(x) = I, \quad Q(x)\alpha_Q(x) + T(x)\beta_Q(x) = I.$$

We shall prove that there exist appropriate $\alpha(x), \beta(x)$ matrix polynomials such that $P(x)Q(x)\alpha(x) + T(x)\beta(x) = I$. We have

$$\begin{aligned} &P(x)Q(x)(\alpha_Q(x)\alpha_P(x)) + T(x)(P(x)\beta_Q(x)\alpha_P(x) + \beta_P(x)) = \\ &P(x)(Q(x)\alpha_Q(x) + T(x)\beta_Q(x))\alpha_P(x) + T(x)\beta_P(x) = \\ &P(x)\alpha_P(x) + T(x)\beta_P(x) = I, \end{aligned}$$

where the first equality holds since $T(x)P(x) = P(x)T(x)$. So we can conclude that also $P(x)Q(x)$ and $T(x)$ are right coprime, and a possible choice for $\alpha(x)$ and $\beta(x)$ is:

$$\alpha(x) = \alpha_Q(x)\alpha_P(x), \qquad \beta(x) = P(x)\beta_Q(x)\alpha_P(x) + \beta_P(x).$$

□

Now we are ready to prove the main result of this section, which provides an ℓ-ification of a matrix polynomial $P(x)$ which is not generally strong. Conditions under which this ℓ-ification is strong are given in Theorem 2.7.

THEOREM 2.5. *Let* $P(x) = \sum_{i=0}^n P_i x^i$, $B_1(x), \dots, B_q(x)$, *and* $W_1(x), \dots, W_q(x)$ *be polynomials in* $\mathbb{C}[x]^{m\times m}$. *Let* $C_i(x) = \prod_{j=1,\, j\neq i}^q B_j(x)$ *and suppose that the following conditions hold:*

1. $P(x) = \prod_{i=1}^q B_i(x) + \sum_{i=1}^q W_i(x)C_i(x)$;
2. *the polynomials* $B_i(x)$ *are regular, commute, i.e.,* $B_i(x)B_j(x) - B_j(x)B_i(x) = 0$ *for any* i, j, *and are pairwise right co-prime.*

Then the matrix polynomial $A(x)$ *defined as*

$$A(x) = D(x) + (e\otimes I_m)[W_1(x), \dots, W_q(x)], \;\; D(x) = \operatorname{diag}(B_1(x), \dots, B_q(x)) \quad (2.3)$$

is equivalent to $P(x)$, *i.e., there exist unimodular* $q\times q$ *matrix polynomials* $E(x)$, $F(x)$ *such that* $E(x)A(x)F(x) = I_{m(q-1)} \oplus P(x)$.

Proof. Define $E_0(x)$ the following (constant) matrix:

$$E_0(x) = \begin{bmatrix} I_m & -I_m & & & \\ & I_m & -I_m & & \\ & & \ddots & \ddots & \\ & & & I_m & -I_m \\ & & & & I_m \end{bmatrix}.$$

A direct inspection shows that

$$E_0(x)A(x) = \begin{bmatrix} B_1(x) & -B_2(x) & & & \\ & B_2(x) & -B_3(x) & & \\ & & \ddots & \ddots & \\ & & & B_{q-1}(x) & -B_q(x) \\ W_1(x) & W_2(x) & \cdots & W_{q-1}(x) & B_q(x)+W_q(x) \end{bmatrix}.$$

Using the fact that the polynomials $B_i(x)$ are right co-prime, we transform the latter matrix into block diagonal form. We start by cleaning $B_1(x)$. Since $B_1(x), B_2(x)$ are right co-prime, there exist polynomials $\alpha(x)$, $\beta(x)$ such that $B_1(x)\alpha(x)+B_2(x)\beta(x) = I_m$. For the sake of brevity, from now on we write α, β, W_i and B_i in place of $\alpha(x), \beta(x)$, $W_i(x)$ and $B_i(x)$, respectively. Observe that the matrix

$$F_1(x) = \begin{bmatrix} \alpha & B_2 \\ -\beta & B_1 \end{bmatrix} \oplus I_{m(q-2)}.$$

is unimodular in view of Lemma 2.3, moreover

$$E_0(x)A(x)F_1(x) = \begin{bmatrix} I_m & & & & \\ -B_2\beta & B_1B_2 & -B_3 & & \\ & & \ddots & \ddots & \\ & & & B_{q-1} & -B_q \\ W_1\alpha - W_2\beta & W_1B_2+W_2B_1 & \cdots & W_{m-1} & B_q+W_q \end{bmatrix}.$$

Using row operations we transform to zero all the elements in the first column of this matrix (by just adding multiples of the first row to the others). That is, there exists a suitable unimodular matrix $E_1(x)$ such that

$$E_1(x)E_0(x)A(x)F_1(x) = \begin{bmatrix} I_m & & & & \\ & B_1B_2 & -B_3 & & \\ & & \ddots & \ddots & \\ & & & B_{q-1} & -B_q \\ & W_2B_1+W_1B_2 & \cdots & W_{q-1} & B_q+W_q \end{bmatrix}.$$

In view of Lemma 2.4, B_1B_2 is right coprime with B_3. Thus, we can recursively apply the same process until we arrive at the final reduction step:

$$E_{q-1}(x)\dots E_0(x)A(x)F_1(x)\dots F_{q-1}(x) = I_{m(q-1)} \oplus \left(\prod_{i=1}^{q} B_i(x) + \sum_{i=1}^{q} W_i(x)C_i(x) \right)$$

where the last diagonal block is exactly $P(x)$ in view of assumption 1. □

Observe that if $m = 1$ and $B_i(x) = x - \beta_i$ then we find that (2.1) provides the secular linearization for $P(x)$.

Now, we can prove that the polynomial equivalence that we have just presented is actually a strong equivalence, under the following additional assumptions

$$\begin{aligned} &\deg B_i(x) = d, \quad i = 1, \dots, q, \\ &n = dq, \quad \deg W_i(x) < \deg B_i(x) \\ &B_i^{\#}(x), \text{ are pairwise right co-prime.} \end{aligned} \tag{2.4}$$

REMARK 2.6. *Recall that, according to Theorem 7.5 of [?] in an algebraically closed field there exists a strong ℓ-ification for a degree n regular matrix polynomial $P(x) \in \mathbb{C}^{m\times m}[x]$ if and only if $\ell | nm$. Conditions* (2.4) *satisfy this requirement, since $\ell = d$ and $d|n$.*

To accomplish this task we will show that the reversed matrix polynomial of $A(x) = D(x) + (e \otimes I_m)W$, $W = [W_1, \dots, W_q]$, has the same structure as $A(x)$ itself.

THEOREM 2.7. *Under the assumptions of Theorem 2.5 and of* (2.4) *the secular ℓ-ification given in Theorem 2.5 is strong.*

Proof. Consider $A^{\#}(x) = x^d A(x^{-1})$. We have

$$A^{\#}(x) = \mathrm{diag}(x^d B_1(x^{-1}), \dots, x^d B_q(x^{-1})) + (e \otimes I_m)[x^d W_1(x^{-1}), \cdots, x^d W_q(x^{-1})].$$

This matrix polynomial is already in the same form as $A(x)$ of (2.3) and verifies the assumptions of Theorem 2.5 since the polynomials $x^d B_1(x^{-1})$ are pairwise right coprime and commute with each other in view of equation (2.4). Thus, Theorem 2.5 implies that $A^{\#}(x)$ is an ℓ-ification for

$$\begin{aligned}\prod_{i=1}^{q} x^d B_i(x^{-1}) &+ \sum_{i=1}^{q} x^d W_i(x^{-1}) \prod_{j\neq i} x^d B_j(x^{-1}) \\ &= x^{dq}\left(\prod_{i=1}^{q} B_i(x^{-1}) + \sum_{i=1}^{q} W_i(x^{-1}) \prod_{j\neq i} B_j(x^{-1})\right) \\ &= x^n P(x^{-1}) = P^{\#}(x),\end{aligned}$$

where the first equality follows from the fact that $n = dq$ in view of equation (2.4). This concludes the proof. □

REMARK 2.8. *Note that in the case where the $B_i(x)$ do not have the same degree, the secular ℓ-ification might not be strong: the finite eigenstructure is preserved but some infinite eigenvalues not present in $P(x)$ will be artificially introduced.*

3. Eigenvectors. In this section we provide an explicit expression of right and left eigenvectors of the matrix polynomial $A(x)$.

THEOREM 3.1. *Let $P(x)$ be a matrix polynomial, $A(x)$ its secular ℓ-ification defined in Theorem 2.5, $\lambda \in \mathbb{C}$ such that $\det P(\lambda) = 0$, and assume that $\det B_i(\lambda) \neq 0$ for all $i = 1, \dots, q$. If $v_A = (v_1^t, \dots, v_q^t)^t \in \mathbb{C}^{mq}$ is such that $A(\lambda)v_A = 0$, $v_A \neq 0$ then $P(\lambda)v = 0$ where $v = -\prod_{i=1}^{q} B_i(\lambda)^{-1} \sum_{j=1}^{q} W_j v_j \neq 0$. Conversely, if $v \in \mathbb{C}^m$ is a nonzero vector such that $P(\lambda)v = 0$, then the vector v_A defined by $v_i = \prod_{j\neq i} B_j(\lambda)v$, $i = 1, \dots, q$ is nonzero and such that $A(\lambda)v_A = 0$.*

Proof. Let $v_A \neq 0$ be such that $A(\lambda)v_A = 0$, so that

$$B_i(\lambda)v_i + \sum_{j=1}^{q} W_j(\lambda)v_j = 0, \quad i = 1, \dots, q. \tag{3.1}$$

Let $v = -(\prod_{i=1}^{q} B_i(\lambda)^{-1}) \sum_{j=1}^{q} W_j(\lambda)v_j$. Combining the latter equation and (3.1) yields

$$v_i = -B_i(\lambda)^{-1}\left(\sum_{j=1}^{q} W_j(\lambda)v_j\right) = \prod_{j=1,\, j\neq i}^{q} B_j(\lambda)v. \tag{3.2}$$

Observe that if $v = 0$ then, by definition of v, one has $\sum_{j=1}^{q} W_j(\lambda)v_j = 0$ so that, in view of (3.1), we find that $B_i(\lambda)v_i = 0$. Since $\det B_i(\lambda) \neq 0$ this would imply that

$v_i = 0$ for any i so that $v_A = 0$ which contradicts the assumptions. Now we prove that $P(\lambda)v = 0$. In view of (3.2) we have

$$P(\lambda)v = \prod_{j=1}^{q} B_j(\lambda)v + \sum_{i=1}^{q} W_i(\lambda) \prod_{j=1,\, j\neq i}^{q} B_j(\lambda)v = \prod_{j=1}^{q} B_j(\lambda)v + \sum_{i=1}^{q} W_i(\lambda)v_i.$$

Moreover, by definition of v we get

$$P(\lambda)v = -\prod_{j=1}^{q} B_j(\lambda)(\prod_{i=1}^{q} B_i(\lambda)^{-1}) \sum_{i=1}^{q} W_i(\lambda)v_i + \sum_{i=1}^{q} W_i(\lambda)v_i = 0.$$

Similarly, we can prove the opposite implication. □

A similar result can be proven for left eigenvectors. The following theorem relates left eigenvectors of $A(x)$ and left eigenvectors of $P(x)$.

THEOREM 3.2. *Let $P(x)$ be a matrix polynomial, $A(x)$ its secular ℓ-ification defined in Theorem 2.5, $\lambda \in \mathbb{C}$ such that* $\det P(\lambda) = 0$, *and assume that* $\det B_i(\lambda) \neq 0$. *If $u_A^t = (u_1^t, \ldots, u_q^t) \in \mathbb{C}^{mq}$ is such that $u_A^t A(\lambda) = 0$, $u_A \neq 0$, then $u^t P(\lambda) = 0$ where $u = \sum_{i=1}^{q} u_i \neq 0$. Conversely, if $u^t P(\lambda) = 0$ for a nonzero vector $u \in \mathbb{C}^m$ then $u_A^t A(\lambda) = 0$, where u_A is a nonzero vector defined by $u_i^t = -u^t W_i(\lambda) B_i(\lambda)^{-1}$ for $i = 1, \ldots, q$.*

Proof. If $u_A^t A(\lambda) = 0$ then from the expression of $A(x)$ given in Theorem 2.5 we have

$$u_i^t B_i(\lambda) + \left(\sum_{j=1}^{q} u_j^t \right) W_i(\lambda) = 0, \quad i = 1, \ldots, q. \tag{3.3}$$

Assume that $u = \sum_{j=1}^{q} u_j = 0$. Then from the above expression we obtain, for any i, $u_i^t B_i(\lambda) = 0$ that is $u_i = 0$ for any i since $\det B_i(\lambda) \neq 0$. This is in contradiction with $u_A \neq 0$. From (3.3) we obtain $u_i^t = -u^t W_i(\lambda) B_i(\lambda)^{-1}$. Moreover, multiplying (3.3) to the right by $\prod_{j=1,\, j\neq i}^{q} B_j$ yields

$$0 = u_i^t \prod_{j=1}^{q} B_j(\lambda) + u^t W_i(\lambda) \prod_{j=1,\, j\neq i}^{q} B_j(\lambda).$$

Taking the sum of the above expression for $i = 1, \ldots, q$ yields

$$0 = \left(\sum_{i=1}^{q} u_i^t \right) \prod_{j=1}^{q} B_j(\lambda) + u^t \sum_{i=1}^{q} W_i(\lambda) \prod_{j=1,\, j\neq i}^{q} B_j(\lambda) = u^t P(\lambda).$$

Conversely, assuming that $u^t P(\lambda) = 0$, from the representation

$$P(x) = \prod_{j=1}^{n} B_j(x) + \sum_{i=1}^{q} W_i(x) \prod_{j=1,\, j\neq i}^{q} B_i(x),$$

defining $u_i^t = -u^t W_i(\lambda) B_i(\lambda)^{-1}$ we obtain

$$\sum_{i=1}^{q} u_i^t = -u^t \sum_{i=1}^{q} W_i(\lambda) B_i(\lambda)^{-1} = -u^t (P(\lambda) \prod_{j=1}^{q} B_j(\lambda)^{-1} - I) = u^t$$

and therefore from (3.3) we deduce that $u_A^t A(\lambda) = 0$. □

The above result does not cover the case where $\det B_i(\lambda) = 0$ for some i.

3.1. A sparse ℓ-ification. Consider the block bidiagonal matrix L having I_m on the block diagonal and $-I_m$ on the block subdiagonal. It is immediate to verify that $L(e \otimes I_m) = e_1 \otimes I_m$, where $e_1 = (1, 0, \dots, 0)^t$. This way, the matrix polynomial $H(x) = LA(x)$ is a sparse ℓ-ification of the form

$$H(x) = \begin{bmatrix} B_1(x) + W_1(x) & W_2(x) & \dots & W_{q-1}(x) & W_q(x) \\ -B_1(x) & B_2(x) & & & \\ & -B_2(x) & \ddots & & \\ & & \ddots & B_{q-1}(x) & \\ & & & -B_{q-1}(x) & B_q(x) \end{bmatrix}$$

4. A particular case. In the previous section we have provided (strong) ℓ-ifications of a matrix polynomial $P(x)$ under the assumption of the existence of the representation

$$P(x) = \prod_{i=1}^{q} B_i(x) + \sum_{i=1}^{q} W_i(x) \prod_{j \neq i} B_j(x) \tag{4.1}$$

and under suitable conditions on $B_i(x)$. In this section we show that a specific choice of the blocks $B_i(x)$ satisfies the above assumptions and implies the existence of the representation (4.1). Moreover, we provide explicit formulas for the computation of $W_i(x)$ given $P(x)$ and $B_i(x)$.

We provide also some additional conditions in order to make the resulting ℓ-ification strong.

ASSUMPTION 1. *The matrix polynomials $B_i(x)$ are defined as follows*

$$\begin{cases} B_i(x) = b_i(x) I & i = 1, \dots, q-1 \\ B_q(x) = b_q(x) P_n + sI & \text{otherwise} \end{cases}$$

where $b_i(x)$ are scalar polynomials such that $\deg b_i(x) = d_i$ and $\sum_{i=1}^{q} d_i = n$; the polynomials $b_i(x)$, $i = 1, \dots, q$, are pairwise co-prime; s is a constant such that $\lambda b_q(\xi) + s \neq 0$ for any eigenvalue λ of P_n and for any root ξ of $b_i(x)$, for $i = 1, \dots, q-1$.

In this case it is possible to prove the existence of the representation (4.1). We rely on the Chinese remainder theorem that here we rephrase in terms of matrix polynomials.

LEMMA 4.1. *Let $b_i(x)$, $i = 1, \dots, q$ be co-prime polynomials of degree $d_1, \dots, d_q$, respectively, such that $\sum_{i=1}^{q} d_i = n$. If $P_1(x)$, $P_2(x)$ are matrix polynomials of degree at most $n-1$ then $P_1(x) = P_2(x)$ if and only if $P_1(x) - P_2(x) \equiv 0 \mod b_i(x)$, for $i = 1, \dots, q$.*

Proof. The implication $P_1(x) - P_2(x) = 0 \Rightarrow P_1(x) - P_2(x) \equiv 0 \mod b_i(x)$ is trivial. Conversely, if $P_1(x) - P_2(x) \equiv 0 \mod b_i(x)$ for every b_i then the entries of $P_1(x) - P_2(x)$ are multiples of $\prod_{i=1}^{q} b_i(x)$ for the co-primality of the polynomials $b_i(x)$. But this implies that $P_1(x) - P_2(x) = 0$ since the degree of $P_1(x) - P_2(x)$ is at most $n-1$ while $\prod_{i=1}^{q} b_i(x)$ has degree n. □

We have the following

THEOREM 4.2. *Let $P(x) = \sum_{i=0}^{n} x^i P_i$ be an $m \times m$ matrix polynomial over an algebraically closed field. Under Assumption 1, set $C_i(x) = \prod_{j \neq i} B_j(x)$. Then there* *exists a unique decomposition*

$$P(x) = B(x) + \sum_{i=1}^{q} W_i(x) C_i(x), \qquad B(x) = \prod_{i=1}^{q} B_i(x), \tag{4.2}$$

where $W_i(x)$ *are matrix polynomials of degree less than* d_i *for* $i = 1, \dots, q$ *defined by*

$$\begin{aligned} W_i(x) &= \frac{P(x)}{\prod_{j=1,\, j\neq i}^{q-1} b_j(x)} (b_q(x) P_n + s I_m)^{-1} \mod b_i(x), \quad i = 1, \dots, q-1 \\ W_q(x) &= \frac{1}{\prod_{j=1}^{q-1} b_j(x)} P(x) - s I_m - s \sum_{j=1}^{q-1} \frac{W_j(x)}{b_j(x)} \mod b_q(x). \end{aligned} \tag{4.3}$$

Proof. We show that there exist matrix polynomials $W_i(x)$ of degree less than d_i such that $P(x) - B(x) \equiv \sum_{i=1}^{q} W_i(x) C_i(x) \mod b_i(x)$ for $i = 1, \dots, q$. Then we apply Lemma 4.1 with $P_1(x) = P(x) - B(x)$ that by construction has degree at most $n-1$, and with $P_2(x) = \sum_{i=1}^{q} W_i(x) C_i(x)$, and conclude that $P(x) = B(x) + \sum_{i=1}^{q} W_i(x) C_i(x)$. Since for $i = 1, \dots, q-1$ the polynomial $b_i(x)$ divides every entry of $B(x)$ and of $C_j(x)$ for $j \neq i$, we find that $P(x) \equiv W_i(x) C_i(x) \mod b_i(x)$, $i = 1, \dots, q$. Moreover, for $i < q$ we have $C_i(x) = \left(\prod_{j\neq i, j<q} b_j(x) I_m\right)(b_q(x) P_n + s I_m)$. The first term is invertible modulo $b_i(x)$ since by assumption $b_i(x)$ is co-prime with b_j for every $j \neq i$. We need to prove that the matrix on the right is invertible modulo $b_i(x)$, that is, its eigenvalues μ are such that $b_i(\mu) \neq 0$. Now, since the eigenvalues of $b_q(x) P_n + s I_m$ have the form $\mu = b_q(x)\lambda + s$, where λ is an eigenvalue of P_n, it is enough to ensure that for every ξ which is a root of $b_i(x)$ the value $\lambda b_q(\xi) + s$ is different from 0 for $i = 1, \dots, q-1$. This is guaranteed by hypothesis, and so we obtain the explicit formula for $W_i(x)$, $i = 1, \dots, q-1$ given by (4.3). It remains to find an explicit expression for $W_q(x)$. We have $W_q(x) C_q(x) = P(x) - \sum_{j=1}^{q-1} W_j(x) C_j(x)$, where the right-hand side is made by known polynomials. This way, taking the latter expression modulo $b_q(x)$ we can compute $W_q(x)$ since $C_q(x) = \prod_{j=1}^{q-1} b_j(x) I_m$ is invertible modulo $b_q(x)$ in view of the co-primality of the polynomials $b_1(x), \dots, b_q(x)$. This way we get the expression of W_q in (4.3). □

Remark 4.3. *Note that in the case where* $P_n = I$ *it is possible to choose* $s = 0$ *so that Equations* (4.3) *take a simpler form.*

We observe that the matrix polynomials $B_i(x)$ which satisfy Assumption 1 verify the hypotheses of Theorem 2.5. Therefore there exists an ℓ-ification of $P(x)$ which can be computed. In view of Theorem 2.7 we have that this ℓ-ification is also strong if the following conditions are satisfied:

1. The $B_i(x)$ have the same degree d. In our case this implies that $\deg b_i(x) = d$ for every $i = 1, \dots, q$.
2. The matrix polynomials $x^d B_i(x^{-1})$ are right coprime. It can be seen that under Assumption 1 this is equivalent to asking that $b_i(0) \neq 0$ for every $i = 1, \dots, q$ and that either $P_n \neq I$ or $\xi^d b_q(\xi) + s \neq 0$ for every ξ root of $b_i(x)$, for $i < q$.

Here we provide an example of an ℓ-ification of degree 2 for a 2×2 matrix polynomial $P(x)$ of degree 4.

Example 4.4. Let

$$P(x) = \begin{bmatrix} x^4 + 2 & -1 \\ x & x^3 - 1 \end{bmatrix}, \quad b_1(x) = x^2 - 2, \quad b_2(x) = x^2 + 2, \quad s = 1.$$

Applying the above formulas we obtain

$$W_1(x) = \begin{bmatrix} \frac{6}{5} & -1 \\ \frac{1}{5}x & -1+2x \end{bmatrix}, \qquad W_2(x) = \begin{bmatrix} -\frac{11}{5} & 0 \\ -\frac{1}{5}x & -1+x \end{bmatrix}.$$

Then we have that $A(x)$ is a degree 2 ℓ-ification for $P(x)$, that is, a quadratization, by setting

$$A(x) = \begin{bmatrix} x^2-2 & 0 & 0 & 0 \\ 0 & x^2-2 & 0 & 0 \\ 0 & 0 & x^2+3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} \frac{6}{5} & -1 & -\frac{11}{5} & 0 \\ \frac{1}{5}x & -1+2x & -\frac{1}{5}x & -1+x \\ \frac{6}{5} & -1 & -\frac{11}{5} & 0 \\ \frac{1}{5}x & -1+2x & -\frac{1}{5}x & -1+x \end{bmatrix}.$$

In the case where $b_i(x)$, $i = 1, \dots, q$ are linear polynomials we have the following:

Corollary 4.5. *If $b_i(x) = x - \beta_i$, $i = 1, \dots, q$, then $q = n$ and*

$$W_i = \frac{P(\beta_i)}{\prod_{j=1,\, j\neq i}^{n-1}(\beta_i - \beta_j)}((\beta_i - \beta_n)P_n + sI_m)^{-1}, \quad i = 1, \dots, n-1,$$

$$W_n = \frac{P(\beta_n)}{\prod_{j=1}^{n-1}(\beta_n - \beta_j)} - sI_m - s\sum_{j=1}^{n-1}\frac{W_j}{\beta_n - \beta_j}.$$

Moreover, if $P(x)$ is monic then with $s = 0$ the expression for W_i turns simply into $W_i = P(\beta_i)/\prod_{j=1,\, j\neq i}^{n}(\beta_i - \beta_j)$, for $i = 1, \dots, n$.

Proof. It follows from Theorem 4.2 and from the property $v(x) \mod x - \beta = v(\beta)$ valid for any polynomial $v(x)$. □

Given n and $q \leqslant n$, let $\ell = \lceil \frac{n}{q} \rceil$. We may choose polynomials $b_i(x)$ of degree d_i in between $\ell - 1$ and ℓ such that $\sum_{i=1}^{q} d_i = n$. This way $A(x)$ is an $mq \times mq$ matrix polynomial of degree ℓ. For instance, if $\ell = 2$ we obtain a quadratization of $P(x)$.

If $P(x)$ is monic, that is $P_n = I$, we can handle another particular case of ℓ-ification by choosing diagonal matrix polynomials $B_i(x)$.

Let $B_i(x) = \mathrm{diag}(d_1^{(i)}(x), \dots, d_m^{(i)}(x)) =: D_i(x)$ be monic matrix polynomials such that the corresponding diagonal entries of $D_i(x)$ and $D_j(x)$ are pairwise co-prime for any $i \neq j$ so that the second assumption of Theorem 2.5 is satisfied. Let us prove that there exist matrix polynomials $W_i(x)$ such that $\deg W_i(x) < \deg D_i(x)$ and

$$P(x) = \prod_{i=1}^{q} D_i(x) + \sum_{i=1}^{q} W_i(x)C_i(x), \quad C_i(x) = \prod_{j=1,\, j\neq i}^{q} D_i(x), \tag{4.4}$$

so that Theorem 2.5 can be applied. Observe that equating the coefficients of x^i in (4.4) for $i = 0, \dots, n-1$ provides a linear system of $m^2 n$ equations in $m^2 n$ unknowns. Equating the jth columns of both sides of (4.4) modulo $d_j^{(i)}(x)$ yields

$$P(x)e_j \mod d_j^{(i)}(x) = \prod_{s=1,\, s\neq i}^{n} d_j^{(s)} W_i(x)e_j \mod d_j^{(i)}(x), \quad i = 1, \dots, m.$$

The above equation allows one to compute the coefficients of the polynomials of degree at most $\deg D_j(x) - 1$ in the jth column of $W_j(x)$ by means of the Chinese remainder theorem.

5. Computational issues. In the previous section we have given explicit formulas for the (strong) ℓ-ification of a matrix polynomial satisfying Assumption 1. Here we describe some algorithms for the computation of the matrix coefficients $W_i(x)$.

In the case where $d = 1$, the equations given in Corollary 4.5 provide a straightforward algorithm for the computation of the matrices W_i for $i = 1, \ldots, n$. In the case where $d_i = \deg b_i(x) > 1$ for some values of i we have to apply (4.3) which involve operations modulo scalar polynomials $b_i(x)$ for $i = 1, \ldots, q$.

The main computational issues in this case are the evaluation of a scalar polynomial modulo a given $b_i(x)$, the evaluation of the inverse of a scalar polynomial modulo $b_i(x)$ and the more complicated task of evaluating the inverse of a matrix polynomial modulo $b_i(x)$.

In general we recall that, given polynomials $v(x)$ and $b(x)$ such that $v(x)$ is co-prime with $b(x)$, there exist polynomials $\alpha(x)$ and $\beta(x)$ such that

$$\alpha(x)v(x) + \beta(x)b(x) = 1. \tag{5.1}$$

This way, we have $\alpha(x) = 1/v(x) \mod b(x)$.

There are algorithms for computing the coefficients of $\alpha(x)$ given the coefficients of $v(x)$ and $b(x)$. We refer the reader to the book [**?**] and to any textbook in computer algebra for the design and analysis of algorithms for this computation. Here, we recall a simple numerical technique, based on the evaluation-interpolation strategy, which can be also directly applied to the matrix case.

Observe that from (5.1) it turns out that $\alpha(\xi_j) = 1/v(\xi_j)$ for $j = 1, \ldots, d$, where ξ_j are the zeros of the polynomial $b(x)$ of degree d. Since $\alpha(x)$ has degree at most $d - 1$, it is enough to compute the values $1/v(\xi_j)$ in order to recover the coefficients of $\alpha(x)$ through an interpolation process. This procedure amounts to d evaluations of a polynomial at a point and to solving an interpolation problem for the overall cost of $O(d^2)$ arithmetic operations. If the polynomial $b(x)$ is chosen in such a way that its roots are multiples of the dth roots of unity then the evaluation/interpolation problem is well conditioned, and it can be performed by means of FFT which is a fast and numerically stable procedure.

The evaluation/interpolation technique can be extended to the case of matrix polynomials. For instance, the computation of the coefficients of the matrix polynomial $F(x) = V(x)^{-1} \mod b(x)$, where $V(x)$ is a given matrix polynomial co-prime with $b(x)I$, can be performed in the following way:

1. compute $Y_k = V(\xi_k)^{-1}$, for $k = 1, \ldots, d$;
2. for any pair (i, j), interpolate the entries $y_{i,j}^{(k)}$, $k = 1, \ldots, d$ of the matrix Y_k and find the coefficients of the polynomial $f_{i,j}(x)$, where $F(x) = (f_{i,j}(x))$.

This procedure requires the evaluation of m^2 polynomials at d points, followed by the inversion of d matrices of order m and the solution of m^2 interpolation problems. The cost turns to $O(m^2d^2)$ ops for the evaluation stage, $O(m^3d)$ ops for the inversion stage, and $O(m^2d^2)$ for the interpolation stage. In the case where the polynomial $b(x)$ is such that its roots are multiple of the d roots of the unity, the evaluation and the interpolation stage have cost $O(m^2 d \log d)$ if performed by means of FFT.

Observe that, in the case of polynomials $b_i(x)$ of degree one, the above procedure coincides with the one provided directly by equations in Corollary (4.5).

6. Numerical issues. Let ω_n be a principal nth root of the unity, define $\Omega_n = \frac{1}{\sqrt{n}}(\omega_n^{ij})_{i,j=1,n}$ the Fourier matrix such that $\Omega_n^*\Omega_n = I_n$ and observe that $\Omega_n e = e_n$ where $e = (1, \ldots, 1)^t$, $e_n = (0, \ldots, 0, 1)^t$. Assume for simplicity $P_n = I_m$. For the linearization obtained with $\beta_i = \omega_n^i$, $i = 1, \ldots, n$, we have, following Corollary 4.5,

$$A(x) = xI_{mn} - \mathrm{diag}(\omega_n^1 I_m, \omega_n^2 I_m, \ldots, \omega_n^n I_m) + (e \otimes I_m)[W_1, \ldots, W_n]$$

with $W_i = \frac{1}{n}\omega_n^i P(\omega_n^i)$. The latter equation follows from $W_i = P(\omega_n^i)/(\prod_{j\neq i}(\omega_n^i - \omega_n^j))$ since $\prod_{j\neq i}(\omega_n^i - \omega_n^j)$ coincides with the first derivative of $x^n - 1 = \prod_{j=1}^n (x - \omega_n^j)$ evaluated at $x = \omega_n^i$, that is $\prod_{j\neq i}(\omega_n^i - \omega_n^j) = n\omega_n^{-i}$. It is easy to verify that the pencil $(\Omega_n^* \otimes I_m)A(x)(\Omega_n \otimes I_m)$ has the form

$$xI_{mn} - F, \quad F = (C \otimes I_m) - \begin{bmatrix} P_0 + I_m \\ P_1 \\ \vdots \\ P_{n-1} \end{bmatrix} (e_n^t \otimes I_m)$$

where $C = (c_{i,j})$ is the unit circulant matrix defined by $c_{i,j} = (\delta_{i,j+1 \mod n})$. That is, F is the block Frobenius matrix associated with the matrix polynomial $P(x)$.

This shows that our linearization includes the companion Frobenius matrix with a specific choice of the nodes. In particular, since Ω_n is unitary, the condition number of the eigenvalues of $A(x)$ coincides with the condition number of the eigenvalues of F. Observe also that if we choose $\beta_i = \alpha\omega_n^i$ with $\alpha \neq 0$, then $(\Omega_n^* \otimes I_m)A(x)(\Omega_n \otimes I_m) = xI - D_\alpha^{-1} F D_\alpha$ for $D_\alpha = \mathrm{diag}(1, \alpha, \ldots, \alpha^{n-1})$. That is, we obtain a scaled Frobenius pencil.

Here, we present some numerical experiments to show that in many interesting cases a careful choice of the $B_i(x)$ can lead to linearizations (or ℓ-ifications) where the eigenvalues are much better conditioned than in the original problem. Here we are interested in measuring the conditioning of the eigenvalues of a pencil built using these different strategies. Recall that the conditioning of an eigenvalue λ of a matrix pencil $xA - B$ can be bounded by $\kappa_\lambda \leqslant \frac{\|v\|\|w\|}{|w^* A v|}$ where v and w are the right and left eigenvectors relative to λ, respectively [?]. This is the quantity measured by the `condeig` function in MATLAB that we have used in the experiments. The above bound can be extended to a matrix polynomial $P(x) = \sum_{i=0}^n P_i x^i$. In particular, the conditioning number of an eigenvalue λ of P can be bounded by $\kappa_\lambda \leqslant \frac{\|v\|\|w\|}{|w^* P'(\lambda) v|}$ where v and w are the right and left eigenvectors relative to λ, respectively, [?]. Observe that, if $xA - B$ is the Frobenius linearization of a matrix polynomial $P(x)$, then the condition number of the eigenvalues of the linearization is larger than the one concerning $P(x)$, since the perturbation to the input data on $xA - B$ can be seen as a larger set with respect to the perturbation to the coefficients of $P(x)$. This is not true, in general, for a linearization in a different basis, as in our case, since there is not a direct correspondence between the perturbations on the original coefficients and the perturbations on the linearization. An analysis of the condition number for eigenvalue problems of matrix polynomials represented in different basis is given in [?].

The code used to generate these examples can be downloaded from `http://numpi.dm.unipi.it/software/secular-linearization/`.

6.1. Scalar polynomials. As a first example, consider a monic scalar polynomial $p(x) = \sum_{i=0}^n p_i x^i$ where the coefficients p_i have unbalanced moduli. In this case, we generate p_i using the MATLAB command `p = exp(12 * randn(1,n+1)); p(n+1)=1;`

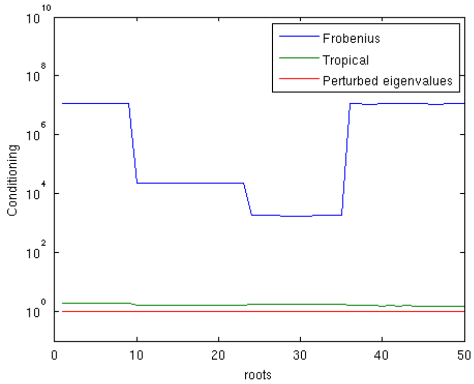

FIGURE 6.1. *Conditioning of the eigenvalues of different linearizations of a degree* 50 *scalar polynomial with random unbalanced coefficients.*

Then we build our linearization by means of the function `seccomp(b,p)` that takes a vector `b` together with the coefficients of the polynomial and generates the linearization $A(x)$ where $B_i(x) = x - \beta_i$ for $\beta_i = \mathtt{b(i)}$. Finally, we measure the conditioning of the eigenvalues of $A(x)$ by means of the Matlab function `condeig`.

We have considered three different linearizations:

- The Frobenius linearization obtained by `compan(p)`;
- the secular linearization obtained by taking as β_i some perturbed values of the roots; these values have been obtained by multiplying the roots by $(1+\epsilon)$ with ϵ chosen randomly with Gaussian distribution $\epsilon \sim 10^{-12} \cdot N(0,1)$.
- the secular linearization with nodes given by the tropical roots of the polynomial multiplied by unit complex numbers.

The results are displayed in Figure 6.1. One can see that in the first case the condition numbers of the eigenvalues are much different from each other and can be as large as 10^8 for the worst conditioned eigenvalue. In the second case the condition number of all the eigenvalues is close to 1, while in the third linearization the condition numbers are much smaller than those of the Frobenius linearization and have an almost uniform distribution.

These experimental results are a direct verification of a conditioning result of [**?**, Sect. 5.2] that is at the basis of the `secsolve` algorithm presented in that paper. These tests are implemented in the function files `Example1.m` and `Experiment1.m` included in the MATLAB source code for the experiments. A similar behavior of the conditioning for the eigenvalue problem holds in the matrix case.

6.2. The matrix case. Consider now a matrix polynomial $P(x) = \sum_{i=0}^{n} P_i x^i$. As in the previous case, we start by considering monic matrix polynomials. As a first example, consider the case where the coefficients P_i have unbalanced norms. Here is the Matlab code that we have used to generate this test:

```
n = 5; m = 64;
P = {};
for i = 1 : n
  P{i} = exp(12 * randn) * randn(m);
end
```

```
P{n+1} = eye(m);
```

We can give reasonable estimates to the modulus of the eigenvalues using the Pellet theorem or the tropical roots. See [**?**, **?**], for some insight on these tools.

The same examples given in the scalar case have been replicated for matrix polynomials relying on the Matlab script published on the website reported above by issuing the following commands:

```
>> P = Example2();
>> Experiment2(P);
```

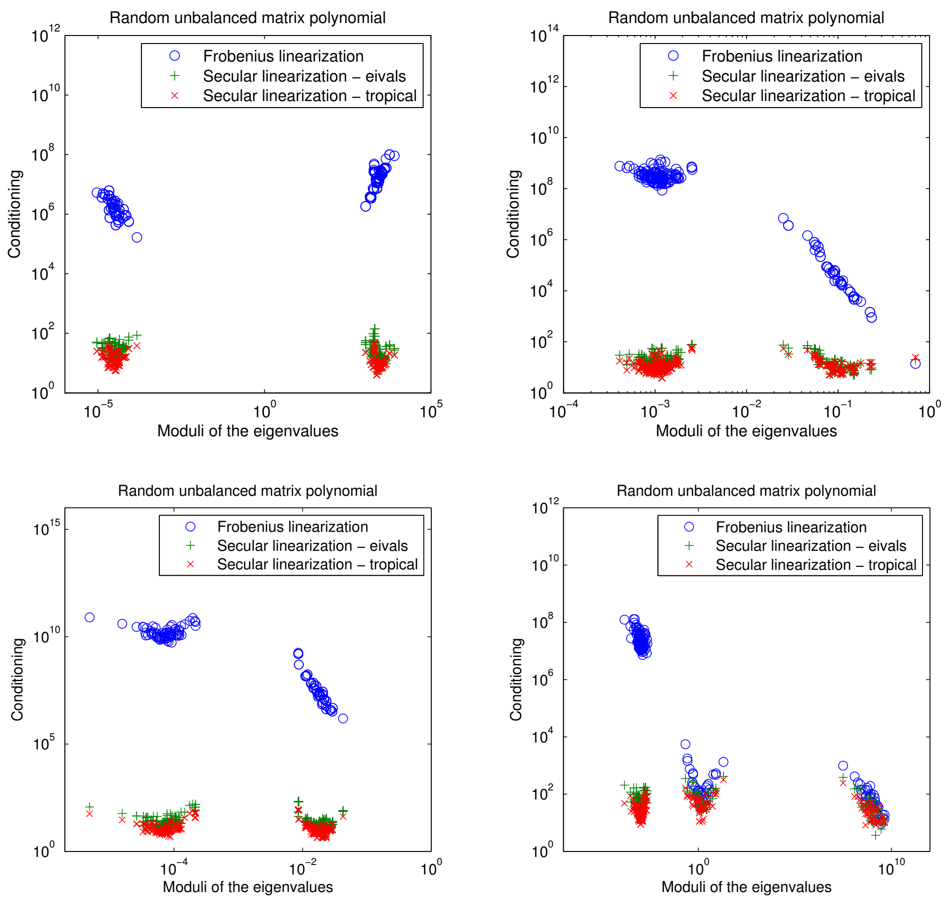

FIGURE 6.2. *Conditioning of the eigenvalues of different linearizations for some matrix polynomials with random coefficients having unbalanced norms.*

We have considered three linearizations: the standard Frobenius companion linearization, and two versions of our secular linearizations. In the first version the nodes β_i are the mean of the moduli of set of eigenvalues with close moduli multiplied by unitary complex numbers. In the second, the values of β_i are obtained by the Pellet estimates delivered by the tropical roots.

In Figure 6.2 we report the conditioning of the eigenvalues, measured with Matlab's `condeig`.

It is interesting to note that the conditioning of the secular linearization is, in every case, not exceeding 10^3. Moreover it can be observed that no improvement is obtained on the conditioning of the eigenvalues that are already well-conditioned. In contrast, there is a clear improvement on the ill-conditioned ones. In this particular case, this class of linearizations seems to give an almost uniform bound to the condition number of all the eigenvalues.

Further examples come from the NLEVP collection of [**?**]. We have selected some problems that exhibit bad conditioning.

As a first example we consider the problem `orr_sommerfeld`. Using the tropical roots we can find some values inside the unique annulus that is identified by the Pellet theorem. In this example the values obtained only give a partial picture of the eigenvalues distribution. The Pellet theorem gives about `1.65e-4` and `5.34` as lower and upper bound to the moduli of the eigenvalues, but the tropical roots are rather small and near to the lower bound. More precisely, the tropical roots are `1.4e-3` and `1.7e-4` with multiplicities 3 and 1, respectively.

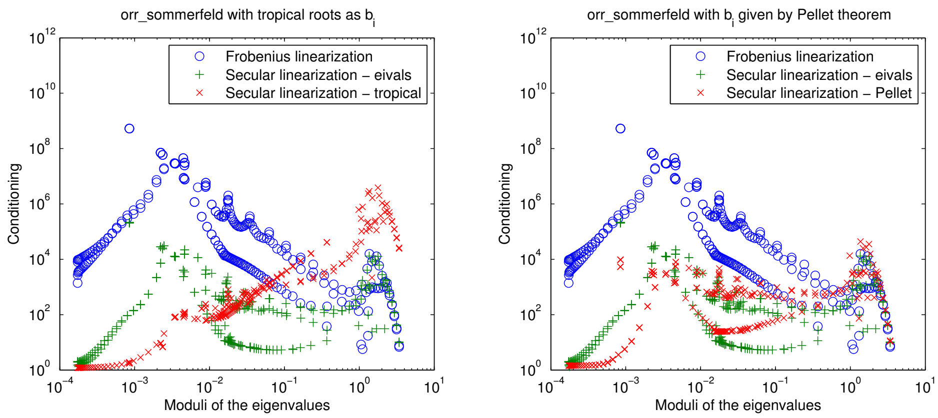

FIGURE 6.3. *On the left we report the conditioning of the Frobenius and of the secular linearization with the choices of b_i as mean of subsets of eigenvalues with close moduli and as the estimates given by the tropical roots. On the right the tropical roots are coupled with estimates given by the Pellet theorem.*

This leads to a linearization $A(x)$ that is well-conditioned for the smaller eigenvalues but with a higher conditioning on the eigenvalues of bigger modulus as can be seen in Figure 6.3 on the left (the eigenvalues are ordered in nonincreasing order with respect to their modulus). It can be seen, though, that coupling the tropical roots with the standard Pellet theorem and altering the b_i by adding a value slightly smaller than the upper bound (in this example we have chosen 5 but the result is not very sensitive to this choice) leads to a much better result that is reported in Figure 6.3 on the right. In the right figure we have used `b = [ 1.7e-4, 1.4e-3, -1.4e-3, 5 ]`. This seems to justify that there exists a link between the quality of the approximations obtained through the tropical roots and the conditioning properties of the secular linearization.

We analyzed another example problem from the NLEVP collection that is called `planar_waveguide`. The results are shown in Figure 6.4. This problem is a PEP of degree 4 with two tropical roots approximately equal to 127.9 and 1.24. Again, it can be seen that for the eigenvalues of smaller modulus (that will be near the tropical root 1.24) the Frobenius linearization and the secular one behave in the same way, whilst

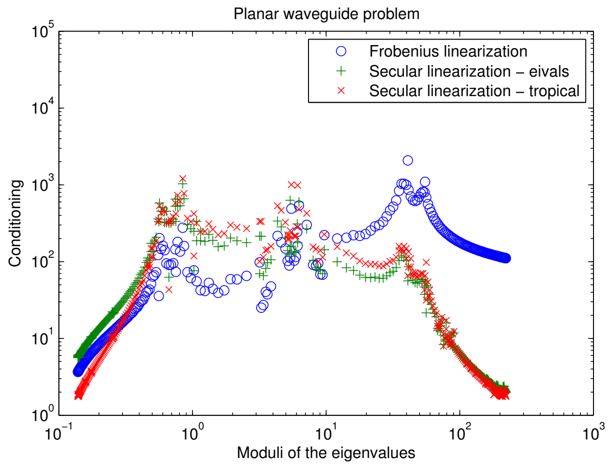

FIGURE 6.4. *Conditioning of the eigenvalues for three different linearizations on the* `planar_waveguide` *problem.*

for the bigger ones the secular linearization has some advantage in the conditioning. This may be justified by the fact that the Frobenius linearization is similar to a secular linearization on the roots of the unity.

Note that in this case the information obtained by the tropical roots seems more accurate than in the `orr_sommerfeld` case, so the secular linearization built using the tropical roots and the one built using the block-mean of the eigenvalues behave approximately in the same way.

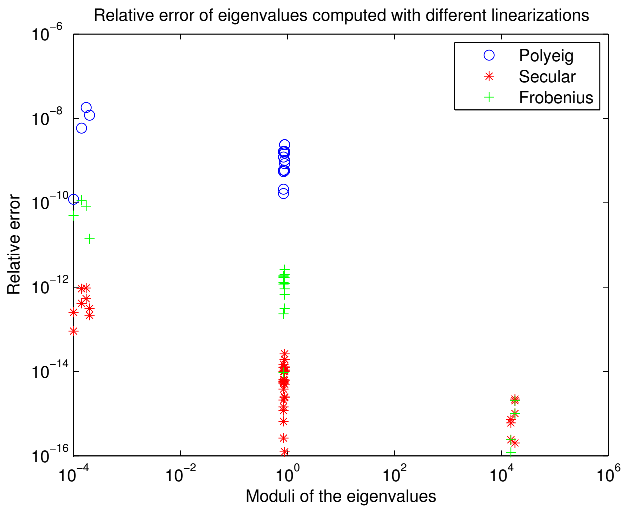

FIGURE 6.5. *The accuracy of the computed eigenvalues using polyeig, the Frobenius linearization and the secular linearization with the b_i obtained through the computation of the tropical roots.*

As a last example, we have tried to find the eigenvalues of a matrix polynomial defined by integer coefficients. We have used polyeig and our secular linearization (using the tropical roots as b_i) and the QZ method. We have chosen the polynomial

$P(x) = P_{11}x^{11} + P_9x^9 + P_2x^2 + P_0$ where

$$P_{11} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ & 1 & 1 & 1 \\ & & 1 & 1 \\ & & & 1 \end{bmatrix}, \; P_9 = 10^8 \begin{bmatrix} 3 & 1 & & \\ 1 & 3 & 1 & \\ & 1 & 3 & 1 \\ & & 1 & 3 \end{bmatrix}, P_2 = 10^8 P_{11}^t, \; P_0 = \begin{bmatrix} 1 & & & \\ & 2 & & \\ & & 3 & \\ & & & 4 \end{bmatrix}.$$

In this case the tropical roots are good estimates of the blocks of eigenvalues of the matrix polynomial. We obtain the tropical roots $1.2664 \cdot 10^4$, 0.9347 and $1.1786 \cdot 10^{-4}$ with multiplicities 2, 7 and 2, respectively. We have computed the eigenvalues with a higher precision and we have compared them with the results of `polyeig` and of `eig` applied to the secular linearization and to the standard Frobenius linearization. Here, the secular linearization has been computed with the standard floating point arithmetic. As shown in Figure 6.5 we have achieved much better accuracy with the latter choice. The secular linearization has achieved a relative error of the order of the machine precision on all the eigenvalues except the smaller block (with modulus about 10^{-4}). In that case the relative error is about 10^{-12} but the absolute error is, again, of the order of the machine precision. Moreover, `polyeig` fails to detect the eigenvalues with bigger modules, and marks them as eigenvalues at infinity. This can be noted by the fact that the circles relative to the bigger eigenvalues are missing in `polyeig` plot of Figure 6.5.

Acknowledgments. We wish to thank the referees for their careful comments and useful remarks that helped us to improve the presentation.

REFERENCES

[1] Maha Al-Ammari and Françoise Tisseur. Standard triples of structured matrix polynomials. *Linear Algebra Appl.*, 437(3):817–834, 2012.

[2] Amir Amiraslani, Robert M. Corless, and Peter Lancaster. Linearization of matrix polynomials expressed in polynomial bases. *IMA J. Numer. Anal.*, 29(1):141–157, 2009.

[3] Timo Betcke, Nicholas J. Higham, Volker Mehrmann, Christian Schröder, and Françoise Tisseur. NLEVP: A collection of nonlinear eigenvalue problems. `http://www.mims.manchester.ac.uk/research/numerical-analysis/nlevp.html`.

[4] Dario A. Bini, Guy Latouche, and Beatrice Meini. *Numerical methods for structured Markov chains*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 2005. Oxford Science Publications.

[5] Dario A. Bini and Vanni Noferini. Solving polynomial eigenvalue problems by means of the Ehrlich-Aberth method. *Linear Algebra and its Applications*, 439(4):1130–1149, 2013.

[6] Dario A. Bini, Vanni Noferini, and Meisam Sharify. Locating the eigenvalues of matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 34(4):1708–1727, 2013.

[7] Dario A. Bini and Victor Y. Pan. *Polynomial and matrix computations. Vol. 1*. Progress in Theoretical Computer Science. Birkhäuser Boston, Inc., Boston, MA, 1994. Fundamental algorithms.

[8] Dario A. Bini and Leonardo Robol. Quasiseparable Hessenberg reduction of real diagonal plus low rank matrices and applications. Submitted.

[9] Dario A. Bini and Leonardo Robol. Solving secular and polynomial equations: A multiprecision algorithm. *Journal of Computational and Applied Mathematics*, 272:276–292, 2014.

[10] Robert M. Corless. Generalized companion matrices in the Lagrange basis. In L. Gonzalez-Vega and T. Recio, editors, *Proceedings EACA*, June 2004.

[11] Robert M. Corless and Stephen M. Watt. Bernstein bases are optimal, but, sometimes, Lagrange bases are better. In *Proceedings of SYNASC, Timisoara*, pages 141–153, 2004.

[12] Fernando De Terán, Froilán M. Dopico, and D. Steven Mackey. Linearizations of singular matrix polynomials and the recovery of minimal indices. *Electron. J. Linear Algebra*, 18:371–402, 2009.

[13] Fernando De Terán, Froilán M. Dopico, and D. Steven Mackey. Fiedler companion linearizations and the recovery of minimal indices. *SIAM J. Matrix Anal. Appl.*, 31(4):2181–2204, 2009/10.

[14] Fernando De Terán, Froilán M. Dopico, and D. Steven Mackey. Fiedler companion linearizations for rectangular matrix polynomials. *Linear Algebra Appl.*, 437(3):957–991, 2012.

[15] Fernando De Terán, Froilán M. Dopico, and D. Steven Mackey. Spectral equivalence of matrix polynomials and the index sum theorem. Technical report, MIMS E-Prints, Manchester, 2013.

[16] Stéphane Gaubert and Meisam Sharify. Tropical scaling of polynomial matrices. In *Positive systems*, volume 389 of *Lecture Notes in Control and Inform. Sci.*, pages 291–303. Springer, Berlin, 2009.

[17] Israel Gohberg, Peter Lancaster, and Leiba Rodman. *Matrix polynomials.* Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1982. Computer Science and Applied Mathematics.

[18] Gene H. Golub. Some modified matrix eigenvalue problems. *SIAM Review*, 15(2):318–334, 1973.

[19] Nicholas J. Higham, D. Steven Mackey, Niloufer Mackey, and Françoise Tisseur. Symmetric linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 29(1):143–159 (electronic), 2006/07.

[20] Nicholas J. Higham, D. Steven Mackey, and Françoise Tisseur. The conditioning of linearizations of matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28(4):1005–1028, 2006.

[21] Nicholas J. Higham, D. Steven Mackey, and Françoise Tisseur. Definite matrix polynomials and their linearization by definite pencils. *SIAM J. Matrix Anal. Appl.*, 31(2):478–502, 2009.

[22] Guy Latouche and Vaidyanathan Ramaswami. *Introduction to Matrix Geometric Methods in Stochastic Modeling.* ASA-SIAM Series on Statistics and Applied Probability. SIAM, Philadelphia, 1999.

[23] D. Steven Mackey, Niloufer Mackey, Christian Mehl, and Volker Mehrmann. Structured polynomial eigenvalue problems: good vibrations from good linearizations. *SIAM J. Matrix Anal. Appl.*, 28(4):1029–1051 (electronic), 2006.

[24] D. Steven Mackey, Niloufer Mackey, Christian Mehl, and Volker Mehrmann. Vector spaces of linearizations for matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 28(4):971–1004 (electronic), 2006.

[25] Marcel F. Neuts. *Structured stochastic matrices of $M/G/1$ type and their applications*, volume 5 of *Probability: Pure and Applied.* Marcel Dekker, Inc., New York, 1989.

[26] Marcel F. Neuts. *Matrix-geometric solutions in stochastic models.* Dover Publications, Inc., New York, 1994. An algorithmic approach, Corrected reprint of the 1981 original.

[27] A.E. Pellet. Sur un mode de séparation des racines des équations et la formule de Lagrange. *Bull. Sci. Math.*, 5:393–395, 1881.

[28] Vaidyanathan Ramaswami. The generality of Quasi Birth-and-Death processes. In S. Chakravarthy A. Alfa, editor, *Advances in Matrix Analytic Methods for Stochastic Models*, pages 93–113, 1998.

[29] Brian T. Smith. Error bounds for zeros of a polynomial based upon Gershgorin's theorems. *J. Assoc. Comput. Mach.*, 17:661–674, 1970.

[30] Leo Taslaman, Françoise Tisseur, and Ion Zaballa. Triangularizing matrix polynomials. *Linear Algebra Appl.*, 439(7):1679–1699, 2013.

[31] Françoise Tisseur. Backward error and condition of polynomial eigenvalue problems. *Linear Algebra and its Applications*, 309(1):339–361, 2000.

[32] Françoise Tisseur and Karl Meerbergen. The quadratic eigenvalue problem. *SIAM Rev.*, 43(2):235–286, 2001.

[33] Françoise Tisseur and Ion Zaballa. Triangularizing quadratic matrix polynomials. *SIAM J. Matrix Anal. Appl.*, 34(2):312–337, 2013.